Automatic Visual Classification of Parking Lot Spaces: A Comparison Between BoF and CNN Approaches

Jhon Edison Goez Mora(✉), Juan Camilo Londoño Lopera, and Diego Alberto Patiño Cortes

Universidad Nacional de Colombia - Medellín Campus, Medellín, Colombia
{jegoez,jclondonol,dapatinoco}@unal.edu.co

Abstract. Computer vision has a wide and diverse range of applications nowadays. A particular one is automatic detection of parking lot occupancy, where a computer has to identify whether a parking lot space is empty or occupied. As in any visual classification problem, detecting parking lot spaces relies on the existence of a representative visual dataset. This problem of binary classification is commonly approached using features with adequate level of invariance to changes in illumination or rotation, that allow feeding these features into classifiers such as the SVM. Most used approaches are based on the use of convolutional neural networks, some times based on pre-trained models which in general have quite high performance. however several of these methods are tested with common experiments that do not take into account the variations that occur when training with different combinations of angles, lighting variations, and weather types. That is why in this paper we present a comparison between two approaches to solve the problem of parking lot classification with two methods: Convolutional Neural Networks and Bag of Features. In this paper we show how to use the standard Bag-of-features model to learn a visual dictionary, and use it to classify empty and occupied spaces. Results are compared with CNN approaches, emphasizing on accuracy, sensitivity analysis, and execution time.

Keywords: Images processing · Invariance · Parking lot
Convolutional Neural Networks · Bag of features

1 Introduction

One common application of computer vision is automatic surveillance. In this area, a camera, in most cases fixed in an environment which is desired to be a lookout in search of particular objects, persons, threats among other visual objects. A particular case of automatic surveillance is parking lot occupancy automatic detection; where a computer has to identify whether a parking lot space is empty or occupied.

J. C. Figueroa-García et al. (Eds.): WEA 2018, CCIS 915, pp. 160–170, 2018.
https://doi.org/10.1007/978-3-030-00350-0_14

Parking lot spaces classification relies on the existence of a representative visual dataset as any visual classification problem. Recent papers focused on the same problem has successfully used the PKLot dataset [6] created in the Federal University of Parana and the Pontifical Catholic University of Parana.

This problem of binary classification is commonly approached using extractors of characteristics that have properties of invariance to changes of illumination or rotation that can later be fed into classifiers such as the SVM (support vector machines). the most common solutions are based on the use of CNN (convolutional neural networks) either based on pre-trained models, sometimes modified for reducing training time. However, several of these methods are tested with common experiments that do not take into account variations that occur when training and testing with different combinations of sets doing an comparing them at running times, accuracy and precision. That is the reason why in this paper we present a comparison between two approaches to solve the problem of parking lot classification: Convolutional Neural Networks and BoF (Bag of Features). Is fine-tuned the VGG16 CNN and compared our own results against several state-of-the-art results. Later, is used Bag-of-features model to learn a dictionary of visual words and used it to classify empty and occupied spaces. The results were compared emphasizing accuracy, sensitivity analysis, and time performance. This comparison also theorizes some of the reasons why both approaches fails on false positives and false negatives.

This document is organized as follows: In Sect. 2 it's present a short review of several approaches to solving the problem of parking lots classification. Sections 3 and 4 describe the general approach we used to solve the binary classification problem. Section 5 shows the results of applying our approach and the one that uses neural networks with various combinations and tests to each method. Finally, the Sect. 6 provides some remarks about this work and present some discussion and future work.

2 Previous Work

A still open discussion in computer vision is to determinate whether the algorithms of parameters selection or the hand selection of features are suitable for a visual classification problem. The algorithms of parameters selection solve the problem to create an informative set of feature adequate. They have proved to obtain good results but require a large data set in order to train the models. On the other way, hand selection of features are fast and also achieve good results. However, it takes time to design an optimal set of feature to a specific problem. In this short review, it's present some examples of both approaches and highlights some of their advantages and weaknesses.

Several studies on the stated problem based on hand selection of features can be found in recent literature [2,8,9]. In those works, authors use a wide variety of feature sets of which a large proportion exploits the color of information of the images through several color spaces. As an example in [4] used the HSV color space and establish as features the histogram of the hue (H) channel. This channel was selected because of its invariance to rotation and low computational cost. In [1] the authors used the LUV color space. The magnitude of the gradient of the image that indicates the intensity of the change in color space at each point and six channels of quantified gradient as the features for determining the parking lot availability. In [12] only the a* and b* from La*b* color space were used because of their invariance to illumination. All these works use classic classifiers such as SVM or KNN (k-nearest neighbors) to determinate the state of the parking lot with successful results of about 90% of accuracy.

There are also some studies in literature based on Deep Learning and CNN aiming to solve the problem of classifying empty/occupied parking lot spaces. Within this approach, the use of pre-trained models is recurrent. As an example in [3] used LeNet-5 model, and AlexNet model. quoting the reference has two convolutional layers followed by max pooling and two fully connected layers. The first layer (conv1) takes a $224 \times 224 \times 3$ input from the original image. Layers conv2 and fc4 have a reduced number of filters and neurons to better suit a binary classification task without over-fitting. Within fc5, the last Gaussian RBF (radial basis function) layer is replaced with a classical inner product layer acting like a 2-way soft max classifier. They also stated that the detection of free spaces is still an open problem because in most cases the developed solution only fits specific environments and are very difficult to generalize. These models are the base for the construction of a less complex net which could Achieve a more robust solution to changes in lighting and perspective.

In [13] the authors used the VggNet model that is designed to recognize over 1000 classes. The model was fine-tuned to solve the binary problem of the identifies free/occupied parking spaces. In addition, they develop an application for smartphones in order to inform users of the availability of parking spaces. In [5] the authors develop own model of CNN. Unlike other related works, discrimination was made in the tests by the type of climate present in the PKLot Dataset. Better performance is obtained in comparison with the original PKLot paper [6]. In most of the related studies with CNN approaches, the authors argued to have better results than the methods that use other classifiers like SVM.

Finally, found two alternative methods [7,10]: In [7] the parking spaces are presented as a set of surfaces to build a 3d cube for later training a set of weak classifiers that are later merged and ponder to solve the problem. On the other side, in [10] surveillance cameras were used to obtain a binary map of the parking space. As features for classification, they used local entropy average and the standard deviation of the average of the local entropy. The reason for this choice is because they related the state of the parking spaces with the local entropy Which is quantified from the uniformity of the grayscale values of the region.

3 Materials and Methods

3.1 Dataset

For testing and validation of the proposed methods for development of the solution is used the PkLot dataset. PKLot is a robust parking lot occupancy dataset that has visual information about three different parking lots with several camera angles, the images are taken at different day hours, and different weather conditions. PKLot image acquisition was done in two parking lots. One of them uses two cameras with different angles but in different periods of time as shown in the Fig. 1(a) and (b). The other parking lot is bigger, an has a different point of view. The images were taken during a period of 30 days excluding the nocturnal hours arguing the low quality obtained during this hours. PkLot images are classified into three types of weather: sunny, cloudy, and rainy. it's were taken in a range of 8 to 25 days. Every single image were segmented by parking spaces and labeled by the availability of spaces. Table 1 show the distribution and the total segmented images for testing with the proposed methods, additionally, the dataset provides the completely segmented images of the three parking lots for all the days and hours in which the test was made and rotates the parking spaces so that they are left with an orientation of 0 or 90 degrees depending on the angle that has the camera for the parking lot. Based on these images is make the training and testing process for the classification of the parking space.

(a) Parking 04 rainy.

(b) Paking 05 sunny.

(c) Parking PUC cloudy.

Fig. 1. Pklot Dataset

3.2 Methods

For the solution of the parking spaces problem, were implemented two methods in order of compared his performance. The first one is CNN, technique that has already implemented in solving this kind of problem. The second one is the method of Bag of Features, an alternative algorithm to the mentioned in the state of art.

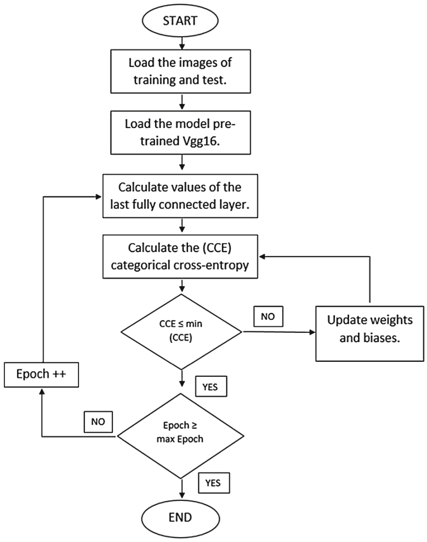
Table 1. PKLot Dataset

Parking lot	Weather condition	Occupied spaces	Empty spaces
UFPR04	Sunny	32166	26334
(28 parking spaces)	Cloudy	11608	27779
	Rainy	2351	5607
UFPR05	Sunny	57584	42306
(45 parking spaces)	Cloudy	33764	23202
	Rainy	6078	2851
PUCPR	Sunny	96762	111672
(100 parking spaces)	Cloudy	42363	90417
	Rainy	55104	27951

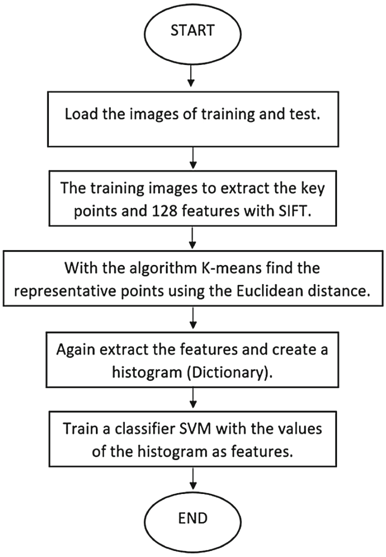
Convolutional Neural Networks. This method is based on the principle of operating of the visual cortex of a biological brain, due to its characteristics poses an advantage for image processing and especially for the comparison process, previously observed the modification of networks such as AlexNet, LeNet and VggNet To work the classification of parking spaces. In this paper, the network Vgg16 is used. That aims to classify 1000 different classes but is modified to become a binary problem (busy, empty) see Fig. 2(a). And a fine-tuning is made modifying the weights of the last layer of the neural network for the specific task of sorting the parking spaces.

Bag of Features. The method consists of creating a set of parking spaces descriptors that will be used as the source of information of a classifier. These descriptors are extracted from the points of interest in the image, such as color changes, corners, edges, among others. Use is made of the SIFT algorithm to extract both the keypoints and the characteristics. These characteristics are extracted for each image resulting in a vector of dimension n x 128 (n is the number of keypoints found in all the images).

With vector of visual words, it's performed a clustering by the method of Euclidean distance for converting the number of clusters selected in the dictionary of visual words. Once the representative points are obtained, it takes new images from the parking lots for compared the "visual words" to the dictionary and so the word recurrence histogram is formed for each image supplied. The histogram has a number of bings equal to the number of clusters. The process described above is done for each class separately, occupied and empty. Later this information is used for the training of an SVM classifier. This process described in the scheme of the Fig. 2(b).



(a) CNN.



(b) BoF.

Fig. 2. Flowchart

4 Methodology

To applicate, the CNN is used the python library Theano to evaluate mathematical expressions and the API Keras, also, is used the pre-trained model Vgg16 and the algorithm of Vgg16 [5] is loaded to do the fine tune of the classes. This algorithm used the categorical cross-entropy as a metric for error minimization and the Adam optimizer to reduce the learning rate during the epochs.The results were compared with the works of the state of art, valid his performance to use as a comparison criterion with the BoF method.

For the Bag of features method, it's use 10000 images that combine all weather types and all the parking lots. Then an iterative algorithm is performed to finding the number of clusterings where is presented best performance. The results of this test determined that the number of clusters more adequate was K = 800 that obtained a performance of 92.04%. The classifier SVM implemented was the Toolbox Balu [11], where is used a Radial Basis Kernel.

With the number of clusters obtained, the tests performed as follows: both methods are exposed to three types of test, one discriminated by the type of parking, climate types are combined and proved to be successful. In the second one test are combined to types of parking and tested the performance of the methods related to climatic variations. Finally is develop an integration of all types of climate and parking using 50% of the images as training and 50% for testing.

5 Experiments and Results

For the compilation used a computer with the next characteristics:

– RAM: 30 GB
– CPU: 8 Cores
– GPU: 8 GB
– HD: SSD

Several experiments are done to test the algorithms. in the first place, the all data base is divided into 70% of the images for doing the training and 30% to do the testing. The Table 2 show the accuracy, precision, and recall obtained for the methods.The better results were for the algorithm CNN. tested in the parking lot PUCPR, with an accuracy, precision, and recall of over 99%. For another side with the results of the BoF method, gets an percents that varies a lot with the type of parking. The worst results are in the parking UFPR04 and the best in the PUCPR which has a fairly high top view that favors the classification.

Table 2. Results of CNN/BoF with all type of weather and parking 70 for train and 30 to test

Parking Lot	Accuracy		Precision		Recall	
	– CNN –	– BoF –	– CNN –	– BoF –	– CNN –	– BoF –
UFPR04 (28 parking spaces)	99.39	76.82	99.60	63.62	99.32	93.11
UFPR05 (45 parking spaces)	99.11	80.84	98.59	63.09	99.24	86.86
PUCPR (100 parking spaces)	99.77	92.04	99.74	92.10	99.84	93.11
All (173 parking spaces)	99.55	85.77	99.46	79.51	99.65	92.11

After performing the test with the total set of images, tests develop discriminated between types of parking and the types of weather. This to find the situations where the algorithms do the task with better performance identifying the invariance versus some changes and identifying the images that represent a challenge.

It trained with each set of parking lots and with the total set of all types of climate, and each trained group is tested with the same training set and with the other parking spaces. Table 3 show the results of the Convolutional Neural Network where obviously the best results are presented when testing with the same set of training and especially when testing with all the set of images. On the other way where there is a considerable decrease in accuracy, it is due to the variations in the parking UFPR04 and UFPR05 that is the same but have very different observation points and angles. The results of the test with the UFPR04 and train with UFPR05 was 94.11%. Also, a possible over-training observed when training with all the images and testing with the parking UFPR05 with an accuracy of 82.2%.

The same proof is made with BoF, the results show a considerable difference with the CNN, the accuracy, and the precision are very low. In the training with the UFPR05 and testing in UFPR04 a precision of 65% and a precision of 38% is presented. This indicates that this parking lot as a reference for training is not very adequate and the alternative method does not achieve a good performance with the small information that is obtained from the parking lot UFPR05.

Table 3. Testing on different parking lot on all type of weathers

Training	Testing	CNN		BoF	
		Accuracy	precision	Accuracy	precision
UFPR04	UFPR04	99.15	99.95	95.26	92.28
	UFPR05	99.68	94.21	90.35	78.63
	PUCPR	97.28	96.74	85.32	77.83
	All	98.05	97.94	88.18	80.67
UFPR05	UFPR04	94.11	98.06	65.10	38.72
	UFPR05	98.98	99.81	89.91	75.95
	PUCPR	93.24	94.89	75.19	57.58
	All	94.77	96.44	76.93	57.09
PUCPR	UFPR04	96.72	98.97	71.63	52.63
	UFPR05	90.37	77.71	77.52	51.91
	PUCPR	99.78	99.92	91.95	91.07
	All	97.11	95.55	85.34	77.07

For the proof with CNN, the same configuration is taken with the training set but for the test is analyzed with each of the types of weather presented in the dataset. This determines with which parking lot presents a better invariance in the face of weather changes and which of the types weather represents a challenge. However it is not a big difference and the precision remains above 90%. the results are observed in the Table 4. For the BoF method, show that the results in accuracy drop considerably obtaining an average of 83.9%.

Once obtained the results with a set of training determined by the three types of parking, we proceed to perform the same tests using the type of climate for the training. For the first experiment with CNN of Table 5, it is obtained that the results of the exactitude stay quite high when training with the weathers and test with each parking lot. the average of the results is 98%, with the best ones being those that were performed with the training of Cloudy tested with each type of parking lot.

The BoF presents on average an accuracy of 81.5%. the type of cloudy weather generates very low precision performance reaching 36% this indicates that many of the parking spaces that were full were identified as empty. While training with the rainy type generates high results over 96% as seen in the

Table 4. Testing different parking lots on different weathers

Training	Testing	CNN		BoF	
		Accuracy	precision	Accuracy	precision
UFPR04	Cloudy	98.99	99.02	86.63	80.62
	Rainy	96.67	95.97	89.51	80.48
	Sunny	97.80	97.71	88.72	80.61
UFPR05	Cloudy	97.67	99.45	71.09	54.27
	Rainy	92.51	98.45	82.29	56.74
	Sunny	93.62	93.73	79.54	61.08
PUCPR	Cloudy	97.39	96.38	84.18	77.73
	Rainy	98.89	98.08	88.19	77.44
	Sunny	96.41	94.39	85.23	76.30

Table 5. Testing on different weathers on different parking lots

Training	Testing	CNN		BoF	
		Accuracy	precision	Accuracy	precision
Cloudy	UFPR04	98.51	99.73	63.86	36.88
	UFPR05	99.10	98.52	76.63	48.52
	PUCPR	99.59	99.70	87.91	79.72
	All	99.29	99.45	81.54	65.83
Rainy	UFPR04	98.73	99.62	75.15	97.93
	UFPR05	98.35	96.39	67.01	98.03
	PUCPR	99.55	99.33	87.13	96.77
	All	99.20	98.89	80.43	97.11
Sunny	UFPR04	98.57	99.88	86.66	84.22
	UFPR05	98.65	99.37	88.29	87.50
	PUCPR	99.67	99.88	92.54	93.27
	All	99.29	99.75	90.59	90.71

Table 5. This method presents some difficulty in recognizing the parking spaces of the different parking lots using only one type of whether to train.

For the last test, the training and test are climate type. With this it is verified which method presents a better invariance to the changes of illumination, brightness and other effects that generate the three types of climate captured in the dataset. In Table 6 observed how distributing the training of the CNN with the types of climate the best results are obtained among all the tests carried out. The average is 99% in accuracy and precision, considering that comparisons are made with rainy days and tested on sunny days these results are quite good and check the effectiveness of the network in the face of changes present by lighting

Table 6. Testing on different weathers on all parking lots

Training	Testing	CNN		BoF	
		Accuracy	precision	Accuracy	precision
Cloudy	Cloudy	99.76	99.84	80.65	69.46
	Rainy	99.44	99.61	84.63	60.63
	Sunny	98.99	99.14	81.58	65.53
Rainy	Cloudy	99.56	99.49	83.40	97.37
	Rainy	99.69	99.61	85.71	96.83
	Sunny	98.80	98.25	77.24	97.18
Sunny	Cloudy	99.51	99.88	90.66	90.93
	Rainy	99.28	99.65	90.16	86.85
	Sunny	99.12	99.70	90.79	91.41

and the previously mentioned factors. In the case of BoF, a result similar to that of the network is obtained, on average an accuracy of 85% is given and at the 87% precision these being the best results as see in Table 6.

6 Conclusions

In each of the proposed methods, there is a small decrease in performance when tested in a different parking lot of which it was trained. however, the results show an improvement in both methods when are trained with the set of the weathers, because despite not having information about variations in lighting it contains information on all types of angle an distances in the PKlot.

The convolutional neuronal network presents a better overall performance against the BoF method; Although the computational cost for training is high, it is a procedure that is performed only in the initial stage, so it is considered that the most appropriate method to implement it is the CNN.

Considering the results obtained in this work, we propose, as future work, to perform practical tests in a real parking lot where a comparison can be made to determine if the computational cost and the execution time inherent to the of the CNN method justifies its implementation by the performance achieved and its advantages over the BoF alternative method.

References

1. Ahrnbom, M., Aström, K., Nilsson, M.: Fast classification of empty and occupied parking spaces using integral channel features. In: Conference on Computer Vision and Pattern Recognition Workshops (2016)
2. Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., Vairo, C.: Deep learning for decentralized parking lot occupancy detection. Expert. Syst. Appl. **72**, 327–334 (2017)

3. Amato, G., Carrara, F., Falchi, F., Gennaro, C., Vairo, C.: Car parking occupancy detection using smart camera networks and deep learning. In: IEEE Symposium on Computers and Communication (2016)
4. Baroffio, L., Bondi, L., Cesana, M., Redondi, A.E., Tagliasacchi, M.: A visual sensor network for parking lot occupancy detection in smart cities. In: 2nd World Forum on Internet of Things (WF-IoT) (2015)
5. Cazamias, J., Marek, M.: Parking space classification using convoluional neural networks (2016)
6. De Almeida, P.R.L., Oliveira, L.S., Britto, A.S., Silva, E.J., Koerich, A.L.: PKLot-a robust dataset for parking lot classification. Expert. Syst. Appl. **42**, 4937–4949 (2015)
7. Huang, C.C., Vu, H.T., Chen, Y.R.: A multiclass boosting approach for integrating weak classifiers in parking space detection. In: 2015 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-TW 2015, vol. 8215, no. c, pp. 314–315 (2015)
8. Huang, C.-C., Vu, H.T.: Vacant parking space detection based on a multi-layer inference framework. IEEE Trans. Circuits Syst. Video Technol. **8215**(c), 1598–1610 (2016)
9. Ichihashi, H., Katada, T., Fujiyoshi, M., Notsu, A., Honda, K.: Improvement in the performance of camera based vehicle detector for parking lot. In: IEEE International Conference on Fuzzy Systems (2010)
10. Mateus, P.A., Maldonado, E.O., Nino, C.L.: Surveillance and management of parking spaces using computer vision. In: 2015 20th Symposium on Signal Processing, Images and Computer Vision, STSIVA 2015 - Conference Proceedings (2015)
11. Mery, D.: BALU: a Matlab toolbox for computer vision, pattern recognition and image processing (2011). http://dmery.ing.puc.cl/index.php/balu
12. True, N.: Vacant Parking Space Detection in Static Images (2007)
13. Valipour, S., Siam, M., Stroulia, E., Jagersand, M.: Neural networks, parking stall vacancy indicator system based on deep convolutional (2016)