



**UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

ANTONIO ALESSANDRO MENEZES FERNANDES

**CONTAGEM DE VEÍCULOS NA VIA ATRAVÉS DE
IMAGENS DE TRÂNSITO**

**QUIXADÁ
2013**

ANTONIO ALESSANDRO MENEZES FERNANDES

**CONTAGEM DE VEÍCULOS NA VIA ATRAVÉS DE IMAGENS DE
TRÂNSITO**

Trabalho de Conclusão de Curso submetido à Coordenação do Curso Bacharelado em Sistemas de Informação da Universidade Federal do Ceará como requisito parcial para obtenção do grau de Bacharel.

Área de concentração: computação

Orientador Prof. Dr. Críston Pereira de Souza

**QUIXADÁ
2013**

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca do Campus de Quixadá

F398c Fernandes, Antonio Alessandro Menezes
 Contagem de veículos na via através de imagens de trânsito/Antonio Alessandro Menezes
 Fernandes. – 2013.
 42f. : il. color., enc. ; 30 cm.

 Monografia (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de
 Sistemas de Informação, Quixadá, 2013.
 Orientação: Prof. Dr. Críston Pereira de Souza
 Área de concentração: Computação

1. Otimização combinatória 2. Aprendizado de máquina 3. Trânsito-congestionamento I. Título.

CDD 519.3

ANTONIO ALESSANDRO MENEZES FERNANDES

CONTAGEM DE VEÍCULOS NA VIA ATRAVÉS DE IMAGENS DE TRÂNSITO

Trabalho de Conclusão de Curso submetido à Coordenação do Curso Bacharelado em Sistemas de Informação da Universidade Federal do Ceará como requisito parcial para obtenção do grau de Bacharel.

Área de concentração: computação

Aprovado em: _____ / dezembro / 2013.

BANCA EXAMINADORA

Prof. Dr. Críston Pereira de Souza
Universidade Federal do Ceará-UFC

Prof. Mcs. Ticiania Linhares Coelho da Silva
Universidade Federal do Ceará-UFC

Prof. Mcs. Ricardo Reis Pereira
Universidade Federal do Ceará-UFC

Aos meus pais, aos meus irmãos
e aos meus amigos...

AGRADECIMENTOS

Agradeço a Deus por ter me concedido a vitória contra o câncer quando eu tinha apenas 8 anos de idade.

Aos meus pais Cesário e Raimunda, por terem acreditado em mim e dado total apoio durante todo o percurso da minha vida pessoal e mais ainda, na minha vida acadêmica.

Aos meus irmãos Sandra, Solange e Silvan, que sempre permaneceram ao meu lado nos momentos bons e maus momentos. Por essa razão, tenho enorme gratidão por vocês.

Ao Prof. Msc. Wladimir Araújo Tavares, pela amizade, pelas palavras de motivação, e por ter contribuído para o enriquecimento da minha bagagem sobre programação nas reuniões do Grupo de Estudos para Maratona de Programação.

Ao Prof. Dr. Críston Pereira de Souza, pela amizade, paciência e incentivo. E através do seu apoio tornou-se possível o desenvolvimento deste trabalho.

Aos meus grandes amigos Geovanny Magalhães e Marcelo Arraes pelas conversas, amizade e ajuda durante o curso.

A todos os professores da Universidade Federal do Ceará – Campus Quixadá, que de alguma forma contribuíram para minha formação acadêmica e profissional.

A todos os meus amigos da Universidade Federal do Ceará – Campus Quixadá, pelo afeto e convívio.

"No meio da dificuldade encontra-se a oportunidade."
(Albert Einstein)

RESUMO

Congestionamentos estão cada vez mais presentes em nossas vidas. Estudos mostram que os congestionamentos se formam quando o número de veículos ultrapassa aproximadamente $1/3$ da capacidade da via. É difícil para os departamentos de trânsito prever os pontos de lentidão nas cidades. Neste trabalho propomos realizar uma estimativa do número de veículos na via através da análise de imagens de trânsito, coletadas por câmeras de registro de infrações. Isso deve ser feito com uma baixa taxa de erro. Para realizar esta tarefa, empregamos neste trabalho algoritmos de visão computacional e aprendizado de máquina, inspirados principalmente na abordagem de Jain, Sharma e Subramanian (2012), embora as características das câmeras utilizadas neste trabalho sejam claramente distintas. Os experimentos mostram que a estratégia proposta apresenta baixa taxa de erro, tanto em vídeos diurnos quanto noturnos.

Palavras chave: Visão computacional, aprendizado de máquina, análise de congestionamento.

ABSTRACT

Congestion are increasingly present in our lives. Studies show that congestion is formed when the number of vehicles exceeds about $1/3$ of the capacity of the channel. It is difficult to predict traffic departments points slowdowns in cities. We propose estimating the number of vehicles on the road by analyzing images from traffic cameras collected for registration violations. This should be done with a low error rate. To accomplish this task, we employ in this paper algorithms of computer vision and machine learning, inspired primarily in Jain, Sharma and Subramanian (2012) approach, although the characteristics of the cameras used in this paper are clearly distinct. The experiments show that the proposed strategy presents a low error rate, in both day and nighttime videos.

Keywords: Computer vision, machine learning, analyzing congestion.

LISTA DE ILUSTRAÇÕES

Figura 1 – Curva de Tráfego.....	22
Figura 2 – Histograma de intensidade (nível de cinza) região marcada do asfalto da via sem veículos.....	24
Figura 3 – Histograma de intensidade (nível de cinza) região marcada do asfalto da via congestionada	24
Figura 4 – Threshold da via durante a noite	25
Figura 5 – Perceptron de Múltiplas Camadas.....	27
Figura 6 – Frame de vídeo durante a noite	29
Figura 7 – Frame de vídeo durante o dia	29
Figura 8 – Veículos antes da oclusão	30
Figura 9 – Oclusão causada pela passagem de um caminhão	31
Figura 10 – Máscara para marcar o asfalto no frame.	31
Figura 11 – Imagem da via antes e depois da aplicação da máscara	32
Figura 12 – Gráfico do percentual de pixels versus o número de veículos dos vídeos diurnos	33
Figura 13 – Gráfico do percentual de pixels versus o número de veículos dos vídeos noturnos	34
Figura 14 – Comparação dos REQM's e EMA's dos algoritmos na abordagem diurna.....	36
Figura 15 – Melhor taxa de aprendizagem do MLP para vídeos diurnos.....	37
Figura 16 – Erro do classificador MLP para os vídeos diurnos	38
Figura 17 – Erros do MLP acima da reta da bisetriz	38
Figura 18 – Erros do MLP abaixo da reta da bisetriz	38
Figura 19 – Escolha do melhor número de vizinhos mais próximos para os vídeos diurnos...	39
Figura 20 – Comparação dos REQM's e EMA's dos algoritmos na abordagem noturna.....	40
Figura 21 – Erro do classificador Regressão Linear para os vídeos noturnos.....	41
Figura 22 – Erro da Regressão Linear acima da reta da bisetriz.....	41
Figura 23 – Erro da Regressão Linear abaixo da reta da bisetriz	41
Figura 24 - Melhor taxa de aprendizagem do MLP para vídeos noturnos	42
Figura 25 – Escolha do melhor número de vizinhos mais próximos para os vídeos noturnos.	43

SUMÁRIO

1	INTRODUÇÃO.....	15
2	REVISÃO BIBLIOGRÁFICA.....	17
2.1	Mobilidade Urbana	17
2.2	Visão Computacional.....	18
2.2.1	<i>Imagem Digital</i>	18
2.2.2	<i>Biblioteca OpenCV</i>	19
2.3	Análise de Congestionamento	20
2.3.1	<i>Curva de Tráfego</i>	21
2.3.2	<i>Abordagens para estimativa automática do número de veículos</i>	22
2.3.2.1	<i>Jain, Sharma e Subramanian (2012) - período diurno</i>	23
2.3.2.2	<i>Jain, Sharma e Subramanian (2012) - período noturno</i>	24
2.4	Algoritmos de previsão numérica	25
2.4.1	<i>IBk</i>	25
2.4.2	<i>Regressão linear</i>	26
2.4.3	<i>Perceptron de múltiplas camadas</i>	27
2.4.4	<i>Avaliação da previsão numérica</i>	27
3	DESENVOLVIMENTO.....	28
3.1	Aquisição dos vídeos	28
3.1.1	<i>Qualidade dos vídeos das câmeras de detecção de infrações</i>	29
3.1.2	<i>Definição da área de marcação do asfalto</i>	31
3.2	Cálculo do somatório do percentual de pixels	32
3.2.1	<i>Cálculo do somatório do percentual de pixels cinza em cada linha</i>	32
3.2.2	<i>Cálculo do somatório do percentual de pixels branco em cada linha</i>	33
3.3	Rotulação dos vídeos	34
3.4	Estimativa do número de veículos	35
4	RESULTADOS	35
4.1	Resultados experimentais para os vídeos diurnos.....	36
4.2	Resultados experimentais para os vídeos noturnos.....	39
5	CONSIDERAÇÕES FINAIS	43
	REFERÊNCIAS	44

1 INTRODUÇÃO

O cenário atual do sistema viário brasileiro apresenta vários pontos de congestionamento nas grandes cidades devido ao aumento da frota de veículos em circulação. Um motivo para tal crescimento pode ser os incentivos dados pelo governo para redução do Imposto sobre Produtos Industrializados (IPI), favorecendo a compra de automóveis e caminhões. Essa medida visa incentivar o mercado automobilístico brasileiro. (GOVERNO..., 2013).

Ações do governo como a redução do IPI favorecem o aumento da frota veicular, principalmente automóveis particulares, tornando o carro como símbolo da mobilidade, independência e do meio de transporte ágil para o deslocamento em zonas urbanas. Outro motivo para o aumento da compra de veículos é por causa do transporte público deficiente. A má qualidade do transporte coletivo e a disputa do espaço da rua com o crescente número de automóveis colaboram com um ambiente urbano poluído e agressivo (ZANDONADE; MORATTI, 2012).

O problema ocorre quando muitas pessoas fazem uso do automóvel em horários de pico, resultando uma disputa por espaço no asfalto e originando lentidões. Essa enorme quantidade de carros parados em filas nas vias colabora para intensificação dos impactos ambientais nas áreas urbanas, com aumento dos níveis de poluição sonora e do ar. Com o excesso de veículos nas vias ocorrem atrasos no deslocamento da população urbana, tanto para quem usa transporte público como transporte particular.

Um dos motivos que originam os congestionamentos é o excesso de veículos em circulação, ou seja, além da capacidade da via. Cada via de tráfego está associada com uma densidade de tráfego, que representa o número máximo de veículos na via sem gerar lentidão (JAIN; SHARMA; SUBRAMANIAN, 2012, tradução nossa).

Os congestionamentos resultam em problemas no trânsito, no meio ambiente e na economia. Podemos citar alguns impactos na economia, como por exemplo: custos com logística, gastos com combustíveis e perdas no consumo de bens. De acordo com a pesquisa realizada pela Fundação Getúlio Vargas, o preço dos congestionamentos na cidade de São Paulo é de um prejuízo que já equivale a 1% de todo o Produto Interno Bruto (PIB) do Brasil (TRÂNSITO..., 2013).

“Os departamentos de controle de tráfego precisam de informações em tempo real para a estimativa de tráfego: com a intensidade do tráfego rodoviário, a ocupação de pista, nível de congestionamento, a estimativa dos tempos de viagem e a detecção de incidente.” (MILLA, 2010, tradução nossa). Câmeras de vídeo são instaladas em vias para fins de análise em tempo real do tráfego de veículos. Essas câmeras fornecem informações visuais do fluxo do tráfego, permitindo assim um melhor monitoramento, tornando mais rápidas as providências para remoção de interferências que dão origem aos congestionamentos. Porém, é necessário o fator humano para observar e identificar anormalidades que geram lentidões a partir de vídeos do fluxo do trânsito capturado por câmeras posicionadas na via.

Jain, Sharma e Subramanian (2012) apresentam um mecanismo de processamento de imagens de câmeras de Circuito Fechado de Televisão (CCTV), automatizando a detecção dos níveis de congestionamento no tráfego rodoviário. Um ponto forte desse trabalho é a análise do congestionamento através de vídeos de trânsito com baixa qualidade de imagens.

A empresa Fotosensores Tecnologia Eletrônica Ltda contribuiu para este estudo fornecendo vídeo de trânsito coletados em sua câmera de teste, similar às câmeras utilizadas na identificação de infrações. Estes vídeos compreendiam tanto o período do dia como o período da noite.

O objetivo geral deste trabalho é produzir um estudo para estimar o número de veículos na via utilizando imagens de vídeos de câmeras de detecção de infrações. Com a estimativa do número de veículos em imagens de trânsito é possível realizar análise de congestionamento como mostra Jain, Sharma e Subramanian (2012) na seção 2.3.1. E tendo como objetivos específicos:

- a) Rotular manualmente, ou seja, contar o número de veículos (exceto motos) em todos os frames de vídeo. O procedimento de rotulação é realizado nos vídeos diurnos e noturnos.
- b) Implementar aplicativo que calcula o atributo percentual de pixels em cada linha para os vídeos diurnos e noturnos;
- c) Utilizar os algoritmos IBk, Regressão Linear e Perceptron de Múltiplas Camadas para estimar o número de veículos em novos frames de vídeo da via;
- d) Utilizar as métricas Erro Quadrático Médio e a Raiz do Erro Quadrático Médio para avaliar os algoritmos de classificação.

As seções deste trabalho estão dispostas da seguinte forma. Na Seção 2 é realizada a revisão bibliográfica. Nesta seção é feito o levantamento das abordagens, algoritmos e ferramentas empregadas neste trabalho. Na Seção 3 descrevemos como foi realizada a aquisição, rotulação dos vídeos, cálculo do percentual de pixels e a estimativa do número de veículos, e na Seção 4 é apresentada uma discussão sobre os resultados alcançados.

2 REVISÃO BIBLIOGRÁFICA

Nesta seção apresentaremos os principais conceitos utilizados para realização desse trabalho. Na Seção 2.1 discutimos mobilidade urbana. Os conceitos sobre visão computacional são apresentados na Seção 2.2. Finalmente, na Seção 2.3 são apresentadas as principais abordagens da literatura para análise automática de congestionamento.

2.1 Mobilidade Urbana

Notamos que nos últimos anos, houve um aumento da renda da população brasileira através de incentivos do governo. Esse aumento na renda proporcionou a motorização de uma parcela significativa da população que até então usava o transporte coletivo, através da aquisição de automóveis e de motos. Assim percebemos um desequilíbrio na balança, aumento pela demanda pelo transporte privado e queda pela demanda pelo transporte coletivo. E ainda mais, nunca na história discutimos tanto sobre os congestionamentos como nesses últimos anos, pelo fato deles estarem cada vez mais presentes em nossas vidas.

Enquanto isso, as cidades brasileiras continuam resolvendo o problema dos congestionamentos apostando na construção de novos viadutos, avenidas e túneis. Ainda é muito pouco o que se investe na mobilidade urbana para possibilitar o planejamento e construção de novas infra-estruturas de transportes coletivos de boa qualidade para população urbana. Se nenhuma atitude for tomada para melhoria dos transportes coletivos, o quadro do transporte coletivo no Brasil continua a agravar-se cada vez mais, e tornando-nos refém dos congestionamentos.

O dicionário Aurélio (FERREIRA, 2001) define mobilidade como qualidade de quem é móvel, ou seja, mover de um lugar para outro. Podemos extrair essa essência para o conceito de mobilidade urbana que está relacionado com o deslocamento pelo ambiente

urbanizado. Para Alves e Junior (2009), o conceito de mobilidade urbana pode ser compreendido como a facilidade de deslocamentos de pessoas e bens dentro de um espaço urbano, e acessibilidade como o acesso da população para realizar suas atividades e deslocamentos.

Segundo a recente lei brasileira, Lei Federal nº. 12.587, de 3 de janeiro de 2012 (BRASIL, 2012), responsável por instituir as diretrizes da Política Nacional de Mobilidade Urbana, define mobilidade urbana como a condição em que se realizam os deslocamentos de pessoas e cargas no espaço urbano. Nesse caso, mobilidade urbana não é necessariamente o deslocamento de pessoas, mas de pessoas e cargas em espaços urbanos por meios físicos.

2.2 Visão Computacional

Conci, Azevedo e Leta (2007) definem a visão computacional como o domínio da ciência da computação que estuda e aplica métodos que permitem aos computadores interpretarem o conteúdo de imagens. A visão computacional é a ciência responsável por emular a visão humana no computador. A área de visão computacional permite que o computador realize extração de informações, identificação e classificação de objetos existentes em imagens digitais. As aplicações mais difundidas da visão computacional são utilizadas para o reconhecimento e rastreamento de objetos, inspeção de peças em linhas de montagem, processamento de vídeo e imagens, orientação e movimentação de robôs, medicina e outras áreas que é necessário extrair informações de imagens.

2.2.1 Imagem Digital

Para poder realizar o processamento de uma imagem é necessário digitalizá-la, ou seja, a imagem será convertida de contínua (real) para sua representação discreta (digital) (CONCI; AZEVEDO; LETA, 2007). A imagem digital consiste numa representação da imagem real de modo que possa ter um uso computacional, podendo assim, ser armazenada na forma de um arquivo digital. Para ser processada pelo computador, uma imagem precisa ser descrita por um número finito de pontos e por um número finito de cores, onde cada ponto da imagem dá-se o nome de pixel, a menor unidade de uma imagem digital.

O computador representa uma imagem digital como sendo uma matriz de N linhas e M colunas. Segundo Conci, Azevedo e Leta (2007), uma imagem digital é descrita por uma

matriz $N \times M$ de valores de pixel como números inteiros positivos, que indicam a intensidade de cor em cada posição (x, y) da imagem. A cor do pixel é representada com um número inteiro de 8 bits. Um byte pode então assumir $2^8 = 256$ valores diferentes, variando de 0 a 255.

Podemos classificar a imagem digital de acordo com o seu canal de cor: imagens monocromáticas e imagens coloridas. Imagens monocromáticas são imagens que possuem apenas um canal de cor. Elas podem ser subdivididas em:

- a) Imagens binárias são imagens na qual os pixels da imagem possuem apenas dois valores, onde geralmente esses valores são 0 ou 1.
- b) Imagem em escala de cinza, que são imagens na qual os pixels podem assumir faixa de valores intermediários de preto ao branco, geralmente de 0 a 255, se for representado por um byte.
- c) Imagens coloridas podem ser representadas com a composição de três canais, em escala de cores de vermelho, verde e azul no sistema RGB.

Os frames do vídeo são imagens digitais, no caso, um vídeo é um conjunto de imagens tiradas em momentos diferentes, o que dá a sensação de movimento. No vídeo, a cada 1 segundo equivale a 24 frames, ou seja, 24 imagens.

2.2.2 *Biblioteca OpenCV*

Para facilitar o desenvolvimento de softwares na área da visão computacional usa-se a biblioteca OpenCV. Bradski e Kaehler (2008) comentam que Open Source Computer Vision Library (OpenCV) é uma biblioteca de visão computacional de código aberto. A biblioteca que utiliza a linguagem C/C++, e tendo versões para Linux, Windows e Mac OS X. OpenCV foi projetado para a eficiência computacional e com um forte foco em aplicações em tempo real.

Neste trabalho a biblioteca de visão computacional OpenCV foi utilizada para auxiliar no desenvolvimento de dois aplicativos. O primeiro aplicativo auxilia no processo de rotulação dos frames dos vídeos. O outro aplicativo calcula o percentual de pixels em cada linha para os frames dos vídeos diurnos e noturnos.

2.3 Análise de Congestionamento

A análise de vídeo automática do tráfego pode ser realizada por aplicações da área da visão computacional. A visão computacional tem como objetivo criar e disponibilizar técnicas para extração de informações relevantes a partir da análise de imagens. A seguir são citados alguns trabalhos que realizam algum tipo de análise de tráfego a partir de imagens de vídeo.

Coifman et al. (1998) desenvolveram um sistema que realiza detecção e rastreamento de veículos através de vídeo em condições de oclusão parcial, transições de iluminação entre a noite/dia e dia/noite, vibração da câmera devido ao vento, e longas sombras geradas pelos os veículos. Coifman et al. (1998) utilizam da abordagem de não detectar o veículo por inteiro, ao invés disso, apenas características do veículo são detectadas.

Mandellos, Keramitsoglou e Kiranoudis (2001) apresentaram um estudo de um sistema que implementa o método de subtração de fundo devidamente adaptado para os propósitos de um sistema de vigilância de tráfego. A subtração de fundo permite destacar os objetos que não fazem parte do cenário da via. A inovação deste estudo encontra-se em um novo algoritmo para subtração de fundo. Esse novo algoritmo é baseado em amostragem estatística da cor do pixel ao longo do tempo. Os experimentos realizados mostraram que o algoritmo proposto é capaz de trabalhar em tempo real devido à sua baixa complexidade. Em todos os casos, o algoritmo conseguiu recortar com precisão os veículos das imagens em várias condições adversas, incluindo congestionamento pesado e mudanças na iluminação.

Lei et al. (2008) apresentam um sistema baseado em vídeo para detecção e contagem de veículos em tempo real. Esse sistema utiliza câmera de vigilância posicionada em certa altura da via para capturar o fluxo do tráfego da via. Os dois principais métodos utilizados neste sistema são: a estimativa de fundo adaptativo e a eliminação de sombra Gaussiana. A primeira permite que uma detecção de movimentos nos frames de vídeo. O último é baseado no sistema de cores HSV, que é capaz de lidar com diferentes tamanhos e intensidade de sombras.

Kanhere et al. (2007) apresentam um sistema que realiza o monitoramento automático de auto-estradas em tempo real. A câmara é montada a uma altura e distância lateral relativamente pequena da via. Em tal caso, devido à perspectiva da câmera, a oclusão é gerada pelas passagens de veículos de grande porte. O sistema apresentado é baseado em uma nova técnica de detecção e de rastreamento (tracking) que utilizasse das frentes dos veículos.

O sistema coleta uma variedade de dados de tráfego, incluindo volume, tempo médio, velocidade média, densidade, classificação dos veículos, e atividade de mudança de pista.

De acordo com Jain, Sharma e Subramanian (2012), um método comum para detecção de congestionamento de tráfego é colocar sensores sob a via e contar o número de vezes que são acionados pela passagem das rodas dos veículos. Jain, Sharma e Subramanian (2012) explicam que essa abordagem sofre quatro problemas principais:

- a) É caro implantar esse método, visto que os sensores precisam ser parcialmente incorporados no asfalto;
- b) Os sensores por estarem expostos no asfalto, estão propensos ao roubo por alguém mal-intencionado;
- c) Sensores precisam ser colocados em vários pontos na via para manter a contagem precisa;
- d) Em um único trecho de via, os sensores precisam ser colocados em intervalos regulares, de modo a estimar a densidade em diferentes segmentos da via.

Jain, Sharma e Subramanian (2012) comentam que o monitoramento do trânsito vem sendo amplamente estudado na área da visão computacional. Existe uma grande variedade de abordagens na literatura de contagem, rastreamento e análise de trânsito. Atualmente o uso de câmeras para o monitoramento do trânsito vem se popularizando nas cidades com alta taxa de veículos em circulação nas ruas.

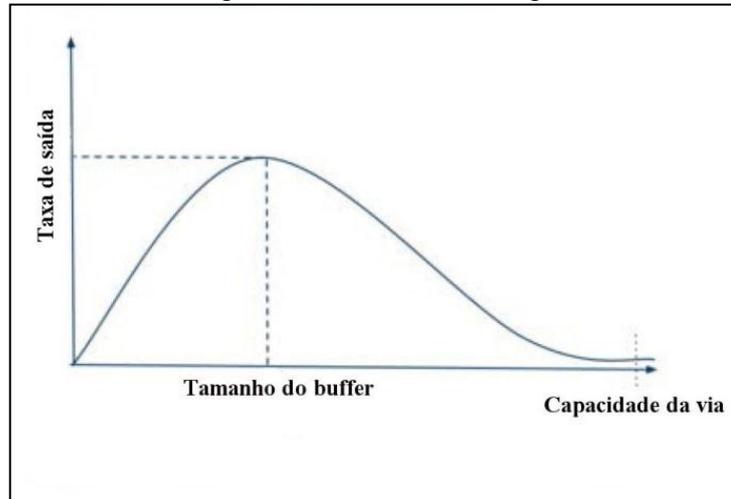
Buch, Velastin e Orwell (2011) comentam que nos últimos anos houve um aumento da análise automática do tráfego urbano. Cujo fato é devido, em parte, ao aumento do número de câmeras e outros sensores, infra-estrutura melhorada, e conseqüentemente, disponibilidade dos dados. Além disso, o avanço das técnicas de processamento do vídeo, em conjunto com o aumento do poder de computação, proporcionou novas aplicações.

2.3.1 Curva de Tráfego

Jain, Sharma e Subramanian (2012) explicam que existem várias métricas que definem as características do tráfego, tais como: velocidade, fluxo e a densidade. Cada via está associada com uma densidade de tráfego, que representa o número de veículos por unidade da via. Qualquer via de trânsito possui uma taxa de saída de veículos, que varia de acordo com a densidade do tráfego. A Curva de Tráfego, na Figura 1, mostra o

relacionamento entre a densidade e a taxa de saída, mostra ainda que o congestionamento começa quando o número de carros ultrapassa $\frac{1}{3}$ da capacidade da via.

Figura 1 – Curva de Tráfego



Fonte: Jain, Sharma e Subramanian (2012).

Existem alguns parâmetros da Curva de Tráfego que descrevem o relacionamento entre a densidade e a taxa de saída da via:

- a) Capacidade da via – é o número máximo de veículos que a via pode suportar;
- b) Tamanho do buffer – é o número de veículos na via em qualquer ponto no tempo;
- c) Taxa de saída – é a taxa de veículos que saem da via.

Neste trabalho estamos interessados em utilizar imagens de trânsito para estimar o número de veículos na via que, de acordo com a Figura 1, é uma maneira indireta de estimar o nível de congestionamento.

2.3.2 Abordagens para estimativa automática do número de veículos

Nesta seção são discutidas duas abordagens para estimativa do número de veículos de Jain, Sharma e Subramanian (2012). Sendo que, uma abordagem é para os vídeos diurnos e a outra para os vídeos noturnos. Essas duas abordagens são utilizadas neste trabalho.

2.3.2.1 Jain, Sharma e Subramanian (2012) - período diurno

Segundo Jain, Sharma e Subramanian (2012), durante o período diurno, quando não existe muito fluxo de tráfego de veículos na estrada, a escala de cinza do asfalto é predominante, independentemente da mudança de iluminação no decorrer do dia. Quando a estrada apresenta um maior fluxo de veículos, a quantidade de cinza visível da via na imagem reduz. Isso é devido ao fato da maioria dos veículos terem cor diferente da cor do asfalto.

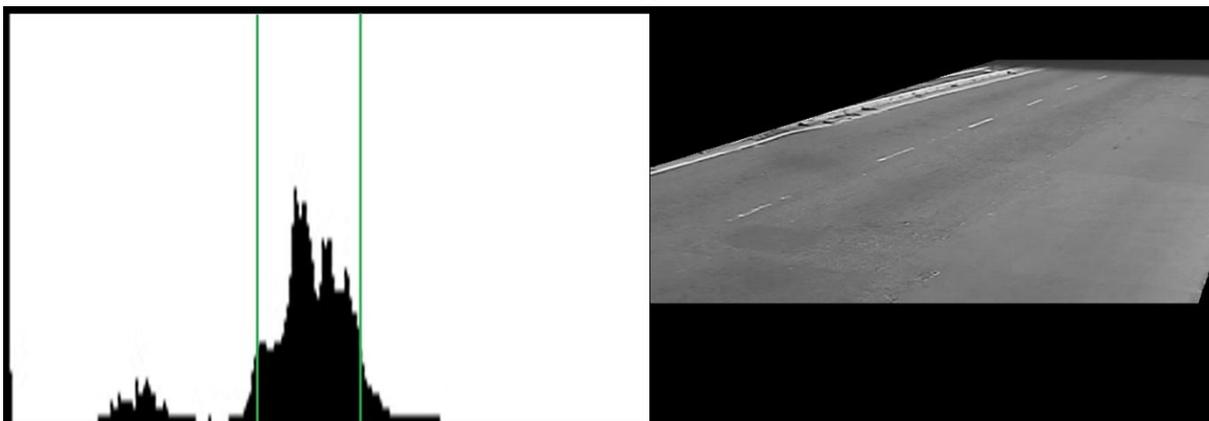
Para realizar a estimativa da densidade do tráfego, é necessário definir uma região do asfalto na imagem (frame de vídeo), por exemplo, um polígono. Após a definição da área, a imagem será convertida para escala de cinza e gerado um histograma da escala de cinza a partir da área definida na imagem. Conci, Azevedo e Leta (2007) definem histograma de uma imagem como um conjunto de números indicando o percentual de pixels naquela imagem, que apresenta determinada variação em luminosidade.

Jain, Sharma e Subramanian (2012) explicam que o asfalto da via encontrasse na faixa de cinza entre 135-165 do histograma. Dependendo do nível do congestionamento da via, o histograma apresenta diferentes picos da área definida na imagem. Em uma imagem com a pista vazia, sem nenhum veículo, o histograma apresenta altos picos. Em comparação com um histograma de uma pista com congestionamento, exhibe picos uniformes, conforme ilustrado nas Figuras 2 e 3. Ao examinar esses picos no histograma, podemos estimar a densidade do tráfego na via de acordo os pixels cinza da faixa entre 135 a 165 que representa o asfalto.

Neste trabalho aplica essa abordagem definida por Jain, Sharma e Subramanian (2012) para estimar o número de veículos em vídeos do período diurno. Entretanto, para os vídeos do dia cedidos pela Fotosensores, à faixa de escala de pixels cinza é definida entre os valores de pixels de 70 até 190. Estes valores foram obtidos empiricamente.

É realizada uma correção de perspectiva na região do asfalto da imagem da via. A correção de perspectiva é realizada calculando o percentual de pixels cinza entre uma faixa de pixels em cada linha da região do asfalto da imagem. O objetivo dessa correção de perspectiva é evitar erros na estimativa do número de veículos, pois um veículo ocupa mais pixels do frame quando está mais próximo da câmera. As Figuras 2 e 3 ilustram o histograma de intensidade da via da região do asfalto.

Figura 2 – Histograma de intensidade (nível de cinza) região marcada do asfalto da via sem veículos



Fonte: Elaborado pelo autor.

Figura 3 – Histograma de intensidade (nível de cinza) região marcada do asfalto da via congestionada



Fonte: Elaborado pelo autor.

2.3.2.2 Jain, Sharma e Subramanian (2012) - período noturno

Detecção de congestionamento durante o período noturno é um problema difícil por causa do reflexo da luz dos faróis dos automóveis. Jain, Sharma e Subramanian (2012) utilizam as luzes e os reflexos dos faróis dos veículos para calcular o número de veículos na via durante a noite. A abordagem consiste em calcular os pixels brancos do frame de vídeo intensidade acima de determinado limiar.

Jain, Sharma e Subramanian (2012) reduziram os ruídos do reflexo das luzes dos faróis como o processo de limiarização (thresholding) que reduz consideravelmente a quantidade de luz falsa presente na imagem, devido às diferentes intensidades das fontes de luz. O threshold é um método que rejeita aqueles pixels de abaixo ou acima de algum valor estabelecido, mantendo os outros valores. Jain, Sharma e Subramanian (2012) descobriram

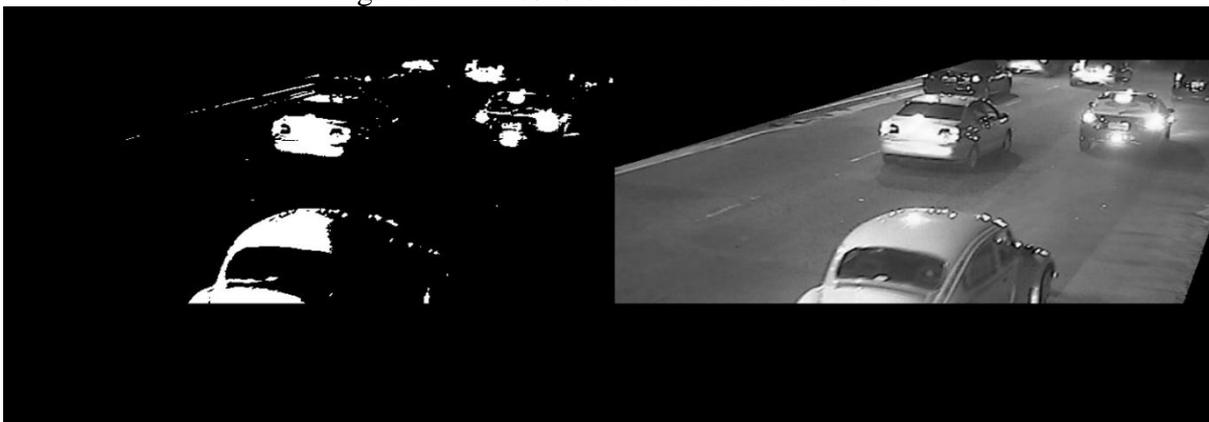
que o threshold de 0,8 é o mais adequado para o conjunto de imagens de vídeo deles sem gerar perda de características das imagens.

Jain, Sharma e Subramanian (2012) usam um somatório que calcula os pixels brancos para estimar a densidade do tráfego como sendo o número de veículos em cada frame de vídeo. Essa somatória usa informações tais com: altura da câmera e altura observada de extensão da via. Essas informações são para corrigir a perspectiva da imagem para estimar a densidade do tráfego.

Esse trabalho também utiliza a abordagem de Jain, Sharma e Subramanian (2012) para estimar o número de veículos no frame para vídeos noturnos. A correção de perspectiva também é utilizada para o cálculo do percentual de pixels brancos em cada linha.

Na Figura 4 é ilustrada a aplicação do threshold na via durante a noite.

Figura 4 – Threshold da via durante a noite



Fonte: Elaborado pelo autor.

2.4 Algoritmos de previsão numérica

Nesta seção são apresentados os algoritmos de previsão numérica utilizados no trabalho. Os algoritmos IBk, Regressão Linear e Perceptron de Múltiplas Camadas são utilizados neste trabalho para realizar a previsão numérica do número de veículos nos frames de vídeo.

2.4.1 IBk

O IBk é um algoritmo que pertence à família de algoritmos utiliza o princípio de que as instâncias do conjunto de treinamento formam um modelo para os dados. Portanto, a

criação de um modelo é bastante simples. Porém, a avaliação de uma nova instância do conjunto de teste é calculada com base na distância mínima entre a instância de entrada e as instâncias memorizadas pelo modelo.

Aha et al. (1991) explica que o IBk é um método de aprendizagem baseado em instâncias (IBL – Instance-Based Learning). Esse tipo de algoritmo é derivado do método de classificação k-Vizinhos Mais Próximo (KNN – K-Nearest Neighbor). Porém, este último é um algoritmo não-incremental e tem como objetivo principal manter uma consistência perfeita como o conjunto inicial de treinamento. Já o algoritmo do tipo IBL é incremental e tem como objetivo maximizar o grau de exatidão sobre novas instâncias do problema.

Neste trabalho utilizamos o algoritmo IBk para estimar o número de carros nas duas abordagens.

2.4.2 *Regressão linear*

“A regressão linear nasce da tentativa de relacionar um conjunto de observações de certas variáveis, designadas genericamente por X_k ($k = 1 \dots p$), com as leituras de certa grandeza Y .” (MATOS, 1995).

Witten, Frank e Hall (2011) definem como sendo um método básico em estatísticas. A ideia é para expressar a classe como uma combinação linear dos atributos, com pesos pré-determinados:

$$x = w_0 + w_1 a_1 + w_2 a_2 + \dots + w_k a_k$$

Onde x é a classe; a_1, a_2, \dots, a_k são os valores dos atributos e w_0, w_1, \dots, w_k são os pesos. Os pesos são calculados a partir dos dados de treinamento. Aqui a notação fica um pouco complexa, porque precisamos de uma forma de expressar os valores de atributos para cada instância de treinamento.

Podemos definir regressão linear como um método que utiliza a relação entre duas ou mais variáveis. De tal forma que cada variável pode ser prevista a partir de outros valores. O caso mais simples de regressão linear é quando temos duas variáveis e a relação entre elas podem ser representadas por uma reta.

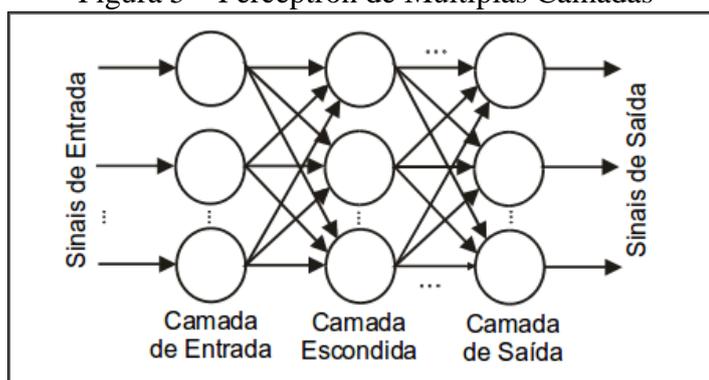
2.4.3 Perceptron de múltiplas camadas

De Souza et al. (2003) explicam que uma rede perceptron de múltiplas camadas (Multilayer Perceptron - MLP) é constituída de uma camada de entrada, uma ou mais camadas escondida ou camadas intermediárias, e uma camada de saída. Cada camada é composta por unidades de processamento chamadas de neurônios. A interligação entre camadas é feita pelas conexões que possuem pesos sinápticos.

Na camada de entrada é onde a informação de entrada é recebida, e em seguida é repassada para as outras camadas. Nesta camada não é realizado nenhum tipo de processamento. As camadas ocultas são responsáveis por transmitir as informações de entrada para a camada de saída. Porém, antes de propagar essas informações para os outros neurônios até chegar à saída da rede, as informações sofrem processamento. “O neurônio calcula sua saída aplicando uma função de limiar à combinação linear dos elementos do vetor de entrada apresentado, juntamente com a entrada fixa, ponderada por seus respectivos pesos sinápticos.” (DE SOUZA et a., 2003). A informação é propagada para os outros neurônios até chegar à camada de saída.

De Souza et al. (2003) explicam que uma rede perceptron de múltiplas camadas é um conjunto de neurônios em camadas com ilustrada na Figura 5.

Figura 5 – Perceptron de Múltiplas Camadas



Fonte: De Souza et al. (2003)

2.4.4 Avaliação da previsão numérica

Podemos usar métricas para avaliar o sucesso da previsão numérica. Em nosso trabalho vamos utilizar duas métricas para avaliar o algoritmo de estimativa do número de

veículos. As métricas usadas na avaliação são: Erro Médio Absoluto (EMA) e Raiz do Erro Quadrático Médio (REQM). Sejam p_1, p_2, \dots, p_n os valores previstos, e a_1, a_2, \dots, a_n os valores observados. O Erro Médio Absoluto (Mean Absolute Error) é definido como (WITTEN; FRANK; HALL, 2011):

$$EMA = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n}$$

A Raiz do Erro Quadrático Médio (Root Mean Squared Error) é definida como:

$$REQM = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}}$$

3 DESENVOLVIMENTO

Esse trabalho apresenta quatro etapas de desenvolvimento: aquisição dos vídeos, cálculo do somatório do percentual de pixels, rotulação dos vídeos, e a estimativa do número de veículos (exceto motos). Nas seções a seguir, cada etapa do desenvolvimento é descrito detalhadamente.

3.1 Aquisição dos vídeos

A captura dos vídeos foi realizada pelo aparelho Sistema de Monitoramento de Trânsito com Câmera Digital (SMTCD), também conhecido como câmera de detecção de infrações de trânsito, da empresa Fotosensores Tecnologia Eletrônica Ltda. A empresa Fotosensores forneceu 17 vídeos do seu aparelho SMTCD, que fica posicionado na margem da Avenida Rogaciano Leite na cidade de Fortaleza, em diferentes horários do dia e da noite. O conjunto dos 17 vídeos é composto por 8 vídeos diurno e 9 vídeos noturno.

A câmera do aparelho SMTCD não apresenta um campo de visão amplo, por estar muito próxima da via. A Figura 6 ilustra o campo de visão do SMTCD durante a noite. Na Figura 7 ilustra a visão durante o dia.

Figura 6 – Frame de vídeo durante a noite



Fonte: Fotosensores.

Figura 7 – Frame de vídeo durante o dia



Fonte: Fotosensores.

Esses vídeos foram utilizados para realização desse estudo. Neste trabalho a análise do congestionamento é feita tanto no período do dia como no período da noite. A seguir, informações sobre os vídeos cedidos:

- a) Formato MPEG;
- b) Resolução de 720x480;
- c) 24 frames por segundo;
- d) Duração de 10 minutos cada.

3.1.1 Qualidade dos vídeos das câmeras de detecção de infrações

Na literatura, os vídeos das câmeras de vigilância apresentam baixa qualidade de imagem. Buch, Velastin e Orwell (2011) afirmam que a qualidade dos vídeos de câmeras de vigilância é geralmente pobre, e a gama de condições operacionais (por exemplo, noite e tempo instável) requer técnicas robustas.

Vídeos com baixa qualidade de imagens contribuem para diminuição da capacidade de isolar características do veículo. As condições ambientais dos cenários afetam também a qualidade dos vídeos, tais como: reflexo da luz do sol, chuva, poeira, fumaça e neblina. Jain, Sharma e Subramanian (2012) argumentam que as câmeras de baixa qualidade para monitoramento do tráfego sofrem com três problemas principais:

- a) Baixa resolução da câmera, resultando em imagens com ruídos;
- b) O tráfego dos veículos limita o campo de visão da câmera;
- c) Iluminação de diversas fontes resulta na distorção da imagem dificultando na identificação dos veículos.

Esse trabalho tem como base o artigo apresentado por Jain, Sharma e Subramanian (2012). Entretanto, a posição da câmera de detecção de infrações da empresa Fotosensores utilizada para gravar os vídeos não é adequada. Mandellos, Keramitsoglou e Kiranoudis (2011) indicam que a câmera deve ser instalada cerca de 10-15m ou mais acima do nível da estrada para minimizar o efeito da oclusão. Ou seja, quando um veículo de grande porte, como o caso de um caminhão ou um ônibus, passa pela frente de outro automóvel, este veículo fica obstruído. A Figura 8 e a Figura 9 ilustram a oclusão provocada por um caminhão durante o dia.

Figura 8 – Veículos antes da oclusão



Fonte: Elaborado pelo autor.

Figura 9 – Oclusão causada pela passagem de um caminhão



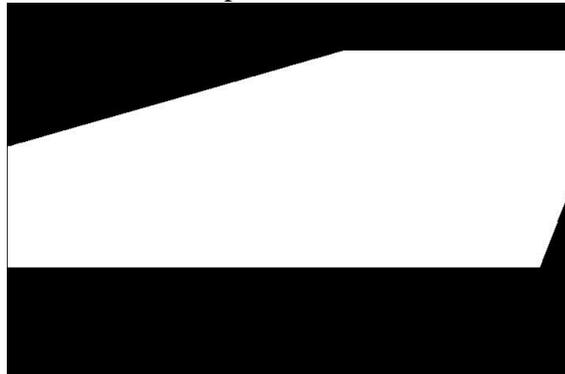
Fonte: Elaborado pelo autor.

Outro problema causado pela curta distância da câmera também pode ser observado nas Figuras 8 e 9. Trata-se da variação de luminosidade da cena causada pelo caminhão baú. Observe que o automóvel está mais escuro na Figura 9, quando comparado com a Figura 8.

3.1.2 Definição da área de marcação do asfalto

Na área do asfalto definimos um polígono, seguindo a abordagem de Jain, Sharma e Subramanian (2012). Essa marcação do asfalto tem como finalidade destacar a área de interesse, e descartar o restante da imagem. Esta marcação serve de máscara para os algoritmos de contagem de pixels cinza/branco, como podemos observar nas Figuras 10 e 11.

Figura 10 – Máscara para marcar o asfalto no frame.



Fonte: Elaborado pelo autor.

Figura 11 – Imagem da via antes e depois da aplicação da máscara.



Fonte: Elaborado pelo autor.

3.2 Cálculo do somatório do percentual de pixels

Para poder estimar o número de veículos em um frame é necessário algum tipo de característica (atributo) da imagem que seja influenciada pela presença de carros. Neste trabalho faz-se uso de dois atributos, um para os vídeos do período diurno e outro para o período noturno. Nas seções a seguir descrevemos como são extraídos esses atributos dos frames dos vídeos do dia e da noite.

3.2.1 Cálculo do somatório do percentual de pixels cinza em cada linha

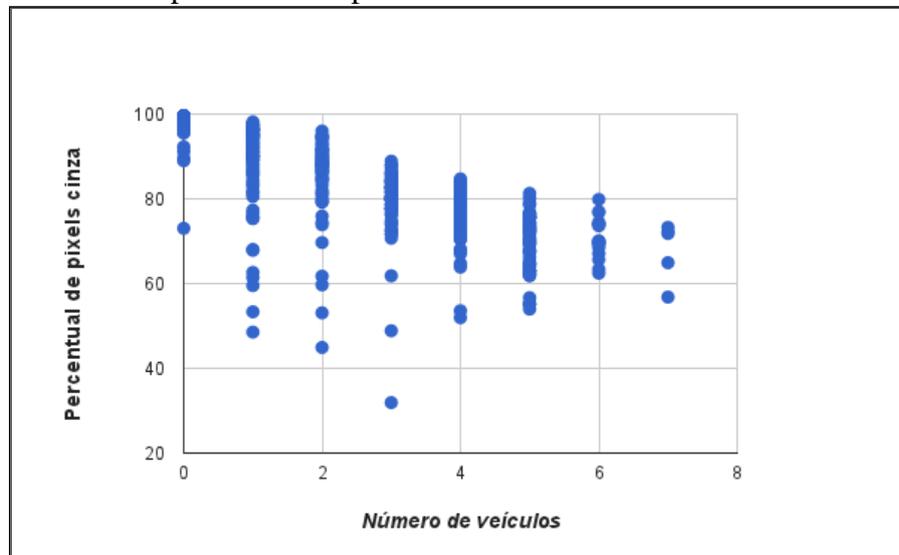
Nesta etapa é gerado um atributo a partir dos frames de vídeos diurnos. Esse atributo é nomeado como sendo o percentual de pixels cinza em cada linha. A intensidade do nível de cinza que representa o asfalto é considerada aqui como sendo na faixa de 70 até 190. Este intervalo foi obtido empiricamente pela avaliação dos vídeos disponíveis. Com este intervalo é possível estimar o percentual de pixels que provavelmente são do asfalto, portanto não estão encobertos por veículos.

Foi desenvolvido um algoritmo com a função de automatizar o cálculo deste atributo. A seguir, os passos deste algoritmo:

- a) Converter o frame atual do vídeo para escala de cinza;
- b) Aplicar a máscara no frame atual na escala de cinza (Figura 10);
- c) Para cada linha da região marcada do frame, calcular o percentual de pixels com intensidade de cinza no intervalo de 70 até 190;
- d) Realizar o somatório dos percentuais de pixels cinza;
- e) E por último, dividir este somatório pelo número de linhas (média aritmética).

Na Figura 12 é ilustrado o gráfico da rotulação dos frames de vídeo e sua relação entre números de veículos e o percentual de pixels cinza.

Figura 12 – Gráfico do percentual de pixels versus o número de veículos dos vídeos diurnos



Fonte: Elaborado pelo autor.

3.2.2 Cálculo do somatório do percentual de pixels branco em cada linha

Para os frames de vídeos noturnos, o atributo é chamado de percentual de pixels brancos em cada linha. O cálculo desse atributo é similar ao cálculo do atributo dos frames do período do dia. A diferença é que ao invés de contar os pixels que possuem intensidade de cinza semelhante à intensidade de cinza do asfalto, agora contamos os pixels brancos originados pelo reflexo dos faróis dos veículos. Da mesma maneira, com este atributo é possível realizar uma estimativa do número de veículos presentes no frame.

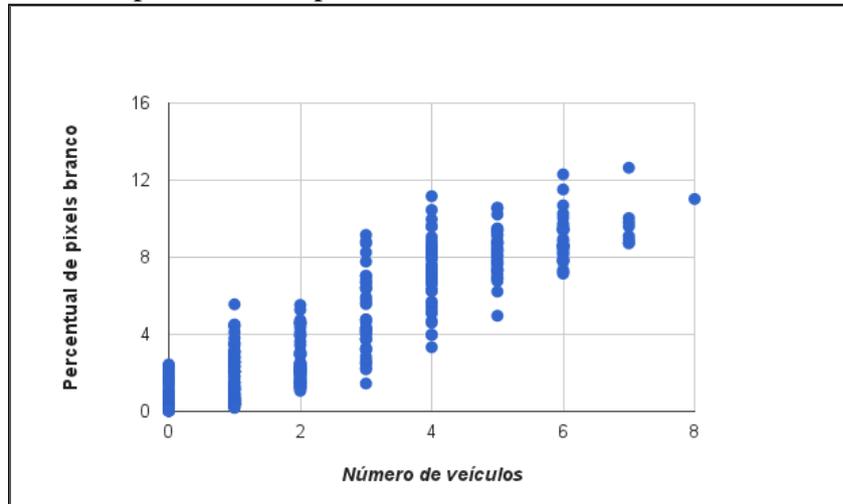
Também foi desenvolvido um algoritmo para realizar o cálculo desse atributo. A seguir, as atividades desempenhadas pelo algoritmo que realiza o cálculo do somatório do percentual de pixels brancos:

- a) Converter o frame atual do vídeo para escala de cinza;
- b) Aplicar a máscara no frame atual na escala de cinza (Figura 10);
- c) Aplicar threshold binário no frame atual: todo pixel com intensidade de cinza acima de 240 passou a ter intensidade 255, e os pixels restantes ficam com intensidade 0 (parâmetro do threshold obtido empiricamente);
- d) Para cada linha da região marcada do frame, calcular o percentual de pixels brancos;

- e) Realizar o somatório dos percentuais de pixels brancos;
- f) E por último, calcular a média aritmética do somatório dos percentuais.

Na Figura 13 é ilustrado o gráfico da rotulação dos frames de vídeo e sua relação entre números de veículos e o percentual de pixels branco.

Figura 13 – Gráfico do percentual de pixels versus o número de veículos dos vídeos noturnos



Fonte: Elaborado pelo autor.

3.3 Rotulação dos vídeos

A empresa Fotosensores Tecnologia Eletrônica Ltda disponibilizou um conjunto de 17 vídeos. Esse conjunto de vídeos foi subdividido em outros dois conjuntos, um referente aos 8 vídeos gravados durante o dia e outro para os 9 vídeos gravados no período da noite. Depois de feita essa separação dos vídeos, a próxima etapa é a realização da rotulação de todos os vídeos.

O processo de rotulação dos vídeos é um procedimento realizado por uma pessoa (rotulador). O rotulador observa cada frame do vídeo e registra o número de veículos presentes (frações de veículos não são computadas na rotulação). O objetivo da rotulação dos frames de vídeos é permitir que se verifique o erro da previsão do número de veículos no frame.

Para facilitar a rotulação dos frames de vídeos do período do dia e da noite, foi implementada uma aplicação que auxilia nesse procedimento. Com o auxílio desta aplicação, ao visualizar o número de carros em um frame, o rotulador pode indicar se houve incremento ou decremento de veículo do frame anterior para o frame atual.

Ao final do processo de rotulação, o algoritmo gera um arquivo texto com as informações da rotulação. As informações são: número do frame e o número de veículos observados no frame.

Esta aplicação de rotulação foi desenvolvida na linguagem C++, utilizando a biblioteca de visão computacional OpenCV para manipulação dos frames. Nesse trabalho foi decidido que a rotulação dos frames é feita a cada 10 segundos, pelo fato de haver muita semelhança entre frames próximos.

Para o treinamento e avaliação do algoritmo de classificação do número de veículos, realizamos uma separação dos vídeos em um conjunto de treino e um conjunto de teste. Inicialmente ordenamos os vídeos com base em seu horário de gravação. Em seguida, atribuímos os vídeos de forma alternada entre os conjuntos de treino e de teste. Desta forma, não testamos o algoritmo de classificação com frames de vídeos utilizados no treinamento. Note que o número de veículos em frames vizinhos são altamente correlacionados. Por outro lado, a atribuição alternada de vídeos permite que o treinamento ocorra com frames coletados em vários horários do dia/noite, e o mesmo ocorra com os frames de teste.

3.4 Estimativa do número de veículos

A estimativa do número de veículos nos frames de vídeos é a etapa final deste trabalho. Nesta etapa, os arquivos do conjunto de treino e de teste de ambas as abordagens são fornecidos aos algoritmos de classificação. Os algoritmos utilizados são o IBk, Regressão Linear e a rede neural Perceptron de Múltiplas Camadas (Multilayer Perceptron - MLP). A ferramenta WEKA foi utilizada, pois fornece implementações desses algoritmos.

O IBk é executado com os valores padrão dos parâmetros do WEKA, exceto o número de vizinhos mais próximos (KNN). O algoritmo de Regressão Linear foi executado com parâmetros padrão. E MLP também foi executado com parâmetros padrão do WEKA, exceto pela variação da taxa de aprendizagem. Os resultados dos algoritmos são apresentados na Seção 4.

4 RESULTADOS

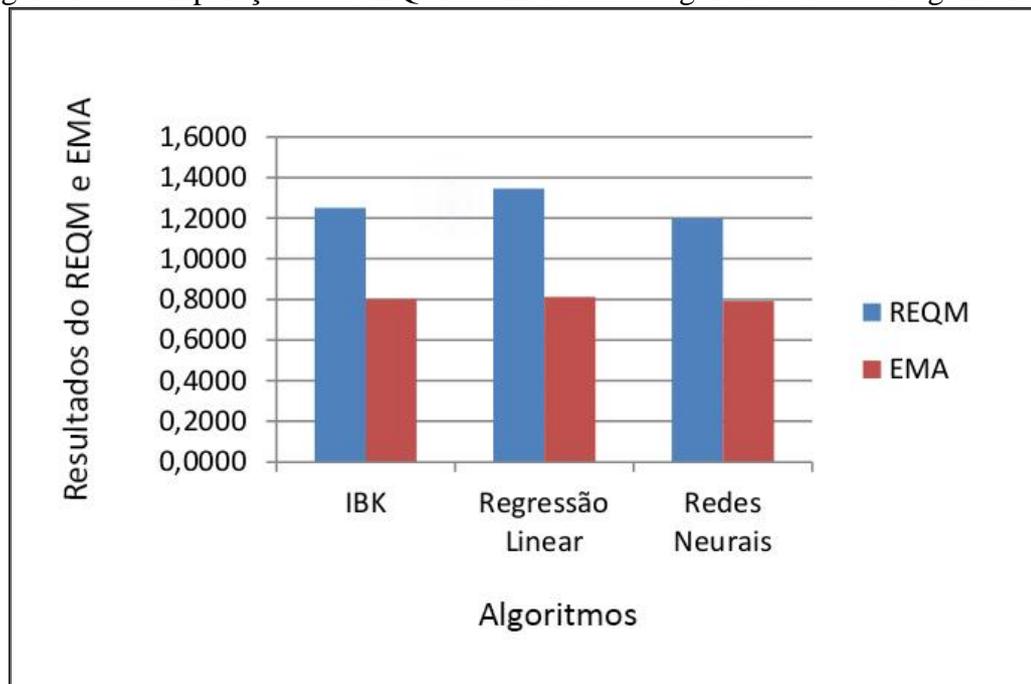
Nesta seção são apresentados os resultados obtidos com os algoritmos IBk, Regressão Linear e a rede neural Perceptron de Múltiplas Camadas (Multilayer Perceptron - MLP). A abordagem de estimativa do número de veículos noturno forneceu resultado

satisfatório, com uma taxa de erro de aproximadamente meio veículo, quando empregamos a Regressão Linear. Já a abordagem de estimativa do número de veículos diurno obteve uma taxa de erro de aproximadamente um veículo, quando empregamos o MLP.

4.1 Resultados experimentais para os vídeos diurnos

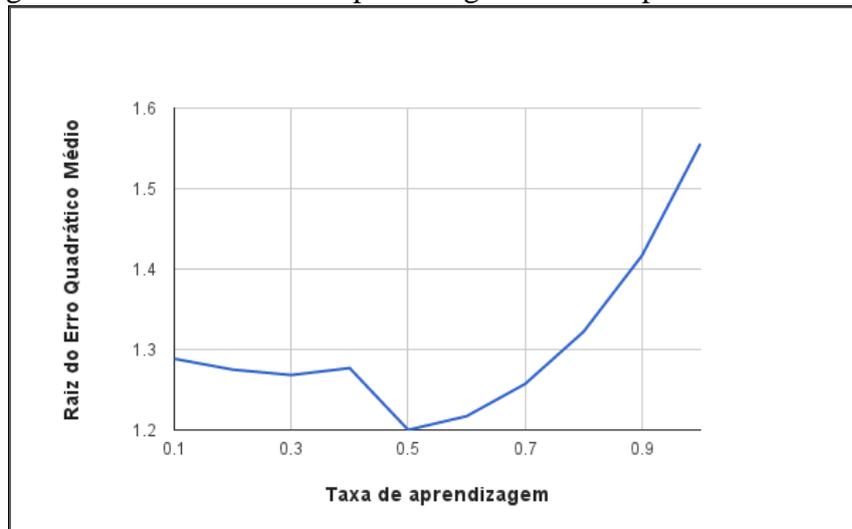
O melhor resultado alcançado foi obtido utilizando a rede neural Perceptron de Múltiplas Camadas (Multilayer Perceptron – MLP), como ilustra a Figura 14. A taxa de aprendizado do MLP foi variada até obter o melhor resultado das métricas da previsão numérica. Esta taxa de aprendizagem é um dos parâmetros da implementação do MLP disponível no WEKA. A rede neural obteve 1.2001 na métrica Raiz do Erro Quadrático Médio (REQM), e 0.7917 na métrica Erro Médio Absoluto (EMA). Esses resultados foram alcançados com a taxa de aprendizagem de 0.5. A Figura 15 ilustra os valores utilizados para taxa de aprendizagem.

Figura 14 – Comparação dos REQM's e EMA's dos algoritmos na abordagem diurna



Fonte: Elaborado pelo autor.

Figura 15 – Melhor taxa de aprendizagem do MLP para vídeos diurnos

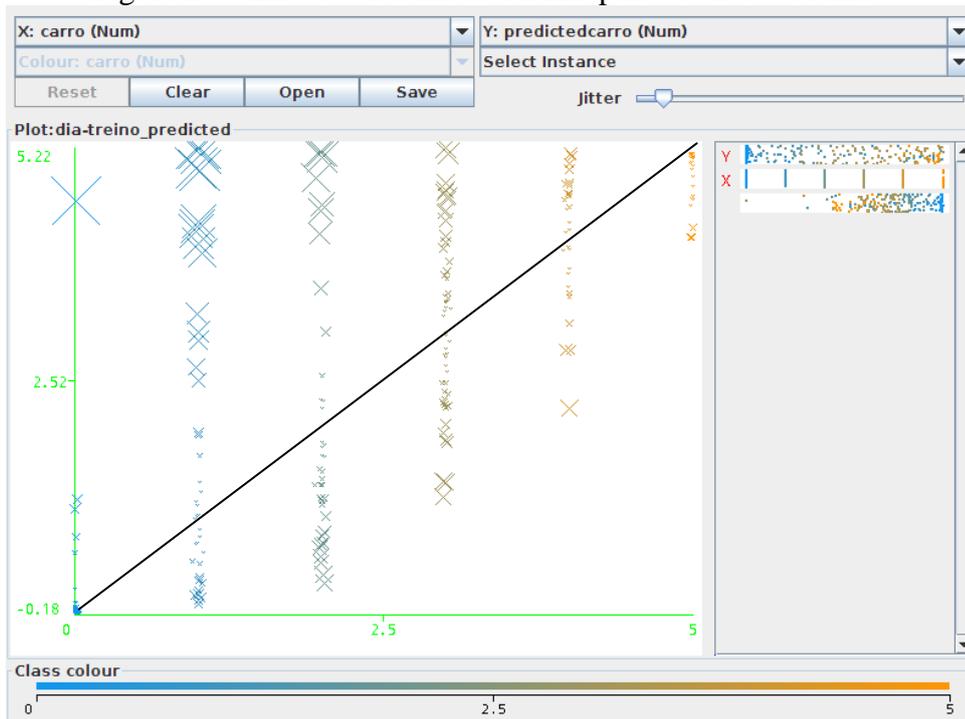


Fonte: Elaborado pelo autor.

Os resultados obtidos não foram melhores devido à proximidade entre a câmera e a via, provocando muita oclusão de veículos. A Figura 16 mostra um gráfico com os erros do classificador, onde o eixo x é a quantidade observada de veículos no frame, e o eixo y é a quantidade de veículos previstos para o frame. Pontos acima da bissetriz dos eixos x e y informam que o classificador estimou mais veículos que os observados. E pontos abaixo da bissetriz informam que o classificador estimou menos veículos do que observados.

Podemos observar na Figura 16 uma tendência do classificador de prever mais veículos que os observados (pontos acima da bissetriz), isto é causado pela ocorrência de oclusão quase total da imagem provocada por grandes veículos (caminhões, ônibus), como ilustrados na Figura 17, pois neste caso o rotulador contou apenas um veículo, e o classificador julgou que haveriam muitos veículos devido à baixa ocorrência de pixels na intensidade de cinza do asfalto. Na Figura 18 temos uma situação que provoca pontos abaixo da bissetriz, que são os frames com carros parcialmente fora da máscara, ou seja, carros parcialmente fora da região considerada. Neste caso temos muitos pixels na intensidade de cinza do asfalto, e muitos carros considerados pelo rotulador.

Figura 16 – Erro do classificador MLP para os vídeos diurnos



Fonte: Elaborado pelo autor.

Figura 17 – Erros do MLP acima da reta da bisettriz



Fonte: Elaborado pelo autor.

Figura 18 – Erros do MLP abaixo da reta da bisettriz

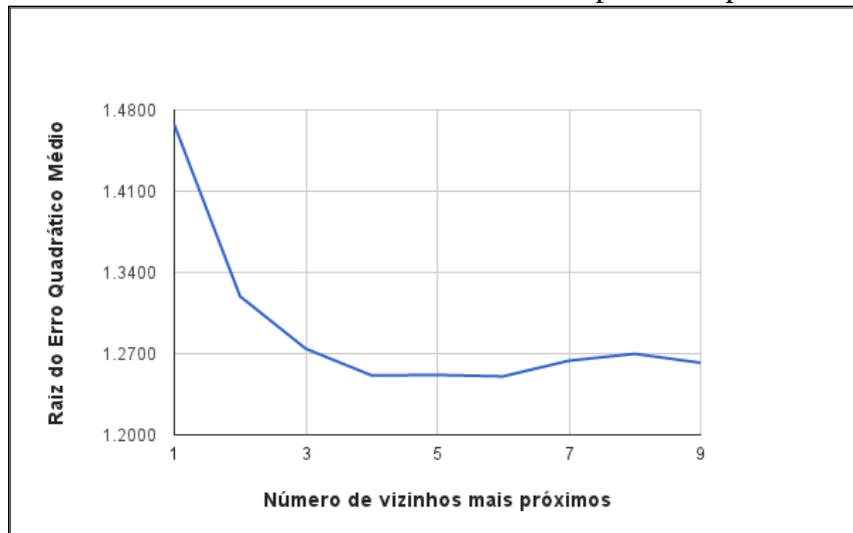


Fonte: Elaborado pelo autor.

A regressão linear obteve o pior resultado entre os algoritmos. O melhor resultado para a Regressão Linear foi de 1.3456 de REQM e de 0.813 para o EMA.

O algoritmo IBk alcançou seu melhor resultado utilizando 5 vizinhos mais próximos. Foram testados vários valores inteiros para o número de vizinhos mais próximos. O resultado das métricas com 4 vizinhos mais próximos foi de 1.2512 para REQM e 0.8004 de EMA. A Figura 19 ilustra os valores de vizinhos mais próximos que foram utilizados.

Figura 19 – Escolha do melhor número de vizinhos mais próximos para os vídeos diurnos

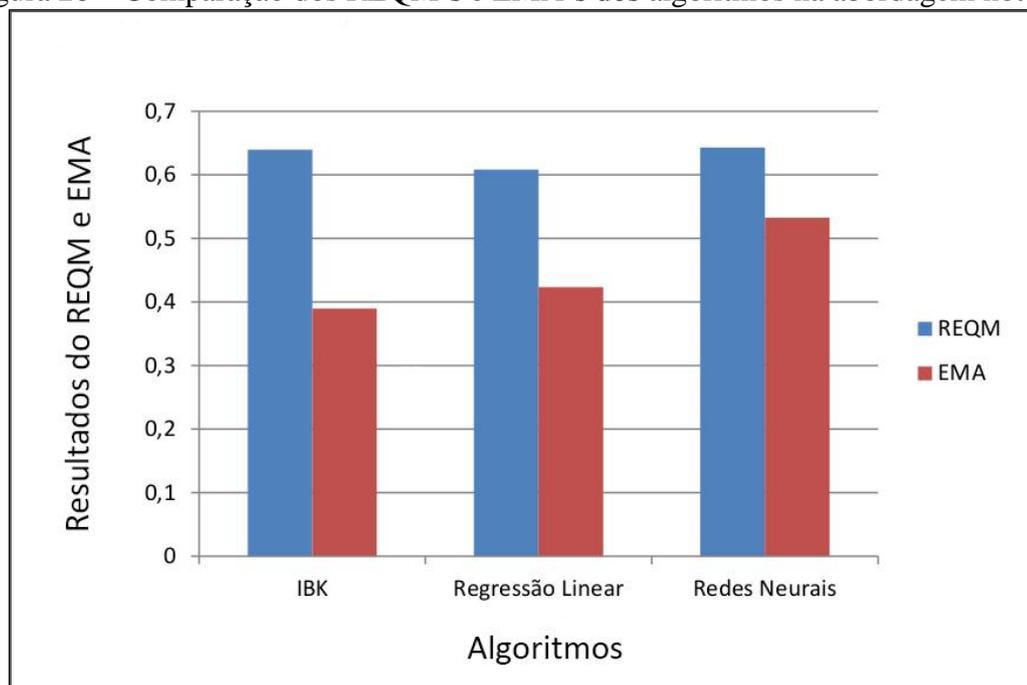


Fonte: Elaborado pelo autor.

4.2 Resultados experimentais para os vídeos noturnos

Os resultados para a Regressão Linear e o IBk foram muito próximos, como ilustra a Figura 20. O IBk obteve um EMA um pouco melhor, mas o REQM foi um pouco pior que o obtido pela Regressão Linear.. Os resultados da Regressão Linear foram: 0.608 para REQM e de 0.4235 para EMA.

Figura 20 – Comparação dos REQM's e EMA's dos algoritmos na abordagem noturna



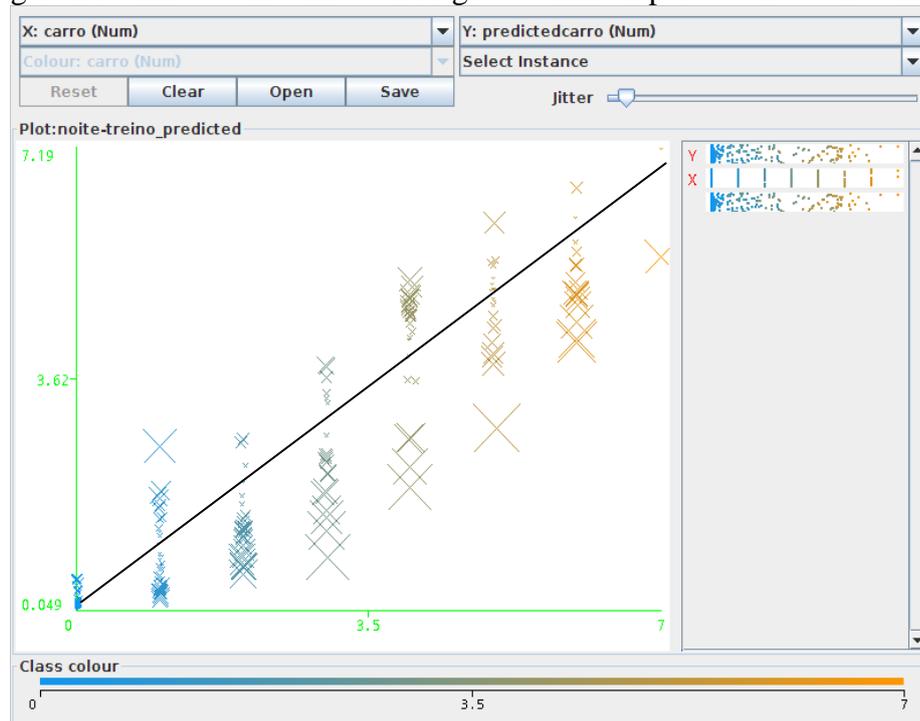
Fonte: Elaborado pelo autor.

Note que o rotulador utilizou apenas números inteiros para indicar a quantidade de veículos, portanto frames com veículos parcialmente na cena produzem um erro médio de meio carro, que não pode ser evitado pelo classificador. Isto permite concluir que o classificador para as imagens noturnas atingiu resultados próximos ao mínimo esperado para o classificador. Por outro lado, temos nos conjuntos de treino e teste 15.59% de frames com 0 veículo e 44.44% de frames com 1 veículo. Ambos equivalem a 60% do conjunto de treino e de teste, o que faz com que um REQM de 0.608 e um EMA de 0.4235 não sejam o ideal.

Na Figura 21 temos um gráfico que permite identificar como estão distribuídos os erros de previsão. No eixo x temos o número de veículos observado no frame, e no eixo y temos a previsão de veículos no frame.

Pontos acima da bissetriz informam que o classificador estimou mais veículos que os observados. Podemos observar na Figura 22 um exemplo com um carro na cena, mas com previsão de dois carros pelo algoritmo. E pontos abaixo da bissetriz informam que o classificador estimou menos veículos do que observados, como ocorrido na cena mostrada na Figura 23. Percebemos que a previsão tende a ficar abaixo da observada (abaixo da bissetriz), o que é provocado pela presença de veículos escuros (com pouco reflexo), conforme ilustrado na Figura 23.

Figura 21 – Erro do classificador Regressão Linear para os vídeos noturnos



Fonte: Elaborado pelo autor.

Figura 22 – Erro da Regressão Linear acima da reta da bissetriz



Fonte: Elaborado pelo autor.

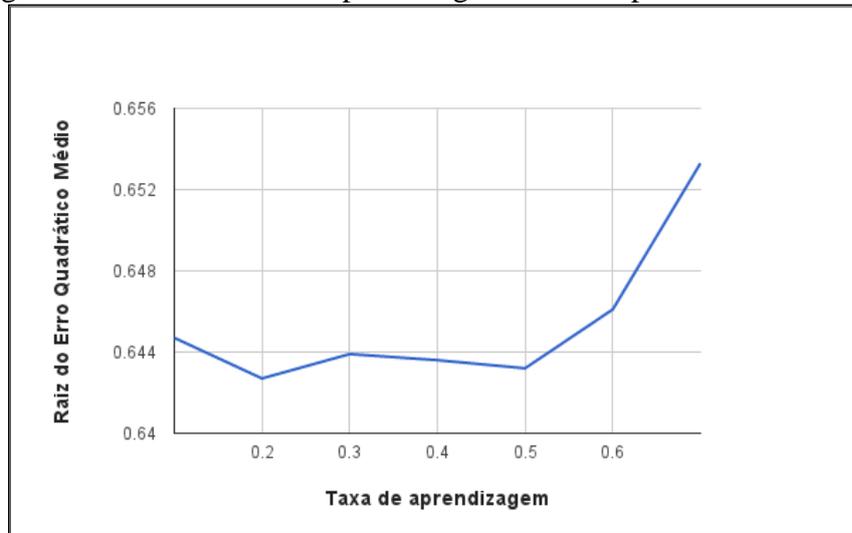
Figura 23 – Erro da Regressão Linear abaixo da reta da bissetriz



Fonte: Elaborado pelo autor.

Para abordagem dos vídeos noturnos, a rede neural Perceptron de Múltiplas Camadas alcançou pior resultado entre os algoritmos. O melhor resultado foi com a taxa de aprendizado de 0.2. A rede neural obteve 0.6427 de REQM e 0.5326 de EMA. A Figura 24 ilustra os valores utilizados para encontrar a melhor taxa de aprendizagem.

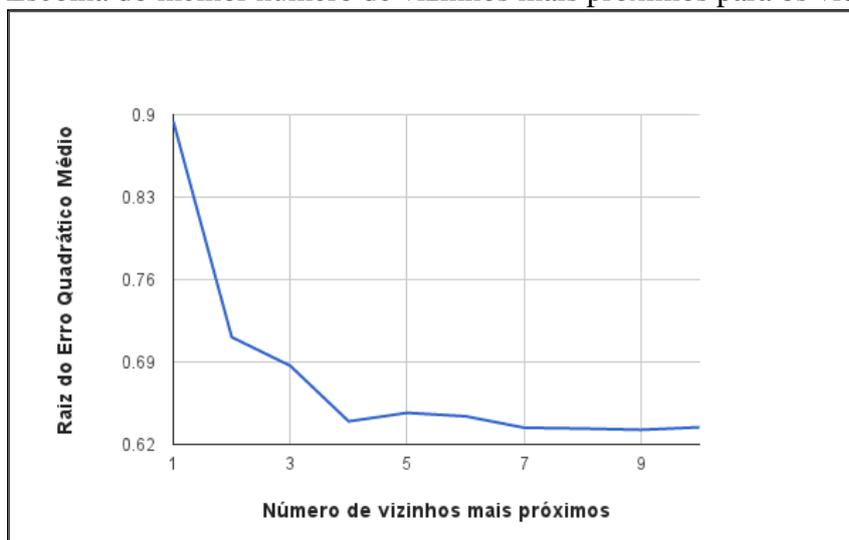
Figura 24 - Melhor taxa de aprendizagem do MLP para vídeos noturnos



Fonte: Elaborado pelo autor.

O IBk alcançou seu melhor resultado utilizando 4 vizinhos mais próximos. O resultado das métricas com um número de 4 vizinhos mais próximos foi de 0.6396 para REQM e 0.3895 de EMA. A Figura 25 ilustra os valores de vizinhos mais próximos que foram utilizados para encontrar o melhor vizinho.

Figura 25 – Escolha do melhor número de vizinhos mais próximos para os vídeos noturnos



Fonte: Elaborado pelo autor.

5 CONSIDERAÇÕES FINAIS

Os congestionamentos estão cada vez mais presentes em nossas vidas influenciando tanto diretamente como indiretamente. Ainda é difícil para os departamentos de trânsito detectar os pontos de lentidão nas vias. Alguns departamentos já utilizam câmeras instaladas em vias com alta taxa de ocorrência de congestionamentos para monitorar as condições de tráfego. Podemos utilizar essas câmeras para realizar detecção automática de congestionamento. Para realizar isso, técnicas da área da visão computacional podem ser empregadas. Entretanto, a detecção de congestionamento utilizando câmeras não é uma tarefa trivial.

O trabalho desenvolvido por Jain, Sharma e Subramanian (2012) serviu como base para tornar possível esse trabalho. O objetivo geral deste trabalho foi produzir um estudo para estimar o número de veículos na via utilizando imagens de vídeos de câmeras de detecção de infrações. Buscamos no decorrer do desenvolvimento deste trabalho atingir os objetivos específicos definidos. Conseguimos rotular manualmente os vídeos diurnos e noturnos. Foi desenvolvido um aplicativo para realizar o cálculo do atributo do percentual de pixels em cada linha para os vídeos diurnos e noturnos. Utilizamos os algoritmos IBk, Regressão Linear e Perceptron de Múltiplas Camadas para estimar o número de veículos. Métricas para a taxa erro da previsão numérica foram utilizadas na avaliação dos algoritmos. Observamos através dos experimentos que para os vídeos diurnos o algoritmo Perceptron de

Múltiplas Camadas obteve o melhor resultado. Para os vídeos noturnos, a Regressão Linear e o IBk alcançaram os melhores resultados.

Na abordagem da estimativa do número de veículos com vídeos diurnos, o resultado atingido com Perceptron de Múltiplas Camadas poderia ser melhor. O resultado alcançado mostra que a curta distância da câmera em relação à via, proporciona o surgimento da oclusão provocada por grandes veículos (caminhões, ônibus, entre outros).

O resultado da abordagem de estimativa do número de veículos com vídeos noturnos também poderia ser melhor. A ocorrência de veículos escuros é um problema nesta abordagem.

Como trabalhos futuros, sugerimos a criação de uma aplicação que utilize o número de veículos previstos para estimar o nível de congestionamento na via. Para identificar o congestionamento, basta verificar se o número de veículos na via ultrapassa aproximadamente 1/3 da capacidade da via.

REFERÊNCIAS

AHA, David W.; KIBLER, Dennis; ALBERT, Marc K. Instance-based learning algorithms. **Machine learning**, v. 6, n. 1, p. 37-66, 1991.

ALVES, P.; RAIJA JUNIOR, A. A. Mobilidade e Acessibilidade Urbanas Sustentáveis: A Gestão da Mobilidade no Brasil. In: **VI Congresso de Meio Ambiente da AUGM. São Carlos. Anais de eventos da UFSCAR 8ª Jornada Científica e Tecnológica**. São Carlo, São Paulo-Brasil. 2009.

BABAK, Olena; DEUTSCH, Clayton V. Statistical approach to inverse distance interpolation. **Stochastic Environmental Research and Risk Assessment**, v. 23, n. 5, p. 543-553, 2009.

BRADSKI, Gary; KAEHLER, Adrian. **Learning OpenCV: Computer vision with the OpenCV library**. O'reilly, 2008.

BRASIL. Lei Federal Nº 12.587, de 3 de janeiro de 2012. Institui as diretrizes da Política Nacional de Mobilidade Urbana e dá outras providências. **Diário Oficial da República Federativa do Brasil**, Brasília, 4 janeiro 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112587.htm>. Acesso em: 25 mai. 2013.

BRINCO, Ricardo. Mobilidade urbana e transporte público: sobre a oportunidade de implantação de sistemas metroviários. **Indicadores Econômicos FEE**, v. 40, n. 1, 2012.

BUCH, Norbert; VELASTIN, Sergio A.; ORWELL, James. A review of computer vision techniques for the analysis of urban traffic. **Intelligent Transportation Systems, IEEE Transactions on**, v. 12, n. 3, p. 920-939, 2011.

COIFMAN, Benjamin et al. A real-time computer vision system for vehicle tracking and traffic surveillance. **Transportation Research Part C: Emerging Technologies**, v. 6, n. 4, p. 271-288, 1998.

CONCI, Aura; AZEVEDO, Eduardo; LETA, Fabiana R. **Computação Gráfica**, v. 2, 2007.

DE SOUZA, B. A. et al. Classificação de faltas via redes neurais artificiais. **V SBQEE**, Aracaju, Brasil, p. 163-8, 2003.

FERREIRA, Aurélio Buarque de Holanda. Miniaurélio século XXI escolar: o minidicionário da língua portuguesa. rev. e ampl. **Rio de Janeiro: Nova Fronteira**, 2001.

FOTOSENSORES TECNOLOGIA ELETRÔNICA LTDA. **Sistema de Monitoramento de Trânsito com Câmera Digital (SMTCD)**. Disponível em: <www.fotosensores.com/SMTCD>. Acessado em: 26 jun. 2013.

GOVERNO prepara redução de IPI para carro até o final de 2013. **Folha de São Paulo**, Brasília, 31 mar. 2013. Disponível em: <<http://www1.folha.uol.com.br/fsp/poder/101420-governo-prorroga-reducao-de-ipi-para-carro-ate-o-final-de-2013.shtml>>. Acesso em: 2 jun. 2013.

JAIN, Vipin; SHARMA, Ashlesh; SUBRAMANIAN, Lakshminarayanan. Road traffic congestion in the developing world. In: **Proceedings of the 2nd ACM Symposium on Computing for Development**. ACM, 2012. p. 11.

KANHERE, Neeraj K. et al. Real-time detection and tracking of vehicle base fronts for measuring traffic counts and speeds on highways. **Transportation Research Record: Journal of the Transportation Research Board**, v. 1993, n. 1, p. 155-164, 2007.

LEI, Manchun et al. A video-based real-time vehicle counting system using adaptive background method. In: **Signal Image Technology and Internet Based Systems, 2008. SITIS'08. IEEE International Conference on**. IEEE, 2008. p. 523-528.

MANDELLOS, Nicholas A.; KERAMITSOGLU, Iphigenia; KIRANOUDIS, Chris T. A background subtraction algorithm for detecting and tracking vehicles. **Expert Systems with Applications**, v. 38, n. 3, p. 1619-1631, 2011.

MARENGONI, Maurício; STRINGHINI, Denise. Tutorial: Introdução à Visão Computacional usando OpenCV. **Revista de Informática Teórica e Aplicada**, v. 16, n. 1, p. 125-160, 2010.

MATOS, Manuel António. Manual Operacional para a Regressão Linear. **Faculdade de Engenharia da Universidade do Porto**, 1995.

MILLA, José Manuel et al. Computer vision techniques for background modelling in urban traffic monitoring. **Urban Transport and Hybrid Vehicles**, 2010.

RESENDE, Paulo Tarso Vilela; SOUSA, Paulo Renato. Mobilidade urbana nas grandes cidades brasileiras: um estudo sobre os impactos do congestionamento. **SIMPOI – SIMPÓSIO DE ADMINISTRAÇÃO DA PRODUÇÃO, LOGÍSTICA E OPERAÇÕES INTERNACIONAIS**, FGV, 2009.

SHEPARD, Donald. A two-dimensional interpolation function for irregularly-spaced data. In: **Proceedings of the 1968 23rd ACM national conference**. ACM, 1968. p. 517-524.

TRÂNSITO lento faz SP perder 1% do PIB do País. **Estadão**, São Paulo, 17 maio 2013. Disponível em: <<http://www.estadao.com.br/noticias/impresso,transito-lento-faz-sp--perder-1-do-pib-do-pais-,1032592,0.htm>>. Acesso em: 1 jun. 2013.

WEKA. **Weka 3: Data Mining Software in Java**. University of Waikato, Hamilton, New Zealand. Disponível em: <<http://www.cs.waikato.ac.nz/~ml/weka/>>. Acessado em: 20 nov. 2013.

WITTEN, Ian H.; FRANK, Eibe; HALL, Mark A. **Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques**. Elsevier, 2011.

ZANDONADE, Patricia; MORETTI, Ricardo. O padrão de mobilidade de São Paulo e o pressuposto de desigualdade. **EURE (Santiago)**, v. 38, n. 113, p. 77-97, 2012.