Journal of Engineering and Applied Sciences 13 (Special Issue 9): 7062-7067, 2018
ISSN: 1816-949X

Cost-Effective Outdoor Car Park System with Convolutional Neural Network on Raspberry Pi

Chin-Kit Ng and Soon-Nyean Cheong
Faculty of Engineering, Multimedia University, 63000 Cyberjaya, Selangor, Malaysia

Abstract: Difficulty in finding a vacant parking space has always been a problem encountered by drivers especially in metropolitan areas. This study proposed a cost-effective vision-based outdoor parking space vacancy detection system, ConvPark to assist vehicle drivers by providing information regarding the availability of parking spaces. The system is designed based on Convolutional Neural Network (CNN) technology and is implemented through a Raspberry Pi to identify the occupancy status of parking spaces live via. an IP camera. This system has been deployed at a university car park for real-time detection of vacant parking spaces. The use of CNN classifier in the proposed system provides superiority in term of automatic image features extraction and robustness against environmental variations as compared to other computer vision-based methods. Evaluation outcomes demonstrated that our proposed system can achieve excellence performance in term of detection accuracy by precisely determining the occupancy status of parking spaces under different environmental conditions.

Key words: Convolutional neural network, outdoor car park system, Raspberry Pi, environmental conditions, extraction, determining

INTRODUCTION

With dramatically increase in the amount of private vehicles and high traffic density especially in metropolitan areas, searching for a parking space has become a big obsession for every driver because it is not only time-consuming but also aggravate the problem of traffic congestion as highlighted by SC (2006). In view of this predicament, several Vacant Parking Spaces Detection (VPSD) systems were proposed, utilizing either sensor-based or computer vision-based method to present information regarding availability and location of parking spaces for drivers. Pavement embedded system and overhead technology are two common implementation of sensor-based systems where different types of sensors including infrared sensors, magnetic field sensors, etc. are used for parking occupancy detection (Chen and Chang, 2011). However, applicability of such systems are limited particularly for outdoor parking lots due to difficulty involved in installation and maintenance procedures. More importantly, such systems are expensive to realize as individual sensors are needed to monitor each parking space. On the contrary, computer vision-based method has the potency to give a more economical solution, thus, making this method to gain ample discussion in recent years.

As proposed by Huang and Wang (2010) in their work existing computer vision-based methods for VPSD are either car-driven or space-driven. For the former method, car features are used in vehicle detection algorithms as described by Ichihashi *et al*. (2010). However, due to perspective distortion, the detection accuracy tends to deteriorate for those cars which located far away from the camera view (Huang and Wang, 2010). On the other hand, space-driven method uses the features of an empty parking space and relies on background subtraction algorithms as proposed by Power and Schoonees (2002) to achieve foreground detection of occupied parking spaces. Implementation of these algorithms normally assume that the background variation is relatively static within a short interval and the appearance of empty parking spaces is homogeneous on the whole (Huang and Wang, 2010). Nonetheless, these assumptions may not be valid for outdoor parking lots where rapid environmental changes in term of luminance level and object occlusions may occur, thereby causing degradation in detection accuracy (Huang and Wang, 2010). By Tschentscher *et al*. (2015) various supervised learning algorithms (i.e., SVM, k-NN, etc.,) are employed to perform classification on the occupancy status of parking space based on color histograms and DoG

Corresponding Author: Chin-Kit Ng, Faculty of Engineering, Multimedia University, 63000 Cyberjaya, Selangor, Malaysia

Table 1: Characteristics of existing VPSD system

Variable	Automatic image feature extraction	Implementation cost	Invariant against environmental variations
Sensor-based method (Chen and Chang, 2011)	X	High	X
Vision-based method (car-driven (Ichihashi *et al*., 2010) and space driven (Tschentscher *et al*., 2015))	X	Low	X
ConvPark	✓	Low	✓

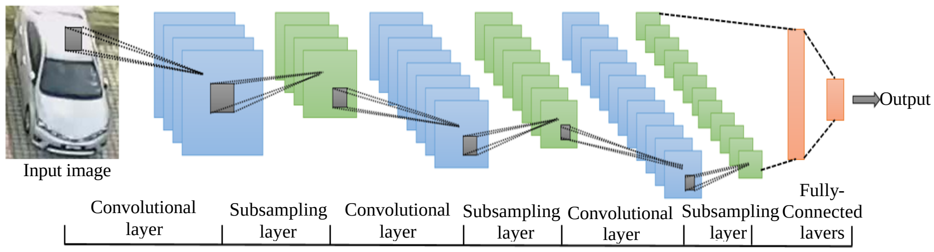

Fig. 1: Network structure of cnn

features. However, additional algorithms to extract out self-engineered image features are required, thereby adding complexities to the system implementation.

Table 1 summarizes the characteristics of existing VPSD system aforementioned. Generally, vision-based methods outperform sensor-based methods in term of implementation cost as camera sensor provides a larger coverage of parking spaces detection while simplifying the installation and maintenance procedures (Kharusi and Bahadly, 2014). However, most of the existing vision-based methods used in outdoor parking spaces do not support automatic image feature extraction (i.e., require self-engineered feature extraction algorithm) or provide treatments against influences of environmental variations, thereby subject to deterioration in detection accuracy. Therefore, in this study, a cost-effective vision-based outdoor vacant parking space detection system (ConvPark) using CNN is presented. The use of CNN classifier in ConvPark offers superiorities in term of automatic image feature extraction and invariant against environmental variations, thereby capable of achieving high detection accuracy.

MATERIALS AND METHODS

Enabling technologies:

Convolutional Neural Network (CNN): CNN is a biologically-inspired variant of Multilayer Perceptron (MLP) which have attained outstanding performance on diverse image recognition standard inclusive of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC 2012) (Krizhevsky *et al*., 2012). Figure 1 shows the network structure of CNN consisting of two alternating layers. The convolutional layer automatically extracts rich features from input feature maps and combine the features into more abstract output feature maps whereas the subsampling layer reduces the resolution of feature maps to provide spatial invariance property. Fully connected ANN classifier in the last few layers are used to generate abstracted classification results. By Chen *et al*. (2014) a variant of CNN is employed for vehicle detection in satellite images and have reported a high detection accuracy of over 90%, thereby assuring the feasibility of CNN for similar application in VPSD which involves detecting the existence of vehicles.

Raspberry Pi: Raspberry Pi is a small, low-priced, low power consumption credit card-sized computer running on a Raspbian operating system. Recent years, the increasing processing power from first generation to the current third generation Raspberry Pi has unfolded its applicability to be used as a popular hardware platform for diverse real-time applications as described by Shinde *et al*. (2015). In addition by Celebre *et al*. (2015) Raspberry Pi has been employed as the processing unit to interface with various hardware appliances for the development of home automation system. Other than that, Raspberry Pi is also being employed as a base station in Wireless Sensor Network (WSN) systems to serve a user-friendly web interface to the respective system (Nikhade, 2015). Clearly, the considerably high processing capability of Raspberry Pi coupled with the abilities to interface with various devices and support network communication has made it an applicable candidate for VPSD system development meanwhile providing a low-cost solution.

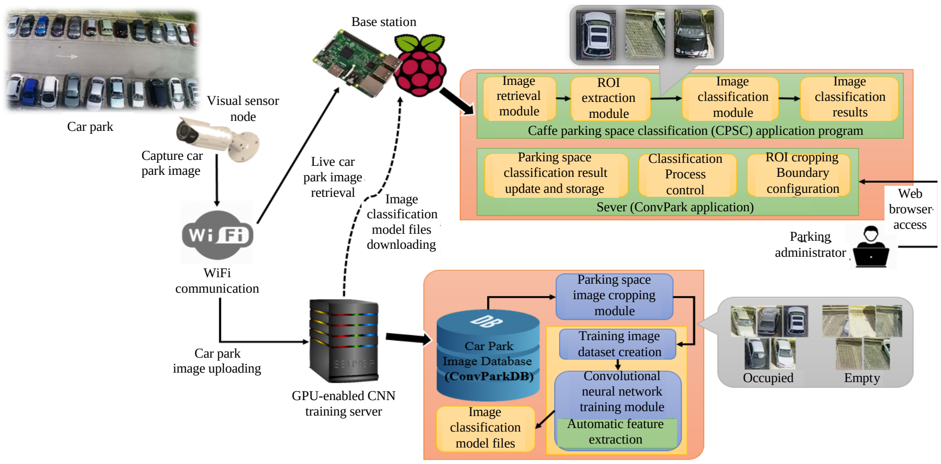

Fig. 2: Convpark system architecture

Convpark system design: System architecture of ConvPark is depicted in Fig. 2 with three main parts: the visual sensor node, the base station and the GPU-enabled CNN training server. Communication between the different parts of the system is established via. Wi-Fi connection.

The visual sensor node is used to capture car park images and store them into a car park image database resides in the CNN training server. Image cropping module within the server will crop out the regions of an individual parking space available in each car park image and categorize them as either occupied or empty. The two classes of parking space images are used to create a training image dataset which is then fed into the CNN training module whereby rich image features are automatically extracted during training process. Upon completion of the training process, an image classification model containing the information of training parameters and results will be generated and accessed by base station.

Base station is responsible to detect vacant parking spaces through the proposed Caffe Parking Space Classification (CPSC) application. Live car park images are first fed into base station through image retrieval module. Subsequently, the ROI extraction module extracts all the regions of an individual parking space within the latest retrieved car park image. The extracted images along with the image classification model files retrieved from the CNN training server are used as inputs to the image classification module to obtain the classification results regarding the occupancy status of parking spaces. Customized ConvPark web application was designed to allow parking administrators to obtain the occupancy information or configure the system anytime and anywhere through their mobile devices.

ConvPark system implementation:

The hardware components: The proposed ConvPark system has been implemented at the staff car park of Faculty of Engineering, Multimedia University, Malaysia. A Vstarcam C7815WIP outdoor HD IP camera that supports IP66 waterproof, WiFi 802.11 b/g/n and 720p video stream functionality is used as the visual sensor to capture car park images. Single board computer namely Raspbery Pi 3 Model B is utilized as the base station. It is equipped with a 16 GB Micro SD card storage and runs on Raspbian Jessie operating system to execute the CPSC application as well as ConvPark web application. A Dell Inspiron 15 7000 laptop that comes with 4GB video memory NVIDIA GeForce GTX 960M graphics card running on Ubuntu MATE 14.04 LTS is employed as the GPU-enabled CNN training server.

Collection of training dataset and testing dataset: Car park images are collected from the IP camera and stored to a car park image database created on the CNN training server at 5 min interval starting from 7 am until 7 pm for 5 consecutive days (Monday-Friday). For all collected images, the regions of 23 parking spaces within the car park are cropped out using an image cropping application program developed with the Python API of OpenCV software library. Around 15k individual parking space images are yielded and categorized into two categories, “Empty” and “Occupied” according to their occupancy status to be used as training dataset. An isolated image sequence which is not part of the training dataset is captured on a typical working day and used as the test data of our experiment to assess the accuracy performance of the system. Figure 3 and 4 depicts image samples of the car park captured at different time instants

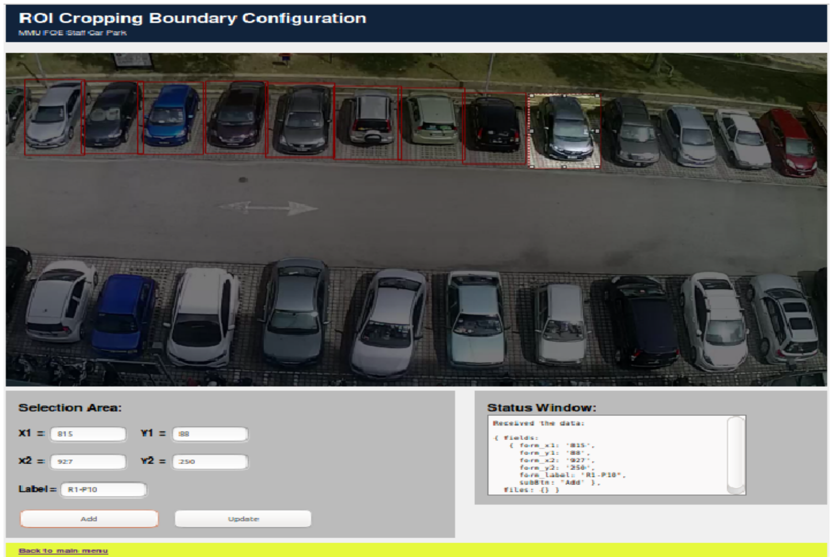

Fig. 3: ROI cropping boundary configuration tool

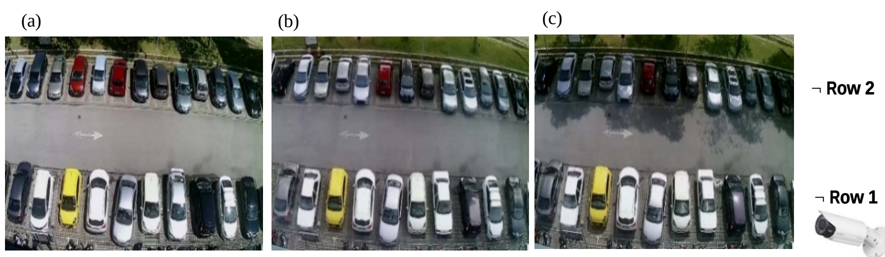

Fig. 4: Car park image samples at different time instants:. a) Represents sunny weather condition; b) Represents cloudy weather condition and c) Represents shadow effect condition

which demonstrate sunny (Fig. 4a) and cloudy weather conditions (Fig. 4b) as well as the one with severe shadow effect (Fig. 4c). Overall, 3298 parking space images are extracted from the image sequence and used as the testing dataset.

Training dataset creation and CNN training: Training dataset creation and CNN training are carried out on the CNN training server using DIGITS, an interactive deep learning GPU training system developed by NVIDIA. A training image dataset are created from the two categories of parking space images with image size scaled to 200×200. Based on the pre-trained AlexNet from Alex (Krizhevsky *et al*., 2012) a customized CNN, ConvPark Net is constructed by dropping two convolutional layers which are not followed by sub-sampling layers (i.e., 3 and conv 4) and a fully-connected layer fc7 in the original AlexNet. To further reduce the computation time for real-time applicability, the number of outputs of all convolutional and fully-connected layers are reduced significantly. The resultant architecture of ConvParkNet which comprised of three alternating convolutional and sub-sampling layers followed by two fully-connected layers is reasonable for our case with much simpler binary classification problem. The network is trained using the created dataset for 30 epochs with base learning rate set to 0.01 and Stochastic Gradient Descent (SGD) solver type chosen. The concept of GPGPU is employed to accelerate the training of ConvParkNet which involves intensive computation process. Training performance of ConvParkNet indicates that the network converges within one training epoch and the validation accuracy saturated at 99.9% at the 20th training epoch. For each training epoch, an image classification model is generated, containing the network parameters of the trained ConvParkNet which are crucial information for image classification.

Caffe Parking Spaces Classification (CPSC): At the base station, using Ffmpeg command line tool, a single live car park image is retrieved every five second from the IP camera video streaming server and overwritten with the same file name to ensure the image used for classification process is showing the latest car park setting. Based on a predefined cropping boundary definition file, a set of 23 parking space images are extracted from the latest retrieved car park image using same method as the one implemented in CNN training server. An image classification program is developed using the Python interface of Caffe deep learning framework to classify parking space images based on the image classification model downloaded from the DIGITS result browser running on CNN training server. The classification of 23 parking spaces for each iteration consumes about 5 sec seconds to complete. Consequently, the classification results regarding the occupancy condition of individual parking space is stored to a JSON file.

ConvPark web application: ConvPark web application is developed using Node.js and is hosted on the base station to provide two key web services for parking administrator. First is the ROI cropping boundary configuration service, developed using the imgAreaSelect jQuery plugin to allow parking administrator to define the cropping boundary for each parking space within the car park image. As shown in Fig. 3, the region of an individual parking space can easily be selected with a bounding box by clicking and dragging on the image. The coordinates of the selected parking space region together with its user-specified label are saved to a buffer and eventually updated to a cropping boundary definition file to be used for parking space image cropping. The second service is classification process control which allows administrator to activate and suspend CPSC application remotely.

RESULTS AND DISCUSSION

Detection accuracy on vacant parking space: The ground truth of testing dataset are manually established in order to evaluate the accuracy of the ConvPark system in identifying vacant parking spaces. Calculation of False Positive Rate (FPR), False Negative Rate (FNR) and overall detection accuracy (ACC) are performed to provide quantitative representations of the assessment results. Table 2 shows the detection accuracy of ConvPark system which demonstrated excellent performance in all the six evaluation aspects. Besides performing assessment on the entire image sequence, the sequence is further divided into three subcategories representing morning, noontime and afternoon periods to evaluate the detection performance under different illumination conditions with varying shadow patterns. Generally, afternoon period exhibits most severe shadow effects whereas noon period has almost no shadow effects. From the ACC of those three periods, the proposed system is proved to be relatively invariant against the influence of shadow effects with over 99% detection accuracy.

To evaluate the robustness of ConvPark system against perspective distortion, the image sequence is separated into two subcategories according to the distance of parking spaces from the IP camera in ascending order, namely first row and second row. Similarly, results in Table 2 indicate that perspective distortion does not significantly deteriorate the detection accuracy.

The assessment results show that ConvPark is robust against the influence of environmental variations as compared to car-driven (Ichihashi *et al.*, 2010) and space-driven approaches (Tschentscher *et al.*, 2015) in computer vision-based method and an excellent detection accuracy of over 99% have been achieved. This is because the hierarchical structure of CNN allows generation of image features which are invariant to various environmental changes (Mallat, 2016). Moreover, implementation of ConvPark does not required any self-engineered image features because CNN feature learning architecture is able to extract and pick up essential features automatically by taking into account the details of training dataset. In addition, compared to sensor-based method (Chen and Chang, 2011) ConvPark offers a cost-effective implementation through the use of low-cost Raspberry Pi and single IP camera to provide a wider area of parking space detection instead of detecting parking space using individual sensor which is more costly.

Table 2: Performance of ConvPark system on detection accuracy of vacant parking spaces

	No. of tested spaces			Assessment results		
Test data	Empty	Occupied	Total	FPR	FNR	ACC
Entire image sequence	611	2687	3298	0.0052	0.0033	0.9951
Morning period	190	638	828	0.0078	0.0053	0.9928
Noon period	79	749	828	0.0000	0.0000	1.0000
Afternoon period	28	778	806	0.0026	0.0357	0.9963
First row	213	1217	1430	0.0000	0.0000	1.0000
Second row	398	1470	1868	0.0095	0.0050	0.9914

CONCLUSION

This study presents an outdoor parking space vacancy detection system, ConvPark which is based on CNN technology. The system is capable of automatic image features extraction and is robust against the influences of environmental variations in outdoor car park, thereby overcomes the limitations of current computer vision-based VPSD system besides providing a cost-effective solution. This system has been successfully implemented and deployed at a university car park for real-time detection of vacant parking spaces. Our experimental results show that the detection accuracy of ConvPark can achieve up to 99% for evaluation over an entire day, over distinct time of a day and over distinct areas of the car park. Currently, the system evaluation only focuses on detection accuracy.

SUGGESTIONS

In near future, user experience evaluation needs to be carried out in order to examine the system applicability and reliability from user perspective, thereby providing visions to further improve the system.

ACKNOWLEDGEMENT

The researcher gratefully acknowledge Multimedia University (Cyberjaya campus) for the financial support to implement the project successfully.

REFERENCES

Celebre, A.M.D., A.Z.D. Dubouzet, I.B.A. Medina, A.N.M. Surposa and R.C. Gustilo, 2015. Home automation using raspberry Pi through Siri enabled mobile devices. Proceedings of the IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), December 9-12, 2015, IEEE, Cebu City, Philippines, ISBN:978-1-5090-0361-7, pp: 1q-6.

Chen, M. and T. Chang, 2011. A parking guidance and information system based on wireless sensor network. Proceedings of the IEEE International Conference on Information and Automation (ICIA), June 6-8, 2011, IEEE, Shenzhen, China, ISBN:978-1-4577-0268-6, pp: 601-605.

Chen, X., S. Xiang, C.L. Liu and C.H. Pan, 2014. Vehicle detection in satellite images by hybrid deep convolutional neural networks. IEEE. Geosci. Remote Sens. Lett., 11: 1797-1801.

Huang, C.C. and S.J. Wang, 2010. A hierarchical bayesian generation framework for vacant parking space detection. IEEE. Trans. Circuits Syst. Video Technol. 20: 1770-1785.

Ichihashi, H., T. Katada, M. Fujiyoshi, A. Notsu and K. Honda, 2010. Improvement in the performance of camera based vehicle detector for parking lot. Proceedings of the IEEE International Conference on Fuzzy Systems, July 18-23, 2010, IEEE, Barcelona, Spain, ISBN:978-1-4244-6919-2, pp: 1950-1956.

Kharusi, H.A. and I.A. Bahadly, 2014. Intelligent parking management system based on image processing. World J. Eng. Technol., 2: 55-67.

Krizhevsky, A., I. Sutskever and G.E. Hinton, 2012. Image net classification with deep convolutional neural networks. Proc. Neural Inf. Process. Syst., 1: 1097-1105.

Mallat, S., 2016. Understanding deep convolutional networks. Phil. Trans. R. Soc. A, 374: 1-16.

Nikhade, S.G., 2015. Wireless sensor network system using Raspberry Pi and zigbee for environmental monitoring applications. Proceedings of the 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), May 6-8, 2015, IEEE, Maharashtra, India, ISBN: 978-1-4799-9854-8, pp: 376-381.

Power, P.W. and J.A. Schoonees, 2002. Understanding background mixture models for foreground segmentation. Proceedings of the Conference on Image and Vision Computing New Zealand, November 26-28, 2002, University of Auckland, Auckland, New Zealand, pp: 267-271.

SC., 2006. Curbing cars, shopping, parking and pedestrian space in Soho. Transportation alternatives, New York, USA.

Shinde, P.A., Y.B. Mane and P.H. Tarange, 2015. Real time vehicle monitoring and tracking system based on embedded Linux board and android application. Proceedings of the IEEE International Conference on Circuit, Power and Computing Technologies (ICCPCT), March 19-20, 2015, IEEE, Nagercoil, India, ISBN:978-1-4799-7075-9, pp: 1-7.

Tschentscher, M., C. Koch, M. Konig, J. Salmen and M. Schlipsing, 2015. Scalable real-time parking lot classification: An evaluation of image features and supervised learning algorithms. Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), July 12-17, 2015, IEEE, Killarney, Ireland, ISBN:978-1-4799-1961-1, pp: 1-8.