Birds Eye Park

Giffy Chris
1001715630
University of Texas at Arlington
giffy.chris@mavs.uta.edu

Paul Lobo
1001686587
University of Texas at Arlington
paul.lobo@mavs.uta.edu

Abstract

In this project, we would like to reduce the time spent and gas burnt trying to find a parking space in public areas such as malls and parks. To do this our solution is a parking lot management tool which classifies occupied and vacant parking spaces using CCTV feeds which are already present in most parking lots. We approached this problem by segmenting cars and parking spaces to find the intersection of union(IOU) of the two.

1. Introduction

There has been a gradual increase in cars from the year 1990 to 2017 ranging from a little less than 200 million in 1990 going all the way to around 270 million cars in 2017 [3] . This creates a huge problem of finding parking spaces at various public parking lots and offices. With the help of CCTV cameras and Computer Vision we propose to solve this problem by identifying the vacant spaces and notifying users the closest available spots.

Many corporate and enthusiasts have worked on this problem using various methodologies, the most common one being installation of sensors at each parking space. This although being the most robust way to solve the problem also adds overhead costs such as the sensors and labor to install them [4].

Other Computer Vision methods include calculating idle time of a vehicle [5] and others use numbering systems to identify each lot, when the numbers aren't visible on the camera it assigns that space as occupied. These systems depend on either historical data or extra effort to number each parking lot.

We propose a method that uses Mask-RCNN to detect object masks of both parking spaces and cars in an image. We then calculate whether each parking space is vacant or not by using the predicted binary masks and finding intersections between each individual parking spaces and all the predicted cars. However due to the limitations of manually annotation data, we also implement a solution where Mask-RCNN is used to directly detect vacant and occupied parking spaces as this kind of annotated data is freely available.

2. Method

2.1. Dataset

We used two different datasets for our experiments. Since most parking lot cameras are static, they generally have a single orientation and images differ mainly in lighting and weather conditions. Due to this we chose the PKLot Dataset[7] which proved most relevant as it provides annotated and segmented data of occupied vs empty parking spaces from 2 parking lots containing varying weather conditions and times of day. It contains 24000+ images which we split using 70 percent for training and 30 percent for testing.

For the first experiment we had to manually annotate a dataset as we could not find a dataset that contained annotations for both cars and parking spaces. However since manual annotating a dataset is very time consuming we do this for just a few samples to show a proof of concept. We then conduct a second experiment using the entire dataset to train the Mask-RCNN algorithm to directly predict occupied vs empty parking spaces.

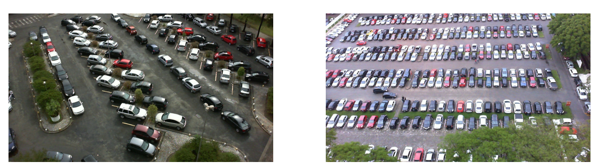

Figure 1. PKLot Dataset Samples

2.2. Implementation

To conduct our experiments we first had to manually annotate our images. The reason for this is that most parking lot datasets only contain annotations whether the parking space is empty or occupied. Since we wanted to take a novel approach we needed a model that is able to predict segments of both the parking spaces as well as the cars. To do this we used the VGG Image Annotator [8]

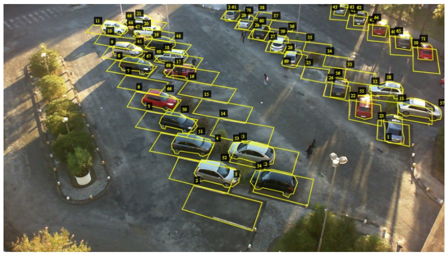

Figure 2. Manually Annotated Images

To detect parking space and cars we used the Mask-RCNN architecture since it can be trained to predict not just the bounding boxes but also the image segments which we can then use to calculate whether each space is vacant or occupied. This approach is better then using just an object detector as we get much more information using segments instead of very inaccurate bounding boxes. An example of this is that many of the parking spaces in this datasets are slanting. A rectangular bounding box includes a lot of area outside the slanting parking space.

The Mask-RCNN architecture that we used uses a Feature Pyramid Network (FPN) and a ResNet101 backbone. We initialized it using weights that were pretrained on the coco dataset. We first train just the last layer and then fine tune the entire network.

To predict segments, the Mask-RCNN architecture requires a binary mask to construct the loss function. It uses a combination of losses including the Bounding box loss, the class loss and the segment loss. We generated masks for parking spaces and cars for the first experiment and empty vs occupied parking spaces for the second.

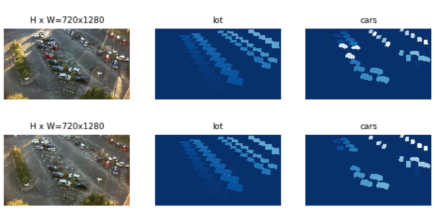

Figure 3. Experiment 1 Mask Generation

Figure 4. Experiment 2 Mask Generation

After predicting the segments we calculate the number of free parking spaces using Algorithm 1 explained in section 2.3.

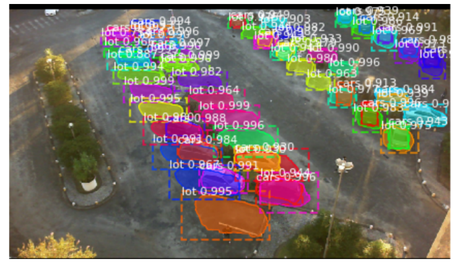

Figure 5. Predicted Segments Cars vs Lots

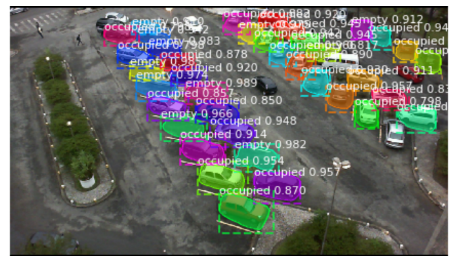

Figure 6. Predicted Segments Occupied vs Empty spaces

2.3. Modified IOU Calculation

Instead of directly predicting occupied vs empty parking spaces we proposed a novel approach that uses a combination of IOU and subtraction to check whether a parking space is vacant or occupied.

Algorithm 1 Calculate No. of Free parking spaces.

$threshold = 0.7$ //we kept it as 0.7 it can be anything from 0 to 1 as per your requirement.
$vacant = 0$
cars = array of all identified car segments in an array of shape same as the original image.
for $lot \quad in \quad identified\_lots$ **do**
  $occupied = element\_wise\_multiplication(lot, cars)$

  $IOU\_pixel = lot - occupied$
  if $sum(IOU\_pixel) > 0.7 * sum(lot)$ **then**
    $vacant+ = 1$
  end if
end for
return $vacant$

However calculating the Intersection Over Union Can be quite inefficient. For example Matching each parking space to a corresponding car would be quite inefficient. Also if we use a combined mask of cars and get the IOU (Intersection over Union), we are not able to get the count of parking spaces.

We propose a solution that uses a combined car mask and each individual parking space mask. We do this by constructing a single mask containing all the pixels where cars are present in the image. Then for each parking space mask we do a element wise multiplication with the car mask. This gives us all the pixels inside that particular lot where a car is present. We then subtract these pixels from the original lot and check the number of remaining pixels. If it is small compared to the total pixels in the lot it means that the lot is occupied, otherwise we can consider it vacant. This is shown in Algorithm 1.

3. Results

For experiment 1 we got the following results. The image on the left contains the combined predicted car masks. We set all the pixels that contain any car to 1. We then use algorithm 1 to decide for each lot whether a car is present in it or not. As we can see that the results are pretty accurate although it was run on a very small dataset.

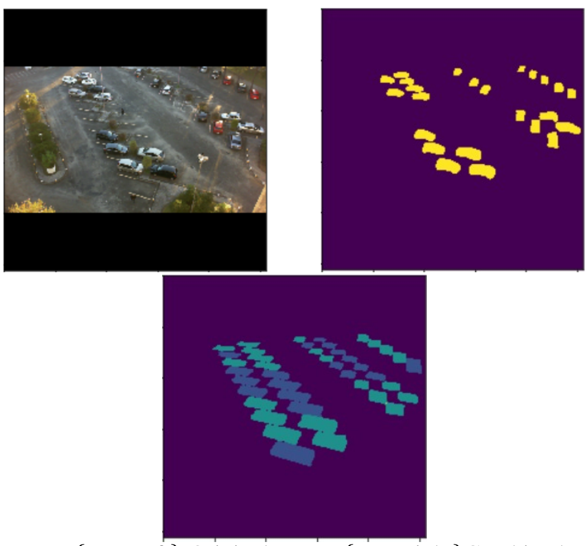

Figure 7. {Top Left} Original Image {Top Right}Combined predicted Car Masks {Bottom} Light Blue: Occupied, Dark Blue: Vacant

For the second experiment on the entire PKLot dataset of 24K images, we trained the Mask RCNN model to predicted occupied vs vacant spaces and we got a Map of 90.17%. This shows us that using Mask RCNN we can get an even better accuracy by using the first approach with a much larger dataset.

4. Conclusion

This project helps show a proof of concept which gives robust predictions of parking spaces without the need of auxiliary data such as markers or sensors in the parking lot.

We propose a novel way to check whether a car is present in a lot by creating an efficient modification to calculate the intersection over union between a combined mask of cars and each individual parking space mask. Although the data is over-fit on the manually annotated dataset we show that with the correct annotations, the Mask RCNN classifier is able to correctly predict whether a parking space is vacant or occupied. We can extend this by increasing the number of images in the manually annotated dataset which will help to increase its performance and generalizing power.

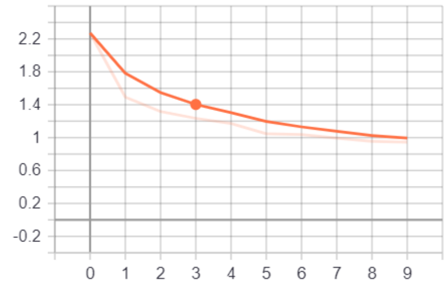

Figure 8. Training Loss

5. Contribution

During the course of the project every task was performed together. Tasks performed were as follows:

- Ideating project topic
- Finding most relevant architecture for training
- Finding an appropriate dataset to work on
- Implementing initial training of the images.
- Re-annotating images to fit our specific needs
- Ideating on the formula to calculate IOU
- Writing the report.

References

[1] Mask R-CNN by Kaiming HeGeorgia, Gkioxari Piotr, Dollar, Ross Girshick paper-link

[2] Mask RCNN implementation by matterport github

[3] Vehicle Population Data Statista.com

[4] Boshch parking sensors sensor

[5] Parking using idle time Medium

[6] Parking using markers youtube

[7] Parking lot Dataset dataset

[8] VGG Image Annotator tool