



UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE PLACAS DE LICENCIAMENTO
VEICULAR UTILIZANDO IMAGENS COM
DIFERENTES RESOLUÇÕES E EM AMBIENTE
NÃO CONTROLADO EMPREGANDO TÉCNICAS
DE SUPER RESOLUÇÃO**

Luiz Carlos da Silva Garcia Junior

Manaus – Amazonas

Setembro de 2019

Luiz Carlos da Silva Garcia Junior

**DETECÇÃO DE PLACAS DE LICENCIAMENTO
VEICULAR UTILIZANDO IMAGENS COM
DIFERENTES RESOLUÇÕES E EM AMBIENTE
NÃO CONTROLADO EMPREGANDO TÉCNICAS
DE SUPER RESOLUÇÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração de Processamento Digital de Imagens.

Orientador: Prof. D.Sc. Waldir Sabino da Silva Júnior

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

G216d Garcia Junior, Luiz Carlos da Silva
Detecção de placas de licenciamento veicular utilizando imagens com diferentes resoluções e em ambiente não controlado empregando técnicas de super resolução / Luiz Carlos da Silva Garcia Junior. 2019
80 f.: il. color; 31 cm.

Orientador: Prof. D.Sc. Waldir Sabino da Silva Júnior
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Processamento digital de imagens. 2. Detecção de placas. 3. Imagens ruidosas. 4. Background complexo. 5. Opencv. I. Silva Júnior, Prof. D.Sc. Waldir Sabino da II. Universidade Federal do Amazonas III. Título

LUIZ CARLOS DA SILVA GARCIA

DETECÇÃO DE PLACAS DE LICENCIAMENTO VEÍCULAR UTILIZANDO IMAGENS COM DIFERENTES RESOLUÇÕES E EM AMBIENTE NÃO CONTROLADO EMPREGANDO TÉCNICAS DE SUPER RESOLUÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovado em 09 de setembro de 2019.

BANCA EXAMINADORA

Prof. Dr. Waldir Sabino da Silva Júnior, Presidente
Universidade Federal do Amazonas

Prof. Dr. Carlos Augusto de Moraes Cruz, Membro
Universidade Federal do Amazonas

Prof. Dr. José Luiz de Souza Pio, Membro
Universidade Federal do Amazonas

Luiz Carlos da Silva Garcia Junior

**DETECÇÃO DE PLACAS DE LICENCIAMENTO
VEICULAR UTILIZANDO IMAGENS COM
DIFERENTES RESOLUÇÕES E EM AMBIENTE
NÃO CONTROLADO EMPREGANDO TÉCNICAS
DE SUPER RESOLUÇÃO**

Banca Examinadora

Prof. D.Sc. Waldir Sabino da Silva Júnior – Presidente e Orientador
Departamento de Eletrônica e Computação – DTEC/UFAM

Prof. D.Sc. Carlos Augusto de Moraes Cruz
Departamento de Eletrônica e Computação – DTEC/UFAM

Prof. D.Sc. José Luiz de Souza Pio
Instituto de Computação - IComp/UFAM

Manaus – Amazonas

Setembro de 2019

Agradecimentos

Primeiramente quero agradecer a Deus por me conceder o dom da vida, saúde plena e pela oportunidade que me fora concedida, tenho gratidão eterna a pessoas que estiveram ao meu lado durante todo o período desta caminhada, a minha mãe, minha irmã e minha amada e companheira Gracy Kelly, vocês me incentivaram em todos os momentos, prestando total compreensão e apoio incondicional, a Raimunda Nonata e dona Graça, muito obrigado por todo apoio e compreensão e cuidados, aos meus amigos Glauco Aguiar e Elvys Araújo que também me acompanharam durante esta caminhada, ao meu orientador Professor Waldir Sabino, por toda ajuda, compreensão e instruções passadas e repassadas, obrigado professor, ao meu coorientador e amigo Newton Bitar, obrigado por toda a ajuda prestada. Aos meus colegas de laboratório que nos últimos dias estiveram comigo. Apenas gratidão a todos que me ajudaram, muito obrigado.

Resumo da Dissertação apresentada à UFAM como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

DETECÇÃO DE PLACAS DE LICENCIAMENTO VEICULAR UTILIZANDO
IMAGENS COM DIFERENTES RESOLUÇÕES E EM AMBIENTE NÃO
CONTROLADO EMPREGANDO TÉCNICAS DE SUPER RESOLUÇÃO

Luiz Carlos da Silva Garcia Junior

Orientador: Waldir Sabino da Silva Júnior

Programa: Pós-Graduação em Engenharia Elétrica

Nesta dissertação, propomos uma solução utilizando técnicas de Processamento Digital de Imagens para detecção de placas de veículos brasileiras, onde consiste em elaborar um método robusto e flexível em relação ao *background* complexo com a aquisição realizada em movimento para realizar a detecção de placas de veículos. A detecção ocorre de forma automática empregando técnicas de PDI, as quais contemplam filtros do domínio espacial e para nos auxiliar, empregamos a biblioteca *open source OPENCV* em *C/C++*. Além disso, como uma das contribuições, abordamos o emprego do algoritmo de Super Resolução Bilateral $TV - L1$ nos vídeos da base de dados com o intuito de contribuir com a detecção do objeto de interesse. Para realização dos experimentos, utilizamos uma base de dados que foi concebida contendo vídeos com três resoluções distintas, a aquisição foi realizada com câmera e objeto de interesse ambos em movimento com um *background* complexo, onde, a partir dos vídeos, empregamos imagens sequencias nos experimentos realizados. Os resultados dos experimentos abrangeram dois cenários, um primeiro cenário não foi empregado o algoritmo de Super Resolução e um segundo, empregamos o algoritmo de Super Resolução, assim, foi possível validar a metodologia proposta demonstrando a capacidade de detecção de placas de veículos frente ao

desafio proposto, no entanto, observamos que a padronização das resoluções das imagens, podem influenciar nos resultados.

Palavras-chave: processamento digital de imagens, detecção de placas de veículos, imagens ruidosas, *background* complexo, *C/C++*, *OPENCV*.

Abstract of Dissertation presented to UFAM as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering

DETECÇÃO DE PLACAS DE LICENCIAMENTO VEICULAR UTILIZANDO
IMAGENS COM DIFERENTES RESOLUÇÕES E EM AMBIENTE NÃO
CONTROLADO EMPREGANDO TÉCNICAS DE SUPER RESOLUÇÃO

Luiz Carlos da Silva Garcia Junior

Advisor: Waldir Sabino da Silva Júnior

Department: Postgraduate in Electrical Engineering

In this dissertation, we propose a solution using Digital Image Processing techniques to detect license plates of Brazilian vehicles, which consists in elaborating a robust and flexible method in relation to the complex textit background with the acquisition made in motion to perform the detection of Vehicle license plates. Detection occurs automatically employing PDI techniques, which include spatial domain filters and to help us, we use the opentext library OPENCV in C/C_{++} . In addition, as one of the contributions, we approached the use of the $TV - L1$ Bilateral Super Resolution algorithm in the database videos in order to contribute to the detection of the object of interest. To perform the experiments, we used a database that was designed containing videos with three different resolutions, the acquisition was performed with camera and object of interest both moving with a complex textit background, where, from the videos, we employed sequential images in the experiments performed. The results of the experiments covered two scenarios, a first scenario was not employed the Super Resolution algorithm and a second, we employed the Super Resolution algorithm, thus, it was possible to validate the proposed methodology by demonstrating the ability to detect license plates on the challenge. However, we propose that the standardization of image resolutions may influence the results.

Keywords: digital image processing, vehicle license plate detection, complex noisy background images, complex background, C/C_{++} , OPENCV.

Sumário

1	Introdução	1
1.1	Contextualização do problema	4
1.2	Contribuições	4
1.3	Objetivos	5
1.3.1	Objetivos específicos	5
1.4	Organização da dissertação	6
2	Fundamentos Teórico	8
2.1	Considerações iniciais sobre processamento de imagem	8
2.2	Conceitos sobre Filtros	10
2.3	Limiarização	14
2.4	Considerações sobre Morfologia Matemática	16
2.5	Técnicas de Super Resolução	18
2.6	Trabalhos Relacionados	19
3	Metodologia	24
3.1	Introdução	24
3.2	Super Resolução bilateral TV-L1	25
3.3	Conversão RGB para escala de cinza	27
3.4	Filtro <i>blur</i>	28
3.5	Filtro <i>Sobel</i>	30
3.6	Limiarização	31
3.7	Filtro Morfológico	32
3.8	Método de preenchimento	32

4	Base de Dados	34
4.1	Outras bases	34
4.2	Placas de veículos brasileiros	39
4.3	Base de dados	41
5	Procedimento Experimental	50
5.1	Experimentos	50
5.2	Setup dos experimentos	51
5.2.1	Combinação de experimentos	52
5.3	Métricas de desempenho	55
5.4	Análises	56
6	Considerações finais	59
6.1	Trabalhos Futuros	60
	Referências Bibliográficas	62

Lista de Figuras

2.1	Sistema de processamento digital de imagens típico e suas etapas de processamento.	9
2.2	Classificação de filtros de gradiente.	12
2.3	Filtragem da uma imagem utilizando-se um filtro <i>blur</i> Gaussiano, onde a esquerda tem-se a imagem original e a direita apresenta-se a filtragem utilizando-se um filtro <i>blur</i> com máscara de dimensões 5×5 . 14	
2.4	Limiarização <i>Otsu</i>	16
3.1	Diagrama de blocos para detecção de placas de veículos em condições ambientais adversas.	25
3.2	Diagrama de blocos correspondente as etapas do algoritmo de Super Resolução	26
3.3	Esquemático empregado no bloco de Super Resolução Bilateral <i>BTVL1</i> . 27	
3.4	Conversão de imagem no modelo de cor RGB para escala de cinza.	28
3.5	Método de emprego do filtro <i>blur</i>	29
3.6	Deteção de borda vertical através do filtro <i>Sobel</i>	30
3.7	Esquema do método <i>otsu</i>	31
3.8	Bloco para filtragem morfológica.	32
3.9	Diagrama proposto do algoritmo <i>flood fill</i>	33
4.1	A figura apresenta algumas imagens das bases de dados <i>Vision calltech</i> . 35	
4.2	A figura apresenta algumas imagens das bases de dados <i>Open ALPR benchmark</i> com placas de veículos americanas e europeias.	36
4.3	A figura apresenta algumas imagens da base de dados de Bangla concebida pelo autor Samiul Azam.	37

4.4	A figura apresenta as tomadas sequenciais de imagens traseiras dos veículos.	38
4.5	A figura apresenta as imagens traseiras de veículos oriundas da base de dados italiana.	38
4.6	(a) Auto escola (b) Aluguel (c) Particular (d) Teste (e) Colecionador (f) Representação	40
4.7	(a) Particular (b) Representação Diplomática (c) Aluguel (d) Teste (e) Representação Oficial (f) Colecionadores	40
4.8	O conjunto de imagens acima, apresenta diversos trechos da avenida Nathan Xavier	42
4.9	A figura ilustra o posicionamento do aparelho celular, dando destaque como o aparelho celular foi fixado através de um dispositivo	42
4.10	Amostra de imagens da base de dados com resolução de 426×240 pixels em condições reais no momento da aquisição.	44
4.11	Amostra de imagens da base de dados com resolução de 854×480 pixels em condições reais no momento da aquisição.	44
4.12	Amostra de imagens da base de dados com resolução de 1280×720 pixels em condições reais no momento da aquisição.	44
5.1	Diagrama de blocos utilizado para cálculo de interseção de das coordenadas manuais e automáticas	55

Lista de Tabelas

2.1	Tabela comparativa entre a metodologia proposta e trabalhos relacionados.	23
4.1	Organização e divisão dos vídeos da base de dados.	43
4.2	Divisão da base de dados em sequências de imagens	46
4.3	Tabela comparativa entre a base de dados	48
5.1	Taxas de acerto por grupo de imagens em dois cenários distintos. . .	54
5.2	Métricas de desempenho correspondente aos grupos de imagem G1 ao G5 em dois cenários distintos.	57
5.3	Métricas de desempenho correspondente aos grupos de imagem G6 ao G10 em dois cenários distintos.	57
5.4	Métricas de desempenho correspondente aos grupos de imagem G11 ao G15 em dois cenários distintos.	57
5.5	Métricas de desempenho correspondente aos grupos de imagem G16 ao G20 em dois cenários distintos.	58
5.6	Métricas de desempenho correspondente aos grupos de imagem G21 ao G25 em dois cenários distintos.	58
5.7	Métricas de desempenho correspondente aos grupos de imagem G26 ao G30 em dois cenários distintos.	58

Capítulo 1

Introdução

Processamento digital de imagens é uma área de pesquisa em evolução contínua que consiste em utilizar técnicas de processamento de imagens através de computadores com o propósito de transformar imagens, permitindo a análise ou ainda extração de características das imagens [1, 2]. Dentre as diversas aplicações, listamos as seguintes, o emprego de técnicas de processamento para analisar defeitos de placas de circuito impresso (PCI) [3], ou ainda, um sistema de detecção de uvas no ambiente [4]. Assim, as aplicações citadas, representam categorias de processamento digital de imagens e dentre essas categorias, destacamos o reconhecimento automático de placas, assim, podemos observar o interesse de diversos pesquisadores em desenvolver novas técnicas ou algoritmos para detecção de placas de veículos em condições adversas que podem ser controladas ou complexas [5–7].

Desta forma, autores Muhammad Rizwan *et al.* [5], detalha em seu trabalho de pesquisa o desenvolvimento de um sistema de detecção multinacional de placas de veículos, empregando um algoritmo específico para encontrar as bordas verticais e de acordo com os autores, utilizaram uma nova técnica de processamento, onde consiste em restringir a área de processamento encontrando as lâmpadas traseiras do veículo, para realizar um novo processamento e localizar a placa. Uma segunda pesquisa, proposta por *Ohnmar Khin* [8], onde, consiste em detectar placas de veículos em condições ambientais reais e diversas, como por exemplo, a pesquisa apresenta imagens inclinada em relação a captura por meio da câmera e para corrigir a inclinação, o autor *Ohnmar Khin* [8] utiliza um algoritmo para detectar o ângulo de inclinação, assim, este algoritmo rotaciona e realiza a correção da inclinação, além

disso, o pesquisador *Ohnmar Khin* [8], também utiliza técnicas de processamento em seu trabalho de pesquisa. A pesquisadora *Veska Georgieva* [9], propõe uma abordagem para diferentes objetos com formas retangulares em imagens reais, de forma que a proposta consiste no aprimoramento da imagem, para realizar a detecção de bordas e eliminar falsos positivos utilizando a transformada de *Hough* [9].

As pesquisas mencionadas anteriormente revelam algumas características comuns dos trabalhos, como por exemplo, em algum dos casos, as imagens das placas de veículos foram tomadas frontalmente e com *background* controlado, o equipamento de aquisição e o objeto de interesse, ambos estão estáticos, proporcionando uma condição estável para outros fatores, como iluminação uniforme, imagens com baixo índice de ruídos, imagens em boa resolução e qualidade, além disso, vale ressaltar que nestes casos, é possível visualizar a placa do veículo.

Assim, nesta dissertação, propomos uma metodologia para detecção de placas de licenciamento veicular, onde a base de dados concebida, apresenta alguns fatores que dificultam a detecção do objeto de interesse, como por exemplo, a base de dados possui um *background* complexo, além disso, o processo de aquisição realizado, foi com a câmera e objeto de interesse(placas de licenciamento veicular) em movimento, outro ponto importante é a iluminação, é perceptível a mudança de iluminação nas imagens, a base de dados também contempla imagens com resoluções distintas, imagens com pouca visibilidade do objeto de interesse e ruídos recorrente do ambiente e durante a aquisição. Deste modo, propomos um método robusto e flexível utilizando o algoritmo Super Resolução aliada a técnicas de processamento digital de imagens com capacidade de detectar placas de veículos frente as dificuldades e variáveis mencionadas.

Nossa proposta, consiste na identificação automática de placas de veículos com *background* complexo e com câmera e objeto de interesse em movimento. A metodologia empregada, inicialmente utiliza o algoritmo de Super Resolução Bilateral $TV - L1$ nas imagens da base de dados e na sequência emprega técnicas de processamento digital de imagens através da biblioteca *Opencv C/C++*, que compreende as seguintes etapas: (i) conversão em escala de cinza; (ii) remoção de ruídos com filtros *blur*; (iii) detecção de bordas verticais utilizando o filtro *Sobel*; (iv) segmentação do objeto de interesse através do método *Otsu*; (v) emprego da operação

morfológica, onde utilizamos a operação de fechamento e definimos o elemento estruturante de acordo com o formato retangular do objeto de interesse e (vi) que compreende o emprego do algoritmo *Flood Fill*.

Os pesquisadores *Muhammad Rizwan Asif* [5] e *Veska Georgieva* [9] utilizaram técnicas de processamento digitais de imagem para realizar detecção de placas de veículos, no entanto, ambos pesquisadores, utilizaram imagens com tomadas frontais, *background* controlado, com uma única resolução para todas as imagens e com excelente visibilidade do objeto de interesse. No entanto, apresentamos uma proposta de detecção de placas de veículos com *background* complexo e inúmeras variáveis que dificultam a eficiência da detecção, além disso, sugerimos um método flexível com capacidade de detectar as placas de veículos em imagens com resoluções distintas.

Inicialmente para validação da metodologia proposta, concebemos uma base de dados com algumas peculiaridades que listamos a seguir, como por exemplo, a aquisição dos vídeos foi realizada com câmera em movimento e objeto de interesse em movimento em um *background* complexo propicio aos ruídos, a base de dados contempla três resoluções diferentes, 426×240 pixels, 854×480 pixels e 1280×720 pixels, no entanto, utilizamos 1542 imagens provenientes dos vídeos adquiridos. Posteriormente, as imagens foram rotuladas manualmente e geramos o rótulo automático através do método proposto ao reconhecer as placas dos veículos, os rótulos correspondem as coordenadas dos quatro vértices da placa do veículo.

Contudo, os experimentos foram realizados em dois cenários distintos, onde, para o primeiro cenário, o método proposto contempla todas as imagens da base de dados utilizada com suas respectivas resoluções de imagens, porém, ao realizar a primeira rodada de experimentos, verificamos que os resultados não foram satisfatórios, pois observamos que não foi possível realizar a detecção de placas de veículos em imagens com a resolução mais baixa (426×240 pixels), no entanto, para um segundo cenário, empregamos o algoritmo de Super Resolução Bilateral $TV - L1$ nas imagens da base de dados, o emprego do algoritmo de SR, nos serviu como estratégia para tornar possível a detecção de placas de veículos em imagens com baixa resolução.

1.1 Contextualização do problema

Segundo Ohnmar Khin e Montri Phothisonothai [8], o reconhecimento de placas veiculares é de extrema importância e pode ser empregada em diversas aplicações considerando inúmeras variações de cenários, como por exemplo, o formato da placa do veículo, o cenário de aquisição da imagem, entre outros. Do mesmo modo, o pesquisador Muhammad Rizwan Asif [5], menciona que o *ALPR*, é uma combinação de processos que compreende a aquisição de imagens, detecção de placas, extração de caracteres e seu reconhecimento, assim, o autor ressalta a importância da execução correta do reconhecimento automático de placas de veículo para outros processos.

Da mesma forma que fora mencionada a importância do desenvolvimento de um método para detecção automática de placas de veículos, foi evidenciada a dificuldade de executar esta ação tendo em vista as condições ambientais diversas, a exemplo de mudanças de iluminação, condições de clima, objetos em movimento que não são veículos, mudança natural do dia para noite e brilho intenso dos veículos. Além disso, citamos algumas condições de captura da imagem, a exemplo da qualidade e resolução da imagem assim como posição e vibração da câmera. Dessa maneira, este conjunto de fatores tornam a detecção de placa dos veículos uma tarefa complexa.

1.2 Contribuições

Os pesquisadores Muhammad Rizwan *et al.* [5], utilizaram técnicas de processamento digital de imagens para realizar a detecção de placas de veículos e apresentaram duas contribuições, dentre elas, o autor descreve o emprego da técnica *coarse-to-fine*, em que consiste localizar as lâmpadas traseiras dos veículos, proporcionando a redução da área de processamento e por fim, utilizar as técnicas de processamento para detecção da placa de veículo em um *background* controlado. Assim, como principal contribuição desta dissertação, cita-se o emprego do algoritmo de Super Resolução Bilateral $TV - L1$, pois as imagens que correspondem ao grupo de baixa resolução, apresentam-se com baixa qualidade e pouca visibilidade da placa do veículo. A aplicação do algoritmo consiste em aumentar a resolução espacial de uma imagem ou vídeo, normalmente a partir de uma sequência de imagens com

baixa resolução. Desta forma, o algoritmo foi utilizado principalmente em vídeos com baixa resolução, assim, o emprego da Super Resolução tem o intuito de aprimorar as taxas de detecção de placas de veículos e utilizar as métricas estabelecidas.

Além disso, esta dissertação apresenta como segunda contribuição a base de dados concebida, pois as imagens das base de dados existentes e disponível, normalmente são imagens com um *background* controlado e com a qualidade da imagem bem superior, resultando em uma excelente visibilidade da placa do veículo, o que difere da proposta desta dissertação, onde, a base de dados empregada, é composta por vídeos *background* complexo e placa do veículo em movimento, assim foi empregado 1542 imagens com resoluções distintas. Além disso, a base de dados servirá de alicerce de estudo para aplicações de trabalhos futuros.

1.3 Objetivos

Estabelecer um método baseado em técnicas de processamento digital de imagens como solução para detecção de placas de veículos com o objeto de interesse em movimento em um *background* complexo.

1.3.1 Objetivos específicos

Os objetivos específicos desta dissertação são os seguintes:

- Analisar um método flexível frente as dificuldades listadas a seguir, mudanças de luz, resoluções de imagens distintas, *background* complexo e ruídos recorrente do ambiente e na aquisição para realizar a detecção de placas de veículos brasileiras;
- Conceber uma base de dados com imagens reais de veículos e câmera em movimento com resoluções diferentes;
- Analisar os resultados diante dos cenários distintos;
- Definir um modelo de pesquisa que sirva como alicerce para trabalhos futuros.

1.4 Organização da dissertação

Esta dissertação está organizada em 6 capítulos, incluindo este capítulo de introdução, a disposição dos demais capítulos está listada a seguir:

- No Capítulo 2, foi revisado os conceitos teóricos empregados nesta dissertação. Inicialmente, conceitos sobre processamentos digitais de imagens, conceito sobre filtros e suas categorias, indicando o emprego dos filtros processamento de imagens. Foi apresentado a limiarização, como técnica de segmentação usada para separar o objeto do fundo. Na sequência, foi abordado algumas considerações sobre morfologia matemática, foi mencionado as técnicas de preenchimento e enchimento através do emprego do algoritmo *flood fill*, que realiza o preenchimento de uma determinada região considerando alguns aspectos iniciais. Além disso, foi citado alguns conceitos sobre técnicas de Super Resolução e descrevemos um exemplo de uma das técnicas e no final deste capítulo apresentamos os trabalhos relacionados a esta dissertação.
- No Capítulo 3, foi apresentada a metodologia proposta para esta dissertação, inicialmente, foi descrito o algoritmo da Super Resolução bilateral $TV - L1$ empregado na metodologia proposta e um segundo algoritmo dotado de passos para realizar a detecção de placas de licenciamento veicular com fundo não controlado e foi exposto um diagrama de blocos para dar uma visão macro ao leitor sobre a nossa proposta e nas seções seguintes deste mesmo capítulo, detalhamos cada item do bloco correspondente a nossa metodologia.
- No Capítulo 4, foi dedicado a base de dados, inicialmente, foi apresentado algumas bases como exemplo e fez-se uma lista contendo algumas características das mesmas, posteriormente, mencionamos os dois tipos de placas brasileiras para veículos encontradas atualmente em território nacional. Além disso, descrevemos de forma detalhada as informações da base de dados utilizada nesta dissertação, contemplando desde o cenário de aquisição, características da base e como esta organizada e quais imagens utilizadas nos experimentos.
- No Capítulo 5, inicialmente mencionamos o setup de experimentos, assim como as ferramentas utilizadas para desenvolver o método proposto. Na sequência,

descrevemos os algoritmos utilizados, especificamos as métricas utilizadas na dissertação, assim como um diagrama de blocos do algoritmo utilizado esta tarefa. Por fim os resultados exibidos, mostram a diferença entre dois cenários, o primeiro cenário não empregamos a SR bilateral $TV - L1$ e para o segundo, contemplamos o emprego do algoritmo SR, os resultados são apresentados em gráficos, onde observamos a diferença dos resultados com o emprego do algoritmo de SR.

- No Capítulo 6, apresentamos as considerações finais trabalho e listamos possíveis trabalhos futuros.

Capítulo 2

Fundamentos Teórico

2.1 Considerações iniciais sobre processamento de imagem

Processamento digital de imagens (PDI) é uma área bastante abrangente cujo principal objetivo é processar dados de entrada, com o auxílio de um computador ou plataforma, através de um conjunto de técnicas, resultando em dados modificados de saída para análise posterior [1, 2]. Há inúmeras aplicações em PDI, como por exemplo, na área médica identifica-se lesões ou regiões do corpo atingidas por câncer [10, 11], ou em automação industrial, onde um sistema de PDI auxilia a inspeção de placas de circuito impresso ou a inspeção de soldas [1, 2]. De uma maneira geral, as aplicações de PDI são implementadas através de um sistema denominado *sistema de processamento digital de imagens*. Neste sistema, há um conjunto de etapas, conforme ilustra a Figura 2.1. Em síntese, as etapas são descritas conforme a seguinte: primeiramente, tem-se a etapa de *digitalização*, na qual uma imagem é capturada por meio de um sensor que a converte em uma imagem digital para processamento posterior. A seguir, tem-se o *pré-processamento* onde utiliza-se técnicas de pré-processamento para, por exemplo, realce ou remoção de ruídos da imagem. Na sequência, tem-se a etapa de *segmentação* onde fragmenta-se a imagem em regiões de interesse. A etapa *representação* é empregada para armazenar e modificar os objetos de interesse extraídos de uma imagem. Por fim tem-se a *interpretação* que é responsável por atribuir significado aos objetos reconhecidos [1, 2].

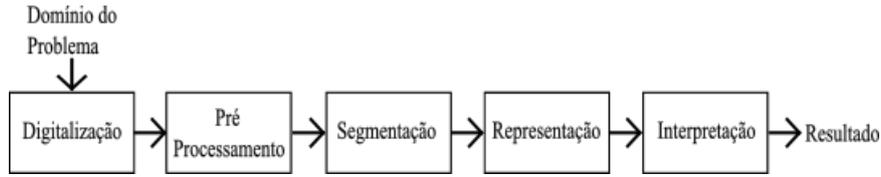


Figura 2.1: Sistema de processamento digital de imagens típico e suas etapas de processamento.

Uma imagem digital é definida por uma função $f(n_1, n_2)$, onde n_1 e n_2 são coordenadas espaciais e a amplitude de f é denominada intensidade luminosa ou nível de cinza, a função f é representada por uma matriz bidimensional onde cada elemento representa um pixel da imagem. Assim, cada pixel possui valor e localização específica [1, 12, 13]. Pode-se categorizar as imagens em: colorida, escala de cinza e binária [14, 15]. As imagens coloridas possuem três canais de oito bits cada, as imagens em escala de cinza tem um canal de oito bits e as binárias tem um canal de um bit.

A formação de cores é composta por dois processos, o aditivo, que gera cores incrementando luz e o subtrativo, que emprega pigmentos para bloquear a luz [16]. Há diferentes modelos de cores, por exemplo: (i) o modelo *RGB* que é composto por canais vermelho, verde e azul correspondentes ao sistema de coordenadas cartesianas expressada em um cubo [17, 18]; (ii) o modelo *HSV* que é determinado por parâmetros matiz (H), saturação (S) e luminância (L), que é representado por uma pirâmide hexagonal invertida, onde os seis vértices hexagonais são as localizações das cores vermelho, verde, ciano, azul, amarelo e magenta [17] e por fim; (iii) o modelo YC_bC_r , onde Y representa a formação de luminância, C_b e C_r guarda informação de crominância, este modelo é frequentemente usado em transmissões de vídeo [18, 19]. Em processamento de imagens, são comuns as conversões de imagens em espaço de cores distintos. Por exemplo, tem-se a conversão *RGB* para escala de cinza, que é efetuada através de uma combinação linear das intensidades dos canais *RGB* ponderada três coeficiente c_R , c_G e c_B , para cada canal, respectivamente, conforme a seguir:

$$y(n_1, n_2) = c_R i_R(n_1, n_2) + c_G i_G(n_1, n_2) + c_B i_B(n_1, n_2) \quad (2.1)$$

A equação (2.1) [15, 20] proporciona a conversão de uma imagem $i(n_1, n_2) = \{i_R, i_G, i_B\}$, composta por três matrizes que representam os canais *RGB*, em uma

imagem $y(n_1, n_2)$ em escala de cinza. Os coeficientes são escalas e os mais comuns são $c_R = 0,299$, $c_G = 0,587$ e $c_B = 0,114$ [18, 19]. Uma segunda alternativa é a conversão de imagens HSV para RGB , segundo a equação a seguir:

$$y_H(n_1, n_2) = \begin{cases} 60^\circ \frac{i_G(n_1, n_2) - i_B(n_1, n_2)}{max - min}, & \text{se } max = i_R(n_1, n_2) \\ 60^\circ \frac{i_B(n_1, n_2) - i_R(n_1, n_2)}{max - min} + 120^\circ, & \text{se } max = i_G(n_1, n_2) \\ 60^\circ \frac{i_G(n_1, n_2) - i_G(n_1, n_2)}{max - min} + 240^\circ, & \text{se } max = i_B(n_1, n_2) \end{cases} \quad (2.2)$$

$$y_S(n_1, n_2) = \frac{max - min}{max}, \text{ se } max \neq 0 \quad (2.3)$$

$$y_V(n_1, n_2) = max \quad (2.4)$$

Semelhante a equação (2.1) mencionada anteriormente, as equações (2.2) até (2.4), permitem converter uma imagem colorida $i = \{i_R, i_G, i_B\}$ para $y = \{y_H, y_S, y_V\}$, onde os valores das matrizes i_R , i_G , i_B variam de 0 a 255, por comodidade, utiliza-se valores normalizados entre em 0 e 1. Os valores das variáveis max e min , nas equações (2.2) até (2.4), referem-se, respectivamente, ao valor máximo e mínimo das matrizes i_R , i_G , i_B . O matiz y_H varia de 0° a 360° , pois cada vértice da pirâmide hexagonal do modelo de cor HSV , possui ângulo de 60° . Desta forma o ângulo vermelho vale 0° , amarelo vale 60° , verde vale 120° e assim por diante. A saturação y_S e luminância y_V estão normalizados em 0 e 1 [1].

2.2 Conceitos sobre Filtros

Considerando-se sinais genericamente, pode-se dizer que os filtros são sistemas que modificam sinais de entrada e produzem sinais de saída. Supondo x e y , respectivamente, como sinais de entrada e saída, e um filtro h , normalmente usa-se a representação $y = h\{x\}$ para o sinal de saída y . Pode-se modificar o sinal x em amplitude, fase e frequência. Normalmente, os filtros são utilizados para remoção e/ou passagem de certas frequências. Adicionalmente, o termo *filtragem* refere-se a operação entre x e h resultando no sinal de saída representado por y [21, 22]. Os filtros ideais são classificados em: (i) passa baixa, que atenua altas frequências, dando suavização ao sinal x ; (ii) passa alta, que atenua as baixas frequências e permite a passagem de altas frequências. O efeito deste filtro consiste em tornar

transições mais nítidas do sinal x ; (iii) passa banda, que permite a passagem de um intervalo de frequências e por fim; (iv) rejeita banda, que rejeita um intervalo de frequências [21–23]. A seguir, comentaremos sobre filtros espaciais e filtros no domínio da frequência.

Outro conceito comum, quando utiliza-se filtros para imagens, refere-se aos *filtros no domínio espacial* (ou *filtros espaciais*) que atuam sobre o plano da imagem, processando diretamente o conjunto de pixels que formam a imagem. Normalmente, esta operação utiliza máscaras, que são sub matrizes da imagem. Supondo uma imagem $f(n_1, n_2)$ e uma máscara $w(m_1, m_2)$, a filtragem consiste na substituição do valor do pixel em uma posição (n_1, n_2) por um novo valor que depende dos pixels vizinhos. Isto é feito deslizando-se a máscara w por toda a imagem f . Este procedimento denomina-se *correlação* ou *convolução*. A diferença entre ambos é a rotação do filtro em 180° [24–26]. Há algumas categorias comuns de filtros espaciais, são eles: (i) filtros de suavização, utilizados para desfoque e redução de ruídos na imagem [24]; (ii) filtros de realce, que realizam a redistribuição das luminâncias para acentuar regiões de interesse e; (iii) filtros de afinamento que proporcionam o realce fino em imagens que foram desfocadas. Nesta categoria destacamos dois tipos, filtros laplacianos aplicados a realce de bordas em imagem baseando-se na segunda derivada [24, 27] e filtros de gradiente que identificam regiões de interesse através do cálculo da primeira derivada [24, 28]. Na Figura (2.2) apresenta-se outras categorias de filtros de gradiente, também denominados por operadores gradientes. Geralmente, esses filtros são empregados para detecção de borda, que é uma das etapas aplicadas a análise de imagens. Por exemplo, o filtro (operador) *Robert*, de fácil aplicação, no qual permite realçar bordas verticais e horizontais por meio de máscaras 2×2 através do cálculo das primeiras derivadas. Há outros filtros (como *Prewitt* e *Sobel*) com características similares [29–31].

Os *filtros no domínio da frequência* são comumente utilizados como filtros em imagens. Nestes casos, emprega-se a *transformada de Fourier* como base de cálculo. Matematicamente, supondo que $f_h(n_1, n_2) = f(n_1, n_2) * h(n_1, n_2)$, onde f é a imagem de entrada (ou original), h é o filtro, f_h é a imagem filtrada e $*$ representa a convolução, então, a representação na frequência resulta em $F_H(k_1, k_2) = F(k_1, k_2)H(k_1, k_2)$. Especificamente, a representação na frequência é feita utilizando-se a Transformada

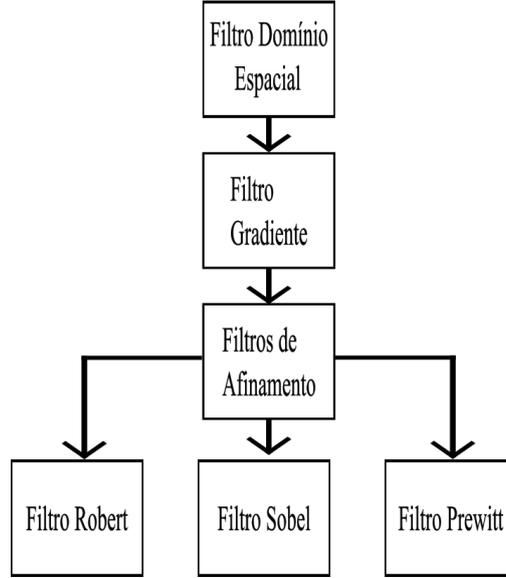


Figura 2.2: Classificação de filtros de gradiente.

Discreta de *Fourier* bidimensional (DFT-2D) direta e inversa [32]. Assim, f_h , f e h são representados por F_H , F e H , e vice versa, utilizando-se as expressões abaixo:

$$X(k_1, k_2) = \sum_{n_1=0}^{M-1} \sum_{n_2=0}^{N-1} x(n_1, n_2) \exp \left[j2\pi \left(\frac{k_1 n_1}{M} + \frac{k_2 n_2}{N} \right) \right] \quad (2.5)$$

$$x(n_1, n_2) = \frac{1}{MN} \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{N-1} X(k_1, k_2) \exp \left[j2\pi \left(\frac{k_1 n_1}{M} + \frac{k_2 n_2}{N} \right) \right] \quad (2.6)$$

onde, nas equações (2.5) e (2.6), X é a representação na frequência e x é a representação espacial. As componentes de frequência horizontal e vertical são representadas por k_1 e k_2 respectivamente. Há alguns filtros no domínio da frequência bastante conhecidos, como por exemplo, os filtros Butterworth, que tem resposta em frequência plana. Neste caso, há dois tipos, os filtros Butterworth passa baixa, que normalmente são usados para remoção de ruídos de alta frequência e os filtros Butterworth passa-alta, utilizados para realce [32]. Por fim, um outro exemplo é o filtro passa baixa *Gaussiano*, com características de largura de banda estreita [33–35].

Os filtros *blur*, também denominados por filtros de borramento (ou borrão), consistem em borrar ou desfocar uma imagem f de forma intencional. Este filtro é implementado utilizando-se diversas funções de desfoque (ou *blurring functions*), geralmente denominadas por h . A filtragem, utilizando-se filtros *blur*, tem o objetivo

de reduzir os ruídos gerados durante a captura da imagem ou ruídos de ambiente¹. Assim, o desfoque atenua as altas frequências e não compromete a compreensão dos objetos contidos na imagem [26, 36]. Há algumas funções de desfoque comuns, como por exemplo: *motion blur*, *gaussian blur* e o *defocus blur* [37, 38]. Para apresentar as funções de desfoque, supõe-se que a imagem resultante desfocada, em escala de cinza, é dada por

$$f_h(n_1, n_2) = f(n_1, n_2) * h(n_1, n_2) + b(n_1, n_2), \quad (2.7)$$

onde f é a imagem original em escala de cinza, h é a função de desfoque, b é o ruído aditivo e f_h é a saída, ou seja, a imagem filtrada. Considerando-se a equação (2.7) e que não há ruído aditivo, então, a função de desfoque para o filtro *motion blur* é definida por meio da equação abaixo [37, 38]:

$$h(n_1, n_2) = \begin{cases} \frac{1}{L}, \text{ se } 0 \leq |n_1| \leq L \cos(\theta) \\ n_2 = L \sin(\theta) \end{cases} \quad (2.8)$$

onde, considerando-se uma imagem f , a equação (2.8) apresenta o ângulo θ que é o movimento relativo entre a câmera e o objeto e variando de 0° a 360° nas dimensões n_1 e n_2 . A variável L , representa o comprimento do desfoque, ambas variáveis atuam diretamente na qualidade da imagem e são escolhidos arbitrariamente.

Uma segunda função de desfoque muito comum e amplamente utilizada em imagens com características diversas é o *Gaussian blur*, pois este realiza o desfoque da imagem através da função *Gaussiana*, que é empregada como um filtro passa baixa proporcionando o borramento na imagem resultando na remoção de ruídos. A função de desfoque *Gaussian blur*, é expressa conforme a seguir [38–40].

$$h(n_1, n_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{n_1^2 + n_2^2}{2\sigma^2}\right) \quad (2.9)$$

onde, na equação (2.9), n_1 é a distância no eixo horizontal, n_2 é a distância no eixo vertical e σ é o desvio padrão e representa a largura do filtro, onde o grau de desfoque está relacionado com este parâmetro, de forma que quanto maior o valor

¹Quando uma imagem foi adquirida através de um meio turbulento, como por exemplo imagens adquiridas em neblinas ou ambientes aquáticos.



Figura 2.3: Filtragem de uma imagem utilizando-se um filtro *blur* Gaussiano, onde a esquerda tem-se a imagem original e a direita apresenta-se a filtragem utilizando-se um filtro *blur* com máscara de dimensões 5×5 .

de σ , maior é a largura do filtro e maior é o grau de desfocagem. Na Figura 2.3, apresenta-se um exemplo de filtragem utilizando-se um filtro *Gaussian blur*. Por fim, o filtro *defocus blur*, que geralmente é utilizado quando uma parte da imagem está em foco e outra não [38, 41].

2.3 Limiarização

A *limiarização* de imagens, também denominada por *thresholding*, é uma das operações de segmentação mais simples e utilizadas, sendo fundamental na identificação de objetos dividindo uma imagem em escala de cinza em duas regiões de interesse: objeto e *background* (ou fundo), resultando em uma imagem dividida/separada por dois níveis de cinza. Quando utiliza-se a limiarização em processamento digital de imagens (PDI) a maior dificuldade é encontrar um valor limiar adequado para cada imagem. Há diversas técnicas de limiarização, a seguir citaremos algumas. Primeiramente, a *limiarização global*, que baseia-se na seleção de um valor de limiar único para toda a imagem, onde a intensidade dos pixels com níveis de cinza são iguais em toda a imagem, permanecendo *background* e objeto, com níveis de cinza distintos. Supondo uma imagem digital em escala de cinza f , a imagem limiarizada f_l será obtida através da seguinte expressão

$$f_l(n_1, n_2) = \begin{cases} 0, & \text{se } f(n_1, n_2) \leq T \\ 1, & \text{se } f(n_1, n_2) > T \end{cases} \quad (2.10)$$

onde, na equação (2.10), T representa o limiar escolhido, 0 representa o *background* e 1 o objeto. Em f_l , será considerado *background* de f , quando o nível de cinza nas coordenadas da imagem f for menor ou igual a T , e será considerado objeto descrito em f , quando o nível de cinza nas coordenadas da imagem f for maior que T . Existem três subcategorias com respeito a limiarização global, são elas: (i) limiarização fixa; (ii) limiarização semi-adaptativa e (iii) limiarização adaptativa [42, 43]. A técnica mais complexa é a limiarização adaptativa, descrita conforme a seguir: a limiarização adaptativa geralmente é empregada em imagens com diferentes condições de iluminação, como por exemplo, regiões com sombra ou iluminação excessiva ou artificial. A imagem limiarizada é determinada, normalmente, considerando-se sub-regiões que são percorridas por uma janela deslizante retangular [44, 45]. Para esta janela, computa-se a saída segundo alguma expressão. Por exemplo, a limiarização adaptativa conhecida por *Niblack* utiliza-se a expressão a seguir [45]

$$f_l(n_1, n_2) = \begin{cases} 1, & \text{se } f(n_1, n_2) > m(n_1, n_2) + k\bar{\sigma}(n_1, n_2) \\ 0, & \text{caso contrário} \end{cases} \quad (2.11)$$

onde, na equação(2.11), descreve-se a limiarização adaptativa calculada dentro da janela. Considera-se $f_l(n_1, n_2)$ a imagem limiarizada, $m(n_1, n_2)$ é a média local, k é um valor constante definida como $-0,2$, σ representa o desvio padrão dos pixels localizados na janela, assim k e σ ajustam o valor do limiar de acordo com a vizinhança [45]. Outro exemplo, denominado por método *Otsu*, é muito utilizado para limiarização. Neste, os pixels da imagem não são processados diretamente. O processamento é realizado utilizando-se o histograma da imagem [1], o qual permite utilizar um paradigma de segmentação em duas categorias (ou classes) conforme o formato (distribuição), que geralmente é bimodal, dos histograma [46–48]. Especificamente, o limiar adaptativo é determinado da seguinte maneira:

$$\sigma(T) = N_B(T)N_O(T)[\mu_B(T) - \mu_O(T)] \quad (2.12)$$

onde, na equação (2.12) descreve o método *Otsu*, N_B e N_O , correspondem aos pesos das classes de fundo e de objeto, μ_B e μ_O , correspondem as médias das classes. A *variância* total, consiste em separar os pixels de cada classe, para determinar a

média de cada classe. A seguir, realiza a diferença entre as médias das classes fundo e objeto e realiza a multiplicação entre os peso das classes. A Figura 2.4 abaixo, apresenta o resultado da limiarização realizada por meio do método *otsu*.



Figura 2.4: Limiarização *Otsu*.

2.4 Considerações sobre Morfologia Matemática

A *morfologia matemática*, que compreende uma importante área de processamento digital de imagens (PDI), baseia-se no processamento considerando-se as formas ou estruturas de objetos contidos na imagem. Neste caso, trata-se do mesmo conceito de objeto utilizado anteriormente. As técnicas fundamentam-se no uso de elementos estruturantes, que compreendem estruturas semelhantes as máscaras utilizadas em filtragem. Normalmente, a morfologia é formulada através da teoria dos conjuntos, na qual os conjuntos representam a forma dos objetos em uma imagem [13, 49–51], porém há outras formulações. Em morfologia, há quatro operações básicas: (i) erosão, (ii) dilatação, (iii) abertura e (iv) fechamento [13, 49, 50] [51]. A erosão, que consiste na redução de um objeto na imagem, removendo detalhes menores que o SE e pontos isolados, é definida conforme a expressão a seguir:

$$[\varepsilon_B(f)](n_1, n_2) = \min_{b \in B} f(n_1 + b, n_2 + b) \quad (2.13)$$

onde, na equação (2.13), $\varepsilon_B(f)$ é a erosão de uma imagem $f(n_1, n_2)$ por um elemento estruturante B . Assim, a saída, ou seja, a imagem erodida, no pixel n_1, n_2 , é o valor mínimo da imagem na janela definida pelo elemento estruturante [51]. A dilatação amplia as regiões de um objeto na imagem, preenchendo lacunas existentes, de

forma que a expansão do objeto é controlada através do formato do SE, é expressa da seguinte maneira:

$$[\delta_B(f)](n_1, n_2) = \max_{b \in B} f(n_1 + b, n_2 + b) \quad (2.14)$$

onde, na equação (2.14), δ_B apresenta a dilatação de uma imagem $f(n_1, n_2)$ por um elemento estruturante B [51]. Assim, a saída, ou seja, a imagem dilatada, no pixel n_1, n_2 , é o valor máximo da imagem na janela definida pelo elemento estruturante [51].

Por fim, existem duas outras operações morfológicas, a abertura e fechamento. A operação de abertura é caracterizada por suavizar o contorno de objetos em uma imagem e também por eliminar saliências finas e pontos isolados. Esta operação resulta da combinação de entre a erosão e dilatação, da seguinte maneira:

$$\gamma_B(f) = \delta_{\check{B}}[\varepsilon_B(f)] \quad (2.15)$$

onde, na equação (2.15), γ_B é a abertura de uma imagem f por um elemento estruturante B . Assim, tem-se a erosão de f por B , seguido pela operação de dilatação com o elemento estruturante refletido, representado por \check{B} . De modo semelhante, o fechamento é definido pela combinação de duas outras operações a dilatação e erosão e consiste em fundir discontinuidades estreitas entre os objetos na imagem alongando formas finas. Assim, o fechamento suaviza o contorno dos objetos em uma imagem. A operação de fechamento é definida na expressão a seguir:

$$\phi_B(f) = \varepsilon_{\check{B}}[\delta_B(f)] \quad (2.16)$$

onde, na equação (2.16), apresenta a operação de fechamento de uma imagem f por um elemento estruturante B que é determinada pela operação de dilatação de f seguido pela operação de erosão com elemento estruturante invertido, representado por \check{B} [51].

2.5 Técnicas de Super Resolução

As técnicas de Super Resolução (SR) compreendem métodos utilizados para aumentar resolução espacial de uma imagem única ou vídeo, adicionando detalhes contidos em uma imagem que está em baixa resolução. Normalmente, os métodos de super resolução combinam informações de várias imagens tiradas da mesma cena para representar detalhes que não foram capturados inicialmente nas imagens originais, pois, durante o processo de aquisição os detalhes da imagem são integrados por diferentes sensores, de modo que cada pixel em cada aquisição possua informações diferentes. Assim, a super resolução busca encontrar a relação entre diferentes capturas que obtiveram detalhes distintos na imagem para então, criar uma nova imagem com mais informações e resultando portanto em uma imagem com resolução mais alta. Normalmente os métodos de super resolução existentes podem ser obtidos através do domínios espacial e técnicas baseadas a análise da frequências de espectro empregando transformada de *Fourier* ou *Wavelet*. Existem diversos método de super-resolução com abordagens distintas na literatura, no entanto mencionaremos apenas uma única técnica listada a seguir: (i) *Iterative Back Projections* [52–54].

O método *Iterative Back Projection* (IBP) pode ser utilizado tanto para imagem única ou pode ser empregado para múltiplas imagens de entrada. Inicialmente, o método começa com uma estimativa inicial da imagem de alta resolução interpolando com uma imagem de baixa resolução de entrada, posteriormente, múltiplas imagens de baixa resolução simuladas geram uma imagem de super resolução reconstruída através do modelo de aquisição de imagens. Assim, caso haja correspondência entre a imagens de super resolução de entrada e imagens de baixa resolução simuladas, as imagens de super resolução reconstruída será igual à imagem de alta resolução original, do contrário, as diferenças entre as duas imagens, imagens de baixa resolução e imagens de baixa resolução de entrada, ambas serão projetadas para imagem de super resolução. Assim, este processo é repetido iterativamente até que as diferenças entre as imagens sejam mínimas suficiente. O método *Iterative*

Back Projection é definido na expressão a seguir [53, 54]:

$$g_{mL}^i = VUW_k f(n_1, n_2)_H^i, \quad (2.17)$$

$$f(n_1, n_2)_H^{i+1} = f(n_1, n_2)_H^i - \lambda \sum_{m=1}^P H_m^{BP} (g_{mL}^i - g_{mL}) \quad (2.18)$$

onde, na equação (2.18), é responsável por calcular a imagem em baixa resolução simulada definida pelo modelo de aquisição da imagem, na equação (2.18), i representa a quantidade de iterações, $f(n_1, n_2)_H^{i+1}$ corresponde a imagem de super resolução reconstruída durante na primeira iteração de $i + 1$, já $f(n_1, n_2)_H^i$ corresponde a imagem reconstruída na iteração, λ é definido como uma constante de valor 0, 1, P corresponde ao número de imagens de baixa resolução de entrada H_m^{BP} corresponde a operação de retroprojeção, as variáveis g_{mL}^i e g_{mL} ambas representam a imagem de baixa resolução simulada e imagem de baixa resolução de entrada respectivamente [53, 54].

2.6 Trabalhos Relacionados

Os autores Muhammad Rizwan *et al.* [5], propuseram um algoritmo de detecção de placas multinacional. O sistema proposto independe do modelo do veículo e baseai-se na localização de suas lâmpadas traseiras. Para isto, utiliza-se o *heuristic energy map*² para detectar a borda vertical na região de interesse, para reduzir a área de processamento, e detectar a região onde há presença de placa descartando o fundo complexo. Os autores tratam esta técnica como uma contribuição e utilizam a denominação *coarse-to-fine*. Outra contribuição emprega uma nova abordagem de histograma. Em termos da metodologia proposta pelos autores, primeiramente: utiliza-se a conversão de modelos de cores, neste caso modelo RGB para o YUV, para destacar as lâmpadas traseiras. Após, aplica-se um *box filter*, que consistem em um filtro de janela $N \times N$ para unir os componentes da lâmpada traseira disjuntas antes da limiarização. Para limiarização, foi empregado o limiar adaptativo, pois as regiões da imagem possuem diferentes condições de iluminação. Ao final deste processo, tem-se a região de interesse da imagem RGB original. A seguir, a

²Heurísticas são processos cognitivos empregados em tomadas de decisões, tendo como estratégias a escolha de informações mais facies e rápidas.

região de interesse é extraída para um novo processamento, onde, para esta nova etapa converte-se a imagem extraída, que é RGB, para uma imagem em escala de cinza. Após, emprega-se um filtro *Sobel* para detecção de bordas e verifica-se os componentes conectados. A última etapa do consiste a aplicação da nova abordagem do histograma, que atua sobre a região da placa analisando a densidade das bordas verticais e horizontais da placa. Os experimentos realizados utilizaram uma base de dados de 855 imagens coloridas de várias nacionalidades e com resolução de 600×800 pixels [5].

A pesquisa proposta por Runmin Wang e Nong Sang [6] apresentaram um sistema de detecção de placas chinesas com capacidade de detectar diferentes modelos de placas, que variam em tamanho, forma e cores. A proposta é baseada no cálculo de gradiente e um classificador em cascata. Os pesquisadores dividiram a metodologia em três módulos principais: *pré-processamento*, que consiste na conversão de modelos de cores de RGB para escala de cinza, em seguida usa-se um filtro de mediana para remoção de ruídos e, por último, emprega-se o gradiente horizontal e vertical, para detecção de placas. O classificador em cascata *Adaboost* [6] foi empregado nesta pesquisa com o objetivo de detectar os verdadeiros e falsos positivos. Uma outra característica deste classificador é que ele possui 12 classificadores forte, 103 operações para o gradiente e mais 132 operações para imagem em escala de cinza. Para a próxima etapa, utiliza-se uma estratégia heurística baseada em um classificador *SVM* [6], responsável por remover falsos positivos através da varredura nas regiões onde há placa. Após a remoção de falsos positivos, a próxima etapa do processamento consiste na utilização do filtro *Sobel* para detecção de bordas, o método *Otsu* binariza a imagem para o uso da operação morfológica de abertura e então é realizado a análise de componentes conectados, que considera a quantidade de caracteres na região da placa. O algoritmo *SIFT* [6], confirma presença de placa através da análise de textura. Para os experimentos utiliza-se uma base de dados com 4087 imagens no total e a resolução é igual a 1902×1088 [6].

Os pesquisadores Alyan Nezhadi *et al.* [7], mencionam que os métodos tradicionais de detecção de placas de veículos baseados em detectores de borda, não apresentam resultados satisfatórios. Para solucionar o reconhecimento automático de placas, os pesquisadores utilizam duas ferramentas após o pré processamento e

detecção de bordas, o filtro *Gaussiano* e a rede *bayssiana*. Assim, os experimentos realizados pelos pesquisadores, confirmaram a eficácia do método de detecção de placas em condições de iluminações adversas e fundos complexo. Os pesquisadores dividiram o método proposto em pré-processamento, onde o primeiro passo é a conversão de imagem *RGB* em escala de cinza. A seguir, aplica-se a equalização do histograma, pois a imagem do veículo é de baixa qualidade e a equalização do histograma proporciona o aumento do contraste na imagem. Na próxima etapa, emprega-se o cálculo de intensidade dos níveis de cinza para obter a região da placa da imagem através da densidade local das bordas. Após, emprega-se um filtro *Sobel* modificado, que identifica arestas verticais na imagem e um filtro *gaussiano* 2D. Posteriormente, realiza-se a limiarização e abertura morfológica da imagem. O último módulo se concentra na rede *bayesiana* com o objetivo de confirmar a região da placa. Para os experimentos, os pesquisadores utilizaram uma base de dados com 400 imagens com resolução de 240×320 pixels [7].

Os pesquisadores Ohnmar Khin e Montri Phothisonothai [8] propuseram um método para detectar a região de diferentes tipos de placas de veículos de *Mianmar*, capturadas a partir de condições ambientais diversas, baseadas na dilatação horizontal e vertical, detecção de ângulo de inclinação e caixa delimitadora. A metodologia proposta consiste na conversão de imagem RGB em escala de cinza, para realizar a limiarização da imagem através do método *Otsu*. A seguir, determina-se o histograma e aplica-se a dilatação morfológica para localizar a região da placa. A próxima etapa consiste na estimativa das regiões binárias através do método de rotulagem, onde o autor utiliza o conceito de conectividade para detecção de placas. A seguir, tem-se a detecção do ângulo e por fim a caixa delimitadora que especifica os objetos detectados na imagem extraindo várias posições da placa. Os pesquisadores contaram com mais de 30 imagens com resolução 4288×2848 pixels com circunstâncias diversas e obtiveram uma taxa de acerto de 99% [8]. Os trabalhos relacionados acima, mencionam as metodologias empregadas, da mesma forma, a proposta de Muhammad Rizwan *et al.* [5] consiste em detectar placas de veículos multinacionais em imagens com resolução de 600×800 pixels utilizando uma de suas contribuições baseando-se na localização das lâmpadas traseiras do veículo para reduzir área de processamento e posteriormente utiliza técnicas de processamento em uma região

específica da imagem que pode conter uma placa, o que difere da proposta desta dissertação que utiliza toda o plano espacial da imagem para realizar as técnicas de processamento e detectar as placas de veículos.

Em uma segunda pesquisa proposta por Runmin Wang e Nong Sang [6], propõe um método baseado no cálculo de gradiente e um classificador em cascata com capacidade de detectar diferentes tipos de placas chinesas, nesta proposta o pesquisador utiliza o gradiente horizontal e vertical para detecção de placas de veículos e as imagens com resolução de 1902×1088 pixels, no entanto, nossa proposta utiliza apenas o filtro *Sobel* para detectar as bordas verticais do objeto de interesse.

A proposta de Alyan Nezhadi *et al.* [7] menciona a ineficiência de métodos tradicionais de detecção de placas e emprega o filtro *Gaussiano* e a rede *bayssiana* após a detecção de bordas, desta forma, o pesquisador assegura uma taxa de detecção aceitável, porém, a metodologia apresentada nesta dissertação utiliza filtros *blur* para realizar a remoção de ruídos e *Sobel* para detectar apenas as bordas verticais da placa do veículo, pois a densidade de bordas verticais é maior que a densidade de bordas horizontais devido aos caracteres presente na placa dos veículos.

Um outro trabalho de pesquisa proposto por Ohnmar Khin e Montri Phothisonothai [8], utiliza como um dos processos principais, um algoritmo para correção de ângulo de inclinação da placa de veículo após a detecção da placa, o método proposto nesta dissertação não utiliza nenhuma correção de ângulo para realizar a detecção das placas de veículos. A metodologia presente nesta dissertação, utiliza técnicas de processamento digital de imagens através da biblioteca *opencv* em imagens com *background* complexo contribuindo para um alto índice de ruídos, além disso, a aquisição das imagens foi realizada com a câmera e veículo em movimento e em três resoluções distintas.

Contudo, esta dissertação lida com imagens com *background* complexo considerando que, a câmera e o veículo estão em movimento, adicionando a esses dois fatores, tem-se as condições do cenário de aquisição, como por exemplo, o fator iluminação, sombra ocasionada pelas copas das árvores, ruas com ondulações e buracos. Assim, esta dissertação tem o objetivo de contribuir com uma solução com capacidade de realizar a detecção de placas de veículos de acordo com os fatores mencionados.

Tabela 2.1: Tabela comparativa entre a metodologia proposta e trabalhos relacionados.

Trab.	Conver. de cores	Detec. de Borda	Limiariz.	Resolução	DB
[45]	RGB to YUV	Filtro Sobel	Adaptativa	600 × 800	BnC, CE
[46]	RGB to Gray	Filtro Sobel	Adaptativa	1902 × 1088	BnC, CE
[47]	Rgb to Gray	Filtro Sobel	-	240x320	BC
[48]	Rgb to Gray	Hist.	Otsu	4288x2848	BC, CE, PI
Prop.	Rgb to Gray	Filtro Sobel	Otsu	1280 × 720 854 × 480 426 × 240	BC, CnE, InU, PI

BC: *background* complexo

BnC: *background* não compl.

CE: câmara estática

CnE: câmara dinâmica

IU: iluminação uniforme

InU: iluminação não unif.

PF: placas frontal

PI: placas inclinada

DB: Data Base

Capítulo 3

Metodologia

3.1 Introdução

Nesta dissertação, apresenta-se uma metodologia, concebida através de um diagrama em blocos, para detecção de placas de licenciamento veicular. Nesta abordagem, foi utilizado imagens reais com condições ambientais adversas, como por exemplo, iluminação não uniforme, *background* complexo e não controlado, onde o veículo e a câmera de aquisição estão em movimento. Na Figura 3.1, apresenta-se o diagrama em blocos proposto para esta dissertação. Em uma breve descrição, primeiramente, utilizamos o método de super resolução bilateral em todos os vídeos da base de dados que contemplam as seguintes resoluções : 1280×720 , 854×480 e 426×240 pixels, o método de super resolução empregado utiliza alguns parâmetros específicos da junto com o algoritmo de fluxo óptico de *Farneback*, posteriormente, realizamos a extração dos *frames* de vídeos de curta duração concebido para construção da base de dados, resultando em uma sequência de imagens correspondente ao respectivo trecho do vídeo da base de dados, onde, esta sequência de imagens é carregada de forma automática no próximo bloco que é responsável por realizar a conversão de imagem do modelo de cor RGB para escala cinza, a seguir, efetua-se a remoção de ruídos através do filtros *blur*, através do emprego de uma função de desfoque *gaussiano*. Após a remoção de ruídos, a imagem está adequada para o emprego do filtro *Sobel* que tem a função de detectar as bordas apenas na direção vertical, utilizando uma única máscara na mesma direção. A seguir, no próximo bloco, realiza-se uma tentativa de segmentar o objeto de interesse, que consiste na

placa de licenciamento veicular, do *background* através do método *Otsu*. Na sequência, emprega-se a operação morfológica de fechamento para detectar as regiões do objeto de interesse e por fim, o utiliza-se o algoritmo *flood fill* para preencher as regiões no formato do objeto de interesse. Nas seções seguintes, aborda-se os blocos com maiores detalhes.

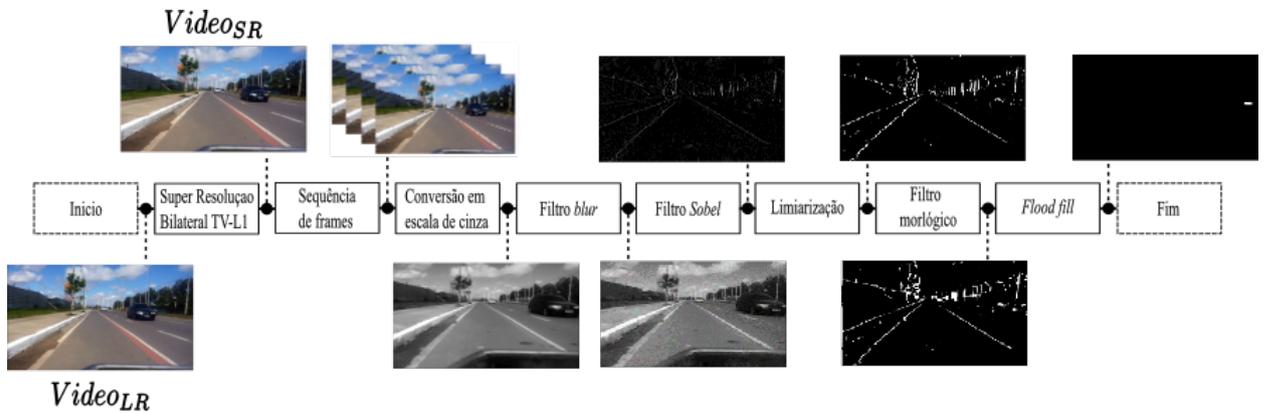


Figura 3.1: Diagrama de blocos para detecção de placas de veículos em condições ambientais adversas.

3.2 Super Resolução bilateral TV-L1

Conforme fora mencionado anteriormente, as técnicas de SR aumentam a resolução espacial de imagens ou vídeos tornando alguns detalhes das imagens mais evidentes através da combinação de informações das sequências de imagens tiradas da mesma cena. Assim, neste bloco empregamos o algoritmo de super resolução Bilateral $TV - L1$ ($BTVL1$) nos vídeos da base de dados utilizados nesta dissertação. O emprego deste bloco, consiste em favorecer detecção do objeto de interesse (placa de licenciamento veicular ou simplesmente placa do veículo), principalmente em imagens com resolução de 426×240 pixels, pois, observamos que a detecção do objeto de interesse era inviável em imagens com baixa resolução, no entanto, ao aplicar o bloco de super resolução, foi constatado que a região de interesse das imagens com resoluções inferiores tornaram-se mais evidentes, proporcionando a detecção da placa de licenciamento veicular após o emprego das técnicas de processamento de imagens.

Assim, a utilização do bloco de SR, resultou em vídeos com resolução de 1920×1080 pixels. Por comodidade, dividimos este bloco SR Bilateral $TV - L1$

(*BTVL1*) em três etapas, conforme ilustramos a seguir:

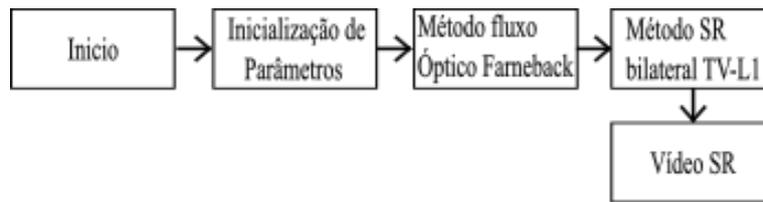


Figura 3.2: Diagrama de blocos correspondente às etapas do algoritmo de Super Resolução

Na Figura 3.2, o primeiro bloco consiste em definir alguns parâmetros iniciais relacionados ao algoritmo de SR, como por exemplo, o valor do fator de escala igual a 4, pois, este fator implica no número de pirâmides de imagens, configuramos a velocidade de reprodução do vídeo SR, na sequência definimos o número de iterações do algoritmo faz em cada nível da pirâmide.

O segundo bloco defini-se o fluxo óptico empregado nesta metodologia. Normalmente as técnicas de fluxo óptico utilizadas representam um padrão de movimento aparente de um objeto, as superfícies e arestas de uma cena visual são resultados causados por movimento relativo entre uma câmera e uma cena. Desta forma os conceito das técnicas de fluxo óptico na visão computacional, está associado as técnicas de detecção de movimento, segmentação de objetos e informações sobre tempo de controle. Assim, foi empregado o *Gunnar-Farneback* como algoritmo de fluxo óptico, pois neste caso particular, a escolha do algoritmo define o padrão de movimento do objeto de interesse nas imagens considerando dois quadros consecutivos de movimento do objeto ou da câmera. Além disso, o algoritmo *Gunnar-Farneback*, foi desenvolvido para produzir densidade de de fluxo óptico e os passos do algoritmo, inicialmente consiste em realizar uma aproximação dos pixels da vizinhança utilizando polinômios quadráticos, posteriormente um novo sinal é construído por um deslocamento global e finalmente, o deslocamento global é calculado através da equação dos coeficientes. O terceiro bloco, empregamos o método de super resolução bilateral $TV - L1$ (*BTVL1*), considerando que os vídeos foram adquiridos com câmera e objeto de interesse em movimento, assim, para esta metodologia específica, o emprego deste método nos proporcionou a redução das variações das imagens, uma vez que, as imagens adquiridas estavam em movimento, pois, o

BTV é empregado no pixel central, além disso, o método proporciona a preservação das áreas de borda e remove ruídos simultaneamente. Ainda neste terceiro bloco, utilizamos alguns parâmetros configurados no bloco anterior, como por exemplo, o fator de escala e número de iterações. Outros parâmetros foram definidos conforme listados a seguir: (i) λ , foi estabelecido como parâmetro de peso a suavidade do vídeo SR de saída, onde, quanto maior este parâmetro, mais suave será o vídeo SR, (ii) α , neste parâmetro definimos a distribuição espacial BTV , (iii) *blur kernel size*, definimos o tamanho do filtro BTV (iv) *blur sigma*, definimos o grau de desfoque do *Gaussian Blur*. (v) *opticalFlow*, definimos o algoritmo de fluxo óptico. Esta última etapa, realiza o processamento do vídeo de entrada e realiza o cálculo da super-resolução em cada quadro do vídeo, resultando em um processo computacional lento. A seguir, apresenta-se o esquemático empregado neste bloco através da Figura 3.3.

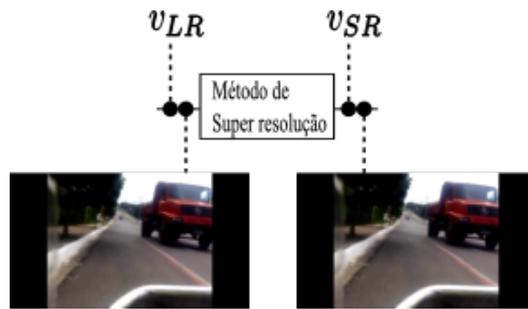


Figura 3.3: Esquemático empregado no bloco de Super Resolução Bilateral $BTVL1$.

A Figura 3.3, ilustra o bloco que emprega o método de Super Resolução bilateral $TV-L1$, onde, a entrada do bloco tem-se o vídeo de baixa resolução que corresponde ao v_{LR} e na saída, o vídeo com Super resolução v_{SR} .

3.3 Conversão RGB para escala de cinza

O presente bloco, além de realizar uma conversão em si, realiza um simples pré-processamento buscando reduzir os efeitos de brilho na imagem devido condições de iluminação excessivas. Assim, este bloco proporciona a diminuição de brilho, resultando em uma imagem com iluminação sem brilho excessivo, o que favorece o desempenho do sistema. Outro aspecto, observamos a melhoria de contraste na placa do veículo, pois observamos que as bordas tornam-se evidenciadas. Adicionalmente,

o tempo de processamento diminui, pois considera-se o que o tempo de processamento de uma imagem com um único canal é mais rápido (imagem em escala de cinza) do que processar uma imagem com três canais (imagem RGB). A descrição matemática é feita como segue: primeiramente, supõe-se, como entrada entrada do bloco, uma imagem colorida no modelo de cores RGB, ou seja, $f_{RGB} = \{f_R, f_G, f_B\}$. A imagem em escala de cinza é obtida conforme a seguir:

$$f_c(n_1, n_2) = 0,114 f_R(n_1, n_2) + 0,587 f_G(n_1, n_2) + 0,299 f_B(n_1, n_2) \quad (3.1)$$

onde, na equação (3.1), f_c é a imagem em escala de cinza. Normalmente, estes valores de coeficientes apresentados na equação são amplamente utilizados, porém há outros coeficientes que podem ser escolhidos. A Figura 3.4, ilustra o bloco utilizado para realizar a conversão do modelo de cor RGB para escala de cinza, assim, a entrada do bloco é a imagem f_{RGB} proveniente da base de dados e o resultado consiste em uma imagem f_c em escala de cinza. O emprego do bloco representa a utilização da equação e os coeficientes de cada canal citado anteriormente.



Figura 3.4: Conversão de imagem no modelo de cor RGB para escala de cinza.

3.4 Filtro *blur*

Neste bloco, foi realizada a remoção de ruídos utilizando-se um filtro *blur Gaussiano*. Considerando-se as imagens que contém o objeto de interesse (placa de licenciamento veicular) e a câmera estão em movimento diante de um *background* complexo, pois, é possível notar ruídos na imagem relacionados ao ambiente e a captura, por exemplo ruído causado por iluminação excessiva ou deficiente, excesso de brilho, ruídos relacionado ao movimento gerado na aquisição, pequenos retângu-

los que se assemelham à placa do veículo. No entanto, observamos que esses tipos de ruídos ocorrem com certa frequência, desta forma, deve-se realizar a filtragem para evitar ou reduzir este prolema. Tendo em vista que o próximo bloco consiste em um detector de borda, é interessante abordar algum benefício do filtro *blur* com relação a melhoria para detecção de borda. De fato, com esta filtragem, há uma redução com respeito a falsas detecções de borda, devido, principalmente aos pequenos retângulos semelhantes à placa e linhas próximas que serão filtrados (borrados). Adicionalmente, uma outra característica do emprego desta filtragem é destacar as linhas que limitam a placa do veículo, pois as espessuras destas linhas serão borradas, o que proporciona uma melhora na detecção de borda. Por último, nota-se que, em muitos métodos de detecção de placas onde utiliza-se imagens reais e verificou-se o emprego do filtro *blur*. A seguir descreve-se matematicamente a obtenção da imagem filtrada:

$$h(n_1, n_2) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-n_1^2 + n_2^2}{2\sigma^2}\right) \quad (3.2)$$

onde, na equação (3.2), $h(n_1, n_2)$ é o filtro empregado na imagem de entrada do bloco, σ representa o tamanho da máscara aplicada em uma janela que é deslocada na imagem, podendo ser 3×3 , 5×5 ou 7×7 , onde quanto maior for o σ , maior será o grau de desfoque causado na imagem.

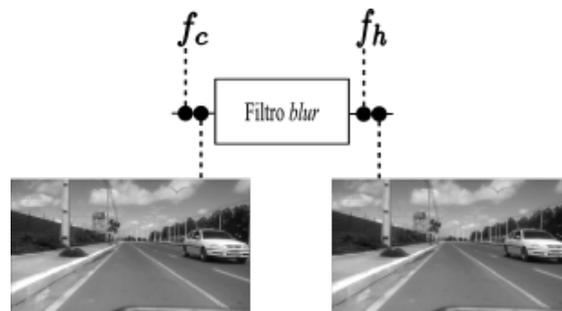


Figura 3.5: Método de emprego do filtro *blur*.

O bloco ilustrado na Figura 3.5 apresenta o emprego do filtro *blur*, onde temos uma imagem de entrada em escala de cinza proveniente do bloco anterior, a saída do bloco corresponde a uma imagem f_h filtrada (borrada) com ruídos suavizados.

3.5 Filtro *Sobel*

Este bloco realiza a detecção de bordas do objeto de interesse (placa de licenciamento veicular) através do filtro *Sobel*. Este filtro é amplamente utilizado em sistemas de detecções de placas, do mesmo modo, muitos pesquisadores mencionam que o número de bordas verticais em uma placa de veículo é maior do que o número de bordas horizontais, pois esta observação inclui os caracteres da placa, proporcionando o emprego de uma particularidade deste filtro, que consiste em efetuar a detecção de bordas em direções separadas (vertical ou horizontal). Neste bloco, o emprego do filtro *Sobel* permitiu realizar a detecção de bordas na direção vertical, onde o método ocorre através da filtragem utilizando máscaras de tamanho $N \times N$. Observou-se que o desempenho melhora quando se emprega uma imagem suavizada para realizar a detecção de bordas, pois a filtragem de alta frequência torna as bordas mais espessas. A seguir, descreve-se matematicamente a obtenção da imagem filtrada:

$$f_d = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 2 & 1 \end{bmatrix} * f_h \quad (3.3)$$

onde, na equação (3.3), f_d é a imagem filtrada, f_h é a imagem de entrada, $*$ representa o operador convolução. Nesta dissertação, foi utilizada a máscara $N \times N$ em uma única direção, a vertical, pois, a escolha da direção resultou em bordas evidentes. A seguir, na Figura 3.6, apresentamos o esquemático empregado neste bloco, onde, a entrada do bloco consiste em uma imagem suavizada f_h e na saída do bloco uma imagem f_d com as bordas verticais da imagem destacadas.

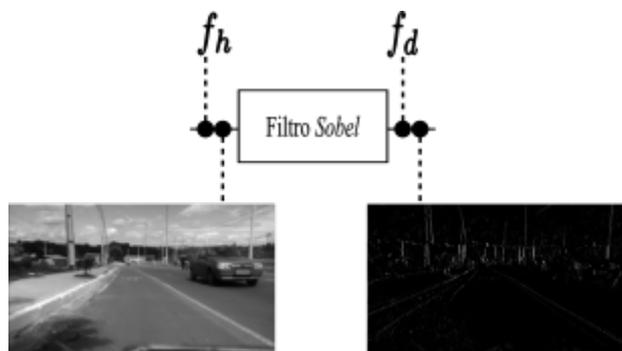


Figura 3.6: Detecção de borda vertical através do filtro *Sobel*.

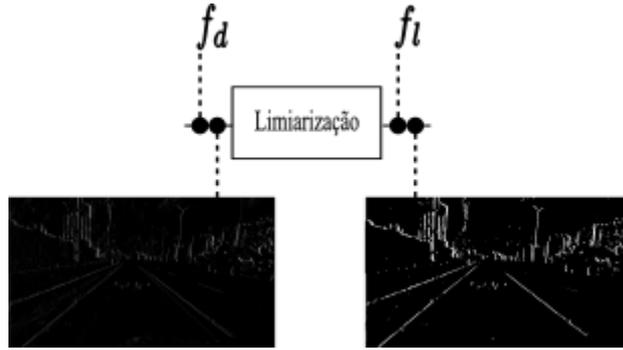


Figura 3.7: Esquema do método *otsu*.

3.6 Limiarização

Na saída do bloco de limiarização, há regiões que podem conter o objeto de interesse (placa de licenciamento veicular), no entanto, o objetivo principal deste bloco é realizar uma tentativa de segmentar o objeto de interesse do *background*, pois, nesta dissertação, não foi possível utilizar alguns métodos de limiarização de baixa complexidade, como por exemplo, limiarização global, que utiliza um único valor de limiar para toda imagem. Isso ocorre devido as características das imagens utilizadas para processamento. Neste caso em específico, tem-se imagens reais, com *background* complexo e não controlado, câmera e objeto de interesse em movimento, imagens ruidosas e com iluminação não uniforme. Assim, de acordo com as características destas imagens, utiliza-se, nesta dissertação, um método de limiarização adaptativa. Neste caso, utiliza-se o método *Otsu*, pois este melhora a separação entre o objeto do *background*, dividindo a imagem em duas classes. O método *Otsu* também se destaca em imagens com diferentes pontos de iluminação e apresenta ótimos resultados de segmentação quando aplicado a imagens ruidosas e se caracteriza por ser amplamente utilizado em imagens reais. A equação (2.12), apresentada na Seção 2.3, descreve matematicamente o método *Otsu* empregado neste bloco para obtenção do limiar empregado na imagem¹. A seguir apresentamos o esquemático utilizado neste bloco.

a Figura 3.7, ilustra o emprego deste bloco de limiarização, onde f_d é uma imagem de entrada no bloco, proveniente do bloco anterior e f_l é a imagem limiarizada. Após o emprego do método *Otsu*, obtemos uma imagem binária com objeto

¹Por comodidade para o leitor, expressa-se novamente a equação do método *Otsu*, que é igual a: $\sigma(T) = N_B(T)N_O(T)[\mu_B(T) - \mu_O(T)]$.

de interesse parcialmente separado do *background*.

3.7 Filtro Morfológico

O emprego da operação morfológica de fechamento utiliza uma imagem binária f_l para realizar um pós processamento nas regiões de interesse, ou seja, nas regiões que podem conter a placa de licenciamento veicular, baseado em seu formato. Neste bloco, a operação de fechamento nos permite remover lacunas em branco entre cada linha de borda vertical e conecta todas as regiões que possuem um alto número de arestas, assim, a detecção da região de interesse foi baseado na escolha do elemento estruturante. Nesta dissertação, utilizou-se um elemento estruturante retangular com dimensões 17×3 . A equação (2.16), apresentada na Seção 2.4, descreve matematicamente o filtro morfológico utilizado².

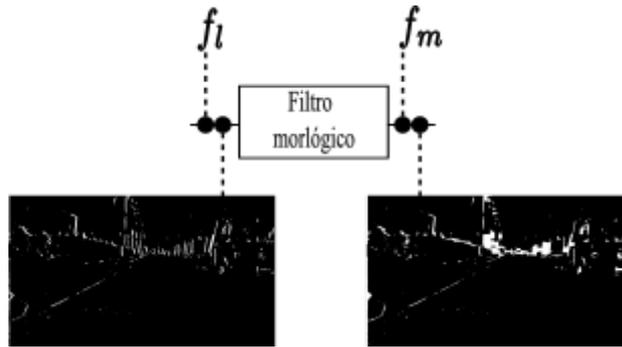


Figura 3.8: Bloco para filtragem morfológica.

3.8 Método de preenchimento

O emprego do algoritmo *flood fill* define se os componentes de uma determinada região estão conectadas, primeiramente se os componentes são conexos através de características, como por exemplo, a vizinhança, se possuem a mesma cor de *backgraound*, textura e intensidade dos níveis de cinza. Uma das propriedades do algoritmo é prever as direções em que estes componentes estão conectados, através da conectividade 4 ou conectividade 8. Assim, neste bloco, os objetos de interesse possuem mesma cor branca de *backgraound*, onde o emprego do algoritmo *flood fill*,

²Por comodidade para o leitor, expressa-se novamente a equação de fechamento, que é igual a: $\phi_B(f) = \varepsilon_{\tilde{B}}[\delta_B(f)]$.

consiste em encontrar um ponto de início na região de interesse, substituindo a região por uma cor de destino, nesse caso, o preenchimento ocorre com a cor branca, e de forma uniforme por toda região do objeto de interesse dentro de um intervalo previamente especificado nas direções de preenchimento, onde para esta metodologia utilizamos a conectividade 8, pois a escolha deste parâmetro proporcionou a detecção do objeto de interesse em decorrência da alta densidade vertical presente no objeto de interesse e este parâmetro eliminando resultados falsos positivos. O resultado final é uma imagem, com o objeto de interesse (placa de licenciamento veicular) detectada. A seguir, ilustramos o método utilizado neste bloco. Na Figura

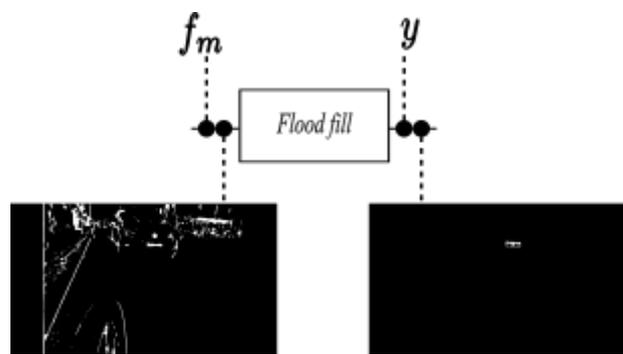


Figura 3.9: Diagrama proposto do algoritmo *flood fill*.

3.9, ilustra o emprego do algoritmo *flood fill* neste bloco, onde f_m é uma imagem de entrada no bloco e proveniente do bloco anterior. A saída do bloco resulta em uma imagem com máscara binária contendo o objeto de interesse destacado.

Capítulo 4

Base de Dados

Neste capítulo, são abordados detalhes de outras bases de dados utilizada em outras pesquisas, serão citados as principais características e na sequência, detalhadas as informações sobre a base de dados empregada nesta dissertação, como por exemplo, explicamos o cenário onde foi adquirida, como se deu o processo de aquisição da base de dados e quais meios utilizados para esta tarefa, descrevemos também como esta base de dados foi armazenada e organizada para uso nesta dissertação e validação do método proposto.

4.1 Outras bases

A base de dados *Vision calltech* foi empregada pelos autores Syed Zain Masood *et al.* [55] para comprovar o objetivo de sua pesquisa sobre detecção de placas que independe do modelo de placa escolhido. A base de dados possui 126 imagens com resolução de 896×592 pixels apresentando imagens com excelente qualidade. Esta base de dados, apresenta em sua composição imagens de veículos com placas dos Estados Unidos com tamanhos $15 \times 30cm$, o padrão americano utiliza um sistema alfanumérico com sete caracteres. As tomadas das imagens foram frontais concentrando-se apenas nas placas traseiras dos veículos, durante a aquisição, mantiveram uma certa distância da traseira do veículo, e foi observado que a iluminação é excelente na região do objeto de interesse (placas de veículos), onde é possível notar claramente os caracteres da placa, pois, a aquisição das imagens foi realizada com dia claro com câmera e veículos estáticos. A figura 4.1 a seguir apresenta algumas

imagens desta base de dados.

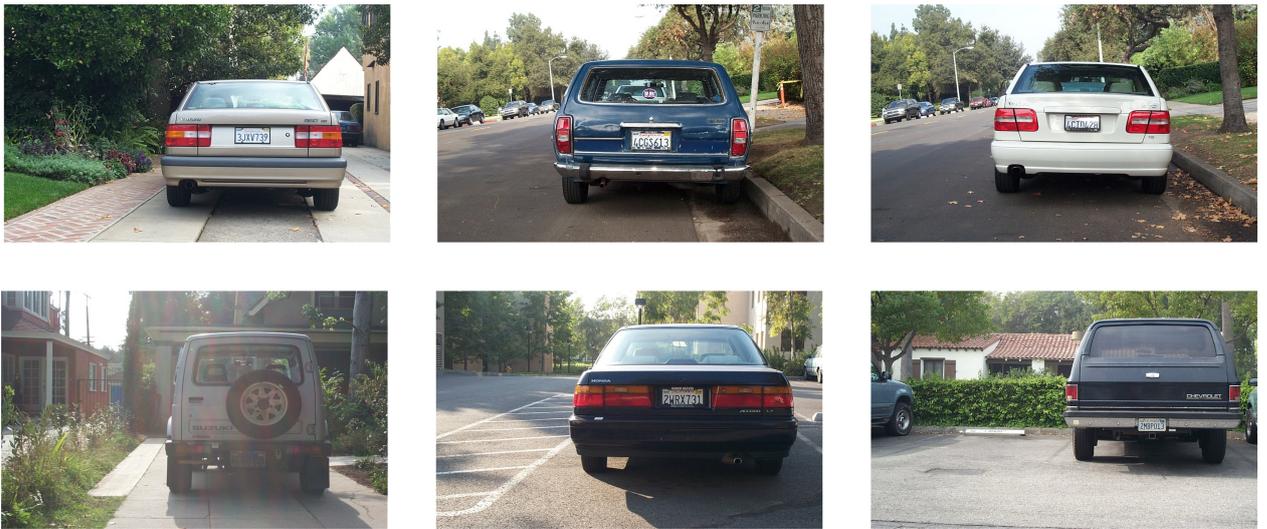


Figura 4.1: A figura apresenta algumas imagens das bases de dados *Vision calltech*.

A base de dados *Open ALPR benchmark* também foi empregada pelos autores Syed Zain Masood *et al.* [55] com o objetivo de realizar a detecção de placas de veículos independente do modelo de placa do veículo escolhido, no entanto, esta segunda quantidade de imagens utilizada na pesquisa, é composta por veículos da União Europeia e dos Estados Unidos, contendo 839 imagens distribuídas entre placas de veículos das duas nacionalidades e com as seguintes resoluções: de 1280×720 , 1920×1080 , 1000×750 , 1004×768 e 640×480 pixels, porém, em todas as resoluções da presente base, as placas dos veículos estão visíveis, ressaltando a qualidade das imagens. Esta base apresenta a aquisição de imagens reais com tomadas frontais e tomadas com inclinação vertical, alguns casos as imagens foram adquiridas com a câmera e veículo estáticos ou em alguns casos, câmera estática e veículos em movimento, diferentemente da base citada anteriormente, a base de dados *Open ALPR benchmark* possui registros de placas de veículos traseiras e dianteiras. Foi observado também as condições de iluminação das imagens nesta base, por exemplo, apresentam-se imagens em ambiente noturnos, iluminação sem luz solar durante o dia e com luz solar também durante o dia. A seguir, apresentamos algumas imagens desta respectiva bases de dados.

Na Figura 4.2, apresenta as imagens que pertencem a base de dados *Open ALPR benchmark* em condições ambientais reais conforme mencionado, podemos observar também a qualidade das imagens e a visibilidade das regiões de interesse

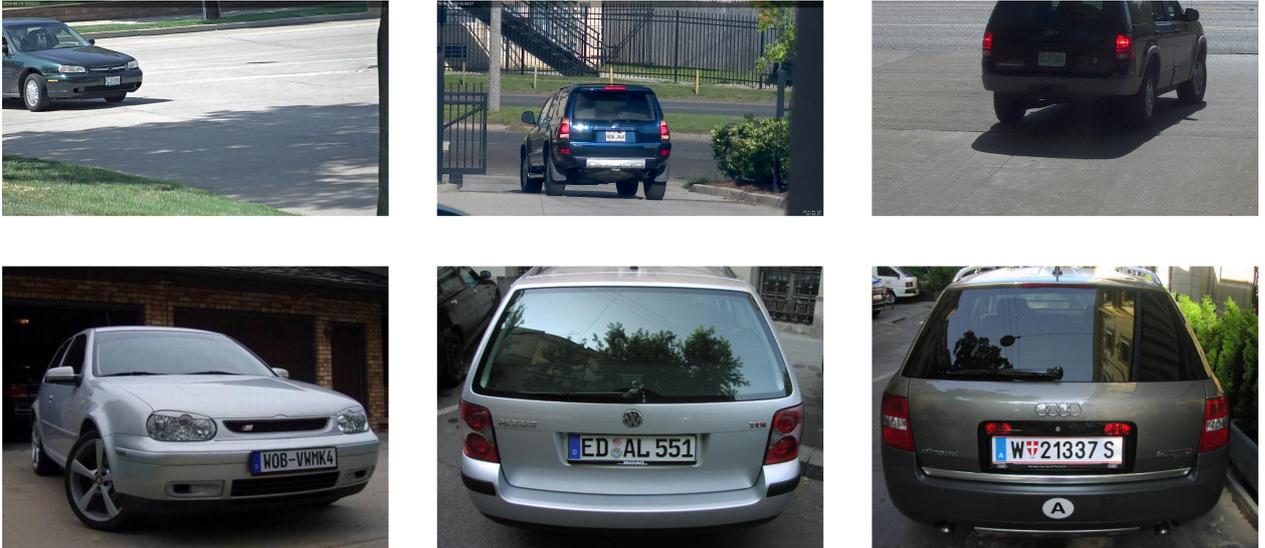


Figura 4.2: A figura apresenta algumas imagens das bases de dados *Open ALPR benchmark* com placas de veículos americanas e europeias.

que contém as placas dos veículos. Às duas bases citadas, *Vision calltech* e *Open ALPR benchmark* foram utilizadas pelo autores Syed Zain *et al* Masood [55].

A base de *Bangla* foi criada e utilizada pelo autor Samiul Azam [56] para detectar placas de veículos em situações perigosas. Esta base contém 525 imagens de placas de veículos de *Bangla*, todas as imagens da base possuem resolução de 640×480 pixels e está dividida da seguinte forma: (i) 105 imagens em chuva forte e com quatro orientações de chuva distintas, (ii) 60 imagens com chuva moderada e quatro orientações de chuva distintas, (iii) 60 imagens com chuva leve com apenas duas orientações de chuvas distintas. As imagens restantes contabilizam 300 e de acordo com o autor não foram afetadas pela chuva e apresentam-se em condições ambientais diversas listadas a seguir: iluminação diurna excelente, imagens em ambientes internos, imagens em nevoeiro, condição desfocada e imagens em ambiente noturno. A câmera utilizada na aquisição da base de *Bangla* foi posicionada a 1,80m de altura do solo e aproximadamente a 3m de distância do veículo, além disso, observamos que as tomadas foram realizadas nas seguintes direções: tomada frontal, tomadas inclinadas verticais e realizaram as aquisições das placas de veículos traseiras e dianteiras. A figura a seguir apresenta algumas imagens da base de dados de *Bangla* em condições ambientais diversas.

Na Figura 4.3, apresenta as imagens da base de *Bangla*, onde observamos as condições de chuva, nevoeiro, ambientes internos, condições de iluminação diurna



Figura 4.3: A figura apresenta algumas imagens da base de dados de Bangla concebida pelo autor Samiul Azam.

excelente conforme mencionado anteriormente e as tomadas das imagens foram realizadas em várias direções.

Uma outra base de dados chamada *cars brad* oriunda do Sul da Califórnia com 526 imagens e com resolução 360×240 pixels e foi utilizada pelo pesquisador Rohith Polishetty [57], porém, esta base foi concebida com a câmera e veículo ambos em movimento, capturando apenas as placas traseiras dos veículos, sem apresentar trepidação da câmera ou qualquer outro tipo de ruído que possa interferir na detecção de placas de veículos. As ruas onde foram realizadas as aquisições não possuem ondulações, buracos ou ainda curvas e apesar de realizar a aquisição em movimento a placa do veículo está visível e com excelentes condições de iluminação. A partir destas sequências de vídeos, foram extraído os *frames* sequenciais e separados para a base de dados. A figura a seguir ilustra algumas imagens desta base de dados.

Na Figura 4.4, apresenta as imagens da base *cars brad* utilizada pelo pesquisador Rohith Polishetty, onde é possível notar as características do ambiente onde essas imagens foram capturadas.

Os autores Syed Tahir *et al.* [58] utilizaram uma base de dados de imagens reais de veículos italianos para um sistema de detecção e reconhecimento de placas usando classificadores neurais profundos. As imagens da base de dados italiana, podem ser baixadas na internet e estão disponíveis publicamente. Esta base é composta por 788 imagens com resolução de 640×480 pixels, as imagens são de excelente

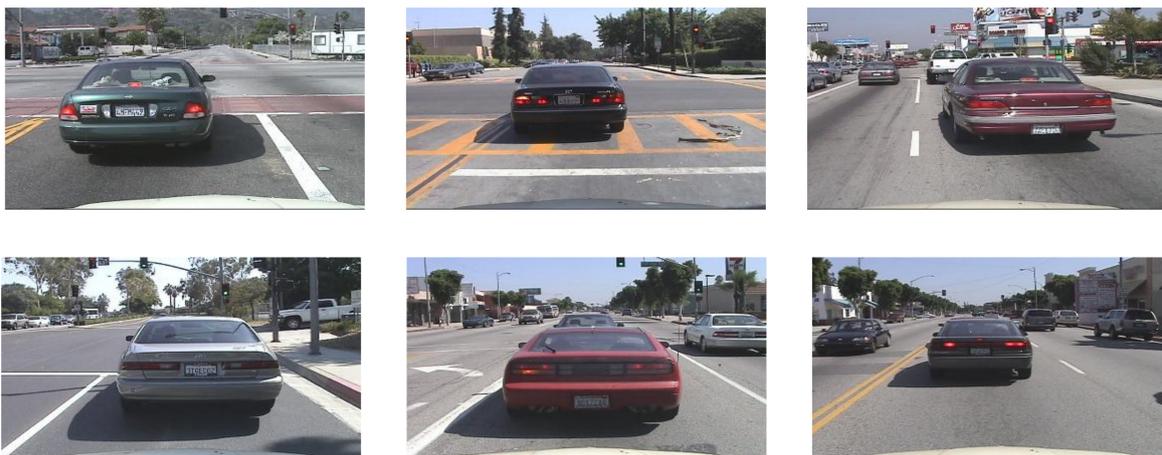


Figura 4.4: A figura apresenta as tomadas sequenciais de imagens traseiras dos veículos.

qualidade com excelente visibilidade da região de interesse onde contem as placas de veículos. As imagens da base de dados italiana foram adquiridas durante a luz do dia, a aquisição foi realizada com câmera e veículos estáticos, onde as tomadas foram frontais e com inclinação vertical. A seguir, apresentamos algumas imagens de veículos italianos desta base.

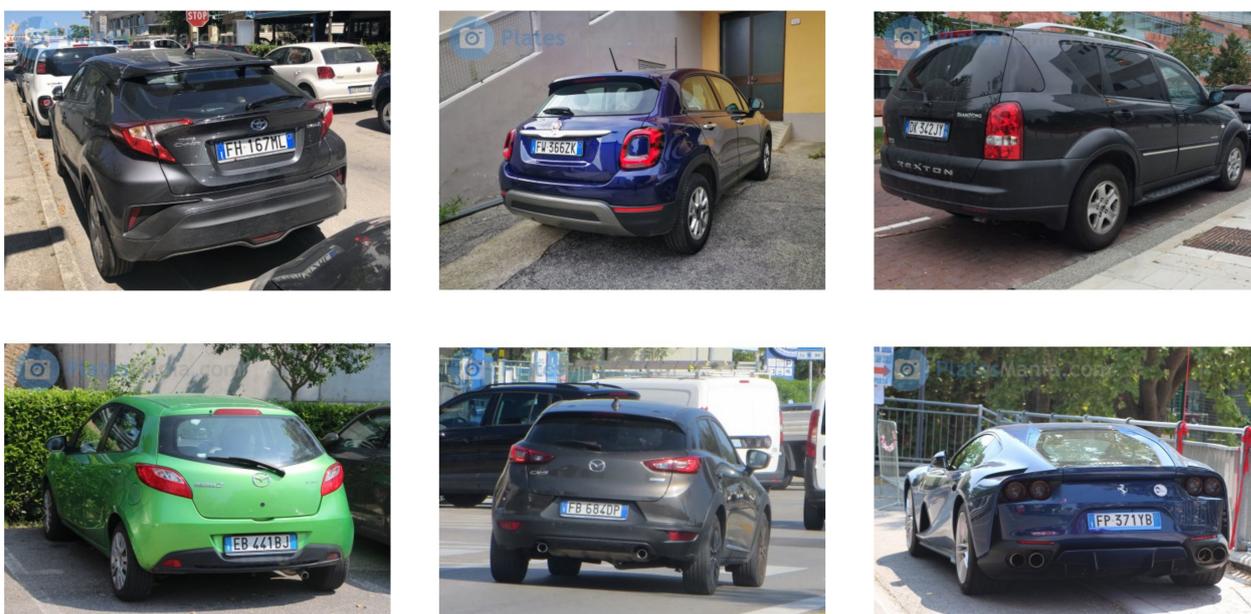


Figura 4.5: A figura apresenta as imagens traseiras de veículos oriundas da base de dados italiana.

4.2 Placas de veículos brasileiros

As placas de licenciamento veicular são utilizadas para identificação e são emitidas pelos Departamentos Estaduais de Trânsito (DETRAN) de cada unidade da federação, seguindo uma padronização para todo território brasileiro. Atualmente, encontramos dois modelos de placas de veículos brasileiros em circulação, inicialmente o modelo criado em 1990, com formato retangular medindo 40 *cm* de largura por 17 *cm* de altura, contém sete caracteres alfanuméricos individualizados, onde o primeiro grupo é composto por três letras resultante de um arranjo de 26 letras e tomada três a três e o segundo grupo de caracteres é composto por quatro números resultante do arranjo de dez algarismos (CONTRAN).

A partir deste modelo criado em 1990, listamos alguns tipos de placas de veículos brasileiros, onde tem-se:

(a) placas com fundo branco e com caracteres vermelho são relacionados a veículos de autoescola;

(b) placas com fundo vermelho e com caracteres brancos são relacionados a veículos de aluguel;

(c) placas com fundo cinza e com caracteres pretos são relacionados a veículos particulares;

(d) placas com fundo verde e com caracteres brancos são relacionados a novos modelos de veículos em teste;

(e) placas com fundo preta e com caracteres cinzas são relacionados a veículos de colecionadores;

(f) placas com fundo azul e com caracteres brancos são relacionados a veículos de representações diplomáticas. Na Figura 4.6, apresentamos os modelos de placas de veículos mencionados. No ano de 2017, iniciaram a implantação das novas placas de veículos com padrão Mercosul em alguns estados brasileiros e posteriormente seguindo para os demais, o novo modelo de placas de veículos possui um fundo branco e é revertida com película retrorreflexiva. Além disso, o novo modelo se aplica a veículos automotores seguindo as mesmas características do modelo antigo, onde a localização das placas nos veículos permanece a mesma, na traseira e dianteira dos veículos, o mesmo formato retangular com as mesmas dimensões de 40 *cm* de largura por 17 *cm* de altura, mantendo em sua composição os setes caracteres alfanuméricos.



Figura 4.6: (a) Auto escola (b) Aluguel (c) Particular (d) Teste (e) Colecionador (f) Representação



Figura 4.7: (a) Particular (b) Representação Diplomática (c) Aluguel (d) Teste (e) Representação Oficial (f) Colecionadores

A seguir, listamos alguns modelos de placas de veículos com padrão Mercosul:

(a) placas com fundo branco e com caracteres pretos referem-se a veículos particulares;

(b) placas com fundo branco e com caracteres dourados referem-se a veículos de representações diplomáticas;

(c) placas com fundo branco e com caracteres vermelho referem-se a veículos com fins comerciais;

(d) placas com fundo branco e com caracteres verdes referem-se a novos veículos que estão em teste;

(e) placas com fundo branco e com caracteres azuis referem-se a veículos de representações oficiais;

(f) placas com fundo branco e com caracteres cinza referem-se a veículos de colecionadores. Na Figura 4.7, apresentamos os novos modelos de placas de veículos padrão Mercosul, conforme foram mencionados.

4.3 Base de dados

A base de dados empregada nesta dissertação é de autoria própria e foi concebida entre os meses de março e setembro do ano de 2016 em dois locais distintos. O primeiro local de aquisição dos vídeos da base de dados, foi na Avenida Nathan Xavier de Albuquerque que é uma via de sentido duplo, localizada no bairro Novo Aleixo na cidade de Manaus - AM, onde cada trecho de sentido da via possui 3 quilômetros de extensão com uma ciclofaixa destinada ao tráfego de ciclistas. Esta ciclofaixa encontra-se posicionada a direita da via acompanhando o respectivo sentido da via, além disso, a ciclofaixa desta avenida encontra-se sinalizada através da pintura de faixas no solo e é separada das vias principais dedicadas aos demais veículos através de separadores do tipo taxinhas com sinalizadores reflexivo. Estes trechos da ciclofaixa, apresentam ainda ondulações no asfalto, buracos, linha reta, curvas e contemplam pequeno aclive e declive, onde, todo este cenário contribui para um *background* complexo. Por ser uma via de grande circulação, normalmente encontramos veículos de pequeno, médio e grande porte circulando pela via. O segundo local da aquisição dos vídeos para base de dados, foi na Avenida Tranquedo Neves, localizada no bairro Parque dez na cidade de Manaus-AM, onde percorremos um único trecho de 100 metros que compreende entre o ponto comercial Centro do Alumínio e Supermercado Veneza, diferentemente da Avenida Nathan Xavier, a Avenida Tranquedo Neves não possui uma ciclofaixa sinalizada na rua com pinturas de solo, com isso, os carros tendem a passar mais próximo da bicicleta. As figuras a seguir, apresentam algumas imagens dos locais referente das aquisições dos vídeos e ilustram algumas características do ambiente de aquisição.

A aquisição da base de dados foi realizada portando um aparelho celular com câmera, onde, o dispositivo foi posicionado e instalado abaixo do banco de uma bicicleta. Além disso, a aquisição dos vídeos foi concebida com três resoluções diferentes listadas a seguir: (i) 1280×720 pixels, (ii) 854×480 pixels, (iii) 426×240 pixels. Desta forma, as tomadas dos vídeos foram efetuadas percorrendo a ciclofaixa em toda extensão da via nos dois sentidos e em horários distintos com condições ambientais diversas. A figura a seguir, ilustra como a câmera foi posicionada na bicicleta para realizar dos vídeos da base de dados.

Na Figura 4.9 ilustra os equipamentos envolvidos na aquisição dos vídeos

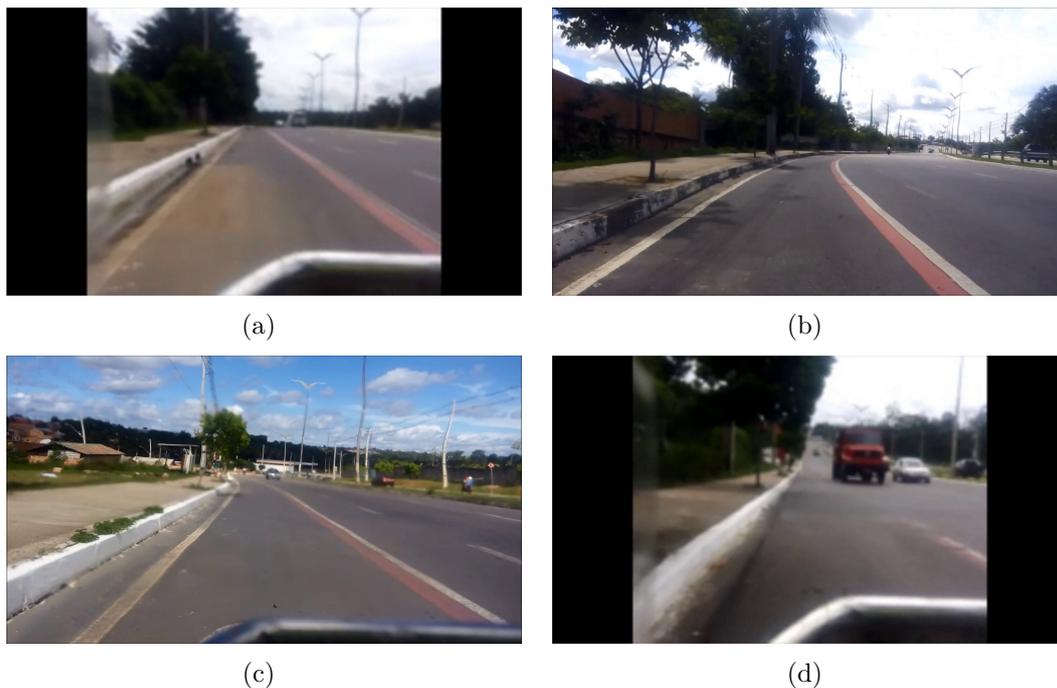


Figura 4.8: O conjunto de imagens acima, apresenta diversos trechos da avenida Nathan Xavier

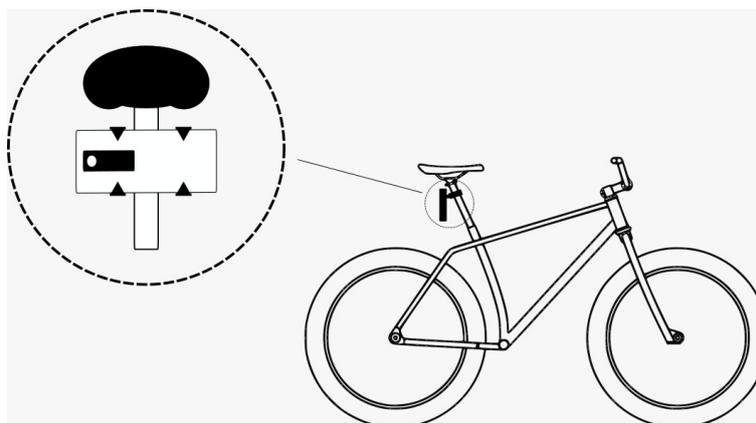


Figura 4.9: A figura ilustra o posicionamento do aparelho celular, dando destaque como o aparelho celular foi fixado através de um dispositivo

para construção da base de dados, além disso, exibe como a câmera foi posicionada e fixada abaixo do banco da bicicleta.

Durante a aquisição dos vídeos, notamos alguns aspectos que resultaram na identificação de algumas características da nossa base de dados que listaremos a seguir: (i) câmera em movimento com objeto de interesse em movimento;

(ii) *background* complexo durante toda a aquisição dos vídeos, pois, ao analisarmos os vídeos, observamos diversos fatores que podem gerar em falsos positivos;

(iii) a trepidação da câmera em decorrência da não uniformidade da rua

ocasionada por ondulações e buracos, conforme mencionado anteriormente;

(iv) a tomada das imagens com inclinação vertical, pois, a ciclofaixa encontra-se posicionada a direita das vias, assim, o objeto de interesse não está posicionado frente a câmera;

(v) condições de iluminação distintas, iluminação diurna, iluminação noturna e iluminação sob condições de tempo nublado sem sol e chuvoso, condições de sombra na via gerada pela copa das árvores;

(vi) vídeos com resoluções de 1280×720 e 854×480 pixels, os vídeos foram adquiridos sob condições de iluminação diurna, pela manhã e pela tarde, porém, a iluminação no objeto de interesse mostrou-se variável dependendo do trecho analisado o objeto de interesse encontra-se visível ou não, assim como, a trepidação da câmera também contribui para a visibilidade do objeto de interesse;

(vii) vídeos adquiridos com resolução de 426×240 pixels, onde neste último caso a aquisição dos vídeos resultaram em imagens de baixa qualidade, sob condições meteorológicas de tempo nublado e em todos os casos o objeto de interesse não se encontra nitidamente visível, além de outros fatores levantados para esta resolução dos vídeos, adiciona-se a trepidação da câmera que contribui para baixa visibilidade do objeto de interesse;

(viii) Por fim, o formato como foi realizado a aquisição dos vídeos, com câmera e veículo em movimento. A seguir, nas Figuras 4.10, 4.11 e 4.12 apresentamos um conjunto de imagens de acordo com cada resolução contida nesta base de dados.

Os vídeos adquiridos possuem tempos de duração distintos, pois, variavam de acordo com o percurso estabelecido e percorrido, assim, resultaram em vídeos longos e fragmentamos em vídeos menores. A tabela a seguir, apresenta organização dos vídeos adquiridos, divididos por resolução e duração em minutos.

Tabela 4.1: Organização e divisão dos vídeos da base de dados.

Vídeos	Resolução	Duração de vídeo
Vídeo 1	1280×720	11m 20s
Vídeo 2	426×240	15m 20s
Vídeo 3	426×240	10m 2s
Vídeo 4	426×240	17m 50s
Vídeo 5	854×480	2m 24s
Vídeo 6	426×240	15m 1s



Figura 4.10: Amostra de imagens da base de dados com resolução de 426×240 pixels em condições reais no momento da aquisição.



Figura 4.11: Amostra de imagens da base de dados com resolução de 854×480 pixels em condições reais no momento da aquisição.



Figura 4.12: Amostra de imagens da base de dados com resolução de 1280×720 pixels em condições reais no momento da aquisição.

Na tabela 4.1, ilustra inicialmente o tempo de duração dos vídeos longos e a resolução de cada vídeo. Os vídeos originais adquiridos na aquisição, foram fragmentados em vídeos de curta duração para realizarmos a extração dos *frames* de interesse e dividir em grupos de imagens sequenciais. Além do mais, como critério de seleção dos *frames* de interesse, consideramos o momento em que o veículo surge na cena e se aproxima da câmera, até o momento em que o veículo sai da cena totalmente. A tabela a seguir, ilustra a próxima organização da base de dados, compreendendo grupos que contém as sequências de imagens e o total de imagens de cada grupo.

Na tabela 4.2, apresentamos a divisão dos grupos com suas respectivas sequências de imagens e separadas por resolução, cada grupo contém uma determinada quantidade de *frames* correspondentes a sequência de imagens oriundas da fragmentação de vídeos de curta duração. A seguir detalharemos as características de cada grupo de imagens.

Os grupos com sequência de imagens com resolução de 1280×720 , correspondem ao vídeo 1 descritos na tabela 4.1, adquiridos na Avenida Nathan Xavier, apresentam algumas características variantes dependendo do trecho analisado, como por exemplo, iluminação diurna e com variações no decorrer das sequências de imagens, em alguns trechos, apresentam baixo contraste e aquisição desfocada, além disso, a imagem desfocada, interfere na visibilidade do objeto de interesse. As condições ambientais da ciclofaixa apresentam diversos fatores que contribuem para um *background* complexo, pois durante a aquisição observamos por exemplo, poças de água na via, notamos a presença de sinalização na ciclofaixas com pinturas de solo, presença de terra, manchas de óleo no asfalto e placas de propaganda. Além de todos os fatores observados, adicionamos a forma como foi realizada a aquisição dos vídeos, com câmera e objeto de interesse em movimento, que é favorável a detecções de falsos positivos. O segundo trajeto realizado para aquisição dos vídeos na mesma Avenida Nathan Xavier, contempla trechos da via em linha reta, acurve com curva, curvas ondulações e buracos, veículos de pequeno, médio e grande porte, mais de um veículo em cena e um último fator observado, é que dependendo do trecho da via no momento da aquisição, é possível observar veículos no outro lado da via.

Descrevemos um segundo grupo de imagens com resolução de 426×240 ,

Tabela 4.2: Divisão da base de dados em sequências de imagens

Resolução dos frames	Grupos	Sequência de imagens	Quantidade de frames
1280 × 720	G1	1 – 52	51
1280 × 720	G2	52 – 102	51
1280 × 720	G3	103 – 153	51
1280 × 720	G4	154 – 199	46
1280 × 720	G5	200 – 250	52
1280 × 720	G6	251 – 301	51
1280 × 720	G7	302 – 352	51
1280 × 720	G8	353 – 403	51
1280 × 720	G9	404 – 454	51
1280 × 720	G10	455 – 505	51
1280 × 720	G11	506 – 556	51
1280 × 720	G12	557 – 606	50
1280 × 720	G13	607 – 657	51
426 × 240	G14	658 – 707	50
426 × 240	G15	708 – 756	49
426 × 240	G16	757 – 805	49
426 × 240	G17	806 – 853	48
426 × 240	G18	854 – 903	50
426 × 240	G19	904 – 952	49
426 × 240	G20	953 – 1000	48
426 × 240	G21	1001 – 1050	50
1280 × 720	G22	1051 – 1098	48
1280 × 720	G23	1099 – 1148	50
1280 × 720	G24	1149 – 1197	49
1280 × 720	G25	1198 – 1245	48
1280 × 720	G26	1246 – 1293	48
854 × 480	G27	1294 – 1343	50
1280 × 720	G28	1344 – 1392	49
854 × 480	G29	1393 – 1451	59
1280 × 720	G30	1452 – 1542	91
Total	-	-	1542

correspondente ao vídeo 2 da tabela 4.1, também os vídeos foram adquiridos na Avenida Nathan Xavier, onde como um dos principais aspectos desse grupo, citamos a qualidade das imagens, pois a baixa resolução e qualidade da imagem interfere diretamente na visibilidade do objeto de interesse. Além do mais, neste grupo, possuímos fatores variantes de iluminação e condições de tempo nublado sem sol, presença de sombras nos veículos e objeto de interesse gerada pelas copas das árvores posicionadas do lado direito da via, no mesmo lado da ciclofaixa. O ambiente externo também contribui para um *background* complexo, visto que, observamos em alguns casos a presença de dois veículos em cena, em que um deles estava na mesma direção do objeto de interesse, ainda nessa sequência de imagens, observamos trechos da ciclofaixa com sinalização no asfalto através de pinturas de solo, manchas no asfalto ocasionadas por óleo e neste lado da via, não há presença de placas com propagandas. Parte do trajeto da via está em linha reta, com aclive seguido de um declive, curvas, ondulações, buracos e veículos com três tipos de porte.

Por fim, descrevemos um terceiro grupo de imagens com resolução de 854×480 , esta sequência de imagens corresponde ao vídeo 3 da tabela 4.1, que foi adquirido na Avenida Tranquedo Neves, conforme fora mencionada anteriormente. Esta terceira sequência possui poucas imagens e foi realizada em um curto trajeto sob condições de iluminação variantes. O ambiente em si, assim como as outras sequências, também dispõe de fatores que contribuem para um *background* complexo, como por exemplo, observamos manchas no asfalto ocasionadas por óleo, a existência de poças de água nas sarjetas e vegetação rasteira em toda a extensão do percurso, no momento da aquisição havia veículos com faróis ligados, observamos também a presença de pedestres nas calçadas e no presente momento da aquisição, apenas veículos de pequeno porte circulavam pela via. O trajeto desta sequência de imagens foi realizado em linha reta e não havia ondulações e buracos, porém em todo o percurso é possível observar os veículos no outro lado da via. A próxima tabela a seguir, ilustra os aspectos da base de dados proposta para uso nesta dissertação e das outras bases mencionadas anteriormente para efeito de comparação.

Na tabela 4.3, apresentamos a comparação entre as bases de dados mencionadas no início desta seção e a base de dados concebida para esta dissertação, observamos alguns pontos que representam a diferença fundamental da base de dados concebida

Tabela 4.3: Tabela comparativa entre a base de dados

Base de dados	Iluminação externa	Iluminação no objeto de interesse	Instabilidade da câmara	Trajetos	Tipos de tomada das imagens	Resolução	Câmara	Visibilidade do objeto de interesse	Tamanho da base
Base de dados proposta	Ruim, boa	Ruim, boa	Sim	linha reta, curva, aiche, declive, ondulações, buracos	Inclinada vertical	1280 × 720 854 × 480 426 × 240	Movimento	Ruim, boa	1542
Base Vision calllech	Boa, excelente	Excelente	Não	Estático	Frontal	896 × 592	Estática	Excelente	126
Base Open ALPR benchmark	Excelente	Boa, excelente	Não	Estático, linha reta, sem buracos, sem ondulações	Frontal, inclinada vertical	1280 × 720, 1920 × 1080, 1000 × 750, 1004 × 768, 640 × 480	Estático	Excelente	839
Base Bangla	Boa, excelente	Boa, excelente	Não	Estático	Frontal, inclinada horizontal	640 × 480	Estático	Excelente	525
Cars Brads	Excelente	Excelente	Não	Linha reta, sem ondulações, buracos	Frontal	360 × 240	Em movimento	Excelente	526
Base Italiana	Excelente	Excelente	Não	Estático	Frontal, inclinada vertical	640 Vines 480	Estático	Excelente	788

para as demais citadas é a questão da câmera em movimento e veículo em movimento em um *background* complexo, apesar de que, a base de dados *Cars Brads* apresente a questão do movimento, no entanto, o trajeto realizado é em linha reta e o asfalto não apresenta buracos ou ondulações, a base de *Bangla*, apresenta condições climáticas com chuva, porém em tomadas frontais e com câmera estática, já a base concebida, apresenta uma resolução mais baixa 426×240 pixels, com condição de tempo nublada sem sol e neste trecho percorrido na aquisição, havia nas imagens presença de sombra, ocasionada pela copa das árvores. A base de dados concebida para esta dissertação contempla 3 tipos de resoluções diferentes sob condições ambientais diversas e foi adquirida em locais diferentes.

Capítulo 5

Procedimento Experimental

Nesta seção, serão abordados os resultados obtidos com método proposto nesta dissertação para detecção de placas em movimento com ambiente não controlado. Apresentamos duas abordagens de resultados distintas, pois, o método proposto foi combinado com algoritmo de super resolução para proporcionar a detecção do objeto de interesse em imagens de baixa resolução. Além disso, reforçamos o fato dos resultados obtidos serem reais.

5.1 Experimentos

Os experimentos foram realizados de acordo com o método proposto apresentado na Seção 3.1 com o propósito de efetuar a detecção de placas de veículos em movimento com *background* complexo de forma automática, no entanto, a avaliação precisa responder aos questionamentos listados a seguir:

1. O método de Super Resolução Bilateral *TVL1*, proporciona a detecção do objeto de interesse?
2. O método proposto, possui a capacidade de detectar placas de veículos considerando imagens com resoluções diferentes?
3. A padronização das resoluções das imagens pode ajudar nas detecções?
4. O método proposto de detecção de placas de veículos em movimento funciona para *background* complexo em todas as imagens contidas na base

Inicialmente o método proposto nesta dissertação, verificou a possibilidade de reconhecimento automático de placas de veículos nas resoluções originais das imagens, no entanto, ao realizar a primeira rodada de experimentos, foi observado que o método proposto conseguiu detectar placas de veículos nas seguintes resoluções, 854×480 e 1280×720 pixels e foi constatado que não foi possível detectar as placas de veículos em imagens com resolução mais baixa, desta forma, utilizamos o algoritmo de Super Resolução Bilateral $TV - L1$ nas imagens da base de dados, resultando em uma única resolução de 1920×1020 pixels em todas as imagens, posteriormente, empregamos novamente o algoritmo proposto para detecção automática de placa de veículos e constatamos que o emprego do método favoreceu a detecção do objeto de interesse.

5.2 Setup dos experimentos

Os experimentos desta dissertação foram implementados utilizando sistema operacional Linux 16.04 *LTS* 64 bits, um computador com processador Intel(R) Core(TM) i3 - 3217 CPU @ 2,0GHz e memória RAM 4GB. Para desenvolvimento dos algoritmos, empregamos a linguagem de programação C/C++ na IDE *Qt Creator* em conjunto com a biblioteca *OpenCV* na versão 3.4.5. O *OpenCV* é uma biblioteca de visão computacional *open source* (código aberto) escrito na linguagem de programação C e C++, oferecendo suporte nos ambientes *Linux*, *Microsoft*, *Mac OS*, plataformas móveis como *iOS* e *Android*, além disso, encontramos versões de interfaces para *Python* e *Java*. A infraestrutura da biblioteca trabalha com um ótimo grau de desempenho, apresentando vantagens ao utilizar processamento *multi-core* quando disponível, pois, garante a otimização das aplicações. O emprego da biblioteca na linguagem C++ proporciona as principais vantagens do padrão orientado a objetos como escrita de código em poucas linhas e torna a aplicação mais legível, assim, optamos pela linguagem C++ no desenvolvimento do projeto desta dissertação.

A base de dados utilizada nesta dissertação foi concebida pelo próprio autor e descrita na Seção 4.3 do Capítulo 4, assim, para os experimentos utilizamos 1542 imagens sequenciais da base de dados, contemplando as três diferentes resoluções

das imagens.

5.2.1 Combinação de experimentos

Para realizarmos os experimentos nesta dissertação, foi empregado inicialmente quatro algoritmos com finalidades distintas e estão descritos a seguir:

(i) O primeiro algoritmo foi empregado nas 1542 imagens sequenciais da base de dados utilizada nesta dissertação, com a finalidade de realizar a marcação das quatro coordenadas dos vértices das placas dos veículos, assim, as coordenadas fornecidas, formam o retângulo correspondente ao formato da placa do veículo, além disso, esta operação é denominada por rótulo manual resultando em uma imagem com marcações em cruz na cor vermelha em cada vértice da placa do veículo;

(ii) O segundo algoritmo, da mesma forma que o anterior, foi empregado nas 1542 imagens sequenciais da base de dados utilizada nesta dissertação e tem o propósito de realizar a detecção automática do objeto de interesse, no entanto, este algoritmo realiza uma segunda tarefa, onde na qual, se encarrega de fornecer de forma automática a informação das quatro coordenadas dos vértices do objeto de interesse detectado, esta operação é denominada por rótulo automático, pois, a união dos vértices resultam na formação do retângulo que corresponde ao formato da placa do veículo;

(iii) O terceiro algoritmo corresponde ao emprego do bloco de Super Resolução bilateral $TV - L1$, nos vídeos da base de dados, principalmente nos vídeos com baixa resolução, para proporcionar a detecção do objeto de interesse que está em movimento com *background* complexo;

(iv) O quarto algoritmo tem o propósito de calcular a interseção das áreas formadas pelas coordenadas fornecidas pelo rótulo manual e rótulo automático, com base nesta informação defini-se as métricas de acerto e erro. As informações deste algoritmo será detalhada posteriormente.

A princípio, em um primeiro cenário de experimentos, nós utilizamos o diagrama do método proposto na Seção 3.1 para todas as 1542 imagens sequenciais da base de dados e foram carregadas neste algoritmo de forma automática, no entanto, neste primeiro cenário não foi empregado o bloco de SR bilateral $TV - L1$, além disso, foi empregado como entrada deste bloco as sequências de imagens provenien-

tes dos vídeos adquiridos para a base de dados. O propósito desta primeira rodada de experimentos era verificar a eficiência de acertos nas detecção das placas de veículos de forma automática, fornecendo ao final de cada teste as coordenadas dos vértices das detecções realizadas em cada imagem. Além disso, para cada imagem empregada neste primeiro experimento, foi empregado os algoritmos (i),(ii) e (iv) conforme descritos anteriormente.

Em um segundo cenário, inicialmente foi empregado o algoritmo (iii) de SR bilateral $TV - L1$ nos vídeos da base de dados, foi realizado a extração dos *frames*, resultando na mesma quantidade de imagens e da mesma forma que fora mencionado no primeiro cenário, foi utilizado 1542 imagens de forma automática como entrada no algoritmo, (ii) do método proposto citado na Seção 3.1, resultando em uma imagem com placa do veículo detectada automaticamente e com as coordenadas dos vértices oriundo da detecção, caracterizando esta tarefa como rótulo automático, além disso, o algoritmo (i) foi utilizado com o propósito de realizar as marcações manuais nas imagens da base de dados, o chamado rótulo manual e o algoritmo (iv), foi utilizado com o propósito de realizar o cálculo de interseção entre as duas coordenadas fornecidas. A tabela a seguir apresenta os resultados preliminares das detecções individuais de cada grupo de imagens com dois cenários apresentados, inicialmente resultados sem o uso da Super Resolução e o segundo resultado com o emprego da Super Resolução.

Na tabela 5.1, apresentamos os resultados preliminares de cada grupo de imagens, onde, inicialmente foi contabilizado apenas as taxas de detecção e desprezou-se os resultados com as métricas, pois, estes resultados serão apresentados posteriormente. Em ambos cenários foi empregado o algoritmo do método proposto na Seção 3.1, além disso, foi mantido os parâmetros dos filtros utilizados, como por exemplo, o filtro *blur* foi empregado a máscara de tamanho 5×5 , o filtro *Sobel* utilizou-se a máscara 3×3 e em uma única direção, devido a alta densidade vertical contida na placa de licenciamento veicular, conforme fora mencionado anteriormente, para o filtro morfológico, empregou-se o elemento estruturante regular com tamanho 17×3 e por fim, definiu-se a conectividade 8 do algoritmo de preenchimento e enchimento.

Tabela 5.1: Taxas de acerto por grupo de imagens em dois cenários distintos.

Grupos de imagem	Sequencia de Imagem	Quantidade de Imagem	Taxa de acesso sem Super Resolução	Taxa de execução com Super Resolução
G1	1 - 52	51	9.8%	11.8%
G2	52 - 102	51	3.9%	5.9%
G3	103 - 153	51	11.5%	11.5%
G4	154 - 199	46	2.2%	4.3%
G5	200 - 250	52	0%	0%
G6	251 - 301	51	11.8%	11.8%
G7	302 - 352	51	0%	0%
G8	353 - 403	51	5.9%	7.8%
G9	404 - 454	51	19.6%	21.6%
G10	455 - 505	51	7.8%	9.8%
G11	506 - 556	51	7.8%	7.8%
G12	557 - 606	50	2%	6%
G13	607 - 657	51	6%	10%
G14	658 - 707	50	0%	8%
G15	708 - 756	49	0%	6%
G16	757 - 805	49	0%	2%
G17	806 - 853	48	0%	2.1%
G18	854 - 903	50	0%	2%
G19	904 - 952	49	0%	0%
G20	953 - 1000	48	0%	2.1%
G21	1001 - 1050	50	0%	10%
G22	1051 - 1098	48	2.1%	4.2%
G23	1099 - 1148	50	6%	10%
G24	1149 - 1197	49	6.1%	4.1%
G25	1198 - 1245	48	2.1%	14.6%
G26	1246 - 1293	48	0%	2.1%
G27	1294 - 1343	50	30%	28%
G28	1344 - 1392	49	4.1%	4.1%
G29	1393 - 1451	59	20.3%	55.9%
G30	1452 - 1542	91	12.1%	17.6%

5.3 Métricas de desempenho

As métricas foram calculadas para os dois cenários mencionados anteriormente, onde, ao realizar o experimento para o primeiro cenário, foi empregado o algoritmo do diagrama do método proposto na Seção 3.1 para todas as 1542 imagens sequenciais da base de dados e foi combinado com o algoritmo responsável pelo cálculo da métricas, da mesma forma, realizou-se este experimento para o segundo cenário, com uma diferença, ao realizar o experimento em um segundo cenário, nós adicionamos o algoritmo de SR bilateral $TV - L1$.

Nesta dissertação, as métricas de desempenho foram calculadas baseando-se nas interseções das áreas dos retângulos formadas pelos rótulos manuais e automáticos respectivamente. O diagrama a seguir, ilustra os passos envolvidos neste algoritmo.

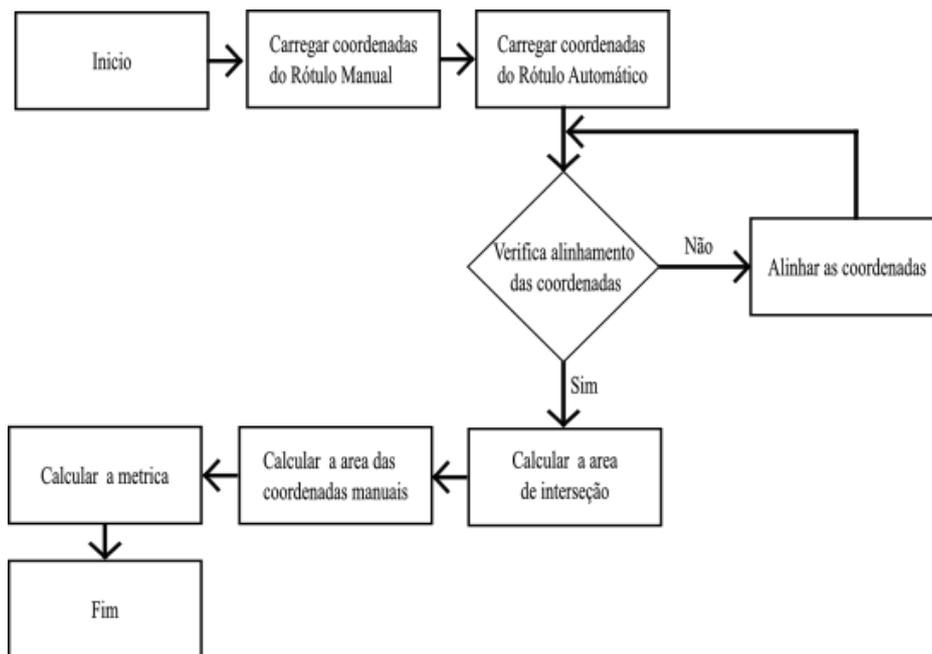


Figura 5.1: Diagrama de blocos utilizado para cálculo de interseção de das coordenadas manuais e automáticas

Na Figura 5.1, foi apresentado o algoritmo utilizado para o cálculo das métricas, inicialmente o algoritmo considera entrada as coordenadas dos vértices manuais e automático, na sequência, verifica se estão alinhadas, caso contrário, o algoritmo alinha as coordenadas automaticamente, a seguir, realiza o cálculo das áreas que correspondem a área de interseção e o área do retângulo formado pelos vértices manuais, por fim realiza o cálculo da métrica. A seguir, apresentamos a equação

utilizada para o cálculo da métrica.

$$\text{Calculo metrica} = \frac{\text{AreaI}}{\text{AreaCM}} \times (100) \quad (5.1)$$

Onde na equação (5.1), AreaI , representa a área da interseção e AreaCM , representa a área das coordenadas manuais fornecidas no rótulo manual. Assim, nesta dissertação as métricas de desempenho foram divididas em quatro categorias listadas a seguir:

- Errado: quando a interseção é igual a zero;
- Ruim: quando a interseção é maior que 70% e menor 80%;
- Boa: quando a interseção é maior que 80% e menor que 90%;
- Excelente: quando a interseção é maior que $\geq 90\%$.

5.4 Análises

Nesta seção, apresentamos os resultados dos experimentos realizado em dois cenários, onde, foi levado em consideração o emprego da Super Resolução na base de dados desta dissertação. Estes resultados foram apresentados contendo as respectivas métricas de desempenho, onde, foi denominado por percentual de interseção da áreas entre os dois retângulos. Assim, os gráficos mostram a diferença dos resultados dos experimentos nos dois cenários apresentados, primeiro cenário, por exemplo, não houve detecções em imagens de baixa resolução, no entanto, ao realizar o segundo experimento em um segundo cenário empregando a SR bilateral $TV - L1$, como fator de impacto nos resultados, foi observado que houve detecções em imagens que anteriormente não houveram detecções, como por exemplo, imagens em baixa resolução e em algumas sequências de imagens isoladas, no entanto, em alguns outros casos, os resultados se mantiveram inalterados para os dois cenários.

Considerando que foi utilizado 1542 imagens sequencias nos dois cenários, os resultados, podem não parecer satisfatórios, no entanto, considera-se que, as detecções ocorreram com imagens sequencias proveniente de vídeos adquiridos em um *background* complexo e a aquisição foi realizada com câmera e objeto de interesse em

movimento, onde, foi observado que as detecções ocorreram em um dado momento da sequência, havendo casos em que, foi realizado as detecções sucessivas e casos em que foi realizado apenas uma única detecção, ou ainda, casos em que não houve nenhuma detecção, como por exemplo o grupo de imagens *G7* e *G19*, pois, para estes casos específicos, o objeto de interesse em questão estava fora do raio de alcance. O conjunto de tabelas a seguir, ilustra os resultados correspondentes a cada grupo de imagens que são equivalente as sequência de imagens proveniente dos vídeos da base de dados e com as respectivas métricas utilizadas e divididas em dois cenários sem Super Resolução e com o emprego da Super Resolução.

Tabela 5.2: Métricas de desempenho correspondente aos grupos de imagem G1 ao G5 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %
G1	5	5	5	5	5	5	6	6	6	6	6	6
G2	3	3	3	3	1	0	3	3	3	2	2	2
G3	2	2	2	2	1	0	1	1	1	1	1	1
G4	1	1	1	1	1	1	2	2	2	1	0	0
G5	1	1	1	1	1	1	0	0	0	0	0	0

Tabela 5.3: Métricas de desempenho correspondente aos grupos de imagem G6 ao G10 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %
G6	6	6	6	6	6	6	6	6	6	6	6	6
G7	0	0	0	0	0	0	0	0	0	0	0	0
G8	3	3	3	3	2	0	4	4	4	4	4	4
G9	10	10	10	10	10	5	11	11	11	10	10	8
G10	4	4	4	3	2	1	5	5	5	4	4	4

Tabela 5.4: Métricas de desempenho correspondente aos grupos de imagem G11 ao G15 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %
G11	4	4	4	4	4	3	3	3	3	3	2	1
G12	1	1	1	1	1	0	3	3	3	3	3	2
G13	3	3	3	3	3	2	5	5	5	5	5	4
G14	0	0	0	0	0	0	4	4	4	4	4	3
G15	0	0	0	0	0	0	3	3	3	3	3	1

Tabela 5.5: Métricas de desempenho correspondente aos grupos de imagem G16 ao G20 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %
G16	0	0	0	0	0	0	1	1	1	1	1	1
G17	0	0	0	0	0	0	1	1	1	1	1	1
G18	0	0	0	0	0	0	1	1	1	1	1	1
G19	0	0	0	0	0	0	0	0	0	0	0	0
G20	0	0	0	0	0	0	1	1	1	1	1	0

Tabela 5.6: Métricas de desempenho correspondente aos grupos de imagem G21 ao G25 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	95
G21	0	0	0	0	0	1	5	5	5	5	3	1
G22	1	1	1	1	0	0	2	2	2	2	1	1
G23	1	1	1	1	1	1	5	5	5	5	5	4
G24	3	3	3	3	2	1	3	3	3	3	3	2
G25	1	1	1	1	1	1	7	7	7	7	7	7

Tabela 5.7: Métricas de desempenho correspondente aos grupos de imagem G26 ao G30 em dois cenários distintos.

Grupos deImagem	Métricas de Desempenho sem Super Resolução						Métricas de Desempenho com Super resolução					
	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %	≥ 70 %	≥ 75 %	≥ 80 %	≥ 85 %	≥ 90 %	≥ 95 %
G26	0	0	0	0	0	0	1	1	1	1	1	1
G27	15	15	15	14	10	1	14	10	10	9	8	1
G28	1	1	1	1	1	1	1	1	1	1	1	1
G29	12	12	12	12	9	1	33	25	22	13	9	1
G30	11	10	9	8	6	3	16	15	14	13	11	8

Capítulo 6

Considerações finais

Nesta dissertação, foi abordado o emprego de técnicas de processamento digital de imagens através da biblioteca *opencv*, para a detecção automática de placas de veículos em movimento e em um *background* complexo, baseando-se em pesquisas de outros autores com o propósito de realizar e empregar um método robusto e flexível.

O Capítulo 2, apresentamos a fundamentação teórica relacionada a pesquisa, mencionamos conceitos sobre filtros e suas categorias, vale ressaltar a importância dos filtros nesta pesquisa, pois representam as técnicas de processamento utilizadas na metodologia desta dissertação, mencionamos ainda conceitos sobre técnicas de Super Resolução, no Capítulo 3, mencionamos sobre o método proposto para esta dissertação, onde procuramos deixar clara a proposta para o leitor através de um diagrama de blocos, que apresenta um panorama geral da proposta aliada ao emprego do algoritmo de Super Resolução. O capítulo 4, apresenta de forma detalhada a base de dados empregada nesta pesquisa, onde, o detalhamento consiste na descrição do cenário, o ambiente em esta base de dados foi concebida, assim como, descrevemos as sequências de imagens envolvidas nesta dissertação, as resoluções das imagens e mencionamos a quantidade de imagens utilizadas nos experimentos. O restante da base de dados pode ser empregado para investigações de trabalhos futuros, servindo de alicerce para pesquisas futuras, pois existem aspectos não abordados nesta dissertação.

Em termos de experimentos, realizamos os testes de validação em dois cenários, o primeiro cenário, observamos que o método proposto apresentou capacidade

de detectar placas de veículos apenas nas sequências de imagens nas seguintes resoluções: 854×480 e 1280×720 pixels, no entanto, o segundo cenário empregamos o algoritmo de Super Resolução Bilateral $TV - L1$, como estratégia para realizar a detecção da placa de veículos em imagens com a seguinte resolução; 426×240 pixels, além disso, o emprego deste método, nos proporcionou uma melhora sutil nas taxas de detecções, pois, foi possível detectar placas de veículos nestas imagens com baixa resolução.

Contudo, observou-se que foi possível estabelecer um método baseado em técnicas de processamentos digital de imagens para realizar a detecção de placas de veículos com o objeto de interesse e câmera em movimento diante das características que compõe a base de dados utilizada nesta pesquisa, além disso, a base de dados concebida para esta pesquisa, será empregada em trabalhos futuros, assim como este trabalho de pesquisa servirá como alicerce para trabalhos futuros, da mesma forma, constatou-se que os experimentos apresentaram resultados em cenários distintos considerando o emprego da Super Resolução, resultando como impacto a detecção do objeto de interesse em imagens com resolução de 426×240 . Por fim, a metodologia proposta nesta dissertação obteve alguns resultados diante da dificuldade de emprego de técnicas de processamento em imagens com *background* complexo propicio a ruídos.

6.1 Trabalhos Futuros

- Aumentar as possibilidades de cenários com outras resoluções superiores, além disso, verificar a eficácia do método proposto para novas placas de veículos com padrão Mercosul;
- Foi investigado a integração do módulo *OCR* para cada placa detectada, pois, verificou-se que é possível realizar a segmentação de cada caractere da placa de licenciamento veicular utilizando um algoritmo de aprendizagem de máquina empregando Redes Neurais para atuar no reconhecimento do caractere. Inicialmente, obtem-se um *patch* de imagem da placa de licenciamento veicular como entrada de dado em uma função *OCR* de segmentação utilizando um histograma equalizado, posteriormente, aplica-se a limiarização nesta imagem

e emprega-se o algoritmo *Find Contours*, onde, para cada contorno detectado, realiza-se uma verificação e remove-se todas as regiões em que o tamanho é menor que a relação de aspecto, em nosso caso específico, a relação de aspecto equivale a 400/130. Assim, se um caractere segmentado, for verificado, é necessário realizar o pré processamento para defini o tamanho e posição do mesmo, pois a função *Find Contours* não retorna os dados desta operação de forma ordenada, para isso, utiliza-se a classe *CharSegment*.

- Foi investigado uma forma de realizar a detecção automática de placas utilizando aprendizagem profunda, inicialmente, emprega-se a arquitetura de detecção de objetos de aprendizagem profunda *Yolo (YouOnlyLookOne)* baseada em redes neurais de convolução e é treinada de ponta a ponta para realizar tarefa de regressão prevendo a detecção da caixa delimitadora do objeto de interesse, a rede apresenta capacidade de processar imagens em tempo real a 45 quadros por segundo. No entanto, para realizar tal tarefa, inicialmente, realiza-se um treinamento com um determinado número de imagens, posteriormente, efetua-se a segmentação de placas, por fim, emprega-se a fase de reconhecimento da placa que representa uma etapa de extrema importância, resultando em imagens com caracteres obtidos ao final da segmentação. O modelo de aprendizagem utilizado no reconhecimento terá a capacidade de ler uma imagem renderizar o caractere correspondente.

Referências Bibliográficas

- [1] PEDRINI, H., SCHWARTZ, W. R., *Análise de imagens digitais: princípios, algoritmos e aplicações*. Thomson Learning, 2008.
- [2] GONZALEZ, R. C., WOODS, R. C., *Processamento digital de imagens ..* Pearson Educación, 2009.
- [3] MELNYK, R., KALYCHAK, Y. I., “Detection of Defects in Printed Circuit Boards by Flood-Fill Algorithm and Distributed Cumulative Histogram”. In: *2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, v. 1, pp. 155–159, 2018.
- [4] DOLEZEL, P., SKRABANEK, P., GAGO, L., “Detection of grapes in natural environment using feedforward neural network as a classifier”. In: *2016 SAI Computing Conference (SAI)*, pp. 1330–1334, July 2016.
- [5] ASIF, M. R., CHUN, Q., HUSSAIN, S., et al., “Multinational vehicle license plate detection in complex backgrounds”, *Journal of Visual Communication and Image Representation*, v. 46, pp. 176–186, 2017.
- [6] WANG, R., SANG, N., HUANG, R., et al., “License plate detection using gradient information and cascade detectors”, *Optik*, v. 125, n. 1, pp. 186–190, 2014.
- [7] ALYANNEZHADI, M. M., HASHEMI, S. M. R., ABOLGHASEMI, V., “License plate detection in complex scenes based on fusion of Gaussian filtering and Bayesian network”. In: *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, pp. 0022–0026, 2017.

- [8] KHIN, O., PHOTHISONOTHAI, M., CHOOMCHUAY, S., “License Plate Detection of Myanmar Vehicle Images Captured from the Dissimilar Environmental Conditions”. In: *2017 International Conference on Advanced Computing and Applications (ACOMP)*, pp. 127–132, 2017.
- [9] GEORGIEVA, V., PETROV, P., DIMITROV, L., “A Multistage Approach for Detection of Objects with Rectangular Forms”. In: *2018 International Conference on High Technology for Sustainable Development (HiTech)*, pp. 1–3, 2018.
- [10] LESZCZYŃSKI, B., GANCARCZYK, A., WRÓBEL, A., et al., “Global and local thresholding methods applied to X-ray microtomographic analysis of metallic foams”, *Journal of Nondestructive Evaluation*, v. 35, n. 2, pp. 35, 2016.
- [11] CHAUBEY, A. K., “Comparison of the local and global thresholding methods in image segmentation”, *World Journal of Research and Review*, v. 2, n. 1, 2016.
- [12] PFLEGER, S. G., OTHERS, “Redução de ruído em vídeos em tempo real baseado na fusão do filtro de Kalman e filtro bilateral”, 2016.
- [13] FEITOSA, F., “Um estudo prático para contagem volumétrica automática de veículos usando Visão Computacional”, *Master’s thesis, Universidade Federal de Goiás*, 2012.
- [14] CHEN, P., LI, Y. X., “The basic concepts and types of image”. In: *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, v. 1, pp. 313–315, 2011.
- [15] BURGER, W., BURGE, M. J., *Digital image processing: an algorithmic introduction using Java*. Springer, 2016.
- [16] CARVALHO, J. N. D., OTHERS, “Seleção e classificação inteligente de mangas por análise de imagens.” 2015.
- [17] CHEN, F.-Q., ZHOU, Y.-P., “Color feature extraction of Hainan Li brocade image based on RGB and HSV”. In: *2015 12th International Computer*

- Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 214–219, 2015.
- [18] AHIRWAL, B., KHADTARE, M., MEHTA, R., “FPGA based system for color space transformation RGB to YIQ and YCbCr”. In: *2007 International Conference on Intelligent and Advanced Systems*, pp. 1345–1349, 2007.
- [19] MAGDA, E. F., COSMIN, S., IOAN, L. A., “Image search algorithms”. In: *2015 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. P–35, 2015.
- [20] FEITOSA, F., “Um estudo prático para contagem volumétrica automática de veículos usando Visão Computacional”, *Master’s thesis, Universidade Federal de Goiás*, 2012.
- [21] RABINER, L. R., “The design of finite impulse response digital filters using linear programming techniques”, *Bell System Technical Journal*, v. 51, n. 6, pp. 1177–1198, 1972.
- [22] NALON, J. A., *Introdução ao processamento digital de sinais*. Grupo Gen-LTC, 2000.
- [23] KUMAR, N., “Optimal design of fir and iir filters using some evolutionary algorithms”, *Master of Technology Thesis in Electrical Engineering, National Institute of Technology, Durgapur, India, (May 2013)*, 2013.
- [24] PATEL, N., SHAH, A., MISTRY, M., et al., “A study of digital image filtering techniques in spatial image processing”. In: *Proceedings of the 2014 International Conference on Convergence of Technology (I2CT)*, pp. 1–6, 2014.
- [25] NEGI, S. S., GUPTA, B., “Survey of various image enhancement techniques in spatial domain using matlab”. In: *International Conference on Advances in Computer Engineering and applications-20 14 (ICACEA-2014), Special Proceedings of International Journal of Computer Applications (IJCA)*, pp. 8–18, 2014.

- [26] FEITOSA, F., “Um estudo prático para contagem volumétrica automática de veículos usando Visão Computacional”, *Master’s thesis, Universidade Federal de Goiás*, 2012.
- [27] MALIK, F., BAHARUDIN, B., “Image Retrieval Using Color Histogram Refinement Method”, *Journal of Applied Sciences*, v. 12, n. 5, pp. 416–427, 2012.
- [28] JAMES, A. P., “Edge detection for pattern recognition: a survey”, *arXiv preprint arXiv:1602.04593*, 2016.
- [29] ACHARJYA, P. P., DAS, R., GHOSHAL, D., “Study and comparison of different edge detectors for image segmentation”, *Global Journal of Computer Science and Technology*, 2012.
- [30] JOSE, A., DIXON, K. D. M., JOSEPH, N., et al., “Performance study of edge detection operators”. In: *2014 International Conference on Embedded Systems (ICES)*, pp. 7–11, 2014.
- [31] GANESAN, P., SAJIV, G., “A comprehensive study of edge detection for image processing applications”. In: *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pp. 1–6, 2017.
- [32] SOUZA DE ANDRADE, F., “FILTRAGEM DE IMAGENS EM ESCALA DE CINZA POR MEIO DE REDE NEURONAL CELULAR ANALOGICA EM TECNOLOGIA CMOS”, *Master’s thesis, Universidade Federal da Bahia*, 2016.
- [33] LEDESMA-CARRILLO, L. M., LOPEZ-RAMIREZ, M., CABAL-YEPEZ, E., et al., “FPGA-based reconfigurable unit for image encryption using orthogonal functions”. In: *2016 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, pp. 168–173, 2016.
- [34] MAKANDAR, A., HALALLI, B., “Image enhancement techniques using high-pass and lowpass filters”, *International Journal of Computer Applications*, v. 109, n. 14, pp. 21–27, 2015.

- [35] SINHA, G. R., “Design and Implementation of Image Enhancement Techniques in Frequency Domain”, 2009.
- [36] PFLEGER, S. G., OTHERS, “Redução de ruído em vídeos em tempo real baseado na fusão do filtro de Kalman e filtro bilateral”, 2016.
- [37] VANKAWALA, F., GANATRA, A., PATEL, A., “A survey on different image deblurring techniques”, *International Journal of Computer Applications*, v. 116, n. 13, pp. 15–18, 2015.
- [38] GAJJAR, R., ZAVERI, T., SHUKLA, A., “Invariants based blur classification algorithm”. In: *2015 5th Nirma University International Conference on Engineering (NUiCONE)*, pp. 1–5, 2015.
- [39] ZHANG, H., LIU, H., SHANG, Z., et al., “The Basic Frequency Domain Recovery Algorithm for Image De-blurring”. In: *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, v. 1, pp. 470–473, 2015.
- [40] FLUSSER, J., FAROKHI, S., HÖSCHL, C., et al., “Recognition of images degraded by Gaussian blur”, *IEEE Transactions on Image Processing*, v. 25, n. 2, pp. 790–806, 2015.
- [41] KARAALI, A., “Spatially varying defocus blur estimation and applications”, 2017.
- [42] KURUVILLA, J., SUKUMARAN, D., SANKAR, A., et al., “A review on image processing and image segmentation”. In: *2016 international conference on data mining and advanced computing (SAPIENCE)*, pp. 198–203, 2016.
- [43] KUMAR, N., “Thresholding in salient object detection: a survey”, *Multimedia Tools and Applications*, v. 77, n. 15, pp. 19139–19170, 2018.
- [44] DENMAN, S., CHANDRAN, V., SRIDHARAN, S., “An adaptive optical flow technique for person tracking systems”, *Pattern recognition letters*, v. 28, n. 10, pp. 1232–1239, 2007.

- [45] SINGH, T. R., ROY, S., SINGH, O. I., et al., “A new local adaptive thresholding technique in binarization”, *arXiv preprint arXiv:1201.5227*, 2012.
- [46] SATAPATHY, S. C., RAJA, N. S. M., RAJINIKANTH, V., et al., “Multi-level image thresholding using Otsu and chaotic bat algorithm”, *Neural Computing and Applications*, pp. 1–23, 2018.
- [47] TRUONG, M. T. N., KIM, S., “Automatic image thresholding using Otsus method and entropy weighting scheme for surface defect detection”, *Soft Computing*, v. 22, n. 13, pp. 4197–4203, 2018.
- [48] YUAN, X.-C., WU, L.-S., PENG, Q., “An improved Otsu method using the weighted object variance for defect detection”, *Applied surface science*, v. 349, pp. 472–484, 2015.
- [49] YUAN, C., LIU, Z., ZHANG, Y., “UAV-based forest fire detection and tracking using image processing techniques”. In: *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 639–643, 2015.
- [50] PANDEY, R. K., MATHURKAR, S., “A Review on Morphological Filter and its Implementation”, *International Journal of Science and Research (IJSR)*, v. 2, n. 1, pp. 69–72, 2017.
- [51] SOILLE, P., *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.
- [52] ZHANG, Y., TAO, M., YANG, K., et al., “Video Superresolution reconstruction using iterative back projection with critical-point filters based image matching”, *Advances in Multimedia*, v. 2015, pp. 4, 2015.
- [53] KHATTAB, M. M., ZEKI, A. M., ALWAN, A. A., et al., “Regularization-based multi-frame super-resolution: A systematic review”, *Journal of King Saud University-Computer and Information Sciences*, 2018.
- [54] VICENTE, A. N., “Método de super-resolução de única figura baseado em aprendizado para imagens de placas brasileiras de baixíssima qualidade”, 2016.

- [55] MASOOD, S. Z., SHU, G., DEHGHAN, A., et al., “License plate detection and recognition using deeply learned convolutional neural networks”, *arXiv preprint arXiv:1703.07330*, 2017.
- [56] AZAM, S., ISLAM, M. M., “Automatic license plate detection in hazardous condition”, *Journal of Visual Communication and Image Representation*, v. 36, pp. 172–186, 2016.
- [57] POLISHETTY, R., ROOPA EI, M., RAD, P., “A next-generation secure cloud-based deep learning license plate recognition for smart cities”. In: *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 286–293, 2016.
- [58] RIZVI, S., PATTI, D., BJÖRKLUND, T., et al., “Deep classifiers-based license plate detection, localization and recognition on GPU-powered mobile platform”, *Future Internet*, v. 9, n. 4, pp. 66, 2017.