# An Approach for License Plate Recognition Based on Temporal Redundancy

Gabriel Resende Gonçalves, David Menotti, William Robson Schwartz
Smart Surveillance Interest Group - Department of Computer Science, Department of Informatics
Universidade Federal de Minas Gerais, Universidade Federal do Paraná
Belo Horizonte - Brazil, Curitiba - Brazil
{gabrielrg, william}@dcc.ufmg.br, menotti@inf.ufpr.br

***Abstract*—Recognition of vehicle license plates is an important task that can be applied for a myriad of real scenarios. Most approaches in the literature first detect an on-track vehicle, locate the license plate, perform a segmentation of its characters and then recognize them using an Optical Character Recognition (OCR) approach. However, these approaches focus on performing these tasks using only a single frame of each vehicle in the video. Therefore, such approaches might have their recognition rates reduced due to noise present in that particular frame. In this work we propose an approach to automatically detect the vehicle on the road and identify its license plate based on temporal redundant information. We also propose a post-processing step that can be employed to improve the accuracy of the system. Experimental results demonstrate that it is possible to improve the vehicle recognition rate in $15.5$ percentage points using our proposal temporal redundancy approach. Furthermore, additional 7.8 percentage points are achieved using the post-processing technique, leading to a final recognition rate of 89.6%. Furthermore, this work also proposes a novel benchmark composed of a dataset designed to focus specifically on the character segmentation step of the ALPR, a new evaluation measure and an evaluation protocol.**

***Keywords*-automatic license plate recognition; pattern recognition; license plate character segmentation; benchmark**

## I. Introduction

Recognition of an on-road vehicle using its license plate is an important task performed by several intelligent transportation systems around the world. This task is known as Automatic License Plate Recognition (ALPR) and plays an important role in many real application scenarios such automatic toll collection, access control in private parking lots, stolen vehicles identification and traffic surveillance. Therefore, many companies and government departments are interested on improving their systems of traffic monitoring, which justifies the need to develop accurate and eficcient approaches to ALPR on uncontrolled environments. Furthermore, there are also other car-related problems that can be improved using modern techniques and new datasets, e.g. vehicle attribute prediction [1] and model classification [2].

ALPR approaches are commonly subdivided into multiple smaller and simpler tasks that are executed sequentially [3]: (i) image acquisition; (ii) vehicle location; (iii) license plate detection; (iv) character segmentation; and (v) optical character recognition (OCR). However, while some approaches have extra stages such as vehicle tracking and frame selection, others skip some of these tasks such as in Prates et al. [4], in which the location of the license plates is performed in the entire scene instead of detecting the vehicle first.

Although some approaches perform vehicle tracking [5], [6], they do not use all captured information to recognize the characters. Instead, they select only a single frame to perform the recognition based on some defined rule [7], [8], rendering the method more sensitive to noise or recognition errors. Therefore, to reduce this problem, we consider multiple frames to solve the vehicle identification problem.

In this work, we propose an entire pipeline to perform real-time ALPR as well as a new temporal redundancy approach that performs the recognition based on multiple frames instead of executing frame selection. To achieve that goal, we perform vehicle tracking using a statistical filtering method to group the detections belonging to the same vehicle. Furthermore, we also propose a two post-processing techniques to improve the results of the recognition/identification using a database of registered license plates and a vehicle appearance classification (VAC).[1]

Our experiments to evaluate the pipeline were performed using a novel dataset composed of $5,200$ samples of $300$ on-track vehicles acquired on an urban road in Brazil. The results demonstrate an improvement of around $15$ percentage points in recognition rate when temporal redundancy information, considering the vehicle tracking is employed. Moreover, we show that we can achieve an additional increase of $7.8$ percentage points when we correct the ALPR results using post-processing step, leading to a final recognition rate of $89.6\%$.

In a automatic license plate recognition system, a precise segmentation is essential to achieve outstanding results (accuracy near 100%) on the next step, the Optical Character Recognition (OCR) [10], [11]. Nonetheless, the License Plate Character Segmentation (LPCS) methods are evaluated considering a large number of different datasets (not always publicly available) and a myriad of evaluation metrics, becoming very hard to compare segmentation approaches. Furthermore, in this work we also propose a technique to perform License Plate Character Segmentation iteratively.

To have a common evaluation environment for the prob-

[1] An extended version of this works is described in Gonçalves et al. [9].

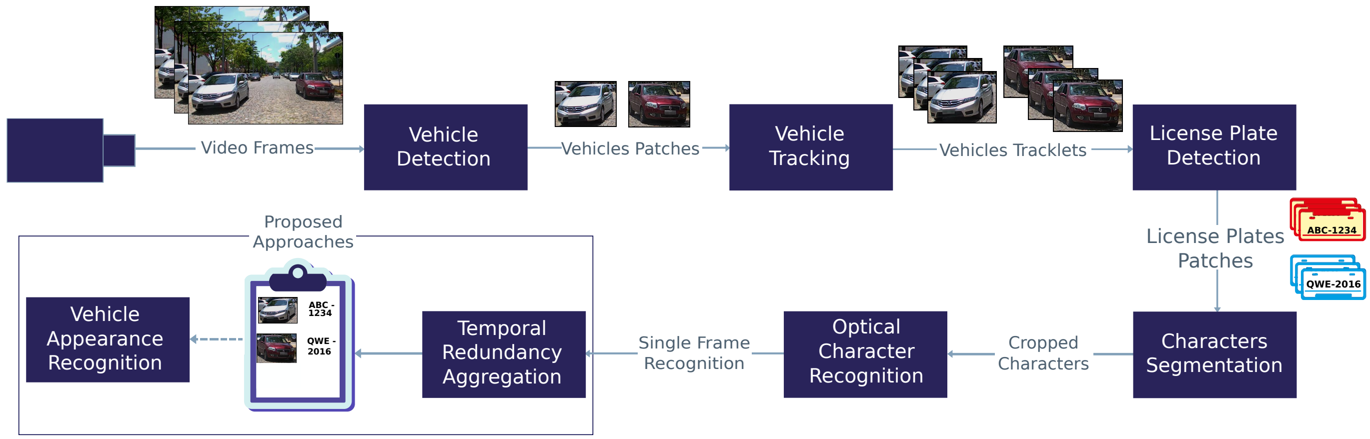


Fig. 1. Sequence of tasks performed by the ALPR. The approaches proposed in this work are highlighted in the rectangle.

lem of license plate character segmentation, this work also proposes a new benchmark to evaluate LPCS approaches. A benchmark is a process to compare multiple approaches under some specific environments. This environment must be fixed for all executions of the process. Our proposed benchmark is composed of a new public dataset and a new evaluation measure developed exclusively to character segmentation problems, which makes it more suitable than the commonly employed Jaccard coefficient.

The main contributions of this work can be pointed as: (i) a new full *Automatic License Plate Recognition* pipeline using temporal redundancy information; (ii) a post-processing technique to improve the final accuracy of the ALPR system; (iii) a iterative approach to perform License Plate Character Segmentation; and (iv) a complete new benchmark to evaluate License Plate Character Segmentation techniques.

Section II describes the proposed real-time ALPR pipeline; a new measure designed to evaluate techniques of character segmentation; a temporal redundancy approach utilized to improve the results of the pipeline using more than one frame; and a post-processing approach that is also conduted to improve the final ALPR results querying a dataset of registered license plates by the vehicle frontal appearance. In Section III, we describe all experiments conducted to evaluate all proposed approaches of this work as well as the proposed LPCS benchmark. Finally, Section IV concludes this work.

## II. Proposed Approaches

As aforementioned, this work proposes a new approach to perform Automatically License Plate Recognition. This section describes the improvements in the pipeline using the redundancy temporal information as well as the post-processing technique to improve the results of the recognition. Figure 1 illustrates the recognition pipeline, described in the next sections.

### A. ALPR Pipeline

Vehicle and license plate detection are crucial tasks on ALPR system. We first detect the vehicle and then its license plate, located inside the vehicle patch. To solve both tasks, we employ a sliding window approach composed of a classifier based on Support Vector Machines (SVM) and Histograms of Oriented Gradient (HOG) [12] as feature descriptors. Afterwards, we track the vehicles over the multiple frames employing the approach described by Kalal et al. [13] to group temporally detections.

Once the license plate has been located, we need to segment the image into multiple patches containing license plate characters (LPCS). For such aim, we developed a straightforward iterative technique to perform LPCS on real scenarios. Starting from a threshold equals 5, we binarize the image as we increase this threshold until we have the number of connected components equals to the number of license plate characters. By doing this, we are trying to avoid the problem where two adjacent characters are touching each other by some noise pixel. Note that when the threshold is too small, we tend to have more connected components due to sliced characters and when the threshold is too large, we have few connected components due to presence of touching characters.

The OCR employed is an one-against-all SVM classifier using HOG features. As a result, we have 36 trained SVMs, one for each character of the Latin alphabet and one for each digit. It is important to note that by knowing the layout of the license plate beforehand (in our case, it has three letters followed by four digits), only the appropriate models can be applied to each character (10 SVM models for digits and 26 SVM models for letters), which reduces the incorrect classification.

### B. Jaccard-Centroid Measure

Since there is no measure in the literature specially designed to evaluate license plate character segmentation approaches, we propose a new measure suitable to this problem, the *Jaccard-Centroid* (JC) coefficient. This measure was inspired

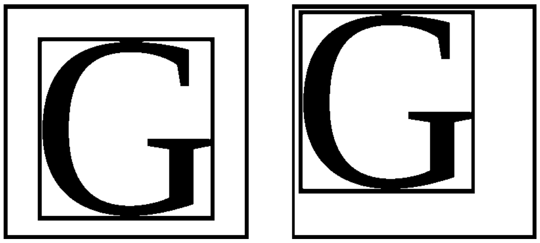


Fig. 2. Illustration of two segmented bounding boxes. Both have the same Jaccard coefficient but one is not well aligned in the centroid, which might difficult the OCR step in the ALPR.

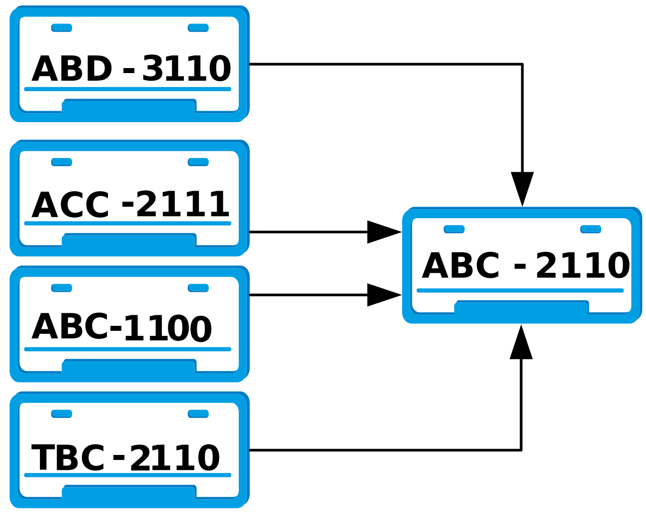


Fig. 3. The proposed approach combine results of multiple frames to improve the vehicle recognition rate.

by the Jaccard coefficient, a widely employed measure to evaluate how well objects are located in images, define by

$$J(A,B) = \frac{A \cap B}{A \cup B}, \quad (1)$$

where $A$ and $B$ are sets constituted by their bounding boxes.

There are two main motivations to create this new measure. First, the Jaccard coefficient is not very suitable to assess whether the location found by an object is well centralized according to the ground truth annotation, which is a very important feature of the segmented character for the further recognition step of the ALPR [10], [11]. For instance, Figure 2 shows two separate bounding boxes with one smaller bounding box inside each. If we consider the inner bounding boxes as the ground truth and the outer boxes as the detection results, they have the same Jaccard coefficients. Second, to the best of our knowledge, most works in the LPCS literature do not employ a standard measure, which makes the comparison of the effectiveness of different techniques a very difficult task.

The *Jaccard-Centroid* (JC) coefficient between two bounding boxes, $JC(A,B)$, is defined as the combination of the Jaccard coefficient and the distance between the centroids of the detected and the desired objects by

$$JC(A,B) = \frac{J(A,B)}{\max(1, C \times \Delta c(A,B))}, \quad (2)$$

where C is a constant and $\Delta c(A,B)$ denotes the distance between the centroids of the detected and the desired objects and is defined by

$$\Delta c(A,B) = \sqrt{(A_x - B_x)^2 + (A_y - B_y)^2}, \quad (3)$$

where $(A_x, A_y)$ and $(B_x, B_y)$ represent their centroid coordinates, respectively. Note that if the centroids are perfectly aligned, the $\Delta c(A,B)$ is zero and the Jaccard-Centroid coefficient will be the same as the Jaccard coefficient.

### C. Temporal Redundancy Aggregation

Since the proposed approach aims at exploring the temporal redundancy information, we hypothesize that the combination of individual results belonging to the same vehicle should improve the recognition of its license plate, as illustrated in Figure 3.

We combine the individual recognition results using two main approaches: (i) majority voting and (ii) average of the classifier confidence. While the former takes all predictions for each frame and assumes that the most predicted character for every license plate position is the correct, the latter averages the classifier confidence and assumes that the class with the highest score is the correct one. In preliminary experiments, we also evaluated the use of the Ranking Aggregation technique proposed by Stuart et al. [14], but the results were not satisfactory.

### D. Post-Processing Techniques

After we recognize the vehicle, it is possible that some vehicle can be misrecognized even using the spatio-temporal information. In this section, we propose a improvement in our method that can be applied when we know the domain of vehicles that can appear in our videos. The main advantage in this method is that we make sure that all recognized vehicles are on the database. We propose this technique based on the fact there are millions of characters combinations that do not correspond to any license plate. For instance, according to the Brazilian Department of Traffic, there are 87 million different license plates in Brazil [2]. However, the combination of three letter followed by four number provide more than 175 milion license plates.

Once we have the vehicle location in multiple frames, we recognize its appearance, which is used then to query the license plate database and retrieve the license plates belonging to vehicles with that appearance. The use of vehicle appearance instead of the recognized license plate itself to select candidates can help the ALPR to discard those candidates that have license plates similar to the correct one but belong to different vehicles models. Therefore, we hypothesize that fewer candidate license plates have to be evaluated, reducing the ALPR recognition error.

[2]http://www.denatran.gov.br/frota2015.htm (in portuguese)

Fig. 4. Two different vehicle models presenting very similar frontal appearance. Voyage (left) vs. Gol (right).

The main challenge of this approach is that several vehicles from the same manufacturer might have the same frontal (or back) appearance, making the distinction of those vehicles a very complex task, even for humans (Figure 4 shows two different models that have very similar frontal appearance). Therefore, we decided to classify vehicles according to their appearance instead of their actual model.

To recognize the vehicle appearance, we employ a standard recognition approach using SVM based on SIFT features and Bag of Visual Words (BoVW) [15]. The only difference between the conventional approaches and the proposed one is in the feature space quantization step of the BoVW. In this work, instead of creating a single global dictionary, we build a dictionary per class and append all codewords generating a large BoVW. Although this approach can generate high-dimensional feature space, it significantly improves the final recognition rate. Furthermore, since our approach considers multiple frames of each vehicle, we recognize the vehicle appearance for each frame and combine all answers using the ranking aggregation technique proposed by Stuart et al. [14].

## III. Experiments

In this section, we briefly discuss the results obtained utilizing the approaches proposed in Section II. First we discuss the results of the LPCS approach using a new proposed benchmark composed of the new measure and a new dataset. Following, we present the results obtained using the new propose ALPR pipeline and the proposed post-processing technique.

### A. LPCS Benchmark Results

We perform an experiment to evaluate our proposed LPCS benchmark using an optical character recognition on characters which were perfectly segmented and characters segmented using the method proposed in Section II. We performed the experiments using a new dataset containing $2,000$ images acquired at the Federal University of Minas Gerais (UFMG) campus. Since the dataset was recorded in Brazil, the license plates have three uppercase letters, one space followed by four numbers, resulting in $14,000$ characters which have been manually annotated with bounding boxes. For more information, we recommend the read of the master thesis described in [16].

To determine the best value of the constant $C$ of the Jaccard-Centroid, we executed the OCR on the $20\%$ best segmented

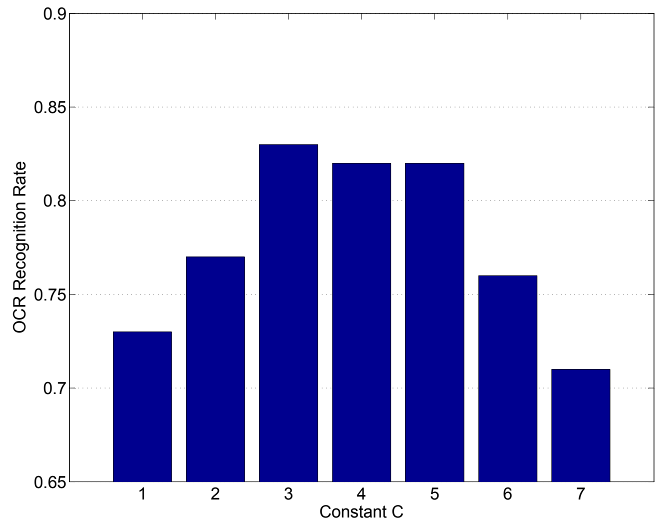


Fig. 5. OCR recognition rates achieved for the first 20% of characters when we vary the value of the constant C from Equation 2.

characters, varying the value of $C$. The best achieved value was 3 as illustrated in Figure 5.

Based on that, all experiments reported on this work were performed using $C = 3$.

We considered two methods available in the literature ae baselines, in which the first aims at improving the quality of degraded images of words [17] and the second performs segmentation by counting the number of black pixels in a binarized license plate based on Connected Component Labeling [18]. In addition, a simple technique that employs prior knowledge regarding the license plate layout and its number of characters was used as the third approach.

Table I shows the average values achieved by the three baselines of segmentation and the proposed iterative approach described in Section II on the testing set of the proposed dataset. According to these results, the baselines approaches do not present promising results, emphasizing the need to the development of new approaches to perform LPCS accurately.

TABLE I
Measure results of segmentation: average values achieved for the three baselines and our proposed approach using three measures.

| Approach | Jaccard | $\Delta c$ | Jaccard-Centroid |
|---|---|---|---|
| Pixel Counting [17] | 0.601 | 2.052 | 0.316 |
| Conn. Component [18] | 0.452 | 1.896 | 0.225 |
| Prior Knowledge-Based | 0.398 | 10.820 | 0.076 |
| Proposed Iterative Approach | 0.601 | 1.433 | 0.419 |

On one hand, the segmentation by the Prior Knowledge-Based approach (expected due to its simplicity) presents the higher average degree of misalignment, represented by the $\Delta c$. As a consequence, this segmentation approach is penalized by the proposed Jaccard-Centroid measure. Therefore, the accuracy of the OCR using the characters segmented by the Prior Knowledge-Based is expected to be reduced due to this misalignment. On the other hand, the connected component

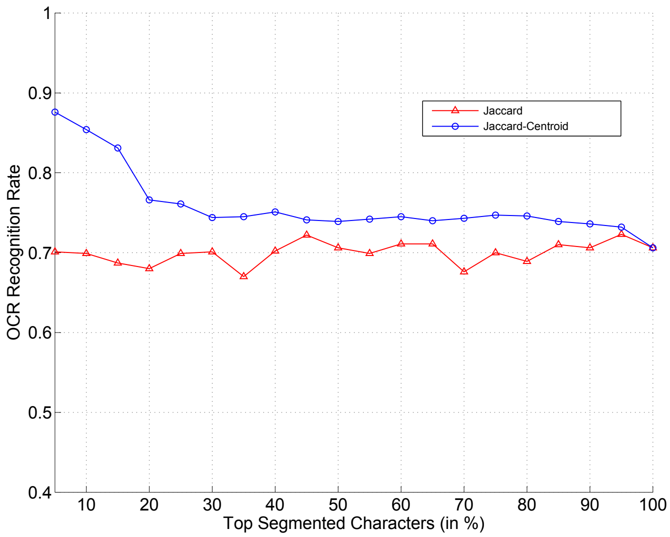


Fig. 6. Recognition rate of OCR as a function of a percentage of the top segmented characters considering Jaccard and Jaccard-Centroid coefficients.

labeling and the pixel counting approaches achieved smaller $\Delta c$ value, causing minor penalization to the Jaccard-Centroid coefficient. The SL*L using Pixel Counting was capable to achieve an average score near $0.60$ by Jacard measure and near $0.30$ by Jaccard-Centroid coefficient, which is the best result of the three proposed baselines.

Our proposed approach was the best evaluated one. It achieves a higher value in Jaccard and it is not much penalized by the $\Delta$ value, which corresponds to low misalignment error. These results supports the hypothesis that our method, despite being straightforward, is the best approach to perform LPCS efficiently.

An accurate segmentation is crucial to an ALPR system once a poor segmentation can lead to a bad final accuracy by the OCR method. To justify it, we performed experiments to evaluate the final accuracy of the OCR when applied to license plate characters segmented with and without a precise segmentation. Given a rank of the characters that are best segmented according to Jaccard and Jaccard-Centroid coefficients, Figure 6 shows the recognition rates of an OCR system when applied to a percentage of the top segmented characters of these ranks. The x-axis represents the proportion of the top characters that were evaluated and the y-axis represents the OCR recognition rate. According to the results, using $5\%$ best segmented characters, the Jaccard-Centroid achieves an OCR recognition rate higher than the one of Jaccard coefficient by around $18$ percentual points. This demonstrates that the proposed Jaccard-Centroid coefficient can assign high values to characters that are easier to be recognized by an OCR better than the Jaccard measure.

### *B. ALPR Results*

In this section, we will describe the experiments performed to evaluate the proposed ALPR pipeline with temporal redundancy. We use two approaches to recognize vehicle using a single frame per vehicle as baselines to evaluate the improvement achieved by the addition of redundancy. Furthermore, this section presents the results achieved when we employ the post-processing technique to perform vehicle appearance classification

TABLE II

RECOGNITION RATES ACHIEVED BY THE PROPOSED APPROACH COMPARED TO THE BASELINE USING MANUAL AND AUTOMATIC CHARACTER SEGMENTATION.

| Approach | Segmentation | |
|---|---|---|
| | Manual | Automatic |
| best frame according to OCR (without redundancy) | 71.3% | 53.5% |
| Bremananth et al. [8] (without redundancy) | 78.3% | 66.3% |
| redundancy with OCR average | 93.6% | 77.9% |
| redundancy with majority voting | 94.6% | 81.8% |

We collected three sets of data to validate the proposed approaches. The first set, used to train vehicle and license plate detectors, contains $650$ images of on-road vehicles used as positive examples to both detectors. The second set, used to evaluate the entire pipeline, contains $300$ on-road moving vehicles extracted from surveillance videos recorded in Brazil. The third set, used for vehicle classification by appearance, contains $1{,}000$ samples divided in $48$ appearance classes corresponding to an average of $20.83$ images per class.

To evaluate the contribution of employing temporal redundancy to the ALPR pipeline, we compare our proposed approach with two baselines: (i) a simple frame selection technique based on the OCR confidence; (ii) the technique proposed in Bremananth et al. [8]. The first baseline is straightforward, the frame selected was the one with the highest average OCR confidence of the seven characters. The second baseline selects the best frame using a machine learning technique that classifies the frame as blurred or non-blurred assuming that the less blurred frame is the most reliable to perform the recognition. We report the results of our approach using two techniques to combine the results: majority voting and average OCR confidence. Furthermore, we perform both automatic and manual segmentation to evaluate the influence of the character segmentation on the final recognition results.

According to the results shown in Table II, the proposed approach using automatic segmentation was able to outperform the best baseline in $11.6$ percentage points (p.p.) using average OCR confidence and $15.5$ p.p. using majority voting, an increase of $17.50\%$ and $23.38\%$, respectively. This fact corroborates the hypothesis that combining the results of multiple vehicle detections can provide better recognition rates than using just a single frame.

As mentioned before, we are proposing a post-processing technique to improve the ALPR results. Once the best results were achieved using majority voting, we utilize the results of this approach as input.

To evaluate our vehicle appearance classification model, we employed a 5-fold cross-validation in the third set of images described earlier. The SVM parameters were set to $\gamma = 10^{-3}$ and $C = 0.5$.

We performed an experiment varying the rank of classes

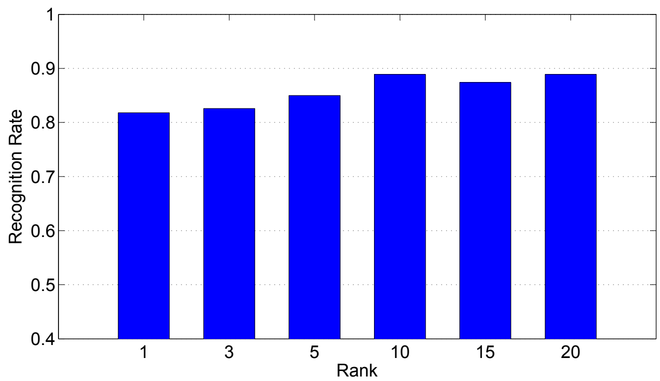


Fig. 7. Percentage of license plates correctly recognized as a function of the amount of license plates evaluated according to rank.

used to predict the license plate, using 900 codewords per class. According to the results shown in Figure 7, the approach achieved 88.9% of recognition rate using the top 10 classes, which is an improvement of 7.1 p.p. compared to the original proposed ALPR approach, as shown in the third row of Table II (81.8%). This supports the claim that classifying a vehicle using its appearance and performing a query on a database can help to improve the ALPR results. Note that the use of more than 10 top predicted classes does not bring significant improvements to the classification.

The proposed temporal redundancy approach was able to significantly outperform the baselines. One can observe that the use of the most reliable frame, using the approach proposed by Bremenath et al. [8], does not provide such high recognition rate as the combination of all images of the same vehicle does. Furthermore, although the results using manual (i.e., perfect) segmentation (Table II) are only theoretical, it is worth noticing the impact of segmentation on the ALPR system. A manual segmentation can improve the results by $12.8 p.p.$ using majority voting and 15.7 p.p. using average OCR confidence, reaching a recognition rate of 94.6%.

## IV. Conclusions

We demonstrated that we can improve the accuracy by 15.5 p.p. using multiple frames to identify the vehicle. In addition, we showed that it is possible to achieve 89.6% of recognition rate using the post-processing proposed approach.

Finally, this work also introduced a new benchmark to the license plate character segmentation (LPCS) problem. This benchmark includes a new dataset with $2,000$ images of 101 different on-road vehicles, spanning a total of $14,000$ alphanumerical symbols (letters and numbers), and a new measure to evaluate the effectiveness of character segmentation approaches called Jaccard-Centroid.

We evaluated our technique and three LPCS approaches as baselines and computed their score on the new dataset. The best result was achieved by our proposed iterative approach. The results demonstrated that the new dataset is very challenging since none of the implemented approaches achieved average values above 0.32 (in a range between 0 and 1) according to the new measure.


## Acknowledgments

The authors would like to thank the Minas Gerais Research Foundation – FAPEMIG (Grants APQ-00567-14 and PPM-00025-15) and the Coordination for the Improvement of Higher Education Personnel – CAPES (DeepEyes Project).


## References


[1] L. Yang, P. Luo, C. C. Loy, and X. Tang, "A large-scale car dataset for fine-grained categorization and verification," in *CVPR*, 2015, pp. 3973–3981.

[2] H.-C. Shih and H.-Y. Wang, "Vehicle identification using distance-based appearance model," in *AVSS*. IEEE, 2015, pp. 1–4.

[3] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 23, no. 2, pp. 311–325, 2013.

[4] R. Prates, G. Cámara-Chávez, W. Schwartz, and D. Menotti, "Brazilian License Plate detection using histogram of oriented gradients and sliding windows," *IJCSIT*, vol. 5, no. 6, pp. 39–52, 2013.

[5] K. Suresh, M. Kumar, and A. Rajagopalan, "Superresolution of license plates in real traffic videos," *ITS, IEEE Transactions on*, vol. 8, no. 2, pp. 321–331, 2007.

[6] N. Sirikuntamat, S. Satoh, and T. Chalidabhongse, "Vehicle tracking in low hue contrast based on camshift and background subtraction," in *JCSSE*. IEEE, 2015, pp. 58–62.

[7] V. Oliveira-Neto, G. Cámara-Chávez, and D. Menotti, "Towards license plate recognition: Comparying moving objects segmentation approaches," in *IPCV*, 2012, pp. 447–453.

[8] R. Bremananth, A. Chitra, V. Seetharaman, and V. S. L. Nathan, "A robust video based license plate recognition system," in *Intelligent Sensing and Information Processing, 2005. Proceedings of 2005 International Conference on*. IEEE, 2005, pp. 175–180.

[9] G. Gonçalves, D. Menotti, and W. Schwartz, "License plate recognition based on temporal redundancy," in *19th International Conference on Intelligent Transportation Systems (ITSC2016)*. IEEE, 2016, (accepted).

[10] D. Menotti, G. Chiachia, A. Falcão, and V. Oliveira-Neto, "Vehicle license plate recognition with random convolutional networks," in *SIBGRAPI*, 2014, pp. 298–303.

[11] L. Araújo, S. Pio, and D. Menotti, "Segmenting and recognizing license plate characters," in *WUW-SIBGRAPI*, 2013, pp. 251–270.

[12] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for human detection," in *CVPR*, 2005, pp. 886–893.

[13] Z. Kalal, K. Mikolajczyk, and J. Matas, "Forward-backward error: Automatic detection of tracking failures," in *ICPR*. IEEE, 2010, pp. 2756–2759.

[14] J. Stuart, E. Segal, D. Koller, and S. Kim, "A gene-coexpression network for global discovery of conserved genetic modules," *Science*, vol. 302, no. 5643, pp. 249–255, 2003.

[15] J. Yang, Y.-G. Jiang, A. G. Hauptmann, and C.-W. Ngo, "Evaluating bag-of-visual-words representations in scene classification," in *Proceedings of the international workshop on Workshop on multimedia information retrieval*. ACM, 2007, pp. 197–206.

[16] G. Gonçalves, D. Menotti, and W. Schwartz, "License plate recognition based on temporal redundancy," 2016, master Thesis in Computer Science.

[17] S. Nomura, K. Yamanaka, T. Shiose, H. Kawakami, and O. Katai, "Morphological preprocessing method to thresholding degraded word images," *Pattern Recognition Letters*, vol. 30, no. 8, pp. 729–744, 2009.

[18] V. Shapiro and G. Gluhchev, "Multinational license plate recognition system: Segmentation and classification," in *Pattern Recognition, (ICPR) International Conf. on*, vol. 4. IEEE, 2004, pp. 352–355.