

# **Detecção de Placas Veiculares : Uma Abordagem Adaptativa para a Localização de Placas Utilizando Descritor HOG**

Raphael Felipe de Carvalho Prates  
Universidade Federal de Ouro Preto

Dissertação submetida ao  
Instituto de Ciências Exatas e Biológicas  
Universidade Federal de Ouro Preto  
para obtenção do título de Mestre em Ciência da Computação



P912d

Prates, Raphael Felipe de Carvalho.

Detecção de placas veiculares [manuscrito]: uma abordagem adaptativa para a localização de placas utilizando descritores HOG / Raphael Felipe de Carvalho Prates – 2014.

99f.: il.; color.; grafs.; tabs.; algoritmos.

Orientador: Prof. Dr. Guillermo Cámara Chávez.

Co-orientador: Prof. Dr. David Menotti Gomes

Dissertação (Mestrado) - Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Departamento de Computação. Programa de Pós- graduação em Ciência da Computação.  
Área de concentração: Visão Computacional.

1. Veículos – Teses. 2. Sistemas de recuperação da informação - Teses. I. Cámara Chávez, Guillermo. II. Gomes, David Menotti. III. Universidade Federal de Ouro Preto. IV. Título.

CDU:004.424.5:629.3

Catálogo: [sisbin@sisbin.ufop.br](mailto:sisbin@sisbin.ufop.br)



*Dedico este trabalho a meus pais, Sandra e Marcelo, ambos exemplos de coragem,  
determinação e bondade para minha vida.*



# **Detecção Automática de Placas Veiculares: Uma Abordagem Adaptativa para a Localização de Placas Veiculares Utilizando Descritor HoG**

## **Resumo**

Um sistema de reconhecimento de placas veiculares é responsável por extrair informações visuais que possibilitem a identificação automática de veículos a partir de uma imagem ou sequência de imagens. Trata-se de uma área com muitos trabalhos publicados recentemente devido a sua aplicação em situações cotidianas como no pagamento automático de pedágios e no monitoramento do tráfego de veículos. Uma das etapas para a identificação dos veículos é a detecção das placas veiculares na imagem. Essa é considerada a fase mais crítica do sistema devido às dificuldades decorrentes de variações ambientais (dia, noite, sol e chuva) e entre diferentes placas (sujeiras, cores, fontes e oclusão). Uma abordagem comum na detecção de placas veiculares é varrer a imagem usando uma janela deslizante com uma proporção entre a largura e a altura semelhante a da placa veicular. No entanto, para lidar com diferentes escalas é necessário fazer esse processo não em uma escala, mas em uma pirâmide de diferentes escalas. Esse método se opõe às exigências de desempenho do sistema que deve ser tempo-real. Assim, neste trabalho é proposto um método de detecção de placas veiculares utilizando janelas deslizantes que visa uma detecção eficaz analisando um número bem inferior de janelas na imagem quando comparado a abordagem em pirâmide. Com esse objetivo, dividimos o processo em duas etapas principais: (1) busca em uma quantidade mínima de escalas utilizando janelas deslizantes de forma a obter uma estimativa inicial da posição da placa e (2) a partir de cada estimativa inicial é realizada uma decomposição em mais escalas localmente, buscando um melhor casamento entre a janela de detecção e a placa real.

Os resultados encontrados comprovam as vantagens do método proposto, são obtidas detecções mais precisas analisando 25% e 75% menos janelas que um método multiescalas convencional para as bases brasileira e grega, respectivamente. A confiabilidade do método é garantida por uma revocação de 98% para coeficiente de *Jaccard* de 50% em duas bases de imagens de diferentes nacionalidades.

# **Vehicle License Plate Detection: An Adaptative Method For Locating Vehicle License Plates Using HoG Descriptor**

## **Abstract**

A License Plate Recognition System is responsible for extracting visual information that enables the automatic identification of vehicles from an image or a sequence of images. It is an area with many works published recently due its application in numerous real-life situations, as automatic toll collection and vehicle traffic monitoring. One of the stages for vehicles identification is the license plates detection in the image. It is considered the most critical step of the system due the difficulty associated to environmental variations (day, night, sunny, cloudy) and among different license plates (dirts, colors, fonts and occlusion). A common approach in license plate detection is to scan the image using a sliding window with a ratio between width and height similar to the license plate. However, to deal with different scales is necessary do this process not at a single scale, but in a pyramid of different scales. This approach is opposed to the performance requirements of the system that must be real-time. So, in this work is proposed a method of detecting vehicle license plates using sliding windows that aims at effectively detecting analyzing very fewer windows in the image than the pyramid approach. With this objective, we divided the process in two main stages: (1) search at a minimum amount of scales using sliding windows aiming to get an initial estimate of the real license plate position and (2) from each initial estimate is conducted a local decomposition at more scales, in order to reach a better matching between the detected window and the real license plate. The results confirm the benefits of the proposed method, more precise detection are reached analyzing 25% and 75% fewer windows than

a common multiscale method. The reliability is ensured by a recall of 98% for Jaccard index of 50% in two databases of different nationality.

## **Declaração**

Esta dissertação é resultado de meu próprio trabalho, exceto onde referência explícita é feita ao trabalho de outros, e não foi submetida para outra qualificação nesta nem em outra universidade.

Raphael Felipe de Carvalho Prates



## Agradecimentos

Agradeço a todos que me ajudam direta ou indiretamente neste trabalho.

Agradeço a *meus pais*, pela paciência, suporte financeiro e amor.

Agradeço a *Luiza Lima*, pelo carinho, companherismo e amor.

Agradeço aos *professores*, pela preocupação.

Agradeço aos *amigos*, pelos conselhos.

Agradeço ao *Guillermo* e ao *David Menotti*, respectivamente, orientador e coordenador.

Agradeço à *UFOP* e *Ouro Preto*, pelos ensinamentos.

Agradeço ao professor *Gustavo Peixoto* (DECOM-UFOP) e ao Laboratório de Visão Computacional da UNICAMP (LIV) pelo auxílio na realização dos experimentos.

Agradeço aos *esquecidos* não mencionados explicitamente aqui.

Agradeço ao *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq).

Muito Obrigado.



# Sumário

<b>Lista de Figuras</b>	<b>xix</b>
<b>Lista de Tabelas</b>	<b>xxiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Considerações Iniciais . . . . .	1
1.2 Justificativa . . . . .	4
1.3 Objetivos . . . . .	6
1.4 Principais Contribuições . . . . .	8
1.5 Organização do Texto . . . . .	8
<b>2 Revisão Bibliográfica</b>	<b>11</b>
2.1 Métodos Baseados na Detecção de Bordas . . . . .	12
2.2 Métodos Baseados em Textura e Formas . . . . .	13
2.3 Métodos Baseados nas Características de Cor . . . . .	15
2.4 Métodos Baseados nas Informações dos Caracteres . . . . .	17
<b>3 Conceitos Fundamentais</b>	<b>19</b>
3.1 Histograma de Gradientes Orientados (HOG) . . . . .	19
3.2 Supressão de Não-Máximos . . . . .	22

3.3	Cobertura das Janelas de Detecção . . . . .	23
3.4	Região Mínima de Busca Local . . . . .	25
3.5	Detecção empregando janelas deslizantes . . . . .	26
3.5.1	Método de Análise em Múltiplas Escalas utilizando Pirâmide de Imagens . . . . .	28
3.6	Considerações Finais . . . . .	29
<b>4</b>	<b>Método Proposto</b>	<b>31</b>
4.1	Introdução . . . . .	31
4.2	Método de Busca Local e Adaptativa em Múltiplas Escalas . . . . .	32
4.2.1	Estimativa da Posição Inicial . . . . .	33
4.2.2	Estimativa do Tamanho . . . . .	34
4.2.3	Refinamento da Localização . . . . .	36
4.2.4	Algoritmo . . . . .	36
4.2.5	Considerações Finais . . . . .	38
<b>5</b>	<b>Resultados e Discussões</b>	<b>41</b>
5.1	Base de Imagens . . . . .	41
5.2	Definição das Métricas de Avaliação e Calibração de Parâmetros . . . . .	42
5.2.1	Definição das Métricas de Avaliação . . . . .	44
5.2.2	Definição das Dimensões da Janela de Detecção . . . . .	47
5.2.3	Definição das Configurações do Descritor HOG . . . . .	49
5.2.4	Parâmetros do Método de Janelas Deslizantes Estático . . . . .	52
5.2.5	Parâmetros do Método de Janelas Deslizantes Adaptativo . . . . .	54
5.3	Método proposto em (Mendes et al. 2011) . . . . .	57
5.4	Configuração Utilizada nos Experimentos . . . . .	58

5.4.1	Parâmetros Base Brasileira . . . . .	59
5.4.2	Parâmetros Base Grega . . . . .	59
5.5	Experimentos . . . . .	60
5.6	Considerações Finais . . . . .	65
<b>6</b>	<b>Conclusões</b>	<b>69</b>
	<b>Referências Bibliográficas</b>	<b>71</b>



# Lista de Figuras

1.1	Representação Esquemática do Funcionamento de um SIAV. . . . .	2
3.1	Representação esquemática dos descritores HOG. . . . .	22
3.2	Diferentes posições relativas para a janela de detecção(borda contínua) e região rotulada(borda pontilhada). A janela de detecção e região rotulada possuem dimensões em pixel de $w \times h$ e $m \times n$ pixels, respectivamente. . . . .	24
3.3	Exemplos de distância entre os centros da janela de detecção (região $D$ em azul pontilhada) e a placa veicular rotulada (região $G$ em vermelho contínua) para diferentes tamanhos de placa veicular. . . . .	27
3.4	Ilustração da região mínima de busca local que deve ser utilizada para que seja possível englobar 100% da placa veicular quando ambas as regiões possuem a mesma dimensão. Nessa, as regiões retangulares $D$ , $G$ e $R$ , corresponde, respectivamente, a janela de detecção, região rotulada e a região mínima de busca local. . . . .	28
3.5	Representação esquemática das janelas deslizantes em uma pirâmide de imagens. . . . .	29
4.1	Visão geral do método proposto. As regiões retangulares destacadas em azul são as nossas principais contribuições. . . . .	33
4.2	Estimativa da posição inicial da placa veicular usando HOG e SVM. . . . .	34
4.3	Processo de estimativa local e em múltiplas escalas do tamanho da placa veicular. . . . .	35
4.4	Refinamento da posição de forma adaptativa. . . . .	37

4.5	Representação esquemática do processo de busca local e adaptativa em múltiplas escalas . . . . .	38
5.1	Exemplos de imagens da base brasileira. . . . .	43
5.2	Exemplos de imagens da base grega. . . . .	44
5.3	Representação de detecções com diferentes valores de Jaccard. A região em vermelho representa a placa veicular rotulada, enquanto que as regiões em azul são as detecções. A primeira linha mostra as regiões no intervalo de $[0, 5; 0, 6[$ , a segunda $[0, 6; 0, 7[$ , a terceira $[0, 7; 0, 8[$ e a última $[0, 8; 0, 9[$ . . . . .	46
5.4	Histograma de altura das placas do conjunto de treinamento. À esquerda, temos o histograma de alturas para a base brasileira e, à direita, para a base grega. . . . .	47
5.5	Histograma de largura das placas do conjunto de treinamento. À esquerda, temos o histograma de larguras para a base brasileira e, à direita, para a base grega. . . . .	48
5.6	Análise da influência de diferentes configurações de janelas de detecção no conjunto de validação. À esquerda, temos os resultados para a base brasileira e, à direita, para base grega. . . . .	49
5.7	Análise da influência de diferentes configurações de blocos e células quanto a revocação para a base brasileira. . . . .	50
5.8	Análise da influência de diferentes configurações de blocos e células quanto a precisão para a base brasileira. . . . .	50
5.9	Análise da influência de diferentes configurações de blocos e células quanto a revocação para a base grega. . . . .	51
5.10	Análise da influência de diferentes configurações de blocos e células quanto a precisão para a base grega. . . . .	51
5.11	Cobertura da base de treino brasileira (esquerda) e grega (direita). Os pontos em verde correspondem a dimensões cobertas, enquanto que os pontos em azul são exemplares da base de treino. As dimensões cobertas foram geradas utilizando coeficiente de <i>Jaccard</i> de <b>80%</b> e considerando deslocamentos unitários das janelas de detecção. . . . .	53

5.12	Avaliação da Influência dos Deslocamentos para as bases brasileira (esquerda) e grega (direita). Cada deslocamento é representado por um valor $w \times h$ pixels que corresponde a ao deslocamento utilizado na horizontal e na vertical, respectivamente. A janela de detecção desloca horizontalmente $w$ pixels até chegar ao final da imagem, retornando ao início, mas com um deslocamento na vertical de $h$ pixels, de forma a cobrir uma região diferente da imagem. . . . .	53
5.13	Cobertura da base de treino brasileira (esquerda) e grega (direita). Os pontos em verde correspondem a dimensões cobertas, enquanto que os pontos em azul são exemplares da base de treino. As dimensões cobertas foram geradas utilizando coeficiente de <i>Jaccard</i> de <b>50%</b> e considerando deslocamentos unitários das janelas de detecção. . . . .	54
5.14	Análise de diferentes valores para deslocamento em valores $w \times h$ pixels, onde $w$ é o deslocamento horizontal e $h$ o deslocamento vertical. Para a determinação de uma região como detectada ou não é utilizado coeficiente de <i>Jaccard</i> de 50%. A imagem à esquerda apresenta os resultados para a base brasileira e, à direita, temos os resultados da base grega. . . . .	55
5.15	Análise de diferentes valores para deslocamento em valores $w \times h$ pixels, onde $w$ é o deslocamento horizontal e $h$ o deslocamento vertical. Esses deslocamentos são empregados na Busca local e adaptativa. O gráfico à esquerda apresenta os resultados para a base brasileira, enquanto que à direita temos os resultados para a base grega. . . . .	56
5.16	Análise de diferentes configurações de parâmetros. O gráfico à esquerda traz os resultados para a base brasileira e, à direita, temos os resultados para a base grega. . . . .	58
5.17	Resultados em termos de Revocação $\times$ <i>Jaccard Index</i> para a base brasileira utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos <b>adaptativo</b> , <b>estático</b> e <b>híbrido</b> . . . . .	61
5.18	Resultados em termos de Precisão $\times$ <i>Jaccard Index</i> para a base brasileira utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos <b>adaptativo</b> , <b>estático</b> e proposto em (Mendes et al. 2011). . . . .	63

5.19	Resultados em termos de Revocação $\times$ <i>Jaccard Index</i> para a base grega utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos <b>adaptativo</b> , <b>estático</b> e <b>híbrido</b> . . . . .	64
5.20	Resultados em termos de Revocação $\times$ <i>Jaccard Index</i> para a base grega utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos <b>adaptativo</b> , <b>estático</b> e <b>híbrido</b> . . . . .	65
5.21	Gráfico de Revocação $\times$ <i>Jaccard Index</i> para o método <b>adaptativo</b> proposto utilizando apenas a estimativa inicial da posição da placa veicular e empregando a etapa de busca local adaptativa em múltiplas escalas. . . . .	66
5.22	Exemplos de resultados obtidos com o método adaptativo (esquerda) e o método híbrido (direita). As regiões em vermelha correspondem às rotulações, enquanto que as regiões em azul são as detectadas. . . . .	67
5.23	Exemplos de resultados obtidos com o método adaptativo (esquerda) e o método estático (direita). As regiões em vermelha correspondem às rotulações, enquanto que as regiões em azul são as detectadas. . . . .	68

# Lista de Tabelas

5.1	Comparação dos parâmetros para o método híbrido . . . . .	58
5.2	Comparação de eficiência dos métodos baseados em janelas deslizantes. .	61



# Lista de Algoritmos

3.1	Método de Supressão de Não-Máximos. . . . .	23
4.1	Método de Busca Local e Adaptativa . . . . .	39



# Capítulo 1

## Introdução

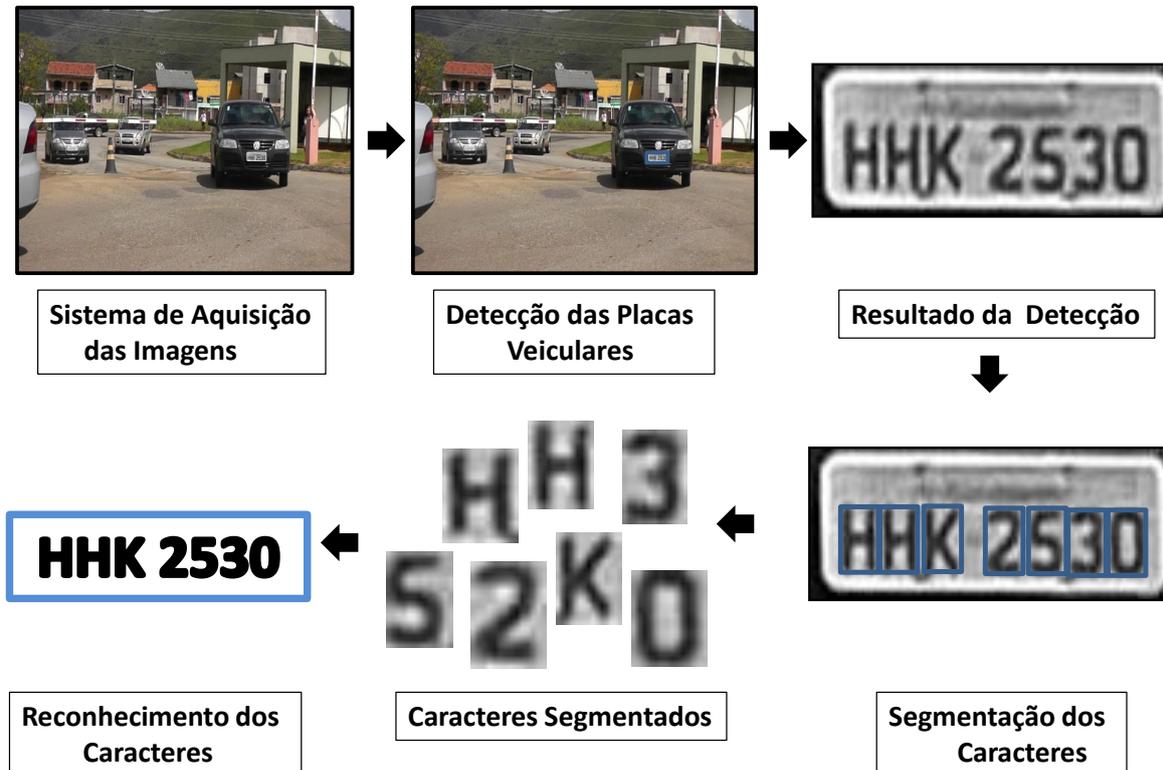
A frota de veículos do tipo automóvel no Brasil já ultrapassa os 42,5 milhões de unidades (Ministério das Cidades, Departamento Nacional de Trânsito - DENATRAN 2014). Além de trazerem comodidade para os proprietários, os veículos também modificam o ambiente urbano, impondo desafios quanto a questões de poluição e segurança, os quais são ainda mais preocupantes nos grandes centros urbanos. O monitoramento desses veículos utilizando técnicas como o processamento de imagens já é considerada uma necessidade e, por isso, tem atraído muitas pesquisas recentemente. Dentro desse contexto se inserem os Sistemas de Identificação Automática de Veículos (SIAV).

### 1.1 Considerações Iniciais

Em um SIAV, objetiva-se a partir de uma imagem ou sequência de imagens ser capaz de reconhecer todos os veículos presentes pelo número da placa veicular. Esses sistemas utilizam de técnicas provenientes de diferentes áreas, como o processamento de imagens, reconhecimento de padrões e a visão computacional. Podemos destacar três módulos principais em um SIAV: (1) detecção de placas veiculares; (2) segmentação de caracteres e (3) reconhecimento dos caracteres, como pode ser visto na Figura 1.1.

Este trabalho aborda o desenvolvimento de um sistema de detecção de placas veiculares, considerada a etapa mais crítica e de maior impacto no desempenho (Du et al. 2013). Entre as aplicações para um SIAV, temos o monitoramento do tráfego em rodovias, pagamento automático de tarifas de pedágios e estacionamentos privativos, controle de acesso de veículos em áreas restritas, identificação de veículos roubados e, mais recente-

mente, a proteção de privacidade em imagens obtidas em ambientes urbanos por sistemas comerciais como o *Google Street View* (Frome et al. 2009).



**Figura 1.1:** Representação Esquemática do Funcionamento de um SIAV.

Apesar de muitos trabalhos reportarem excelente desempenho, inclusive alguns operando em aplicações reais, tratam-se de métodos que operam muito bem em ambientes controlados, mas que têm o desempenho prejudicado à medida que se aumenta a complexidade. Assim, a detecção de placas veiculares em ambientes abertos é ainda um desafio. As principais dificuldades estão relacionadas às variações nas placas (diferentes cores, estado de conservação e até mesmo pela presença de sujeiras) e ambientais (diferentes condições de iluminação, perspectiva e distância da placa veicular variáveis, condições climáticas e a identificação de múltiplos veículos).

A detecção de placas veiculares pode ser realizada com base nas informações provenientes dos caracteres e de bordas da placa veicular. Muitos métodos se baseiam na textura resultante da diferença de cor dos caracteres em relação ao plano de fundo das placas veiculares. Diferentes descritores de textura e formas vêm sendo empregados,

tais como os descritores do tipo *Haar*, Transforma de *Wavelet* Discreta, *Scale-Invariant Features* (SIFT) e Histograma de Gradientes Orientados (*Histogram of Oriented Gradients - HOG*) (Du et al. 2013). Sendo o último aquele que tem apresentado os melhores resultados (Wang et al. 2013).

Visando a detecção de placas de diferentes tamanhos, uma abordagem comum é a utilização de uma pirâmide de imagens em diferentes escalas, abordagem conhecida como multiescalas. Cada imagem da pirâmide é percorrida por uma janela retangular e para cada uma dessas regiões é extraído um conjunto de atributos que são classificados utilizando técnicas de Inteligência Artificial (IA) como placas ou não-placas. Os classificadores mais empregados são o *Adaboost* (*ADAPTative BOOSTing*) e a Máquina de Vetores de Suporte (*Support Vector Machine - SVM*). Nessa abordagem clássica, a utilização do HOG resulta em um custo computacional proibitivo devido a enorme quantidade de janelas analisadas.

O método desenvolvido neste trabalho parte de duas premissas: (1) é possível obter uma boa aproximação inicial da posição da placa veicular utilizando uma quantidade mínima de escalas com janelas deslizantes e, (2) a partir dessa aproximação inicial consegue-se obter uma localização mais precisa da placa veicular utilizando um método local e adaptativo em mais escalas. Assim, nosso método pode ser dividido em dois módulos principais: (1) busca de uma estimativa inicial da posição da placa veicular utilizando poucas escalas e (2) para cada um dos candidatos iniciais é realizada uma busca local e adaptativa em mais escalas de forma a garantir uma localização e tamanho mais precisos da placa.

Em ambas as etapas, é utilizado um mesmo classificador SVM linear treinado a partir de descritores HOG. Dessa forma, pretendemos garantir uma detecção mais precisa da placa que os métodos em múltiplas escalas convencionais e que analise uma quantidade reduzida de candidatos, diminuindo o tempo necessário para a detecção. Uma comparação dos resultados obtidos em duas bases de dados é realizada com os métodos multiescalas clássico e um método baseado na detecção de bordas descrito em (Mendes et al. 2011).

## 1.2 Justificativa

De um algoritmo de detecção de placas veiculares, espera-se que seja capaz de responder a duas questões: onde se encontra a placa veicular e qual sua dimensão. Além disso, o processo de detecção deve ser eficaz e eficiente. Alguns métodos propostos na literatura conseguem resolver esse problema de forma parcial e/ou restrita a determinados ambientes.

Um dos maiores desafios encontrados é a detecção de placas com diferentes tamanhos. Uma abordagem tradicional, para problemas de detecção em geral, é o emprego de uma pirâmide de imagens em diferentes escalas. Cada uma das escalas é analisada por uma janela deslizante de tamanho fixo que foi treinada em uma única escala, denominada escala canônica (Benenson et al. 2012). Trata-se uma técnica com alto impacto no custo computacional e que requer algum pós-processamento de forma a lidar com múltiplas detecções para um mesmo objeto.

Na literatura, temos muitos trabalhos que buscam combinar atributos simples em um arranjo de classificadores em cascata de forma a obterem métodos de tempo-real. Um bom exemplo são as cascatas de classificadores *Adaboost* baseados em atributos do tipo *Haar*, os quais pela sua simplicidade e devido a representação imagem integral (Viola & Jones 2001) conseguem alcançar os requisitos de tempo-real, mas que têm sua eficiência prejudicada pela alta taxa de falsos positivos (Zhang et al. 2010).

Diferentes trabalhos da literatura abordam o emprego de descritores de forma robustos como o HOG e SIFT. Nesses trabalhos costuma-se dividir o processo de detecção em duas etapas: (1) geração de candidatos e (2) verificação de candidatos. Porém, por preocupações de desempenho, tais descritores só são utilizados na etapa de verificação de candidatos. Nessa etapa, regiões que não são placas, mas que foram classificadas como positivas pela etapa anterior devem ser eliminadas. O grande problema dessa abordagem é que algumas regiões que são placas veiculares podem ter sido descartadas na etapa prévia devido a simplicidade dos descritores empregados.

O ideal é utilizar descritores com maior poder de discriminação em todos os estágios da detecção. Uma questão que surge, é como utilizar esses descritores sem que haja uma perda significativa do desempenho computacional. Existe uma iniciativa nesse sentido: a detecção de placas veiculares apenas depois da detecção do veículo. No primeiro estágio, é feita uma localização da dianteira ou traseira do veículo e a partir dessas regiões realizar uma busca pela placa veicular. Para cada problema de detecção (veículo e placas), são

treinados classificadores SVM utilizando descritores HOG (Li 2010). Além de aumentar a complexidade na fase de treino, tal método também realiza uma busca pela placa veicular em uma região extensa da imagem, no caso, toda a traseira ou dianteira do veículo e qualquer erro na identificação dos veículos é propagado para a detecção da placa veicular.

Sabe-se que, no problema de detecção de placas veiculares, a maior parte da imagem é constituída por plano de fundo, enquanto que apenas uma parcela bem pequena é a região que desejamos como resultado. Assim, apesar de seus pontos fracos, o trabalho de Li (2010) aponta para uma forma de se possibilitar a utilização de descritores mais complexos em uma abordagem de detecção multiescalas: a redução do espaço de soluções. A partir de uma estimativa inicial da posição da placa veicular (por exemplo, dianteira ou traseira do veículo), é possível focar uma busca em múltiplas escalas apenas naquele local, já descartando a análise desnecessária de outras janelas e reduzindo o custo computacional. Além disso, espera-se que a quantidade de escalas analisadas não seja uma quantidade fixa e que se adapte as necessidades do ambiente.

Apesar da maioria dos trabalhos que tratam o problema de detecção como uma classificação binária utilizarem multiescalas, em (Dlagnekov 2004), temos o trabalho pioneiro que utilizou atributos *Haar* simples com *Adaboost* na detecção de placas veiculares em uma única escala, alcançando uma taxa de detecção de 95,6% e uma taxa de 5,7% falsos positivos. Esse trabalho é importante por mostrar que mesmo com atributos simples podemos obter uma boa estimativa da posição das placas veiculares na imagem mesmo em uma única escala. Porém, para obter uma posição e tamanho precisos para as etapas posteriores de reconhecimento é necessária uma análise em múltiplas escalas (Wang et al. 2013).

Dessa forma, são necessários estudos com o objetivo de utilizar a informação advinda da detecção em uma quantidade mínimas de escalas com a finalidade de se reduzir a quantidade de janelas analisadas. Além disso, devido à redução na quantidade de janelas analisadas torna-se possível empregar descritores com maior poder de discriminação, tais como HOG e SIFT sem inviabilizar o sistema. Em (Wang et al. 2013), temos uma comparação entre o poder de discriminação dos descritores HOG e SIFT para o problema de verificação de placas veiculares, mostrando que o HOG apresenta resultados superiores. Esse trabalho justifica a utilização do HOG em nosso trabalho em detrimento do SIFT.

## 1.3 Objetivos

A detecção de placas veiculares em ambientes abertos é ainda um desafio devido a complexidade e variedade de placas e condições ambientais. Uma forma de tratar o problema de complexidade é utilizando descritores com maior poder de discriminação. Esses descritores são geralmente empregados em uma abordagem multiescalas, sendo extraídos em regiões retangulares de tamanho fixo que varrem a imagem tentando localizar a posição das placas veiculares. Trata-se um método com elevado custo computacional, inviabilizando a sua aplicação prática.

O objetivo principal deste trabalho é desenvolver um método de detecção de placas veiculares em múltiplas escalas utilizando descritores HOG que apresente um custo computacional inferior a abordagem clássica. Trata-se de um importante passo rumo ao desenvolvimento de um SIAV em ambientes abertos em tempo-real. Com essa finalidade, uma diferente forma de busca no espaço de possíveis soluções deve ser elaborada, alcançando uma maior eficiência sem que haja uma perda na eficácia.

Trabalhos existentes na literatura sugerem que é possível obter taxas de detecção excelentes utilizando descritores simples do tipo *Haar* em uma única escala (Dlagnekov 2004). Além disso, devido a maior dimensionalidade e a informação espacial, os descritores HOG apresentam maior poder de discriminação que os descritores do tipo SIFT e *Haar* (Yoon et al. 2012). Assim, uma análise do desempenho dos classificadores HOG na detecção de placas veiculares utilizando uma quantidade mínima de escalas será conduzido de forma a verificar a aplicabilidade do método para a obtenção de uma estimativa inicial da posição da placa veicular.

As etapas de segmentação e reconhecimento de caracteres têm seu desempenho influenciado pela forma como a detecção da placa veicular é conduzida. Uma detecção em poucas escalas fornece uma estimativa inicial da posição da placa veicular, informação insuficiente para uma determinação precisa da posição e tamanho das placas veiculares. Assim, uma decomposição em mais escalas local deve ser conduzida de forma a se obter as informações da placa veicular com a confiança necessária. Essa etapa é crítica na eficiência do método. Quanto menos escalas tiverem que ser analisadas para se chegar a localização precisa, mais eficiente será o método.

Os métodos encontrados na literatura tratam o problema de detecção de placas em múltiplas escalas utilizando uma quantidade fixa de escalas (Li 2010), abordagem aqui denominada de busca **Estática**. Idealmente, deve-se analisar uma ampla faixa de es-

calas, cobrindo uma maior faixa de tamanho de placas veiculares. Assim, utilizando uma quantidade fixa de escalas acabamos por restringir a dimensão das placas a serem detectadas por questões de eficiência.

No entanto, caso seja possível utilizar alguma informação de forma a direcionar a busca em diferentes escalas, em uma abordagem denominada neste trabalho de **Adaptativa**, a detecção de placas será independente da dimensão e ainda não analisará escalas desnecessárias. Por exemplo, sabendo previamente que a placa tem uma dimensão superior a escala canônica (escala na qual o descritor foi treinado), pode-se focar a busca apenas em escalas superiores. Na abordagem estática, mesmo a placa tendo uma dimensão superior, todas as escalas, superiores e inferiores são analisadas.

No entanto, uma questão a ser respondida na abordagem **adaptativa** é qual informação utilizar para guiar essa busca. Em (Yoon et al. 2012), os autores utilizam o HOG em conjunto com um classificador SVM e julgam a afinidade de uma região ser uma placa veicular pela resposta fornecida pelo SVM ao vetor de atributos HOG extraídos para aquela região. Regiões com maiores resposta são mantidas como placa veicular, enquanto que as demais são descartadas em um processo de verificação de candidatos. A partir desse indício, será feita uma análise da aplicabilidade da resposta do classificador SVM utilizando atributos HOG para direcionar o processo de busca **adaptativa** em múltiplas escalas.

A fim de avaliar o desempenho do método, duas bases de dados de placas veiculares será empregada. Essas bases foram coletadas de ambientes abertos, contendo todas os desafios reportados na literatura quanto a variabilidade entre placas e as complexidades ambientais. Buscando comprovar a viabilidade do método proposto, será conduzida uma comparação com uma abordagem clássica multiescalas utilizando atributos HOG. Uma segunda comparação será conduzida com um método que utiliza informações de bordas e operações morfológicas para a detecção de placas veiculares presente em (Mendes et al. 2011).

Esta dissertação resultou na seguinte publicação:

- Prates, R. F., Cámara-Chávez, G. , Schwartz, W. R. , Menotti, D. (2013). Brazilian License Plate Detection Using Histogram of Oriented Gradients And Sliding Windows. *International Journal of Computer Science and Information Technology (IJCSIT)*.

## 1.4 Principais Contribuições

Neste trabalho é abordado o problema de detecção de placas veiculares utilizando descritores HOG. Podemos destacar como nossas principais contribuições:

- Análise da acurácia do descritor HOG no problema de detecção de placas veiculares utilizando janelas deslizantes em uma pirâmide de imagens em diferentes escalas.
- É proposto um método inédito de detecção de placas veiculares utilizando descritores HOG com janelas deslizantes. Nesse método buscamos utilizar a informação da resposta do descritor em uma quantidade mínima de escalas como um indicativo para uma busca local e adaptativa em um número maior de escalas.
- O método proposto é comparado com uma abordagem presente na literatura, em (Mendes et al. 2011), que utiliza as informações de bordas em conjunto com operações morfológicas e análise de componentes conexos. Assim, mostramos seu melhor desempenho na localização mais precisa da posição e tamanho da placa veicular.

## 1.5 Organização do Texto

O texto desta dissertação encontra-se organizado da seguinte forma:

No Capítulo 2, é apresentada uma revisão bibliográfica do estado da arte, focando nas principais técnicas de detecção de placas veiculares encontradas na literatura: detecção de bordas, análise de textura e formas, características de cor e as informações dos caracteres.

No Capítulo 4, é descrito o método desenvolvido neste trabalho. Assim, é exposto o descritor HOG empregado, bem como as técnicas de janelas deslizantes **estática** e **adaptativa**. São apresentados os conceitos de cobertura da janela de detecção e da área mínima de busca local, os quais servem de suporte para a escolha dos parâmetros do método proposto. A Supressão de não-máximos também é detalhada.

No Capítulo 5, estão os resultados obtidos com o algoritmo desenvolvido na base de imagens. O método é comparado com um método presente na literatura em (Mendes et al. 2011) e com uma implementação utilizando janelas deslizantes de forma **estática**.

No Capítulo 6, conclusões sobre o assunto são discutidas e os trabalhos futuros são definidos.



## Capítulo 2

### Revisão Bibliográfica

A detecção de placas veiculares é considerada a etapa mais desafiadora nos SIAV. Assim, muitos estudos têm se focado nessa tarefa com o intuito de se elevar o desempenho do sistema como um todo. Muitos dos trabalhos desenvolvidos até então apresentam resultados satisfatórios em situações restritas, tais como iluminação uniforme, pequena perspectiva de captura das imagens e planos de fundo mais simples.

O desenvolvimento de um método de detecção de placas veiculares que tenha um desempenho satisfatório em situações complexas ainda é um desafio. Nessa etapa, objetiva-se a partir de uma imagem fornecida como entrada, retornar uma ou mais regiões da mesma contendo placas veiculares. Para tanto, são considerados atributos e o formato retangular das placas veiculares. Analisando informações de cor, textura, características locais de forma, presença de bordas e caracteres em regiões retangulares da imagem é possível determinar quais são placas veiculares (Du et al. 2013). Neste capítulo é apresentada uma revisão bibliográfica dos trabalhos sobre detecção de placas veiculares. Os trabalhos serão categorizados de acordo com os atributos empregados.

É importante ressaltar que uma comparação direta dos resultados obtidos pelos métodos é impossibilitada pela utilização de diferentes bases de imagens, as quais geralmente não são disponibilizadas devido a proteção de privacidade dos proprietários dos veículos. Além disso, para cada tipo de aplicação as dificuldades empregadas são muito distintas. Por exemplo, em estacionamentos privativos as imagens das placas são capturadas a poucos metros do sistema de aquisição enquanto que no monitoramento do tráfego em rodovias, as placas são detectadas a dezenas de metros.

## 2.1 Métodos Baseados na Detecção de Bordas

Os métodos de detecção de placas veiculares baseados nas informações das bordas são muito comuns devido aos seus benefícios: simplicidade e rapidez (Du et al. 2013). A ideia por trás dessas técnicas é explorar a diferença entre a região que compreende a placa e o pára-choque do veículo, além da rica informação de borda e textura fornecida pela transição de cor entre os caracteres e o plano de fundo da placa veicular.

As informações verticais de bordas são as mais empregadas, uma vez que o pára-choques dos carros são ricos em bordas horizontais e costumam confundir os métodos de detecção (Zheng et al. 2005). O operador *Sobel* é a ferramenta mais utilizada para a extração das bordas, apesar do algoritmo *VEDA* proposto em (Al-Ghaili et al. 2008) ser mais rápido e alcançar uma maior taxa de detecção. Outra abordagem também encontrada na literatura recente são as janelas deslizantes concêntricas, na qual um ponto de borda é identificado quando a razão entre o desvio padrão de duas regiões retangulares concêntricas, centradas nesse ponto, é maior que um limiar (Deb et al. 2009).

Ao se realizar a detecção de bordas verticais, algumas regiões do plano de fundo e do próprio carro acabam se destacando, principalmente se as bordas nas regiões das placas forem pouco nítidas. Nesses casos, se faz necessário um pré-processamento da imagem. Em (Zheng et al. 2005), é utilizado um algoritmo para aumentar o contraste nas placas veiculares com base na variância da intensidade das regiões. Tal método apresenta a desvantagem de não ser muito específico, realçando também outras regiões da imagem, enquanto que técnicas baseadas na densidade de bordas são mais precisas e resultam em maiores taxas de detecção (Abolghasemi & Ahmadyfard 2009, Lalimi et al. 2013).

Outra alternativa para lidar com as variações de contraste, é realizar uma equalização histogrâmica de forma auto-regulada utilizando as informações geométricas das regiões detectadas nas imagens (Guo & Liu 2008). Quando as bordas da placa não são muito visíveis, métodos baseados em blocos são preferidos por buscarem por regiões com maior presença de bordas ao invés de se focarem nas bordas das placas (Lee et al. 2004).

Devido às propriedades geométricas da placa, a transformada de *Hough* também já foi estudada para a detecção de placas veiculares, identificando na imagem linhas horizontais e verticais com até 30° de inclinação. A principal desvantagem dessas técnicas está no alto custo de memória e processamento. Resultados de detecção de 98,8% das placas podem ser encontrados em (Duan et al. 2005) utilizando transformada de *Hough* em conjunto com algoritmo de contorno para uma redução do tempo de processamento.

Em alguns casos, pode-se optar por uma abordagem híbrida. Nessa técnica são envolvidas as informações de borda, atributos geométricos de componentes conexos e morfologia para eliminação de falsos positivos, alcançando assim taxas de detecção superiores (Bai & Liu 2004, Mendes et al. 2011). Em (Lalimi et al. 2013), os autores conseguem uma acurácia da detecção de 92,5% em uma base complexa utilizando filtragem de bordas indesejadas com um filtro passa-baixas baseado em regiões, além da eliminação de falsos positivos empregando morfologia e informações geométricas como área, densidade de bordas e a proporção entre altura e largura das regiões.

Em aplicações nas quais a captura da imagem é realizada bem próxima a placa veicular, métodos baseados em bordas são muito indicados pela rapidez (no geral, métodos mais custosos como a transformada de *Hough* são exceções). Mas, ainda assim, uma filtragem das bordas detectadas com base em seu tamanho e orientação se faz necessária de forma a obter uma alta taxa de detecção, muitas vezes superior a 90% (Al-Ghaili et al. 2013, Abolghasemi & Ahmadyfard 2009). No entanto, à medida que a complexidade do ambiente aumenta, o desempenho desses métodos se degrada devido à sensibilidade a ruídos (Wang et al. 2013).

Neste trabalho, não foram utilizados métodos baseados em informações globais de bordas, uma vez que desejamos obter uma detecção precisa da placa veicular e com a menor quantidade de falsos positivos, mesmo em ambientes abertos. Além disso, nossa base se caracteriza por possuir placas com distâncias médias a longas do sistema de aquisição, não sendo a mais indicada para os métodos baseados na detecção de bordas. No entanto, uma comparação com um método baseado em bordas será conduzido, objetivando mostrar a superioridade da técnica proposta.

## 2.2 Métodos Baseados em Textura e Formas

Os métodos baseados em textura, da mesma forma que os métodos de detecção de bordas, são baseados na presença de caracteres nas placas veiculares. A fim de serem visíveis ao olho humano, eles apresentam características de cor diferentes do plano de fundo das placas, constituindo uma rica informação de textura.

A transformada discreta de *Wavelet* é uma das ferramentas empregadas para se capturar as informações de textura, mais precisamente as sub-bandas de frequência *High-Low* (HL) e *Low-High* (LH) por trazerem informações importante sobre as bordas verticais

e horizontais, respectivamente. Em (Wang et al. 2011), os autores, após uma filtragem das linhas com base em seu tamanho, aplicam a detecção das placas utilizando o canal HL para a identificação de regiões candidatas e o LH para a verificação.

Os descritores do tipo *Haar* foram empregados com sucesso na detecção de faces e na extração de caracteres, motivando a sua aplicação na detecção de placas veiculares devido a sua capacidade de discriminação entre diferentes texturas. Sua primeira utilização nesse problema ocorreu em (Dlagnekov 2004), onde é utilizado um classificador do tipo *Adaboost* de forma a selecionar um conjunto de 100 diferentes atributos *Haar*, construindo um classificador forte. Esse trabalho foi importante por demonstrar a aplicabilidade desses atributos na detecção de placas veiculares, mas apresenta desvantagens quanto a alta taxa de falsos positivos e o custo computacional, uma vez que não utilizou a cascata de classificadores.

Após esse trabalho pioneiro, muitos outros surgiram buscando solucionar alguns dos seus pontos fracos. Para conseguirem detectar placas com diferentes tamanhos na imagem, as imagens passaram a ser analisadas em diferentes escalas. De forma a atenderem aos requisitos temporais, a maioria dos trabalhos optou pela estrutura em cascata de classificadores, em sua maioria *Adaboost*. Além de mais rápida, tal estratégia reduz a taxa de falsos positivos. No entanto, o treinamento nesses casos é bastante demorado e a quantidade de camadas necessárias para tornar a taxa de falsos positivos aceitável torna o método computacionalmente proibitivo.

Uma forma encontrada para lidar com a elevada quantidade de falsos positivos é utilizando outros descritores em conjunto com os descritores do tipo *Haar* de forma a analisarem diferentes características da imagem e descartarem uma maior quantidade de falsos positivos nas primeiras camadas. Em (Zhang et al. 2010), os autores utilizaram descritores de textura mais complexos como a matriz de co-ocorrência em escala de cinza e o descritor de Tamura, obtendo uma considerável melhora na precisão do método. Outros trabalhos utilizaram diferentes descritores com a mesma finalidade: características retangulares (Xu et al. 2006), atributos globais baseados na densidade e na variância da densidade das bordas verticais (Zheng et al. 2013), informações de cor utilizando representação HSI (Yao & Yi 2014), HOG (Wang et al. 2013, Zheng et al. 2006) e SIFT (Ho et al. 2009).

Além disso, foi empiricamente mostrado que quando os atributos do tipo *Haar* são calculados a partir das informações de bordas da imagem (Wang et al. 2013) ou apenas as bordas verticais (Zheng et al. 2013) consegue-se um maior poder de discriminação do

que utilizando a intensidade dos *pixels* em tons de cinza.

Diferentes algoritmos para o treinamento do classificador *Adaboost* foram também considerados, tais como *Discrete Adaboost* (DAB), *Real Adaboost* (RAB) e o *Gentle Adaboost* (GAB). Tais algoritmos interferem diretamente no resultado da detecção por determinarem a forma como os pesos dos exemplos de treino são rearranjados. Assim, é mostrado empiricamente em (Cui D. and Gu, D. and Cai, H. and Sun, J. 2009) que entre os algoritmos, aquele que apresenta uma maior taxa de detecção com um menor número de falsos positivos é o GAB. Além disso, nesse trabalho os autores ainda mostram que o desempenho desses classificadores se encontram longe do necessário para as aplicações reais devido a alta taxa de falsos positivos.

Um método comum para eliminação de falsos positivos é a realização de uma etapa de verificação de candidatos. Nessa etapa podem ser utilizadas informações geométricas, análise de componentes conexos e/ou descritores de textura com maior poder de discriminação. Em (Zheng et al. 2006, Yoon et al. 2012), os autores empregam descritores HOG para a verificação de candidatos, elevando a acurácia do sistema.

Neste trabalho, iremos explorar os descritores HOG, os quais são baseados em informações de forma locais da imagem. Essa escolha se deve aos resultados superiores em comparação aos descritores do tipo *Haar* e SIFT reportados na literatura. No entanto, ao invés de utilizar o HOG apenas na última etapa, a verificação de candidatos, iremos empregar o HOG em todo o processo de detecção. Partimos do pressuposto de que utilizando descritores com maior poder de discriminação em todas as etapas do processo estamos evitando a não-localização das placas veiculares nas etapas iniciais. De forma a lidar com o alto custo computacional, uma abordagem diferente de busca no espaço de possíveis soluções será utilizado.

## 2.3 Métodos Baseados nas Características de Cor

Existe uma padronização nas características de cor das placas veiculares, as quais variam de país para país, mas que podem ser utilizadas para a detecção de placas veiculares em um âmbito nacional. Como essas cores costumam se diferenciar com relação aos demais objetos presentes na imagem, elas podem ser utilizadas para a detecção da placa veicular. A ideia é localizar na imagem regiões retangulares com propriedades de cor semelhante às placas veiculares.

Uma forma de lidar com as variações de cor em função das condições ambientais é a utilização da lógica difusa (*fuzzy*). Em (Chang et al. 2004), os autores introduzem um detector de bordas baseado em informações de cor na representação RGB. As bordas são calculadas a partir de um conjunto específico de transições (branco-preto, vermelho-branco e verde-branco) devido às características de cor das placas veiculares de Taiwan. A pertinência de determinada borda à placa veicular é determinada utilizando lógica difusa. Em (Wang et al. 2008), é proposta uma técnica de detecção de placas usando a informação do espaço de cor HSI em conjunto com funções de pertinência em lógica difusa. Todos os parâmetros são ajustados em um processo de aprendizado, tornando o método altamente adaptativo.

Outra técnica bastante utilizada para filtrar e detectar regiões da imagem contendo placas veiculares é a *Mean Shift* (Jia et al. 2005, Jia et al. 2007, Pan & Li 2010). Em (Jia et al. 2007), o *Mean Shift* é empregado para segmentar a imagem em regiões de interesse com base nas informações de cor, as quais são depois analisadas com relação as suas propriedades geométricas e densidade de bordas. Em (Pan & Li 2010), os autores buscam resolver uma das principais desvantagens do *Mean Shift*, o custo computacional, propondo o *Fast Mean Shift* para a detecção de placas veiculares.

Podem ser encontrados na literatura outros trabalhos que utilizam das informações de cor na detecção de placas veiculares, como o algoritmo CAMShift combinado com SVM para análise de textura das cores (KwangIn et al. 2002), algoritmos genéticos de forma a definir diferentes limiares de cores para as regiões de placas veiculares (Yohimori et al. 2004) e um modelo hexagonal do baricentro das cores é utilizado para a localização das placas veiculares, uma vez que é mais robusto às variações ambientais, permitindo o estabelecimento de limiares de detecção mais confiáveis (Wan et al. 2011).

Os métodos baseados em cores apresentam a vantagem de detectarem placas com deformações e diferentes inclinações. No entanto, esses métodos são sensíveis às variações ambientais e apresentam resultados ruins quando existem regiões na imagem, como as demais partes do veículo, que apresentam características de cores semelhantes as das placas veiculares. Como em nossa base de treino existe uma quantidade expressiva de veículos de coloração prateada (38%), cor utilizada na placa veicular dos veículos particulares, a utilização das informações de cor se torna inviável.

## 2.4 Métodos Baseados nas Informações dos Caracteres

Uma outra abordagem para a detecção de placas veiculares é a busca na imagem por regiões contendo caracteres. Dado que uma região contém caracteres, são avaliadas as demais características de forma a classificá-la como placa ou não-placa veicular.

Em (Alegria & Girão 2006), é proposto um método que realiza a detecção das placas veiculares a partir da análise de componentes conexos em uma imagem localmente binarizada. Aqueles componentes com propriedades geométricas semelhantes aos caracteres e que cuja transformada de *Hough* forma duas linhas horizontais e paralelas são empregados na detecção da placa veicular. Dessa forma, os autores chegam a uma acurácia de 91%. No entanto, eles lidam com imagens extraídas bem próximas a placa veicular, reduzindo a influência dos ruídos ambientais.

Uma binarização múltipla das imagens para a detecção das regiões de caracteres é explorado em (Lee & Choi 2010) em conjunto com uma análise de compactação das regiões. As regiões obtidas são ao final classificadas em caracteres ou não utilizando uma rede neural e um conjunto de 11 atributos baseados em suas posições relativas e informações locais de intensidade.

Em (Matas & Zimmermann 2005), os autores propõem um método para detecção de placas veiculares com base na análise de regiões extremas. Estas regiões são caracterizadas como qualquer componentes conexo encontrado na imagem após uma binarização da mesma utilizando um limiar. Descritores são calculados em tempo computacional linear e de forma incremental para diferentes limiares. As regiões descritas são classificadas utilizando redes neurais em regiões de interesse ou não. As regiões de interesse são então combinadas de acordo com o tamanho e usando transformada de *Hough*. Baseado nesse trabalho, em (Donoser et al. 2007), temos um detector de placas veiculares baseado em Regiões Extremas de Máxima Estabilidade (*Maximally Stable Extremal Regions - MSER*). Os autores utilizam duas variações do MSER, respectivamente MSER+ e MSER-, para a obtenção de regiões semelhantes a placas e caracteres, respectivamente. Uma análise relativa dessas regiões permite a detecção das placas sem nenhum prévio treinamento.

Em (Lim & Y.H. 2010), é apresentada uma técnica de detecção de placas veiculares utilizando MSER em conjunto com descritores SIFT. Um classificador do tipo *Core Vector Machine*(CVM) é utilizado para determinar se um ponto de interesse corresponde ou não a um caractere. As regiões retornadas pelo MSER e que possuem 2 ou mais pontos

de interesse classificados como caracteres são utilizadas para localizar a região da placa veicular. Outra abordagem é a modelagem do problema como sendo uma detecção a partir da identificação de componentes. As regiões prováveis de serem caracteres são obtidas utilizando MSER e suas relações são modeladas utilizando Campos Randômicos Condicionais (*Conditional Random Field - CRF*) (Li et al. 2013).

Em (Ho et al. 2009), é empregado inicialmente 36 classificadores *Adaboost* utilizando descritores do tipo *Haar* para a detecção de regiões candidatas a caracteres. Essas são filtradas em uma etapa de verificação de candidatos utilizando descritores do tipo SIFT. Esse mesmo descritor também é utilizado em (Zhou et al. 2012) na detecção de caracteres com uma representação denominada *Principal Visual Word* (PVW).

Os métodos que realizam a detecção da placa a partir dos caracteres são interessantes por já realizarem simultaneamente a detecção das placas e a segmentação dos caracteres. Esses métodos, no entanto, são sensíveis a outros textos presentes na imagem e requerem modelos que relacionam a estrutura rígida dos caracteres nas placas veiculares, tornando-os mais complexos. Assim, neste trabalho foi mantido o foco na placa, a qual fornece mais informações locais de forma para uma detecção utilizando descritores HOG.

# Capítulo 3

## Conceitos Fundamentais

Neste capítulo, serão descritos os principais conceitos relacionados ao nosso trabalho. Inicialmente, é apresentado os descritores HOG, os quais são empregados de forma a capturar as informações locais de forma das placas veiculares, permitindo a sua discriminação em relação aos demais objetos presentes na imagem. Apresentamos então como é realizada a eliminação de múltiplas detecções de um mesmo objeto utilizando a supressão de não-máximos.

É descrita aqui também a cobertura das janelas de detecção, a qual é empregada para definir a quantidade de escalas necessárias de serem utilizadas de forma a realizar a detecção com determinado coeficiente de *Jaccard*. Até onde pesquisamos, é a primeira vez que a definição das escalas é feita baseado nas características geométricas e não apenas de forma empírica. Por fim, abordamos o conceito de região mínima de busca local. Uma região na qual é garantida presença de 100% da região contendo a placa veicular, sendo essa região onde se realiza a decomposição local em múltiplas escalas.

### 3.1 Histograma de Gradientes Orientados (HOG)

Nesta Seção, será apresentado o HOG, descritor empregado neste trabalho para a discriminação com base nas características locais de forma entre regiões de placas e não-placas. Esse descritor foi proposto por Dalal & Triggs (2005) para o problema de detecção de pedestres e posteriormente empregado em uma vasta gama de problemas de detecção (Zhang et al. 2008, Laborda et al. 2012).

Conforme citado em (Dalal & Triggs 2005), os descritores HOG são reminescentes dos descritores SIFT, histograma de bordas e contexto de forma. Tratam-se de informações locais da direção das bordas, as quais são calculadas a partir da imagem gradiente em células densamente arranjadas na imagem. Os gradientes podem ser computados de diversas maneiras, sendo o filtro *Sobel 1D* calculado a partir das máscaras  $[-1, 0, 1]$  e  $[-1, 0, 1]^T$ , para gradiente horizontal e vertical, respectivamente, o que apresentou melhores resultados (Dalal & Triggs 2005). A forma de cálculo dos gradientes horizontais e verticais é apresentados nas equações

$$dx = I(x + 1, y) - I(x - 1, y) \quad (3.1)$$

e

$$dy = I(x, y + 1) - I(x, y - 1), \quad (3.2)$$

respectivamente. Nessas Equações,  $I(x, y)$  corresponde a intensidade do *pixel* na posição  $x$  e  $y$  da imagem.

Quando se opta por trabalhar com imagens coloridas na representação RGB, os gradientes para cada um dos canais deve ser calculado separadamente e escolhe-se por aquele de maior magnitude para representar o gradiente do pixel em cada posição da imagem.

A partir desses gradientes, a orientação e o módulo são calculados pelas equações

$$m(x, y) = \sqrt{dx^2 + dy^2} \quad (3.3)$$

e

$$\theta(x, y) = \arctan\left(\frac{dy}{dx}\right), \quad (3.4)$$

respectivamente.

Na Figura 3.1, é apresentada uma representação esquemática da extração dos descritores HOG. Nesta, temos a imagem da placa veicular onde são representados dois blocos (regiões em amarelo e vermelho) formados por um agrupamento de quatro células. Utilizando os gradientes computados da imagem, para cada célula é construído um his-

tograma de orientações. A construção do histograma depende da determinação de uma faixa de ângulos, 0 a 180° ou 0 e 360°, além de um número de intervalos ou *bins*. Para cada *pixel* o voto é acumulado em determinado *bin* de acordo com a orientação do gradiente. Uma interpolação trilinear pode ser empregada a fim de propagar os votos em função da orientação e posição para os *bins* vizinhos. O peso do voto pode ser a própria magnitude do gradiente ou uma função dessa magnitude.

Diferentes configurações de célula podem ser empregadas de acordo com as características do histograma de orientações dos objetos que se deseja detectar. As configurações mais comuns são Retangular HOG (R-HOG) e o Circular HOG (C-HOG), sendo o primeiro, o de mais simples implementação e que apresentou melhores resultados (Dalal & Triggs 2005). As células presentes na Figura 3.1 são um exemplo de células do tipo R-HOG.

A finalidade de se unir as células em blocos é aplicar uma normalização local por blocos e realizar a sobreposição entre blocos vizinhos. Na Figura 3.1, pode-se perceber que os blocos compartilham uma mesma célula devido a sobreposição. Em (Dalal & Triggs 2005), a sobreposição de 50% entre blocos vizinhos é citada como significativa para um melhor desempenho. Além disso, a normalização torna o descritor mais robusto a mudanças de iluminação e presença de sombras na imagem. É importante ressaltar que devido a normalização por blocos, a informação de gradiente da mesma célula será diferente no descritor de cada um dos blocos em que ela contribui.

Diferentes formas de normalização podem ser empregadas, sendo avaliadas em (Dalal & Triggs 2005) quatro diferentes normalizações: *L1-norm*, *L1-sqrt*, *L2-norm* e *L2-hys*. Os resultados mostram que utilizando a *L2-norm*, conforme equação

$$v' = \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}, \tag{3.5}$$

é possível se alcançar resultados 5% superiores a *L1-norm*. Nessa,  $v$  é um vetor contendo todos os descritores HOG do bloco não-normalizados e  $\varepsilon$  é um constante muito pequena.

Os descritores HOG podem apresentar excelentes resultados na detecção de placas veiculares devido a sua capacidade de captarem as informações de forma presentes na imagem. Além disso, devido a forma como os histogramas são calculados, eles apresentam um certo grau de invariância a transformações geométricas e fotométricas.

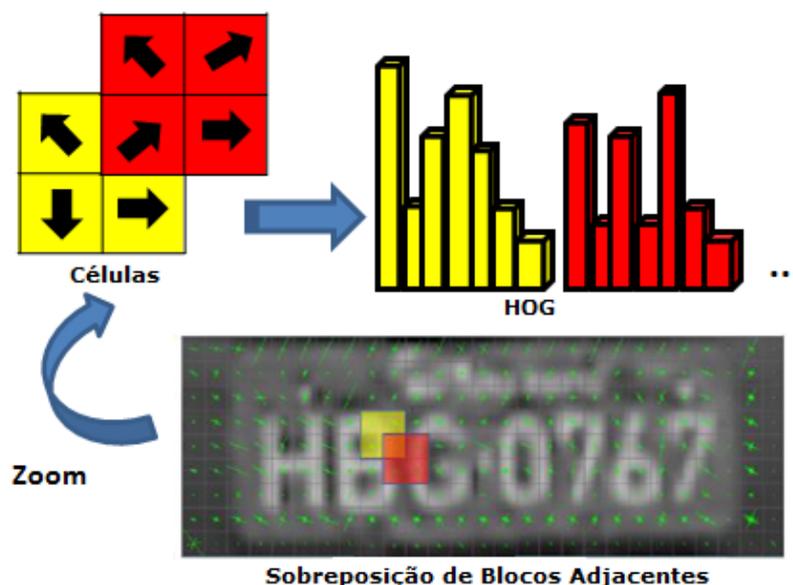


Figura 3.1: Representação esquemática dos descritores HOG.

## 3.2 Supressão de Não-Máximos

Devido à análise em múltiplas escalas utilizando janelas deslizantes, é comum que ocorram janelas vizinhas que cobrem a mesma região da placa veicular com diferentes sítis entre elas, regiões representadas na Figura 4.1(d). Essas regiões, quando fornecidas ao classificador, costumam ser todas classificadas como placa veicular. Assim, o classificador costuma detectar múltiplas janelas que correspondem ao mesmo objeto de interesse.

Múltiplas detecções de um mesmo objeto são desnecessárias. Desejamos saber dentre essas janelas, qual que apresenta um melhor casamento com o objeto a ser detectado, eliminando as demais. Uma forma de se tentar inferir essa informação é empregando a resposta do classificador. Regiões com maior resposta são assumidas como sendo as de maior casamento e as de menor resposta, o contrário.

A forma de se calcular o quanto essas regiões se sobrepõem também é importante no processo de supressão. Uma métrica bastante utilizada para calcular o casamento de duas regiões é o coeficiente de *Jaccard* ou *Jaccard index* (Dollár et al. 2012). O coeficiente de *Jaccard* de duas regiões  $A$  e  $B$  pode ser calculado segundo a equação:

$$J(A, B) = \frac{A \cap B}{A \cup B}. \quad (3.6)$$

O Algoritmo 3.1, é empregado para realizar a supressão das regiões com base nas resposta do classificador. Para isso, inicialmente, as janelas devem ser ordenadas em ordem decrescente utilizando os valores de respostas. Além disso, um valor de limiar deve ser definido, de forma que seja caracterizada a sobreposição entre duas regiões. Nas linhas 3-7 é realizado o processo de supressão, de forma que as regiões de maior resposta suprimem aquelas de menor resposta que possuem um valor de coeficiente de *Jaccard* superior a um limiar. Nas linhas 8-11, as janelas resultantes do processo de supressão são obtidas e posteriormente retornadas pelo método.

```

Entrada: Um conjunto  $J$  de Janelas de Detecção.
Saída: Um conjunto  $J'$  de Janelas de Detecção.
1 Inicialmente, as janelas de detecção presentes em  $J$  são ordenadas em ordem decrescente com base na resposta do classificador, seu atributo  $Suprimida$  recebe o valor  $false$  e deve ser estabelecido um valor para o limiar.;
2 cont  $\leftarrow 0$  ;
3 para cada  $j \leftarrow 1$  até  $Tamanho(J)$  faça
4   se  $J(j).Suprimida == false$  então
5     para cada  $i \leftarrow j + 1$  até  $Tamanho(J)$  faça
6       se  $JaccardIndex(J(j), J(i)) \geq limiar$  então
7          $J(i).Suprimida \leftarrow true$ ;
8 para cada  $j \leftarrow 1$  até  $Tamanho(J)$  faça
9   se  $J(j).Suprimida == false$  então
10     $J'(cont) = J(j)$ ;
11    cont ++ ;

```

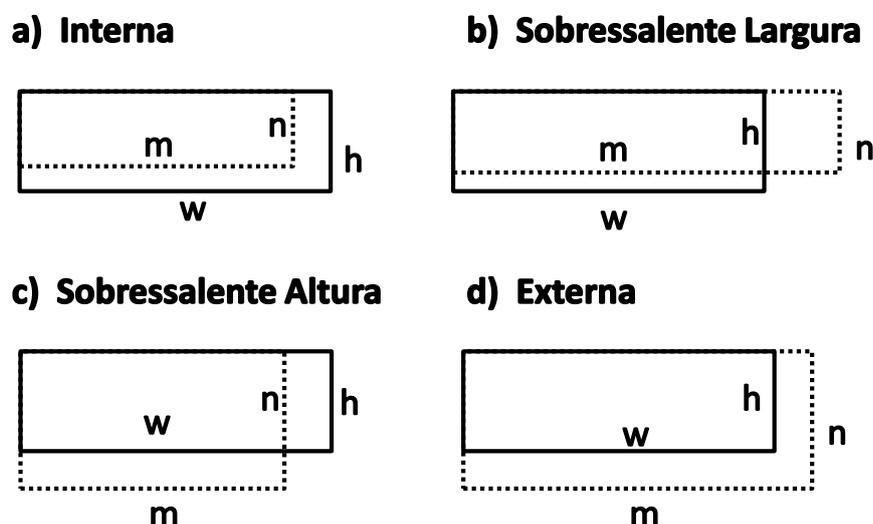
**Algoritmo 3.1:** Método de Supressão de Não-Máximos.

### 3.3 Cobertura das Janelas de Detecção

Na análise em múltiplas escalas, pode-se considerar que , em cada escala analisada, estamos tornando passíveis de detecção determinada faixa de tamanhos de placas veiculares. Assim, sabendo a dimensão das placas do conjunto de treino e utilizando um limiar de coeficiente de *Jaccard* é possível descobrir quantas e quais são as escalas que devemos

utilizar de forma que 100% dessas placas sejam passíveis de detecção. Essas regiões passíveis de detecção são ditas cobertas pela janela de detecção nas escalas empregas.

Na Figura 3.2, são ilustradas quatro possíveis posições relativas entre a janela de detecção (bordas contínuas) e a placa veicular rotulada (bordas pontilhadas). Essas posições são : Interna, Externa, Sobressalente Largura e Sobressalente Altura . Para cada uma dessas situações, diferentes equações são utilizadas para o cálculo do coeficiente de *Jaccard*.



**Figura 3.2:** Diferentes posições relativas para a janela de detecção(borda contínua) e região rotulada(borda pontilhada). A janela de detecção e região rotulada possuem dimensões em pixel de  $w \times h$  e  $m \times n$  pixels, respectivamente.

Pela equação

$$J(A,B) = \frac{m \times n}{w \times h}, \quad (3.7)$$

podemos calcular o coeficiente de *Jaccard* quando a região rotulada é interna à janela de detecção, representada na Figura 3.2(a). Quando se trata de um caso em que a região rotulada é maior que a janela de detecção em todas as dimensões, consideramos a posição relativa como externa, Figura 3.2(d), sendo o cálculo do coeficiente de *Jaccard* dado por

$$J(A,B) = \frac{w \times h}{m \times n}. \quad (3.8)$$

Nessas Equações, as dimensões da janela de detecção e região rotulada são, respectivamente,  $w \times h$  e  $m \times n$  pixels.

Para as demais situações, quando a região rotulada é maior que a placa de detecção em uma das dimensões apenas, largura ou altura, temos as posições relativas representadas pelas Figuras 3.2(b) e (c), respectivamente. O cálculo do coeficiente de *Jaccard* nesses casos é realizado utilizando

$$J(A,B) = \frac{w \times n}{w \times h + (m - w) \times n} \quad (3.9)$$

e

$$J(A,B) = \frac{m \times h}{w \times h + (n - h) \times m}, \quad (3.10)$$

respectivamente. Nessas equações, as dimensões da janela de detecção e região rotulada são, respectivamente,  $w \times h$  e  $m \times n$  pixels.

Utilizando os valores de coeficiente de *Jaccard* entre uma região rotulada e a janela de detecção é possível obter a cobertura da base obtida com determinado conjunto de escalas. Assim, a calibração dos parâmetros pode ser feito em função da cobertura necessária da base.

### 3.4 Região Mínima de Busca Local

O processo de busca local em múltiplas escalas depende da definição de uma região em torno da estimativa da posição inicial. Essa região será denominada de região mínima de busca local se atender a duas exigências: (1) deve englobar toda a área da placa rotulada que se deseja detectar e (2) deve ser a menor possível.

Com o intuito de se determinar esta área, foi inicialmente definido um limiar mínimo de coeficiente de *Jaccard* entre a janela de detecção e a placa veicular rotulada de 50%, de forma que ela seja caracterizada como uma estimativa inicial. A escolha desse limiar se deve a ser o mais utilizado em diversos outros trabalhos para caracterizar uma detecção (Dollár et al. 2012). Observe no entanto, que não consideramos essas regiões como sendo detecções finais, mas apenas como uma boa estimativa inicial. Para placa para ser detectada com todos os seus caracteres íntegros, existe a necessidade de coeficiente de *Jaccard* superiores, como será mostrado posteriormente.

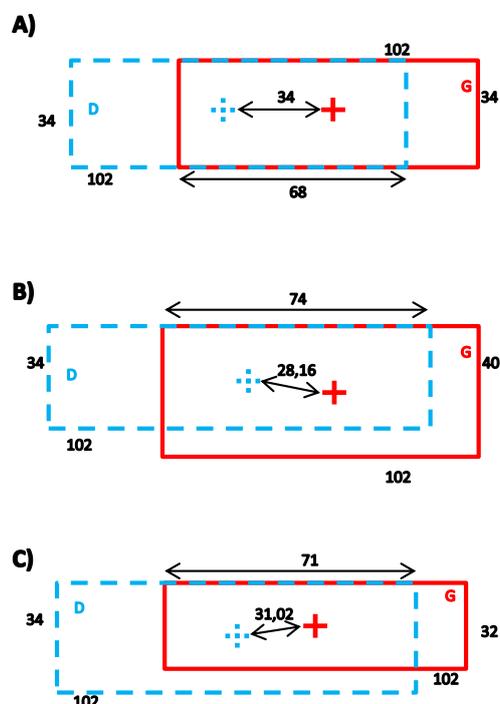
A partir dessa estimativa inicial, precisamos estabelecer uma quantidade de *pixels* na horizontal e vertical em torno dela a ser adicionado de forma que mesmo no pior caso, a região a ser detectada esteja 100% contida nela. Inicialmente, caracterizamos qual seria o pior caso. Na Figura 3.3, são representadas duas regiões, sendo a primeira região representada por retângulo de dimensões fixas de  $102 \times 34$  *pixels* e bordas pontilhadas em azul (janela de detecção), enquanto que a segunda tem diferentes dimensões e é representada por bordas contínuas e vermelhas (placa veicular rotulada). Nessa figura, temos ainda representada a largura da interseção e as distâncias entre os centros, representados por uma cruz, em *pixels* que resultam em um coeficiente de *Jaccard* de 50%.

Analisando as informações das Figuras 3.3 (a) a (c), podemos perceber que o pior caso, ou seja, aquele com maior deslocamento da janela de detecção com relação ao objeto real ocorre quando esses possuem a mesma dimensão. Nessa situação, pode-se mostrar que para que as regiões tenham um coeficiente de *Jaccard* de 50%, a distância entre os centros na horizontal deve ser de um terço da largura, no caso 34 *pixels*. Qualquer região rotulada maior 3.3 (b) ou menor 3.3 (c) que a janela de detecção precisa se aproximar do centro dessa de forma a compensar as variações nos valores de união e interseção entre as regiões, mantendo assim o coeficiente de *Jaccard* em 50%. Realizando uma análise semelhante, mas com as janelas ao longo de uma mesma linha vertical, pode-se mostrar que é a distância máxima entre os centros nesse caso é de um terço da altura.

Generalizando para uma janela de detecção de tamanho  $w \times h$ , podemos afirmar que a distância máxima com relação ao centro da região rotulada para um valor de coeficiente de *Jaccard* de 50% será de  $\frac{w}{3}$  e  $\frac{h}{3}$  *pixels* na horizontal e vertical, respectivamente. A Figura 3.4 mostra a janela de detecção de dimensão  $w \times h$  *pixels*, representada pela região de cor azul e bordas pontilhadas, juntamente com a região da placa rotulada em vermelho de dimensão  $w' \times h'$  *pixels* e a região mínima de detecção  $R$ , com bordas contínuas e cor azul. Nessa, podemos perceber que adicionando  $\frac{w}{3}$  e  $\frac{h}{3}$  *pixels* em cada lado ao longo das direções horizontal e vertical, respectivamente, é possível englobar 100% da região da placa veicular, mesmo no pior caso.

### 3.5 Detecção empregando janelas deslizantes

O processo de detecção por janelas deslizantes depende da definição do deslocamento horizontal e vertical da janela. Partindo do canto superior esquerdo a janela irá se mo-

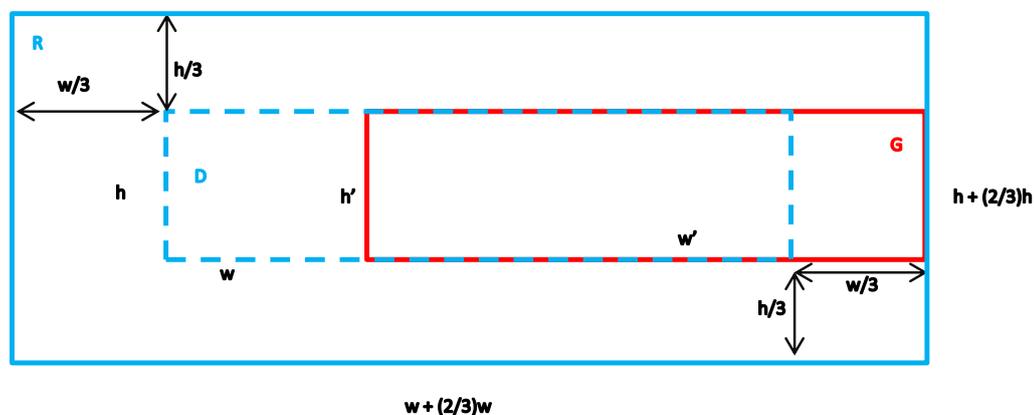


**Figura 3.3:** Exemplos de distância entre os centros da janela de detecção (região *D* em azul pontilhada) e a placa veicular rotulada (região *G* em vermelho contínua) para diferentes tamanhos de placa veicular.

ver utilizando o deslocamento horizontal até a última posição possível na linha atual, retornando ao começo, mas com um deslocamento vertical de forma a cobrir uma área diferente na imagem. Quanto menor o deslocamento, mais janelas são analisadas, aumentando a probabilidade de se detectar a placa veicular, no entanto, elevando-se o custo computacional.

Devido ao formato rígido das placas veiculares, a relação entre sua largura e altura, denominada razão de aspecto (*Aspect Ratio - AR*), possui um valor fixo e padronizado por lei. Assim, empregando o *AR* padrão e as características geométricas das placas da base de treino, é definido o tamanho da janela de detecção. A janela de detecção possui sempre um tamanho fixo, sendo a imagem reescalada de acordo com a escala analisada.

Um último passo na detecção é fornecer o conjunto de atributos extraídos para um classificador binário, definindo se a região em questão trata-se ou não de uma placa



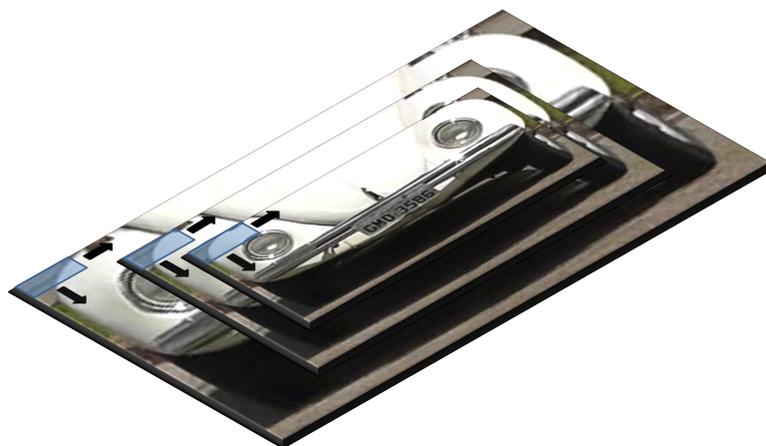
**Figura 3.4:** Ilustração da região mínima de busca local que deve ser utilizada para que seja possível englobar 100% da placa veicular quando ambas as regiões possuem a mesma dimensão. Nessa, as regiões retangulares  $D$ ,  $G$  e  $R$ , corresponde, respectivamente, a janela de detecção, região rotulada e a região mínima de busca local.

veicular. Dessa forma, obtemos um conjunto de janelas detectadas como candidatas a placas veiculares, as quais serão fornecidas como entrada para a etapa de supressão de não-máximos.

### 3.5.1 Método de Análise em Múltiplas Escalas utilizando Pirâmide de Imagens

Como as placas veiculares podem aparecer na imagem em diferentes distâncias com relação ao sistema de aquisição, a detecção utilizando informações locais deve ser feita em múltiplas escalas. O procedimento comum é a utilização de uma pirâmide de imagens, conforme a Figura 3.5. Para cada escala, essas janelas de varredura extraem um vetor de informações, denominados atributos da região, utilizando os descritores HOG. O tamanho desse descritor depende das configurações das células, blocos e número de bins utilizados.

O grande problema dessa abordagem está na quantidade de janelas analisadas, o que torna o custo computacional proibitivo para a utilização de descritores mais robusto como o HOG em aplicações de tempo-real. Além disso, como cada escala determina um tamanho de placa a ser detectado, caso a pirâmide seja formulada de forma esparsa, ou



**Figura 3.5:** Representação esquemática das janelas deslizantes em uma pirâmide de imagens.

seja, com maior intervalo entre escalas consecutivas, o casamento entre a placa veicular e a janela de detecção pode ser insuficiente para o reconhecimento do veículo, por cortar determinados caracteres. Trata-se de um método pouco eficiente, pois independente da placa ser maior ou menor, todas as escalas, inferiores e superiores, são analisadas. Essa técnica de análise de imagens em número fixo de escalas independente do tamanho da placa veicular é denominada neste trabalho de **estática**.

Para a detecção precisa do tamanho da placa veicular utilizando informações locais extraídas a partir de janelas deslizantes, é imprescindível a busca em múltiplas escalas. No entanto, o ideal é que seja possível reduzir a quantidade de janelas analisadas, empregando a análise em múltiplas escalas em regiões específicas da imagem, as quais possuem maior probabilidade de conter uma placa veicular. Além disso, a busca em diferentes escalas deve ser feita de forma **adaptativa**, descartando direções que se mostram improváveis de conter a placa veicular.

## 3.6 Considerações Finais

Neste capítulo, foram apresentados os principais conceitos relacionados ao método proposto. Foi inicialmente apresentado o descritor utilizado, o HOG. O método clássico de detecção de placas veiculares utilizando descritores HOG em uma análise em uma pirâmide de imagens também foi descrito. Para a definição da quantidade de escalas ana-

lisadas, definimos um conjunto de quatro posições relativas entre a janela de detecção e a região rotulada: interna, externa e sobressalentes em altura e largura. Para cada uma dessas configurações são apresentadas equações, as quais permitem definir qual faixa de tamanhos de janelas estamos cobrindo utilizando uma janela de detecção de dimensão  $w \times h$  *pixels*. Essa análise, até onde nosso estudo abrangeu, é inédita.

Ainda, definimos uma região mínima de busca local, região essa que utilizamos para a partir de uma placa detectada com coeficiente de *Jaccard* mínimo de 50%, realizar uma busca local em múltiplas escalas. Como mostramos, essa região deve ser formada pela adição de  $\frac{w}{3}$  e  $\frac{h}{3}$  *pixels* em cada lado ao longo das direções horizontal e vertical da região detectada. Foi dada uma motivação ao método adaptativo proposto, cujo principal objetivo é realizar uma detecção com elevado coeficiente de *Jaccard* analisando menos janelas que o método estático, o que é comprovado nos resultados.

# Capítulo 4

## Método Proposto

Neste capítulo, serão descritos os módulos principais que compõem nosso sistema. O método de busca local e adaptativa em múltiplas escalas proposto busca uma detecção mais precisa (maior coeficiente de *Jaccard*) da placa veicular analisando uma quantidade bastante inferior de janelas quando comparado aos métodos tradicionais. Esse método é baseado em três principais etapas: (1) estimativa inicial da posição analisando uma quantidade mínima de escalas, (2) estimativa inicial do tamanho em uma quantidade maior de escalas localmente e (3) refinamento da detecção em um processo adaptativo.

### 4.1 Introdução

O método proposto trabalha com imagens em tons de cinza, objetivando diminuir os ruídos provenientes da variação de iluminação. O processo de busca na imagem por regiões contendo placas veiculares utiliza de janelas deslizantes que varrem a imagem do canto superior esquerdo ao inferior direito. Inicialmente de uma forma grosseira, buscando apenas uma estimativa inicial da posição da placa veicular, para depois determinar uma estimativa para o tamanho da placa veicular em uma decomposição local em múltiplas escalas e, finalmente, um refinamento da localização de forma adaptativa.

Para cada região coberta pela janela de detecção na imagem, um vetor de atributos é extraído: os descritores HOG da região. O classificador SVM é então utilizado para determinar a pertinência daquela região como placa veicular. Nosso método é denominado **adaptativo** devido a busca por regiões contendo placas veiculares ser guiada pelas respostas do classificador SVM. Assim, dependendo do tamanho em que a placa veicu-

lar aparece na imagem, a quantidade de janelas analisadas é diferente. Dessa forma, buscamos garantir uma detecção mais precisa e eficiente.

A Figura 4.1 é empregada para descrever os principais módulos que compõem o método proposto. A imagem fornecida como entrada (Figura 4.1(a)) passa por uma etapa inicial de pré-processamento na qual é convertida para tons de cinza (Figura 4.1(b)). Os gradientes da imagem são então calculados. Uma janela deslizante percorre toda a imagem em uma quantidade reduzida de escalas classificando cada região em placa ou não-placa, conforme representado na Figura 4.1(c). Nessa etapa são empregados um classificador SVM previamente treinado e os descritores HOG computados com base nas informações do gradiente.

Na Figura 4.1(d), é mostrado que podem existir múltiplas detecções para um mesmo objeto, as quais são evitadas utilizando uma supressão de não-máximos, conforme Figura 4.1(e). Para cada uma das regiões resultantes da supressão de não-máximos, que são as estimativas iniciais para a posição da placa veicular, é aplicada uma decomposição local em múltiplas escalas, representada na Figura 4.1(f). Nesta etapa esperamos obter uma boa estimativa inicial para as dimensões da placa veicular, a qual é refinada em um processo de busca local, representado na Figura 4.1(g). O resultado desta etapa é uma ou mais regiões retangulares que se espera conter uma placa veicular, conforme Figura 4.1(h).

## 4.2 Método de Busca Local e Adaptativa em Múltiplas Escalas

O método proposto neste trabalho busca tornar o processo de busca utilizando janelas deslizantes em múltiplas escalas mais eficiente, de forma que seja possível empregar descritores mais robustos em todo o processo de detecção. Para isso, o processo de busca é integrado ao processo de classificação, utilizando as informações de pertinência da janela de detecção fornecidas pelo classificador. Essa resposta será responsável por guiar o processo de detecção, tornando-o adaptativo. Na continuação, iremos descrever as três principais etapas do método proposto: (1) estimativa do tamanho, (2) estimativa da posição e (3) refinamento.

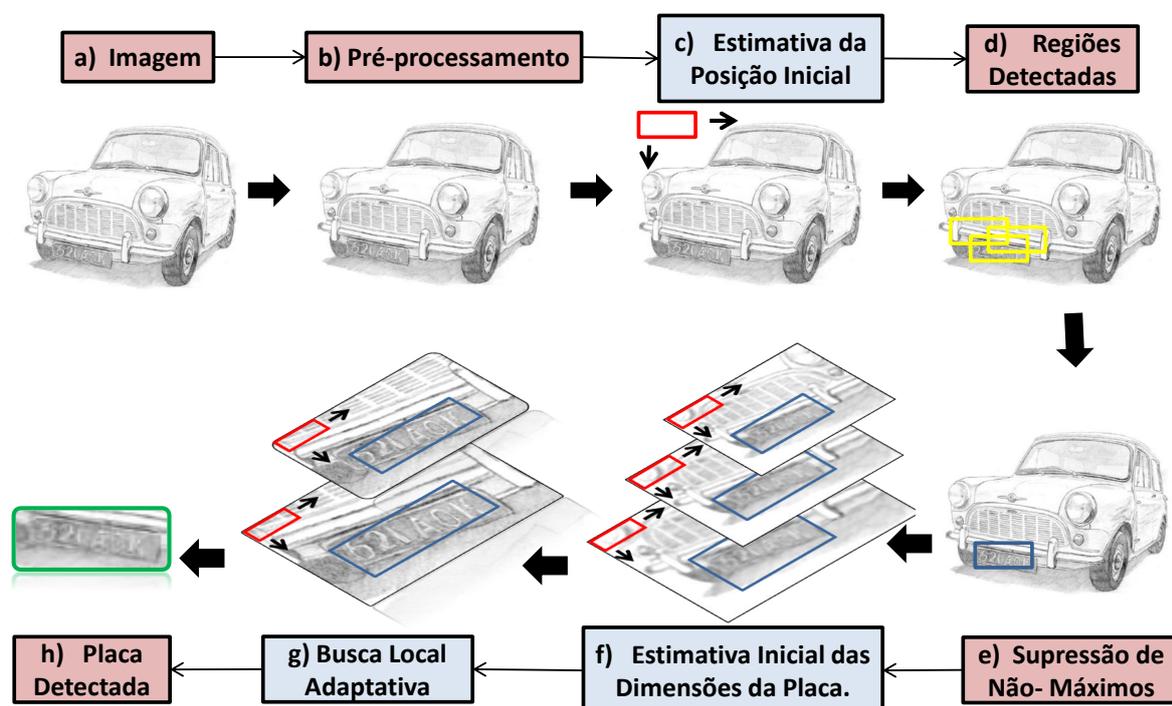


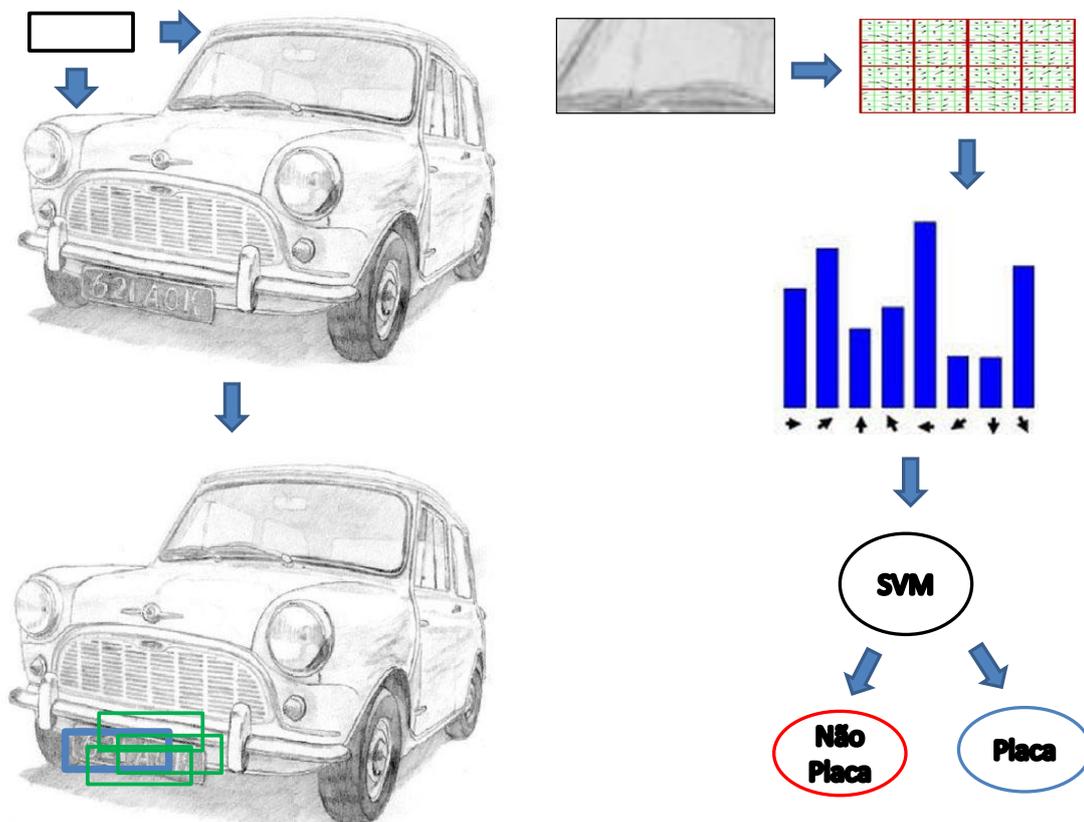
Figura 4.1: Visão geral do método proposto. As regiões retangulares destacadas em azul são as nossas principais contribuições.

### 4.2.1 Estimativa da Posição Inicial

Inicialmente, as imagens são analisadas utilizando as janelas de detecção em uma quantidade reduzida (idealmente, uma única escala) de escalas e com um espaçamento maior entre janelas vizinhas, conforme Figura 4.2. Para cada região, são calculados os descritores HOG em células arranjadas densamente, constituindo o vetor de atributos da região. A resposta do classificador SVM é calculada para todas as regiões cobertas pela janela de detecção utilizando o seu respectivo vetor de atributos. Regiões acima de determinado limiar de resposta do classificador são consideradas como uma boa estimativa inicial para a posição da placa veicular, essas regiões aparecem como janelas de detecção em cores verde e azul na Figura 4.2.

Caso algumas regiões detectadas apresentem um casamento grande entre elas, representado por um alto valor de coeficiente de *Jaccard*, essas passam por uma fase de supressão de não-máximos. Assim, regiões de resposta menor são suprimidas por aquelas de resposta maior. Todas as regiões detectadas nessa etapa e não-suprimidas, passam para uma próxima fase em que é realizada a busca adaptativa em uma quantidade maior de escalas. Um exemplo de supressão pode ser visto na Figura 4.2, onde as regiões

detectadas em verde, após um processo de supressão são eliminadas, restando apenas a detecção em azul.

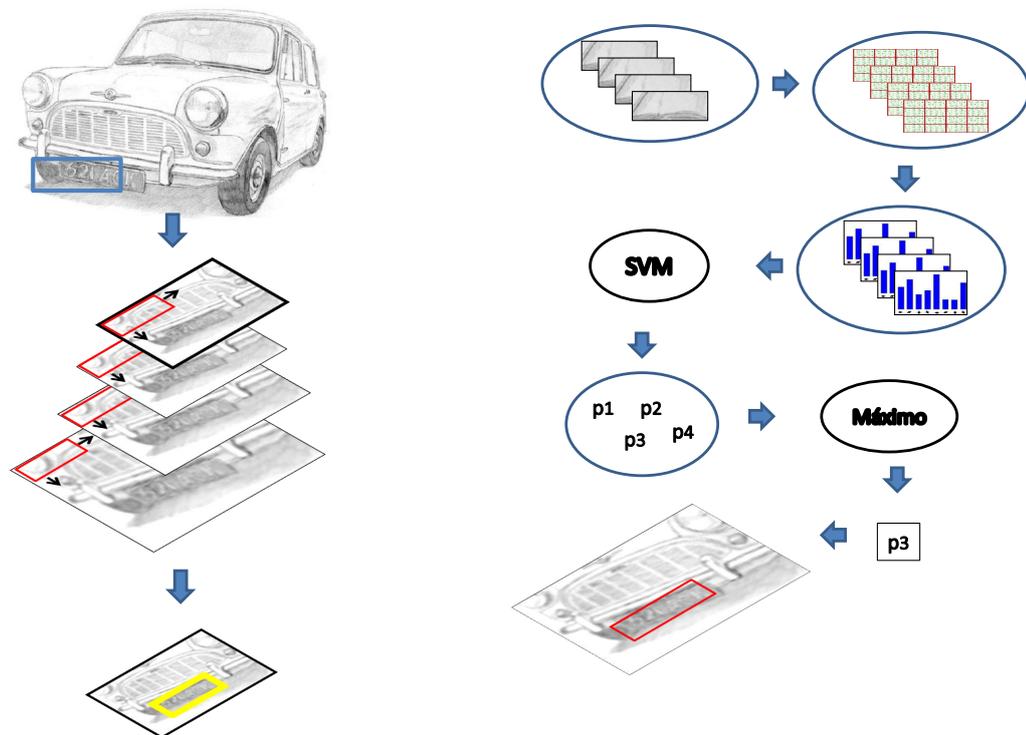


**Figura 4.2:** Estimativa da posição inicial da placa veicular usando HOG e SVM.

### 4.2.2 Estimativa do Tamanho

Como apenas a informação de posição não é suficiente para o reconhecimento, buscamos também obter o tamanho da placa veicular. Para isso, localmente é criada uma pirâmide de imagens em diferentes escalas, representada na Figura 4.3. Nesse caso, o processo é ilustrado para quatro escalas, mas essa quantidade pode ser maior ou menor, dependendo das características da base.

Nessa etapa, procuramos uma boa estimativa inicial para o tamanho da placa veicular em cada região detectada na etapa anterior. O processo de detecção com janelas deslizantes é aplicada em cada uma dessas escalas buscando encontrar a região de máxima



**Figura 4.3:** Processo de estimativa local e em múltiplas escalas do tamanho da placa veicular.

resposta do classificador SVM, representada pela janela em amarelo na Figura 4.3. Tanto a região quanto a escala em que essa região ocorre são utilizadas pelas etapas seguintes. É importante ressaltar que apesar de serem ilustradas apenas quatro escalas, essa quantidade pode ser maior ou menor, dependendo das características da base de imagens.

Na Figura 4.3, é demonstrado ainda o processo de determinação da região de máxima resposta do classificador SVM. Para cada região coberta pela janela de detecção, são extraídos os descritores HOG, constituindo o vetor de atributos da região. Esses vetores são fornecidos ao classificador, o qual retorna a pertinência de cada região. Utilizando esse conjunto de valores de pertinência, é possível calcular o máximo, que corresponde a uma região na imagem que pode ser diretamente relacionada a uma escala. Esse conjunto região e escala é a estimativa de tamanho da placa veicular.

A região com bordas em preto mais grossas, na Figura 4.3, corresponde a escala retornada como a melhor estimativa inicial para o tamanho da placa veicular. A partir dessa estimativa, um processo de refinamento adaptativo é realizado para a detecção precisa em termos de posição e dimensão da placa. Esse processo será descrito na

próxima Seção.

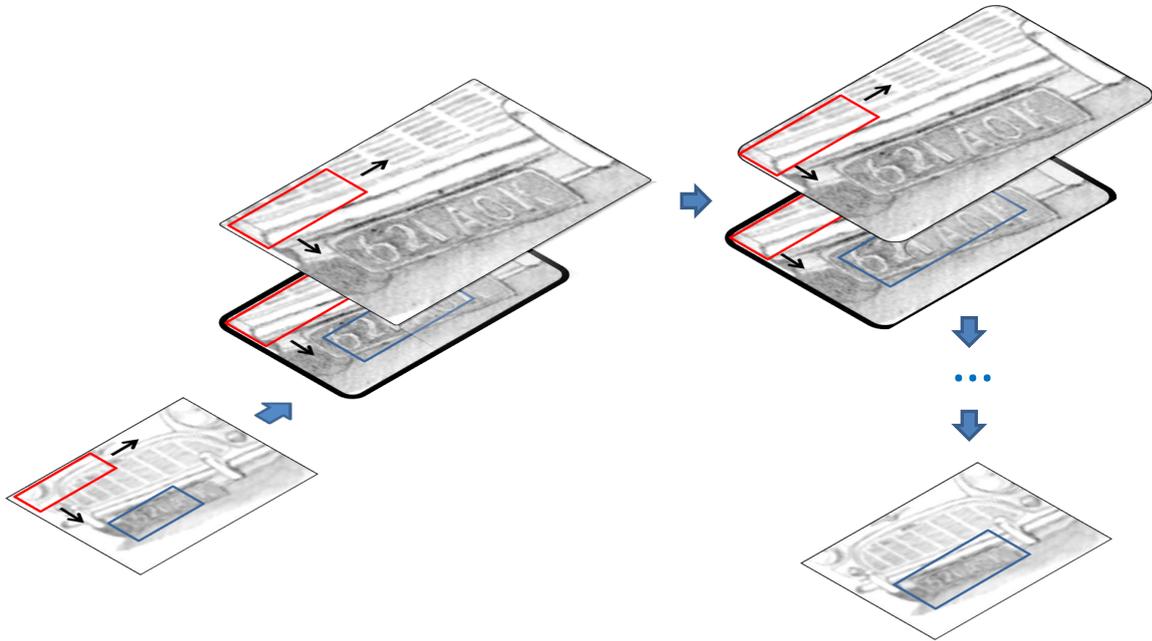
### 4.2.3 Refinamento da Localização

Como o processo de estimação do tamanho parte de uma quantidade fixa de escalas na qual a imagem é decomposta localmente, essa ainda não pode ser dita como uma boa detecção da placa veicular. Escalas fixas limitam o tamanho das placas a serem detectadas e desejamos que nosso método detecte placas independente da dimensão. Assim, é necessário um refinamento da posição da placa veicular utilizando a melhora na resposta do SVM para guiar o processo em diferentes escalas. Dessa forma, objetivamos obter uma detecção precisa da placa veicular, ou seja, com altos valores para coeficiente de *Jaccard*.

Na Figura 4.4, temos a imagem retornada como a região local que obteve a região de máxima resposta para o classificador SVM na etapa prévia de estimativa do tamanho da placa veicular. Partindo dessa imagem são obtidas duas imagens, uma para escala inferior e outra para a escala superior. Cada uma dessas escalas é analisada e a maior resposta obtida é comparada com a detecção obtida até então utilizando o valor de pertinência retornado pelo classificador SVM. O processo é repetido enquanto houverem melhoras significativas na detecção. Uma melhora é significativa quando a diferença entre a nova solução e a antiga é maior que um limiar. Assim, nas regiões próximas as estimativas iniciais são obtidas novas detecções, as quais são consideradas as detecções finais do método.

### 4.2.4 Algoritmo

O Algoritmo 4.1 é empregado para descrever a implementação do método de busca local e adaptativa em múltiplas escalas, representado na Figura 4.5 pelas etapas (e) e (f). O método recebe como entrada as estimativas iniciais para a posição da placa veicular e retorna como resposta um conjunto de mesmo tamanho contendo as respostas refinadas. Um conjunto de escalas a serem analisadas inicialmente ( $S$ ), bem como um fator de reescalamiento da imagem ( $f$ ) e um região de busca ( $r$ ) devem ser previamente definidos com base nos dados das placas do conjunto de validação. Para a definição de  $S$  e  $f$  é utilizada a análise de cobertura da base de treinamento, conforme descrito na Seção 3.3 e para a definição de  $r$  é empregado o conceito de região mínima de busca local, definido na Seção 3.4.



**Figura 4.4:** Refinamento da posição de forma adaptativa.

Nas linhas 6-9, é realizado o processo de estimativa inicial do tamanho da placa veicular com base na análise de todas as escalas  $s$  pertencentes ao conjunto  $S$ . A escolha da melhor escala é realizada com base na melhor resposta entre todas as janelas analisadas dentro da região  $r$ , valor fornecido pela função *RetornaMelhor*.

Após definida a melhor estimativa inicial, o processo de busca adaptativa é iniciado, compreendendo as linhas 11-25 do Algoritmo 4.1. Nessa etapa, é utilizada a estimativa inicial fornecida pela etapa prévia, e com base no fator de reescalamo  $f$ , é calculada a escala superior e inferior da região local da imagem, definida pela região  $r$ . Em cada uma dessas regiões reescaladas, as respostas de todas as janelas são calculadas e a melhor resposta é utilizada com indicativo do melhor caminho para se seguir. O processo continua enquanto forem sendo encontradas melhorias significativas. Uma melhoria é considerada significativa quando é superior a um limiar *diff*.

Ao final, é utilizado o método *Melhor* de forma a retornar a janela de detecção que apresentou a melhor resposta de pertinência. Essa é considerada a detecção do método proposto, a qual se espera que tenha coeficiente de *Jaccard* superior quanto comparado ao método em multiescalas e o proposto em (Mendes et al. 2011), o que é comprovado pelos experimentos.

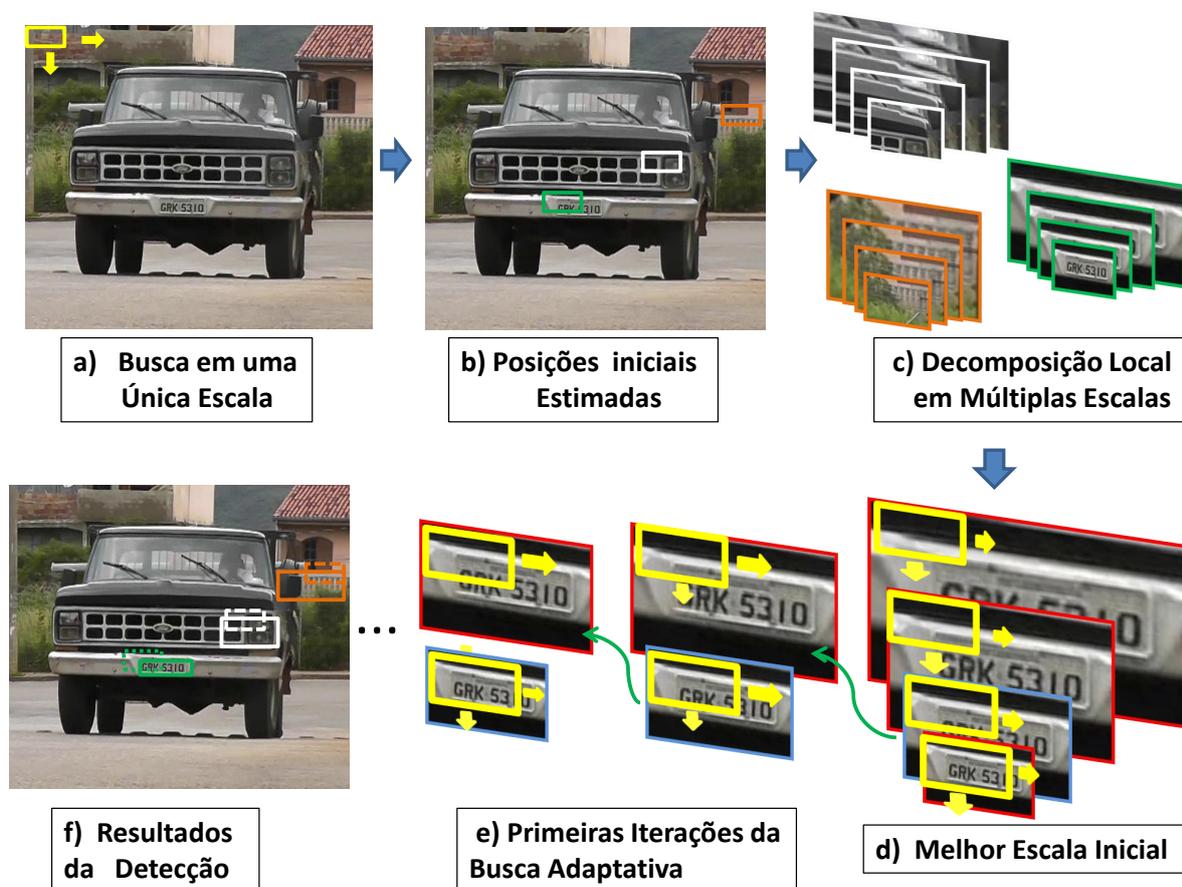


Figura 4.5: Representação esquemática do processo de busca local e adaptativa em múltiplas escalas

#### 4.2.5 Considerações Finais

Neste capítulo, foi apresentado o método **adaptativo** proposto neste trabalho. A Figura 4.5 é empregada para ilustrar as principais etapas desse método. Trata-se de um método baseado em três etapas: (1) obtenção de uma estimativa inicial para a posição da placa veicular, (2) determinação de uma estimativa inicial para o tamanho da placa veicular e (3) refinamento da localização.

A primeira etapa, Figura 4.5(a), é conduzida em uma quantidade mínima de escalas e usando a imagem inteira, enquanto que a segunda, Figura 4.5(c), necessita de um conjunto maior de escalas e é conduzido em uma região local em torno da estimativa inicial para a posição. Na terceira etapa, Figura 4.5(e), é aplicado um processo adaptativo de refinamento da localização da placa veicular com base na resposta do classificador SVM. Para a definição das escalas é empregado o conceito de cobertura da janela de detecção,

**Entrada:** Um conjunto  $J$  de Janelas de Detecção Superiores a um Limiar do Classificador

**Saída:** Um conjunto  $J'$  de Janelas Detecção

```

1 bestscale  $\leftarrow -1$  ;
2 best  $\leftarrow -inf$  ;
3 cont  $\leftarrow 0$  ;
4 Inicializando as Variaveis S, f e r, respectivamente, conjunto de escalas,
   fator de reescalamento do método Adaptativo e a região da busca local;
5 para cada elemento  $j \in J$  faça
6   para cada elemento  $s \in S$  faça
7     se RetornaMelhor( $s, j$ )  $\geq$  best então
8       best  $\leftarrow$  RetornaMelhor( $s, j, r$ ) ;
9       bestscale  $\leftarrow s$  ;
10    melhora  $\leftarrow$  true ;
11    enquanto melhora faça
12      melhora  $\leftarrow$  false ;
13      up  $\leftarrow$  bestscale  $\times$  f ;
14      down  $\leftarrow$  bestscale  $\div$  f ;
15      upbest  $\leftarrow$  RetornaMelhor( $up, j, r$ ) ;
16      downbest  $\leftarrow$  RetornaMelhor( $down, j, r$ ) ;
17      se upbest  $>$  downbest e upbest - best  $>$  diff então
18        best  $\leftarrow$  upbest;
19        bestscale  $\leftarrow$  up;
20        melhora  $\leftarrow$  true ;
21      senão
22        se downbest - best  $>$  diff então
23          best  $\leftarrow$  downbest;
24          bestscale  $\leftarrow$  down;
25          melhora  $\leftarrow$  true ;
26     $j' \leftarrow$  Melhor( $j, bestscale$ ) ;
27     $J'(cont) \leftarrow j'$  ;
28    cont ++ ;

```

**Algoritmo 4.1:** Método de Busca Local e Adaptativa

enquanto que para a definição da região na qual é aplicado os procedimentos referentes a segunda e terceira etapa é utilizado o conceito de região mínima de busca local.



# Capítulo 5

## Resultados e Discussões

Neste capítulo, apresentaremos inicialmente as bases de imagens utilizada nos experimentos. Em seguida, a calibração dos parâmetros do método proposto será descrita com base na avaliação das características do conjunto de treino e com testes preliminares realizados na base de validação. Finalmente, os resultados obtidos com o método proposto serão comparados com um algoritmo presente na literatura.

### 5.1 Base de Imagens

Na continuação, apresentaremos as bases de imagens de veículos utilizada neste trabalho. A opção de utilizar duas bases, uma brasileira e uma grega, foi realizada com o objetivo de se avaliar a robustez do método proposto. Tratam-se de base completamente diferentes, sendo uma obtida de carros em movimento, com pouca variação de perspectiva, enquanto a outra é obtida de carros estacionados, com uma grande variação de perspectiva. Ambas as bases encontram-se disponíveis *online* em (Mendes-Júnior et al. 2011).

A base de imagens brasileira é composta por um conjunto de 377 imagens, com resolução de  $800 \times 600$  *pixels*. Foi realizada uma divisão da base em conjunto de treinamento, validação e teste, sendo o primeiro constituído por uma amostra de 221 imagens e os demais por 78 cada. A base grega é composta por 345 imagens, sendo repartidas em conjuntos de 204 para treino, 70 para teste e 71 para validação.

A Figura 5.1 é utilizada para mostrar alguns exemplos da base de dados brasileira.

As imagens foram obtidas durante o dia, sendo em sua maioria bem iluminadas. Na primeira linha da Figura 5.1, podemos perceber que algumas imagens foram capturadas no final da tarde, sendo por isso mais escuras.

Há alguns casos em que, devido a posição do sol e a existência de árvores no entorno, foram formadas sombras, como pode ser notado na imagem inferior direita da Figura 5.1. Outra característica da base é que as imagens são tiradas com moderada variação da distância com relação ao sistema de aquisição, a maior e a menor imagem do conjunto de treino têm, respectivamente, 6400 e 2100 *pixels*<sup>2</sup>. A variação de perspectiva é mínima, uma vez que as imagens são obtidas com a dianteira ou traseira do veículo em posição frontal com relação ao sistema de aquisição.

Apesar das características já mencionadas, a base de imagens brasileira utilizada não pode ser considerada fácil, pois, como podemos observar na Figura 5.1, além dos veículos, existem vegetações, cercas de arame, residências e intensa movimentação de pedestres.

Na Figura 5.2, podemos perceber que a base grega também impõe desafios, porém relacionados a diferentes fatores. Nessa base, as imagens variam bastante de iluminação por serem obtidas tanto em ambientes fechados (imagens da primeira linha), quanto em ambientes abertos (imagens da segunda linha). Além disso, temos uma variação bem maior da perspectiva e da distância da placa veicular com relação ao sistema de aquisição. No entanto, ela apresenta menos ruídos de fundo quando comparada a base brasileira, sendo mais propícia para métodos suscetíveis a ruídos.

## 5.2 Definição das Métricas de Avaliação e Calibração de Parâmetros

Na continuação desta Seção, serão inicialmente descritas as métricas de avaliação empregadas: Revocação e Precisão. As curvas de que relacionam a revocação e a precisão com a quantidade de Falsos Positivos Por Imagem (FPPI), respectivamente, Revocação  $\times$  FPPI e Precisão  $\times$  FPPI serão descritas.

Posteriormente, serão definidos os protocolos de ajuste dos parâmetros empregados nos métodos de detecção de placas veiculares implementados: método **estático** e o método **adaptativo** proposto. Ambos compartilham de uma mesma configuração do



Figura 5.1: Exemplos de imagens da base brasileira.

descritor HOG e de um mesmo tamanho de janela de detecção. A principal diferença está no deslocamento da janela de detecção e na quantidade de escalas analisadas. A parametrização da busca local e adaptativa em múltiplas escalas é exclusiva para o método **adaptativo**. Para o método descrito em (Mendes et al. 2011), foi realizada apenas um refinamento de alguns parâmetros, uma vez que o código se encontra disponível *online* (Mendes-Júnior et al. 2011) e a base empregada em nossos experimentos é uma das utilizadas pelos autores.

Todos os experimentos que serão apresentados foram executados em um computador *Pentium Core i7* com 2.10 GHz, memória RAM de 8GB. Para implementação foi utilizada a biblioteca OpenCV 2.4.6 (Bradski 2000), utilizando a linguagem C++. Essa implementação se encontra disponível *online* em (Raphael-Prates 2013).



Figura 5.2: Exemplos de imagens da base grega.

### 5.2.1 Definição das Métricas de Avaliação

Neste trabalho, será utilizado um protocolo de avaliação similar ao desenvolvido em (Wang et al. 2013). O coeficiente de *Jaccard* entre duas regiões  $A$  e  $B$ , representado por  $J(A, B)$ , e descrito na Equação 3.6 será empregado para definir o melhor casamento entre uma região retangular  $r$  com relação ao conjunto de retângulos  $R$ , segundo a Equação 5.1.

$$m(r; R) = \max \{J(r, r_0) | r_0 \in R\}. \quad (5.1)$$

A revocação e a precisão são definidas em função do melhor casamento pelas Equações 5.2 e 5.3, respectivamente. Onde  $T$  e  $E$  representam os conjuntos de regiões rotuladas e regiões detectadas, respectivamente. O termo  $m_t(a, B)$  é igual a 1 quando o valor de

melhor casamento dado pela Equação 5.1 é superior ao limiar  $t$  e 0, caso contrário.

$$\text{Revocação} = \left( \sum_{r \in T} m_t(r, E) \right) / |T| \quad (5.2)$$

$$\text{Precisão} = \left( \sum_{r \in E} m_t(r, E) \right) / |E| \quad (5.3)$$

A revocação, descrita pela Equação 5.2, é uma proporção entre a quantidade de placas veiculares capturadas com coeficiente de *Jaccard* superior a um limiar estabelecido e o total de placas presentes na base de imagens. A precisão, descrita na Equação 5.3, por outro lado, representa a proporção entre a quantidade de placas detectadas com um limiar superior ao limiar estabelecido e a quantidade total de detecções retornadas pelo método.

No entanto, a análise isolada da revocação não é interessante por não informar a quantidade de falsos positivos obtidos, ou seja, regiões que foram classificadas como placas veiculares, mas na prática não o são. Assim, a análise da revocação será feita em conjunto com a quantidade de FPPI em curvas do tipo semilogarítmicas. Quanto maior revocação e menor a quantidade de FPPI, melhor será a acurácia do classificador. Essa curva é construída ordenando em ordem decrescente todas as janelas de detecção com base nos valores de resposta do classificador SVM e aplicando 11 diferentes limiares que variam no intervalo de  $[-1,0;1,0]$  com intervalos de 0,2. Para cada limiar é calculada a revocação utilizando a Equação 5.2 para um valor fixo de coeficiente de *Jaccard*, por exemplo 50%.

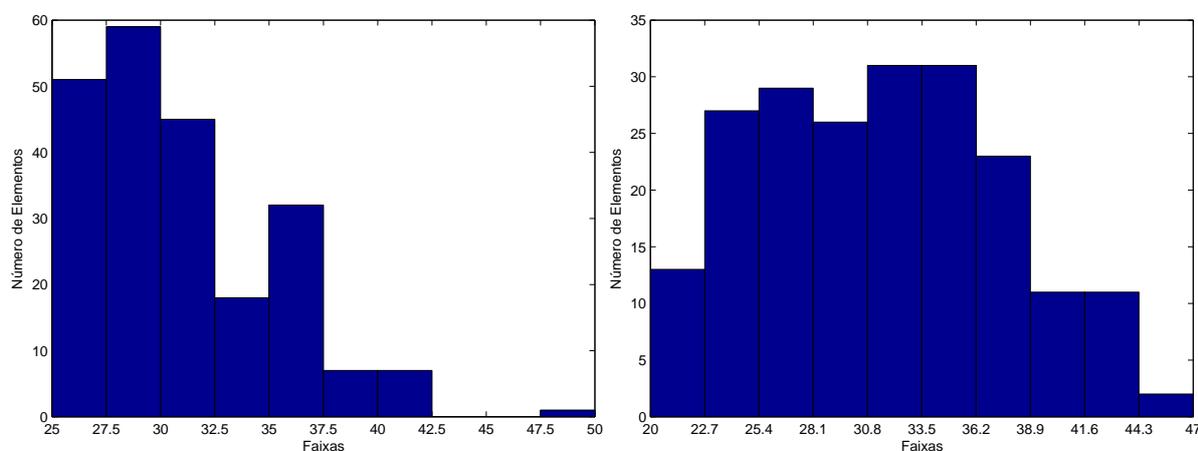
Um dos parâmetros que deve ser analisado com o intuito de determinar a qualidade da detecção é o valor do coeficiente de *Jaccard*. Na Figura 5.3, é apresentado em cada linha um exemplo de quatro detecções dentro de uma mesma faixa de coeficiente de *Jaccard*, sendo a região detectada representada pelo retângulo de borda azul, enquanto que a região da placa veicular rotulada, pelo retângulo de borda vermelha. Assim, na primeira linha, são apresentados os resultados cujo coeficiente de *Jaccard* se encontram no intervalo de  $[0,5;0,6[$ , para as demais linhas temos, respectivamente, os intervalos  $[0,6;0,7[$ ,  $[0,7;0,8[$  e  $[0,8;0,9[$ .

Pela Figura 5.3, pode-se concluir que com valores de coeficiente de *Jaccard* inferiores a 70%, alguns caracteres podem aparecer cortados, como na primeira e segunda linhas, impedindo o reconhecimento. Além disso, as detecções que podem consideradas ideais, com pouca informação de fundo e com 100% de confiabilidade de que os caracteres são passíveis de reconhecimento, ocorrem apenas com *Jaccard* superior a 80%, representadas na quarta linha. Por isso, a avaliação dos métodos será realizada considerando tanto o coeficiente de *Jaccard* usualmente empregado de 50% quanto o de 80%.



**Figura 5.3:** Representação de detecções com diferentes valores de Jaccard. A região em vermelho representa a placa veicular rotulada, enquanto que as regiões em azul são as detecções. A primeira linha mostra as regiões no intervalo de  $[0, 5; 0, 6[$ , a segunda  $[0, 6; 0, 7[$ , a terceira  $[0, 7; 0, 8[$  e a última  $[0, 8; 0, 9[$ .

Com o intuito de analisar a qualidade da detecção obtida, será feita uma análise da precisão e revocação para diferentes valores de coeficiente de *Jaccard*, contruindo curvas de Precisão  $\times$  *Jaccard Index* e Revocação  $\times$  *Jaccard Index*. Nessas curvas, pretende-se analisar se os valores de precisão e revocação se mantêm elevados à medida que se



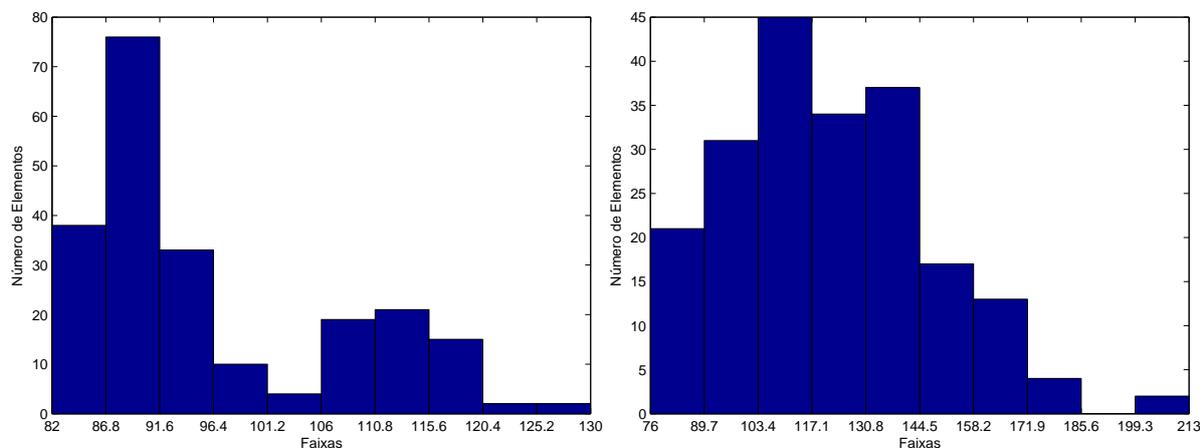
**Figura 5.4:** Histograma de altura das placas do conjunto de treinamento. À esquerda, temos o histograma de alturas para a base brasileira e, à direita, para a base grega.

aumenta as exigências de casamento entre a região detectada e a placa veicular rotulada. Essas curvas são contruídas para um limiar fixo do classificador SVM, o qual deve ser definido previamente. Assim, para todas as detecções não-suprimidas e cuja resposta se encontra acima desse limiar é calculado o valor do coeficiente de *Jaccard*. Essas regiões são então ordenadas em ordem crescente de coeficiente de *Jaccard* e um conjunto de 10 limiares igualmente espaçados no intervalo (0;1) são utilizados para se determinar a precisão (Equação 5.3) ou a revocação (Equação 5.2).

## 5.2.2 Definição das Dimensões da Janela de Detecção

A calibração dos parâmetros do descritor HOG foi realizada com base nas informações do conjunto de treino das placas. Assim, inicialmente buscamos caracterizar a base de treino quanto as dimensões das placas presentes, as quais estão diretamente associadas ao tamanho da janela de detecção. A Figura 5.4 é empregada para caracterizar a altura das placas presentes na base de treinamento, sendo o histograma à esquerda para a base brasileira e à direita para a base grega. Para cada faixa de altura em *pixels*, é calculada a frequência das placas que se enquadram naquela região, construindo assim um histograma de altura das placas. Os valores de média e desvio padrão para a altura são  $30,8 \pm 4,15$  *pixels*, para a base brasileira e  $31,5 \pm 6,04$  *pixels* para a grega.

Um histograma de largura das placas veiculares também foi construído, como pode ser visto na Figura 5.5, sendo representado á esquerda os dados para a base brasileira e



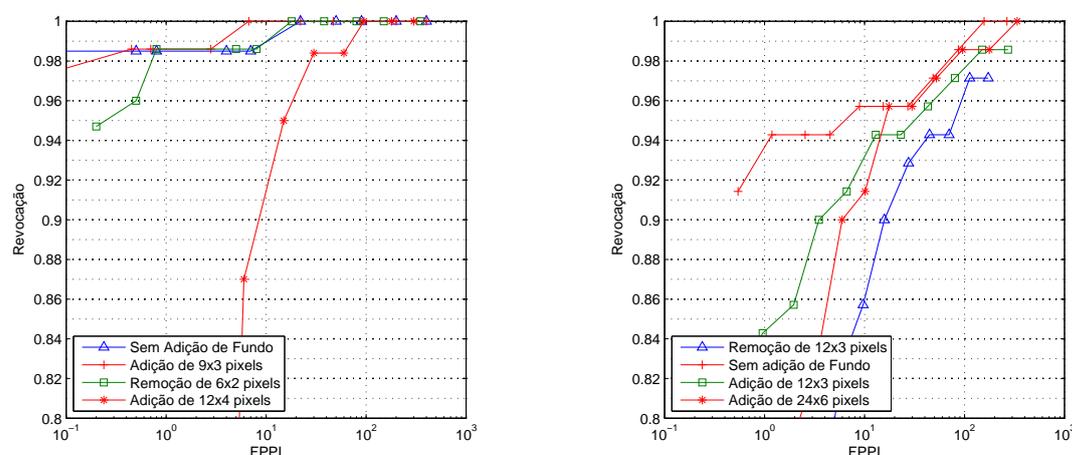
**Figura 5.5:** Histograma de largura das placas do conjunto de treinamento. À esquerda, temos o histograma de larguras para a base brasileira e, à direita, para a base grega.

à direita para a base grega. Os valores média e desvio padrão foram nesse caso de  $96,2 \pm 11,51$  e  $121,7 \pm 25,26$  para a base grega e brasileira, respectivamente. O *aspect ratio* para a base brasileira é de  $3,1 \pm 0,16$ , valor bem próximo ao especificado na legislação brasileira que é de aproximadamente 3. Para a base grega, o *aspect ratio* obtido foi de  $3,9 \pm 0,62$ .

Baseado nos dados geométricos obtidos, definimos o tamanho inicial da janela de detecção em  $102 \times 34$  *pixels*, valores próximos aos valores médios de largura e altura, e que ainda mantém o *aspect ratio* padrão para a base brasileira. Essa é a dimensão do objeto que desejamos detectar, sendo utilizada para o cálculo do coeficiente de *Jaccard* entre a região de detecção e a região da placa (região rotulada). Uma análise análoga foi realizada para a base grega, chegando aos valores de  $120 \times 30$  *pixels*.

Em (Dalal & Triggs 2005), os autores mostram que a adição de fundo à janela de detecção é responsável por um aumento na acurácia da detecção. Assim, um estudo da quantidade de *pixels* a serem adicionados à janela de detecção de  $102 \times 34$  e  $120 \times 30$  *pixels* se faz necessário. Para isso, são definidas configurações iniciais para o descritor HOG, as quais permaneceram fixas, de forma que apenas o impacto da adição de fundo possa ser estudado. Definimos de forma empírica uma configuração inicial de  $6 \times 6$  *pixels* para as células agrupadas em blocos de  $2 \times 2$  células com uma análise **estática** de 11 escalas.

Na Figura 5.6, é apresentada a Revocação  $\times$  FPPI para diferentes configurações de



**Figura 5.6:** Análise da influência de diferentes configurações de janelas de detecção no conjunto de validação. À esquerda, temos os resultados para a base brasileira e, à direita, para base grega.

fundo da imagem utilizando o conjunto de validação da base brasileira (esquerda) e grega (direita). Por essa figura, podemos perceber que a remoção de *pixels* em ambos os casos apresenta um impacto negativo na revocação do método. Para a base brasileira, a adição de fundo representou uma melhora na taxa de detecção, enquanto que para a base grega, manter a janela com mesma dimensão é a melhor escolha.

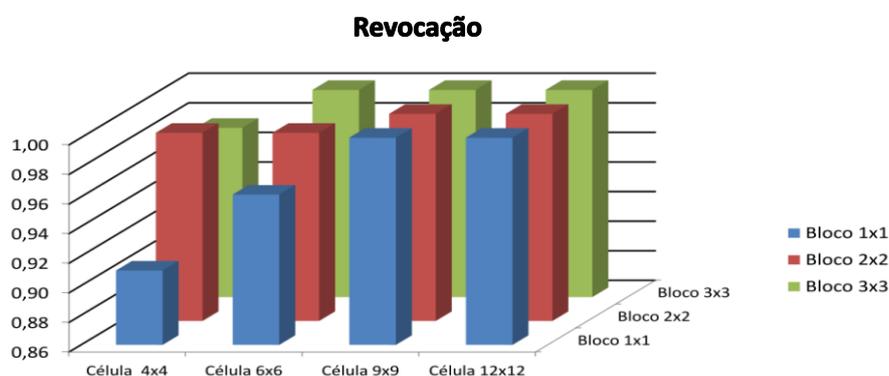
Assim, para a base brasileira será utilizada uma janela de detecção de  $120 \times 40$  *pixels*, correspondendo ao objeto de interesse ( $102 \times 34$  *pixels*) juntamente com os *pixels* de fundo (9 *pixels* na horizontal e 3 *pixels* na vertical em cada lado). Para a base grega, nenhum fundo é adicionado, resultando em uma janela de detecção de  $120 \times 30$  *pixels*.

### 5.2.3 Definição das Configurações do Descritor HOG

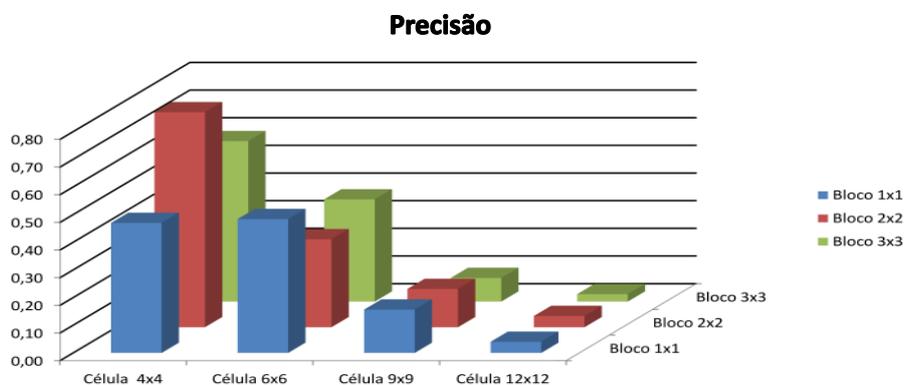
Com o intuito de avaliar diferentes configurações de células e blocos, mantemos a análise **estática** de 11 escalas, ajustadas empiricamente e escolhemos um limiar para o classificador SVM de forma que a quantidade de Falsos Positivos Por Imagem (FPPI) fosse menor que 1. Para a base brasileira, são analisadas as configurações dos descritores HOG relacionadas a janela de detecção de  $120 \times 40$  *pixels*. Para a base grega, o mesmo procedimento é repetido para uma janela de detecção de  $120 \times 30$  *pixels*. Observe, que devido as diferentes características das bases, foi necessária a parametrização individual de dois classificadores, um para cada base.

### Configuração do Descritor - Base Brasileira

A Figura 5.7 traz uma análise da revocação em função de diferentes tamanhos de células ( $4 \times 4$ ,  $6 \times 6$ ,  $9 \times 9$  e  $12 \times 12$  *pixels*) e diferentes blocos ( $1 \times 1$ ,  $2 \times 2$  e  $3 \times 3$  células por bloco). Uma comparação semelhante, mas em termos de precisão pode ser encontrada na Figura 5.8. Considerando que a configuração a ser escolhida deve possuir uma elevada taxa de detecção, representada por uma alta revocação, com uma baixa taxa de FPPI, indicada por uma elevada precisão, foi feita a escolha da configuração de células de  $4 \times 4$  *pixels* agrupadas em blocos de  $2 \times 2$  células.



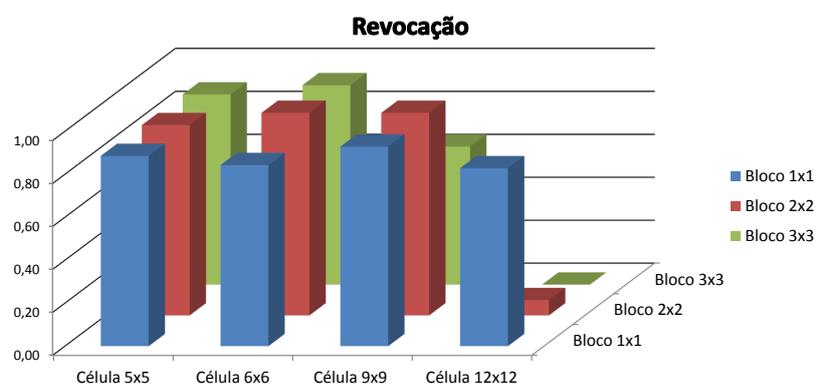
**Figura 5.7:** Análise da influência de diferentes configurações de blocos e células quanto a revocação para a base brasileira.



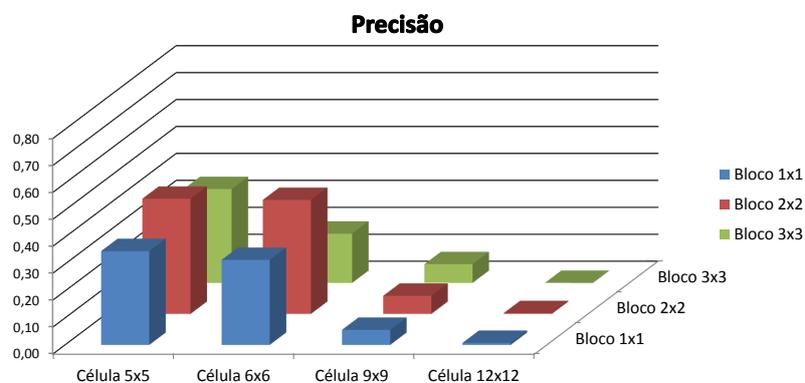
**Figura 5.8:** Análise da influência de diferentes configurações de blocos e células quanto a precisão para a base brasileira.

### Configuração do Descritor - Base Grega

A Figura 5.9 é utilizada para mostrar os resultados quanto revocação para diferentes combinações de células e blocos. Por essa, podemos perceber que para células maiores, por exemplo,  $12 \times 12$  *pixels*, os resultados são inferiores. Os melhores resultados de revicacão foram obtidos com células de  $6 \times 6$  e  $9 \times 9$  *pixels* organizadas em blocos de  $2 \times 2$  células, alcançando 94% de revocacão. No entanto, analisando a Figura 5.10, podemos perceber que a precisão da primeira (42%) é bem superior a segunda (7%). Assim, a configuração com células de  $6 \times 6$  *pixels* organizadas em blocos de  $2 \times 2$  células foi a escolhida.



**Figura 5.9:** Análise da influência de diferentes configurações de blocos e células quanto a revocação para a base grega.



**Figura 5.10:** Análise da influência de diferentes configurações de blocos e células quanto a precisão para a base grega.

### 5.2.4 Parâmetros do Método de Janelas Deslizantes Estático

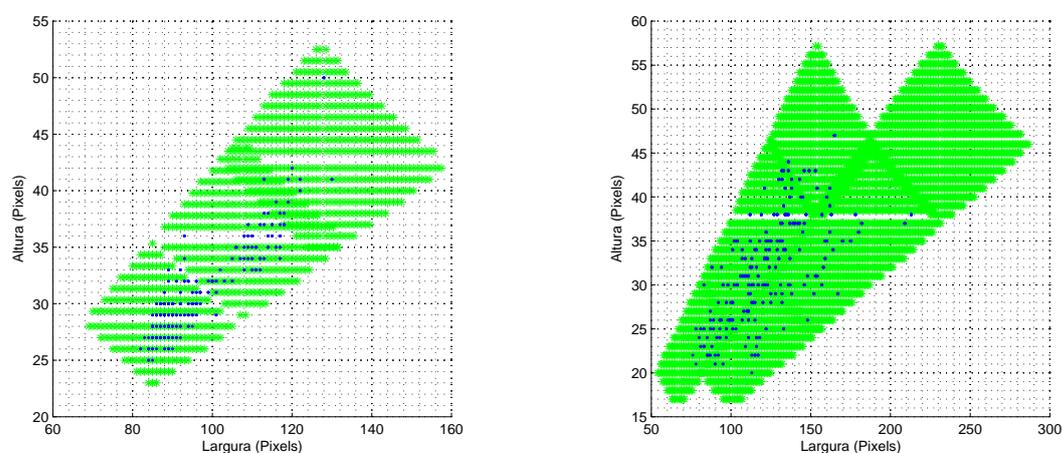
Foram analisados os parâmetros diretamente associados ao método **estático** : deslocamento entre duas janelas de detecção vizinhas e a quantidade de escalas. Para essas avaliações foram utilizadas as configurações obtidas nas análises anteriores. Janela de detecção de  $120 \times 40$  *pixels*, com blocos de  $2 \times 2$  células, sendo cada célula de  $4 \times 4$  *pixels*, para a base brasileira. Para a base grega, é empregada uma janela de detecção de  $120 \times 30$  *pixels*, com blocos de  $2 \times 2$  células, cada célula de dimensão  $6 \times 6$  *pixels*. Para a definição da quantidade de escalas a serem analisadas, foi empregado o conceito de cobertura da janela de detecção, descrito na Seção 3.3.

Na Figura 5.11, temos a cobertura para ambas as bases, cada ponto em verde corresponde a uma faixa de valores da janela que seriam cobertas com coeficiente de *Jaccard* superior a 80%, enquanto que os pontos em azul correspondem a dimensões das imagens presentes na base de treino. Essa Figura foi obtida utilizando um conjunto de três escalas (0,80;0,95;1,20) e um único AR de 1,0 para a base brasileira. Enquanto que para a base grega, de forma a lidar com as variações de perspectiva, dois valores de AR tiveram que ser empregados (0,80 e 1,20). Foi também necessária a análise de 5 escalas (0,65; 0,80; 1,0; 1,20; 1,50). Como pode ser visto, com essas configurações é possível cobrir 100% das imagens presentes no conjunto de treino.

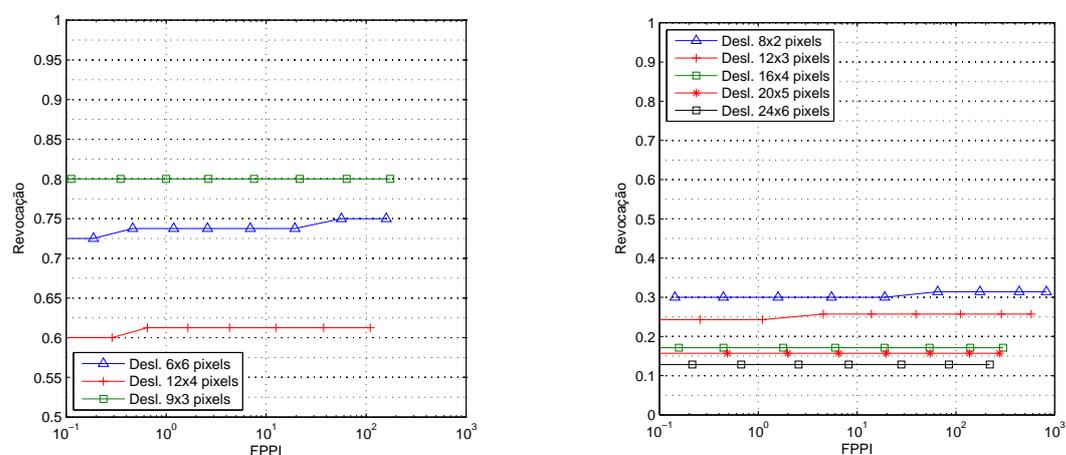
É importante ressaltar que essa cobertura só é garantida no caso de se utilizar deslocamentos unitários na vertical e horizontal, mas que se espera que generalize de tal forma que em deslocamentos superiores seja possível obter excelentes taxas de detecção para valores de coeficiente de *Jaccard* próximos a 80%.

Na Figura 5.12, é realizada a análise da influência dos deslocamentos horizontais e verticais na revocação do método **estático**. Para isso, são utilizadas as escalas obtidas pela análise da cobertura das base com coeficiente de *Jaccard* de 80%. Esses resultados foram obtidos utilizando o conjunto de validação, formado por 78 imagens, no caso das placas brasileira, e 71 imagens, para as gregas.

Como esperado, não são todas as placas que são passíveis de serem detectadas com coeficiente de *Jaccard* superior ou igual a 80% devido a influência dos deslocamentos. Mas, consegue-se revocação de 80% para a base brasileira, já para a base grega os resultados foram bastante inferiores, alcançando cerca de 30%. Devido aos melhores resultados apresentados, a configuração com deslocamento  $9 \times 3$  e  $8 \times 2$  *pixels* foram escolhidas para serem utilizadas nos testes das bases brasileira e grega, respectivamente. Confi-

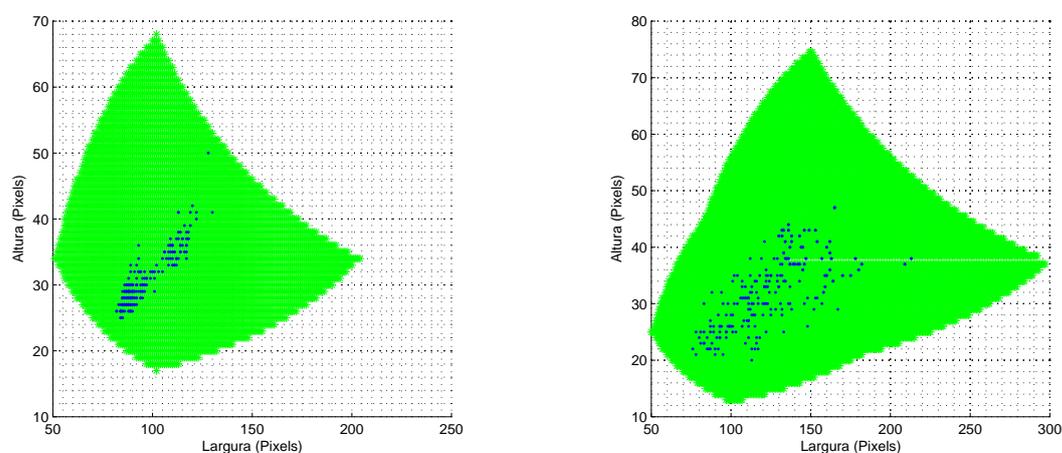


**Figura 5.11:** Cobertura da base de treino brasileira (esquerda) e grega (direita). Os pontos em verde correspondem a dimensões cobertas, enquanto que os pontos em azul são exemplares da base de treino. As dimensões cobertas foram geradas utilizando coeficiente de *Jaccard* de 80% e considerando deslocamentos unitários das janelas de detecção.



**Figura 5.12:** Avaliação da Influência dos Deslocamentos para as bases brasileira (esquerda) e grega (direita). Cada deslocamento é representado por um valor  $w \times h$  pixels que corresponde a ao deslocamento utilizado na horizontal e na vertical, respectivamente. A janela de detecção desloca horizontalmente  $w$  pixels até chegar ao final da imagem, retornando ao início, mas com um deslocamento na vertical de  $h$  pixels, de forma a cobrir uma região diferente da imagem.

gurações com deslocamentos mais próximos de 1 *pixel* poderiam garantir revocações maiores, mas ao mesmo tempo teriam um impacto expressivo no custo computacional, inviabilizando os testes.



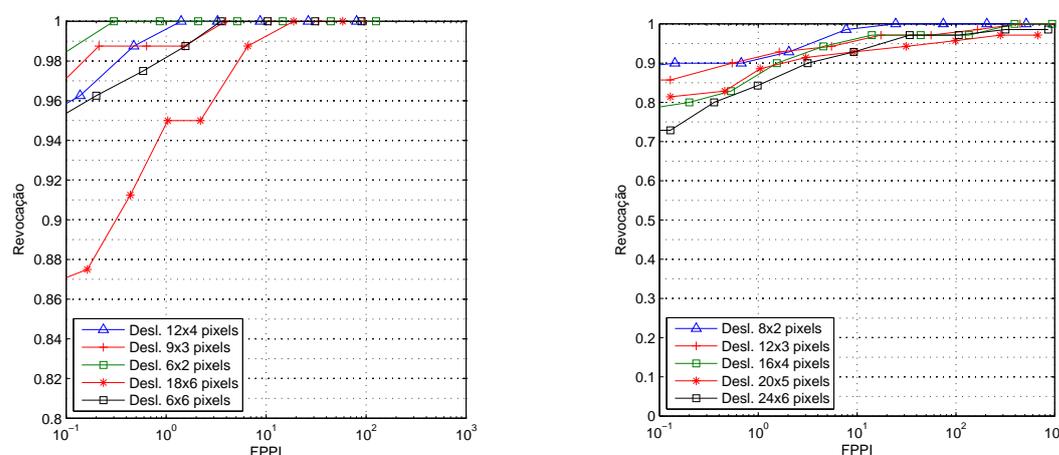
**Figura 5.13:** Cobertura da base de treino brasileira (esquerda) e grega (direita). Os pontos em verde correspondem a dimensões cobertas, enquanto que os pontos em azul são exemplares da base de treino. As dimensões cobertas foram geradas utilizando coeficiente de *Jaccard* de 50% e considerando deslocamentos unitários das janelas de detecção.

### 5.2.5 Parâmetros do Método de Janelas Deslizantes Adaptativo

Na primeira etapa do método adaptativo, estamos interessados em obter uma estimativa inicial para a posição da placa veicular. A quantidade de escalas a serem analisadas é a mínima necessária, pois não é necessário nessa etapa uma estimativa com elevados valores de coeficiente de *Jaccard*. Assim, foi definida a quantidade de escalas necessária para uma cobertura da base com coeficiente de *Jaccard* de 50%. Dessa forma, estamos deixando para a etapa posterior de refinamento o ajuste da posição e tamanho da janela de detecção.

A Figura 5.13 é empregada para mostrar as dimensões cobertas com coeficiente de *Jaccard* de 50% (pontos verdes), e as dimensões das placas veiculares presentes na base de treino (pontos azuis). Para a base brasileira (esquerda), foi possível cobrir todas as janelas do conjunto de treino usando apenas uma escala (1,0), enquanto que para a base grega foram necessárias duas escalas (0,8 e 1,2). Porém, devemos ainda ajustar os deslocamentos horizontais e verticais para a detecção da estimativa inicial da posição da placa veicular.

Na Figura 5.14, é apresentada uma análise da Revocação  $\times$  FPPI para diferentes valores de deslocamento horizontal e vertical em *pixels*. Para a essa análise, é utilizado um coeficiente de *Jaccard* de 50%, uma vez que o interesse é encontrar uma boa estimativa



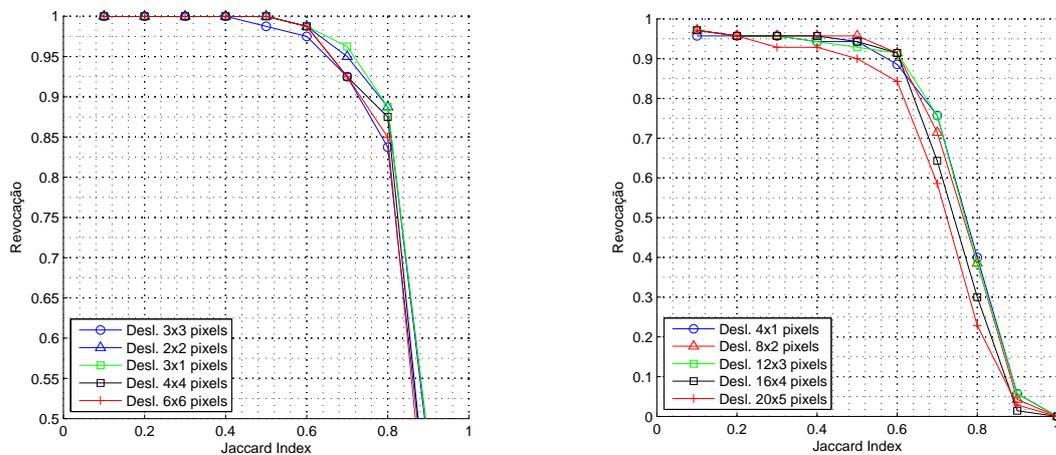
**Figura 5.14:** Análise de diferentes valores para deslocamento em valores  $w \times h$  pixels, onde  $w$  é o deslocamento horizontal e  $h$  o deslocamento vertical. Para a determinação de uma região como detectada ou não é utilizado coeficiente de Jaccard de 50%. A imagem à esquerda apresenta os resultados para a base brasileira e, à direita, temos os resultados da base grega.

inicial da posição da placa veicular. Assim, é escolhida a configuração que permite uma detecção de 100% da base com a menor quantidade de FPPI possível, correspondendo a  $6 \times 2$  e  $8 \times 2$  pixels para as bases brasileira e grega, respectivamente.

Partindo da estimativa inicial, o método **adaptativo** proposto neste trabalho realiza uma busca local e em múltiplas escalas com o intuito de obter uma boa estimativa inicial para o tamanho da placa veicular. Serão utilizadas as mesmas escalas e AR que foram definidos para a busca estática no processo de busca local em múltiplas escalas. Os parâmetros a serem calibrados se reduzem então a: dimensões da região local a ser pesquisada, fator multiplicativo de reescalonamento da imagem e o deslocamento horizontal e vertical utilizado na varredura da região local.

Para a determinação da região local a ser analisada no método de busca local, foi utilizada a análise da região mínima de busca local, que é a menor região que deve ser analisada de forma que se tenha uma garantia que 100% da área da placa veicular esteja contida, conforme descrito na Seção 3.4.

Para o cálculo do coeficiente de Jaccard, é desprezada a adição de fundo, ou seja, a janela considerada é o próprio objeto de interesse. Portanto, como a região mínima de busca local deve ser formada pela adição de um terço da largura na horizontal e um terço da altura na vertical em pixels para cada lado, temos uma dimensão de  $170 \times 56$  e  $200 \times 50$  pixels, centradas na detecção inicial, para as bases brasileira e grega, respec-



**Figura 5.15:** Análise de diferentes valores para deslocamento em valores  $w \times h$  pixels, onde  $w$  é o deslocamento horizontal e  $h$  o deslocamento vertical. Esses deslocamentos são empregados na Busca local e adaptativa. O gráfico à esquerda apresenta os resultados para a base brasileira, enquanto que à direita temos os resultados para a base grega.

tivamente.

A Figura 5.15 traz os resultados em Revocação  $\times$  Jaccard Index para diferentes configurações de deslocamentos da janela de detecção no processo de busca local e adaptativa em multiescalas para ambas as bases. Observe, que os melhores resultados obtidos para coeficiente de Jaccard de 80% na base brasileira (gráfico à esquerda), são para deslocamentos de  $2 \times 2$  e  $3 \times 1$  pixels, correspondendo a aproximadamente 88% de revocação. A configuração que será utilizada nos testes para os deslocamentos locais será a de  $3 \times 1$  pixels devido ao seus resultados superiores ou iguais as demais em toda as faixas de coeficiente de Jaccard. Por analogia, podemos concluir que a melhor configuração para a base grega é de  $4 \times 1$  pixels.

Os demais parâmetros diretamente associados ao método **adaptativo** são o fator de reescalonamento da imagem e a quantidade mínima de melhora entre duas iterações consecutivas para que o método continue. Esses parâmetros foram ajustados empiricamente para um fator de reescalonamento de 10% e uma melhora mínima de 0,05 entre duas iterações consecutivas.

### 5.3 Método proposto em (Mendes et al. 2011)

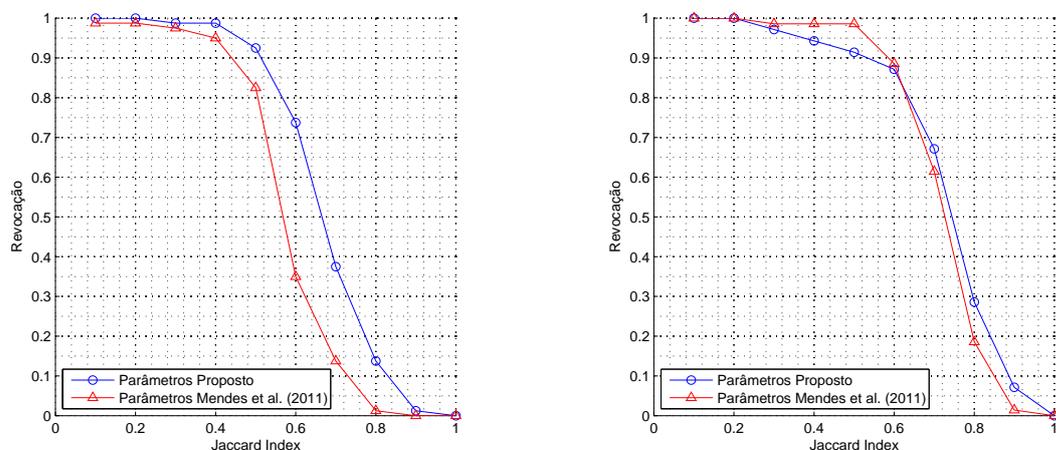
Objetivando uma comparação justa com um método presente na literatura, foi feita a opção pela comparação com o método presente em (Mendes et al. 2011), uma vez que tanto as bases quanto o código estão disponíveis *online* em (Mendes-Júnior et al. 2011).

Esse método é baseado na detecção de bordas horizontais da imagem e aplica uma sequência de operações morfológica de forma a realçar as regiões que contêm as placas veiculares. Uma análise de componentes conexos também é empregada pelos autores de forma a eliminar regiões cujas características geométricas não condizem com os parâmetros obtidos em uma validação cruzada utilizando 5 conjuntos. Devido a junção de diferentes técnicas, este método será referido como método **híbrido**. Trata-se de um método com revocação de  $94,43\% \pm 0,50\%$  e  $96,52\% \pm 0,18\%$  nas bases brasileira e grega, respectivamente. No entanto, os autores não utilizam o coeficiente de *Jaccard*, mas sim um conjunto de métricas é proposto para avaliar a detecção.

Os parâmetros utilizados em (Mendes et al. 2011) são muito bem detalhados. Assim, é possível utilizar os parâmetros fornecidos pelos autores para os testes a serem realizados neste trabalho. No entanto, de forma a lidar com a variação na base devido as divisões diferentes dos conjuntos empregados, será realizada uma análise empírica dos parâmetros, usando os valores disponível em (Mendes et al. 2011) apenas como uma referência para um ajuste mais preciso.

A Figura 5.16 é empregada para mostrar os resultados superiores obtidos no conjunto de validação após um refinamento da calibração sugerida em (Mendes et al. 2011) utilizando nossas bases de treino. Por exemplo, os parâmetros empregados pelos autores consegue uma revocação de 86,8% na base brasileira (esquerda), enquanto que realizando um ajuste empírico adicional no conjunto de treino é possível alcançar uma taxa de detecção de 95%, ambos resultados para coeficiente de *Jaccard* de 50%. É importante ressaltar que na base grega (direita), apesar de a detecção utilizando os parâmetros sugeridos em (Mendes et al. 2011) serem superiores aos nossos para coeficiente de *Jaccard* de 50%, foi optado por utilizar a nossa calibração pelo melhor desempenho em valores maiores de coeficiente de *Jaccard*.

A Tabela 5.2 é empregada para descrever os parâmetros utilizados, a coluna *Original* se refere aos parâmetros informados em (Mendes et al. 2011), enquanto que a coluna *Proposto* refere-se a nossa calibração. Uma descrição desses parâmetros vai além do escopo deste trabalho e pode ser encontrada em (Mendes et al. 2011).



**Figura 5.16:** Análise de diferentes configurações de parâmetros. O gráfico à esquerda traz os resultados para a base brasileira e, à direita, temos os resultados para a base grega.

**Tabela 5.1:** Comparação dos parâmetros para o método híbrido

	Parâmetros	
	Proposto	Original
<i>MIN_HEIGHT_CHAR (pixels)</i>	15	11
<i>MIN_HEIGHT_VLP (pixels)</i>	15	11
<i>MAX_HEIGHT_CHAR (pixels)</i>	53	43
<i>MIN_WIDTH_VLP (pixels)</i>	70	72
<i>MAX_DISTANCE_CHARS (pixels)</i>	15	20
<i>MAX_DISTANCE_PARTS (pixels)</i>	5	5
<i>HEIGHT_MEAN_FILTER (pixels)</i>	8	6
<i>WIDTH_MEAN_FILTER (pixels)</i>	16	25

## 5.4 Configuração Utilizada nos Experimentos

Na continuação, serão resumidos os parâmetros utilizados para cada base (grega e brasileira) em cada um dos métodos : **adaptativo**, **estático**.

Para o método híbrido, foi realizado um refinamento nos parâmetros fornecidos em (Mendes et al. 2011). Esses parâmetros estão disponíveis na Tabela 5.2, no campo

*Proposto*, e serão utilizados neste trabalho por terem conseguido resultados superiores no conjunto de validação quando comparados com os parâmetros originais, como pode ser visto na Figura 5.16. Para coeficiente de *Jaccard* de 0,6, é possível observar que com os parâmetros proposto conseguimos resultados superiores em mais de 30% com relação aos parâmetros originais para a base brasileira (esquerda) e praticamente iguais para a base grega.

### 5.4.1 Parâmetros Base Brasileira

O método adaptativo e estático utilizam uma mesma configuração para o descritor HOG: janela de detecção de  $120 \times 40$  *pixels* com 9 e 3 *pixels* de fundo em cada lado na horizontal e vertical, respectivamente. Nessa janela, temos um conjunto de blocos formados por  $2 \times 2$  células de dimensão igual a  $4 \times 4$  *pixels*. Os histogramas de cada uma das células é calculado com 9 bins regularmente espaçados na faixa de 0 a  $180^\circ$ . Os atributos extraídos de cada um dos blocos são normalizados utilizando *norma - L1* e são concatenados em um vetor de 10089 atributos.

A principal diferença dos métodos estático e adaptativo está na forma como o espaço de soluções é explorado. No método estático, é realizada uma busca em 3 escalas (0, 80; 0, 95; 1, 20) com deslocamentos de  $9 \times 3$  *pixels*, horizontal e vertical, respectivamente. Já no método adaptativo, inicialmente, é conduzida uma busca em uma única escala (1, 0) utilizando deslocamentos de  $6 \times 2$  *pixels* para a obtenção de uma estimativa inicial da posição. Uma estimativa inicial para o tamanho é obtida em uma busca em três escalas (0, 80; 0, 95; 1, 20) com deslocamentos de  $3 \times 1$  *pixels* em uma região centrada na placa veicular de dimensão  $170 \times 56$  *pixels*. Após essa etapa, é iniciado o processo adaptativo que reescala a imagem para valores inferiores e superiores utilizando um fator de 10% e continua enquanto as melhoras forem superiores a 0,05.

### 5.4.2 Parâmetros Base Grega

Para a base grega, são utilizados para o descritor HOG os seguintes parâmetros: janela de detecção de  $120 \times 30$  *pixels* sem adição de fundo. Os blocos são formados por  $2 \times 2$  células de dimensão igual a  $6 \times 6$  *pixels*. Os histogramas de cada uma das células é calculado com 9 bins regularmente espaçados na faixa de 0 a  $180^\circ$ . Uma normalização do tipo *norma - L1* é empregada no vetor final de 3159 atributos.

No método estático, é realizada uma busca em cinco escalas (0, 65; 0, 80; 1, 0; 1, 20; 1, 50) e dois AR (0, 80; 1, 20) com deslocamentos de  $8 \times 2$  *pixels*, horizontal e vertical, respectivamente. Para o método adaptativo é conduzida uma busca em duas escala (0, 80; 1, 20) utilizando deslocamentos de  $8 \times 2$  *pixels* obtendo as estimativas iniciais de posição. Uma estimativa inicial para o tamanho é obtida em uma busca local utilizando as mesmas escalas e AR do método estático, porém com deslocamentos de  $4 \times 1$  *pixels* em uma região centrada na placa veicular de dimensão  $200 \times 50$  *pixels*. O processo adaptativo utiliza um fator de reescalonamento de 10% e uma melhoria mínima de 0,05.

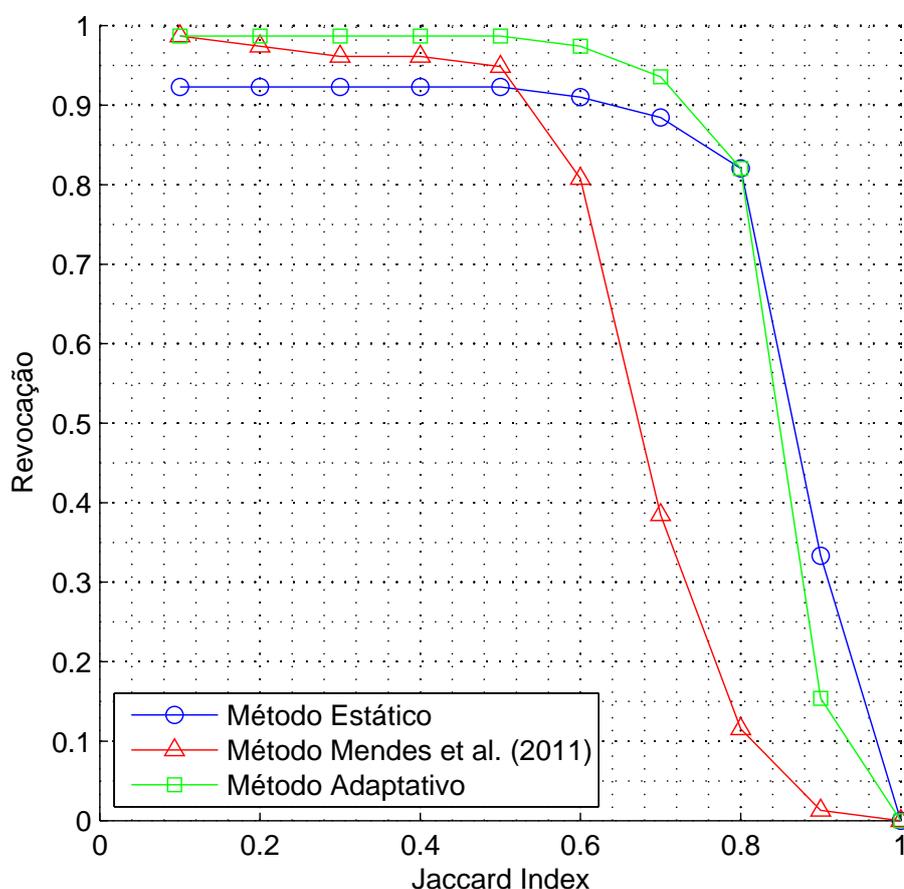
## 5.5 Experimentos

Nesta Seção serão apresentados os resultados obtidos utilizando os conjuntos de testes das bases brasileira e grega. Para uma análise da qualidade da detecção, serão empregados gráficos de revocação e precisão em função do coeficiente de *Jaccard*.

### Resultados - Base Brasileira

Na Figura 5.17, são apresentados os resultados de Revocação  $\times$  *Jaccard Index* para os diferentes métodos analisados: **estático**, **adaptativo** e o **híbrido**. Analisando essa Figura, podemos perceber que os métodos baseado em descritores HOG conseguem garantir uma detecção mais precisa, uma vez que o desempenho não se degrada tão rapidamente quanto o do método híbrido. Além disso, é possível afirmar que para coeficiente de *Jaccard* de 80%, os métodos apresentam desempenho praticamente igual, revocação de 82%. Para coeficiente inferiores o método adaptativo possui resultados em média 5% melhores, enquanto que para valores superiores o método estático foi ligeiramente superior.

Na Figura 5.17, se utilizarmos o limiar de coeficiente de *Jaccard* de 50%, obtemos revocações de 98%, 95% e 92%, para os métodos adaptativo, estático e híbrido, respectivamente. Apesar do desempenho do método adaptativo, no geral, ser pouco superior ao método estático, enquanto o método estático precisa analisar 42608 janelas por imagem, o método adaptativo analisa  $32653,3 \pm 314,6$  janelas (valor em média  $\pm$  desvio padrão), como pode ser visto na Tabela 5.2. Assim, conseguimos um ganho de eficiência de aproximadamente 25%, sem perda de eficácia do método.



**Figura 5.17:** Resultados em termos de Revocação  $\times$  *Jaccard Index* para a base brasileira utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos **adaptativo**, **estático** e **híbrido**.

**Tabela 5.2:** Comparação de eficiência dos métodos baseados em janelas deslizantes.

	Base de Imagens	
	Brasileira	Grega
Total windows (#)	42608	285600
Adaptativo Porcentagem do Total (%)	76,6( $\pm 0,01$ )	26,21( $\pm 0,04$ )
Estático Tempo(seg.)	20,3( $\pm 0,97$ )	64,2( $\pm 1,56$ )
Adaptativo Tempo (seg.)	10,5( $\pm 0,8$ )	15,8( $\pm 0,83$ )

Os resultados quanto a Precisão  $\times$  *Jaccard Index* são apresentados na Figura 5.18. Por essa, podemos perceber que apesar dos métodos baseados em descritores HOG serem

bem próximos quanto a revocação, o método adaptativo apresenta um desempenho bem superior ao método estático quanto a precisão. Foi observado que algumas regiões de fundo da imagem que em uma única escala não seriam confundidas com placas, passaram a ser na análise em múltiplas escalas.

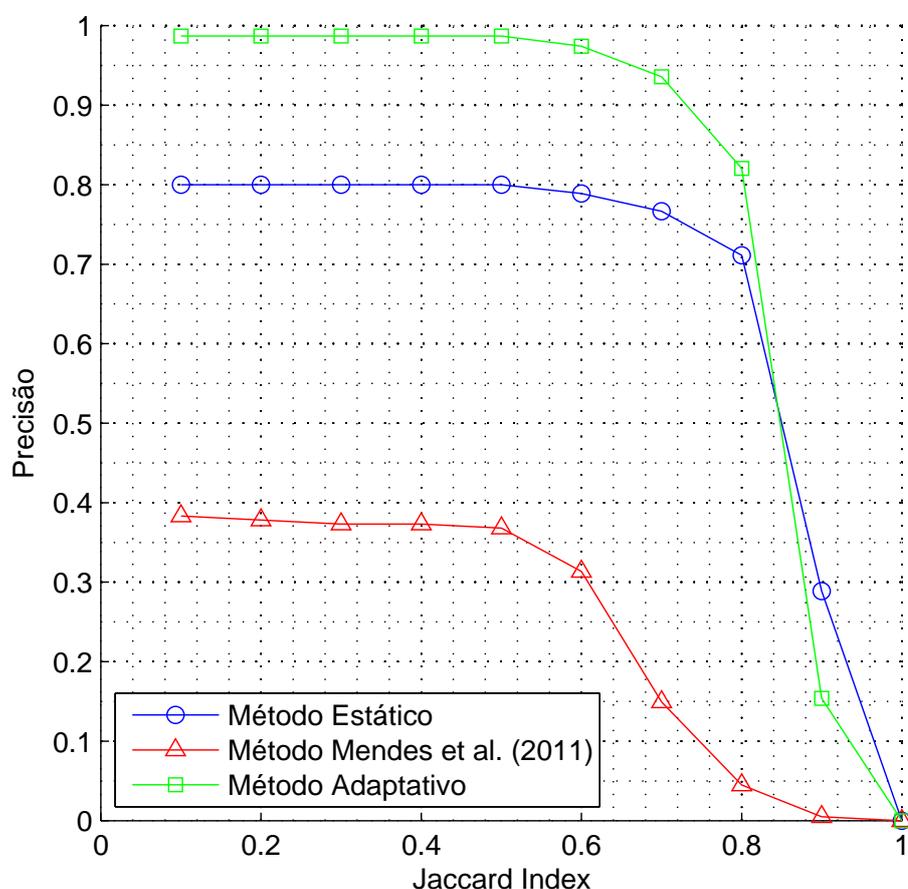
Ambos os métodos apresentaram resultados superiores quando comparados ao método híbrido para a precisão. Como já mencionado, os métodos baseados na detecção de bordas apresentam a desvantagem de serem muito sensíveis aos ruídos, justificando a grande quantidade de falsos positivos e a menor precisão. Para coeficiente de *Jaccard* de 80%, o método adaptativo apresenta uma precisão de 82%, enquanto que o método estático e híbrido têm 71,5% e 5%, respectivamente. É importante ressaltar que o método híbrido apresenta a possibilidade de utilizar um módulo de verificação de candidatos, o qual foi desativado nesta análise de forma a obter a máxima revocação possível, prejudicando a precisão.

### Resultados - Base Grega

A Figura 5.19 mostra os resultados de Revocação  $\times$  *Jaccard Index* para a base grega. Pode-se perceber a partir desse que o desempenho dos métodos baseados em descritores apresentam uma queda de desempenho quando comparados a base brasileira para coeficiente de *Jaccard* de 80%. Essa piora pode ser atribuída as diferentes perspectivas e inclinações das placas nessa base, as quais dificultam tanto a rotulação quanto a detecção. No entanto, podemos afirmar que o método adaptativo proposto neste trabalho apresenta resultados de revocação superiores com relação aos demais comparados. Além disso, para coeficiente de *Jaccard* de 50%, o método adaptativo apresenta uma excelente taxa de detecção de, aproximadamente, 99%.

Na Figura 5.20, temos os resultados de Precisão  $\times$  *Jaccard Index* para a base grega. Nesse caso, como a quantidade de ruídos de fundo com características de bordas semelhantes as placas veiculares é menor, o método híbrido se destaca em relação aos demais. O método adaptativo apresenta maior precisão que o método estático, o que pode ser justificado pela menor quantidade de janelas analisadas.

A Tabela 5.2 é utilizada para mostrar a análise de eficiência dos métodos que utilizam janelas deslizantes, adaptativo e estático. Como podemos perceber, existe um ganho considerável na eficiência, mas ambos os métodos se encontram distantes do tempo necessário para uma aplicação, que seria da ordem de milissegundos. O método híbrido

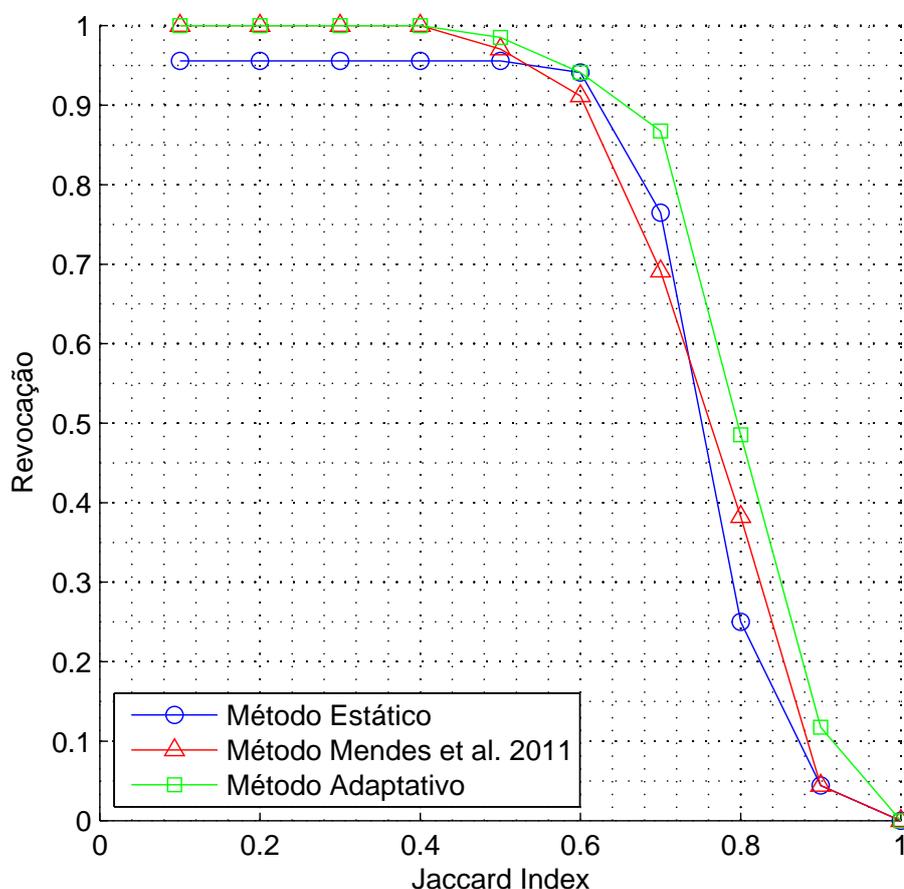


**Figura 5.18:** Resultados em termos de Precisão  $\times$  *Jaccard Index* para a base brasileira utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos **adaptativo**, **estático** e proposto em (Mendes et al. 2011).

é o que mais se aproxima das exigências de tempo real, levando menos de um segundo por imagem. O ganho maior de eficiência na base grega se deve a enorme quantidade de janelas necessárias de analisar em uma abordagem estática, resultado da variação maior da dimensão das placas.

## Discussões

A Figura 5.21, é empregada para mostrar a influência do processo de busca local e adaptativa em múltiplas escalas no aumento da qualidade da detecção. Sem a utilização da etapa adaptativa, a taxa de detecção de placas com *Jaccard* superior 80% seria de 50%, enquanto que empregando o processo de adaptativo, temos um aumento em mais

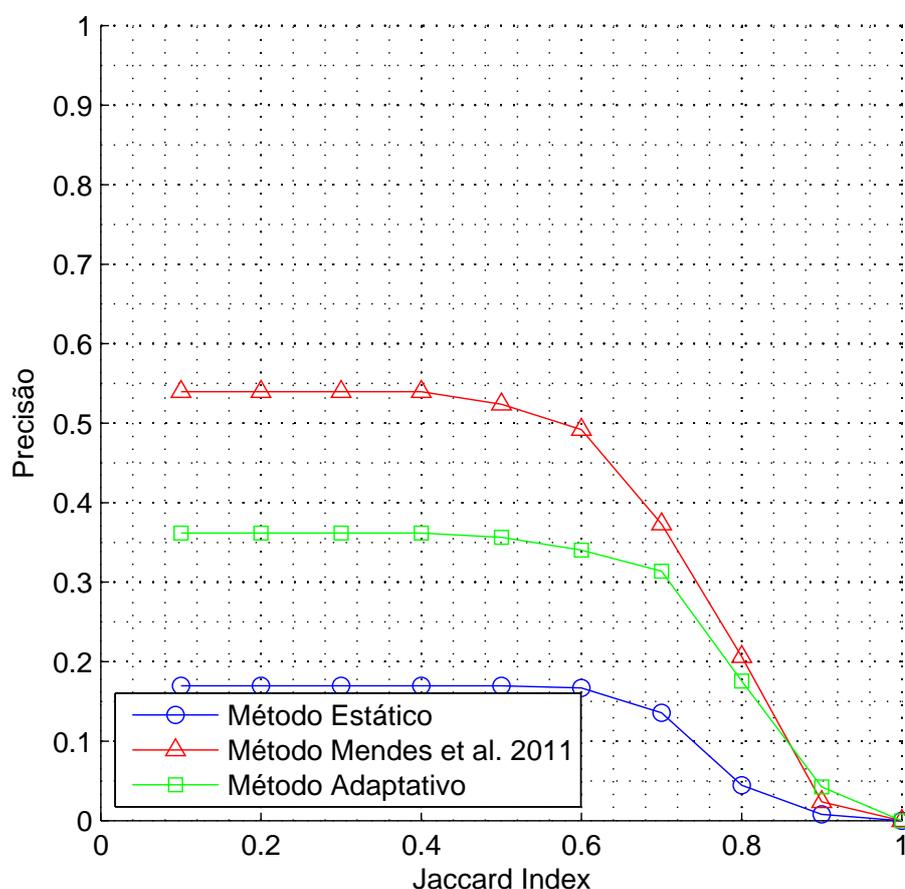


**Figura 5.19:** Resultados em termos de Revocação  $\times$  *Jaccard Index* para a base grega utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos **adaptativo**, **estático** e **híbrido**.

de 30%, detectando 82% das placas. Como mostrado anteriormente, na Figura 5.3, esse aumento no coeficiente de *Jaccard* é essencial, uma vez que com valores de *Jaccard* inferiores a 70%, não se pode garantir o reconhecimento da placa veicular.

A Figura 5.22 traz alguns exemplos de detecção para os métodos adaptativo e híbrido. Como podemos ver pelas imagens, ambos os métodos foram capazes de determinar uma região contendo a placa veicular. No entanto, o método híbrido apresenta a desvantagem de ter um elevado número de FPPI, além de detectar uma quantidade muito grande de informação além da placa veicular, o que dificulta as etapas posteriores de segmentação dos caracteres e reconhecimento.

Na Figura 5.23, podemos realizar a comparação visual entre os resultados do método estático e adaptativo. Ambos os métodos conseguiram boas detecções nos exemplos

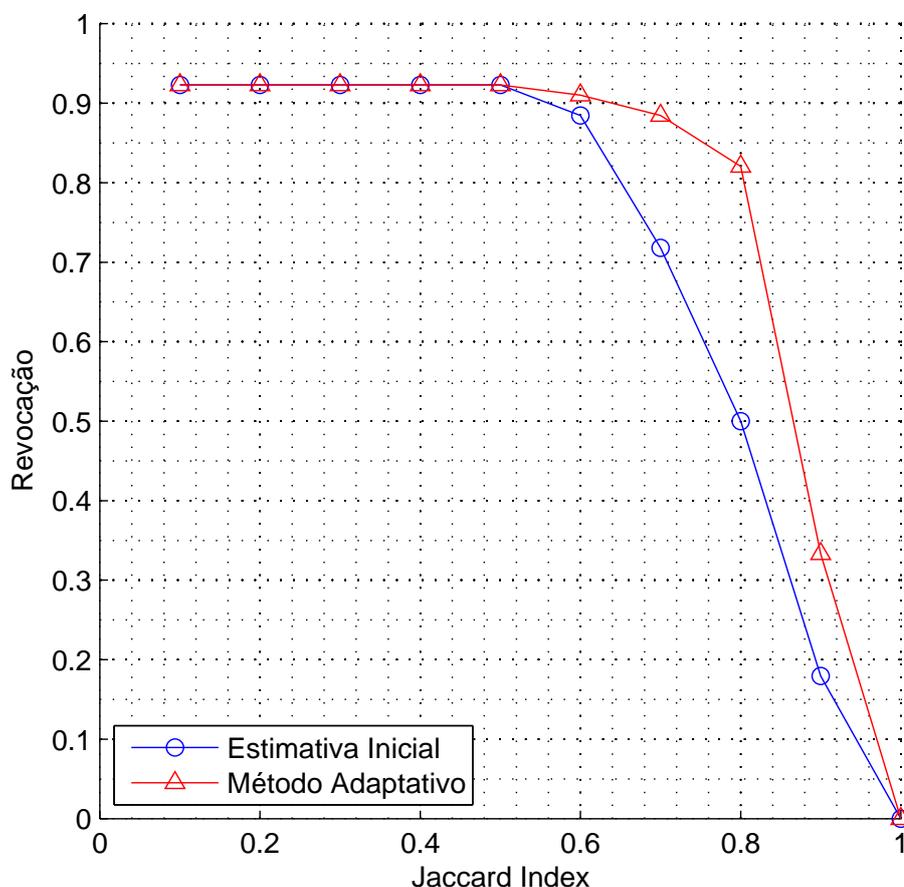


**Figura 5.20:** Resultados em termos de Revocação  $\times$  *Jaccard Index* para a base grega utilizando as configurações de parâmetros que apresentaram os melhores resultados durante a calibração para os métodos **adaptativo**, **estático** e **híbrido**.

mostrados, sendo a diferença entre as regiões detectadas pelo método estático (direita) e o adaptativo (esquerda) de poucos *pixels*. No entanto, devido a análise em múltiplas escalas da imagem inteira e não apenas localmente, o método estático acaba por detectar mais FPPI, representado na imagem pelas janelas de detecção (regiões com contorno azul) fora da placa veicular.

## 5.6 Considerações Finais

Neste capítulo, apresentamos as bases de imagens utilizadas nos experimentos, as quais apresentam diversos outros elementos capazes de confundirem os métodos de detecção. Os resultados obtidos com o método híbrido comprovam a dificuldade da base brasileira



**Figura 5.21:** Gráfico de Revocação  $\times$  *Jaccard Index* para o método **adaptativo** proposto utilizando apenas a estimativa inicial da posição da placa veicular e empregando a etapa de busca local adaptativa em múltiplas escalas.

devido a presença de um número elevado de falsos positivos. Por outro lado, a maior quantidade de escalas necessárias e o pior resultado para valores maiores de coeficiente de *Jaccard* reforçam a maior variação de dimensão e perspectiva da base grega.

Os descritores HOG são mais indicados pelos resultados superiores de revocação em ambas as bases. Uma análise da qualidade da detecção empregando o coeficiente de *Jaccard* é empregada, mostrando que para valores de *Jaccard* inferiores a 70%, as detecções não garantem que todos os caracteres da placa estejam íntegros. Assim, foi estabelecido um limiar de *Jaccard* de 80% como ideal. Para esse limiar, tanto o método adaptativo quanto o estático apresentam a mesma taxa de detecção de 82% na base brasileira, mas a quantidade de janelas analisadas pelo primeiro é 25% menor, comprovando sua eficiência. Para a base grega, todos os métodos têm um desempenho ruim, abaixo dos 50% de revocação. No entanto, ainda assim o desempenho do método adaptativo



**Figura 5.22:** Exemplos de resultados obtidos com o método adaptativo (esquerda) e o método híbrido (direita). As regiões em vermelha correspondem às rotulações, enquanto que as regiões em azul são as detectadas.

é superior aos demais com uma análise de aproximadamente 26,21% do total analisado no método estático. Os resultados superiores encontrados pelo método adaptativo em ambas as bases comprova a eficiência dos métodos de determinação dos parâmetros utilizando a análise de cobertura das janelas de detecção e a área mínima de busca local. No entanto, o resultado abaixo das expectativas na base grega indica a necessidade de se estudar técnicas para a detecção de placas com diferentes perspectivas.



**Figura 5.23:** Exemplos de resultados obtidos com o método adaptativo (esquerda) e o método estático (direita). As regiões em vermelha correspondem às rotulações, enquanto que as regiões em azul são as detectadas.

# Capítulo 6

## Conclusões

Neste trabalho é apresentado um método inédito de detecção de placas veiculares em diferentes escalas e de forma adaptativa de empregando os descritores HOG. Nossa abordagem pode ser dividido em duas fases. Inicialmente, é realizada a estimativa da posição das placas veiculares empregando janelas deslizantes em uma quantidade mínima de escalas. A partir dessa posição inicial, precisamos determinar a posição e o tamanho precisos da placa veicular, fazendo uso de uma busca local e adaptativa em mais escalas. Assim, garantimos um maior grau de casamento quando comparado a outros métodos da literatura, facilitando as etapas de segmentação e reconhecimento de caracteres. Mais importante ainda, essa detecção é realizada com uma redução de 25% e 75% na quantidade de escalas analisadas quando comparado a uma abordagem em múltiplas escalas clássica nas bases brasileira e grega, respectivamente.

Nosso método é comparado a dois trabalhos presentes na literatura: (1) um híbrido que utiliza morfologia matemática em conjunto com as informações de bordas verticais (Mendes et al. 2011); (2) uma detecção em múltiplas escalas utilizando janelas deslizantes e descritores HOG. As diferentes abordagens foram avaliadas em duas bases de dados disponível online e os resultados obtidos mostram que o método proposto consegue manter resultados superiores de revocação e precisão à medida que as exigências de grau de casamento aumentam. O híbrido, apesar de apresentar um menor custo computacional, tem o desempenho rapidamente degradado à medida que maiores graus de casamento são desejados. Além disso, o híbrido apresenta uma menor precisão quando comparado à abordagem proposta em ambientes abertos. Isso pode ser justificado pelo menor poder de discriminação de regiões com textura semelhante a das placas veiculares.

Este trabalho é importante para os SIAV, mais precisamente na detecção de placas veiculares, por propor um método no qual se utiliza os descritores HOG, que têm sido reportados na literatura como aqueles que apresentam maior poder de discriminação entre regiões de placa e não-placa, mas em uma abordagem adaptativa de busca em múltiplas escalas. Dessa forma, tal sistema vai ao encontro das exigências de eficácia e eficiência. No entanto, apesar dos resultados encontrados ainda serem bem superiores a abordagem em múltiplas escalas clássica, eles ainda se encontram distantes de um sistema com as exigências de tempo-real necessárias (em média 10,5 segundos para uma imagem de  $800 \times 600$  *pixels*).

Em trabalhos futuros pretendemos lidar com as exigências temporais acelerando o cálculo dos descritores HOG e da resposta do classificador SVM utilizando Unidades de Processamento Gerais (*General Process Units* - GPU) (Prisacariu & Reid 2009).

## Referências Bibliográficas

- Abolghasemi, V. & Ahmadyfard, A. (2009). An edge-based color-aided method for license plate detection, *Image and Vision Computing (IVC)* **27**(8): 1134–1142.
- Al-Ghaili, A. M., Mashohor, S., Ismail, A. & Ramli, A. R. (2008). A new vertical edge detection algorithm and its application, *International Conference on Computer Engineering and Systems*, pp. 204–209.
- Al-Ghaili, A. M., Mashohor, S., Ramli, A. R. & Ismail, A. (2013). Vertical-edge-based car license plate detection method, *IEEE Transactions on Vehicular Technology*, Vol. 62, pp. 26–38.
- Alegria, F. & Girão, P. S. (2006). Vehicle plate recognition for wireless traffic control and law enforcement, *Industrial Technology, 2006. ICIT 2006. IEEE International Conference on*, IEEE, pp. 1800 – 1804.
- Bai, H. & Liu, C. (2004). A hybrid license plate extraction method based on edge statistics and morphology, *International Conference on Pattern Recognition (ICPR)*, pp. 831–834.
- Benenson, R., Mathias, M., Timofte, R. & Gool, L. J. V. (2012). Pedestrian detection at 100 frames per second., *International Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 2903–2910.
- Bradski, G. (2000). The OpenCV Library, *Dr. Dobb's Journal of Software Tools* .
- Chang, S.-L., Chen, L.-S., Chung, Y.-C. & Chen, S.-W. (2004). Automatic license plate recognition, *IEEE Transactions on Intelligent Transportation Systems* **5**(1): 42–53.
- Cui D. and Gu, D. and Cai, H. and Sun, J. (2009). License Plate Detection Algorithm Based on Gentle AdaBoost, *IEEE International Conference on Robotics and Biomimetics*, pp. 1962–1966.

- Dalal, N. & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection, *International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1, pp. 886–893.
- Deb, K., Chae, H.-U. & Jo, K.-H. (2009). Vehicle license plate detection method based on sliding concentric windows and histogram, *Journal of Computers* **4**(8): 771–777.
- Dlagnekov, L. (2004). License plate detection using adaboost, *Technical report*, Department of Computer Science and Engineering, University California of San Diego.
- Dollár, P., Wojek, C., Schiele, B. & Perona, P. (2012). Pedestrian detection: An evaluation of the state of the art, *Pattern Analysis and Machine Intelligence* **34**: 743 – 761.
- Donoser, M., Arth, C. & Bischof, H. (2007). Detecting, tracking and recognizing license plates, *Asian Conference on Computer Vision (ACCV)*, Springer Berlin Heidelberg, pp. 447–456.
- Du, S., Ibrahim, M., Shehata, M. S. & Badawy, W. M. (2013). Automatic license plate recognition (ALPR): A state-of-the-art review, *IEEE Transactions on Circuits and Systems for Video Technologie*. pp. 311–325.
- Duan, T. C., Du, T. H., Phuoc, T. V. & Hoang, N. V. (2005). Building An Automatic Vehicle License-plate Recognition System, *International Conference in Computer Science (RIVF)*, pp. 59–63.
- Frome, A., Cheung, G., Abdulkader, A., Zennaro, M., Wu, B., Bissacco, A., Adam, H., Neven, H. & Vincent, L. (2009). Large-scale privacy protection in google street view, *IEEE International Conference on Computer Vision*, pp. 2373–2380.
- Guo, J.-M. & Liu, Y.-F. (2008). License plate localization and character segmentation with feedback self-learning and hybrid binarization techniques, *IEEE Transactions on Vehicular Technology* **57**(3): 1417–1424.
- Ho, W. T., Lim, H. W. & Tay, Y. (2009). Two-stage license plate detection using gentle adaboost and sift-svm, *Asian Conference on Intelligent Information and Database Systems (ACIIDS 2009)*, pp. 109–114.
- Jia, W., Zhang, H. & He, X. (2005). Mean shift for accurate number plate detection, *International Conference on Information Technology and Applications (ICITA)*, IEEE Computer Society, pp. 732–737.

- Jia, W., Zhang, H. & He, X. (2007). Region-based license plate detection, *Journal of Network and Computer Applications* **30**(4): 1324–1333.
- KwangIn, K., Keechul, J. & JinHyung, K. (2002). Color texture-based object detection: An application to license plate localization, *Pattern Recognition with Support Vector Machines* **2388**: 293–309.
- Laborda, J. A., Salgado, L. & Marinas, J. (2012). Hog-like gradient-based descriptor for visual vehicle detection, *Intelligent Vehicles Symposium*, IEEE, pp. 223–228.
- Lalimi, M. A., Ghofrani, S. & McLernon, D. (2013). A vehicle license plate detection method using region and edge based methods, *Computers & Electrical Engineering* **39**(3): 834–845.
- Lee, D.-H. & Choi, J.-H. (2010). Precise detection of car license plates by locating main characters, *Journal of the Optical Society of Korea* **14**(4): 376–382.
- Lee, H.-J., Chen, S.-Y. & Wang, S.-Z. (2004). Extraction and recognition of license plates of motorcycles and vehicles on highways, *International Conference on Pattern Recognition (ICPR)*, Vol. 4, pp. 356–359.
- Li, B., Tian, B., Li, Y. & Wen, D. (2013). Component-based license plate detection using conditional random field model, *IEEE Transactions on Intelligent Transportation Systems* **14**(4): 1690–1699.
- Li, X. (2010). *Vehicle license plate detection and recognition*, Master's thesis, Faculty of the Graduate School at the University of Missouri.
- Lim, H. & Y.H., T. (2010). Detection of license plate characters in natural scene with msr and sift unigram classifier, *IEEE Conference on Sustainable Utilization and Development in Engineering and Technology*, IEEE, pp. 95–98.
- Matas, J. & Zimmermann, K. (2005). Unconstrained licence plate and text localization and recognition, *IEEE Transaction on Intelligent Transportation Systems*, 2005, IEEE, pp. 225–230.
- Mendes-Júnior, P. R. et al. (2011). Vehicle license plate location (VLPL) algorithms, <https://github.com/pedrormjunior/vlpl>.
- Mendes, P., Neves, J., Tavares, A. & Menotti, D. (2011). Towards an automatic vehicle access control system: License plate location, *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2916–2921.

- Ministério das Cidades, Departamento Nacional de Trânsito - DENATRAN (2014). Frota de veículos - automóveis, <http://www.denatran.gov.br/frota2014.htm>. Acesso em 18/05/2014.
- Pan, L. & Li, S. (2010). A new license plate extraction framework based on fast mean shift, *International Conference on Image Processing and Pattern Recognition in Industrial Engineering (SPIE)*, pp. 782007-1 –782007-9.
- Prisacariu, V. A. & Reid, I. (2009). fasthog - a real-time gpu implementation of hog, *Technical Report 2310/09*, University of Oxford, Department of Engineering Science.
- Raphael-Prates, C. (2013). Vehicle license plate location (VLPL) algorithms, <https://github.com/pratesufop/vlpl>.
- Viola, P. & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features, *International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. I-511 – I-518.
- Wan, X., Liu, J. & Liu, J. (2011). A vehicle license plate localization method using color barycenters hexagon model, *International Conference on Digital Image Processing (ICDIP 2011)*, Vol. 8009, pp. pp.800920-1 –800920-5.
- Wang, F., Man, L., Wang, B., Xiao, Y., Pan, W. & Lu, X. (2008). Fuzzy-based algorithm for color recognition of license plates, *Pattern Recognition Letters* **29**(7): 1007–1020.
- Wang, R., Sang, N., Huang, R. & Wang, Y. (2013). License plate detection using gradient information and cascade detectors, *Optik - International Journal for Light and Electron Optics* pp. 1–10. in press.
- Wang, Y.-R., Lin, W.-H. & Horng, S.-J. (2011). A sliding window technique for efficient license plate localization based on discrete wavelet transform, *Expert Systems with Applications (ESWA)* **38**(4): 3142–3146.
- Xu, X., Wang, Z., Zhang, Y. & Liang, Y. (2006). A method of multi-view vehicle license plates location based on rectangle features, *International Conference on Signal Processing*, pp. 1502–1505.
- Yao, Z. & Yi, W. (2014). License plate detection based on multistage information fusion, *Information Fusion* **18**(0): 78 – 85.

- Yohimori, S., Mitsukura, Y., Fukumi, M., Akamatsu, N. & Pedrycz, N. (2004). License plate detection system by using threshold function and improved template matching method, *Conference of the North American Fuzzy Information Processing Society*.
- Yoon, J., Kang, B.-N. & Kim, D. (2012). Car license plate detection under large variations using covariance and hog descriptors, *in* G. Bebis, R. Boyle, B. Parvin, D. Koracin, C. Fowlkes, S. Wang, M.-H. Choi, S. Mantler, J. P. Schulze, D. Acevedo, K. Mueller & M. E. Papka (eds), *ISVC (2)*, Vol. 7432 of *Lecture Notes in Computer Science*, Springer, pp. 636–647.
- Zhang, W., Sun, J. & Tang, X. (2008). Cat head detection - how to effectively exploit shape and texture features, *European Conference on Computer Vision (ECCV)*, Vol. 4, pp. 802–816.
- Zhang, X., Shen, P., Xiao, Y., Li, B., Hu, Y., Qi, D., Xiao, X. & Zhang, L. (2010). License plate-location using adaboost algorithm, *IEEE International Conference on Information and Automation (ICIA)*, pp. 2456–2461.
- Zheng, D., Zhao, Y. & Wang, J. (2005). An efficient method of license plate location, *Pattern Recognition Letters* **26**(15): 2431–2438.
- Zheng, K., Zhao, Y., Gu, J. & Hu, Q. (2006). License plate detection using haar-like features and histogram of oriented gradients, *International Symposium on Industrial Electronics*, Vol. 3, pp. 16–20.
- Zheng, L., He, X., Samali, B. & Yang, L. T. (2013). An algorithm for accuracy enhancement of license plate recognition, *Journal of Computer and System Sciences* **79**(2): 245–255.
- Zhou, W., Li, H., Lu, Y. & Tian, Q. (2012). Principal visual word discovery for automatic license plate detection, *IEEE Transactions on Image Processing* **21**(9): 4269–4279.