A Systematic Review on Computer Vision-Based Parking Lot Management Applied on Public Datasets

Paulo Ricardo Lisboa de Almeida[a,*], Jeovane Honório Alves[a],
Rafael Stubs Parpinelli[b], Jean Paul Barddal[c]

[a]*Department of Informatics, Federal University of Paraná (UFPR), Curitiba (PR), Brazil*
[b]*Graduate Program in Applied Computing, Santa Catarina State University (UDESC),
Joinville (SC), Brazil*
[c]*Graduate Program in Informatics (PPGIa), Pontifícia Universidade Católica do Paraná
(PUCPR), Curitiba, Brazil*

Abstract

Computer vision-based parking lot management methods have been extensively researched upon owing to their flexibility and cost-effectiveness. To evaluate such methods authors often employ publicly available parking lot image datasets. In this study, we surveyed and compared robust publicly available image datasets specifically crafted to test computer vision-based methods for parking lot management approaches and consequently present a systematic and comprehensive review of existing works that employ such datasets. The literature review identified relevant gaps that require further research, such as the requirement of dataset-independent approaches and methods suitable for autonomous detection of position of parking spaces. In addition, we have noticed that several important factors such as the presence of the same cars across consecutive images, have been neglected in most studies, thereby rendering unrealistic assessment protocols. Furthermore, the analysis of the datasets also revealed that certain features that should be present when developing new benchmarks, such as the availability of video sequences and images taken in more diverse conditions, including nighttime and snow, have not been incorporated.

*Corresponding author
  Email addresses: `paulo@inf.ufpr.br` (Paulo Ricardo Lisboa de Almeida),
`jhalves@inf.ufpr.br` (Jeovane Honório Alves), `rafael.parpinelli@udesc.br`
(Rafael Stubs Parpinelli), `jean.barddal@ppgia.pucpr.br` (Jean Paul Barddal)

---

1. Introduction

Over the last few years, many authors proposed computer vision-based approaches to address problems related to parking lot management. These problems focus on processing images from parking lots. They include different goals, such as (i) the automatic detection of parking spaces positions, e.g., defining for each parking space the bounding box that delimits the object, (ii) the individual parking space classification, determining whether a specific parking spot is occupied by a vehicle or not, and (iii) detecting and counting vehicles in images. The tasks mentioned above are often the core components of Smart Parking solutions. They aim at providing, among others, automated parking lot management, e.g., dynamic pricing according to the number of cars in the parking lot; and parking guidance for drivers, e.g., a route to the nearest parking space available. Smart Parking solutions are essential as a system that accurately guides the driver to the nearest parking spot available can save both time and fuel (Polycarpou et al., 2013; Paidi et al., 2018).

In this study, publicly available parking lot image datasets were surveyed. In addition, the computer vision-based works that have employed such datasets to address parking lot management problems were evaluated. The scope of this study was limited to computer vision-based approaches as they are advantageous over individual sensors. For example, in contrast to magnetometers and ultrasonic sensors, a single camera can monitor a wide parking area eliminating the need of requiring a sensor per parking spot. Furthermore, cameras reduce installation and maintenance costs and can aid in additional tasks, such as abnormal behavior and theft detection (Paidi et al., 2018; Li et al., 2019; Varghese & Sreelekha, 2020).

To the best of our knowledge, this work is the first dataset-centered review of computer vision-based approaches to address problems related to parking lot

management. It is relevant to mention that sensors other than cameras may be used for automatic parking lot management. Such sensors are beyond the scope of this work. Recent comprehensive reviews of different sensors and approaches for managing parking lots can be found in (Polycarpou et al., 2013; Fraifer & Fernström, 2016; Paidi et al., 2018; Barriga et al., 2019).

Reproducibility is among the most important guidelines to be followed in any research. Thus, this review focused on works that utilize at least one robust and publicly available parking lot image dataset, thereby increasing the reproducibility of experiments performed. Moreover, we also propose a criterion that the parking lot image datasets must encompass real-world challenges, avoiding trivial and unrealistic problems. Owing to this filtering process, the selected datasets that fulfilled the this criterion were further described and compared. We expect this review to aid researchers in (i) the development of new computer vision-based parking lot management methods, and (ii) with the proposal of novel robust datasets that can be employed to validate such methods. Furthermore, we reviewed the existing vision-based approaches, which use at least one of the surveyed datasets, addressing the following problems:

- Classification of individual parking spaces;

- Automatic detection of parking spaces positions;

- Vehicles counting or detection;

As the reviewed approaches use publicly available datasets, these are easier to compare and verify. Moreover, the surveyed approaches were also categorized and compared in this work. Consequently, the results obtained were used to identify well-studied solutions for problems such as the individual parking slots classification. In addition, research gaps that researchers could further investigate, for example, the automatic detection of parking spaces and problems generated by camera angle changes, were also identified. The complete contributions of this work are as follows:

- We propose certain criteria to define a parking lot image dataset as robust;

- We bring forward a review of existing robust parking lot image datasets;

- We review, categorize and compare the results of state-of-the-art works that use the surveyed datasets;

- After analyzing the state-of-the-art methods and results, we identify the research gaps that researchers should address in the future.

This paper is divided as follows. Section 2 brings forward a discussion on existing surveys and their main shortcomings, which are used as guidance for determining our research method. The research procedure used in this work, including the criteria established to select the public parking lot images studies and datasets, is brought forward in Section 3. Details about the selected datasets (PKLot, CNRPark-EXT, and PLds) are given in Section 4. By following the research method presented in Section 3, we found 66 works that use the datasets mentioned above. These works are presented in Section 5, where the approaches are categorized according to the following tasks: individual parking spot classification, automatic parking space detection, and car detection and counting. In Section 6, we summarize and discuss the datasets and the reviewed works. Furthermore, we also present our findings, such as that most authors focus on the classification of individual parking spaces. Finally, the conclusions and envisioned future works are presented in Section 7.

2. Motivation

In this section, we outline the research method used. To justify this research method and the scope of our work, we first highlight and discuss nine surveys and reviews that cite the datasets analyzed in this work (Meduri & Estebanez, 2018; Mahmud et al., 2020; Paidi et al., 2018; Enríquez et al., 2017; Kawade et al., 2020; Barriga et al., 2019; Chen et al., 2019; Zantalis et al., 2019; Diaz Ogás et al., 2020). In Meduri & Estebanez (2018), a brief review of Deep Learning (DL) based approaches for parking lots is given. Nevertheless, the authors considered only six works. They stated how questionable the generalization of

4

these works is since the samples' conditions, e.g., weather, lightning, occlusions, and car size, are reasonably similar. The authors also point out that no solution detected the parking spaces automatically. Mahmud et al. (2020) surveyed some works related to the individual parking spaces classification problem and, similarly to Meduri & Estebanez (2018), concluded that generalization issues and automatic parking spot detection are open problems.

Most reviews focus on discussing the different sensors available for individual parking space monitoring and user software, e.g., smartphone applications, developed to manage the parking lots or guide drivers to the parking spaces (Paidi et al., 2018; Enríquez et al., 2017; Kawade et al., 2020; Barriga et al., 2019). The sensors often surveyed in these works include ultrasonic, Radio-Frequency Identification (RFID), magnetometers, microwave radars, and camera sensors.

The authors in Paidi et al. (2018) claim that individual parking spaces monitoring in open areas is still an open problem in the literature. In Enríquez et al. (2017) is presented a survey about infrastructure-based, vision-based, and crowd-sensing-based solutions for on- and off-street solutions. They argued that vision-based systems have problems with illumination, occlusions, and generalization capabilities. On the other hand, infrastructure-based solutions have higher costs. In contrast, crow-sensing solutions require many contributors, which are not always available.

In Kawade et al. (2020), a survey is presented, including hardware, i.e., ultrasonic and infrared sensors, and computer vision-based works. However, only a limited number of papers were analyzed. They stated that installation and maintenance were problems found in sensor-based works, as occlusion was in computer vision-based ones. DL-based approaches for smart cities' problems were surveyed in Chen et al. (2019). These problems include human mobility, traffic flow, traffic surveillance, and parking lot management. The authors conclude that, besides its importance, DL-based solutions suffer from problems such as high computational cost and the lack of training datasets. Likewise, Zantalis et al. (2019) surveyed different techniques for smart city problems, including Smart Parking. In Diaz Ogás et al. (2020), the authors present a systematic re-

view of different types of Smart Parking systems, such as guidance, reservation, and crowdsourcing.

The works of Enríquez et al. (2017); Paidi et al. (2018); Barriga et al. (2019); Chen et al. (2019); Zantalis et al. (2019); Diaz Ogás et al. (2020); Kawade et al. (2020) offer a broader vision of the parking lot management systems, including different sensors and user-level software. However, these works do not present an in-depth discussion about vision-based parking lot management approaches. In addition, they did not discuss publicly available datasets developed to verify these approaches. The only surveys that have focused on vision-based parking lot management problems are Meduri & Estebanez (2018); Mahmud et al. (2020). Nevertheless, these works are not systematic reviews and considered only a few works in the analysis. Therefore, the systematic mapping of vision-based solutions that can be reproduced using public datasets proposed here has its relevance justified.

3. Research Procedure

In this section, we bring forward the research procedure adopted to conduct a systematic literature review on computer vision-based parking lot management systems and public datasets. First, we have defined the following research questions (RQs) to guide the identification and assessment of relevant works:

- RQ_1: Which are the primary parking lot management problems dealt with by computer vision-based state-of-the-art solutions?

- RQ_2: What are the primary computer vision-based techniques employed in the state-of-the-art to solve parking lot management problems?

- RQ_3: What are the open problems in computer vision-based techniques not covered by the state-of-the-art or that still require more research?

With the scope defined, we discuss the planning and conduction of the review. We first find the publicly available image parking lot datasets and then

6

collect related works. We applied three keywords in the well-known Scopus[1] and Web of Science (WoS)[2] peer-reviewed citation database search engines to accomplish this. The keywords are *parking lot dataset*, *parking lot images*, and *parking lot database*. We refined the search to include only the works that propose datasets containing parking lot images. Finally, we reviewed only robust datasets, which must fulfill the following criteria.

First, the datasets' publicity is the most critical restriction considered in this work. Public data enable researchers to reproduce experiments, results and create their experiments without collecting and labeling novel datasets. The cameras must be installed at fixed points since this is an expected feature in the real world (we do not consider, for instance, images collected via drone). The datasets must contain images collected on different days of the week, periods, and weather conditions to include the expected variability in a real scenario. Images collected from different camera angles are necessary to test classifiers' generalization power. For instance, a classifier is trained using images obtained from a camera angle different from the test images. Finally, the ground truth is imperative to check and compare the results when using the datasets. Although restrictive, we consider the established criteria important for advancing the research about vision-based parking lot management.

As a result, three datasets were selected: the PKLot (Almeida et al., 2013; Almeida et al., 2015), the CNRPark-EXT (Amato et al., 2016, 2017), and the Parking Lot dataset (PLds) (Martín Nieto et al., 2019). Certain datasets despite being interesting, do not meet these restrictions, including the QuickSpotDB (Màrmol & Sevillano, 2016), CARPK (Li et al., 2019) and UAVDT (Du et al., 2018) datasets.

Following the selection of datasets, we proceeded further by surveying works that used these datasets to evaluate vision-based parking lot management approaches. Therefore, we used the Scopus and WoS search engines to identify

---

[1]Scopus Website: www.scopus.com (Accessed on Aug 23, 2021).

[2]WoS Website: www.webofknowledge.com (Accessed on Aug 23, 2021)

cross-referenced works. Hence, we applied a snowballing approach (Wohlin, 2014), as the primary references of each dataset were used to find all the other works that cited each of them. These references are referred to as the primary references set. Further, the closure requirements, such as the objective, inclusion, and exclusion criteria, are as follows:

- Objective Criteria (OC)

    - Document type: Articles

    - Availability: Any

- Inclusion Criteria (IC)

    - IC_1: Include works that proposed the surveyed datasets.

- Exclusion Criteria (EX)

    - EX_1: Remove non-English works;

    - EX_2: Remove works that only mention the datasets as related work without using the datasets;

    - EX_3: Remove surveys or systematic reviews.

Based on the above rules, an initial number of 248 unique works were found. The IC_1 included 5 (Almeida et al., 2013; Almeida et al., 2015; Amato et al., 2016, 2017; Martín Nieto et al., 2019) works. The exclusion criteria EX_1, EX_2, EX_3 removed 10, 155, and 22 works, respectively. Thus, a total of 66 works were identified for analysis.

4. Parking Lot Datasets

In this section, we bring forward an analysis on existing parking lot datasets available in the literature, i.e., PKLot, CNRPark, CNRPark-EXT, and PLds.

4.1. PKLot Dataset

The first version of the PKLot dataset, containing images of one parking lot and one camera angle, was released in Almeida et al. (2013). The current version of the dataset was proposed by Almeida et al. (2015), and contains 12,417 images collected from two different parking lots and three camera angles. The images were collected using a camera installed on the fourth and fifth floors of one building from the Federal University of Parana (UFPR) to generate the UFPR04 and UFPR05 subsets. Both UFPR04 and UFPR05 subsets contain images from the same parking lot yet collected under different camera angles and different days. The third subset of parking lot images was taken from the 10th floor of a building from the Pontifical Catholic University of Parana (PUCPR) and present a different camera angle and parking lot.

(a) UFPR04: Rainy   (b) UFPR05: Sunny   (c) PUCPR: Cloudy

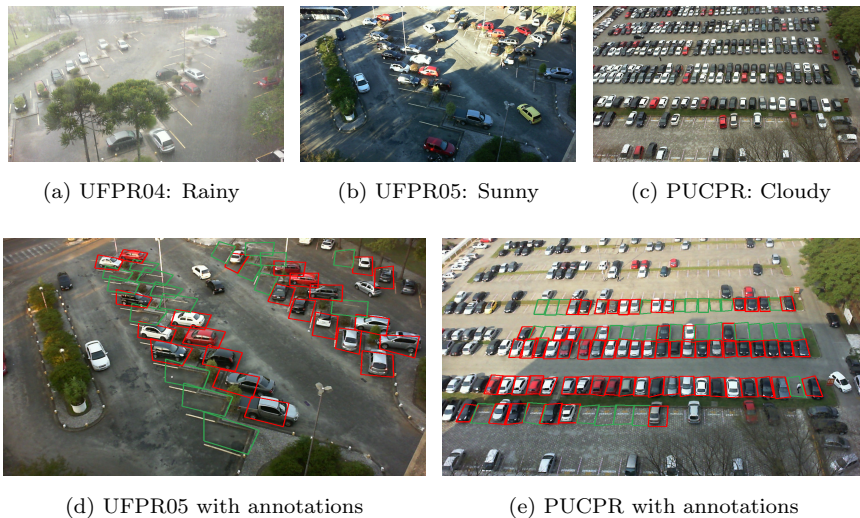(d) UFPR05 with annotations   (e) PUCPR with annotations

Figure 1: PKLot image examples. Figures (a), (b), and (c) show examples of different parking lots and weather conditions. Figures (d) and (e) show examples of images with the location and status (red for occupied and green for empty) of the parking spaces drawn.

The dataset contains 695,851 manually labeled individual parking spaces, such that 337,780 (48.6%) are occupied, and 358,071 (51.4%) are empty. Cropped images of the individual parking spaces are available altogether with the orig-

inal dataset. All images are 1280×720 pixels in size and stored in the JPEG format. An Extensible Markup Language (XML) file containing four points of polygons representing each monitored parking space is available for each image. Additionally, a rotated rectangle representing the same information but with right angles (making it easier to crop the images) is available. Moreover, each parking space' status (empty/occupied) is also available in the XML files.

Most of the parking spaces available in UFPR04 and UFPR05 parking lots are labeled in the dataset. In contrast, for the PUCPR, only 100 parking spaces were manually labeled. In comparison, approximately 300 parking spaces are visible to the human eye in such images. Images were collected under a 5 min interval during the daytime and labeled according to sunny, rainy, and cloudy weather. Figure 1 shows image examples from the PKLot dataset for different parking lots and weather conditions.

4.2. CNRPark and CNRPark-EXT Datasets

The CNRPark dataset, proposed in Amato et al. (2016), contains images collected from a single parking lot using two different camera angles. Images were taken considering a 5 min interval during daytime in sunny conditions on two different days for each camera angle. The dataset contains 12,584 parking spaces, where 4,181 (33.2%) are free, and 8,403 (66.8%) are occupied. Only the segmented parking spaces are publicly available, where all images were resized to 150×150 pixels. In addition to certain parking spaces being angled relative to the camera, the images were segmented using non-rotated squares, which may have resulted in the inclusion of undesired areas or exclusion of certain regions of the parking spaces or cars. Examples of the CNRPark dataset can be seen in Figure 2.

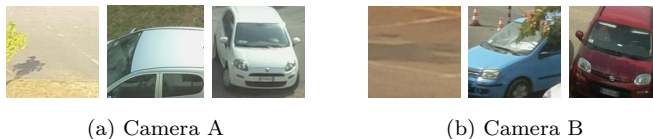(a) Camera A   (b) Camera B

Figure 2: CNRPark image examples.

The CNRPark was extended to create the CNRPark-EXT dataset in Amato et al. (2017). The authors used nine cameras to capture images from a single parking lot at different angles. Images were acquired at a 30 min interval. The cameras are not synchronized with each other. Similar to the PKLot dataset, the authors separated the images according to the weather. The images are 1000×750 pixels in size, with a total of 4,278 JPEG parking lot images available. The dataset also encompasses Comma-separated Values (CSV) files that indicate the monitored parking spaces' coordinates, represented by non-rotated squares (see Figures 3d and 3e). Cropped images of the individual parking spaces are also available, where the images were resized to 150×150 pixels. When considering the images of the original CNRPark altogether with the CNRPark-EXT datasets, there are 157,549 labeled parking spaces available, where 69,839 (44,3%) are free, and 87,710 (55,7%) are occupied.

The CNRPark-EXT is considered an extension of the original CNRPark dataset by the authors, and only Amato et al. (2016) used the CNRPark images without the CNRPark-EXT dataset. Thus, we have considered CNRPark- EXT concatenated with the original CNRPark as a single dataset. Figure 3 shows certain examples of the images contained in the CNRPark-EXT dataset. They were acquired from different camera angles and under different weather conditions, and have been presented along with certain annotated image examples.

4.3. PLds Dataset

The Parking Lot dataset (PLds) was proposed in Martín Nieto et al. (2019), and contains 1280×960 pixel images collected from three different camera angles at the Pittsburgh International Airport parking lot. Images were taken under several time intervals, ranging from a few seconds to several minutes, i.e., 15 s to 30 min[3]. Several climate conditions are available in the dataset, including snow and rain. Light conditions also include nighttime images.

---

[3]We considered the time tag available in the top left corner of the images.

(a) Camera 3: Rainy   (b) Camera 5: Sunny   (c) Camera 9: Overcast

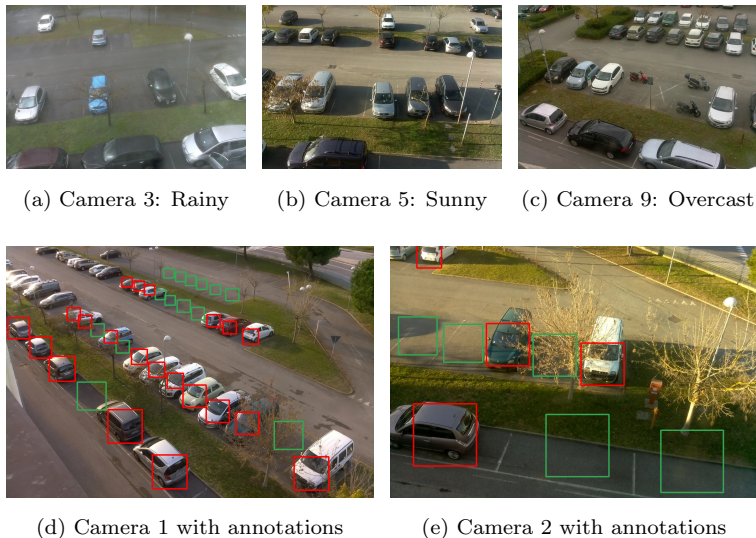(d) Camera 1 with annotations   (e) Camera 2 with annotations

Figure 3: CNRPark-EXT image examples. Figures (a), (b), and (c) show examples of different parking lots and weather conditions. Figures (d) and (e) show examples of images with the location and status (red for occupied and green for empty) of the parking spaces drawn.

The images are stored in JPEG format and the annotations for each image has been provided by the authors in XML files. In addition, non-rotated bounding boxes for parked cars are provided (annotations for empty parking spots are not available). An appealing feature of this dataset is that a subset of PLds containing 100 images is synchronized between two different camera angles. Image examples of the PLds dataset, including different weather conditions and annotated images, are shown in Figure 4.

The dataset contains 8,340 images[4]. In the original dataset, the images are not labeled according to the climate nor to luminosity conditions. As a byproduct contribution of this review, we manually classified the images between day/night and climate conditions (see the summary of the images after this classification in Section 6.1). We made the list of the images classified ac-

---

[4]We did not consider the sequence of images isshk_1955 to isshk_2146, and isshk_2598 to isshk_2681 since the images seem to repeat.

(a) Rainy	(b) Sunny	(c) Snow

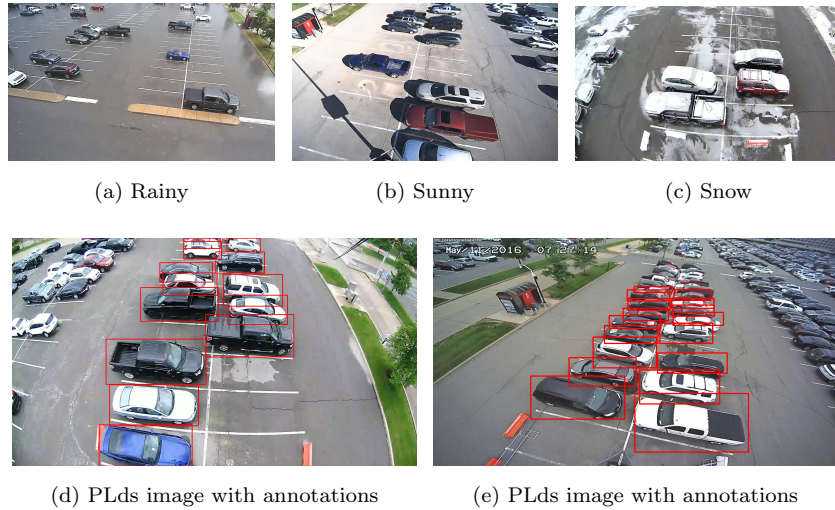|	
(d) PLds image with annotations	(e) PLds image with annotations

Figure 4: PLds image examples. Figures (a), (b), and (c) shows examples of different weather conditions and camera angles. Figures (d) and (e) show examples of images with cars annotations.

cording to climate and luminosity publicly available at `github.com/paulorla/datasets/tree/main/PLds`.

5. State of the Art Review

In this section, we present the review of the works regarding computer vision-based approaches for parking lot management. Specifically, we divide our analysis and discussion based on the different approaches regarding parking lot management. In practice, we first discuss methods for the individual parking spaces classification (Section 5.1), followed by approaches for the automatic parking space detection (Section 5.2), and finally, methods for car detection and counting (Section 5.3).

5.1. Individual parking spaces classification

The individual parking spot classification task can be modeled as the binary problem of classifying individual parking spaces based on whether occupied or

vacant. This problem is exemplified in Figure 5. Each image containing a simple parking spot is fed to a classifier to define its status as vacant, as exemplified in Figure 5a, or occupied, as in Figure 5b. The methods were split into two major groups: methods based on Feature Extraction (Section 5.1.1) and DL methods (Section 5.1.2).
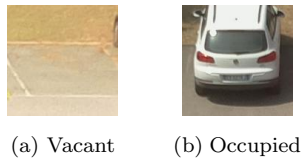

(a) Vacant   (b) Occupied

Figure 5: Individual parking spaces classification example (images from CNRPark-EXT).

5.1.1. Feature Extraction-based Methods

The training phase of the methods based on feature extraction follows the scheme depicted in Figure 6. As an *input image*, we consider the entire image of the parking lot, acquired via a camera. The *image pre-processing* step is used to segment the complete image in individual parking spaces. Some authors may also use techniques to make the image more suitable for feature extraction during the *image pre-processing* step, such as image scaling and histogram equalization. One or more feature vectors may be extracted from the images in the *feature extraction step*, such as the Local Phase Quantization (LPQ) (Ojansivu & Heikkilä, 2008), the Local Binary Patterns (LBP) (Ojala & Pietikäinen, 1999), or the Histogram of Oriented Gradients (HOG) (Dalal & Triggs, 2005).

A classifier, such as an Support Vector Machine (SVM) or Multilayer Perceptron (MLP), is then used in the *model training* step. The feature vectors extracted from the images, and the ground-truth of each image (indicating the correct label of each parking space), are fed to train the classifier during this step. Thus, this final step results in the creation of the trained model, which can be used to classify unseen images.

The use of texture-based features, such as LBP, LPQ, and Quaternionic Local Ranking Binary Pattern (QLRBP) (Lan et al., 2016), is common when
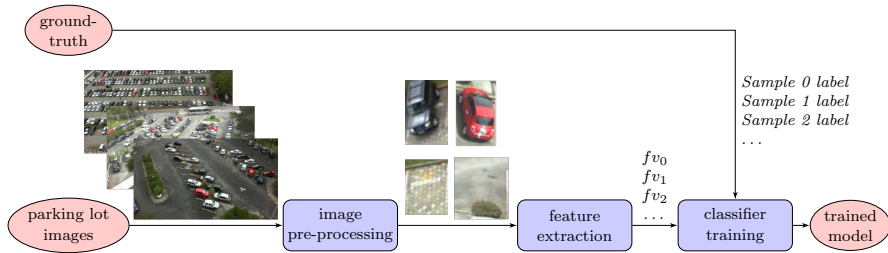
14

Figure 6: Feature Extraction based methods high-level scheme. In the illustration, fv_i refers to the feature vector extracted from the sample (parking space image) i.

dealing with the individual parking spaces classification problem. In Almeida et al. (2013) is proposed the use of LPQ and LBP textures as feature vectors and SVMs as classifiers. In this work, the first version of the PKLot dataset was introduced. The work and the dataset were extended in Almeida et al. (2015), where the entire PKLot dataset was released. Ensembles of SVMs trained using LPQ/LBP as features were used for classification.

Owing to the possible changes caused by luminosity, camera shifts, and parking lot area changes, the authors in Almeida et al. (2018, 2020) considered the individual parking spaces classification from the perspective of a concept drift. Therefore, the authors in Almeida et al. (2018) considered their custom framework for dealing with concept drifts (called Dynse) and employed LBP features. For the experiments, the PKLot dataset was used. The parking lot images are presented day by day in an ordered fashion. Therefore, all current-day instances must be classified according to 100 randomly sampled instances (images) from the previous day used for training. In Almeida et al. (2020), the authors assessed several datasets, including the PKLot, to search for concept drifts' evidence. Using the same features and data split of Almeida et al. (2018), the authors showed that a static or naïve classifier yielded results that are worse than approaches tailored to address concept drifts in the PKLot dataset.

In Suwignyo et al. (2018) the QLRBP was employed as texture features from the color images of the parking spaces. The authors use k-nearest neighbors (k-NN) and SVMs as classifiers. For the tests, 6,000 individual parking spaces

15

of the UFPR04 subset were used. Hammoudi et al. (2018a,b, 2019, 2020) also proposed the use of LBP-based features to classify parking lot images. In these four quite similar works, the authors used a k-NN classifier and small image subsets (3,000 to 6,000 segmented images) of the PKLot for the tests. However, the manner in which the authors grouped the images into subsets is not clear. Further, the authors in Hammoudi et al. (2019) also included SVM classifiers for the tests and tested the changes within parking lots employing a subset of the PKLot and CNRPark-EXT to conduct the tests.

In Dizon et al. (2017) LBP and HOG were used as features descriptors for a linear SVM classifier. The authors also employed a background subtraction approach with Adaptive Median Filter (AMF). The results reported in this study indicated that a classifier trained using only the HOG exhibited good results in the UFPR04 subset from PKLot. Thike & Thein (2019) used Uniform Local Binary Pattern (ULBP) in a complemented image combined with Mean Squared Error (MSE), which is used to classify a parking space based on a threshold applied to the MSE output. The authors tested 1,000 images from the PKLot dataset from different weathers (it is unclear how images were selected).

In the work of Dornaika et al. (2019), SVM and k-NN classifiers are trained with textural features extracted from different scales of the images. The authors used subsets of the PKLot and CNRPark datasets. In addition, a custom-built dataset, including images from the CNRPark and ImageNet (Deng et al., 2009) was used in the tests. Irfan et al. (2020) proposed Gray-Level Co-Occurrence Matrixes (GLCM) as texture features. The test images are classified as occupied or empty according to their similarity to the train images. The authors used only 60 images from the PUCPR subset of the PKLot for the tests. A combination of color and texture features is employed in Mago & Kumar (2020). In addition, the authors put to the test Neural Networks, SVMs, k-NNs, and Naïve Bayes classifiers. The authors employed the PKLot for the tests, but no details about the testing procedure were given.

The use of the pixel values under different color spaces is another approach commonly employed by authors, such as Baroffio et al. (2015); Ahrnbom et al.

(2016); Hadi & George (2019). Baroffio et al. (2015) compute the histograms in Hue, Saturation and Value (HSV) color space directly in smart cameras. The histograms are sent to a central, which uses it as features for an SVM classifier with a linear kernel (also seen in Bondi et al. (2015), from the same authors). The PKLot dataset was used in the tests. The authors claim that energy and bandwidth can be saved by pre-processing the images inside the cameras and sending only the feature vectors or compressed images to a central. Ahrnbom et al. (2016) tested SVM and Logistic Regression (LR) classifiers trained using feature channels, i.e., the individual color channels of an image. Tests were conducted using the PKLot dataset.

After pre-processing the images using Discrete Wavelet Transform (DWT), grayscale conversion, and binary thresholding, the authors in Vítek & Melničuk (2018) employed a simple image average as their feature descriptor. A threshold is applied in the feature descriptor for classification. PKLot and CNRPark-EXT datasets were used for the tests. In Hadi & George (2019), the chromatic gradient analysis of the images is used to classify the individual parking spaces. In addition, the authors proposed an adaptive weather analysis technique to improve the results. Moreover, the classification of parking was based on a threshold, and the PUCPR subset of the PKLot was used to conduct the tests. However, details regarding the test procedure and quantitative results are absent.

Amato et al. (2019a) proposed an approach based on background modeling and the Canny edge detector. The approach was developed to be deployed on smart cameras. The CNRPark-EXT dataset was used during the tests. Raj et al. (2019) also employed the Canny Edge detector but combined with a transformation to the LUV color space to generate the features.A random forest classifier is used for classification.

Bag of features representations of the features is employed in Varghese & Sreelekha (2020) and Mora et al. (2018). Varghese & Sreelekha (2020) proposed an approach that uses an SVM classifier and a bag of features representation. This representation combines the Speeded Up Robust Features (SURF) (Bay

et al., 2008) descriptor and color features to classify the individual parking spaces of the PKLot and CNRPark-EXT datasets. Mora et al. (2018) also proposed an approach using a bag of features for classifying the individual parking spaces. The authors use the SIFT (Lowe, 1999) algorithm to extract the features and use an SVM with a radial basis kernel as a classifier. The authors evaluated their method with the PKLot, considering camera angle, climate, and parking lot changes. The authors also propose a DL-based approach, further described in Section 5.1.2.

The HOG descriptor's used as features and the SVM classifier is explored by Bohush et al. (2018) and by Vítek & Melničuk (2018). In Bohush et al. (2018), a subset of the PKLot dataset containing 2,135 images is used for the tests. However, the authors do not make clear how the images were selected. The authors also propose a segmentation method based on classical image processing methods, further described in Section 5.2. In Vítek & Melničuk (2018), a lightweight approach developed to be deployed on Smart Cameras is proposed. Information about the parking angle is concatenated with the HOG feature vector to improve the results. Tests were made using the PKLot and two private datasets.

5.1.2. Deep Learning Based Methods

DL-based methods follow a workflow similar to feature-based approaches, except with the feature extraction and classifier training parts being joined together in a *representation learning* block. Here, feature engineering is not employed since DL models aim to learn the representation of the parking spaces. Also, the classifier is generally a part of the DL model. This workflow can be seen in Figure 7.

The *image-processing* step follows the same concept described for feature-based methods (Section 5.1.1). It may also include data augmentation approaches, aiming to increase the number of samples and their variability. The *representation learning* in this problem can be divided into (i) transfer learning of well-known convolutional networks for classification, such as LeNet (LeCun
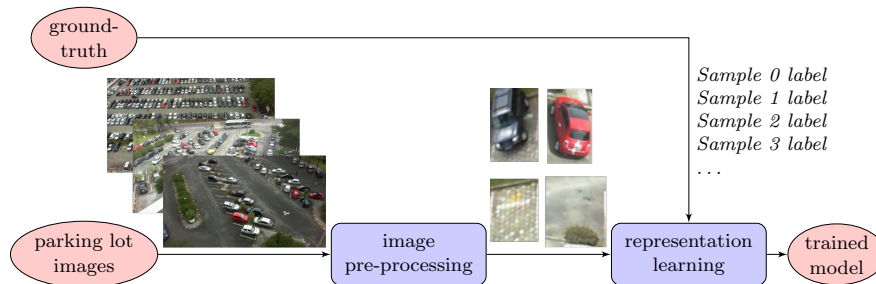
18

Figure 7: DL-based methods high level scheme.

et al., 1998) and AlexNet (Krizhevsky et al., 2012); (ii) the proposal of a custom convolutional model, generally based on these well-known networks; and (iii) the use of DL networks for object detection or segmentation, such as Faster-RCNN (Ren et al., 2015) and Mask R-CNN (He et al., 2017). It is also noteworthy that in (Acharya et al., 2018), authors used the SVM classifier to replace the softmax function of neural networks.

Transfer learning (Yosinski et al., 2014), where a network is first trained in a generic dataset and then fine-tuned in a parking lot dataset, is a common approach. LeNet and AlexNet networks are popular due to their compactness. Both are used in Nyambal & Klein (2017) to classify parking lots. The authors use the PKLot and a private dataset, although it is unclear how the tests were performed for the PKLot. The usage of AlexNet is also seen in Di Mauro et al. (2016). It focuses on optimizing a model with only a few samples, either PKLot or a private dataset.

The authors in Ding & Yang (2019) proposed to add residual blocks in the Yolov3 (Redmon & Farhadi, 2018) to extract more granular features. The modified network is used to classify the parking lot's images. Vehicle images from the PASCAL VOC (Everingham et al., 2010) and COCO (Lin et al., 2014) datasets are used to train this network. Then, it is fine-tuned using some images of the PUCPR subset of the PKLot. Images of PUCPR were also used during the tests. The authors do not clarify how the images were split between the training and testing sets.

19

Many works proposed lightweight models based on well-known convolutional networks, such as LeNet, AlexNet, and VGGNet (Simonyan & Zisserman, 2014). These custom models are primarily convolutional networks similar to the original networks but with fewer layers, i.e., shallow networks. These models are usually developed for low-power and restricted processing capabilities devices, such as smart cameras. Works such as Amato et al. (2016, 2017); Polprasert et al. (2019); Valipour et al. (2016); Acharya et al. (2018); Bura et al. (2018); Merzoug et al. (2019); Rahman et al. (2020); Manjur Kolhar (2021) can be grouped in these lightweight versions, where most authors use the PKLot dataset for the tests. Private datasets were included in Acharya et al. (2018); Polprasert et al. (2019); Merzoug et al. (2019); Bura et al. (2018).

The authors in Amato et al. (2016) proposed the mAlexNet, based on the AlexNet network, and executed experiments in CNRPark. Amato et al. (2017) used the extended version of the dataset, the CNRPark-EXT. Their mAlexNet can cope with the parking lot and camera angle variations with tiny accuracy drops in many scenarios (the authors also included some tests in the PKLot). Similarly, Amato et al. (2019a) employed the mAlexNet to classify the parking spaces using smart-cameras and used the CNRPark-EXT dataset for the tests. Rahman et al. (2020) also employed mAlexNet but changed the kernel size of the first layer. No significant difference in the final results was found. Nguyen et al. (2021) evaluated the mAlexNet, AlexNet, and MobileNet (Howard et al., 2017) networks with a width multiplier of 0.5 (which presented better performance) using Camera A of the CNRPark and a private dataset. They aimed an approach for low-cost hardware, but the processing cost (and time) after image capture was left unclear.

Bura et al. (2018) also proposed a lightweight version of the AlexNet network in their work, wherein they used subsets of the PKLot, CNRPark-EXT, and a private dataset to conduct the tests. Nevertheless, information regarding the availability of the private dataset and the manner in which the images were selected for the subsets is absent.

A lighter version of the AlexNet network is also proposed in Ali & Mohamed

(2021). The authors removed one of the convolutional layers of the original network, together with some minor modifications. The PKLot dataset was used for the tests.

The authors in Valipour et al. (2016) trained a VGGNet-F (Simonyan & Zisserman, 2014) model with the ImageNet (Deng et al., 2009) dataset and fine-tuned it with PKLot images, which resulted in better generalization capabilities (considering camera and parking lot changes) when compared with the feature extraction methods used in Almeida et al. (2015). Acharya et al. (2018) proposed an approach wherein a VGGNet-F model is trained for feature extraction and inputs features to an SVM classifier. The authors employed 5-fold cross-validation in the PKLot images during the test phase. Zhang et al. (2019) proposed a modified version of the VGG16 network and a custom network to classify whether the parking spaces are empty or occupied. The authors also proposed applying image transformations to get a top view of the parking lots and employed the PKlot dataset during the tests.

The original VGG16 network is used by Mora et al. (2018) and Dhuri et al. (2021). Mora et al. (2018) employed the PKLot in the tests, considering camera angles, climate conditions, and parking lot changes. The authors also propose a method based on Bag of Features, discussed in Section 5.1.1. In Dhuri et al. (2021) the PKLot and CNRPark-EXT datasets are used to test scenarios containing parking lot changes. The authors used a small custom dataset containing 1,000 cropped images of individual parking spaces as well.

The MobileNetV2 (Sandler et al., 2018) was used to create a lightweight network for parking spaces classification in Merzoug et al. (2019). The authors employed the PKLot, CNRPark-EXT, and a private dataset during the tests. However, the authors' testing procedure is not clear. The results were reported as the number of detected vehicles without any ground truth for comparison. A shallower version of the ResNet50 (He et al., 2016) residual network is used to classify the individual parking spaces in Gregor et al. (2019). The authors used the PKLot and a private dataset for the tests. The training/testing instances were split using stratified sampling. The original ResNet50 network is employed

in Baktir & Bolat (2020). They employed small subsets of the PKLot and CNRPark-EXT datasets for the tests.

In Chen et al. (2020), a lightweight version of the Yolov3 network that uses the MobileNetV2 extraction layer is employed for classification. The authors consider that the images may come from a video stream source. Thus the parking spot is considered occupied if the bounding box of a car detected by the network overlaps a parking space during a time window of n images. The CNRPark-EXT and a private dataset were used during the tests.

Custom models were proposed by Di Mauro et al. (2016); Jensen et al. (2017); Di Mauro et al. (2017); Thomas & Bhatt (2018); Nurullayev & Lee (2019); Shah et al. (2020). A fine-tuned AlexNet network is used in Di Mauro et al. (2016) and Di Mauro et al. (2017), using images from the same parking lot and camera angle (from PKLot and a private dataset). A pseudo-label model for semi-supervised learning is employed in Di Mauro et al. (2016) to compare with AlexNet, beating the former accuracy. Jensen et al. (2017) proposed a model with a fixed input size of 40x40 pixels in the PKLot dataset using the original test protocol proposed for the PKLot. In Nurullayev & Lee (2019) is proposed a Convolutional Neural Network (CNN)-based method using dilated convolutions that achieved promising results, which skips pixels in the convolution kernel. According to the authors, it increases the classifier's ability to learn the global context of the images. Thomas & Bhatt (2018) also used a custom model, but they did not specify its structure or parameters.

The authors in Shah et al. (2020) proposed a custom network called Fully-Multitask Convolutional Neural Network. The proposed approach depends on the Mask R-CNN for the extraction of masked regions. The authors claim that the proposed approach is capable of counting and automatically detecting the parking spaces. However, no details about these features nor results are given (thus, this work is not considered in Sections 5.2 and 5.3). For the individual parking spaces classification, the authors report quantitative results only for the training and validation phases in the PKLot dataset (results for the test phase are given in a plot according to the training epoch, making its analysis difficult).

DL models for detection and segmentation were employed in Martín Nieto et al. (2019) and Sairam et al. (2020) to aid vehicle detection. Martín Nieto et al. (2019) used Faster R-CNN (Ren et al., 2015) and multiple cameras from their proposed PLds dataset. They applied a homographic transformation and perspective correction to transform the plane of each camera to a common plane and correct the positions of the detected cars, respectively. Thereafter, the cars detected by the different cameras were fused to classify the parking spots. The Faster R-CNN was also used in Khan et al. (2019), where the authors focused on tests involving different camera angles and parking lot changes using the PKLot dataset.

More recently, Sairam et al. (2020) proposed a method based on the Mask R-CNN (He et al., 2017) network. It was used to extract individual vehicles and to detect the proportion of the parking space the vehicles are occupying to differentiate between cars and two-wheel vehicles. The Mask R-CNN is also used in Agrawal & Urolagin (2020). The network is employed to detect the individual parking spaces (details in Section 5.2) and classify the spots between occupied and empty. The authors trained their model using the COCO (Lin et al., 2014), and COWC datasets (Mundhenk et al., 2016), and tested using the PKLot dataset.

In Mettupally & Menon (2019), the Mask R-CNN network is trained to classify the individual parking spaces. The authors included the images' timestamps and orientation to improve the classification results and used a custom subset of the PKLot dataset containing 6,100 segmented images for the tests.

Generative Adversarial Networks (GANs) (Isola et al., 2017) are employed in Li et al. (2017) to detect occupied automatically and vacant parking spaces using a team of drones. The GAN is trained using the labeled images of the PKLot dataset, wherein the division between and training and testing subsets was carried out according to the original PKLot protocol. In the test phase, the images rotated at a random angle between $[-10, +10]$ degrees were included to simulate the data obtained from the drones.

5.2. Automatic Parking Space Detection

The automatic parking space detection task focuses on automatically detecting the coordinates of each parking space, e.g., obtaining a bounding box, regardless of its status (occupied or empty). This task is exemplified in Figure 8, where the yellow polygons (that define each parking space) coordinates should be automatically defined. It is a challenging task since parking spaces are similar to roads, i.e., how can a model discriminate between a parking space and a road segment? The presence of cars may hinder correct detection, especially for methods that rely on the painted demarcations in the parking lots that delimits the parking spaces.

Figure 8: Parking spaces locations example(image from CNRPark-EXT).

An automatic approach for detecting parking spaces using classical image processing methods is proposed in Bohush et al. (2018). The approach uses perspective transformation in the entire input image. It makes the parking spaces rectangular and parallel to the axes (the approach may not be suitable for parking areas where the cars park in angled parking spaces). Otsu's binarization and morphological operations are used for parking space detection, where painted lines must delimit the parking spaces. The authors did not make clear the results achieved by their proposed method. Zhang et al. (2019) also proposed an approach for automatic parking space detection using perspective transformation and classical image processing approaches, such as the Canny and Gaussian edge detectors. The authors did not present quantitative results.

Further, the authors in Vítek & Melničuk (2018) employed a grid-based approach, where for each block of the grid, HOG features are extracted. There-

after, a classifier is used to classify each block as a car or street. Blocks classified as cars are merged into parking spaces according to their neighbor blocks. Adjacent blocks are classified as cars that are considered as belonging to the same parking spot. The method seems to detect cars and not necessarily parking spaces, e.g., a car may be just passing by the parking lot. The authors only report some images with qualitative results.

By assuming that the car park area is rectangular (forming a parking grid), the authors in Martín Nieto et al. (2019) can automatically define each parking spot. Given an aerial image of the parking lot (computed using a homography matrix and a regular image collected by a camera), the corners of the parking grid, and the number of rows and columns, the method can automatically define each parking space of the parking grid.

The GAN-based approach by Li et al. (2017) generates parking spaces from the PKLot dataset using the manually-made masks to train the network. Although parking spaces are segmented, only an evaluation of individual parking spaces is made. Evaluation of the automatic detection was not executed. Agrawal & Urolagin (2020) proposed using the Mask R-CNN to identify the positions where the cars stay parked. The authors use this information to extract the parking space positions by assuming that areas where cars stay parked for long periods, can be considered parking spaces. Unfortunately, the authors do not describe the details of the parking space detection approach. Also, they do not give quantitative results.

The authors in Padmasiri et al. (2020) used the ResNet (He et al., 2016) and Faster-RCNN networks to detect the parking spaces in the PKLot dataset automatically. The authors reported results using the Average Precision (AP) metric. However, images from the same parking lot were used for both training and testing, which may lead to biased results, as discussed in Section 6.3.

A two-step automatic parking space detection approach is proposed by (Patel & Meduri, 2020). First, an object detector, Faster R-CNN (same work with this detector is found in Kirtibhai Patel & Meduri (2020)) or YOLOv4(Bochkovskiy et al., 2020), is employed for car detection (trained in the CARPK dataset (Hsieh

et al., 2017)). Then, the bounding boxes detected are used for car tracking. For car tracking, the idea is to find bounding boxes with a stationary car for some time/frames. They evaluated the proposed two-step approach using only three busy days from CNRPark-EXT for each weather condition, introducing bias in the tests. A parking space is marked if a car is parked for at least one hour and a half.

5.3. Car Detection and Counting

In this Section, methods that aim to detect and count individual cars in the images are presented. For car detection, bounding boxes or segmentation masks may be generated and evaluated for each car, as exemplified in Figure 9. In contrast, we have a regression problem for the car counting task. In this case, we are only interested in the discrete final number of cars present in the image (seven in the example in Figure 9). Like in the individual parking spaces classification problem, an image from the parking lot is the input. Different from it, no pre-processing to extract the individual parking spaces is applied. Proposed works aim to achieve a high car detection precision and reduce the difference between the prediction and the actual number of cars in the images.
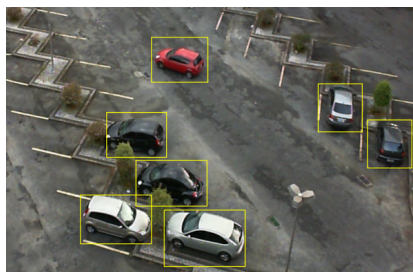

Figure 9: Car Detection example. Image containing 7 cars (image from PKLot).

Common datasets used to evaluate methods that deal with this task include the PUCPR+ (an extension of the PUCPR subset from the PKLot) (Hsieh et al., 2017; Li et al., 2019; Gabzdyl, 2020), the original PUCPR subset of the PKLot (Laradji et al., 2018), the complete PKLot dataset (Varghese &

Sreelekha, 2020), the CARPK dataset (Hsieh et al., 2017; Gabzdyl, 2020), and drone-acquired image datasets (Hsieh et al., 2017; Li et al., 2019).

Most authors use DL-based approaches for car detection. They also count the number of detected instances (e.g., bounding boxes) to obtain the final prediction for the number of cars in the images (Hsieh et al., 2017; Laradji et al., 2018; Li et al., 2019; Amato et al., 2019a,b). The authors in Hsieh et al. (2017) aimed to detect and count aerial images from drones but used 5-fold cross-validation (which may overestimate the reported results). In contrast, Li et al. (2019) generated the anchors for the training phase adaptively.

In Laradji et al. (2018) a new loss function with a convolutional model for car detection and counting is proposed. They also used annotations that roughly contain the objects of interest, which are, according to the authors, easier to label manually. The Yolov3 was used in Amato et al. (2019b) to count the number of cars in the images. The authors used the PUCPR+ and CARPK datasets in the tests using the test protocols proposed by Hsieh et al. (2017). Similarly, Amato et al. (2019a) (also in Ciampi et al. (2018)) used the Mask R-CNN for counting by density in a CNRPark-EXT counting version called Counting CNRPark-EXT. Since the original dataset does not have car masks, they initially trained Mask R-CNN to correctly generate mask predictions of cars (using about 10% of the training subset). Then, they retrained Mask R-CNN with the generated masks (and some cases that were manually corrected).

The authors in Varghese & Sreelekha (2020) and Sharma & Pandey (2021) proposed detection-only approaches. In the work of Varghese & Sreelekha (2020) a background subtraction-based method is used for hypothesis generation of the possible areas where cars park. A shadow model is employed to reduce the amount of noise. Then a classifier is used to verify if the segmented areas contain cars. A custom CNN is proposed in Sharma & Pandey (2021). The authors claim that the custom network is lightweight and can be processed in a regular CPU.

The approaches discussed in Gabzdyl (2020); Stahl et al. (2018); Dobeš et al. (2020) count the number of objects in the images globally, without employing a
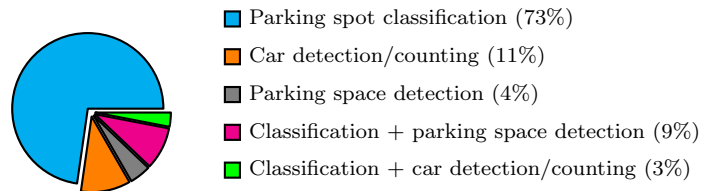
car detection step. Gabzdyl (2020) proposed a tree-like CNN-based car counting approach. The first ten layers from a VGG16 network are used for feature extraction. The following convolutional layers are used to generate a density prediction, then used to count cars. Point-wise and dilated convolutions were employed in this part.

The authors in Stahl et al. (2018) proposed an approach where the image is divided into several patches and fed to a modified R-FCN (Dai et al., 2016) network. The method includes an inclusion-exclusion layer to detect objects counted more than once since an object may appear in more than one patch. The approach only needs the number of objects in the training images for the training phase. It was evaluated under several object counting benchmarks, including the PUCPR+ dataset. A modified version of the Stacked Hourglass Network (Newell et al., 2016) is used in Dobeš et al. (2020). The modified network can identify the best scale of the input images to count the cars. The network input is a pyramid of gradually downsampled images and, for the training phase, the network needs labels in the form of point annotations. The authors used the PUCPR+ altogether with other car counting benchmarks.

6. Discussion

In this section, the datasets reviewed and works that used them are summarized and discussed. Figure 10a shows that most works (73%) focus on the individual parking spaces classification only. As one can observe in Figure 10b, the PKLot dataset is the most popular dataset when considering the surveyed works, being used in 88% (70% + 18%) of the works, while the CNRPark-EXT and the PLDs datasets are used in 29% (11% + 18%) and 1% of the works, respectively. The difference between the usage ratios can be partially explained by the publication dates of these datasets, as PKLot, CNRPark-EXT, and PLds datasets were released in 2015, 2017, and 2019, respectively.

In Figure 10b, it is also possible to notice that only 18% of the surveyed works use more than one publicly available dataset for the tests. When comparing

(a) Main tasks considered in the surveyed works.

- Parking spot classification (73%)
- Car detection/counting (11%)
- Parking space detection (4%)
- Classification + parking space detection (9%)
- Classification + car detection/counting (3%)

(b) Overall datasets usage.

- PKLot (70%)
- PKLot + CNRPark-EXT (18%)
- CNRPark-EXT (11%)
- PLds (1%)

(c) Feature extraction and Deep Learning usage.

- Feature extraction (35%)
- Deep Learning (61%)
- Feature extraction + DL (4%)

Figure 10: General findings considering all surveyed works.

the proportion of approaches based on Feature Extraction and Deep Learning shown in Figure 10c, it is possible to check that DL-based approaches are more prevalent in general.

An overview of our findings and recommendations regarding the surveyed datasets and research gaps for the individual parking space classification, parking space detection, and car detection/counting tasks are the following:

1. The current publicly available datasets lack some features. Future datasets should be robust as defined in this work and may contain:

   - Video sequences;

   - Images in nighttime and snow conditions;

   - Labels, other than the parking spaces, such as segmentation masks for the objects (e.g., cars, obstacles, pedestrians).

2. There is a lack of standard protocols for testing approaches:

   - New protocols should include the data split procedure, evaluation metrics, and specific challenges (e.g., generalization problems, automatic parking space detection problems);

   - The protocols should take into consideration realistic scenarios.

3. Many of the surveyed works are not reproducible:

   - Authors should use publicly available datasets;

   - Standard test protocols, as suggested in Item 2, should be used.

4. The individual parking spaces classification under camera or parking lot change scenarios is an open problem:

   - Authors should consider a multiple datasets perspective, for instance, training in the PKLot dataset and testing the CNRPark-EXT.

5. The automatic parking space detection is an open problem:

   - Quantitative metrics must be used to report the results;

   - Standard test protocols should be created (as in Item 2);

   - Authors should consider multiple datasets (as in Item 4).

These findings are discussed more deeply in Sections 6.1, 6.2, 6.3, and 6.4. In Section 6.1, we summarize and discuss the surveyed datasets. In Sections 6.2, 6.3, and 6.4, we present, compare, and discuss the authors' results in the individual parking spaces classification, parking space detection, and car detection/counting tasks, respectively. These discussions are used to provide a thoughtful understanding of the vision-based parking lot management problems that already have a factual basis for solutions and open problems that require more research and attention from the scientific community.

6.1. Parking Lot Datasets

Table 1 shows the main features of the datasets discussed in this paper. Images are classified according to the climate condition, regardless if they were

acquired under day or nighttime. For the overcast climate condition during nighttime, we considered the images right after rains when the floor is visibly wet and may contain water puddles.

Table 1: Main features of the datasets

PKLot – 1280 × 720 pixels, 12,417 daytime images

Camera/	# Of Days (# Of Images)				spaces/	Labeled Parking Spaces		
Park. Lot	Clear Sky	Overcast	Rain	Snow	image	Occupied	Empty	Total
UFPR04	20 (2,098)	15 (1,408)	14 (285)	-	28	46,125 (43.6%)	59,718 (56.4%)	105,843
UFPR05	25 (2,500)	19 (1,426)	8 (226)	-	45	97,426 (58.8%)	68,359 (41.2%)	165,785
PUCPR	24 (2,315)	11 (1,328)	8 (831)	-	100	194,229 (45.8%)	229,994 (54.2%)	424,223
Total	69 (6,913)	45 (4,162)	30 (1,342)	-	-	337,780 (48.6%)	358,071 (51.4%)	695,851

CNRPark-EXT – 1000 × 750 pixels, 4,278 daytime images

Camera A	2 (-)	0 (-)	0 (-)			3,621 (58.7%)	2,550 (41.3%)	6,171
Camera B	2 (-)	0 (-)	0 (-)			4,781 (74.6%)	1,632 (25.4%)	6,413
Camera 1	10 (198)	7 (137)	6 (124)	-	35	9,308 (59.2%)	6,407 (40.8%)	15,715
Camera 2	10 (201)	7 (139)	6 (124)	-	10	2,641 (64.5%)	1,454 (35.5%)	4,095
Camera 3	10 (198)	7 (138)	6 (124)	-	22	5,370 (56.7%)	4,101 (43.3%)	9,471
Camera 4	10 (197)	7 (137)	6 (123)	-	38	9,357 (56.4%)	7,219 (43.6%)	16,576
Camera 5	10 (204)	7 (143)	6 (127)	-	47	11,256 (54.0%)	9,582 (46.0%)	20,838
Camera 6	10 (211)	7 (154)	6 (130)	-	45	10,646 (52.9%)	9,462 (47.1%)	20,108
Camera 7	10 (211)	7 (154)	6 (130)	-	47	10,519 (49.8%)	10,595 (50.2%)	21,114
Camera 8	10 (206)	7 (151)	6 (127)	-	55	12,847 (53.3%)	11,237 (46.7%)	24,084
Camera 9	10 (205)	7 (154)	6 (131)	-	29	7,363 (56.8%)	5,601 (43.2%)	12,964
Total[1]	12 (1,831)	7 (1,307)	6 (1,140)	-	-	87,709 (55.7%)	69,840 (44.3%)	157,549

PLds – 1280 × 960 pixels, 5,215 daytime images (62.5%), 3,125 (37.5%) nighttime images

isshk	45 (2,280)	22 (635)	21 (428)	8 (321)	-	32,254 (100%)	-	32,254
vxusd/vmlix[2]	46 (2,077)	37 (1,018)	27 (485)	2 (96)	-	26,719 (100%)	-	26,719
qridr	18 (435)	15 (513)	6 (52)	-	-	-	-	-
Total[1]	69 (4,792)	47 (2,166)	34 (965)	10 (417)	-	58,973 (100%)	-	58,973

[1] Different cameras can collect images in the same days, thus the total number of days may not be equal to the sum of the days for each camera;   [2] We detected that the vxusd and vmlix refer to the same camera angle.

Table 1 indicates that the PKLot is the dataset with the most parking lot images (12,417) and individual parking spaces samples available (695,851). Another interesting feature of the PKLot, unavailable in other surveyed benchmarks, is having more than one parking lot (i.e., PUCPR and UFPR). Regarding the CNRPark-EXT, one of the most appealing features is its vast number of camera angles available, which can be used in camera angle change tests. In Table 1, it is also possible to check that both the PKLot and CNRPark-EXT datasets are reasonably well balanced between the occupied and empty classes.

When considering the PLds dataset, we observe that it is the only surveyed dataset that contains images collected in snow and nighttime conditions. Another attractive feature of the PLds dataset is the presence of images synchro-

nized between different cameras. However, only 100 synchronized images are available. The main drawback of the PLds dataset is the absence of annotated empty parking spaces.

Although all the datasets combined give 25,035 parking lot images and 912,373 individual parking spaces images, only 417 (1.7%) images represent snow conditions, and 3,125 (12%) images were collected under nighttime. Both snow and nighttime condition images are from the PLDs dataset and were collected in a single parking lot. None of the surveyed datasets contained video sequences. The only labels available were the climate condition, the positions of the parking spaces and their statuses (for the PKLot and CNRPark-Ext), and the parked cars bounding boxes (for the PLds).

These findings indicate that new datasets are needed, mainly containing conditions such as nighttime and snow climate. These datasets may also contain video sequences and labels regarding objects other than the parking spaces, such as segmentation masks for the objects in the scene (e.g., vehicles, pedestrians, obstacles). These datasets could help create more general methods for the tasks described in this paper and other tasks, such as object tracking.

When considering the test protocols, the authors of all discussed datasets suggested that the same day's images may belong to just one set - training or test set. This rule avoids pictures related to the same car parked in the same space for hours to appear in the training and the testing sets simultaneously, creating trivial and unrealistic problems – See an example in Figure 11. Almeida et al. (2015) suggest that the PKLot dataset must be split using 50% of the days for training and the other 50% for tests. Amato et al. (2017) gives their train and test images for the CNRPark-EXT dataset, containing 74.8% and 25.2% of the images, respectively. The authors of the PLds dataset (Martín Nieto et al., 2019) also made available the images already partitioned between the train and test sets, containing 76% and 24% of the images, respectively.

In Almeida et al. (2015) and Amato et al. (2017), the test protocols include camera angle and parking lot changes that may represent more challenging and realistic situations. The general ideas of the tests suggested by the authors of

32

(a) 17:12   (b) 17:32   (c) 17:52   (d) 18:12   (e) 18:32

Figure 11: Car parked for several minutes in a parking space (PKLot).

the datasets are next described. These scenarios (single parking lot, camera angle change, and parking lot change) are considered in Section 6.2. There, the surveyed techniques regarding the individual parking spot classification problem are compared.

- Single parking lot: images for both training and test sets come from the same parking lot and are collected using the same camera angle;

- Camera angle change: images must be from the same parking lot, but the train set must contain images collected using a different camera angle from the test set's images;

- Parking lot change: the test set must contain images from a parking lot different from the train set.

6.2. Individual Parking Spot Classification

Table 2 shows a summary of the individual parking spot classification methods based on feature extraction approaches. As for DL-based methods, a summary can be seen in Table 3. Works that do not clearly describe the test details are marked as not reproducible in these tables. This marking indicates that not enough details are given for other researchers to reproduce the experiments. Some of the main reasons encountered in the works that led us not to consider the results reproducible include the use of images in private datasets, use of not well-described subsets of the public datasets, and incomplete evaluation protocols, e.g., the protocol used to split the dataset between train and test is not given. Even though all surveyed methods use at least one public dataset, surprisingly, 56% of them were deemed as not reproducible.

33

Different authors may use distinct metrics to assess their methods, such as accuracy, Area Under the ROC Curve (AUC-ROC), and Area Under the Precision-Recall Curve (AUC-PR). This pattern imposed a challenge to summarize the methods, and we chose to show all results in terms of accuracy rates. The accuracy may not be the most suitable metric to assess the parking spaces classification task, e.g., due to class imbalances in the test sets. Nevertheless, we use it to show the results since most authors reported their results using this metric. On the other hand, for specific scenarios in which accuracy was not used, we report the results following the metric adopted in the original studies depicted in Tables 2 and 3. Authors that did not present quantitative results in their works are not discussed in this section.

The results in Tables 2 and 3 are reported according to the authors' tests, including scenarios where the images from the training and test subsets come from the same parking lot and camera angle, scenarios containing camera angle changes, and scenarios when the train and test parking lots are different – see Section 6.1 for more details. For most works, we reported a range in the results in the format $x - y$, where x is the worst result achieved in a given scenario, and y is the best result. For instance, in the *Single Parking lot* test, if a method achieved 88%, 97%, and 93% of accuracy when tested in the UFPR04, UFPR05, and PUCPR PKLot's subsets, respectively, the reported result is $88 - 97\%$. It is worth mentioning that the results reported in Tables 2 and 3 are not directly comparable since the authors may vary the datasets and experimental protocols.

Notice in Tables 2 and 3 that many authors, such as Almeida et al. (2013); Di Mauro et al. (2016); Suwignyo et al. (2018); Dornaika et al. (2019); Vítek & Melničuk (2018); Acharya et al. (2018); Polprasert et al. (2019); Farag et al. (2020), may have used test protocols that led to biased results. Since a car may remain parked in the same spot in a parking lot for long periods, it is expected that the same car to appear in several consecutive images. Any test approach that, for instance, randomly splits the data for training and testing may include images of the same car in both sets (training and testing) acquired at different times, leading to a trivial and unrealistic problem. Figure 11 shows an example

Table 2: Feature extraction based individual parking spaces classification approaches overview.

Authors, year	Classifier and Features	Repro-ducible	Accuracy (%)			Datasets Used
			Single Parking Lot	Angle Change	Parking Lot Change	
Almeida et al. (2013)	Ensembles of SVM and LBP/LPQ	yes	99.8%[1]	-	-	PKLot (UFPR04)
Almeida et al. (2015)	Ensembles of SVM and LBP/LPQ	yes	99.3 - 99.6%	85.8 - 88.3%	84.2 - 89.8%	PKLot
Baroffio et al. (2015)	SVM and HSV Histogram	yes	87.0 - 96.0%	-	-	PKLot
Ahrnbom et al. (2016)	SVM/LR and Feature Channels	yes	0.999 - 1.0[3]	0.978 - 0.996[3]	0.940 - 0.988[3]	PKLot
Suwignyo et al. (2018)	k-NN/SVM and QLRBP	yes	99.2 - 99.4%[1]	-	-	PKLot
Dizon et al. (2017)	Background subtraction with AMF and linear SVM/LBP with HOG	no	1.8%/21.0%[4]	-	-	PKLot (UFPR04)[2]
Almeida et al. (2018)	Dynse Framework and LBP	yes	92.2%	-	-	PKLot
Bohush et al. (2018)	SVM and HOG	no	99.7%[1]	-	-	PKLot (Subset)
Vítek & Melničuk (2018)	SVM and HOG	no	90.7 - 93.2%[1]	-	83 - 96%	PKLot[2]
Hammoudi et al. (2018a)	k-NN and LBP-based features	no	98% - 98.6%[1]	-	-	PKLot(Subset)
Hammoudi et al. (2018b)	k-NN and LBP-based features	no	88.65[1]	-	-	PKLot(Subset)
Mora et al. (2018)	SVM and SIFT (Bag of Features)	yes	76.8 - 92.0%[1]	94.1 - 99.7%	90.4 - 97.3%	PKLot
Amato et al. (2019a)	Background modeling and Canny	yes	99.4%	-	-	CNRPark-EXT
Dornaika et al. (2019)	SVM/k-NN and textural features	yes	88.7 - 99.6%[1]	-	63.0 - 87.8%	PKLot, CNRPark-EXT
Raj et al. (2019)	Canny and color features with Random Forest	yes	98.31%[1]	-	-	PKLot
Hammoudi et al. (2019)	k-NN/SVM and LBP-based features	no	85.6%[1]	-	71.0 - 80.5%	PKLot, CNRPark-EXT (Subsets)
Thike & Thein (2019)	Threshold and ULBP	no	79%[1]	-	-	PKLot
Varghese & Sreelekha (2020)	SVM and SURF	yes	91.1 - 99.7%	-	81.7%	PKLot, CNRPark-EXT
Farag et al. (2020)	Threshold and color average	no	79.0 - 90.0%[1]	-	-	PKLot, CNRPark-EXT
Irfan et al. (2020)	Euclidean Similarity and GLCM	no	95.0%	-	-	PKLot(PUCPR)[2]
Hammoudi et al. (2020)	k-NN and LBP-based features	no	87.0%[1]	-	-	PKLot(UFPR04)
Mago & Kumar (2020)	Neural Networks, k-NN, SVM, Naïve Bayes, textures and color features	no	99.7%[1]	-	-	PKLot
Almeida et al. (2020)	Approaches for concept drift and LBP features	yes	86.7 - 90.4%	-	-	PKLot

[1] The test procedure may have led to biased results, or it is not clear.   [2] Authors included private datasets in the tests.   [3] Results in AUC-ROC.   [4] Results in False Negative Rate (FNR)/False Positive Rate (FPR).

Table 3: Deep learning based individual parking spaces classification approaches overview.

Authors, year	Network	Repro-ducible	Accuracy (%) Single Parking Lot	Angle Change	Parking Lot Change	Datasets Used
Amato et al. (2016)	mAlexNet	yes	89.8 - 99.6%	86.3 - 90.7%	82.9 - 90.4%	PKLot, CNRPark-EXT (Cam. A and B)
Valipour et al. (2016)	VGGNet-F	yes	99.6 - 100%	85.5 - 98.2%	92.2 - 99.2%	PKLot
Di Mauro et al. (2016)	AlexNet	no	99.2 - 99.8%[1]	-	-	PKLot[2]
Di Mauro et al. (2017)	AlexNet	no	99.9 - 100%[1]	-	-	PKLot[2]
Amato et al. (2017)	mAlexNet	yes	90.1 - 98.1%	93.3 - 93.7%	92.7 - 98.3%	PKLot, CNRPark-EXT
Jensen et al. (2017)	Custom CNN	yes	99.7 - 99.9%	95.5 - 96.0%	96.7 - 98.7%	PKLot
Nyambal & Klein (2017)	AlexNet and LeNet	no	98.0 - 99.0%	-	-	PKLot[2]
Li et al. (2017)	GAN	yes	94.7 - 97.5%[3]	56 - 94%[3]	-	PKLot
Acharya et al. (2018)	VGGNet-F(feat. extr.)/SVM	no	99.7%[1]	-	-	PKLot[2]
Bura et al. (2018)	Custom AlexNet	no	99.5%[1]	-	-	PKLot, CNRPark-EXT (Subsets)[2]
Thomas & Bhatt (2018)	Custom CNN	no	100%[1]	-	-	PKLot
Amato et al. (2019a)	mAlexNet	yes	99.6%	-	-	CNRPark-EXT
Nurullayev & Lee (2019)	CarNet	yes	95.6 - 98.8%	95.2 - 97.6%	94.4 - 98.4%	PKLot, CNRPark-EXT
Polprasert et al. (2019)	mAlexNet	no	88%[1]	-	-	PKLot[2]
Gregor et al. (2019)	Custom ResNet	no	99.9%[1]	-	-	PKLot[2]
Martín Nieto et al. (2019)	Faster R-CNN	yes	0.919[4]	-	-	PLds
Zhang et al. (2019)	Custom VGG16 and Custom CNN	no	99.1 - 99.9%[1]	-	-	PKLot
Ding & Yang (2019)	Yolov3	no	93.3%[3]	-	-	PKLot
Khan et al. (2019)	Faster R-CNN	yes	-	82.5 - 96.4%	90.5 - 99.9%	PKLot
Mettupally & Menon (2019)	Mask R-CNN	no	91.9% [1]	-	-	PKLot
Agrawal & Urolagin (2020)	Mask R-CNN	yes	-	95%	95%	PKLot
Baktir & Bolat (2020)	ResNet50	no	99.5 - 99.8%[1]	-	97.4 - 99.3%	PKLot, CNRPark-EXT (Subsets)
Sairam et al. (2020)	Mask R-CNN	no	92%[1]	-	-	PKLot
Chen et al. (2020)	Yolov3	no	99%[1]	-	-	CNRPark-EXT[2]
Rahman et al. (2020)	CmAlexNet	no	99.12%[1]	-	-	CNRPark[1]
Ali & Mohamed (2021)	Custom AlexNet	no	92.9 - 98.9%[1]	94.9 - 96.9%	96.0 - 99.0%	PKLot
Dhuri et al. (2021)	VGG16	yes	95.7 - 100%	97.1%	87.2%	PKLot, CNRPark
Nguyen et al. (2021)	mAlexNet, AlexNet and MobileNet	no	95.5 - 97.3%[1]	-	-	CNRPark[2]
Manjur Kolhar (2021)	mAlexNet	no	90.1 - 99.5%[1]	-	-	PKLot, CNRPark-EXT

[1] The test procedure may have led to biased results or, it is not clear.
[2] Authors included private datasets in the tests. [3] Results in Precision. [4] Results in AUC-PR.

where the same car remained parked in the same spot for several minutes.

All consecutive images acquired from a parked car should be exclusively selected as part of the training or testing sets in a more realistic scenario. The test protocols originally suggested in the datasets discussed in Section 6.1 consider this problem. Protocols that ignore the fact that a car can stay parked in the same parking spot for a long time into consideration may have over-optimistic results. We reported and highlighted results from these approaches in Tables 2 and 3.

In Table 4, the results reported by the different authors are averaged. Only works deemed reproducible and not marked as possibly containing biased results were considered for this average. As one can observe, on average, the accuracy reported for the single parking lot tests is of 96.1%. Nevertheless, when considering the camera and parking lot change scenarios, the results are not as good, on average, reaching accuracies of 92.7% and 91.8%, respectively. In Almeida et al. (2015), where the PKLot is proposed, the authors recognized that the classifiers should be less dependent on the training set and proposed test protocols for this type of scenario. Nevertheless, as one can observe in Figure 12a, most authors focus only on classifying images without considering any change, which is a more manageable problem.
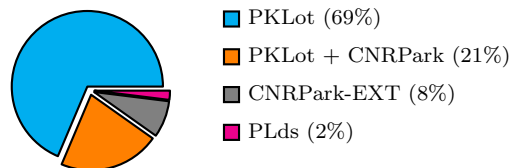
Table 4: Averaged accuracy rates for the feature extraction and deep learning-based approaches. The numbers in parenthesis indicate the number of works considered in the averages.

Approach	Single Park. Lot	Angle Change	Park. Lot Change
Feature extraction-based	94.4% (6)	87.0% (1)	84.4% (2)
Deep learning-based	97.6% (7)	93.4% (8)	93.7% (8)
Average	96.1%(13)	92.7%(9)	91.8%(10)

The datasets usage analysis presented in Figure 12b also shows that only 21% of the works employ more than one dataset in the testing procedure. Still considering Figure 12b, it is possible to verify that the PKLot is the most popular dataset, used in 90% of the works (69% + 21%). The CNRPark-EXT

(a) Types of tests.

(b) Datasets usage.

Figure 12: Types of tests and datasets usage for the individual parking spaces classification problem.

dataset is used in 29% of the works (8% + 21%). The PLds is used in 2% of the works. PKLot and CNRPark-EXT, and PLds datasets were released in 2015, 2017, and 2019, respectively. As mentioned earlier, the high skew in the datasets usage may be (in part) due to the year of publication of each dataset.

When comparing the average results for approaches based on feature extraction and deep learning in Table 4, the results favor the DL-based approaches. However, the comparison made in Table 4 should be taken with care since only a handful of works were used to compute the average. For example, only one feature extraction-based work was considered for the camera angle change scenario. When considering the number of works published, the results in Figure 13a show that the publications are almost evenly split between feature extraction and DL-based approaches.

Figure 13b shows that the most popular descriptor is the LBP and its variants when considering the feature extraction-based works. The most popular classifier is the SVM, as one can observe in Figure 13c, and when considering the DL-based approaches, the most used network is the AlexNet and the networks derived from it, as shown in Figure 13d.

(a) Feature Extraction vs. Deep Learning.

(b) Number of approaches using each feature extractor.

(c) Number of approaches using each classifier.

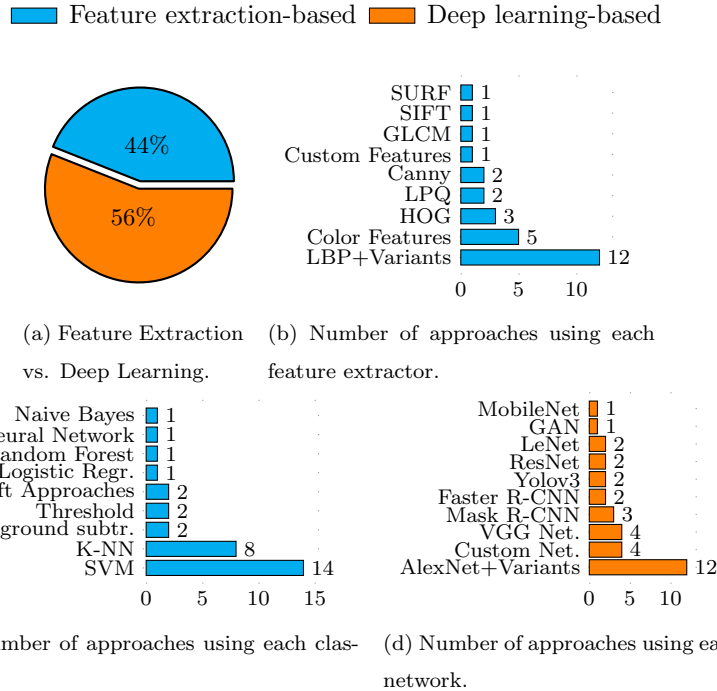(d) Number of approaches using each network.

Figure 13: Individual Parking Spaces Classification Problem – Proportion and number of works taking into consideration different approaches. Works that present multiple approaches or use more than one feature/classifier may be computed more than once.

These findings suggest that research gaps exist for the individual parking spaces classification problem. One research gap is the generalization problem. Future researchers may use multiple datasets to avoid biases, using different datasets for training (e.g., PKLot) and testing (e.g., CNPark-Ext). Standard test protocols considering multiple datasets and standard metrics are also needed to make the comparison between works easier and fairer.

6.3. Automatic Parking Space Detection

As one can observe in Section 5.2, only a few automatic parking space detection works were found. Most discussed approaches did not address the problem, presenting it as an intermediary step for other problems. Except to Padmasiri et al. (2020); Kirtibhai Patel & Meduri (2020); Patel & Meduri (2020), no

author discussed in Section 5.2 made available quantitative results, making it impossible to evaluate and compare the different works.

For the automatic parking space detection, the GAN-based approach proposed in Li et al. (2017) may be highlighted. Although only reporting the individual parking spot classification precision, it generates parking space masks that could be evaluated with appropriate metrics, such as Intersection over Union (IoU) or AP (Everingham et al., 2010). Other works, such as Vítek & Melničuk (2018); Martín Nieto et al. (2019), tried to generate grid-segmentation or fixed-size masks. Nevertheless, a qualitative analysis of the generated parking masks shows that these approaches may be tied to specific parking image properties, such as camera angles and distances.

In Padmasiri et al. (2020) the authors reported results ranging from 59.2 - 63.6 when considering the AP50 metric and 3.16 - 4.75 for the AP75. The authors considered the parking spaces as right angle (90 degrees) bounding boxes, which may not suit angled parking lots, such as the parking lots UFPR04 and UFPR05 of the PKLot dataset. The authors also used images from the same parking lot for both the training and test, which may have led to biased results.

Kirtibhai Patel & Meduri (2020); Patel & Meduri (2020) reported automatic parking space detection results using Faster RCNN and YOLOv4 for the car detection step. Then, they detect parking spaces using a car tracking step. Recall/accuracies of 74.5%/84.79% and 72.27%/76.62%, respectively, were obtained in the CNRPark-EXT dataset. The authors detected right-angle bounding boxes, which may not work in angled parking spaces. Experiments were applied only on three busy days, which may present bias in the test scenario. Also, CNRPark-EXT has only square ground-truths, which do not include the entire parking spaces.

This analysis highlights that automatic parking space segmentation approaches are still needed. It is also imperative that approaches developed for this task present quantitative results. For the quantitative measurement of the results, the authors could use well-known metrics, such as the F1-score, IoU, or AP (Everingham et al., 2010). The metrics used must not rely on true negatives

since the entire parking lot is a true negative, except the parking space polygons. Metrics that require true negatives, such as accuracy, may lead to biased results.

Standard test protocols are also necessary for this task. Robust test protocols should, for instance, use different datasets for the training and testing phases. Using different datasets is necessary since approaches developed to automatically extract the parking spaces should perform without any fine-tuning for the specific parking lot.

The human effort to label training samples from the parking lot is more significant than manually marking each parking space. So how can we expect to use labeled images from the test parking lot to train a method developed for this task, as done in (Li et al., 2017; Bohush et al., 2018; Padmasiri et al., 2020; Kirtibhai Patel & Meduri, 2020; Patel & Meduri, 2020)? With the same reasoning, methods should not rely on camera parameters, such as the distance from the lens to the parking lot, since these parameters are difficult to acquire in the real world.

6.4. Car Detection and Counting

In this section, a discussion about car detection and counting is presented. Only works that present quantitative results are considered. Overviews of car detection, e.g., detect the bounding boxes of the cars; and counting, e.g., the final number of cars in the images, are showed in Tables 5 and 6, respectively. As done in Section 5.1, some results are reported in the $x - y$ format, where x is the worst result achieved in a given scenario, and y is the best result.

Some works employed both car detection and counting, appearing in both tables. Since different metrics and datasets were employed when considering different authors in Tables 5 and 6, a direct comparison of the works is unfeasible.

The average Mean Absolute Error (MAE) of the works presented in Table 6 (MAE is the most common metric in Table 6) is 7.3, which is a relatively small error. However, this result should be considered with care as most of the works employed small subsets of the datasets for the tests. For instance, the PUCPR+

Table 5: Overview of car detection approaches

Authors	Approach	Metric	Reported Results	Datasets Used
Hsieh et al. (2017)	Custom CNN	AR	57.5 - 62.5%	PKLot (PUCPR+)
Laradji et al. (2018)	LC-FCN8	F1-Score	99%	PKLot (PUCPR)
Varghese & Sreelekha (2020)	Bag of Features	Precision / Recall	100% / 97.2%	PKLot (single day)
Li et al. (2019)	Custom CNN	AP@.5 and AP@.7	70.3 - 92.9% / 44.9 - 61.4%	PKLot (PUCPR+)
Sharma & Pandey (2021)	Custom CNN	Mean Average Precision (mAP)	0.85[1]	PKLot

[1] The test procedure may have led to biased results or it is not clear.

Table 6: Overview of car counting approaches.

Authors	Approach	Metric	Reported Results	Datasets Used
Hsieh et al. (2017)	Custom CNN	MAE / RMSE	22.8 - 23.8 / 34.5 - 36.8[1]	PKLot (PUCPR+)
Laradji et al. (2018)	LC-FCN8	MAE	0.2	PKLot (PUCPR)
Amato et al. (2019a); Ciampi et al. (2018)	Mask R-CNN	MAE / RMSE	1.0 / 2.1	CNRPark-EXT
Stahl et al. (2018)	R-FCN-based	MAE	15.1	PKLot (PUCPR+)
Amato et al. (2019b)	Custom CNN	MAE / RMSE	1.8 - 3.7 / 2.7 - 5.1	PKLot (PUCPR+)
Li et al. (2019)	Custom CNN	MAE / RMSE	3.7 - 9.0 / 5.1 - 9.0	PKLot (PUCPR+)
Gabzdyl (2020)	VGG-encoder / Custom CNN	MAE / RMSE	7.5 / 8.8	PKLot (PUCPR+)
Newell et al. (2016)	Stacked Hourglass (mod.)	MAE / RMSE	2.32 / 3.21	PKLot (PUCPR+)

[1] The test procedure may have led to biased results or it is not clear.

(Hsieh et al., 2017) subset of the PKLot dataset contains only one day of data.

When considering the datasets, the PUCPR subset of the PKLot dataset and its extended version, PUCPR+ (Hsieh et al., 2017) is popular for the car detection task. It was used in all but two of the reported works in Tables 5 and 6. Many of the presented works, such as Hsieh et al. (2017); Amato et al. (2019b); Gabzdyl (2020), also employed the CARPK (Hsieh et al., 2017), an interesting dataset containing images captured by a drone. However, this dataset is beyond the scope of this work.

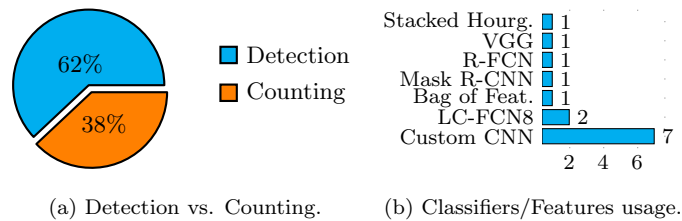(a) Detection vs. Counting.   (b) Classifiers/Features usage.

Figure 14: Types of tasks and approaches used for the car detection and counting problem. Works that deal with detection and counting tasks may be computed twice.

Figure 14a shows that 38% of the approaches may also be used to detect cars. Once the cars are detected in the images, getting the final number of cars becomes a trivial task. Figure 14b shows the number of works using each strategy. As one can observe, the use of custom CNNs is the most popular strategy for car detection and counting task.

Similar to the automatic parking space detection task, discussed in Section 6.3, there is a lack of common protocols and standard metrics for the car detection and counting task. As discussed in Hsieh et al. (2017), another problem in this task is that the datasets must contain a demarcation for all cars present per image in its ground truth. These requirements are not met by any of the datasets discussed in this work. Most authors thus use a subset (often containing all cars manually labeled by the authors of each work) of the original datasets to alleviate this problem. This approach is suboptimal since the subsets are often small, e.g., one day of images, and may not represent the possible

conditions encountered in the real world. It indicates that the datasets' ground-truth discussed in this work could be improved by adding all cars' positions in the ground-truth.

7. Conclusion and Future Works

This paper presented a systematic review of the vision-based approaches that address parking lot management problems. To the best of our knowledge this is the first such vision-centered review. First, considering the reproducibility of results, the public datasets related to vision-based parking lot management were analyzed. Criteria were established based on which the available datasets were filtered to obtain those relevant to this problem. Three image-based datasets passed the established criteria: PKLot, CNRPark-EXT, and PLds datasets. When considering the reviewed works that use these datasets, we concluded that the PKLot is the most popular dataset, used in 88% of the surveyed works. This was followed by the CNRPark-EXT and PLDs datasets, employed by 28% and 1% of the works, respectively.

Although we consider both image and video-based datasets relevant in this work, we found that only the image-based datasets met our restrictions. Thus, robust video-based parking lot datasets can be a future contribution to the scientific community. Video-based datasets could be helpful in various ways, for instance, tracking cars or to identify suspicious behavior. In addition, only a few samples from the analyzed datasets were from night-time and or had snow weather. Thus, datasets containing more of these and other scenarios may be a contribution to future research. Further, datasets containing labels, for example, bounding boxes, for all cars available in each image could also be interesting for future developments, particularly for car counting.

When considering the works that used the surveyed datasets, we discovered that the authors focused on three main tasks: the individual parking spaces classification (between occupied and empty), the automatic parking spaces detection, that is, automatically detect the parking space locations in the images,

and the car detection and counting, for example, counting the number of vehicles in the images.

Overall, 73% of the reviewed works proposed approaches exclusively to address the individual parking spaces classification problem. However, this task has been well-addressed considering that both the train and test images originate from the same parking lot and camera angle, reaching accuracy rates of 96.1% on average. However, we concluded that only 35% of the works that deal with the individual parking spaces classification consider parking lot or camera angle changes.

On average, the accuracy reached in the surveyed works in scenarios containing camera angle and parking lot changes was 92.7% and 91.8%, respectively. Thus, with this in mind, future works should focus on camera and parking lot changes, that is, scenarios where the training images originate from a parking lot or camera angle different from the test images.

Further, only 13% of the surveyed works dealt with the automatic parking spot detection problem, and only three authors made the quantitative results available. This absence of quantitative results can act as an encouragement for significant research efforts for this task. In addition, by reducing the human labor in labeling the parking spaces positions in new parking lots (or when the camera angle changes), automatic parking space detection approaches could lead to more robust and easier to deploy vision-based systems crafted to monitor parking spaces areas.

Of all the surveyed works, 14% considered the car detection and counting task. Most of these works used the PUCPR subset of the PKLot dataset for the tests. Many of these approaches prefer using image datasets from non-fixed cameras, for instance, ones using drones, such as the CARPK (Li et al., 2019). However, although interesting, these datasets are beyond the scope of this work.

There is a lack of standard protocols for testing approaches for all tasks discussed in this work, rendering the comparison of works and reproduction of the experiments a challenge. Future test protocols should consider which evaluation metrics should be used, the manner in which the data should be

split, which datasets exhibit which challenges, and so forth. The proposal and usage of well-defined test protocols could provide better insights and lead to more robust developments.

When comparing handcrafted feature descriptors and deep learning-based approaches, 35% of the surveyed works employed the former, and 61% the latter (4% of the methods used both). For the individual parking spaces classification problem, the method suited to obtain results in not clear. However, on average, deep learning methods tend to generate slightly better accuracies. Moreover, it is crucial to notice that most works lack information about, for instance, the computational cost of the proposed methods. High computational costs could result in certain approaches being prohibitive for certain applications, such as in embedded systems. For the car detection and counting task, all but one of the surveyed approaches employed deep learning-based approaches, clearly indicating that the authors prefer this technique.

Moreover, we also want to point out that many authors included small private datasets in the tests, despite the availability of public datasets. Although we recommend the development and usage of new datasets, the usage of small datasets may lead to biased results. In addition, by making these datasets private, authors are undermining the scientific community's ability to reproduce the results. Further, authors should also consider the specific properties of parking lot image datasets, such as the occurrence of a specific car in several images, to avoid trivial and unrealistic test protocols.

We hope that new research works and datasets can be developed based on the points discussed in this work. As made evident in this work, new developments should consider making the train and test data publicly available. These proposals would improve experiment reproducibility and contribute to the scientific community with new datasets.

References

Acharya, D., Yan, W., & Khoshelham, K. (2018). Real-time image-based parking occupancy detection using deep learning. In P. S. Khoshelham K. (Ed.), *Research@Locate 2018* (pp. 33–40). CEUR-WS volume 2087.

Agrawal, T., & Urolagin, S. (2020). Multi-angle parking detection system using mask r-cnn. In *Proceedings of the 2020 2nd International Conference on Big Data Engineering and Technology* (pp. 76–80). Association for Computing Machinery volume 2020.

Ahrnbom, M., Astrom, K., & Nilsson, M. (2016). Fast classification of empty and occupied parking spaces using integral channel features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1609–1615). IEEE volume 2016.

Ali, K., & Mohamed, G. (2021). Roadside parking spaces image classification using deep learning. In M. R. Senouci, M. E. Y. Boudaren, F. Sebbak, & M. Mataoui (Eds.), *Advances in Computing Systems and Applications* (pp. 323–333). Cham: Springer International Publishing.

Almeida, P., Oliveira, L. S., Silva, E., Britto, A., & Koerich, A. (2013). Parking space detection using textural descriptors. In *2013 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 3603–3608). IEEE volume 2013.

Almeida, P. R., Oliveira, L. S., Britto, A. S., & Sabourin, R. (2018). Adapting dynamic classifier selection for concept drift. *Expert Systems with Applications*, *104*, 67 – 85.

Almeida, P. R., Oliveira, L. S., Britto Jr, A. S., Silva Jr, E. J., & Koerich, A. L. (2015). Pklot–a robust dataset for parking lot classification. *Expert Systems with Applications*, *42*, 4937–4949.

Almeida, P. R. L. d., Oliveira, L. S., Souza Britto, A. d., & Paul Barddal, J. (2020). Naïve approaches to deal with concept drifts. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1052–1059). doi:10.1109/SMC42975.2020.9283360.

Amato, G., Bolettieri, P., Moroni, D., Carrara, F., Ciampi, L., Pieri, G., Gennaro, C., Leone, G. R., & Vairo, C. (2019a). A wireless smart camera network for parking monitoring. In *2018 IEEE Globecom Workshops (GC Wkshps)* (pp. 1–6). Institute of Electrical and Electronics Engineers Inc. volume 2019.

Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., & Vairo, C. (2017). Deep learning for decentralized parking lot occupancy detection. *Expert Systems with Applications*, *72*, 327–334.

Amato, G., Carrara, F., Falchi, F., Gennaro, C., & Vairo, C. (2016). Car parking occupancy detection using smart camera networks and deep learning. In *2016 IEEE Symposium on Computers and Communication (ISCC)* (pp. 1212–1217). Institute of Electrical and Electronics Engineers Inc. volume 2016-August.

Amato, G., Ciampi, L., Falchi, F., & Gennaro, C. (2019b). Counting vehicles with deep learning in onboard uav imagery. In *2019 IEEE Symposium on Computers and Communications (ISCC)* (pp. 1–6). Institute of Electrical and Electronics Engineers Inc. volume 2019-June.

Baktir, A. B., & Bolat, B. (2020). Determining the occupancy of vehicle parking areas by deep learning. In *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)* (pp. 1–4). IEEE.

Baroffio, L., Bondi, L., Cesana, M., Redondi, A. E., & Tagliasacchi, M. (2015). A visual sensor network for parking lot occupancy detection in smart cities. In *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)* (pp. 745–750). Institute of Electrical and Electronics Engineers Inc. volume 2015.

Barriga, J., Sulca, J., León, J., Ulloa, A., Portero, D., Andrade, R., & Yoo, S. G. (2019). Smart parking: A literature review from the technological perspective. *Applied Sciences*, *9*, 4569.

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (surf). *Computer vision and image understanding*, *110*, 346–359.

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, .

Bohush, R., Yarashevich, P., Ablameyko, S., & Kalganova, T. (2018). Extraction of image parking spaces in intelligent video surveillance systems. *Machine Graphics and Vision*, *27*, 47–62.

Bondi, L., Baroffio, L., Cesana, M., Redondi, A., & Tagliasacchi, M. (2015). Ez-vsn: an open-source and flexible framework for visual sensor networks. *IEEE Internet of Things Journal*, *3*, 767–778.

Bura, H., Lin, N., Kumar, N., Malekar, S., Nagaraj, S., & Liu, K. (2018). An edge based smart parking solution using camera networks and deep learning. In *2018 IEEE International Conference on Cognitive Computing (ICCC)* (pp. 17–24). Institute of Electrical and Electronics Engineers Inc. volume 2018.

Chen, L.-C., Sheu, R.-K., Peng, W.-Y., Wu, J.-H., & Tseng, C.-H. (2020). Video-based parking occupancy detection for smart control system. *Applied Sciences*, *10*, 1079.

Chen, Q., Wang, W., Wu, F., De, S., Wang, R., Zhang, B., & Huang, X. (2019). A survey on an emerging area: Deep learning for smart city data. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *3*, 392–410.

Ciampi, L., Amato, G., Falchi, F., Gennaro, C., & Rabitti, F. (2018). Counting vehicles with cameras. In D. N. T. Bergamaschi S., Maurino A. (Ed.), *SEBD* (pp. 1–8). CEUR-WS volume 2161.

Dai, J., Li, Y., He, K., & Sun, J. (2016). R-FCN: object detection via region-based fully convolutional networks.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (pp. 886–893 vol. 1). San Diego, CA, USA: IEEE volume 1.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248–255). IEEE Miami, FL, USA: IEEE.

Dhuri, V., Khan, A., Kamtekar, Y., Patel, D., & Jaiswal, I. (2021). Real-time parking lot occupancy detection system with vgg16 deep neural network using decentralized processing for public, private parking facilities. In *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)* (pp. 1–8). doi:`10.1109/ICAECT49130.2021.9392506`.

Di Mauro, D., Battiato, S., Patanè, G., Leotta, M., Maio, D., & Farinella, G. M. (2016). Learning approaches for parking lots classification. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 410–418). volume 10016 LNCS.

Di Mauro, D., Moltisanti, M., Patane, G., Battiato, S., & Farinella, G. (2017). Park smart. In *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1–5). Institute of Electrical and Electronics Engineers Inc. volume 2017.

Diaz Ogás, M. G., Fabregat, R., & Aciar, S. (2020). Survey of smart parking systems. *Applied Sciences*, *10*, 3872.

Ding, X., & Yang, R. (2019). Vehicle and parking space detection based on

improved yolo network model. In *Journal of Physics: Conference Series* (p. 012084). Institute of Physics Publishing volume 1325.

Dizon, C. C., Magpayo, L. C., Uy, A. C., & Tiglao, N. M. C. (2017). Development of an open-space visual smart parking system. In *2017 International Conference on Advanced Computing and Applications (ACOMP)* (pp. 77–82). IEEE IEEE.

Dobeš, P., Špaňhel, J., Bartl, V., Juránek, R., & Herout, A. (2020). Density-based vehicle counting with unsupervised scale selection. In *2020 Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1–8). doi:`10.1109/DICTA51227.2020.9363401`.

Dornaika, F., Hammoudi, K., Melkemi, M., & Phan, T. (2019). An efficient pyramid multi-level image descriptor: application to image-based parking lot monitoring. *Signal, Image and Video Processing*, *13*, 1–7.

Du, D., Qi, Y., Yu, H., Yang, Y., Duan, K., Li, G., Zhang, W., Huang, Q., & Tian, Q. (2018). The unmanned aerial vehicle benchmark: Object detection and tracking. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 370–386).

Enríquez, F., Soria, L. M., Álvarez-García, J. A., Velasco, F., & Déniz, O. (2017). Existing approaches to smart parking: An overview. In *International Conference on Smart Cities* (pp. 63–74). Springer Springer.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, *88*, 303–338.

Farag, M. S., El Din, M. M., & El Shenbary, H. (2020). Deep learning versus traditional methods for parking lots occupancy classification. *Indonesian Journal of Electrical Engineering and Computer Science*, *19*, 964–973.

Fraifer, M., & Fernström, M. (2016). Investigation of smart parking systems and their technologies. In *Thirty Seventh International Conference on Informa-*

tion Systems. IoT Smart City Challenges Applications (ISCA 2016), Dublin, Ireland (pp. 1–14).

Gabzdyl, D. (2020). Counting vehicles in image and video.

Gregor, M., Pirník, R., & Nemec, D. (2019). Transfer learning for classification of parking spots using residual networks. *Transportation Research Procedia*, *40*, 1327–1334.

Hadi, R. A., & George, L. E. (2019). Vision-based parking lots management system using an efficient adaptive weather analytic technique. In *2019 12th International Conference on Developments in eSystems Engineering (DeSE)* (pp. 522–525). IEEE IEEE.

Hammoudi, K., Cabani, A., Melkemi, M., Benhabiles, H., & Windal, F. (2018a). Towards a model of car parking assistance system using camera networks: Slot analysis and communication management. In *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)* (pp. 1248–1255). IEEE IEEE.

Hammoudi, K., Melkemi, M., Dornaika, F., Benhabiles, H., Windal, F., & Taoufik, O. (2018b). A comparative study of 2 resolution-level lbp descriptors and compact versions for visual analysis. In *Advanced Multimedia and Ubiquitous Engineering* (pp. 221–227). Springer.

Hammoudi, K., Melkemi, M., Dornaika, F., Phan, T. D. A., & Taoufik, O. (2019). Computing multi-purpose image-based descriptors for object detection: powerfulness of lbp and its variants. In *Third International Congress on Information and Communication Technology* (pp. 983–991). Springer Springer.

Hammoudi, K., Taha, M. A., Benhabiles, H., Melkemi, M., Windal, F., El Assad, S., & Queudet, A. (2020). Image-based ciphering of video streams and

object recognition for urban and vehicular surveillance services. In *Fourth International Congress on Information and Communication Technology* (pp. 519–527). Springer Springer.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961–2969). IEEE.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778). IEEE IEEE.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. `arXiv:1704.04861`.

Hsieh, M.-R., Lin, Y.-L., & Hsu, W. H. (2017). Drone-based object counting by spatially regularized regional proposal network. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4145--4153). Springer Springer.

Irfan, A., Nur, N., & Wajidi, F. (2020). Parking slot detection using glcm and similarity measure. In *IOP Conference Series: Materials Science and Engineering* (p. 012094). IOP Publishing volume 875.

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125--1134).

Jensen, T. H., Schmidt, H. T., Bodin, N. D., Nasrollahi, K., & Moeslund, T. B. (2017). Parking space occupancy verification-improving robustness using a convolutional neural

network. In *International Conference on Computer Vision Theory and Applications* (pp. 311--318). SCITEPRESS volume 6.

Kawade, S. S., Kate, V. D., Kotasthane, C. S., Dholay, S. R., & Gajbhiye, C. R. (2020). Survey of vacancy detection techniques in parking lots. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)* (pp. 1--7). IEEE IEEE.

Khan, G., Farooq, M. A., Tariq, Z., & Khan, M. U. G. (2019). Deep-learning based vehicle count and free parking slot detection system. In *2019 22nd International Multitopic Conference (INMIC)* (pp. 1--7). IEEE IEEE.

Kirtibhai Patel, R., & Meduri, P. (2020). Faster r-cnn based automatic parking space detection. In *2020 The 3rd International Conference on Machine Learning and Machine Intelligence* (pp. 105--109).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097--1105).

Lan, R., Zhou, Y., & Tang, Y. Y. (2016). Quaternionic local ranking binary pattern: A local descriptor of color images. *IEEE Transactions on Image Processing*, *25*, 566--579.

Laradji, I. H., Rostamzadeh, N., Pinheiro, P. O., Vazquez, D., & Schmidt, M. (2018). Where are the blobs: Counting by localization with point supervision. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 547--562).

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. et al. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*, 2278--2324.

Li, W., Li, H., Wu, Q., Chen, X., & Ngan, K. N. (2019).
Simultaneously detecting and counting dense vehicles from drone
images. *IEEE Transactions on Industrial Electronics*, *66*, 9651--9662.

Li, X., Chuah, M. C., & Bhattacharya, S. (2017). Uav assisted
smart parking solution. In *2017 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 1006--1013). IEEE IEEE.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P.,
Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft
coco: Common objects in context. In *European conference on computer vision* (pp. 740--755). Springer Springer.

Lowe, D. G. (1999). Object recognition from local scale-invariant
features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (pp. 1150--1157 vol.2). IEEE volume 2.

Mago, N., & Kumar, S. (2020). Role of computers in material
science and design of classification model to search for the
vacancy in outdoor parking lots. *Materials Today: Proceedings*, .

Mahmud, R., Saif, A. S., & Gomes, D. (2020). A comprehensive
study of real-time vacant parking space detection towards the
need of a robust model. *AIUB Journal of Science and Engineering
(AJSE)*, *19*, 99--106.

Manjur Kolhar, A. A. (2021). Multi criteria decision making
system for parking system. *Computer Systems Science and Engineering*, *36*, 101--116. doi:10.32604/csse.2021.014915.

Màrmol, E., & Sevillano, X. (2016). Quickspot: a video analytics
solution for on-street vacant parking spot detection. *Multimedia
Tools and Applications*, *75*, 17711--17743.

Martín Nieto, R., García-Martín, Á., Hauptmann, A. G., &
Martínez, J. M. (2019). Automatic vacant parking places

management system using multicamera vehicle detection. *IEEE Transactions on Intelligent Transportation Systems*, *20*, 1069--1080.

Meduri, P., & Estebanez, A. (2018). A brief review of convolutional neural networks based solutions for smart parking systems. In *2018 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 454--457). IEEE.

Merzoug, M. A., Mostefaoui, A., & Benyahia, A. (2019). Smart iot notification system for efficient in-city parking. In *Proceedings of the 15th ACM International Symposium on QoS and Security for Wireless and Mobile Networks* (pp. 37--42).

Mettupally, S. N. R., & Menon, V. (2019). A smart eco-system for parking detection using deep learning and big data analytics. In *2019 SoutheastCon* (pp. 1--4). IEEE IEEE.

Mora, J. E. G., Lopera, J. C. L., & Cortes, D. A. P. (2018). Automatic visual classification of parking lot spaces: A comparison between bof and cnn approaches. In *Workshop on Engineering Applications* (pp. 160--170). Springer Springer.

Mundhenk, T. N., Konjevod, G., Sakla, W. A., & Boakye, K. (2016). A large contextual dataset for classification, detection and counting of cars with deep learning. In *European Conference on Computer Vision* (pp. 785--800). Springer Springer.

Newell, A., Yang, K., & Deng, J. (2016). Stacked hourglass networks for human pose estimation. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (pp. 483--499). Cham: Springer International Publishing.

Nguyen, T., Tran, T., Mai, T., Le, H., Le, C., Pham, D., & Phung, K.-H. (2021). An adaptive vision-based outdoor car parking lot

monitoring system. In *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)* (pp. 445--450). IEEE.

Nurullayev, S., & Lee, S.-W. (2019).   Generalized parking occupancy analysis based on dilated convolutional neural network. *Sensors*, *19*, 277.

Nyambal, J., & Klein, R. (2017).   Automated parking space detection using convolutional neural networks.   In *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)* (pp. 1--6). IEEE IEEE.

Ojala, T., & Pietikäinen, M. (1999).   Unsupervised texture segmentation using feature distributions.   *Pattern recognition*, *32*, 477--486.

Ojansivu, V., & Heikkilä, J. (2008).   Blur insensitive texture classification using local phase quantization.   In A. Elmoataz, O. Lezoray, F. Nouboud, & D. Mammass (Eds.), *Image and Signal Processing* (pp. 236--243).   Berlin, Heidelberg: Springer Berlin Heidelberg.

Padmasiri, H., Madurawe, R., Abeysinghe, C., & Meedeniya, D. (2020).   Automated vehicle parking occupancy detection in real-time. In *2020 Moratuwa Engineering Research Conference (MER-Con)* (pp. 1--6). IEEE.

Paidi, V., Fleyeh, H., Håkansson, J., & Nyberg, R. G. (2018). Smart parking sensors, technologies and applications for open parking lots: a review.   *IET Intelligent Transport Systems*, *12*, 735--741.

Patel, R., & Meduri, P. (2020). Car detection based algorithm for automatic parking space detection.   In *2020 19th IEEE Interna-*

tional Conference on Machine Learning and Applications (ICMLA) (pp. 1418--1423). IEEE.

Polprasert, C., Sruayiam, C., Pisawongprakan, P., & Teravetchakarn, S. (2019). A camera-based smart parking system employing low-complexity deep learning for outdoor environments. In *2019 17th International Conference on ICT and Knowledge Engineering (ICT&KE)* (pp. 1--5). IEEE IEEE.

Polycarpou, E., Lambrinos, L., & Protopapadakis, E. (2013). Smart parking solutions for urban areas. In *2013 IEEE 14th International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)* (pp. 1--6). IEEE IEEE.

Rahman, S., Ramli, M., Arnia, F., Sembiring, A., & Muharar, R. (2020). Convolutional neural network customization for parking occupancy detection. In *2020 International Conference on Electrical Engineering and Informatics (ICELTICs)* (pp. 1--6). IEEE.

Raj, S. U., Manikanta, M. V., Harsitha, P. S. S., & Leo, M. J. (2019). Vacant parking lot detection system using random forest classification. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 454--458). IEEE IEEE.

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *CoRR*, *abs/1804.02767*.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91--99).

Sairam, B., Agrawal, A., Krishna, G., & Sahu, S. P. (2020). Automated vehicle parking slot detection system using deep

learning. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 750--755). IEEE IEEE.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510--4520). IEEE.

Shah, k. i., Abbas, S., Khan, M., Athar, A., Khan, M., Fatima, A., & Ahmad, G. (2020). Autonomous parking-lots detection with multi-sensor data fusion using machine deep learning techniques. *Cmc -Tech Science Press*, *66*, 1595--1612. doi:10.32604/cmc.2020.013231.

Sharma, R., & Pandey, R. (2021). Swd: Low-compute real-time object detection architecture. In *Innovations in Computational Intelligence and Computer Vision* (pp. 512--520). Springer.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *Computer Science Computer Vision and Pattern Recognition*, .

Stahl, T., Pintea, S. L., & van Gemert, J. C. (2018). Divide and count: Generic object counting by image divisions. *IEEE Transactions on Image Processing*, *28*, 1035--1044.

Suwignyo, M. A., Setyawan, I., & Yohanes, B. W. (2018). Parking space detection using quaternionic local ranking binary pattern. In *2018 International Seminar on Application for Technology of Information and Communication* (pp. 351--355). IEEE IEEE.

Thike, L. L., & Thein, T. L. L. (2019). Parking space detection using complemented-ulbp background subtraction. In *2019 IEEE 8th*

Global Conference on Consumer Electronics (GCCE) (pp. 894--896). IEEE IEEE.

Thomas, T., & Bhatt, T. (2018). Smart car parking system using convolutional neural network. In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 172--174). IEEE IEEE.

Valipour, S., Siam, M., Stroulia, E., & Jagersand, M. (2016). Parking-stall vacancy indicator system, based on deep convolutional neural networks. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)* (pp. 655--660). IEEE IEEE.

Varghese, A., & Sreelekha, G. (2020). An efficient algorithm for detection of vacant spaces in delimited and non-delimited parking lots. *IEEE Transactions on Intelligent Transportation Systems*, *21*, 4052--4062.

Vítek, S., & Melničuk, P. (2018). A distributed wireless camera system for the management of parking spaces. *Sensors*, *18*, 69.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: Association for Computing Machinery. URL: https://doi.org/10.1145/2601248.2601268. doi:10.1145/2601248.2601268.

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems* (pp. 3320--3328).

Zantalis, F., Koulouras, G., Karabetsos, S., & Kandris, D. (2019). A review of machine learning and iot in smart transportation. *Future Internet*, *11*, 94.

Zhang, W., Yan, J., & Yu, C. (2019). Smart parking system based
on convolutional neural network models. In *2019 6th International
Conference on Information Science and Control Engineering (ICISCE)* (pp.
561--566). IEEE IEEE.