MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
IVAN FRANKO NATIONAL UNIVERSITY OF LVIV

Faculty of Applied Mathematics and Informatics
Department of Applied Mathematics

Course Project

DETECTING PARKING OCCUPANCY USING MACHINE LEARNING

Author: student of group PMP-62
specialty 113 – Applied Mathematics
____________ John Joshua Junior

Supervisor:
____________ As. Prof Yuriy Muzychuk

Lviv – 2021

TABLE OF CONTENTS

ABSTRACT

Detecting parking occupancy proffers a potential in reducing the congestion in crowded areas, by providing real-time indications of the availability of parking spaces. Current systems are mostly implemented for indoor environments using costly sensor-based techniques. Consequently, the increasing demand for parking space in outdoor environments, using low-cost image-based detection methods have become the focus of research and development in recent times. Driven by the notable performance of Convolutional Neural Networks (CNNs) in various image category recognition tasks, this study presents a thorough parking occupancy detection framework, using a modified Mask RCNN with Intersection over Union (IoU) to detect the occupancy of outdoor parking spaces from images. The model was trained and tested by the features learned from the Mask RCNN.

ABBREVATIONS

CNN - Convolutional Neural Networks

ReLU – Rectified Linear Unit

RCNN – Region Based Convolutional Neural Networks

IoU – Intersection over Union

HSV – Hue Saturation Value

RGB – Red, Green and Blue

CCTV - Closed Circuit Television

YOLO – You Only Look Once

OpenCV – Open Computer Vision

COCO – Common Objects in Context

FCN – Fully Convolutional Network

1. INTRODUCTION

1.1.Problem statement

In recent years, with rapid economic development and increase in standard of living, the ownership of automobiles has also increased annually. This development has brought about stationary traffic and other problems. Stationary traffic refers to the parking of vehicles, which includes short-term and long term parking. Short-term parking happens with passengers getting on and off or loading and unloading of goods within a short period, while long-term parking involves doing these things with a longer period of time in the parking lots. The problems of stationary traffic are mainly manifested in the following aspects: insufficient control of land use indicators, planning and layout, insufficient parking spaces in public buildings, a severe road occupation occurrence, and fewer parking spaces in parking lots. The above situation depicts that parking occupancy have become a very acute problem. As a result, detecting parking occupancy in order of priority has become the top of the list in resolving parking issues.

1.2.Aim of Research

The rapid increase in the number of motor vehicles in the world is due to the fast economic and social development. In contrast to this, the construction of parking lots has been relatively slow, thereby causing parking difficulties to become an increasingly prominent problem. The availability of parking spaces nowadays has become an issue that should not to be neglected, this is mostly because it consumes time and energy, amongst other factors. This research paper focuses on mitigating the problems associated with finding an available parking space using machine learning.

1.3. Dataset Description

1.3.1. Video from Lviv Parking lot and streets

This dataset will be created by capturing series of images from the various parking lots in Lviv and from various street parking space in Lviv Oblast, Ukraine. The images would either be captured using professional camera or via smartphone at various time intervals under various weather conditions.

1.4. Practical Value of Results

The benefits of the detecting parking occupancy for the region/city include improved traffic flow, less congestion, and better mobility and living conditions. It further helps in significantly curbing air and noise pollution.

Parking occupancy detection also has several social benefits, some of these are direct and indirect social benefits. Direct social benefits include increase in personal safety at night, reduction in accidents and car damage, minimizes theft, increases safety for pedestrians, provides accessible parking for all users including disabled and parents with children. While some of the indirect benefits includes reduction in anxiety levels among drivers, thus reducing the stress levels and avoiding unnecessary over-speeding and honking etc. These benefits even extend to local businesses by easing parking difficulties, which results in increased customer footfall. These also becomes of maximum benefit to the citizens through reduction in circling, which leading to savings in time and fuel, less congestion and general improvement in quality of life.

1.5.Related and existing research

Many research work has been carried out in related field, such as: Sensor-based approach used to detect a moving object entering the parking lot, where the size of the object and the time it takes to cross the gate where the sensor is located infers the occupation of a spot by the car [1]. It was also discovered that their use is broad but limited when budget is an issue, due to the cost linked to the installation and maintenance of sensors.

Another research in this field is that of the use of binary classifiers for Convolutional Neural Network (CNN) for the smart parking system [2]. The data acquired is an RGB image converted to HSV which then extracts the value V. This value was grayed and its outline sharpened. This process results are converted into binary images and entered into the CNN. However, the research does not explain how the system can detect each empty or filled parking spaces.

Another Research was carried out, and the method used to detect parking spaces availability was using a drone camera which has already been proposed [6]. At first, the parking area was detected by giving four points in each parking area. After that, they made corrections on the result images, so as to transform the lines created into a straight line to save its coordinates. In determining the availability of parking spaces, a line model was used where the already parked car formed a straight line. It so happens that if there is a parking space available, then the line will be disconnected, then the disconnected line will be marked as an available parking space. This research had an effective method for drones only, so even if this method was applied to a camera, it will still have deficiencies, where lines cannot be formed to show parking space availability.

In yet another research, the CNN method used smaller size Alexnet (mini Alexnet) to detect parking space availability [7]. The primary objective of this research, was for the detection of parking spaces to become automatic, using a CCTV camera made by the researcher. Using mini Alexnet and dataset made by the researcher, the research succeeded in detecting parking spaces with accuracy above 90%. Also, the mini Alexnet could automatically find the parking spaces from the whole image with whether or not the researcher makes a mark to find out the location of the parking spaces. Unfortunately, the research method did not explain how to retrieve each parking space’s location, therefore some questions were still left unanswered.

A recent research used a method called You Only Look Once (YOLO) V3 and Lite Alexnet, for the detection of parking spaces availability [8]. The research conducted automatic parking spaces detection by utilizing the boundary box of YOLO V3. The classification of parking space availability was done using CNN with Alexnet architecture which modified the number of layers and parameters called Lite Alexnet. The research had a constraint which is needed to help mark the form of lines on the parking spaces, so as to automatically find the parking spaces from the whole image.

In a more recent research, an approach using the Mask R-CNN model was used for counting vehicles in parking areas [4]. The methodology did not require a manually entered information on parking spot locations, thus allowing for a simple 'plug-and-play' installation. The proposed car detection allowed the input image have various perspectives, illuminations, and conclusions. The decline in the performance is slight and responded well to changes in perspective and illumination in the research. However, the research underestimated the number of vehicles in parking locations.

Finally, the use of the Mask R-CNN method for the detection of parking space in another research [3]. The research involved comparing the bounding box from detection with a manually annotated bounding box, in order to determine the classification of the parking space status. The research could detect parking spots correctly, about 90%. Although the research scenario had no obstruction in the parking lot so much that the research obtained high results, the research scenario is still seen as limited because of its inability to initialize manual annotations for each parking space.

However, in this project, we intend to make improvements upon the above listed references. We will auto detect parking lot regions, modify the Mask RCNN model and apply a threshold for the Intersection over Union (IoU) algorithm to adequately detect free parking space.

2. METHODOLOGY

This research involves 3 major stages which includes identification of region of interest, detection of cars and detection of free parking spots using by computing the IoU as seen below.

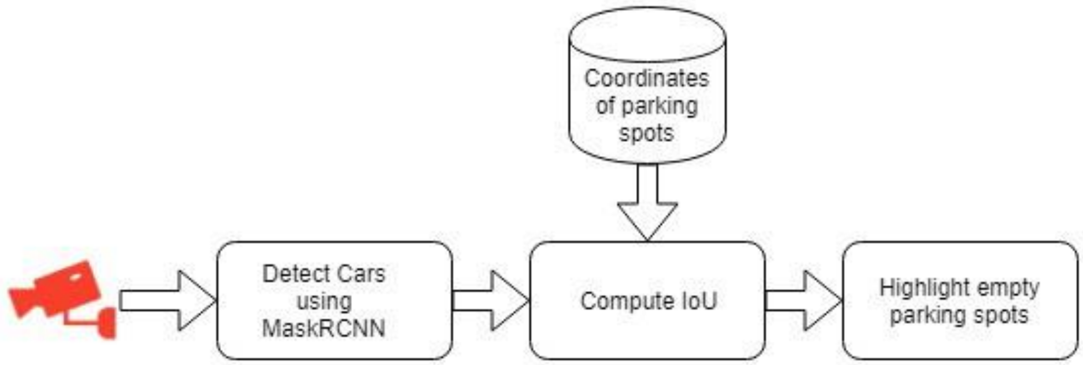

Figure 2.1: Detecting parking occupancy

On each frame of the input video, we will first use the Mask-RCNN object detection model to detect the cars and their bounding boxes. After getting the bounding boxes from the Mask-RCNN, we will compute the Intersection over Union (IoU) on each pair of the bounding boxes and parking spot coordinates. If the IoU value for any parking spot is greater than a certain threshold, we will consider that parking spot as occupied.

Operational definition of Terms/Terminologies

- Convolutional Neural Network (CNN): A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. It uses kernel convolution which is a process where small matrix of numbers (called kernel or filter) are passed over our image, thus transforming the image based on the values from filter. Feature map values are calculated according to the equation (1), where the input image is denoted by f and our kernel by h. The indexes of rows and columns of the result matrix are marked with m and n respectively.

$$G[m,n] = (f * h)[m,n] = \sum_{i} \sum_{j} h[i,j] f[m-i, n-j] \qquad (2.1)$$

After placing filter over a selected pixel, each value from kernel are multiplied in pairs with corresponding values from the image, then everything is summed. Sometimes filter does not fit perfectly fit the input image. For this scenario, padding is applied. This means adding zeros to the borders of the image matrix.

Pooling is then used to reduce the number of parameters thus reducing the size of the image, hence decreases the complexity and computations as shown in figure 2.

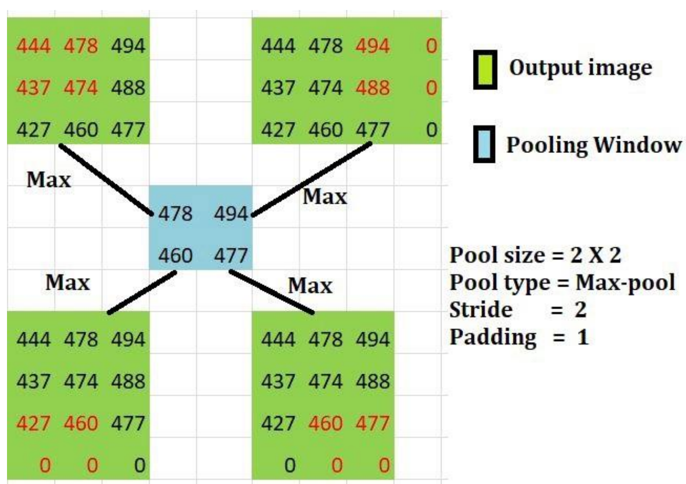

Figure 2.2: Padding and Pooling to an Image

The resulting output passes through an activation function ReLu which adds non-linearity to the convolutional network as shown in equation (2.2).

$$ReLU(x) = \begin{cases} 0 & if\ x < 0 \\ x & if\ x \geq 0 \end{cases} \tag{2.2}$$

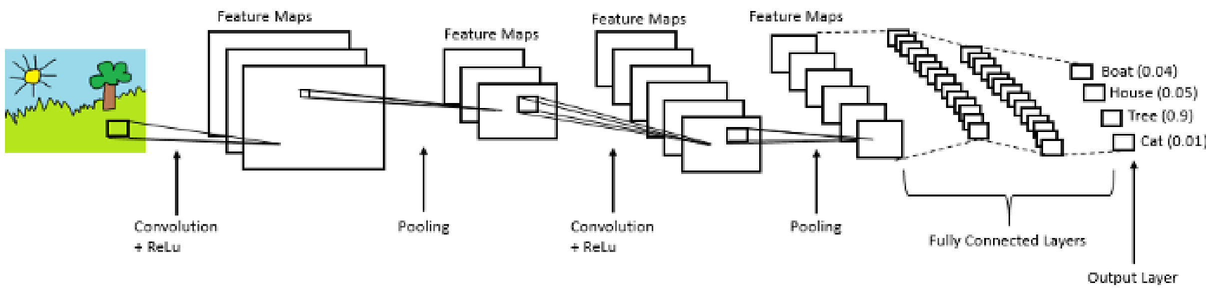

Figure 2.3: Convolution Neural Network CNN

- Region Based Convolutional Neural Networks (RCNN): The R-CNN first extracts many (e.g., 2000) region proposals from the input image (e.g., anchor boxes can also be considered as region proposals), labeling their classes and bounding boxes (e.g., offsets). Then a CNN is used to perform forward propagation on each region proposal to extract its features. Next, features of each region proposal are used for predicting the class and bounding box of this region proposal.

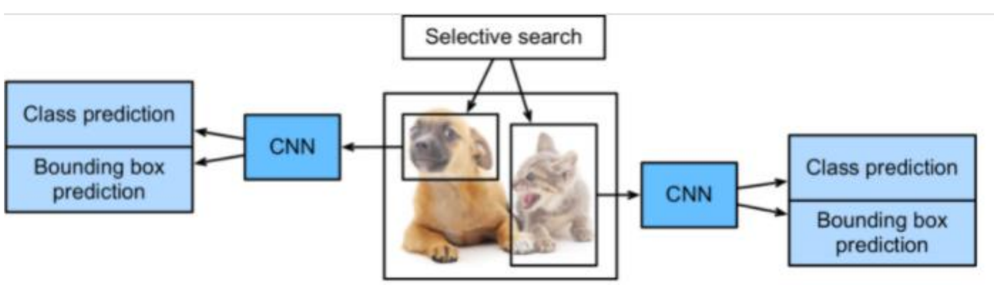

Figure 2.4: The R-CNN model

- Fast R-CNN: The main performance bottleneck of an R-CNN lies in the independent CNN forward propagation for each region proposal, without sharing computation. Since these regions usually have overlaps, independent feature extractions lead to much repeated computation. One of the major improvements of the *fast R-CNN* from the R-CNN is that the CNN forward propagation is only performed on the entire image [9].

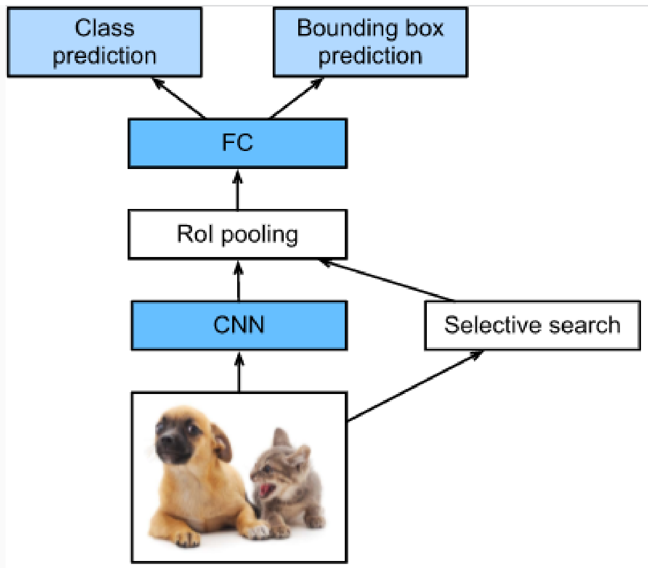

Figure 2.5: The fast R-CNN model.

- Faster R-CNN: To be more accurate in object detection, the fast R-CNN model usually has to generate a lot of region proposals in selective search. To reduce region proposals without loss of accuracy, the *faster R-CNN* proposes to replace selective search with a *region proposal network* [10].

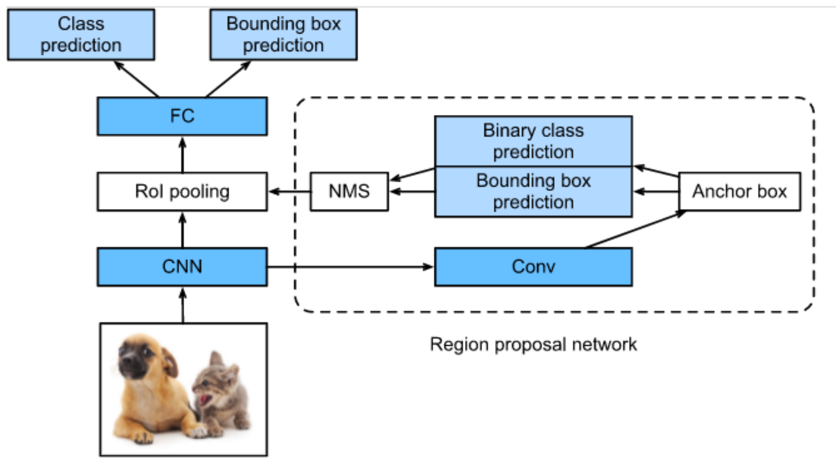

Figure 2.6: The faster R-CNN model.

- FCN – Fully Convolutional Network: FCN is a network that does not contain any "Dense" layers (as in traditional CNNs) instead it contains 1x1 convolutions that perform the task of fully connected layers (Dense layers).
- Mask R-CNN: Mask RCNN is a deep neural network aimed to solve instance segmentation problem in machine learning or computer vision. It can separate different objects in an image or a video. It takes as input an image, then outputs the object bounding boxes, classes and masks of objects in the image. Mask R-CNN extends Faster R-CNN to solve instance segmentation tasks. It achieves this by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. In principle, Mask R-CNN is an intuitive extension of Faster R-CNN, but constructing the mask branch properly is critical for good results. Most importantly, Faster R-CNN was not designed for pixel-to-pixel alignment between network inputs and outputs. This is evident in how RoIPool, the de facto core operation for attending to instances, performs coarse spatial quantization for feature extraction. To fix the misalignment, Mask R-CNN utilizes a simple, quantization-free layer, called RoIAlign, that faithfully preserves exact spatial locations.

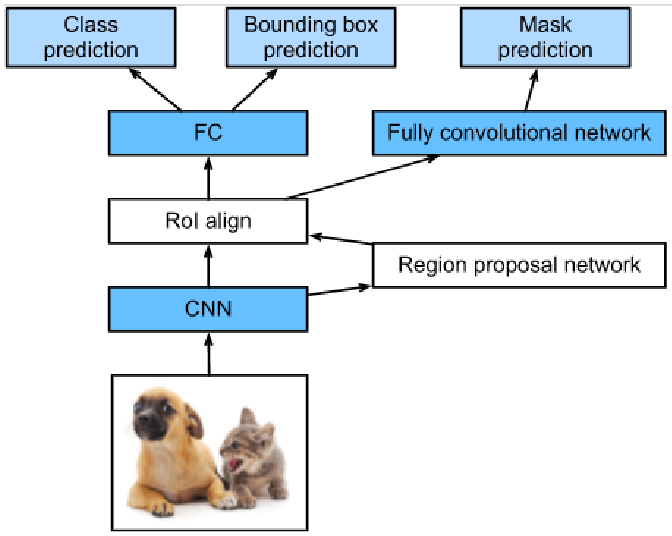

Figure 2.7:mask R-CNN

Secondly, Mask R-CNN decouples mask and class prediction; it predicts a binary mask for each class independently, without competition among classes, and relies on the network's RoI classification branch to predict the category.

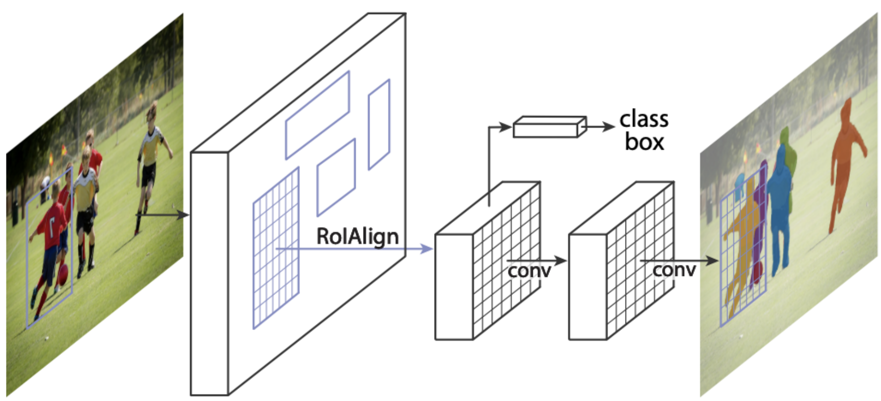

Figure 2.8: Mask RCNN

- OpenCV: OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications.

- Common Object in Context COCO: is a large-scale object detection, segmentation, and captioning dataset.

We explain each of the stages involved in the carrying out this project successfully. These include;

2.1.Identification of Region of Interest

The very first step in a parking space detection system is to identify the region of interest which is the total parking spots. This is by taking a frame/image from our video/stream of the parking location and we will mark the parking regions. The frame/image should be the parking spot of interest at full capacity. Then Mask RCNN is applied to extract the bounding boxes of the parked vehicle. The bounding boxes would be our point of reference to detecting if a slot is empty.

2.2. Detecting Vehicles in a Video

To detect cars in a video we will use the Mask-RCNN. It is a convolutional neural network trained on millions of images and videos from several datasets, including the COCO dataset, to detect various objects and their boundaries. Mask-RCNN is built on the top of the Faster-RCNN object detection model. Here we will use the implementation of Mask-RCNN by matterport because M-RCNN has very good accuracy and matterport's implementation is very easy to use.

2.3.Calculating Intersection Over Union (IoU)

Intersection over Union is an evaluation metric. This calculates the similarity (Jaccard similarity coeefficient) between known parking box area and current vehicle box area as shown in Equation (1). In other words, it is the ratio of the area of intersection and area of union as shown in figure 2.9.

$$IoU = J(A,B) = \frac{|A \cap B|}{A \cup B} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{2.1}$$

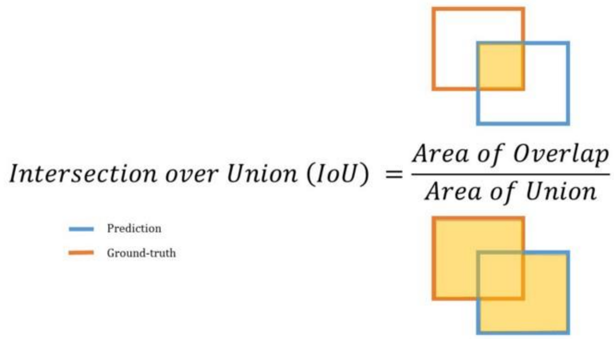

Figure 2.9: Calculation of IoU

We will compute the IoU for every pair of parking spot coordinates and bounding box of cars. If the IoU for a pair is higher than a certain threshold, we will consider that parking spot as occupied

3. RESULTS

a. Identification of Region of Interest: M-RCNN is used on every frame of the video and it returned a dictionary that containing the bounding box coordinates, masks of detected objects, confidence score for each prediction, and class_ids of detected objects. Using the class_ids we filtered out the bounding boxes of the cars, trucks, and buses as seen below.

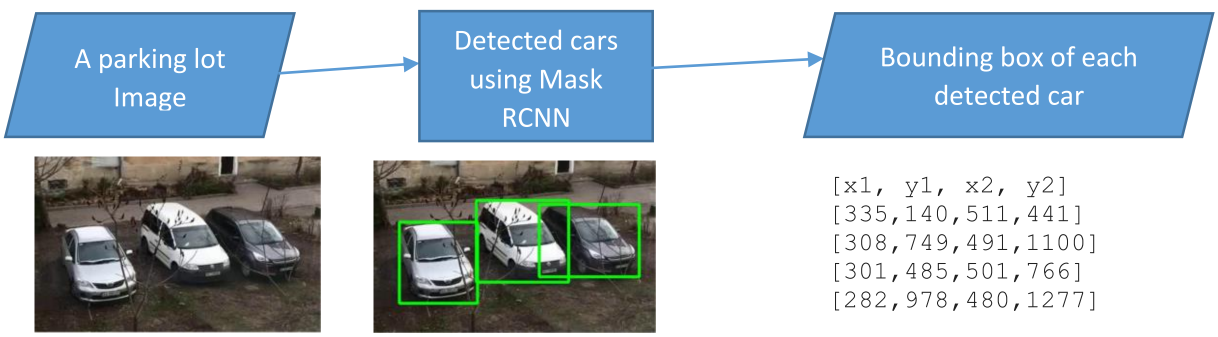

Figure 3.1: Detecting Vehicles Using Mask RCNN for a full parking spot

b. Detecting Vehicles when parking slot in not full capacity

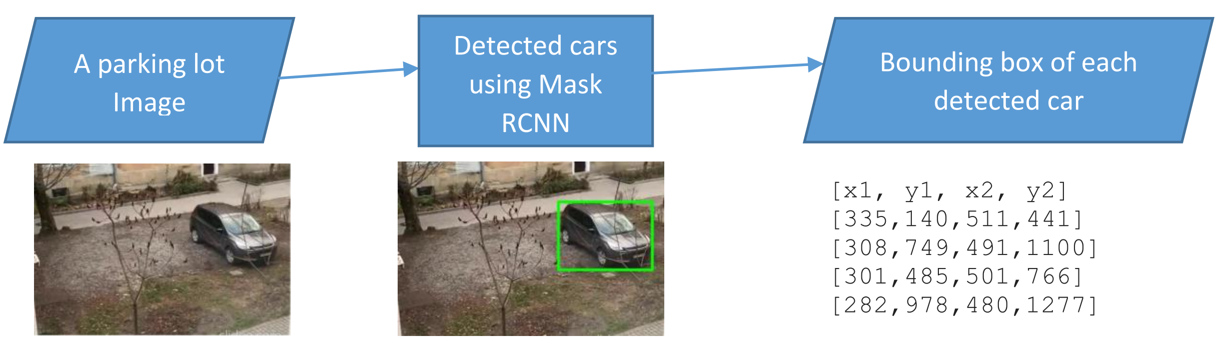

Figure 3.2: Detecting Vehicles Using Mask RCNN when the parking lot is not full

c. Applying IoU threshold to get free parking slot

Here we applied an IoU threshold of 0.15 to detect free parking space as seen below

Figure 3.3: Detected parking space

3.1. Performance metrics evaluation

In this research, we measured the performance of detection results quantitatively. We used various IoU value to calculate parking space detection accuracy, sensitivity and specificity.

3.2. Evaluation Criteria

For the evaluation we used three measures: overall accuracy, sensitivity, and specificity, as defined in Equations (3.1), (3.2), and (3.3) respectively. In the equations, TP (True Positive) is the number of occupied sub-images classified as occupied, TN (True Negative) is the number of unoccupied sub-images classified as unoccupied, FP (False Positive) is the number of unoccupied sub-images classified as occupied, and FN (False Negative) is the number of occupied sub-image classified as unoccupied

$$Overall\ accuracy\ = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.1}$$

$$Sensivity\ (True\ Positive\ Rate) = \frac{TP}{TP+TN} \tag{3.2}$$

$$Specificity\ (True\ Negative\ Rate) = \frac{TN}{TP+TN} \tag{3.3}$$

3.3. Evaluation Results

The test was done on a sample of 17 distinct frames/images extracted from a video capturing a mini-parking space at dormitory No. 8 of the IFNUL

The table and figure shows the plot of the accuracy at various IoU threshold values.

Table 3.1: IoU Threshold, Accuracy, Sensitivity and Specificity

IoU Threshold	Accuracy	Sensitivity	Specificity
0.10	0.70	1.0	0.0
0.15	0.70	1.0	0.0
0.20	0.70	1.0	0.0
0.25	0.70	1.0	0.0
0.30	0.60	0.67	0.33
0.35	0.70	0.57	0.43
0.40	0.70	0.57	0.43
0.45	0.70	0.57	0.43

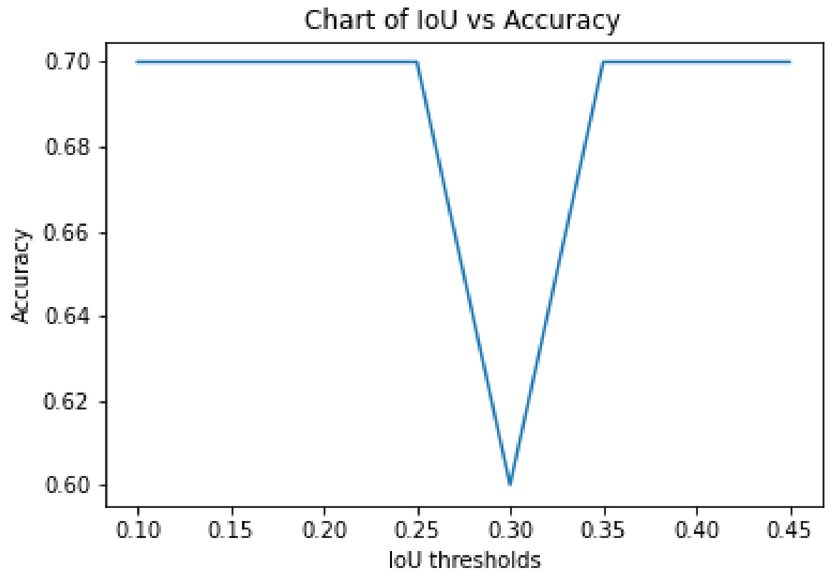

Figure 3.4: IoU vs Accuracy

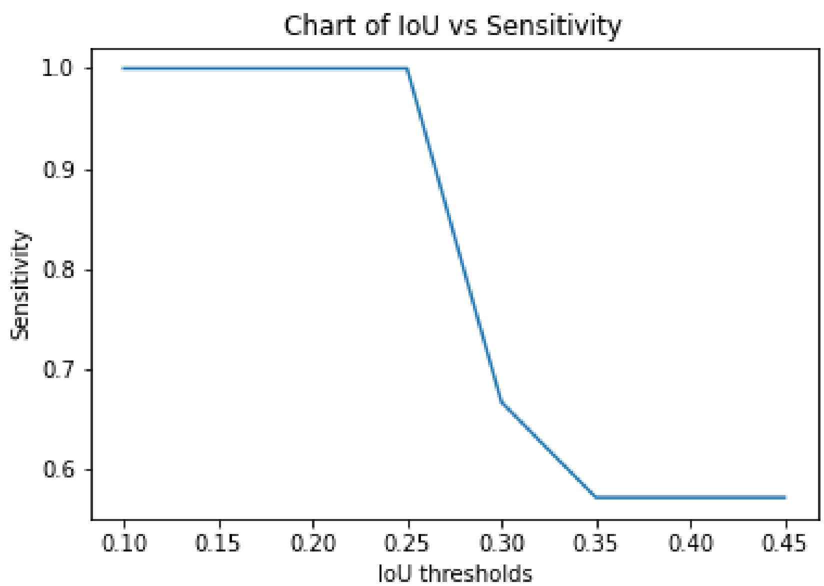

Figure 3.5: IoU vs Sensitivity

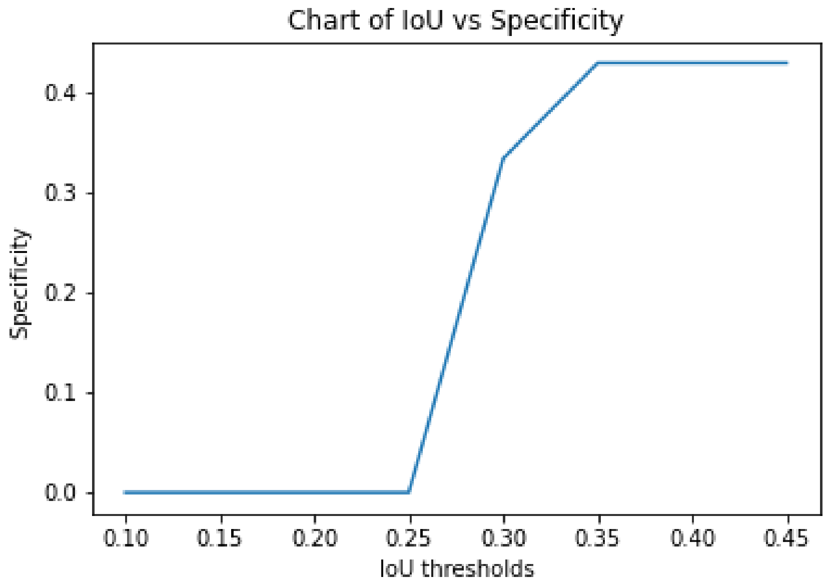

Figure 3.6: IoU vs Specificity

From this, we can see that the detector performed best at IoU threshold values less than 0.25. We can also see how sensitive the IoU value affects our result.

4. CONCLUSION

This research developed a system for detecting parking space occupancy using machine learning through video data analysis from the outdoor parking area. This involved two main stages; the first stage is identifying the region of Interest by marking the parking position on the image of a full parking lot. The second stage is feature extraction based on the region of interest (ROI) of each marking on the parking area. Experiments for parking space detection are evaluated using the accuracy of IoU. Using Mask R-CNN as the proposed object detection method provides an accurate automatic parking space occupancy detection and can reduce unnecessary maintenance installation costs. The major solution does not require additional information for manual segmentation of video data like other studies, making it more flexible.

REFERENCES

[1] S. Lee, D. Yoon, and A. Ghosh, "Intelligent parking lot application using wireless sensor networks," in Collaborative Technologies and Systems, 2008. CTS 2008. International Symposium on, pp. 48–57, May 2008

[2] T. Thomas and T. Bhatt, "Smart Car Parking System Using Convolutional Neural Network", In: Proc. of Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2018, No. Icirca, pp. 172–174, 2018.

[3] J. Ahmad, Z. Lewis, P. Duraisamy, and T. McDonald, "Parking Lot Monitoring using MRCNN", In: Proc. of 2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019, pp. 10–13, 2019

[4] L. Ciampi, G. Amato, F. Falchi, C. Gennaro, and F. Rabitti, "Counting Vehicles with Cameras," CEUR Workshop In: Proc. of., Vol. 2161, 2018.

[5] Parking Lot Dataset [Electronic resource]. - [Cited 12 November 2021]. - Available at: https://www.kaggle.com/blanderbuss/parking-lot-dataset

[6] C. F. Peng, J. W. Hsieh, S. W. Leu, and C. H. Chuang, "Drone-based vacant parking space detection", In: Proc. of - 32nd IEEE Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2018, vol. 2018-Janua, pp. 618–622, 2018.

[7] G. Amato, F. Carrara, F. Falchi, C. Gennaro, C. Meghini, and C. Vairo, "Deep learning for decentralized parking lot occupancy detection", Expert Syst. Appl., Vol. 72, pp. 327–334, 2017

[8] E. Tanuwijaya and C. Fatichah, "Modification of Alexnet Architecture For Detection of Car Parking Availability In Video CCTV", J. Ilmu Komput. dan Inf., Vol. 13, No. 2, pp. 47–55, 2020.

[9] Girshick, R. (2015). Fast r-cnn. Proceedings of the IEEE international conference on computer vision (pp. 1440–1448).

[10] Ren et al., 2015, Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: towards real-time object detection with region proposal networks. Advances in neural information processing systems (pp. 91–99).