

# Early Detection of Vacant Parking Spaces Using Dashcam Videos

**Ming-Che Wu, Mei-Chen Yeh**

Department of Computer Science and Information Engineering

National Taiwan Normal University, Taipei, Taiwan

Email: myeh@csie.ntnu.edu.tw

## Abstract

A major problem in metropolitan areas is finding parking spaces. Existing parking guidance systems often adopt fixed sensors or cameras that cannot provide information from the driver's point of view. Motivated by the advent of dashboard cameras (dashcams), we develop neural-network-based methods for detecting vacant parking spaces in videos recorded by a dashcam. Detecting vacant parking spaces in dashcam videos enables early detection of spaces. Different from conventional object detection methods, we leverage the monotonicity of the detection confidence with respect to the distance away of the approaching target parking space and propose a new loss function, which can not only yield improved detection results but also enable early detection. To evaluate our detection method, we create a new large dataset containing 5,800 dashcam videos captured from 22 indoor and outdoor parking lots. To the best of our knowledge, this is the first and largest driver's view video dataset that supports parking space detection and provides parking space occupancy annotations.

## Introduction

Finding a vacant parking space in a congested area or a large parking lot is time consuming and frustrating for drivers, especially at peak hours or when the target parking lot is crowded or almost full. Furthermore, the distraction caused by searching for a parking space is one of the main causes of parking lot accidents because drivers are focused on finding a vacant spot and often neglect to watch for other drivers and pedestrians who are on the move. This situation is dangerous when the parking lot is congested with vehicular traffic and pedestrians. An intelligent parking guidance system can mitigate much of this hassle by promptly providing accurate information concerning car park space availability to drivers.

Current parking guidance systems determine the availability of parking spaces using the sensors installed throughout an entire parking lot. However, deploying sensors in a large parking lot can be costly and is limited in indoor parking lots. Alternatively, some studies (Huang and Wang 2010; Ichihashi et al. 2009; Almeida et al. 2015; Amato et al. 2017) have explored the application of image analysis techniques for parking space detection by using images taken

from a fixed camera positioned at the top of a parking area. A vision-based system has advantages in terms of coverage, cost, and versatility; and it can provide many value-added services such as parking space guidance and video surveillance (Enríquez et al. 2017). This would be helpful when surveillance cameras are available; however, many parking areas do not have such a camera monitoring system.

For this study, we are interested in automatically detecting vacant parking spaces by analyzing the content of videos captured during the driving process. We propose to take advantage of the cheap and widely available dashcams to collect data at scale and train a robust detection model. Prior studies concerning parking lot monitoring have often assumed that the locations of the parking spots are already known and the cameras are fixed, making conventional vision-based parking space detection methods not directly applicable for unconstrained dashcam videos.

Detecting a vacant parking space as early as possible is a key goal for drivers. This practical necessity requires vacant parking spaces to be detected before they are completely seen. Most existing methods (Viola and Jones 2004; Girshick et al. 2014) for object detection have been designed for offline processing and are thus limited in their ability to signal the observation of objects of interest as early as possible. For vacant parking space detection, it is necessary to predict the availability of a parking space when it is partially observed or even completely occluded, a fact that has been ignored in the training processes of existing object detectors.

Learning to detect vacant parking spaces in dashcam videos is a practical but challenging task because it must handle typical visual recognition problems such as light conditions, viewpoint changes, the presence of shadows, and occlusions. Early detection of vacant parking spaces is even more challenging because a vacant parking space is often less visible when it is relatively distant from the driver. We follow the trend of applying deep learning techniques to address this problem. In particular, we simulate the sequential frame-by-frame data arrival for a training time series and create a vacant parking space detector that can correctly classify partially observed instances. We train a detector to predict the availability of parking spaces located "ahead" of a driver even when the target parking space has not yet been observed. This requires monotonicity of the detection reliability with respect to the distance from the driver to the park-

ing space—the detection score should increase if the car is approaching an empty parking spot (and vice versa). We introduce a new loss function specifically designed for early detection and determine that it is superior to conventional loss functions such as cross-entropy and mean squared error with respect to the timeliness and accuracy of the detection.

For training and evaluating the detection methods, we introduce a new dashcam video dataset containing 5,800 short dashcam videos with high video quality (1080p resolution) and parking space occupancy annotations. The new dataset provides the first and the largest driver’s view videos, capturing diverse and unconstrained scenes of 22 parking lots. In summary, this study makes the following contributions:

- We demonstrate that early detection of vacant parking spaces can be achieved using dashcam videos that implement deep learning.
- We introduce the first dashcam video dataset for studying early detection of parking spaces (available online).
- We propose a new loss function for early detection and demonstrate that constraining the subsequent detection scores can considerably improve the detection results.
- We demonstrate the potential for using a vast amount of dashcam videos to improve intelligent parking guidance.

In the following sections, we first review related studies. Section 3 presents the new video dataset, and the proposed method is presented in Section 4. Finally, experiments and results are discussed in Section 5.

## Related Work

### Parking Space Detection

Pioneering works focus on infrastructure, where sensors (infrared, ultrasonic, radar etc.) are installed to monitor the parking places (Chunhe and Jilin 2004; Wolff et al. 2006; Schmid et al. 2011). The main drawback of sensor based approaches is the cost for developing and maintaining the system because a large amount of sensors units are required to cover the entire parking lot.

With the advancement of computer vision techniques, there has been a growing interest in developing smart camera solutions. Vision based systems can be grouped according to the imagery data used to detect empty spaces. Most existing works use static cameras. Early studies used different types of features and classification algorithms to model empty spaces (Lin, Chen, and Liu 2006; Wu et al. 2007; Huang and Wang 2010). Under different lighting conditions texture descriptors are more robust than color features (Sastre et al. 2007; Almeida et al. 2015). Ichihashi et al. in particular dealt with various luminance conditions using images captured during different hours (Ichihashi et al. 2009). Recently del Postigo et al. used background subtraction to detect and track vehicles and created a transience map to detect the parking and leaving of vehicles (del Postigo, Torres, and Menéndez 2015). All of these methods assume that the positions of parking spots are known. Given a parking region, they solve the classification problem whether it contains a vehicle.

A few works use aerial images. Peng et al. applied the RANSAC scheme to estimate the homograph relation between an image and the reference parking space and extracted three types of features to determine the occupancy condition (Peng et al. 2018). A drone-based method was proposed in (Hsieh, Lin, and Hsu 2017), which detected and counted cars in drone videos.

Despite the importance of these contributions, parking space detection using visual information is still an open problem. Most techniques are tailored and fine-tuned to specific contexts and scenarios, and thus the adaptation of one solution to a different place is not easy. Moreover, none of these works use imagery data from the driver’s point of view. To the best of our knowledge, this is the first work that uses dashcam videos for detecting parking spaces.

### Dashcam Video Dataset

Recently dashcam videos have been used for studying visual recognition and localization tasks. For example, KITTI (Geiger et al. 2013) is a well-established real-world benchmark dataset for studying vision-based self-driving tasks, including stereo, optical flow, visual odometry, 3D object detection and 3D tracking. KITTI contains videos captured by driving around the city of Karlsruhe, in rural areas and on highways. The cityscapes dataset (Cordts et al. 2016) focuses on semantic understanding of urban street scenes. It consists of videos captured in 50 cities. Among the video data 5,000 images have dense pixel annotations, and 20,000 images have coarse polygonal annotations.

Despite that these datasets offer voluminous dashcam video data recorded under real world driving situations, they are mostly irrelevant to the parking scenario. The proposed dataset is one of the first dashcam video datasets for studying parking space detection.

### Dataset

A dashcam is a low-cost vision sensor, usually mounted inside a vehicle to record street-level visual observation from the driver’s point of view. We collected the dataset by taking high-quality dashcam videos at 22 different places (10 indoor and 12 outdoor) during 5 months, capturing parallel, perpendicular, angled and on-street parking spaces. The video contents are unscripted and diverse in various scenes, containing different lighting conditions, the presence of shadows, and occlusions.

We manually annotated the temporal locations of vacant parking spaces in each video. Consider that an effective detector should alert the driver before the target parking space is reached, we create 5,800 clips from the annotated videos, where each clip consists of 150 frames (5 seconds, 30 fps). These clips consist of 2,900 positive clips (1,450 indoor and 1,450 outdoor) containing the moment of reaching a vacant parking spot in the last second, and 2,900 negative clips (1,450 indoor and 1,450 outdoor) containing no empty parking spaces. More specifically, a clip is labeled “positive” if the fifth second of this clip captures a vacant parking spot aside the driver, regardless of the occupancy of the other parking spaces in the first 120 frames. A clip is labeled “neg-

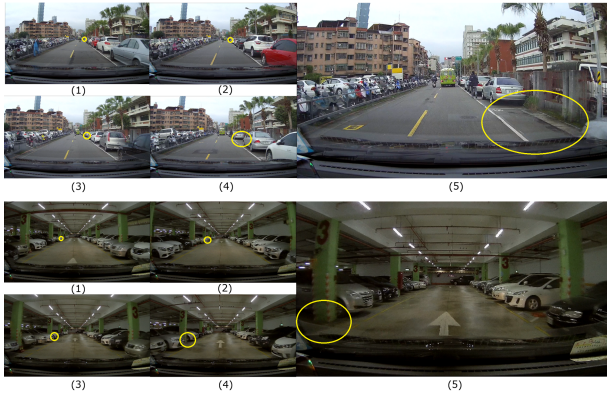


Figure 1: Two positive samples in the dashcam video dataset. The number below each frame denotes the timestamp. The frame rate is 1 fps for the visualization purpose. The vacant parking spaces are highlighted in yellow.

ative” if it does not capture any vacant parking spaces. Figure 1 shows two positive samples in this dataset, in which the number below each frame denotes the timestamp (1 fps for the visualization purpose). The vacant parking spaces are highlighted in yellow.

Compared with existing parking space datasets, our dataset has the following characteristics:

- It contains voluminous dashcam video data recorded under real world driving situations at 22 different parking places.
- It allows future evaluation of systems for both indoor and outdoor scenarios.
- It provides visual information from the driver’s point of view.

The dataset is available at <https://bit.ly/2PmXfv7>.

## Approach

In this section, we formally define the early detection problem and then present the proposed method. The method is built upon convolutional neural networks (CNN) (LeCun et al. 1990) and recurrent neural networks (RNN) (Giles, Kuhn, and Williams 1994).

### Problem Definition

The objective of this study is to use observations in a dashcam video to predict the availability of a parking space ahead. This application is fundamentally an image understanding task: to detect objects from images, where object categories are “car” and “background.” However, we are interested in detecting objects “ahead,” rather than detecting them in the given frame. This is different to the definition of the general visual recognition problem, whose goal is to detect objects of interest in the given image.

We mimic the human capability of finding a parking space before it is completely seen, and train a detector to learn the transition of “seeing” the parking spot. Therefore, early detection is formulated as a binary classification problem:

Given a one-second video clip, we aim at predicting whether a parking space is available ahead.

Let  $D = (X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)$  be the set of training videos, where  $X_i = [x_1, x_2, \dots, x_m]$  represents a sequence of  $m$  1-s clip ( $m = 5$  in this study) and  $y_i$  denotes the label of  $X_i$ . Specifically,  $y_i$  is 1 if  $X_i$  is positive, and 0 otherwise. Recall that  $X_i$  is positive if the  $m$ -th second of this clip captures a vacant parking spot. Therefore, each  $x_i$  ( $i = 1, \dots, m$ ) in a positive video are considered positive because a parking space will soon appear in this video. A detector—characterized by  $\theta$  (will be detailed later in this section)—receives a 1-s clip  $x_i$  and returns the detection score  $f(x_i, \theta)$ . We trained  $\theta$  from  $D$  by minimizing the detection loss. While in testing, the model is given an observation of  $x_t$  one at a time following the order of the video frames. The goal is to detect a vacant parking space as early as possible given the observations  $(x_1, x_2, \dots, x_t) | t < m$  before reaching the vacant packing spot at frame  $m$ . This formulation ensures the detection of parking spaces at the shortest possible time.

### Detection Loss

Intuitively, the detection should be more reliable when the driver is approaching an empty parking spot. The penalty of a failed detection at a frame very close to the vacant parking space should be higher than the penalty at a frame far away from it. The first term of the loss function we use on each 5-s training video clip  $X_i$  is defined as:

$$L_{conf}(X_i) = \frac{1}{m-1} \sum_{t=2}^m \left( \max \left( 0, l \times (f(x_{t-1}) - f(x_t)) \right) \right)^2, \quad (1)$$

where the variable  $l$  has two possible values:

$$l = \begin{cases} 1 & \text{if } y_i = 1 \\ -1 & \text{if } y_i = 0. \end{cases}$$

For brevity we use  $f(x_i)$  instead of  $f(x_i, \theta)$  in the equation to denote the score of 1-s clip  $x_i$ . The main idea is to impose constraints that require the detection score of a frame at time  $t$  to be higher than the score of any other frame that has been seen previously—for processing a positive training sample. Similarly, the scores should be decreasing if  $X_i$  is negative (i.e.  $y_i = 0$ ). The design of  $L_{conf}$  provides a mechanism to achieve the monotonicity of the score function  $f(x_i, \theta)$ . The detection score of a hardly visible parking space (when the driver is far from it) cannot exceed that of a visible one. A parking space has the highest detection score when it is completely seen.

The second term of the loss function measures the prediction accuracy:

$$L_{acc}(X_i) = \frac{1}{m} \sum_{t=1}^m (y_i - f(x_t))^2, \quad (2)$$

This is basically a standard mean squared loss. The final loss is the sum of these losses across all training videos:

$$\sum_{X_i \in D} (L_{conf}(X_i) + \alpha \cdot L_{acc}(X_i)), \quad (3)$$

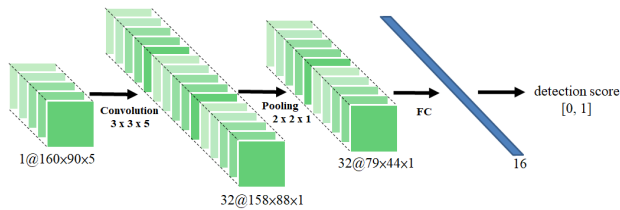


Figure 2: Structure of the first model based on 3D-CNN.

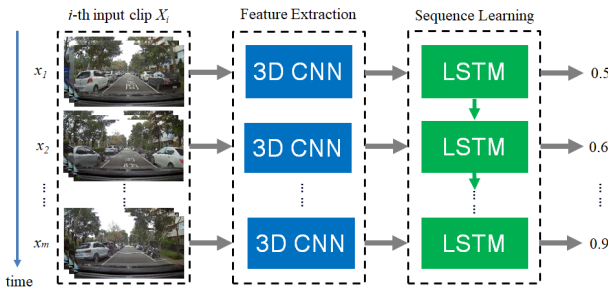


Figure 3: Illustration of the second model combining 3D CNN and RNN.

where the parameter  $\alpha$  controls the tradeoff between them ( $\alpha$  was empirically set to 1 in this study). Using the proposed loss function enables the detection of vacant parking spaces before they are completely observed. During training, the detector simulates the sequential arrival of dashcam video data and is trained to recognize not-fully-observable parking spaces. We train a single detector to recognize parking spaces under different visibilities.

### Network Architecture

We implemented two detectors using deep learning methods. We will compare the performance of these two methods in the experiment.

The first network model used in this study is a spatial temporal 3D convolutional neural network (3D CNN) (Ji et al. 2013). CNNs have been shown great promise in many visual recognition tasks (Krizhevsky, Sutskever, and Hinton 2012; Simonyan and Zisserman 2015). Because a video clip consists of a stack of images, we extend CNNs by adding the time dimension to process multiple images simultaneously. The network architecture is shown in Fig. 2. We capture the appearance changes from the input 1-s clip via the convolutional layer. Because the input data has a third dimension (i. e. time), the convolutional layer and the pooling layer of the network consist of three-dimensional kernels. We set the kernel size of the convolution layer to  $3 \times 3 \times 5$  because 5 frames are sampled from a 1-s input clip, providing complementary yet redundant motion information.

The second network model is built upon RNN with Long Short-Term Memory (LSTM) cells (Hochreiter and Schmidhuber 1997). RNN is a special type of network which takes a sequence of observations as input and returns a sequence of hidden representations and outputs. It is particularly suitable for analyzing sequential data, because the recurrence oper-

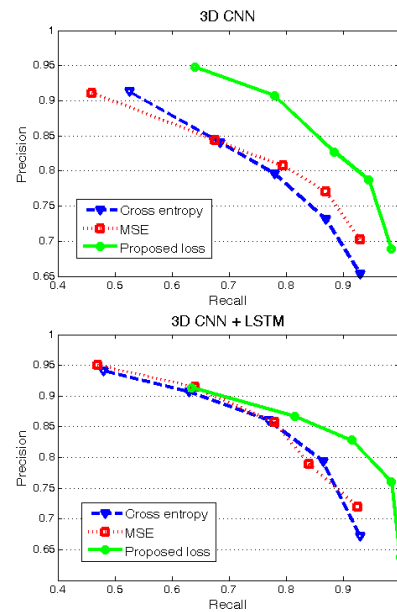


Figure 4: Performance comparison of the cross entropy loss, MSE and the proposed loss: (top) 3D CNN (bottom) 3D-CNN+LSTM.

ation allows response to depend on a time-evolving state. However, it also suffers from the vanishing gradients problem; and therefore, training a RNN to capture long-term dependencies is difficult. A practical way to address this problem is to use LSTM cells (Hochreiter and Schmidhuber 1997), which are introduced to memorize and maintain information over time.

The architecture of our second method, referred to as 3D-CNN+LSTM, is shown in Fig. 3. It incorporates 3D CNN in a RNN framework. Feature maps obtained via 3D CNN are fed into the LSTM layer to model the dependency of each output on the previous hidden representations and outputs. The number of network parameters of 3D CNN and 3D-CNN+LSTM are 1,781,217 and 7,121,425, respectively.

## Experiments

In this section, we first describe the experimental setup and the evaluation methods. Then, we describe the baseline methods for comparison and report the experimental results.

### Setup

The dashcam video dataset introduced in Section 3 is used in the experiments. We randomly split the dataset into training, validation and testing: 5,000 training clips, 400 validation clips, and 400 testing clips. Training and testing clips capture completely different places.

The evaluation is performed on the basis of the correctness of classifying a 5-s video clip. Given a test video, a method is required to predict the probability of the availability of a parking space ahead in each second. When the probability is greater than or equal to 0.5, the frame provides a visual evidence of possible vacancy ahead.

Threshold $T$	1	2	3	4	5
Cross Entropy	71.75%	77.50%	79.00%	84.50%	73.75%
MSE	76.75%	80.50%	80.25%	77.50%	70.75%
Proposed	77.00%	84.50%	85.00%	85.00%	81.50%

Table 1: Classification accuracy rates of different loss functions using 3D CNN.

Threshold $T$	1	2	3	4	5
Cross Entropy	73.75%	82.00%	82.25%	78.25%	72.50%
MSE	78.25%	80.75%	82.50%	79.00%	72.25%
Ours	71.50%	83.75%	86.25%	84.50%	78.75%

Table 2: Classification accuracy rates of different loss functions using 3D-CNN+LSTM.

A detection is made when sufficient visual evidence has become available. When the number of accumulated positives is above a threshold  $T$ —which controls the sensitivity of the detector—, the method claims that there will be a parking space ahead. If the video is a positive video, this is a true positive (TP) detection; if the video is negative, this is a false positive (FP) detection. On the other hand, when the number of accumulated positives is below  $T$ , the method claims that there will be no vacant parking space ahead. If the video is a positive video, this is a false negative (FN) prediction; otherwise this is a true negative (TN) detection. For each threshold  $T$ , we compute the precision =  $\frac{TP}{TP+FP}$ , recall =  $\frac{TP}{TP+FN}$  and report the precision-recall curves. We also report the classification accuracy =  $\frac{TP+TN}{TP+FP+TN+FN}$ .

## Results

We compare the proposed detection loss to two baselines: the cross entropy loss and the mean squared loss (MSE). We test these loss functions using the 3D CNN and the 3D-CNN+LSTM architectures. We report the precision-recall curves in Fig. 4 and the classification performances in Table 1 and Table 2.

We first observe that these two network models have comparable performances. Using LSTM slightly improves the classification performance when the cross entropy loss and MSE are used. Nevertheless, the 3D-CNN+LSTM model is much more complex and intense than the 3D CNN model.

The proposed loss function consistently outperforms the baselines with all measures. For example, under a rigorous setting ( $T = 5$ ) the proposed method outperforms cross entropy by 7.75% and outperforms MSE by 10.75% in classification accuracy, experimented using the 3D CNN model. The improvement can also be observed when using the 3D-CNN+LSTM model, although the performance gap is less significant.

Because the proposed loss constrains the subsequent detection scores, which in turn models the dependency of output scores, the performance difference between 3D CNN and 3D-CNN+LSTM trained using the proposed loss is negligible, as displayed in Fig. 5. Equipped with the proposed loss, a simple model (3D CNN) achieves a slight improvement over a complex model (3D-CNN+LSTM).

Time to Parking Lot	3 s	2 s	1 s	0 s
Cross Entropy	63.5%	68.5%	75.5%	82.0%
MSE	67.0%	72.5%	78.0%	83.0%
Ours	80.0%	87.5%	93.0%	97.0%

Table 3: Performance on early detection. Using 3D-CNN+LSTM and the proposed loss can predict a vacant parking space 3 seconds before reaching it with 80% recall.

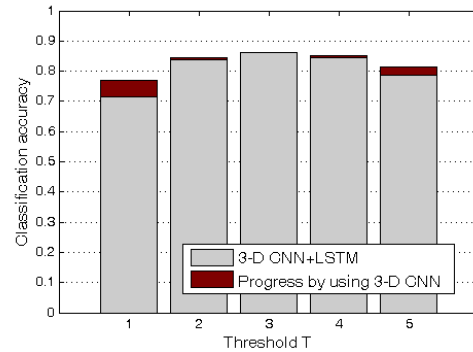


Figure 5: Performance comparison of 3D CNN and 3D-CNN+LSTM using the proposed loss.

Finally, we discuss the performance of early detection. In this experiment we implemented a detector which alerted the driver when two consecutive evidences were observed. Table 3 shows the recall rates of the 3D-CNN+LSTM model trained to optimize different loss functions. The detector combining 3D-CNN+LSTM and the proposed loss has the 86% classification accuracy. This combination predicts a parking space 3 seconds before reaching it with 80% recall. The recall rate increases (from 80% to 97%) with respect to the time to approaching a vacant parking space.

## Conclusions

We propose to detect vacant parking spaces using dashcam videos. A new video dataset has been collected to train and validate such a system. To address this challenging task, we

have implemented two baseline network models based on deep learning. We have further proposed a new detection loss to constrain the detection scores in subsequent frames. Our method can predict an empty parking spot 3 seconds before reaching it with 82.5% precision and 80% recall.

The current method detects the occurrence of a vacant parking space alone. One future research direction we are pursuing is to extend the method for parking space localization, which further determines the location (e.g. which side on the road and how far) of the parking space. Another direction is to develop the method in a low-cost dashboard camera system.

## References

- Almeida, P. R.; S.Oliveira, L.; Jr., A. S.; Jr., E. J. S.; and L.Koerich, A. 2015. Pklot - a robust dataset for parking lot classification. *Expert Systems with Applications* 42:4937–4949.
- Amato, G.; Carrara, F.; Falchi, F.; Gennaro, C.; Meghini, C.; and Vairo, C. 2017. Deep learning for decentralized parking lot occupancy detection. *Expert Systems with Applications* 72:327–334.
- 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)*.
- Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; and Schiele, B. 2016. The cityscapes dataset for semantic urban scene understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- del Postigo, C. G.; Torres, J.; and Menéndez, J. M. 2015. Vacant parking area estimation through background subtraction and transience map analysis. *IET Intelligent Transport Systems* 9:835–841.
- Enríquez, F.; Soria, L. M.; Álvarez García, J. A.; Velasco, F.; and Déniz, O. 2017. Existing approaches to smart parking: An overview. In *Proc. of the International Conference on Smart Cities*, 63–74.
- Geiger, A.; Lenz, P.; Stiller, C.; and Urtasun, R. 2013. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*.
- Giles, C. L.; Kuhn, G. M.; and Williams, R. J. 1994. Dynamic recurrent neural networks: Theory and applications. *IEEE Transactions on Neural Networks* 5(2):153–156.
- Girshick, R.; Donahue, J.; Darrell, T.; and Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587.
- Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural Computation* 9(8):1735–1780.
- Hsieh, M.-R.; Lin, Y.-L.; and Hsu, W. H. 2017. Drone-based object counting by spatially regularized regional proposal network. In *Proc. of the IEEE Conference on Computer Vision (ICCV)*.
- Huang, C.-C., and Wang, S.-J. 2010. A hierarchical bayesian generation framework for vacant parking space detection. *IEEE Transactions on Circuits and Systems for Video Technology* 20:1770–1785.
- 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*.
- Ji, S.; Xu, W.; Yang, M.; and Yu, K. 2013. 3d convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(1):221–231.
- Krizhevsky, A.; Sutskever, I.; and Hinton, G. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, 1097–1105.
- LeCun, Y.; Boser, B. E.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W. E.; and Jackel, L. D. 1990. Handwritten digit recognition with a back-propagation network. In *Advances in Neural Information Processing Systems*.
- Lin, S.-F.; Chen, Y.-Y.; and Liu, S.-C. 2006. A vision-based parking lot management system. In *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*.
- Peng, C.-F.; Hsieh, J.-W.; Leu, S.-W.; and Chuang, C.-H. 2018. Drone-based vacant parking space detection. In *Proc. of the International Conference on Advanced Information Networking and Applications Workshops (WAINA)*.
- Sastre, R. J. L.; Jimenez, P. G.; Acevedo, F. J.; and Bascon, S. M. 2007. Computer algebra algorithms applied to computer vision in a parking management system. In *Proc. of the IEEE International Symposium on Industrial Electronics*.
- 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)*.
- Simonyan, K., and Zisserman, A. 2015. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556v6*.
- Viola, P., and Jones, M. J. 2004. Robust real-time face detection. *International Journal of Computer Vision* 57(2):137–154.
- 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*.
- 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*.