



## UM CLASSIFICADOR VISUAL AUTÔNOMO DE DENSIDADE DE TRÁFEGO URBANO

Gabriel Marques Rosario

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientador: Felipe Maia Galvão França

Rio de Janeiro  
Setembro de 2012

UM CLASSIFICADOR VISUAL AUTÔNOMO DE DENSIDADE DE TRÁFEGO  
URBANO

Gabriel Marques Rosario

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO  
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE  
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE  
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A  
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE  
SISTEMAS E COMPUTAÇÃO.

Examinada por:

---

Prof. Felipe Maia Galvão França, Ph.D.

---

Prof. Ricardo Cordeiro de Farias, Ph.D.

---

Prof. Paulo Cezar Martins Ribeiro, Ph.D.

---

Prof. Massimo De Gregorio, Ph.D.

RIO DE JANEIRO, RJ – BRASIL  
SETEMBRO DE 2012

Rosario, Gabriel Marques

Um Classificador Visual Autônomo de Densidade de Tráfego Urbano/Gabriel Marques Rosario. – Rio de Janeiro: UFRJ/COPPE, 2012.

XIII, 96 p.: il.; 29, 7cm.

Orientador: Felipe Maia Galvão França

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2012.

Referências Bibliográficas: p. 81 – 83.

1. Sensoriamento de tráfego. 2. Computação visual.  
3. Sistemas Inteligentes de Transporte. 4. Redes  
Neurais Sem Peso. I. França, Felipe Maia Galvão.  
II. Universidade Federal do Rio de Janeiro, COPPE,  
Programa de Engenharia de Sistemas e Computação. III.  
Título.

# Agradecimentos

Um agradecimento aos meus pais, Janete e José, que sempre estiveram ao meu lado, oferecendo apoio de todas as maneiras possíveis. À minha irmã, Flávia, por ser essa pessoa incrível e atenciosa.

Aos amigos de longa data, Daniel Maia, Felipe Pollola, Rafael Cidade, Raul Palermo e os irmãos Ricardo, Flávio e Pedro Barreto, amigos que até hoje caminham ao meu lado, sempre dispostos a me ajudar.

Aos amigos que fiz na UFRJ e na COPPETEC, por todo o companheirismo e compreensão nesses longos anos: Gizelle Gaspar, Lúcio Paiva, Glauber Menezes, Alexandre Papanis, Vitor Gamboa, Carlos Gomes, Daniel Schneider, Rafael Jardim e muitos outros.

A todos os meus familiares, em especial aqueles com quem mais tenho contato. Nathália, Nádia, Lethícia e Ademário Santana; Fátima, Lourdes e Américo Rosário; e Lady, muito obrigado a todos vocês por serem a família que eu preciso.

Ao professor Felipe França, por todo o apoio, os ensinamentos e as conversas bem humoradas.

E, por fim, aos professores Ricardo Farias, Paulo Cezar e Massimo de Gregorio, por terem aceitado o convite de fazer parte desse trabalho.



Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## UM CLASSIFICADOR VISUAL AUTÔNOMO DE DENSIDADE DE TRÁFEGO URBANO

Gabriel Marques Rosario

Setembro/2012

Orientador: Felipe Maia Galvão França

Programa: Engenharia de Sistemas e Computação

O conceito de Sistemas Inteligentes de Transporte, oriundo da revolução na computação e nas comunicações ocorrida nos anos 90, vem sendo cada vez mais discutido em Engenharia de Tráfego como um meio de se aprimorar os sistemas de transporte existentes através das facilidades oferecidas pelas novas ferramentas tecnológicas. Uma das aplicações possíveis consiste em um sistema de controle de sinalização adaptativo capaz de tomar decisões, em tempo real, baseado no comportamento do tráfego em uma cidade e assim controlar a temporização dos semáforos de forma inteligente, melhorando o fluxo de tráfego e reduzindo congestionamentos. Porém, tal solução requer um meio de se extrair das vias observadas informações como taxa de fluxo ou densidade, para que então suas decisões possam ser tomadas. Levando-se em conta a grande difusão de câmeras comuns de tráfego e o preço razoável com que se obtém um alto poder computacional atualmente, este trabalho propõe uma alternativa rápida, eficiente e de baixo custo às soluções comerciais para determinação da situação em uma via de tráfego. Baseada puramente na análise de imagens, a solução desenvolvida, que trabalha de maneira autônoma, requisitando o mínimo de interação humana, é capaz de determinar, segundo a segundo, a densidade do fluxo de tráfego na via analisada. Além disso, foi desenvolvido um segundo método, utilizando redes neurais sem peso, para determinação da situação presente na via baseado na relação entre o fluxo e a densidade.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

## AN AUTONOMOUS VISUAL CLASSIFIER OF URBAN TRAFFIC DENSITY

Gabriel Marques Rosario

September/2012

Advisor: Felipe Maia Galvão França

Department: Systems Engineering and Computer Science

The concept of Intelligent Transportation Systems, which originated in the revolution in computing and communications occurred in the '90s, is being increasingly discussed in Traffic Engineering as a means to enhance existing transportation systems through the facilities offered by new technological tools. One of its possible applications consists in an adaptive signal control system capable of making decisions, in real time, based on the behavior of traffic in a city and then control, in an intelligent manner, the traffic signals timing, improving traffic flow and reducing congestion. However, this solution requires a means of extracting, from the traffic facilities being observed, information like flow rate or density, so it can, then, make its decisions. Taking into account the widespread use of common traffic cameras and the reasonable price at which a high computational power can be obtained, this work proposes a fast, efficient and low cost alternative to commercial solutions in determining the situation in a traffic facility. Purely based on image analysis, the developed solution, which works autonomously, requiring minimal human interaction, is able to determine, second after second, the density of a traffic stream. Moreover, a second method was developed using weightless neural networks for determining the current situation of a traffic facility based on the relationship between flow and density.

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	2
1.2 Trabalhos correlatos . . . . .	3
1.3 Objetivo do trabalho . . . . .	4
1.4 Estrutura do texto . . . . .	4
<b>2 Fundamentos de Engenharia de Tráfego</b>	<b>5</b>
2.1 Introdução . . . . .	5
2.2 Tipos de vias de tráfego . . . . .	6
2.2.1 Fluxo ininterrupto . . . . .	6
2.2.2 Fluxo interrompido . . . . .	6
2.3 Parâmetros de fluxo de tráfego . . . . .	7
2.3.1 Volume e Taxa de Fluxo . . . . .	7
2.3.2 Velocidade e Tempo de Viagem . . . . .	9
2.3.3 Densidade e Ocupância . . . . .	10
2.4 Relações entre Taxa de Fluxo, Velocidade e Densidade . . . . .	10
<b>3 Fundamentos de Processamento de Imagens</b>	<b>13</b>
3.1 Introdução . . . . .	13
3.2 Modelo de cores RGB . . . . .	14
3.3 Filtros morfológicos . . . . .	15
3.4 Segmentação . . . . .	17
3.4.1 Segmentador Adaptativo Pixel a Pixel . . . . .	18
<b>4 Fundamentos de Agrupamento</b>	<b>22</b>
4.1 Introdução . . . . .	22
4.2 K-means . . . . .	23

<b>5</b>	<b>Fundamentos de Redes Neurais Sem Peso</b>	<b>25</b>
5.1	Introdução à Computação Neural . . . . .	25
5.2	Redes Neurais . . . . .	26
5.2.1	Nó Neural . . . . .	26
5.2.2	Rede Associativa . . . . .	29
5.2.3	Modelo McCulloch e Pitts . . . . .	32
5.3	Redes Neurais Sem Peso . . . . .	34
5.3.1	Nó RAM . . . . .	34
5.3.2	Discriminador RAM . . . . .	36
5.3.3	WISARD . . . . .	37
5.3.4	GRAM e VGRAM . . . . .	38
<b>6</b>	<b>Classificador de Densidade</b>	<b>40</b>
6.1	Introdução . . . . .	40
6.2	Classificador de Densidade . . . . .	42
6.2.1	Detecção de objetos . . . . .	42
6.2.2	Detecção de movimento . . . . .	44
6.2.3	Área de Interesse . . . . .	46
6.2.4	Determinação de densidade . . . . .	48
6.2.5	Determinação de velocidade . . . . .	51
6.2.6	Determinação de densidade com o PBAS . . . . .	52
6.2.7	Determinação de densidade com uma VGRAM . . . . .	53
6.2.8	Determinação de <i>Fluxo x Densidade</i> . . . . .	53
6.2.9	Observações . . . . .	55
<b>7</b>	<b>Resultados experimentais</b>	<b>57</b>
7.1	Introdução . . . . .	57
7.2	Determinação de densidade . . . . .	58
7.2.1	Método Privilegiado . . . . .	62
7.2.2	Método Padrão . . . . .	69
7.2.3	Método Automático por Agrupamento . . . . .	72
7.3	Determinação de densidade com o PBAS . . . . .	75
7.4	Determinação de Densidade com uma VGRAM . . . . .	75
7.5	Determinação de <i>Fluxo x Densidade</i> . . . . .	76
<b>8</b>	<b>Conclusão</b>	<b>79</b>
8.1	Trabalhos futuros . . . . .	80
	<b>Referências Bibliográficas</b>	<b>81</b>

<b>A Resultados</b>	<b>84</b>
A.1 Método Privilegiado . . . . .	84
A.2 Método Padrão . . . . .	86
A.3 Método Automático por Agrupamento . . . . .	88
A.4 Método Privilegiado (Média) . . . . .	90
A.5 Método Padrão (Média) . . . . .	92
A.6 Método Autom. por Agrupamento (Média) . . . . .	94

# Lista de Figuras

2.1	Relações entre Fluxo, Velocidade e Densidade . . . . .	11
3.1	Canais de cores do RGB . . . . .	15
3.2	Erosão e dilatação em uma imagem . . . . .	16
3.3	Um elemento de estruturação em forma de um quadrado 3x3 . . . . .	16
3.4	Abertura e fechamento em uma imagem . . . . .	17
4.1	Exemplo de agrupamento utilizando o <i>K-means</i> . . . . .	23
4.2	Agrupamento ineficaz . . . . .	24
5.1	Nó simples . . . . .	27
5.2	Exemplos de redes neurais . . . . .	30
5.3	Uma rede associativa . . . . .	31
5.4	Nó do modelo MCP . . . . .	33
5.5	Nó RAM . . . . .	35
5.6	Um discriminador RAM . . . . .	36
6.1	Interface do sistema . . . . .	41
6.2	Definindo a região de interesse . . . . .	41
6.3	Diferença de imagens binárias . . . . .	43
6.4	Diferença de imagens RGB . . . . .	43
6.5	Diferença de imagens utilizando intervalo de tolerância . . . . .	45
6.6	Diferença de imagens em uma via vazia e em uma congestionada . . . . .	46
6.7	Frequência observada em uma via ao longo de um dia . . . . .	47
6.8	Diferença de imagens consecutivas com movimento . . . . .	48
6.9	Diferença de imagens usando a imagem de via vazia . . . . .	48
6.10	Diferença de imagens com iluminação diferente . . . . .	49
6.11	Iluminações diferentes ao longo do dia . . . . .	50
6.12	Utilização do fechamento . . . . .	51
6.13	Comparação entre a diferença de imagens e o PBAS . . . . .	52
6.14	Detalhes da VGRAM . . . . .	54
6.15	Intervalos de classificação . . . . .	54

6.16	Padrão de entrada . . . . .	55
7.1	Exemplo de teste da densidade . . . . .	59
7.2	Exemplos de imagens para cada densidade . . . . .	61
7.3	Situações que induzem o erro da classificação . . . . .	62
7.4	Exemplos da sobreposição entre fronteiras (7 iterações) . . . . .	66
7.5	Intervalos de transição do primeiro teste . . . . .	66
7.6	Intervalos de transição utilizados . . . . .	68
7.7	Intervalos de transição utilizados para o Método Padrão . . . . .	71
7.8	Intervalos de transição utilizados para o Método Automático por Agrupamento . . . . .	74

# Lista de Tabelas

5.1	Função do nó ou-exclusivo . . . . .	27
5.2	Exemplo de saídas indefinidas . . . . .	28
5.3	Função do nó antes de aplicada a regra de disparo . . . . .	29
5.4	Função do nó depois de aplicada a regra de disparo . . . . .	29
5.5	Função dos nós da rede associativa . . . . .	32
5.6	Função do nó MCP . . . . .	33
5.7	Exemplo do espalhamento na VGRAM . . . . .	39
6.1	Exemplo de intervalos de densidade . . . . .	51
6.2	Média do tempo de execução . . . . .	53
7.1	Média das diferenças para cada quantidade de carros (fechamento com 7 iterações) . . . . .	60
7.2	Mapeamento Densidade-Quantidade de carros . . . . .	60
7.3	Valores usados na classificação média . . . . .	62
7.4	Média das diferenças para cada quantidade de carros . . . . .	63
7.5	Mapeamento Densidade-Diferença para o Método Privilegiado . . . . .	63
7.6	Resultados dos testes para o Método Privilegiado . . . . .	64
7.7	Detalhes do teste para 11 iterações do fechamento (Privilegiado) . . . . .	64
7.8	Detalhes dos testes para diferentes iterações do fechamento (Privilegiado) . . . . .	65
7.9	Resultados dos testes para o Método Privilegiado (Média) . . . . .	65
7.10	Detalhes do teste para 11 iterações do fechamento (Privilegiado, Média) . . . . .	65
7.11	Desvio padrão das diferenças para cada quantidade de carros (11 iterações) . . . . .	67
7.12	Detalhes do primeiro teste com intervalos de transição (Privilegiado) . . . . .	67
7.13	Resultados dos testes com intervalos de transição (Privilegiado) . . . . .	67
7.14	Detalhes dos testes com intervalos de transição (Privilegiado) . . . . .	68
7.15	Mapeamento Densidade-Diferença genérico para o Método Padrão . . . . .	69
7.16	Diferenças máximas observadas . . . . .	70
7.17	Mapeamento Densidade-Diferença para o Método Padrão . . . . .	70



7.18	Resultados dos testes para o Método Padrão . . . . .	70
7.19	Detalhes do teste para 11 iterações do fechamento (Padrão) . . . . .	70
7.20	Resultados dos testes para o Método Padrão (Média) . . . . .	71
7.21	Detalhes do teste para 11 iterações do fechamento (Padrão, Média) . . . . .	71
7.22	Resultados dos testes com intervalos de transição (Padrão) . . . . .	72
7.23	Detalhes dos testes com intervalos de transição (Padrão) . . . . .	72
7.24	Mapeamento Densidade-Diferença para o Método Automático por Agrupamento . . . . .	73
7.25	Resultados dos testes para o Método Automático por Agrupamento . . . . .	73
7.26	Detalhes do teste para 11 iterações do fechamento (Agrupamento) . . . . .	73
7.27	Resultados dos testes para o Método Automático por Agrupamento (Média) . . . . .	74
7.28	Detalhes do teste para 11 iterações do fechamento (Agrupamento, Média) . . . . .	74
7.29	Resultados dos testes com intervalos de transição (Agrupamento) . . . . .	75
7.30	Detalhes dos testes com intervalos de transição (Agrupamento) . . . . .	75
7.31	Resultados dos testes do PBAS para o Método Padrão . . . . .	75
7.32	Mapeamento Densidade-Quantidade de carros para os testes . . . . .	76
7.33	Resultados dos testes da VGRAM utilizando o PBAS . . . . .	76
7.34	Resultados dos testes da VGRAM utilizando comparação de imagens com 9 iterações do fechamento . . . . .	76
7.35	Distribuição dos padrões para cada intervalo . . . . .	77

# Capítulo 1

## Introdução

Desde os anos 90, quando teve início uma revolução na computação e nas comunicações, o conceito de *Sistemas Inteligentes de Transporte* vem sendo cada vez mais discutido em Engenharia de Tráfego. As facilidades trazidas pelos avanços tecnológicos (tanto custo computacional reduzido e miniaturização de processadores, como uso largamente difundido de dispositivos como celulares e GPS) fez com que o foco se voltasse para as maneiras com que os sistemas de transporte pudessem ser aprimorados, baseando-se em todo o trabalho teórico realizado até então, mas se utilizando das vantagens oferecidas pelas relativamente baratas e difundidas ferramentas tecnológicas.

Sistemas Inteligentes de Transporte incluem uma vasta gama de tecnologias e aplicações, que podem ser agrupadas em cinco categorias: *Sistemas Avançados de Informação ao Viajante*, que disponibilizam informações aos motoristas em tempo real, como rotas de tráfego, existência de congestionamentos e obras na pista; *Sistemas Avançados de Gestão de Transporte*, que incluem dispositivos de controle do tráfego, como semáforos, placas de aviso digitais e *Centros de Operações de Tráfego*; *Sistemas de Coleta Automática de Taxas*, como a cobrança de pedágio eletrônica; *Sistemas Avançados de Transporte Público*, que permitem aos usuários do transporte público saberem a exata localização dos veículos (seja via *internet*, seja no próprio ponto de espera); e *Sistemas Inteligentes de Transporte Totalmente Integrados*, integrações que permitem, por exemplo, a comunicação tanto dos veículos com os sensores da via e os semáforos, como entre os próprios veículos [1]. Entre os benefícios oferecidos por todo esse aparato estão o aumento na segurança do trânsito, o aumento no desempenho operacional das vias, especialmente na redução de congestionamentos, os benefícios ambientais, como a redução da emissão de dióxido de carbono, além do aumento de produtividade.

Sistemas Avançados de Gestão de Transporte, que enfocam os dispositivos de controle do tráfego, como semáforos e placas de aviso digitais (que fornecem aos motoristas informações em tempo real sobre as condições do tráfego), também in-

cluem os Centros de Operações de Tráfego, que consistem em centros de gerenciamento centralizado de tráfego controlados por cidades e estados em todo o mundo. Esses centros utilizam as tecnologias de informação para conectar seus dispositivos de tráfego disponíveis, sejam eles sensores e equipamentos das vias, veículos sonda, câmeras, placas de aviso, entre outros, de forma a criar uma visão integrada do fluxo de tráfego e para detectar acidentes, eventos climáticos perigosos e outros perigos possíveis.

Em muitos países, como os Estados Unidos, a grande maioria dos cruzamentos sinalizados utilizam planos de temporização estáticos, desatualizados, baseados em dados coletados anos ou décadas atrás [2]. De fato, estima-se que, de todos os congestionamentos nas maiores rodovias americanas, de 5 a 10 por cento seja atribuído a um plano de sinalização precário. Um controle de sinalização adaptativo, capaz de detectar a presença de veículos (seja de forma autônoma ou por comunicação enviada pelo próprio veículo, através de comunicações dedicadas de curto alcance), aprimoraria a temporização de semáforos, melhorando o fluxo de tráfego e reduzindo congestionamentos [1].

## 1.1 Motivação

Com a revolução nas comunicações, os sensores magnéticos (*loop detectors*), que sempre constaram como o método de detecção de tráfego mais utilizado pelas agências de tráfego, encontraram concorrência em dispositivos como câmeras comuns ou infravermelhas que, aliadas a *softwares* capazes de criar “detectores virtuais”, observam contagem e ocupância em determinados pontos ou áreas da via, além de velocidades [3]. Utilizando tal ferramenta, o profissional de transporte pode desenhar os pontos a serem observados em uma imagem da via e contar com o *software* para processar os dados. Ao passo em que os sensores magnéticos detectam a presença de um veículo pela variação da indutância em seu circuito [4], um detector virtual realiza essa detecção através da análise visual da região determinada pelo usuário, se utilizando de técnicas de processamento de imagens.

Mesmo sendo uma solução mais barata que sensores magnéticos (que, além de caros, podem apresentar dificuldade de instalação em determinados cruzamentos), as soluções para sensoriamento inteligente disponíveis no mercado envolvem *hardware* proprietário, como câmeras, placas e bastidores (*racks*) computacionais, que representam certo custo de implementação e manutenção. Além disso, essas soluções carecem de documentação sobre sua eficiência, dificultando sua avaliação.

Com isso em mente, e levando-se em conta a grande difusão de câmeras comuns de tráfego e o preço razoável com que se obtém um alto poder computacional atualmente, esse trabalho teve como motivação a criação de uma maneira rápida,

eficiente e de baixo custo de se determinar a situação em uma via de tráfego. Tal solução permitiria a tomada de medidas em tempo real por um sistema inteligente de controle de tráfego (Sistema Avançado de Gestão de Transporte), ou até mesmo a simples atualização automática da situação das vias em um *website*, por exemplo, disponibilizando tais informações para motoristas (Sistema Avançado de Informação ao Viajante).

## 1.2 Trabalhos correlatos

Diversas soluções para o sensoriamento de vias encontram-se disponíveis no mercado. A SWARCO, empresa que adquiriu a Weiss Electronics, pioneira no desenvolvimento de sensores magnéticos, disponibiliza, além dessa tecnologia, soluções baseadas em detecção por vídeo, por radar e por sensores infravermelhos, todas com *hardware* proprietário [5]. Um dos seus dispositivos mais completos, o *TDC4* [6], garante classificação de veículos, velocidades individuais de cada um, além de contagem e outras propriedades do tráfego, utilizando quatro tecnologias de detecção diferentes. A medição de velocidade é feita com radar pela análise da frequência das microondas refletidas (utilizando o efeito Doppler), um sensor ultrassônico determina o perfil de altura do veículo, e as zonas de infravermelho passivo determinam a posição de cada um, ao mesmo tempo em que o detector por vídeo realiza seu processamento de forma independente.

A AutoScope [7], empresa pioneira nos sistemas de vídeo, de acordo com [3], também possui uma gama de soluções, todas contando com *hardware* próprio, como câmeras, bastidores e placas. Entre seus produtos se destaca o *Autoscope Duo* [8], um dispositivo com câmera que mescla a detecção por vídeo com a precisão de um sensor por radar, garantindo uma performance mais confiável e maior acurácia devido à redundância obtida ao se utilizar dois sistemas independentes. Como vantagem da utilização do radar encontra-se sua imunidade a artefatos de iluminação e condições meteorológicas (como neblina), que costumam afetar sistemas de vídeo.

Soluções com câmeras infravermelhas também são encontradas no mercado, como as câmeras da FLIR [9], que garantem uma visão clara da via ininterruptamente. Obstáculos para sistemas de imagens, como reflexos solares, sombras, asfalto molhado, neve e neblina são ignorados pelo sistema, que mescla visão térmica com de luz visível. Detecção de incidentes, monitoramento de fluxo de tráfego e contagem de carros são divulgados como suas funcionalidades.

No campo acadêmico, ATKOCIUNAS *et al.* [10] desenvolveram um sofisticado sistema de rastreamento de veículos em que foram utilizados, a título de comparação, dois métodos para definição de objetos: extração de contornos de objetos e detecção de movimento. Utilizando um arsenal de métodos e técnicas de processamento de

imagens, a solução desenvolvida consegue, em ambos os casos, realizar um rastreamento fiel nos casos apresentados, necessitando apenas de um conjunto de imagens da pista vazia inicialmente.

A medição de velocidade também é apresentada como funcionalidade do sistema. Porém, embora precisa, é estimada a partir da comparação do rastreamento realizado com sucessivos vídeos de viagens controladas com velocidade fixa e conhecida. Essa medição, portanto, requisita um dispendioso processo de calibragem para cada via analisada, o que não é muito desejável.

### 1.3 Objetivo do trabalho

Em virtude do discutido anteriormente, o objetivo desse trabalho consiste, portanto, em se criar uma solução automática para determinação dos parâmetros densidade e velocidade em uma via a partir de imagens, de forma que requisite o mínimo de interação humana (o que inclui a calibragem inicial) e gere respostas que possam ser utilizadas em diferentes aplicações, seja em coleta de dados para planejamento, seja para utilização em tempo real. Os parâmetros densidade e velocidade serão explicados em detalhes no capítulo 2.

### 1.4 Estrutura do texto

Este trabalho encontra-se organizado da seguinte forma: no capítulo 2, são apresentados os conceitos fundamentais de Engenharia de Tráfego que servem de base para a solução desenvolvida, como taxa de fluxo, densidade e velocidade; no capítulo 3, são apresentadas as técnicas de processamento de imagens utilizadas neste trabalho, como filtros morfológicos e segmentação de imagens; no capítulo 4, são apresentados conceitos de Agrupamento (*clustering*), assim como o algoritmo *K-means*; no capítulo 5, são apresentados os conceitos fundamentais de Redes Neurais Sem Peso, também essenciais para este trabalho; no capítulo 6, a solução desenvolvida é apresentada e explicada em detalhes; na seção 7, são discutidos os resultados experimentais obtidos; e, por fim, no capítulo 8, são expostas as conclusões do trabalho, bem como os possíveis trabalhos futuros baseados na solução.

# Capítulo 2

## Fundamentos de Engenharia de Tráfego

Neste capítulo, serão apresentados os conceitos de Engenharia de Tráfego relacionados à análise das características de fluxos de tráfego, como vistos em [3]. Na seção 2.1, será feita uma breve introdução sobre o tópico; na seção 2.2, serão descritos os tipos de vias de tráfego existentes; na seção 2.3, serão apresentados os parâmetros de fluxo de tráfego mais utilizados pelos engenheiros de tráfego; e na seção 2.4, serão mostradas as relações entre esses parâmetros e as características observáveis a partir delas.

### 2.1 Introdução

Ao contrário da mecânica de fluidos, em que o fluxo de um líquido (ou gás) por tubos de determinadas características se comporta de forma completamente previsível, de acordo com leis estabelecidas, o comportamento do tráfego em uma via envolve muitos elementos variáveis, tornando difícil sua previsão em termos exatos. Tanto os motoristas como os veículos que formam uma corrente de tráfego possuem respectivamente comportamentos e características variáveis, o que resulta em comportamentos diferentes por parte de cada veículo na via. Além disso, duas correntes de tráfego, mesmo em situações semelhantes, jamais se comportarão da mesma forma porque o comportamento do motorista varia de acordo com as características locais da via e os hábitos regionais de direção. O desafio de um engenheiro de tráfego torna-se, portanto, planejar e projetar para um meio que envolve tanto restrições físicas como as complexas características comportamentais humanas.

Felizmente, embora as características exatas variem, tendências consistentes do comportamento dos motoristas, e por consequência, do tráfego, podem ser observadas. Por exemplo, motoristas em uma via projetada para determinada velocidade

máxima viajarão dentro de um intervalo que varia de um pouco abaixo para um pouco acima desse valor, ao passo que uma minoria viajará a uma velocidade muito abaixo ou muito acima da estabelecida.

De forma a entender a variabilidade inerente nas características das correntes de tráfego e definir os intervalos comuns de comportamento, parâmetros chave são definidos e medidos, descrevendo assim em termos quantitativos as correntes de tráfego. Os engenheiros de tráfego então analisam, avaliam e eventualmente planejam melhorias em vias de tráfego baseados nesses parâmetros e em seus conhecimentos dos intervalos comuns de comportamento. Esses parâmetros, portanto, constituem a medida de realidade do engenheiro de tráfego, uma forma de descrever e entender as correntes de tráfego.

## 2.2 Tipos de vias de tráfego

Vias de tráfego são amplamente separadas em duas categorias principais: fluxo ininterrupto e fluxo interrompido.

### 2.2.1 Fluxo ininterrupto

Presentes principalmente em autoestradas, onde não há cruzamentos, semáforos, placas de sinalização de “pare” ou “dê a preferência” ou outros fatores interruptivos externos, as vias de fluxo ininterrupto são aquelas em que não há interrupções externas ao fluxo de tráfego. O acesso nessas vias é muito limitado, não havendo cruzamentos, calçadas ou quaisquer outras formas de acesso direto aos terrenos adjacentes. Dessa forma, o acesso completo da via garantido aos motoristas permite que as características do fluxo de tráfego sejam influenciadas apenas pelas interações entre os veículos, a pista e o ambiente ao redor.

Embora um fluxo puramente ininterrupto exista apenas em autoestradas, ele também pode existir em trechos de rodovias onde existam longas distâncias entre interrupções fixas. Como regra geral, fluxos ininterruptos podem existir em situações em que a distância entre sinais de trânsito e/ou outras interrupções fixas significativas seja maior que 3,2 km.

É importante notar que o termo *fluxo ininterrupto* refere-se ao tipo de via, e não à qualidade das operações na via. Dessa forma, mesmo uma autoestrada com grande retenção em horário de pico ainda assim estará operando em um fluxo ininterrupto.

### 2.2.2 Fluxo interrompido

Vias de fluxo interrompido são aquelas que incorporam em seu projeto e operação interrupções externas fixas. Praticamente todas as ruas e rodovias são vias de fluxo

interrompido, com seus semáforos, cruzamentos, placas, calçadas e acessos impactando o tempo de viagem.

O grande diferencial entre vias de fluxo ininterrupto e interrompido é o impacto no tempo. Enquanto em uma via ininterrupta a via esteja a todo o momento disponível aos veículos, em uma dada via interrompida o movimento sofre constante paralização, resultado do limite de acesso aos trechos da via imposto pelas interrupções. Ao invés de um fluxo de tráfego em movimento contínuo, há um fluxo que é periodicamente parado e reiniciado.

O fluxo interrompido é, portanto, mais complexo que o ininterrupto. Embora muitos dos parâmetros de fluxo de tráfego apresentados na seção 2.3 se apliquem a ambos os tipos de vias, eles serão elucidados de forma simplificada com foco no fluxo ininterrupto, como visto na literatura [3].

## 2.3 Parâmetros de fluxo de tráfego

Parâmetros de fluxo de tráfego são divididos em duas categorias. Ao passo que os parâmetros macroscópicos descrevem o fluxo de tráfego como um todo, os parâmetros microscópicos descrevem o comportamento de veículos em um fluxo de tráfego individualmente ou em pares.

Os três principais parâmetros macroscópicos são volume e taxa de fluxo (2.3.1), velocidade (2.3.2) e densidade (2.3.3). Os principais parâmetros microscópicos são a velocidade individual dos veículos, o *gap* (distância entre dois veículos sucessivos em uma faixa de tráfego, medida a partir de um ponto de referência comum, como o para-choque frontal ou as rodas dianteiras) e o *headway* (intervalo de tempo entre os instantes em que veículos sucessivos passam por um ponto sobre a faixa, também medido a partir de pontos de referência comuns aos veículos).

Esta seção se limita a apresentar os parâmetros macroscópicos, que constituem o foco do trabalho.

### 2.3.1 Volume e Taxa de Fluxo

Volume de tráfego é definido como o número de veículos que passam por um ponto em uma rodovia, ou em uma faixa ou direção de uma rodovia, durante um intervalo de tempo especificado. A unidade de medida é simplesmente “veículos”, embora seja normalmente expressa como “veículos por unidade de tempo”, sendo as unidades de tempo mais comuns a “por dia” e “por hora”.

A taxa de fluxo, comumente definida em unidades de “veículos por hora”, representa fluxos que existem por períodos de tempo menores que uma hora. O cálculo é simples: um volume de 200 veículos observados por um período de 15 minutos (1/4



de uma hora) é expresso como uma taxa de  $200/0,25 = 800$  veículos/hora, mesmo que 800 veículos não fossem observados caso a hora completa fosse analisada. O valor 800 veículos/hora então representa a taxa de fluxo que existe para um intervalo de 15 minutos.

### **Volumes diários**

A medição de volume diário é importante para se estabelecer tendências ao longo do tempo e para propósitos de planejamento em geral. Ao se documentar tendências anuais na utilização de uma rodovia, tem-se como planejar melhorias ou até novas vias para acomodar o aumento de demanda em épocas específicas. A análise desses dados pode sugerir, por exemplo, que determinada via serve principalmente para propósitos recreacionais, por ser mais utilizada nos finais de semana de verão, e que, portanto, durante a elaboração de melhorias, esse deve ser o período focado.

### **Volumes horários**

Como o volume em uma via varia consideravelmente ao longo das 24 horas do dia, a medição do volume diário não é de muita ajuda para os engenheiros de tráfego no que diz respeito aos períodos do dia em que há ocorrência de fluxo máximo. O termo *hora de pico* refere-se à hora do dia com o maior volume horário, e a medição de volume nessa hora provê o conhecimento necessário às decisões detalhadas de controle e projeto.

Ao contrário do volume diário, que geralmente não é diferenciado por direção ou faixa, representando assim valores totais para uma via na localização designada, o volume na hora de pico é normalmente determinado como um volume direcional (ou seja, cada direção é contada separadamente). Rodovias e controles devem ser projetados para servir adequadamente ao volume de tráfego da hora de pico na direção do pico de fluxo, e, como o tráfego viajando em uma direção na hora de pico matinal vai trafegar pela direção oposta durante o pico noturno, ambos os lados da via devem ser projetadas de forma a acomodar o pico de fluxo direcional.

### **Volumes por fração horária**

Embora volumes de tráfego horário sejam importantes para muitas formas de projeto e análise de tráfego, a variação do tráfego dentro do período de uma hora também é de suma importância. A natureza do fluxo de tráfego é frequentemente ligada a oscilações de curto prazo na demanda de tráfego. Uma via pode ter capacidade suficiente para servir a demanda da hora de pico, mas picos de fluxo de curta duração podem exceder a capacidade e criar congestionamento na via.

Por exemplo, uma via com capacidade para comportar 2.000 veículos em uma hora comporta apenas 500 veículos a cada 15 minutos (pois a capacidade é constante). Fracionando a hora de pico em intervalos de 15 minutos e analisando a taxa de fluxo em cada um deles, pode-se notar que seus valores variam e, mesmo que no total de veículos o volume observado esteja dentro do comportado, caso em determinada fração dessa hora haja uma demanda acentuada, cuja taxa de fluxo exceda 2.000 veículos/hora (digamos, 550 veículos na segunda fração da hora, o que equivale a uma taxa de 2.200 veículos/hora), haverá retenção no tráfego, o que afetará o tráfego dos intervalos seguintes.

Devido a esse tipo de impacto, torna-se necessário projetar vias e analisar condições de tráfego para o período de taxa máxima de fluxo na hora de pico. Para propósitos práticos, o intervalo de 15 minutos é considerado o período de tempo mínimo em que condições de tráfego são estatisticamente estáveis.

### 2.3.2 Velocidade e Tempo de Viagem

Velocidade é definida como a distância percorrida por unidade de tempo, enquanto o tempo de viagem é o tempo que se leva para percorrer um trecho definido da rodovia. Ambos os parâmetros são, portanto, inversamente proporcionais.

Em um típico fluxo de tráfego, cada veículo viaja com velocidades diferentes. Visto que a análise está sendo feita a nível macroscópico, torna-se necessário o cálculo de uma velocidade média, a fim de caracterizar o fluxo de tráfego como um todo.

Uma das maneiras de determinar a velocidade média para um fluxo de tráfego é calcular a velocidade média de todos os veículos que passam por um ponto na rodovia (ou faixa) por determinado período de tempo. Ou seja, encontra-se a velocidade individual de cada veículo e calcula-se a média dos resultados. Esse cálculo é denominado *time mean speed*.

A segunda maneira, denominada *space mean speed*, consiste em calcular a velocidade média de todos os veículos ocupando determinado trecho da rodovia (ou faixa) por certo período de tempo. Isso significa que, após descobrir o tempo de viagem médio para um veículo percorrer o trecho em questão, computa-se a velocidade média.

A *time mean speed* equivale à *space mean speed* acrescida da variância da *space mean speed* dividida pela *space mean speed* [11], como visto na equação 2.1. Se todos os veículos possuírem a mesma velocidade, então a *time mean speed* e a *space mean speed* serão iguais.

$$v_t = v_s + S^2/v_s \quad (2.1)$$

### 2.3.3 Densidade e Ocupância

Densidade representa o número de veículos que ocupam determinada extensão da rodovia ou faixa, frequentemente expressa como “veículos por quilômetro” ou “veículos por quilômetro por faixa”. Como se torna necessário um ponto de observação elevado sobre o trecho a ser estudado, o que dificulta sua obtenção, a densidade é frequentemente calculada utilizando-se os valores medidos de velocidade e taxa de fluxo.

A densidade, porém, é talvez o mais importante dos três parâmetros de fluxo de tráfego, pois é a medida mais diretamente relacionada com a demanda de tráfego. Demanda não ocorre como uma taxa de fluxo, mesmo que engenheiros de tráfego utilizem esse parâmetro como a medida principal de demanda. Tráfego é gerado a partir de vários acessos, que injetam certo número de veículos em um espaço confinado de rodovia, gerando assim uma densidade de veículos. Motoristas se utilizam de velocidades consistentes com a distância a que estão dos outros veículos. A velocidade e a densidade combinadas, portanto, indicam a taxa de fluxo observada.

Densidade também é uma importante medida de qualidade do fluxo de tráfego. Por se tratar da medida de proximidade dos veículos, um fator que influencia a liberdade de movimentos, a densidade tem relação com o conforto psicológico dos motoristas.

Embora a densidade seja difícil de medir diretamente, existem alternativas mais diretas que a utilização dos valores de velocidade e taxa de fluxo. Detectores modernos podem medir ocupância, que é um parâmetro relacionado, definido como a proporção do tempo em que um detector é “ocupado”, ou coberto, por um veículo em determinado período de tempo.

## 2.4 Relações entre Taxa de Fluxo, Velocidade e Densidade

As três medidas macroscópicas, que indicam o estado de determinado fluxo de tráfego, se relacionam da seguinte forma:

$$v = S * D$$

em que  $v$  é a taxa de fluxo (em “veículos por hora” ou “veículos por hora por faixa”),  $S$  é a *space mean speed* (em “quilômetros por hora”) e  $D$  é a densidade (em “veículos por quilômetro” ou “veículos por quilômetro por faixa”). *Space mean speed* e densidade são medidas que fazem referência a um trecho em comum da faixa ou rodovia, enquanto a taxa de fluxo é uma medida pontual.

A relação entre as medidas sugere que uma dada taxa de fluxo ( $v$ ) possa ser obtida por infinitas combinações de pares de velocidade ( $S$ ) e densidade ( $D$ ) que tenham o mesmo produto. Porém, existem relações adicionais entre os pares dessas variáveis que restringem o número de combinações que podem e que de fato ocorrem em situações reais. A figura 2.1 ilustra a forma geral dessas relações. É importante notar que trata-se de uma representação genérica; a forma exata das relações varia de local para local e até com o passar do tempo em um mesmo local.

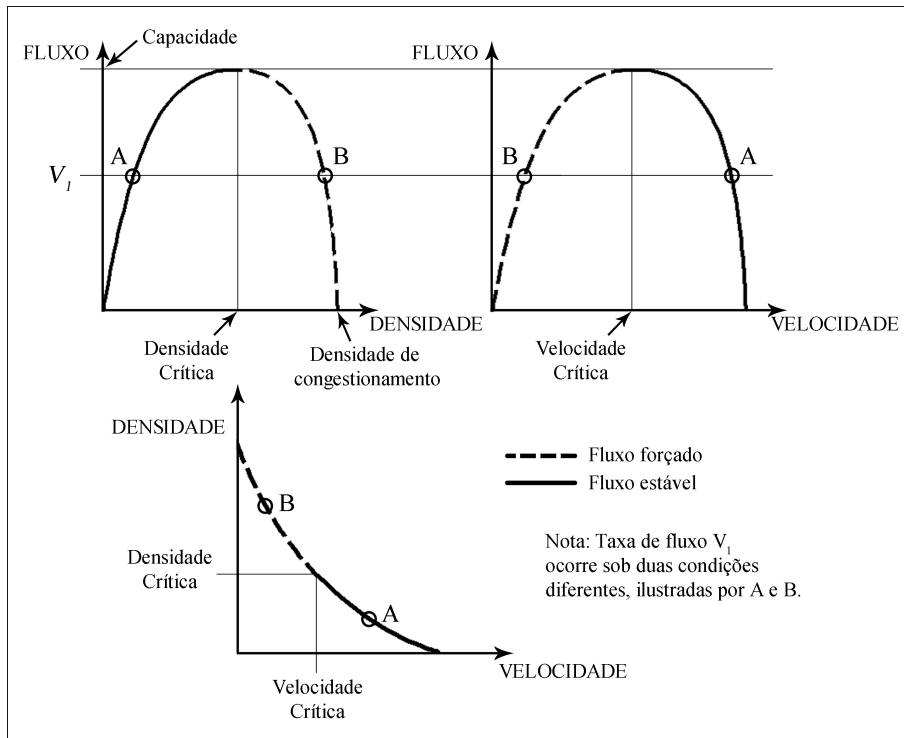


Figura 2.1: Relações entre Fluxo, Velocidade e Densidade

Como visto na figura, uma taxa de fluxo de zero veículo por hora ocorre sob duas condições diferentes. Quando não há veículos na rodovia, a densidade é zero veículo por quilômetro e, portanto, nenhum veículo pode ser observado passando por um ponto. Nessa condição, a velocidade é imensurável e referida como *velocidade de fluxo livre*, um valor teórico que existe como uma extensão matemática da relação entre velocidade e fluxo (ou velocidade e densidade). Em termos práticos, velocidade de fluxo livre pode ser pensada como a velocidade que um único veículo pode atingir quando não há outros veículos na pista e o motorista está dirigindo tão rápido quanto possível, levando em conta a geometria da rodovia e o ambiente ao redor.

Um fluxo de zero veículo por hora também ocorre quando há tantos veículos na pista que todo o movimento cessa. Nessa densidade altíssima, chamada de *densidade de congestionamento*, nenhum fluxo é observado, visto que nenhum veículo é capaz de passar no ponto observado.

Entre ambos os extremos encontra-se o ponto de pico que, tanto na curva de

fluxo e velocidade como na de fluxo e densidade, representa a taxa máxima de fluxo, ou seja, a capacidade da rodovia em questão. Embora pareça a situação ideal, operar no limite da capacidade da via é muito instável, visto que, em um fluxo sem lacunas utilizáveis, a mínima perturbação causada por um veículo entrando na via ou mudando de faixa, ou a simples freada de um motorista, pode causar uma reação em cadeia que não pode ser amortecida. Essa perturbação se propaga adiante até que lacunas suficientes no fluxo de tráfego permitam que ela seja efetivamente dissipada.

A porção tracejada das curvas representa fluxo instável ou forçado, que se traduz no fluxo lento, de fila, que se forma atrás de um local de colapso do trânsito. Um colapso ocorre em qualquer ponto onde a taxa de fluxo de chegada excede a capacidade da via. Tem-se como exemplo rampas de acesso em autoestradas. Eventos menos previsíveis, como acidentes, também são responsáveis por colapsos. O trecho sólido das curvas representa fluxo estável, que pode ser mantido durante determinado período de tempo.

Com exceção da taxa de fluxo operando na capacidade da via, qualquer outra taxa de fluxo existirá obrigatoriamente sob uma das seguintes condições: alta velocidade e baixa densidade relativas (na porção estável das relações de fluxo); ou baixa velocidade e alta densidade relativas (na porção instável das relações de fluxo). Devido a essas duas condições possíveis para um mesmo valor, volume e taxa de fluxo não podem descrever completamente as condições de fluxo, ou serem usados como medidas de qualidade do fluxo de tráfego. Valores de velocidade e/ou densidade, no entanto, poderiam definir pontos únicos em qualquer uma das relações da figura.

Um ponto importante a ser destacado é que a figura 2.1 retrata as relações entre os parâmetros quando observados em uma via de fluxo ininterrupto. Como já mencionado, os parâmetros de fluxo de tráfego apresentados são focados nos fluxos sem interrupção externa, embora também possam ser aplicados em vias de fluxos interrompidos. Com certas adaptações, as observações acerca da figura 2.1 podem ser expandidas para abranger ambos os tipos de vias.

Em um fluxo interrompido, por exemplo, além das duas condições citadas para as taxas de fluxo não equivalentes à capacidade da via, é possível acrescentar uma terceira e uma quarta condições: a de que ambas as densidade e velocidade são baixas ou de que a densidade é baixa e a velocidade é zero. Tais situações podem ocorrer, respectivamente, quando os poucos carros na via reduzem a velocidade para dar prioridade a uma via adjacente ou os poucos carros na via estão parados em um sinal vermelho.

Embora este trabalho esteja focado em uma solução para ambos os tipos de fluxo, os conceitos aplicados às vias de fluxo ininterrupto fornecem uma boa base de entendimento sobre o assunto.

# Capítulo 3

## Fundamentos de Processamento de Imagens

Neste capítulo, serão apresentados os conceitos de Processamento de Imagens utilizados neste trabalho. Na seção 3.1 será feita uma breve introdução sobre o tópico; na seção 3.2, será definido o modelo de cores RGB; na seção 3.3 serão explicados os filtros morfológicos; e na seção 3.4, será explicado o processo de segmentação de imagens.

### 3.1 Introdução

A observação visual presta um papel importantíssimo na ciência desde o seu surgimento, quando a única maneira de se documentar os resultados de um experimento era através de descrição verbal e desenhos manuais. Com a invenção da fotografia, os resultados puderam ser documentados de maneira objetiva: as imagens podiam ser usadas para documentação, descrição qualitativa e ilustração do fenômeno observado.

Porém, as imagens de um experimento também eram capazes de revelar novas descobertas através de sua análise. Uma extensa análise por inúmeras imagens de câmaras de bolhas de hidrogênio levaram ao descobrimento de diversas partículas elementares da física, por exemplo. Esses procedimentos de avaliação manual, contudo, consumiam muito tempo e, mesmo com a criação de dispositivos mecânicos semi ou totalmente automáticos, o alto grau de especialização desses dispositivos contribuiu para que a avaliação quantitativa das imagens não se popularizasse [12]. O mesmo ocorreu, de certa forma, com a computação em seus estágios iniciais. O custo computacional era alto e as aplicações, além de especializadas, não disponibilizavam da velocidade de processamento desejada.

Com o acelerado progresso das tecnologias de vídeo e computação, em pouco

tempo os computadores pessoais passaram a ter capacidade de processar dados de imagens de forma rápida e barata. Como resultado, tanto *software* como *hardware* se tornaram a solução padrão para processamento de imagens, sequências de imagens e visualizações em três dimensões, e essa tecnologia se tornou disponível para qualquer cientista ou engenheiro. Por conta disso, o processamento de imagens expandiu de poucas aplicações especializadas para uma ferramenta científica padrão. Atualmente, técnicas de processamento de imagens são utilizadas em praticamente todas as ciências naturais e disciplinas técnicas.

## 3.2 Modelo de cores RGB

Modelos de cores são modelos matemáticos que descrevem como as cores podem ser representadas numericamente. Seu propósito é facilitar a especificação de cores em uma forma padrão amplamente aceita. Um modelo consiste, em essência, em uma especificação de um sistema de coordenadas em três dimensões e um subespaço dentro desse sistema em que cada cor é representada por um único ponto [13].

Utilizando o princípio do Método de Maxwell, de reprodução de todas as cores do espectro visível pela mistura, em diferentes intensidades, de feixes de luz das cores primárias vermelha, verde e azul [14], o modelo de cores RGB figura como o modelo mais adequado para gráficos computacionais. Isso se dá em virtude da natureza das telas, tanto de computadores como de televisores, que utilizam o vermelho, o verde e o azul para criar a cor desejada. Dessa forma, a escolha do modelo de cor RGB simplifica a arquitetura e o projeto do sistema [13].

Nas imagens RGB, as intensidades de cada cor primária em cada ponto da imagem (chamado de *pixel*) são armazenadas em *canais*, como ilustrado na figura 3.1. Em imagens de 8 *bits* por canal (o padrão atualmente), os três canais se traduzem em 24 *bits* (8 *bits* x 3 canais) de informação de cor por *pixel*. Dessa forma, a imagem pode reproduzir 16,7 milhões de cores por *pixel* [15]. Como 8 *bits* representam 256 valores diferentes, a intensidade de cada cor pode variar entre 0 e 255. O branco puro, então, é obtido quando os três componentes da imagem possuem o valor de 255 (máxima intensidade) e o preto puro, quando os três são 0. Todas as outras cores são obtidas variando-se esses valores. Um vermelho vibrante, por exemplo, possui os valores 246 para o vermelho, 20 para o verde e 50 para o azul.



Figura 3.1: Canais de cores do RGB

### 3.3 Filtros morfológicos

A filtragem morfológica consiste em uma teoria desenvolvida na década de 60 para a análise e processamento de imagens discretas. Mais comumente utilizada em imagens binárias, seu método se resume à transformação de uma imagem baseada na relação de cada *pixel* com sua vizinhança, sendo a natureza dessa transformação definida por um operador morfológico [16]. Entre os propósitos da filtragem morfológica incluem-se a remoção de ruídos e correção na forma dos objetos.

A convenção adotada para morfologia dita que, em uma imagem binária (ou seja, uma imagem em que todo *pixel* só pode assumir os valores preto ou branco), os *pixels* brancos representam os objetos em primeiro plano e os *pixels* pretos representam o fundo da imagem[16]. Os operadores morfológicos são então utilizados para realçar a diferença entre ambos, indicando a maneira como a imagem será transformada, sendo os dois mais importantes a *erosão* e a *dilatação*.

Na erosão, todo *pixel* de objeto considerado como de fronteira é transformado em um *pixel* de fundo ao passo que, na dilatação, todo *pixel* de fundo considerado como de fronteira é transformado em um *pixel* de objeto, como ilustrado de forma ampliada na figura 3.2. A erosão encolhe os objetos, podendo quebrar um objeto em muitos outros, enquanto a dilatação os expande, podendo unir muitos objetos em um só [17].

Embora as transformações na figura 3.2 tenham sido realizadas nos *pixels* mais externos (ou seja, nos *pixels* de objeto que fazem fronteira imediata com *pixels* de fundo ou vice-versa), os *pixels* que de fato são transformados pela filtragem são definidos pelo *elemento de estruturação*, instrumento fundamental da filtragem





Figura 3.2: Erosão e dilatação em uma imagem

morfológica. Um elemento de estruturação consiste em uma configuração de *pixels* (ou forma) em que uma origem é definida, como ilustrado na figura 3.3, sendo as formas mais utilizadas um quadrado, um círculo ou um losango, todos com a origem localizada ao centro.

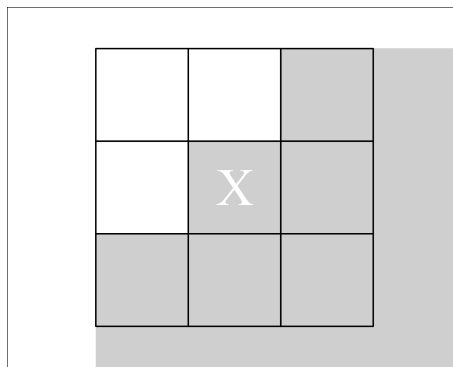


Figura 3.3: Um elemento de estruturação em forma de um quadrado 3x3

Aplicar o filtro morfológico consiste então em varrer toda a imagem, sondando cada *pixel* com o elemento de estruturação e determinando se ele deve ou não ser alterado. Quando a origem do elemento de estruturação está alinhada com um *pixel*, o conjunto de *pixels* englobados pelo elemento é analisado. Na erosão, caso o *pixel* na origem seja de objeto e haja pelo menos um *pixel* de fundo no conjunto englobado, esse *pixel* é considerado como de fronteira e, portanto, é transformado em um pixel de fundo. De forma análoga, um *pixel* de fundo é considerado de fronteira quando um de objeto é detectado na região englobada pelo elemento de estruturação. Em uma erosão, o *pixel* sendo analisado na figura 3.3 seria transformado em um *pixel* branco. A figura 3.2, por sua vez, ilustra uma erosão e uma dilatação utilizando-se um quadrado 3x3 como elemento de estruturação (o que equivale dizer que os *pixels* de fronteira são apenas aqueles que “tocam” os *pixels* de cor diferente).

Por fim, outros dois operadores importantes, derivados da erosão e dilatação, são a *abertura* e o *fechamento*. A abertura é definida como uma erosão seguida de uma

dilatação, sendo o oposto de um fechamento, que consiste em uma dilatação seguida de uma erosão, como ilustrado na figura 3.4. A abertura remove pequenas ilhas e filamentos dos *pixels* de objeto, enquanto o fechamento remove ilhas e pequenos filamentos dos *pixels* de fundo. Essas técnicas se mostram muito úteis no tratamento de imagens com ruídos, em que alguns *pixels* contém o valor binário incorreto [17].

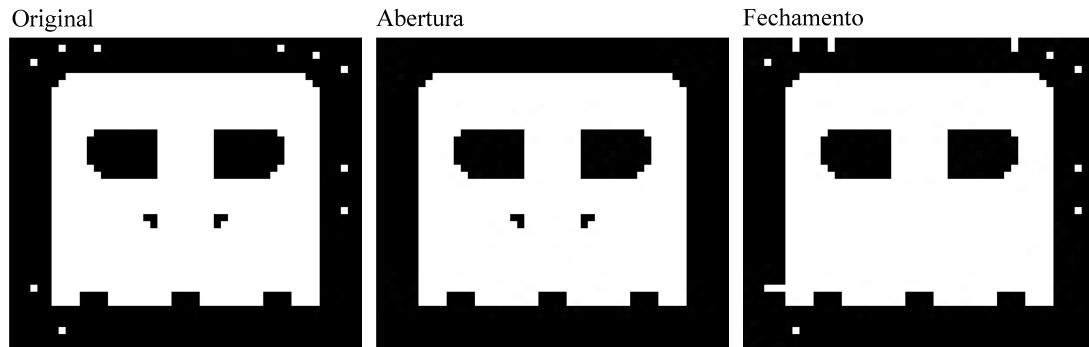


Figura 3.4: Abertura e fechamento em uma imagem

### 3.4 Segmentação

De forma geral, é conhecido como *Segmentação* o processo de particionamento de uma imagem digital em muitos segmentos, ou conjuntos de *pixels*, de forma a simplificar a representação da imagem, tornando-a mais significativa e fácil de analisar [19]. Tal resultado é obtido determinando um rótulo para cada *pixel* em uma imagem, de forma que *pixels* com mesmo rótulo compartilhem de determinadas características visuais.

Em muitas aplicações envolvendo análise visual, uma importante etapa de pré-processamento da imagem consiste na delimitação dos objetos em movimento no primeiro plano em contraste com o plano de fundo predominantemente estático. Conhecida como *Segmentação de Fundo* e utilizada em aplicações como vigilância visual e jogos virtuais, sua ideia geral consiste em gerar automaticamente uma máscara binária que divida o conjunto de *pixels* que pertencem ao primeiro plano daqueles que pertencem ao plano de fundo.

No caso mais simples, uma imagem contendo apenas o fundo estático pode ser comparada com a imagem atual sendo analisada e, a partir dessa comparação, os *pixels* com maior variação podem ser determinados como de primeiro plano. Porém, embora funcione bem em casos específicos, diversas vezes a imagem contendo apenas o fundo é impossível de ser obtida, o plano de fundo apresenta movimentação sutil, ou até mesmo a iluminação sofre alteração gradual. Por conta desses fatores, diversos métodos sofisticados vêm sendo desenvolvidos para modelar o plano de fundo, sendo

o *Segmentador Adaptativo Pixel a Pixel* (*Pixel-Based Adaptive Segmenter*, ou PBAS) [20] um dos mais bem conceituados atualmente.

### 3.4.1 Segmentador Adaptativo Pixel a Pixel

Baseado em diversos métodos existentes em conjunto com ideias de teoria de sistemas de controle, o PBAS apresenta um ganho de performance significativo em relação à maioria das abordagens que atualmente representam o estado da arte, de acordo com as avaliações realizadas com a base de dados *Change Detection Challenge* [21]. Seu diferencial em relação aos outros métodos reside na existência de parâmetros ajustados adaptativamente em tempo de execução para cada *pixel* em separado, em contraste com os parâmetros estáticos comumente utilizados.

No PBAS, cada *pixel*  $x_i$  é modelado como um conjunto de valores de fundo recentemente observados. Em suma, a decisão se um *pixel* pertence ou não ao primeiro plano é feita analisando-se a imagem atual e o modelo de fundo  $B(x_i)$  e levando-se em conta o limiar  $R(x_i)$ , um parâmetro com valor específico para cada *pixel*. Além disso, o modelo de fundo é atualizado com o passar do tempo (permitindo o acompanhamento de mudanças no plano de fundo) e essa atualização é feita de acordo com o parâmetro de aprendizado  $T(x_i)$ , cujo valor também é específico para cada *pixel*. A alteração dinâmica realizada em tempo de execução nos valores de ambos os limiares, que leva em consideração a situação presente no plano de fundo, consiste na essência do PBAS.

A seguir serão descritos o processo de decisão de segmentação e a atualização do modelo de fundo, seguidos do detalhamento do método de atualização de ambos os limiares,  $R(x_i)$  e  $T(x_i)$ .

#### Decisão de segmentação

O modelo de fundo  $B(x_i)$  é definido como um conjunto de  $N$  valores de pixels recentemente observados:

$$B(x_i) = \{B_1(x_i), \dots, B_k(x_i), \dots, B_N(x_i)\}$$

Um pixel  $x_i$  é definido como pertencendo ao plano de fundo caso seu valor  $I(x_i)$  esteja mais próximo que determinado limiar de decisão  $R(x_i)$  a pelo menos  $\#_{min}$  dos  $N$  valores de fundo. Ou seja, em uma máscara de segmentação de primeiro plano, em que  $F = 1$  significa primeiro plano, essa máscara pode ser calculada como:

$$F(x_i) = \begin{cases} 1 & \text{se } \#\{dist(I(x_i), B_k(x_i)) < R(x_i)\} < \#_{min} \\ 0 & \text{caso contrário} \end{cases}$$

A decisão, portanto, envolve dois parâmetros: o limiar de distância  $R(x_i)$ , que é definido separadamente para cada *pixel* e alterado dinamicamente, e o número mínimo  $\#_{min}$ , que é um parâmetro global fixo.

Como a imagem analisada é uma imagem com três canais de cores, cada canal é tratado independentemente e todos os algoritmos são, portanto, executados em paralelo para cada um deles. A segmentação final  $F(x_i)$  resulta de uma operação OU das três segmentações  $F^R(x_i)$ ,  $F^G(x_i)$  e  $F^B(x_i)$ , ou seja, basta um deles ser classificado como de primeiro plano para o *pixel* ser classificado como tal.

### Atualização do Modelo de Fundo

A atualização do modelo de fundo  $B$ , como já mencionado, permite o acompanhamento de mudanças no fundo, como alteração na iluminação, aparição de sombras e objetos de fundo em movimento (como árvores). Como regiões de primeiro plano não podem ser usadas para atualização, o modelo de fundo é atualizado apenas para os *pixels* que atualmente se encontram no plano de fundo (ou seja,  $F(x_i) = 0$ ). Atualizar significa que, para determinado índice  $k \in 1 \dots N$  (escolhido aleatoriamente), o valor de modelo de fundo correspondente,  $B_k(x_i)$ , é substituído pelo valor atual do *pixel*,  $I(x_i)$ . Isso permite que o valor atual do *pixel* seja “aprendido” pelo modelo de fundo. Essa atualização, porém, é realizada apenas com a probabilidade  $p = 1/T(x_i)$ ; caso contrário, nenhuma atualização é realizada. Dessa forma, o parâmetro  $T(x_i)$  define a taxa de atualização. Quanto maior  $T(x_i)$ , menor a probabilidade de um *pixel* ser atualizado.

Também é atualizado (com probabilidade  $p = 1/T(x_i)$ ) um *pixel* de vizinhança  $y_i \in N(x_i)$  escolhido aleatoriamente. Assim, o modelo de fundo  $B_k(y_i)$  nesse *pixel* de vizinhança é substituído pelo valor atual desse *pixel*,  $V(y_i)$ .

Em geral, um *pixel*  $x_i$  é atualizado apenas quando classificado como de fundo. Porém, um *pixel* de vizinhança  $y_i$ , que pode ser de primeiro plano, também pode ser atualizado. Isso significa que alguns pixels de primeiro plano que estejam nas fronteiras podem ser gradualmente inseridos no modelo de fundo. Com esse método, todo objeto de primeiro plano eventualmente é “engolido” após determinado tempo, dependendo do parâmetro de atualização  $T(x_i)$ . A vantagem dessa propriedade é que, com o tempo, objetos de primeiro plano erroneamente classificados são corretamente inseridos no plano de fundo.

### Atualização do Limiar de Decisão

Em uma sequência de vídeo, é comum haver regiões de fundo com muito movimento (como a superfície de corpos de água ou árvores ao vento), assim como regiões com pouca ou nenhuma alteração (como construções). Idealmente, para áreas dinâmicas,

o limiar  $R(x_i)$  deve ser incrementado para não incluir objetos de fundo no primeiro plano. Para regiões estáticas,  $R(x_i)$  deve ser baixo, de forma que pequenos desvios sejam considerados no primeiro plano. Portanto, o limiar  $R(x_i)$  precisa se adaptar automaticamente para as circunstâncias que se apresentam, e, para tal, as dinâmicas do fundo precisam ser medidas.

Além de se armazenar um conjunto dos valores de *pixel* observados recentemente (modelo de fundo  $B(x_i)$ ), o mesmo é feito para um conjunto  $D(x_i) = \{D_1(x_i), \dots, D_N(x_i)\}$  de distâncias mínimas de decisão. Sempre que uma atualização de  $B_k(x_i)$  é realizada, a distância mínima observada,  $d_{min}(x_i) = \min_k \text{dist}(I(x_i), B_k(x_i))$  é inserida nesse conjunto:  $D_k(x_i) \leftarrow d_{min}(x_i)$ . Assim, um histórico de distâncias mínimas de decisão é criado. A média desses valores,  $\bar{d}_{min}(x_i) = 1/N \sum_k D_k(x_i)$ , é então uma medida das dinâmicas de fundo.

Por exemplo, assumindo um plano de fundo completamente estático,  $\bar{d}_{min}(x_i)$  será zero. Já para um plano de fundo dinâmico, sempre haverá um (pequeno) desvio do valor sendo observado em comparação com os anteriormente observados, e portanto  $\bar{d}_{min}(x_i)$  será maior.

Com essa estimativa das dinâmicas do plano de fundo, e levando em conta que  $R_{inc/dec}$  e  $R_{scale}$  são parâmetros fixos, o limiar de decisão pode ser dinamicamente adaptado para:

$$R(x_i) = \begin{cases} R(x_i) \cdot (1 - R_{inc/dec}) & \text{se } R(x_i) > \bar{d}_{min}(x_i) \cdot R_{scale} \\ R(x_i) \cdot (1 + R_{inc/dec}) & \text{caso contrário} \end{cases}$$

Para um  $\bar{d}_{min}(x_i)$  constante, o limiar de decisão  $R(x_i)$  se aproxima do produto  $\bar{d}_{min}(x_i) \cdot R_{scale}$ . Portanto, um aumento repentino na dinâmica do fundo leva a um lento incremento de  $R(x_i)$  em direção a um limiar de decisão maior. O controle da taxa de atualização apresentado consiste em uma maneira robusta de se tratar diferentes valores de dinâmica de fundo.

## Atualização da Taxa de Aprendizado

Como já mencionado, independente do estado de primeiro plano  $F(x_i)$ , eventualmente todo objeto será fundido com o plano de fundo de acordo com o parâmetro de aprendizado  $T(x_i)$ . Para atenuar o problema, um segundo controle dinâmico para  $T(x_i)$  é utilizado, de forma que a probabilidade do aprendizado do plano de fundo seja lentamente incrementado quando o *pixel* é de fundo e lentamente decrementado quando o *pixel* é de primeiro plano. O problema disso é que um primeiro plano erroneamente classificado demora para se tornar corretamente plano de fundo.

Pode-se supor que, na maioria das vezes em que *pixels* são classificados erroneamente como de primeiro plano, os mesmos pertencem a regiões do plano de fundo

com muito movimento. Portanto, a intensidade do ajuste no controle pode ser adaptada utilizando-se o estimador dinâmico  $\bar{d}_{min}(x_i)$ , definindo-se assim:

$$T(x_i) = \begin{cases} T(x_i) + \frac{T_{inc}}{\bar{d}_{min}(x_i)} & \text{se } F(x_i) = 1 \\ T(x_i) - \frac{T_{dec}}{\bar{d}_{min}(x_i)} & \text{se } F(x_i) = 0 \end{cases}$$

Para ambos os casos existem parâmetros fixos diferentes,  $T_{inc}$  e  $T_{dec}$ , pois na maior parte do tempo os *pixels* são de fundo e, portanto, escolhendo-se parâmetros independentes para o plano de fundo e o primeiro plano, há um ajuste equilibrado de  $T(x_i)$ . Além disso, limites superiores e inferiores  $T_{lower} < T < T_{upper}$  são definidos, de forma que os valores de  $T(x_i)$  não os ultrapassem. O controle mostrado garantirá que, no caso de plano de fundo altamente dinâmico (ou seja, grande  $\bar{d}_{min}(x_i)$ ), o parâmetro de aprendizado  $T(x_i)$  se manterá constante ou sofrerá pouca alteração. Nesse caso de fundo muito dinâmico, o primeiro plano erroneamente detectado não permanecerá assim por muito tempo, porque a probabilidade de atualização  $p = 1/T(x_i)$  não chegará a zero tão rápido. Em um caso ideal, de um plano de fundo completamente estático, a classificação de primeiro plano é bem consistente, e portanto  $T(x_i)$  aumenta rapidamente, validando o modelo de fundo, de forma que retenha menos atualizações.

# Capítulo 4

## Fundamentos de Agrupamento

Neste capítulo, serão apresentados os conceitos de Agrupamento utilizados neste trabalho. Na seção 4.1, será realizada uma breve introdução sobre o tópico; e na seção 4.2, será apresentado o algoritmo *K-means*.

### 4.1 Introdução

Cérebros humanos são bons em perceber regularidades em dados, sendo a separação de objetos em grupos, de acordo com suas similaridades, uma das maneiras de se expressar essa capacidade [22]. Por mais raças de cães ou gatos que existam, as pessoas tendem a agrupar todos sob os nomes “cão” e “gato”. Uma pessoa não precisa ter conhecimento prévio sobre uma raça de cachorro para identificar o “novo” animal como sendo um cão. Embora nunca tenha visto o animal, ela consegue reconhecer aspectos que o assemelham aos exemplos de cães que já tenha visto. Da mesma forma, dados podem ser reunidos em grupos que representem suas similaridades.

À operação de se reunir em grupos objetos que de alguma forma se assemelham se dá o nome de *Agrupamento* (*clustering*). O agrupamento consiste em dividir um conjunto de observações em grupos não sobrepostos, ou *clusters*, em que todos os elementos de um grupo apresentam maior semelhança entre si do que com os elementos de outros grupos. “Maior semelhança”, nesse sentido, significando maior proximidade em relação à medida de distância adotada. Quando um conjunto de dados encontra-se agrupado, todas as suas observações pertencem a algum grupo, e todo grupo pode ser caracterizado por uma única referência, geralmente a média das observações do grupo [23].

Entre as utilizações do agrupamento se encontram classificações biológicas (separação de animais e plantas em grupos e espécies) e redução de tamanho e complexidade de dados (ao invés de armazenar todos os pontos de um grupo, armazena-se apenas o ponto de referência do grupo, tornando muito mais rápida a manipulação

dos dados). Um exemplo comum de agrupamento consiste na consolidação de um conjunto de notas de testes, expressados em porcentagem, em cinco grupos, um para cada letra (A, B, C, D e F). As notas dos testes são as observações, e o ponto de referência de cada grupo é a média das notas do grupo. As letras podem então ser pensadas como substituições simbólicas dos pontos de referências numéricos.

## 4.2 K-means

O algoritmo de agrupamento *K-means*, muito utilizado em mineração de dados, agrupa  $N$  pontos de um espaço de  $I$  dimensões em  $K$  grupos através de repetidas iterações de duas etapas (atribuição e atualização) [22]. Nele, cada grupo é representado por um ponto de referência, seu *centroide*, que consiste na média de todos os elementos do grupo.

A execução do *K-means* é bem simples. Determinada inicialmente a quantidade desejada de grupos, os centroides de cada grupo são então gerados em localizações aleatórias. Logo em seguida, cada ponto é atribuído ao grupo cujo centroide encontra-se mais próximo (etapa de atribuição). Após todos os pontos serem agrupados, o centroide de cada grupo é recalculado (etapa de atualização), levando em conta a posição de todos os pontos pertencentes ao grupo. O algoritmo executa sucessivamente essas duas etapas até convergir (ou seja, até o novo centroide calculado para cada grupo ser exatamente igual ao da iteração anterior).

A figura 4.1 ilustra um exemplo de agrupamento. Primeiramente são determinados os pontos de referência iniciais, representados pelos pontos com preenchimento. Os pontos então são atribuídos aos grupos, de acordo com o ponto de referência mais próximo, e logo em seguida é feita a média dos pontos, atualizando os centroides. As etapas de atribuição e atualização são repetidas até ser atingida a convergência.

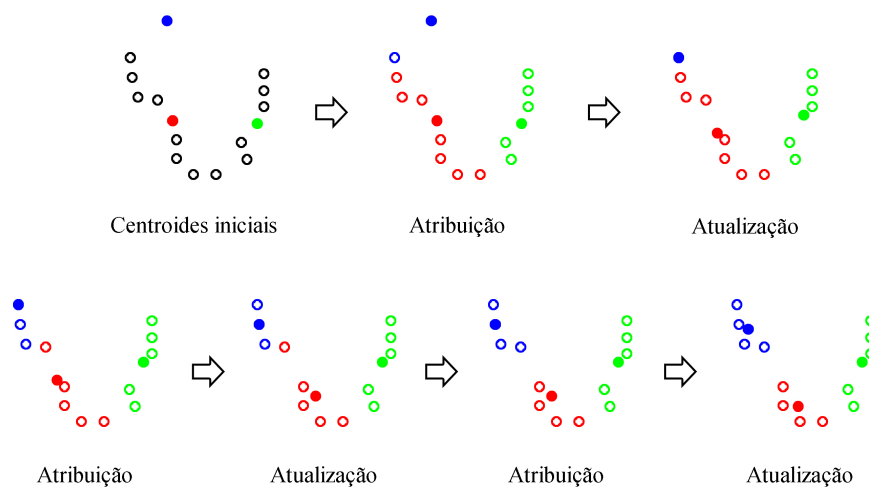


Figura 4.1: Exemplo de agrupamento utilizando o *K-means*



Como o *K-means* é um algoritmo heurístico, não há garantia de convergência para o ótimo global, sendo seu resultado muitas vezes dependente dos grupos iniciais. Como a escolha dos centroides iniciais é aleatória, resultados insatisfatórios podem ser obtidos, como mostrado na figura 4.2. Esse ótimo local, evidentemente, não representa o melhor agrupamento possível.

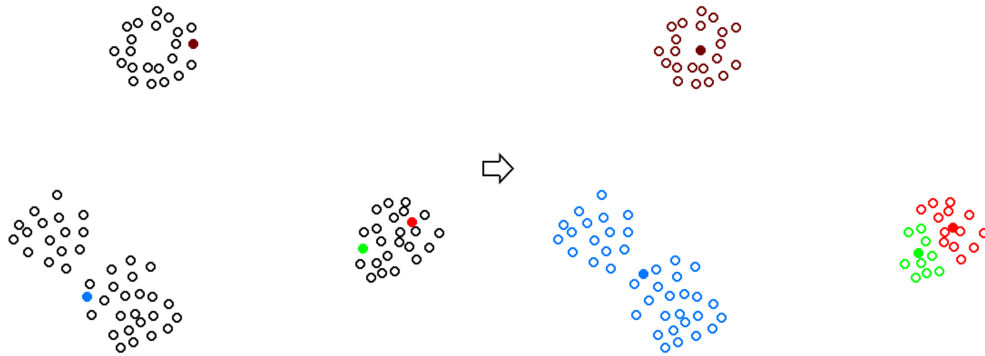


Figura 4.2: Agrupamento ineficaz

Da mesma forma, como o número de grupos é predeterminado, uma escolha inapropriada pode gerar resultados ruins. Caso as observações da figura 4.2 fossem agrupadas em dois grupos, novamente se obteria um agrupamento não ideal (visto que pode-se constatar visualmente a existência de pelo menos três grupos).

Essas limitações do algoritmo, embora indesejadas, também são responsáveis por sua eficiência, tornando fácil sua implementação e possível sua aplicação em conjuntos de dados muito grandes. E por conta de sua velocidade (salvo casos especiais [24]), diversas execuções com condições iniciais alternativas podem ser realizadas, contornando esses problemas. Devido às vantagens citadas, o *K-means* tem sido utilizado em diversas áreas, como segmentação do mercado, visão computacional e geostatística, geralmente servindo como uma etapa de pré-processamento para outros algoritmos (por exemplo, encontrando uma configuração inicial).

# Capítulo 5

## Fundamentos de Redes Neurais Sem Peso

Neste capítulo, serão apresentados os conceitos de Redes Neurais Sem Peso de relevância para esse trabalho, como vistos em [25] e em diversos trabalhos relacionados. Na seção 5.1, uma breve introdução elucidará as necessidades que culminaram no surgimento da computação neural; na seção 5.2, serão apresentados os conceitos fundamentais de Redes Neurais importantes para o entendimento das Redes Neurais Sem Peso; e na seção 5.3, as Redes Neurais Sem Peso serão de fato apresentadas.

### 5.1 Introdução à Computação Neural

Desde os seus primórdios, o computador sempre foi, em sua essência, uma máquina de realizar instruções simples, como cálculos aritméticos e lógicos, em série. Não importando o quão sofisticada fosse sua tarefa, o quão avançada estivesse sua capacidade de processamento ou o quão complexas fossem as novas instruções originadas com as linguagens de alto-nível, o trabalho da máquina sempre se resumiu na execução de longas listas de instruções elementares.

À série de etapas a serem realizadas para a obtenção de um objetivo se dá o nome de *algoritmo*, que é a base da cultura computacional convencional [25]. Todo programa segue rotinas, fórmulas, para alcançar os resultados desejados pelo humano que o projeta. Mesmo em aplicações de Inteligência Artificial que, quando realizadas por humanos, seriam consideradas atividades que requerem inteligência, como jogar xadrez, há algoritmos sendo seguidos passo-a-passo. As regras a serem seguidas, os cálculos a serem feitos para determinar a melhor jogada, entre muitos outros detalhes, estão todos explicitados nos algoritmos.

Algoritmos são, evidentemente, representações do conhecimento humano. Os algoritmos para um computador jogar xadrez não passam de representações do co-

nhhecimento humano com relação a regras e estratégias de xadrez. Da mesma forma, algoritmos para solução de problemas matemáticos refletem o conhecimento humano nessa área, e assim por diante.

Algoritmos precisam, portanto, de humanos para inventá-los e implementá-los nos computadores, e esses, por sua vez, precisam dos algoritmos para realizar suas atividades. Assim, computadores convencionais limitam-se a realizar apenas as tarefas para o qual humanos consigam encontrar algoritmos, e esse simples fato distingue o processamento de informação em computadores convencionais do processamento de informação no cérebro humano: enquanto as pessoas são capazes de desenvolver seus procedimentos através do aprendizado (experiência), computadores precisam esperar que um humano os alimente com os algoritmos necessários para realização de uma tarefa [25].

Existem, porém, tanto tarefas cuja solução é desconhecida (o que inviabiliza a utilização da computação convencional, visto que não há algoritmos para realizá-las), como tarefas em que o desempenho do cérebro é insuperável por algoritmos, como as relacionadas à análise visual. Não existem regras que se equiparem ao desempenho do cérebro em tarefas importantes como reconhecimento de um amigo ou uma rápida reação a uma situação perigosa enquanto dirigindo, por exemplo. O rosto de uma pessoa que tenha sido previamente identificada pode, durante um novo reconhecimento, apresentar sinais de idade, mudanças na feição, uma barba ou corte de cabelo diferente, alterações que, embora em um conjunto de regras possam ser contornadas com a adição de novas regras, resultam em uma gama aparentemente infinita de regras a serem consideradas. Já a mesma tarefa para o cérebro é considerada fácil, bastando, para sua realização, recorrer à experiência armazenada em sua rede de neurônios interconectados.

A “computação” realizada nos neurônios ocorre claramente de forma diferente do método algorítmico, baseado em regras, e identificar como essa rede é capaz de armazenar e utilizar experiência é o objetivo dos pesquisadores de redes neurais. O objetivo da computação neural é, portanto, descobrir como máquinas podem armazenar e utilizar experiência, caso seja necessário projetar dispositivos com habilidades impossíveis de serem obtidas com programação [25].

## 5.2 Redes Neurais

### 5.2.1 Nó Neural

O neurônio consiste na estrutura básica do sistema nervoso. Com seu longo prolongamento denominado axônio, que se ramifica em diversas terminações, o neurônio se conecta a outros neurônios, para onde, através dessa estrutura, envia mensagens em

forma de pulsos elétricos. Esse contato entre neurônios ocorre nas sinapses, que em sua maioria se encontram localizadas nos dendritos (numerosos e curtos prolongamentos do corpo celular), e são responsáveis, portanto, pelo recebimento de pulsos. De forma simplificada, o neurônio recebe pulsos de outros neurônios e, de acordo com essa “entrada”, emite ou não pulsos para outros neurônios.

De forma análoga, um componente eletrônico representando um nó adaptável bem simples, como ilustrado na figura 5.1, recebe sinais de entrada e emite um sinal de saída. Para fins de exemplo, valores de voltagens ou correntes são desconsiderados, levando-se em conta apenas que determinada conexão pode assumir os valores 0 ou 1 (podendo significar, por exemplo, ausência ou presença de um pulso elétrico).

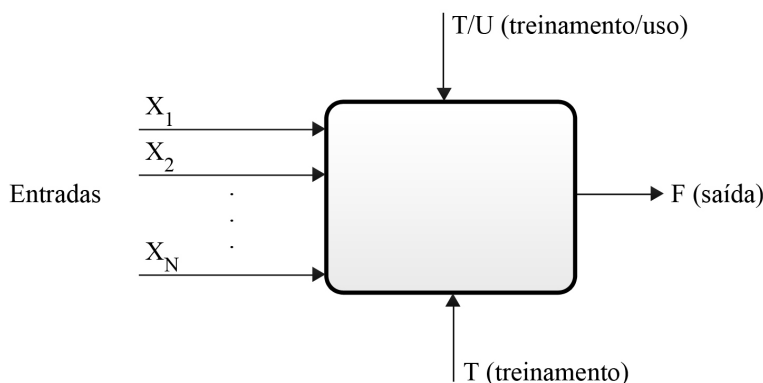


Figura 5.1: Nó simples

As  $N$  entradas (representadas por  $X_1, X_2, \dots, X_N$ ) podem tanto estar conectadas a saídas de outros nós como a alguma fonte externa, e seu conjunto de valores forma o *padrão de entrada* do nó. Para cada padrão de entrada possível (ou seja, para cada combinação de 0 e 1 nas entradas), o nó emite uma saída correspondente, de acordo com sua *função*.

Um exemplo de função para um nó com apenas duas entradas que implementa um ou-exclusivo (o nó responde com 1 apenas quando as entradas diferem) é mostrado na tabela 5.1. Quando os valores tanto em  $X_1$  como em  $X_2$  forem 0, a saída será 0. Quando apenas um deles for igual a 1, a saída será 1, e quando ambos forem 1, a saída será 0.

$X_1$ :	0	0	1	1
$X_2$ :	0	1	0	1
F:	0	1	1	0

Tabela 5.1: Função do nó ou-exclusivo

A determinação da saída em resposta a determinada entrada ocorre quando o nó se encontra em modo de *uso*. O outro modo de operação possível, conhecido

como *treinamento*, é responsável por realizar alterações na função do nó. O modo de operação do nó, que é selecionado através do terminal  $T/U$  (por exemplo, valor 0 para uso e 1 para treinamento), determina então o que o nó fará com a entrada recebida. No caso do treinamento, o nó associa o valor encontrado na entrada  $T$  com o padrão de entrada, alterando assim a função do nó, caso necessário. Isso significa que, quando esse mesmo padrão de entrada for recebido pelo nó em modo de uso, o valor de saída associado (aprendido) vai ser emitido em  $F$ .

Além dos valores 0 e 1 para  $F$  na função do nó, há também o valor “0/1”, representando um valor de saída indefinido. Para tal condição, o nó emite na saída 0 ou 1 de forma aleatória. Essa condição existe quando um padrão de entrada em particular não foi aprendido (treinado), ou se existe alguma espécie de conflito em relação à saída desejada para determinado padrão. Por exemplo, um nó treinado para responder 1 a determinado padrão de entrada que também é treinado para responder 0 ao mesmo padrão terá sua saída indefinida. Da mesma forma, um nó treinado apenas para responder 0 ao padrão 000 e 1 para o padrão 111 terá saídas indefinidas para todos os outros padrões de entrada, como ilustrado na tabela 5.2.

$X_1$ :	0	0	0	0	1	1	1	1
$X_2$ :	0	0	1	1	0	0	1	1
$X_3$ :	0	1	0	1	0	1	0	1
$F$ :	0	0/1	0/1	0/1	0/1	0/1	0/1	1

Tabela 5.2: Exemplo de saídas indefinidas

Essas indecisões das saídas podem ser resolvidas por *regras de disparo*, mecanismo inspirado no funcionamento dos neurônios. É dito que um neurônio “dispara” quando o mesmo emite um trem de pulsos elétricos à taxa de 100 por segundo. Porém, o foco dos estudos no cérebro é principalmente se o neurônio dispara ou não, desconsiderando-se condições intermediárias. De forma análoga, o nó eletrônico usado como exemplo pode ser considerado “disparando” quando a saída é 1 e sem disparo quando a saída é 0.

Uma regra de disparo determina como calcular se o nó deve ou não disparar para determinado padrão, sendo importante notar que é usada para todos os padrões de entrada, não apenas para aqueles não aprendidos durante o treinamento. A regra de disparo que utiliza a distância de Hamming figura como uma das mais aplicadas em redes neurais.

A distância de Hamming corresponde, quando se comparando duas *strings* de mesmo tamanho, ao menor número de substituições necessárias para transformar uma na outra [26], ou seja, indica o grau de similaridade entre ambas. Levando-se em conta que diversos padrões serão treinados tanto para a saída 0 como para a saída 1, a regra de disparo que utiliza a distância Hamming determina que um disparo

será dado apenas quando o padrão de entrada tiver mais elementos em comum com o padrão “mais próximo” do conjunto de padrões com a saída 1 do que com o “mais próximo” dos padrões com a saída 0. Nos casos de empate, o padrão continua com a saída indefinida.

Como exemplo, tem-se o nó de três entradas cujo conjunto de padrões treinados com saída 1 é 111 e 101 e o conjunto de padrões treinados com a saída 0 é 000 e 001, como indicado na tabela 5.3.

X <sub>1</sub> :	0	0	0	0	1	1	1	1
X <sub>2</sub> :	0	0	1	1	0	0	1	1
X <sub>3</sub> :	0	1	0	1	0	1	0	1
F:	0	0	0/1	0/1	0/1	1	0/1	1

Tabela 5.3: Função do nó antes de aplicada a regra de disparo

O padrão 010, que se diferencia do padrão 000 apenas no segundo elemento, do padrão 001 nos segundo e terceiro elementos, do 101 nos três elementos e do 111 nos primeiro e terceiro elementos, recebe portanto a saída 0, por estar “mais próximo” de um padrão com essa saída. Assim, a regra de disparo determina que o nó não disparará para o padrão 010. Em contrapartida, o padrão 011 está equidistante de padrões de ambos os conjuntos (001 e 111, ambos com diferença de um elemento), e assim sua saída continua determinada como 0/1. A tabela 5.4 demonstra a nova configuração da função do nó após a aplicação da regra de disparo.

X <sub>1</sub> :	0	0	0	0	1	1	1	1
X <sub>2</sub> :	0	0	1	1	0	0	1	1
X <sub>3</sub> :	0	1	0	1	0	1	0	1
F:	0	0	0	0/1	0/1	1	1	1

Tabela 5.4: Função do nó depois de aplicada a regra de disparo

A diferença entre a função antes e depois de aplicada a regra de disparo é chamada de *generalização do nó devido à regra de disparo*. Em outras palavras, a regra de disparo garante ao nó o senso de similaridade entre padrões de entrada, permitindo que ele responda de forma satisfatória a padrões não vistos durante o treinamento (padrões similares, mas não idênticos, aos treinados).

## 5.2.2 Rede Associativa

Ao dispositivo de armazenagem de experiência formado por nós neurais, como o apresentado na seção anterior, se dá o nome de *Rede Neural Artificial*. Sua estrutura, isto é, a maneira como estão dispostas suas entradas, seus nós e suas saídas, embora possa apresentar diferentes configurações, basicamente apresenta um dos dois formatos: *feed-forward* ou *feedback*. A diferença entre ambas é a presença de

um *loop*, de uma maneira da informação de saída voltar para a rede, como ilustrado na figura 5.2. É importante notar que, embora não representados, os terminais  $T$  e  $T/U$  devem ser sempre assumidos como existentes e operantes em cada nó da rede.

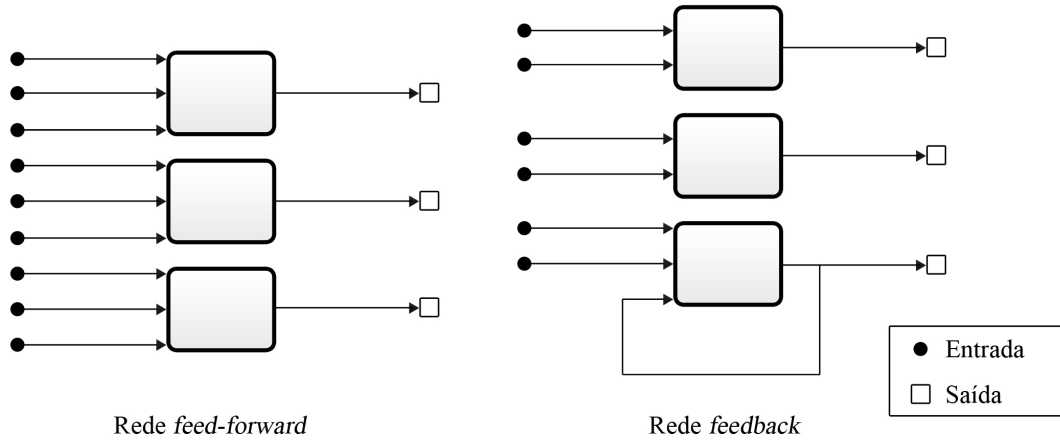


Figura 5.2: Exemplos de redes neurais

A rede *feed-forward*, também conhecida como rede associativa, é assim denominada por associar o padrão apresentado nos terminais de entrada com o padrão nos terminais de saída, como exemplificado pela figura 5.3. Depois de treinada com os padrões de entrada e suas respectivas respostas, a rede é capaz de classificar novos padrões a ela apresentados, indicando seu grau de similaridade com aqueles anteriormente vistos.

O exemplo consiste em uma rede simples de três nós capaz de reconhecer padrões em forma de “T” ou “H”. A rede é treinada para disparar em todas as saídas ( $F_1$ ,  $F_2$  e  $F_3$ ) quando o padrão em forma de “T” é apresentado nas entradas  $X_{11}$  (primeira entrada do primeiro nó) a  $X_{33}$  (terceira entrada do terceiro nó) e para não disparar quando o mesmo é feito com o padrão “H”. A regra de disparo que utiliza a distância de Hamming é então aplicada para generalização dos nós, resultando nas funções mostradas na tabela 5.5.

A figura 5.3 apresenta alguns resultados de testes realizados na rede. Por exemplo, o padrão de teste (a) provê a entrada 110 para o Nó 1, gerando, por “proximidade”, 1 em  $F_1$ , enquanto  $F_2$  e  $F_3$  são os mesmos para o padrão de treino “T”. Da mesma forma, (b) é visto como um padrão “H”. Já (c) produz 101 ou 111, devido ao “0/1” em  $F_2$ . A rede indica então que o padrão (c) parece mais com “T” do que com “H”, mas que a alternância em  $F_2$  pode ser interpretada como um sinal de incerteza.

Essa mesma tarefa de classificação poderia ser realizada por um algoritmo simples: “dado um padrão de entrada, calcule a distância de Hamming do mesmo em relação aos padrões ‘T’ e ‘H’ e classifique-o de acordo com a menor distância”.

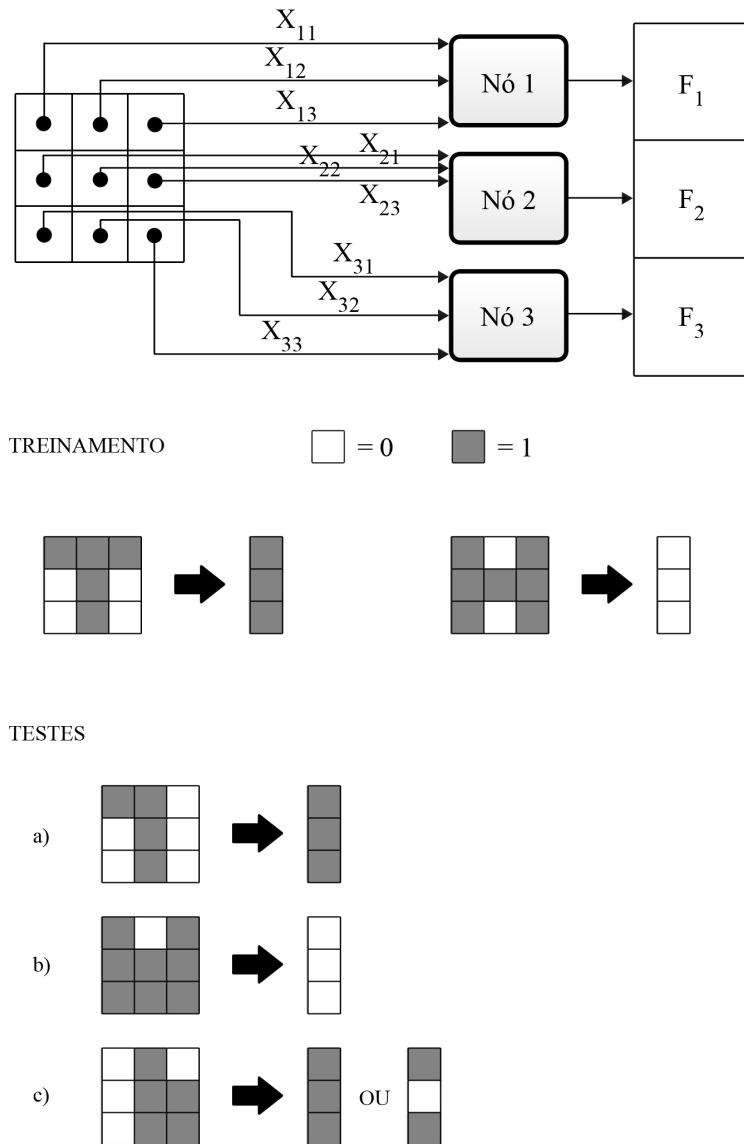


Figura 5.3: Uma rede associativa

Porém, para imagens grandes e cuja definição de classes seja vaga, além de requisitar incontáveis exemplos para o treinamento, o algoritmo se mostra inflexível, pois regiões sem aparente importância para a classificação podem influenciar o resultado.

Tem-se como exemplo a tarefa de se diferenciar um rosto sorridente de um rosto sério. Em uma análise por completo da imagem para determinar a menor distância, caso no treinamento tenham sido utilizadas uma imagem de rosto sorridente e careca e outra de um rosto sério e com cabelo, o algoritmo é capaz de classificar uma imagem de um rosto sorridente e com cabelo como sendo um rosto sério baseado simplesmente na existência de cabelo (pois essa região gera uma maior diferença entre as imagens).

Porém, ao se quebrar o problema em diversas decisões (uma por nó), a rede gera



X <sub>11</sub> :	0	0	0	0	1	1	1	1
X <sub>12</sub> :	0	0	1	1	0	0	1	1
X <sub>13</sub> :	0	1	0	1	0	1	0	1
F <sub>1</sub> :	0	0	1	1	0	0	1	1
X <sub>21</sub> :	0	0	0	0	1	1	1	1
X <sub>22</sub> :	0	0	1	1	0	0	1	1
X <sub>23</sub> :	0	1	0	1	0	1	0	1
F <sub>2</sub> :	1	0/1	1	0/1	0/1	0	0/1	0
X <sub>31</sub> :	0	0	0	0	1	1	1	1
X <sub>32</sub> :	0	0	1	1	0	0	1	1
X <sub>33</sub> :	0	1	0	1	0	1	0	1
F <sub>3</sub> :	1	0	1	1	0	0	1	0

Tabela 5.5: Função dos nós da rede associativa

diversas pequenas regras armazenadas nos próprios nós. O fato de a rede levar em conta simultaneamente a opinião de todos os nós para chegar a uma conclusão é o que garante sua flexibilidade. Ou seja, a regra de disparo sendo utilizada repetidamente para pequenos trechos da imagem é flexível, ao contrário da aplicação do método à imagem como um todo.

### 5.2.3 Modelo McCulloch e Pitts

O contato entre o axônio do neurônio emissor e o dendrito do neurônio receptor ocorre na fenda sináptica (ou sinapse), como já mencionado. As sinapses são conhecidas por influenciar a emissão de pulsos por parte do seu neurônio, sendo denominadas *excitatórias* aquelas que, quando disparadas (ou seja, quando recebem um pulso do neurônio a elas conectadas), encorajam o neurônio a emitir um pulso, e *inibitórias* as que, quando disparadas, desencorajam a sua emissão. Acredita-se que a todo o momento o neurônio decida se deve ou não realizar seu disparo baseado apenas na soma das atividades sinápticas, tanto as excitatórias como as inibitórias, que estejam ocorrendo em determinado instante, disparando sempre que essa soma ultrapassa certo limiar. Embora haja muito a se descobrir sobre o funcionamento dos neurônios, é sabido que o peso das sinapses em um neurônio (isto é, a carga de influência de cada uma) é variável, e é essa variabilidade que confere ao neurônio sua adaptabilidade, sua capacidade de aprendizado [25].

Observando esse mecanismo, McCulloch e Pitts desenvolveram o modelo MCP, que utiliza um nó mais sofisticado que o apresentado anteriormente, uma versão aproximada de um neurônio real. Nesse modelo, cada entrada possui um peso indicando sua carga de influência sobre o disparo do neurônio, como ilustrado na figura 5.4.

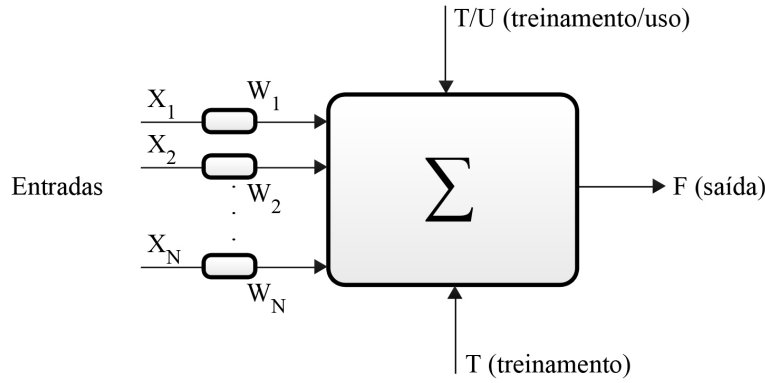


Figura 5.4: Nó do modelo MCP

O estado de um axônio, ou seja, a existência ou não de um disparo pelo seu neurônio, é representado por  $X$ , que pode assumir os valores 1 ou 0 (disparo e ausência de disparo, respectivamente). Já o peso da sinapse é representado por  $W$ , que pode assumir valores entre  $-1$  e  $1$ , sendo os valores negativos aqueles constantes nas sinapses inibitórias. O efeito que uma determinada sinapse exerce sobre um neurônio então é o produto  $XW$ , indicando que, sempre que uma sinapse recebe um pulso do neurônio a ela conectado, seu peso influencia a decisão do neurônio. A regra de disparo consiste então no neurônio realizando a soma de todos os efeitos e a comparando a um limiar  $T$ , sendo o disparo realizado sempre que o limiar for excedido. Em termos matemáticos, o neurônio dispara (saída  $F$  com valor 1) sempre que a seguinte inequação for verdadeira:

$$X_1W_1 + X_2W_2 + \dots + X_NW_N > T$$

Por exemplo, para um nó MCP com apenas duas entradas,  $X_1$  e  $X_2$ , e a função representada pela tabela 5.6, tem-se a seguinte forma geral da regra de disparo:

$$F = 1 \text{ se } X_1W_1 + X_2W_2 > T$$

$X_1$ :	0	0	1	1
$X_2$ :	0	1	0	1
$F$ :	0	1	1	1

Tabela 5.6: Função do nó MCP

Quando ambos  $X_1$  e  $X_2$  possuem o valor 1,  $F$  também é 1, indicando que o limiar deve ser excedido nesse caso. Portanto:

$$W_1 + W_2 > T$$

Da mesma forma, quando apenas uma das entradas tem valor 1, temos:

$$W_1 > T$$

$$W_2 > T$$

E, por fim, como  $F$  é 0 no caso de ambos possuírem o valor 0:

$$0 < T$$

Como os valores de  $W_1$  e  $W_2$  só podem ser entre  $-1$  e  $1$ , o valor de  $T$  deve estar no intervalo  $-2$  e  $2$ , de forma a acomodar  $W_1 + W_2$ . A inequação  $0 < T$ , por sua vez, restringe o valor de  $T$  a um valor maior que 0, e assim por diante. Levando em conta as restrições impostas por cada inequação, chega-se aos valores possíveis para  $W_1$ ,  $W_2$  e  $T$ .

Porém, supondo que a saída  $F$  seja 0 quando  $X_1$  e  $X_2$  são 1 (que corresponde à função da tabela 5.1), a função não pode ser satisfeita, visto que o problema se torna linearmente inseparável. De fato, em um nó MCP de duas entradas, apenas 14 das 16 possíveis tabelas-verdade podem ser obtidas e, quanto maior o número de entradas, maior o número de tabelas-verdade não obtíveis.

A flexibilidade desse modelo vem do fato dos pesos e limiar serem variáveis, se adaptarem a cada novo padrão treinado, identificando os trechos do padrão de entrada de maior influência sobre o resultado. Contudo, um problema inerente ao MCP é que, a cada treinamento, os novos valores devem continuar sendo válidos para os padrões já treinados (ou o comportamento do nó para esses padrões será alterado), o que restringe seus possíveis valores.

Enfim, embora não se tenha conhecimento completo sobre a maneira como o cérebro funciona, o modelo MCP, com sua mecânica aproximada de um neurônio real, tornou-se a base da computação neural, sendo utilizado até hoje no que se refere a Redes Neurais. Dispositivos mais complexos, como o *perceptron*, por exemplo, utilizam o MCP.

## 5.3 Redes Neurais Sem Peso

### 5.3.1 Nó RAM

Ao contrário das Redes Neurais tradicionais, que se baseiam em neurônios com pesos e limiar, como o nó MCP, Redes Neurais Sem Peso baseiam-se em nós de *Memória de Acesso Aleatório* (*Random Access Memory*, ou RAM) [27]. Seu nó, portanto, funciona exatamente como a memória utilizada em computadores, como ilustrado

na figura 5.5. Quando no modo de escrita, o nó recebe um dado de entrada pelo terminal  $E$  e o armazena internamente no endereço indicado pelos terminais  $I$ , que podem assumir, cada um, apenas os valores 0 ou 1. Quando no modo de leitura, o nó apresenta no terminal  $S$  o dado previamente armazenado na posição indicada pelos terminais  $I$ .

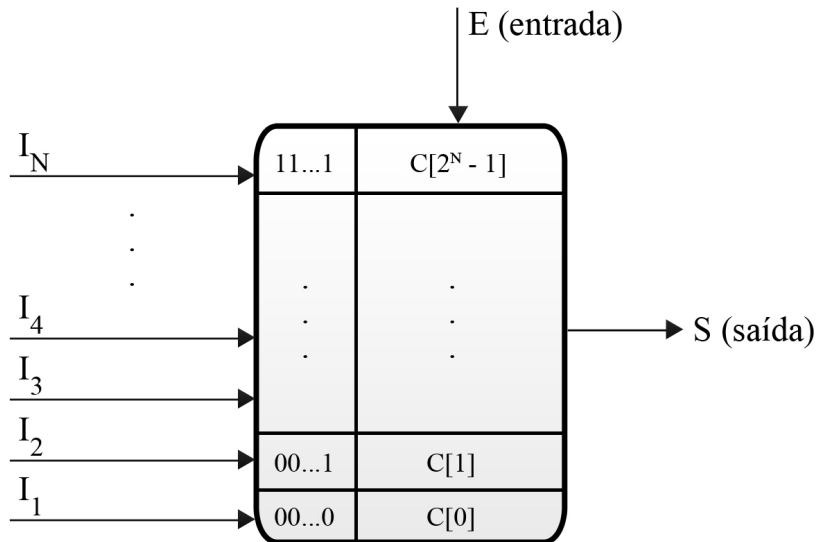


Figura 5.5: Nó RAM

Por exemplo, caso um nó com três entradas  $I$  esteja armazenando um dado e  $I_1$  e  $I_2$  possuíam o valor 0 e  $I_3$  possuía o valor 1, o dado será armazenado na posição 001. Da mesma forma, esse dado será recuperado no modo de leitura quando as entradas  $I$  indicarem a posição 001. Esse nó, portanto, é capaz de armazenar oito dados diferentes, pois esse é o número de combinações de 0 e 1 possíveis. Generalizando esse conceito, um nó com  $N$  entradas  $I$  é capaz de endereçar e, por conseguinte, armazenar dados em  $2^N$  posições.

O treinamento em uma Rede Neural Sem Peso então ocorre de maneira bem intuitiva: dado um padrão de entrada (binário) mapeado para as entradas  $I$  de um nó, a resposta esperada é armazenada na posição correspondente. Durante o teste, ao alimentar o nó com o padrão a ser testado, o nó responde com o dado armazenado na posição indicada pelo padrão. Devido a esse mapeamento direto do padrão com sua saída, esse nó é capaz de computar todas as funções binárias possíveis, ao contrário dos nós com peso, que, como já mencionado, só conseguem computar funções linearmente separáveis [27].

Além dessa flexibilidade, o mapeamento direto em um nó RAM confere à rede uma alta velocidade de treinamento, contrastando com o complexo processo de treinamento em uma rede neural convencional, em que os pesos precisam ser alterados

para cada novo padrão treinado. É de se notar também que, enquanto o recálculo dos pesos possa vir a alterar o comportamento do nó MCP com relação aos padrões treinados anteriormente, o mesmo não ocorre em um nó RAM. Por fim, o custo de implementação de uma rede sem peso é bem baixo se comparado às redes convencionais, visto que utiliza componentes amplamente disponíveis no mercado e a preços acessíveis.

Um último aspecto importante é que a generalização, que pode ser entendida como a habilidade de produzir a saída correta para padrões de entrada não vistos durante o treinamento [28], não existe no nó RAM convencional, pois as saídas corretas só são armazenadas nas posições do nó vistas durante o treinamento. Porém, como será visto adiante, a generalização em uma Rede Neural Sem Peso tanto pode ser implementada pela própria rede como em cada nó através de métodos como a distância de Hamming.

### 5.3.2 Discriminador RAM

Um *discriminador RAM* é definido como a mais simples rede de nós RAM com propriedades de generalização. Como ilustrado na figura 5.6, sua estrutura consiste em uma camada de  $K$  nós RAM com  $N$  entradas cada que recebe padrões binários de  $KN$  bits, estando cada nó mapeado a um trecho diferente do padrão de entrada.

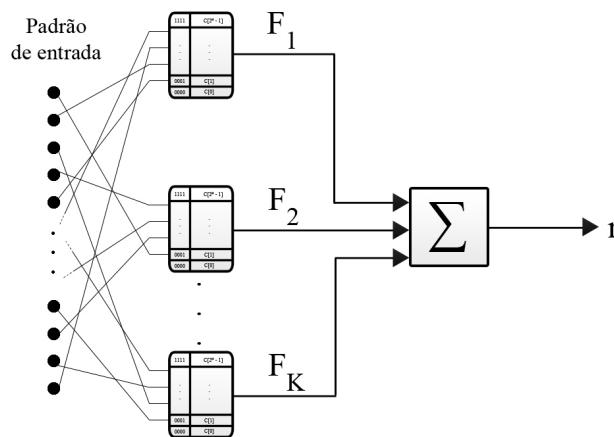


Figura 5.6: Um discriminador RAM

Em um típico discriminador, cada nó RAM, que inicialmente armazena o valor 0 em todas as suas posições de memória, é treinado para responder com 1 apenas para os padrões vistos no treinamento. Ou seja, cada nó é treinado para responder com 1 sempre que, ao analisar seu respectivo trecho do padrão de entrada, se deparar com os padrões vistos em seu treinamento. Um padrão de entrada sendo testado então

gera uma resposta  $r$  do discriminador, que se traduz no número de nós RAM que responderam com 1.

As diferentes respostas dadas por um discriminador a padrões de testes podem ser interpretadas como indicadores do grau de semelhança desses padrões com os padrões vistos no treinamento. O padrão testado não precisa ser exatamente idêntico aos padrões utilizados no treinamento para ser classificado como pertencendo à mesma classe; basta um número considerável de “trechos” desse padrão ser parecido com um visto no treinamento para que a similaridade entre ambos seja notada. Por conta disso, esse dispositivo de pontuação garante a essas redes tanto generalização como tolerância a ruídos, as mesmas propriedades encontradas nos modelos de redes neurais baseados em pesos sinápticos [27].

### 5.3.3 WISARD

Quando há mais de duas classes a serem determinadas, mais redes RAM são usadas, sendo cada uma treinada para cada classe. Esse é o caso da *WISARD*, rede sem peso que utiliza discriminadores. Depois de treinados os discriminadores, cada um com exemplos da sua classe correspondente, um padrão de entrada é testado em todos ao mesmo tempo. O padrão de entrada então é considerado da classe cujo discriminador correspondente obtiver maior pontuação (isto é, aquele com maior quantidade de nós respondendo com 1).

Na *WISARD*, as entradas dos nós são mapeadas aleatoriamente com o padrão de entrada, como pode ser visto na figura 5.6. Dada a cadeia de *bits* que forma o padrão de entrada (que pode ser a binarização de uma imagem, por exemplo), cada uma das entradas de um nó é associada exclusivamente a um *bit* dessa cadeia, e essa maneira como a rede está conectada à entrada é chamada de *mapeamento de entrada*, ou *padrão de conectividade* [28]. Embora seja aleatório, o mapeamento de entrada é um parâmetro fixo da rede, sendo utilizado igualmente para todos os diferentes discriminadores.

Embora o nó RAM convencional seja eficiente, seu conceito apresenta um indesejado problema de ambiguidade. Quando acessada por um padrão de teste, uma posição da RAM que contenha o valor 0 pode tanto significar que o padrão é um contraexemplo da classe referente ao discriminador do nó, como pode significar que o padrão pertence à classe, mas não fora visto no treinamento. Em ambos os casos, o nó responde com 0, o que não é desejável no segundo caso. Para contornar esse problema, nós mais elaborados, que trabalham com probabilidades, foram desenvolvidos [28].

### 5.3.4 GRAM e VGRAM

O *Nó Lógico Probabilístico* (*Probabilistic Logic Node*, ou PLN) difere de um nó RAM pelo fato de armazenar na posição de memória endereçada um número de dois *bits* (ao contrário de apenas um *bit*). Esse número, ao invés de representar a resposta a ser dada pelo nó, como acontece em um nó RAM, representa a probabilidade do nó disparar (ou seja, de responder 1). Em outras palavras, um nó PLN consiste em um nó RAM aprimorado com um gerador de saída probabilístico [28].

Os valores armazenados em um nó PLN, isto é, os valores que as probabilidades de disparo podem ter, são 0, 1 e  $u$  (representando “não definido”). A saída para 0 e 1 são, obviamente, 0 e 1, e, para o valor  $u$ , o nó escolhe, de forma aleatória e com igual probabilidade, um dos dois valores. O valor  $u$ , que é armazenado em todas as posições de memória de todos os nós antes do treinamento, representa então a ignorância da rede, permitindo a ela demonstrar sua hesitação quando confrontada com uma situação de incerteza (como a mencionada na seção 5.3.3).

Como apenas as posições acessadas pelos padrões de treino são modificadas durante o treinamento, a PLN sofre com excesso de hesitação quando a quantidade de padrões de treino é muito pequena. Isso ocorre porque muitas posições de memória continuam com valor  $u$ . Outra característica das PLN é que, assim como o nó RAM convencional, sua generalização ocorre apenas em nível de rede; os nós individuais não possuem propriedades de generalização. Devido a esses fatores, um novo nó baseado no PLN foi desenvolvido, adicionando ao seu algoritmo de aprendizado uma *fase de espalhamento* (*spreading phase*), realizada logo após a fase de treinamento.

A intenção desse “espalhamento” é gerar uma “área de influência” em torno dos padrões treinados, englobando as posições de memória mais próximas e as assimilando como pertencentes à mesma classe do padrão em questão. Dado um número máximo de iterações definido pelo usuário, o processo se repete sucessivamente até atingir esse limite. No primeiro passo, apenas as posições com valor  $u$  que estão a uma distância de Hamming de 1 de pelo menos um dos padrões armazenados recebem o mesmo valor desse mais próximo (aqueles que estejam à mesma distância de padrões de classes diferentes, no que se caracteriza como uma contradição, continuam com o valor  $u$ ). No segundo passo, apenas as posições com distância de Hamming de 2 são afetadas, e assim por diante.

O algoritmo de espalhamento, em outras palavras, generaliza a informação presente nos padrões de treino, fazendo com que padrões não vistos na fase de treinamento, mas que enderecem áreas “sob influência”, sejam classificados como pertencendo à classe correspondente. Caso uma região contraditória seja acessada, a saída é aleatoriamente escolhida entre 0 e 1 (devido ao  $u$  presente). Como, devido a esse algoritmo, o nó individual adquire a capacidade de generalizar, ele é chamado de

*Memória de Acesso Aleatório Generalizante (Generalizing Random Access Memory, ou GRAM) [29].*

Apesar de serem simples, de treinamento rápido e de fácil implementação, as redes sem peso podem apresentar um custo de armazenamento extremamente alto, se levado em conta que a quantidade de posições de memória endereçáveis em um nó cresce exponencialmente com a quantidade de entradas do mesmo ( $2^N$ , sendo  $N$  o número de entradas). As redes *GRAM Virtuais (Virtual GRAM, ou VGRAM)* contornam esse problema ao requisitar capacidade de armazenagem apenas para os dados dos padrões de treino. Ao contrário do nó GRAM, que armazena um valor de saída para cada posição endereçável, o nó VGRAM armazena apenas os pares entrada-saída vistos durante o treinamento, garantindo uma economia de memória. O algoritmo de espalhamento é então usado apenas durante os testes. Assim que um novo padrão é testado, cada nó compara a entrada desse padrão com as entradas nos pares entrada-saída armazenados e, através da distância de Hamming, identifica o par mais próximo, decidindo assim a saída correspondente. No caso de empate entre pares com mesma distância, a saída é escolhida aleatoriamente entre esses pares.

Um nó VGRAM com apenas três entradas,  $X_1$ ,  $X_2$  e  $X_3$  e saída  $F$  possui três pares entrada-saída treinados, como visto na tabela 5.7. Assim que o padrão de entrada 100 é testado, o nó analisa suas entradas armazenadas, comparando cada uma ao novo valor apresentado. Como a entrada do par 1 difere tanto em  $X_1$  e  $X_3$ , a distância de Hamming para esse par é 2. Realizando o mesmo cálculo para os pares restantes, descobre-se que a entrada do par 3 é a mais próxima da entrada apresentada, com distância de 1. Assim, o valor do par 3 é dado como resposta para o padrão de entrada testado, no caso, 0.

	$X_1$	$X_2$	$X_3$	F
Par 1	0	1	0	1
Par 2	0	1	1	0
Par 3	1	0	1	0
Teste	1	0	0	<b>0</b>

Tabela 5.7: Exemplo do espalhamento na VGRAM

É importante notar que, embora haja ganho em armazenamento em uma rede VGRAM, há um aumento no tempo de resposta da rede, visto que ela passa a realizar mais cálculos durante a execução. Mesmo assim, a economia de memória proporcionada em conjunto com a crescente capacidade computacional da atualidade tornam a VGRAM uma poderosa Rede Neural Sem Peso.



# Capítulo 6

## Classificador de Densidade

Neste capítulo, será apresentada a solução desenvolvida nesse trabalho. A seção 6.1 apresenta o sistema, explicando de forma geral o seu funcionamento; e a seção 6.2 vai mais a fundo, explicando em detalhes cada passo do seu processamento.

### 6.1 Introdução

Como visto no capítulo 2, a taxa de fluxo, a velocidade e a densidade figuram como os parâmetros de fluxo macroscópicos mais utilizados pelos engenheiros de tráfego. Devido às relações entre esses parâmetros (figura 2.1), é possível, com o conhecimento de apenas dois deles, se determinar o valor do terceiro parâmetro. Dessa forma, o comportamento do fluxo de tráfego como um todo (em uma via) pode ser compreendido sabendo-se apenas os valores de dois dos parâmetros macroscópicos.

Tendo isso em mente, o Classificador de Densidade foi desenvolvido para, da forma mais automática possível (ou seja, com o mínimo de interferência humana), identificar o comportamento do fluxo de tráfego na via observada. Dada uma sequência constante de imagens geradas por uma câmera fixa em determinada via, o sistema é capaz de, segundo a segundo, determinar a densidade e a velocidade relativas da via, seja seu fluxo interrompido ou ininterrupto, assim como, de tempos em tempos, indicar a situação da mesma em relação ao fluxo e à densidade.

De forma mais específica, o sistema analisa uma imagem da via por segundo (independente da quantidade de imagens geradas pela câmera a cada segundo) e responde se a densidade é inexistente (ausência de carros), baixa, média ou alta em relação à densidade máxima suportada pela via, e se a velocidade média dos carros é zero, baixa, média ou alta, como ilustrado na figura 6.1. Além disso, a cada cinco minutos o sistema indica em qual região do gráfico de *Fluxo x Densidade* (figura 2.1) a via observada se encontra. Os valores gerados pelo sistema, apesar de não quantitativos, fornecem uma maneira eficiente de se determinar congestionamentos de tráfego, vias vazias, vias com tráfego bom, com fluxo intenso, entre

outras situações, podendo ser empregados em sistemas automatizados de controle de tráfego.



Figura 6.1: Interface do sistema

Desenvolvido em C++ na IDE Microsoft Visual Studio 2010 [30], o Classificador de Densidade utiliza a biblioteca OpenCV 2.1 [31] para manipulação de imagens e, em conjunto com a biblioteca MFC 10.0 [32], para gerenciar a interface do sistema.

Embora o foco do trabalho tenha sido um sistema completamente isento de interação humana, uma ação por parte do usuário torna-se necessária para seu correto funcionamento: a determinação da área de interesse, a ser realizada para cada via observada. Antes de o sistema iniciar a análise das imagens em uma via, o usuário demarca os pontos que formam a região da via a ser analisada, como ilustrado na figura 6.2, e indica que seu processamento já pode ser iniciado.



Figura 6.2: Definindo a região de interesse

Para identificar a situação presente na via, o Classificador de Densidade realiza basicamente duas operações: identificação dos veículos nas imagens e aferição da densidade a partir dessa informação. Para realização da segmentação das imagens (destacando os veículos na via), o sistema apresenta dois métodos alternativos: um próprio, bem simples, mas rápido, e um já existente, sofisticado, porém mais lento (PBAS). Já a classificação da densidade a partir das imagens geradas pode ser realizada tanto pelo método algorítmico implementado, como pela VGRAM utilizada.

Além da classificação de densidade descrita, o sistema dispõe de uma VGRAM encarregada de detectar a região do gráfico de *Fluxo x Densidade* em que a via se encontra.

Todos os métodos citados serão vistos na próxima seção, em que serão explicados em detalhes os procedimentos realizados pelo sistema, desde o tratamento das imagens de entrada até a determinação dos resultados.

## 6.2 Classificador de Densidade

### 6.2.1 Detecção de objetos

O primeiro passo para se determinar o comportamento do fluxo de tráfego em uma via consiste na detecção dos carros que por ela trafegam. Tanto para a determinação da densidade como para a determinação da velocidade, torna-se necessária uma forma de se perceber a presença de veículos na pista.

Levando-se em conta que as imagens geradas por câmeras de tráfego são estáticas, uma abordagem baseada na diferença de imagens foi experimentada. A natureza fixa das imagens, que resulta em um fundo constante quando observadas imagens consecutivas, permite a descoberta de novos elementos (carros, ônibus, pessoas) quando esses surgem entre uma imagem e outra.

Dadas duas imagens de mesmas dimensões, todos os *pixels* de uma imagem possuem um *pixel* correspondente na outra. Em outras palavras, o iésimo *pixel* da primeira imagem,  $P_i$ , possui um *pixel* correspondente na segunda,  $P'_i$ , localizado nas mesmas coordenadas (altura e largura da imagem). Ambas as imagens são então consideradas idênticas quando, para todo *pixel*  $P_i$ , seu correspondente  $P'_i$  possui a mesma cor.

Utilizando-se desses princípios, o método de comparação de imagens consiste na comparação de cores, *pixel a pixel*, de duas imagens, resultando em uma representação visual da diferença. Na imagem resultante, os *pixels* brancos representam igualdade de cor entre os *pixels* correspondentes ao passo que todo *pixel* preto representa diferença, como visto na figura 6.3, que ilustra a diferença entre duas imagens binárias (seus *pixels* só podem assumir duas cores, branco ou preto).

Embora funcione de modo satisfatório em imagens binárias, esse método encontra problemas em imagens coloridas, como as imagens RGB, pois, para dois *pixels* serem considerados idênticos, as intensidades devem ser exatamente as mesmas em cada canal de cor (vermelho, verde e azul). De fato, mesmo em imagens aparentemente idênticas ao olho humano, a diferença de pelo menos uma unidade na intensidade de apenas um dos canais de cor de um *pixel* (o que se traduz em uma nova cor) já é suficiente para diferenciá-lo do seu correspondente. A imagem 6.4 ilustra essa

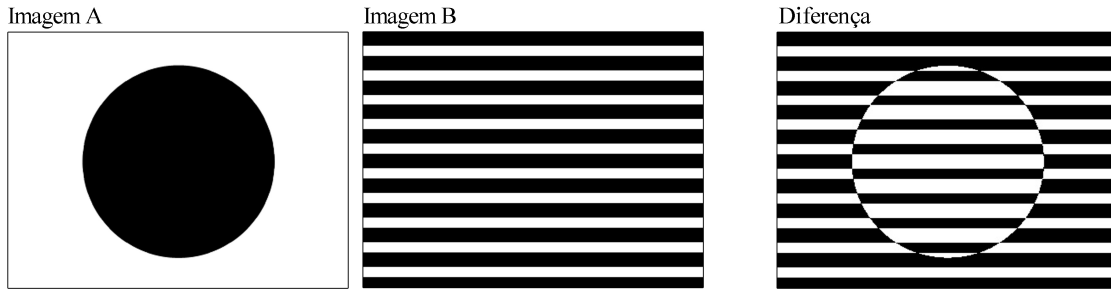


Figura 6.3: Diferença de imagens binárias

situação; mesmo em regiões aparentemente semelhantes há uma grande diferença de cores sob uma perspectiva computacional.



Figura 6.4: Diferença de imagens RGB

Ainda que, à primeira vista, o método de comparação aparente ser ineficaz para imagens coloridas, resultados interessantes podem ser obtidos com a adição de intervalos de tolerância à comparação das cores. Sejam  $R_i$ ,  $G_i$  e  $B_i$  as respectivas intensidades dos canais vermelho, verde e azul do *pixel*  $P_i$  em uma imagem RGB de 8 *bits* (que limita o valor de cada canal para um entre 0 e 255) e  $R'_i$ ,  $G'_i$  e  $B'_i$  seus correspondentes em  $P'_i$ . De acordo com o método de comparação,  $P_i$  e  $P'_i$  possuem a mesma cor apenas quando  $R_i = R'_i$ ,  $G_i = G'_i$  e  $B_i = B'_i$ , o que se mostra problemático, como já visto. Porém, ao se criar intervalos de tolerância em torno dos valores de cada canal de  $P_i$ , permitindo assim que ambos os *pixels* sejam considerados iguais todas as vezes em que os valores de  $P'_i$  estiverem dentro desses intervalos, imagens satisfatórias da diferença são obtidas. Sendo  $d$  a distância do valor de um canal de  $P_i$  até os limites superior e inferior do intervalo a ser gerado,  $P'_i$  será considerado igual a  $P_i$  quando:

$$R_i - d < R'_i < R_i + d$$

$$G_i - d < G'_i < G_i + d$$

$$B_i - d < B'_i < B_i + d$$

A imagem 6.5 ilustra a comparação de duas imagens para diferentes valores de

*d.* Nota-se que, para intervalos de tolerância mais brandos, o ruído na imagem resultante some, prevalecendo apenas os *pixels* de maior contraste entre as imagens comparadas, o que revela uma maneira eficaz de perceber os carros na via. Até mesmo os veículos trafegando em outras vias (como as vistas ao fundo e o viaduto) são delineados, como pode ser observado na parte superior das imagens de diferença.

O método, como descrito, figura como a base da solução desenvolvida, sendo usado tanto para determinação da densidade, como para a determinação da velocidade, como será visto posteriormente. Por se mostrar um valor razoável,  $d = 50$  foi utilizado.

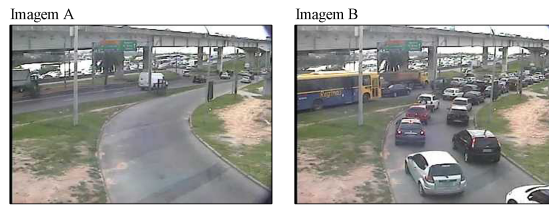
Por fim, é importante notar que, embora a grande maioria das câmeras gerem imagens estáticas, operadores de tráfego de uma cidade podem movimentar remotamente esses dispositivos, de modo a capturar, por exemplo, detalhes de acidentes ou outras situações anormais. Nesses casos, um reposicionamento para a posição inicial torna-se necessário devido à natureza do sistema.

## 6.2.2 Detecção de movimento

Com a utilização da diferença de imagens, a detecção de movimento torna-se uma tarefa tão trivial quanto determinar que há movimento quando existem *pixels* pretos na imagem da diferença e que há ausência de movimento quando a imagem da diferença é completamente branca. De fato, dado que o método de comparação é utilizado em imagens consecutivas, geradas segundo a segundo pela câmera posicionada na via, a aparição de um objeto na imagem da diferença indica claramente a presença de movimento.

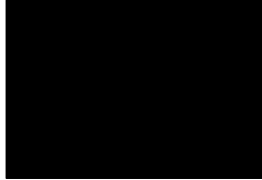
Uma peculiaridade da detecção de movimento por meio do método de comparação consiste no fato de que duas situações diferentes podem ocasionar uma imagem da diferença completamente branca: a total ausência de veículos na pista em ambas as imagens e a presença de veículos que não se locomovem de um segundo para o outro (um congestionamento, por exemplo). A figura 6.6 ilustra essas situações. Nota-se que, embora haja diferença nas vias em torno, na via em destaque não há *pixels* indicando diferença. A maneira de determinar qual das duas situações se apresenta em determinado momento leva em conta o cálculo da densidade na via, que será discutido posteriormente.

Como pode ser observado na figura 6.6, as vias ao redor da via em destaque influenciam a comparação de imagens e conseqüentemente a detecção de movimento no trecho que se quer analisar. Por conta disso, uma maneira de eliminar esses ruídos se faz necessária, de forma a isolar a região de interesse da análise, permitindo que uma correta aferição de velocidade e densidade seja realizada.



Diferença

d = 0



d = 10



d = 20



d = 30



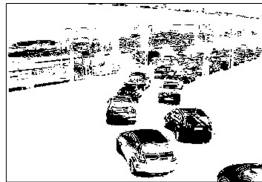
d = 40



d = 50



d = 60



d = 70



Figura 6.5: Diferença de imagens utilizando intervalo de tolerância

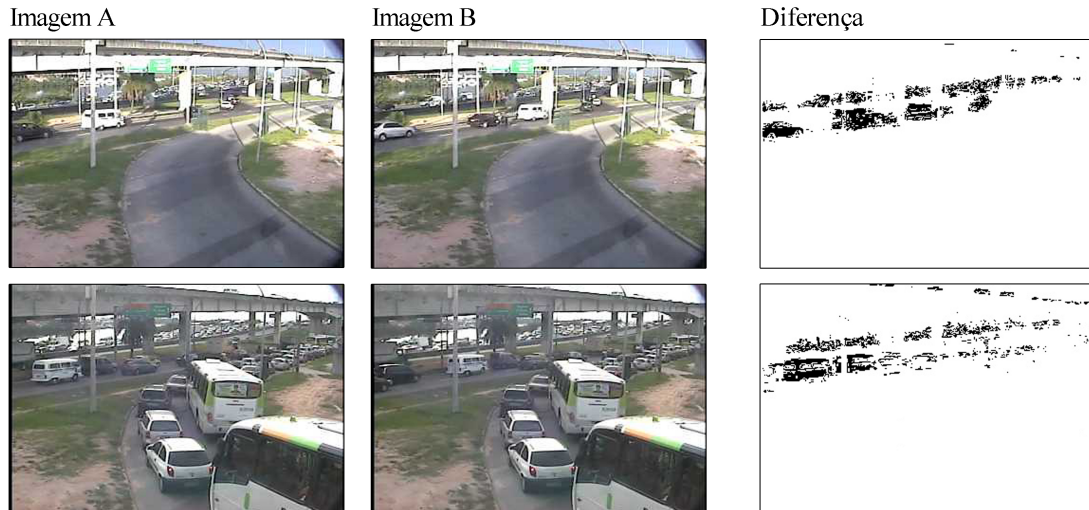


Figura 6.6: Diferença de imagens em uma via vazia e em uma congestionada

### 6.2.3 Área de Interesse

Em virtude da necessidade de se excluir a influência de outras vias sobre a análise da via de interesse, aliada ao objetivo de tornar a solução o mais automatizada possível, um método de detecção automática da região de interesse, baseado na diferença de imagens, foi desenvolvido. Levando em conta que as câmeras são posicionadas de forma a dar destaque à via de maior interesse, bastaria, depois de detectadas todas as regiões onde ocorrem fluxos de tráfego, tomar aquela com maior presença na imagem como sendo o foco da análise.

Seja a imagem da frequência uma imagem em tons de cinza (também conhecida como *preta e branca*) completamente branca (significando que todo *pixel* possui o valor 255 em seu único canal de cor) em que, para cada imagem da diferença analisada, o valor de determinado *pixel* seja subtraído em uma unidade sempre que o *pixel* correspondente for preto. Ao final da análise de todas as imagens da diferença, as regiões mais escuras representarão a maior frequência de ocupação, sendo 0 o valor que indica a máxima frequência. Para evitar saturação (todos os *pixels* nas regiões de interesse ficando 0 pelo simples fato do número de imagens da diferença ser exorbitante), um decaimento é aplicado de tempos em tempos à imagem da diferença, o que significa adicionar determinado valor fixo a todos os seus *pixels* periodicamente, realizando um leve “clareamento” na imagem.

A figura 6.7 ilustra a imagem da frequência gerada ao longo de um dia para a via observada. Nota-se na imagem que há frequência acentuada até em regiões fora da pista, resultado da natureza dos veículos altos, como ônibus e caminhões, que, em conjunto com a perspectiva da imagem, extrapolam os limites da via. Pode-se notar, também, que as frequências das vias adicionais se fundem com a da via de



interesse, dificultando a decisão do quê deve ser considerado pertencente à região de interesse.



Figura 6.7: Frequência observada em uma via ao longo de um dia

Portanto, a intenção inicial, de se descobrir todos os fluxos existentes em uma imagem e em seguida optar por aquele de maior presença, se mostrou inviável, dado que diferentes vias podem convergir na imagem da frequência e não há uma maneira de se diferenciar quando são vias separadas e quando fazem parte de uma mesma via. Por conta dessa impossibilidade de se perceber de forma automática a região de interesse, a determinação do que constitui a região de interesse ficou a cargo do usuário.

Por esse motivo, foi criada uma interface em que o usuário define a região da imagem a que a análise deve se restringir, como mostrado na figura 6.2. Após definida a região de interesse, o sistema então passa a descartar toda informação não contida ali.

Entre as vantagens oferecidas pela nova abordagem figura a possibilidade de uma mesma câmera ser usada para análise de diferentes vias, bastando para isso o usuário delimitar cada uma em uma região de interesse. O sistema pode até mesmo ser usado para analisar calçadas nas áreas próximas a um semáforo, de forma a identificar a densidade do corpo de pedestres esperando para atravessar a rua.

Embora sem muita utilidade para a solução desenvolvida, a técnica criada para a geração da imagem da frequência possui possíveis aplicações em outros problemas. Representando uma maneira de se determinar automaticamente a região de interesse em uma sequência de imagens estáticas, essa técnica poderia ser usada para, por exemplo, determinar as pistas mais usadas em uma via, as áreas mais frequentemente ocupadas em uma praça de alimentação, ou até mesmo os caixas automáticos mais utilizados em um banco.



## 6.2.4 Determinação de densidade

A determinação da densidade relativa, que representa o número de veículos que ocupam determinada extensão da rodovia em comparação com a quantidade máxima suportada no mesmo trecho, torna-se possível pelas imagens de diferença a partir do momento em que se delimita a região de interesse. Dado que, para determinada imagem, se tenha o conhecimento da quantidade de *pixels* que formam a região de interesse, a razão dos *pixels* de diferença sobre essa quantidade total resulta em uma boa estimativa da área ocupada na via.

Porém, como pode ser observado na figura 6.8, a imagem da diferença de duas imagens consecutivas, que de forma eficaz indica a presença de movimento, peca ao exibir duas vezes o mesmo veículo, de maneira sobreposta, resultando em uma imagem confusa. De fato, a diferença entre uma imagem da via vazia com a do instante sendo analisado gera uma imagem da diferença mais condizente com o estado atual da rodovia, como pode ser observado na figura 6.9. Por conter informações mais precisas para a análise da densidade, a diferença utilizando-se imagens da pista vazia torna-se de grande importância para essa análise.



Figura 6.8: Diferença de imagens consecutivas com movimento



Figura 6.9: Diferença de imagens usando a imagem de via vazia

De forma a sempre ter a imagem de pista vazia mais recente para realizar a diferença com a imagem corrente, o sistema trata de armazenar esse tipo de imagem sempre que possível. Quando a diferença de duas imagens consecutivas é completamente branca (indicando ou ausência de carros ou presença de carros com ausência

de movimento), a comparação com a imagem de pista vazia é feita para se determinar qual a situação presente. Uma vez que essa nova diferença resulte em uma imagem completamente branca, a nova imagem é então armazenada como sendo a imagem de pista vazia mais recente. O termo *pista vazia* se refere apenas à região de interesse, não importando, portanto, a situação das vias adicionais ou até mesmo dos trechos da via analisada que não estejam dentro dos limites de interesse.

Um aspecto importante dessa abordagem é que a primeira imagem de pista vazia precisa ser assinalada pelo usuário. Realizada logo após a delimitação da região de interesse, essa atividade direciona o sistema a corretamente determinar novas imagens de pista vazia, justamente por uma imagem pré-existente dessa natureza ser a única maneira de se diferenciar uma via vazia de uma muito congestionada, sem movimento.

Um problema aparente desse método reside na comparação de imagens com iluminações diferentes. A comparação de uma imagem de pista vazia obtida pela manhã com uma imagem qualquer obtida à noite, por exemplo, não funciona, pois a diferença de iluminação induz até duas imagens de pista vazia a apresentarem diferença, como ilustrado na figura 6.10. Porém, a natureza do método de comparação de imagens, com seu intervalo de tolerância, permite ao sistema acompanhar a gradual mudança de iluminação ao longo de um dia, adquirindo corretamente as imagens de pista vazia durante as manhãs, tardes e noites. Outro fator que também contribui para o acompanhamento de pista vazia é o fato da porcentagem da diferença ser discretizada. Ou seja, valores menores que 1% são arredondados para 0%, indicando ausência de diferença, o que faz com que pequenos ruídos não sejam considerados. Para os casos de pistas que apresentam fluxo acentuado e constante durante o dia, regiões de interesse menores são mais indicadas, de forma a permitir ao sistema que calibre corretamente suas imagens de pista vazia ao longo do dia. A figura 6.11 ilustra alguns exemplos da variação de iluminação processada pelo sistema ao longo de um dia.



Figura 6.10: Diferença de imagens com iluminação diferente

Mesmo representando de forma eficiente a presença dos carros na via, a imagem da diferença com imagens de pista vazia por si só não satisfaz a análise de densidade.



Figura 6.11: Iluminações diferentes ao longo do dia

Como pode ser visto na figura 6.9, a silhueta dos carros não é completa; existem diversas regiões brancas em seus interiores, que influenciam o cálculo da densidade. Uma frota de carros escuros ou muito claros, que exibam bastante contraste com o asfalto, resultará em uma densidade maior que uma frota de mesmo tamanho de carros com cores próximas as do asfalto, devido ao fato de apresentarem mais *pixels* pretos na imagem de diferença.

Para contornar esse problema, o filtro morfológico de fechamento (seção 3.3) é aplicado nas imagens da diferença antes de calculada a densidade. Utilizando-se o filtro com um elemento de estruturação em forma de um quadrado 3x3 e sete iterações (significando que são realizadas sete dilatações e depois sete erosões), resultados satisfatórios são obtidos, como os apresentados na figura 6.12. Cabe salientar que, como na convenção da morfologia *pixels* brancos são os de objeto e *pixels* pretos, de fundo (justamente o contrário do ocorrido nas imagens da diferença), as operações realizadas no sistema são invertidas (ou seja, é utilizada uma abertura para se obter resultado semelhante a um fechamento em imagens convencionais).

Definida a melhor maneira de se gerar a imagem da diferença, resta ao sistema apenas determinar qual a densidade representada pelo valor obtido. Embora seja trivial determinar a inexistência de densidade (imagens com 0% de diferença), a determinação de quais valores de diferença representam cada densidade (baixa, média ou alta) requer um processo mais elaborado. Esse assunto será discutido em detalhes no capítulo 7. Como exemplo, a tabela 6.1 ilustra um determinação simplificada dos intervalos.

Por fim, o conjunto de técnicas apresentadas resulta em uma constante determinação de densidade cuja eficiência será discutida no capítulo 7. Representando o componente principal do sistema, a análise de densidade figura também como

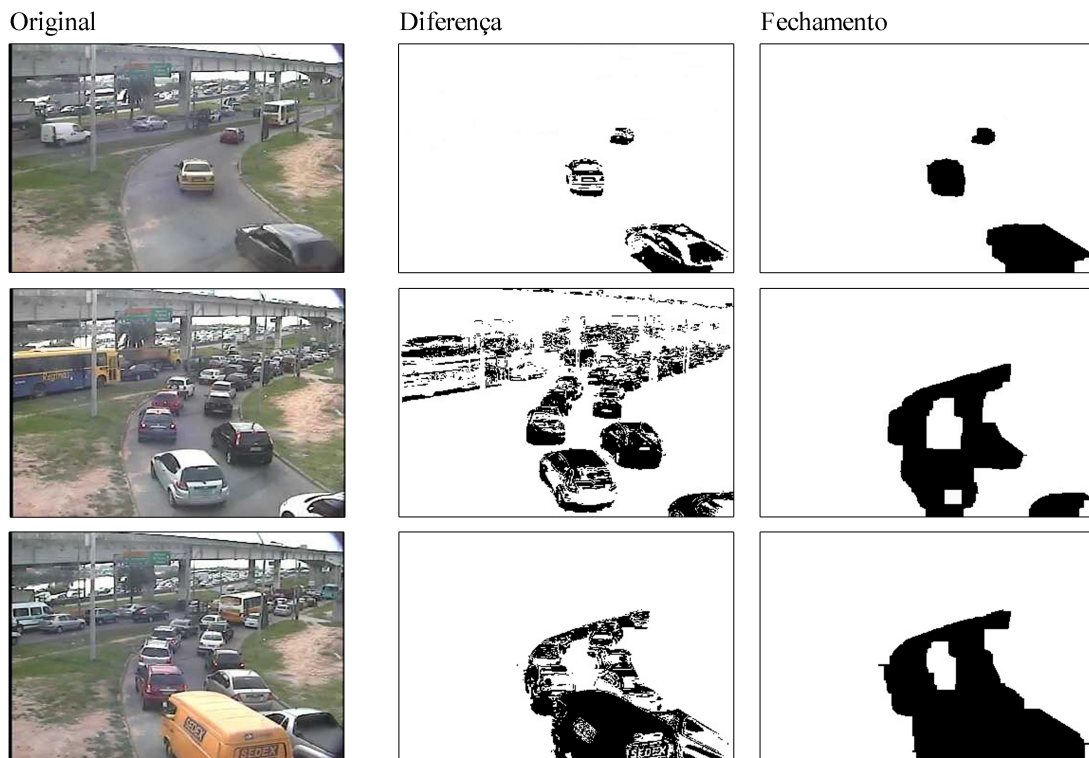


Figura 6.12: Utilização do fechamento

Densidade	Diferença
Inexistente	0%
Baixa	1% a 32%
Média	33% a 65%
Alta	66% a 100%

Tabela 6.1: Exemplo de intervalos de densidade

ferramenta importante de auxílio a outros procedimentos, como a determinação de velocidade, a ser visto adiante. A utilização da imagem de pista vazia, em especial, se mostra de muita utilidade não só para a densidade.

### 6.2.5 Determinação de velocidade

Como não há uma forma direta de mensuração da velocidade, carro a carro, através das imagens obtidas a cada segundo, a forma de determiná-la leva em consideração as relações observadas na figura 2.1. Baseando-se na detecção de movimento e na análise da densidade, é possível se obter valores aproximados da situação na via.

Quando a diferença de imagens consecutivas demonstra ausência de movimento, é correto assumir que a velocidade é zero, seja por falta de veículos na via, seja por causa de congestionamento, e, portanto, o sistema assim a determina. Já nos casos em que há movimento, a presença de densidade baixa influi em uma velocidade



alta, pois, mesmo que os carros não estejam em alta velocidade, eles se encontram próximos do ponto de velocidade de fluxo livre (seção 2.4). Uma densidade alta, por sua vez, influi em uma velocidade baixa, como visto na figura 2.1.

Embora simples, a determinação da velocidade oferece valores aproximados que, em conjunto com a determinação da densidade, permitem uma análise bem sucedida do comportamento do fluxo em uma via. Por exemplo, situações em que a densidade em uma via é alta por um intervalo de tempo prolongado podem ser diferenciadas pelo simples número de vezes em que a situação de velocidade zero ocorre. Em uma situação de tráfego inevitavelmente intenso, o ideal é que os veículos trafeguem sem parar, mesmo que sua velocidade seja bem baixa em todo o percurso, e a determinação da velocidade é capaz de mostrar isso.

### 6.2.6 Determinação de densidade com o PBAS

Como mencionado na seção 3.4, métodos simples de segmentação, como o desenvolvido nesse trabalho, não são robustos. Embora funcionem bem em muitos casos, seu correto funcionamento pode ser prejudicado em determinadas situações. Por exemplo, em uma via congestionada cuja iluminação mude rapidamente, o método desenvolvido não será capaz de acompanhar a nova iluminação (por falta de uma imagem de pista vazia indicando a mudança). Por conta disso, o PBAS (seção 3.4.1) também foi implementado para ser utilizado na geração das imagens de diferença a serem usadas na determinação de densidade, em alternativa à comparação de imagens.

A figura 6.13 ilustra a semelhança entre a segmentação gerada pelo diferença de imagens (com aplicação de 9 iterações do fechamento) e a segmentação gerada pelo PBAS, ambas para a região de interesse previamente apresentada (figura 6.2). Além da robustez, outra vantagem do PBAS é que não há necessidade do usuário indicar a pista vazia inicial, devido à sua modelagem do plano de fundo, como vista na seção 3.4.1.

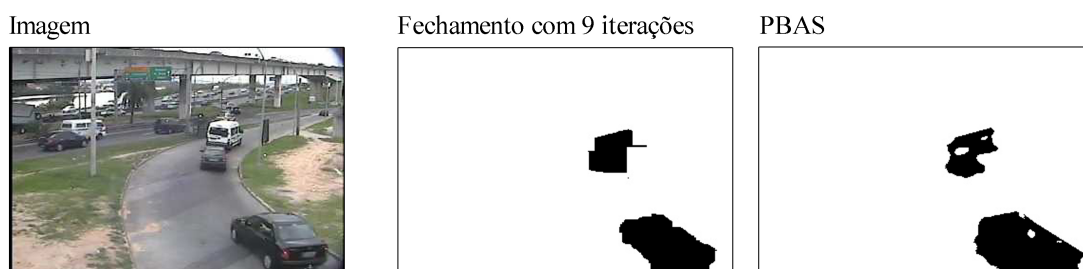


Figura 6.13: Comparação entre a diferença de imagens e o PBAS

Como desvantagem da utilização do PBAS no lugar da comparação de imagens

encontra-se seu tempo de execução elevado. Como pode ser visto na tabela 6.2, uma iteração da comparação de imagens (que envolve a comparação de imagens consecutivas, assim como da imagem sendo analisada com a de pista vazia, além de processamentos como recorte da área de interesse e aplicação do fechamento) é realizada muito mais rapidamente que uma iteração do PBAS (que envolve todo o processo descrito na seção 3.4.1 para cada *pixel* da imagem).

Método	Tempo execução (ms)
Comparação (9 it.)	28
PBAS	849

Tabela 6.2: Média do tempo de execução

### 6.2.7 Determinação de densidade com uma VGRAM

Da mesma forma que foi implementada uma alternativa para o método de geração das imagens de diferença, uma alternativa para a aferição da densidade, baseada em uma rede neural sem peso, também foi desenvolvida. Escolhida por preferência pessoal, a VGRAM (seção 5.3.4) previamente treinada com diversas imagens de diferença e suas respectivas densidades (inexistente, baixa, média ou alta) classifica, a cada segundo, a densidade presente na via a partir da imagem de diferença gerada.

Como peculiaridade da VGRAM implementada se destaca o armazenamento, em cada neurônio, da quantidade de vezes que determinado padrão de entrada foi treinado, assim como a classe correspondente, como visto na figura 6.14. Dessa forma, para um mesmo padrão de entrada, as diferentes classes e suas frequências observadas são registradas, sendo futuramente utilizadas nos testes como critério de escolha. Quanto maior o número de vezes que uma classe é vista no treinamento para determinado padrão de entrada, maior a probabilidade dela ser escolhida no teste.

Como desvantagem em relação aos outros métodos encontra-se a necessidade de classificação manual prévia de imagens da via, o que torna o processo inicial de calibragem dispendioso. Para cada nova câmera, sua respectiva VGRAM deve ser treinada com exemplos de imagens de diferença da via e suas densidades correspondentes.

### 6.2.8 Determinação de *Fluxo x Densidade*

Em alternativa aos métodos de determinação de densidade apresentados, um outro método de sensoriamento da via foi elaborado. Baseado na relação entre a taxa de fluxo e a densidade, citada na seção 2.4, seu objetivo consiste em identificar, no gráfico de *Fluxo x Densidade*, em que situação se encontra a via analisada, como

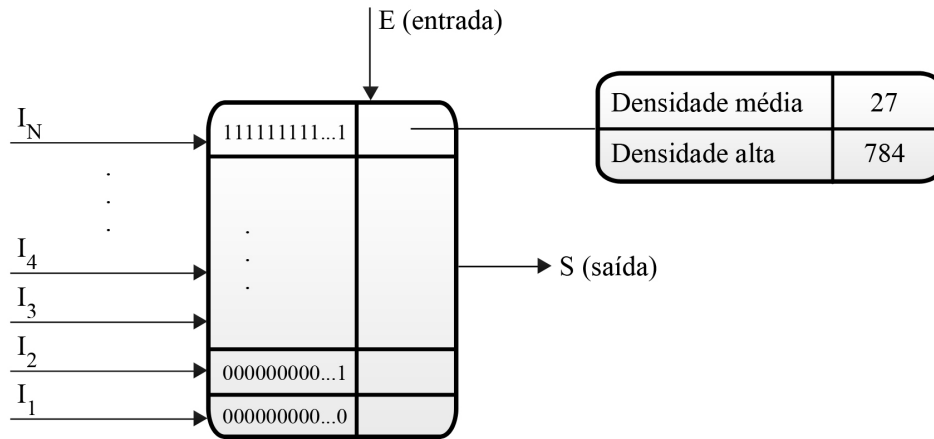


Figura 6.14: Detalhes da VGRAM

ilustrado na figura 6.15. Como pode ser visto, foram estabelecidos cinco intervalos de classificação: fluxo zero com densidade zero (V), fluxo estável (W), fluxo próximo da capacidade da via (X), fluxo forçado (Y) e fluxo zero com densidade máxima (Z).

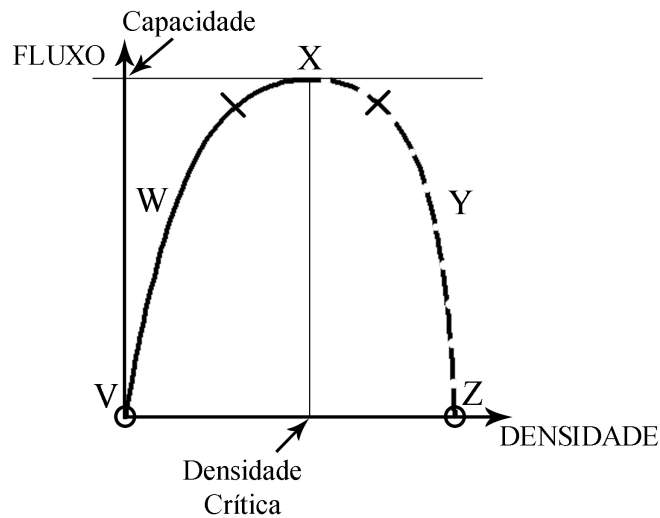


Figura 6.15: Intervalos de classificação

Ao contrário do método original, que determina a densidade a cada segundo, sua classificação se baseia em períodos de tempo, sendo cinco minutos o intervalo adotado. Como o sistema analisa uma imagem por segundo, isso representa 300 imagens sendo analisadas por período.

Para determinação do estado atual da via observada, são obtidas, para cada uma das imagens analisadas, as informações de existência ou inexistência de densidade (seção 6.2.4) e existência ou inexistência de movimento (seção 6.2.2). Como é impossível uma imagem apresentar ausência de densidade e presença de movimento,

as informações obtidas podem ser agrupadas em apenas três categorias: ausência de densidade com ausência de movimento (00), presença de densidade com ausência de movimento (X0), e presença de densidade com presença de movimento (X1). A frequência com que cada categoria é observada no período é então utilizada na geração do padrão de entrada, como representado na figura 6.16, em que cada *bit* 1 representa uma dezena (ou sua fração). Dessa forma, um valor entre 31 e 40, por exemplo, é representado por 11110...0. Por fim, a VGRAM com 18 nós de 5 entradas é alimentada com o padrão obtido, e o intervalo correspondente é determinado.

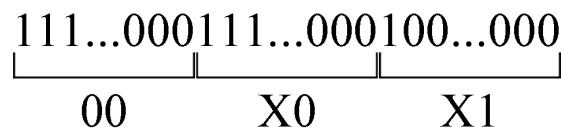


Figura 6.16: Padrão de entrada

### 6.2.9 Observações

Embora não completamente automático, como idealizado, o sistema ainda assim apresenta um grau de automação elevado, dependendo do usuário apenas para a delimitação da área de interesse e da indicação inicial de pista vazia (sendo desnecessária essa última quando se utilizando o PBAS). Após realizada a calibragem inicial, a análise se desenvolve por conta própria indefinidamente, sem necessidade de intervenção humana.

Devido a essa busca pela mínima interação humana, a solução foi desenvolvida de forma a requisitar que esse processo de calibragem inicial seja realizado apenas uma vez para cada nova região de interesse. Dados como diferença máxima observada, região de interesse e imagem de pista vazia são armazenados em disco, dispensando os processos necessários em futuras execuções. No caso da imagem de pista vazia, o sistema armazena uma imagem a cada fração do dia (de hora em hora, por exemplo) de forma a permitir que o usuário reinicie o sistema a qualquer hora do dia e ainda assim possa ser encontrada uma imagem de pista vazia apropriada para a imagem atual. As VGRAM são salvas em disco com o término da execução e carregada a cada nova execução, mantendo assim todo o seu treinamento adquirido.

Como limitação do sistema encontram-se as situações em que há presença de certos fenômenos meteorológicos adversos, como neblina. Embora o sistema não tenha sido testado para esses casos, a influência negativa das imagens geradas sob condições de pouca ou nenhuma visibilidade da via pode ser comparada à causada por imagens defeituosas.



Por fim, a eficiência do sistema, os resultados obtidos com os testes em uma situação real, podem ser vistos no capítulo 7. Detalhes como os diferentes métodos para determinação dos intervalos de densidade também serão apresentados.

# Capítulo 7

## Resultados experimentais

Neste capítulo, serão discutidos os resultados experimentais da solução desenvolvida. Na seção 7.1, será feita uma breve introdução aos testes, assim como uma apresentação do conjunto das fontes utilizadas; na seção 7.2, será explicada a metodologia utilizada para o teste da determinação de densidade, assim como seus resultados; na seção 7.3, será demonstrada a utilização do PBAS no lugar da comparação de imagens; na seção 7.4, serão apresentados os resultados da classificação de densidade com uma VGRAM; e na seção 7.5, serão apresentados a metodologia e os resultados para a determinação de *Fluxo x Densidade*.

### 7.1 Introdução

De forma a avaliar a eficiência da solução desenvolvida, foram identificados os componentes (e suas alternativas) passíveis de testes: a detecção de objetos, a determinação de densidade e a determinação de *Fluxo x Densidade*. Para a detecção de movimento, que consiste apenas em perceber se há movimento na via, não foi constatada necessidade de testes, visto que se trata de um resultado direto da detecção de objetos. Da mesma forma, a determinação de velocidade, por ser resultado direto da determinação de densidade e da detecção de movimento, não foi testada.

A detecção de objetos, por consistir em um método que determina se há ou não objetos na via, pode ser analisada junto com a determinação de densidade, visto que entre as classificações possíveis encontra-se a inexistência de densidade. Ou seja, testando-se a determinação de densidade, que identifica tanto a existência ou não de densidade, a eficiência da detecção de objetos pode ser analisada.

Com o intuito de possibilitar a realização dos testes, foram obtidas imagens da via próxima ao Hospital Universitário [33], nos entornos da Ilha do Fundão, ao longo de um dia útil. Esse intervalo garantiu que diferentes situações relevantes para fins de teste fossem contempladas, como tráfego na hora de pico, durante a noite, com transição da tarde para a noite, entre outras. Diversas situações atípicas também

são representadas pelas imagens utilizadas, como pessoas atravessando a via, carros trafegando no sentido contrário e funcionários públicos trabalhando na via. Nos testes de determinação de *Fluxo x Densidade*, as imagens obtidas ao longo de um dia útil na rampa de acesso à Linha Amarela [33] também foram utilizadas, de maneira a complementar os testes.

As imagens com o aviso de indisponibilidade do sistema foram retiradas. Consistindo apenas em uma figura indicando manutenção, essas imagens não possuíam utilidade para os testes. Após esse tratamento inicial, restaram, da fonte principal, 34.320 imagens disponíveis para o teste, e 35.381 da secundária (Linha Amarela).

A determinação de densidade utilizando a comparação de imagens foi exaustivamente testada, sendo suas metodologias e resultados encontrados na seção 7.2. Na seção 7.3, é demonstrado como a utilização do PBAS em alternativa à comparação de imagens gera resultados semelhantes. Os resultados da utilização da VGRAM para determinação da densidade podem ser vistos na seção 7.4, ao passo que os resultados da determinação de *Fluxo x Densidade* se encontram na seção 7.5.

## 7.2 Determinação de densidade

Para o teste independente da determinação de densidade, foram removidas as 26 imagens defeituosas da fonte principal, resultando em 34.294 imagens para teste. Dessa quantidade, 5.365 consistem em repetições de imagens, o que não representa problemas, visto que o sistema corretamente identifica imagens idênticas e que, embora a classificação de uma imagem repetida seja a mesma da imagem anterior, os testes contam apenas a classificação dada para a imagem original. Sendo assim, o número de classificações que de fato são computadas nos testes da densidade é de 28.928 (um a menos do esperado, visto que a primeira imagem testada, a de pista vazia inicial, não recebe classificação).

De forma a computar a taxa de acerto da classificação de densidade, de cada imagem analisada é extraída a quantidade de carros e essa quantidade é traduzida em uma das quatro densidades informadas pelo sistema. Ao mesmo tempo, o sistema calcula a porcentagem de diferença da imagem e converte esse valor para a densidade correspondente (o que constitui seu processamento normal). Logo em seguida, os dois valores de densidade obtidos são comparados, determinando o erro ou o acerto, como visto na figura 7.1.

Como a fonte utilizada consiste apenas nas imagens, sem maiores informações sobre cada uma, um longo processo de classificação manual foi realizado, classificando-as pela quantidade de carros presentes na região delimitada pela área de interesse ilustrada na figura 6.2. Os carros foram contados de acordo com o seguinte critério: veículos de passeio contavam como um veículo, ônibus e outros veículos de mesmo

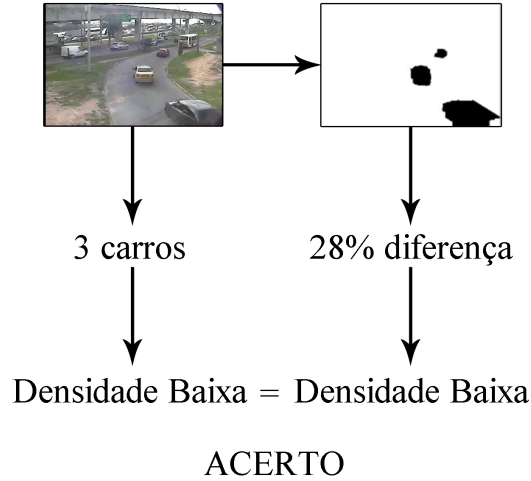


Figura 7.1: Exemplo de teste da densidade

porte, como dois (medidas comumente usadas por engenheiros de tráfego [34]), e motos solitárias como um, ao passo que motos acompanhadas de veículos, como zero. O motivo para a classificação diferenciada das motos é que uma moto no meio do tráfego congestionado não encontra dificuldade em seguir seu caminho, e, portanto, afeta o tráfego minimamente (salvo alguns casos excepcionais). Porém, uma via com uma única moto não pode ser considerada vazia pelo simples fato de haver um motorista atravessando a via e, portanto, requisitando um “semáforo verde”. Por exemplo, em um sistema que controle o semáforo em uma rua, um motoqueiro não pode ser obrigado a esperar um carro aparecer para então o sinal ficar verde.

Tendo em vista que a solução desenvolvida não lida com valores quantitativos, o mapeamento dessas contagens de carros para as classificações utilizadas no sistema precisou ser criado. Baseando-se nas médias das diferenças para cada quantidade de carros (tabela 7.1) e na percepção visual da disposição dos carros e seu preenchimento na via, o mapeamento mostrado na tabela 7.2 foi gerado. A figura 7.2 ilustra esse mapeamento com alguns exemplos de imagens para cada densidade.

O primeiro problema encontrado com o mapeamento refere-se à ocasional similaridade entre as fronteiras das densidades. Por exemplo, algumas imagens de quatro carros (densidade média) possuem preenchimento da via parecido com algumas imagens de três carros (densidade baixa), assim como algumas imagens de dez carros possuem preenchimento menor que algumas de nove. Por conta desse fator, um teste mais abrangente, menos rigoroso que o teste inicial, também foi realizado.

É importante notar também que imagens com pessoas atravessando a pista vazia são classificadas como contendo nenhum carro, o que, embora correto, já permite prever que a diferença gerada pela presença desses pedestres afetará os resultados do sistema (classificando como baixa densidade uma imagem sem carros, por exemplo). O fato do sistema não diferenciar pessoas de veículos consiste em uma limitação

Qtd.	Diferença
0	0%
1	5%
2	17%
3	16%
4	28%
5	32%
6	39%
7	43%
8	43%
9	46%
10	52%
11	60%
12	56%
13	60%
14	66%
15	60%

Tabela 7.1: Média das diferenças para cada quantidade de carros (fechamento com 7 iterações)

Densidade	Quantidade
Inexistente	0
Baixa	1 a 3
Média	4 a 9
Alta	10 em diante

Tabela 7.2: Mapeamento Densidade-Quantidade de carros

relevante, que merece foco em trabalhos futuros.

Outro fator disruptivo notado no mapeamento se refere às imagens com ônibus solitários. Como as imagens utilizadas são de uma curva, um ônibus apenas (contando como dois carros), ao fazer a curva, preenche quase toda a região de interesse, gerando uma diferença alta e induzindo a solução a classificá-la como alta densidade. A figura 7.3 ilustra ambas as situações problemáticas.

Como última etapa para tornar possível a realização dos testes, as faixas de diferenças para cada densidade precisaram ser definidas. Após ser calculada a diferença da imagem sendo analisada em relação à imagem de pista vazia mais recente, o sistema precisa converter essa porcentagem de diferença obtida em uma das quatro densidades possíveis. Porém, embora seja trivial determinar a ausência de densidade (diferença de 0%), são necessários métodos mais elaborados para a determinação dos intervalos das outras três densidades.

Ao longo do desenvolvimento da solução, três métodos de determinação do mapeamento Densidade-Diferença foram desenvolvidos. Os testes foram, portanto, realizados nos três a fim de determinar o mais eficiente. O primeiro dos métodos a

## Densidade

Baixa



Média



Alta



Figura 7.2: Exemplos de imagens para cada densidade

ser testado denomina-se *Privilegiado*, por fazer uso da informação da quantidade de carros em cada imagem. Embora esse método não possa ser utilizado em situações reais, apenas em testes, seus resultados podem ser usados como parâmetro de comparação para os outros métodos. O segundo método consiste no *Padrão*, em que os intervalos são definidos de forma uniforme (cada um corresponde a um terço do intervalo de valores de diferença observados) e o terceiro consiste no *Automático por Agrupamento*, em que as diferenças observadas ao longo de um período (por exemplo, um dia), são agrupadas em três grupos, de acordo com a distribuição das diferenças, definindo assim de forma automática os intervalos.

Por fim, como alternativa à classificação segundo a segundo (realizada a cada nova imagem), foram realizados testes da classificação média. A cada intervalo de 10 segundos, o sistema calcula a média das classificações que seriam geradas e compara com a média das classificações reais, utilizando para o cálculo os valores mostrados na tabela 7.3.



Figura 7.3: Situações que induzem o erro da classificação

Densidade	Valor
Inexistente	0
Baixa	1
Média	2
Alta	3

Tabela 7.3: Valores usados na classificação média

### 7.2.1 Método Privilegiado

O Método Privilegiado de determinação do mapeamento Diferença-Densidade, por se utilizar de informação não disponível ao sistema em tempo de execução (a quantidade de carros em cada imagem), tem seu uso limitado para fins de teste. Mesmo assim, seu valor é incontestável, visto que seus resultados servem de base comparativa para os outros métodos.

Embora imagens classificadas com determinada quantidade de carros possam exibir uma vasta gama de valores de diferença (por exemplo, um ônibus, que conta como dois carros, exibe valores de diferença altíssimos em determinadas situações), uma tendência de proporcionalidade na relação entre a quantidade de carros e a diferença gerada pode ser constatada, como já observado na tabela 7.1. Para os testes do Método Privilegiado, foram observadas todas as imagens (com exceção das repetições) e suas respectivas diferenças para cada quantidade distinta de iterações do filtro morfológico de fechamento, como visto na tabela 7.4 (o valor zero indica a não utilização do filtro).

A partir da análise dos valores de fronteira da tabela 7.4, foram determinados os intervalos para cada densidade, como mostrado na tabela 7.5. O método com que foram definidos os intervalos consiste em, dados o valor máximo observado para uma densidade e o valor mínimo da densidade seguinte (tabela 7.4), determinar o ponto médio entre ambos. Tomando como exemplos a densidade baixa para 7 iterações, que possui 17% como máximo (referente a dois carros) e a densidade média, que

Qtd.	Diferença (iterações)					
	0	3	5	7	9	11
0	0%	0%	0%	0%	0%	0%
1	3%	4%	4%	5%	5%	5%
2	11%	15%	16%	17%	18%	19%
3	9%	14%	15%	16%	18%	19%
4	17%	24%	26%	28%	30%	31%
5	19%	27%	30%	32%	35%	36%
6	23%	33%	36%	39%	42%	44%
7	24%	35%	39%	43%	47%	49%
8	24%	35%	39%	43%	47%	50%
9	24%	36%	41%	46%	50%	53%
10	27%	41%	47%	52%	57%	62%
11	33%	48%	55%	60%	66%	70%
12	29%	44%	50%	56%	62%	67%
13	32%	47%	54%	60%	66%	72%
14	36%	52%	59%	66%	72%	78%
15	28%	45%	53%	60%	68%	75%

Tabela 7.4: Média das diferenças para cada quantidade de carros

possui 28% como mínimo, tem-se 22% como valor equidistante (arredondado para baixo). Portanto, esse 22% é o último valor da densidade baixa e 23%, o primeiro da densidade média (tabela 7.5).

Densidade	Intervalo (iterações)					
	0	3	5	7	9	11
Zero	0%	0%	0%	0%	0%	0%
Baixa	1%-14%	1%-19%	1%-21%	1%-22%	1%-24%	1%-25%
Média	15%-25%	20%-38%	22%-44%	23%-49%	25%-53%	26%-57%
Alta	26%-100%	39%-100%	45%-100%	50%-100%	54%-100%	58%-100%

Tabela 7.5: Mapeamento Densidade-Diferença para o Método Privilegiado

Os resultados obtidos com os intervalos gerados pelo Método Privilegiado podem ser vistos na tabela 7.6. A taxa de acerto (*% acerto*) referencia a totalidade das classificações, enquanto a taxa de acerto com densidade (*% acerto dens.*) exclui os testes sem densidade, ou seja, leva em conta apenas os acertos e erros das classificações de densidade baixa, média e alta.

O detalhamento por densidade do teste com 11 iterações do fechamento pode ser visto na tabela 7.7. Indicando o número de vezes em que determinada densidade deveria ser a classificada (*Total*), a tabela também mostra a quantidade de classificações corretas, assim como a taxa de acerto. Nota-se que o nível de acerto para densidade média foi bem abaixo do restante, tendência que pôde ser constatada também no restante dos testes para o Método Privilegiado. As informações



<b>Iterações</b>	<b>% acerto</b>	<b>% acerto dens.</b>
0	88,06%	64,23%
3	88,82%	66,63%
5	89,09%	67,49%
7	89,19%	67,83%
9	89,43%	68,6%
11	89,66%	69,31%

Tabela 7.6: Resultados dos testes para o Método Privilegiado

detalhadas sobre os resultados de todos os testes podem ser encontradas no apêndice A.

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>% acerto</b>
Zero	19638	19840	98,98%
Baixa	4724	6631	71,24%
Média	952	1710	55,67%
Alta	623	747	83,4%

Tabela 7.7: Detalhes do teste para 11 iterações do fechamento (Privilegiado)

A taxa de acerto da classificação de ausência de densidade, como mostrada na tabela 7.7, manteve-se constante em todos os testes da classificação a cada segundo realizados. Visto que a densidade inexistente é identificada pela ausência da diferença, seu intervalo é sempre o mesmo (diferença de 0%) e, portanto, seus resultados independem tanto do processamento feito nas imagens de diferença como dos intervalos de diferença determinados. Na classificação média, como o acerto depende não apenas de uma imagem, mas sim do conjunto, é esperado que não haja essa constância.

Um segundo ponto a ser destacado sobre a classificação de ausência de densidade é sua eficiência, comprovada pela alta taxa de acerto. Como já mencionado, os casos raros de pessoas trafegando pela pista vazia induzem o sistema a classificar a densidade como baixa, quando na verdade a imagem fora classificada manualmente como contendo zero carro, o que explica os poucos casos de erro na classificação. Mas isso é uma peculiaridade, um erro oriundo da incapacidade do sistema de diferenciar pessoas de veículos. O que essa alta taxa de acerto comprova é a eficácia da detecção de objetos pela diferença de imagens, que também permite ao sistema acompanhar as mudanças de iluminação ao longo do dia.

Outra constatação que se pode fazer com os resultados obtidos é o ganho que se tem com a utilização do fechamento. Analisando-se o detalhamento por densidade do fechamento com diferentes iterações (tabela 7.8), nota-se que até um fechamento com poucas iterações incrementa a taxa de acerto para o método utilizado. Comparando-se a tabela os resultados para 0 iterações com 11 iterações, fica bem evidente essa

melhoria.

<b>Classificação</b>	<b>% acerto 0 it.</b>	<b>% acerto 3 it.</b>	<b>% acerto 11 it.</b>
Zero	98,98%	98,98%	98,98%
Baixa	70,5%	70,59%	71,24%
Média	37,13%	46,02%	55,67%
Alta	70,68%	78,71%	83,4%

Tabela 7.8: Detalhes dos testes para diferentes iterações do fechamento (Privilegiado)

Os resultados da classificação média, como vistos na tabela 7.9, mostram uma taxa de acerto um pouco melhor que a dos testes para classificação segundo a segundo. Como pode ser notado na tabela 7.10, a discrepância entre as taxas de acerto por densidade continuam presentes.

<b>Iterações</b>	<b>% acerto</b>	<b>% acerto dens.</b>
0	91,86%	61,37%
3	93,08%	68,5%
5	93,35%	69,88%
7	93,49%	70,57%
9	93,64%	70,11%
11	93,75%	71,03%

Tabela 7.9: Resultados dos testes para o Método Privilegiado (Média)

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>% acerto</b>
Zero	2906	2994	97,06%
Baixa	158	198	79,79%
Média	110	169	65,08%
Alta	41	68	60,29%

Tabela 7.10: Detalhes do teste para 11 iterações do fechamento (Privilegiado, Média)

Devido à natureza sobreposta dos valores de diferença observados para as faixas de densidade, que resulta em fronteiras fracamente delimitadas (como ilustrado pela figura 7.4, em que uma imagem de quatro carros possui valor de diferença menor que uma imagem de três carros), um teste da classificação a cada segundo mais abrangente, com tolerância, foi realizado. Como não há uma distinção bem definida entre os limites de cada faixa de densidade, um valor de diferença classificado erradamente como determinada densidade, mas que esteja dentro do “intervalo de transição” entre a densidade aferida e a sua densidade real, deveria ser considerado como classificado corretamente. Tomando como exemplo a figura 7.4, a classificação da imagem de quatro carros (densidade média) como sendo de densidade baixa deveria ser considerada um acerto nos casos em que sua diferença se encontre no intervalo de transição entre as densidades baixa e média.

## Densidade



Figura 7.4: Exemplos da sobreposição entre fronteiras (7 iterações)

Para a realização inicial desse teste, amplos intervalos de transição baseados no desvio padrão das diferenças observadas para cada quantidade de carros (tabela 7.11) foram estabelecidos, como ilustrado na figura 7.5. Por ter obtido os melhores resultados nos testes anteriores, as diferenças do fechamento com 11 iterações foram as utilizadas. A definição dos limites inferior e superior de cada intervalo de transição levou em conta as maiores variações de diferença encontradas em cada faixa de densidade. Por exemplo, para estabelecer o limite inferior do intervalo entre as densidades baixa e média, o limite inferior do intervalo das imagens de quatro carros (que possuem média de 31 e desvio padrão de 18) foi utilizado, assim como para o limite superior foi utilizada a maior variação encontrada na faixa de densidade baixa (que, no caso, é obtida pelos valores das imagens de dois carros).

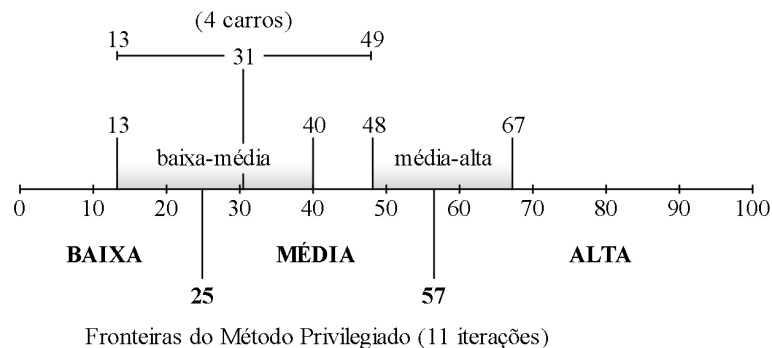


Figura 7.5: Intervalos de transição do primeiro teste

Os resultados desse primeiro teste com os intervalos de transição podem ser vistos na tabela 7.12. Como esperado, houve um aumento considerável na taxa de acerto da

Qtd.	Média	Variância	Desvio padrão
0	0%	0	0
1	5%	45,04	6,71
2	19%	449,13	21,19
3	19%	187,68	13,7
4	31%	340,14	18,44
5	36%	283,79	16,84
6	44%	225,37	15,01
7	49%	173,51	13,17
8	50%	216,46	14,71
9	53%	215,6	14,68
10	62%	209,62	14,48
11	70%	195,40	13,98
12	67%	108,33	10,41
13	72%	62,84	7,92
14	78%	38,54	6,21
15	75%	19,16	4,37

Tabela 7.11: Desvio padrão das diferenças para cada quantidade de carros (11 iterações)

solução, resultado direto do tamanho exagerado dos intervalos de transição. Mesmo não sendo ideais, esses intervalos servem como ponto de partida para a busca por intervalos menores que otimizem a solução. Como ilustrado na figura 7.6, os testes seguintes foram realizados com intervalos cada vez menores e, finalmente, para o quarto teste, em virtude da menor taxa de acerto para a classificação de densidade média, os limites dessa faixa de densidade foram levemente aumentados para se averiguar seus efeitos nos resultados, que podem ser vistos nas tabelas 7.13 e 7.14.

Classificação	Acertos	Total	% acerto
Zero	19638	19840	98,98%
Baixa	6081	6631	91,7%
Média	1497	1710	87,54%
Alta	688	747	92,1%

Tabela 7.12: Detalhes do primeiro teste com intervalos de transição (Privilegiado)

Teste	% acerto	% acerto dens.
1	96,46%	90,95%
2	94,64%	85,17%
3	93,89%	82,8%
4	94,45%	84,56%

Tabela 7.13: Resultados dos testes com intervalos de transição (Privilegiado)

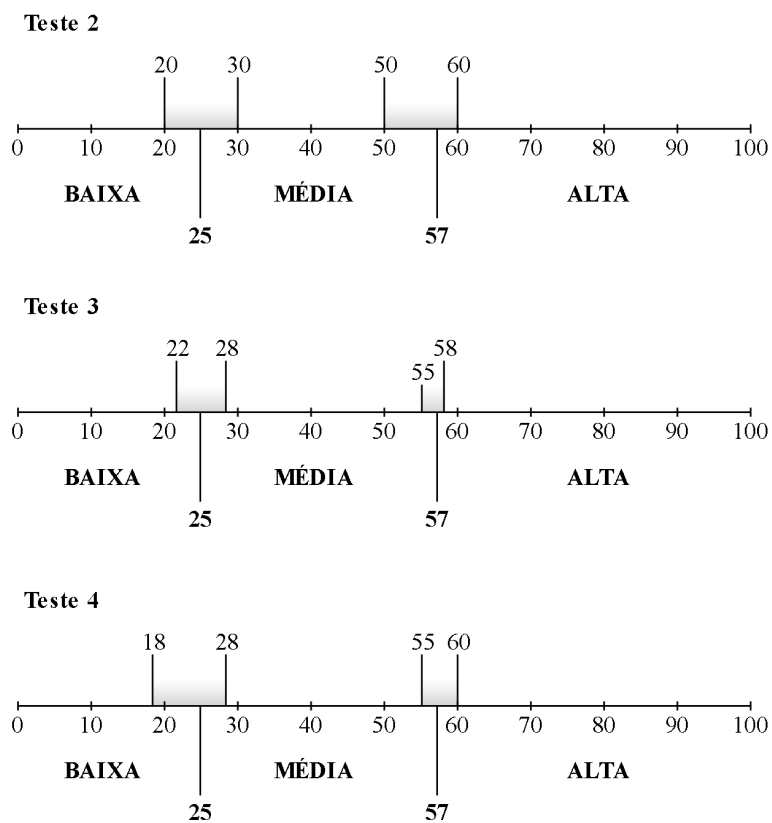


Figura 7.6: Intervalos de transição utilizados

<b>Classificação</b>	<b>% acerto T2</b>	<b>% acerto T3</b>	<b>% acerto T4</b>
Zero	98,98%	98,98%	98,98%
Baixa	88,41%	87,46%	87,46%
Média	69,82%	62,69%	72,04%
Alta	91,56%	87,41%	87,41%

Tabela 7.14: Detalhes dos testes com intervalos de transição (Privilegiado)

Como pode ser notado, mesmo com intervalos de transição reduzidos, a solução ainda obtém boas taxas de acerto. Os intervalos utilizados no último teste, por apresentarem tamanhos razoáveis e boa taxa de acerto, foram tomados como base para os testes com intervalos de transição dos outros métodos.

De forma geral, com os testes do Método Privilegiado realizados, tem-se, em teoria, os melhores resultados obteníveis pela solução, visto que fora utilizada informação privilegiada. Portanto, a eficácia dos outros métodos poderá ser comparada de forma eficiente.

## 7.2.2 Método Padrão

O segundo método desenvolvido, o Método Padrão, consiste em dividir o intervalo de valores de diferença possíveis em três partes iguais, como mostrado na tabela 7.15. Como já mencionado, a determinação da densidade é baseada na razão da área preenchida sobre a área total da região de interesse, sendo esse preenchimento obtido pela diferença da imagem sendo analisada com a de pista vazia mais recente. Mas como não há certeza se a máxima quantidade de carros na via é capaz de ocupar a região de interesse por completo, para que o Método Padrão possa ser utilizado, o sistema precisa manter um registro da maior porcentagem de ocupação observada até o momento. Dessa forma, assim que uma nova análise da diferença é realizada, o sistema compara esse valor com o valor máximo observado, determinando em qual intervalo ela se encontra e assim determinando sua densidade.

Densidade	Diferença
Inexistente	0%
Baixa	1%-32% * Max
Média	33%-65% * Max
Alta	66%-100% * Max

Tabela 7.15: Mapeamento Densidade-Diferença genérico para o Método Padrão

Todas as classificações realizadas antes da diferença máxima ser observada obviamente darão resultados não satisfatórios. Porém, ao final de um dia, provavelmente o valor máximo suportado pela via terá sido alcançado, não se alterando muito nos dias seguintes (levando em conta que esse processamento tenha sido feito em um dia típico de pico para o tráfego na via; por exemplo, um dia útil para vias de centros comerciais).

Os valores máximos de diferença observados nas imagens de teste para cada número de iterações do fechamento podem ser vistos na tabela 7.16. Os testes do Método Padrão foram realizados já com esses valores desde seu início, e o mapeamento correspondente pode ser visto na tabela 7.17.

<b>Iterações</b>	<b>Dif. máxima</b>
0	66%
3	79%
5	86%
7	91%
9	97%
11	99%

Tabela 7.16: Diferenças máximas observadas

<b>Densidade</b>	<b>Intervalo (iterações)</b>					
	<b>0</b>	<b>3</b>	<b>5</b>	<b>7</b>	<b>9</b>	<b>11</b>
Zero	0%	0%	0%	0%	0%	0%
Baixa	1%-20%	1%-25%	1%-27%	1%-29%	1%-31%	1%-31%
Média	21%-42%	26%-51%	28%-55%	30%-59%	32%-63%	32%-64%
Alta	43%-100%	52%-100%	56%-100%	60%-100%	64%-100%	65%-100%

Tabela 7.17: Mapeamento Densidade-Diferença para o Método Padrão

A tabela 7.18 apresenta os resultados dos testes para a classificação segundo a segundo, assim como a comparação com os resultados do Método Privilegiado. Como pode ser observado, os resultados são bem parecidos com os do Método Privilegiado, apresentando leves incrementos nas taxas de acerto. O detalhamento por densidade, como visto na tabela 7.19, mostra que houve um pequeno ganho no acerto das densidade baixa e média, mas uma perda considerável na densidade alta.

<b>Iterações</b>	<b>% acerto</b>	<b>% acerto dens.</b>	<b>% a. dens. Priv.</b>
0	88,13%	64,45%	64,23%
3	88,88%	66,84%	66,63%
5	89,33%	68,28%	67,49%
7	89,65%	69,3%	67,83%
9	89,95%	70,25%	68,6%
11	90,3%	71,36%	69,31%

Tabela 7.18: Resultados dos testes para o Método Padrão

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>% acerto</b>
Zero	19638	19840	98,98%
Baixa	4938	6631	74,46%
Média	1034	1710	60,46%
Alta	514	747	68,8%

Tabela 7.19: Detalhes do teste para 11 iterações do fechamento (Padrão)

Os resultados para a classificação média foram bem abaixo dos obtidos pelo Método Privilegiado, como pode ser constatado pelas tabelas 7.20 e 7.21. O acerto geral, embora alto, é mascarado pelo fato da taxa de acerto ser alta para densidade

zero, cuja quantidade de testes é muito maior. Analisando-se os acertos com densidade apenas, pode-se perceber o fraco desempenho do método usado, como pode ser confirmado pela tabela 7.21. Ao contrário dos resultados obtidos pela classificação a cada segundo, que indicam equivalência entre os métodos, os resultados para a classificação média demonstram a inferioridade do Método Padrão em comparação com o Método Privilegiado para essa modalidade.

Iterações	% acerto	% acerto dens.	% a. dens. Priv.
0	91,48%	45,51%	61,37%
3	92,65%	54,94%	68,5%
5	92,85%	57,24%	69,88%
7	92,97%	57,7%	70,57%
9	93,11%	58,85%	70,11%
11	93,49%	62,52%	71,03%

Tabela 7.20: Resultados dos testes para o Método Padrão (Média)

Classificação	Acertos	Total	% acerto
Zero	2934	2994	97,99%
Baixa	148	198	74,74%
Média	94	169	55,62%
Alta	30	68	44,11%

Tabela 7.21: Detalhes do teste para 11 iterações do fechamento (Padrão, Média)

Para o teste com intervalos de transição, foram utilizados os intervalos ilustrados na figura 7.7. Como as fronteiras entre as densidades do Método Privilegiado diferem das do Método Padrão, a conversão dos intervalos de transição foi feita de acordo com as distâncias relativas às fronteiras. Sendo assim, como no Método Privilegiado o limite inferior do primeiro intervalo difere de sete unidades da fronteira entre as densidades baixa e média, essa mesma distância é assegurada para o limite inferior desse mesmo intervalo para o Método Padrão, e o mesmo procedimento é aplicado ao restante dos limites.

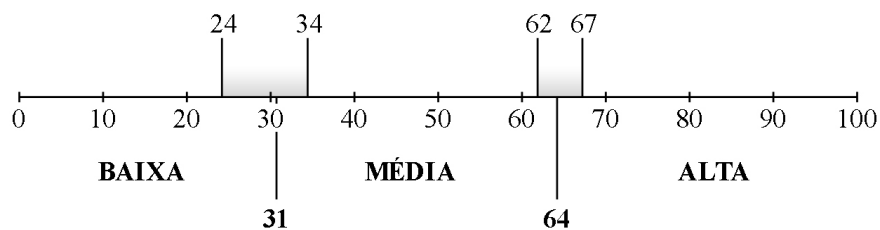


Figura 7.7: Intervalos de transição utilizados para o Método Padrão



Como pode ser observado nas tabelas 7.22 e 7.23, o Método Padrão obteve resultados melhores que o Privilegiado, com exceção da taxa de acerto para a densidade alta.

<b>Método</b>	<b>% acerto</b>	<b>% acerto dens.</b>
Padrão	94,97%	86,21%
Privilegiado	94,45%	84,56%

Tabela 7.22: Resultados dos testes com intervalos de transição (Padrão)

<b>Classificação</b>	<b>% acerto Padrão</b>	<b>% acerto Priv.</b>
Zero	98,98%	98,98%
Baixa	89,88%	87,46%
Média	75,67%	72,04%
Alta	77,77%	87,41%

Tabela 7.23: Detalhes dos testes com intervalos de transição (Padrão)

### 7.2.3 Método Automático por Agrupamento

O terceiro método de determinação do mapeamento Densidade-Diferença consiste em analisar todas as diferenças observadas ao longo de um período (como um dia) e separá-las em grupos, utilizando o método de agrupamento *k-means* (capítulo 4). Os intervalos da diferença são então definidos a partir desses grupos.

Como as observações manipuladas são de apenas uma dimensão e a quantidade de grupos é bem definida, os centroides iniciais são posicionados sobre a metade de cada terça parte das observações (no caso, 1/6, 3/6 e 5/6 do intervalo compreendendo o valor mínimo e o valor máximo observado). Em virtude do objetivo de definir os intervalos apenas para as densidades baixa, média e alta, todas as diferenças de 0% são descartadas do processo.

Após realizado o agrupamento, os intervalos são obtidos de forma análoga ao que é feito no Método Privilegiado: dados o valor máximo de um grupo e o mínimo do grupo seguinte, determina-se o ponto equidistante como fronteira entre ambos. Dessa forma, para o teste sendo realizado, obteve-se os intervalos de diferença mostrados na tabela 7.24. Ao ser definido o intervalo da densidade alta, todos os valores acima do valor máximo observado são incluídos, o que explica o valor máximo dessa densidade como 100%.

Como pode ser constatado pelos resultados mostrados na tabela 7.25, o Método Automático por Agrupamento teve um desempenho geral pior que os outros dois métodos para a classificação a cada segundo. Analisando os detalhes por densidade (tabela 7.26), nota-se que esse método, por conter uma faixa de densidade alta mais

Densidade	Intervalo (iterações)					
	0	3	5	7	9	11
Zero	0%	0%	0%	0%	0%	0%
Baixa	1%-12%	1%-16%	1%-17%	1%-18%	1%-19%	1%-21%
Média	13%-30%	17%-38%	18%-40%	19%-44%	20%-47%	22%-51%
Alta	31%-100%	39%-100%	41%-100%	45%-100%	48%-100%	52%-100%

Tabela 7.24: Mapeamento Densidade-Diferença para o Método Automático por Agrupamento

Iterações	% acerto	% acerto dens.	% a. dens. Priv.
0	87,84%	63,54%	64,23%
3	88,47%	65,52%	66,63%
5	88,36%	65,19%	67,49%
7	88,35%	65,16%	67,83%
9	88,31%	65,03%	68,6%
11	88,56%	65,83%	69,31%

Tabela 7.25: Resultados dos testes para o Método Automático por Agrupamento

extensa que os outros dois métodos, passou a cometer muito mais acertos quando classificando imagens dessa densidade. Em contrapartida, por esse mesmo motivo, houve uma quantidade maior de erros nas classificações das imagens das outras densidades, em uma espécie de remanejamento de acertos. Porém essa troca não se mostrou vantajosa, visto que o método obteve uma taxa de acerto geral mais baixa que os demais. Com isso, constata-se que, dos três métodos, o Método Padrão é o que provê resultados mais equilibrados entre as densidades, exibindo um desempenho geral equivalente ao do Método Privilegiado.

Classificação	Acertos	Total	% acerto
Zero	19638	19840	98,98%
Baixa	4468	6631	67,38%
Média	839	1710	49,06%
Alta	676	747	90,49%

Tabela 7.26: Detalhes do teste para 11 iterações do fechamento (Agrupamento)

Embora os resultados para a classificação segundo a segundo não tenham sido animadores, o mesmo não pode ser dito para a classificação média. Como pode ser visto na tabela 7.27, o Método Automático por Agrupamento obteve resultados melhores que o Método Privilegiado para acertos com densidade apenas, ao mesmo tempo em que obteve equilíbrio entre as taxas de acerto por densidade, como mostrado na tabela 7.28.

Para o teste com intervalos de transição, o mesmo procedimento utilizado para o Método Padrão, de conversão dos intervalos, foi aplicado para o Método Automático por Agrupamento, resultando nos intervalos ilustrados na figura 7.8. O desempenho

Iterações	% acerto	% acerto dens.	% a. dens. Priv.
0	92,33%	64,59%	61,37%
3	93,03%	70,34%	68,5%
5	93,05%	72,18%	69,88%
7	93,32%	73,79%	70,57%
9	93,14%	72,18%	70,11%
11	93,37%	72,64%	71,03%

Tabela 7.27: Resultados dos testes para o Método Automático por Agrupamento (Média)

Classificação	Acertos	Total	% acerto
Zero	2886	2994	96,39%
Baixa	155	198	78,28%
Média	107	169	65,31%
Alta	54	68	79,41%

Tabela 7.28: Detalhes do teste para 11 iterações do fechamento (Agrupamento, Média)

desse método foi bem inferior aos outros dois métodos, como pode ser visto nas tabelas 7.29 e 7.30. Novamente, a ênfase na faixa de densidade alta resulta em taxas de acerto altas para essa densidade em detrimento do acerto das outras densidades e do acerto geral.

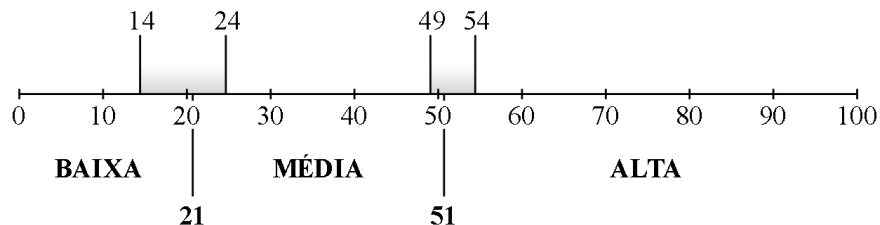


Figura 7.8: Intervalos de transição utilizados para o Método Automático por Agrupamento

Os resultados dos testes mostram então que, em situações reais, para a classificação a cada segundo, o Método Padrão é mais indicado, ao passo que, para a classificação média, o Método Automático por Agrupamento se mostra o mais adequado. Pelos testes com intervalos de transição, conclui-se que uma maneira de tornar a classificação mais confiável seria adicionar classificações intermediárias, como densidades “baixa-média” e “média-alta”, indicando que a densidade da via encontra-se em uma zona intermediária entre duas faixas de densidade. Assim se conclui a primeira etapa dos testes, que consistia em testar de forma independente a classificação de densidade.

Método	% acerto	% acerto dens.
Agrupamento	93,43%	81,32%
Padrão	94,97%	86,21%
Privilegiado	94,45%	84,56%

Tabela 7.29: Resultados dos testes com intervalos de transição (Agrupamento)

Classificação	% acerto Agrup.	% acerto Padrão	% acerto Priv.
Zero	98,98%	98,98%	98,98%
Baixa	85,11%	89,88%	87,46%
Média	61,98%	75,67%	72,04%
Alta	91,96%	77,77%	87,41%

Tabela 7.30: Detalhes dos testes com intervalos de transição (Agrupamento)

### 7.3 Determinação de densidade com o PBAS

De forma a mostrar a semelhança entre a segmentação do PBAS e as imagens de diferença geradas pela comparação de imagens, um teste simples, com o Método Padrão e sem intervalos de confiança, foi realizado. Como pode ser visto na tabela 7.31, os resultados de ambos os métodos são bem parecidos.

Método	% acerto	% acerto dens.
PBAS	88,08%	68,69%
5 iterações	89,33%	68,28%
11 iterações	90,3%	71,36%

Tabela 7.31: Resultados dos testes do PBAS para o Método Padrão

### 7.4 Determinação de Densidade com uma VGRAM

Para o teste da VGRAM de densidade, foi realizada validação cruzada com todas as imagens de diferença geradas a partir das 28.928 imagens da fonte principal. A contagem realizada previamente, determinando a quantidade de carros em cada uma das imagens, foi usada para determinar a classe de cada imagem de diferença, da mesma forma como visto na tabela 7.2. Embora esse mapeamento tenha sido utilizado para o treinamento, durante o teste foram utilizados intervalos de transição, como vistos nos testes anteriores, devido à natureza sobreposta das fronteiras entre as densidades. Essa pequena expansão nos valores de fronteira, gerando os intervalos de transição, resultou no mapeamento sobreposto que pode ser visto na tabela 7.32.

O método de validação cruzada utilizado foi o *k-fold* com cinco partições. Dessa forma, o conjunto de padrões de entrada (cadeias de bits representando as imagens de

Densidade	Quantidade
Inexistente	0
Baixa	1 a 5
Média	3 a 10
Alta	8 em diante

Tabela 7.32: Mapeamento Densidade-Quantidade de carros para os testes

diferença, que são binárias) foi separado em cinco subconjuntos de mesmo tamanho e, em cada um dos cinco testes, um subconjunto diferente foi selecionado como conjunto de testes enquanto todo o restante foi utilizado para o treinamento da VGRAM.

Para geração das imagens a serem usadas no treinamento, foram utilizados tanto a comparação de imagens com 9 iterações do fechamento, como o PBAS, para fins de comparação. Foram analisados também os desempenhos para diferentes tamanhos de neurônio (ou seja, o número de *bits* de entrada em cada um). Como pode ser visto nas tabelas 7.33 e 7.34, a VGRAM obtém melhores resultados com neurônios de 32 *bits* e há um pequeno ganho com a utilização das imagens de diferença com fechamento.

Tamanho	% acerto	% acerto dens.
4	71,83%	59,26%
16	78,54%	63,85%
32	84,64%	80,56%
64	84,55%	80,27%

Tabela 7.33: Resultados dos testes da VGRAM utilizando o PBAS

Tamanho	% acerto	% acerto dens.
4	71,57%	60,13%
16	78,85%	71,56%
32	84,19%	84,17%
64	84,09%	83,94%

Tabela 7.34: Resultados dos testes da VGRAM utilizando comparação de imagens com 9 iterações do fechamento

## 7.5 Determinação de *Fluxo x Densidade*

Como visto na seção 2.3.1, a taxa de fluxo consiste no número de veículos que passam por determinado ponto, faixa ou direção de uma rodovia durante um intervalo de tempo especificado. Portanto, o passo inicial para realização dos testes consistiu na classificação manual das taxas de fluxo encontradas nas imagens utilizadas. Como

o período de observação adotado é de cinco minutos, as fontes utilizadas (imagens da via próxima ao Hospital Universitário e da rampa de acesso à Linha Amarela) foram separadas em trechos de 300 imagens e um longo processo de contagem de carros que ultrapassam os limites da região de interesse foi realizado. Além disso, foram extraídas as frequências de cada categoria (00, X0 e X1) para cada trecho.

Durante essa análise, foi constatado que, para ambas as fontes, taxas de fluxo maiores que 100 carros/período (1200 carros/hora) ocasionavam situações próximas de fluxo ininterrupto de veículos, com eventuais paradas ou pistas vazias. Por conta disso, os trechos observados com taxa de fluxo maior ou igual a esse valor foram considerados como pertencendo ao intervalo X (figura 6.15). Para taxas de fluxo menores que 100 carros/período, a diferenciação entre W e Y foi feita com base na análise das frequências das categorias. Ou seja, trechos com quantidade variável de imagens de movimento (X1) e pista vazia (00) foram considerados do intervalo W e trechos com imagens predominantemente de presença de densidade e ausência de movimento (X0) foram considerados do intervalo Y. Os poucos trechos contendo 300 instâncias de pista vazia foram consideradas do intervalo V e, por falta de amostra (como pode ser visto na tabela 7.35), um padrão contendo 300 instâncias de X0 foi gerado, para ser utilizado como única amostra do intervalo Z.

<b>Intervalo</b>	<b>Hospital</b>	<b>Linha Amarela</b>
V	0	3
W	85	84
X	6	1
Y	5	8
Z	0	0

Tabela 7.35: Distribuição dos padrões para cada intervalo

Foi realizada então uma validação cruzada com todos os 193 padrões de entrada e suas respectivas classificações, sendo o método utilizado o *k-fold* com cinco partições. Ao final dos testes, uma taxa de acerto média de 92,24% com desvio padrão de 2,78 foi obtida.

Analisando-se os erros de classificação ocorridos, nota-se alguns casos em que trechos do intervalo W com baixíssimo fluxo (como, por exemplo, 7 carros/período, ou 84 carros/hora) foram classificados como V, devido ao alto número de imagens de pista vazia, e alguns casos em que trechos desse mesmo intervalo, mas com uma alta frequência de movimento e taxa de fluxo próxima de 90, foram classificados como X, entre outros. Isso mostra que os erros de classificação cometidos pela rede neural ocorrem basicamente para padrões que se situam nas fronteiras entre intervalos.

De forma a testar a universalidade das distribuições que formam o padrão de entrada utilizado (e, por consequência, a aplicabilidade do método desenvolvido em qualquer via observável), foi realizado um experimento com os trechos da fonte do

Hospital Universitário como padrões de treinamento e os da Linha Amarela, como padrões de teste. A execução do teste 100 vezes resultou em um acerto médio de 92,81% com desvio padrão de 3,8. O experimento inverso também foi realizado, e, após 100 iterações, foi obtida uma taxa de acerto de 87,2% com desvio padrão de 4,43.

# Capítulo 8

## Conclusão

Neste trabalho, foi desenvolvido o Classificador de Densidade, uma solução para determinação de densidade e fluxo em uma via a partir de imagens geradas por câmeras estáticas. Baseado na diferença de imagens, o sistema é capaz de determinar o comportamento do fluxo em uma via a cada segundo, servindo de ferramenta de auxílio a engenheiros de tráfego e sistemas de transporte inteligente. Entre as funcionalidades desenvolvidas, encontra-se a capacidade de se adaptar continuamente, dispensando ao máximo a interação humana.

Experimentos foram realizados em imagens reais de trânsito com diversas configurações possíveis. Embora incapaz de diferenciar pessoas de veículos, o sistema apresentou taxas de acerto razoáveis nos testes.

Como possíveis aplicações da solução desenvolvida encontram-se a obtenção de dados de uma via em determinados períodos de tempo e a alimentação, segundo a segundo, de informações sobre o fluxo da via para um sistema inteligente de controle de tráfego. O Classificador de Densidade pode ser usado para se obter diversas informações sobre uma via, como horários de pico, temporadas de aumento no fluxo e intervalos do dia em que o fluxo é inexistente, informações úteis para os responsáveis em planejar e projetar melhorias no trânsito. Um sistema complexo, que controle semáforos em tempo real, baseado nas condições do trânsito, pode ser alimentado constantemente pela solução.

Até mesmo um sistema simples de controle de semáforos, responsável pela sinalização de um cruzamento, por exemplo, poderia se beneficiar desse sistema. Durante a madrugada, em que a quantidade de carros nas ruas é bem pequena, motoristas constantemente atravessam sinais vermelhos, seja por motivos de segurança (medo de assaltos e sequestros), seja por outros motivos. Com esse sistema, um carro se aproximando de um semáforo seria prontamente detectado e, caso nenhum carro fosse detectado na via transversal, teria a passagem liberada pelo sinal verde, sem precisar parar. Da mesma forma, um carro se aproximando de um semáforo e percebendo que o sinal vermelho não muda, teria a certeza da presença de um carro



prestes a cruzar a pista.

## 8.1 Trabalhos futuros

Além dos sistemas inteligentes mencionados, que por si só representam trabalhos futuros desejáveis, outros trabalhos derivados também são almejados. Por exemplo, como notado pelos testes, umas das limitações do sistema é sua incapacidade de diferenciar carros de pessoas trafegando na pista. Para o correto funcionamento do sistema em uma via urbana, com intenso tráfego humano, a região de interesse não deve englobar a faixa de pedestres e, principalmente, o uso dessa faixa deve ser incentivado a ponto dos pedestres evitarem atravessar fora dela (o que muitas vezes se torna inviável). Com isso em vista, um aprimoramento interessante para o sistema seria um identificador de objetos, utilizando uma rede neural, para diferenciar pessoas de veículos.

A técnica para se determinar automaticamente a região de interesse (seção 6.2.3), embora não tenha apresentado muita utilidade para o trabalho, possui aplicações em outros problemas. Como trabalho futuro interessante, inclui-se a determinação de frequência de ocupação, através da diferença de imagens, em diversos ambientes diferentes, como campos de futebol (para inferir regiões mais danificadas do gramado ou análise tática de uma partida), estacionamentos (para determinar as vagas mais utilizadas ou a frequência com que vagas para deficientes são utilizadas), ou até mesmo praças de alimentação.

# Referências Bibliográficas

- [1] EZELL, S. “Intelligent Transport Systems”. Disponível em [http://www.itif.org/files/2010-1-27-ITS\\_Leadership.pdf](http://www.itif.org/files/2010-1-27-ITS_Leadership.pdf), Janeiro 2010.
- [2] ATKINSON, R., CASTRO, D. “Digital Quality of Life: Understanding the Personal and Social Benefits of the Information Technology Revolution”. Disponível em <http://ssrn.com/abstract=1278185>, Outubro 2008.
- [3] ROESS, R. P., PRASSAS, E. S., MCSHANE, W. R. *Traffic Engineering*. 3 ed. New Jersey, Person Prentice Hall, 2004.
- [4] *Traffic Detector Handbook*, 2 ed. Institute of Transportation Engineers, Washington D.C., 1997.
- [5] “SWARCO Traffic Detectors”. Disponível em <http://www.swarco.com/en/Products-Services/Traffic-Management/Urban-Traffic-Management/Traffic-Detectors>, Julho 2012.
- [6] “SWARCO TDC4 Traffic Detectors”. Disponível em <http://www.swarco.com/en/Products-Services/Traffic-Management/Urban-Traffic-Management/Traffic-Detectors/TDC4-Traffic-detectors>, Julho 2012.
- [7] “AutoScope”. Disponível em <http://autoscope.com>, Julho 2012.
- [8] “AutoScope Duo”. Disponível em [http://autoscope.com/products/duo\\_tech.htm](http://autoscope.com/products/duo_tech.htm), Julho 2012.
- [9] “FLIR Thermal Vision”. Disponível em <http://www.flir.com/cvs/americas/en/security/view/?id=44284>, Julho 2012.
- [10] ATKOCIUNAS, E., BLAKE, R., JUOZAPAVICIUS, M., et al. “Image Processing in Road Traffic Analysis”, *Nonlinear Analysis: Modelling and Control*, v. 10, n. 4, pp. 315–332, Dezembro 2005.

- [11] WARDROP, J. G. “Some Theoretical Aspects of Road Traffic Research”, *Proceedings of the Institute of Civil Engineers*, v. 1, n. 36, pp. 325–362, Janeiro 1952.
- [12] JÄHNE, B. *Digital Image Processing*. Springer, 2005.
- [13] Intel® *Integrated Performance Primitives for Intel® Architecture - Volume 2: Image and Video Processing*. Intel, 2011. Disponível em: <[http://software.intel.com/sites/products/documentation/hpc/composerxe/en-us/2011Update/ippxe/ipp\\_manual\\_lnx/index.htm](http://software.intel.com/sites/products/documentation/hpc/composerxe/en-us/2011Update/ippxe/ipp_manual_lnx/index.htm)>.
- [14] HUNT, R. W. G. *The Reproduction of Colour*. 6 ed. Chichester, UK, John Wiley & Sons, Lt, 2004.
- [15] *Using ADOBE PHOTOSHOP CS5*. Adobe, Dezembro 2011. Disponível em: <[http://help.adobe.com/en\\_US/photoshop/cs/using/photoshop\\_cs5\\_help.pdf](http://help.adobe.com/en_US/photoshop/cs/using/photoshop_cs5_help.pdf)>.
- [16] LAGANIÈRE, R. *OpenCV 2 Computer Vision Application Programming Cookbook*. Packt Publishing, 2011.
- [17] SMITH, S. W. *The Scientist and Engineer’s Guide to Digital Signal Processing*. California Technical Publishing, 2003.
- [18] SEZGIN, M., SANKUR, B. “Survey over image thresholding techniques and quantitative performance evaluation”, *Journal of Electronic Imaging*, v. 13, n. 1, pp. 146–165, Janeiro 2004.
- [19] SHAPIRO, L. G., STOCKMAN, G. C. *Computer Vision*. Prentice-Hall, Fevereiro .
- [20] HOFMANN, M., TIEFENBACHER, P., RIGOLL, G. “Background Segmentation with Feedback: The Pixel-Based Adaptive Segmenter”, *Proceedings of IEEE Workshop on Change Detection*, 2012.
- [21] JODOIN, P. M., PORIKLI, F., KONRAD, J., et al. “Change Detection Challenge”. Disponível em <http://www.changedetection.net>, 2012.
- [22] MACKAY, D. *Information Theory, Inference and Learning Algorithms*. University of Cambridge, Setembro 2003.
- [23] FABER, V. “Clustering and the Continuous k-Means Algorithm”, *Los Alamos Science*, v. 22, pp. 138–144, 1994.

- [24] VATTANI, A. “K-means requires exponentially many iterations even in the plane”, *Discrete and Computational Geometry*, v. 45, n. 4, pp. 596–616, 2011.
- [25] ALEKSANDER, I., MORTON, H. *An Introduction to Neural Computing*. 1 ed. London, Chapman & Hall, 1991.
- [26] HAMMING, R. W. “Error detecting and error correcting codes”, *Bell System Technical Journal*, v. 29, pp. 147–160, 1950.
- [27] ALEKSANDER, I., DE GREGORIO, M., FRANÇA, F., et al. “A brief introduction to weightless neural systems”, *Advances in Computational Intelligence and Learning, Proceedings of ESANN 2009*, pp. 299–305, 2009.
- [28] LUDERMIR, T., CARVALHO, A., BRAGA, A., et al. “Weightless Neural Models: a review of current and past works”, *Neural Computing Surveys*, v. 2, pp. 41–61, 1998.
- [29] ALEXANDER, I. “Ideal neurons for neural computers”. In: *Parallel Processing in Neural Systems and Computers*, pp. 225–228, Elsevier Science, 1989.
- [30] “Microsoft Visual Studio 2010”. Disponível em <http://www.microsoft.com/visualstudio/pt-br/home-produtos>, Julho 2012.
- [31] “Open Source Computer Vision (OpenCV)”. Disponível em <http://opencv.org/>, Julho 2012.
- [32] “Microsoft Foundation Class Library (MFC)”. Disponível em <http://msdn.microsoft.com/en-us/library/583ya1kc%28v=vs.100%29.aspx>, Outubro 2010.
- [33] “CET Fundação”. Disponível em <http://diseg.pu.ufrj.br/>, Novembro 2010.
- [34] *Manual de Estudos de Tráfego*. Departamento Nacional de Infra-Estrutura de Transportes, 2006. Disponível em: <[http://ipr.dnit.gov.br/manuais/manual\\_estudos\\_trafego.pdf](http://ipr.dnit.gov.br/manuais/manual_estudos_trafego.pdf)>.

# Apêndice A

## Resultados

### A.1 Método Privilegiado

#### Iterações do Fechamento: 0

Total de classificações: 28928

Total de acertos: 25476

Taxa de acerto: 88,06%

Total de acertos com densidade apenas: 5838

Taxa de acerto com densidade apenas: 64,23%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4675	6631	70,5%
Média	635	1710	37,13%
Alta	528	747	70,68%

#### Iterações do Fechamento: 3

Total de classificações: 28928

Total de acertos: 25694

Taxa de acerto: 88,82%

Total de acertos com densidade apenas: 6056

Taxa de acerto com densidade apenas: 66,63%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4681	6631	70,59%
Média	787	1710	46,02%
Alta	588	747	78,71%

**Iterações do Fechamento: 5**

Total de classificações: 28928

Total de acertos: 25772

Taxa de acerto: 89,09%

Total de acertos com densidade apenas: 6134

Taxa de acerto com densidade apenas: 67,49%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4695	6631	70,8%
Média	842	1710	49,23%
Alta	597	747	79,91%

**Iterações do Fechamento: 7**

Total de classificações: 28928

Total de acertos: 25803

Taxa de acerto: 89,19%

Total de acertos com densidade apenas: 6165

Taxa de acerto com densidade apenas: 67,83%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4677	6631	70,53%
Média	890	1710	52,04%
Alta	598	747	80,05%

**Iterações do Fechamento: 9**

Total de classificações: 28928

Total de acertos: 25873

Taxa de acerto: 89,43%

Total de acertos com densidade apenas: 6235

Taxa de acerto com densidade apenas: 68,6%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4725	6631	71,25%
Média	895	1710	52,33%
Alta	615	747	82,32%

**Iterações do Fechamento: 11**

Total de classificações: 28928

Total de acertos: 25937

Taxa de acerto: 89,66%

Total de acertos com densidade apenas: 6299

Taxa de acerto com densidade apenas: 69,31%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4724	6631	71,24%
Média	952	1710	55,67%
Alta	623	747	83,4%

**A.2 Método Padrão****Iterações do Fechamento: 0**

Total de classificações: 28928

Total de acertos: 25496

Taxa de acerto: 88,13%

Total de acertos com densidade apenas: 5858

Taxa de acerto com densidade apenas: 64,45%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	5027	6631	75,81%
Média	742	1710	43,39%
Alta	89	747	11,91%

**Iterações do Fechamento: 3**

Total de classificações: 28928

Total de acertos: 25713

Taxa de acerto: 88,88%

Total de acertos com densidade apenas: 6075

Taxa de acerto com densidade apenas: 66,84%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4972	6631	74,98%
Média	916	1710	53,56%
Alta	187	747	25,03%

**Iterações do Fechamento: 5**

Total de classificações: 28928

Total de acertos: 25844

Taxa de acerto: 89,33%

Total de acertos com densidade apenas: 6206

Taxa de acerto com densidade apenas: 68,28%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4960	6631	74,8%
Média	984	1710	57,54%
Alta	262	747	35,07%

**Iterações do Fechamento: 7**

Total de classificações: 28928

Total de acertos: 25936

Taxa de acerto: 89,65%

Total de acertos com densidade apenas: 6298

Taxa de acerto com densidade apenas: 69,3%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4962	6631	74,83%
Média	1002	1710	58,59%
Alta	334	747	44,71%

**Iterações do Fechamento: 9**

Total de classificações: 28928

Total de acertos: 26023

Taxa de acerto: 89,95%

Total de acertos com densidade apenas: 6385

Taxa de acerto com densidade apenas: 70,25%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4969	6631	74,93%
Média	1011	1710	59,12%
Alta	405	747	54,21%



**Iterações do Fechamento: 11**

Total de classificações: 28928

Total de acertos: 26124

Taxa de acerto: 90,3%

Total de acertos com densidade apenas: 6486

Taxa de acerto com densidade apenas: 71,36%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4938	6631	74,46%
Média	1034	1710	60,46%
Alta	514	747	68,8%

**A.3 Método Automático por Agrupamento****Iterações do Fechamento: 0**

Total de classificações: 28928

Total de acertos: 25413

Taxa de acerto: 87,84%

Total de acertos com densidade apenas: 5775

Taxa de acerto com densidade apenas: 63,54%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4510	6631	68,01%
Média	962	1710	56,25%
Alta	303	747	40,56%

**Iterações do Fechamento: 3**

Total de classificações: 28928

Total de acertos: 25593

Taxa de acerto: 88,47%

Total de acertos com densidade apenas: 5955

Taxa de acerto com densidade apenas: 65,52%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4471	6631	67,42%
Média	896	1710	52,39%
Alta	588	747	78,71%

**Iterações do Fechamento: 5**

Total de classificações: 28928

Total de acertos: 25563

Taxa de acerto: 88,36%

Total de acertos com densidade apenas: 5925

Taxa de acerto com densidade apenas: 65,19%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4409	6631	66,49%
Média	856	1710	50,05%
Alta	660	747	88,35%

**Iterações do Fechamento: 7**

Total de classificações: 28928

Total de acertos: 25560

Taxa de acerto: 88,35%

Total de acertos com densidade apenas: 5922

Taxa de acerto com densidade apenas: 65,16%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4396	6631	66,29%
Média	858	1710	50,17%
Alta	668	747	89,42%

**Iterações do Fechamento: 9**

Total de classificações: 28928

Total de acertos: 25548

Taxa de acerto: 88,31%

Total de acertos com densidade apenas: 5910

Taxa de acerto com densidade apenas: 65,03%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	19638	19840	98,98%
Baixa	4387	6631	66,15%
Média	842	1710	49,23%
Alta	681	747	91,16%

**Iterações do Fechamento: 11**

Total de classificações: 28928

Total de acertos: 25621

Taxa de acerto: 88,56%

Total de acertos com densidade apenas: 5983

Taxa de acerto com densidade apenas: 65,83%

Classificação	Acertos	Total	Taxa de acerto
Zero	19638	19840	98,98%
Baixa	4468	6631	67,38%
Média	839	1710	49,06%
Alta	676	747	90,49%

**A.4 Método Privilegiado (Média)****Iterações do Fechamento: 0**

Total de classificações: 3429

Total de acertos: 3150

Taxa de acerto: 91,86%

Total de acertos com densidade apenas: 267

Taxa de acerto com densidade apenas: 61,37%

Classificação	Acertos	Total	Taxa de acerto
Zero	2883	2994	96,29%
Baixa	154	198	77,77%
Média	80	169	47,33%
Alta	33	68	48,52%

**Iterações do Fechamento: 3**

Total de classificações: 3429

Total de acertos: 3192

Taxa de acerto: 93,08%

Total de acertos com densidade apenas: 298

Taxa de acerto com densidade apenas: 68,5%

Classificação	Acertos	Total	Taxa de acerto
Zero	2894	2994	96,65%
Baixa	157	198	79,29%
Média	99	169	58,57%
Alta	42	68	61,76%

**Iterações do Fechamento: 5**

Total de classificações: 3429

Total de acertos: 3201

Taxa de acerto: 93,35%

Total de acertos com densidade apenas: 304

Taxa de acerto com densidade apenas: 69,88%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2897	2994	96,76%
Baixa	157	198	79,29%
Média	103	169	60,94%
Alta	44	68	64,7%

**Iterações do Fechamento: 7**

Total de classificações: 3429

Total de acertos: 3206

Taxa de acerto: 93,49%

Total de acertos com densidade apenas: 307

Taxa de acerto com densidade apenas: 70,57%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2899	2994	96,82%
Baixa	160	198	80,8%
Média	107	169	63,31%
Alta	40	68	58,82%

**Iterações do Fechamento: 9**

Total de classificações: 3429

Total de acertos: 3211

Taxa de acerto: 93,64%

Total de acertos com densidade apenas: 305

Taxa de acerto com densidade apenas: 70,11%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2906	2994	97,06%
Baixa	159	198	80,3%
Média	105	169	62,13%
Alta	41	68	60,29%

**Iterações do Fechamento: 11**

Total de classificações: 3429

Total de acertos: 3215

Taxa de acerto: 93,75%

Total de acertos com densidade apenas: 309

Taxa de acerto com densidade apenas: 71,03%

Classificação	Acertos	Total	Taxa de acerto
Zero	2906	2994	97,06%
Baixa	158	198	79,79%
Média	110	169	65,08%
Alta	41	68	60,29%

**A.5 Método Padrão (Média)****Iterações do Fechamento: 0**

Total de classificações: 3429

Total de acertos: 3137

Taxa de acerto: 91,48%

Total de acertos com densidade apenas: 198

Taxa de acerto com densidade apenas: 45,51%

Classificação	Acertos	Total	Taxa de acerto
Zero	2939	2994	98,16%
Baixa	145	198	73,23%
Média	48	169	28,4%
Alta	5	68	7,35%

**Iterações do Fechamento: 3**

Total de classificações: 3429

Total de acertos: 3177

Taxa de acerto: 92,65%

Total de acertos com densidade apenas: 239

Taxa de acerto com densidade apenas: 54,94%

Classificação	Acertos	Total	Taxa de acerto
Zero	2938	2994	98,12%
Baixa	147	198	74,24%
Média	78	169	46,15%
Alta	14	68	20,58%

**Iterações do Fechamento: 5**

Total de classificações: 3429

Total de acertos: 3184

Taxa de acerto: 92,85%

Total de acertos com densidade apenas: 249

Taxa de acerto com densidade apenas: 57,24%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2935	2994	98,02%
Baixa	146	198	73,73%
Média	88	169	52,07%
Alta	15	68	22,05%

**Iterações do Fechamento: 7**

Total de classificações: 3429

Total de acertos: 3188

Taxa de acerto: 92,97%

Total de acertos com densidade apenas: 251

Taxa de acerto com densidade apenas: 57,7%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2937	2994	98,09%
Baixa	147	198	74,24%
Média	87	169	51,47%
Alta	17	68	25%

**Iterações do Fechamento: 9**

Total de classificações: 3429

Total de acertos: 3193

Taxa de acerto: 93,11%

Total de acertos com densidade apenas: 256

Taxa de acerto com densidade apenas: 58,85%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2937	2994	98,09%
Baixa	147	198	74,24%
Média	91	169	53,84%
Alta	18	68	26,47%

**Iterações do Fechamento: 11**

Total de classificações: 3429

Total de acertos: 3206

Taxa de acerto: 93,49%

Total de acertos com densidade apenas: 272

Taxa de acerto com densidade apenas: 62,52%

Classificação	Acertos	Total	Taxa de acerto
Zero	2934	2994	97,99%
Baixa	148	198	74,74%
Média	94	169	55,62%
Alta	30	68	44,11%

**A.6 Método Autom. por Agrupamento (Média)****Iterações do Fechamento: 0**

Total de classificações: 3429

Total de acertos: 3166

Taxa de acerto: 92,33%

Total de acertos com densidade apenas: 281

Taxa de acerto com densidade apenas: 64,59%

Classificação	Acertos	Total	Taxa de acerto
Zero	2885	2994	96,35%
Baixa	157	198	79,29%
Média	106	169	62,72%
Alta	18	68	26,47%

**Iterações do Fechamento: 3**

Total de classificações: 3429

Total de acertos: 3190

Taxa de acerto: 93,03%

Total de acertos com densidade apenas: 306

Taxa de acerto com densidade apenas: 70,34%

Classificação	Acertos	Total	Taxa de acerto
Zero	2884	2994	96,32%
Baixa	158	198	79,79%
Média	106	169	62,72%
Alta	42	68	61,76%

**Iterações do Fechamento: 5**

Total de classificações: 3429

Total de acertos: 3191

Taxa de acerto: 93,05%

Total de acertos com densidade apenas: 314

Taxa de acerto com densidade apenas: 72,18%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2877	2994	96,09%
Baixa	155	198	78,28%
Média	106	169	62,72%
Alta	53	68	77,94%

**Iterações do Fechamento: 7**

Total de classificações: 3429

Total de acertos: 3200

Taxa de acerto: 93,32%

Total de acertos com densidade apenas: 321

Taxa de acerto com densidade apenas: 73,79%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2879	2994	96,15%
Baixa	158	198	79,79%
Média	112	169	66,27%
Alta	51	68	75%

**Iterações do Fechamento: 9**

Total de classificações: 3429

Total de acertos: 3194

Taxa de acerto: 93,14%

Total de acertos com densidade apenas: 314

Taxa de acerto com densidade apenas: 72,18%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2880	2994	96,19%
Baixa	153	198	77,27%
Média	108	169	63,9%
Alta	53	68	77,94%



**Iterações do Fechamento: 11**

Total de classificações: 3429

Total de acertos: 3202

Taxa de acerto: 93,37%

Total de acertos com densidade apenas: 316

Taxa de acerto com densidade apenas: 72,64%

<b>Classificação</b>	<b>Acertos</b>	<b>Total</b>	<b>Taxa de acerto</b>
Zero	2886	2994	96,39%
Baixa	155	198	78,28%
Média	107	169	63,31%
Alta	54	68	79,41%