Se-PSD: A sequence-based parking-slot detection approach for indoor parking

Jiaxin Zhang

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611730, China

jx_zhang@std.uestc.edu.cn

Abstract. Accurate parking slot detection is crucial for autonomous vehicles to navigate automatic parking. In recent years, significant progress has been made in research in this field, with some methods achieving high accuracy and efficiency on specific datasets. However, existing methods still face some challenges in practical applications, such as: when the ground reflection covers the parking-slot marking lines, the recognition accuracy of existing methods will significantly decrease; The generalization performance of existing methods is not strong enough, and the recognition accuracy will decrease after changing the place or camera model. To address these limitations, this paper introduces Se-PSD, a novel parking slot detection method utilizing image sequences. Se-PSD analyzes a series of images to predict individual marking point locations, shapes, and orientations. Finally, through geometric rules, parking-slots can be found on the last image of the image sequence. Se-PSD prioritizes generalizability without sacrificing accuracy compared to existing methods. While real-time performance may be slightly impacted, the relaxed time constraints of automatic parking applications make Se-PSD a promising solution.

Keywords: Autonomous Parking, Parking-slot Detection, LSTM.

1. Introduction

Automatic parking is an important part of autonomous driving technology and one of the key technologies to achieve fully autonomous driving. Automatic parking systems can help vehicles automatically complete parking operations, reduce the driver's burden [1], improve parking efficiency, and lay the foundation for the popularization of autonomous driving. A key issue in automatic parking is how to use onboard sensors to detect and locate parking-slots surround the vehicle effectively and correctly.

There are two main types of parking-slot detection algorithms, free-space-based and vision-based. Free-space-based approaches detect parking-slots by measuring the distance between adjacent vehicles. These algorithms typically use a variety of range sensors, such as ultrasonic radar [2], laser scanners [3], and short-range radars [4]. The main advantage of space-based algorithms is their low cost and simple implementation. However, these approaches are highly relied on the presence of other parked vehicles as a reference. Therefore, they may fail to work in empty parking lots or other scenarios where there are no other vehicles present. Vision-based approaches overcome the limitations of free-space-based

DOI: 10.54254/2755-2721/75/20240522

approaches by using cameras to capture images of the surrounding environment. These approaches then employ image recognition and deep learning techniques to identify parking-slots in the images.

2. Related works

2.1. Vision-based parking-slot detection

Vision-based parking-slot detection is a method of using cameras to recognize ground parking-slot lines, which is intuitive and in line with the human habit of finding parking-slot. Compared to empty spaces, parking-slot lines can provide more accurate parking information. In addition, with the popularization of bird-eye's view systems [5], almost all vehicles are equipped with wide-angle cameras [6], providing a hardware foundation for vision-based parking-slot detection.

Research in this field began with semi-automatic method. Xu et al. [7] first proposed a segmentation method based on the color of parking-slot lines and estimated the contour of parking-slots using two perpendicular lines. Although this method is easy to use, it cannot recognize the type of parking-slot and cannot solve the problem of oblique parking-slot detection. On this basis, Jung et al. [8] proposed a semi-automatic two-point parking-slot detection method that can recognize various types of parkin-slots. Du and Tan [9] have developed a reverse parking system that utilizes a ridge detector to extract centerlines of parking-slot lines. However, the above methods [7, 8, 9] are semi-automated, they rely on human drivers to give prompts, which limits their application in practice.

Limitations of semi-automatic methods spurred research in fully automated vision-based parking slot detection. These methods fall into two categories: line-based and point-based. Line-based methods primarily rely on detecting lane markings. Common approaches include edge detection (Sobel filters, Canny edge detectors [10]) or line segment detection (LSD [11]). Extracted lines are then fitted using techniques like Hough transform [10],, Radon transform [5], RANSAC [10], or customized clustering algorithms [12]. Finally, the geometric relationship between lines determines the parking slot's entry line, leading to slot detection.

Point-based methods utilize various strategies for identifying marking points within parking-slots. Suhr et al. [13] employed the Harris corner detector, followed by template matching, to determine the shape, direction, and relative position of the parking slot. Li et al. [14] leveraged a boosting decision tree and a Gaussian line filter to detect and determine the direction of the entry line, respectively. Zhang et al. [15] introduced DeepPS, a CNN-based method, outperforming prior approaches using only low-level features. This finding underscores the effectiveness of CNNs for this task. Building upon DeepPS, Zhang et al. [16] proposed DMPR-PS, a one-stage CNN method that directly regresses the position, type, and direction of directional marking points from the image. This approach leverages simple combination and geometric judgment to achieve state-of-the-art results on the ps2.0 [15] dataset.

2.2. Recurrent neural network and long short-term memory

Our proposed parking-slot detection method, Se-PSD, leverages sequential image data. In deep learning, Recurrent Neural Networks (RNNs) [17] and Long Short-Term Memory (LSTM) networks [18] are adept at handling such sequences. Let's briefly explore these architectures. RNNs are powerful models that can learn from sequential data. Unlike traditional neural networks, they can retain information across processing steps, allowing them to capture temporal dependencies. However, traditional RNNs struggle with long-term dependencies. Information from earlier time steps can fade over time as it passes through recurrent connections. To address this limitation, LSTMs were introduced. These networks incorporate a memory gating mechanism, enabling them to effectively learn and remember long-term dependencies.

2.3. Our motivations and contributions

While the directional marking point (DMP)-based DMPR-PS method achieves high accuracy and efficiency on the ps2.0 dataset, it exhibits limitations in underground parking lots. Strong reflections from epoxy floors and insufficient camera tolerance can obscure marking points, leading to detection failures. Additionally, DMPR-PS suffers from weak generalization, as shown in Figure 1. Its performance may decline on datasets with different collection sensors or lighting conditions, such as the BeVIS underground parking SLAM dataset [19] by Zhang et al., where scenes overlap with ps2.0 but differ in image brightness and clarity. It's important to remember that parking-slot detection is a continuous process. Image sequences acquired during a vehicle's movement contain richer information than individual frames.

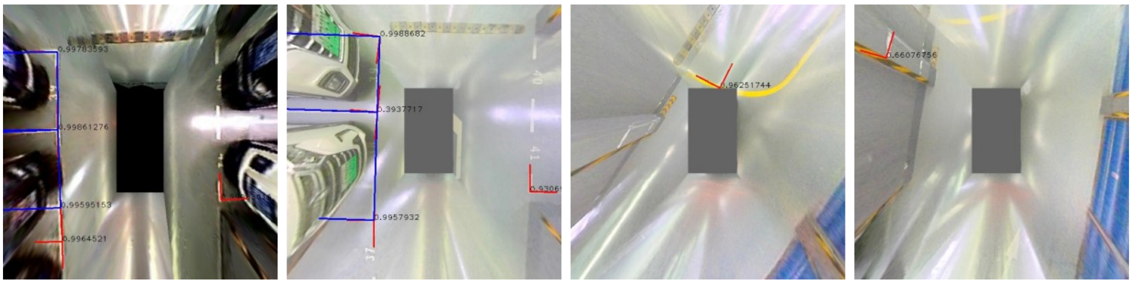

Figure 1. Failure of DMPR-PS in the same place and image from different sensor.

We introduce Se-PSD, a novel parking-slot detection method specifically designed to address the limitations of DMPR-PS in challenging underground parking environments. Unlike DMPR-PS, Se-PSD leverages a sequence of images as input, enabling it to capture temporal information and potentially handle dynamic parking lot scenes. While Se-PSD utilizes the same directional marking points from DMPR-PS [16] for spatial awareness, it employs a distinct approach that prioritizes generalizability. This approach, while potentially sacrificing some operational efficiency compared to DMPR-PS, allows Se-PSD to achieve significantly stronger performance on the BeVIS dataset [19]. Notably, Se-PSD maintains its high accuracy on the benchmark ps2.0 dataset, demonstrating its effectiveness in both controlled and real-world scenarios. This unique combination of generalizability and accuracy across diverse datasets positions Se-PSD as a promising solution for real-time parking-slot detection in various applications, including autonomous driving, and smart parking management systems.

3. Method

3.1. Problem description

The parking-slot detection problem aims to identify and localize parking-slots within a sequence of images. Formally, given an image sequence $\mathrm{S_{image}} = \{\mathrm{I_1}, \mathrm{I_2}, \cdots, I_n\}$, where I_i represents the i-th image in the sequence and n is the total number of images (n = 6 in this paper), the objective is to predict the position and direction of each parking slot within the final image, I_n. This information, along with established geometric rules, allows for the final detection and localization of parking slots within the last image (I_n) of the sequence.

Figure 2. Two shapes of directional marking points.

3.2. Directional marking point

Like DMPR-PS [16], Se-PSD relies on directional marking points to identify parking-slots. Directional marking points are represented as $P = \{x, y, s, \theta\}$, where x, y is the coordinate of the marking point, s is the shape of the marking point, and θ is the direction of the marking point. There are two shapes of directional marking points, "T" and "L", corresponding to two different corner points of the parking-slot

DOI: 10.54254/2755-2721/75/20240522

as shown in Figure 2. By incorporating these attributes, directional marking points provide rich contextual details about the parking-slot, enabling Se-PSD to effectively detect parking-slots.

3.3. Se-PSD Model

The Se-PSD model is a sequence-based approach for parking-slot detection. The Se-PSD model consists of two main components (as shown in Figure 3): a CNN-based image feature extractor and a LSTM-based sequence feature extractor. The image feature extractor is a pretrained CNN model, which is used to extract features from the input image sequence S_{image}. The sequence feature extractor is based on LSTM network, which is capable of processing sequence inputs and capturing temporal dependencies in the sequence data. And then the feature map is regressed to the position, shape, and direction of parking-slot marking points through three fully connected layers. Finally, the parking-slots are infered through the geometric relationship between parking-slot marking point-pairs.

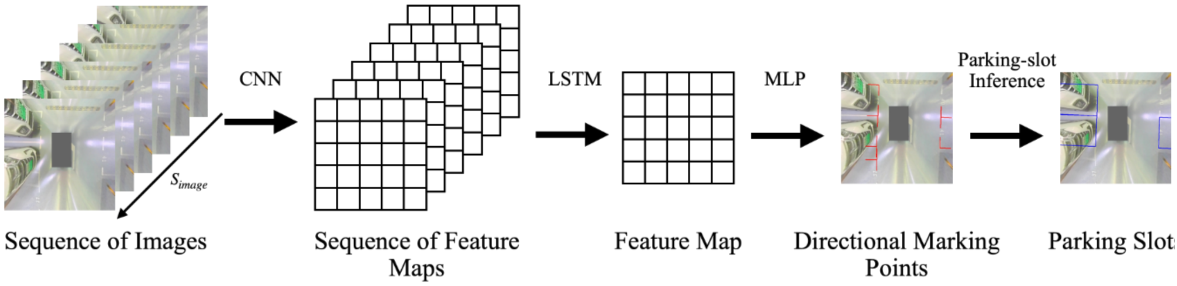

Figure 3. The pipeline of our proposed method Se-PSD.

3.4. Training

We implemented Se-PSD using the PyTorch deep learning framework and leveraged an NVIDIA A100 GPU for efficient model training. To optimize the training process and achieve optimal convergence, we employed the Adam optimizer. This powerful optimizer dynamically adjusts the learning rate for each parameter during training, ensuring efficient exploration of the search space and facilitating faster convergence. Initially, the learning rate starts at a low value of 0.0001, allowing the model to fine-tune on the initial parameters effectively. Following 10 epochs (training iterations), the learning rate is gradually increased to 0.0005, encouraging exploration of the parameter space and potentially leading to improved performance. After this initial increase, the learning rate undergoes a controlled decay of 50% every 20 epochs. This decay schedule prevents overfitting by gradually reducing the influence of each update as training progresses, allowing the model to focus on learning generalizable features. The entire training process culminates in a total of 90 epochs, ensuring sufficient exposure to the training data and allowing the model to learn robust representations.

To further enhance Se-PSD's generalization capabilities and improve its robustness to diverse real-world scenarios, we incorporated various data augmentation techniques during training. These techniques artificially expand the training dataset by generating additional variations of existing training images, effectively forcing the model to learn more generalizable features, and reducing the risk of overfitting to the specific training data. The employed data augmentation techniques included: Random horizontal flipping, Random vertical flipping, Normal and reversed image sequences, Replication for discontinuous images (for scenarios involving discontinuous image sequences, individual images were replicated n times to create artificial sequences of length n).

By combining the Adam optimizer with a carefully designed learning rate schedule and comprehensive data augmentation techniques, we ensured efficient training of the Se-PSD model, fostering its ability to learn robust and generalizable features for accurate parking-slot detection in real-world environments.

DOI: 10.54254/2755-2721/75/20240522

4. Experimental results and discussion

To evaluate the performance of Se-PSD, we employed a comprehensive validation strategy. After training the model on the benchmark ps2.0 dataset's training set, we assessed its accuracy on both the ps2.0 testing set and the BeVIS underground parking SLAM dataset, an open-source resource commonly used for parking-slot detection evaluation. We compared Se-PSD against several established parking-slot detection methods on both datasets.

On the ps2.0 dataset, all methods achieved accuracy exceeding 98%, demonstrating the effectiveness of the existing approaches on this controlled dataset. However, the true test lies in generalizability to unseen scenarios. When evaluated on the BeVIS dataset, which presents more diverse and challenging conditions, the performance of several methods, including DMPR-PS, understandably declined in terms of accuracy and recall. Notably, Se-PSD maintained a significant accuracy advantage over DMPR-PS on BeVIS, highlighting its superior generalization capability. This observation strongly supports the notion that image sequences inherently hold more information than individual images, and Se-PSD's ability to effectively extract temporal features from these sequences contributes to its robust performance across diverse parking environments.

Table 1. Performance comparison of parking-slot detection in ps2.0 test set and BeVIS.

Method	Dataset			
	ps2.0 [15]		BeVIS [16]	
	Precision	Recall	Precision	Recall
PSD_L [14]	98.41%	86.96%	95.22%	83.94%
DeepPS [15]	98.99%	99.13%	95.76%	93.43%
DMPR-PS [16]	**99.42**%	**99.37**%	95.88%	94.14%
Se-PSD	99.40%	99.36%	**98.28**%	**97.56**%

We evaluated the processing speed of Se-PSD using PyTorch on an NVIDIA A100 GPU. The average processing time for a single image sequence was approximately 50 milliseconds (ms). Interestingly, when processing a continuous stream of image sequences, the overlapping nature of convolutional neural network (CNN) feature extraction across sequences leads to a significant efficiency gain. In this scenario, the average processing time per sequence drops to 20ms, which is roughly three times faster than DMPR-PS, which requires 6ms per image under the same hardware configuration. It's important to note that real-time parking-slot detection requirements are typically not exceptionally strict, and Se-PSD comfortably satisfies these requirements even when processing individual sequences.

5. Conclusion

This paper introduces Se-PSD, a novel parking-slot detection method that leverages sequential image data. A key feature of Se-PSD is its incorporation of a Long Short-Term Memory (LSTM) network, enabling it to extract temporal information from image sequences. Unlike prior methods that rely solely on individual frames, Se-PSD offers improved generalization performance while maintaining accuracy on par with existing techniques. While Se-PSD does incur a slight trade-off in real-time processing speed, the relaxed real-time constraints in autonomous parking applications render this a manageable compromise. This makes Se-PSD a promising solution for real-world deployment.

However, we acknowledge that Se-PSD currently faces limitations in handling heavily obstructed parking-slots and detecting slanted parking slots. These challenges will be addressed in future work, which will involve exploring more robust feature extraction techniques and incorporating additional image analysis methods to achieve comprehensive parking-slot detection under various conditions.

DOI: 10.54254/2755-2721/75/20240522

References

[1] M. Wada, K. S. Yoon, and H. Hashimoto. 2003. Development of advanced parking assistance system. IEEE Transactions on Industrial Electronics, vol. 50, no. 1, pp. 4–17.
[2] W. Park, B. Kim, D. Seo, D. Kim, and K. Lee. 2008. Parking space detection using ultrasonic sensor in parking assistance system. in 2008 IEEE Intelligent Vehicles Symposium, pp. 1039–1044.
[3] J. Zhou, L. E. Navarro-Serment, and M. Hebert. 2012. Detection of parking spots using 2d range data. 15th International IEEE Conference on Intelligent Transportation Systems, pp. 1280–1287.
[4] R. Dube ́, M. Hahn, M. Schu ̈tz, J. Dickmann, and D. Gingras. 2014.Detection of parked vehicles from a radar-based occupancy grid. 2014 IEEE Intelligent Vehicles Symposium Proceedings, pp. 1415–1420.
[5] C. Wang, H. Zhang, M. Yang, X. Wang, L. Ye, and C. Guo. 2014. Automatic parking based on a bird's eye view vision system. Advances in Mechanical Engineering, vol. 6, p. 847406. https://doi.org/10.1155/2014/847406
[6] Z. Zhang. 2000. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330–1334.
[7] J. Xu, G. Chen, and M. Xie. 2000. Vision-guided automatic parking for smart car. Proceedings of the IEEE Intelligent Vehicles Symposium 2000 (Cat. No.00TH8511), pp. 725–730.
[8] H. G. Jung, Y. H. Lee, and J. Kim. 2010. Uniform user interface for semiautomatic parking slot marking recognition. IEEE Transactions on Vehicular Technology, vol. 59, no. 2, pp. 616–626.
[9] X. Du and K. K. Tan. 2015. Autonomous reverse parking system based on robust path generation and improved sliding mode control. IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 3, pp. 1225–1237.
[10] M. Sonka, V. Hlavac, and R. Boyle. 2007. Image Processing, Analysis, and Machine Vision. Thomson-Engineering.
[11] R. Grompone, J. Jakubowicz, J. Morel, and G. Randall. 2012. Lsd: a line segment detector. Image Processing On Line, vol. 2, pp. 35–55. https://doi.org/10.5201/ipol.2012.gjmr-lsd.
[12] S. Lee and S.-W. Seo. 2016. Available parking slot recognition based on slot context analysis. IET Intelligent Transport Systems, vol. 10, no. 9, pp. 594–604.
[13] C. Harris and M. Stephens. 1988. A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference, pp. 147–151.
[14] L. Li, L. Zhang, X. Li, X. Liu, Y. Shen, and L. Xiong. 2017. Vision-based parking-slot detection: a benchmark and a learning-based approach. 2017 IEEE International Conference on Multimedia and Expo (ICME), pp. 649–654.
[15] L. Zhang, J. Huang, X. Li, and L. Xiong. 2018. Vision-based parking-slot detection: a dcnn-based approach and a large-scale benchmark dataset. IEEE Transactions on Image Processing, vol. 27, no. 11, pp. 5350–5364.
[16] J. Huang, L. Zhang, Y. Shen, H. Zhang, S. Zhao, and Y. Yang. 2019. Dmpr-ps: a novel approach for parking-slot detection using directional marking- point regression," in 2019 IEEE International Conference on Multimedia and Expo (ICME) , pp. 212–217.
[17] J. L. Elman. 1990. Finding structure in time. Cognitive Science, vol. 14, no. 2, pp. 179–211.
[18] S. Hochreiter and J. Schmidhuber. 1997. Long short-term memory. Neural Comput., vol. 9, no. 8, p. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
[19] X. Shao, Y. Shen, L. Zhang, S. Zhao, D. Zhu, and Y. Zhou. 2023. Slam for indoor parking: a comprehensive benchmark dataset and a tightly coupled semantic framework. ACM Trans. Multimedia Comput. Commun. Appl., vol. 19, no. 1. https://doi.org/10.1145/3510856