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

Christopher Singh

St. John's University

Division of Computer Science

christopher.singh15@stjohns.edu

Christoforos Christoforou

St. John's University

Division of Computer Science

christoc@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 time leads to increased congestion and frustration for the driver because they are forced to continue their search for a parking spot. Our approach is able to solve this issue and provide the driver with useful information through the use of transfer learning methodologies. The main contribution of this paper is to examine and improve on previously implemented transfer learning methods in order to better increase the detection accuracy. This paper differs from previous attempts such that it considers all environmental factors such as weather and time of day. Other models are not able to handle these conditions with a high accuracy and subsequently falter. When compared to previous attempts, our implementation focuses solely on the reliance of transfer learning. The results indicate that our model is capable of identifying vacant parking spaces under all conditions with competitive accuracies. The proposed model is able to surpass the accuracy of the latest attempt at solving this issue.

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. 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 (Chunhe, Y 2004). The need for a parking spot detector would directly reduce the amount of accidents, decrease driver tension and improve motorist moral.

The current state of parking space detection systems is such that the existing systems implement sensors and convolutional neural networks (CNN) throughout the entire parking lot and support both ultrasound and infrared methodologies for determining vacancy (Amato, G 2020). The main drawback with this setup is that the current CNN model uses minimal hidden layers with a high number of neurons per layer which results in longer wait times. This paper aims to take advantage of the hidden layers by reducing the amount of neurons needed. Aside from CNNs, the attempt made by (Wolff, J 2009) relied on parking sensors which require high costs for both the installation and maintenance process. Due to this high cost, the demand for a computer vision-based system has increased because it is highly scalable and affordable to implement (Almeida, P. 2015).

On this basis, the purpose of our work is to introduce transfer learning without the help of sensors to document unoccupied parking spaces especially at night and in precipitation. Previous attempts made by (Amato, G 2020) and (Oresti, GI) ignore weather conditions and is not an ideal solution. Both (Lee, CH 2013) and (Sastre, RJ 2007) fail to create a dataset which captures dark night images in reduced lighting. This ultimately leads to a reduction in accuracy for parking spots at night. Our imaging model dataset is derived from a camera-based dataset in which images are taken throughout the day and at various times of day to take into account environmental factors. Then two well-known pre-trained convolutional neural networks are implemented: namely ResNet50 and MobileNet. In doing so, the following applications can be realized from this implementation:

- Both ResNet50 and MobileNet models provide a simple and fast solution that leads to an improvement in accuracy over that of a self-defined network architecture.

- The performance analysis for both models in terms of accuracy and its loss function are very close and provide encouraging results.

- After 100 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 space detector domain. The methodology and experimentation setup sections document the data pre-processing tasks completed along with the training and testing of both models. The results section provides insight on the model performance and how well it was able to predict vacant parking spots. Additionally, an algorithmic comparison is performed and reported against the latest model with the same dataset. 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 failed 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 intervals 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. A further experiment was made by (Hsieh, Lin, and Hsu 2017) in which they used drones to capture car data and relay the available vacancies. Their proposal to identify parking spots were accurate but the main drawback was that the drone needed to be operated by a physical person. The other disadvantage was that their model did not take into account weather factors such as rain or snow due to safety issues with drones.

The latest state of the art attempt was made by (Amato, Carrara. 2020) in which they used a CNN along with image segmentation to identify empty spaces in which a car can possibly fit. However, their model does not take into account non-ideal situations such as night-time where there is reduced light. Even with all of these previous implementations, the ability to detect parking spots is still an open issue which needs some fine-tuning. This paper aims to solve this problem of detecting parking spaces in all conditions.

Methodology

The parking spot detector was implemented based on an open source public 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. This dataset contains approximately 3300 images and showcases a wide variety of parking lots with multiple angles. The image resolution was 1296 x 972 pixels. The first steps were to apply basic data pre-processing steps to the images. This included resizing the image dataset and then normal-

izing 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. Also appropriate callbacks to both networks were configured. Both ResNet50 and MobileNet models are composed 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. This transfer learning approach was taken since both models were successfully capable of object recognition. The main motivation for using MobileNet is that it provides both a reduced network size and number of parameters which results in a low-latency convolution neural network. The ResNet model was chosen because it reduced the effect of vanishing gradients while accelerating the time required to train the model.

The MobileNet model was able to increase the precision of the network while decreasing the amount of parameters used. This network consisted 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. 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.

ResNet's biggest advantage is that it avoids inherent and negative outputs while increasing the network depth. ResNet uses residual connections to reduce the vanishing gradient problem. 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. Both models are summarized below:

	MobileNet	ResNet50
Layers	31	56
Trainable Parameters	3,469,890	32,187,138
Batch Size	32	32
Epochs	100	100

Table 1: Network Architecture Of Both Models

Results

Both the MobileNet and ResNet50 models were analyzed using cross entropy loss. We observed a large difference in the learning process where the ResNet50 model took 9 seconds per epoch while MobileNet took 95 seconds. The reason that MobileNet network took a longer time even though ResNet50 is a larger network is because it uses both depthwise and point-wise convolutional layers to extract features.

The performance of both models were examined using a cross-entropy loss function. This was used to calculate the difference in the probability distribution 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 both models decrease in terms of their loss function. ResNet50 experienced a dramatic decrease while MobileNet's loss was more stable. The respective graphs are displayed below:
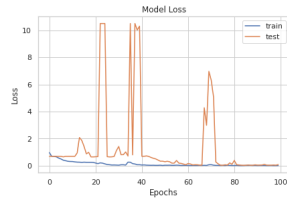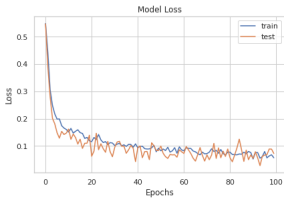

Figure 1: MobileNet Loss   Figure 2: ResNet50 Loss

In order to thoroughly examine the effectiveness of our model, we compared our results with the latest attempt made by (Amato 2020). We re-sized and trained our images onto the same LeNet-5 architecture that was used by Amato. The only change that was made was resizing the image dataset since LeNet-5 requires an image size of 32x32. In this experiment, our model was superior in both accuracy and loss.

The table below summarizes the final accuracies and loss values for all of the models. Both MobileNet and ResNet50 outperformed LeNet-5 and are very competitive with respect to classification accuracy.

	MobileNet	ResNet50	LeNet-5
Accuracy	0.9602	0.9806	0.8667
Loss	0.0770	0.0843	0.7861

Table 2: Cross Entropy Loss Function Results

Furthermore, we evaluated both of our models in terms of their respective precision, recall and f-measure scores with values obtained from the confusion matrix. These performance metrics are summarized in the table below:

	Precision	Recall	F-Measure
ResNet50 Lot Full	1.00	1.00	1.00
ResNet50 Vacancy	1.00	1.00	1.00
MobileNet Lot Full	0.04	0.08	0.06
MobileNet Vacancy	0.09	0.05	0.06

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 dwindle when handling the same dataset even though it is approximately 2% less accurate than ResNet. Below are outputs of images that the models were able to predict:
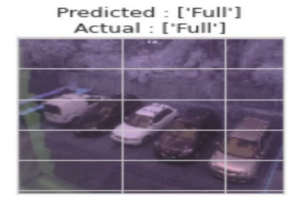

Figure 3: MobileNet Prediction   Figure 4: ResNet50 Prediction

A further analysis was made to determine the effectiveness of our model solution. We examined the p-value of the results with respect to the null hypothesis which claims that convolutional neural networks cannot precisely determine the vacancy of a parking spot. In this analysis, we assumed the level of significance to be less than or equal to 0.05. Any p-value which is less than this threshold should be accepted and the null hypothesis is ultimately rejected.

This analysis was performed between both MobileNet and ResNet50 in order to determine the better model. The results of the p-value calculations show that both models are statistically significant. The values are documented below:

	MobileNet	ResNet50
Significance Level	0.05	0.05
p-value	0.024	0.018

Table 4: p-value results of both MobileNet and ResNet50

Since the p-value is less than the level of significance, the null hypothesis can be rejected and the alternative hypothesis can be accepted which in turn means that CNNs are an effective solution. Moreover, the ResNet model outperformed the MobileNet model since it had a lower p-value. This difference in p-value from both models are significant in proving that ResNet is the superior solution since both models produced almost identical accuracies.

Discussion

The latest state of the art model proposed by Amato and Carrara (2020) relied on using transfer learning as well. They implemented the LeNet-5 architecture which is a fully connected CNN with minimal trainable parameters. Their LeNet-5 model focused on using a hyperbolic tangent formula instead of relu as their choice of an activation function. As a result, they were able to create a model with an accuracy of 0.983 based on their own labelled dataset. The

model that we were able to build closely mirrors the accuracy of the LeNet-5 model. In order to further analyze our results, we were able to recreate their experiment and perform an algorithmic comparison with our dataset since their model was open source to the public.

When running the LeNet-5 model against our resized dataset, the accuracy of detecting vacant parking spots is reduced to 0.867. This reduction in accuracy is significant because it was trained on a smaller image size since the requirement of LeNet-5 is rather small. A possible explanation for such a dramatic change in accuracy is that their model does not take into account night images. Both MobileNet and ResNet50 models were able to outperform the LeNet-5 architecture and produce a highly effective solution to determine the availability of a given parking space.

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. 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 back-propagation neural network and would require similar training and testing labels. The use of back-propagation 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.

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 positive 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 as well as a better p-value score. Moreover, our model was able to provide more accurate results than the LeNet-5 model which is currently the latest developed model. It can be concluded that transfer learning should be the preferred method of research over that of user-defined networks.

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. In final analysis, this is how the issue of finding a parking space can be effectively solved.

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 inter-space 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: *2020 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 lots. In: *17th International Conference on Pattern Recognition, vol. 3*.