Detection Of Vacant Parking Spaces Through The Use Of Convolutional Neural Networks

Dr. Christoforos Christoforou
St. John's University
Division of Math, Science and Computer Science
christoc@stjohns.edu

Christopher Singh
St. John's University
Division of Math, Science and Computer Science
christopher.singh15@stjohns.edu

Abstract

This paper focuses on the application of computer vision and convolutional neural network techniques in the automotive industry to reduce the amount of time required to locate a vacant parking spot and to reduce driving time. The main motivation for a vacant parking spot detector is such that today's drivers are facing major difficulties in finding available spots in largely populated cities. This often times leads to increased congestion and frustration for the driver because they are forced to continue their search for a parking spot. Additionally, the distraction that is generated by excessively searching for a vacant parking spot can lead to an increase in the amount of accidents because motorists shift their attention from the road and focus on looking for a spot. This neglect for other drivers and pedestrians can lead to fatal situations. The need for a parking spot detector would reduce the amount of accidents, decrease driver tension and improve motorist moral. A parking space detector would prevent motorists from driving aimlessly around a parking lot without the guarantee that they would find a spot. Today's malls and security cameras are currently able to detect the amount of free parking spots in a given lot or garage level and then relay the information count to the driver. This advises the driver whether or not they should enter a given parking lot. This would be useful in large scale parking lots such as sports stadiums or an amusement park. Our approach is able to solve this issue and provide the driver with useful information.

Introduction

One of the main problems that motorists face today is finding a parking space in either a parking lot or on the street, a majority of the time leading to traffic congestion and driver frustration. This has the possibility to turn into a dangerous situation in which people can be injured. A parking detector system can alleviate this condition by providing information to the driver regarding the availability of a parking space.

The current state of parking space detection systems is such that the existing systems implement sensors throughout the entire parking lot that supports both ultrasound and infra-red methodologies for determine vacancy. The main drawback with this setup is that these sensors require high costs for both the installation process and maintenance of the sensor especially for parking lots that house a large volume of spots. Due to this high cost, the demand for a computer vision-based system has increased because it is highly scalable and affordable to implement. It can serve a wide range of parking lot configurations without the need of any sensors.

On this basis, the purpose of our work is to introduce and outline a novel system for documenting unoccupied parking spaces. It can provide a solution that accurately determines the availability of a parking spot in a given lot. Our imaging model dataset is derived from a camera-based dataset in which images are taken throughout the day and at various time of day. This is done so to take into account environmental factors such as sunlight and weather. Data normalization techniques are implemented to ensure a standard image dataset for training and testing. Then two well-known pre-trained convolutional neural networks are implemented: namely ResNet50 and MobileNet. The results of both models are then analyzed to determine the more accurate model. In doing so, the following applications can be realized from this implementation:

- The training and fitting of the model through both ResNet50 and MobileNet leads to an improvement in accuracy over that of a self-defined network architecture. This is due to the multiple layers in the network.
- The performance analysis for both models in terms of accuracy and its loss function are very close and provide promising results.
- After n-iterations of epochs, both the accuracy and the loss function tend to stabilize. If we were to train more, then we would run into the situation of over-fitting the model.

In the following sections, we first review previously related work on the parking domain. The methodology and experimentation setup sections document the data preprocessing tasks completed along with training and testing both models. The results section provides insight on the model performance and how well it was able to predict new images. Lastly, potential future application and final conclusions are discussed.

Related Work

Previous attempts to create a detector have successfully resulted in a usable models that are able to determine vacancy. However, the previous implementations had specific limitations and restrictions. Moreover, weather seemed to played a main factor in the accuracy of previous models as well.

The first attempt made by (Chunhe and Jilin 2004; Wolff et al. 2006;Schmid et al. 2011) relied heavily on infrastructure and ultrasonic sensors to capture data. This was not an ideal solution because it was very costly for large scale parking lots. The implementation found in (Lin, Chen, and Liu 2006; Wu et al. 2007) focused on using imagery classification data from static cameras to determine the empty space in an image. The main issue here was that their model faltered during different lighting conditions. This issue was when resolved by (Ichihashi et al. 2009) where they took into account different lighting conditions by introducing data from various parts of the day. They implemented a Principal Component Analysis algorithm which was built on K-Means to classify single parking spaces.

Furthermore, (Huang and Wang, 2010) created a Bayesian model to extract the region of interest for each parking spot and then classified the availability. Another proposed solution was made by (Lin et al., 2006) in which a background subtraction algorithm was implemented and detected objects in the foreground. This was not ideal because their model faced challenges when classifying parking spots that had no background object. The latest attempt was made by (Hsieh, Lin, and Hsu 2017) in which they used drones to capture car data and relay the available vacancies. Even with all of these previous implementations, the ability to detect parking spots is still an open issue which needs some fine-tuning.

Methodology

The parking spot detector was implemented based on an open source imaging dataset which was taken by surveillance cameras at various parts of the day and in different weather conditions. The dataset captures the vacancy of parking spots in which the status is classified as either available or full. This was a preliminary labelled dataset in which the images contained both vacant and full parking structures. The first steps were to apply basic data pre-processing steps to the images. This included resizing the image dataset and then normalizing the output. All of the images were first cropped and then resized in order to provide uniformity. The image vectors were then normalized to the preferred range of neural network models. The final pre-processing step was to implement a one-hot encoder in order to convert the labels into a numeric format suitable for supervised machine learning.

At this point, the dataset was split into both training and testing sets with a ratio of 70:30 respectively. We assumed a random state of 42 for this process. Both ResNet50 and MobileNet models are configured such that the inputs for fitting require training and testing datasets.

Experimentation Setup

The convolutional neural network setup for both MobileNet and ResNet50 implementations used in this research experiment are designed based on the existing layers that are already a part of the pre-defined model. We expanded on the implementation of both models by adding additional network layers in an attempt to improve the classification accuracy. We were able to mimic the same image shape, kernel size, and activation type in order to maintain consistency across both models when reporting the performance. It was critically important that both models ended up with a similar amount of trainable parameters.

The main purpose of implementing a MobileNet model was to increase the precision of the network while decreasing the amount of parameters used. This network consists of 28 Convolutional layers followed by a Fully Connected layer and lastly a SoftMax layer. A Batch Normalization layer is added after each Convolutional layer because it re-scales and re-centers the image. Additional Dense layers were added on to the base model in order to match the number of parameters in ResNet50. The driving force behind MobileNet's powerful architecture is that it implements depth-wise convolution followed by a point-wise convolution. The depth-wise convolution layer uses the actual number of channels as a kernel size while the point-wise convolution always uses a 1 x 1 matrix to change the dimension.

The architecture of the ResNet50 model is separated into 4 main stages of convolution with 50 total layers in the network. The network begins with performing a convolutional transformation followed by a Max Pooling layer before entering the Residual Block layers. Each stage in the network consists of 3 residual blocks which house 3 Convolutional layers. At the end of each block, a Max Pooling layer is used to calculate the largest value from each layer in the feature map. It is worth noting that as the model transitions from stage to stage, the input shape is reduced by half each time. Aside from the base ResNet model, we implemented additional Dense, Dropout and Average Pooling layers to improve the out-of-box accuracy.

	MobileNet	ResNet50
Layers	36	60
Trainable Parameters	25,936,343	25,936,738
Batch Size	32	32
Epochs	100	100

Table 1: Summary Of Both Models

Results

Both the MobileNet and ResNet50 models were analyzed using cross entropy loss as the baseline. The Adam optimizer was implemented to update the network weights after each iteration. We observed a large difference in the learning process in which the ResNet50 model took approximately 9 seconds per epoch while MobileNet took on average 95 seconds. The model performances were evaluated based on their respective precision, recall and f-measure scores.

The performance of both models were examined using the cross-entropy loss function to determine which category a given input image should be classified as. The cross-entropy was used to calculate the difference in the probability distribution between the categorical labels and is defined as:

$$Loss(p, q) = -\sum_{x \in \mathcal{X}}^{N} p(x)\, \log q(x)$$

At the end of 100 epochs, we saw a both models decrease in terms of their loss function. ResNet50 experienced a more dramatic decrease to practically no loss while MobileNet's loss was more stable. The respective graphs are displayed below.

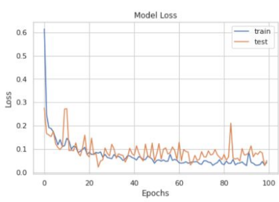

Figure 1: MobileNet Loss

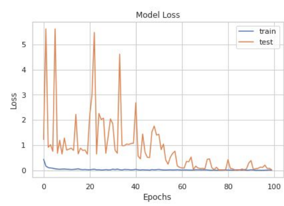

Figure 2: ResNet50 Loss

The table below summarizes the final loss value for the models. Both MobileNet and ResNet50 are very competitive with respect to classification accuracy.

	MobileNet	ResNet50
Accuracy	98.17%	99.59%
Loss	4.79%	2.24%

Table 2: Cross Entropy Loss

Furthermore, we evaluated both models in terms of their respective Precision, Recall and F-measure scores with values obtained from the confusion matrix. The precision represents the measure of correctly predicted observations divided by the total amount of predictions and is defined as:

$$Precision = \frac{TP}{(TP + FP)}$$

The recall that was computed represents the measure of correct observations divided by the total amount of observations and is defined as:

$$Recall = \frac{TP}{(TP + FN)}$$

The F-measure score provided insight as to the accuracy of the model by applying the mean to both precision and recall methods and is implemented as:

$$F - Measure = 2 * \frac{precision * recall}{(precision + recall)}$$

In the above formulas, TP is defined as true positives, FP is defined as false positives and FN is defined as false negatives all of which are observations classified by the model.

	Precision	Recall	F-Measure
ResNet50 Full	1.00	1.00	1.00
ResNet50 Free	1.00	1.00	1.00
MobileNet Full	0.04	0.07	0.05
MobileNet Free	0.06	0.03	0.04

Table 3: Evaluation Metrics Of MobileNet and ResNet50

The final evaluations clearly show that although both MobileNet and ResNet50 were extremely accurate, ResNet50 was able to outperform MobileNet in terms of precision, recall and f-measure. ResNet50 performed better in classifying both vacant and occupied parking space labels with nearly a perfect model. MobileNet seems to falter when handling the same dataset even though it is approximately 1% less accurate than ResNet. The main inference that can be drawn from these initial results is that MobileNet was not capable of handling an imbalanced dataset whereas ResNet50 excelled. Below are outputs of images that the models were able to correctly predict:

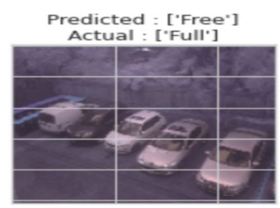

Figure 3: MobileNet Prediction

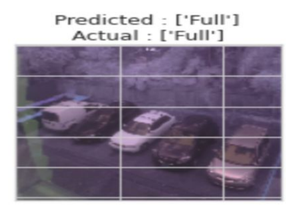

Figure 4: ResNet50 Prediction

Future Work

This novel parking space detector has plenty of room for future improvements and implements. One future application for both models would be to implement them on a large scale parking structure instead of a minimal lot. Another area of interest would be the usage of the system for indoor or underground parking lots where the time of day is essentially constant throughout the training process. A further improvement to the detection system would be to increase the amount of features used to determine the occupancy of a parking spot. This can be achieved through the use of a backpropogation neural network and would require similar training and testing labels. The use of backpropogation would allow the network to compute a gradient descent value in terms of model weight.

The biggest improvement that can be implemented would be the ability to take into account whether or not a parking space is free without the aid of parking lines. Ideally, the model would be able to predict on parking spaces on regular streets which do not have visible parking lines. This would be demonstrated through the use of a Hough transformation in which image features can be extracted. The features would then be analyzed to find line patterns of similar length and angle.

Minor improvements can be made to both CNN models in order to reduce the training time and the total number of parameters. The MobileNet model can be improved by introducing additional Max Pooling layers after each Convolutional layer. ResNet50 can be improved by adding additional residual blocks and a SoftMax layer at the end of the model.

Final Conclusion

In this paper, we presented a novel computer vision solution to finding an available parking space through the use of convolutional neural networks. Both MobileNet and ResNet50 models showed promising results when dealing with labelled data. Our results show that ResNet50 is more capable of predicting than that of MobileNet. This can be verified with a very low loss percentage. The growing need for a parking spot detector is critical because it reduces the driver's responsibilities which leads to a reduction in the amount of accidents. Our models show that this issue can be partially resolved by providing a solution which applies to only parking lots. The future research direction that we are pursing involves the recognition of vacant parking spaces on city streets which are bi-directional.

References

Almeida, P. R.; S.Oliveira, L.; Jr., A. S.; Jr., E. J. S.; andL.Koerich, A. 2015. Pklot - a robust dataset for parking lot classification. *Expert Systems with Applications*.

Chunhe, Y., and Jilin, L. 2004. A type of sensor to detect occupancy of vehicle berth in carpark. In *Proc. of the International Conference on Signal Processing (ICSP)*.

Wolff, J.; Heuer, T.; Gao, H.; Weinmann, M.; Voit, S.; and Hartmann, U. 2006. Parking monitor system based on magnetic field sensor. In *Proc. of the IEEE Intelligent Transportation Systems Conference*.

Schmid, M. R.; Ates, S.; Dickmann, J.; von Hundelshausen,F.; and Wuensche, H.-J. 2011. Parking space detection with hierarchical dynamic occupancy grids. In *Proc. of the IEEE Intelligent Vehicles Symposium (IV)*.

Lin, S.-F.; Chen, Y.-Y.; and Liu, S.-C. 2006. A vision-based parking lot management system. *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*.

Wu, Q.; Huang, C.; Wang, S.-Y.; Chiu, W.-C.; and Chen,T. 2007. Robust parking space detection considering interspace correlation. In . *Proc. of the IEEE International Conference on Multimedia and Expo*.

Sastre, R. J. L.; Jimenez, P. G.; Acevedo, F. J.; and Bascon,S. M. 2007. Computer algebra algorithms applied to com-puter vision in a parking management system. In *Proc. of the IEEE International Symposium on Industrial Electronics*.

Lee, C.H., Wen, M.G., Han, C.C., Kou, D.C.: An automatic monitoring approach for unsupervised parking lots in outdoors. In: *39th Annual International Carnahan Conference*.

Fabian, T. (2008). An algorithm for parking lot occupation detection. In *Computer Information Systems and Industrial Management Applications, 2008. CISIM '08.7th*.

Schanz, A., Spieker, A., Kuhnert, D.: Autonomous parking in subterranean garages: a look at the position estimation. In: *IEEE Intelligent Vehicle Symposium*.

M. Tschentscher, C. Koch, M. Knig, J. Salmen and M. Schlipsing, "Scalable real-time parking lot classification: An evaluation of image features and supervised learning algorithms", *International Joint Conference on Neural Networks (IJCNN)*.

Amato, G., Carrara, F., Falchi, F., Gennaro, C., Vairo, C.: Car parking occupancy detection using smart camera networks and deep learning. In: *2016 IEEE Symposium on Computers and Communication (ISCC)*.

Ichihashi, H.; Notsu, A.; Honda, K.; Katada, T.; and Fujiyoshi, M. 2009. Vacant parking space detector for outdoor parking lot by using surveillance camera and fcm classifier.In *Proc. of the IEEE International Conference on Fuzzy Systems*.

oresti, G.L., Micheloni, C., Snidaro, L.: Event classification for automatic visual-based surveillance of parking lots. In: *17th International Conference on Pattern Recognition, vol. 3*.