

---

# YOLO-Vacant Parking Spot Detector

---

**Andrew Park**

Electrical and Computer Engineering  
A17406465

**Minsang Kim**

Electrical and Computer Engineering  
A16636382

## Abstract

Describe the overall picture of your project here. Give a short description about the different parts you are going to cover in the proposal, but do not include too many details here.

We do not have requirements for the code. The only requirement is that it should be from you, not copy paste from online repo.

## 1 Introduction

Finding a vacant parking space in a given lot is a tedious and costly task: a person must drive through the whole parking lot in order to find a vacant space. This ineffective process will waste gasoline and time while searching for a parking space. Furthermore, this will cause congestion in the parking lot. Therefore, by displaying the occupancy of spaces in a given parking lot will significantly reduce the time cost, traffic congestion, and gasoline consumption. YOLO-Parking will aid this wasteful process by displaying vacant and occupied parking spaces. Although there are existing methods to detect the occupancy of a parking space, such as car detector on every parking space and other object detector architectures that detects the occupancy, most of them are expensive or slow in detecting parking spots[2]. Furthermore, there are several traditional models to detect the instance segmentation of a vacant parking spot, but they are slow in detection for videos.

In this paper, we propose a new model to detect vacant parking spots: You Only Look Once (YOLO). YOLO architecture is one of the fastest architecture that accurately performs object detection [3]. Additionally, several cameras mounted in the parking lot to detect vacancy of a space is relatively cheaper than placing a car detector on every parking space in a parking lot [2].

## 2 Related Work

**R-CNN** R-CNN [2] is a widely used object detecting CNN model with relatively high accuracy. The first step is to regional proposal using selective search methods to propose bounding box. Then, for each region proposal, the model warped the proposed regions to use them as inputs for a convolutional neural network. The CNN classifies each input proposed region using SVM specific to a class [1]. However, the training rate is slow due to large number of parameters to learn, and the inference rate is also slow. In one of the previous approaches in detecting parking spots, Martin Marek proposed R-CNN: the model first pool proposed parking spot patches in the original image and then a CNN produces a binary output whether the space in each patch is occupied or not [2]. The authors provided two methods for pooling for the first step in R-CNN: the interpolation of pixels to construct quadrilateral bounding boxes and construction of minimal bounding boxes [2].

**Faster R-CNN FPN** The input image first undergoes a feature pyramid network (FPN). Then, the authors extract the features of the parking spots and pass them through respective classification head for classification. However, the input image must be resized, thus limiting the output image from being full resolution [2]. Both R-CNN and R-CNN is still slow in inference for object detection in videos. Faster R-CNN can use one of the two methods of pooling mentioned above [2].

These two models are merely the baseline of parking vacancy object detector using CNN. In this paper, we intend to improve the model in inference time by using YOLO model.

### 3 Method

In this section, you should describe your method (Do not need to be the exact method you describe in project proposal). Besides, You can include the following things(You do not have to cover all of them):

- the detailed structure of your method, including the formulations, network architecture figure.
- training algorithm, testing algorithm.
- what is the new proposed techniques compared to previous work, and the reason and strength of choosing the method.

### 4 Experiments

In this section, You should include the following things:

- the datasets you use:
  - the brief introduction of the dataset
  - the data format
  - other information related to your experiments
- your results
- ablation study on training your networks, how does the method work with more or less data, with/without some components (optional)

### 5 Supplementary Material

You should also include a video recording a presentation (with motivation, approach, results) for this project.

### References

- [1] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation, 2014.
- [2] Martin Marek. Image-based parking space occupancy classification: Dataset and baseline. *CoRR*, abs/2107.12207, 2021.
- [3] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016.