

**LICENSE PLATE RECOGNITION BASED ON
TEMPORAL REDUNDANCY**

GABRIEL RESENDE GONÇALVES

LICENSE PLATE RECOGNITION BASED ON
TEMPORAL REDUNDANCY

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADOR: WILLIAM ROBSON SCHWARTZ
COORIENTADOR: DAVID MENOTTI GOMES

Belo Horizonte
Agosto de 2016

GABRIEL RESENDE GONÇALVES

LICENSE PLATE RECOGNITION BASED ON
TEMPORAL REDUNDANCY

Dissertation presented to the Graduate Program in Ciência da Computação of the Universidade Federal de Minas Gerais in partial fulfillment of the requirements for the degree of Master in Ciência da Computação.

ADVISOR: WILLIAM ROBSON SCHWARTZ
CO-ADVISOR: DAVID MENOTTI GOMES

Belo Horizonte

August 2016

© 2016, Gabriel Resende Gonçalves.
Todos os direitos reservados.

Gonçalves, Gabriel Resende

G6351 License Plate Recognition Based on Temporal
Redundancy / Gabriel Resende Gonçalves. — Belo
Horizonte, 2016
xx, 55 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de
Minas Gerais

Orientador: William Robson Schwartz
Coorientador: David Menotti Gomes

1. Computação. 2. Aprendizado de máquina. 3.
Reconhecimento de placas de veículos. 4. Visão
Computacional. 4. Reconhecimento de padrões.
I.Orientador. II. Coorientador. III. Título.

CDU 519.6*84(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

License plate recognition based on temporal redundancy

GABRIEL RESENDE GONÇALVES

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. WILLIAM ROBSON SCHWARTZ - Orientador
Departamento de Ciência da Computação - UFMG

PROF. DAVID MENOTTI GOMES - Coorientador
Departamento de Informática - UFPR

PROF. FLAVIO LUIS CARDEAL PÁDUA
Departamento de Computação - CEFET/MG

PROF. THIAGO OLIVEIRA DOS SANTOS
Departamento de Ciência da Computação - UFES

Belo Horizonte, 26 de agosto de 2016.

Acknowledgments

I would like to thank the Brazilian National Research Council – CNPq (Grants #477457/2013-4 and #307010/2014-7), 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).

Resumo

O Reconhecimento de placas veiculares é uma tarefa importante que pode ser aplicada a diversos cenários reais. A maioria das abordagens na literatura primeiro detectam um veículo na pista, localizam sua placa, segmentam os seus caracteres para, finalmente, reconhecê-los utilizando uma técnica de *Optical Character Recognition* (OCR). No entanto, essas abordagens focam em realizar esses passos utilizando apenas um quadro de cada veículo capturado no vídeo. Conseqüentemente, essas técnicas podem ter sua acurácia reduzida devido aos ruídos que podem estar presentes nesse quadro. Por outro lado, neste trabalho nós propomos uma abordagem para localizar o veículo e reconhecer sua placa utilizando a informação de redundância temporal ao invés de selecionar um quadro para realizar o processo. Nós também propomos duas técnicas de pós-processamento que podem ser utilizadas para melhorar a acurácia da abordagem inicial através de consultas à um banco de dados de placas veiculares (por exemplo, o Departamento de Trânsito (Detran) do governo possui uma lista de todas as placas de carros com seus respectivos modelos). Nossos resultados experimentais demonstraram que é possível aumentar a acurácia do método em 15.5 pontos percentuais (p.p.) (um aumento de 23.38%) usando a abordagem de redundância temporal. Ademais, é possível incrementar a acurácia ainda mais em 7.8 pontos percentuais utilizando as duas técnicas de pós-processamento propostas, levando a uma taxa de reconhecimento final de 89.6% em um dataset de 5,200 quadros contendo 300 veículos gravados no campus da Universidade Federal de Minas Gerais. Além disso, esse trabalho também propõe uma nova base de dados, a ser potencialmente utilizado como padrão de estratégia para avaliar técnicas de segmentação de caracteres, sendo composta por 2,000 placas de carros brasileiras (resultando em 14,000 caracteres alfanuméricos), um protocolo de avaliação e uma nova medida de avaliação chamada coeficiente Jaccard-Centroid.

Palavras-chave: aprendizagem de máquina, reconhecimento de placas de veículos, visão computacional, reconhecimento de padrões, segmentação de caracteres, protocolo de avaliação.

Abstract

Recognition of vehicle license plates is an important task applied to 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 the characters 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 techniques might have their recognition rates reduced due to noise present in that particular frame. On the other hand, in this work we propose an approach to automatically detect the vehicle on the road and identify (locate/recognize) its license plate based on temporal redundant information instead of selecting a single frame to perform the recognition. We also propose two post-processing steps that can be employed to improve the accuracy of the system by querying a license plate database (e.g., the Department of Motor Vehicles database containing a list of all issued license plates and car models). Experimental results demonstrate that it is possible to improve the vehicle recognition rate in 15.5 percentage points (p.p.) (an increase of 23.38%) of the baseline results, using our proposal temporal redundancy approach. Furthermore, additional 7.8 p.p. are achieved using the two post-processing approaches, leading to a final recognition rate of 89.6% on a dataset with 5,200 frame images of 300 vehicles recorded at Federal University of Minas Gerais (UFMG). In addition, this work also proposes a novel benchmark, designed specifically to evaluate character segmentation techniques, composed of a dataset of 2,000 Brazilian license plates (resulting in 14,000 alphanumeric symbols) and an evaluation protocol considering a novel evaluation measure, the Jaccard-Centroid coefficient.

Keywords: machine learning, automatic license plate recognition, computer vision, pattern recognition, license plate character segmentation, benchmark.

List of Figures

1.1	Example of the Brazilian license plate standard. It is composed by two rows: in the first one, the acronym of state followed by its origin city (blurred in the image); in the second row, under the first one, there are three letters, a hyphen and four digits to identify the vehicle.	2
3.1	Sequence of tasks performed by the ALPR. The proposed approaches are highlighted in the rectangle.	13
3.2	A sample of a frame in the dataset. Each frame might have more than one vehicle. More details of the dataset are presented in Section 5.2.1	14
3.3	Kalman filter model applied to ALPR.	15
3.4	Samples of the license plate considering different threshold values, 1 and 10 at the top images and 20 and 30 at the bottom images.	16
3.5	The proposed approach combines results from multiple frames to improve the vehicle recognition rate.	18
3.6	Four different vehicle models presenting two very similar frontal appearance. Top: Voyage (left) vs. Gol (right). Bottom: Prisma (left) vs. Onix (right).	19
3.7	Illustration of the tree-based search. Note that the number of candidate license plate is reduced on each iteration.	20
4.1	Example of different license plate colors in the dataset (the plates were blurred due to privacy constraints).	24
4.2	Example of the annotation provided with each image on the dataset.	25
4.3	Frequency distribution of letters in our dataset.	26
4.4	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.	28

4.5	The graphic of the Jaccard Coefficient has a plateau when one box is completely inside the other one. However, the Jaccard-Centroid measure, with $C = 2$, does not has this plateau.	29
5.1	OCR recognition rates achieved for the first 20% of characters when we vary the value of the constant C in Equation 4.2.	32
5.2	SL*L preprocessing. At the top, there is an example of an image binarized without the SL*L preprocessing and at the bottom, there is an image binarized using the SL*L processing method.	33
5.3	The three shadow types (CST1, CST2, and CST3 - in gray) reduced by employing the approach.	33
5.4	Percentage of individual characters correctly segmented as a function of the Jaccard-Centroid coefficient.	36
5.5	Percentage of correctly segmented license plates (all seven characters were segmented in the plate) as a function of the Jaccard-Centroid coefficient.	37
5.6	Recognition rate of OCR as a function of percentage of the top segmented characters considering Jaccard and Jaccard-Centroid coefficients.	39
5.7	Examples of images from the dataset used for vehicle classification.	41
5.8	Recognition rates as a function of the top rank positions.	44
5.9	Percentage of license plates correctly recognized as a function of the amount of license plates evaluated according to rank.	45

List of Tables

4.1	Comparison between the proposed and other datasets available in the literature, regarding different aspects. The proposed dataset is the only one to provide high resolution images with annotation of the individual characters in the license plate, essential to evaluate LPCS approaches.	27
5.1	Measure results of segmentation: values achieved for the three baselines and our proposed approach using three measures.	35
5.2	Recognition rates of the OCR using both segmentation approaches (manual and automatic) and two classifiers (Radial SVM and oRF).	38
5.3	Accuracy of the subtasks utilized on the proposed ALPR pipeline.	42
5.4	Recognition rates (per plate) achieved by the proposed approach compared to the baseline using manual and automatic character segmentation.	43
5.5	Final Recognition rates (per plate) achieved by the proposed approaches when we apply the post-processing techniques.	46

Contents

Acknowledgments	ix
Resumo	xi
Abstract	xiii
List of Figures	xv
List of Tables	xvii
1 Introduction	1
1.1 Motivation	3
1.2 Main Goals	4
1.3 Contributions	4
1.4 Work Organization	5
2 Related Work	7
2.1 Car-Related Problems	7
2.2 Automatic License Plate Recognition (ALPR)	8
2.2.1 Vehicle and License Plate Detection	8
2.2.2 License Plate Character Segmentation	9
2.2.3 Optical Character Recognition	10
2.2.4 Complete ALPR Pipeline	11
3 Automatic License Plate Recognition Approach	13
3.1 Vehicle and License Plate Detection	14
3.2 Vehicle Tracking	15
3.3 License Plate Character Segmentation	15
3.4 Optical Character Recognition	17
3.5 Temporal Redundancy Aggregation	17

3.6	Post-Processing Techniques	17
3.6.1	Vehicle Appearance Classification	18
3.6.2	Tree-Based Query	20
4	License Plate Character Segmentation Benchmark	23
4.1	Dataset	23
4.2	Jaccard-Centroid Measure	26
5	Experimental Evaluation	31
5.1	LPCS Benchmarking Results	31
5.1.1	Parameter Setting	31
5.1.2	Baselines	32
5.1.3	Individual Character Segmentation Evaluation	34
5.1.4	Full License Plate Segmentation Evaluation	36
5.1.5	Optical Character Recognition Evaluation	38
5.2	Proposed ALPR Results	40
5.2.1	Datasets	40
5.2.2	Experimental Setup	41
5.2.3	ALPR Subtasks Results	41
5.2.4	Temporal Redundancy Aggregation	42
5.2.5	Post-Processing Approaches	43
5.2.6	Discussion	46
6	Conclusions and Future Works	47
	Bibliography	49

Chapter 1

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, known as Automatic License Plate Recognition (ALPR), plays an important role in many real application scenarios such as 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 efficient approaches to ALPR on uncontrolled environments. Nowadays, researchers are still studying new approaches to perform ALPR in an efficient way [Rao, 2015; Shih and Wang, 2015]. In addition, there are also other car-related problems that can be improved using modern techniques and new datasets, e.g., simultaneous recognition of multiple vehicles in low-light environments and in high speed highways with low quality samples; vehicle attribute prediction; and car verification.

ALPR approaches are commonly subdivided into multiple smaller and simpler tasks that are executed sequentially [Du et al., 2013]: (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. [2014b], in which the location of the Brazilian license plates (see Figure 1.1) is performed in the entire image plane instead of detecting the vehicle first.

Although some approaches perform vehicle tracking [Suresh et al., 2007; Sirikuntamat et al., 2015], they do not use all captured information to recognize the characters. Instead, based on some defined rule [Oliveira-Neto et al., 2012; Bremananth et al., 2005], they select only a single frame to perform the recognition, rendering the method more sensitive to noise or recognition errors. Therefore, to reduce this problem, we



Figure 1.1. Example of the Brazilian license plate standard. It is composed by two rows: in the first one, the acronym of state followed by its origin city (blurred in the image); in the second row, under the first one, there are three letters, a hyphen and four digits to identify the vehicle.

consider multiple frames to solve the license plate recognition problem.

There are several works in the literature that do not propose methods to solve all subtasks of ALPR [Sivaraman and Trivedi, 2014; Prates et al., 2014a]. Instead, they provide methods to solve only one of them at a time. For instance, in most cases, methods to perform License Plate Character Segmentation are evaluated assuming that the plate and vehicle detection do not have any kind of noise.

This thesis proposes techniques designed in two main directions. The first aims at improving the recognition rate of an ALPR system. The second was developed to be used by the research community to evaluate license plate character segmentation techniques.

One of the proposed approaches comprises the entire a real-time ALPR pipeline based on a 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 two post-processing techniques to improve the results of the recognition/identification using a database of registered license plates and vehicle models. The first is based on vehicle appearance classification (VAC) and the second is based on a search tree containing valid license plates. For both approaches, we can perform a query on a dataset searching for registered license plates and their models to match the identified vehicles and then fix license plates incorrectly identified.

A precise segmentation is essential to achieve outstanding results (accuracy near 100%) on the Optical Character Recognition (OCR) [Menotti et al., 2014; Araújo et al., 2013] in an automatic license plate recognition system. Nonetheless, the License Plate Character Segmentation (LPCS) methods in the literature 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. Therefore,

to have a common evaluation environment for the problem of license plate character segmentation, this work also proposes a new benchmark to evaluate LPCS approaches. Our proposed benchmark is composed of a new public dataset and a new evaluation measure developed exclusively to LPCS problem, which makes it more suitable than the commonly employed Jaccard coefficient.

Our approach was designed to deal with Brazilian license plates. Thus, the two main datasets developed in this work contains licenses plates from Brazil. The first contains 2,000 license plates images (therefore, 14,000 characters) used in the proposed benchmark to evaluate LPCS techniques. The second is composed of 5,200 frame samples of 300 on-track vehicles extracted from surveillance videos on an urban road. As one can see in Figure 1.1, the Brazilian license plate standard is composed of three letters, a hyphen, and four digits.

The ALPR results demonstrate an improvement of around 15 percentage points in the recognition rate when temporal redundancy information is used and vehicle tracking is employed when compared to the baseline. Moreover, we show that we can achieve an additional increase of 7.8 percentage points when we correct the ALPR results using post-processing steps, leading to a final recognition rate of 89.6%, in contrast to 66.3%, achieved by the baseline approach.

1.1 Motivation

Automatic License Plate Recognition is fundamental to perform very important tasks in an automatic way in the real world. Currently, such tasks are only successfully on controlled environments [Du et al., 2013]. Therefore, many companies and government departments are interested on improving their systems of traffic monitoring which justifies the need to develop accurate and efficient approaches to ALPR for uncontrolled environments, and the proposition of improved evaluation benchmarks to evaluate methods focused on ALPR.

An ALPR system is subdivided in multiple simple tasks that are executed in sequence and an inaccurate task may compromise the entire system performance. Therefore, we believe that is possible to improve the system accuracy by upgrading the techniques used in each subtasks. For instance, we can help the vehicle recognition exploiting vehicle appearance using modern machine learning approaches that work with high dimensionality in a computer with high processing power.

1.2 Main Goals

This work aims at developing a new efficient way to identify on-road vehicles in real-time as well as new methods to perform and evaluate ALPR subtasks techniques. We can point out the main objectives of this work as follows.

1. We intend to show that we can utilize multiple frames to improve the recognition of an on-road vehicle. To this purpose, we will try to recognize the same vehicle multiple times and combine the results using some defined rule, i.e. majority voting and average responses.
2. We intend to demonstrate that we can improve the recognition of an ALPR pipeline once we have access to a database containing all possible license plates. We perform the database query using the vehicle appearance and a dynamically tree-search in order to retrieve registered license plates that are alike to the recognized one. Therefore, we can check whether the recognized license plate is valid, since there is a large number of license plates that do not correspond to any vehicle.
3. We will try to demonstrate that our proposed benchmark is more suitable to evaluate License Plate Character Segmentation techniques than the conventional ones. We intend to do this by showing that the highest-rated methods according to our benchmark provide segmented character easily to be recognized. Our benchmark contains a new measure, called Jaccard-Centroid, that is more suitable to evaluate LPCS techniques than the commonly employed Jaccard measure. This is necessary since the Jaccard measure does not consider the alignment of the detected and ground-truth bounding-boxes, which is highly-important to a good segmentation technique aiming a success in the next step, the character recognition.

1.3 Contributions

The main contributions of this work can be pointed out as follows:

- a new fully *Automatic License Plate Recognition* pipeline using temporal redundancy information to combine recognitions belonging to the same vehicle;
- two post-processing techniques that query a license plate database and are aimed to improve the final accuracy of the ALPR system;

- an straightforward iterative approach to perform license plate character segmentation;
- a new measure to evaluate character segmentation techniques that is more suitable than the normally employed Jaccard technique;
- a complete new benchmark to evaluate license plate character segmentation techniques.

During the development of this work, a technical paper entitled “License plate recognition based on temporal redundancy” containing part of the contributions of this thesis (ALPR pipeline and post-processing approaches) was published in the proceedings of the 19th International Conference on Intelligent Transportation Systems (ITSC2016) [Goncalves et al., 2016b]. Furthermore, another work concerning the entire benchmark of LPCS entitled "Benchmark for license plate segmentation" was published in the Journal of Electronic Imaging [Goncalves et al., 2016a].

1.4 Work Organization

In Chapter 2, we review works that are related to Automatic License Plate Recognition as well as works proposing other important car-related problems. Chapter 3 describes the proposed real-time ALPR pipeline; a temporal redundancy approach utilized to improve the results of the pipeline using more than one frame; and two post-processing approaches that are also conceived to improve the final ALPR results querying a dataset of registered license plates. In Chapter 4, we present in details the new benchmark developed to evaluate LPCS techniques. In Chapter 5, we describe all experiments conducted to evaluate all proposed approaches of this work as well as the proposed LPCS benchmark. Also, all achieved results are described and discussed in details. Finally, Chapter 6 concludes this work and discusses some perspectives for future works.

Chapter 2

Related Work

In this chapter, we present a literature overview describing some works related to the main goals of this master thesis. The reviewed papers are divided in two groups: general car-related problems and works related to automatic license plate recognition. The latter is subdivided into four main categories: works presenting approaches to other car-related problems, works related to vehicle and license plate detection, license plate character segmentation, optical character recognition and works concerning an entire ALPR pipeline.

2.1 Car-Related Problems

Nowadays, many researchers have turned their attention to vehicle-related problems, such as vehicle classification according to a set of characteristics, other than license plate identification [Shih and Wang, 2015; Dong et al., 2014; Hsieh et al., 2015; Tang et al., 2015; Yang et al., 2015]. Shih and Wang [2015] proposed a technique to recognize the vehicle using its appearance instead of the license plate characters. They employed Chamfer distance transform to create the Vehicle Appearance Model based on the training samples and compared with the testing samples using points extracted by Speeded-Up Robust Features (SURF) [Bay et al., 2006]. Dong et al. [2014] presented an approach to classify vehicles into six categories (bus, microbus, minivan, SUV, sedan and truck) using a Convolutional Neural Network (CNN). Their model achieved a high accuracy on two datasets by learning discriminative features using a CNN as well.

Tang et al. [2015] proposed a technique to recognize the vehicle model using Haar-like features and an Adaboost algorithm to locate the vehicle in the image and a Gabor wavelet transform combined with Local Binary Pattern (LBP) to extract features for the model classification. Their approach was able to achieve 91.6% of recognition

rate using a dataset with 227 images of eight different vehicle models. Hsieh et al. [2015] classified the vehicle according to color applying a correction to reduce the effect of the lightning change. They state that there is a major problem with color-based classifications, which is the presence of shading colors (i.e. white vs silver and gray vs white). However, the authors propose a new tree-based classifier to overcome this problem. This was created to classify the sample into chromatic and non-chromatic using their non-chromatic strengths. They showed that this could increase the accuracy of the system significantly.

Recently, Yang et al. [2015] introduced a new large dataset called CompCars containing 136,727 car images. The authors showed that there are still many car-related problems that are not well-explored by the research community, i.e. fine-grained classification, vehicle attribute prediction and car verification. Despite the dataset is publicly available, our post-processing step based on vehicles classification cannot use it because all cars in our experiments must have a corresponding class on the dataset in order to make sure that all vehicles in the test set have a possible correct prediction by the machine learning classifier. Therefore, in this master thesis, we propose a set with 1,000 vehicles divided into 48 classes according to their frontal appearance. In addition, the datasets provided by Dong et al. [2014] and Tang et al. [2015] have very few classes compared to ours, which makes them less suitable for filtering unlikely license plate candidates as we are proposing on our post-processing technique.

2.2 Automatic License Plate Recognition (ALPR)

In this section, we present some works related to License Plate Recognition and its subtasks. It is divided into four parts: detection of vehicles and license plates; segmentation of license plate characters; optical character recognition; and full ALPR pipeline.

2.2.1 Vehicle and License Plate Detection

The preliminary tasks performed in ALPR is vehicle and license plate detection, usually solved using approaches such as connected components labeling (CCL) [Caner et al., 2008; Wen et al., 2011], template matching [Betke et al., 2000], background separation [Jazayeri et al., 2011], and more often machine learning techniques [Zhang et al., 2006; Prates et al., 2014b; Wen et al., 2015]. In Caner et al. [2008], a binarization is applied to the image and the detection is made by analyzing the components area in terms of pixels. Betke et al. [2000] detects the vehicle or the license plate by matching a

pre-established template with multiple patches from the test image. Also, they detect highway scene features and crop the vehicle based on these features. In Jazayeri et al. [2011], the goal is to compare two consecutive frames in order to discover which part of the image is in movement. Therefore, they assume that everything static on the image is background and can be removed from the analysis. In the approaches using machine learning, a window is slid on the image and classified as whether containing or not a license plate (or a vehicle) according to feature descriptors extracted from each image location. The work described in Sivaraman and Trivedi [2014] compares three methods to perform vehicle detection using active learning. Furthermore, there are others important works in the field such as the one by Chen et al. [2011] that proposes a new system to perform night time vehicle detection and by Kembhavi et al. [2011] which proposes an approach to detect vehicles on aerial cameras using Partial Least Squares.

In our work, we also utilize machine learning techniques to detect a vehicle and its license plate utilizing a sequence of two classifications. We first detect the vehicle in the frame and after we try to locate its license plate using only the vehicle region in order to avoid to search the license plate in absurd regions.

2.2.2 License Plate Character Segmentation

Since one focus of this work is license plate character segmentation (LPCS) and on datasets for this problem, this section focuses on two aspects. First, we review some works related to techniques of character segmentation. Then, we present works that propose character segmentation evaluation datasets used in different contexts.

Du et al. [2013] classify the license plate character segmentation techniques into five main categories: based on pixel connectivity, pixel projection, prior knowledge of the characters, characters contours and based on the combination of these features.

Since the license plate images might contain artifacts such as skew transform, shadows and blurring, generated during the image acquisition process, one of the most challenging tasks in ALPR is the character segmentation [Araújo et al., 2013; Soumya et al., 2014; Wang et al., 2013; Xing-lin and Yun-lou, 2012]. Araújo et al. [2013] proposes a technique to segment characters using CCL and showed that the OCR results are greatly affected by the character segmentation step. For instance, while they achieved recognition rates of 95.59% for manually segmented license plates, only 71.15% was obtained when automatic segmentation was performed. Such behavior is corroborated by our work (see Table 5.4).

The approach proposed by Soumya et al. [2014] performed character segmenta-

tion by counting the black pixels in the horizontal and vertical direction of each license plate region. Wang et al. [2013] employed a sequence of techniques to improve the segmentation based on vertical projection and an A* pathfinding algorithm. In [Xinglin and Yun-lou, 2012], the authors proposed a technique to segment the characters using prior knowledge regarding the shape and the font used in the license plate considering license plates containing English (Latin) and Chinese characters. In addition, image quality improvement can help to achieve better segmentation results as shown by Chuang et al. [2014] that employs a super-resolution technique to this purpose.

Some statistical-based and machine learning approaches also have been employed on LPCS. For instance, Fan et al. [2012] used likelihood maximization to find the best parameters values of the license plate features and its characters and Franc and Hlaváč [2005] proposed a technique using Hidden Markov Models to create a relationship between the license plate input and the correct segmentation of its characters, Nagare [2011] and Guo and Liu [2008] employed supervised machine learning techniques to aid the character segmentation phase of the ALPR. Our proposed technique is a straightforward one as will be explained in the Chapter 3.

There are works proposing datasets to evaluate several aspects of text recognition and document analysis. For instance, in Antonacopoulos et al. [2009], the authors propose a dataset to evaluate techniques of document layout analysis. That dataset contains 1,240 images from websites, newspaper pages, magazines pages. The UNIPEN dataset was proposed in Guyon et al. [1994] and is composed by over 23,000 images of words and handwritten characters. Yao et al. [2012] proposed a dataset of real images to evaluate approaches to perform text detection containing 500 real images in various sizes. On the other hand, in this work we propose a large dataset designed specifically to evaluate segmentation of license plate characters instead of evaluating general approaches of text segmentation. Moreover, we also propose a new suitable measure to be used within the dataset.

2.2.3 Optical Character Recognition

This section presents the last step of the ALPR, which is the optical character recognition to identify each of the letters and digits of the license plate. For this goal, there are works producing outstanding results that use deep learning techniques, such as Convolutional Network, to perform character recognition in noisy images [Menotti et al., 2014; Netzer et al., 2011; Sermanet et al., 2012; Goodfellow, I. J. and Bulatov, Y. and Ibarz, J. and Arnaud, S. and Shet, V., 2014]. It is important to note that the majority of those works perform the OCR taking into account 36 classes (26 for letters

and 10 for numbers).

The work described in [Netzer et al., 2011] compares two deep learning techniques over two hand-designed descriptors using optimized classifiers in SVHN database. They show that the deep learning approaches can provide much better results than the conventional techniques. In [Sermanet et al., 2012], the authors applied a multi-stage characteristic approach combined with different pooling methods in a traditional convolutional neural network. Furthermore, Goodfellow, I. J. and Bulatov, Y. and Ibarz, J. and Arnoud, S. and Shet, V. [2014] proposed an approach that, at the same time, can locate, perform segmentation and recognize texts in a image using deep learning. Although most works employ learning-based techniques to perform the OCR step, there are also works producing promising results that use template matching to perform it [Araújo et al., 2013; Shuang-Tong and Wen-ju, 2005].

When a specific license plate layout is known a priori, the lexicon size can be diminished (for instance, plates with 3 letters and 4 digits in a sequence) and the classification accuracy can be improved. Therefore, since it is the case of the Brazilian license plates, we utilized this information in order make our pipeline more effective. Note that in the license plate recognition scenario, an OCR has to work as close as possible to the optimality (100% of recognition rate) since a single mistake in the character recognition task may imply in an incorrect identification of the vehicle.

2.2.4 Complete ALPR Pipeline

Although many works propose approaches to solve a single subtask at a time, there are works proposing techniques to perform the entire ALPR pipeline [Donoser et al., 2007; Guo and Liu, 2008; Kocer and Cevik, 2011; Ozbay and Ercelebi, 2005; Qadri and Asif, 2009; Wang et al., 2010].

The approach presented in Donoser et al. [Donoser et al., 2007] utilizes analysis of Maximally Stable Extremal Region (MSER) to detect the license plate, track the vehicle and segment its characters. The characters are recognized using a SVM-based OCR. Furthermore, they also combine multiple detections in order to make the recognition robust to noises presented in a single frame. The work proposed by Wang et al. [2010] proposed a technique able to locate the license plate using horizontal scans of contrast changes, segment the plate using lateral histogram analysis and recognize the characters using an Artificial Neural Network for Italian license plates. Kocer and Cevik [2011] proposed a work to locate the region of the image with the most transition points assuming that it corresponds to the license plate. The characters are then segmented using a blob coloring method and the characters are recognized using a multi

layered perceptron.

Guo and Liu [2008] detects the license plate using template matching, segments the characters using vertical and horizontal projections and recognizes Dutch license plates using Hotelling transform and Euclidean distance. In Ozbay and Ercelebi [2005], the authors detect the license plate using edge detection and smearing algorithms, segment the characters utilizing filters and morphological operation and recognize the characters using a template matching OCR. Finally, in Qadri and Asif [2009], their input is a cropped vehicle-rear image. Since the handled license plate is yellow with black characters, they segment the characters identifying the yellow pixels and recognize them using a template-matching-based OCR.

Different from the aforementioned works, we propose an approach to perform the entire ALPR pipeline using a sequence of classifications. We track the vehicle and combine multiple recognitions of the same car in order to make the approach robust to noise present in one single frame.

Chapter 3

Automatic License Plate Recognition Approach

This work proposes a methodology to perform automatic license plate recognition (ALPR). It is important to mention that in this work the entire pipeline is implemented using C++ and the library OpenCV 3.0 assisted by the Smart Surveillance Framework [Nazare et al., 2014]. This section describes the improvements in the pipeline using the temporal redundancy information as well as the two post-processing techniques to improve the results of the recognition. Figure 3.1 illustrates the recognition pipeline, described in the next sections.

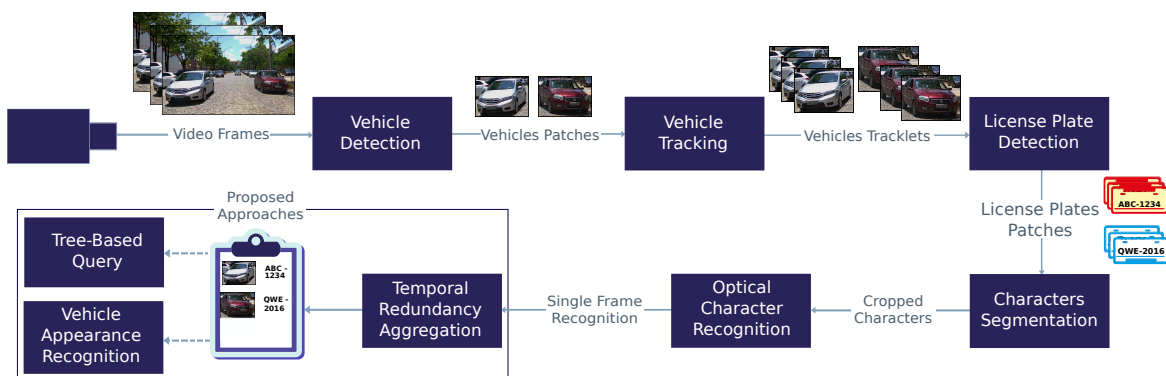


Figure 3.1. Sequence of tasks performed by the ALPR. The proposed approaches are highlighted in the rectangle.

3.1 Vehicle and License Plate Detection

Vehicle and license plate detection are crucial tasks on ALPR system. As one can see in Figure 3.2, each frame may contain multiple vehicles. Thereby, the approach has to be capable of detecting more than one vehicle at time. Therefore, we first detect the vehicle and then its license plate, located inside the vehicle bounding box.



Figure 3.2. A sample of a frame in the dataset. Each frame might have more than one vehicle. More details of the dataset are presented in Section 5.2.1

To solve both tasks, we employ a sliding window approach composed of a classifier based on Oblique Random Forest (oRF) [Jordao and Schwartz, 2016] and Histograms of Oriented Gradient (HOG) [Dalal and Triggs, 2005] as feature descriptors. Oblique Random Forests utilize a modified version of the conventional decision trees. Instead of a single feature-threshold (calculated by the entropy) to split the data in each node orthogonally, these decision trees utilize a weak classifier to perform an oblique split on the data. In our problem, we utilize one Partial Least Squares classifier on each node of each tree. On the other hand, we utilize HOG as feature descriptor. HOG describes the form of the image by counting the orientations of its gradients. It performs the counting in blocks in an attempt to introduce small translation, rotation and scale invariance to the final feature vector.

3.2 Vehicle Tracking

Before performing the license plate detection, we employ a tracking approach to group detections belonging to the same vehicle. For this purpose, we use a Kalman filter, which has been widely utilized to perform object tracking with some noises on transitions [Shantaiya et al., 2015]. The filter is based on two models: a transition model and a measurement model. In our context, the former explains the vehicle movement in the scene (speed, acceleration). The latter describes the observation obtained by the sensors. In this case, they are the cameras and the observation is the vehicle detection on each frame. The Kalman filter is expressed as

$$X_k = F(X_{k-1}, U_{k-1}) + W_k, \quad (3.1)$$

where X_k is the current state, F is the transition model applied to the last state (X_{k-1}) and the sensor observation (U_{k-1}), and W_k is the presented noise. This stage outputs a tracklet to each vehicle present in the sequence of frames. The model of the Kalman filter applied to our methodology is illustrated in Figure 3.3.

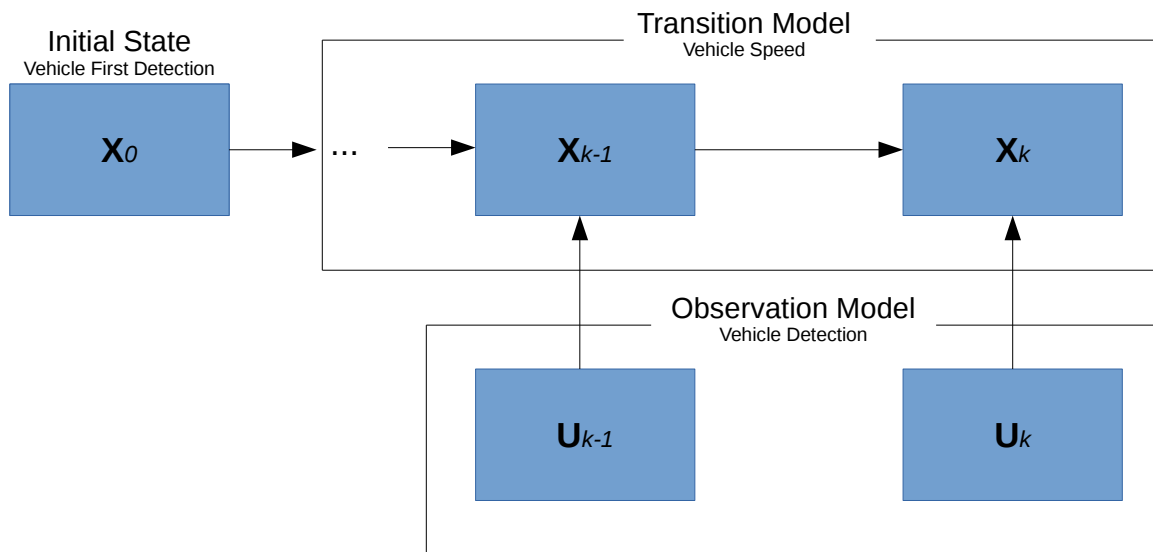


Figure 3.3. Kalman filter model applied to ALPR.

3.3 License Plate Character Segmentation

Once the license plate has been located, we need to segment it into multiple patches containing the characters. The License Plate Character Segmentation (LPCS) stage is

one of the most challenging tasks of an ALPR system due to several noise effects that can appear on a license plate image.

We developed a straightforward iterative technique to perform LPCS on real scenarios. It is composed of two steps: (i) thresholding and (ii) estimation of connected components. A similar idea was used in Matas and Zimmermann [2005] to find threshold for images containing cars and license plates.

In our LPCS approach, instead of using a single threshold to perform license plate binarization using the Otsu method [Gonzalez, 2009], we consider a set of different values. Starting from a threshold equals 10, we threshold 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 avoid the problem where two adjacent characters are touching each other due to noisy pixels, once a thresholding operation starting from small thresholds tends to set most pixels belonging to the background to the maximum value, resulting in fewer noise connecting two adjacent characters. At each iteration, we discard connected components that are too large and too small to be a character according to the width and the area of the component. We also merge all connected components that have overlap on the x -axis since in Brazilian license plate standard, there are only one text without split in two lines.

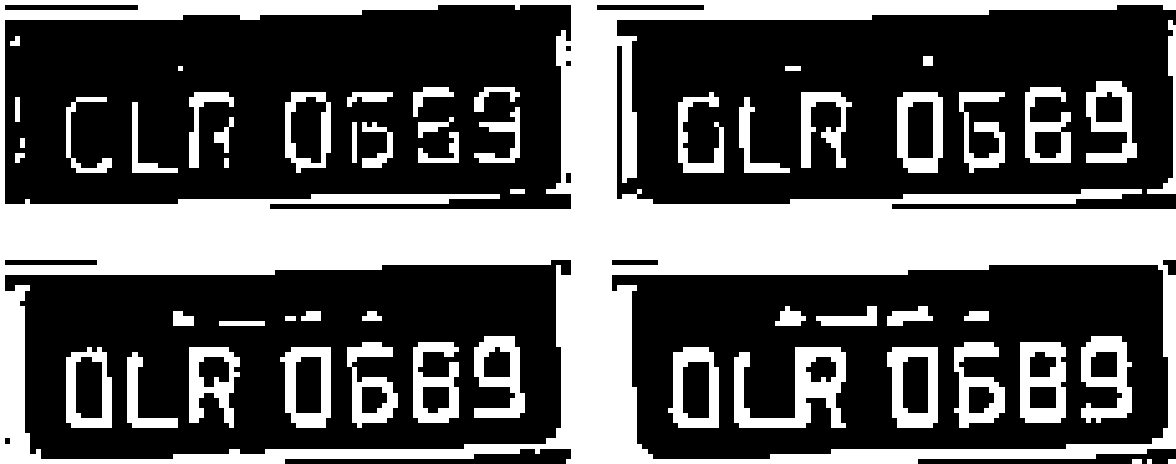


Figure 3.4. Samples of the license plate considering different threshold values, 1 and 10 at the top images and 20 and 30 at the bottom images.

Figure 3.4 illustrates the thresholding process. 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.

Although this technique is able to provide accurate segmentation, it is not perfect

for every license plate detected. Nonetheless, we minimized the impact of the incorrect segmentation using the temporal redundancy information on the recognition step (see Section 3.5).

3.4 Optical Character Recognition

Our OCR system is an one-against-all version of the support vector machines (SVM) classifier. Its input is the bounding box of a previously character segmented (described in the last section). As a result, we have 36 (26 for letters and 10 for digits) trained SVMs, one for each character of the Latin alphabet and one for each digit. Furthermore, once we know the license plate layout a priori (in our case, it has three letters followed by four digits), the SVM models can be easily trained, since it is possible to know whether the character is a letter or a number based on the position of the license plate character. For instance, an SVM to recognize the letter 'O' does not have any image of the number '0' as negative example, which reduces the incorrect classification. We also tried to employ the use of the Google's text recognizer called Tesseract Smith [2007] in order to recognize the license plate as a text but the results were not satisfactory.

3.5 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 3.5.

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. In preliminary experiments, we also evaluated the use of the Ranking Aggregation technique proposed by Stuart et al. [2003], but the results were not promising.

3.6 Post-Processing Techniques

In this section, we detail two post-processing techniques employed to fix incorrectly recognized license plates. These approaches are based on the assumption that we

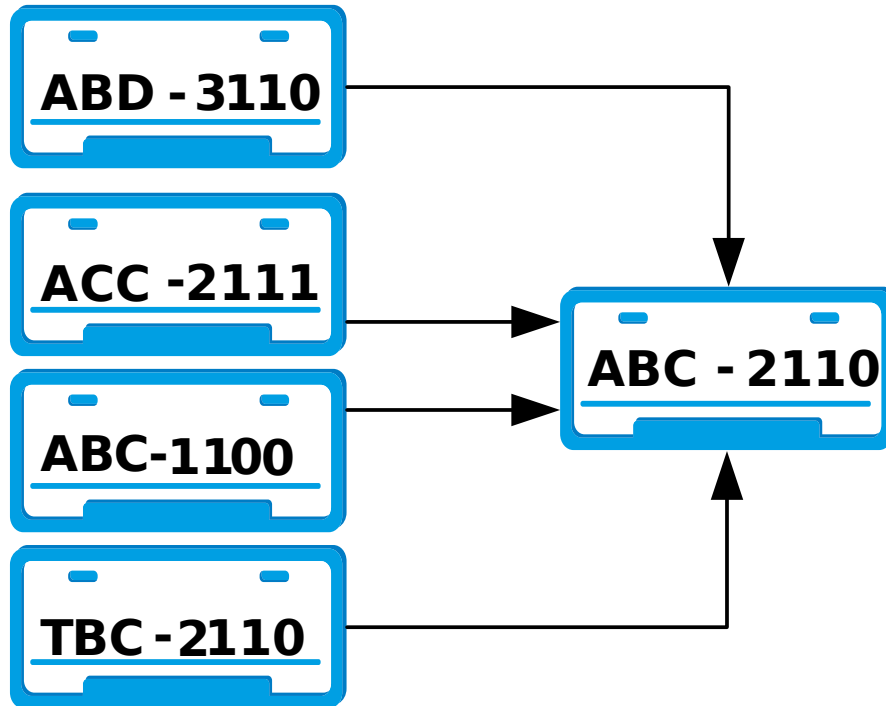


Figure 3.5. The proposed approach combines results from multiple frames to improve the vehicle recognition rate.

have access to a database containing all possible vehicles that could appear on the surveillance video.

After we recognize the vehicle, it is possible that some vehicle might be misrecognized even using the spatio-temporal information. Hence, we propose two improvements in our method that can be applied when we know the possible vehicles that can appear in our videos. The main advantage in these methods is that we make sure that all recognized vehicles are on the database. We propose these techniques based on the fact that 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 Brazil¹. However, the combination of three letter followed by four number would allow more than 175 million license plates to be issued.

3.6.1 Vehicle Appearance Classification

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

¹<http://www.denatran.gov.br/frota2015.htm>



Figure 3.6. Four different vehicle models presenting two very similar frontal appearance. Top: Voyage (left) vs. Gol (right). Bottom: Prisma (left) vs. Onix (right).

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.

The main challenge of this approach is that several vehicles from the same manufacturer have the same frontal (or back) appearance, making the distinction of those vehicles a very complex task, even for humans. For instance, Figure 3.6 shows two examples of different models that have very similar frontal appearance. Therefore, we decided to classify vehicles according to their frontal 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) [Yang et al., 2007]. The only difference between the conventional approaches and the proposed 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. [2003]. It is important to mention that the technique is highly database-dependent. In our case, all license plates have an appearance associated with. Moreover, as will be described in Chapter 5, our dataset was manually divided into 48 distinct appearances (classes).

3.6.2 Tree-Based Query

Once the license plate has been recognized by the temporal redundancy ALPR, we sort the recognized characters by the OCR confidence (given classifier score). In the case of SVM, the confidence is defined by the sample distance to the calculated hyperplane. From the most to the least confident character (therefore, 7 iterations), we filter those candidate license plates that do not have that same character on that particular position. If we find a group having only a single license plate in some iteration, we assume that this is the correct license plate. Otherwise, if we find a group that does not have any license plate at some iteration, we return one level of the filtering and choose the license plate that is the most likely to be the correct one using the OCR confidence.

We implement this technique using a tree. In this case, the root node contains all possible license plates and the number of plates is reduced at every level of the

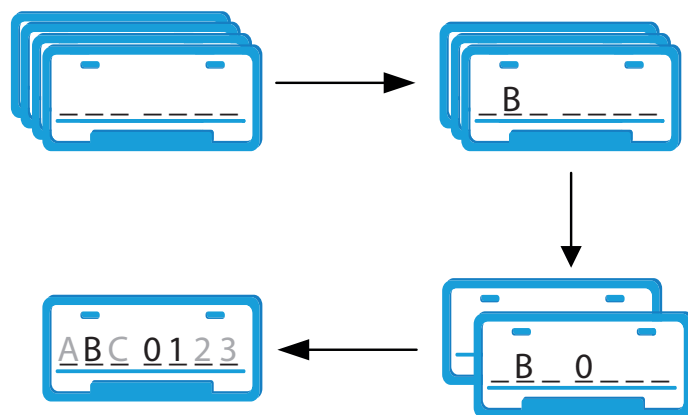


Figure 3.7. Illustration of the tree-based search. Note that the number of candidate license plate is reduced on each iteration.

tree until convergence to a single license plate at a leaf node. The edge connecting two nodes represents the filtering of the license plate from the parent to the child by a specific character. However, it is not feasible to generate the entire tree due to its high branching factor. Instead, we can use the OCR confidence to dynamically build the tree using only the required nodes, ignoring branches with low confidence characters. This process is illustrated in Figure 3.7

Chapter 4

License Plate Character Segmentation Benchmark

This work also proposes a benchmark for the License Plate Character Segmentation (LPCS) problem. This benchmark is composed of a new public dataset (to the best of our knowledge, this is the first dataset focused on the license plate character segmentation task), described in Section 4.1 and a novel evaluation measure to evaluate LPCS approaches, described in Section 4.2.

4.1 Dataset

To be able to evaluate techniques of license plate character segmentation, we compiled a large set of images of on-track vehicles and their license plates into a novel dataset. This dataset, called *SSIG-SegPlate*¹, contains 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 and a hyphen followed by four numbers, resulting in 14,000 characters (alphanumeric symbols) which have been manually annotated with bounding boxes.

The images of the dataset were acquired with a digital camera in Full-HD and are available in the Portable Network Graphics (PNG) format with size of 1920×1080 pixels each. The average size of each file is 4.08 Megabytes (resulting in 8.60 Gigabytes for the entire dataset). In addition, since there are some approaches that track the car to utilize redundant information to improve the recognition results, we decided to make a dataset with multiples frames per car. In this dataset, there are, on average,

¹Available at <http://www.ssig.dcc.ufmg.br>

19.8 images per vehicle (with a standard deviation of 4.14).

The Brazilian license plate has the size of 40×13 cm, resulting in an aspect ratio of 3.08. In the dataset, the license plates have sizes varying from 68×21 to 221×77 pixels. On average, the license plates have the size of 120×42 pixels (aspect ratio of 2.86), which is very close to actual value. In addition, each character of the Brazilian license plate has height of 6.3 cm and the width varying according to the character. In the dataset, the characters in the license plate have their heights varying from 11 to 43 pixels, with an average of 21.19 pixels.

Our dataset is composed of images from multiple vehicle types, among them are passenger vehicles (1762), buses and trucks (118), police cars (14), and service vehicles (106). This variability is important since the license plate color is not the same for all of them, as illustrated in Figure 4.1. For instance, while the license plate for buses and cabs is red with white characters, it is gray with black characters for passenger vehicles. In addition, older cars might have license plates characters from a different text font and some of them might be difficult to read due to the dirt that may be present. Such large variance makes the proposed dataset very challenging and suitable to evaluate LPCS methods on conditions very similar to real environments.



Figure 4.1. Example of different license plate colors in the dataset (the plates were blurred due to privacy constraints).

In addition to the images, we also make available information regarding the position of each plate, its characters and the correct label of the plate characters, allowing a quantitative evaluation of both the plate segmentation and recognition methods. Such information is on a text file with the same name of the image. An example of this file is shown in Figure 4.2. The first row contains the characters of the plate, the second contains the license plate location in the image and from the fourth row on, it lists the location of each symbol in the plate (three letters and four digits). All positions are in

the following format: the x and y coordinates of the upper left corner followed by the height and width size of the bounding box, respectively.

```

text: ABC-0000
position_plate: 1390 782 145 49
position_chars:
  char0: 1399 799 16 25
  char1: 1416 798 17 26
  char2: 1434 799 16 24
  char3: 1459 799 17 24
  char4: 1481 799 7 24
  char5: 1499 799 6 24
  char6: 1511 799 16 24

```

Figure 4.2. Example of the annotation provided with each image on the dataset.

Due to the Brazilian license plate allocation policy, the first letter of the license plate is not uniformly distributed in all locations. Hence, the frequency is unbalanced according to the State where the plate was issued. It means that one letter can appear much more often than others depending on the location. For instance, in Rio de Janeiro there are more license plates beginning with the letters K and L, in Tocantins there are more license plates with M, Rio Grande do Sul has more with I and J, and so on. Our dataset was recorded only on the State of Minas Gerais, therefore some letters appear more than others, as can be seen in Figure 4.3. The letter H appears almost one thousand times, while the letters E and T occur less than one hundred times. Although the distribution is unbalanced, we believe that it does not influence on the segmentation task because the character recognition is not being addressed at this stage of the ALPR process.

We also define a protocol to evaluate segmentation techniques. We split our dataset into three sets: training, testing and validation. Instead of using a regular division of 60% of the dataset to training (model estimation), 20% to validation (parameter optimization) and 20% to testing (reporting final performance), we decided to provide more images for testing, resulting in the following splits: 40% of the dataset to training, 20% to validation and 40% to testing. We keep more images on the testing set because the majority of LPCS approaches do not rely on learning techniques, i.e., do not require model estimation. This way, we are able to evaluate those methods with a large number of test images to make the reported results more statistically significant.

Table 4.1 provides a comparison between well-known datasets of vehicles and our proposed dataset. All of these are publicly available to the research community.

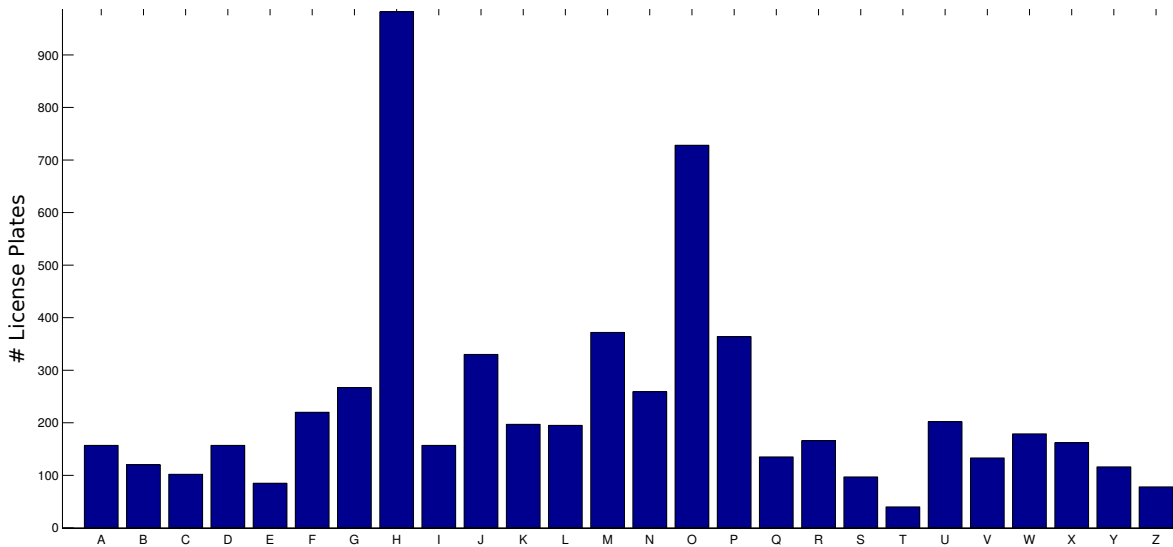


Figure 4.3. Frequency distribution of letters in our dataset.

As it is possible to note, the datasets have multiple purposes and do not provide neither the labels of the vehicles license plates, i.e., their identification, e.g. ABC-1234, nor their character annotation (the bounding boxes of characters composing the license plate), essential to perform fine evaluation of LPCS methods. Some of them do not provide an evaluation protocol, essential to allow a fair comparison among different algorithms. In addition, many provide low resolution images, not suitable to be employed in tasks such as LPCS. The proposed dataset overcomes the majority of these undesired characteristics found in the currently available datasets.

4.2 Jaccard-Centroid Measure

Since there is no measure in the literature specifically 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 by the Jaccard coefficient, a widely employed measure to evaluate how well objects are located in images, define by

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (4.1)$$

where A and B are sets represented by their bounding boxes and $|\cdot|$ stands for the cardinality of a set.

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

Table 4.1. Comparison between the proposed and other datasets available in the literature, regarding different aspects. The proposed dataset is the only one to provide high resolution images with annotation of the individual characters in the license plate, essential to evaluate LPCS approaches. 1. Ferencz & Malik Dataset Ferencz et al. [2004], 2. Caltech Fink [2001], 3. VOC2006 Pascal Everingham et al. [2010], 4. BIT-Vehicle Dataset Dong et al. [2014], 5. UIUC Dataset Agarwal et al. [2004], 6. Krause Cars Dataset Krause et al. [2013], 7. EPFL Car Dataset Ozuysal et al. [2009]

Dataset	# Images	High Resolution	License Plate Labeled	Characters Annotation	Evaluation Protocol	Purpose
1. Ferencz & Malik Dataset	4,000	No	No	No	Yes	Car Detection
2. Caltech	526	No	No	No	No	Object Recognition
3. VOC2006 Pascal	56	No	No	No	Yes	Object Recognition
4. BIT-Vehicle Dataset	900	Yes	No	No	No	Vehicle Type Classification
5. UIUC Dataset	828	No	No	No	Yes	Vehicle Recognition
6. Krause Cars Dataset	16,185	No	No	No	Yes	Vehicle Type Classification
7. EPFL Car Dataset	2,000	Yes	No	No	No	Vehicle Pose Estimation
SSIG SegPlate	2,000	Yes	Yes	Yes	Yes	LPCS

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 [Menotti et al., 2014; Araújo et al., 2013]. 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.

To achieve high character recognition accuracy, the segmentation task must provide characters that are well-segmented easily to be recognized. Menotti et al. [2014] stated that a character is easily recognizable by most supervised learning techniques if the character is centralized on the bounding box. However, the Jaccard coefficient does not consider the alignment of the objects. For instance, Figure 4.4 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.

Note that, in Figure 4.4, we obtain the same Jaccard coefficients, even when the reverse case is considered (outer bounding boxes are the ground truth). Nonetheless, the detection on the left example is expected to be easily recognizable by an OCR since the two bounding boxes are aligned according to their center, i.e., the distance between their centroids is small. Therefore, to capture the alignment precision, it is necessary to combine the Jaccard coefficient and the distance between the centroids of detected and ground-truth bounding boxes, which is precisely the focus of the proposed Jaccard-Centroid coefficient.

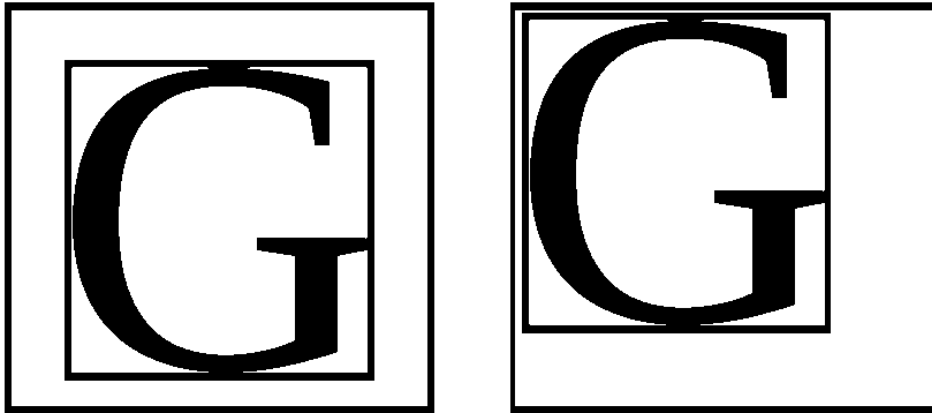


Figure 4.4. Illustration of two segmented bounding boxes. Both have the same Jaccard coefficient but one is not well aligned in the centroid, which might difficult the OCR step in the ALPR.

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

$$JC(A, B) = \frac{J(A, B)}{\max(1, C \times \Delta c(A, B))}, \quad (4.2)$$

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

$$\Delta c(A, B) = \sqrt{(A_x - B_x)^2 + (A_y - B_y)^2}, \quad (4.3)$$

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

The denominator of Equation 4.2 can be considered a penalty term for the Jaccard coefficient. The minimum value is 1 when the misalignment, weighted by the constant C , is less than 1. The best value for constant C was determined experimentally to maximize the recognition rate achieved by the OCR (please, refer to Chapter 5).

The curves in Figure 4.5, obtained by sliding a window diagonally over the ground truth annotation, illustrate the difference between the Jaccard and the Jaccard-Centroid coefficients. While the curve representing the Jaccard measure has a plateau on the top (different alignments lead to the same value), the Jaccard-Centroid measure presents a peak when the centers of the bounding boxes are perfectly aligned, as desirable for performing the OCR.

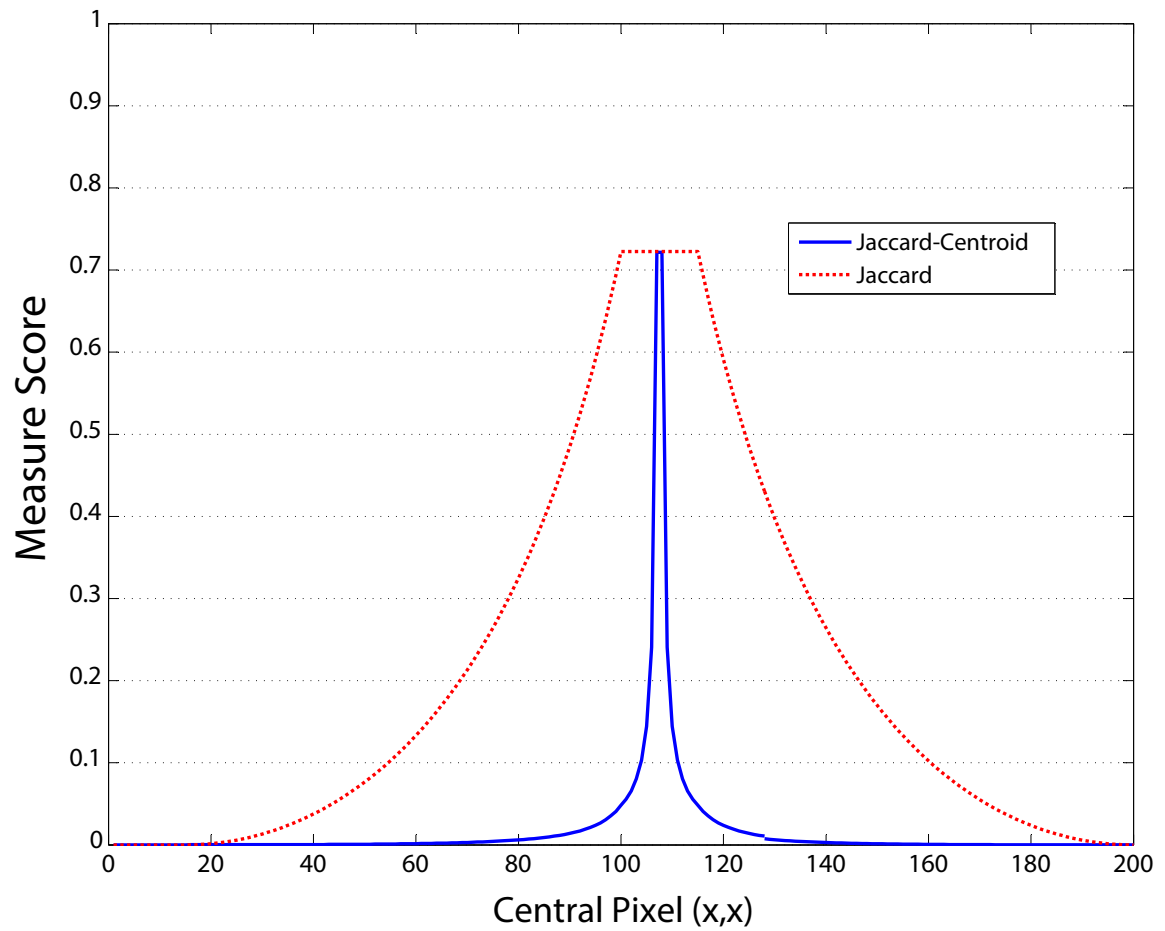


Figure 4.5. The graphic of the Jaccard Coefficient has a plateau when one box is completely inside the other one. However, the Jaccard-Centroid measure, with $C = 2$, does not have this plateau.

Chapter 5

Experimental Evaluation

This chapter presents the results achieved using the techniques described in Chapters 3 and 4. Section 5.1 presents the results of the LPCS benchmark and Section 5.2 presents the experiments performed to evaluate the Automatic License Plate Recognition approach.

5.1 LPCS Benchmarking Results

We perform three main experiments to evaluate our proposed LPCS benchmark: i) individual character segmentation, ii) complete license plate character segmentation to assess whether the baseline approaches are able to segment all characters of a given license plate; and iii) OCR on characters perfectly segmented and on characters segmented using the iterative method proposed in Chapter 3. We performed the experiments in the entire pipeline using the dataset described in Chapter 4.

5.1.1 Parameter Setting

As classifiers for the OCR, we used an One Against-All version of the Oblique Random Forest classifier and the OCR described in Chapter 3. As feature descriptor, we employed the histogram of oriented gradients (HOG) [Dalal and Triggs, 2005] using 9 bins, 4 blocks and 16×16 cell size with 50% of stride (8 pixels).

To determine the best value of the constant C of the Jaccard-Centroid, we executed the OCR considering the 20% best segmented characters on the validation set, varying the value of C . The best achieved value was 3, as illustrated in Figure 5.1. Based on that, all experiments reported on this thesis were performed using $C = 3$.

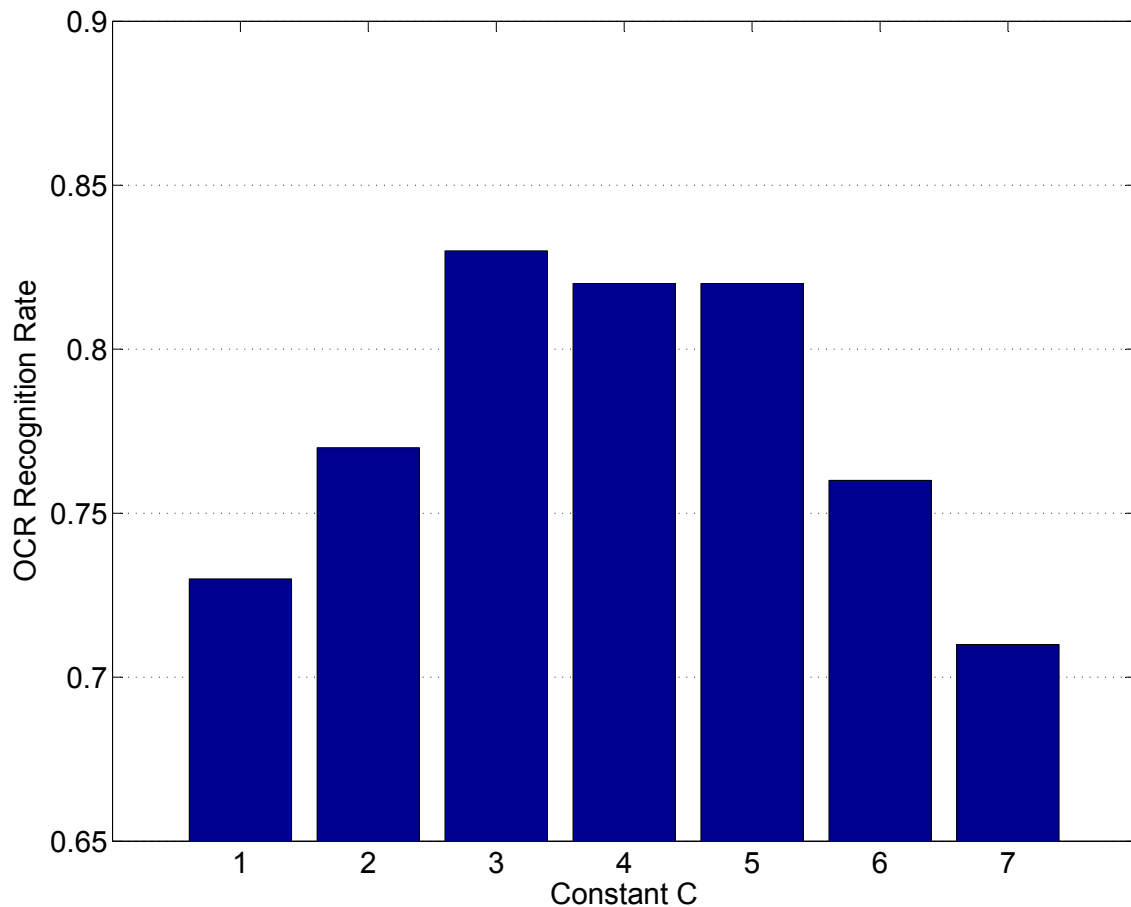


Figure 5.1. OCR recognition rates achieved for the first 20% of characters when we vary the value of the constant C in Equation 4.2.

5.1.2 Baselines

This section describes the three LPCS techniques used as baselines. We considered two methods available in the literature. The first aims at improving the quality of degraded images of words [Nomura et al., 2009] (Section 5.1.2.1) and the second performs segmentation by counting the number of black pixels in a binarized license plate based on Connected Component Labeling [Shapiro and Gluhchev, 2004] (Section 5.1.2.2). 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 (Section 5.1.2.3).

5.1.2.1 Pixel Counting Approach

This approach takes into account a very specific preprocessing method called Shadow Location and Lightening (SL*L) to improve the quality of degraded images containing text. This method consists of a sequence of mathematical morphological operations



Figure 5.2. SL*L preprocessing. At the top, there is an example of an image binarized without the SL*L preprocessing and at the bottom, there is an image binarized using the SL*L processing method.

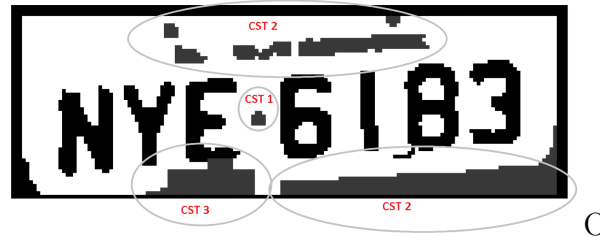


Figure 5.3. The three shadow types (CST1, CST2, and CST3 - in gray) reduced by employing the approach.

applied to locate the shadowed regions on the image and lightening them to remove the noise before applying the final thresholding process by the Otsu method [Otsu, 2007]. Figure 5.2 shows the difference when the SL*L preprocessing approach is employed.

The approach begins with a binarization of the image and the application of a thickening operation. Then, it locates regions with three types of shadows to reduce their effect. These types are named Critical Shadow Type 1 (CST1), 2 (CST2), and 3 (CST3). The CST1 is the shadow that can occur between the characters, CST2 is the shadow that does not occur between two characters and does not touch them, and the CST3 is the shadow that does not occur between two characters but touches one. Figure 5.3 illustrates these three types on a license plate converted to grayscale.

The CSTs are detected using the pruning algorithm based on

$$S \ominus B = \left(\bigcup_i^n (S * B^i) \oplus H \right) \cap X, \quad (5.1)$$

where X is the binary image, S is the skeleton of X , B and H are structuring elements

and n is the number of iterations, the operation $*$ denotes hit-or-miss transform and \oplus is the dilatation operation [Serra, 1986]. After applying the pruning process, the image presents an enclosing boundary surrounding the shadowed regions, highlighting these regions so that a noiseless image is obtained. Finally, to perform the segmentation, the approach thresholds the image using a global thresholding technique and counts the white pixels on both directions to locate the segmentation points. For further information, please refer to the original paper where the method was proposed [Nomura et al., 2009].

5.1.2.2 Connected Component Labeling Approach

The approach proposed by Shapiro and Gluhchev [2004] is straightforward. An adaptive thresholding is performed in the image, followed by connected components labeling and then a greedy selection process is performed to choose the best characters based on their size, similar to the step performed in the segmentation technique proposed in this thesis (Section 3.3).

Each connect component is analyzed based on its height with respect to the license plate. This is done analyzing the size of its connected component. Based on the real proportion of the height of a character regarding the height of a Brazilian license plate, which is 45%, we use the proportion height range of [40%, 50%] to accept a connected component as a character. Since the license plate actual proportion is 45%, we decided to use 40% and 50% as its minimum and maximum heights, respectively.

5.1.2.3 Prior Knowledge-Based Approach

For this simple approach, we consider a technique to segment the seven characters of the license plate using only the information regarding the real shape of the Brazilian License Plate and its characters. The Brazilian license plate has seven characters and an hyphen between the third and the forth - to separate the letters from the digits. Therefore, we consider this hyphen as one character and divide the license plate equally into eight horizontal regions. In addition, we eliminate 15% from the top of the license plate and 5% from the bottom to crop only the portion containing the characters.

5.1.3 Individual Character Segmentation Evaluation

Table 5.1 shows the average values achieved by the three baselines of segmentation and the proposed iterative approach described in Chapter 3 on the testing set of the proposed dataset. According to these results, the baseline approaches do not present

promising results neither the proposed, emphasizing the need to the development of new approaches to perform LPCS accurately.

Table 5.1. Measure results of segmentation: values achieved for the three baselines and our proposed approach using three measures.

Approach	Jaccard	Δc	JC
Pixel Counting [Nomura et al., 2009]	0.601	2.052	0.316
Conn. Component [Shapiro and Gluhchev, 2004]	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 presents the higher average degree of misalignment, represented by the Δc (expected due to its simplicity). As a consequence, this segmentation approach is penalized by the proposed Jaccard-Centroid measure (its value is 0.352 lower than the value computed using the Jaccard coefficient). 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 labeling and the pixel counting approaches achieved smaller Δc value, causing minor penalization to the Jaccard-Centroid coefficient. The SL*L using Pixel Counting was capable to achieve an average scores near 0.60 by Jaccard measure and near 0.30 by Jaccard-Centroid coefficient, which is the best result of the three proposed baseline.

Our proposed approach achieved the best results. It achieved a higher value in Jaccard and it is not much penalized by the Δc value, which corresponds to low misalignment error. These results support the hypothesis that our method, despite being straightforward, is the best approach to perform LPCS efficiently.

We also analyzed the number of characters that were satisfactory segmented as a function of the Jaccard-Centroid coefficient. Figure 5.4 shows curves of the effectiveness (the correctly segmented characters) of our proposed method and each evaluated baseline approach as a function of the threshold on the Jaccard-Centroid measure. That is, for a given threshold value (ranging from 0.05 to 1), we compute the percentage of characters that have obtained a Jaccard-Centroid measure equal or higher than the threshold.

For the analysis, we consider 0.4 as a Jaccard-Centroid threshold to obtain a satisfactory segmentation (this value was determined empirically), otherwise, the character might not be well-centered and the OCR will not properly work. According to the results shown in Figure 5.4, none of the baselines approaches is accurate enough to be employed in a feasible ALPR system. The approach using Pixel Counting was capable

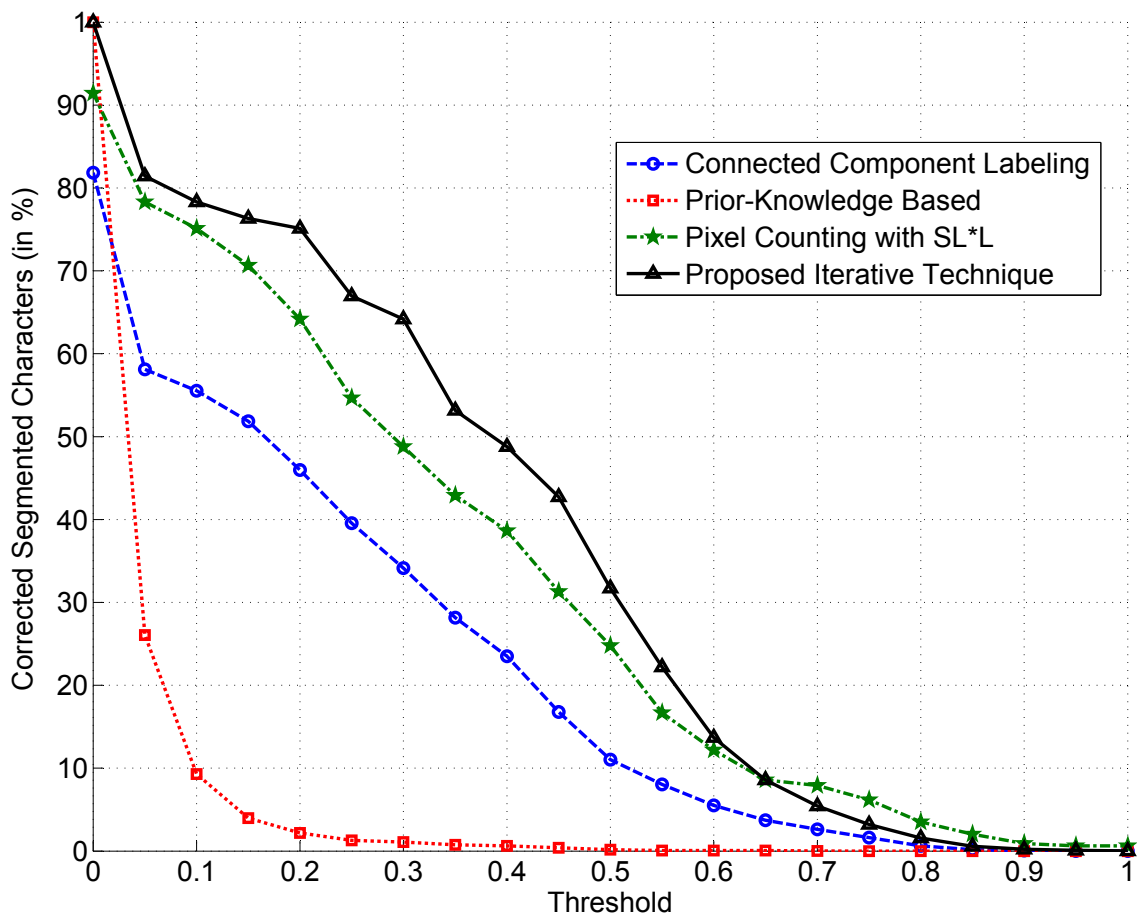


Figure 5.4. Percentage of individual characters correctly segmented as a function of the Jaccard-Centroid coefficient.

of segmenting satisfactorily around 40% of the characters, while the technique using connected component labeling and the approach using Prior-Knowledge were able to segment only 25% and 3%, respectively. Our approach was capable to segment around 50% of all license plate characters, achieving the best results of all evaluated methods. Nonetheless, our full ALPR method is feasible for a real system since we diminish this problems with the temporal redundancy.

5.1.4 Full License Plate Segmentation Evaluation

In this section, we evaluate the segmentation of the entire license plate to analyze its relation with the Jaccard-Centroid coefficient. We used the values of our measure applied to the seven characters of one license plate to determine whether the license plate segmentation would be plausible for recognition or not. This is an important evaluation since all characters must be found in the plate and each of them must be well

located/segmented so that the plate can be properly recognized by OCR techniques.

To perform the evaluation, we analyze the Jaccard-Centroid coefficient values by varying a threshold. That is, if the character of the license plate with the lowest Jaccard-Centroid coefficient is higher or equal than the current threshold, the license plate is considered as correctly segmented. Figure 5.5 shows how the correctly segmented license plates percentage varies as a function of the Jaccard-centroid coefficient for each baseline approach. According to the curves, it is possible to see that none of the approaches present high accuracies for higher threshold values.

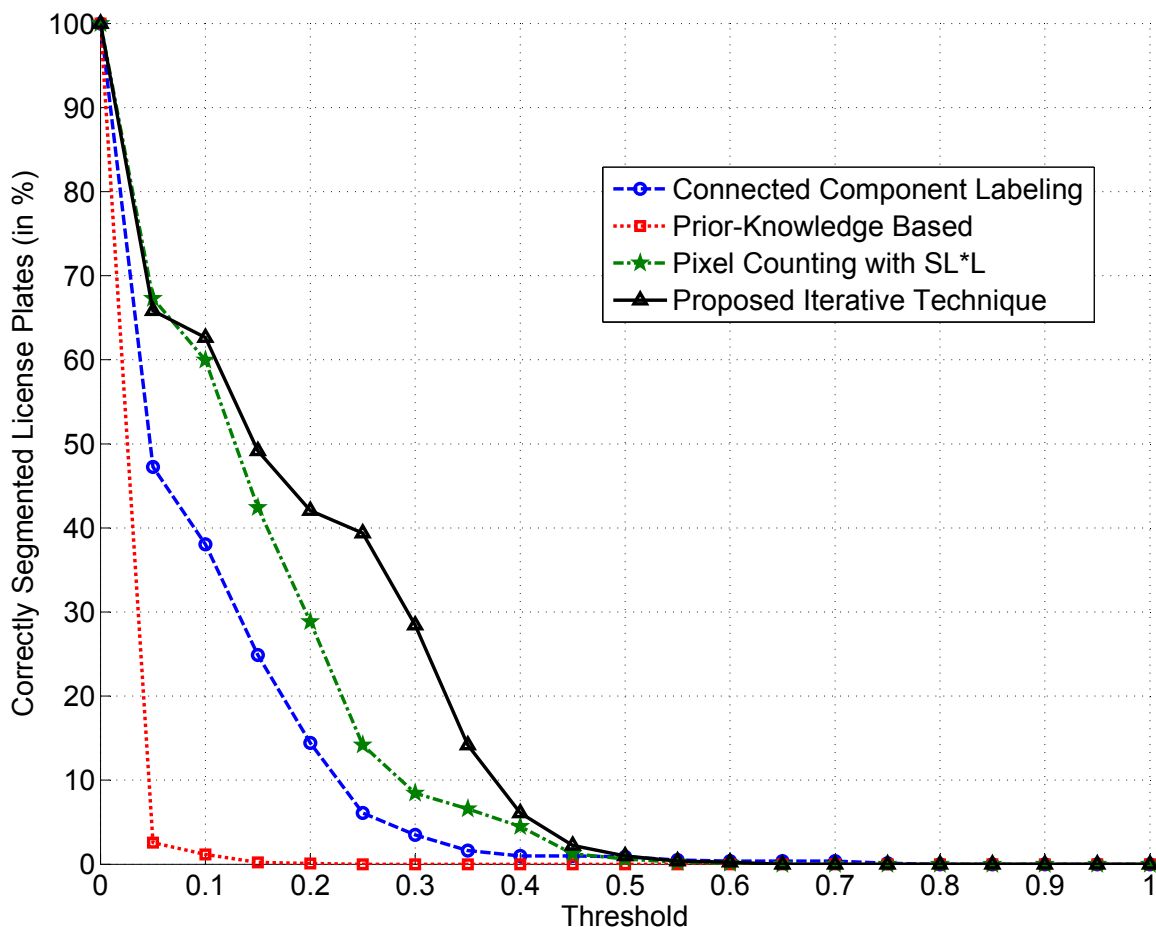


Figure 5.5. Percentage of correctly segmented license plates (all seven characters were segmented in the plate) as a function of the Jaccard-Centroid coefficient.

Considering the same threshold used in the previous section (0.4), the prior knowledge based approach is not able to segment any license plate, confirming that this method is too simple to perform LPCS. Besides, the connected component labeling approach achieved a segmentation rate close to 2% on the mentioned threshold which also entails a very low performance. The approach using SL*L and pixel counting achieved

Table 5.2. Recognition rates of the OCR using both segmentation approaches (manual and automatic) and two classifiers (Radial SVM and oRF).

Approach	Manual Segmentation		Segmentation by Pixel Counting	
	Letters	Numbers	Letters	Numbers
Radial SVM	0.919	0.962	0.275	0.552
oRF	0.947	0.969	0.586	0.782

6% segmentation rate, which would not be considered satisfactory for any real application. Our approach was capable to achieve 8% of segmentation rate which, despite being the best result of all evaluated techniques, it is also not satisfactory enough to be employed in real scenarios. Such results reinforce the fact that our dataset is challenging and suitable to evaluate the robustness of LPCS techniques.

According to the results shown in Figures 5.4 and 5.5, it is possible to see that, even though the approach is capable of segmenting 50% of the characters, it is not able to segment all characters in more than 8% of license plates. This fact shows that in almost all license plates, there is at least one character that is not well segmented by any approach, which is critical for the license plate recognition once the OCR requires an acceptable segmentation for correctly recognize all characters - if a single character was not well-segmented, the identification of the license plate is compromised. Nonetheless, later in this chapter, we show that the use of multiple frames of the same vehicle can help to suppress the error of the segmentation phase, leading to an acceptable final recognition rate using our proposed segmentation approach.

5.1.5 Optical Character Recognition Evaluation

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 that, we performed experiments to evaluate the final accuracy of the OCR when applied to license plate characters segmented with and without a precise segmentation.

Table 5.2 shows the accuracy of the two previously mentioned learning based OCR systems obtained when applied to: (i) manually segmented characters and (ii) automatic segmented characters. According to the results, there is a large influence of the character segmentation on the final results of the OCR. For instance, the OCR recognition rate can decrease on 0.644 and 0.410 points in the worst-case, for letters and numbers respectively, justifying therefore the need to have a precise segmentation system in the ALPR pipeline.

Given a rank of the characters that are best segmented according to Jaccard

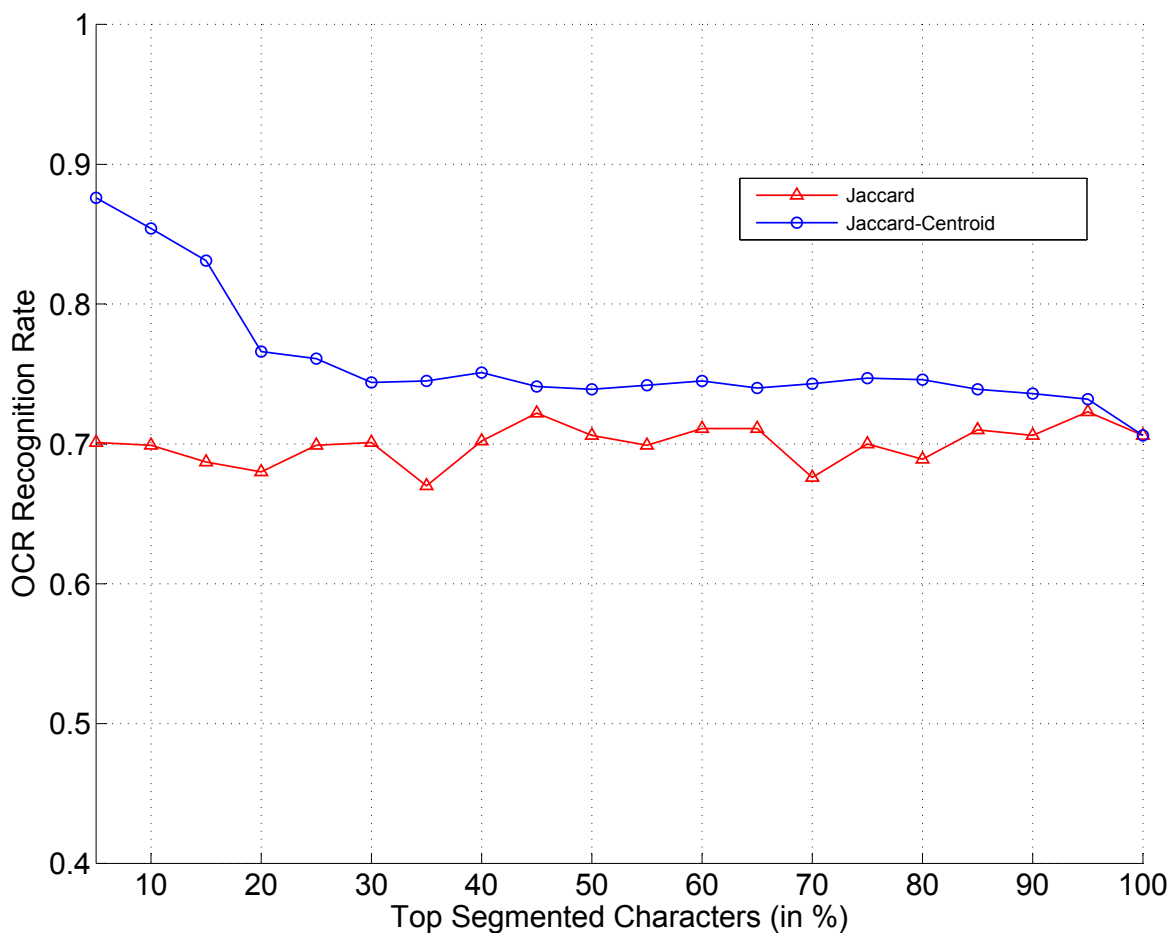


Figure 5.6. Recognition rate of OCR as a function of percentage of the top segmented characters considering Jaccard and Jaccard-Centroid coefficients.

and Jaccard-Centroid coefficients, Figure 5.6 shows the recognition rates of an OCR system when applied to a percentage of the top segmented characters of these ranks. All characters were segmented using the best proposed baseline (pixel counting). 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 percentage 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.

5.2 Proposed ALPR Results

In this section, we 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. Although they perform the recognition using only one frame, all baselines evaluate all frames to determine the best one. Therefore, one can note that is not unfair to compare them with our approach. Furthermore, this section presents the results achieved when we employ the post-processing techniques to perform vehicle appearance classification and the tree-search.

5.2.1 Datasets

We collected three sets of data to validate the proposed approaches. Each one has a different purpose. 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 5,200 frames, with size of 1920×1080 pixels, extracted from surveillance videos with 300 on-road moving vehicles (17.33 frames per vehicle on average) recorded in Brazil. The vehicles license plates have size of 120×42 pixels and aspect ratio of 2.86 on average. The third set, used for vehicle classification by appearance, contains 1,000 samples divided in 48 classes corresponding to an average of 20.83 vehicles per class. The latter set was acquired using a crawler in Google Images. Even though we could have used the dataset proposed in Yang et al. [2015], we chose to collect our own samples due to the fact that all Brazilian vehicles used in our experiments must present a corresponding appearance class within our dataset, which is not available in their data. Some examples of this dataset are shown in Figure 5.7. Finally, we generate 80 million random valid license plates to test our tree-based approach on a real scenario.

Although we developed our method using images of vehicles with Brazilian license plate models, we can also use the proposed approach on different models. For this purpose, we only have to train the license plate detector with examples of the new model and adapt the character segmentation technique to work properly with the model concerned. Furthermore, in the case of license plate being located in the vehicle rears, we can also train the vehicle detection to recognize the back of the vehicle using new appropriate examples.



Figure 5.7. Examples of images from the dataset used for vehicle classification.

5.2.2 Experimental Setup

All parameters of the ALPR pipeline were experimentally calibrated. To vehicles and license plates detections, we utilized 20 trees in Oblique Random Forest with 8 factors for each PLS model. Each one of the 650 training vehicle images used as examples have size of 128×96 pixels. In the vehicle detection, we utilized blocks of size 16×16 pixels with 4 cells and 8 bins. In the case of the license plate, each sample has the size of 88×32 pixels and the HOG parameters are the same used in the vehicle detection. We search for vehicles and license plates sliding a window over the image. We used 6 scales and 15% slide window dimension stride on both directions to detect vehicles and 12 scales combined with 8% stride to detect the license plates.

For the OCR, the SVMs were trained with a radial-basis kernel (RBF) and the parameters were calibrated empirically. We perform a grid search evaluating one parameter at a time. Since the parameter search space is not convex, we can not guarantee that the values are the best possible values to be used. In the case of letters, the gamma parameter was 10^{-5} and the parameter C was 10^{-2} . In case of digits, the parameter gamma was 2.45^{-3} and C was 10^{-2} as well. We used a dataset with 10,000 license plate number and 7,800 license plate letter to train our models gathered from the LPCS dataset described in Section 5.1. Each character has the size of 32×32 pixels and the features was extracted using HOG with block of size 16×16 pixels with 4 cells each one and 8 bins.

5.2.3 ALPR Subtasks Results

We evaluated the performance of the vehicle and license plate detection as well as the accuracy of the OCR system in the collected vehicles dataset. On one hand, the vehicle

and license plate detection were evaluated by the number of corrected detections using 0.4 as jaccard threshold and 5-fold cross-validation. This value was shown to be a good threshold to determine whether a license plate was correctly detected by Prates et al. [2014b]. We decided to utilize the same for vehicle detection. On the other hand, we measure the OCR accuracies using the ratio of correctly recognized characters. Table 5.3 shows the results.

Task	Accuracy
Vehicle Detection	95.4%
License Plate Detection	88.3%
Letters OCR	94.7%
Numbers OCR	96.9%

Table 5.3. Accuracy of the subtasks utilized on the proposed ALPR pipeline.

Although these subtasks can achieve high accuracy values when evaluated individually, the final vehicle recognition rate is influenced when we perform all of them in sequence. This happens due to a cascading noise effect between all subtasks. Because of that, some ALPR approaches achieve very low recognition rates when applied to real scenarios as one can see in Table 5.4.

5.2.4 Temporal Redundancy Aggregation

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. [2005]. 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 5.4, 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 com-

Table 5.4. Recognition rates (per plate) 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. [Bremananth et al., 2005] (without redundancy)	78.3%	66.3%
redundancy with OCR average	93.6%	77.9%
redundancy with majority voting	94.6%	81.8%

binning the results of multiple vehicle detections can provide better recognition rates than using just a single frame.

5.2.5 Post-Processing Approaches

Once the best results were achieved using majority voting, we utilize the results of this approach as input to both post-processing techniques.

The first post-processing technique consists in: (i) performing a classification on the vehicle frontal appearance; (ii) consulting the database and retrieving all possible license plates for the predicted class; and (iii) choosing the most similar one to the recognized plate. The second post-processing approach is based on: (i) sorting the characters by the highest OCR score; (ii) perform queries on a dataset by the candidate license plates containing the same characters on the same positions; and (iii) select the best candidate filtered in the queries.

5.2.5.1 Vehicle Appearance Classification

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 $C = 0.5$.

Figure 5.8 illustrates the achieved results of the proposed classifier model for different number of codewords per class. It is possible to see that there are no improvements on the classification when we use more than 900 visual words per class (final dimensionality of 43,200). In the best case, the model was capable of predicting correctly around 48% of the test vehicle images in rank-1. Nonetheless, the model returned the correct class in 80% and 91% of the cases using ranking 10 and 20, respectively. Therefore, the use of ranks higher than 1 can reduce the search space significantly without degrading much the recognition rate. The model was capable to recognize all license plates only using the first 35 classes in the best case.

We performed an experiment varying the rank of classes used to predict the license plate, using 900 codewords per class. According to the results shown in Figure 5.9,

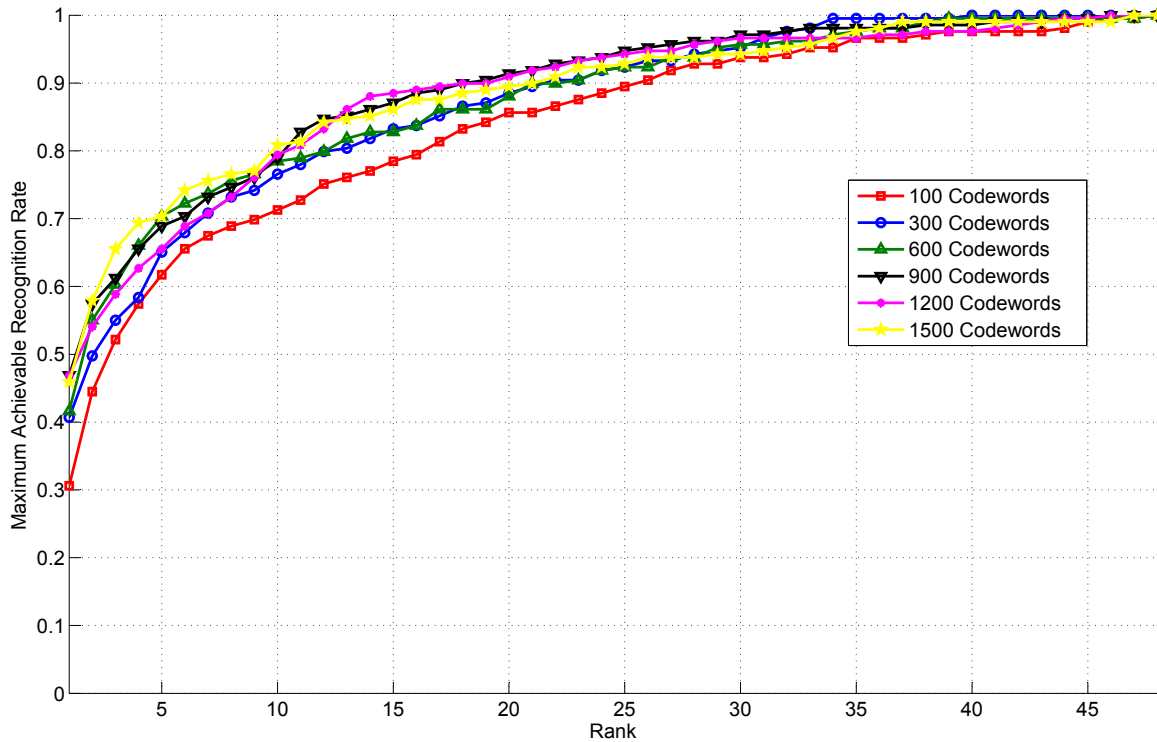


Figure 5.8. Recognition rates as a function of the top rank positions.

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 fourth row of Table 5.4 (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 in Figure 5.9, that the use of more than 10 top predicted classes does not bring significant improvements to the classification. We believe that this happen because the use of many classes provides more candidate license plates. Hence, it does not filter similar license plate belonging to complete different vehicles satisfactorily.

5.2.5.2 Tree-Based Query

To execute the experiment using the tree-based approach, we generate a database containing 80 million random license plates to simulate a real vehicles scenario. The approach was capable of improving the results obtained using only the temporal redundancy information in 4.8 percentage points, leading to a recognition rate of 86.2%. This demonstrates that, once we have access to the database of all registered vehicles (i.e. the Department of Motor Vehicles database), we can correct erroneous recognitions,

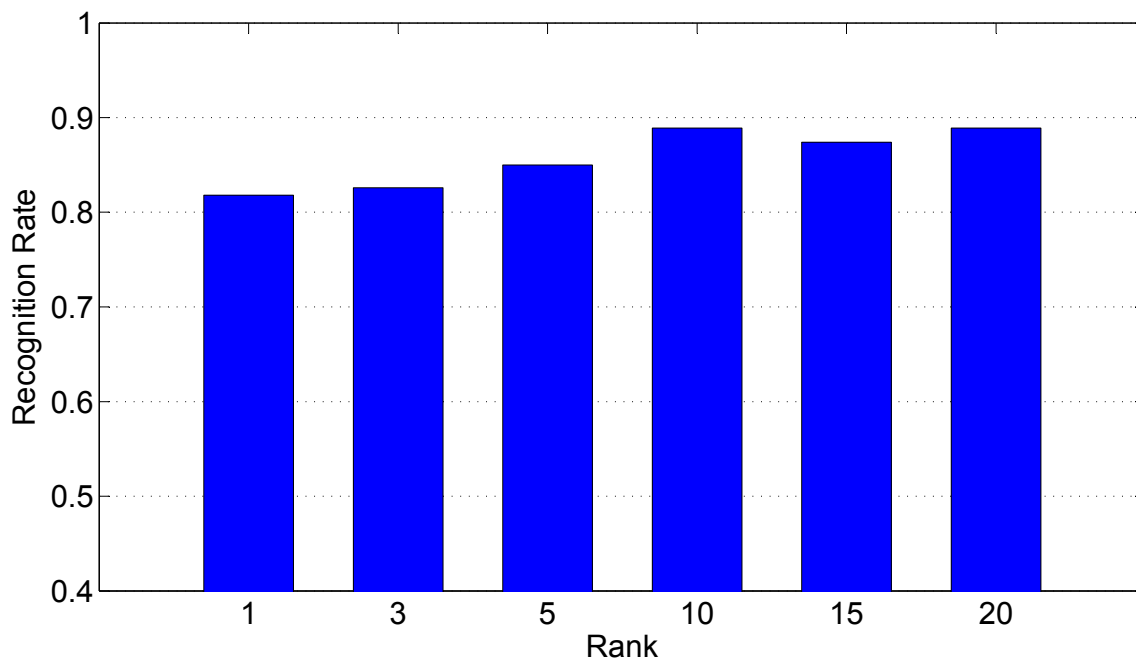


Figure 5.9. Percentage of license plates correctly recognized as a function of the amount of license plates evaluated according to rank.

even when this database is very large.

5.2.5.3 Combined Results

In this last experiment, we combine the three approaches proposed in this thesis. First, we recognized the vehicle combining multiple frames employing the temporal redundancy approach. Then, we performed the vehicle appearance classification to filter those candidate license plates to be used in the next step. We executed the tree-based query approach in the set of license plates filtered by the VAC model. The combined approach achieved a recognition rate of 89.6%, an increase of 7.8 p.p. (9.35%) compared to the results obtained considering only temporal redundancy using the proposed dataset of 300 vehicles. Finally, in addition to the approaches results summarized in Table 5.5, if we assume a normal distribution regarding the quantity of wrong characters on the misrecognized license plates, we have 94.3% of character recognition accuracy, which is highly coherent with the reported accuracy of our SVM-based Optical Character Recognition technique.

Table 5.5. Final Recognition rates (per plate) achieved by the proposed approaches when we apply the post-processing techniques.

Approach	Recognition Rate
Bremanath et al.	53.5%
redundancy with majority voting	81.8%
redundancy + vehicle appearance (VAC) query	88.9%
redundancy + tree-based search	86.2%
redundancy + VAC + tree-based	89.6%

5.2.6 Discussion

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 Bremananth et al. [2005], 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 5.4) 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%.

Both post-processing approaches were able to improve the results of the temporal redundancy approach by querying a dataset of all possible license plates. It is important to point out that the vehicle appearance classification is computationally expensive due to the high dimensionality of the feature vector. Therefore, it should be used in systems with high computational processing power, otherwise, it may compromise the ALPR system, once such system should be able to run in real-time. Furthermore, when we combined both approaches, we observed a gain of 7.8 p.p. compared to the proposed temporal redundancy approach, which is a significant improvement and justifies the combined use of both post-processing approaches.

Chapter 6

Conclusions and Future Works

In this work, we proposed a new approach to perform ALPR exploring temporal redundancy information from detected vehicles. We also proposed two post-processing techniques to improve the final recognition accuracy of the ALPR pipeline by querying a license plate database. The former approach classifies the vehicle according to its appearance and verifies whether the recognized plate corresponds to a valid license plate of a vehicle with the correspondent frontal appearance. The latter performs a tree-based search on the database to verify whether the recognized license plate is valid or not. Both approaches can be used by an agent/system that has access to the enrolled vehicles (their license plates) in the scenario, e.g., the Department of Motor Vehicle of a country/state.

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

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 scores 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. Furthermore, if we consider 0.4 as a satisfactory Jaccard-Centroid threshold to determine whether the characters in the plate were correctly segmented (from our experience with OCR, near or perfect

recognition accuracy can be achieved only when Jaccard-Centroid is equal or greater than 0.4), none of the approaches was capable of segmenting all characters in more than 10% of the license plates.

As future directions, we plan to employ a Vehicle *Model* Classification trained with more classes and a larger dataset to make the filtering process more effective. Also, we intend to collect more images to create an extension of the benchmark dataset SSIG-SegPlate1 with more than 10,000 images of on-road vehicles and at least 1,000 samples of each character to perform an extensive analysis of OCR techniques in the ALPR context.

Bibliography

- Agarwal, S., Awan, A., and Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, 26(11):1475–1490.
- Antonacopoulos, A., Bridson, D., Papadopoulos, C., and Pletschacher, S. (2009). A realistic dataset for performance evaluation of document layout analysis. In *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conf. on*, pages 296–300. IEEE.
- Araújo, L., Pio, S., and Menotti, D. (2013). Segmenting and recognizing license plate characters. In *Workshop of Undergraduate Work - Conference on Graphics, Patterns and Images*, pages 251–270.
- Bay, H., Tuytelaars, T., and Van Gool, L. (2006). Surf: Speeded up robust features. In *European conference on computer vision*, pages 404–417. Springer.
- Betke, M., Haritaoglu, E., and Davis, L. (2000). Real-time multiple vehicle detection and tracking from a moving vehicle. *Machine Vision and Applications*, 12(2):69–83.
- Bremananth, R., Chitra, A., Seetharaman, V., and Nathan, V. S. L. (2005). A robust video based license plate recognition system. In *Intelligent Sensing and Information Processing. International Conference on*, pages 175–180. IEEE.
- Caner, H., Gecim, S., and Alkar, A. (2008). Efficient embedded neural-network-based license plate recognition system. *Vehicular Technology, IEEE Transactions on*, 57(5):2675–2683.
- Chen, Y.-L., Wu, B.-F., Huang, H.-Y., and Fan, C.-J. (2011). A real-time vision system for nighttime vehicle detection and traffic surveillance. *Industrial Electronics, IEEE Transactions on*, 58:2030–2044.

- Chuang, C.-H., Tsai, L.-W., Deng, M.-S., Hsieh, J.-W., and Fan, K.-C. (2014). Vehicle licence plate recognition using super-resolution technique. In *International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 411–416.
- Dalal, N. and Triggs, B. (2005). Histograms of Oriented Gradients for human detection. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 886–893.
- Dong, Z., Pei, M., He, Y., Liu, T., Dong, Y., and Jia, Y. (2014). Vehicle type classification using unsupervised convolutional neural network. In *International Conference on Pattern Recognition (ICPR)*, pages 172–177. IEEE.
- Donoser, M., Arth, C., and Bischof, H. (2007). Detecting, tracking and recognizing license plates. In *Asian Conference on Computer Vision*, pages 447–456. Springer.
- Du, S., Ibrahim, M., Shehata, M., and Badawy, W. (2013). Automatic license plate recognition (ALPR): A state-of-the-art review. *Circuits and Systems for Video Technology, IEEE Transactions on*, 23(2):311–325.
- Everingham, M., Van Gool, L., Williams, C. K., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338.
- Fan, Z., Zhao, Y., Burry, A. M., and Kozitsky, V. (2012). License plate character segmentation using likelihood maximization. US Patent App. 13/464,357, Google Patents.
- Ferencz, A. D., Learned-Miller, E. G., and Malik, J. (2004). Learning hyper-features for visual identification. In *Advances in Neural Information Processing Systems*, pages 425–432.
- Fink, M. (2001). Caltech object category datasets. <http://www.vision.caltech.edu/archive.html>. Accessed on: 2015-03-25.
- Franc, V. and Hlaváč, V. (2005). License plate character segmentation using hidden markov chains. In *Joint Pattern Recognition Symposium*, pages 385–392. Springer.
- Goncalves, G. R., da Silva, S. P. G., Menotti, D., and Schwartz, W. R. (2016a). Benchmark for license plate character segmentation. *Journal of Electronic Imaging*, 25(5):1–11.
- Goncalves, G. R., Menotti, D., and Schwartz, W. R. (2016b). License Plate Recognition based on Temporal Redundancy. In *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–5.

- Gonzalez, R. C. (2009). *Digital image processing (3rd Edition)*. Pearson Education India.
- Goodfellow, I. J. and Bulatov, Y. and Ibarz, J. and Arnoud, S. and Shet, V. (2014). Multi-digit number recognition from street view imagery using deep convolutional neural networks. In *International Conference on Learning Representation*. arXiv preprint arXiv:1312.6082.
- Guo, J.-M. and Liu, Y.-F. (2008). License plate localization and character segmentation with feedback self-learning and hybrid binarization techniques. *Vehicular Technology, IEEE Transactions on*, 57(3):1417–1424.
- Guyon, I., Schomaker, L., Plamondon, R., Liberman, M., and Janet, S. (1994). Unipen project of on-line data exchange and recognizer benchmarks. In *Pattern Recognition, Computer Vision & Image Processing., IAPR International. Conf. on*, volume 2, pages 29–33.
- Hsieh, J.-W., Chen, L.-C., Chen, S.-Y., Chen, D.-Y., Alghyaline, S., and Chiang, H.-F. (2015). Vehicle color classification under different lighting conditions through color correction. *Sensors Journal, IEEE*, 15(2):971–983.
- Jazayeri, A., Cai, H., Zheng, J., and Tuceryan, M. (2011). Vehicle detection and tracking in car video based on motion model. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):583–595.
- Jordao, A. and Schwartz, W. R. (2016). Oblique random-forest based on partial least squares applied to pedestrian detection. In *International Conference on Image Processing (ICIP)*, pages 375–382. IEEE.
- Kembhavi, A., Harwood, D., and Davis, L. (2011). Vehicle detection using partial least squares. *TPAMI*, pages 1250–1265.
- Kocer, E. and Cevik, K. (2011). Artificial neural networks based vehicle license plate recognition. *Procedia Computer Science*, 3:1033–1037.
- Krause, J., Deng, J., Stark, M., and Li, F.-F. (2013). Collecting a large-scale dataset of fine-grained cars. at <https://www.d2.mpi-inf.mpg.de/sites/default/files/fgvc13.pdf> (Accessed in 06/07/2016).
- Matas, J. and Zimmermann, K. (2005). Unconstrained licence plate and text localization and recognition. In *International Conference on Intelligent Transportation Systems (ITSC)*, pages 225–230. IEEE.

- Menotti, D., Chiachia, G., Falcão, A., and Oliveira-Neto, V. (2014). Vehicle license plate recognition with random convolutional networks. In *Conference on Graphics, Patterns and Images*, pages 298–303.
- Nagare, A. P. (2011). License plate character recognition system using neural network. *International Journal of Computer Applications*, 25(10):36–39.
- Nazare, Antonio C., J., Ferreira, R., and Schwartz, W. R. (2014). Scalable feature extraction for visual surveillance. In *Iberoamerican Congress on Pattern Recognition (CIARP)*, volume 8827 of *Lecture Notes in Computer Science*, pages 375–382. Springer International Publishing.
- Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., and Ng, A. Y. (2011). Reading digits in natural images with unsupervised feature learning. In *Workshop on Deep Learning and Unsupervised Feature Learning - International Conference on Neural Information Processing Systems*, pages 1–9.
- Nomura, S., Yamanaka, K., Shiose, T., Kawakami, H., and Katai, O. (2009). Morphological preprocessing method to thresholding degraded word images. *Pattern Recognition Letters*, 30(8):729–744.
- Oliveira-Neto, V., Cámara-Chávez, G., and Menotti, D. (2012). Towards license plate recognition: Comparing moving objects segmentation approaches. In *International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICIP)*, pages 447–453.
- Otsu, N. (2007). A threshold selection method from gray-level histograms. *Transactions on Systems, Man, and Cybernetics*, 9:62–66.
- Ozbay, S. and Ercelebi, E. (2005). Automatic vehicle identification by plate recognition. *WASET*, pages 222–225.
- Ozuysal, M., Lepetit, V., and Fua, P. (2009). Pose estimation for category specific multiview object localization. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 778–785. IEEE.
- Prates, R., Cámara-Chávez, G., Schwartz, W., and Menotti, D. (2014a). An adaptive vehicle license plate detection at higher matching degree. In *Iberoamerican Congress on Pattern Recognition (CIARP)*, pages 454–461. Springer.

- Prates, R., Cámara-Chávez, G., Schwartz, W., and Menotti, D. (2014b). Brazilian license plate detection using histogram of oriented gradients and sliding windows. *arXiv preprint arXiv:1401.1990*.
- Qadri, M. T. and Asif, M. (2009). Automatic number plate recognition system for vehicle identification using optical character recognition. In *International Conference on Education Technology and Computer*, pages 335–338. IEEE.
- Rao, Y. (2015). Automatic vehicle recognition in multiple cameras for video surveillance. *The Visual Computer*, 31(3):271–280.
- Sermanet, P., Chintala, S., and LeCun, Y. (2012). Convolutional neural networks applied to house numbers digit classification. In *International Conference on Pattern Recognition (ICPR)*, pages 3288–3291.
- Serra, J. (1986). Introduction to mathematical morphology. *Computer vision, graphics, and image processing*, 35(3):283–305.
- Shantaiya, S., Verma, K., and Mehta, K. (2015). Multiple object tracking using kalman filter and optical flow. *EJAET*, pages 34–39.
- Shapiro, V. and Gluhchev, G. (2004). Multinational license plate recognition system: Segmentation and classification. In *International Conference on Pattern Recognition*, volume 4, pages 352–355. IEEE.
- Shih, H.-C. and Wang, H.-Y. (2015). Vehicle identification using distance-based appearance model. In *AVSS*, pages 1–4. IEEE.
- Shuang-Tong, T. and Wen-ju, L. (2005). Number and letter character recognition of vehicle license plate based on edge hausdorff distance. In *International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT)*, pages 850–852.
- Sirikuntamat, N., Satoh, S., and Chalidabhongse, T. (2015). Vehicle tracking in low hue contrast based on camshift and background subtraction. In *JCSSE*, pages 58–62. IEEE.
- Sivaraman, S. and Trivedi, M. (2014). Active learning for on-road vehicle detection: A comparative study. *Machine Vision and Applications*, pages 1–13.
- Smith, R. (2007). An overview of the tesseract ocr engine.

- Soumya, K., Babu, A., and Therattil, L. (2014). License plate detection and character recognition using contour analysis. *International Journal of Advanced Trends in Computer Science and Engineering*, 3(1):15–18.
- Stuart, J., Segal, E., Koller, D., and Kim, S. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science*, 302(5643):249–255.
- Suresh, K., Kumar, M., and Rajagopalan, A. (2007). Superresolution of license plates in real traffic videos. *Transactions on Intelligent Transportation Systems*, 8(2):321–331.
- Tang, Y., Zhang, C., Gu, R., Li, P., and Yang, B. (2015). Vehicle detection and recognition for intelligent traffic surveillance system. *Multimedia Tools and Applications*, online:1–16.
- Wang, M.-L., Liu, Y.-H., Liao, B.-Y., Lin, Y.-S., and Horng, M.-F. (2010). A vehicle license plate recognition system based on spatial/frequency domain filtering and neural networks. In *International Conference on Computational Collective Intelligence*, pages 63–70. Springer.
- Wang, R., Wang, G., Liu, J., and Tian, J. (2013). A novel approach for segmentation of touching characters on the license plate. In *International Conference on Graphic and Image Processing*, pages 876847–876847.
- Wen, X., Shao, L., Fang, W., and Xue, Y. (2015). Efficient feature selection and classification for vehicle detection. *Transactions on Circuits and Systems for Video Technology*, pages 508–517.
- Wen, Y., Lu, Y., Yan, J., Zhou, Z., von Deneen, K., and Shi, P. (2011). An algorithm for license plate recognition applied to intelligent transportation system. *Transactions on Intelligent Transportation System*, 12(3):830–845.
- Xing-lin, F. and Yun-lou, F. (2012). A new license plate character segmentation algorithm based on priori knowledge constraints. *Journal of Chongqing Technology and Business University (Natural Science Edition)*, 8:11.
- Yang, J., Jiang, Y.-G., Hauptmann, A. G., and Ngo, C.-W. (2007). Evaluating bag-of-visual-words representations in scene classification. In *International Workshop on Multimedia Information Retrieval*, pages 197–206. ACM.

- Yang, L., Luo, P., Loy, C. C., and Tang, X. (2015). A large-scale car dataset for fine-grained categorization and verification. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3973–3981.
- Yao, C., Bai, X., Liu, W., Ma, Y., and Tu, Z. (2012). Detecting texts of arbitrary orientations in natural images. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conf. on*, pages 1083–1090. IEEE.
- Zhang, H., Jia, W., He, X., and Wu, Q. (2006). Learning-based license plate detection using global and local features. In *International Conference on Pattern Recognition (ICPR)*, pages 1102–1105.