A Self-supervised Parking Spot Monitoring System using Google Coral Edge TPU

Dawson Drake and Wei Tang
Klipsch School of Electrical and Computer Engineering
New Mexico State University, Las Cruces, NM, USA
wtang@nmsu.edu

***Abstract*—This paper reports a vision-based parking spot monitoring system implemented on Google Coral Edge TensorFlow Processing Unit. The system utilizes videos from available surveillance cameras to automatically learn the parking spot locations after a few parking events in each spot. Then it can detect vacant and occupied parking spots in the parking lot in real time. This is achieved using the combination of temporal difference images and machine learning models. The temporal difference image detects moving objects while the machine learning models identify potential vehicles. By combining the two methods with size and speed filters, the system is able to successfully detect the parking and leaving events for each parking spot near the camera with an accuracy of 92.31%. Without additional information, the system also learns newly available parking spots and unavailable prior parking spots by calculating the confidence level based on recent records of parking events. The system is also able to reconstruct the 3D environment of the parking lot. This method greatly reduces the efforts of manual assignments of parking spots in prior reported systems.**

***Index Terms*—Self-supervised, Temporal Difference, Parking Spot Monitoring, Machine Learning**

I. Introduction

Detecting vacant parking spots has been an important application in automobile transportation. Drivers can save a lot of time if they can be directed to the nearest available parking spots using path-finding algorithms [1]–[3]. Sensing technologies have been actively studied and applied in indoor parking garages or buildings [4], [5]. For example, in [6] and [7] the authors propose systems that utilize ultrasonic sensors to detect whether a parking space is vacant or occupied. However, considering the cost of deployment and maintenance, it is difficult to install individual sensors for each parking space when the size of the parking lot is large. Considering that most open-space parking lots have equipped surveillance cameras, an image-based parking spot monitoring system is the best choice for detecting vacant parking sport in such parking lots.

The image-based parking spot monitoring system usually requires a manual drawing of each parking space in the parking lot before deployment [8], [9]. Then the system analyzes the key features of cars, such as windows or wheels in the image to decide if a certain parking spot is occupied [10]–[12]. The primary challenge in such an image-based parking spot monitoring system comes from the complexity of the parking environment. First, each parking lot has a different layout, so the system is subject to manual adjustment. Second, the image contains many moving subjects such as the shade of the trees, clouds, pedestrians, or birds flying by the camera. Third, due to the finite height of the monitoring camera, the parking spot monitoring system may fail when a small vehicle is partially or fully blocked by a large vehicle in the image. Moreover, when certain parking spots become unavailable or new parking spots are created in the parking lot, the system may need new calibration, which is also possibly necessary if the camera was moved by wind or other natural forces. The new calibration may cost additional time and resources if the system requires frequent updates.

In order to alleviate such problems, we proposed a new method combining temporal difference image sensing [13] and machine-learning-based vehicle detection. The machine learning model identifies vehicles while the temporal difference image sensing [14], [15] tracks moving objects [16], [17] and the moment when a vehicle stops or starts moving. In addition, 3D reconstruction is applied to better assist the parking spot monitoring system. We also inserted size and speed filters to reduce the effects of other moving objects such as motorcycles, pedestrians, shades, and animals. Furthermore, we introduce a confidence level to decide if the layout of the parking lot changes. Since our system monitors the parking lot and generates a reconstruction based on what it observes, there is no need for any prior knowledge of the parking space layout to be provided to the system. This makes the system self-supervised so it does not require initial manual drawing and can adapt adjustments automatically. These methods help to reduce the frequency calibration or adjustment of the system, which allows for more flexibility in the environment.

II. System Design

The proposed system utilizes video inputs from the surveillance camera of the parking lot. The input image is processed by the machine learning model and temporal difference method. Figure 1 shows a general flow diagram of the system's reconstruction process. As seen in the figure the system consists of three main parts including vehicle detection, vehicle tracking, and the 3-D reconstruction of the parking lot. The first step is to process the video feed using temporal difference imaging and a machine learning model which results in a set of bounding boxes around vehicles that are determined to be moving. The results of these two processes are then filtered based on metrics such as the bounding box size and shape to remove any extraneous detection. The bounding boxes from the detection step are then used in the vehicle tracking step to

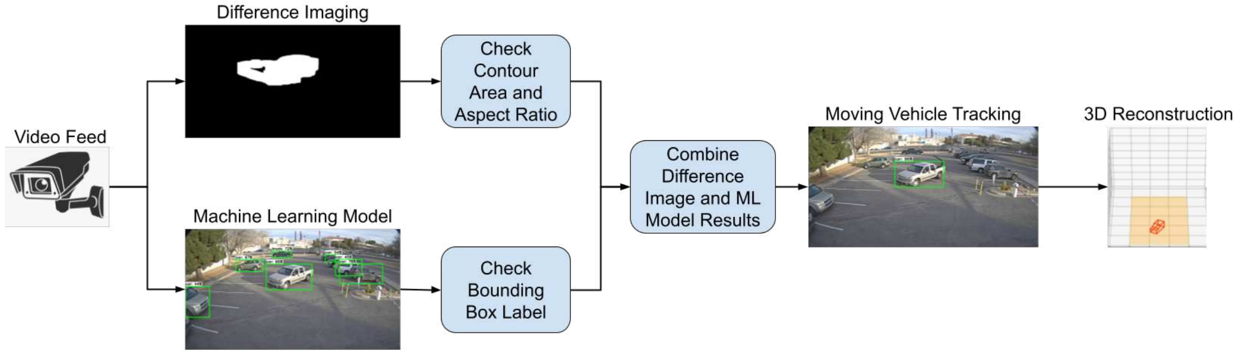

Fig. 1. System Sensing and Signal Processing Flow Diagram.

determine vehicle locations and generate the path the vehicle has traveled through the parking lot. The final step is to use the vehicle location and path information to identify possible parking space locations and reconstruct the parking lot in a digital 3-D environment. Detailed algorithms in the system are described in the following subsections.

A. *Machine Learning Model*

The machine learning models are used to process the video frames and perform object detection to obtain information about vehicles, especially the bounding box locations in the image. We tested several pre-trained models including the *SSD(Single Shot Multibox Detector) Mobilenet v1* model along with the Tensorflow API, the *Mask R-CNN for Object Detection and Segmentation* model, and the *YOLO-tiny v4* [18] model. The testing was performed using a large sample of parking lot images. We choose *Mobilenet v1* model since it is compatible with the Edge TPU and provides satisfactory object detection accuracy. We applied a size filter based on the size of the bounding boxes generated by the machine-learning model to exclude remote vehicles. If the bounding box pixel area was below a certain threshold then the boxes would be discarded, leaving only the vehicles that can be seen clearly for the detection process. We assume that there will be another camera that can view remote spaces more clearly. These cameras can then work together to detect all parking spaces in the parking lot. The machine learning step in Fig. 1 shows the results of processing one frame to detect the objects in the frame and performing the filtering described above. Since the machine learning model looks at each frame individually, it has no knowledge of how each vehicle is moving through the frame, so both moving and parked vehicles are detected.

B. *Temporal Difference Image Processing*

To learn about parking spots, the system needs to distinguish between moving and parked vehicles and track only the vehicles which are moving. This is achieved using temporal difference image processing between the consecutive images in the video. The temporal difference image is calculated by subtracting the previous frame from the current frame. If the difference between the previous and current values from the same pixel location reaches a pre-defined threshold, the pixel is marked as white, else the pixel is left as black. The temporal difference image is good at drawing the contour of the moving object but may generate noise pixels from the background. An erosion filter is applied to remove the background noise pixels. The goal of using a temporal difference image is to find the moving vehicle. However, when computing the difference image, not all of the pixels that make up the vehicle change value due to most cars being a solid color, therefore the entire vehicle is not always part of the difference image. This causes the issue of a single vehicle possibly being segmented into multiple regions. Therefore, dilation filters are applied to remove these noise pixels and amalgamate the temporal difference pixels that belong to the same moving vehicle.

Similar to the filtering process for the machine learning step, the system also performs basic filtering in the difference imaging step. First, the area of all the contours is calculated and any contour with an area below a certain threshold is ignored. This removes any small objects that may be moving in the frame as well as any noisy pixels that may have caused a contour to appear. The second filter examines the aspect ratio of each contour. The main purpose of this filter is to remove pedestrians moving throughout the frame. In general, a vehicle has an aspect ratio where the width is greater than or equal to the height depending on the vehicle orientation, while pedestrians have a much narrower aspect ratio where the height is larger than the width.

C. *Combination of Detection Methods*

By combining machine learning detection and temporal difference processing, the system detects moving vehicles, learns the parking spot, and monitors if a certain parking spot is occupied. The machine learning model is good at distinguishing between different objects. However, it cannot determine the moving objects from one frame. Temporal difference imaging on the other hand cannot distinguish between different objects but can easily determine where something is moving in the frame. One important issue of how these methods work together is by looking at the shadow of a vehicle. As shown in Fig. 1 the temporal different imaging detects the shadow of the truck since it moves as the truck moves. This would be an issue if we only used the temporal different imaging since tracking the shadow of the vehicle would cause the detection to be inaccurate. With help from the machine learning model, the system can more accurately detect the moving truck without including its shadow.

In order to identify the bounding boxes that are moving, the system computes the percentage of white pixels from the

	Shih, et al. [5]	Amato, et al. [19]	Patel, et al. [20]	Ramasamy, et al. [6]	This Work
Detection Method	Background Subtraction	Convolutional Neural Network	Machine Learning Object Detection	Sensor Polling	Machine Learning and Difference Imaging
Equipment Used	Wide Angle Camera	Surveillance Camera	Surveillance Camera	Multiple Ultrasonic Sensors	Surveillance Camera
Accuracy	99.67 %	91.1 %	97.69 %	-	92.31 %

TABLE I
COMPARISON WITH OTHER PUBLICATIONS

temporal difference imaging contained within each bounding box from the machine learning models and compared the value to the total number of pixels in the box. Only when the value is above a certain threshold (80%), the system keeps the box for the vehicle tracking step.

D. Vehicle Tracking and 3D Reconstruction

Once the bounding boxes of the moving vehicles have been obtained, the system starts tracking the moving vehicles in the parking lot. A path is generated when a vehicle is detected and then the path is extended as the vehicle moves. The system computes the distance between the most recent path position and the center of the bound boxes. If a path is within a specified distance of one of the vehicles in the frame, the system concludes that the path belongs to the current vehicle. If there is not a vehicle within this range then the system assumes that the vehicle has stopped. To determine if a vehicle has parked, each path is checked to see how long it has been since it was last updated. If the path has not changed in a given time frame then the system considers this location to be where the vehicle has parked. The vehicle tracking results can be used to generate a real-time 3D reconstruction of the parking lot, as shown in Fig. 1. Assuming the parking lot is a plane and all vehicles move on the same plane along with the fact that the vehicle does not change its physical size during movement, the system can then obtain an estimation of the plane of the parking lot.

III. System Testing and Results

Both temporal difference imaging and machine learning models were implemented on a Google Coral Edge TPU board as shown in Fig. 2. In order to develop and test our system, we were able to obtain prerecorded security camera footage of a parking lot on a university campus. The footage provided us with several video clips of vehicles entering and parking within the parking lot. We selected the footage to include a variety of different environments such as different times of day, different weather conditions, activity levels, and vehicle types, to ensure the system performed nominally in each situation. The system is able to accurately reconstruct the parking lot in a 3D environment. Table I presents the detection accuracy of the parking spot and compares the result with recent references. Our system achieved acceptable accuracy with low hardware complexity, which enables the implementation of Edge devices.

	Count	Rate
True Positive	516	92.307%
True Negative	257	99.227%
False Positive	2	0.772%
False Negative	43	7.692%

TABLE II
SYSTEM DETECTION RESULTS.

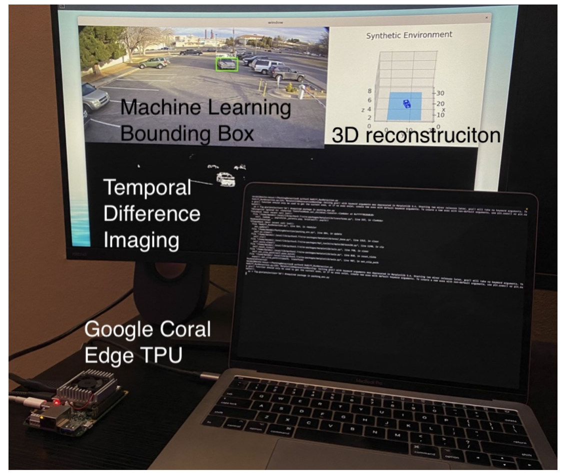

Fig. 2. Experimental Setup using Google Coral Edge TPU.

Table II shows the detection metrics of this method on 819 frames. The proposed method proves to be effective in correctly detecting moving vehicles having a true positive rate(sensitivity) of 92.307% and a true negative rate(specificity) of 99.227%. The system keeps track of the number of detection for each space in order to adjust the confidence level that the detected space is actually a parking space. The confidence level for each space is also affected by other factors such as the amount of time since the last detection for that space. Combining information together, the system is able to dynamically update the parking lot layout as it changes over time. Not only does the confidence level give the system a way to determine any false detections, but it also gives the ability to remove spaces that may be currently unavailable for any car to park in. For example, a space may become unavailable if it is under construction or if it is used for another reason besides parking. Since our system tracks how many vehicles have parked in a space and how long it has been since the last detection in that space, it can determine the likelihood a space is truly available.

IV. Conclusion

This paper presents a vision-based open-space parking lot monitoring system that detects occupied and available parking spots using videos from surveillance cameras. The system combines the advantages of the temporal difference image sensing and machine learning object detection. A series of image filters are implemented to increase the detection accuracy. The system is also able to reconstruct the real-time 3D environment based on vehicle tracking results. The detection accuracy is 92.31% which is comparable with other references. The self-supervised feature saves time for manual calibration in initial deployment and maintenance.

References

[1] Z. Liu, Y. Yang, D. Li, X. Li, X. Lv, and X. Chen, "Design and implementation of the optimization algorithm in the layout of parking lot guidance," in *2020 16th International Conference on Computational Intelligence and Security (CIS)*, 2020, pp. 144–148.

[2] R. Chen, X. Hu, and W. Mu, "Research on parking lot management system based on parking space navigation technology," in *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, 2020, pp. 773–777.

[3] K. Kaarthik, A. Sridevi, and C. Vivek, "Image processing based intelligent parking system," in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, 2017, pp. 1–4.

[4] K. Choeychuen, "Available car parking space detection from webcam by using adaptive mixing features," in *2012 Ninth International Conference on Computer Science and Software Engineering (JCSSE)*, 2012, pp. 12–16.

[5] S.-E. Shih and W.-H. Tsai, "A convenient vision-based system for automatic detection of parking spaces in indoor parking lots using wide-angle cameras," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2521–2532, 2014.

[6] M. Ramasamy, S. G. Solanki, E. Natarajan, and T. M. Keat, "Iot based smart parking system for large parking lot," in *2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA)*, 2018, pp. 1–4.

[7] R. Vishnubhotla, P. S. Rao, A. Ladha, S. Kadiyala, A. Narmada, B. Ronanki, and S. Illapakurthi, "Zigbee based multi-level parking vacancy monitoring system," in *IEEE International Conference on Electro/Information Technology*, 2010, pp. 1–4.

[8] R. Fusek, K. Mozdřeň, M. Šurkala, and E. Sojka, "Adaboost for parking lot occupation detection," in *Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013*, R. Burduk, K. Jackowski, M. Kurzynski, M. Wozniak, and A. Zolnierek, Eds. Heidelberg: Springer International Publishing, 2013, pp. 681–690.

[9] J. Liu, M. Mohandes, and M. Deriche, "A multi-classifier image based vacant parking detection system," in *2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS)*, 2013, pp. 933–936.

[10] H. T. Vu and C.-C. Huang, "A multi-task convolutional neural network with spatial transform for parking space detection," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1762–1766.

[11] C.-C. Huang and H. T. Vu, "Vacant parking space detection based on a multilayer inference framework," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 9, pp. 2041–2054, 2017.

[12] J. Nyambal and R. Klein, "Automated parking space detection using convolutional neural networks," in *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, 2017, pp. 1–6.

[13] Y. Liu, X. Yu, S. Chen, and W. Tang, "Object localization and size measurement using networked address event representation imagers," *IEEE Sensors Journal*, vol. 16, no. 9, pp. 2894–2895, 2016.

[14] I. White, E. Curry, D. K. Borah, S. J. Stochaj, and W. Tang, "An optical spatial localization algorithm using single temporal difference image sensor," *IEEE Sensors Letters*, vol. 3, no. 3, pp. 1–4, 2019.

[15] I. White, D. K. Borah, and W. Tang, "Robust optical spatial localization using a single image sensor," *IEEE Sensors Letters*, vol. 3, no. 6, pp. 1–4, June 2019.

[16] H. Stuckey, A. Al-Radaideh, L. Escamilla, L. Sun, L. G. Carrillo, and W. Tang, "An optical spatial localization system for tracking unmanned aerial vehicles using a single dynamic vision sensor," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 3093–3100.

[17] H. Stuckey, A. Al-Radaideh, L. Sun, and W. Tang, "A spatial localization and attitude estimation system for unmanned aerial vehicles using a single dynamic vision sensor," *IEEE Sensors Journal*, vol. 22, no. 15, pp. 15 497–15 507, 2022.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[19] G. Amato, F. Carrara, F. Falchi, C. Gennaro, and C. Vairo, "Car parking occupancy detection using smart camera networks and deep learning," in *2016 IEEE Symposium on Computers and Communication (ISCC)*, 2016, pp. 1212–1217.

[20] R. Patel and P. Meduri, "Car detection based algorithm for automatic parking space detection," in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2020, pp. 1418–1423.