

## YOLO BASED TINY VEHICLE DETECTION

DR.V.ANANTHA KRISHNA<sup>1</sup>, RAVIRALA AKARSHA<sup>3</sup>, BHATRAJU DEEPIKA<sup>4</sup>,  
TABASSUM FATHIMA<sup>5</sup>

<sup>1</sup>Professor, Department of Computer Science and Engineering, Sridevi Women's Engineering College,  
Hyderabad, India

Email: [Krishnaanthav@gmail.com](mailto:Krishnaanthav@gmail.com),

<sup>2,3,4</sup>B.Tech Student, Department of Computer Science and Engineering, Sridevi Women's Engineering  
College, Hyderabad, India

**Abstract:** To solve real-life problems for different smart city applications, using deep Neural Network, such as parking occupancy detection, requires fine-tuning of these networks. For large parking, it is desirable to use a central-plane camera located at a high distance that allows the monitoring of the entire parking space or a large parking area with only one camera. However, if we use our own data different from that of the general-purpose datasets, such as COCO and Image Net, we have a large margin for improvisation. We propose a modified, yet lightweight, deep object detection model based on the YOLO-v5 architecture. The proposed model can detect large, small, and tiny objects. Specifically, we propose the use of a multi-scale mechanism to learn deep discriminative feature representations at different scales and it is automatically determine the most suitable scales for detecting objects in a scene (i.e., in our case vehicles). The proposed multi-scale module reduces the number of trainable parameters compared to the original YOLO-v5 architecture. The experimental results also demonstrate that precision is improved by a large margin. The results show a small reduction from 7.28 million parameters of the YOLO-v5-S profile to 7.26 million parameters in our model. In addition, we reduced the detection speed by inferring 30 fps compared to the YOLO-v5-L/X profiles. In addition, the tiny vehicle detection performance was significantly improved compared to the YOLO.

**Index Terms** - Smart city applications, Deep Neural Network, Parking occupancy detection, YOLO-v5 architecture, Object detection model, Multi-scale mechanism, Discriminative feature representations, Trainable parameters, Detection speed, Tiny vehicle detection.

**Keywords:** - Smart city applications, Deep Neural Network, Parking occupancy detection, YOLO-v5 architecture, Object detection model, Multi-scale mechanism, Discriminative feature representations, Trainable parameters, Detection speed, Tiny vehicle detection.

### 1. INTRODUCTION

The ever-increasing city population has reached a point where the management of city resources has

become a critical and important problem for large cities. Following this objective, one of the most time-consuming processes for drivers is finding a parking spot. A driver will travel extra kilometers per year to find an available slot, which not only has a direct impact on the driver's time consumption but also on

environmental pollution using deep Neural Network, such as parking occupancy detection, requires fine-tuning of these networks. For large parking, it is desirable to use acenital-plane camera located at a high distance that allows the monitoring of the entire parking space or a large parking area with only one camera. Today's most popular object detection models, such as YOLO, achieve good precision scores at real-time speed.

## 2. LITERATURE REVIEW

The development of smart cities has become a focal point in urban planning, aiming to enhance efficiency, sustainability, and overall quality of life. Smart parking systems, as an integral component of smart urban infrastructure, play a crucial role in addressing traffic congestion, reducing air pollution, and improving the overall urban mobility experience. This literature survey explores key contributions in the field of smart parking solutions, with a focus on leveraging deep learning and IoT-based analytics. The review encompasses studies from various researchers and covers aspects ranging from survey methodologies to practical implementations.

Atitallah et al. (2020) present a comprehensive review in "Leveraging deep learning and IoT big data analytics to support smart cities development." The authors emphasize the potential of deep learning techniques for enhancing smart city functionalities. The paper provides insights into the integration of deep learning and big data analytics to support the development of smart cities, offering a theoretical foundation for subsequent research in this domain [1].

Lin et al. (2017) contribute to the field with their survey on smart parking solutions in "A survey of smart parking solutions." This work delves into different methodologies and technologies employed in smart parking systems, providing a holistic view of the state-of-the-art in the domain. The authors address the challenges and opportunities associated with smart parking and highlight potential directions for future research [2].

Faraji and Jafari Naozar (2019) explore the impact of smart parking on city management and air pollution reduction in "Smart parking: An efficient approach to city's smart management and air pollution reduction." This study emphasizes the environmental benefits of efficient parking management and establishes a connection between smart parking systems and air quality improvement, providing a holistic perspective on the broader implications of smart city technologies [3].

Cats et al. (Year) offer a methodology for measuring parking occupancy in urban centers in "Survey methodology for measuring parking occupancy." The authors investigate the impacts of on-street parking pricing schemes on parking occupancy. This work contributes valuable insights into the practical challenges of implementing smart parking solutions and provides a basis for assessing the effectiveness of pricing strategies [4].

Amato et al. (Year) focus on decentralized parking lot occupancy detection using deep learning in "Deep learning for decentralized parking lot occupancy detection." The authors present an innovative approach to parking occupancy detection, emphasizing the potential of deep learning techniques. The study provides a technological

perspective on addressing the challenges of decentralized parking management in smart cities [5].

Koumetio Tekouabou et al. (Year) enhance parking availability prediction using IoT and ensemble-based models in "Improving parking availability prediction in smart cities with IoT and ensemble-based model." The authors propose a data-driven approach to enhance the accuracy of parking availability prediction in smart cities. The integration of IoT and ensemble models contributes to the optimization of smart parking systems [6].

In addition to the research articles, several influential books on machine learning and deep learning provide essential background knowledge for understanding the methodologies employed in smart parking solutions. These include "The Hundred-Page Machine Learning Book" by Andriy Burkov [7], "Hands-On Machine Learning with Scikit-Learn, Keras, and Tensor Flow" by Aurélien Géron [8], "Pattern Recognition and Machine Learning" by Christopher Bishop [9], and "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville [10].

In conclusion, this literature survey presents a comprehensive overview of key contributions in the field of smart parking solutions, emphasizing the integration of deep learning and IoT-based analytics for smart city development. The reviewed studies collectively contribute to our understanding of the challenges, methodologies, and potential future directions in the implementation of efficient and sustainable smart parking systems.

### 3. METHODOLOGY

In previous work they introduced a solution for detecting parking lot occupancy using smart cameras. The solution is based on a deep Convolutional Neural Network (CNN) designed specifically for smart cameras. They compare this solution with state-of-the-art approaches using two visual datasets: PK Lot and CNR Park-EXT. Additionally, a new dataset called CNR Park-EXT has been created in the context of their research, which is publicly available to the scientific community. In another study they used two methods for parking lot occupancy detection. The first method involves converting the colored lot image to gray scale, then to black / white and computing the mean of the resulted image. The image is then classified as occupied or empty according to a threshold. These cond methods depend on deep learning techniques and is a deep learning network consisting of 11 layers, 3 of them are convolution layers with different kernel size.

#### Drawbacks:

1. The grayscale-based method, relying on simple thresholding, might not be scalable to handle parking lots with large numbers of parking spaces, as it may not generalize well to diverse parking lot layouts and object sizes.
2. The deep CNN model typically requires substantial computational resources, making it less suitable for real-time or resource-constrained applications, such as deployment on smart cameras with limited processing power.
3. The existing methods might struggle to accurately detect tiny objects, like small cars

or motorcycles, due to their reliance on fixed-size filters and pooling operations in the CNN architecture.

We propose a modified, yet lightweight, deep object detection model based on the YOLOv5 architecture. Our model can detect large, small, and tiny objects. Specifically, introducing a new first layer in the state-of-the-art model, that is, YOLOv5. This new first layer was used to extract discriminative features from different scales. The proposed work suggested replacing in the focus layer of YOLO-v5 with a multi-scale layer based on an Efficient Neural Network (ENet). The introduction of this new layer in YOLO-v5 resulted in a significant out performance of the baseline model. To avoid redundant low-level features in the backbone of YOLO, we assess the impact of spatial and channel information using attention modules on the performance of tiny object detection. The proposed attention modules can capture meaningful information of the “where” and “what” for tiny object detection. The proposed multi-scale module reduces the number of trainable parameters compared to the original YOLO-v5 architecture.

#### Benefits:

1. The proposed modified YOLO-v5 architecture, with the new first layer and multi scale layer, is designed to handle large, small, and tiny objects, leading to improved accuracy in parking lot occupancy detection.
2. Our work aims to maintain a light weight model despite introducing modifications. This makes it more suitable for real-time

applications and deployment on resource constrained devices like smart cameras.

3. The introduction of attention modules in the modified YOLO-v5 allows the model to focus on meaningful spatial and channel information, leading to better performance in detecting tiny objects in parking lots.
4. The proposed multi-scale module in YOLO-v5 helps reduce the number of trainable parameters compared to the original architecture. This can potentially lead to faster training and reduced memory requirements.

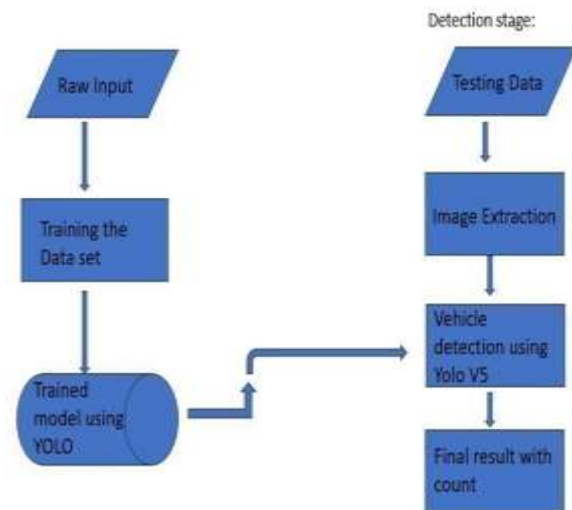


Fig 1 System Architecture

#### Modules:

The modules are:

- Import libraries: Import necessary Python libraries for the project.

- Import dataset: Load dataset for analysis and model training.
- Image processing: Preprocess images for optimal model input and analysis.
- Feature selection: Choose relevant features for effective model learning and prediction.
- Building the model - YOLO: Construct the YOLO (You Only Look Once) object detection model.
- Training the model: Train the YOLO model using the imported dataset.
- Predict: Apply the trained model to make predictions on new data.

#### 4. IMPLEMENTATION

Yolo V5 - YOLOv5 uses a Convolutional Neural Network (CNN) backbone to form image features. These features are combined in the model neck and sent to the head. The model head then interprets the combined features to predict the class of an image.

#### YOLO

YOLO is an algorithm that uses neural networks to provide real-time object detection. This algorithm is popular because of its speed and accuracy. It has been used in various applications to detect traffic signals, people, parking meters, and animals. This article introduces readers to the YOLO algorithm for object detection and explains how it works. It also highlights some of its real life applications.

Introduction to object detection

Object detection is a phenomenon in computer vision that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals. Object detection consists of various approaches such as fast RCNN, Retina Net, and Single-Shot Multi Box Detector (SSD). Although these approaches have solved the challenges of data limitation and modeling in object detection, they are not able to detect objects in a single algorithm run. YOLO algorithm has gained popularity because of its superior performance over the fore mentioned object detection techniques.

What is YOLO?

YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images.

YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.

This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously. The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO and YOLOv3.

Why the YOLO algorithm is important

YOLO algorithm is important because of the following reasons:

Speed

This algorithm improves the speed of detection because it can predict objects in real time.

High accuracy

YOLO is a predictive technique that provides accurate results with minimal background errors.

How the YOLO algorithm works

YOLO algorithm works using the following three techniques:

- Residual blocks
- Bounding box regression
- Intersection over Union (IOU)

Residual blocks

First, the image is divided into various grids. Each grid has a dimension of  $S \times S$ . The following image shows how an input image is divided into grids.



Fig 2 Input Images

5. EXPERIMENTAL RESULTS

Dataset Description:

The dataset is specifically curate for smart city applications, focusing on parking occupancy detection using a modified YOLO-v5 architecture. Captured through a cenital-plane camera positioned at a high distance, allowing comprehensive monitoring of entire parking spaces or large parking areas with a single camera. The dataset emphasizes diversity in parking scenarios, encompassing different lighting conditions, vehicle sizes, and parking space layouts.



Fig 3 Anaconda Prompt



Fig 4 Home Page



Fig 5 Registration page



Fig 9 Random selection of images

Fig 6 Login Page



Fig 7 Classification Performance



Fig 10 Result



Fig 8 Upload any images

## 6. CONCLUSION

The conclusion of our proposed work is that the proposed modification of the YOLOv5 model, which targets tiny car objects from a cenital view, outperformed the original YOLOv5 model for this specific application with a good precision. The proposed model also slightly improved the Recall and mAP values while maintaining the same number of trainable parameters. However, the adapted variation yields a decrease in speed compared to the small and medium profiles of the YOLO-v5 network, although it exceeds the large and extreme profiles. The single-branch solution out speeds the YOLO-v5 small profile and was almost as precise as the multi-branch solution. We also mentions that ongoing work aims to develop a reliable tracker based on the developed detector, and future work aims to use the developed detector and tracker in low-end terminals, such as a

Field-Programmable Gate Array (FPGA) or an NVIDIA Jet son Nano Developer Kit for real-time parking monitoring.

### FUTURE ENHANCEMENTS

YOLO-based tiny vehicle detection can be enhanced in several ways. one approach is to improve the accuracy and speed of detection by optimizing the network architecture and training process. Another way is to incorporate more diverse datasets to enhance the model's ability to detect tiny vehicles in different scenarios. Additionally, continuous research and advancements in computer vision techniques can contribute to improving the performance of YOLO-based detection. These enhancements aim to make the models more accurate and efficient in detecting tiny vehicles. The proposed model can detect large, small, and tiny objects. Specifically, we propose the use of a multi-scale mechanism to learn deep discriminative feature representations at different scales and automatically determine the most suitable scales for detecting objects in a scene.

Object detection is a phenomenon in computer vision that involves the detection of various objects in digital images or videos. Some of the objects detected include people, cars, chairs, stones, buildings, and animals. Object detection consists of various approaches such as fast RCNN, Retina Net, and Single-Shot Multi Box Detector(SSD). Although these approaches have solved the challenges of data limitation and modeling in object detection, they are not able to detect objects in a single algorithm run. YOLO algorithm has gained popularity because of its superior performance over the aforementioned object detection techniques.

These enhancements aim to make the models more accurate and efficient in detecting tiny vehicles. The proposed model can detect large, small, and tiny objects. Specifically, we propose the use of a multi-scale mechanism to learn deep discriminative feature representations at different scales and automatically determine the most suitable scales for detecting objects in a scene.

### REFERENCES

- [1] S.B. Atitallah, M.Driss, W.Boulila, and H.B.Ghézala, "Lever aging deep learning an diot big data analytics to support the smart cities development: Review and future directions," *Computer Science Review*, vol. 38, p. 100303, 2020
- [2] T. Lin, H. Rivano, and F. Le Mouël, "Asurvey of smart parking solutions," *IEEE Transactions on Intelligent Transportation Systems*, vol.18,no.12, pp.3229–3253, 2017.
- [3] S.J. Farajiand M. Jafari Naozar, "Smart parking: An efficient approach to city's smart management and air pollution reduction," *Journal ofAir Pollution* 53–72, Feb. 2019.
- [4] O.Cats, C.Zhang, and A.Nissan, "Survey methodology for measuring parking occupancy: Impacts of an on-street parking pricing scheme in an urban center," *Transport Policy*, vol.
- [5] G.Amato, F.Carrara, F.Falchi, C.Gennaro, C.Meghini, and C.Vairo, "Deep learning for decentralized parking lot occupancy detection," *Expert Systems with Applications*
- [6] S.C. Koumetio Tekouabou, E.A. Abdellaoui Alaoui, W.Charif,andH.Silkan, "Improving parking

availability prediction in smart cities with iot and ensemble-based model, "Journal M.A. Merzoug, A. Mostefaoui, G. Gianini, and E. Damiani, "Smart connected parking lots based on secured multimedia iot devices," Computing, 2021.

[7] The Hundred-Page Machine Learning Book by Andriy Burkov:

[8] Hands-On Machine Learning with Scikit-Learn, Keras, and Tensor Flow by Aurélien Géron:

[9] Pattern Recognition and Machine Learning" by Christopher

[10] Deep Learning" by Ian Good fellow, Yoshua Bengio, and Aaron Courville: