UNIVERSIDADE FEDERAL DO PARÁ

INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Renato Hidaka Torres

Investigação em Aprendizado de Máquina Sobre Distração dos Motoristas Devido à Utilização de Telefone Celular

Belém

2019

Renato Hidaka Torres

Investigação em Aprendizado de Máquina Sobre Distração dos Motoristas Devido à Utilização de Telefone Celular

Tese de doutorado apresentada no Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas e Naturais como requisito parcial para obtenção do título de doutor em ciência da computação.

Universidade Federal do Pará

Orientador: Gustavo Pessin

Belém
2019

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

T693i Torres, Renato Hidaka
Investigação em Aprendizado de Máquina Sobre Distração dos Motoristas Devido à Utilização de Telefone Celular / Renato Hidaka Torres. — 2019.
182 f. : il. color.

Orientador(a): Prof. Dr. Gustavo Pessin
Tese (Doutorado) - Programa de Pós-Graduação em Ciência da Computação, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, 2019.

1. Distração. 2. Motorista. 3. Smatphone. 4. Sensores. 5. Aprendizado de máquina. I. Título.

CDD 006.3

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

RENATO HIDAKA TORRES

INVESTIGAÇÃO EM APRENDIZADO DE MÁQUINA SOBRE DISTRAÇÃO DE MOTORISTAS DEVIDO À UTILIZAÇÃO DE TELEFONE CELULAR

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal do Pará como requisito para obtenção do título de Doutor em Ciência da Computação, defendida e aprovada em 25/10/2019, pela banca examinadora constituída pelos seguintes membros:

Prof. Dr. Gustavo Pessin
Orientador – PPGCC/UFPA

Prof. Dr. Denis Lima do Rosário
Membro Interno – PPGCC/UFPA

Prof. Dr. Ronnie Cley de Oliveira Alves
Membro Interno – PPGCC/UFPA

Prof. Dr. Geraldo Pereira Rocha Filho
Membro Externo – UnB

Prof. Dr. Orlando Shigeo Ohashi Junior
Membro Externo – UFRA

Profª. Drª. Patrícia Amâncio Vargas
Membro Externo – Heriot-Watt University/UK

Visto:
Prof. Dr. Nelson Cruz Sampaio Neto
Coordenador do PPGCC/UFPA

Dedicado esta tese à minha família.

Agradecimentos

Ao Mestre dos mestres. Obrigado Jesus!

A minha família pelo amor, carinho e apoio incondicional. Em especial à minha esposa Júlia Torres, minha filha Lívia Akemi, meu sobrinho Heitor Gabriel, meus pais Sérgio e Ilda Torres, aos meus avós maternos Maria Piedade e Massayuki Hidaka, aos meus avós paternos, in memorian, Raimundo e Nancy Torres e ao meu tio Walter Dantas. Amo muito vocês!

Ao professor e amigo Gustavo Pessin, por sua orientação e pelo comprometimento com esta pesquisa.

Ao professor e amigo Orlando Ohashi, por sua orientação e pelo comprometimento com esta pesquisa.

Ao professor e amigo Nelson Sampaio Neto, por sua orientação e supervisão no estágio de docência.

Aos voluntários que participaram desta pesquisa. Em especial ao amigo Jailton Wagner por realizar a coleta de dados durante as nossas longas viagens para Paragominas.

Aos professores do programa de Pós-Graduação de Ciência da Computação da UFPA, pelo conhecimento adquirido nas disciplinas cursadas.

Aos membros da Banca Examinadora, pelas contribuições no trabalho.

"Foco é poder; distração é sofrimento"
(Jô Furlan)

Resumo

A distração de motorista é um problema de escala mundial que tem contribuído para um número cada vez maior de acidentes e mortes no trânsito. Dos motivos que levam a distração, a utilização do telefone celular tem chamado a atenção devido ao crescente aumento de usuários que utilizam enquanto estão dirigindo. Leis e campanhas educativas são políticas públicas adotadas no mundo todo para coibir o manuseio do telefone celular por motoristas. Entretanto, relatórios da Organização Mundial da Saúde demonstram que essas medidas, embora necessárias, não são suficientes, pois os motoristas continuam utilizando os smartphones enquanto estão dirigindo. Esse comportamento demonstra que a necessidade de utilização do smartphone é maior do que a consciência dos riscos que esse ato pode ocasionar. Por esse motivo, pesquisas e soluções tecnológicas estão sendo desenvolvidas para assistir a distração dos motoristas. As soluções desenvolvidas capturam informações e realizam tomadas de decisões para prevenir possíveis falhas humanas. Nesta tese, propõem-se três abordagens de construção de modelos de aprendizado de máquina, cujo objetivo é classificar motoristas e passageiros durante a utilização do telefone celular. Na primeira abordagem, foi desenvolvida uma rede neural convolucional a partir da metodologia de observação do comportamento do condutor. As imagens capturadas por meio de uma câmera acoplada ao veículo foram utilizadas no desenvolvimento e análise do modelo. Na segunda abordagem, foram desenvolvidos diferentes modelos de aprendizado de máquina para classificar motoristas e passageiros durante a escrita de mensagens de texto. Nesse experimento, foi utilizada a metodologia de observação dos efeitos do comportamento do condutor. Dados sensoriais foram coletados e utilizados na construção e validação dos modelos. Por fim, na terceira abordagem, foram desenvolvidos diferentes modelos de aprendizado de máquina para classificar motoristas e passageiros durante a leitura de mensagens/notificações. Nessa abordagem também foi utilizada a metodologia de observação dos efeitos do comportamento do condutor. Os resultados obtidos nesta tese demonstram que as abordagens propostas são eficientes para o desenvolvimento de soluções não intrusivas que ofereçam a possibilidade de segurança e conveniência para os passageiros e motoristas.

Palavras-chave: Distração, Motorista, Smatphone, Sensores, Aprendizado de máquina.

Abstract

Driver distraction is a worldwide problem that climbs the number of road traffic deaths and injuries. Of the distracted reasons, distracted driving by mobile phone is a growing risk factor due to the number of users who perform this secondary event. Laws and education campaigns are public policies adopted worldwide to control cell phone use while driving. However, reports from the World Health Organization show that these measures are necessary, but not sufficient, because drivers continue to use smartphones while driving. This behavior shows that the need to use a smartphone is greater than the damage recognition that this action may cause. For this reason, research and technological solutions are developed to assist driver distraction. Solutions capture information and make decisions to prevent potential human failures. In this thesis, we propose three approaches to building machine learning models, whose objective is to classify drivers and passengers while using the mobile phone. In the first approach, we model a convolutional neural network from the conductor behavior observation methodology. The images captured through a camera attached to the vehicle were used in the development and analysis of the model. In the second approach, we built different machine learning models to classify drivers and passengers in the texting while driving event. In this experiment, we used the methodology for observing the effects of driver behavior. Sensory data were collected and used in the construction and validation of the models. Finally, in the third approach, we model and analyze different machine learning models to classify drivers and passengers in the reading while driving event. In this approach, we also used the methodology of observation of the effects of driver behavior. The results obtained in this thesis show that the proposed approaches are efficient for the development of non-intrusive solutions that offer the possibility of safety and convenience for passengers and drivers.

Keywords: Distraction, Driver, Smatphone, Sensors, Machine Learning.

Sumário

1 INTRODUÇÃO 27
1.1 Definição do problema 32
1.2 Objetivo 33
1.3 Hipótese da tese 34
1.4 Publicações e projetos 35
1.5 Organização da Tese 36

2 REVISÃO SISTEMÁTICA 37
2.1 Motivação 37
2.2 Planejamento 40
2.2.1 Identificação da necessidade da revisão 41
2.2.2 Desenvolvimento do protocolo 41
2.3 Condução 43
2.3.1 Visão geral dos resultados 44
2.3.2 Questão de pesquisa 1 46
2.3.3 Questão de pesquisa 2 50
2.3.4 Questão de pesquisa 3 52
2.3.5 Questão de pesquisa 4 55
2.4 Considerações finais 60

3 SURVEY DO PERFIL DE DISTRAÇÃO DOS MOTORISTAS . . . 63
3.1 Motivação 63
3.2 Condução do survey 66
3.3 Análise dos resultados 68
3.4 Considerações finais 83

4 FUNDAMENTOS DOS MODELOS DE APRENDIZADO DE MÁQUINA 85
4.1 Modelos supervisionados 85
4.1.1 Logistic Regression 85
4.1.2 Árvore de Decisão 87
4.1.3 Random Forest 88
4.1.4 Máquina de Vetor de Suporte 90
4.1.5 Adaptive Boosting e Gradient Boosting 92
4.1.6 Redes Neurais Convolucionais 94
4.1.6.1 Convolução 94

4.1.6.2 Função de ativação . . . 98
4.1.6.3 Pooling . . . 101
4.2 Métricas de avaliação dos modelos . . . 103
4.2.1 K-fold Cross-Validation . . . 104
4.2.2 Precision, Recall e F1-score . . . 106
4.2.3 Kolmogorov-Smirnov test . . . 108

5 ANÁLISE DA DISTRAÇÃO DO MOTORISTA A PARTIR DA OBSERVAÇÃO DO COMPORTAMENTO DO CONDUTOR . . . 111
5.1 Motivação . . . 111
5.2 Metodologia . . . 113
5.2.1 Preparação da base de dados . . . 114
5.2.2 Desenvolvimento do modelo . . . 117
5.2.3 Avaliação do modelo . . . 119
5.3 Considerações finais . . . 122

6 ANÁLISE DA DISTRAÇÃO DO MOTORISTA DEVIDO À COMPOSIÇÃO DE MENSAGENS DE TEXTO . . . 125
6.1 Motivação . . . 125
6.2 Metodologia . . . 127
6.2.1 Preparação da base de dados . . . 129
6.2.2 Configuração dos parâmetros . . . 131
6.2.3 Treinamento dos modelos . . . 132
6.2.4 Avaliação dos modelos . . . 133
6.2.5 Verificação da significância estatítica . . . 135
6.3 Considerações finais . . . 137

7 ANÁLISE DA DISTRAÇÃO DO MOTORISTA DEVIDO À LEITURA DE MENSAGENS DE TEXTO . . . 139
7.1 Motivação . . . 140
7.2 Metodologia . . . 141
7.2.1 Configuração dos parâmetros . . . 148
7.2.2 Treinamento dos modelos . . . 153
7.2.3 Avaliação dos modelos . . . 155
7.2.4 Verificação da significância estatística . . . 157
7.2.5 Análise dos modelos em função da importância dos atributos . . . 159
7.3 Considerações finais . . . 163

8 CONCLUSÕES . . . 165

REFERÊNCIAS 171

APÊNDICE A – ARTIGOS FILTRADOS NA REVISÃO SISTEMÁTICA 181

Lista de ilustrações

Figura 1. Protocolo da revisão sistemática adaptada de Kitchenam. 40
Figura 2. Relação entre método e tipo de estudo pesquisados. 44
Figura 3. Quantidade de publicações filtradas por repositório. 45
Figura 4. Quantidade de publicações por ano. 46
Figura 5. Países que publicaram os trabalhos selecionados nesta revisão. 46
Figura 6. Quantidade de publicações por método de estudo e repositório. 48
Figura 7. Quantidade de publicações por tipo de estudo e repositório. 50
Figura 8. Mapa de nuvem dos sensores utilizados nos artigos filtrados. 53
Figura 9. Informações socioeconômicas do estado do Pará (CNT, 2018). 65
Figura 10. Informações gerais dos participantes. 68
Figura 11. Análise da distração por tipo de veículo. 69
Figura 12. Análise da distração por gênero. 70
Figura 13. Análise dos tipos de distração filtrados por idade. 71
Figura 14. Análise da distração dos motoristas devido à realização de chamadas telefônicas filtradas por gênero e idade. 73
Figura 15. Análise da distração dos motoristas devido à escrita de mensagens de texto filtradas por gênero e idade. 74
Figura 16. Análise da distração dos motoristas devido ao envio de mensagens de áudio filtradas por gênero e idade. 75
Figura 17. Análise da distração dos motoristas devido à leitura de mensagens de texto filtradas por gênero e idade. 76
Figura 18. Análise da distração dos motociclistas devido à realização de chamadas telefônicas filtradas por gênero e idade. 77
Figura 19. Análise da distração dos motociclistas devido à escrita de mensagens de texto filtradas por gênero e idade. 78
Figura 20. Análise da distração dos motociclistas devido ao envio de mensagens de áudio filtradas por gênero e idade. 79
Figura 21. Análise da distração dos motociclistas devido à leitura de mensagens de texto filtradas por gênero e idade. 80
Figura 22. Respostas da 9^a pergunta do *survey*: Estando com o veículo em movimento, você utiliza o celular com maior frequência por qual motivo? . . 82
Figura 23. Respostas da 10^a pergunta do *survey*: Você conhece pessoas que utilizam o telefone celular enquanto estão dirigindo ? 82
Figura 24. Arquitetura do modelo de regressão logística (RASCHKA; MIRJALILI, 2017). 86

Figura 25. Exemplo de construção de uma árvore de decisão (RASCHKA; MIRJALILI, 2017). 88
Figura 26. Exemplo de vetor de suporte de uma SVM (RASCHKA; MIRJALILI, 2017). 90
Figura 27. Ilustração da redefinição dos pesos das amostras utilizadas no algoritmo de Boosting (RASCHKA; MIRJALILI, 2017). 92
Figura 28. Convolução exemplificada pelo lançamento de bola. Fonte: http://colah.github.io/posts/2014-07-Understanding-Convolutions 95
Figura 29. Exemplo de convolução bidimensional. Fonte: (GOODFELLOW; BENGIO; COURVILLE, 2016) 96
Figura 30. Exemplo da relação peso e valor de entrada em uma rede totalmente conectada. 97
Figura 31. Exemplo da relação peso e valor de entrada, em uma rede convolucional. Fonte: http://colah.github.io/posts/2014-07-Conv-Nets-Modular/ . . . 97
Figura 32. Abstração das funções de *kernel* de uma CNN. Fonte: (CHOLLET, 2017) 98
Figura 33. Comportamento da função sigmoid. 99
Figura 34. Comportamento da função tangente hiperbólica. 100
Figura 35. Comportamento da função ReLu. 101
Figura 36. Aproximação da função quadrática pela função ReLu. 101
Figura 37. Exemplo de variação de tamanho, translação e rotação em imagens digitais. Fonte: (CHAUDHARI; GARG; KOTECHA, 2017). 102
Figura 38. Comportamento da função *max pooling* no espaço bidimensional. . . . 103
Figura 39. Ilustração do procedimento de validação cruzada de k-fold (RASCHKA, 2018). 104
Figura 40. Exemplo de uma matriz de confusão. 106
Figura 41. Exemplo da busca pela distância vertical máxima. 109
Figura 42. Tipos de distração do motorista analisados. Imagens da base de dados utilizada no experimento. 114
Figura 43. Comportamento esperado na direção segura. Imagem da base de dados utilizada no experimento. 114
Figura 44. Metodologia utilizada na aquisição dos dados e extração das características do primeiro experimento. 115
Figura 45. Pré-processamento realizado na base de imagens. 115
Figura 46. Fluxo do experimento. 116
Figura 47. Possíveis comportamentos durante a generalização do problema. Fonte: (RASCHKA; MIRJALILI, 2017) 117
Figura 48. Arquitetura da CNN desenvolvida no segundo experimento. 118
Figura 49. Curva de aprendizado do modelo em função do erro. 120

Figura 50. Matriz de confusão construída com os dados separados, para testar a capacidade de generalização do modelo. 120
Figura 51. Aplicativo construído para capturar os dados sensoriais. 127
Figura 52. Metodologia utilizada na aquisição dos dados e extração das características do segundo experimento. 128
Figura 53. Fluxo do segundo experimento. 130
Figura 54. Curva de aprendizado dos modelos avaliados pela validação cruzada de 10-folds. 133
Figura 55. Matriz de confusão construída a partir dos dados de teste. 134
Figura 56. Construção dos vetores utilizados no teste de significância. 135
Figura 57. Estimativa de densidade dos modelos. 136
Figura 58. Metodologia utilizada na aquisição dos dados e extração das características. 141
Figura 59. Aplicativo construído para aquisição dos dados: Na primeira tela (**canto superior direito**), o usuário deve digitar o nome do arquivo de *log*. Os sensores são ativados quando o usuário pressiona o botão START. Para cada mensagem lida, o usuário deve pressionar os botões BEGIN/END para definir a janela de dados do sensor. 143
Figura 60. Viagens feitas por voluntários para coletar dados sensoriais. Figuras (**a**-**c**) foram viagens realizadas por motoristas e as Figuras (**d**-**f**) foram viagens realizadas por passageiros. 145
Figura 61. Distribuição do tamanho das janelas coletadas pelos usuários. Cada janela de dados representa o intervalo que o usuário levou para pressionar os botões de BEGIN/END do aplicativo. 147
Figura 62. Gráfico violino da distribuição dos atributos construídos a partir dos dados sensoriais. 149
Figura 63. Fluxo do terceiro experimento. 151
Figura 64. Arquitetura da CNN desenvolvida no terceiro experimento. 153
Figura 65. *Boxplot* da primeira fase do experimento. 155
Figura 66. Gráfico radar da fase de teste (**esquerda**) e da fase de validação (**direita**). 155
Figura 67. Desempenho dos modelos com os atributos coletados pelo acelerômetro. 157
Figura 68. Desempenho dos modelos com os atributos coletados pelo giroscópio. . 157
Figura 69. Desempenho dos modelos com os atributos coletados pelo magnetômetro. 158
Figura 70. Importância dos atributos computada pelo modelo *Random Forest*. . . 160
Figura 71. Gráfico radar com resultados da terceira avaliação. 161

Lista de tabelas

Tabela 1. Lista de quantidade de publicações por país 47
Tabela 2. Modelos de aprendizado de máquina utilizados pelos artigos desta revisão. 55
Tabela 3. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos SVM. 57
Tabela 4. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos de redes neurais. 58
Tabela 5. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos de regressão logística. 59
Tabela 6. *Surveys* filtrados na revisão da literatura apresentada no Capítulo 2. . 64
Tabela 7. Registro de acidente de trânsito nas rodovias federais do estado do Pará em 2018 (PRF, 2018). 66
Tabela 8. Questões que foram construídas para o *survey*. 67
Tabela 9. Matriz de confusão de um classificador hipotético. 105
Tabela 10. Distribuição das amostras por classes. 113
Tabela 11. Distribuição das amostras, por classes, dos conjuntos de treinamento e validação. 115
Tabela 12. Computação das métricas *precision*, *recall*, *F1-score*. 122
Tabela 13. Atributos da base de dados 129
Tabela 14. Distribuição das amostras por classes dos conjuntos de treinamento e validação. 129
Tabela 15. Configuração dos parâmetros do *grid search* do segundo experimento. . 131
Tabela 16. Configuração dos hiperparâmetros da árvore de decisão 132
Tabela 17. Métricas calculadas a partir da matriz de confusão. 135
Tabela 18. Resultado do teste Kolmogorov-Smirnov aplicado 137
Tabela 19. Número de mensagens lidas por cada voluntário. A subdivisão da coluna do condutor ou passageiro indica que o voluntário realizou mais de uma viagem com esse perfil. 144
Tabela 20. Lista de sensores utilizados durante a coleta de dados. 146
Tabela 21. Lista de atributos extraídos do conjunto de dados. 147
Tabela 22. Configuração dos parâmetros do *grid search* do terceiro experimento. . 150
Tabela 23. Subconjuntos de Ω usados como dados de entrada na segunda fase do experimento. 156
Tabela 24. Teste de hipótese *t-Student's* aplicado aos modelos. 158
Tabela 25. Teste de hipótese t-Student's aplicado aos modelos construídos nas três fases do experimento. 160
Tabela 26. Parte 1: Lista de artigos filtrados. 181

Tabela 27. Parte 2: Lista de artigos filtrados. 185
Tabela 28. Parte 3: Lista de artigos filtrados. 188

Lista de abreviaturas e siglas

ABNT	Associação Brasileira de Normas Técnicas
OMS	Organização Mundial da Saúde
ADAS	Sistema Avançado de Assistência ao Motorista
CNN	Rede Neural Convolucional
GPS	Sistema de Posicionamento Global
IMU	Unidade de medição inercial
SVM	Support Vector Machine
ANN	Rede Neural Artificial
KNN	K-Nearest Neighbor
RNN	Recurrent Neural Network
ELM	Extreme Learning Machine
PRF	Polícoa Rodoviária Federal

1 Introdução

Em 2008, em uma pesquisa realizada em 179 países, a Organização Mundial de Saúde (OMS) constatou que no ano de 2004 os acidentes de trânsito ocasionaram, aproximadamente, 1.3 milhões de mortes e entre 20 a 50 milhões de feridos. Tais valores ranquearam o acidente de trânsito como a nona principal causa de morte no mundo, sendo esta a única causa, dentre as dez, que não é uma doença. Também foi observado que o acidente de trânsito é a principal causa de morte de pessoas com faixa etária entre 15 e 29 anos e a terceira principal causa de morte entre pessoas com faixa etária entre 30 e 44 anos. Segundo a OMS, se nada fosse feito o acidente de trânsito poderia chegar à patamares de 2.4 milhões de mortes por ano, se tornando a quinta principal causa de morte no mundo (CHAN, 2009).

Diante dos dados relatados, a Organização Mundial de Saúde (OMS), em conjunto com as Nações Unidas, lançou em 2010 o programa "Década de Ações para a Segurança no Trânsito - 2011 a 2020" (KI-MOON, 2010). O programa tem o objetivo de reduzir em 50% o número de mortos em decorrência de acidentes de trânsito e salvar 5 milhões de vidas. Para alcançar estes objetivos, a OMS incentiva planos nacionais e regionais que visam atender aos cinco pilares definidos pelo programa, que são: (1) gestão da segurança viária; (2) vias mais seguras e mobilidade; (3) veículos mais seguros; (4) conscientização dos usuários; e (5) resposta ao acidente.

Depois do lançamento do programa, a OMS passou a publicar relatórios com a finalidade de apresentar o progresso das medidas desenvolvidas nos cinco pilares definidos no programa. O relatório também tem o objetivo de divulgar diversas estatísticas em relação à quantidade, características, causas e consequências dos acidentes de trânsito que ocorreram no período. No último relatório publicado em 2018 pela OMS (CHAN, 2018), observou-se que os números não mudaram significativamente desde o lançamento do programa. Pelo contrário, hoje o acidente de trânsito é a principal causa de morte de crianças e adultos jovens entre 5 e 29 anos.

No Brasil, de acordo com os dados do relatório da segurança viária nacional (AMBEV, 2017), em 2015, a taxa de óbitos por 100 mil habitantes atingiu a marca de 19,2%. Esse foi o melhor resultado desde 2004. Por outro lado, o número de feridos aumentou 29% em relação ao ano de 2010. Apesar de ter tido uma redução de 6,3% na taxa de óbitos por 100 mil habitantes, o Brasil (23,4) supera a média mundial (17,4) e fica na frente de países como a China (18,8), Índia (16,6) e Estados Unidos (10,6), que possuem uma população maior do que a brasileira. Em número absoluto, o Brasil (43.075) é o terceiro país com maior número de óbitos decorrentes de acidente de trânsito, ficando

atrás somente da Índia (137.572) e da China (62.945) (CHAN, 2018).

De acordo com os dados apresentados é possível perceber que a segurança viária ainda é um problema mundial e, particularmente, brasileiro. Das medidas já adotadas para reduzir a quantidade de acidentes e mortes nas estradas, destacam-se as leis e as campanhas educativas. Segundo (CHAN, 2018), dos dez países mais populosos do mundo, o Brasil é o que apresenta as leis mais efetivas de segurança viária. As legislações para fiscalizar o uso de bebidas alcoólicas e outras drogas (Lei 13.546, de 19 de dezembro de 2017), transporte de crianças (Resolução CONTRAN 277 de 28 de maio de 2008), uso de capacete (Resolução CONTRAN 453 de 26 setembro de 2013) e cinto de segurança (Lei 9.503, de 23 de setembro de 1997) foram consideradas efetivas pela OMS.

Considerando o cenário brasileiro, nota-se que a existência das leis e das campanhas de conscientização dos motoristas são medidas necessárias, mas não suficientes. Não são suficientes uma vez que se observou um aumento significativo tanto no número de acidentes absolutos (21,6%) quanto na proporção para 100 mil habitantes (15,4%) (AMBEV, 2017). Este cenário é um indicativo de que também são necessárias outras medidas para fortalecer a segurança viária.

Dadas as possibilidades, melhorar a infraestrutura das vias parece ser uma medida promissora para prevenir acidentes e mortes resultantes de acidentes de trânsito. No Brasil, em 2015, 52% dos acidentes em rodovias federais ocorreram em rodovias simples, 41% em rodovias duplicadas e 7% em rodovias múltiplas (PRF, 2015). Este cenário demonstra que, quanto melhor o tipo da via, menor o número de acidentes. No contexto da infraestrutura, a condição e a sinalização das vias também são fatores importantes para garantir a segurança dos usuários. Ademais, outra medida que vem sendo adotada diz respeito ao aperfeiçoamento tecnológico dos veículos. Automóveis equipados com *Global Positioning System* (GPS), *Adaptive Cruise Control*, *Collision Prevention*, *Speed Bump Detection*, *Brake Assist*, *Lane Departure Warning System*, *Electronic Stability Control* e *Automotive Night Vision* estão se tornando realidade no mercado e são conhecidos como *Advance Driver Assistance System* (ADAS). Tecnologias como essas tem a finalidade de interagir com o ambiente e prevenir possíveis falhas humanas. Um ponto ainda negativo é que, devido ao elevado custo, carros equipados com essas tecnologias são, em sua maioria, adquiridas por uma pequena parcela da população (CHAN, 2015).

Em decorrência do alto custo dos ADAS, pesquisadores estão procurando soluções alternativas de menor custo. Vários pesquisadores aproveitaram a consolidação da internet das coisas e estão utilizando os recursos de *smartphones* para desenvolver soluções que interajam com o ambiente e diminuam a probabilidade de falhas humanas. Trabalhos como os de (ZADEH; GHATEE; EFTEKHARI, 2017; MUKHERJEE; MAJHI, 2016; DEVAPRIYA; BABU; SRIHARI, 2015; YEN; HUO; SUN, 2014) são exemplos de *Advance Driver Assistance System* que foram desenvolvidos utilizando os sensores de um *smartphone*.

Além do desenvolvimento de ADAS, os *smartphones* também estão sendo utilizados para coletar dados e entender o comportamento do motorista. Análise e reconhecimento do perfil do motorista, sonolência e distração são temas em investigação (MEIRING; MYBURGH, 2015; MUSICANT; LOTAN, 2016; LI et al., 2017; JUNIOR et al., 2017).

No que diz respeito à distração do motorista, este é um tema multidisciplinar que está em constante investigação. De forma geral, a distração de motorista caracteriza qualquer atividade secundária que desvia a sua atenção da atividade primária. Dentre as tarefas secundárias, pode-se destacar a utilização do telefone, comer, beber, conversar com o passageiro, ajustar a rádio ou o dispositivo multimídia, utilização do GPS e a interação com a navegação de bordo. Segundo (CARNEY et al., 2016), a utilização do telefone (11,9%) é a segunda tarefa secundária que possui o maior potencial para distração. No *survey* aplicado por Carney, apenas a conversa com o passageiro (14,6%) apresentou maior potencial de distração do que a utilização do telefone celular. A distração decorrente da utilização do telefone celular pode ser explicada pelos avanços tecnológicos desses dispositivos. O *smartphone* se tornou mais avançado e inclui recursos antes nunca vistos, como conectividade de alta velocidade, mensagens instantâneas, além das redes sociais. Essa nova e poderosa ferramenta muda a maneira como as pessoas interagem e se socializam, e cria o que os psicólogos estão chamando de vício em *smartphones* (MAYA; NAZIR, 2016).

No Brasil, dados do IBGE apontam que os *smartphones* são os dispositivos móveis mais utilizados para acessar a internet (IBGE, 2016). A última versão do Portal da Estatísticas (PORTAL, 2018) também mostrou que os dispositivos móveis representavam 48% do tráfego da *web* em todo o mundo. Esse tráfego da *web* pode estar relacionado ao tipo de aplicativo utilizado pelos usuários. Aplicativos como *Facebook*, *WhatsApp*, *Instagram* e *Messenger* são os quatro principais em número de *downloads* entre os usuários do Android e são aplicativos de alta interatividade social (ANDROIDRANK, 2018).

Além do tipo de aplicativo utilizado, a frequência de uso também é um fator relevante para a distração. No *survey* encomendado pela AT&T e conduzido pela Braun Research, foi revelado que 70% dos motoristas que possuem um telefone celular o utilizam em algum momento enquanto estão dirigindo (AT&T, 2015). Na pesquisa realizada pela *AAA Foundation for Traffic Safety*, aproximadamente 70% dos motoristas entre 16 e 18 anos afirmaram que já falaram ao telefone enquanto dirigiam, 42% afirmaram já terem lido alguma mensagem de texto ou *e-mail* e 32% já ter digitado alguma mensagem de texto enquanto dirigiam (CARNEY et al., 2016).

Na Índia, a *SaveLIFE Foundation* (SLF), que é uma organização não governamental comprometida a melhorar a segurança no trânsito e o atendimento médico emergencial em toda a Índia, analisou os padrões, efeitos e a percepção do uso do telefone celular de motoristas em toda a Índia. O *survey* foi aplicado nas cidades de Delhi, Chennai, Jaipur, Bengaluru, Mangaluru, Kanpur, Mumbai e Kolkata e envolveu o total de 1749 participantes

(SAVELIFE, 2017). Em cada cidade, condutores de quatro tipos de veículos, quais sejam: moto, carro, ônibus/caminhão e táxi indiano, foram entrevistados. Os resultados mostram que 47% dos motoristas afirmam receber chamadas telefônicas enquanto estão dirigindo e que 60% dos motoristas não param em um local seguro antes de atender uma ligação telefônica. Além disso, 20% dos entrevistados relataram que sofreram ou quase sofreram um acidente de trânsito devido ao uso do telefone celular, e 34% dos motoristas afirmam que tendem a dar freadas bruscas ao utilizar um *smartphone* enquanto dirigem. Esses dados evidenciam que a distração do motorista, em decorrência da utilização do telefone celular, é um problema existente em diversos países.

Diferentemente do cinto de segurança e da bebida alcoólica, cujos usos podem ser facilmente detectados, a distração do motorista é um grande desafio, devido às dificuldades inerentes à medição de atenção do motorista. A atenção de um motorista pode ser desviada de forma visual, auditiva, cognitiva ou física. Dependendo da tarefa secundária que o motorista esteja realizando, a sua atenção pode ser desviada de mais de uma forma. Por exemplo, ao realizar uma chamada telefônica, o motorista pode desviar sua atenção física - se estiver segurando o telefone com uma das mãos; auditiva - por estar em uma ligação telefônica; e cognitiva - por estar envolvido no assunto que esteja sendo dialogado (WHO, 2011). Além das várias possibilidades de distração, a identificação da distração do motorista também é considerada complexa, devido à frequência com que um evento secundário pode ocorrer. Segundo (DINGUS et al., 2016), os motoristas realizam alguma atividade com potencial de distração em mais de 51% do tempo em que estão dirigindo. Outro motivo que também torna complexa a detecção da distração diz respeito ao tempo que o evento perdura. Por exemplo, ler ou digitar uma mensagem de texto, enquanto está dirigindo, é uma atividade com elevado potencial de distração, mas que, geralmente, perdura por poucos segundos.

Apesar da janela do evento secundário ser pequena, os riscos da distração crescem significativamente ao utilizar um telefone celular. De acordo com (HOSKING; YOUNG; REGAN, 2009), a quantidade de tempo que um motorista fica sem olhar para a pista aumenta em 400% se ele estiver digitando uma mensagem de texto. Ao escrever uma mensagem de texto, o motorista está desviando sua atenção física, cognitiva e visual. Nesse contexto, é possível inferir que quanto mais formas de distrações concomitantes e maior for o tempo da distração, maior será o risco de acidentes provocados pelo motorista.

Levando em consideração o cenário apresentado, pode-se perceber que a distração do motorista decorrente da utilização do telefone celular é um problema complexo que precisa ser investigado. Sendo assim, essa tese apresenta a investigação da distração do motorista em três situações: (1) a partir da observação do comportamento do condutor, apresentado no Capítulo 5, (2) analisando a distração do motorista devido à composição de mensagens de texto, apresentada no Capítulo 6 e (3) analisando a distração do motorista

devido à leitura de mensagens de texto, apresentada no Capítulo 7.

No primeiro estudo, o objetivo foi demonstrar como o modelo de aprendizado de máquina pode ser utilizado para resolver o problema da distração do motorista a partir da metodologia de observação do comportamento do condutor. Para contemplar esse objetivo, foi desenvolvida uma arquitetura de Rede Neural Convolucional(CNN) para, a partir de imagens capturadas, classificar o tipo de distração do motorista. A CNN desenvolvida levou em consideração a distração devido à realização de chamadas telefônicas e composição/leitura de mensagens de texto. O conjunto de métricas utilizadas nesse estudo permitiu analisar a capacidade de generalização do modelo.

No segundo estudo, o objetivo foi demonstrar como o modelo de aprendizado de máquina pode ser utilizado para resolver o problema da distração do motorista devido à composição de mensagens de texto. Além de detectar a distração, esse estudo também teve o objetivo de resolver o problema da inconveniência para o passageiro. Assim, os modelos desenvolvidos foram projetados para classificar motoristas e passageiros durante a composição de mensagens de texto. Nesse estudo, foi utilizada a metodologia de observação dos efeitos do comportamento do motorista. Os dados foram coletados utilizando os sensores GPS, giroscópio, magnetômetro, acelerômetro e *touch screen*. Diferentes modelos de aprendizado de máquina foram desenvolvidos e comparados. O conjunto de métricas utilizadas permitiu analisar a capacidade de generalização dos modelos, bem como a significância estatística entre eles. Na análise dos resultados, nós demonstramos quais foram os melhores modelos construídos.

No terceiro estudo, o objetivo foi demonstrar como o modelo de aprendizado de máquina pode ser utilizado para resolver o problema da distração do motorista devido à leitura de mensagens de texto. Na revisão sistemática realizada no Capítulo 2, não foi filtrado nenhum trabalho que analisasse somente a distração em decorrência da leitura de mensagens de texto. Os trabalhos que analisaram esse tipo de distração consideraram a escrita e leitura de mensagem de texto como um único evento. Entretanto, como demonstrado no *survey* do Capítulo 3, pode-se observar que a distração devido à leitura de mensagens de texto é um evento secundário que ocorre com maior frequência quando comparado à distração decorrente da escrita de mensagens de texto. Nesse sentido, este estudo utilizou a metodologia de observação dos efeitos do comportamento do motorista para construir modelos de aprendizado de máquina capazes de classificar passageiros e motoristas durante a leitura de mensagens de texto. Diferentes modelos de aprendizado de máquina foram desenvolvidos e comparados. Assim como no segundo estudo, o conjunto de métricas utilizadas permitiu analisar a capacidade de generalização dos modelos, bem como a significância estatística entre eles. Na análise dos resultados, nós demonstramos quais foram os melhores modelos construídos.

1.1 Definição do problema

Considerando a distração do motorista, ainda são necessárias muitas pesquisas para conseguir detectar e combater a utilização do telefone celular enquanto se está dirigindo. Segundo a (NHSTA, 2017), Administração Nacional de Segurança Rodoviária Norte Americana, primeiramente é necessário sistematizar o monitoramento da utilização do telefone celular para poder diagnosticar fatores, como as regiões onde aparelho é mais utilizado, por quais grupos de pessoas, em qual período e em quais circunstâncias. A partir desta avaliação, soluções preventivas poderão ser adotadas efetivamente.

A distração do motorista devido à utilização do telefone celular pode ocorrer em decorrência da realização de uma ligação telefônica, da composição de uma mensagem de texto, ou pela leitura de mensagens e/ou notificações. A melhor forma de evitar este problema seria a conscientização dos usuários e o cumprimento das leis. Entretanto, devido ao elevado índice de acidentes nota-se a necessidade de outras medidas para prevenir a distração do motorista devido ao uso do telefone celular. Soluções de bloqueio do dispositivo ao detectar a movimentação do veículo podem ser eficientes, contudo, geram um problema que denominamos de inconveniência para o passageiro. Nas soluções de bloqueio por detecção de movimentação, todos os dispositivos são bloqueados, independentemente de ser motorista ou passageiro. Nesse sentido, além de construir uma solução eficiente que iniba a utilização do telefone celular pelos motoristas, também é necessário que esta mesma solução permita que os passageiros o utilizem.

Os estudos que objetivam detectar a distração dos motoristas, em geral, possuem duas vertentes: analisam o comportamento do condutor ou analisam os efeitos causados pelo comportamento do condutor. As pesquisas que analisam o comportamento do condutor buscam encontrar padrões de comportamento observando características como o olhar, as mãos, o posicionamento da cabeça ou observando o próprio comportamento do motorista de uma forma global. Já as pesquisas que analisam os efeitos causados pelo comportamento do condutor visam detectar a mudança dos padrões de uma direção considerada normal. Segundo (REGAN; LEE; YOUNG, 2009) e (SWOV, 2017), a utilização do telefone celular pode influenciar negativamente a performance do motorista em várias situações. Maior tempo para reagir a eventos inesperados, freadas bruscas, distância inapropriada e reação retardada para sinais luminosos são alguns exemplos de baixo desempenho que o motorista pode apresentar enquanto estiver dirigindo e utilizando um dispositivo móvel.

As pesquisas que analisam o comportamento do condutor, costumam utilizar equipamentos específicos para a captura dos dados e para observação do motorista, tais como: câmera, *eye tracker*, *wearable sensor network*, *electrocardiogram* e simuladores (RAJENDRA; DEHZANGI, 2017; LIU et al., 2016; SAHAYADHAS et al., 2015a). Os trabalhos que analisam os efeitos causados pelo comportamento do condutor costumam utilizar simuladores ou sensores específicos para capturar os dados do veículo. Nessa linha

de pesquisa, existe a vertente que utiliza os sensores dos *smartphones* para realizar a captura dos dados e o monitoramento do veículo. Os trabalhos que seguem essa metodologia justificam que as soluções de prevenção de distração utilizando os recursos do próprio *smartphone* são mais baratas por não requererem nenhum equipamento adicional e, por esse motivo, poderem ser utilizadas por um maior público. Outra vantagem é o fato dessa abordagem não ser intrusiva, em comparação a utilização de câmeras, por exemplo.

Partindo da hipótese de que prevenir a distração do motorista implica diminuir a probabilidade de acidentes e, consequentemente, salvar vidas, esta tese propõe o estudo de modelos de aprendizado de máquina para identificar e proporcionar uma segestão de prevenção da distração do motorista devido à utilização do telefone celular enquanto estiver dirigindo. Além da possibilidade de prevenção da distração, as soluções desenvolvidas também têm o objetivo de proporcionar a conveniência para o passageiro. Ou seja, os modelos de aprendizado de máquina foram construídos para diferenciar a utilização do celular por motoristas e passageiros. A escolha pela utilização de modelos de aprendizado de máquina se deu pela particularidade do problema a ser resolvido. O evento de distração do motorista pode ocorrer frequentemente e em poucos segundos. Nesse sentido, considera-se que, a partir dos dados coletados nesse intervalo de tempo, a tarefa de detecção da distração e classificação entre motoristas e passageiros é uma atividade complexa e apropriada para esse tipo de abordagem.

1.2 Objetivo

Esta tese tem como objetivo principal construir e analisar diferentes modelos de aprendizado de máquina, na tarefa de classificar motoristas e passageiros durante a distração devido à utilização do telefone celular. Essa classificação se faz necessária para proporcionar o que chamamos de conveniência para o passageiros. Para tanto, foram realizados experimentos utilizando a abordagem da observação do comportamento do condutor e a abordagem da observação dos efeitos do comportamento do condutor. Nas duas metodologias exploradas, foi definido que a coleta dos dados seria realizada por recursos disponíveis no próprio *smartphone*. Esse requisito foi estabelecido por acreditar que as soluções provenientes de *smartphones* são mais baratas e podem ser utilizadas por um maior público. Foram realizados três experimentos, cuja coleta de dados foi realizada por câmera, GPS, acelerômetro, magnetômetro, giroscópio, e *touch screen*. A definição dos sensores e dos modelos de aprendizado utilizados foram determinadas a partir da revisão sistemática conduzida nesta tese.

Além da coleta e rotulação dos dados, a utilização de modelos de aprendizado de máquina apresentam desafios na configuração, na análise, interpretação e comparação dos modelos. Nesse sentido, em cada estudo, nós demonstramos como os parâmetros foram

configurados. Para a análise, interpretação e comparação dos modelos, foi definido um conjunto de métricas as quais permitiram identificar os modelos e configurações mais adequadas para cada tipo de problema abordado.

Outro objetivo desta tese diz respeito à investigação do perfil de distração dos motoristas paraenses em decorrência da utilização do telefone celular. Segundo a pesquisa divulgada pelo Ministério da Saúde (SAÚDE, 2019), Belém é a capital que possui o maior percentual (24%) de condutores que utilizam o telefone celular enquanto estão dirigindo. Nesse sentido, o Capítulo 3 apresenta os resultados do *survey* que investigou o perfil de distração dos motoristas paraenses.

Para alcançar os objetivos desta tese, as seguintes tarefas foram planejadas e realizadas:

- Realização da revisão sistemática da literatura para analisar as formas de coleta e métodos utilizados para detectar e prevenir a distração do motorista devido ao uso do telefone celular.
- Aplicação do *survey* para diagnosticar o perfil de distração dos motoristas paraenses em relação à utilização do telefone celular.
- Realização de um experimento utilizando a metodologia de análise do comportamento do condutor para detectar os três tipos de distração do motorista provenientes da utilização do telefone celular.
- Realização de um experimento utilizando a metodologia de análise dos efeitos causados pelo comportamento do condutor, para detectar a digitação de mensagens de texto e classificação de passageiros e motoristas.
- Realização de um experimento utilizando a metodologia de análise dos efeitos causados pelo comportamento do condutor, para detectar a leitura de mensagens de texto e classificação de passageiros e motoristas.

1.3 Hipótese da tese

Para a realização dos objetivos desta tese, a hipótese principal abordada nesta pesquisa é a de que os padrões sensoriais observados durante a condução e a utilização do telefone celular possam ser suficientes para a construção de modelos de aprendizados de máquina, que classifiquem a utilização do telefone celular por motoristas e passageiros de forma não intrusiva.

As hipóteses secundárias são:

- Ao dirigir e utilizar um telefone celular, acredita-se que os padrões de direção de um condutor se alterem. Ou seja, se um motorista estiver distraído, espera-se que a mudança de comportamento como, por exemplo, aceleração, desaceleração e freadas sejam percebidas pelos sensores.

- Considerando os dados sensoriais, acredita-se que o uso de abordagens computacionais que utilizem limiar não são suficientes para classificar motoristas e passageiros. Nesse caso, espera-se demonstrar que os algoritmos de aprendizado de máquina podem ser utilizados para solucionar esse problema.

- A distração devido à leitura de mensagens/notificação pode acionar os mesmos gatilhos da distração devido à composição de mensagens de texto. Nesse caso, espera-se demonstrar que a distração devido à leitura de mensagens/notificação costuma ser mais praticada e que, por esse motivo, justifica-se a proposta de uma solução específica para tratar esse problema.

- O problema da distração devido a utilização do telefone celular é um problema complexo que pode ser resolvido a partir da abordagem de modelos de aprendizado de máquina. Nesse caso, espera-se demonstrar, a partir de uma revisão da literatura, que o aprendizado de máquina é uma possibilidade para detectar e classificar a distração do motorista.

1.4 Publicações e projetos

O desenvolvimento das atividades desta tese resultou nas seguintes publicações:

- *A Deep Learning Approach to Detect Distracted Drivers Using a Mobile Phone* na *26th International Conference on Artificial Neural Networks* em 2017. Qualis B1.

- *Exploiting Machine Learning Models to Avoid Texting While Driving* na *International Joint Conference on Neural Networks* em 2019. Qualis A1.

- *A Machine-Learning Approach to Distinguish Passengers and Drivers Reading While Driving* na *Sensors 2019, Volume 19, Issue 14* em 2019. Qualis A1.

- Aprovação do projeto de extensão: Investigação e medidas educativas para prevenir a distração dos motoristas em decorrência da utilização do telefone celular, no edital 03/2019 PROEXTENSÃO-IFPA. Esse projeto permitirá a ampliação e publicação do *survey* apresentado no capítulo 3.

- Publicação da base de dados *Passengers and Drivers reading while driving* no repositório *Figshare*[1].

[1] <https://figshare.com/articles/Passengers_and_Drivers_reading_while_driving/8313620>

1.5 Organização da Tese

Esta tese de doutorado está dividida em oito capítulos:

- O Capítulo 1 contextualiza a proposta do projeto de pesquisa, apresentando a motivação, a problemática e os objetivos pretendidos.
- O Capítulo 2 faz um levantamento dos trabalhos relacionados e aborda os principais métodos e teorias utilizados na captura, análise e classificação da distração do motorista devido à utilização do telefone celular.
- O Capítulo 3 apresenta os resultados da condução do *survey* que verificou o perfil de distração dos motoristas paraenses em decorrência da utilização do telefone celular.
- O Capítulo 4 apresenta a fundamentação teórica dos modelos de aprendizado de máquina utilizados nesta tese.
- O Capítulo 5 apresenta o experimento que utiliza a metodologia de observação do comportamento do condutor para classificar os três tipos de distração devido à utilização do telefone celular.
- O Capítulo 6 apresenta o experimento que utiliza a metodologia de observação dos efeitos do comportamento do motorista para classificar passageiros e motoristas durante a distração devido á composição de mensagens de texto.
- O Capítulo 7 apresenta o experimento que utiliza a metodologia de observação dos efeitos do comportamento do motorista para classificar passageiros e motoristas durante a distração devido á leitura de mensagens de texto.
- Por fim, no capítulo 8, é possível encontrar as conclusões desta tese, entre elas, as principais contribuições, as publicações obtidas, as limitações e propostas de trabalhos futuros.

2 Revisão sistemática

Este capítulo descreve a revisão sistemática conduzida nessa tese. O objetivo da revisão compreende a verificação do estado da arte e identificação de oportunidades de pesquisa na temática da distração do motorista em decorrência da utilização do telefone celular. Os tipos de pesquisa; métodos; sensores e técnicas foram as questões investigadas nessa revisão.

O decorrer do capítulo está dividido em quatro principais seções. Na primeira seção é discutida a motivação da revisão sistemática. Na segunda, apresentamos o planejamento e o protocolo adotado da revisão. Na terceira seção são apresentados os filtros e a análise de cada questão da revisão sistemática. Por fim, na última seção, são apresentadas as considerações finais desse estudo.

2.1 Motivação

Os estudos que investigam a distração do motorista em consequência da utilização do telefone celular podem ser exemplificados em observação microscópica e macroscópica. Segundo (STOLEN; GRANDE; ALLAN, 2004), o conceito de observação micro e macroscópica é utilizado no estudo da termodinâmica.

Na termodinâmica, a abordagem microscópica considera que o sistema é formado por uma grande quantidade de moléculas, que tem características de mudanças frequentes devido à colisão entre elas, mas em um determinado momento têm a posição, velocidade e energia definida. Assim, o comportamento é analisado a partir da observação individual das moléculas. A visão microscópica ajuda a obter uma compreensão mais profunda das leis da termodinâmica. No entanto, é mais complexa e demorada.

O estudo macroscópico é utilizado na termodinâmica clássica. Nesta vertente, uma certa quantidade de matéria é observada sem levar em consideração as particularidades do nível molecular. Em outras palavras, essa linha da termodinâmica está preocupada com o comportamento global das moléculas. O estado ou condição do sistema pode ser completamente descrito por valores medidos de pressão, temperatura e volume, que são chamados de variáveis macroscópicas. Conforme observado por (STOLEN; GRANDE; ALLAN, 2004), por não estar interessada nas individualidades das moléculas, a metodologia macroscópica requer uma explanação matemática mais simples do que a metodologia microscópica.

Essas duas formas de observação também são utilizadas na análise do comportamento de motoristas. Identificar padrões comportamentais do motorista é de significativa

importância para áreas que investigam o tráfego de veículos, segurança rodoviária, perfis de agressividade, de sonolência e de distração do motorista. Na análise comportamental do motorista, assim como na termodinâmica, a observação macro e microscópica também se caracterizam como observações globais e individuais. Ao analisar o tráfego de veículos, câmeras podem ser utilizadas ao longo da rodovia para a observação macroscópica. Neste cenário, a velocidade média dos veículos; quantidade de veículos por minuto; tempo médio de congestionamento e período de maior fluxo são exemplos de observações que podem ser realizadas macroscopicamente.

Na perspectiva microscópica, o tráfego é analisado a partir dos padrões individuais dos veículos. Dados sensoriais são utilizados para esta finalidade e permitem a análise contínua de características como a aceleração, desaceleração, fluxo livre, momento estacionário, curvas e mudanças de faixa. A identificação dessas características aprofunda o estudo realizado a nível macroscópico e possibilita o desenvolvimento de soluções preventivas. Na linha das soluções microscópicas, em 2006, um estudo de caso realizado na Califórnia, mostrou que um *Automated Workzone Information Systems* (AWIS) proporcionou uma redução de 50% no tempo médio de atraso de uma viagem (LEE; KIM, 2006), diminuindo, consequentemente, os riscos dessa viagem.

Soluções como AWIS se tornaram populares graças ao aumento do uso de dispositivos móveis e da ascendência das redes de sensoriamento participativo (RSP). Para (BURKE et al., 2006), um sensoriamento participativo pode ser definido como um processo distribuído de coleta de dados pessoais e sobre diversos aspectos da cidade. Tal processo requer a participação ativa das pessoas para compartilhar, voluntariamente, informação contextual ou tornar seus dados sensoriais disponíveis. Aplicações como *Waze*[1] e *Moovit*[2] são casos de sucesso que utilizam as redes de sensoriamento participativo no contexto de tráfego de veículos.

Em relação à distração do motorista, as observações micro e macroscópica podem ser classificadas de acordo com as características analisadas. A observação macroscópica analisa o comportamento global do motorista sem levar em consideração partes específicas do seu corpo. Em geral, este tipo de observação pode ser realizado utilizando câmeras acopladas ao veículo, espalhadas ao longo da rodovia ou até mesmo a olho nu. As câmeras espalhadas ao longo da rodovia caracterizam-se como observação de ponto fixo. Assim como na análise de tráfego de veículos, este tipo de observação possui o campo de observação limitado. Por estarem distantes do alvo, as câmeras utilizadas para esta finalidade precisam ser de alta resolução. Neste sentido, a observação de ponto fixo costuma ser realizada a olho nu e, em geral observações dessa natureza tem a finalidade de coletar dados para relatórios estatísticos.

[1] Aplicação para smartphones: <https://www.waze.com/pt-BR/>
[2] Aplicação para smartphones: <https://moovitapp.com/>

Além de poder coletar dados estatísticos, a análise macroscópica com câmeras acopladas ao veículo pode utilizar algoritmos inteligentes para detectar e coibir as distrações de motoristas. Nesta metodologia, a questão da intrusão pode ser considerada um problema. Parte da literatura afirma que o fato de o motorista estar sendo observado pode gerar a rejeição dessas soluções. Entretanto, segundo o *survey* aplicado pela SaveLife, 68% dos entrevistados aceitam a ideia da utilização de câmeras para o monitoramento de distração de motorista (SAVELIFE, 2017). Este dado revela que a intrusão das câmeras não é um problema para a maioria dos motoristas. Nesse contexto, por serem soluções simples, eficientes e aceitáveis, a metodologia macroscópica com câmeras acopladas ao veículo tem ganhado destaque na literatura.

As pesquisas que se preocupam com as questões de intrusão têm a alternativa da metodologia microscópica. A observação microscópica utiliza redes sensoriais para identificar padrões de distração a partir de características específicas do comportamento do motorista. Analisar o olhar, o movimento da cabeça e das mãos, são exemplos de observação microscópica. Apesar de apresentarem soluções não intrusivas, uma desvantagem desta metodologia é que, em geral, necessitam de *hardwares* específicos para a observação das características, o que eleva os custos de aplicação efetiva. Entretanto, uma vez que os dados são coletados e tradados, modelos de alta precisão são desenvolvidos para a análise e detecção dos padrões de distração do motorista.

Além disso, como alternativa às observações do comportamento micro e macroscópico, temos ainda a observação dos efeitos do comportamento distraído. Nessa metodologia assume-se a hipótese de que os efeitos do comportamento de um motorista distraído alteram o seu padrão de direção. Sendo assim, ao invés de utilizar sensores para observar o comportamento do motorista, a observação dos efeitos do comportamento distraído utiliza sensores para analisar o comportamento do veículo. Mudanças nos padrões de velocidade, aceleração, desaceleração, direção, freadas e mudança de faixa, são exemplos de parâmetros analisados por trabalhos que utilizam a metodologia de observação dos efeitos do comportamento distraído. Assim como na observação microscópica, a observação dos efeitos do comportamento distraído também apresenta a vantagem da coleta de dados não intrusiva. Outra vantagem apresentada por essa metodologia diz respeito ao baixo custo do *hardware* necessário para a coleta de dados. Em vários trabalhos pesquisados na literatura, observou-se que muitos utilizam os sensores do próprio *smartphone* para a realização da coleta de dados.

Apesar das vantagens dessa metodologia, devemos destacar que existe uma desvantagem de complexidade contextual. Por analisar apenas as mudanças do comportamento do veículo, é difícil inferir qual o tipo de distração que ocasionou a mudança observada. Por exemplo, assumindo que foi observada uma desaceleração e mudanças significativas na velocidade do veículo em um intervalo de cinco segundos, não se pode inferir se a causa

deste comportamento decorreu de uma distração ou de um sinistro. Um modelo hipotético poderia inferir uma possível distração do motorista.

Em uma perspectiva contextual não se pode determinar qual a distração exata que ocasionou a mudança dos padrões observados. Na verdade, não se pode nem depreender que o motivo do padrão observado foi ocasionado por uma distração. Por exemplo, a desaceleração e a redução de velocidade poderia ser reflexo de um congestionamento. As pesquisas que utilizam a metodologia da observação dos efeitos do comportamento distraído tendem a desenvolver modelos complexos que consigam extrair padrões contextuais.

Levando em consideração as vantagens e desvantagens das metodologias, esta seção tem o objetivo de apresentar o levantamento dos trabalhos que propuseram soluções para a análise ou detecção de distração do motorista. Para a revisão realizada neste trabalho, foi usado o guia de revisão sistemática de Kitchenham (KITCHENHAM, 2007) com algumas adaptações e simplificações. A Figura 1 apresenta o fluxo das atividades da revisão utilizada neste trabalho.

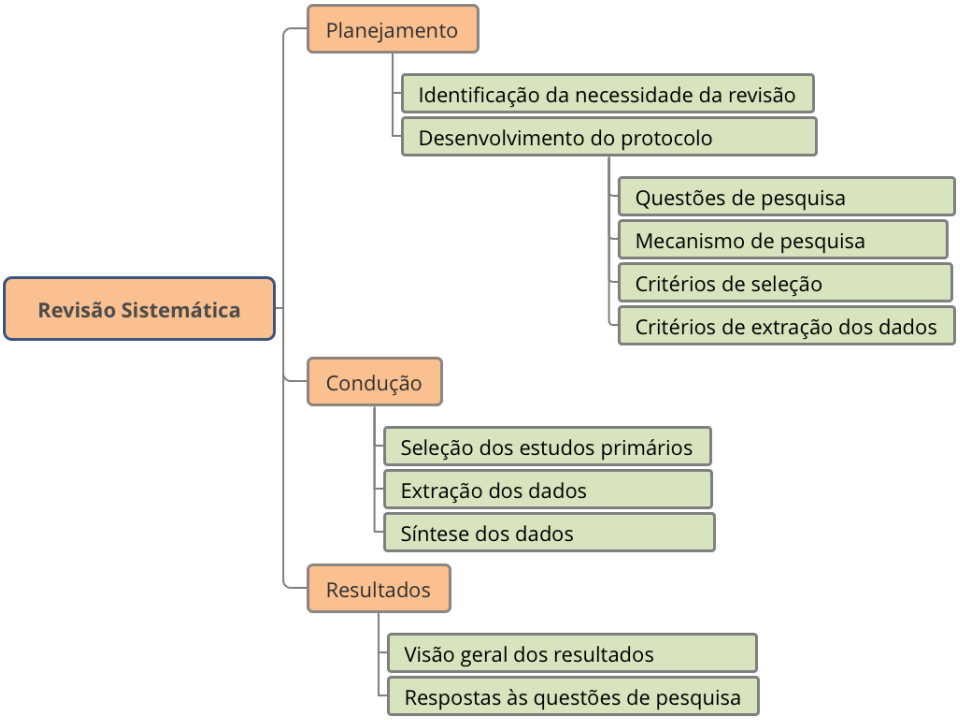


Figura 1. Protocolo da revisão sistemática adaptada de Kitchenam.

2.2 Planejamento

Na etapa de planejamento é fundamental identificar a necessidade da revisão sistemática e definir o protocolo de sua condução.

2.2.1 Identificação da necessidade da revisão

A distração do motorista em decorrência da utilização do telefone celular é um problema que vem crescendo e se tornando fator motivador de acidentes e mortes no trânsito. Por esse motivo, diversos pesquisadores estão investigando os motivos e soluções para distração do motorista. Como reflexo dessas pesquisas podemos perceber os avanços tecnológicos dos veículos. Atualmente, diversos carros já contam com tecnologias de sensoriamento para as soluções de *Advance Driver Assistance System* (ADAS). Como observado por (HANSEN et al., 2017), embora exista grande interesse em migrar para veículos totalmente automatizados e autônomos, fatores como desempenho tecnológico, barreiras de custo, segurança pública, questões de seguro, implicações legais e regulamentações governamentais indicam que o primeiro passo para combater a insegurança no trânsito seja o desenvolvimento de veículos multifuncionais.

A linha de raciocínio de (HANSEN et al., 2017) sugere considerarmos que os veículos multifuncionais equipados com ADAS representam um passo necessário para a segurança, comunicação e autonomia dos veículos. Nesse contexto, por entender que esta tese está inserida no contexto dos trabalhos que pesquisam os motivos e soluções da distração do motorista, se faz necessário a revisão da literatura para identificar o estado da arte e as oportunidades de pesquisas.

2.2.2 Desenvolvimento do protocolo

Questões de pesquisa

A condução da revisão sistemática foi realizada para responder as seguintes questões de pesquisa:

- **Questão 1:** Quais são os tipos de métodos adotados para a análise ou detecção da distração do motorista?
- **Questão 2:** Quais são os tipos de estudos realizados para a análise ou detecção da distração do motorista?
- **Questão 3:** Caso existam, quais são os sensores utilizados para a coleta de dados?
- **Questão 4:** Quais são as técnicas utilizadas para a análise ou detecção da distração do motorista?

Mecanismo de pesquisa

A investigação de trabalhos primários foi realizada nos seguintes repositórios:

- *ACM Digital Library: http://dl.acm.org/*

- *ScienceDirect: http://www.sciencedirect.com/*
- *IEEE Xplore Digital Library: http://ieeexplore.ieee.org/*

Em cada repositório, os trabalhos foram pesquisados utilizando as seguintes palavras chaves:

- *Driver Behavior;*
- *Driving Style Recognition;*
- *Driver Style Classification;*
- *Driver Distraction;*
- *Distracted Driving Detection;*
- *Driving Maneuver Detection;*
- *Smartphone-Based Driver;*
- *Driver Inattention*

Critérios de seleção

A consolidação da lista de trabalhos candidatos desta revisão sistemática foi realizada por dois filtros descritos a seguir. Os filtros tiveram o objetivo de eliminar os trabalhos fora dos escopo estabelecido ou que não atenderam as questões de pesquisa.

No primeiro filtro foram adotados os seguintes critérios de seleção de trabalhos:

- Ser um estudo publicado nos últimos dez anos (de 2007 até 2017);
- Presença das palavras-chave no título ou *abstract*;
- Leitura do *abstract* e avaliação do assunto conforme objetivo e as questões de pesquisa definidas nesta revisão sistemática.

No segundo filtro foi realizada a leitura completa dos trabalhos para verificar os seguintes critérios de seleção de trabalhos:

- Se é um estudo primário;
- Se é um estudo experimental;
- Se é um estudo observacional;
- Se é um estudo que analisa ou propõe soluções para a distração do motorista em decorrência da utilização do telefone celular.

Critérios de extração dos dados

Após a aplicação dos dois filtros, os trabalhos remanescentes foram utilizados para responder as questões de pesquisa estabelecidas no protocolo. Os seguintes dados foram extraídos de cada trabalho:

- Título;
- Autores;
- Ano de publicação;
- País;
- Método de pesquisa;
- Tipo de estudo realizado;
- Sensores utilizados para a coleta de dados;
- Algoritmo ou técnica utilizada.

2.3 Condução

A condução da revisão, que compreende a condução do protocolo e a análise dos resultados, levou quatro meses não contínuos para a conclusão, durante o primeiro semestre de 2018. A busca nos repositórios e a aplicação dos filtros foi iniciada em agosto de 2017 e finalizada em dezembro de 2017.

Seleção dos estudos primários

Utilizando as palavras-chave definidas no protocolo, foram obtidos 292 artigos nos três repositórios definidos. Após a aplicação do primeiro e segundo filtro, restaram, respectivamente, 170 e 86 artigos. Os 86 artigos selecionados foram utilizados para a extração e síntese dos dados. Na condução da revisão, foi utilizado o *software* JabRef[3] que é um gerenciador de referências bibliográficas de código aberto.

Extração dos dados

A extração dos dados se deu pela leitura dos trabalhos selecionados e preenchimento de um formulário para posterior síntese e apresentação. Conforme definido no protocolo, para cada trabalho selecionado, foram extraídos o título, autores, país, método de pesquisa,

[3] JabRef - http://www.jabref.org/

tipo de estudo realizado, sensores utilizados para a coleta de dados e os algoritmos ou técnicas utilizada para a experimentação ou sintetização dos dados observados.

Quanto ao método de estudo realizado, o objetivo foi classificar os trabalhos selecionados em três submétodos, quais sejam: estudos de observação do comportamento macroscópico; estudo de observação do comportamento microscópico; ou estudo de observação dos efeitos do comportamento distraído. Como observado na Figura 2, o método de estudo pode ser relacionado com o tipo de estudo realizado. O tipo de estudo dos trabalhos selecionados foram classificados como *survey*, experimental e observacional. Os estudos experimentais e observacionais foram realizados de forma naturalística ou por simulação. No contexto da distração do motorista, um estudo naturalístico consiste na coleta de dados realizadas no fluxo de trânsito real. Já um estudo por simulação, consiste na coleta de dados realizados de forma controlada em ambientes virtuais.

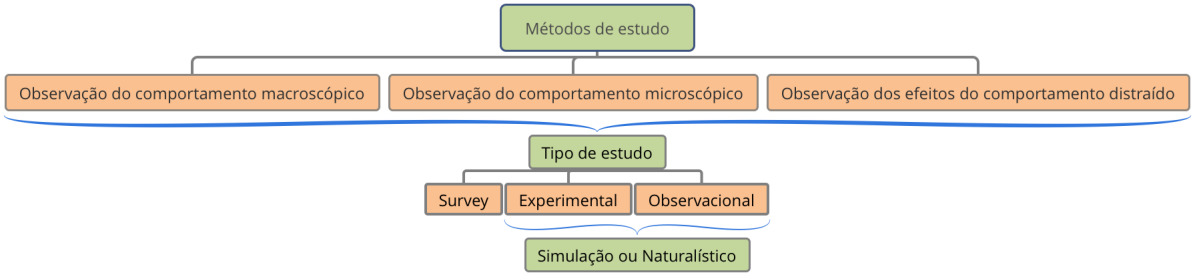


Figura 2. Relação entre método e tipo de estudo pesquisados.

Síntese dos dados

Esta etapa da revisão sistemática consiste na transformação dos dados extraídos em informações relevantes para a resolução das questões de pesquisa definidas no protocolo. Assim, tabulação de dados, geração de gráficos, correlação e análise dos dados são realizadas com o objetivo de identificar o estado da arte e as oportunidades de pesquisa existentes na área investigada.

2.3.1 Visão geral dos resultados

Esta seção apresenta a visão geral da síntese realizada na condução da revisão sistemática. Os dados gerais têm como objetivo contextualizar os trabalhos filtrados para que seja possível responder as questões de pesquisa.

A Figura 3 apresenta a quantidade de publicações filtradas por repositório. Dos 86 trabalhos selecionados, podemos observar que a maioria pertence aos repositórios da *Science Direct* e IEEE. O volume de trabalhos nesses repositórios se explica pela existência de jornais e conferências específicas para a área de segurança e tecnologias de veículos. Dentre as revistas do repositório *Science Direct*, destacam-se *Journal of*

Safety Research, *Accident Analysis & Prevention* e *Journal of Transport & Health* como as revistas mais procuradas para publicação na área. Dos periódicos e conferências da IEEE, destacam-se *IEEE Transactions on Intelligent Transportation Systems*, *IEEE Intelligent Vehicles Symposium* e *International Conference on Intelligent Transportation Systems* como os periódicos e conferências mais procurados para publicações na linha de distração de motorista.

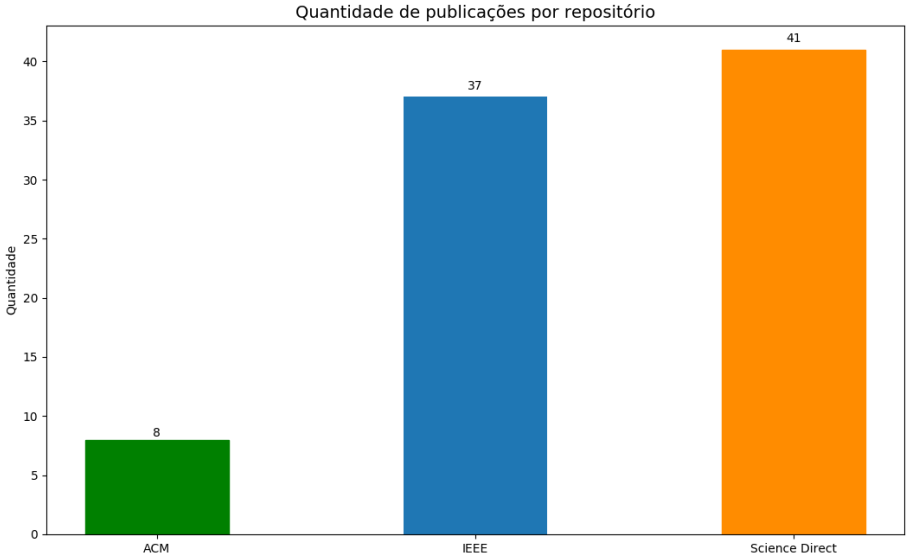


Figura 3. Quantidade de publicações filtradas por repositório.

Na Figura 4, podemos observar a quantidade de publicações, por ano, no período de 2007 à 2017. Analisando o gráfico, é notório o volume de publicações após o ano de 2011. Esse comportamento pode ser explicado pelo lançamento do programa "Década de Ações para a Segurança no Trânsito - 2011 a 2020"(KI-MOON, 2010). Nesse programa a OMS lançou o desafio de diminuir em 50% o número de mortos em decorrência de acidentes de trânsito e salvar 5 milhões de vidas em uma década. Sendo assim, observou-se que o referido desafio estimulou o financiamento de pesquisas nas áreas de gestão da segurança viária; infraestrutura e mobilidade; tecnologias para veículos mais seguros; conscientização dos usuários e resposta ao acidente.

Na Figura 5, temos uma perspectiva global dos países que publicaram os 86 artigos selecionados nesta revisão. Para o levantamento deste quantitativo, foi levado em consideração o país onde os pesquisadores desenvolveram e aplicaram seus trabalhos. No total, 27 países realizaram ao menos uma publicação. Um fato importante é que todos os continentes estão pesquisando e publicando sobre a temática da distração de motorista. Essa observação reforça a importância da área que está sendo investigada nesta tese. Os cinco países com maior número de publicações foram os Estados Unidos com 29 publicações,

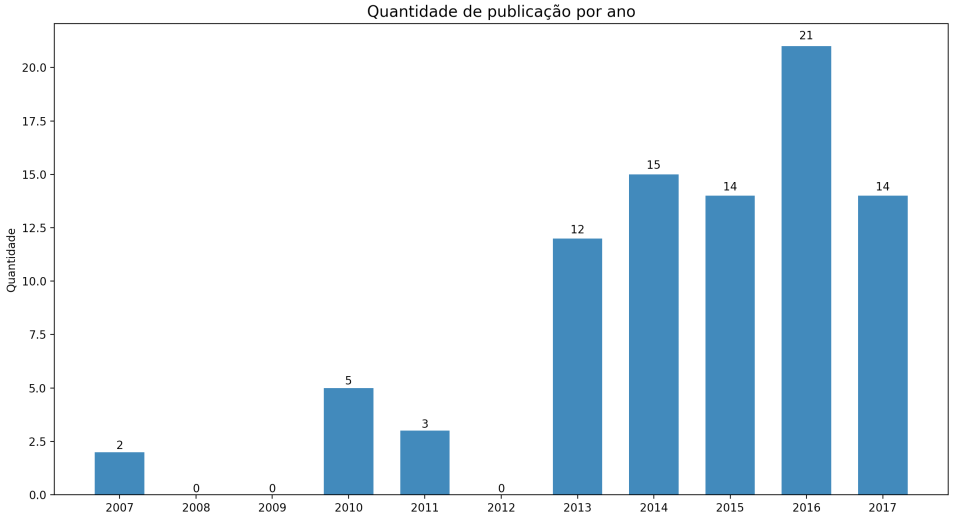


Figura 4. Quantidade de publicações por ano.

Figura 5. Países que publicaram os trabalhos selecionados nesta revisão.

China, Índia, Austrália e Suécia, cada um com 6 publicações. A Tabela 1 apresenta a quantidade de publicações de todos os 27 países.

2.3.2 Questão de pesquisa 1

Questão 1: Quais são os tipos de métodos adotados para a análise ou detecção da distração do motorista?

A Figura 6 apresenta a sintetização para responder a primeira questão de pesquisa. Conforme definido no protocolo da revisão, o objetivo era verificar os trabalhos que utilizaram os métodos de estudo de observação do comportamento macroscópico, observação

Tabela 1. Lista de quantidade de publicações por país

País	**Publicações**
Estados Unidos	29
Austrália	6
China	6
Índia	6
Suécia	6
Alemanha	5
Coréia do Sul	4
Espanha	3
Finlândia	2
Grécia	2
África do Sul	1
Bangladesh	1
Brasil	1
Canadá	1
Catar	1
Colômbia	1
França	1
Irã	1
Itália	1
Japão	1
Nigéria	1
Paquistão	1
Reino Unido	1
Singapura	1
Tailândia	1
Taiwan	1
Vietnã	1

do comportamento microscópico e observação dos efeitos do comportamento distraído.

Como já mencionado neste capítulo, a observação do comportamento macroscópico analisa o comportamento global do motorista, sem levar em consideração partes específicas do seu corpo. Pesquisas que analisaram o comportamento de motoristas a olho nu ou por

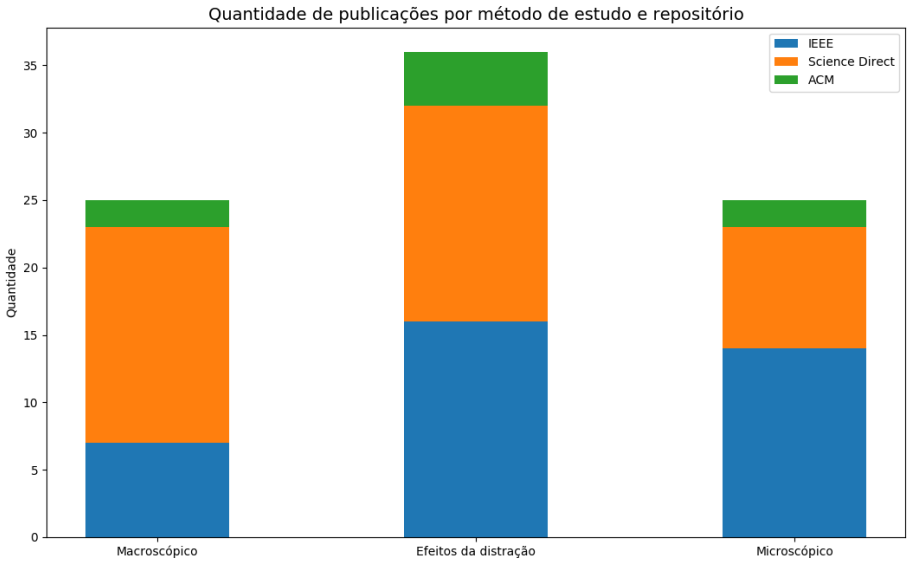


Figura 6. Quantidade de publicações por método de estudo e repositório.

câmeras são exemplos de estudos macroscópicos.

Trabalhos como os de (SABZEVARI et al., 2016), (HUTH; SANCHEZ; BRUSQUE, 2015), (KIDD et al., 2016) e (ORTIZ; RAMNARAYAN; MIZENKO, 2017) são exemplos de artigos selecionados que realizaram a observação macroscópica a olho nu. Trabalhos com essa metodologia costumam ter uma quantidade significativa de observações. Por exemplo, a pesquisa realizada por (SABZEVARI et al., 2016) realizou 7979 observações em três diferentes cidades do Irã. Outra característica dos trabalhos de observação macroscópica a olho nu é que, em geral, costumam utilizar ferramentas estatísticas para analisar e correlacionar os parâmetros observados.

Os trabalhos que realizam a observação macroscópica utilizando câmeras, além de contarem com utilização de ferramentas estatísticas, também utilizam métodos de classificação para supervisionar a distração do motorista. Trabalhos como os de (BILLAH; RAHMAN; MAHBUBUR, 2016), (KUTILA et al., 2007), (BERRI et al., 2014) e (DINH et al., 2014) são exemplos de artigos selecionados que realizam a observação macroscópica com auxílio de câmera e que utilizam métodos de classificação para supervisionar a distração do motorista.

No que diz respeito aos trabalhos que utilizam a observação microscópica, estes têm como objetivo analisar, de forma não intrusiva, partes específicas do corpo do motorista para supervisionar a distração. Trabalhos que utilizam essa metodologia costumam utilizar câmeras e *hardwares* específicos para a coleta de dados. Pesquisas como as de (LIU et al., 2016), (SAHAYADHAS et al., 2015a), (LI; BUSSO, 2013a) e (LARSSON, 2016) são exemplos de trabalhos microscópicos que utilizaram *hardwares* específicos para supervisionar a

distração do motorista. A pesquisa de (LI; BUSSO, 2013a), a título de exemplo, utilizou um *Eye Tracker* para capturar as expressões faciais do motorista. Utilizando os dados coletados pelo *Eye Tracker*, (LI; BUSSO, 2013a) desenvolveu modelos de aprendizado de máquina (Support Vector Machine - SVM e K-Nearest Neighbor - KNN) para monitorar a distração do motorista.

Já os estudos que utilizaram a observação do efeito do comportamento distraído representaram a maioria dos trabalhos selecionados nesta revisão. Esse quantitativo pode ser explicado uma vez que estudos dessa natureza costumam utilizar sensores como acelerômetro, GPS e giroscópio para monitorar o comportamento do veículo. A presença desses sensores em *smartphones* justifica a facilidade de realização de estudos experimentais que monitorem a distração do motorista, a partir da observação dos efeitos do comportamento distraído.

Pesquisas como as de (LIU et al., 2017), (YE et al., 2017), (BO et al., 2013) e (CHOUDHARY; VELAGA, 2017c) são exemplos de artigos que utilizaram a observação do efeito do comportamento distraído para monitorar a distração do motorista. Podemos citar o trabalho de (YE et al., 2017) como um dos artigos selecionados que utilizaram os dados de sensores de *smartphones* para analisar os efeitos da distração. Na ocasião (YE et al., 2017) desenvolveu um modelo supervisionado de rede neural totalmente conectada para classificar e detectar a distração do motorista.

Levando em consideração os dados levantados, a resposta da primeira questão de pesquisa demonstra que as três metodologias investigadas são utilizadas pelos pesquisadores que investigam a distração do motorista. Cada metodologia apresenta suas particularidades, e a utilização de uma não exclui a necessidade das outras duas.

A metodologia macroscópica normalmente é utilizada quando se quer analisar a distração do motorista a partir de uma perspectiva global de observação do comportamento. Pesquisas que utilizam essa metodologia desenvolvem modelos com menor complexidade no tratamento das variáveis.

Os trabalhos que utilizam a metodologia microscópica, normalmente, preferem essa metodologia à macroscópica porque estão preocupados com as questões de intrusão. Na metodologia microscópica somente partes específicas do corpo do motorista são supervisionadas. Quando comparada com as pesquisas macroscópicas, as pesquisas que utilizam a abordagem microscópica desenvolvem modelos com maior complexidade no tratamento das variáveis.

Os estudos que analisam a distração do motorista a partir da observação dos efeitos do comportamento distraído também possuem a vantagem da não intrusão. Porém, diferentemente da abordagem microscópica que necessita de *hardwares* específicos, a metodologia do efeito do comportamento distraído normalmente utiliza os sensores de *smartphones*

para coletar os dados do veículo. Os modelos desenvolvidos com essa metodologia tendem a ser complexos e apresentam a dificuldade de contextualização da informação. Por observar apenas o comportamento do veículo, é difícil contextualizar se os padrões observados realmente são oriundos de uma distração. Para obter resultados fidedignos, é comum que essa metodologia seja utilizada em combinação com a metodologia macroscópica ou microscópica.

2.3.3 Questão de pesquisa 2

Questão 2: Quais são os tipos de estudos realizados para a análise ou detecção da distração do motorista?

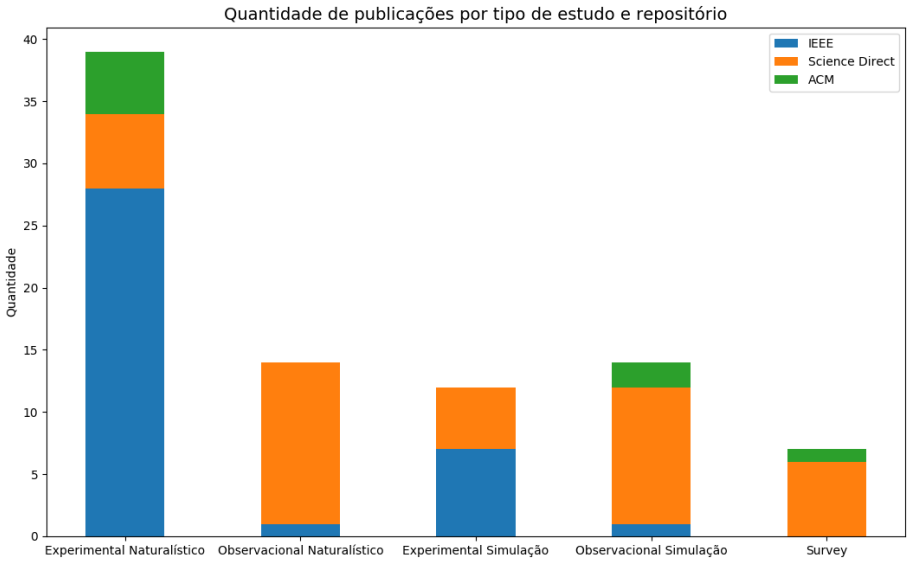


Figura 7. Quantidade de publicações por tipo de estudo e repositório.

A Figura 7 ilustra os cinco tipos de estudos utilizados pelos artigos selecionados nesta revisão. Podemos perceber que os estudos experimentais, sejam eles naturalísticos ou simulados, são, em sua maioria, publicados no repositório da IEEE. Já os *surveys* e estudos observacionais são, predominantemente, publicados no repositório da *ScienceDirect*. Esta constatação pode ser justificada em virtude dos tipos de revistas e conferência de cada repositório.

No que se refere à pesquisa experimental, segundo (GIL, 2002) este tipo de pesquisa consiste, essencialmente, em determinar um objeto de estudo, selecionar as variáveis capazes de influenciá-lo e definir as formas de controle e de observação dos efeitos que a variável produz no objeto. No contexto desta revisão, classificamos como trabalhos experimentais aqueles estudos que coletaram dados e desenvolveram algum método de análise ou controle para a distração de motorista investigada.

Dos trabalhos experimentais, alguns foram realizados de forma naturalísticas e outros por simulação. Segundo a (WHO, 2011), nos estudos naturalísticos os voluntários permitem que seu comportamento seja registrado durante um período de condução normal. Os trabalhos que optam por essa metodologia defendem a ideia de que o estudo naturalístico denota maior credibilidade por estar coletando dados em condições normais. Entende-se por condições normais o fluxo e interferência não programadas que podem ocorrer nas rodovias.

A parte da literatura que opta pela simulação argumenta que, além de garantir a segurança dos participantes, a simulação também tem a vantagem de proporcionar maior controle das variáveis observadas. Se um estudo está interessado em analisar a distração do motorista, na simulação, pode-se definir a configuração do ambiente e dos parâmetros observados. Variáveis como fluxo do trânsito, pedestres e condições climáticas podem ser manipuladas para a melhor experimentação.

Entendemos que tanto os experimentos por simulação quanto os experimentos naturalísticos tem a sua parcela de importância para a investigação da distração do motorista. Analisando a Figura 7, o que se tem percebido é que os estudos naturalísticos estão sendo mais utilizados pelos pesquisadores. Essa tendência pode ser explicada pelos motivos da fidedignidade e custo.

A questão da fidedignidade está associada a capacidade dos modelos desenvolvidos conseguirem bons desempenhos em um ambiente não controlado. Se um pesquisador consegue desenvolver um estudo que apresente bons resultados em um ambiente naturalístico, então os seus resultados indicam maior credibilidade.

Em relação ao custo, observamos que os estudos naturalísticos experimentais coletam seus dados utilizando sensores como GPS, acelerômetro, giroscópio, magnetômetro e câmera. Hoje, tais sensores tornaram-se acessíveis uma vez que estão disponíveis na maioria dos telefones celulares. Essa característica faz com que o custo dos estudos experimentais seja bem menor do que o custo dos estudos por simulação.

A simulação e o naturalístico também foram utilizados nos estudos classificados como observacionais. Segundo (GIL, 2002), estudos desta natureza tem como preocupação central identificar os fatores que determinam ou que contribuem para a ocorrência do fenômeno observado. No caso dos trabalhos filtrados nesta revisão, classificamos como estudos observacionais, as pesquisas que realizaram estudos naturalísticos ou de simulação para identificar os fatores que determinam ou que contribuem para a ocorrência da distração do motorista.

No contexto da distração de motorista, para diferenciar um estudo experimental do observacional, levamos em consideração a existência ou não de uma proposta de detecção, classificação ou prevenção da distração do motorista. Se um trabalho propõe

e avalia a eficiência de uma solução para detectar, classificar ou prevenir a distração do motorista, então classificamos esse trabalho como de cunho experimental. Por outro lado, classificamos como estudo observacional, as pesquisas que realizam investigações para entender os motivos e consequências da distração do motorista.

Além dos estudos experimentais e observacionais, também identificamos os estudos que realizaram *survey*. Segundo (PRODANOV; FREITAS, 2013), esse tipo de pesquisa ocorre quando envolve a interrogação direta das pessoas cujo comportamento desejamos conhecer através de algum tipo de questionário. Em geral, procedemos à solicitação de informações a um grupo significativo de pessoas acerca do problema estudado para, em seguida, mediante análise quantitativa, obtermos as conclusões correspondentes aos dados coletados.

No entendimento de (GIL, 2002), o *survey* é um tipo de estudo que deve ser considerado quando existe a necessidade de se realizar um estudo descritivo, como por exemplo, preferência eleitoral, comportamento do consumidor, pesquisa de opiniões e atitudes. No filtro realizado nesta revisão, observamos que o *survey* foi o tipo de estudo menos explorado pelos pesquisadores. Entretanto, há de se destacar que realizar o levantamento de opiniões e atitudes acerca da distração do motorista é de suma importância para o desenvolvimento de soluções preventivas.

Por exemplo, na Espanha, no *survey* conduzido por (GRAS et al., 2007), mais de 60% dos entrevistados admitem utilizar o telefone celular enquanto estão dirigindo. Por outro lado, no Catar, (SHAABAN, 2013) constatou-se que menos de 20% dos entrevistados afirmaram utilizar o telefone celular enquanto estão dirigindo. Essa diferença demonstra que as diferenças culturais influenciam no comportamento dos motoristas. Nesse sentido, realizar um *survey* para entender as opiniões e atitudes a partir de uma perspectiva local, parece ser de grande valia para o desenvolvimento de soluções customizadas. No caso da distração do motorista, se um *survey* identifica que uma parte significativa dos entrevistados costuma ler mensagens de texto enquanto estão dirigindo, então, realizar pesquisas e desenvolver soluções para contornar essa situação aparenta ser importante para a região onde o *survey* foi realizado.

2.3.4 Questão de pesquisa 3

Questão 3: Caso existam, quais são os sensores utilizados para a coleta de dados?

Na Figura 8 podemos observar um mapa de nuvem enfatizando os principais sensores utilizados pelas pesquisas filtradas nesta revisão. Dos sensores utilizados, destacam-se a câmera, acelerômetro, GPS, IMU, *speed control*, etc. A lista completa de sensores utilizados em cada trabalho pode ser vista na Tabela 28 do Apêndice A.

Nesta pesquisa, dos 86 trabalhos selecionados, 13 não utilizaram nenhum tipo de

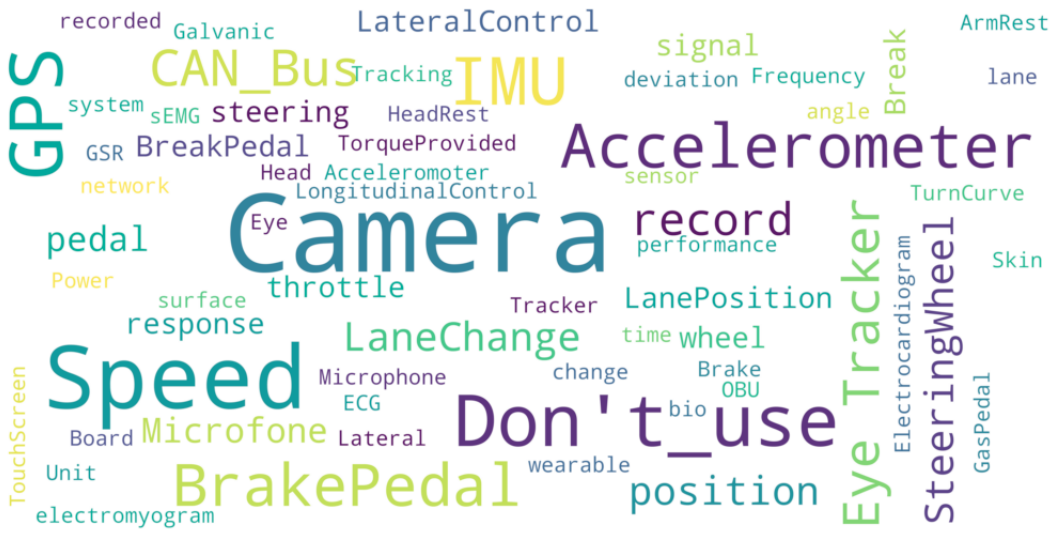


Figura 8. Mapa de nuvem dos sensores utilizados nos artigos filtrados.

sensor. Desses 13, 7 são trabalhos do tipo *survey* e 6 são do tipo estudo observacional naturalístico. O fato de os trabalhos do tipo *survey* não utilizarem sensores pode ser explicado pela própria natureza do trabalho. Considerando que um *survey* é um tipo de estudo que deve ser considerado quando existe a necessidade de se realizar um estudo descritivo, observamos que todos os *surveys* filtrados por essa revisão foram conduzidos a partir da aplicação de questionários.

Por outro lado, dos 13 trabalhos observacionais naturalísticos, verificamos que os 6 trabalhos que não utilizaram nenhum tipo de sensor foram realizados a partir da observação a olho nu. Pesquisas dessa natureza tem como característica a observação de ponto fixo e, geralmente, são do tipo macroscópica. No caso específico desses 6 trabalhos, todos realizaram estudos de observação macroscópica com o objetivo de analisar e contextualizar a distração do motorista.

Dos trabalhos que utilizaram a simulação, sejam eles observacionais ou experimentais, todos utilizaram um conjunto de sensores para capturar os dados gerados pelos motoristas. Os sensores para controlar a velocidade, aceleração, freio, mudança de faixa e câmeras, foram os sensores mais utilizados nesse tipo de estudo. De uma forma geral, as pesquisas observacionais por simulação utilizaram esse conjunto de sensores para coletar dados dos motoristas e analisar suas reações.

Por exemplo, a pesquisa realizada por (CHOUDHARY; VELAGA, 2017a) realizou simulações para observar quais os impactos que a utilização do telefone celular ocasionam no perfil de direção de motoristas. Na ocasião, (CHOUDHARY; VELAGA, 2017a) realizou ensaios para examinar os efeitos que as ligações telefônicas e composição de mensagens de texto simples e complexas podem afetar no desempenho dos motoristas. Na referida pesquisa, (CHOUDHARY; VELAGA, 2017a) utilizou os sensores de controle de freio,

velocidade, aceleração e deslocamento lateral.

Ainda no contexto da simulação, diferentemente das pesquisas observacionais, as pesquisas experimentais utilizam os dados coletados pelos sensores para a construção de modelos de classificação ou de predição. Nessa linha, (HAQUE et al., 2016b) utilizou os dados sensoriais para construir um modelo híbrido de árvore de decisão e regressão logística. Esse modelo foi utilizado para analisar a tomada de decisão dos motoristas distraídos ao se depararem com o sinal amarelo.

Em relação aos trabalhos experimentais naturalísticos, observamos que geralmente utilizam o conjunto de sensores disponíveis nos *smartphones* para capturarem dados e para o desenvolvimento de modelos. Os estudos naturalísticos que analisam os efeitos do comportamento distraído costumam capturar dados a partir de sensores como acelerômetro, magnetômetro, giroscópio, *touch screen* e GPS. Trabalhos como os de (LIU et al., 2017), (YE et al., 2017), (KARATAS; LI; LIU, 2016) e (TCHANKUE; WESSON; VOGTS, 2013) são exemplos de pesquisas que utilizaram esses sensores para o desenvolvimento de modelos de aprendizagem com a capacidade de classificar ou detectar a distração do motorista em contextos específicos.

Nos trabalhos naturalísticos que analisam a distração de forma macroscópica, além desses sensores, é comum também observarmos a utilização de câmeras para examinar o comportamento do motorista. Nesse contexto, (YOU et al., 2013) desenvolveu uma máquina de vetor de suporte para detectar a sonolência e distração de motoristas. Na ocasião (YOU et al., 2013) capturou os dados de entrada utilizando os sensores acelerômetro, giroscópio e magnetômetro de um *smartphone* e também utilizou a câmera frontal e traseira para analisar o comportamento do motorista.

As pesquisas que analisam a distração de motorista de forma microscópica utilizam sensores como *eye tracking*, *bio-signal*, *electrocardiogram* e *galvanic skin response*. Tais sensores tem como características coletarem dados de partes específicas do corpo. Trabalhos como os de (RAJENDRA; DEHZANGI, 2017), (XIAO; FENG, 2016), (PRZYBYLA et al., 2015) e (YUN et al., 2014) são exemplos de pesquisas que utilizaram esses sensores para a análise microscópica da distração de motorista.

Em particular, a pesquisa realizada por (RAJENDRA; DEHZANGI, 2017) utilizou o sensor *galvanic skin response* para capturar atividades eletrodérmicas em diversos momentos de distração. Os participantes foram submetidos a tarefas de distração de realizar chamada telefônica, digitar mensagem de texto, conversar com o passageiro, ouvir música e testes de memória. Utilizando os dados coletados, (RAJENDRA; DEHZANGI, 2017) desenvolveu uma máquina de vetor de suporte para realizar a classificação binária entre motoristas distraídos e não distraídos.

De todos os trabalhos filtrados nessa revisão da literatura, observamos que, in-

dependentemente do tipo de estudo realizado, os sensores presentes nos *smartphones* costumam ser os sensores mais utilizados nas pesquisas que investigam a distração de motorista. Acelerômetro, magnetômetro, giroscópio, GPS e câmera são sensores presentes em pesquisas macroscópicas, microscópicas e de efeitos da distração do motorista.

2.3.5 Questão de pesquisa 4

Questão 4: Quais são as técnicas utilizadas para a análise ou detecção da distração do motorista?

O objetivo de investigar as técnicas utilizadas para a análise ou detecção da distração do motorista está voltado, principalmente, para os trabalhos de cunho experimental que utilizaram algum modelo de aprendizado de máquina. A Tabela 2 apresenta todas as técnicas de aprendizado de máquina utilizadas nos diferentes trabalhos filtrados nesta revisão. Conforme observado na Tabela 2, dos trabalhos que desenvolveram algum modelo de aprendizado de máquina, modelos fundamentados em *support vector machine*, redes neurais e regressão logísticas foram as escolhas mais frequentes.

Tabela 2. Modelos de aprendizado de máquina utilizados pelos artigos desta revisão.

Modelos de utilizados	**Quantidade de trabalhos**
Support Vector Machine (SVM)	13
Artificial Neural Network (ANN)	7
Logistic Regression	7
K-Nearest Neighbor (KNN)	4
Adaptive Boosting (AdaBoost)	2
Convolutional Neural Network (CNN)	2
Decision Tree	2
Naive Bayesian	2
Recurrent Neural Networks (RNN)	2
Extreme Learning Machine (ELM)	1

Na Tabela 3 podemos observar os trabalhos experimentais que desenvolveram modelos de *support vector machine*. Dos 13 trabalhos, 11 realizaram experimentação naturalística e 2 realizaram os experimentos simulados. Esse cenário reforça a hipótese de custo e credibilidade dos experimentos naturalísticos já discutidos na seção 2.3.2.

Entre os sensores utilizados para a coleta de dados, a câmera está presente em quase todos os trabalhos que desenvolveram modelos de SVM. Sua utilização predominou em pesquisas macroscópica e microscópica. Nas pesquisas microscópicas também podemos perceber a utilização do *eye tracker*. Trabalhos como os de (XIAO; FENG, 2016) e

(CHUANG et al., 2014) utilizaram câmeras para analisar a distração do motorista a partir de padrões de olhares distraídos.

Em particular, (XIAO; FENG, 2016) utilizou a câmera frontal de um *smartphone* para detectar a pupila do motorista bem como rastrear a direção do olhar. Nesse estudo (XIAO; FENG, 2016) definiu regiões seguras para onde o motorista pudesse estar olhando. A correlação do olhar e da segurança foi definida em conjunto com a imagem capturada pela câmera traseira do *smartphone*. Dada a direção do olhar e a imagem capturada, (XIAO; FENG, 2016) desenvolveu uma SVM para determinar se a região do olhar condizia com uma distração ou não.

Em relação aos trabalhos que desenvolveram modelos de redes neurais, podemos observar na Tabela 4 que estes, predominantemente, exploraram o método de estudo que analisa os efeitos da distração do motorista. Pesquisas dessa natureza normalmente utilizam sensores que capturam os dados do veículo para que assim possam identificar padrões que distingam a condução normal de uma condução distraída. Sensores como acelerômetro, giroscópio, magnetômetro, *throttle position, brake pedal position* e *speed controller* costumam ser utilizados em estudos dessa natureza.

Os trabalhos de (IM et al., 2014), (OU; LIU; SHIH, 2013), (TANGO; BOTTA, 2013) e (ERSAL et al., 2010) são exemplos de estudos experimentais simulados que desenvolveram modelos de redes neurais para investigar a distração do motorista, pelo método de observação dos efeitos do comportamento distraído. Nessa linha, (IM et al., 2014) utilizou sensores para coletar a velocidade, aceleração e movimentação do volante de experimentos simulados. A partir dos dados coletados, (IM et al., 2014) projetou uma arquitetura de rede neural binária para distinguir os padrões de comportamento distraído do comportamento normal.

O estudo de (YE et al., 2017) foi outro trabalho que desenvolveu uma rede neural para analisar a distração do motorista. A rede neural construída por (YE et al., 2017) foi projetada para classificar as distrações em decorrência de uma chamada telefônica, composição de mensagem de texto e conversa com o passageiro. Os resultados desse estudo naturalístico demonstraram maior exatidão nos tipos de distração analisadas, quando comparadas com os resultados do trabalho de (IM et al., 2014).

A Tabela 5 apresenta os trabalhos filtrados que fizeram uso da regressão logística para a análise da distração do motorista. Podemos observar que esta técnica foi utilizada tanto para estudos experimentais, quanto para estudo observacionais. Também podemos observar na Tabela 27, do Apêndice A, que, em geral, esta técnica foi utilizada em conjunto com outras técnicas de aprendizado de máquina.

A utilização da regressão logística combinada com outros modelos de aprendizado de máquina pode acontecer por meio de comparação ou de composição de modelos. Se

Tabela 3. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos SVM.

Id	**Tipo**	**Sensor**	**Título**
10	Naturalístico	Câmera e *CAN BUS*	*Driver Distraction Detection With A Camera Vision System*
18	Naturalístico	Câmera e IMU	*CarSafe App: Alerting Drowsy and Distracted Drivers using Dual Cameras on Smartphones*
20	Naturalístico	Câmera	*A Pattern Recognition System for Detecting Use of Mobile Phones While Driving*
38	Naturalístico	IMU	*Toward Detection of Unsafe Driving with Inertial Head-Mounted Sensors*
47	Simulação	*speed, accelerometer, throttle and brake pedal position*	*Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks*
48	Naturalístico	*microphone*, câmera e *brake pedals*	*Detecting Drivers? Mirror-Checking Actions and Its Application to Maneuver and Secondary Task Recognition*
61	Naturalístico	Câmera	*Estimating Gaze Direction of Vehicle Drivers using a Smartphone Camera*
67	Naturalístico	Câmera, Microfone, *CAN Bus*	*Analysis of Facial Features of Drivers Under Cognitive And Visual Distractions*
68	Naturalístico	*Eye Tracker*	*Driver Distraction Detection Using Semi-Supervised Machine Learning*
71	Naturalístico	*wearable sensor network*, *Galvanic Skin Response* (GSR)	*Detection of Distraction under Naturalistic Driving Using Galvanic Skin Responses*
77	Simulação	Câmera, *speed, brake pedal position, steering wheel, lane change*	*Real-Time Detection System of Driver Distraction Using Machine Learning*
81	Naturalístico	*Eye Tracker* e câmera	*Detection of Drivers Visual Attention Using Smartphone*
84	Naturalístico	*Eye Tracker*, câmera	*Video-Based Detection and Analysis of Driver Distraction and Inattention*

Tabela 4. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos de redes neurais.

Id	Tipo	Título	Método
35	Naturalístico	*Detection of driver engagement in secondary tasks from observed naturalistic driving behavior*	Efeitos da distração
41	Simulação	*Risk prediction model for drivers in-vehicle activities: Application of task analysis and back-propagation neural network*	Efeitos da distração
47	Simulação	*Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks*	Efeitos da distração
49	Simulação	*Driver Distraction Detection By In-Vehicle Signal Processing*	Efeitos da distração
53	Naturalístico	*Driver Identification and Driver's Emotion Verification using KDE and MLP Neural Networks*	Efeitos da distração
77	Simulação	*Real-Time Detection System of Driver Distraction Using Machine Learning*	Efeitos da distração
85	Naturalístico	*The Study of Driver Distraction Characteristic Detection Technology*	Macroscópica

utilizada para a comparação de modelos, por ser um modelo simples, geralmente a regressão logística é utilizada como o modelo *baseline*. Ou seja, se um estudo está comparando o desempenho de diferentes modelos de aprendizado de máquina frente à resolução de um problema, espera-se que os outros modelos tenham desempenho superior ao modelo de regressão logística. Se esse não for o caso, então constata-se que, pela facilidade em interpretação dos resultados e de processamento, o modelo de regressão logística é a melhor escolha.

Se for utilizada para a composição de modelos, a regressão logística pode ser utilizada como parte de um modelo *ensemble* ou como subparte de uma arquitetura que utiliza diferentes modelos de aprendizado de máquina para finalidades específicas. Por exemplo, (LI; BUSSO, 2013b) realizou experimentos naturalísticos para analisar a distração do motorista, a partir da extração de características faciais. Nesse estudo, (LI; BUSSO, 2013b) desenvolveu diferentes modelos de aprendizado de máquina para distinguir as diferenças das distrações visuais e cognitivas, a partir de características faciais. Nesse estudo (LI; BUSSO, 2013b) utilizou a regressão logística para filtrar as expressões faciais que, de fato, fossem importantes para serem submetidas como dados de entrada dos modelos de aprendizagem.

Outro trabalho que também utilizou a regressão logística de forma combinada foi o de (HAQUE et al., 2016b). Nesse caso, (HAQUE et al., 2016b) criou um modelo *ensemble* utilizando modelos de regressão logística e árvore de decisão para analisar os efeitos da

Tabela 5. Lista de artigos que analisaram a distração do motorista a partir do desenvolvimento de modelos de regressão logística.

Id	Tipo	Título	Método
9	Observacional Naturalístico	*The influence of roadway situation, other contextual factors, and driver characteristics on the prevalence of driver secondary behaviors*	Macroscópico
25	Simulação	*Decisions and actions of distracted drivers at the onset of yellow lights*	Efeitos da distração
28	Observacional Simulação	*Simulation of texting impact on young drivers behavior and safety on motorways*	Efeitos da distração
29	Simulação	*Mutual interferences of driving and texting performance*	Efeitos da distração
31	Simulação e *Survey*	*Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour*	Efeitos da distração
67	Naturalístico	*Analysis of Facial Features of Drivers Under Cognitive And Visual Distractions*	Microscópico
72	Observacilnal Simulação	*A study of young adults examining phone dialing while driving using a touchscreen vs. a button style flip-phone*	Microscópico

distração do motorista, ao passar pelo sinal amarelo. O modelo *ensemble* foi construído de forma iterativa, tal que a árvore de decisão era responsável em discriminar as variáveis indicadoras para o modelo de regressão, e o modelo de regressão era responsável em inferir os dados significativos para podar a árvore de decisão. A iteração do modelo proposto por (HAQUE et al., 2016b) era processada até se obter uma especificação justificável do modelo de regressão logística, teoricamente e praticamente atraente.

Das demais técnicas, vale ressaltar os trabalhos de (LE et al., 2016) e (VORA; RANGESH; TRIVEDI, 2017) que desenvolveram modelos de redes neurais convolucionais. A partir da publicação das redes *AlexNet* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), as redes neurais convolucionais tem sido utilizadas e modificadas para vários problemas no escopo da visão computacional, e tem demonstrado melhorias significativas na tarefa de classificação de imagens.

Nesse contexto, (LE et al., 2016) propôs uma *Faster R-CNN* (FRCNN) para determinar se um motorista está usando um telefone celular, bem como detectar se suas mãos estão no volante. Na arquitetura desenvolvida, (LE et al., 2016) primeiro utilizou a FRCNN para detectar as regiões de interesse como as mãos, celular e o volante. Em seguida, a informação geométrica foi extraída para determinar se as mãos estão corretamente no volante ou se o motorista está manuseando o telefone celular. No referido trabalho, (LE et

al., 2016) afirma que além do bom desempenho, o desenvolvimento da FRCNN possibilitou o monitoramento da distração de motorista de forma não intrusiva.

Também podemos observar na Tabela 27, do Apêndice A, que outras técnicas foram utilizadas nos estudos experimentais. Dentre elas, podemos destacar a utilização das cadeias de *Markov* nos trabalhos de (BO et al., 2013), (SCHMITT et al., 2016) e (BO et al., 2017), e também o desenvolvimento de algoritmos *threshold* em 9 outros trabalhos. No contexto desta revisão, classificamos como algoritmos *threshold*, os métodos que realizam detalhado tratamento e análise dos dados e, então, definem valores de referência estáticos ou dinâmicos para uma determinada tomada de decisão.

Por exemplo, (WANG et al., 2016) realizou um estudo para inferir a utilização do telefone celular por motoristas e passageiros. No estudo em questão, (WANG et al., 2016) analisou os padrões da aceleração centrípeta no momento em que o veículo realiza uma curva. Ao analisar esta variável, (WANG et al., 2016) observou que os valores da aceleração centrípeta se diferem quando os dados são coletados pelo celular do passageiro ou do motorista. Assim, a ideia central desse estudo consistiu em analisar a aceleração centrípeta em diferentes cenários. Na análise levou-se em consideração diferentes tipos de curva, de veículo, de velocidade e das posições de coleta dentro do veículo. Após os experimentos, (WANG et al., 2016) concluiu que, se a aceleração centrípeta é menor do que a aceleração central, ao realizar uma curva para a esquerda, ou maior ao realizar uma curva para a direita, então o dispositivo de coleta de dados encontra-se do lado esquerdo do veículo.

Na revisão realizada, podemos perceber que os estudos que optaram pela proposta de um algoritmo *threshold* necessitam analisar os dados minuciosamente, para encontrar um padrão de inferência. Contudo, tal padrão revela-se altamente dependente do contexto. No cenário da distração de motorista, a dependência de contexto pode ser considerada um fator de implicação para a generalização da solução. Entendemos que a opção por técnicas mais sofisticadas como, por exemplo, as de aprendizado de máquina, são mais adequadas para os estudos que investigam a distração do motorista.

2.4 Considerações finais

Nos últimos anos, a distração de motorista tem se apresentado como um dos notáveis motivos para a ocorrência de mortes e acidentes de trânsito. Das possibilidades de distração, a distração do motorista em consequência da utilização do telefone celular chama a atenção, devido à frequente utilização e potencialidades de distração desse dispositivo. Constatados os riscos de saúde pública que podem ser ocasionados pela utilização dos telefones celulares enquanto se está dirigindo, diversos países criaram leis que restringem ou proíbem seu uso. Contudo, como observado nos relatórios publicados pela OMS, as campanhas educativas e

a existência de leis restritivas apresentam-se como medidas insuficientes para combater a utilização do telefone celular durante a condução de veículos.

Diante do cenário apresentado, diversas pesquisas estão sendo realizadas para investigar as causas e efeitos da distração do motorista, bem como para desenvolver soluções de prevenção. Nesse contexto, entendendo que esta tese tem como objetivo investigar e propor soluções para distração do motorista, em função da utilização do telefone celular, realizou-se a revisão da literatura para identificar os avanços na área, bem como identificar as oportunidades de pesquisa.

A partir da revisão realizada, classificamos os trabalhos filtrados em: *surveys*, pesquisas observacionais e experimentais. Os trabalhos que realizaram *survey* aplicaram questionários com o objetivo de entender o perfil do motorista e os motivos da sua distração. Os resultados dos *surveys* demonstram que o perfil dos motoristas é influenciado por variáveis como sexo, idade, contexto social e geográfico. A particularidade do perfil de motorista por região geográfica revela que as pesquisas de prevenção da distração devem levar em consideração essa variável. Por exemplo, se em uma determinada região foi observado que os jovens do sexo masculino são os motoristas potencialmente suscetíveis para a distração, devido à utilização do telefone celular, então campanhas educativas ou até mesmo soluções computacionais voltadas para esse público, podem ser priorizadas.

No que tange ao cenário das pesquisas observacionais, percebemos que estas tem como objetivo entender e contextualizar o comportamento do motorista distraído. Para isso, vários trabalhos realizaram observações naturalísticas ou simuladas. As pesquisas naturalísticas dessa natureza possuem, como característica, a observação em ponto fixo e em grande quantidade. Já as pesquisas que utilizam o simulador, por possuírem controle sobre as variáveis, realizam observações em contextos específicos. Nesse tipo de pesquisa trabalhos realizaram observações do comportamento distraído em diferentes cenários, por exemplo, ao passar pelo sinal amarelo, no momento de uma freada, ao realizar uma curva ou no momento de uma mudança de faixa. Observações como essas geram informações relevantes para entender o contexto e os reflexos ocasionados pela direção distraída.

As informações geradas, tanto nos trabalhos de *survey* quanto nos trabalhos observacionais, são informações importantes para a realização dos trabalhos experimentais. Os trabalhos de pesquisa experimental foram realizados a partir dos métodos macroscópico, microscópico e da análise dos efeitos do comportamento distraído. Os trabalhos que experimentaram a partir do método macroscópico tiveram como característica o desenvolvimento de soluções que analisaram o comportamento do motorista, em uma perspectiva global. Devido a esta característica, a câmera acoplada ao veículo foi o sensor mais utilizado nesse tipo de experimento. Parte da literatura costuma criticar as soluções macroscópicas devido as suas características de intrusão. Entretanto, os resultados demonstram que as soluções macroscópicas são eficientes, e que a maior parte dos motoristas aceitariam

utilizar tecnologias macroscópicas de prevenção da distração.

As pesquisas experimentais microscópicas apresentam-se como alternativa às pesquisas macroscópicas para eliminar a crítica da intrusão da coleta de dados. Identificamos que as pesquisas que utilizam o método microscópico, além de utilizarem as câmeras, também utilizam sensores como *eye tracker*, microfone, *galvanic skin response* e *electrocardiogram*. Tais sensores são utilizados para capturar informações do motorista, de forma não intrusiva. Dos trabalhos filtrados, observamos que as pesquisas microscópicas que investigam a distração do motorista a partir da análise do olhar são mais frequentes. Em relação ao método desenvolvido, as pesquisas microscópicas são mais complexas, quando comparadas com as pesquisas macroscópicas. Métodos de pré-processamento e refinamento dos dados coletados são mais explorados naquele tipo de pesquisa. Outra característica notada em pesquisas microscópicas foi a preocupação em contextualizar a distração investigada. Além de analisar microscopicamente o comportamento do motorista, os métodos propostos também estão preocupados em correlacionar esse comportamento com o contexto do trânsito.

A correlação da distração com o contexto do trânsito também é objeto de estudo das pesquisas que estudam a distração do motorista a partir da análise dos efeitos da distração. Estudos dessa natureza costumam utilizar sensores para coletar as informações do veículo e, então, investigar os padrões existentes quando a condução está sendo realizada de forma distraída. Sensores como acelerômetro, giroscópio, magnetômetro e GPS foram os sensores mais utilizados pelos artigos filtrados nesta revisão. Estudos naturalísticos que analisam os efeitos do comportamento distraído ganharam popularidade, em virtude dos principais sensores utilizados neste método de pesquisa estarem encapsulados nos *smartphones*. Nesse contexto, os *smartphones*, um dos principais motivos da distração, também se tornaram uma das principais ferramentas para a coleta de dados.

De uma maneira geral, ao finalizar esta revisão, concluímos que a distração do motorista é um tema de saúde pública que está sendo estudado em diversos países do mundo. Diferentes tipos de estudos, métodos e tecnologias estão sendo empregados. Em relação às técnicas utilizadas nos estudos experimentais, constatamos que as técnicas de aprendizado de máquina possuem potencial importância para essa linha de pesquisa, porém ainda não foram exploradas suficientemente. Identificada essa lacuna, o presente trabalho tem como objetivo desenvolver e comparar diferentes técnicas de aprendizado de máquina, para investigar e contextualizar o problema da distração do motorista.

3 Survey do perfil de distração dos motoristas

Este capítulo apresenta a condução do *survey* realizado no estado do Pará. O referido *survey* tem como objetivo investigar os perfis de distração dos motoristas paraenses em relação à utilização do telefone celular. As distrações devido à realização de chamadas telefônicas; composição de mensagens de texto; envio de mensagens de áudio e leitura de mensagens de texto foram levadas em consideração nessa pesquisa. Os resultados demonstram que os motoristas paraenses possuem o hábito de utilizar o telefone celular enquanto estão dirigindo e, por esse motivo, comprovam que somente as campanhas educativas e as legislações proibitivas não são medidas suficientes para prevenir a distração dos motoristas em decorrência da utilização do telefone celular.

O decorrer do capítulo está dividido em quatro principais seções. Na primeira seção é discutida a motivação para a realização desse *survey*. Na segunda, apresentamos o planejamento e a forma como o *survey* foi conduzido. Na terceira seção são apresentados os dados quantitativos e as respectivas análises. Por fim, na última seção, são apresentadas as considerações finais desse estudo.

3.1 Motivação

Na revisão da literatura apresentada no Capítulo 2, foram filtrados 9 artigos que realizaram um *survey* para investigar a distração do motorista. Os estudos desses artigos foram conduzidos nos Estados Unidos (1), Canadá (1), Espanha (1), Colômbia (1), Austrália (1), Catar (1), Índia (2) e Finlândia (1). A Tabela 6 lista os 9 *surveys* filtrados na revisão.

Esses *surveys* tiveram como objetivo investigar as causas, consequências e tipos de distrações dos motoristas que ocorrem devido à utilização do telefone celular. Nós observamos que os *surveys* realizados em países distintos apresentaram resultados diferentes que podem ter sido influenciados pela cultura local. Por exemplo, na Espanha, no *survey* conduzido por (GRAS et al., 2007), mais de 60% dos entrevistados admitem utilizar o telefone celular enquanto estão dirigindo. Por outro lado, no Catar, segundo (SHAABAN, 2013), constatou-se que menos de 20% dos entrevistados utilizam o telefone celular enquanto estão dirigindo. Esta diferença demonstra que as diferenças culturais influenciam no comportamento dos motoristas. Nesse sentido, realizar um *survey* para entender as opiniões e atitudes a partir de uma perspectiva local, parece ser importante para o desenvolvimento de soluções customizadas. No caso da distração do motorista, se um

Tabela 6. *Surveys* filtrados na revisão da literatura apresentada no Capítulo 2.

País	Título	Autor
Estados Unidos	*Patterns of distracted driving behaviors among young adult drivers: Exploring relationships with personality variables*	(BRAITMAN; BRAITMAN, 2017)
Canadá	*Self-reported engagement in driver distraction: An application of the Theory of Planned Behaviour*	(CHEN et al., 2016)
Espanha	*Mobile phone use while driving in a sample of Spanish university workers*	(GRAS et al., 2007)
Colômbia	*Mobile phone use while driving: A hybrid modeling approach*	(MÁRQUEZ; CANTILLO; ARELLANA, 2015)
Austrália	*Exploring Boredom Proneness as a Predictor of Mobile Phone Use in the Car*	(SCHROETER et al., 2015)
Catar	*Investigating Cell Phone Use While Driving in Qatar*	(SHAABAN, 2013)
Índia	*Modeling safety risk perception due to mobile phone distraction among four wheeler drivers*	(RAJESH et al., 2017)
Índia	*Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour*	(CHOUDHARY; VELAGA, 2017b)
Finlândia	*Context-sensitive distraction warnings – Effects on drivers' visual behavior and acceptance*	(TUOMO; HANNU; JAKKE, 2016)

survey identifica que uma parte significativa dos entrevistados costuma ler mensagens de texto enquanto estão dirigindo, então, realizar pesquisas e desenvolver soluções para contornar essa situação é significativo para a região onde o *survey* foi realizado.

Considerando a importância da investigação local, esta tese conduziu um *survey* no estado do Pará, localizado na região norte do Brasil, para investigar quais os perfis de distração dos motoristas, devido à utilização do telefone celular, são mais frequentes nessa região. Segundo os dados do IBGE (IBGE, 2016) e da Confederação Nacional de Transporte (CNT, 2018), o estado do Pará possui 144 municípios, uma área de 1.247.955 km^2, aproximadamente oito milhões de habitantes, seis mil quilômetros de rodovias pavimentadas e quase dois milhões de veículos. A Figura 9 ilustra as informações socioeconômicas do estado do Pará.

O tipo e as condições das rodovias são dois fatores que, se somadas à distração do motorista, podem aumentar significativamente a probabilidade de acidentes de trânsito.

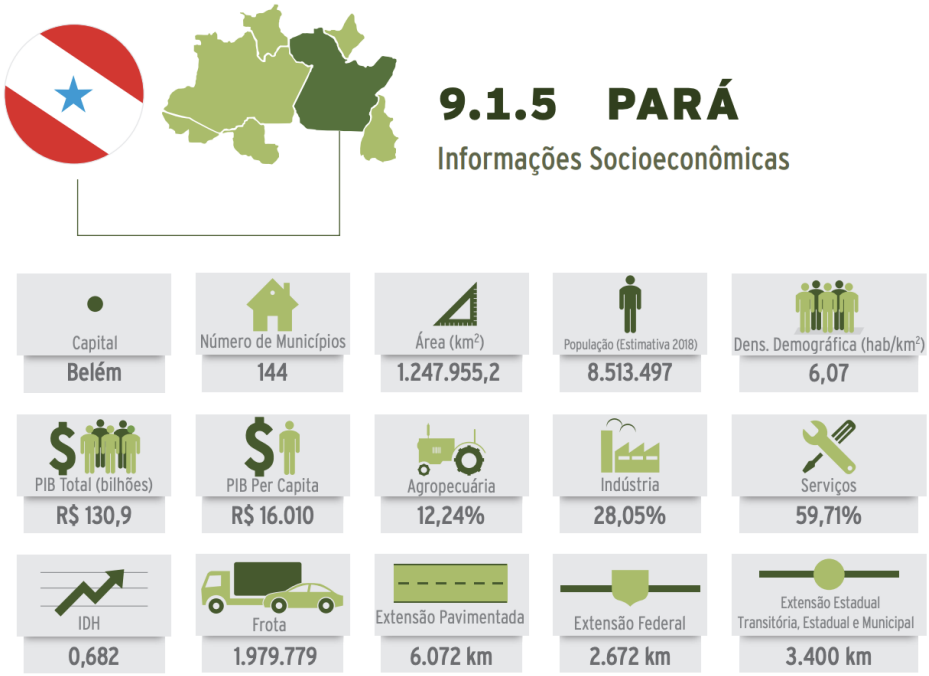


Figura 9. Informações socioeconômicas do estado do Pará (CNT, 2018).

No Pará, segundo os dados da Confederação Nacional de Transporte (CNT, 2018), 98% das rodovias são de pistas simples de mão dupla e quase 40% das rodovias tiveram o seu estado de conservação avaliados como ruim ou péssimo. Se somarmos a avaliação regular, esse quantitativo chega a quase 77%. Esses dados demonstram que as rodovias do estado do Pará são precárias e podem contribuir diretamente no quantitativo de acidentes e mortes de trânsito. Sobre esse quantitativo, segundo o relatório que retrata a segurança viária no Brasil (AMBEV, 2017), no ano de 2015, o estado do Pará, para cada 100 mil habitantes, obteve a proporção de 21 óbitos e 143 feridos. Em ambos os casos, o Pará superou a média nacional que foi de 19 óbitos e 99 feridos para cada 100 mil habitantes.

Sobre a distração do motorista, segundo os dados da Polícia Rodoviária Federal (PRF, 2018), dos 313 acidentes de trânsito registrados nas rodovias federais do estado do Pará em 2018, aproximadamente 29% foram devido à falta de atenção ao dirigir. A desatenção do motorista foi o motivo mais frequente, ficando à frente da desobediência às normas de trânsito (23%), ingestão de álcool (7%) e ultrapassagem indevida (2%). A Tabela 7 apresenta os cinco principais motivos de acidentes de trânsito registrados pela PRF no estado do Pará em 2018.

Outro dado importante que justifica a realização do *survey* diz respeito ao uso do telefone celular de uma maneira geral. Segundo dados do Ministério da Saúde, 19% da população das capitais brasileiras afirma que faz o uso do telefone celular enquanto dirige (SAÚDE, 2019). Das capitais que apresentaram maior índice, Belém, a capital do estado do Pará, foi a campeã com 24% dos motoristas afirmando que utilizam o telefone

Tabela 7. Registro de acidente de trânsito nas rodovias federais do estado do Pará em 2018 (PRF, 2018).

Motivo	Porcentagem
Falta de atenção à condução	29.39%
Desobediência às normas de trânsito pelo condutor	23.32%
Não guardar distância de segurança	9.58%
Ingestão de álcool	7.35%
Defeito mecânico no veículo	3.83%

celular enquanto dirigem. Uma vez que o Brasil é um dos países com o maior índice de acidentes e mortes devido à acidentes de trânsito, e sabendo que o estado do Pará é um dos estados brasileiros com consideráveis índices de distração devido ao uso do telefone celular, justifica-se a realização do *survey* para investigar os principais perfis de distração dos motorista provenientes da utilização do telefone celular no estado do Pará.

3.2 Condução do survey

A coleta de dados deste *survey* foi realizada no período de junho a agosto de 2018. Nesse período, nós obtivemos 692 respostas das quais 450 foram de participantes residentes do estado do Pará. Utilizamos o *linkedin* como meio de divulgação do *survey*. O questionário foi enviado para 1000 pessoas e obtivemos respostas de aproximadamente 69% dos participantes. A Tabela 8 apresenta as questões respondidas pelos participantes.

Das respostas obtidas, 65% foram de participantes residentes no estado do Pará. Por esse motivo, na análise dos resultados, nós levamos em consideração apenas as respostas dessa unidade federativa. A partir dessa restrição, a análise do *survey* teve como objetivo identificar o perfil de distração dos motoristas paraenses. A Figura 10 apresenta as informações básicas dos participantes. Como pode ser observado, quase 85% dos participantes são motoristas que dirigem carros e 15% são motoristas que dirigem motos. Em relação ao gênero, quase 75% dos participantes são do sexo masculino e 25% são do sexo feminino. A distribuição da idade revela que a maioria dos participantes está na faixa etária de 26 a 35 anos. Os motoristas mais novos representam um pouco mais de 20% dos entrevistados. Os motoristas mais experientes que estão na faixa etária de 36 a 45 anos e os que possuem mais de 45 anos, representaram 25% e 8%, respectivamente.

Se levarmos em consideração os motoristas com idade entre 18 a 35 anos, esses representam mais de 65% dos entrevistados. Segundo a Organização Mundial da Saúde (CHAN, 2018), esses motoristas estão inseridos no grupo de pessoas cuja principal causa de morte é o acidente de trânsito. Ademais, independentemente da idade, os dados do

Tabela 8. Questões que foram construídas para o *survey*.

Perguntas	Alternativas
1- Qual a sua faixa etária?	* 18-25 * 26-35 * 36-45 * mais de 45
2- Qual o seu gênero?	* masculino * feminino
3- Qual a sua unidade federativa?	Lista de todas as unidade federativas do Brasil
4- Qual o tipo de veículo que você dirige com maior frequência?	* carro * moto * caminhão * ônibus
5- Estando com o veículo em movimento, com que frequência você atende uma ligação telefônica?	* sempre * às vezes * nunca
6- Estando com o veículo em movimento, com que frequência você envia uma mensagem de texto?	* sempre * às vezes * nunca
7- Estando com o veículo em movimento, com que frequência você envia uma mensagem de áudio?	* sempre * às vezes * nunca
8- Estando com o veículo em movimento, com que frequência você lê uma mensagem de texto?	* sempre * às vezes * nunca
9- Estando com o veículo em movimento, você utiliza o celular com maior frequência para:	* ligação telefônica * ler mensagem de texto * escrever mensagem de texto * mandar áudio * não utilizo o celelular
10- Você conhece pessoas que utilizam o telefone celular enquanto estão dirigindo ?	* muitas (+10) * algumas (5-9) * poucas (1-4) * nenhuma

relatório da OMS revelam que o acidente de trânsito é a 8^a principal causa de morte no mundo. No que diz respeito ao tipo de veículo, segundo os dados do relatório que

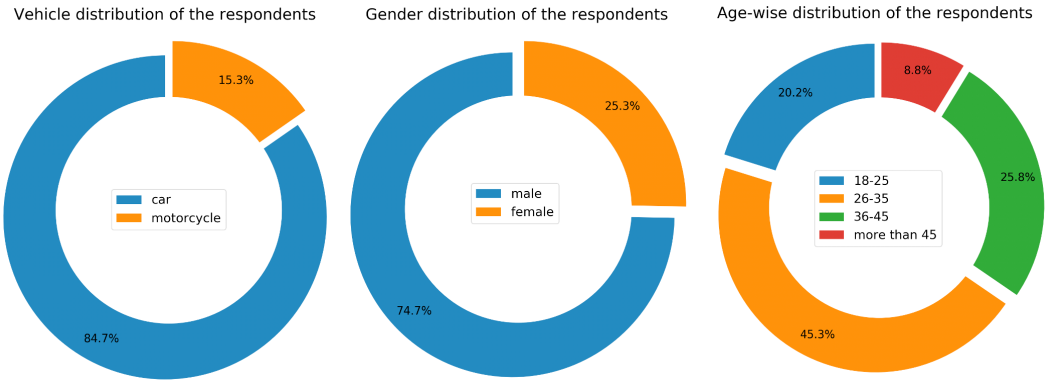


Figura 10. Informações gerais dos participantes.

retrata a segurança viária no Brasil (AMBEV, 2017), no estado do Pará, os motoristas de motocicleta são as principais vítimas de acidentes de trânsito.

3.3 Análise dos resultados

Nós utilizamos as informações básicas para analisar os quatro perfis de distração investigados. Em uma primeira análise, verificamos como cada perfil de distração se apresenta em relação ao tipo de veículo, gênero e idade. A Figura 11 apresenta a análise dos quatro tipos de distração filtradas, por tipo de veículo. Para simplificar, chamaremos de motoristas os motoristas que dirigem carro e de motociclistas os que dirigem moto. Podemos perceber que em todas as distrações investigadas, os motoristas são mais distraídos do que os motociclistas. Se somarmos a utilização frequente e a esporádica, em todos os casos, mais de 50% os motoristas admitem utilizar o telefone celular enquanto estão dirigindo. Na distração proveniente de uma chamada telefônica, mais de 67% dos motoristas admitem realiza-la enquanto estão dirigindo. Em relação ao *texting while driving*, foi observado que esse é o tipo de distração menos praticado pelos dois perfis de condutores. Sobre esse tipo de distração, quase 80% dos motociclistas afirmaram não pratica-la. Tal resultado demonstra, entre outros fatores, uma possível consciência dos motociclistas paraenses sobre os efeitos dos perigos que o evento de *texting while driving* pode proporcionar. Por outro lado, apesar de ser a distração menos praticada pelos motoristas, ainda assim, mais de 50% admite utilizar o telefone celular para digitar uma mensagem de texto, enquanto estão dirigindo.

Sobre a leitura de mensagens de texto, apesar de ser um evento que provoca as mesmas distrações que a composição de mensagens de texto, podemos perceber que o percentual de condutores que se distraem por esse motivo é maior. No caso dos motoristas, mais de 67% afirmaram já terem lido uma mensagem de texto enquanto dirigiam. Entre os motociclistas, o percentual é de um pouco mais de 33%, entretanto, os dados revelam que, tanto para motoristas quanto para motociclistas, a distração em decorrência da leitura de

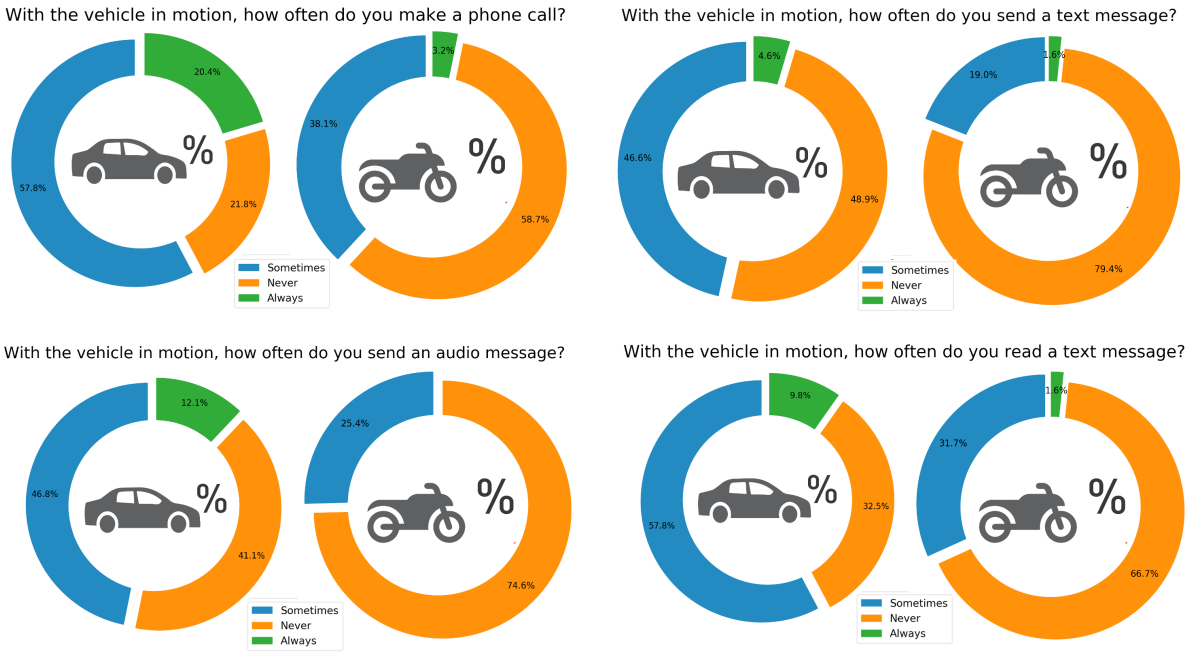


Figura 11. Análise da distração por tipo de veículo.

mensagens de texto é o segundo tipo de distração mais praticado por eles. A distração devido à chamada telefônica é a única que aparece com maior percentual. Esses dados demonstram que, apesar das distrações física, cognitiva e visual provocadas pela leitura de mensagens de texto, a maioria dos condutores ignoram os riscos e priorizam a utilização dos *smartphones*.

A Figura 12 apresenta o perfil das quatro distrações filtradas por gênero. Podemos perceber que a diferença de gênero não apresentou mudanças significativas nos perfis de distrações investigados. A maior diferença de perfil pode ser observada nas distrações devido à leitura de mensagem de texto e envio de mensagem de áudio. Na leitura de mensagem de texto, pode-se observar que quase 66% dos homens leem mensagens de texto enquanto estão dirigindo. No caso das mulheres, esse percentual cai para aproximadamente 52%. Em relação ao envio de mensagem de áudio, esse cenário se inverte, pois 60% das mulheres admitem enviar mensagem de áudio enquanto estão dirigindo enquanto que este percentual, para os homens, é de 52%. A distração devido à chamada telefônica foi a de maior percentual tanto para homens quanto para mulheres. No caso dos homens, aproximadamente 74% dos entrevistados admitem realizar uma chamada telefônica enquanto estão dirigindo. Quanto as mulheres, 68% responderam que realizam, em algum momento, chamadas telefônicas enquanto estão dirigindo. A única distração que obteve menos de 50% de admissão foi a de composição de mensagens de texto. Para esse tipo de distração, 46% dos homens responderam que escrevem mensagens de texto enquanto estão dirigindo. Para as mulheres, esse percentual foi 45%.

Em relação ao gênero, se analisarmos somente as pessoas que responderam que

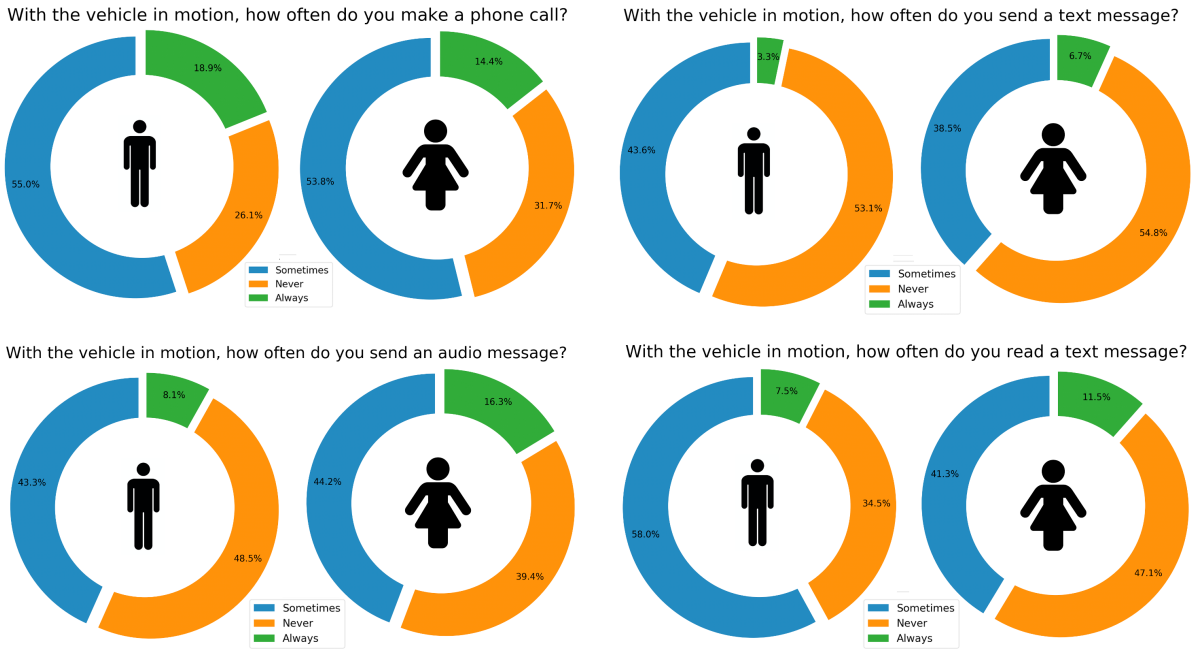


Figura 12. Análise da distração por gênero.

utilizam o celular frequentemente, podemos perceber que as mulheres se distraem mais do que os homens. Com exceção da chamada telefônica, as mulheres apresentaram maior percentual de uso frequente do celular. Essa análise é importante uma vez que, quanto maior for o tempo e a frequência do evento secundário, maiores são os riscos de acidente de trânsito, em consequência da distração do motorista. Apesar de os homens apresentarem maior percentual de uso eventual do celular, as mulheres, por apresentarem maior percentual de uso frequente, acabam se enquadrando em uma posição de maior periculosidade no trânsito.

A Figura 13 apresenta o perfil das quatro distrações filtradas por idade. Podemos perceber que, para todas as idades a distração devido à realização de uma chamada telefônica e da leitura de mensagem de texto foram as mais praticadas. Na distração devido à realização de uma chamada telefônica, com exceção da faixa etária de 18 a 25 anos, todos apresentaram percentual de utilização acima de 70%. Para os motoristas mais jovens, a taxa de admissão do uso do celular para realizar ligação telefônica foi de 64%. Em relação à leitura de mensagem de texto, 44% dos motoristas mais experientes admitem praticar esse tipo de distração enquanto estão dirigindo. Em todas as outras faixas etárias, esse percentual foi maior do que 60%. Assim como nos outros filtros, a distração menos praticada foi a de composição de mensagens de texto. A única faixa etária que apresentou mais de 50% nesse tipo de distração foi a dos motoristas de 26 a 35 anos. Esse percentual pode ser alarmante, uma vez que uma parcela dos motoristas de faixa etária entre 26 a 35 anos está no grupo de pessoas que mais morrem devido o acidente de trânsito.

As Figuras 14 a 21 apresentam a análise dos perfis de distração, a partir da

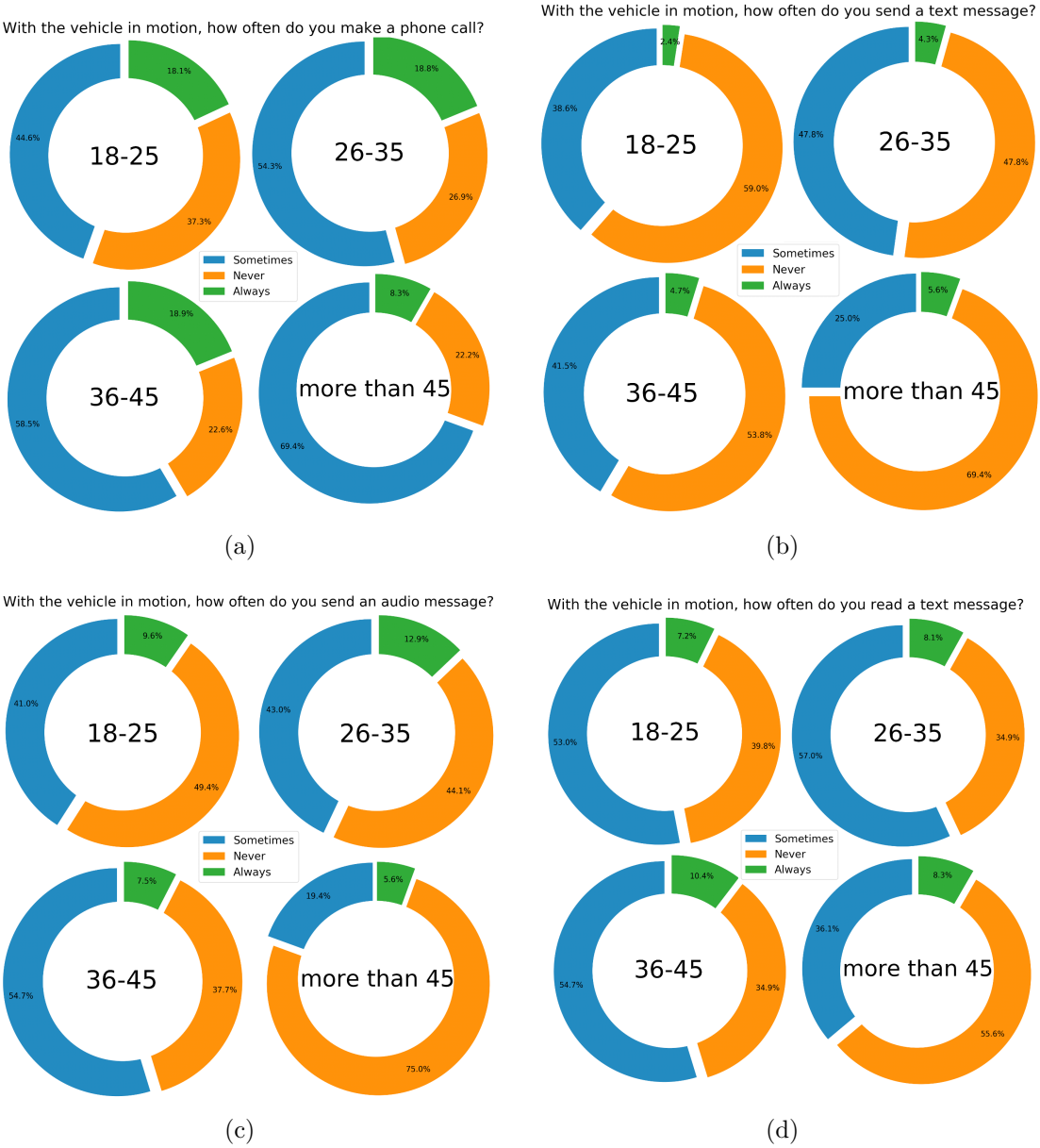


Figura 13. Análise dos tipos de distração filtrados por idade.

composição dos três filtros: idade, gênero e veículo. Durante as entrevistas, não obtivemos nenhuma resposta de motociclista com mais de 45 anos. Por esse motivo, os respectivos gráficos desse filtro não constam nas Figuras 18 a 21. A partir de uma visão geral, independentemente de idade e gênero, podemos perceber que os motociclistas são mais prudentes do que os motoristas, quando se trata de distração devido à utilização do telefone celular. Os motociclistas se distraírem menos é um fato importante, já que estão mais propensos a sofrerem acidentes por estarem dirigindo um veículo com recursos mínimos de segurança. Quanto menores forem os eventos secundários, menores são as chances de os motociclistas sofrerem acidentes de trânsito. Analisando as Figuras 18 a 21, pode-se notar que os motociclistas jovens são os que mais se distraem devido à utilização do telefone celular. A maioria dos motociclistas jovens do gênero masculino responderam que utilizam

o celular com maior frequência para ler mensagens de texto. Dentre os motociclistas jovens do gênero feminino, foi observado que a realização de uma chamada telefônica é o maior motivo de distração para esse perfil. Os motociclistas com faixa etária entre 36 a 45 anos são os mais prudentes em relação à utilização do telefone celular, senso as mulheres ainda mais precavidas do que os homens.

Ao analisar os gráficos das Figuras 14 a 17, podemos observar que o cenário da distração dos motoristas que dirigem carro é bem diferente do cenário dos motociclistas. Na distração em decorrência da chamada telefônica, os motoristas homens com faixa etária entre 36 a 45 anos são os que mais se distraem. Somando o uso eventual e o uso frequente, mais de 88% dos motoristas com esse perfil responderam que realizam uma chamada telefônica enquanto estão dirigindo. No caso das mulheres, as motoristas com faixa etária entre 26 a 35 anos são as que mais realizam chamadas telefônicas enquanto estão dirigindo. Dessas motoristas, 25% responderam que fazem ligações frequentemente e 54% eventualmente. Para ambos os gêneros, os motoristas jovens foram os mais prudentes nessa modalidade de distração, pois 26% deles responderam não realizar chamadas telefônicas durante a condução do veículo. Para as mulheres, esse percentual foi de 41%.

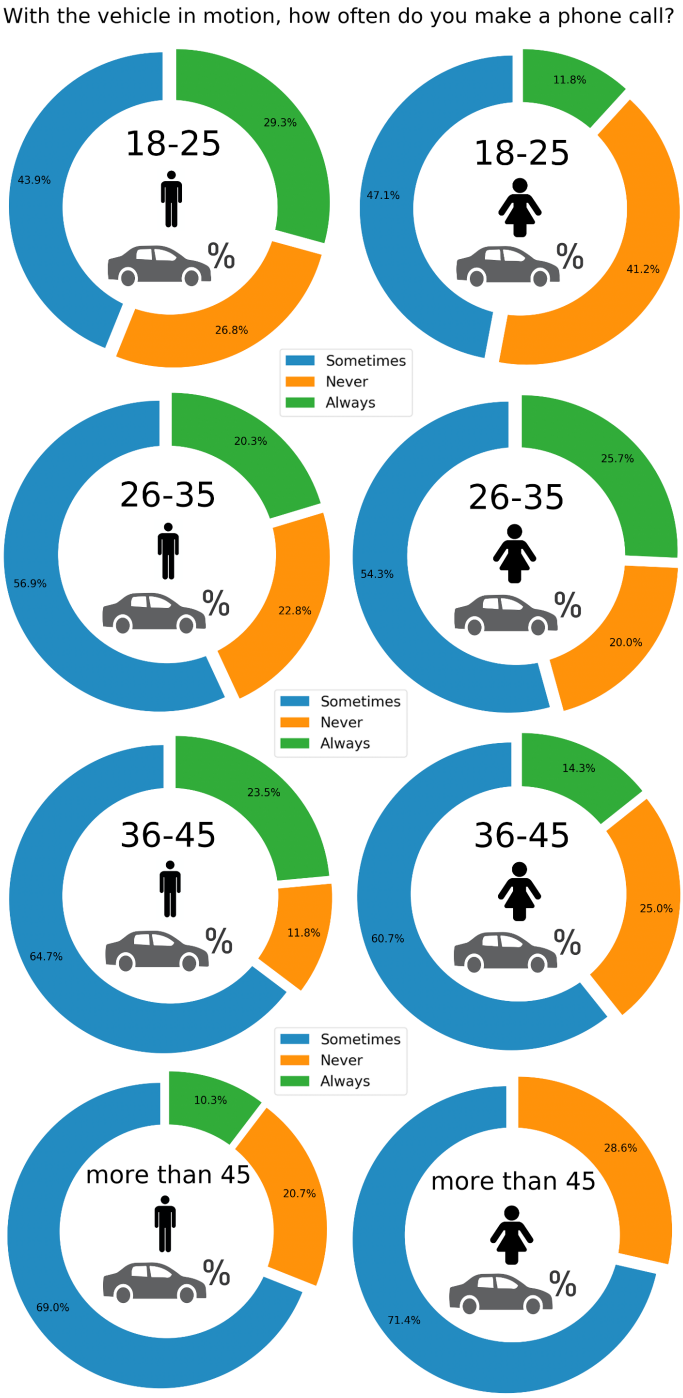


Figura 14. Análise da distração dos motoristas devido à realização de chamadas telefônicas filtradas por gênero e idade.

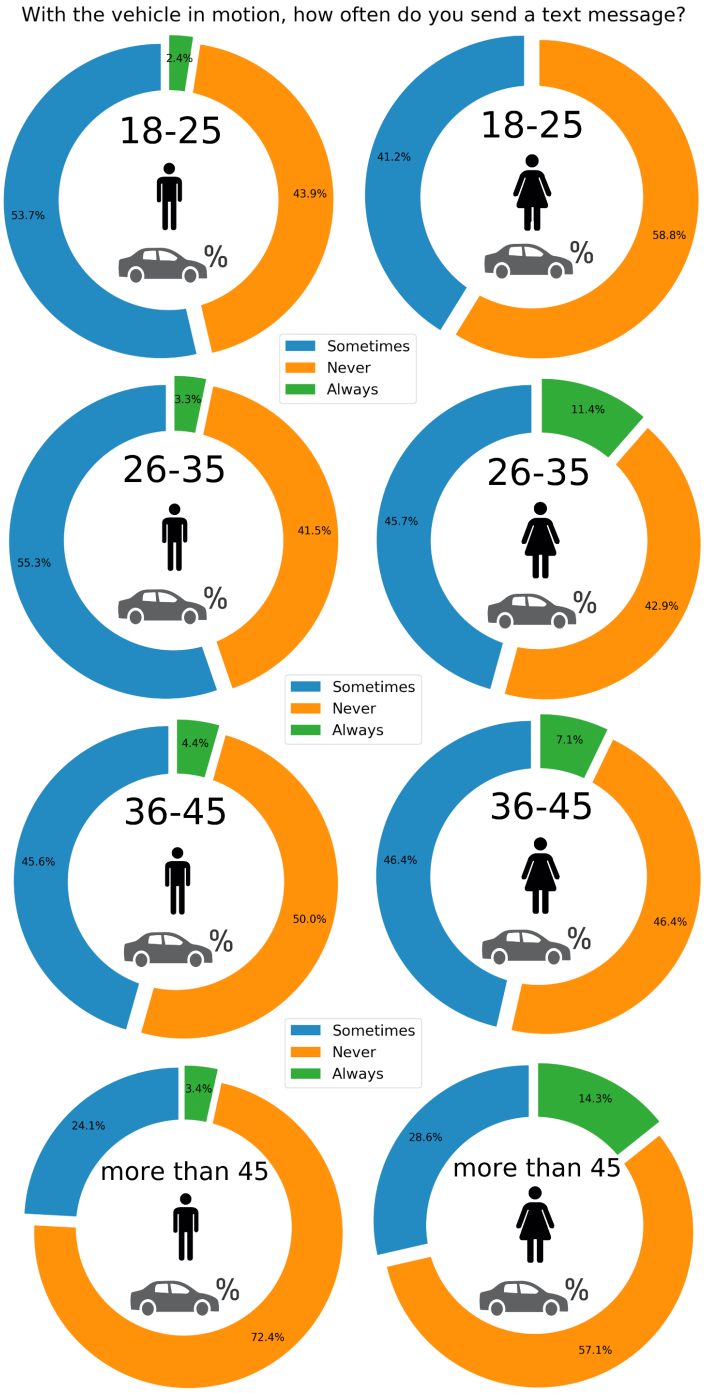


Figura 15. Análise da distração dos motoristas devido à escrita de mensagens de texto filtradas por gênero e idade.

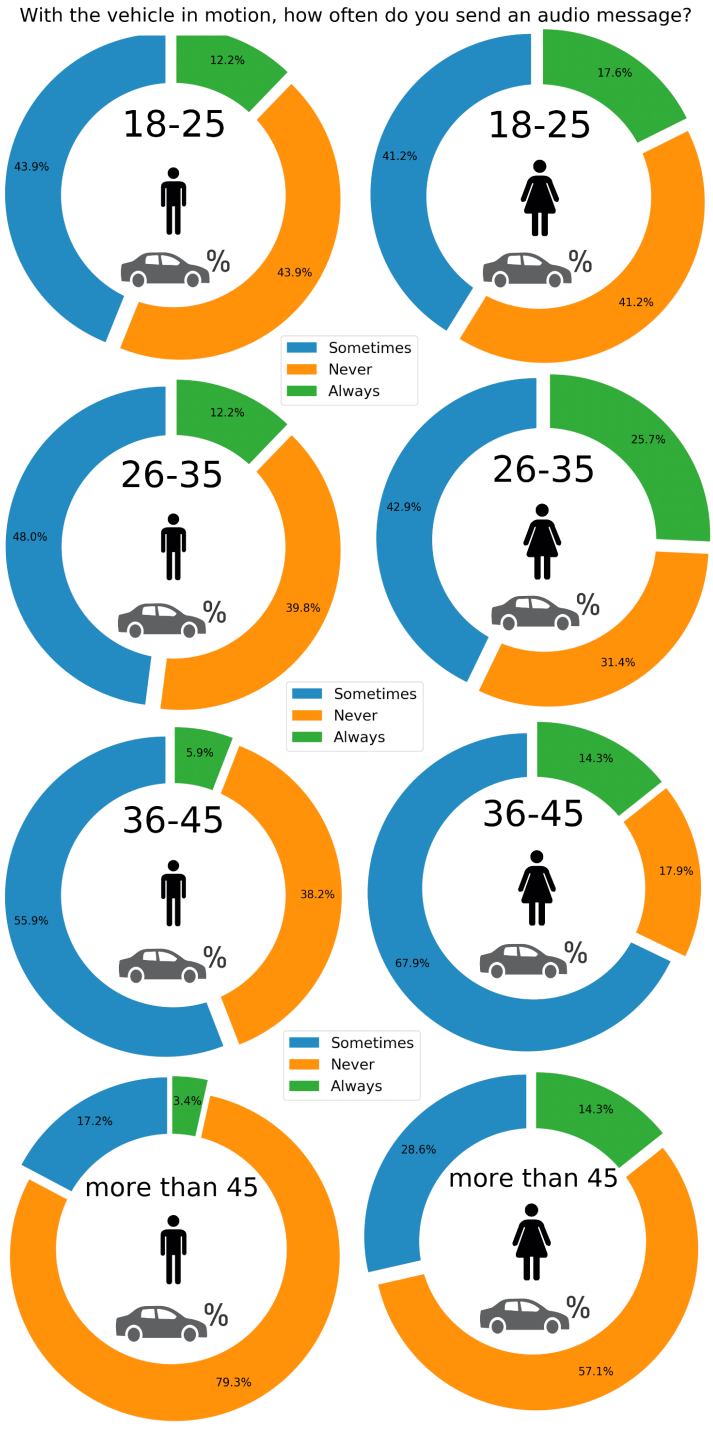


Figura 16. Análise da distração dos motoristas devido ao envio de mensagens de áudio filtradas por gênero e idade.

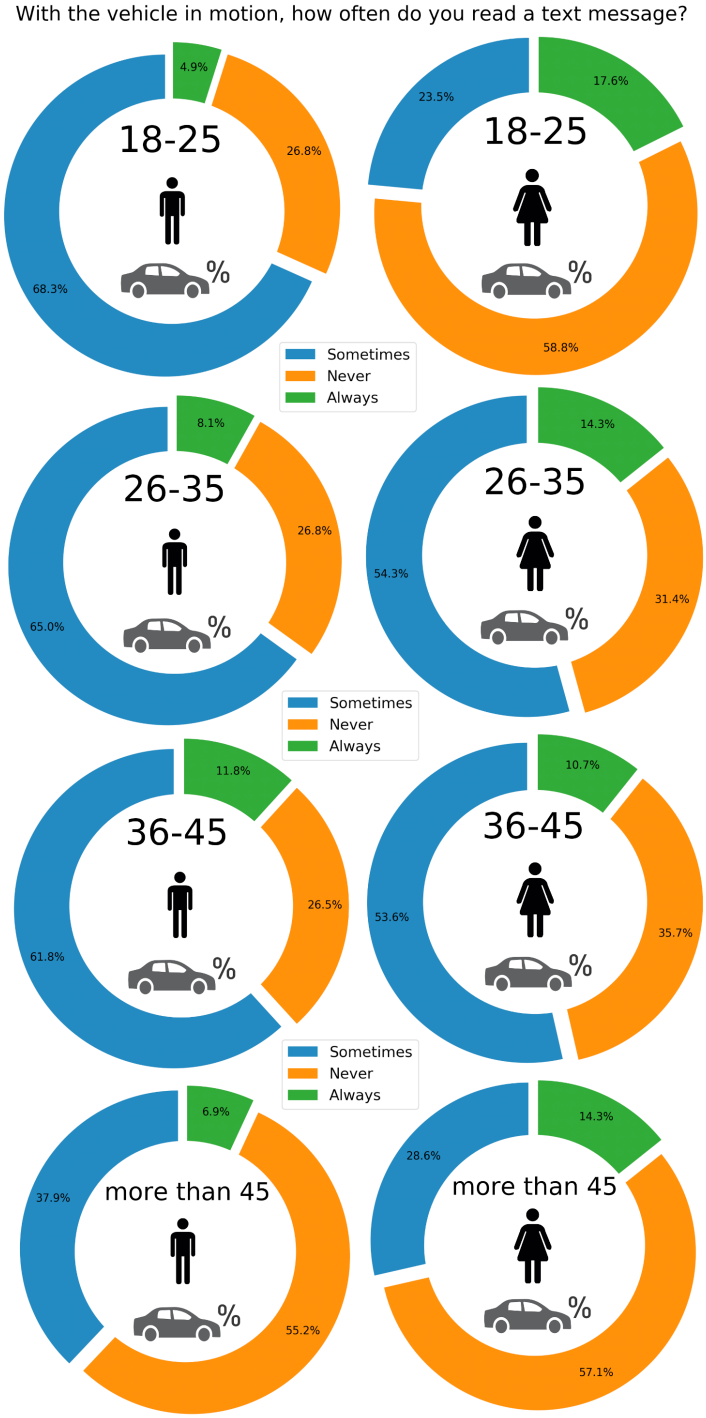


Figura 17. Análise da distração dos motoristas devido à leitura de mensagens de texto filtradas por gênero e idade.

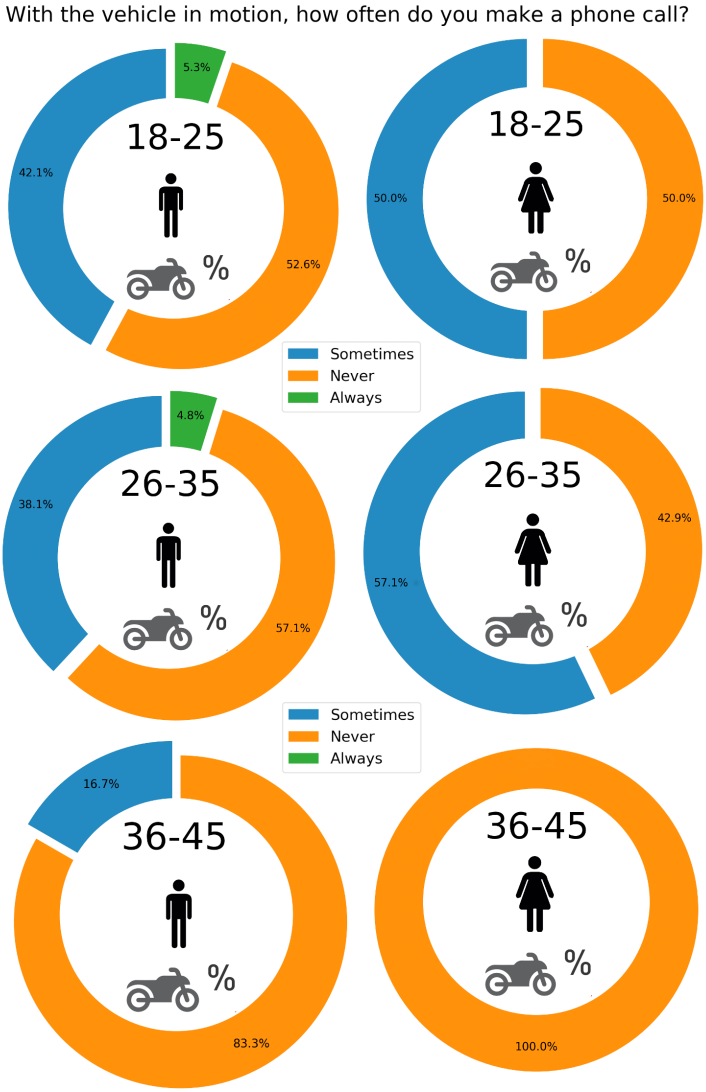


Figura 18. Análise da distração dos motociclistas devido à realização de chamadas telefônicas filtradas por gênero e idade.

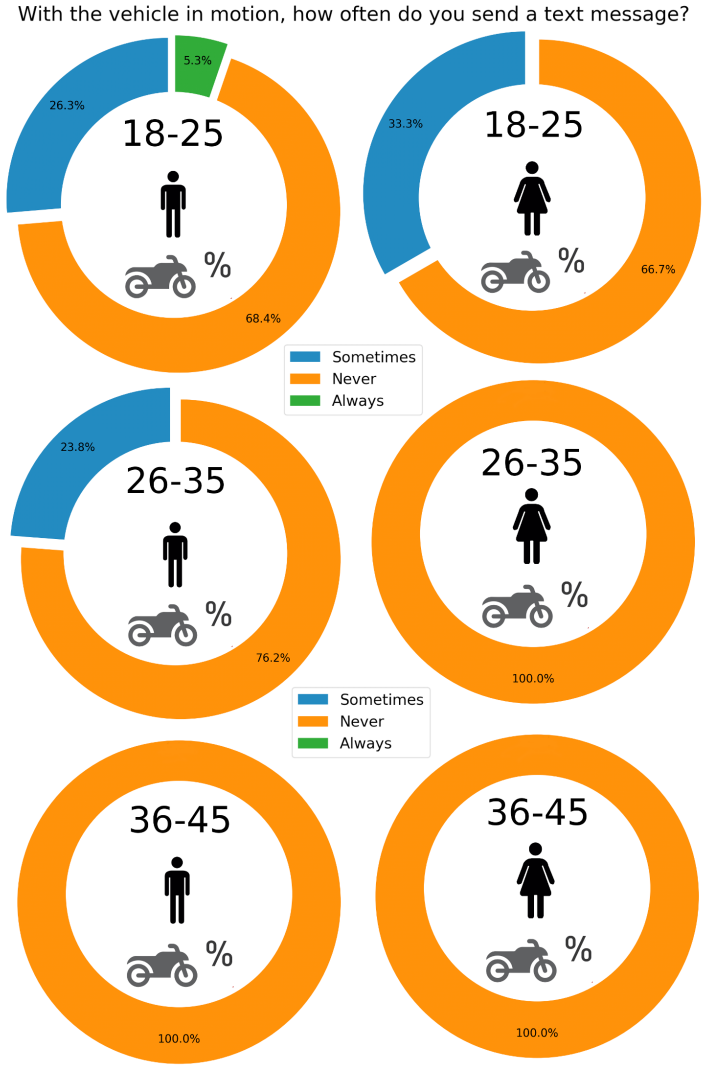


Figura 19. Análise da distração dos motociclistas devido à escrita de mensagens de texto filtradas por gênero e idade.

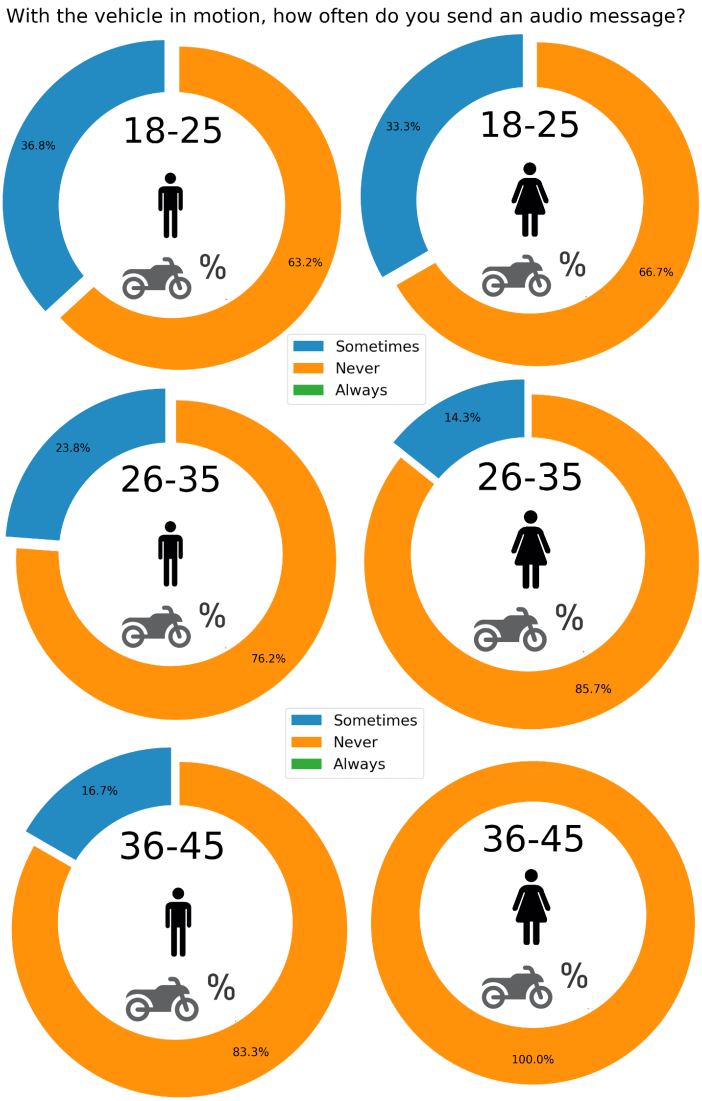


Figura 20. Análise da distração dos motociclistas devido ao envio de mensagens de áudio filtradas por gênero e idade.

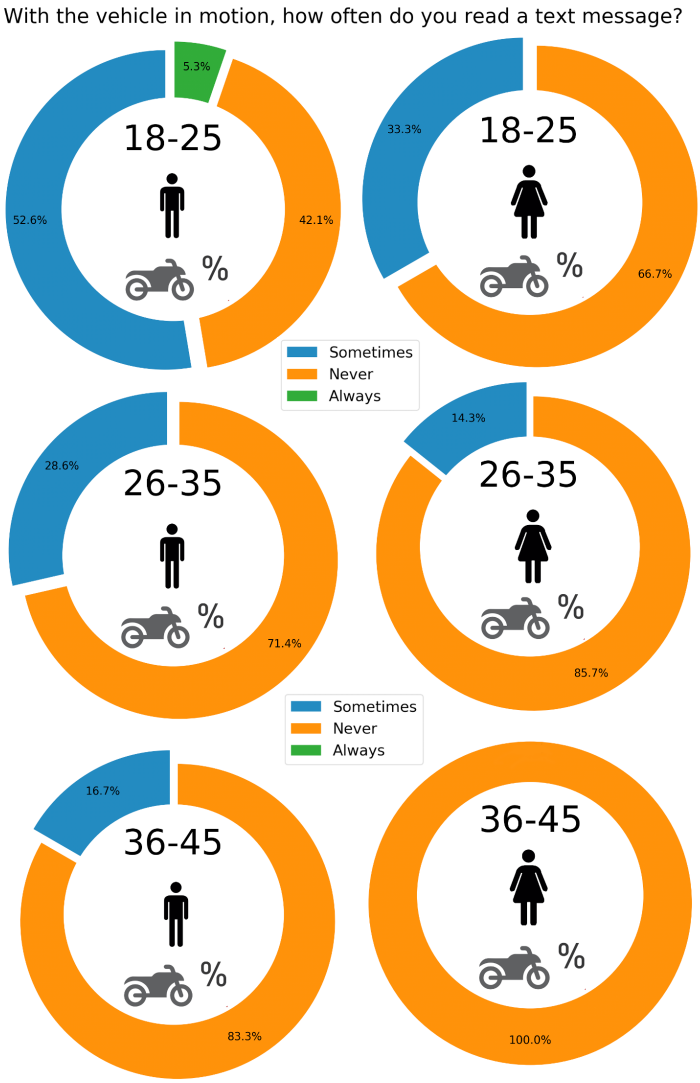


Figura 21. Análise da distração dos motociclistas devido à leitura de mensagens de texto filtradas por gênero e idade.

Em relação ao evento de *texting while driving*, aproximadamente 50% dos homens e mulheres na faixa etária entre 18 a 45 anos afirmaram digitar mensagens de texto enquanto estão dirigindo. Os motoristas com mais de 45 anos são os que menos utilizam o telefone celular para essa finalidade. Conforme os dados apresentados na Figura 15, 27% dos homens e 42% das mulheres afirmaram digitar mensagens de texto enquanto estão dirigindo. De todos os tipos de distrações investigados, a digitação de mensagens de texto foi o evento secundário menos praticados pelos motoristas. Esses dados comprovam que motoristas e motociclistas estão cientes dos perigos que este tipo de evento secundário pode proporcionar.

Na distração devido ao envio de mensagens de áudio, podemos observar na Figura 16 que as mulheres, independentemente da idade, se distraem mais que os homens. As mulheres com faixa etária entre 26 a 35 anos são as que mais se distraem utilizando o telefone celular para esta finalidade. Segundo as respostas obtidas, mais de 82% das mulheres nessa faixa etária admitem mandar mensagens de áudio eventualmente ou frequentemente enquanto estão dirigindo. Nesse tipo de distração, homens e mulheres com mais de 45 anos ou com menos de 25 anos são os que menos utilizam o telefone celular para essa finalidade. 79% dos homens, com mais de 45 anos, responderam que nunca utilizam o telefone celular para mandar mensagens de áudio enquanto estão dirigindo. Entre os homens com menos de 25 anos, esse percentual foi de quase 44%. Em relação às mulheres com mais de 45 anos, 57% das entrevistadas responderam que não praticam esse tipo de distração. Entre as mulheres com menos de 25 anos, esse percentual foi de 41%. Os motoristas mais jovens estarem entre os que menos praticam esse tipo de distração é um dado relevante, uma vez que os motoristas dessa faixa etária estão entre os que mais sofrem acidentes de trânsito com vítimas fatais. Minimizar a distração dos motoristas jovens implica diminuir os riscos de acidentes de trânsito para esse grupo.

Na Figura 17, podemos observar o perfil de distração dos motoristas em decorrência da leitura de mensagens de texto. Para esse evento secundário, os motoristas homens, independentemente da idade, se distraem mais do que as mulheres. 74% dos motoristas com idade entre 18 a 45 anos de idade responderam que leem mensagens de texto enquanto estão dirigindo. Entre os homens, depois da chamada telefônica, essa foi a segunda distração mais praticada por eles. Analisando o perfil feminino, observamos que, aproximadamente, 65% das mulheres com idade entre 26 a 45 anos costumam ler mensagens de texto enquanto estão dirigindo. Para as mulheres jovens ou com mais de 45 anos, esse percentual cai para 40%. O fato de o evento de *read while driving* ser o segundo tipo de distração mais praticado pelos homens, é um dado preocupante. Durante a leitura de uma mensagem de texto, o condutor se distrai fisicamente, visualmente e cognitivamente. O envolvimento de diferentes tipos de distração prejudica o desempenho do condutor e pode aumentar os riscos de acidentes.

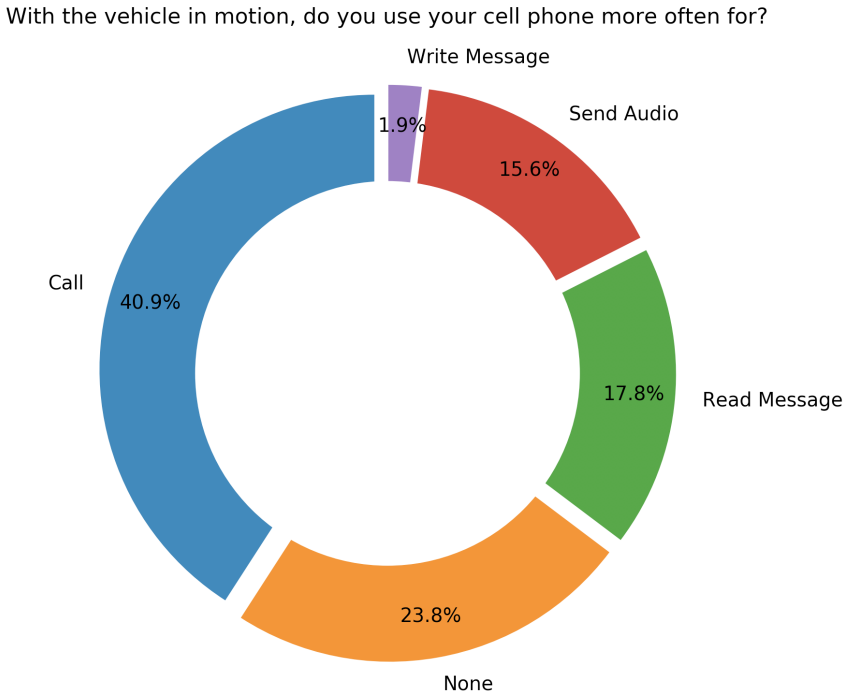


Figura 22. Respostas da 9^a pergunta do *survey*: Estando com o veículo em movimento, você utiliza o celular com maior frequência por qual motivo?

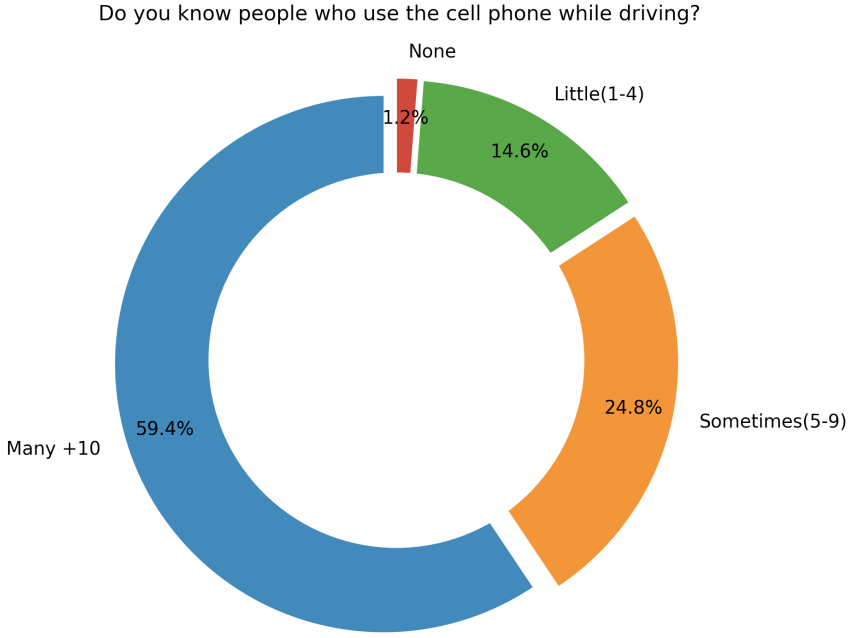


Figura 23. Respostas da 10^a pergunta do *survey*: Você conhece pessoas que utilizam o telefone celular enquanto estão dirigindo ?

A 9^a pergunta do *survey* teve como objetivo apontar, entre os quatro tipos de distrações investigadas, qual o evento secundário mais praticado pelos participantes. A Figura 22 apresenta o percentual das respostas. Considerando todos os motivos de distrações, concluímos que 76% dos entrevistados utilizam o telefone celular enquanto estão dirigindo. 40% responderam que utilizam o telefone celular com maior frequência para realizar uma chamada telefônica, 17% para ler mensagens de texto, 15% para enviar mensagens de áudio e pouco menos de 2% para escrever mensagens de texto. Ao interpretar esse gráfico, é importante notar que a utilização do telefone celular para uma determinada atividade não exclui a possibilidade de utilização em outro evento secundário. Por exemplo, como vimos nos gráficos apresentados nas Figuras 14 a 17, a leitura de mensagens de texto foi o segundo evento secundário mais praticado pelos motoristas do gênero masculino. Portanto, no gráfico da Figura 22, os homens que responderam que utilizam o telefone celular, com maior frequência, para realizar uma chamada telefônica provavelmente também leem mensagens de texto, porém, com menor frequência.

Na Figura 23, podemos observar a resposta da última pergunta do *survey*. O objetivo era saber se as pessoas entrevistadas conheciam outras pessoas que utilizavam o telefone celular enquanto estavam dirigindo. Menos de 2% dos entrevistados responderam que não conhecem pessoas que utilizam o telefone celular enquanto estão dirigindo. Mais de 59% responderam que conhecem mais de 10 pessoas que se distraem no trânsito devido à utilização do telefone celular, 24% responderam que conhecem entre 5 a 9 pessoas, e 14% responderam que conhecem entre 1 a 4 pessoas. Se consideramos esses dados e também o fato de que menos de 24% dos entrevistados responderam que não utilizam o telefone celular enquanto estão dirigindo, podemos concluir que a distração dos motorista em decorrência da utilização do telefone celular é um problema real no estado do Pará e que as leis e medidas educativas não estão sendo suficientemente eficazes para combater este problema. As respostas e análises desse *survey* confirmam que soluções como as apresentadas nos Capítulos 5, 6 e 7 são importantes para combater e prevenir a distração dos motoristas devido à utilização do telefone celular.

3.4 Considerações finais

Este estudo apresentou a condução de um *survey* cujo objetivo era investigar o perfil de distração dos motoristas paraenses, em decorrência da utilização do telefone celular. Dois fatores motivaram a realização desse estudo. O primeiro diz respeito às especificidades da cultura local. Nos *surveys* filtrados na revisão sistemática apresentada no Capítulo 2, observamos que o comportamento dos motoristas se diferem de acordo com a cultura do seu país. Por exemplo, na Espanha, no *survey* conduzido por (GRAS et al., 2007), mais de 60% dos entrevistados admitem utilizar o telefone celular enquanto estão dirigindo. Por outro lado, no Catar, segundo (SHAABAN, 2013), constatou-se que

menos de 20% dos entrevistados utilizam o telefone celular enquanto estão dirigindo. Nesse sentido, o *survey* apresentado nesse capítulo teve como objetivo verificar a influência da cultura paraense, em relação à distração do motorista.

O segundo fator que motivou a condução desse *survey* refere-se aos dados divulgados pelo Ministério da Saúde. Segundo o (SAÚDE, 2019), Belém, a capital do estado do Pará, é a capital que possui o maior percentual de motoristas que utilizam o telefone celular enquanto dirigem. Assim, o *survey* apresentado nesse capítulo buscou investigar os tipos de distrações mais praticados pelos motoristas paraenses. Quatro tipos de distrações foram investigados, quais sejam: 1- realização de chamadas telefônica; 2- composição de mensagens de textos; 3- envio de mensagens de áudio; 4- leitura de mensagens de texto.

Os resultados obtidos demonstram que mais de 70% dos entrevistados admitem utilizar o telefone celular, em algum momento, enquanto estão dirigindo. Dos motivos de distração investigados, a chamada telefônica e a leitura de mensagem de texto foram os mais frequentes em diversos filtros realizados. A distração foi investigada levando em consideração o gênero do entrevistado; o tipo de veículo; a idade; e a composição dessas três categorias. Os filtros revelaram certas especificidades, por exemplo, podemos constatar que os motoristas tendem a se distrair mais que os motociclistas. Esses e outros dados analisados nesse *survey*, sugerem que as campanhas educativas e a legislações de trânsito que proíbem o uso do telefone celular são medidas necessárias mas insuficientes para prevenir a distração do motorista, em decorrência da utilização do telefone celular.

4 Fundamentos dos modelos de aprendizado de máquina

Este capítulo tem como objetivo descrever a fundamentação teórica dos modelos de aprendizado de máquina abordados nessa tese. Ao todo, foram abordados 7 diferentes modelos, quais sejam: 1- *Losgistic Regression*; 2- Árvore de Decisão; 3- *Random Forest*; 4- Máquina de Vetor de Suporte; 5- *Adaptive Boosting*; 6- *Gradient Boosting*; 7- Rede Neural Convolucional. Todos os modelos abordados nesta tese são modelos de aprendizado de máquina supervisionado, uma subcategoria do aprendizado de máquina que utiliza conjunto de dados rotulados para a construção dos modelos. Seja X um conjunto de variáveis independentes e Y um conjunto de variáveis dependentes, um conjunto de dados rotulados L é definido por N pares de valores de entrada e valores de saída $(x_1, y_1), ..., (x_N, y_N)$, onde $x_i \in X$ e $y_i \in Y$. Dessa forma, a tarefa de um modelo supervisionado implica o aprendizado da função $\varphi : X \rightarrow Y$ a partir do conjunto $L = (X_N, Y_N)$. Após a construção do modelo, espera-se que a função $\varphi : X \rightarrow Y$ seja capaz de mapear valores $(x \in X, y \in Y) \not\subset L$. Ou seja, que o modelo construído seja capaz de classificar corretamente novos dados de entrada não utilizados durante a sua construção. Se o modelo atingir resultados satisfatórios, então dizemos que esse modelo possui a capacidade de generalização para o domínio do problema. Para analisar a capacidade de generalização dos modelos, este capítulo também apresenta a fundamentação do conjunto de métricas utilizadas nos experimentos descritos nos Capítulos 5, 6 e 7.

4.1 Modelos supervisionados

Nesta seção será descrita a fundamentação dos sete modelos de aprendizado de máquina construídos durante as experimentações desta tese.

4.1.1 Logistic Regression

A regressão logística é um dos algoritmos mais utilizados para resolução de problemas de classificação. Na revisão da literatura realizada no Capítulo 2, verificamos que a regressão logística foi o segundo modelo mais utilizados para estudar a distração do motorista. Os trabalhos de (YANNIS et al., 2016), (HE et al., 2015) e (HAQUE et al., 2016a) aplicaram a regressão logística para investigar a distração do motorista em decorrência do uso dos telefones celulares.

Em nosso estudo, consideramos a regressão logística como um modelo de referência.

Após a construção e avaliação de outros modelos de aprendizado de máquina, caso nenhum modelo demonstre melhor desempenho, então, pela simplicidade do modelo de regressão, este será o escolhido para distinguir motoristas e passageiros em momentos de distração pelo uso do telefone celular.

Considerando X um conjunto de variáveis independentes e Y um conjunto de variáveis dependentes, a regressão logística é um modelo probabilístico discriminativo. Ele modela a distribuição de probabilidade posterior $P(y|x)$, onde $y \in Y$ é a variável de saída e $x \in X$ são os dados de entrada. A Figura 24 mostra uma arquitetura típica do modelo de regressão logística. Em um problema de classificação binária, a saída da função sigmoide (eq. 4.1) é interpretada como a probabilidade de uma amostra particular pertencer à classe positiva, ou seja, $\phi(z) = P(y = 1|x; w)$ onde z é a combinação linear dos pesos e dos dados de entrada $z = w^T x$.

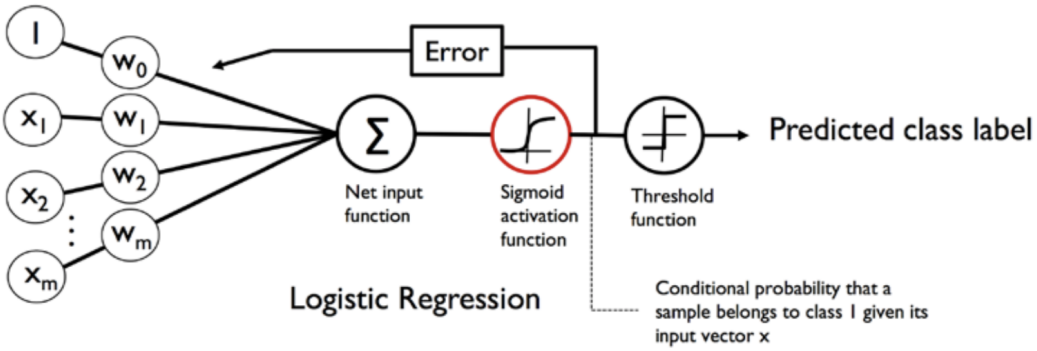


Figura 24. Arquitetura do modelo de regressão logística (RASCHKA; MIRJALILI, 2017).

$$\phi(z) = \frac{1}{1 + e^{-z}} \tag{4.1}$$

Na regressão logística, quando $z = w^T x$ é muito grande, então $\phi(z)$ tende a ser próximo de 1. Se $z = w^T x$ é muito pequeno, então $\phi(z)$ tende a ser próximo de 0. Em um problema de classificação binária, o seguinte *threshold* pode ser utilizado:

$$\phi(z) = \begin{cases} 1 & \text{se } \phi(z) \geq 0.5 \\ 0 & \text{caso contrário} \end{cases} \tag{4.2}$$

Durante a construção de um modelo de regressão logística, o objetivo é a minimização do erro para que se consiga a melhor classificação do conjunto de treinamento e, posteriormente, dos dados inéditos. A minimização do erro implica o ajuste dos pesos de w. Nesse caso, utiliza-se a equação 4.3 para a atualização dos valores de w, a partir do

método gradiente descendente. A utilização do método gradiente descendente aplicado ao modelo de regressão logística pode ser consultada em (RASCHKA; MIRJALILI, 2017).

$$J(w) = -\sum_i y^{(i)} log\left(\phi\left(z^{(i)}\right)\right) + \left(1 - y^{(i)}\right) log\left(1 - \phi\left(z^{(i)}\right)\right) \tag{4.3}$$

4.1.2 Árvore de Decisão

O modelo de árvore de decisão costuma ser escolhido quando a interpretabilidade é mais importante para o escopo do problema do que o próprio resultado da classificação. Uma árvore de decisão é um modelo $\varphi : X \to Y$ que prevê a classe y associada a x caminhando do nó raiz até uma determinada folha. O caminho da raiz para a folha representa as regras construídas pelo modelo para a realização da classificação. Em cada nó interno no caminho da raiz para a folha, o filho sucessor é escolhido com base em uma divisão do espaço de entrada. A divisão do espaço de entrada é um processo iterativo aplicado em cada nó, até que as folhas sejam puras. Uma folha pura significa que todas as amostras deste nó pertencem à mesma classe. O algoritmo utilizado no processo de ramificação da árvore é fundamentado no Ganho de Informação (IG).

O objetivo do IG é dividir os nós nos recursos mais informativos para que as amostras em cada nó pertençam à mesma classe. A equação (4.4) mostra o cálculo do IG de uma árvore binária. Aqui, f é o recurso para realizar a divisão, I é a medida de impureza, D_p, D_{left} e D_{right} são o conjunto de dados do nó pai, do nó filho da esquerda e do nó filho da direita, N_p é o número total de amostras no nó pai, N_{left} e N_{right} são o número de amostras no nó filho da esquerda e no nó filho da direita.

$$IG(D_p, f) = I(D_p) - \frac{N_{left}}{N_p} I(D_{left}) - \frac{N_{right}}{N_p} I(D_{right}) \tag{4.4}$$

Para uma árove N-ária, o Ganho de Informação (IG) pode ser calculado conforme a equação (4.5)

$$IG(D_p, f) = I(D_p) - \sum_{j=1}^{m} \frac{N_j}{N_p} I(D_j) \tag{4.5}$$

Segundo (RASCHKA; MIRJALILI, 2017), as duas medidas de impureza mais utilizadas no cálculo do IG são: *Entropy* (I_H) e *Gini impurity* (I_G).

$$I_H(t) = -\sum_{i=1}^{c} p(i|t)\ log_2\ p(i|t) \tag{4.6}$$

$$I_G(t) = 1 - \sum_{i=1}^{c} p(i|t)^2 \tag{4.7}$$

No cálculo da impureza, $p(i|t)$ representa a proporção das amostras que pertencem à classe c para um determinado nó t. Na equação 4.6, $I_H(t) = 0$, se todas as amostras de um nó pertencerem à mesma classe. $I_H(t)$ é igual a 1 se a distribuição das amostras for uniforme entre as classes. Em relação a equação 4.7, esta pode ser interpretada como um critério para minimizar a probabilidade de erro de classificação. Semelhante à entropia, a *Gini impurity* é máxima quando as amostras são perfeitamente distribuídas entre as classes.

A Figura 25 exemplifica como uma árvore de decisão pode ser interpretada. Com esse tipo de interpretação é possível analisar a decisão que o modelo tomou para gerar cada nó da árvore. Na prática, poderemos entender totalmente de onde as classificações e erros estão vindo, com que tipo de dados o modelo se sairia bem e como a saída é influenciada pelos valores de entrada. Quando este tipo de análise é importante para o escopo do problema, então o modelo de árvore de decisão se apresenta como uma ótima solução.

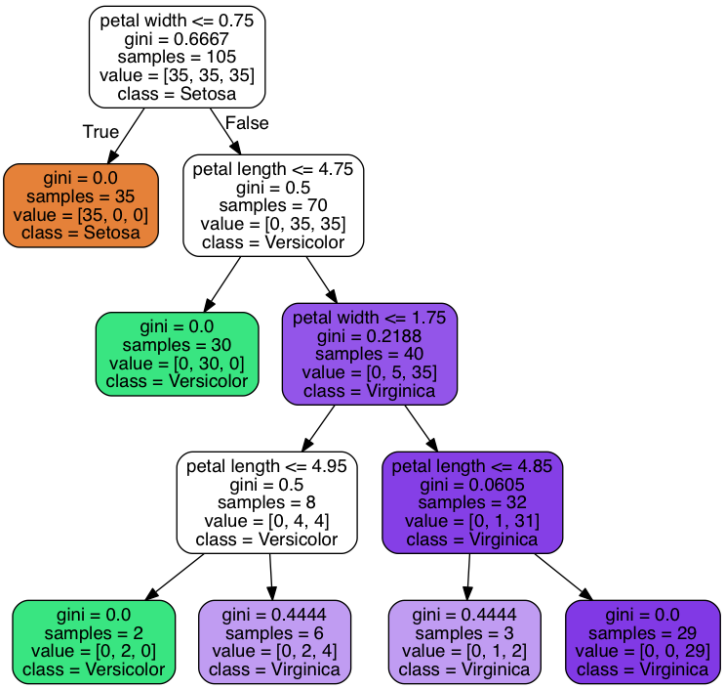


Figura 25. Exemplo de construção de uma árvore de decisão (RASCHKA; MIRJALILI, 2017).

4.1.3 Random Forest

De acordo com (BREIMAN, 2001), um modelo *Random Forest* consiste em uma coleção de classificadores de árvore de decisão $\{h_k(x, \Theta_k), k = 1, ...\}$, onde $\{\Theta_k\}$ são vetores aleatórios distribuídos independentemente. Os modelos que são construídos a partir da combinação de outros modelos, são chamados de modelos *ensemble*. Assim, o modelo

Random Forest é um modelo *ensemble*, no qual cada árvore $h_k(x, \Theta_k)$ representa um voto unitário para a classificação do problema em função do dado de entrada x. A classe com mais votos torna-se a previsão do modelo.

Segundo (RASCHKA; MIRJALILI, 2017), um modelo *Random Forest* pode ser construído da seguinte maneira:

1. Selecione, aleatoriamente, com reposição, n amostras do conjunto de treinamento para definir Θ_k.
2. Construa a árvore de decisão $h_k(x, \Theta_k)$. Para cada nó:
   - Selecione, aleatoriamente, d *features* sem reposição.
   - Divida o nó usando a *feature* que fornece a melhor divisão de acordo com a função objetivo, por exemplo, maximizando o ganho de informações conforme a equação 4.7.
3. Repita o passo 1 e 2 n vezes sendo n o número de árvores utilizadas no modelo.
4. Compute a previsão de cada árvore para atribuir o rótulo de classe por maioria de votos.

Considerando $h = \{h_k(x, \Theta_k), k = 1, .., n\}$ uma *random forest*, e A uma previsão do classificador $h_k(x, \Theta_k) \subset h$, nós podemos definir $P(A)$ como a proporção de classificadores h_k $1 \leq k \leq n$ para o qual o evento A ocorre. Assim é possível definir a função marginal de uma *random forest* conforme a seguinte equação:

$$m(x, y) = P_k(h_k(x) = y) - max_{j \neq y} P_k(h_k(x) = j) \; \forall \; 1 \leq k \leq n \tag{4.8}$$

Na função marginal, é possível perceber que se $m(x, y) > 0$, então o conjunto de classificadores realizaram a classificação correta. Por outro lado, quando $m(x, y) < 0$ significa que a maioria dos classificadores do conjunto realizaram a classificação incorreta. Segundo (BREIMAN, 2001), quando o número de árvores em uma *random forest* tende ao infinito, o erro é minimizado, uma vez que a função marginal da maioria dos classificadores tende a ser maior que zero. Esse comportamento demonstra a força da utilização do modelos *ensemble*.

Se compararmos o modelo *Random Forest* com o modelo de Árvore de Decisão, podemos notar que o modelo *Random Forest* tem a desvantagem da interpretabilidade. Por ser um aglomerado de árvores construídas a partir de seleções aleatórias, a plotagem dessas árvores e sua interpretação torna-se inviável. Portanto, se a análise do modelo for determinante no escopo do problema, a utilização da Árvore de Decisão é mais recomendada do que o modelo *Random Forest*.

4.1.4 Máquina de Vetor de Suporte

A Máquina de Vetor de Suporte (SVM) está entre os modelos de aprendizado de máquina mais utilizados pela comunidade científica. Isso pode ser comprovado na revisão de literatura realizada no Capítulo 2, na qual o modelo SVM foi o mais utilizado na tarefa de detecção e classificação da distração dos motoristas. O algoritmo SVM trabalha com a ideia de encontrar um hiperplano em um espaço N-dimensional que classifique os pontos de dados. Como pode ser visto na Figura 26, existem muitos hiperplanos que podem ser traçados para separar o ponto de dados. Nesse caso, o objetivo do algoritmo SVM é encontrar um plano que tenha a distância máxima entre pontos de dados de classes diferentes. Maximizar a distância da margem possibilita que dados futuros possam ser classificados com maior confiança. Como pode ser observado na Figura 26, os pontos utilizados para determinar a distância máxima entre os pontos de dados de classes distintas, são denominados como vetores de suporte. Sendo assim, a essência do algoritmo SVM está em encontrar os vetores de suporte que determinam a distância máxima entre os pontos de dados de classes distintas.

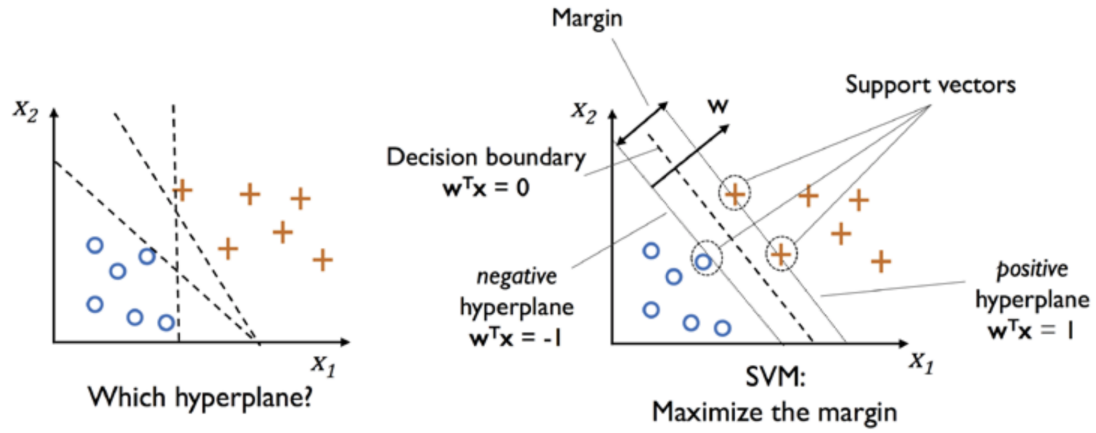


Figura 26. Exemplo de vetor de suporte de uma SVM (RASCHKA; MIRJALILI, 2017).

Os classificadores que separam os dados por meio de um hiperplano são denominados lineares. A equação de um hiperplano é determinada pela equação 4.9 onde w é um vetor normal ao hiperplano, ou seja, é perpendicular em relação ao hiperplano. Considerando b e w, pode-se definir $\frac{b}{\|w\|}$ como a distância do hiperplano em relação a origem, tal que $\| w \|$ corresponde a normalização euclidiana conforme a equação 4.10.

$$f(x) = w^T x + b = 0 \tag{4.9}$$

$$\| w \| = \sqrt{\sum_{j=1}^{m} w_j^2} \tag{4.10}$$

Como pode ser observado na Figura 26, a equação 4.9 divide o espaço dos dados de X em duas regiões e dá origem as equações 4.11 e 4.12. Sendo assim, segundo a explicação de (BURGES, 1998), a margem entre os vetores de suporte é dada pela equação 4.13, onde x_1 e x_2 são, respectivamente, os vetores de suporte do lado positivo e negativo.

$$w^T x + b \geq 1 \;\; paray = 1 \tag{4.11}$$

$$w^T x + b \leq -1 \;\; paray = -1 \tag{4.12}$$

$$(x_1 - x_2) \left(\frac{w}{\parallel w \parallel} \frac{(x_1 - x_2}{\parallel x_1 - x_2 \parallel} \right) \tag{4.13}$$

Como x_1 e x_2 são vetores de suporte, tem-se que $w^T x_1 + b = 1$ e $w^T x_2 + b = -1$. A diferença entre essas duas equações nos diz que $w(x_1 - x_2) = 2$. Substituindo esse resultado na equação 4.13, podemos definir a margem entre os vetores de suporte como:

$$\frac{2(x_1 - x_2)}{\parallel w \parallel \parallel x_1 - x_2 \parallel} \tag{4.14}$$

Segundo (BURGES, 1998), a equação 4.14 pode ser simplificada para $\frac{2}{\|w\|}$. A partir dessas considerações, pode-se perceber que a maximização da margem em relação a equação 4.9 pode ser obtida pela minimização de $\parallel w \parallel$. O tratamento matemático da otimização desse problema pode ser consultado em (BURGES, 1998).

Em relação a utilização do modelo SVM para problemas não lineares, é necessária a transformação dos dados de entrada a partir de uma função ϕ. A escolha apropriada de ϕ faz com que o conjunto de treinamento de características não lineares possa ser classificado por uma SVM. Dependendo da função ϕ escolhida, é possível que a computação se torne inviável devido à alta dimensionalidade dos dados. Porém, independentemente da dimensão, o que se quer é uma forma eficiente para realizar o cálculo de produtos escalares entre os dados transformados por ϕ. Para resolver esse problema, o modelo SVM faz uso das funções de *Kernel*. Uma função de *Kernel* K é uma função que recebe dois pontos do conjunto de entrada e computa o produto escalar no espaço transformado por ϕ. Segundo (RASCHKA; MIRJALILI, 2017), a função Radial Basis Function (RBF) definida pela equação 4.15 é uma das funções de *kernel* mais utilizadas na literatura.

$$\kappa(x^{(i)}, x^{(j)}) = exp\left(-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\sigma^2} \right) \tag{4.15}$$

4.1.5 Adaptive Boosting e Gradient Boosting

Os algoritmos de *Boosting* são algoritmos de aprendizado de máquina que utilizam a abordagem gulosa para criar regras de previsão altamente precisa. Em um problema de classificação, nas primeiras iterações, o algoritmo cria regras de previsão ϵ com desempenho próximo de $0.5 + \lambda$ para um λ não muito grande (SCHAPIRE; FREUND, 2012). Para cada iteração, a estratégia gulosa redefine a importância das amostras de tal forma que as amostras classificadas incorretamente na iteração anterior sejam as de maior importância na próxima iteração. Assim, os ajustes das classificações incorretas de cada iteração permite a construção de regras de previsão ϵ com desempenho $0.5 + \lambda$ para um λ cada vez maior.

Na Figura 27 é possível perceber como se dá o processo de redefinição da importância das amostras para a construção das regras de previsão do algoritmo de *Boosting*. O exemplo ilustra um problema de classificação binária, na qual os dados de entrada possuem os atributos x_1 e x_2. Os números em cada quadro representam a iteração do algoritmo. Na primeira iteração, todas as amostras possuem a mesma importância. Considerando a importância atribuída, a linha tracejada do quadro 1 representa a regra de previsão construída para classificar as duas classes do problema, as quais estão representadas por círculos e triângulos.

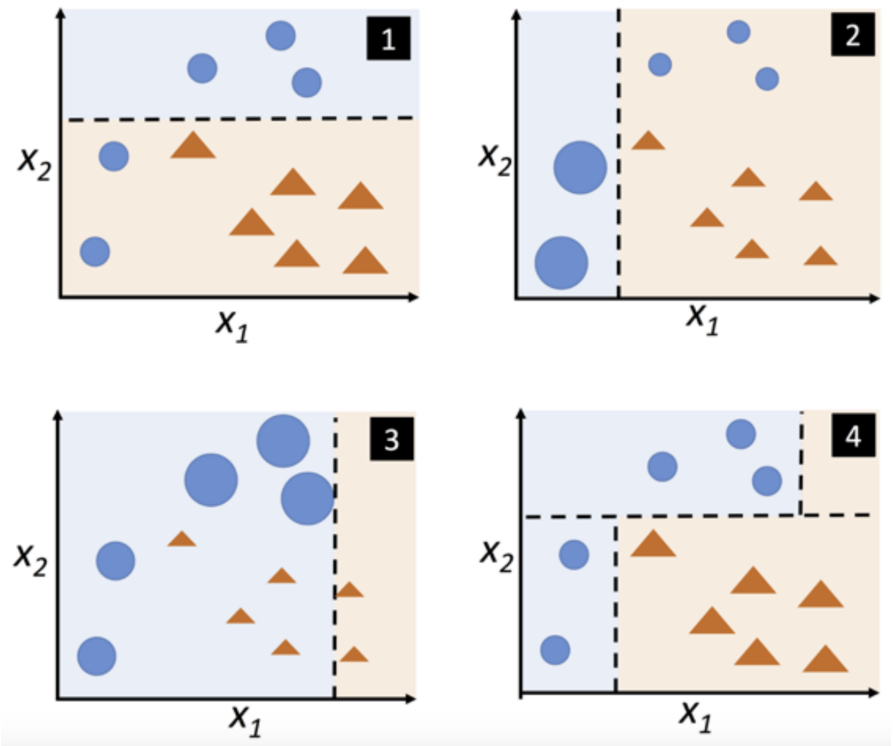


Figura 27. Ilustração da redefinição dos pesos das amostras utilizadas no algoritmo de Boosting (RASCHKA; MIRJALILI, 2017).

Na segunda rodada, a importância das amostras é recalculada. As amostras classificadas incorretamente na rodada anterior devem ser as mais importantes para a rodada

atual. No quadro 2, pode-se observar que os dois círculos classificados incorretamente na rodada 1 estão maiores na atual rodada. Essa diferença de tamanho representa a maior importância dessas duas amostras. A regra de previsão construída na segunda rodada priorizará as amostras de treinamento que possuem os maiores pesos para classificá-las corretamente.

Na próxima rodada, novamente, as importâncias são recalculadas e a regra de previsão deve ser construída para classificar corretamente as amostras de maior peso. Esse processo se repete de forma iterativa até a última rodada. Nesse exemplo, consideramos que o algoritmo possui apenas três rodadas. O quadro 4 representa a combinação das regras de previsão construídas nas três rodadas. No quadro 4 é possível observar que a combinação dos classificadores fracos formam uma classificador forte que consegue separar as classes do problema corretamente.

De um modo mais formal, o algoritmo de Boosting descrito a partir da ilustração da Figura 27 pode ser definido pelas seguintes etapas:

1. Inicializar o vetor de peso w de maneira uniforme, onde $\sum_i w_i = 1$.

2. Para $j = 1, ..., m$:

   a) Treinar o classificador fraco para criar a regra de previsão $C_j = \varphi : (X, w) \rightarrow Y$.

   b) Testar a regra de previsão $\widehat{y} = C_j(X, w)$.

   c) Calcular a taxa de erro ponderada: $\varepsilon = w \cdot (\widehat{y} \neq y)$

   d) Calcular a importância da regra de previsão: $\alpha_j = \frac{1}{2} log\left(\frac{1-\varepsilon}{\varepsilon}\right)$.

   e) Atualizar o vetor de peso: $w = w \times exp(-\alpha_j \times \widehat{y} \times y)$

   f) Normalizar o vetor de peso: $w = w / \sum_i w_i$

3. Combinar as regras de previsão para construir o classificador: $\widehat{y} = sign\left(\sum_{j=1}^{m} \alpha_j \times C_j(X)\right)$

Assim como no modelo *Adaptive Boosting*, o *Gradient Boosting* também é construído a partir da combinação de classificadores fracos para formar um classificador forte. A equação 4.16 descreve a estratégia incremental utilizada pelo algoritmo *Gradient Boosting*. Na primeira iteração, o algoritmo cria um classificador para construir a primeira regra de previsão $f_0(x)$. A partir da segunda iteração, a regra de previsão sempre é construída utilizando a regra de previsão antecessora e o coeficiente de aprendizado η. Assim, conforme

a equação 4.17, o modelo final do algoritmo é determinado pela minimização do erro a partir da função L, a qual é uma função gradiente.

$$\begin{aligned}
\widehat{y}_0 &= f_0(x) \\
\widehat{y}_1 &= f_0(x) + \eta f_1(x) = \widehat{y}_0 + \eta f_1(x) \\
\widehat{y}_2 &= f_0(x) + \eta f_1(x) + \eta f_2(x) = \widehat{y}_1 + \eta f_2(x) \\
\widehat{y}_m &= \widehat{y}_{m-1} + \eta f_m(x)
\end{aligned} \tag{4.16}$$

$$f(x) = arg\min_{\widehat{y}} \sum_{i=0}^{N} L\left(y_i, \widehat{y}_m\right) \tag{4.17}$$

4.1.6 Redes Neurais Convolucionais

A rede neural convolucional é uma categoria de rede neural que tem demonstrado eficiência em áreas como classificação de objetos, reconhecimento de faces, reconhecimento de ações e reconhecimento de fala. Arquiteturas como *LeNet-5, AlexNet, ZFNet, VGGNet, GoogleNet e ResNet* mostram a evolução das redes neurais convolucionais e a tendência de arquiteturas cada vez mais profundas. Segundo (GU et al., 2018), CNNs mais profundas podem proporcionar melhor desempenho, uma vez que aumentam a não linearidade da rede. Por outro lado, a profundidade também aumenta a complexidade da rede e dificulta a sua otimização. Independente da arquitetura e profundidade, é possível observar que existe um padrão dos componentes utilizados na construção de uma CNN. Segundo (GOODFELLOW; BENGIO; COURVILLE, 2016), uma camada convolucional é formada pela 3-upla <convolução, função de ativação, *pooling*> que será descrita a seguir.

4.1.6.1 Convolução

Convolução é um teorema matemático aplicado sobre duas funções f e g para obter uma terceira função h, definida da seguinte forma:

$$(f * g)(c) = h(c) = \sum_{a} f(a) \cdot g(c - a) \tag{4.18}$$

Para exemplificar, considere o cenário da Figura 28. Neste cenário, uma bola é lançada de uma altura $h1$ e percorre uma distância A com probabilidade $f(A)$. A partir do ponto de parada, a bola é relançada de uma altura $h2$ e percorre uma distância B, com probabilidade $g(B)$. Assim, a distância total percorrida pela bola é $c = a + b$. A probabilidade da bola percorrer a distância c é dada por $f(A)g(B)$, ou seja, a probabilidade de percorrer a distância A e a probabilidade de percorrer a distância B.

Assumindo $c = 8$, é possível obter diferentes valores para a e b que satisfaçam a equação $a + b = 8$. Logo, para obter a função de probabilidade de c, é necessário levar

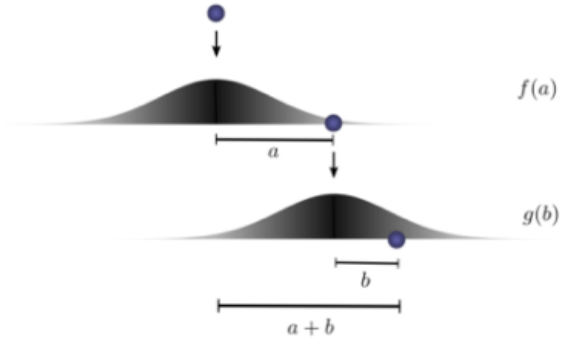


Figura 28. Convolução exemplificada pelo lançamento de bola. Fonte: http://colah.github.io/posts/2014-07-Understanding-Convolutions

em consideração todas as possibilidades de $f(A)g(B)$. Portanto, a probabilidade da bola percorrer uma distância c pode ser definida como:

$$(f * g)(c) = h(c) = \sum_{a+b=c} f(a) \cdot g(b) \tag{4.19}$$

Realizando a substituição $b = c - a$ temos a convolução:

$$(f * g)(c) = h(c) = \sum_{a} f(a) \cdot g(c - a) \tag{4.20}$$

No contexto das redes neurais, muitas aplicações utilizam valores de entrada multidimensionais. As aplicações que trabalham com a manipulação de imagens são exemplos de redes que manipulam entradas de duas dimensões. Neste contexto, a convolução é dada por:

$$(f * g)(c_1, c_2) = h(c_1, c_2) = \sum_{a_1, a_2} f(a_1, a_2) \cdot g(c_1 - a_1, c_2 - a_2) \tag{4.21}$$

A convolução bidimensional pode ser vista como uma janela deslizante de uma função sobre a outra. No caso da rede neural convolucional, a convolução implica na função *kernel* deslizando sobre a matriz dos valores de entrada. A Figura 29 ilustra um exemplo de convolução bidimensional realizada por uma CNN.

Conforme observado por (BISHOP, 2006), as funções de *kernel* implicam conectividade esparsa e compartilhamento de pesos. A conectividade esparsa é caracterizada pela utilização de *kernels* muito menores do que as matrizes de entrada. O conceito de conectividade esparsa vem da teoria dos grafos. Na teoria dos grafos, um grafo G é dito esparso quando o seu número de arestas $|E'|$ é muito menor do que o número de arestas possíveis $|E|$, ou seja, $|E'| \ll |E|$.

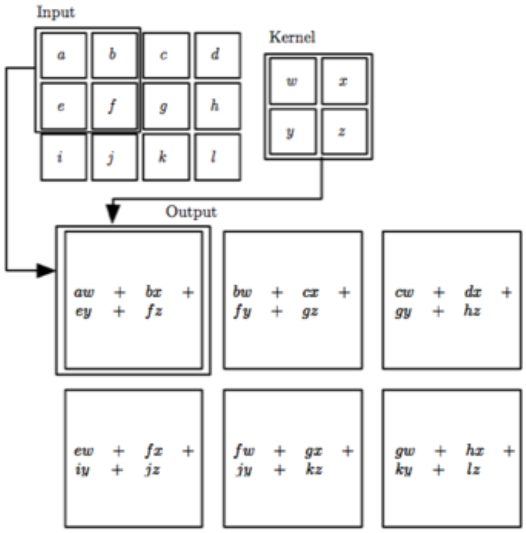


Figura 29. Exemplo de convolução bidimensional. Fonte: (GOODFELLOW; BENGIO; COURVILLE, 2016)

Projetando as redes neurais como um grafo, podemos considerar os valores de entrada e os neurônios como vértices e a conectividade entre eles como arestas direcionadas. Admitindo que o direcionamento das arestas parte somente dos valores de entrada para os neurônios, o número máximo de conexões em uma rede de N valores de entrada e M neurônios é dado pelo produto NM. A equação 4.22 representa a computação de cada neurônio da rede.

$$s_j = \sum_{j}^{M} \sum_{i}^{N} x_i w'_j \tag{4.22}$$

Nesta equação os valores de entrada são representados por x_i e os pesos das arestas por w'_j. Observe na Figura 30 que os pesos w'_j podem ser pesos diferentes para cada valor de entrada. Por exemplo, o peso w'_1 pode ser igual a 0.25 para a entrada x_1 e 0.1 para a entrada x_2. O fato dos pesos possuírem valores diferentes para entradas diferentes demonstra que não há compartilhamento de pesos em redes totalmente conectadas. O comportamento do não compartilhamento reflete na necessidade de atualização de todos os pesos, para cada iteração de treinamento.

No caso das redes que trabalham com convolução, o número de conexões é bem menor do que o produto NM e, por isso, é considerada esparsa. Em uma rede convolucional, o número de conexões pode ser definida pelo produto NK sendo N o tamanho dos valores de entrada e K o tamanho do *kernel*. Apesar do número de conexões também ser representada por um produto, em uma rede convolucional, K é muito menor que M. Conforme observador por (GOODFELLOW; BENGIO; COURVILLE, 2016), o fato de $K \ll M$ não altera a ordem da complexidade do algoritmo $O(NM)$, mas traz ganhos de performance significativos em relação ao consumo de memória e processamento.

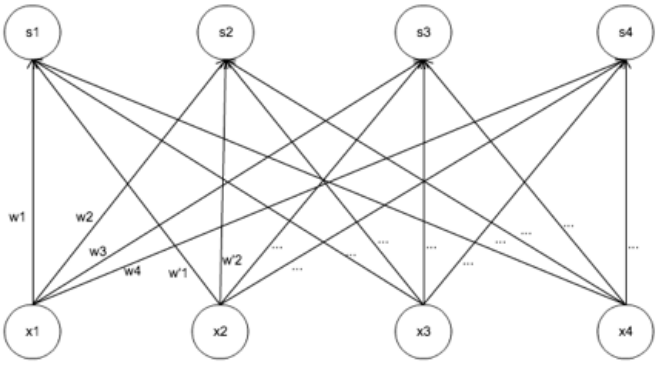


Figura 30. Exemplo da relação peso e valor de entrada em uma rede totalmente conectada.

A Figura 31 exemplifica as conexões de uma rede neural convolucional. Nesta figura podemos perceber que somente alguns valores de entrada estão conectados à função de *kernel* A. A função A deve ser entendida como uma função de *kernel*, assim como a apresentada na Figura 29. Neste contexto, os valores do *kernel* representam os pesos w'_j da convolução.

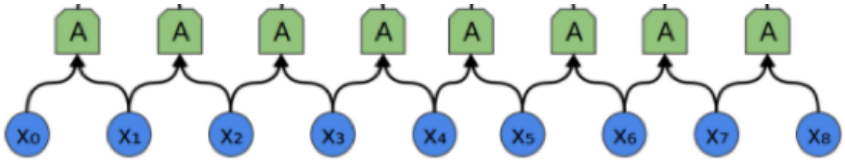


Figura 31. Exemplo da relação peso e valor de entrada, em uma rede convolucional. Fonte: http://colah.github.io/posts/2014-07-Conv-Nets-Modular/

A repetição de A pela rede representa o deslizamento da função de *kernel* por todos os valores de entrada. O deslizamento da função de *kernel* implica compartilhamento dos pesos. Como ressaltado por (BISHOP, 2006), o compartilhamento dos pesos caracteriza um ganho de performance, uma vez que somente os *kernels* serão atualizados na fase de treinamento da rede. Além disso, a utilização do mesmo *kernel* possibilita a paralelização das operações que, normalmente, são realizadas nas GPUs.

Segundo (BUDUMA; LOCASCIO, 2017), as funções de *kernel* das redes neurais convolucionais funcionam, essencialmente, como detectores de características. Cada *kernel* tem a função de aprender características específicas do dado de entrada. O aprendizado de cada *kernel* se dá pelos ajustes dos pesos, durante o treinamento da rede. Em problemas de classificação de imagens ou de reconhecimento de características, as camadas convolucionais normalmente aprendem a detectar bordas, textura, formas e características específicas do contexto do problema. A título de exemplo, a Figura 32 apresenta uma abstração do funcionamento das camadas convolucionais.

Cada quadrado da Figura 32 representa uma função de *kernel* realizando uma convolução. Segundo (CHOLLET, 2017), o aprendizado de características específicas por

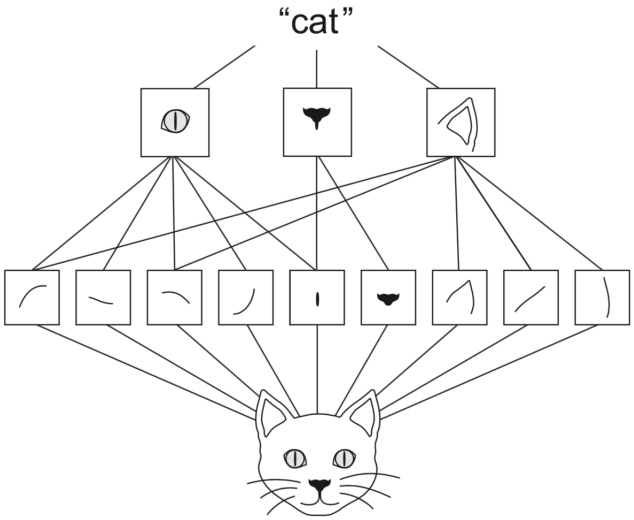


Figura 32. Abstração das funções de *kernel* de uma CNN. Fonte: (CHOLLET, 2017)

cada *kernel* possibilita o aprendizado hierárquico. Uma primeira camada de convolução aprende pequenos padrões locais, como bordas; uma segunda camada de convolução aprende padrões mais específicos, e assim por diante. Para (CHOLLET, 2017), o aprendizado hierárquico permite que as CNN's aprendam, eficientemente, conceitos visuais cada vez mais complexos e abstratos. No exemplo da Figura 32, os filtros da primeira camada convolucional aprenderam formas genéricas do objeto. Já na segunda camada, a combinação das características aprendidas possibilitou o aprendizado de formas mais específicas.

4.1.6.2 Função de ativação

Nas redes neurais, as funções de ativação tem o objetivo de determinar a ativação ou não de um neurônio. Um neurônio ativado implica propagação do seu sinal para a próxima camada da rede. Se o neurônio for um neurônio de saída, então, a sua função de ativação determina a saída da rede. De acordo com a equação 4.23, em uma rede neural, a computação de um neurônio é dada pela soma do produto dos valores de entrada e seus pesos correspondentes. Ao valor resultante é aplicada uma função de ativação que determina o valor de saída do neurônio.

$$z = f(w_0x_0 + w_1x_1 + ... + x_nw_n) = f(\sum_{i=0}^{n} w_ix_i) = f(W^TX) \tag{4.23}$$

Dependendo da função de ativação utilizada, além de determinar a ativação dos neurônios, a função de ativação também tem o objetivo de introduzir propriedades não lineares na rede. Com a apropriada função de ativação, um modelo de rede neural consegue resolver problemas não lineares como os que envolvem a manipulação de imagem, vídeo, áudio e reconhecimento de voz.

Segundo (GOODFELLOW; BENGIO; COURVILLE, 2016), a convolução de uma CNN é um sistema linear. Neste sentido, múltiplas convoluções, em cascata, também formariam um sistema linear. Logo, para que uma CNN possa resolver problemas não linearmente separáveis, é necessária a utilização de funções de ativação não lineares, entre cada camada de convolução.

Outra característica importante em relação à escolha da função de ativação é que ela deve ser diferenciável. Segundo (HAYKIN, 1998), esta característica é necessária, uma vez que, possibilita a utilização do método de treinamento *backpropagation*. O algoritmo *backpropagation* foi proposto por Rumelhart, Hinton e Williams na década de oitenta e, até hoje, é um dos métodos mais utilizados e eficientes para o treinamento de redes neurais (RUMELHART; HINTON; WILLIAMS, 1986). De maneira geral, o treinamento de uma rede neural busca ajustar os pesos dos neurônios, tal que, para cada valor de entrada, seja produzida a saída esperada. Para cada iteração de treinamento, o erro entre a saída produzida e a saída esperada é calculado e, em seguida, o erro é propagado para as camadas inferiores. Como demonstrado por (RUMELHART; HINTON; WILLIAMS, 1986), o fato das funções serem diferenciáveis é o que possibilita a propagação do erro e, consequentemente, o ajuste dos pesos.

Das funções de ativação não lineares utilizadas nos neurônios de uma rede neural, (KARLIK; VEHBI, 2011) cita a função *sigmoid* (ver equação 4.24) como uma das mais utilizadas na literatura. Esta é uma função assimétrica que produz somente valores reais não negativos no intervalo entre 0 e 1. Devido à característica da função, recomenda-se a sua utilização em problemas de classificação binária. A Figura 33 ilustra o comportamento da função.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4.24}$$

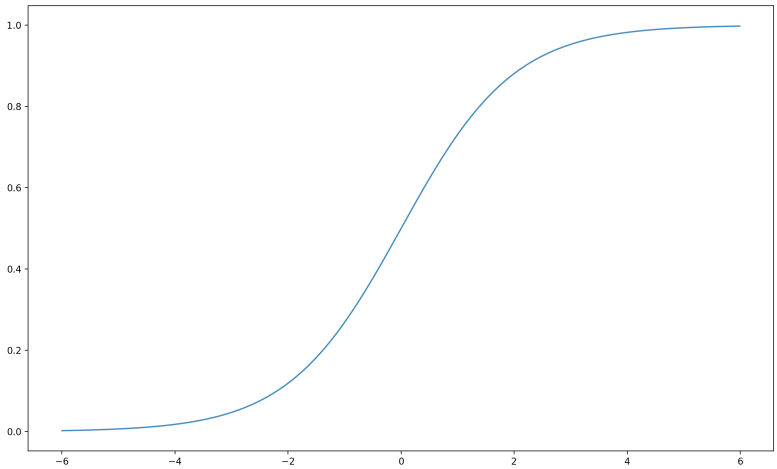


Figura 33. Comportamento da função sigmoid.

O fato da função *sigmoid* produzir somente valores reais não negativos, no intervalo entre zero e um, pode ser indesejável para algumas aplicações. Levando isto em consideração, a função tangente hiperbólica (ver equação 4.25), que é uma versão escalonada da função *sigmoid*, pode ser utilizada para resolver este problema. Na Figura 34, é possível perceber que a função tangente hiperbólica é simétrica em relação à origem e produz valores de saída reais no intervalor entre menos um e um.

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{4.25}$$

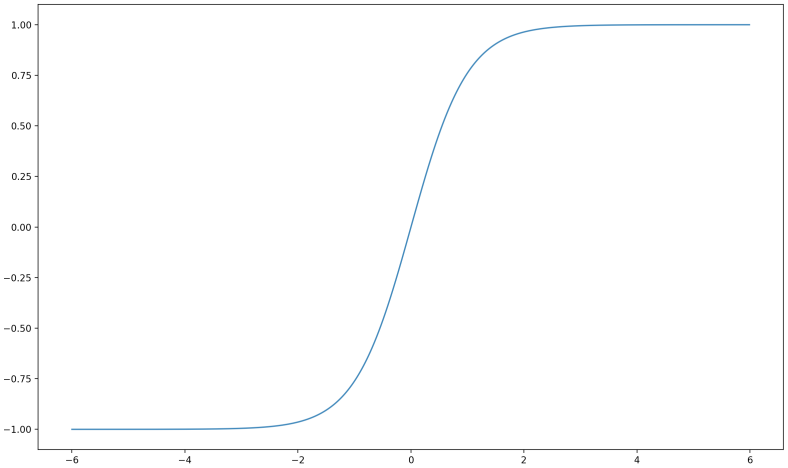


Figura 34. Comportamento da função tangente hiperbólica.

No contexto da rede neural convolucional, estudos como (GU et al., 2018), (ALCANTARA, 2017) e (LECUN; BENGIO; HINTON, 2015) apontam que a função Rectified Linear Unit (ReLu) e suas variações possuem bom desempenho para os neurônios das camadas de convolução. A função ReLU (ver equação 4.26) calcula o valor máximo dos valores de entrada e transforma todos os valores de entrada negativos para zero. A Figura 35 ilustra o comportamento da função.

$$f(x) = max(0, x) \tag{4.26}$$

Segundo (MAAS; HANNUN; NG, 2013), devido à função max, a ReLU é, computacionalmente, mais eficiente do que as funções de ativação *sigmoid* e *than*. (GU et al., 2018) também ressalta que a convergência dos valores negativos permite a não ativação de todos os neurônios e, consequentemente, uma representação esparsa do modelo. A representação esparsa proporcionada pela função ReLU permite que as CNNs sejam treinadas de forma eficiente, sem a necessidade de pré-treinamento. Essas particularidades constatam a importância da função ReLU e das suas variações no contexto das CNNs.

Além da eficiência computacional, justifica-se a utilização das funções ReLU, uma vez que, a sua composição possibilita a aproximação de funções mais complexas. Para

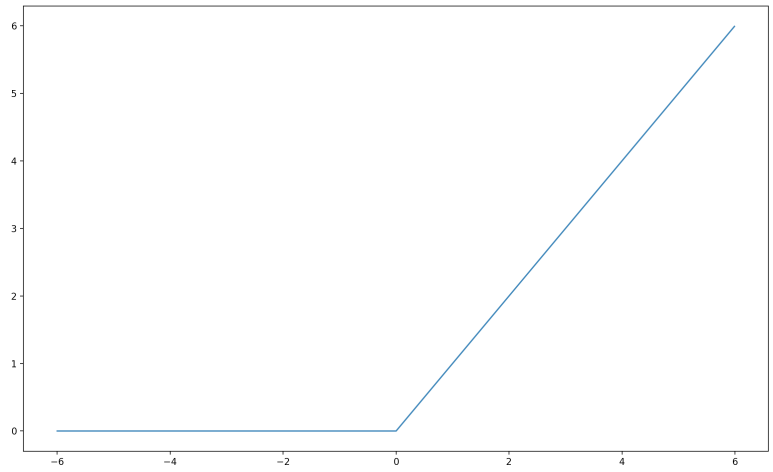


Figura 35. Comportamento da função ReLu.

exemplificar, considere a função quadrática $f(x) = x^2$. Neste caso, poderíamos utilizar a seguinte função ReLU, para aproximação de $f(x)$ (ver Figura 36):

$$h(x) = ReLU(x) + ReLU(-x) + ReLU(2x - 2) + ReLU(-2x + 2) \tag{4.27}$$

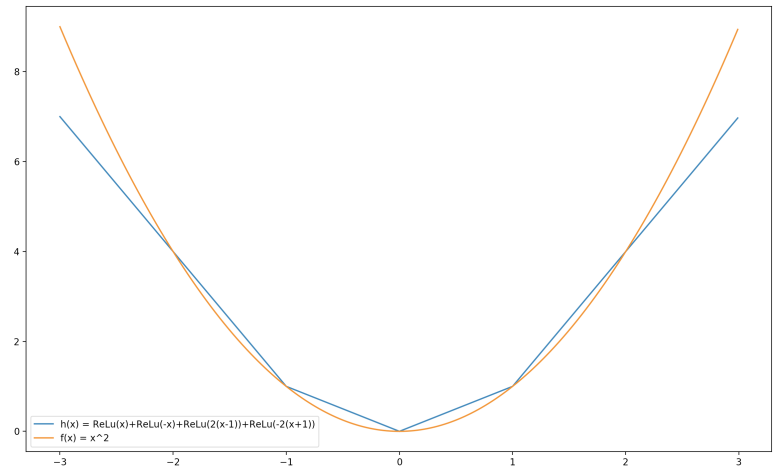


Figura 36. Aproximação da função quadrática pela função ReLu.

4.1.6.3 Pooling

Conceitualmente, as funções de pooling têm como objetivo tornar o mapeamento de características invariantes, mesmo que existam variações nos valores de entrada. Segundo (GOODFELLOW; BENGIO; COURVILLE, 2016), o conceito de invariância pode ser útil em aplicações que verificam se uma determinada característica está presente, mas sem levar em consideração o espaço onde esta característica se encontra. Por exemplo, ao determinar se uma imagem contém um rosto, não é necessário saber a localização dos

olhos com uma precisão perfeita de pixels. O que a aplicação precisa reconhecer é que há um olho do lado direito e outro do lado esquerdo da face. Em aplicações desta natureza, as funções de *pooling* são extremamente interessantes, uma vez que eliminam as variações do mapeamento de características extraídas na convolução.

No contexto das imagens digitais, a variância pode ocorrer na luminosidade, rotação, translação e dimensionalidade. A Figura 37 ilustra as possibilidades de variações nas imagens digitais. Observando as imagens desta figura, é possível perceber que as aplicações, altamente dependentes da distribuição dos pixels das imagens, teriam dificuldade em trabalhar com imagens variantes.

Figura 37. Exemplo de variação de tamanho, translação e rotação em imagens digitais. Fonte: (CHAUDHARI; GARG; KOTECHA, 2017).

Segundo (LECUN; BENGIO; HINTON, 2015), a função de *pooling* trabalha como uma função de agrupamento, cujo objetivo é fundir características semanticamente semelhantes. De uma maneira geral, essas funções analisam um conjunto de vizinhos e extraem uma característica que os representem. Por exemplo, a função *max pooling* proposta por (ZHOU; CHELLAPPA, 1988) realiza essa representatividade extraindo o valor máximo entre os vizinhos observados (ver Figura 38). Além da *max pooling*, (GU et al., 2018) destaca que as funções *lp pooling*, *mixed pooling*, *stochastic pooling*, *spectral pooling*, *spatial pyramid pooling* e *multi-scale orderlss pooling* também são funções utilizadas nas CNNs e possuem suas vantagens e desvantagens, dependendo do contexto aplicado.

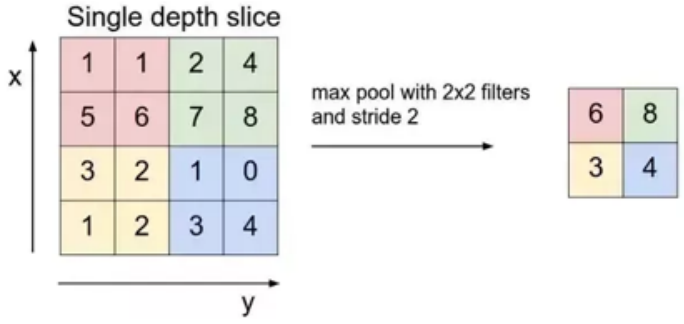


Figura 38. Comportamento da função *max pooling* no espaço bidimensional.

4.2 Métricas de avaliação dos modelos

Fazer previsões sobre dados futuros é, frequentemente, o principal problema que queremos resolver em aplicações de aprendizado de máquina. No entanto, dependendo do objetivo, estimar o desempenho de um modelo não é uma tarefa trivial. Em problemas que envolvem a construção de modelos de aprendizado de máquina, é comum realizar o desenvolvimento e a comparação de vários modelos para ver qual o algoritmo que consegue resolver o problema da melhor forma. De forma simples, nós poderíamos pensar na construção e comparação desses modelos da seguinte forma:

1. Separar o conjunto de dados em conjunto de treinamento e conjunto de teste.
2. Utilizar o conjunto de dados de treinamento para a construção dos modelos.
3. Utilizar os modelos construídos para realizar a previsão dos dados do conjunto de teste.
4. Para cada modelo, contar o número de previsões incorretas realizadas no conjunto de teste, para calcular a *accuracy* do modelo.
5. Comparar a *accuracy* dos modelos para definir qual o melhor.

Essa forma de construção e comparação possui alguns problemas que podem comprometer a melhor escolha de um modelo de aprendizado de máquina. O primeiro problema que pode comprometer a escolha final é a forma como os dados foram separados. Considerando um problema de classificação n-ária, a separação da base de dados como descrita no passo 1 só será eficiente se a base de dados possuir uma quantidade de amostras significativas e balanceadas para cada classe. Além disso, teria que garantir que a separação fosse realizada de forma estratificada, ou seja, que as amostras no conjunto de teste fossem representativas para cada classe. Uma forma de resolver esse problema de separação dos dados de treinamento e validação implica a utilização do *K-fold Cross-Validation*.

4.2.1 K-fold Cross-Validation

A ideia principal da validação cruzada é que cada amostra do conjunto de dados tem a oportunidade de ser testada. A validação cruzada de *k-fold* é um caso especial em que fazemos a iteração sobre um conjunto de dados k vezes. Em cada rodada, dividimos o conjunto de dados em k subconjuntos mutuamente exclusivos, de mesmo tamanho e estratificados. Um subconjunto é usado para validação e os restantes $k - 1$ são mesclados em um subconjunto de treinamento para a avaliação do modelo. A Figura 39 ilustra o processo de validação cruzada de *5-fold*. Após k iterações, o desempenho da validação cruzada é dada pela média aritmética das estimativas de desempenho dos k subconjuntos de validação.

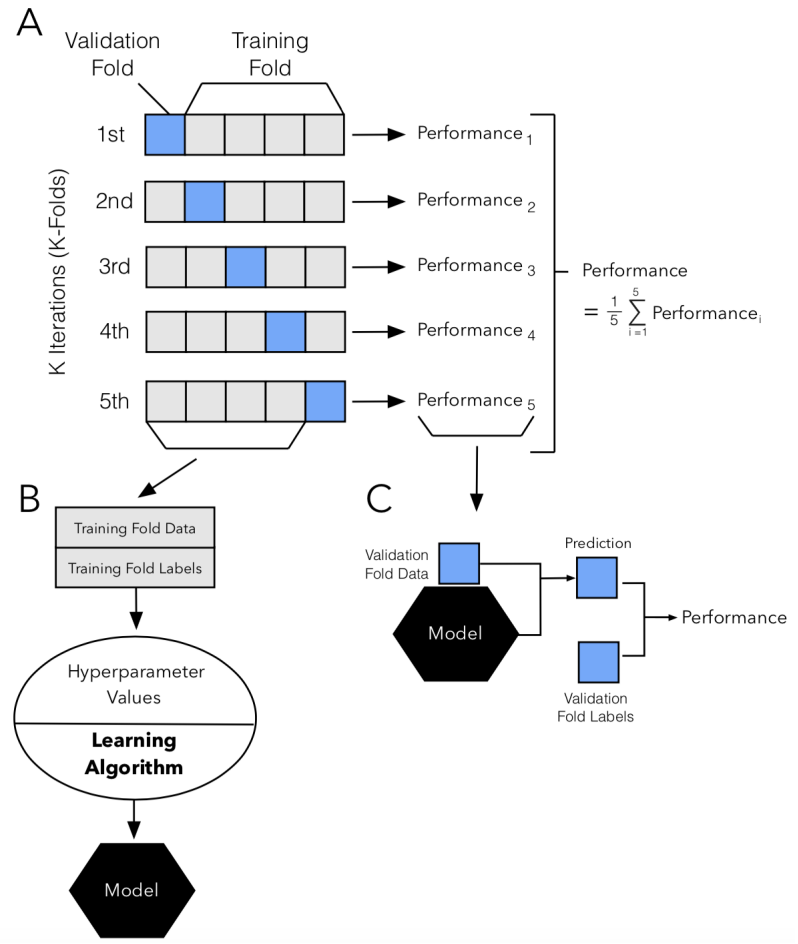


Figura 39. Ilustração do procedimento de validação cruzada de k-fold (RASCHKA, 2018).

A validação cruzada de *k-fold* usa todos os dados para realizar o treinamento e teste do modelo. Essa característica elimina o problema da má distribuição dos dados de treinamento e teste e também possibilita que o modelo em construção seja avaliado com a maior quantidade e variabilidade de amostras possíveis. Outro problema que o *K-fold Cross-Validation* resolve é a questão do tamanho da base de dados. Se uma base de dados pequena for separada em conjunto de treinamento e teste, é provável que a base de

treinamento não seja suficientemente representativa para que o modelo consiga realizar a generalização do problema.

O segundo problema que pode comprometer a escolha final é a forma de comparação dos modelos. Como descrito no passo 5, os modelos são comparados pela *accuracy*. A *accuracy* é calculada pela proporção do número de previsões realizadas corretamente e a quantidade de amostras no conjunto de teste. Em termos formais, a *accuracy* pode ser definida conforme a equação 4.28.

$$\text{ACC} = 1 - \text{ERR} \tag{4.28}$$

Nessa equação, o erro ERR é calculado de acordo com a equação 4.29.

$$\text{ERR} = \frac{1}{n}\sum_{i=1}^{n} L(\widehat{y}_i, y_i) \tag{4.29}$$

Em um problema de classificação, a função $L(.)$ é definida conforme a equação 4.30

$$L(\widehat{y}_i, y_i) = \begin{cases} 1 & \text{se } \widehat{y}_i = y_i \\ 0 & \text{se } \widehat{y}_i \neq y_i \end{cases} \tag{4.30}$$

A *accuraccy* avalia o desempenho global do classificador. Nesse sentido, quando a quantidade de amostras por classe é desbalanceada, as boas taxas de *accuracy* podem não refletir o real desempenho do modelo. Para exemplificar este cenário, suponha um problema de classificação das classes A e B, em que o número de amostras da classe A é igual a 10 e o número de amostras da classe B é igual a 90. A Tabela 9 representa a matriz de confusão obtida a partir de um classificador hipotético. Diante do cenário apresentado, a *accuracy* seria de 0.88. Entretanto, apesar da boa *accuracy* obtida, podemos observar que o classificador hipotético não conseguiu aprender os padrões da classe A. Isso pode ser evidenciado pela quantidade de falsos negativos desta classe.

Tabela 9. Matriz de confusão de um classificador hipotético.

Classes	**A**	**B**
A	1	9
B	3	87

Para a melhor avaliação dos modelos, é evidente que, além da visão global da métrica *accuracy*, também há a necessidade de outras métricas para avaliar as especificidades de cada classe. Dessa forma, apresentamos a estratégia de divisão dos dados, assim como as métricas escolhidas para estimar a performance dos modelos de aprendizado de máquina construídos a aplicados nos experimentos descritos nos Capítulos 5, 6 e 7.

4.2.2 Precision, Recall e F1-score

Em um problema de classificação, é possível observar o desempenho do classificador a partir de uma matriz de confusão. Uma matriz de confusão é uma matriz quadrada, na qual o número de linhas e colunas correspondem ao número de classes do problema. A linha da tabela representa a classe prevista, enquanto a coluna representa a classe real. Considerando um problema de classificação binária, podemos denominar as célula da matriz, conforme a Figura 40.

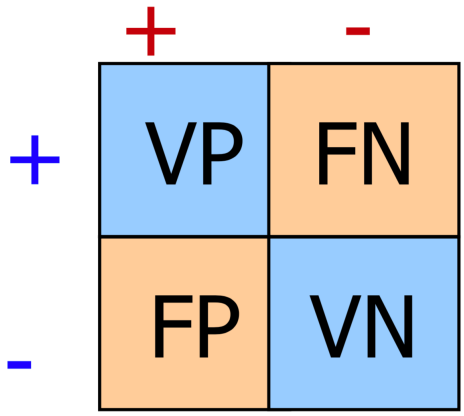


Figura 40. Exemplo de uma matriz de confusão.

Assim, VP são os verdadeiros positivos e correspondem ao número de exemplos da classe positiva classificados corretamente. VN são os verdadeiros negativos e correspondem ao número de exemplos da classe negativa classificados corretamente. FP são os falsos positivos e correspondem ao número de exemplos da classe negativa classificados incorretamente como positivos. FN são os falsos negativos e correspondem ao número de exemplos da classe positiva classificados incorretamente como negativos.

Para um problema n-ário com $n > 2$, nós podemos definir as células VP, VN, FP e FN da seguinte forma:

- $VP = M_{ii}$ onde i corresponde a classe analisada.
- $FP = M_{ij}$ onde j corresponde a classe analisada e $i \neq j$.
- $FN = M_{ji}$ onde j corresponde a classe analisada e $i \neq j$.
- VN são todos os outros valores da matriz não definidos pelas regras anteriores.

A partir dessa definição, é possível utilizar as métricas de *precision* (4.31), *recall* (4.32) e *F1-score* (4.33) para qualquer problema de classificação.

$$\text{precision} = \frac{VP}{(VP + FP)} \tag{4.31}$$

$$\text{recall} = \frac{VP}{(VP + FN)} \tag{4.32}$$

$$\text{F1-score} = 2 \times \frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})} \tag{4.33}$$

A métrica *precision*, para uma classe C_x, verifica a proporção de amostras corretamente classificadas como C_x, em função das amostras que foram classificadas como C_x. Quanto mais próxima de zero for a precisão de uma classe, maior será a quantidade de falsos positivos desta classe. Em um problema de classificação, boas taxas de precisão indicam a corretude do classificador. Ou seja, que o classificador conseguiu aprender quais as amostras não pertencem a uma determinada classe C_x.

Em relação a métrica *recall*, essa verifica a proporção das amostras que foram classificadas como C_x, em função das amostras que deveriam ser classificadas como C_x. A *recall* obtida de uma classe C_x é inversamente proporcional a quantidade de falsos negativos desta classe. Quanto mais próxima de 1 for a *recall*, menor será a quantidade de falsos negativos. Em um problema de classificação multiclasses, boas taxas de *recall* indicam a completude do classificador, ou seja, que o classificador conseguiu generalizar que as amostras de uma classe C_x são realmente da classe C_x.

Como observado na equação 4.33, a métrica *F1-score* utiliza as métricas *precision* e *recall* em seu cálculo. A métrica *F1-score* pode ser interpretada como a média harmônica das medidas de precisão e revocação. Quanto mais próximo de 1 for o *F1-score*, melhor será balanceamento da completude e corretude do modelo.

Outra métrica que também foi utilizada no terceiro experimento descrito no Capítulo 7 foi a *Cohen's kappa*, que é uma estatística utilizada para testes de confiabilidade entre avaliadores. Essa métrica pode variar de -1 a $+1$, onde valores ≤ 0 indicam que a concordância esperada é aleatória, e 1 representa a concordância perfeita entre os avaliadores. Na perspectiva da aprendizagem de máquina, essa métrica pode ser utilizada para medir a proximidade das instâncias rotuladas pelos classificadores em relação aos dados rotulados na base de dados. O kappa pode ser calculado conforma a equação 4.34.

$$kappa = \frac{p_o - p_e}{1 - p_e} \tag{4.34}$$

Na equação 4.34, p_o corresponde à concordância observada e p_e é a probabilidade hipotética de concordância ao acaso. Considerando a matriz de confusão da Figura 40, p_o e p_e podem ser calculados conforme as equações 4.35 e 4.36.

$$p_o = \frac{(VP + VN)}{(VP + VN + FP + FN)} \tag{4.35}$$

$$p_e = p_x + p_y \text{ onde,}$$

$$p_x = \frac{(VP + FN)}{(VP + VN + FP + FN)} \tag{4.36}$$

$$p_y = \frac{(VN + FP)}{(VP + VN + FP + FN)}$$

4.2.3 Kolmogorov-Smirnov test

No segundo e terceiro experimento apresentado no Capítulo 6, nós construímos e comparamos diferentes modelos de aprendizado de máquina. Com o objetivo de verificar a significância estatística entre os modelos, aplicou-se o teste *Kolmogorov-Smirnov* que é um teste não paramétrico que mede a distância entre duas distribuições de probabilidade. Assumindo que cada modelo de aprendizado de máquina gera uma distribuição dos dados de saída em função da distribuição dos dados de entrada, o que queremos verificar com o teste *Kolmogorov-Smirnov* é o quão próximas estão as distribuições de probabilidade geradas por cada classificador. A escolha pela a aplicação do teste *Kolmogorov-Smirnov* se justifica devido a não observação da distribuição normal em todos os modelos construídos nos experimentos.

O teste *Kolmogorov-Smirnov* é conduzido como um teste de hipótese estatística. A hipótese nula, H_0, determina que as duas distribuições que estão sendo testadas possuem a mesma distribuição. O teste busca evidências de que a hipótese nula possa ser rejeitada. Se a probabilidade de as amostras serem de diferentes distribuições excederem um nível de confiança, a hipótese H_0 é rejeitada em favor da hipótese H_1, de que as duas amostras são de distribuições diferentes. Como observado na Figura 41, o que o teste busca é a distância vertical máxima entre as funções de distribuição cumulativa empírica das duas amostras. A distribuição empírica de uma amostra é a proporção dos valores de amostra que são menores ou iguais a um determinado valor.

Suponha que X_i são n observações independente e identicamente distribuídas. A função de distribuição cumulativa empírica F_n é dada pela equação 4.37, onde I é uma função indicadora a qual é igual a 1 se X_i for menor ou igual a x, e 0, caso contrário.

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty,x]}(X_i) \tag{4.37}$$

Para duas amostras com funções empíricas de distribuição cumulativa $F_n(x)$ e $G_m(x)$, a estatística do teste de *Kolmogorov-Smirnov* se dá pela diferença absoluta máxima

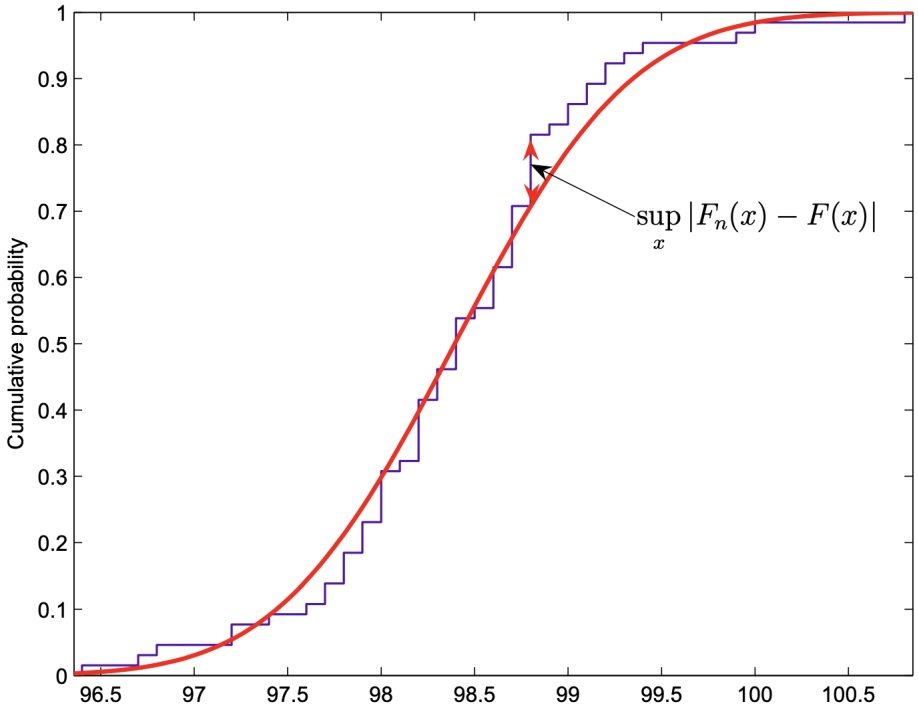


Figura 41. Exemplo da busca pela distância vertical máxima.

entre $F_n(x)$ e $G_m(x)$ para o mesmo x, ou seja, a maior distância vertical entre os gráficos.

$$D = \sup_{-\infty<x<\infty} |F_x(x) - G_m(x)| \tag{4.38}$$

5 Análise da distração do motorista a partir da observação do comportamento do condutor

Este capítulo tem como objetivo descrever a pesquisa e o desenvolvimento do modelo de aprendizado de máquina construído para classificar a distração do motorista em decorrência da utilização do telefone celular. As distrações provenientes de chamadas telefônicas e composição de mensagens de texto foram analisadas a partir da metodologia de observação do comportamento do motorista. O experimento realizado nesse capítulo está inserido no contexto da investigação da observação macroscópica. Nesse sentido, imagens de uma câmera acoplada ao veículo foram utilizadas para treinar e avaliar a capacidade de generalização da Rede Neural Convolucional desenvolvida.

As principais contribuições desse capítulo são:

- Desenvolvimento de um modelo de observação macroscópica que alcance o estado da arte, no que diz respeito a classificação da distração do motorista em decorrência da utilização do telefone celular.
- Justificativa da arquitetura da Rede Neural Convolucional desenvolvida.
- Metodologia de avaliação da capacidade de generalização do modelo.

O decorrer do capítulo está dividido em três principais seções. Na primeira seção é discutida a motivação do método proposto. Na segunda, apresentamos a metodologia da proposta, as avaliações realizadas e a análise dos resultados. Por fim, na última seção, são apresentadas as considerações finais desse estudo.

5.1 Motivação

Detectar e classificar a distração do motorista é um evento importante para a prevenção de acidentes. Nesse sentido, diferentes metodologias podem ser aplicadas para essa finalidade. Como visto no Capítulo 2, as metodologias de observação do comportamento macroscópico, de observação do comportamento microscópico e de observação dos efeitos do comportamento distraído são abordadas na literatura. Nesse estudo, a observação do comportamento macroscópico foi escolhida por ser a forma de observação mais simples e, consequentemente, por requerer menos recursos para a captura de dados. No contexto da

distração do motorista, os estudos observacionais macroscópicos normalmente analisam o comportamento do condutor a partir de câmeras. No estudo em questão, as imagens de uma câmera acoplada ao veículo foram utilizadas como dados de entrada do modelo desenvolvido. Considerando que a observação macroscópica implica a observação do motorista, o modelo desenvolvido com essa metodologia também pode ser utilizado para outras finalidades, que não sejam a de prevenção de acidentes. Por exemplo, no contexto de seguro de veículos, em vez de considerar as estatísticas baseadas em grupos como idade, gênero e estado civil, as empresas seguradoras poderiam utilizar o modelo de observação macroscópica para desenvolver soluções inteligentes como *Usage-Based Insurance* (UBI) ou *Pay-How-You-Drive* (PHYD). A aplicação dessas soluções proporcionaria o aperfeiçoamento dos serviços e tornaria o seguro de automóvel mais barato, ao recompensar os motoristas com bons comportamentos de condução.

Nesse estudo, a motivação para o desenvolvimento de uma Rede Neural Convolucional (CNN) se deu pelo desempenho que essa rede tem demonstrado em áreas correlatas. Estudos como os de (SHARMA; JAIN; MISHRA, 2018; RAWAT; WANG, 2017; LECUN; BENGIO; HINTON, 2015) demonstram que as CNNs apresentam bons desempenhos, quando trabalham com classificação de imagens. De um modo geral, a rede neural convolucional é uma categoria de rede neural que tem demonstrado eficiência em áreas como classificação de objetos, reconhecimento de faces, de ações e de fala. Trabalhos como (SHARMA; JAIN; MISHRA, 2018; KAMENCAY et al., 2017; ZHU et al., 2017; ABDEL-HAMID et al., 2014) apresentam arquiteturas de CNN para as aplicações das áreas mencionadas.

Na revisão sistemática apresentada no Capítulo 2, foram filtrados apenas dois trabalhos que desenvolveram modelos de CNN (LE et al., 2016; VORA; RANGESH; TRIVEDI, 2017). Assim como a proposta dessa pesquisa, o trabalho realizado por (LE et al., 2016), também utilizou a observação macroscópica para detectar a distração do motorista. Na ocasião, (LE et al., 2016) propôs uma CNN para determinar se um motorista está usando um telefone celular, bem como detectar se suas mãos estão no volante. Na arquitetura desenvolvida por (LE et al., 2016), a CNN foi projetada para detectar a distração a partir das regiões de interesse. O fato de poucos trabalhos que desenvolveram CNN terem sido filtrados, também é outro fator que motivou o desenvolvimento desse estudo. Diferentemente da CNN desenvolvida por (LE et al., 2016), nesse estudo, a CNN proposta tem como objetivo classificar a distração do motorista a partir da observação global do comportamento do motorista.

5.2 Metodologia

A modelagem e experimentação desse estudo foi realizada a partir de uma base de dados disponibilizada pela *State Farm*[1], que é uma grande empresa de seguros e serviços financeiros dos Estados Unidos. A base de dados foi construída a partir da observação macroscópica utilizando uma câmera acoplada ao veículo. No total, foram coletadas 11.745 imagens, das cinco classes investigadas:

- C_0: Motorista falando ao telefone com a mão esquerda;
- C_1: Motorista falando ao telefone com a mão direita;
- C_2: Motorista lendo/digitando mensagem de texto com a mão esquerda;
- C_3: Motorista lendo/digitando mensagem de texto com mão direita;
- C_4: Motorista dirigindo de forma segura.

As Figuras 42 e 43 ilustram, respectivamente, as quatro classes de distrações investigadas e o comportamento esperado na direção segura. Como pode ser observado, a direção segura é caracterizada quando o motorista está com as duas mão no volante e olhando para frente. A Tabela 10 apresenta a distribuição das amostras por cada classe do problema.

Tabela 10. Distribuição das amostras por classes.

	C0	C1	C2	C3	C4	Total
Total	2489	2267	2317	2346	2326	11745

A Figura 52 demonstra a metodologia de aquisição dos dados e extração das características utilizada nesse experimento. Uma câmera foi utilizada para monitorar o comportamento do motorista durante a condução do veículo. Originalmente, as imagens coletadas possuíam dimensões de 640×480 *pixels*. Para diminuir a carga de processamento, as imagens foram redimensionadas para matrizes quadradas de 150 *pixels*. Também foi realizada a normalização *Min-Max*[2] das matrizes. Por se tratar do desenvolvimento de uma CNN, optou-se pela não aplicação de outras técnicas de pré-processamento de imagem por acreditar que os filtros das camadas convolucionais desempenham este papel. A Figura 45 descreve a fase de pré-processamento realizada neste estudo.

[1] https://www.statefarm.com/

[2] $x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$

Figura 42. Tipos de distração do motorista analisados. Imagens da base de dados utilizada no experimento.

Figura 43. Comportamento esperado na direção segura. Imagem da base de dados utilizada no experimento.

5.2.1 Preparação da base de dados

A Figura 46 descreve a metodologia utilizada para o desenvolvimento e análise da CNN proposta nesse experimento. Como pode ser observado, a metodologia é dividida em três etapas: divisão da base de dados, treinamento do modelo, avaliação do modelo. Por se tratar do desenvolvimento de um modelo de aprendizado de máquina supervisionado, na primeira etapa, foi realizada a separação da base de dados em conjuntos de treinamento e teste. Em seguida, o conjunto de imagens de treinamento foi subdivido em conjunto de treinamento e de validação, utilizando a estratégia da validação cruzada. A Tabela 11

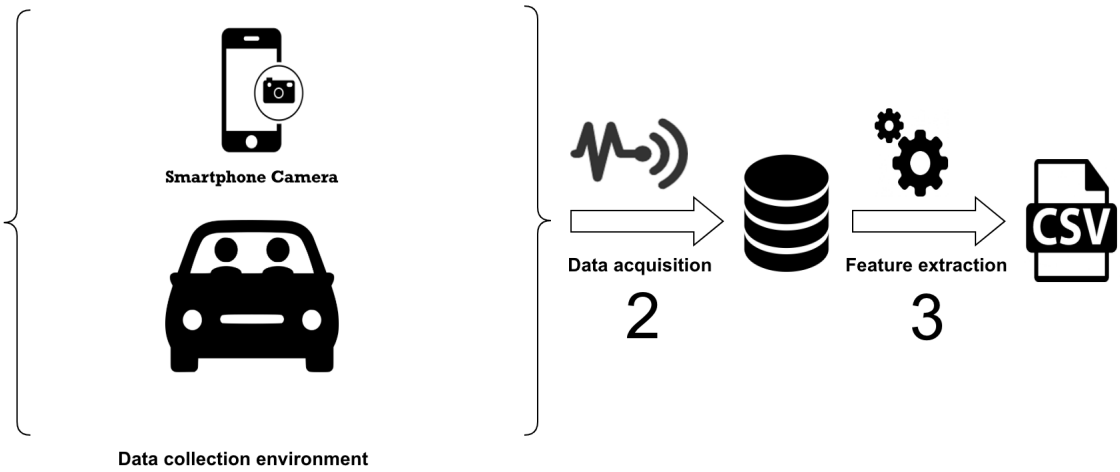


Figura 44. Metodologia utilizada na aquisição dos dados e extração das características do primeiro experimento.

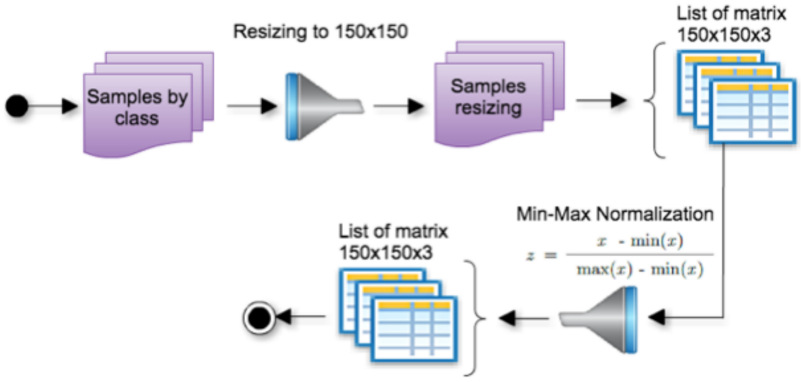


Figura 45. Pré-processamento realizado na base de imagens.

ilustra a quantidade de amostras, por classe, nos conjuntos de treinamento e teste.

Tabela 11. Distribuição das amostras, por classes, dos conjuntos de treinamento e validação.

	C0	C1	C2	C3	C4	Total
Treinamento	2088	1864	2065	1936	1888	9841
Teste	401	403	252	410	438	1904
Total	2489	2267	2317	2346	2326	11745

A utilização desses três conjuntos de dados é importante para treinar, ajustar e avaliar o modelo desenvolvido. Segundo (RIPLEY; HJORT, 1995), o conjunto de treinamento é utilizado para o aprendizado do modelo. Já o conjunto de validação é o conjunto responsável por fornecer diretrizes para os ajustes do modelo em desenvolvimento. Por

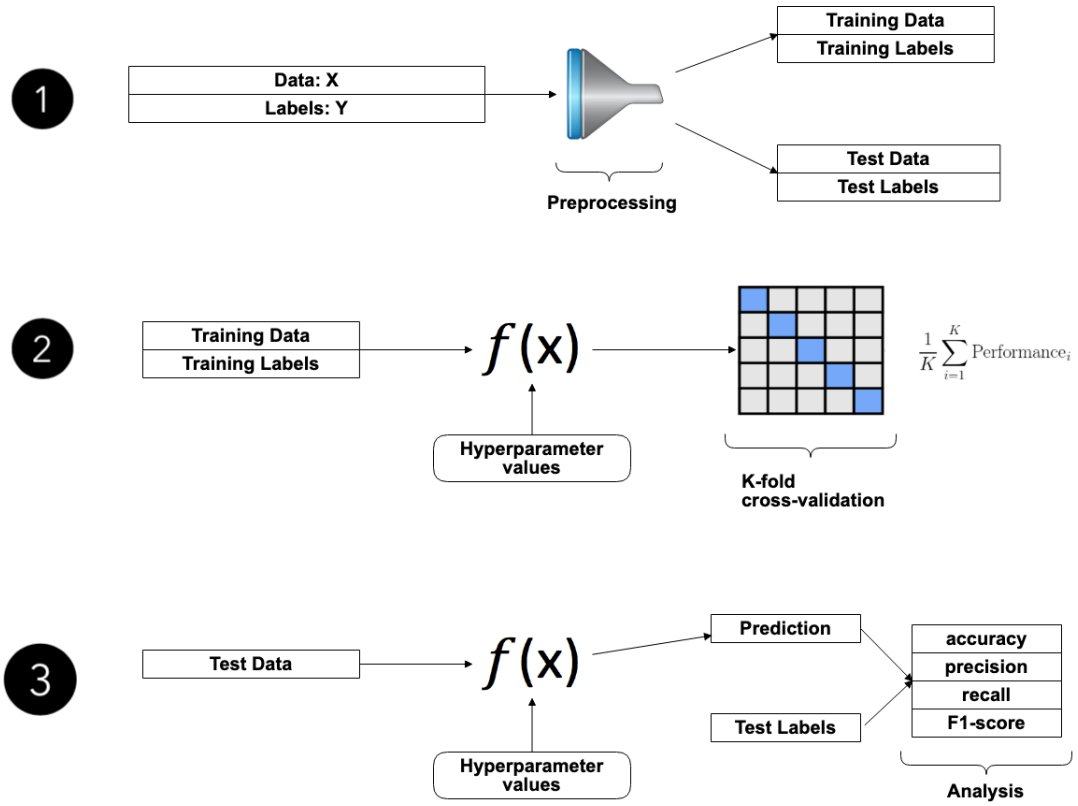


Figura 46. Fluxo do experimento.

.

fim, para (RIPLEY; HJORT, 1995), o conjunto de teste é o conjunto utilizado somente para avaliar o modelo já desenvolvido.

É importante destacar que os conjuntos de imagens de treinamento, teste e validação não possuem interseção, ou seja, não existem imagens em comum nesses conjuntos. A separação dos dados é necessária para avaliar a capacidade de generalização do modelo. Se um modelo treinado com um conjunto de dados consegue classificar, corretamente, amostras inéditas, então, assume-se que este modelo possui a capacidade de generalização.

No modelo de aprendizado supervisionado, a generalização está relacionada com o conceito de *overfitting*. Se o desempenho do modelo com os dados rotulados for, drasticamente, superior ao seu desempenho com os dados inéditos, assume-se que este modelo não possui capacidade de generalização e, por isso, apresentou *overfitting*. Conforme apresentado na Figura 47, o comportamento de *overfitting* pode ser verificado pela análise da curva de aprendizado. Considerando a métrica *accuracy*, o *overfitting* é caracterizado quando a *accuracy* do treinamento apresenta boa performance e as curvas de treinamento e validação estão distantes entre si. Este comportamento demonstra que o modelo apresentou alta variância em relação aos dados.

Além do *overfitting*, existem situações nas quais o modelo desenvolvido fica aquém da *accuracy* esperada, tanto na fase de treinamento, quanto na fase de validação. Quando isto

acontece, dizemos que o modelo apresentou *underfitting*. Na Figura 47, o comportamento de *underfitting* pode ser visualizado no gráfico do canto superior esquerdo.

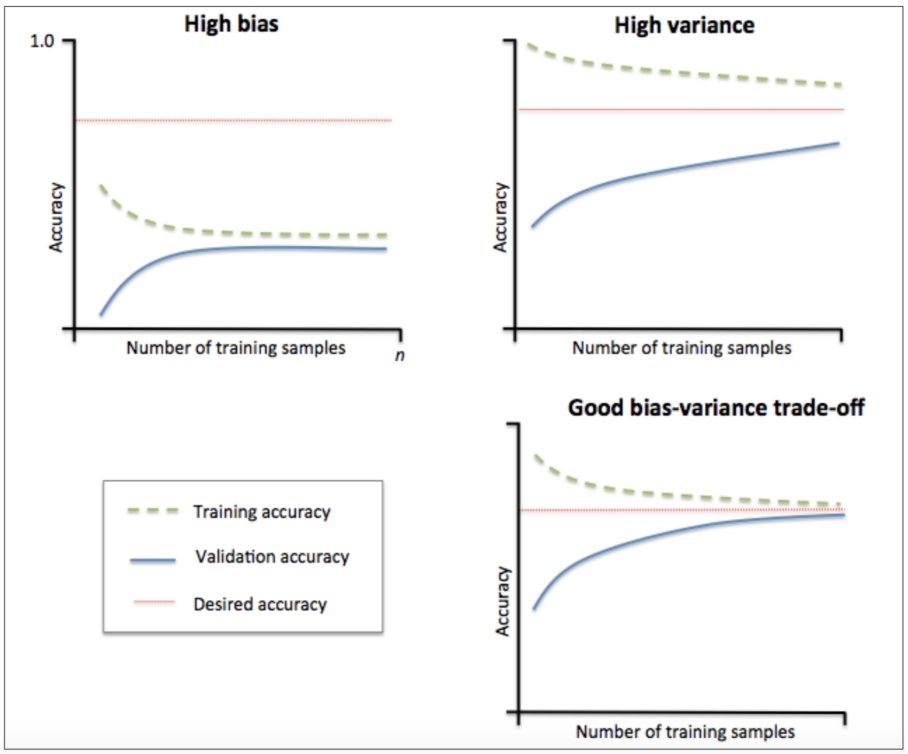


Figura 47. Possíveis comportamentos durante a generalização do problema. Fonte: (RASCHKA; MIRJALILI, 2017)

.

Quando se está desenvolvendo um modelo de aprendizado de máquina, espera-se que o modelo desenvolvido possua a capacidade de generalização, mesmo para dados inéditos. Tecnicamente, estamos almejando que o modelo desenvolvido possua baixo *bias* e variância. Uma das formas de verificar esse comportamento se dá pela separação dos dados em conjuntos disjuntos de treinamento e validação. Para esses dois conjuntos, se o comportamento observado for similar ao do gráfico inferior direito da Figura 47, então, o modelo desenvolvido, provavelmente, terá a capacidade de generalização esperada.

5.2.2 Desenvolvimento do modelo

A segunda etapa da metodologia apresentada na Figura 46 refere-se ao desenvolvimento e treinamento do modelo. Considerando a fundamentação apresentada no Capítulo 4, a Figura 48 ilustra a arquitetura da rede neural convolucional desenvolvida neste estudo. Cada camada de convolução foi formada pela 3-úpla <convolução, função de ativação, *pooling*>. Nas duas primeiras camadas, foram utilizados 32 *kernels* de dimensão 3×3, a função de ativação ReLU e a função *maxpooling*. Na terceira camada, a única diferença em relação às camadas anteriores foi o número de *kernels* utilizados na convolução. Foram

utilizados 64 *kernels* de dimensão 3×3 e a função de ativação e *pooling* continuaram sendo a ReLU e *maxpooling*, respectivamente.

Acrescentou-se à última camada oculta da rede uma *Multilayer perceptron* (MLP) totalmente conectada com 64 neurônios e, por se tratar de um problema de classificação com cinco classes, a camada de saída da rede foi formada por cinco neurônios. As funções de ativação utilizadas na última camada oculta e na camada de saída foram, respectivamente, a função ReLU e *sigmoid*.

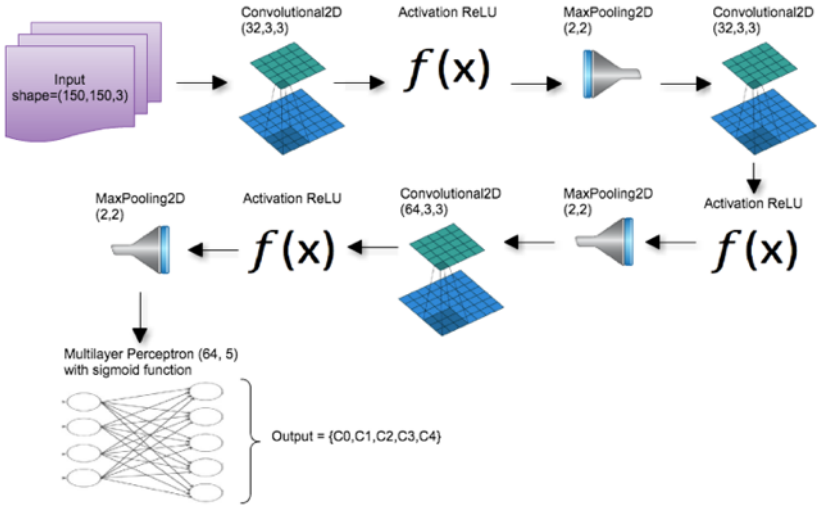


Figura 48. Arquitetura da CNN desenvolvida no segundo experimento.

A construção desta arquitetura levou em consideração as observações realizadas por (SZEGEDY et al., 2016). Na construção de uma CNN, o autor recomenda e demonstra que a utilização de *kernels* de pequena dimensão é mais eficiente do que a utilização de *kernels* maiores. Além de diminuir a carga de processamento, (SZEGEDY et al., 2016) também ressalta que a utilização de múltiplos filtros de pequena dimensão conseguem equiparar a representatividade de filtros maiores. O conceito de representatividade está relacionado à capacidade da convolução conseguir detectar mudanças estruturais no objeto analisado.

Nesse experimento, optou-se pela utilização de 32 filtros de dimensão 3×3 nas duas primeiras camadas de convolução e de 64 filtros de mesma dimensão na terceira camada. A justificativa para aumentar a quantidade de filtros na terceira camada está pautada no comportamento da convolução. Como observado por (CHOLLET, 2017), as primeiras camadas de convolução são utilizadas para filtrar características abstratas, como identificação de borda, cores e algumas formas simples que dependem do contexto do objeto analisado. Já as camadas de convolução mais profundas trabalham no reconhecimento de características mais complexas. Por exemplo, no estudo de reconhecimento facial realizado por (KAMENCAY et al., 2017), foi observado que as últimas camadas de convolução são responsáveis pelo reconhecimento das partes que caracterizam uma face, tais como nariz, boca e olhos.

Em relação ao treinamento da CNN, o ajuste dos pesos foram calculados a partir da função de custo *cross-entropy* 5.1. Nesta equação, temos que:

$$L(\Theta) = -\frac{1}{n}\sum_{i}^{n}\sum_{j}^{m} y_{ij} log(p_{ij}) \tag{5.1}$$

- n corresponde ao número de classificações realizadas.
- m corresponde ao número de classes.
- y refere-se aos rótulos das amostras.
- p refere-se ao valor de predição das amostras.

A Figura 49 apresenta o erro médio de treinamento e teste de cada época. Como pode ser observado, o treinamento do modelo foi realizado em 13 épocas. Em cada época, o modelo foi treinado e avaliado com a validação cruzada de *10-folds*. Conforme definido no Capítulo 4, a validação cruzada *K-fold* é um método de avaliação que divide o conjunto de treinamento em K subconjuntos mutuamente exclusivos de mesmo tamanho e, então, a cada iteração, separa um *fold* para validação e utiliza os outros $k-1$ *folds* para o treinamento do modelo. K iterações devem ser realizadas para que todos os *folds* sejam utilizados, para avaliar o desempenho do modelo.

Em cada época, foram realizadas 10 iterações de treinamento. Como o número de amostras por classe era diferente, a formação dos *folds* foi realizada de forma balanceada. Cada *fold* possuía 984 amostras, respeitando a proporção de amostras por classe. Ao término de cada época, foi calculado o erro médio do treinamento e validação. O erro de cada *fold* foi calculado conforme a equação 5.2, que verifica a proporção de amostras classificadas, erroneamente, dentro do *fold*.

$$\text{Erro}_{fold} = (\frac{1}{984}\sum_{i}^{984} I(y_i \neq f(x_i))) \tag{5.2}$$

Na Figura 49, é possível perceber que, com o passar das épocas, o erro de treinamento e validação foram diminuindo e ficando cada vez mais próximos de zero. A minimização do erro e a proximidade das curvas sugerem que o modelo está adquirindo a capacidade de generalização e, consequentemente, eliminando a possibilidade de *overfitting*. Na última época de treinamento, a taxa de erro de treinamento e validação foram, respectivamente, de 0.0159 e 0.0074.

5.2.3 Avaliação do modelo

Na terceira fase da metodologia, o modelo foi submetido ao processo de teste. Assim, 1904 amostras foram separadas para testar a eficácia do modelo. A realização

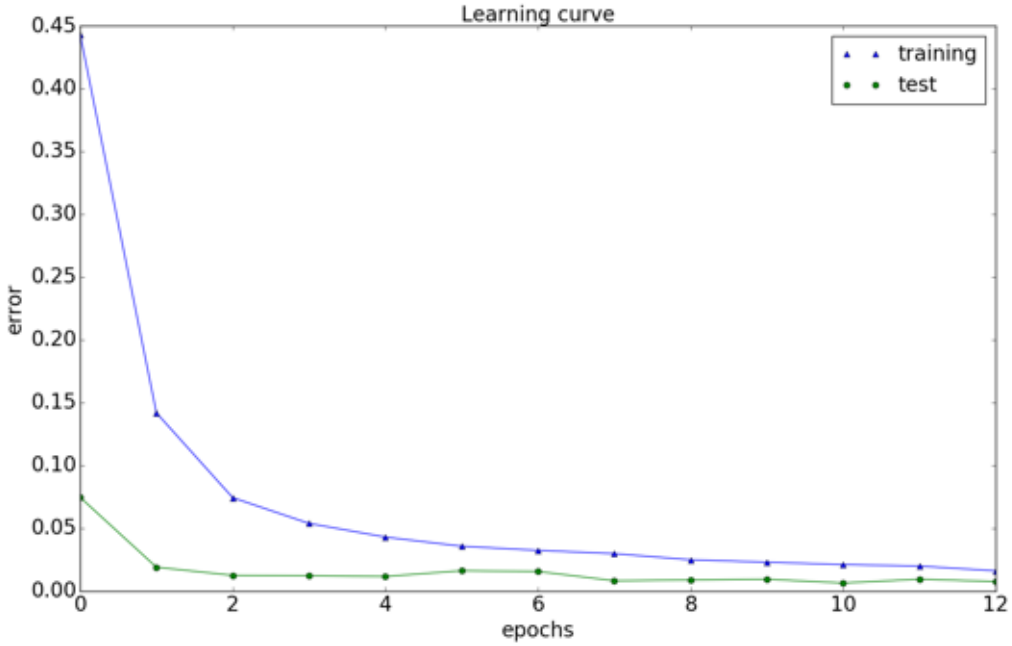


Figura 49. Curva de aprendizado do modelo em função do erro.

do teste teve o objetivo de certificar os resultados obtidos no treinamento e validação do modelo. O resultado do teste pode ser observado na matriz de confusão exibida na Figura 50. Na diagonal principal da matriz estão as amostras que foram corretamente classificadas pela CNN. Nas demais posições estão as amostras erroneamente classificadas.

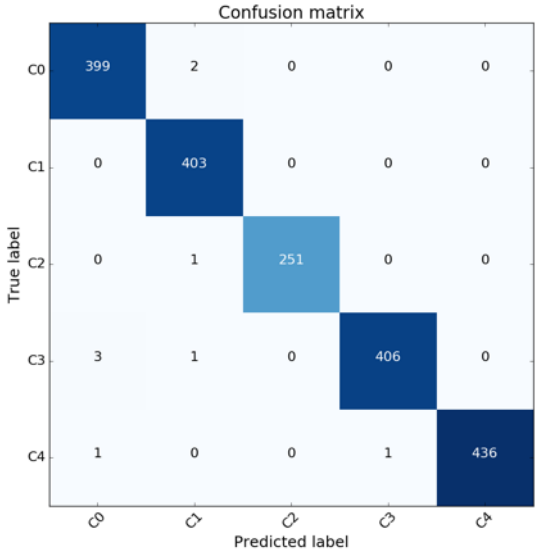


Figura 50. Matriz de confusão construída com os dados separados, para testar a capacidade de generalização do modelo.

Na matriz de confusão, os valores de cada classe podem ser definidos como verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN). Os valores verdadeiros positivos de uma classe C_x correspondem às amostras cor-

retamente classificadas como C_x. Já os valores verdadeiros negativos de uma classe C_x referem-se às amostras corretamente não classificadas como C_x. Por exemplo, na matriz de confusão multiclasses apresentada na Figura 50, os verdadeiros negativos da classe C_0 estão representados pelos valores da diagonal principal, excluindo a célula C_{00}.

Ainda utilizando a classe C_0 como exemplo, os valores falsos negativos desta classe são os valores que pertencem à classe C_0 mas foram incorretamente classificados como pertencentes a outras classes. Na matriz de confusão, os valores falsos negativos de uma classe C_x estão dispostos horizontalmente. No caso da classe C_4, que corresponde ao motorista dirigindo de forma segura, podemos observar que houveram somente duas amostras falsas negativas. As amostras foram erroneamente classificadas como pertencentes às classes C_1 e C_3. A classe C_1 refere-se ao motorista falando ao telefone com a mão direita, enquanto que a classe C_3 refere-se ao motorista lendo/digitando mensagem de texto com a mão direita.

Levando em consideração que o modelo desenvolvido tem o objetivo de detectar a distração do motorista em decorrência da utilização do celular, os falsos negativos da classe C_4 significam a inconveniência do sistema. Ou seja, um falso negativo da classe C_4 implica um motorista dirigindo de forma segura, sendo alertado de uma distração que não existe. Neste caso, o alerta pode gerar uma nova distração desnecessária, que poderia ser evitada.

Em relação aos falsos positivos de uma classe C_x, estes estão dispostos verticalmente na matriz de confusão. Os falsos positivos de uma classe C_x são os valores que não pertencem à classe C_x, mas foram classificados como pertencentes. No contexto deste trabalho, os falsos positivos da classe C_4 são críticos, pois implicam distrações não detectadas. Observando a matriz de confusão da Figura 50, pode-se perceber que classe C_4 não possui nenhum falso positivo. Este comportamento nos informa que o modelo desenvolvido não classificou nenhuma distração como direção segura. Em um caso prático, a ausência de falsos positivos da classe C_4 sugere que todas as distrações dos motoristas seriam alertadas pelo sistema.

Com o objetivo de certificar a eficácia do modelo, utilizou-se os valores da matriz de confusão para calcular a *accuracy*, *precision recall* e *F1-score* do modelo. A *accuracy* do modelo foi superior a 99%. Analisando este valor, podemos perceber que o modelo conseguiu alcançar o desempenho esperado, que já havia demonstrado na fase de treinamento e avaliação do modelo. A boa performance obtida com os dados inéditos da fase de teste ratifica a capacidade de generalização da CNN.

A Tabela 12 apresenta a *precision*, *recall* e *F1-score* de cada classe do problema. Analisando os valores apresentados, podemos perceber que a CNN conseguiu boas taxas de corretude e completude em todas as classes. A não disparidade entre os valores de *precision*, *recall* e *accuracy* nos indica que a *accuracy* da fase de avaliação realmente reflete a capacidade de generalização do problema. Se, porventura, fosse observada baixa

performance de *precision* ou *recall* para alguma classe do problema, indicaria que a CNN não havia conseguido a generalização para aquela classe específica. Este comportamento nos indicaria que, apesar da boa *accuracy*, ainda assim, o sistema não seria eficiente em casos reais. Como este não foi o caso, podemos concluir que a CNN desenvolvida está apta para detectar a distração do motorista, em decorrência da utilização do celular.

Tabela 12. Computação das métricas *precision*, *recall*, *F1-score*.

Métricas	C_0	C_1	C_2	C_3	C_4
Precision	0.99	0.99	1	0.99	1
Recall	0.99	1	0.99	0.99	0.99
F1-score	0.99	0.99	0.99	0.99	0.99

Nesse estudo, além de detectar a distração, para garantir a segurança de passageiros e motoristas, também foi objetivo classificar o tipo de distração. Nesse caso, era necessário maximizar a precisão e revocação de todas as classes do problema. Conforme observado na Tabela 12 e na matriz de confusão da Figura 50, podemos perceber que as taxas de *precision* e *recall* foram todas maximizadas e que a quantidade de falsos positivos e falsos negativos foi pequena. Este comportamento, somado à analise anterior, demonstra que a CNN desenvolvida é eficiente para classificar a distração do motorista em decorrência da utilização do celular.

5.3 Considerações finais

Este estudo realizou um experimento naturalístico com o objetivo de detectar e classificar a distração do motorista em decorrência da utilização do telefone celular. As distrações de digitar uma mensagem de texto e realizar uma chamada telefônica foram analisadas macroscopicamente, a partir de imagens coletadas por uma câmera fixada no painel do veículo. Levando em consideração os dados disponíveis, este estudo desenvolveu uma arquitetura de rede neural convolucional, cujo objetivo foi classificar a distração do motorista ou detectar se a condução estava sendo realizada em segurança. Segundo as amostras da base de dados, a direção segura era caracterizada quando o motorista estava com as duas mãos no volante e com olhar fixo para frente.

A avaliação da CNN foi realizada pela observação da capacidade de generalização do modelo. Para isso, os dados foram divididos em conjuntos disjuntos de treinamento, teste e avaliação. Durante a fase de treinamento e avaliação, verificamos se o modelo apresentava características de *overfitting*. Uma vez constatada a não existência de *overfitting*, certificamos que o modelo desenvolvido está apto para trabalhar com dados desconhecidos.

Na fase de teste, analisamos o desempenho do modelo de forma global e local

para cada classe do problema. Tal avaliação foi realizada com as métricas de *accuracy*, *precision*, *recall*e *F1-score*. A aplicação da métrica *accuracy* analisou se o modelo, em uma perspectiva global, possuía a capacidade de classificar as distrações corretamente. As métricas *precision*, *recall* e *F1-socre* foram utilizadas para verificar o desempenho individual do modelo para cada classe de distração. Nessa etapa da análise, verificamos que o modelo possui bom desempenho para todas as classes de distração e também para a classe correspondente à direção segura.

De acordo com as tarefas planejadas para comtemplar o objetivo desta tese, este experimento corresponde a realização da seguinte tarefa:

- Realização de um experimento utilizando a metodologia de análise do comportamento do condutor para detectar os três tipos de distração do motorista provenientes da utilização do telefone celular.

Os resultados deste experimento foram publicados em 2017 na *26th International Conference on Artificial Neural Networks* com o título *A Deep Learning Approach to Detect Distracted Drivers Using a Mobile Phone*.

6 Análise da distração do motorista devido à composição de mensagens de texto

Este capítulo tem como objetivo descrever a pesquisa e o desenvolvimento dos modelos de aprendizado de máquina construídos para classificar passageiros e motoristas durante a distração proveniente da composição de mensagens de texto. Essa pesquisa foi conduzida a partir da metodologia de observação dos efeitos do comportamento do condutor. As pesquisas que utilizam essa metodologia partem da ideia de que os padrões de direção de um motorista distraído são diferentes dos padrões de direção de um motorista que está dirigindo com atenção. A hipótese é de que um motorista distraído tende a apresentar variações nos padrões de velocidade, aceleração, desaceleração e direção do veículo. Além disso, por se tratar de um estudo que investiga a distração devido à composição de mensagens de texto, nós também assumimos a hipótese de que os padrões de digitação de um motorista distraído são diferentes dos padrões de digitação de um passageiro. A fim de observar esses padrões, um estudo naturalístico foi realizado para coletar dados utilizando os sensores acelerômetro, magnetômetro, giroscópio, GPS e *touch screen* de um *smartphone*. Os resultados demonstram que os modelos desenvolvidos alcançam o estado da arte do tema investigado.

As principais contribuições desse capítulo são:

- Construção da base de dados naturalística com dados sensoriais de motoristas e passageiros, referentes a composição de mensagens de texto.
- Desenvolvimento de modelos de observação do efeito do comportamento que alcance o estado da arte, no que diz respeito a classificação de motoristas e passageiros durante a distração proveniente da composição de mensagens de texto.
- Metodologia de avaliação da capacidade de generalização dos modelos.

O decorrer do capítulo está dividido em três principais seções. Na primeira seção é discutida a motivação do método proposto. Na segunda, apresentamos a metodologia da proposta, as avaliações realizadas e a análise dos resultados. Por fim, na última seção, são apresentadas as considerações finais desse estudo.

6.1 Motivação

Segundo (HOSKING; YOUNG; REGAN, 2009), a quantidade de tempo que um motorista fica sem olhar para a pista aumenta em 400% se ele estiver digitando uma

mensagem de texto. Das possíveis distrações que o uso do telefone celular pode ocasionar, esta é considerada a de maior risco. Como já mencionado no Capítulo 1, ao escrever uma mensagem de texto, o motorista está desviando sua atenção física, cognitiva e visual. Levando em consideração os riscos causados por esse tipo de distração, seria apropriado que esse problema fosse resolvido apenas por campanhas educativas e o cumprimento das leis. Entretanto, os dados estatísticos apresentados no Capítulo 1 mostram que essas medidas não são suficientes. Portanto, podemos concluir que os usuários consideram o *texting while drving* um evento necessário, mesmo estando cientes dos riscos agregados.

Para resolver este problema, entendemos que é necessária a aplicação de soluções computacionais não intrusivas. No contexto da distração do motorista, uma solução não intrusiva é aquela que não necessita da interação do usuário para funcionar. Ou seja, em uma solução não intrusiva as chances do usuário burlar o seu funcionamento são minimizadas. Por exemplo, na solução apresentada no experimento do Capítulo 5, nós utilizamos como dados de entrada as imagens de uma câmera acoplada ao veículo. Na referida solução, o usuário conseguiria burlar o sistema, simplesmente, impedindo a captura das imagens. Em soluções intrusivas como a apresentada no primeiro experimento, ainda há a necessidade da conscientização do usuário para que o sistema seja eficiente.

Segundo a revisão de literatura descrita no Capítulo 2, as soluções não intrusivas costumam utilizar os sensores do próprio *smartphone* para coletar os dados de entrada da solução. A revisão aponta que os sensores acelerômetro, giroscópio, GPS, magnetômetro e *touch screen* costumam ser utilizados nas soluções de *texting while drving*. Contudo, a utilização desses sensores não garante que a solução seja não intrusiva. Detectar a digitação de uma mensagem, enquanto o veículo estiver em movimento, não é uma tarefa complexa utilizando esse conjunto de sensores. Porém, proporcionar a conveniência para os usuários, poderia tornar a solução intrusiva. Por exemplo, ao detectar a digitação de uma mensagem de texto, o sistema poderia requisitar uma confirmação para certificar se a mensagem está sendo digitada por um passageiro ou pelo motorista. Nesse caso, o sistema poderia ser burlado pelo motorista, o que característica o sistema intrusivo.

Segundo (LIU et al., 2017), a complexidade do desenvolvimento da solução não intrusiva para prevenir o *texting while drving* é, justamente, conseguir a eficiência e conveniência do sistema. Nesse sentido, este experimento realizou um estudo para analisar o desempenho de diferentes modelos de aprendizado de máquina na tarefa de detectar e classificar o evento de *texting while drving* de forma não intrusiva. Para conseguir a eficiência e conveniência, os modelos de aprendizado de máquina analisados foram construídos para detectar a digitação de mensagem de texto e classificar se essa digitação é realizada por motorista ou passageiro.

6.2 Metodologia

A coleta de dados desse experimento foi realizada de forma naturalística, utilizando a observação microscópica e os efeitos do comportamento do motorista. Os dados foram coletados utilizando os sensores acelerômetro, giroscópio, GPS e *touch screen*. A premissa para o desenvolvimento dos modelos é a de que os padrões de direção e digitação de um motorista se alteram, caso ele esteja realizando o evento de *texting while drving*. Por outro lado, se o motorista estiver dirigindo em segurança e o passageiro estiver digitando uma mensagem de texto, acredita-se que os padrões de direção e digitação não são alterados. Nesse sentido, os dados coletados pelos sensores acelerômetro, giroscópio e GPS capturam os padrões de direção enquanto que o sensor *touch screen*, o padrão de digitação.

A coleta dos dados sensoriais foi realizada a partir de um aplicativo desenvolvido para esta finalidade. A Figura 51 apresenta a única tela do aplicativo desenvolvido. Os sensores são ativados no primeiro *touch* realizado pelo usuário e os dados são gravados quando o usuário clica na tecla *SEND*. Para cada mensagem digitada, o aplicativo registra o *log* dos dados sensoriais.

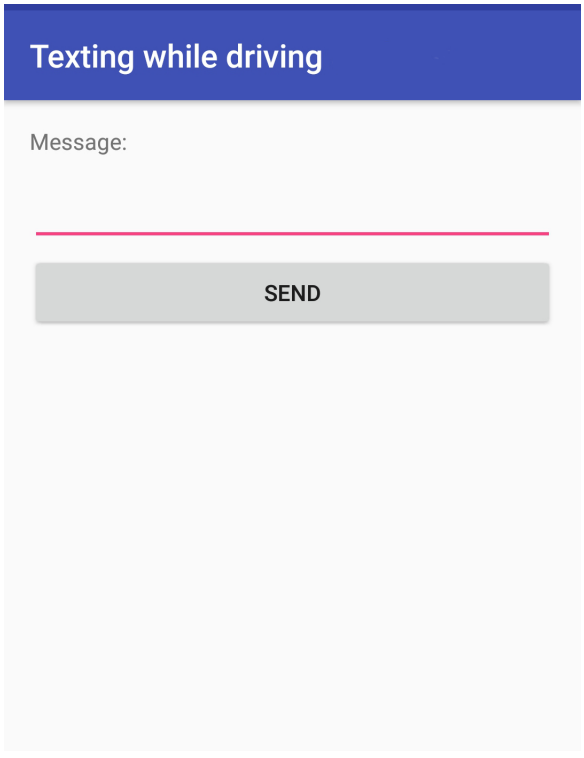


Figura 51. Aplicativo construído para capturar os dados sensoriais.

A Figura 52 demonstra a metodologia de aquisição dos dados e extração das características utilizada nesse experimento. A coleta dos dados foi realizada na cidade de Belém - Brasil, onde 13 pessoas participaram dos experimentos. Todos os participantes tinham idade entre 23 e 29 anos, e 8 eram do sexo masculino e 5 do sexo feminino. A faixa etária escolhida justifica-se, pois as lesões causadas pelo trânsito são a principal causa

de morte entre pessoas de 15 a 29 anos de idade (CHAN, 2015) e, conforme destacado por (EHSANI; LI; SIMONS-MORTON, 2015), aproximadamente 62% dos motoristas adolescentes admitem o envio de mensagens de texto enquanto dirigem.

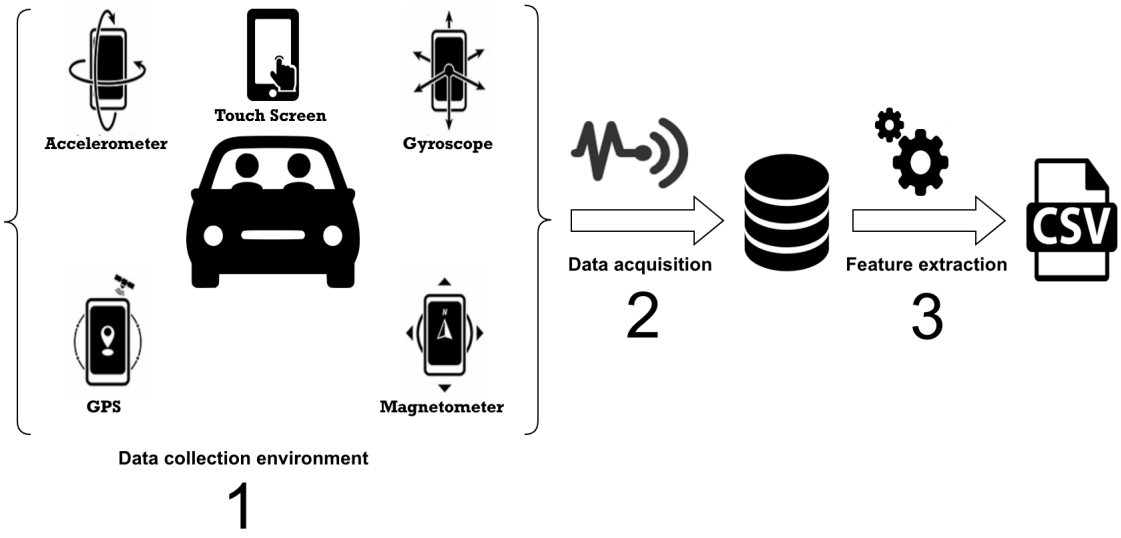


Figura 52. Metodologia utilizada na aquisição dos dados e extração das características do segundo experimento.

A coleta de dados foi dividida em duas fases. Na primeira etapa, todos os participantes assumiram o papel de passageiro. Na segunda etapa, os participantes assumiram o papel de condutores. Em ambas as fases, os participantes foram instruídos a enviar mensagens de texto apenas enquanto o carro estava em movimento. Para garantir a segurança dos motoristas, todos eles foram acompanhados por um copiloto que avisou sobre possíveis riscos de colisão. O conjunto completo de dados apresenta 404 tuplas para motoristas e 404 para passageiros.

Para serem caracterizados como dados naturalísticos, os participantes puderam escolher qualquer trajetória na cidade e também as mensagens que desejavam enviar. Em média, cada participante encaminhou trinta mensagens em uma viagem com duração de 30 a 60 minutos. A única restrição era a de que os participantes deveriam enviar as mesmas mensagens nas duas fases da coleta de dados. O conjunto de dados empregado pode ser obtido em <https://goo.gl/FTM5e8>.

A base de dados foi construída a partir dos dados sensoriais coletados pelos voluntários. Para cada mensagem, a base de dados foi elaborada com os seguintes atributos: média e desvio padrão do tempo entre cada *touch*; média e desvio padrão da velocidade; média e desvio padrão dos 3 eixos do acelerômetro; e média e desvio padrão dos 3 eixos do giroscópio. Os dados dos sensores do acelerômetro e do giroscópio foram coletados em uma frequência de 5Hz, e os dados do GPS foram coletados a uma taxa de 1Hz. O toque foi obtido em cada interação do usuário usando a unidade de tempo em microssegundos.

Na Tabela 13, pode-se observar os atributos da base de dados.

Tabela 13. Atributos da base de dados

Database attributes	
mean time between touches	standard deviation time between touches
mean speed	standard deviation speed
mean acceleration X-axis	standard deviation acceleration X-axis
mean acceleration Y-axis	standard deviation acceleration Y-axis
mean acceleration Z-axis	standard deviation acceleration Z-axis
mean gyroscope X-axis	standard deviation gyroscope X-axis
mean gyroscope Y-axis	standard deviation gyroscope Y-axis
mean gyroscope Z-axis	standard deviation gyroscope Z-axis

Ainda na extração de características, foi realizada a normalização dos atributos. Para esta etapa, utilizou-se a equação de padronização 6.1.

$$x_{new} = \frac{x - \mu}{\sigma} \tag{6.1}$$

6.2.1 Preparação da base de dados

A Figura 46 descreve a metodologia utilizada para o desenvolvimento e análise dos modelos desse experimento. Como pode ser observado, a metodologia é dividida em cinco etapas: 1- divisão da base de dados, 2- configuração dos parâmetros, 3- treinamento, 4- avaliação e 5- verificação da significância estatítica. Por se tratar do desenvolvimento de um modelo de aprendizado de máquina supervisionado, na primeira etapa, foi realizada a separação da base de dados em conjuntos de treinamento e teste. Em seguida, o conjunto de dados de treinamento foi subdivido em conjunto de treinamento e de validação, utilizando a estratégia da validação cruzada. A Tabela 14 ilustra a quantidade de amostras, por classe, nos conjuntos de treinamento e teste.

Tabela 14. Distribuição das amostras por classes dos conjuntos de treinamento e validação.

	$C_{\text{passageiro}}$	$C_{\text{motorista}}$	Total
Treinamento	280	280	560
Teste	124	124	248
Total	404	404	808

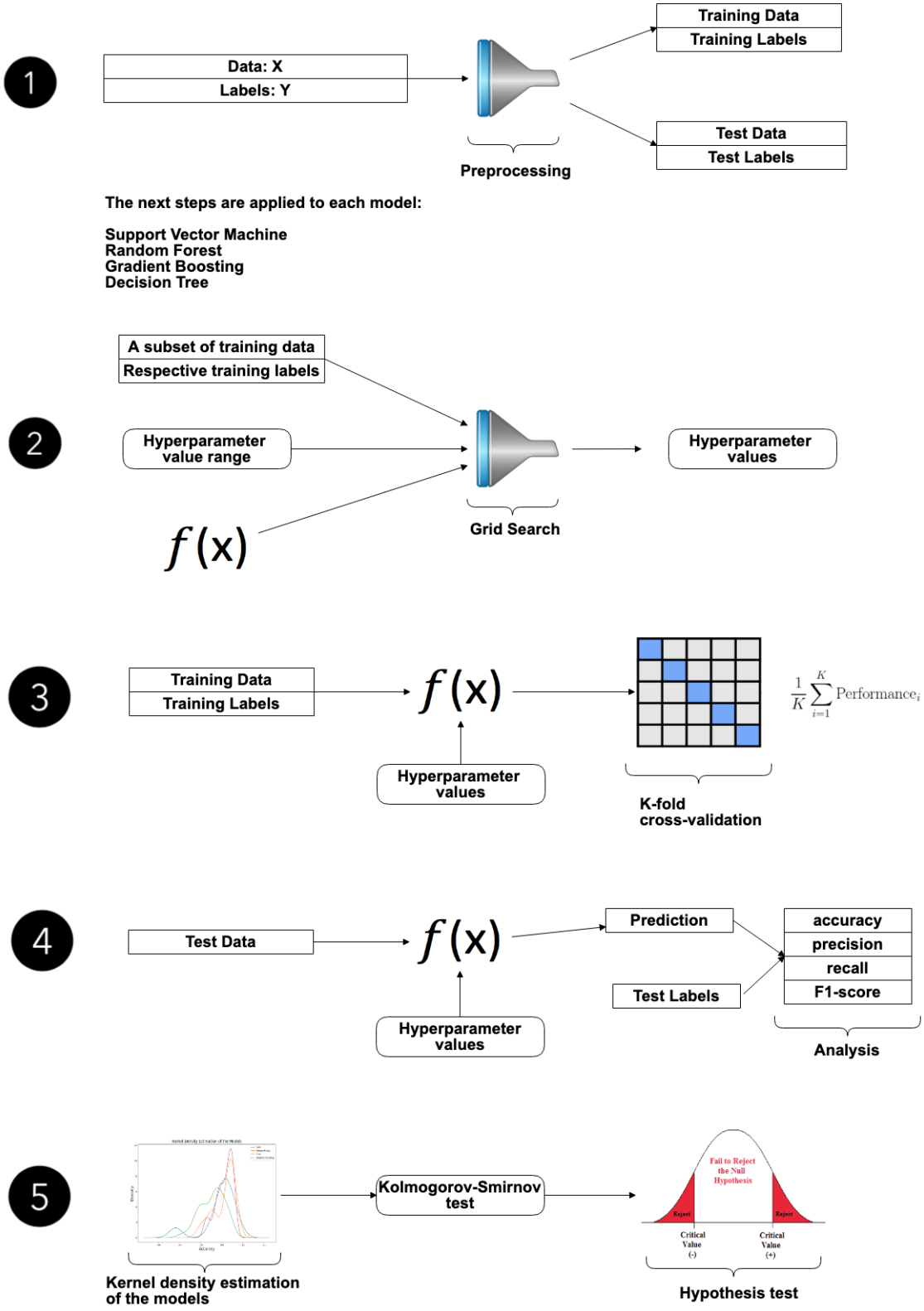


Figura 53. Fluxo do segundo experimento.

.

6.2.2 Configuração dos parâmetros

A próxima etapa do experimento consistiu na escolha dos modelos e na definição de seus parâmetros. Os modelos *Support Vector Machine*, *Decision Tree*, *Random Forest* e *Gradient Boosting* foram escolhidos para esse experimento devido à observação da sua aplicabilidade em áreas correlatas (JUNIOR et al., 2017; BURTON et al., 2017; VLAHOGIANNI; BARMPOUNAKIS, 2017; ZYLIUS, 2017; MANASSEH; SENGUPTA, 2013). Nesse sentido, assumimos a hipótese de que tais modelos também podem apresentar bons resultados no contexto do *texting while driving*.

A construção da maioria dos modelos de aprendizado de máquina implica a configuração de parâmetros. Tais parâmetros podem ser divididos em duas categorias. Os parâmetros aprendidos pelo modelo durante a fase de treinamento, e os parâmetros configurados previamente ao treinamento do modelo. Os parâmetros configurados previamente são denominados hiperparâmetros. No aprendizado de máquina, os hiperparâmetros podem ser configurados manualmente a partir da experiência do projetista, por meio de heurísticas ou por meio de técnicas de força bruta como o *grid search*. No *grid search*, nós devemos definir os hiperparâmetros e o intervalo de valores que são combinados no ajuste do modelo. Em resumo, na técnica de pesquisa de grade, cada combinação de valores representa um modelo que deve ser validado. Nesse experimento, nós usamos a técnica de *grid search* para configurar os hiperparâmetros dos modelos.

Tabela 15. Configuração dos parâmetros do *grid search* do segundo experimento.

Model	N estimators	Criterion	Min sample split	Max depth	Max features
Random Forest	200		2	2	$\sqrt{n}$
	500	**Gini impurity**	**5**	**5**	$\boldsymbol{log_2(n)}$
	1000	Entropy	20	10	0.2
	1500		50	15	0.5

Model	N estimators	Learning rate	Min sample split	Max depth
Gradient Boosting	100	0.5	[80 – 100]	2
	125	0.4		3
	150	**0.3**		**4**
	175	0.2	**90**	5
	200	0.1		6
		0.01		

Model	C	Kernel	Gamma
Support Vector Machine	0.0001		0.0001
	0.001		**0.001**
	0.01		0.01
	0.1	**rbf**	0.1
	1.0	linear	1.0
	10.0		10.0
	100.0		100.0
	1000.0		1000.0

A Tabela 15 apresenta os hiperparâmetros e a faixa de valores definidos para cada modelo. Os valores em negrito destacam os melhores valores encontrados para cada modelo. Como a pesquisa da grade foi implementada com a biblioteca *scikit-learn*[1], a escolha dos intervalos de valores baseou-se nos valores padrão da biblioteca. Para cada hiperparâmetro, definimos intervalos de valores menores e maiores que o valor padrão da biblioteca.

O modelo de árvore de decisão não foi submetido à pesquisa de grade porque o usamos como modelo *baseline*. Assim, os hiperparâmetros deste modelo foram configurados manualmente com os valores padrão da biblioteca *scikit-learn*. A árvore de decisão foi escolhida como o modelo *baseline* devido à facilidade de interpretação do modelo. A Tabela 16 mostra os valores dos hiperparâmetros da árvore de decisão.

Tabela 16. Configuração dos hiperparâmetros da árvore de decisão

Model	Criterion	Splitter	Min sample split	Min sample leaf	Max features
Decision	Gini impurity	Best	2	1	16

6.2.3 Treinamento dos modelos

Uma vez definidos os hiperparâmetros dos modelos, o próximo passo consistiu na utilização da base de dados de treinamento para que os modelos pudessem aprender os parâmetros autoajustáveis. Conforme observado no terceiro passo da Figura 53, o treinamento foi realizado a partir da validação cruzada de 10-*folds*. Para analisar a performance de cada modelo, nós utilizamos a curva de aprendizado. A curva de aprendizado é uma ferramenta de diagnóstico que pode nos ajudar a diagnosticar se um algoritmo de aprendizado tem um problema de *overfitting* (alta variância) ou *underfitting* (alto viés).

A Figura 54 ilustra a curva de aprendizado dos modelos resultantes da validação cruzada de 10-*folds*. O eixo x representa o tamanho das amostras de treinamento e o eixo y a *accuracy* média alcançada por cada modelo. As linhas com pontos circulares representam a *accuracy* média alcançada no conjunto de treinamento da validação cruzada. As linhas com pontos quadrados representam a *accuracy* média alcançada no conjunto de avaliação da validação cruzada.

Em uma curva de aprendizado, o *underfitting* é percebido se as curvas de treinamento e validação estiverem abaixo de uma *accuracy* esperada. Já o *overfitting* é notado quando a curva de treinamento apresenta bons níveis de *accuracy* e as curvas de validação apresentam uma *accuracy* abaixo do esperado, o que caracteriza as curvas distantes umas das outras. Na Figura 54, nós podemos ver que nenhum dos modelos desenvolvidos apresentou problema de *underfitting*.

[1] https://scikit-learn.org/stable/

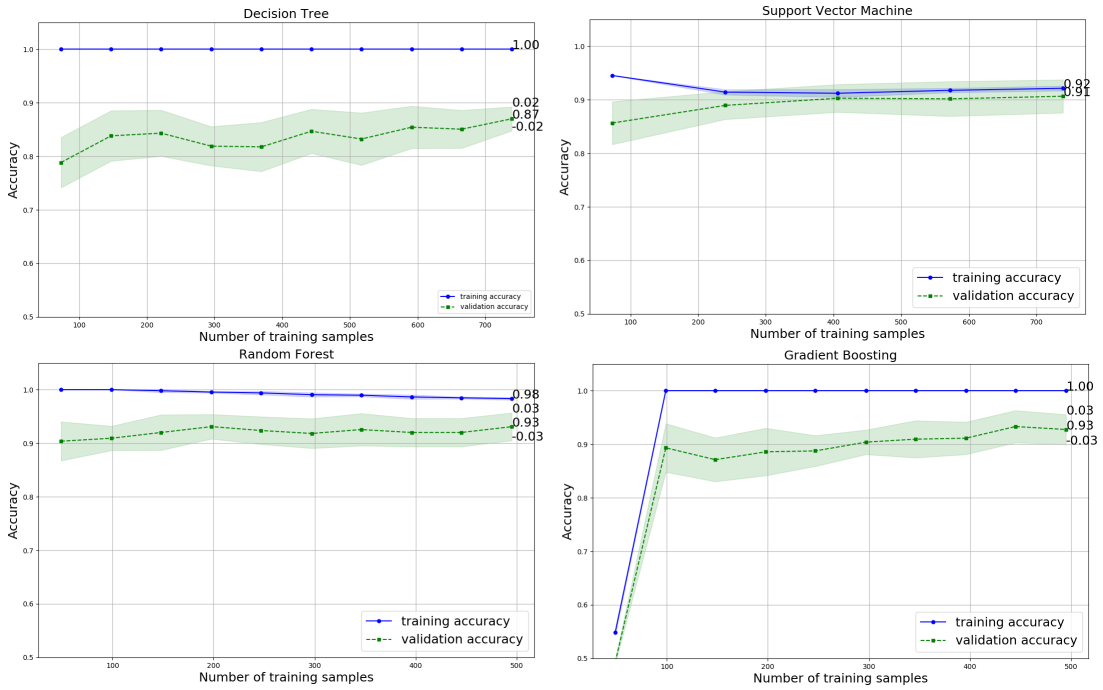


Figura 54. Curva de aprendizado dos modelos avaliados pela validação cruzada de 10-folds.

Ao analisar as curvas, podemos notar que a árvore de decisão foi o modelo que apresentou as curvas mais distantes. Esse comportamento indica um possível *overfitting*. Apesar de não apresentar a melhor *accuracy*, o modelo *Support Vector Machine* foi o que apresentou as curvas mais próximas. Esse comportamento demonstra que o modelo atingiu níveis satisfatórios de generalização do problema. Os outros dois modelos, *Random Forest* e *Gradient Boosting*, também apresentaram curvas próximas e foram os modelos com melhor *accuracy* no conjunto de validação.

6.2.4 Avaliação dos modelos

Para certificar o desempenho observado na fase de treinamento, o próximo passo compreendeu a utilização da base de teste para gerar a matriz de confusão de cada modelo. Como já mencionado no Capítulo 4, as células da matriz de confusão podem ser divididas em verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN). Nesse experimento, os falsos positivos correspondem aos passageiros que foram classificados, erroneamente, como motoristas e os falsos negativos correspondem aos motoristas que foram classificados, erroneamente, como passageiros. A minimização dos falsos positivos e falsos negativos correspondem, respectivamente, às garantias de eficiência e conveniência da solução.

Na Figura 55 pode-se observar a matriz de confusão dos quatro modelos. A partir dessas matrizes, nós calculamos as métricas *accuracy*, *precision*, *recall* e *F1-score*. No contexto desse experimento, a precisão representa a proporção de motoristas que foram

corretamente classificados, entre as amostras que foram classificadas como motoristas. Já a revocação representa a proporção de motoristas que foram classificados corretamente como motoristas, dentre as amostras que deveriam ser classificadas como motoristas.

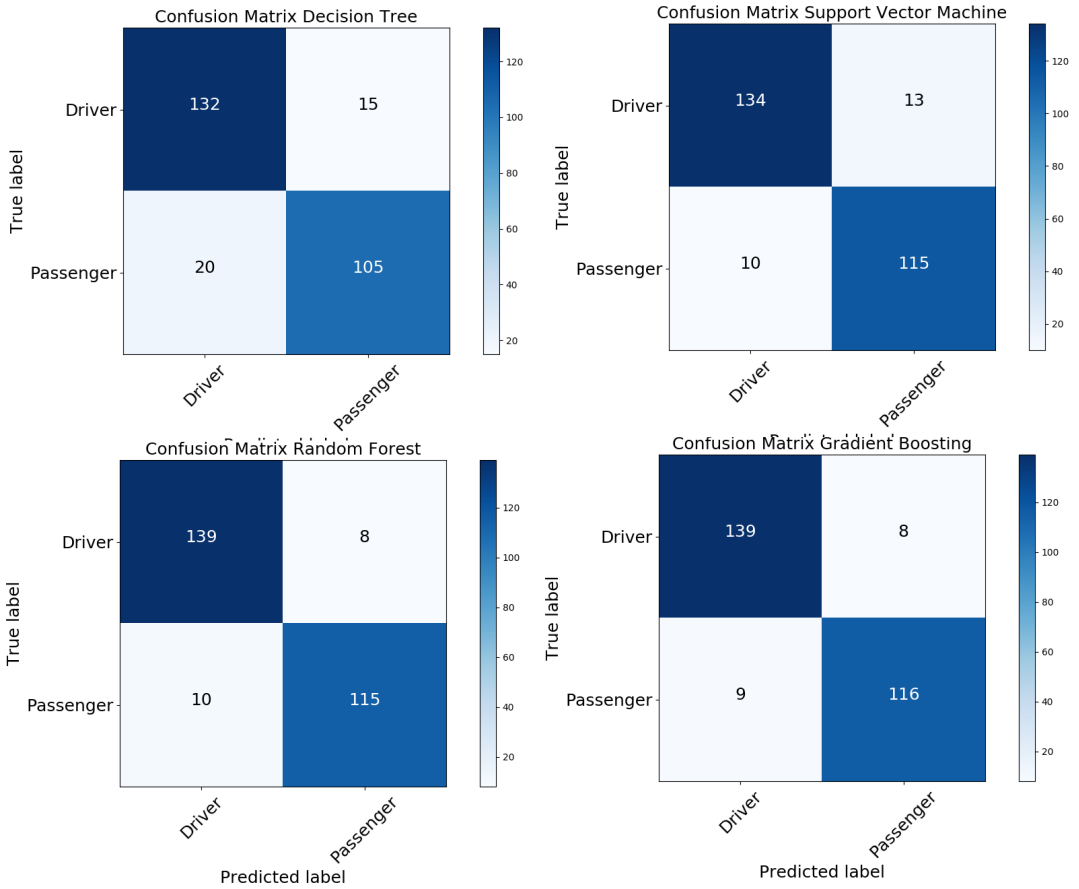


Figura 55. Matriz de confusão construída a partir dos dados de teste.

Analisando a Tabela 17, podemos ver que a precisão dos modelos são próximas. Isso significa que, entre as amostras classificadas como motoristas, a proporção de verdadeiros positivos é praticamente a mesma. Supondo um cenário em que o celular do motorista é bloqueado, automaticamente, se estiver enviando mensagens de texto enquanto dirige, nos modelos *Support Vector Machine*, *Random Forest* ou *Gradient Boosting*, teríamos aproximadamente 7% dos passageiros com o celular bloqueado indevidamente. Neste contexto, melhorar a precisão do modelo significa diminuir a inconveniência para os passageiros. No modelo de árvore de decisão, esse valor seria de, aproximadamente, 14%.

Sobre o métricas *recall*, podemos notar, na Tabela 17, que os modelos *Random Forest* e *Gradient Boosting* foram, significativamente, melhores. Esse comportamento indica que esses modelos apresentam maior confiabilidade para o motorista. A partir da *recall*, se fizermos $1 - R$, teremos a proporção de motoristas que foram incorretamente classificados como passageiros. Por exemplo, como a *recall* do modelo *Gradient Boosting* foi 0.945, temos menos de 6% dos motoristas classificados como passageiros. Em um contexto de

bloqueio do *smartphone*, menos de 6% dos motoristas não teriam o *smartphone* bloqueado durante as mensagens de texto, enquanto dirigiam.

Tabela 17. Métricas calculadas a partir da matriz de confusão.

Models	Accuracy	Precision	Recall	F1-score
Decision Tree	0.87	0.868	0.898	0.882
Support Vector Machine	0.91	0.93	0.911	0.92
Random Forest	0.93	0.932	0.945	0.938
Gradient Boosting	0.93	0.939	0.945	0.941

6.2.5 Verificação da significância estatítica

Considerando os valores das métricas calculadas, podemos ver que os modelos *Random Forest* e *Gradient Boosting* apresentaram desempenhos similares e foram superiores aos demais modelos. Para verificar se existe alguma diferença entre os modelos, aplicamos um teste de significância. O teste de significância foi construído a partir da observação da *accuracy* dos modelos. Como a *accuracy* dos modelos foi praticamente a mesma na fase de validação e teste, optamos por utilizar os valores da *accuracy* da fase de validação. Sabendo que a *accuracy* dos modelos foi coletada em dez rodadas de validação cruzada, cada modelo foi representado por um vetor de tamanho igual a 100. A Figura 56 apresenta o processo de construção desses vetores.

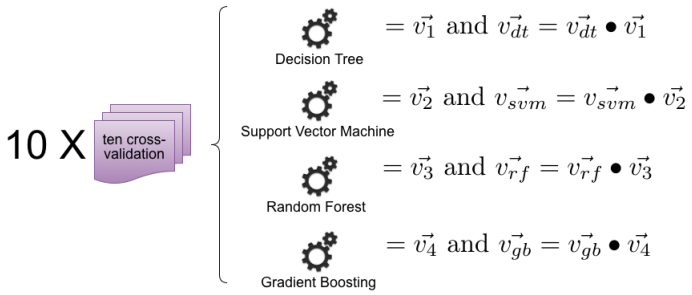


Figura 56. Construção dos vetores utilizados no teste de significância.

Para realizar o teste de significância, analisou-se a densidade dos modelos, a partir dos vetores gerados. A Figura 57 ilustra a estimativa da densidade dos modelos. Observando o comportamento das densidades, podemos ver que as funções de densidade dos modelos não são perfeitamente caracterizadas como uma distribuição gaussiana. Para ter certeza desse comportamento, nós aplicamos o teste de hipótese *normatest* da biblioteca *scipy*[2].

A função *normaltest* da biblioteca *scipy* é baseada nas pesquisas de (D'AGOSTINO, 1971) e (D'AGOSTINO; PEARSON, 1973). Essa função testa a hipótese nula de que

[2] <https://www.scipy.org/>

uma amostra é proveniente de uma distribuição normal. Após a aplicação do teste, se a hipótese nula for rejeitada, podemos certificar, com intervalo de confiança de 95%, que os modelos não apresentam uma distribuição gaussiana. O fato de as distribuições não serem gaussianas implica que não podemos testar a significância dos modelos usando, por exemplo, o teste *Student's t-test* e o teste *Welch's t-test*. Nesse caso, foi realizado o teste de significância estatística utilizando teste *Kolmogorov-Smirnov* o qual foi apresentado no Capítulo 4.

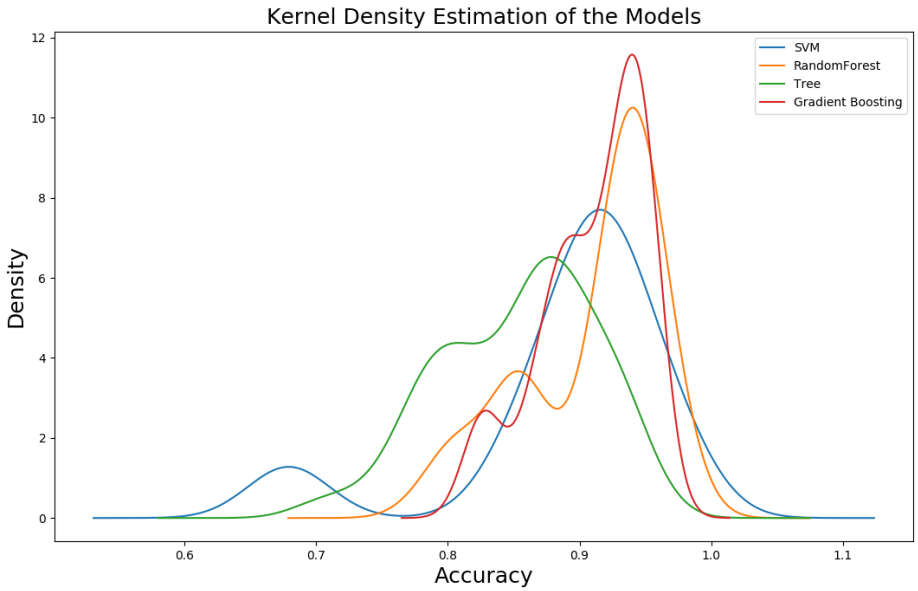


Figura 57. Estimativa de densidade dos modelos.

O teste foi aplicado para verificar se os modelos desenvolvidos são, significativamente, diferentes. A hipótese nula do teste (H0) considera que os modelos são iguais e a hipótese alternativa (H1) considera que os modelos são diferentes. Se aceitarmos a hipótese nula, significa que não há diferença significativa entre os modelos comparados. Ao aplicar o teste de *Kolmogorov-Smirnov*, se obtivermos um valor de $p\text{-}value \leq 0.05$, significa que podemos rejeitar a hipótese nula e que os modelos de aprendizado de máquina são significativamente diferentes, com 95% de confiança.

Nós aplicamos o teste de *Kolmogorov-Smirnov* para todas as combinações dois a dois dos modelos desenvolvidos. Como desenvolvemos quatro modelos de aprendizado de máquina, foram realizados seis testes de hipótese. A Tabela 18 mostra o *p-value* e a descrição se a hipótese H0 foi rejeitada ou aceita.

Como podemos ver, apenas o teste de hipótese que compara a significância dos modelos *Random Forest* e *Gradient Boosting* não foi rejeitado. O fato de os outros testes terem sido rejeitados nos diz que os outros modelos têm funções de densidade de probabilidade significativamente diferentes. Analisando o resultado do teste *Kolmogorov-*

Smirnov, a curva de aprendizado e as métricas utilizadas na avaliação dos modelos, podemos concluir que os modelos *Random Forest* e *Gradient Boosting* desenvolvidos neste trabalho são equivalentes. Ou seja, considerando os parâmetros utilizados nos dados de entrada e os hiperparâmetros dos modelos, a escolha do modelo *Random Forest* ou *Gradient Boosting* é indiferente à segurança dos motoristas e ao conforto dos passageiros, no que diz respeito ao evento de *texting while driving*.

Tabela 18. Resultado do teste Kolmogorov-Smirnov aplicado

Model 1	Model 2	p-value	H0
Decision Tree	SVM	0.000	reject
Decision Tree	Random Forest	0.000	reject
Decision Tree	Gradient Boosting	0.000	reject
SVM	Random Forest	0.000	reject
SVM	Gradient Boosting	0.031	reject
Random Forest	**Gradient Boosting**	0.069	**accept**

6.3 Considerações finais

Nesse experimento, foram desenvolvidos quatro modelos de aprendizado de máquina (Árvore de Decisão, *Support Vector Machine*, *Random Forest* e *Gradient Boosting*) para detectar a digitação de texto em um veículo em movimento. O estudo foi construído assumindo a hipótese de que o desenvolvimento de soluções não intrusivas são necessárias para evitar a distração de *texting while driving*. Entendemos que há essa necessidade uma vez que, mesmo com massivas campanhas educativas e leis proibitivas, esse tipo de distração ainda é um problema que contribui para um número cada vez maior de acidentes de trânsito.

Para realizar a otimização e a avaliação dos modelos, esse experimento realizou uma coleta de dados naturalística no município de Belém. Treze voluntários participaram da fase de coleta dos dados. Cada voluntário assumiu o papel de passageiro e motorista. Para caracterizar uma coleta naturalística, os voluntários tiveram a autonomia para escolher a rota de viagem, bem como as mensagens que digitariam. Para cada mensagem digitada, dados sensoriais foram coletados, utilizando os sensores acelerômetro, giroscópio, GPS e *touch screen* de um *smartphone*.

Os modelos de aprendizado de máquina foram desenvolvidos para classificar se a digitação da mensagem de texto é realizada por um passageiro ou por um motorista. A formulação deste problema está embasada na hipótese de que os padrões de direção e digitação de um motorista se alteram se ele estiver realizando o evento de *texting*

while driving. A classificação de motoristas e passageiros, a partir dos dados sensoriais, possibilita que a solução seja não intrusiva. Nós entendemos que uma solução é não intrusiva, se as chances do usuário burlar o seu funcionamento forem minimizadas. Por exemplo, se a solução necessitasse de uma câmera para capturar os dados de entrada, o usuário conseguiria, facilmente, burlar o funcionamento do sistema impedindo a captura das imagens.

Os resultados da avaliação dos modelos demonstram que os algoritmos *Random Forest* e *Gradient Boosting* apresentaram a melhor performance para resolver o problema proposto. Os algoritmos foram construídos e treinados utilizando a estratégia de validação cruzada de 10-*folds*. Para a avaliação da capacidade de generalização de cada modelo, nós realizamos a análise da curva de aprendizado e também calculamos as métricas *accuracy*, *precision*, *recall* e *F1-score*, a partir da matriz de confusão proveniente do conjunto de dados de teste. Além da eficiência dos modelos, nós também analisamos a significância dos modelos desenvolvidos. A significância foi medida pelo teste de *Kolmogorov-Smirnov* e objetivou verificar se os modelos desenvolvidos eram significativamente diferentes. A conclusão foi que os modelos *Random Forest* e *Gradient Boosting* não apresentam diferenças significativas.

De acordo com as tarefas planejadas para contemplar o objetivo desta tese, este experimento corresponde à realização da seguinte tarefa:

- Realização de um experimento utilizando a metodologia de análise dos efeitos causados pelo comportamento do condutor, para detectar a digitação de mensagens de texto e classificação de passageiros e motoristas.

Os resultados deste experimento foram publicados em 2019 na *International Joint Conference on Neural Networks* com o título *Exploiting Machine Learning Models to Avoid Texting While Driving*.

7 Análise da distração do motorista devido à leitura de mensagens de texto

Este capítulo tem como objetivo descrever a pesquisa e o desenvolvimento dos modelos de aprendizado de máquina construídos para classificar passageiros e motoristas durante a distração proveniente da leitura de mensagens de texto. Essa pesquisa foi conduzida a partir da metodologia de observação dos efeitos do comportamento do condutor. As pesquisas que utilizam essa metodologia partem da ideia de que os padrões de direção de um motorista distraído são diferentes dos padrões de direção de um motorista que está dirigindo com atenção. A hipótese é de que um motorista distraído tende a apresentar variações nos padrões de velocidade, aceleração, desaceleração e direção do veículo. Na revisão sistemática apresentada no Capítulo 2, os trabalhos que analisaram esse tipo de distração consideraram a escrita e leitura de mensagem de texto como um único evento. Entretanto, como demonstrado no *survey* do Capítulo 3, pode-se observar que a distração devido à leitura de mensagens de texto é um evento secundário que ocorre com maior frequência quando comparado com a distração devido à escrita de mensagens de texto. Nesse sentido, este estudo apresenta uma nova metodologia para classificar passageiros e motoristas durante a leitura de mensagens de texto.Um estudo naturalístico foi realizado para coletar dados utilizando os sensores acelerômetro, magnetômetro, giroscópio, GPS de um *smartphone*. Os resultados demonstram que os modelos desenvolvidos possuem a capacidade de generalização e conseguem distinguir a leitura de mensagens de texto realizadas por passageiros e motoristas somente pela observação do efeito do comportamento do condutor.

As principais contribuições desse capítulo são:

- Construção da base de dados naturalística com dados sensoriais de motoristas e passageiros, referentes à leitura de mensagens de texto.

- Desenvolvimento de modelos de observação do efeito do comportamento que consiga classificar passageiros e motoristas, no que diz respeito à distração proveniente da leitura de mensagens de texto.

- Análise da influência dos dados de entrada para o desempenho dos modelos construídos.

- Metodologia de avaliação da capacidade de generalização dos modelos.

O decorrer do capítulo está dividido em três principais seções. Na primeira seção é discutida a motivação do método proposto. Na segunda, apresentamos a metodologia da proposta, as avaliações realizadas e a análise dos resultados. Por fim, na última seção, são apresentadas as considerações finais desse estudo.

7.1 Motivação

Na revisão da literatura realizada no Capítulo 2, podemos observar que a composição de mensagens de texto é considerada a distração com maior risco de acidentes de trânsito, devido à quantidade de distrações envolvidas. Um motorista, ao digitar uma mensagem de texto, está distraído fisicamente - por estar digitando, visualmente - por estar lendo o que está digitando, e cognitivamente - por estar pensando no conteúdo que está sendo digitado. Por esse motivo, pesquisas que investigam soluções de *texting while driving* foram filtradas na revisão. Contudo, segundo os dados do *survey* conduzido pela (AT&T, 2015), a distração devido à leitura de mensagens de texto é mais frequente do que a distração ocasionada pela composição de mensagens de texto. Esses dados também foram constatados no *survey* apresentado no Capítulo 3. O maior número de distrações devido à leitura de mensagens de texto nos faz acreditar que os usuários acham que esse tipo de distração oferece menores riscos, por ocorrer em um curto espaço de tempo. Entretanto, assim como a digitação de mensagens de texto, a distração devido à leitura de mensagens de texto também pode ocasionar a distração física, visual, e cognitiva.

Por entender que a distração de *reading while driving* também oferece riscos de acidentes de trânsito, esse experimento propõe o desenvolvimento de soluções não intrusivas para detectar e classificar a distração devido à leitura de mensagens de texto. Assim como no experimento apresentado no Capítulo 6, a questão da não intrusão também foi levada em consideração, com o objetivo de minimizar as chances de o usuário burlar a eficiência do sistema. Para isso, o estudo foi realizado utilizando o método de observação dos efeitos do comportamento distraído. Acredita-se que o perfil de direção do motorista seja alterado se este estiver realizando o evento de *reading while driving*. O efeito do comportamento distraído foi analisado a partir de dados sensoriais. Nesse experimento, foram utilizados os sensores: acelerômetro, giroscópio, magnetômetro e GPS de um *smartphone*.

A tarefa de distinguir passageiros e motoristas no evento de *reading while driving* é uma tarefa desafiadora. A primeira razão vem da pequena janela de tempo em que o evento ocorre. Quando um motorista segura o telefone para ler uma mensagem, a distração visual, física e cognitiva deve durar alguns segundos. Podemos considerar que a tarefa de detectar mudanças no comportamento do motorista, neste curto espaço de tempo, é um problema complexo. O segundo desafio deste problema diz respeito à variabilidade do comportamento do motorista, observada a partir dos sensores. Ao se distrair lendo uma mensagem de

texto, diferentes reações podem ocorrer, desde as reações mais notáveis, como uma freada repentina, até as reações mais leves, como uma ligeira desaceleração ou mudança de direção angular. Nesse sentido, para se definir regras para todas as possibilidades de comportamento do motorista, seria necessária uma quantidade significativa de representações de dados e, mesmo assim, a possibilidade de erro ainda seria muito alta.

Considerando a não linearidade do problema e a necessidade de generalização, devido às várias possibilidades de comportamento do motorista, entendemos que a questão da distração devido à leitura de mensagens de texto pode ser resolvida por modelos de aprendizado de máquina. Neste contexto, este experimento utilizou os dados coletados por sensores de um *smartphone* para desenvolver modelos de aprendizado máquina que classifiquem passageiros e motoristas no evento de *reading while driving*. Em resumo, este trabalho avalia a influência dos sensores Ω e da capacidade de generalização dos modelos de aprendizado de máquina Ψ, na tarefa de distinguir passageiros e motoristas durante o evento de *reading while driving* Φ, ou seja, avaliamos como Φ pode ser resolvido por $\Psi(\Omega)$ e também qual a influência de $\omega \in \Omega \; \forall \; \psi \in \Psi$.

7.2 Metodologia

O conjunto de dados naturalísticos foi coletado em 2018 nas cidades de Belém e Paragominas, cidades localizadas no estado do Pará, norte do Brasil. A coleta de dados envolveu 18 voluntários que usaram um aplicativo para capturar dados sensoriais enquanto liam mensagens de texto. No total, foram coletadas 11 mil amostras, metade das amostras representaram as leituras realizadas pelos motoristas e a outra metade representaram as leituras realizadas pelos passageiros. A Figura 58 ilustra o processo de aquisição de dados e extração de características realizadas neste experimento.

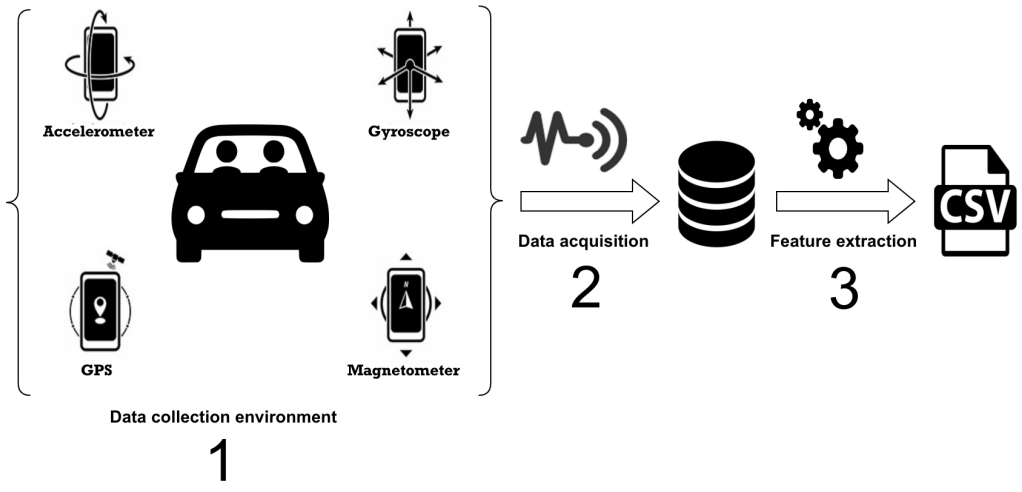


Figura 58. Metodologia utilizada na aquisição dos dados e extração das características.

Para rotular cada mensagem, nós criamos um aplicativo *Android* para exibir mensagens curtas e aleatórias para o usuário. Enquanto o usuário lê uma mensagem, os dados são coletados, usando os sensores de acelerômetro, GPS, magnetômetro e giroscópio. A Figura 59 ilustra as quatro telas do aplicativo. Na primeira tela, no canto superior esquerdo, o usuário deve pressionar o botão *START* para iniciar a coleta de dados. Em seguida, na tela superior direita, o usuário deve definir o nome do arquivo que armazenará os dados do sensor. Para ler uma mensagem gerada pelo aplicativo, o usuário deve pressionar o botão *BEGIN*; Neste ponto, os dados do sensor começam a ser coletados. No final da leitura, o usuário deve pressionar o botão *END*, que informa que a coleta dos dados do sensor pode ser interrompida e armazenada.

Para cada mensagem lida, o usuário deve repetir o procedimento de pressionar os botões *BEGIN*/*END* para armazenar os dados dos sensores. Cada mensagem que o usuário lê representa uma instância dos dados coletados. Um arquivo tem várias instâncias de dados. A rotulagem de dados foi determinada por arquivo, ou seja, quando o usuário cria um arquivo, aceita-se que todas as mensagens lidas são realizadas por apenas um perfil de usuário que pode ser de motorista ou de passageiro. Um usuário pode criar quantos arquivos ele precisar.

Na fase de coleta de dados, cada participante criou pelo menos dois arquivos por categoria. No total, 58 arquivos foram criados, 36 arquivos contendo instâncias de dados coletados por motoristas e 22 arquivos contendo instâncias de dados coletadas por passageiros. Os participantes tiveram autonomia para decidir o horário adequado para ler uma mensagem de texto. Os únicos pré-requisitos estabelecidos foram: (1) Ler a mensagem de texto apenas quando o carro estiver em movimento, e (2) Ler a mensagem de texto, segurando o *smartphone* com uma ou ambas as mãos.

O primeiro pré-requisito visa descartar as mensagens de texto lidas quando o carro estiver parado. Além disso, como a classificação de passageiros e motoristas foi baseada nos dados sensoriais do veículo, nenhuma diferenciação seria observada neste contexto. O segundo pré-requisito foi estabelecido para eliminar a interação do usuário quando o smartphone estava acoplado ao veículo. Entendemos que um *smartphone* fixo tem características de distração semelhantes às de um computador de bordo e, portanto, as coleções de dados nessa configuração não foram levadas em consideração nesse experimento.

A coleta de dados teve início em junho de 2018 e terminou em dezembro de 2018. Durante esse período, os 18 participantes fizeram 58 viagens para coletar dados. Todos os participantes coletaram dados assumindo o papel de motorista e passageiro. Na Tabela 19, nós podemos observar o número de mensagens lidas por cada voluntário. Cada linha representa a coleta realizada por um voluntário. A coluna de passageiros informa o número de mensagens lidas por um voluntário com o perfil do passageiro e a coluna do motorista informa o número de mensagens lidas por um voluntário com o perfil do motorista. A

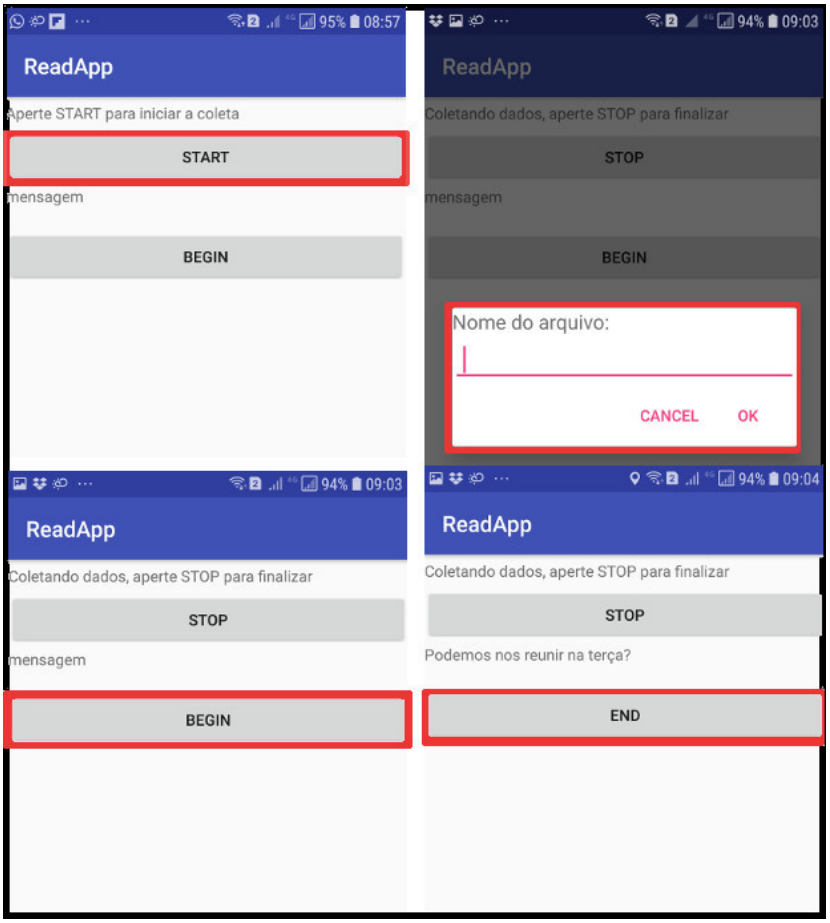


Figura 59. Aplicativo construído para aquisição dos dados: Na primeira tela (**canto superior direito**), o usuário deve digitar o nome do arquivo de *log*. Os sensores são ativados quando o usuário pressiona o botão START. Para cada mensagem lida, o usuário deve pressionar os botões BEGIN/END para definir a janela de dados do sensor.

subdivisão da coluna de passageiros e motoristas informa o número de viagens realizadas por um voluntário. Por exemplo, o voluntário 1 realizou uma viagem com o perfil do passageiro e três viagens com o perfil do motorista. Na viagem como passageiro, o voluntário 1 leu 103 mensagens. Em três viagens como motorista, o voluntário 1 leu, respectivamente, 26, 38 e 59 mensagens.

O número de mensagens coletadas em uma viagem foi diferente entre os participantes. A distância percorrida e as condições de tráfego foram dois fatores importantes que influenciaram o número de dados coletados em cada viagem. A quantidade de viagens também foi diferente entre os participantes. Por ser uma coleta de dados naturalística, um participante tinha autonomia para decidir quantas viagens gostaria de fazer. Examinando a Tabela 19, podemos observar que a maioria dos participantes optou por realizar um maior número de viagens como motorista.

Na Figura 60, nós podemos ver seis exemplos de viagens feitas pelos participantes.

Tabela 19. Número de mensagens lidas por cada voluntário. A subdivisão da coluna do condutor ou passageiro indica que o voluntário realizou mais de uma viagem com esse perfil.

Participant	**Passenger**	**Driver**
1	156	179
2	124	230
3	350	345
4	354	128
5	138	129
6	219	214
7	355	355
8	301	165 127
9	141	66 159
10	309	164 303
11	103	26 38 59
12	124	96 155 23
13	1975	452 354 380
14	152	30 56 81 50
15	87 119	55 27 98 20
16	35 136	31 170 144
17	165 55	283 126
18	81 21	182

A Figura 60a–c são viagens executadas pelos motoristas e a Figura 60d–f são viagens executadas pelos passageiros. Por se tratar de um estudo naturalístico, os participantes puderam escolher a rota da viagem. Na rota escolhida, com o carro em movimento, sempre que possível, o participante lia uma mensagem para coletar dados dos sensores. Como podemos ver na Figura 60, os dados foram coletados em diferentes rotas com diferentes direções. Essa característica é importante para não enviesar a coleta de alguns sensores.

Outro ponto importante que também deve ser destacado é como os participantes devem ler a mensagem. O objetivo era ter o menor controle possível sobre a forma como os dados seriam coletados. Por esse motivo, os participantes foram instruídos a manter os *smartphones* como de costume. Alguns participantes seguravam o smartphone com a mão direita, outros com a mão esquerda, outros com as duas mãos. Alguns seguravam o smartphone na horizontal e outros na vertical. Aceitar essas diferenças é importante para testar a capacidade de generalização dos modelos de aprendizado de máquina construídos

neste experimento.

(a) (b)

(c) (d)

(e) (f)

Figura 60. Viagens feitas por voluntários para coletar dados sensoriais. Figuras (**a**-**c**) foram viagens realizadas por motoristas e as Figuras (**d**-**f**) foram viagens realizadas por passageiros.

Como mencionado anteriormente, a coleta de dados foi realizada utilizando os sensores acelerômetro, giroscópio, magnetômetro e GPS de um *smartphone*. A escolha destes sensores foi motivada pela sua utilização em trabalhos relacionados. Como observado por (ALLUHAIBI; AL-DIN; MOYAID, 2018), vários pesquisadores estão investigando o comportamento do motorista a partir de dados coletados por *smartphones*. A pesquisa realizada por (JUNIOR et al., 2017) é um exemplo de trabalho relacionado que usou o mesmo conjunto de sensores selecionados para a coleta de dados do nosso trabalho. Em sua pesquisa, (JUNIOR et al., 2017) analisou a influência do conjunto de sensores na tarefa de classificar o perfil do comportamento do motorista.

Como visto na revisão de literatura apresentada do Capítulo 2, há estudos que também analisam a distração usando dados coletados por outros sensores. Por exemplo, os sensores *eye tracking*, *wearable sensor network*, *galvanic skin response*, *head tracking system*, *electrocardiogram (ECG)* e *surface electromyogram (sEMG)* são usados em trabalhos correlatos (SAHAYADHAS et al., 2015b; WOLLMER et al., 2011; RAJENDRA; DEHZANGI, 2017; REIMER; MEHLER; DONMEZ, 2014). Usando sensores mais específicos, esses estudos visam analisar a distração do motorista de diferentes perspectivas. Análise cognitiva, movimentos oculares, movimentos da cabeça, reações corporais e batimentos cardíacos são levados em consideração para inferir os motivos e os tipos de distrações do motorista. Pesquisas que buscam esse nível de detalhe, geralmente, realizam experimentos em simuladores. Apesar de apresentarem resultados importantes, os estudos que utilizam *hardwares* específicos têm sua abrangência limitada. Por esse motivo, para fornecer uma solução abrangente, nós optamos por usar apenas os sensores dos telefones celulares.

Nesse experimento, com exceção do sensor GPS, os sensores foram configurados para coletar dados a uma frequência de 5Hz. Para esses sensores, todos os valores correspondentes aos eixos x, y e z foram considerados durante a coleta de dados. No caso do sensor GPS, a frequência foi de 1Hz e os dados coletados correspondem às coordenadas geográficas de latitude e longitude. Todos os eixos dos sensores foram coletados, uma vez que um dos objetivos deste trabalho é verificar a influência desses sensores como insumo para os modelos de aprendizado de máquina. A Tabela 20 lista os 11 tipos de dados coletados. Devido à configuração da frequência dos sensores, para cada mensagem de texto lida, a janela de dados tem tamanhos diferentes. O tamanho da janela de dados está diretamente relacionado ao tempo que o participante leva para ler uma mensagem. Esse tempo pode estar associado ao tamanho da mensagem e aos motivos de distração secundários.

Tabela 20. Lista de sensores utilizados durante a coleta de dados.

accelerometer-x	accelerometer-y	accelerometer-z
gyroscope-x	gyroscope-y	gyroscope-z
magnetometer-x	magnetometer-y	magnetometer-z
GPS-lat	GPS-long	

A Figura 61 mostra que o tamanho médio da janela dos motoristas é maior que as janelas dos passageiros. Esse comportamento pode ser explicado devido ao foco da tarefa executada. Quando um motorista está dirigindo e decide ler uma mensagem de texto, a leitura se torna uma tarefa secundária. Sua atenção visual, física e cognitiva será compartilhada entre a tarefa de dirigir e a de ler a mensagem de texto. Por outro lado, quando um passageiro decide ler uma mensagem de texto, esta é provavelmente sua tarefa principal. Sua concentração é total e, por esse motivo, tende a realizar uma leitura mais rápida do que um motorista. Considerando a análise da janela de dados, foi realizada a

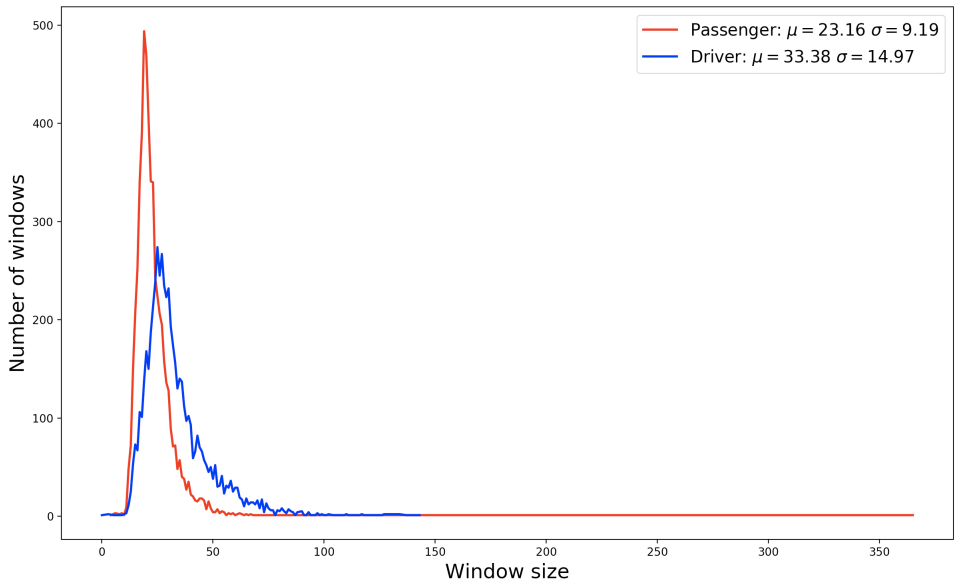


Figura 61. Distribuição do tamanho das janelas coletadas pelos usuários. Cada janela de dados representa o intervalo que o usuário levou para pressionar os botões de BEGIN/END do aplicativo.

extração de características da média e desvio padrão de cada eixo dos sensores utilizados. No caso das coordenadas de latitude e longitude, nós as usamos para calcular a velocidade em m/s de cada janela de dados. A Tabela 21 mostra os 19 atributos computados dos quatro sensores utilizados nesse experimento.

Considerando que queremos avaliar como Φ pode ser resolvido por $\Psi(\Omega)$ e também qual a influência de $\omega \in \Omega \; \forall \; \psi \in \Psi$. Nesta formulação, o conjunto Ω consiste nos atributos da Tabela21.

Tabela 21. Lista de atributos extraídos do conjunto de dados.

μ-acc-x	σ-acc-x	μ-acc-y	σ-acc-y
μ-acc-z	σ-acc-z	μ-gyro-x	σ-gyro-x
μ-gyro-y	σ-gyro-y	μ-gyro-z	σ-gyro-z
μ-mag-x	σ-mag-x	μ-mag-y	σ-mag-y
μ-mag-z	σ-mag-z	*speed*	

A Figura 62 mostra um gráfico de violino[1] da distribuição dos dados de cada atributo. Em cada parcela, o lado esquerdo representa a distribuição dos dados coletados pelos motoristas e o lado direito a distribuição dos dados coletados pelos passageiros. Para simplificar a visualização, aplicamos a normalização MinMax[2] a todos os atributos. Desta forma, todos os atributos foram redimensionados para o intervalo de valores entre 0 e 1. As

[1] Um gráfico de violino é uma versão híbrida de um gráfico *box plot* e de um gráfico de densidade

[2] $x' = \frac{x - x_{min}}{x_{max} - x_{min}}$

plotagens da Figura 62a correspondem aos atributos coletados pelos sensores acelerômetro e GPS. Para esses atributos, podemos observar que a densidade dos dados dos motoristas e passageiros é quase simétrica, além disso, os interquartis também são próximos. A única exceção são os dados do atributo *mean-accX*. Neste atributo, a densidade de dados do motorista é mais dispersa quando comparada aos dados do passageiro.

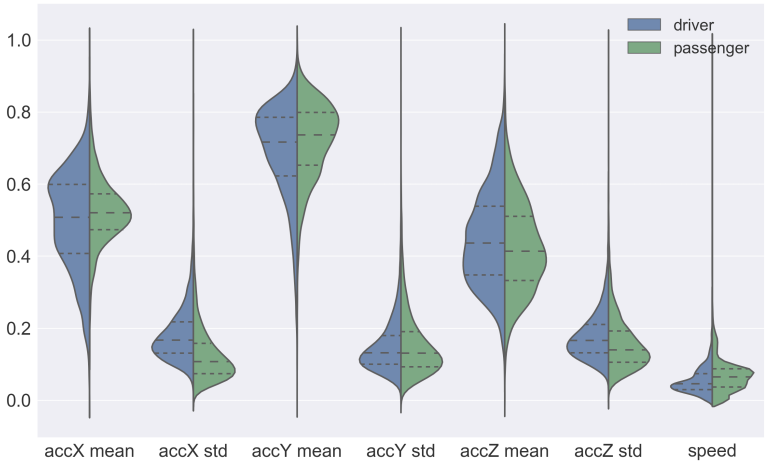
Na Figura 62b podemos ver os gráficos dos atributos coletados pelo sensor giroscópio. Nestas parcelas, os atributos computados pela média têm densidades e interquartis muito próximas. Esse comportamento sugere que os dados coletados por passageiros e motoristas são semelhantes. Para os atributos calculados pelo desvio padrão, podemos ver que a densidade de dados do motorista é um pouco mais dispersa do que a densidade de dados do passageiro. Analisando as parcelas dos atributos coletados pelo sensor magnetômetro apresentado na Figura 62c, podemos ver que estes são os dados que têm maior diferença em densidade e interquartis. Os atributos calculados pela média têm densidade dispersa e assimétrica, assim como diferentes interquartis entre os dados dos passageiros e dos condutores.

Analisando a distribuição dos dados, podemos inferir que os atributos correspondentes ao magnetômetro são os melhores para tentar distinguir passageiros e motoristas. O magnetômetro mede a força do campo magnético aplicado ao dispositivo em micro-Tesla (μT) e funciona como um ímã.

Apesar da assimetria na distribuição dos dados do magnetômetro, podemos ver que a diferença nos dados dos motoristas e passageiros não é muito significativa. Esse comportamento sugere que a classificação de motoristas e passageiros, usando apenas dados de sensores, não é uma tarefa trivial. A análise estatística ou algoritmos de limiar provavelmente teriam dificuldade nesta tarefa. Por essa razão, neste experimento, pretendemos avaliar o desempenho de diferentes modelos de aprendizado de máquina, na tarefa de distinguir passageiros e motoristas durante a leitura de mensagens de texto.

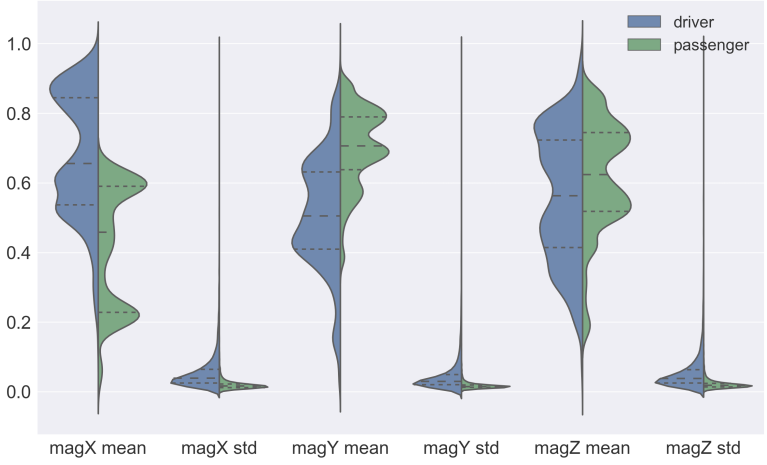
7.2.1 Configuração dos parâmetros

A Figura 63 descreve a metodologia utilizada para o desenvolvimento e análise dos modelos desse experimento. Como pode ser observado, a metodologia é dividida em cinco etapas: 1- divisão da base de dados, 2- configuração dos parâmetros, 3- treinamento, 4- avaliação e 5- verificação da significância estatítica. Os passos 1 e 2 foram realizados utilizando a mesma metodologia do experimento apresentado no Capítulo 6. Os dados foram divididos em conjuntos disjuntos de treinamento e teste, e os hiperparâmetros de cada modelo foram definidos a partir da técnica de *grid search*. A Tabela 22 apresenta os parâmetros de configuração e o intervalo de valores que foram predeterminamos para cada modelo. A pesquisa em grade foi realizada utilizando 1000 amostras estratificadas do conjunto de treinamento. Os valores em negrito representam a melhor configuração dos



(a) Gráfico violino da distribuição dos atributos construídos a partir dos sensores acelerômetro e GPS.

(b) Gráfico violino da distribuição dos atributos construídos a partir do sensor giroscópio.



(c) Gráfico violino da distribuição dos atributos construídos a partir do sensor magnetômetro.

Figura 62. Gráfico violino da distribuição dos atributos construídos a partir dos dados sensoriais.

modelos após o processamento da pesquisa em grade. Nesse experimento, nós consideramos a regressão logística o nosso modelo de referência, por esse motivo nós não o submetemos ao processo de pesquisa em grade. Neste caso, para a regressão logística, nós usamos as configurações padrão da biblioteca *scikit-learn*[3].

Nesse experimento foram construídos sete modelos de aprendizado de máquina para compor o conjunto Ψ. Levamos em conta distintas estratégias de aprendizagem supervisionadas. Nós escolhemos os modelos de aprendizagem levando em consideração a revisão da literatura apresentada no Capítulo 2 e também pela eficiência desses algoritmos em trabalhos relacionados. Nesse experimento, foram desenvolvidos os modelos *Support Vector Machines*, Árvore de Decisão, Regressão Logística, *Random Forest*, *AdaBoost*, *Gradient Boosting* e *Convolutional Neural Network*.

Tabela 22. Configuração dos parâmetros do *grid search* do terceiro experimento.

Model	**C**	**Kernel**	**Gamma**
Support Vector Machine	prange (0.001, 1000.0, 10) **100.0**	**rbf** linear	prange (0.001, 1000.0, 10) **0.001**

Model	**Criterion**	**Max depth**	**Max features**	**Min samples split**
Decision Tree	Gini impurity **Entropy**	range (2, 40, 2) **10**	$\sqrt{n}$ $log_2(n)$	range (2, 40, 2) **16**

Model	**N estimators**	**Criterion**	**Min sample split**	**Max depth**	**Max features**
Random Forest	range (75, 200, 25) **175**	**Gini impurity** Entropy	range (5, 100, 5) **80**	range (2, 10, 2) **6**	$\sqrt{n}$ $log_2(n)$

Model	**N estimators**	**Learning rate**
Adaptive Boosting	range (75, 200, 25) **150**	range (0.1, 0.5, 0.1) **0.1**

Model	**N estimators**	**Learning rate**	**Min sample split**	**Max depth**
Gradient Boosting	range (75, 200, 25) **100**	range (0.1, 0.5, 0.1) **0.4**	range (5, 100, 5) **80**	range (2, 10, 2) **4**

No modelo SVM, definimos os parâmetros *C*, *Kernel* e *Gamma*. O parâmetro *C* informa o quanto você deseja evitar classificar erroneamente cada exemplo de treinamento. Um grande valor do parâmetro *C* faz com que o modelo use um hiperplano de menor margem, se esse hiperplano realizar boas classificações dos dados de treinamento. Um pequeno valor do parâmetro *C* faz com que o modelo procure por maiores margens de separação do hiperplano tornando a superfície de decisão suave. O parâmetro *Kernel* determina como os hiperplanos serão construídos. Se uma função linear for usada, o

[3] <https://scikit-learn.org/>

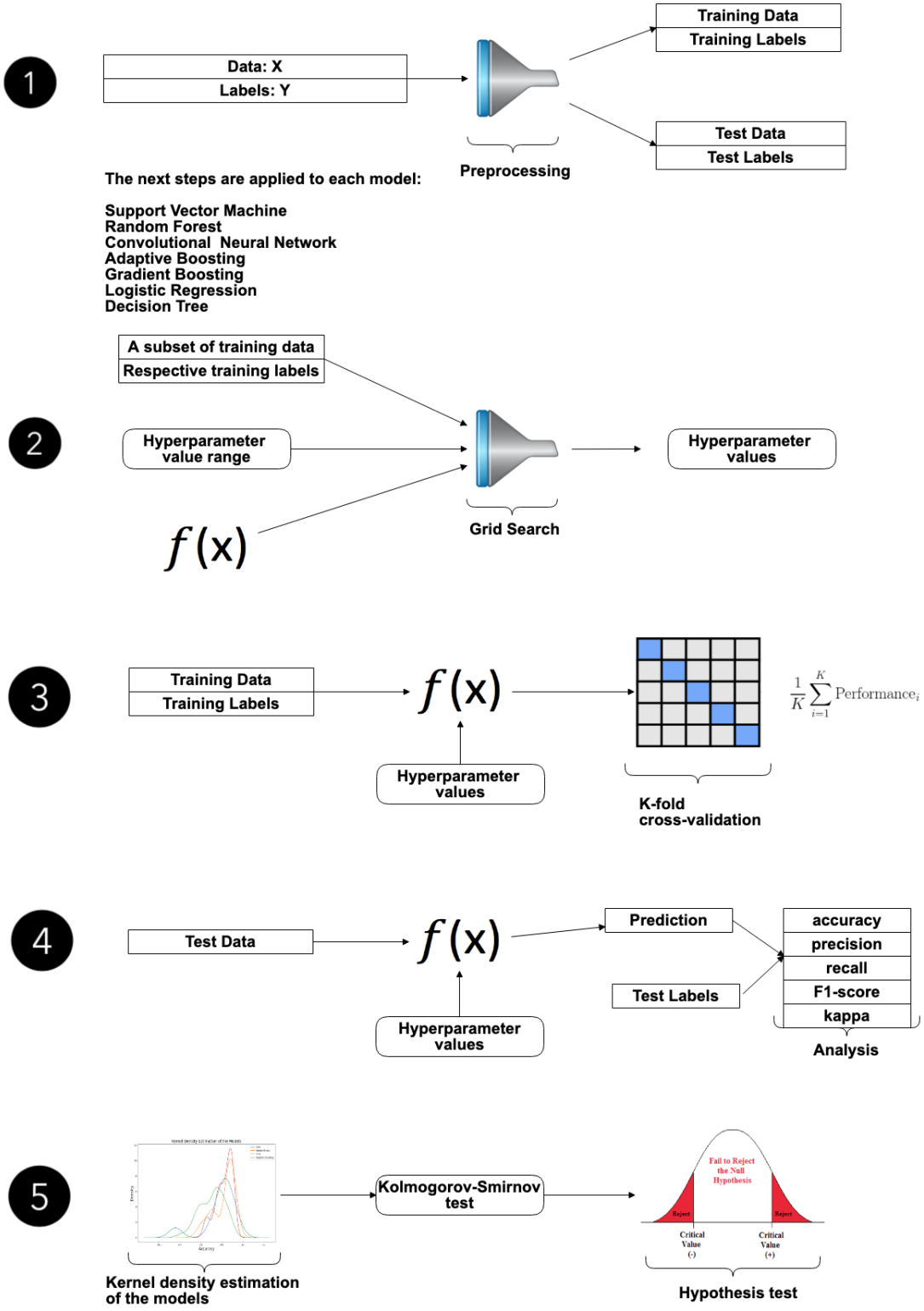


Figura 63. Fluxo do terceiro experimento.

.

modelo tratará os dados de entrada como linearmente separáveis. Para problemas não lineares, o RBF Gaussiano é amplamente utilizado quando não há conhecimento prévio sobre os dados. Com relação ao parâmetro *Gamma*, ele define até a influência de um único

exemplo de treinamento. Se o valor de *Gamma* for pequeno, a influência de um único exemplo de treinamento na construção dos hiperplanos é menor. Na pesquisa de grade realizada, podemos observar que o modelo SVM utiliza um hiperplano de menor margem; considerou que os dados de entrada não são linearmente separáveis, e que a influência de uma única amostra de treinamento não deve ser significativa para a construção dos hiperplanos.

Nos modelos Árvore de Decisão e *Random Forest*, os mesmos parâmetros foram configurados. A única exceção foi o parâmetro *N estimators*, que é uma particularidade do modelo *Random Forest* e determina o número de árvores na floresta. No parâmetro *Criterion*, é definida a função para medir a qualidade das divisões que são realizadas durante a construção das árvores. Os parâmetros *Max depth*, *Max features* e *Min sample split* são usados para determinar a configuração da árvore. O parâmetro *Max depth* determina a profundidade máxima da árvore. Se nenhum valor for determinado para esse parâmetro, então a profundidade da árvore é explorada até que todas as folhas sejam puras ou até que todas as folhas contenham menos amostras do que o valor de configuração do parâmetro *Min sample split*. O parâmetro *Max features* determina o número máximo de amostras que devem ser consideradas pelo algoritmo de ramificação da árvore. O parâmetro *Min sample split* determina o número mínimo de amostras que devem ser consideradas para que um nó seja uma folha. Se um nó possui menos amostras do que o valor de *Min sample split*, então esse nó é uma folha.

Na pesquisa de grade realizada, podemos ver que a Árvore de Decisão e o *Random Forest* foram configurados para construir árvores rasas. Em relação ao número mínimo de amostras necessárias para dividir um nó interno, o *Random Forest* usou um grande número de amostras para executar essa tarefa. No modelo *Random Forest* que construímos, 80 amostras foram necessárias para dividir um nó. No parâmetro Max depth, que diz respeito ao número de recursos a serem considerados ao procurar a melhor divisão, os dois modelos foram configurados com $\sqrt{n}$ em que n é o número da amostra de treinamento.

Nos modelos *Adaptive Boosting* e *Gradient Boosting*, nós realizamos a configuração dos parâmetros *N estimators* e *Learning rate*. O parâmetro *N estimators* define o número de estágios de reforços a serem executados. Os modelos de *Boosting* são robustos para evitar o *overfitting* quando o valor de *N estimators* é grande. O parâmetro *Learning rate* tem como objetivo definir a contribuição de cada classificador. Quanto mais próximo de zero for a taxa de aprendizado, os classificadores, individualmente, terão menor contribuição para o conjunto, ou seja, para uma pequena taxa de aprendizado, o conjunto levará em conta o desempenho conjunto dos classificadores. Para o modelo de *Gradient Boosting*, os parâmetros *Min sample split* e *Max depth* também foram configurados. Esses parâmetros têm o mesmo propósito da configuração realizada nos modelos *Random Forest* e Árvore de Decisão.

No caso da CNN, nós realizamos a configuração de forma híbrida. O número de camadas convolucionais e o número de filtros em cada camada foram definidos empiricamente. O tamanho de cada filtro foi especificado usando o processo de pesquisa em grade. Na pesquisa em grade, para encontrar a melhor configuração dos filtros convolucionais, nós determinamos que os filtros deveriam ter tamanho mínimo e máximo entre 1 e 5. Para cada camada, todos os filtros têm o mesmo tamanho. Uma vez que nós definimos 4 camadas convolucionais, 625 combinações foram testadas para encontrar a melhor arquitetura. A Figura 64 ilustra a arquitetura da CNN construída.

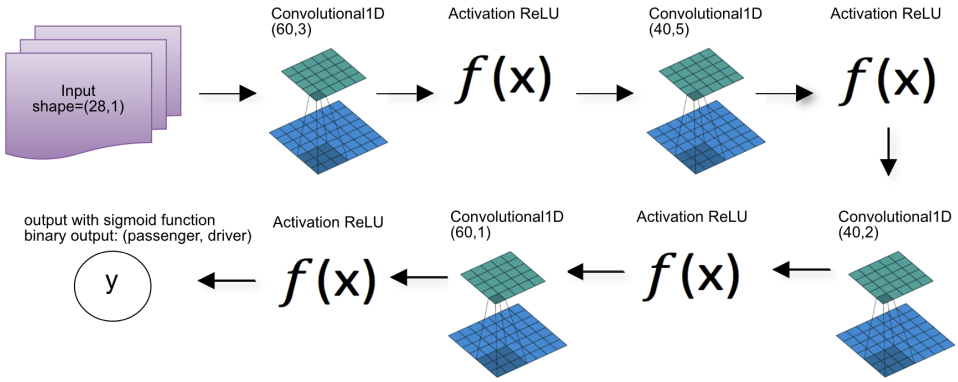


Figura 64. Arquitetura da CNN desenvolvida no terceiro experimento.

7.2.2 Treinamento dos modelos

Para analisar como os sensores $\omega \in \Omega$ podem influenciar o desempenho dos modelos $\psi \in \Psi$, nós dividimos os experimentos em três fases. Na primeira fase, usamos todos os atributos do conjunto Ω para treinar, avaliar e testar o desempenho dos sete modelos de aprendizado de máquina que pertencem ao conjunto Ψ. Na segunda fase, avaliamos o desempenho dos modelos $\psi \in \Psi$ usando apenas os atributos de cada sensor. O objetivo foi verificar se algum modelo apresenta bom desempenho de classificação utilizando as características de um único sensor. Na última fase do experimento, avaliamos o desempenho dos modelos $\psi \in \Psi$ a partir das características mais significativas destacadas pela análise de importância dos atributos do modelo *Gradient Boosting*. Nesta terceira fase, o objetivo foi verificar se é possível usar um subconjunto do Ω para obter um modelo $\psi \in \Psi$ que seja estatisticamente equivalente ao melhor modelo obtido no experimento da primeira fase. Se isso for possível, significa que podemos reduzir a quantidade de informações necessárias para distinguir os passageiros e os motoristas no evento de *reading while driving*.

Em todas as fases experimentais, reservamos dois terços do conjunto de dados para realizar o treinamento e a validação do modelo, usando uma metodologia de validação cruzada de 10-*folds*. Além disso, para garantir a estabilidade dos modelos, em cada fase, repetimos dez vezes o processo de treinamento e validação. Para avaliar o desempenho dos modelos, utilizamos as métricas *accuracy*, *precision*, *recall*, *F1-score* e *Cohen kappa*.

As métricas *accuracy*, *precision*, *recall*, *F1-score* já foram utilizadas nos experimentos anteriores e possuem a mesma interpretação para esse experimento. A *accuracy* verifica a proporção de amostras classificadas corretamente. Quando a *accuracy* atinge seu valor máximo (= 1), significa que o modelo avaliado não produziu falsos positivos e falsos negativos. A *accuracy* avalia o desempenho global do classificador. Quando temos classes desequilibradas ou um problema multiclasse, as taxas boas da *accuracy* podem não refletir o desempenho real do modelo. Para uma melhor avaliação do modelo, há também a necessidade de métricas como *precision* e *recall* para avaliar as especificidades de cada classe. Em um problema de classificação, boas taxas de precisão e revocação indicam, respectivamente, a corretude e completude do classificador.

Nesse experimento, nós também utilizamos a métrica *Cohen kappa*, que é uma estatística utilizada para testes de confiabilidade entre modelos. Conforme definido no Capítulo 4, os valores dessa métrica podem variar de -1 a $+1$, onde valores ≤ 0 indicam que a concordância esperada é aleatória, e 1 representa a concordância perfeita entre os avaliadores. Na perspectiva da aprendizagem de máquina, essa métrica pode ser utilizada para medir a proximidade das instâncias rotuladas pelos classificadores, em relação aos dados rotulados na base de dados.

As Figuras 65 e 66 mostram o desempenho dos modelos na primeira fase do experimento. O *boxplot* foi construído a partir das 100 performances de validação obtidas com as 10 repetições da validação cruzada de *10-folds*. Podemos observar que os modelos *Gradient Boosting* e CNN obtiveram o melhor desempenho em todas as métricas analisadas. O modelo Árvore de Decisão apresentou um comportamento similar ao nosso modelo base (*Logistic Regression*). Além disso, para quase todas as métricas, a Árvore de Decisão apresentou a cauda mais significativa. Esse comportamento nos diz que esse modelo foi o que apresentou maior instabilidade na classificação de motoristas e passageiros. Os conjuntos de modelos *Random Forest*, *AdaBoost* e SVM tiveram um bom desempenho, mas não foram tão eficazes em comparação aos modelos *Gradient Boosting* e CNN. No caso específico do modelo SVM, podemos observar que ele apresentou pouca disparidade na métrica *recall* e uma diferença considerável na métrica de precisão. Esse comportamento nos mostra que o modelo SVM é melhor para garantir a segurança dos motoristas do que para evitar a inconveniência dos passageiros.

A Figura 66 mostra o gráfico radar da fase de teste (esquerda) e da fase de validação (direita). Como podemos ver, em todas as métricas os modelos mantiveram seu desempenho na fase de teste e validação. Esse comportamento demonstra que nenhum modelo se especializou apenas no conjunto de treinamento, e cada um com seu nível de desempenho foi capaz de realizar a generalização do problema. Sobre a capacidade de generalização, podemos classificar os modelos em três níveis. No primeiro nível estão os modelos *Gradient Boosting* e CNN como os modelos com maior capacidade de generalização.

No segundo nível estão os modelos *AdaBoost*, *Random Forest* e *SVM* como modelos com capacidade de generalização intermediária. Finalmente, no terceiro nível, temos os modelos Árvore de Decisão e Regressão Logística como os modelos com a menor capacidade de generalização do problema.

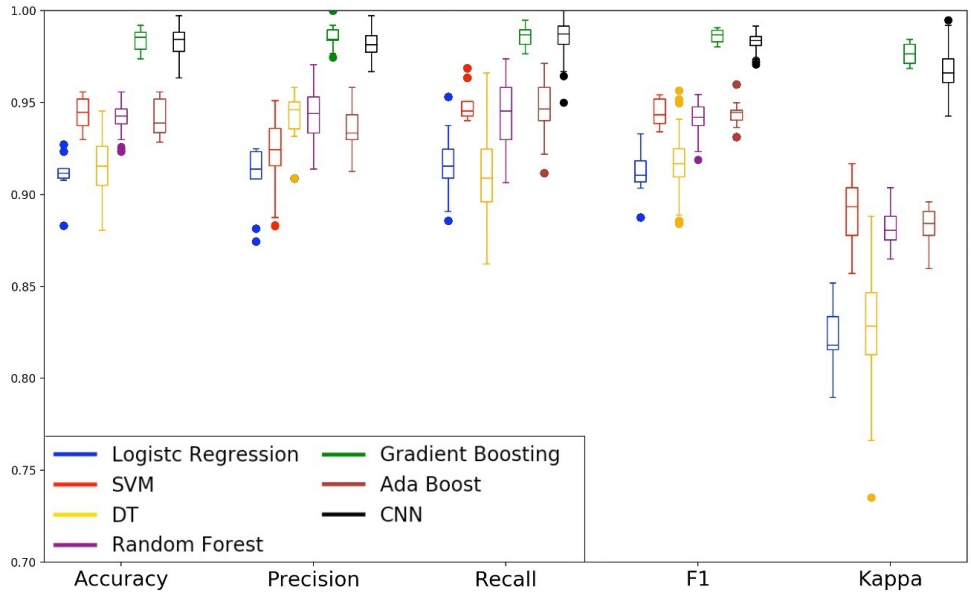


Figura 65. *Boxplot* da primeira fase do experimento.

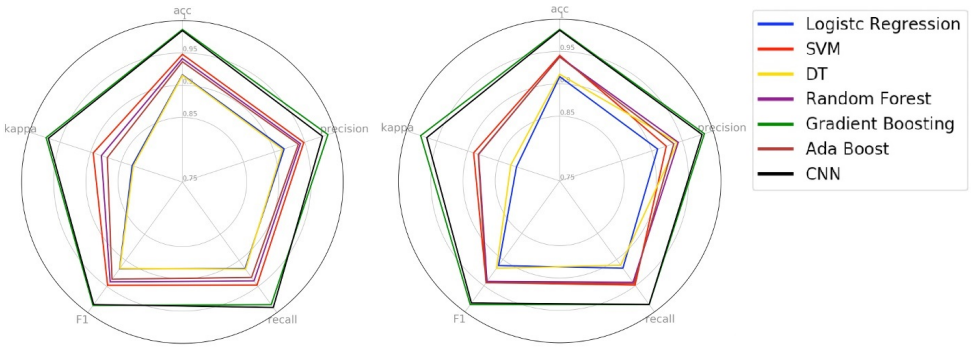


Figura 66. Gráfico radar da fase de teste (**esquerda**) e da fase de validação (**direita**).

7.2.3 Avaliação dos modelos

A próxima fase do experimento consistiu na análise de desempenho dos modelos, levando em consideração apenas os parâmetros coletados no mesmo sensor. Nós dividimos o conjunto de sensores Ω em três subconjuntos. O primeiro subconjunto contém os parâmetros coletados pelo sensor acelerômetro. O segundo subconjunto inclui os parâmetros obtidos a partir do sensor giroscópio. O terceiro subconjunto contém os parâmetros coletados pelo sensor magnetômetro. A Tabela 23 mostra os parâmetros pertencentes a cada subconjunto. Neste estágio do experimento, não utilizamos o parâmetro de velocidade calculado a partir dos dados coletados pelo sensor de GPS, por se tratar da única informação proveniente

desse sensor. Neste caso, analisamos que somente a velocidade não é informação suficiente para distinguir os passageiros dos motoristas.

Tabela 23. Subconjuntos de Ω usados como dados de entrada na segunda fase do experimento.

Subset 1: Accelerometer sensor					
μ-acc-x	σ-acc-x	μ-acc-y	σ-acc-y	μ-acc-z	σ-acc-z
Subset 2: Gyroscope sensor					
μ-gyro-x	σ-gyro-x	μ-gyro-y	σ-gyro-y	μ-gyro-z	σ-gyro-z
Subset 3: Magnetometer sensor					
μ-mag-x	σ-mag-x	μ-mag-y	σ-mag-y	μ-mag-z	σ-mag-z

As Figuras 67-69 mostram o desempenho dos modelos para cada subconjunto de parâmetros especificados na Tabela 23. Para cada subconjunto, plotamos o *boxplot* dos dados de validação. Nesta fase, o treinamento dos modelos também foi realizado executando dez vezes o método de validação cruzada de *10-folds*.

Analisando os *boxplots* das Figuras 67 e 68, podemos ver que o desempenho dos modelos foi menor quando comparado ao experimento que utilizou todos os atributos do conjunto Ω. No contexto do comportamento do motorista, a combinação dos sensores acelerômetro e giroscópio nos permite identificar mudanças nas acelerações, curvas e padrões de frenagem. No entanto, ao analisar os dados desses sensores separadamente, podemos observar que o uso independente do acelerômetro e do giroscópio não é suficiente para distinguir os passageiros e os motoristas no caso de leitura durante a condução.

(MCHUGH, 2012) sugere que os valores de kappa inferiores a 0.81 estão fora do intervalo de concordância desta métrica. Assim, podemos observar que na análise dos subconjuntos 1 e 2, todos os modelos estão com valores de kappa abaixo do esperado.

A Figura 69 mostra o desempenho dos modelos para os dados do sensor do magnetômetro. Podemos ver no *boxplot* que o desempenho dos modelos de *Gradient Boosting* e CNN foram semelhantes aos verificados no experimento realizado com todos os atributos do conjunto Ω. Em relação ao modelo Árvore de Decisão, embora seja o modelo que apresenta a cauda mais substancial em todas as métricas, podemos notar uma melhora considerável em seu desempenho quando comparado ao desempenho do primeiro experimento. Por outro lado, o modelo SVM, que estava no segundo estágio de desempenho no primeiro experimento, apresentou resultados próximos ao modelo de referência, desta vez.

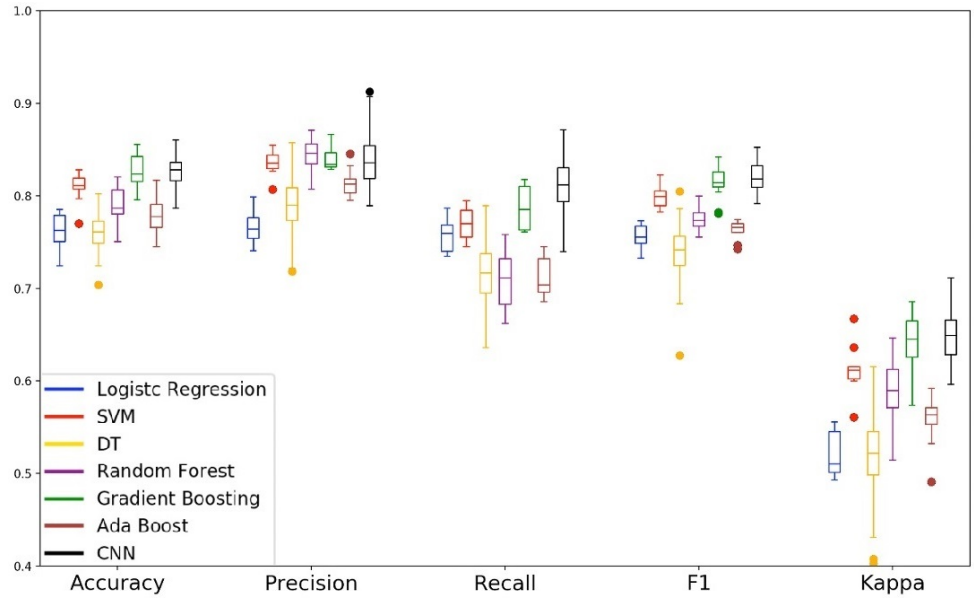


Figura 67. Desempenho dos modelos com os atributos coletados pelo acelerômetro.

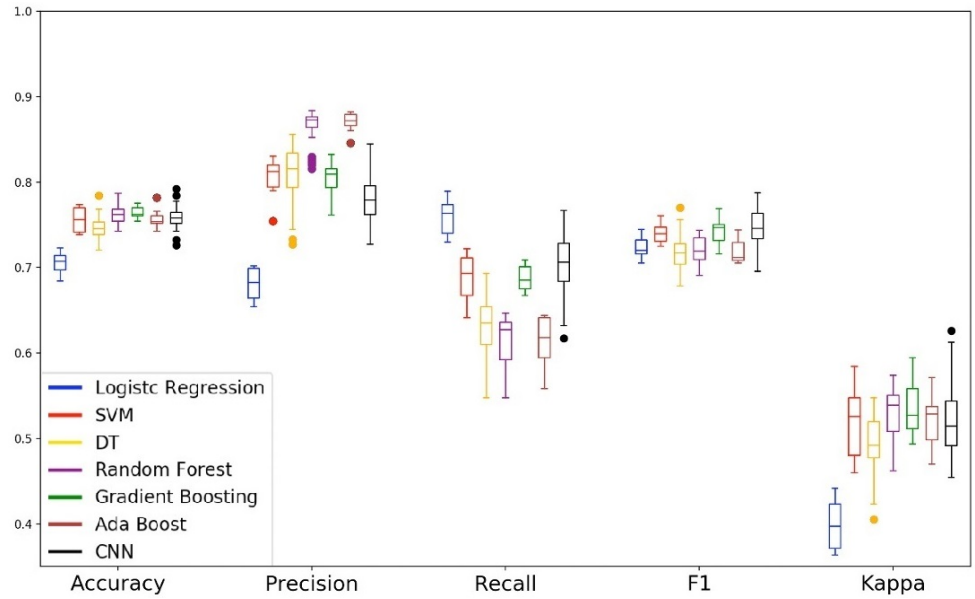


Figura 68. Desempenho dos modelos com os atributos coletados pelo giroscópio.

7.2.4 Verificação da significância estatística

Nessa fase do experimento, observou-se a similaridade do desempenho dos modelos *Gradient Boosting* e CNN. Para certificar esta hipótese, nós aplicamos um teste de significância estatística para verificar a equivalência entre os modelos. Chamamos de GB_1 e CCN_1 os modelos *Gradient Boosting* e *Convolutional Neural Network* treinados com todos os atributos do conjunto Ω e de GB_2 e CNN_2 os modelos treinados apenas com o subconjunto de recursos derivados do sensor magnetômetro. Primeiro, geramos a estimativa da densidade dos modelos e aplicamos o teste normal para garantir que todos os modelos tenham uma distribuição normal com 95% de confiança. Assumindo que todas as

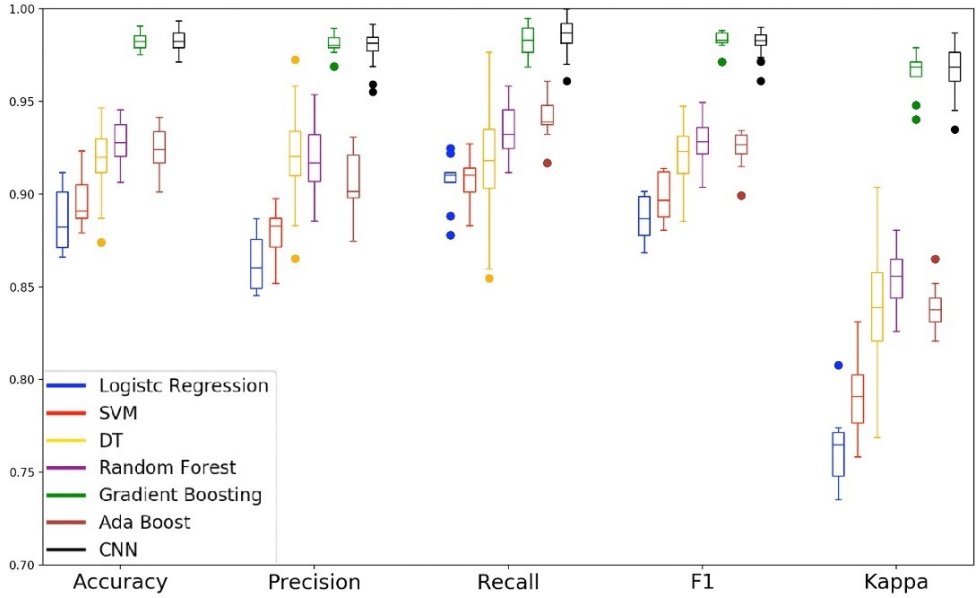


Figura 69. Desempenho dos modelos com os atributos coletados pelo magnetômetro.

distribuições são normais, aplicamos o teste *Student's* t-test. A hipótese nula (H0) do teste considera que os modelos são iguais e a hipótese alternativa (H1) considera que os modelos são diferentes. Ao aplicar o teste de *Student*, se obtivermos um valor de $p \ leq 0.05$, significa que podemos rejeitar a hipótese nula (H0) e considerar que os modelos de aprendizado de máquina são significativamente diferentes com 95% de confiança. Aplicamos o teste para todas as combinações dois a dois dos modelos. Como estamos comparando quatro modelos, realizamos seis testes de hipótese. A Tabela 24 mostra o valor de p para cada teste e a descrição se a hipótese H0 foi rejeitada ou aceita.

Tabela 24. Teste de hipótese *t-Student's* aplicado aos modelos.

Model 1	**Model 2**	***p*-Value**	**H0**
GB_1	GB_2	0.030	reject
CNN_1	CNN_2	0.576	accept
CNN_1	GB_1	0.123	accept
CNN_2	GB_2	0.327	accept
CNN_1	GB_2	0.841	accept
CNN_2	GB_1	0.215	accept

Podemos ver na Tabela 24 que somente a comparação dos modelos *Gradient Boosting* tiveram sua hipótese rejeitada. A rejeição da hipótese H0 na comparação entre GB_1 e GB_2 nos mostrou que o teste de *Student* não é distributivo. Por exemplo, o fato de GB_1 ser equivalente a CNN_1 e CNN_1 ser equivalente a GB_2 não implica equivalência de GB_1 e GB_2. Se levarmos em conta o desempenho dos modelos, a quantidade de informação

necessária e a equivalência estatística, podemos dizer que o modelo CNN_2 é o mais adequado para realizar a classificação de condutores e passageiros no evento de *reading while driving*. CNN_2 é o melhor modelo, pois é estatisticamente equivalente a todos os outros, apresenta um desempenho semelhante para todas as métricas analisadas e requer apenas informações do sensor do magnetômetro para realizar a classificação.

7.2.5 Análise dos modelos em função da importância dos atributos

Na terceira fase do experimento, usamos o modelo *Random Forest* para calcular a importância de cada atributo do conjunto Ω. A ideia é que o cálculo utilizado para determinar o nó de uma árvore possa ser usado para avaliar a importância relativa desse nó sobre a previsibilidade da variável de destino. Os recursos usados no topo da árvore são mais importantes para a decisão final da previsão. De acordo com (LOUPPE, 2014), nas árvores aleatórias, a importância de um atributo X_j pode ser avaliado somando-se a impureza ponderada $p(t)\Delta_i(s_t, t)$ para todos os nós t, em que X_j é usado e obtendo a média de todas as árvores M na floresta. A equação 7.1 calcula a redução média na impureza de um atributo. Na equação, X_j representa o atributo que queremos calcular a importância; M representa o número de árvores na floresta; t é um nó da árvore; $p(t)$ é a proporção de amostras que atingem t; e $\Delta_i(s_t, t)$ é a impureza de t que é calculada por alguma medida de impureza, como o *Gini index* ou *Shannon entropy*.

$$Imp(X_j) = \frac{1}{M} \sum_{m=1}^{M} \sum_{t \in m} p(t)\Delta_i(s_t, t) \tag{7.1}$$

Podemos ver a importância de cada atributo do conjunto Ω na Figura 70. Os seis atributos mais importante são: {*1: gyroz_std, 2: magx_mean, 3: magx_std, 4: gyroz_mean, 5: magz_mean, 6: magy_mean*}. Considerando que no segundo experimento fomos capazes de construir um modelo que usa seis atributos e apresenta um desempenho equivalente ao modelo que usa todos os atributos do conjunto Ω, neste terceiro experimento o nosso objetivo é verificar se é possível construir um modelo equivalente usando 6 atributos ou menos. Assim, treinamos e comparamos o desempenho dos modelos levando em conta os seis atributos mais importantes. Em seguida, repetimos esse procedimento para os 5 atributos mais importantes e depois para os 4 mais importantes. Nós não avaliamos o subconjunto *top 3* porque observamos a queda de desempenho nos experimentos *top 6*, *top 5* e *top 4*.

Na Figura 71 podemos observar o gráfico radar dos modelos. De uma perspectiva global, podemos notar que o desempenho do modelo diminui a medida que o conjunto de atributos também diminui. No gráfico de radar dos subconjuntos *top 6* e *top 5*, podemos observar a hierarquia de desempenho que também é vista nos experimentos anteriores. No entanto, na análise do subconjunto *top 4*, a hierarquia de desempenho é inexistente. Esse

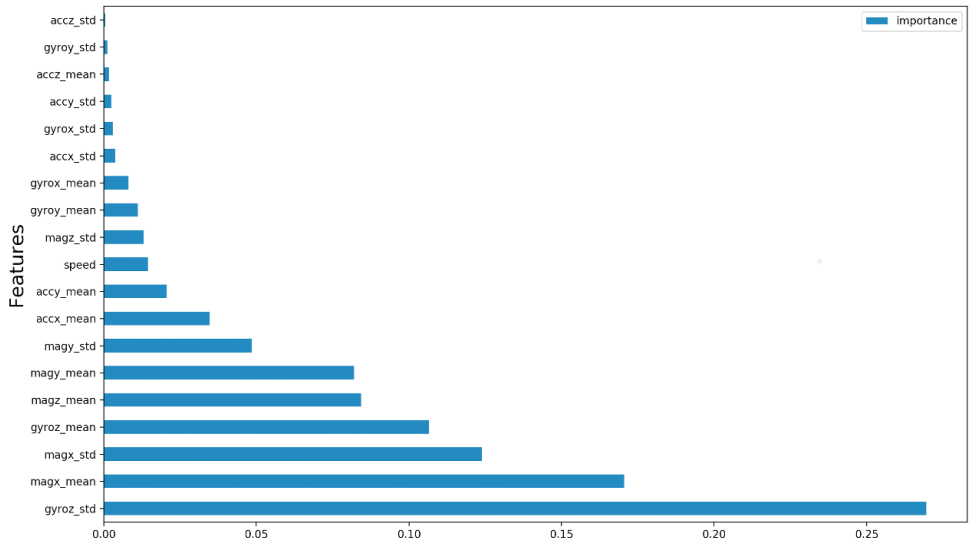


Figura 70. Importância dos atributos computada pelo modelo *Random Forest*.

comportamento, e a queda de desempenho em todas as métricas, mostram que os modelos não podem generalizar o problema usando apenas os atributos do subconjunto *top 4*. No gráfico radar, podemos ver que os modelos GB e CNN mostram um comportamento similar em todas as cinco métricas analisadas. Para certificar essa semelhança, aplicamos o teste estatístico para verificar a equivalência entre os modelos. Além dessa comparação, nós também realizamos uma correlação estatística com os melhores modelos dos experimentos anteriores.

Tabela 25. Teste de hipótese t-Student's aplicado aos modelos construídos nas três fases do experimento.

Model 1	**Model 2**	***p*-Value**	**H0**
CNN_{top6}	CNN_1	0.092	accept
CNN_{top6}	CNN_2	0.154	accept
CNN_{top6}	GB_1	0.998	accept
CNN_{top6}	GB_2	0.012	reject
GB_{top6}	GB_1	0.075	accept
GB_{top6}	GB_2	0.476	accept
GB_{top6}	CNN_1	0.766	accept
GB_{top6}	CNN_2	0.675	accept
GB_{top6}	CNN_{top6}	0.032	reject

A Tabela 25 mostra os resultados do teste de equivalência estatística aplicado aos modelos. Usamos a estimativa de densidade da *accuracy* na comparação dos modelos. Depois de certificar que todas as distribuições são normais, aplicamos o *Student's t*-test.

Figura 71. Gráfico radar com resultados da terceira avaliação.

A hipótese nula (H0) considera que os modelos são iguais e a hipótese alternativa (H1) considera que os modelos são diferentes. Assim como no experimento da segunda fase, aqui, também usamos 95% de confiança, ou seja, se obtivermos um valor de $p \leq 0.05$, significa que podemos rejeitar a hipótese nula e os modelos podem ser considerados significativamente diferentes com 95% de confiança. Na 25, chamamos os dois melhores modelos do subconjunto *top 6* de GB_{top6} e CNN_{top6}. Os nomes dos melhores modelos do primeiro e segundo experimentos foram mantidos como GB_1, CNN_1, GB_2 e CNN_2.

Embora os modelos CNN_{top6} e GB_{top6} mostrem desempenho semelhante para todas as métricas analisadas, o teste indica que, estatisticamente, esses modelos são diferentes. No entanto, ao comparar o desempenho desses modelos com os melhores modelos obtidos nas duas primeiras fases dos experimentos, podemos ver que ambos são estatisticamente equivalentes em quase todas as comparações. A única exceção foi a comparação entre os modelos CNN_{top6} e GB_2. O modelo GB_{top6}, que é um modelo de *Gradient Boosting*, foi considerado equivalente a todos os melhores modelos dos dois primeiros experimentos. Assim, levando em conta a análise dos três experimentos, podemos verificar as seguintes evidências:

- O problema da classificação de passageiros e motoristas, no caso de leitura durante a condução, pode ser resolvido de uma forma não intrusiva, a partir de uma abordagem de aprendizagem máquina. Esta conclusão foi demonstrada a partir do desempenho dos modelos desenvolvidos.

- Considerando todos os recursos do conjunto Ω, o uso do modelo CNN ou GB é estatisticamente equivalente na tarefa de classificar os passageiros e os motoristas no evento de *reading while driving*.

- O subconjunto de Ω, contendo apenas recursos coletados pelo sensor do acelerômetro ou pelo sensor giroscópio, não foi suficiente para construir modelos eficientes para a classificação de passageiros e motoristas no evento de *reading while driving*.

- O subconjunto de Ω, contendo apenas recursos coletados pelo sensor magnetômetro, proporcionou a construção de modelos de aprendizado de máquina eficientes para classificar motoristas e passageiros no caso de leitura durante a condução. De acordo com a análise e teste estatístico, a Rede Neural Convolucional foi o melhor modelo para trabalhar com este subconjunto.

- O subconjunto de Ω, contendo as seis principais características derivadas do modelo *Random Forest*, proporcionou o desenvolvimento de modelos eficientes para classificar passageiros e motoristas no caso de leitura durante a condução. De acordo com a análise e teste estatístico, o modelo *Gradient Boosting* foi o melhor modelo para trabalhar com esse subconjunto.

- Para $K \leq 5$, o subconjunto de Ω contendo as principais características de K derivadas do modelo *Random Forest* não foi suficiente para construir modelos eficientes para a classificação de passageiros e motoristas no evento de *reading while driving*.

7.3 Considerações finais

Estudos mostram que a distração do motorista é um problema crescente e que desencadeia lesões e mortes no trânsito. Neste contexto, esta pesquisa apresentou uma análise de aprendizado de máquina para resolver o problema da distração devido à leitura de mensagens de texto, de forma não intrusiva. Para analisar os efeitos da distração, coletamos as informações usando os sensores GPS, acelerômetro, magnetômetro e giroscópio de um *smartphone*. Criamos um aplicativo para coletar e rotular o conjunto de dados. O aplicativo gerou mensagens aleatórias que deviam ser lidas pelo usuário. Durante a leitura, conseguimos armazenar os valores dos sensores. No total, construímos um conjunto de dados com 11.000 amostras.

Nós dividimos o banco de dados em três conjuntos disjuntos: treinamento, validação e teste. Os conjuntos de treinamento e validação foram divididos e aplicados no processo de validação cruzada de *10-folds*. Depois de cada fase do experimento, usamos o conjunto de testes para verificar a capacidade de generalização dos modelos. Três experimentos foram realizados. O objetivo foi verificar a influência dos atributos Ω e a capacidade de generalização dos modelos de aprendizado de máquina Ψ, na tarefa de distinguir passageiros e motoristas durante a distração devido à leitura de mensagens de texto Φ, ou seja, nós avaliamos como Φ pode ser resolvido por $\Psi(\Omega)$ e também a influência de $\omega \in \Omega \; \forall \; \psi \in \Psi$.

No primeiro experimento, verificamos a capacidade de generalização dos modelos, levando em consideração todos os atributos do conjunto Ω. Os resultados deste experimento mostraram que os modelos CNN e GB tiveram o melhor desempenho. No segundo experimento, analisamos o desempenho dos modelos, considerando como entrada apenas os atributos coletadas pelo mesmo sensor. Assim, observamos como os modelos se comportaram, tomando como dados de entrada os recursos do acelerômetro, magnetômetro e sensor de giroscópio. Os resultados mostram que os modelos não generalizaram o problema apenas com as informações dos sensores acelerômetro ou do giroscópio. Em relação aos dados do sensor magnetômetro, os resultados mostram que a CNN teve a melhor generalização. Os testes estatísticos também apontaram a equivalência entre o modelo CNN e os dois melhores modelos do experimento que levaram em consideração todo o conjunto Ω.

No terceiro experimento, utilizamos o modelo *Random Forest* para verificar a importância das características. Considerando os atributos mais importantes, a ideia era encontrar um subconjunto de Ω que pudesse ser usado para construir modelos equivalentes ou mais eficientes do que os do primeiro experimento. Como já havíamos encontrado um modelo equivalente no segundo experimento e esse modelo usava apenas seis recursos como dados de entrada, neste terceiro experimento procurávamos encontrar um subconjunto de atributos ≤ 6. Os resultados mostraram que os subconjuntos ≤ 5 não foram suficientes para construir um modelo com desempenho similar aos melhores modelos do primeiro experimento. No entanto, com o subconjunto dos seis atributos mais importantes

obtidos pelo modelo *Random Forest*, foi possível construir um modelo *Gradient Boosting* estatisticamente equivalente aos melhores modelos adquiridos nos experimentos anteriores.

Tendo em conta os resultados apresentados nesta pesquisa, podemos concluir que a CNN e o GB são modelos eficientes para distinguir passageiros e motoristas em caso de leitura durante a condução. O fato de construirmos modelos que funcionam com dados não intrusivos indica a possibilidade de criar soluções comerciais que possam fornecer segurança ao motorista e conveniência para os passageiros. De acordo com as tarefas planejadas para contemplar o objetivo desta tese, este experimento corresponde a realização da seguinte tarefa:

- Realização de um experimento utilizando a metodologia de análise dos efeitos causados pelo comportamento do condutor para detectar a leitura de mensagens de texto e classificação de passageiros e motoristas.

Os resultados deste experimento foram publicados em 2019 na *Sensors 2019, Volume 19, Issue 14* com o título *A Machine-Learning Approach to Distinguish Passengers and Drivers Reading While Driving*.

8 Conclusões

É inegável que o acidente de trânsito é um problema em escala mundial que tem provocado milhares de mortes nos últimos anos. Das várias causas de acidentes de trânsito, destaca-se a questão da distração dos motoristas como um dos principais motivos, devido ao crescente uso do telefone celular por condutores. Estudos como (BRAITMAN; BRAITMAN, 2017; CHEN et al., 2016; RAJESH et al., 2017) demonstram que a utilização do telefone celular tem contribuído significativamente na distração dos motoristas e, consequentemente, nos índices de acidentes e mortes no trânsito. Para minimizar este problema, campanhas educativas e leis proibitivas do uso de telefone celular estão sendo empregadas em diferentes países do mundo. Entretanto, de acordo com os dados divulgados pelo OMS (CHAN, 2018), a utilização do telefone celular é um problema pertinente em diferentes países. A melhor forma de resolver a questão da distração em decorrência da utilização do telefone celular seria a conscientização dos usuários. Como as estatísticas demonstram que isso ainda não é plenamente possível, soluções tecnológicas estão sendo desenvolvidas para esta finalidade. As soluções assistivas conhecidas como *Advance Driver Assistance System* (ADAS) já são uma realidade em alguns países. Essas tecnologias têm a finalidade de interagir com o ambiente e prevenir possíveis falhas humanas. *Adaptive Cruise Control*, *Collision Prevention*, *Speed Bump Detection*, *Brake Assist*, *Lane Departure Warning System*, *Electronic Stability Control* e *Automotive Night Vision* são exemplos de ADAS.

Apesar dos ADAS surgirem como contribuições importantes para a segurança dos motoristas e passageiros, essa realidade ainda é exclusiva para os países desenvolvidos (CHAN, 2015). Nos países subdesenvolvidos, onde os índices de acidentes e mortes no trânsito costumam ser maiores, as tecnologias assistivas de prevenção de acidente de trânsito ainda não são uma realidade. Com o objetivo de resolver esse problema, pesquisadores estão utilizando as tecnologias do próprio *smartphone* para construir soluções de prevenção ao acidente de trânsito. Nessa perspectiva, o presente trabalho utilizou os sensores de um *smartphone* para coletar dados e analisar o comportamento e os efeitos do comportamento dos motoristas, a partir de modelos de aprendizado de máquina. Foram desenvolvidos e analisados diferentes modelos de aprendizado de máquina com o objetivo de detectar a distração do motorista devido ao uso do telefone celular. As distrações devido à realização de chamadas telefônicas, composição de mensagens de texto e leitura de mensagens/notificações foram levadas em consideração nos experimentos realizados. Durante o desenvolvimento dos modelos, observou-se a questão da intrusão. No contexto da distração do motorista, uma solução não intrusiva é aquela que não necessita da interação do usuário para funcionar. Ou seja, em uma solução não intrusiva as chances do usuário

burlar o seu funcionamento são minimizadas.

Nós entendemos que as soluções precisam ser não intrusivas já que está comprovado que nem todos os usuários possuem a consciência para evitar o uso do telefone celular enquanto estão dirigindo. Essa comprovação é evidenciada uma vez que as leis de proibição do telefone celular são ignoradas pelos motoristas. Para garantir as características da não intrusão, nós modelamos os algoritmos para adquirir a capacidade de generalizar e classificar motoristas e passageiros durante um evento distraído. Na condução dos experimentos, nós avaliamos diferentes modelos de aprendizado de máquina, na tarefa de classificar motoristas e passageiros durante os eventos de *texting while driving* e *reading while driving*. Esses dois eventos secundários foram investigados pois são as distrações que proporcionam maiores riscos de acidentes de trânsito. Tanto a leitura quanto a escrita de mensagens de texto faz com que o motorista fique distraído fisicamente, cognitivamente e visualmente. Apesar de ativarem as mesmas distrações, na revisão da literatura apresentada no Capítulo 2, só foram encontrados trabalhos que investigam e desenvolvem soluções não intrusivas para o evento de *texting while driving*. Nesse caso, entendemos que o experimento realizado para analisar a questão do *reading while driving* é um estudo pioneiro nessa área. Em síntese, podemos destacar as seguintes contribuições desta tese:

- Condução de uma revisão sistemática para investigar os métodos, algoritmos e sensores utilizados nas soluções de prevenção da distração do motorista, em decorrência da utilização do telefone celular.
- Aplicação da metodologia de observação do comportamento do condutor para desenvolver uma solução de *deep learning* que classifique dois tipos de distração do motorista: chamada telefônica e escrita de mensagem de texto.
- Aplicação da metodologia de observação dos efeitos do comportamento do condutor para desenvolver e analisar diferentes modelos de aprendizado de máquina que classifique, de forma não intrusiva, motoristas e passageiros durante o evento de *texting while driving*.
- Aplicação da metodologia de observação dos efeitos do comportamento do condutor para desenvolver e analisar diferentes modelos de aprendizado de máquina que classifique, de forma não intrusiva, motoristas e passageiros durante o evento de *reading while driving*.
- Condução de um *survey* para investigar o perfil de distração dos motoristas paraenses em decorrência da utilização do telefone celular.

O desenvolvimento das atividades desta tese resultou nas seguintes publicações:

- *A Deep Learning Approach to Detect Distracted Drivers Using a Mobile Phone* na *26th International Conference on Artificial Neural Networks* em 2017. Qualis B1.
- *Exploiting Machine Learning Models to Avoid Texting While Driving* na *International Joint Conference on Neural Networks* em 2019. Qualis A1.
- *A Machine-Learning Approach to Distinguish Passengers and Drivers Reading While Driving* na *Sensors 2019, Volume 19, Issue 14* em 2019. Qualis A1.

Como trabalhos futuros, as seguintes atividades podem ser realizadas:

Utilização dos modelos desenvolvidos para a construção de softwares que autem como serviço nos smartphones.

Nesse contexto, podemos definir que o software como serviço é um programa que fica em segundo plano de execução, monitorando os demais aplicativos utilizados pelo usuário. A ideia é que os sensores do telefone celular sejam ativados para coletar dados em tempo real, os quais serão utilizados pelo modelo de aprendizado de máquina para realizar a classificação dos motoristas e passageiros. Dada a classificação realizada, o software como serviço poderia atuar, de forma apropriada, para garantir a conveniência do passageiro e a segurança do condutor. Por exemplo, assumindo que o modelo classifica que o motorista está utilizando o telefone celular, o software como serviço poderia prevenir a distração a partir do bloqueio do dispositivo.

Devido a complexidade contextual para saber qual o aplicativo que está rodando em primeiro plano, outra possibilidade seria a implementação dos modelos de aprendizado de máquina no próprio *backend* de aplicativos como *whatsapp*, *messenger* e *instagram*. Tal implementação facilitaria o entendimento do contexto dos dados sensoriais coletados, uma vez que não seria necessário a programação de rotinas de sistemas operacionais. No próprio aplicativo, seria mais fácil saber se o mesmo está sendo utilizado para leitura ou composição de mensagens de texto. Nesse caso, os dados sensoriais já seriam submetidos para o modelo correto de classificação de motoristas e passageiros.

Avaliação e ajustes dos modelos desenvolvidos para classificar motoristas e passageiros em outros contextos de distração

Nesta tese, foram classificados motoristas e passageiros no contexto da distração decorrente da leitura e composição de mensagens de texto. Além dessas distrações, na pesquisa realizada e apresentada no capítulo 3, foi observado que a distração proveniente

do envio de mensagens de áudio também costuma ser praticada durante a condução de veículos. Sendo assim, como trabalhos futuros, pode-se avaliar e, se necessário, ajustar os modelos de aprendizado de máquina apresentados nessa tese, para classificar motoristas e passageiros nesse tipo de distração.

Outra possibilidade de pesquisa implica a investigação dos modelos de aprendizados de máquina no contexto da distração devido à leitura de mensagens de texto, porém sem a distração física. A base de dados construída e utilizada no estudo apresentado no capítulo 7 levou em consideração que os motoristas e passageiros utilizam o telefone celular para ler mensagens de texto segurando o telefone celular com uma ou as duas mãos. Entretanto, os usuários também costuma utilizar o telefone celular com um suporte para acoplar ao painel do veículo. Nesse caso, a leitura de mensagens e notificações são realizadas com menor interação física. Devido ao contexto apresentado, a investigação da distração devido à leitura de mensagens de texto sem interação física é uma possibilidade de pesquisa na qual pode ser avaliado a aplicabilidade dos modelos de aprendizado de máquina apresentados nesta tese.

Avaliação de outras técnicas de aprendizado de máquina, no contexto das soluções não intrusivas, para prevenir a distração de motorista

Outra possibilidade de pesquisa, diz respeito à investigação de outras técnicas de aprendizado de máquina, no contexto do objetivo desta tese. Nos três experimentos abordados nesta tese, foram utilizados modelos de aprendizado de máquina supervisionados. Foram construídos e ajustados os parâmetros de modelos de Árvore de Decisão, Regressão Logística, Máquina de Vetor de Suporte, Random Forest, Adaptive Boosting, Gradient Boosting e Redes Neurais Convolucionais. Os modelos foram escolhidos levando em consideração a revisão sistemática realizada e apresentada no capítulo 2. Contudo, outros ajustes e outros modelos podem ser investigados. Por exemplo, a abordagem de modelos não supervisionados pode ser uma vertente de pesquisa futura desta tese.

Aplicação do *survey* para investigar o perfil de distração dos motoristas em outros estados brasileiros

A aplicação do survey apresentado no capítulo 3 em outros estados do brasileiros, também é uma possibilidade de trabalhos futuros. Tal aplicação é importante para diagnosticar se o perfil de distração se altera dependendo da região geográfica. Uma vez observada a mudança de comportamento, poder-se-ia pensar em soluções específicas para cada região investigada.

A partir das sugestões de atividades futuras, pode-se perceber que o problema da distração do motorista devido à utilização do celular ainda é um problema suscetível de

soluções e que precisa ser investigado. Contudo, os resultados obtidos no experimentos desta tese trazem contribuições e avanços significativos para este tema. O desenvolvimento de modelos de aprendizado de máquina que possibilitem a construção de soluções não intrusivas para os eventos de *texting while driving* e *reading while driving* foram as principais contribuições dessa tese. Ademais, a realização do *survey* também proporcionou, pela primeira vez, a investigação do perfil de distração dos motoristas paraenses devido à utilização do telefone celular.

Referências

ABDEL-HAMID, O. et al. Convolutional neural networks for speech recognition. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, IEEE Press, Piscataway, NJ, USA, v. 22, n. 10, p. 1533–1545, out. 2014. ISSN 2329-9290. Disponível em: <http://dx.doi.org/10.1109/TASLP.2014.2339736>. Citado na página 112.

ALCANTARA, G. Empirical analysis of non-linear activation functions for deep neural networks in classification tasks. *CoRR*, abs/1710.11272, 2017. Disponível em: <http://arxiv.org/abs/1710.11272>. Citado na página 100.

ALLUHAIBI, S.; AL-DIN, M.; MOYAID, A. Driver behavior detection techniques: A survey. *International Journal of Applied Engineering Research (IJAER)*, v. 13, p. 8856–8861, 2018. Citado na página 145.

AMBEV. *Retrato da Segurança Viária no Brasil - 2017*. [S.l.], 2017. Disponível em: <https://www.ambev.com.br/conteudo/uploads/2017/09/Retrato-da-Seguran%C3%A7a-Vi%C3%A1ria_Ambev_2017.pdf>. Citado 4 vezes nas páginas 27, 28, 65 e 68.

ANDROIDRANK. *Android application ranking: All applications*. [S.l.], 2018. <https://www.androidrank.org/app/ranking?hl=en>. Citado na página 29.

AT&T. *Smartphone Use While Driving Grows Beyond Texting to Social Media, Web Surfing, Selfies*. [S.l.], 2015. Disponível em: <http://about.att.com/story/smartphone_use_while_driving_grows_beyond_texting.html>. Citado 2 vezes nas páginas 29 e 140.

BERRI, R. A. et al. A pattern recognition system for detecting use of mobile phones while driving. In: *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*. [S.l.: s.n.], 2014. v. 2. Citado na página 48.

BILLAH, T.; RAHMAN, S.; MAHBUBUR, M. Tracking-based detection of driving distraction from vehicular interior video. In: *2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. [S.l.: s.n.], 2016. Doi: 10.1109/AVSS.2016.7738077. Citado na página 48.

BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738. Citado 2 vezes nas páginas 95 e 97.

BO, C. et al. Detecting Driver #x2019;s Smartphone Usage via Nonintrusively Sensing Driving Dynamics. *IEEE Internet of Things Journal*, v. 4, n. 2, p. 340–350, 2017. Citado na página 60.

BO, C. et al. You're driving and texting: Detecting drivers using personal smart phones by leveraging inertial sensors. In: *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*. [S.l.: s.n.], 2013. (MobiCom' 13). Doi: 10.1145/2500423.2504575. Citado 2 vezes nas páginas 49 e 60.

BRAITMAN, K. A.; BRAITMAN, A. L. Patterns of distracted driving behaviors among young adult drivers: Exploring relationships with personality variables. *Transportation*

Research Part F: Traffic Psychology and Behaviour, v. 46, p. 169 – 176, 2017. Citado 2 vezes nas páginas 64 e 165.

BREIMAN, L. Random forest. *Springer Machine Learning*, v. 45, p. 5–32, 2001. Citado 2 vezes nas páginas 88 e 89.

BUDUMA, N.; LOCASCIO, N. *Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms*. 1st. ed. [S.l.]: O'Reilly Media, Inc., 2017. ISBN 1491925612, 9781491925614. Citado na página 97.

BURGES, C. J. C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.*, Kluwer Academic Publishers, Norwell, MA, USA, v. 2, n. 2, p. 121–167, jun. 1998. ISSN 1384-5810. Citado na página 91.

BURKE, J. et al. Participatory sensing. 2006. Doi: 10.1.1.122.3024. Citado na página 38.

BURTON, A. et al. Driver identification and authentication with active behavior modeling. In: *12th International Conference on Network and Service Management (CNSM)*. [S.l.: s.n.], 2017. Doi: 10.1109/CNSM.2016.7818453. Citado na página 131.

CARNEY, C. et al. Using naturalistic driving data to examine teen driver behaviors present in motor vehicle crashes, 2007-2015. *AAA Foundation for Traffic Safety*, 2016. Disponível em: <https://aaafoundation.org/wp-content/uploads/2017/12/TeenCrashCausationII.pdf>. Citado na página 29.

CHAN, M. *Global status report on road safety: time for action*. [S.l.], 2009. Disponível em: <http://apps.who.int/iris/bitstream/handle/10665/44122/9789241563840_eng.pdf>. Citado na página 27.

CHAN, M. *Global Status Report on Road Safety 2015*. [S.l.], 2015. Disponível em: <http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/>. Citado 3 vezes nas páginas 28, 128 e 165.

CHAN, M. *World Health Organization, Global status report on road safety 2018*. [S.l.], 2018. Licence: CC BY- NC-SA 3.0 IGO. Citado 4 vezes nas páginas 27, 28, 66 e 165.

CHAUDHARI, K.; GARG, D.; KOTECHA, K. An enhanced approach in image mosaicing using orb method with alpha blending technique. *International Journal of Advanced Research in Computer Science*, v. 8, n. 5, 2017. Doi: https://doi.org/10.26483/ijarcs.v8i5.3487. Citado 2 vezes nas páginas 20 e 102.

CHEN, H.-Y. W. et al. Self-reported engagement in driver distraction: An application of the Theory of Planned Behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 38, p. 151–163, 2016. Citado 2 vezes nas páginas 64 e 165.

CHOLLET, F. *Deep Learning with Python*. 1st. ed. Greenwich, CT, USA: Manning Publications Co., 2017. ISBN 1617294438, 9781617294433. Citado 4 vezes nas páginas 20, 97, 98 e 118.

CHOUDHARY, P.; VELAGA, N. R. Analysis of vehicle-based lateral performance measures during distracted driving due to phone use. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 44, p. 120–133, 2017. Citado na página 53.

CHOUDHARY, P.; VELAGA, N. R. Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour. *Accident Analysis & Prevention*, v. 106, p. 370–378, 2017. Citado na página 64.

CHOUDHARY, P.; VELAGA, R. Modelling driver distraction effects due to mobile phone use on reaction time. *Transportation Research Part C: Emerging Technologies*, v. 77, 2017. Doi: 10.1016/j.trc.2017.02.007. Citado na página 49.

CHUANG, M. C. et al. Estimating Gaze Direction of Vehicle Drivers Using a Smartphone Camera. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2014. p. 165–170. Doi: 10.1109/CVPRW.2014.30. Citado na página 56.

CNT. *Pesquisa CNT de rodovias 2018: relatório gerencial*. [S.l.], 2018. Disponível em: <https://cms_pesquisarodovias.cnt.org.br>. Citado 3 vezes nas páginas 19, 64 e 65.

D'AGOSTINO, R.; PEARSON, E. S. Tests for departure from normality. empirical results for the distributions of b2 and $\sqrt{b1}$. *Biometrika*, Oxford University Press, Biometrika Trust, v. 60, n. 3, p. 613–622, 1973. Citado na página 135.

D'AGOSTINO, R. B. An omnibus test of normality for moderate and large size samples. *Biometrika*, Oxford University Press, Biometrika Trust, v. 58, n. 2, p. 341–348, 1971. Citado na página 135.

DEVAPRIYA, W.; BABU, C.; SRIHARI, T. Advance driver assistance system (adas) - speed bump detection. *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, December. 2015. Doi: 10.1109/ICCIC.2015.7435753. Citado na página 28.

DINGUS, T. et al. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America*, v. 113, n. 10, March. 2016. Doi: 10.1073/pnas.1513271113. Citado na página 30.

DINH, T. H. et al. Image segmentation based on histogram of depth and an application in driver distraction detection. In: *2014 13th International Conference on Control Automation Robotics Vision (ICARCV)*. [S.l.: s.n.], 2014. Doi: 10.1109/ICARCV.2014.7064437. Citado na página 48.

EHSANI, J.; LI, K.; SIMONS-MORTON, B. G. Teenage drivers portable electronic device use while driving. In: *2015 Driving Assessment Conference*. [S.l.: s.n.], 2015. Doi: 10.17077/drivingassessment.1575. Citado na página 128.

ERSAL, T. et al. Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks. *IEEE Transactions on Intelligent Transportation Systems*, v. 11, n. 3, p. 692–701, set. 2010. ISSN 1524-9050. Doi : 10.1109/TITS.2010.2049741. Citado na página 56.

GIL, A. *Como elaborar projetos de pesquisa*. 4. ed. [S.l.]: S ao Paulo, 2002. Citado 3 vezes nas páginas 50, 51 e 52.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>. Citado 5 vezes nas páginas 20, 94, 96, 99 e 101.

GRAS, M. E. et al. Mobile phone use while driving in a sample of Spanish university workers. *Accident Analysis & Prevention*, v. 39, n. 2, 2007. Doi: 10.1016/j.aap.2006.08.006. Citado 4 vezes nas páginas 52, 63, 64 e 83.

GU, J. et al. Recent advances in convolutional neural networks. *Pattern Recognition*, v. 77, p. 354–377, 2018. Doi: https://doi.org/10.1016/j.patcog.2017.10.013. Citado 3 vezes nas páginas 94, 100 e 102.

HANSEN, J. et al. Driver modeling for detection and assessment of distraction: Examples from the utdrive test bed. *IEEE Signal Processing Magazine*, v. 34, n. 4, 2017. Doi: 10.1109/MSP.2017.2699039. Citado na página 41.

HAQUE, M. et al. Decisions and actions of distracted drivers at the onset of yellow lights. *Accident Analysis & Prevention*, v. 96, p. 290–299, 2016. Citado na página 85.

HAQUE, M. M. et al. Decisions and actions of distracted drivers at the onset of yellow lights. *Accident Analysis & Prevention*, v. 96, p. 290–299, nov. 2016. ISSN 0001-4575. Doi : 10.1016/j.aap.2015.03.042. Citado 3 vezes nas páginas 54, 58 e 59.

HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998. ISBN 0132733501. Citado na página 99.

HE, J. et al. Mutual interferences of driving and texting performance. *Computers in Human Behavior*, v. 52, p. 115–123, 2015. Citado na página 85.

HOSKING, S.; YOUNG, K.; REGAN, M. The effects of text messaging on young novice driver performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, v. 51, n. 4, August. 2009. Doi: 10.1177/0018720809341575. Citado 2 vezes nas páginas 30 e 125.

HUTH, V.; SANCHEZ, Y.; BRUSQUE, C. Drivers' phone use at red traffic lights: A roadside observation study comparing calls and visual-manual interactions. *Accident Analysis & Prevention*, v. 74, 2015. Doi: 10.1016/j.aap.2014.10.008. Citado na página 48.

IBGE. *Pesquisa Nacional por Amostra de Domicílios: Síntese de Indicadores*. [S.l.], 2016. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv98887.pdf>. Citado 2 vezes nas páginas 29 e 64.

IM, S. et al. Driver distraction detection by in-vehicle signal processing. In: *2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*. [S.l.: s.n.], 2014. p. 64–68. Doi: 10.1109/CIVTS.2014.7009479. Citado na página 56.

JUNIOR, J. et al. Driver behavior profiling: An investigation with different smartphone sensors and machine learning. *Transportation Research Part F: Traffic Psychology and Behaviour*, April. 2017. Doi: 10.1371/journal.pone.0174959. Citado 3 vezes nas páginas 29, 131 e 145.

KAMENCAY, P. et al. A new method for face recognition using convolutional neural network. *Advances in Electrical and Electronic Engineering*, v. 15, n. 4, 2017. Doi: https://doi.org/10.26483/ijarcs.v8i5.3487. Citado 2 vezes nas páginas 112 e 118.

KARATAS, C.; LI, H.; LIU, L. Toward detection of unsafe driving with inertial head-mounted sensors. In: *Proceedings of the Eighth Wireless of the Students, by the Students, and for the Students Workshop*. New York, NY, USA: ACM, 2016. (S3), p. 45–47. Doi: 10.1145/2987354.2987369. Citado na página 54.

KARLIK, B.; VEHBI, A. Performance analysis of various activation functions in generalized mlp architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems (IJAE)*, v. 1, n. 4, p. 111–122, 2011. Citado na página 99.

KI-MOON, B. *Global Plan for the Decade of Action for Road Safety 2011-2020*. [S.l.], 2010. Disponível em: <http://www.who.int/roadsafety/decade_of_action/plan/plan_english.pdf>. Citado 2 vezes nas páginas 27 e 45.

KIDD, D. et al. The influence of roadway situation, other contextual factors, and driver characteristics on the prevalence of driver secondary behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 41, 2016. Doi: 10.1016/j.trf.2016.06.004. Citado na página 48.

KITCHENHAM, B. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. [S.l.], 2007. Disponível em: <https://www.elsevier.com/___data/promis_misc/525444systematicreviewsguide.pdf>. Citado na página 40.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. USA: Curran Associates Inc., 2012. (NIPS'12), p. 1097–1105. Disponível em: <http://dl.acm.org/citation.cfm?id=2999134.2999257>. Citado na página 59.

KUTILA, M. et al. Driver distraction detection with a camera vision system. In: *2007 IEEE International Conference on Image Processing*. [S.l.: s.n.], 2007. v. 6. Doi: 10.1109/ICIP.2007.4379556. Citado na página 48.

LARSSON, P. Speech feedback reduces driver distraction caused by in-vehicle visual interfaces. In: *Proceedings of the Audio Mostly 2016*. [S.l.: s.n.], 2016. (AM' 16). Doi: 10.1145/2986416.2986435. Citado na página 48.

LE, T. H. N. et al. Multiple Scale Faster-RCNN Approach to Driver #x2019;s Cell-Phone Usage and Hands on Steering Wheel Detection. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. [S.l.: s.n.], 2016. p. 46–53. Doi: 10.1109/CVPRW.2016.13. Citado 3 vezes nas páginas 59, 60 e 112.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature International Journal of Science*, v. 521, p. 436–444, 2015. Doi:10.1038/nature14539. Citado 3 vezes nas páginas 100, 102 e 112.

LEE, E.; KIM, C. Automated work zone information system (awis) on urban freeway rehabilitation: California implementation. *Transportation Research Board 85th Annual Meeting*, January. 2006. Doi: 10.3141/1948-09. Citado na página 38.

LI, G. et al. Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 74, January. 2017. Doi: 10.1016/j.trc.2016.11.011. Citado na página 29.

LI, N.; BUSSO, C. Analysis of facial features of drivers under cognitive and visual distractions. In: *2013 IEEE International Conference on Multimedia and Expo (ICME)*. [S.l.: s.n.], 2013. Doi: 10.1109/ICME.2013.6607575. Citado 2 vezes nas páginas 48 e 49.

LI, N.; BUSSO, C. Analysis of facial features of drivers under cognitive and visual distractions. In: *2013 IEEE International Conference on Multimedia and Expo (ICME)*. [S.l.: s.n.], 2013. p. 1–6. Citado na página 58.

LIU, T. et al. Driver distraction detection using semi-supervised machine learning. *IEEE Transactions on Intelligent Transportation Systems*, v. 17, n. 4, 2016. Doi: 10.1109/TITS.2015.2496157. Citado 2 vezes nas páginas 32 e 48.

LIU, X. et al. Drive now, text later: Nonintrusive texting-while-driving detection using smartphones. *IEEE Transactions on Mobile Computing*, v. 16, n. 1, 2017. Doi: 10.1109/TMC.2016.2539171. Citado 3 vezes nas páginas 49, 54 e 126.

LOUPPE, G. Understanding random forest from theory to practice. *PhD dissertation, University of Lieg Faculty of Applied Sciences Department of Electrical Engineering and Computer Science*, 2014. Citado na página 159.

MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. Rectifier nonlinearities improve neural network acoustic models. In: *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*. [S.l.: s.n.], 2013. Citado na página 100.

MANASSEH, C.; SENGUPTA, R. Predicting driver destination using machine learning techniques. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. [S.l.: s.n.], 2013. Doi: 10.1109/ITSC.2013.6728224. Citado na página 131.

MÁRQUEZ, L.; CANTILLO, V.; ARELLANA, J. Mobile phone use while driving: A hybrid modeling approach. *Accident Analysis & Prevention*, v. 7, p. 73 – 80, 2015. Citado na página 64.

MAYA, S.; NAZIR, H. Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Elsevier Computers in Human Behavior*, v. 57, p. 312–325, 2016. Citado na página 29.

MCHUGH, M. Interrater reliability: the kappa statistic. *US National Library of Medicine National Institutes of Health*, v. 22, p. 76–82, 2012. Citado na página 156.

MEIRING, G.; MYBURGH, H. A review of intelligent driving style analysis systems and related artificial intelligence algorithms. *Sensors (Basel, Switzerland)*, v. 15, n. 12, December. 2015. Doi: 10.3390/s151229822. Citado na página 29.

MUKHERJEE, A.; MAJHI, S. Characterisation of road bumps using smartphones. *European Transport Research Review*, v. 13, n. 8, June. 2016. Doi: 10.1007/s12544-016-0200-1. Citado na página 28.

MUSICANT, O.; LOTAN, T. Can novice drivers be motivated to use a smartphone based app that monitors their behavior? *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 42, n. 3, October. 2016. Doi: 10.1016/j.trf.2015.10.023. Citado na página 29.

NHSTA, N. T. L. C. *Investigation and prosecution of distracted driving cases (Report No. DOT HS 812 407)*. [S.l.], 2017. Disponível em: <https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/812407-distracteddrivingreport.pdf>. Citado na página 32.

ORTIZ, N.; RAMNARAYAN, M.; MIZENKO, K. Distraction and road user behavior: An observational pilot study across intersections in washington, d.c. *Journal of Transport & Health*, 2017. Doi: 10.1016/j.jth.2017.05.362. Citado na página 48.

OU, Y.-K.; LIU, Y.-C.; SHIH, F.-Y. Risk prediction model for drivers' in-vehicle activities Application of task analysis and back-propagation neural network. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 18, p. 83–93, maio 2013. ISSN 1369-8478. Doi: 10.1016/j.trf.2012.12.013. Citado na página 56.

PORTAL, T. S. *Mobile internet traffic as percentage of total web traffic in November 2018*. [S.l.], 2018. <https://www.statista.com/statistics/306528/share-of-mobile-internet-traffic-in-global-regions/>. Citado na página 29.

PRF, P. R. F. *Statistical database of injuries and deaths on federal highways*. [S.l.], 2015. Disponível em: <https://www.prf.gov.br/portal/dados-abertos/acidentes/acidentes>. Citado na página 28.

PRF, P. R. F. *Statistical database of injuries and deaths on federal highways*. [S.l.], 2018. Disponível em: <https://www.prf.gov.br/portal/dados-abertos/acidentes/acidentes>. Citado 3 vezes nas páginas 23, 65 e 66.

PRODANOV, C.; FREITAS, E. *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico*. 2. ed. [S.l.]: Rio Grande do Sul, 2013. Citado na página 52.

PRZYBYLA, J. et al. Estimating risk effects of driving distraction: A dynamic errorable car-following model. *Transportation Research Part C: Emerging Technologies*, v. 50, p. 117–129, jan. 2015. ISSN 0968-090X. Doi: 10.1016/j.trc.2014.07.013. Citado na página 54.

RAJENDRA, V.; DEHZANGI, O. Detection of distraction under naturalistic driving using Galvanic Skin Responses. In: *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. [S.l.: s.n.], 2017. p. 157–160. Citado 3 vezes nas páginas 32, 54 e 146.

RAJESH, R. et al. Modeling safety risk perception due to mobile phone distraction among four wheeler drivers. *IATSS Research*, v. 41, n. 1, p. 30–37, 2017. Citado 2 vezes nas páginas 64 e 165.

RASCHKA, S. Model evaluation, model selection, and algorithm selection in machine learning. *CoRR*, 2018. Disponível em: <http://arxiv.org/abs/1811.12808>. Citado 2 vezes nas páginas 20 e 104.

RASCHKA, S.; MIRJALILI, V. *Python Machine Learning Second Edition*. 2. ed. [S.l.]: Packt Publishing Ltd, 2017. Birmingham B3 2PB, UK. Citado 10 vezes nas páginas 19, 20, 86, 87, 88, 89, 90, 91, 92 e 117.

RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput.*, v. 29, n. 9, p. 2352–2449, 2017. ISSN 0899-7667. Citado na página 112.

REGAN, M.; LEE, J.; YOUNG, K. *Driver distraction : theory, effects, and mitigation*. [S.l.]: Taylor & Francis Group, 2009. Citado na página 32.

REIMER, B.; MEHLER, B.; DONMEZ, B. A study of young adults examining phone dialing while driving using a touchscreen vs. a button style flip-phone. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 23, p. 57–68, 2014. Citado na página 146.

RIPLEY, B. D.; HJORT, N. L. *Pattern Recognition and Neural Networks*. 1st. ed. New York, NY, USA: Cambridge University Press, 1995. ISBN 0521460867. Citado 2 vezes nas páginas 115 e 116.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. In: RUMELHART, D. E.; MCCLELLAND, J. L.; GROUP, C. P. R. (Ed.). Cambridge, MA, USA: MIT Press, 1986. cap. Learning Internal Representations by Error Propagation, p. 318–362. ISBN 0-262-68053-X. Disponível em: <http://dl.acm.org/citation.cfm?id=104279.104293>. Citado na página 99.

SABZEVARI, T. et al. An observational study of secondary task engagement while driving on urban streets in iranian safe communities. *Accident Analysis & Prevention*, v. 96, November 2016. Doi: 10.1016/j.aap.2016.07.020. Citado na página 48.

SAHAYADHAS, A. et al. A physiological measures-based method for detecting inattention in drivers using machine learning approach. *Biocybernetics and Biomedical Engineering*, v. 35, n. 3, p. 198–205, 2015. Citado 2 vezes nas páginas 32 e 48.

SAHAYADHAS, A. et al. A physiological measures-based method for detecting inattention in drivers using machine learning approach. *Biocybernetics and Biomedical Engineering*, v. 35, n. 3, p. 198 – 205, 2015. Citado na página 146.

SAÚDE, M. da. *Um em cada cinco brasileiros afirma dirigir usando o celular*. 2019. <http://saude.gov.br/noticias/agencia-saude/45539-um-em-cada-cinco-brasileiros-afirma-dirigir-usando-o-celular>. Accessed: 24-06-2019. Citado 3 vezes nas páginas 34, 65 e 84.

SAVELIFE. *Distracted driving in India a study on mobile phone usage, pattern & behavior*. [S.l.], 2017. Disponível em: <http://savelifefoundation.org/wp-content/uploads/2017/04/Distracted-Driving-in-India_A-Study-on-Mobile-Phone-Usage-Pattern-and-Behaviour.pdf>. Citado 2 vezes nas páginas 30 e 39.

SCHAPIRE, R.; FREUND, Y. *Boosting Foundations and Algorithms*. 1. ed. [S.l.]: The MIT Press, 2012. Cambridge, Massachusetts. Citado na página 92.

SCHMITT, F. et al. Predicting lane keeping behavior of visually distracted drivers using inverse suboptimal control. In: *2016 IEEE Intelligent Vehicles Symposium (IV)*. [S.l.: s.n.], 2016. p. 412–418. Doi: 10.1109/IVS.2016.7535419. Citado na página 60.

SCHROETER, R. et al. Exploring Boredom Proneness As a Predictor of Mobile Phone Use in the Car. In: *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*. New York, NY, USA: ACM, 2015. (OzCHI '15), p. 465–473. Citado na página 64.

SHAABAN, K. Investigating Cell Phone Use While Driving in Qatar. *Procedia - Social and Behavioral Sciences*, v. 104, 2013. Doi: 10.1016/j.sbspro.2013.11.201. Citado 4 vezes nas páginas 52, 63, 64 e 83.

SHARMA, N.; JAIN, V.; MISHRA, A. An analysis of convolutional neural networks for image classification. *Procedia Computer Science*, v. 132, p. 377–384, 2018. ISSN 1877-0509. Doi: https://doi.org/10.1016/j.procs.2018.05.198. Citado na página 112.

STOLEN, S.; GRANDE, T.; ALLAN, N. *Chemical Thermodynamics of Materials: Macroscopic and Microscopic Aspects*. [S.l.]: John Wiley and Sons, Ltda., 2004. Citado na página 37.

SWOV. *Use of the mobile phone while driving*. [S.l.], 2017. Disponível em: <https://www.swov.nl/en/facts-figures/factsheet/use-mobile-phone-while-drivingf>. Citado na página 32.

SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,*. [s.n.], 2016. Disponível em: <http://arxiv.org/abs/1512.00567>. Citado na página 118.

TANGO, F.; BOTTA, M. Real-Time Detection System of Driver Distraction Using Machine Learning. *IEEE Transactions on Intelligent Transportation Systems*, v. 14, n. 2, p. 894–905, jun. 2013. ISSN 1524-9050. Doi: 10.1109/TITS.2013.2247760. Citado na página 56.

TCHANKUE, P.; WESSON, J.; VOGTS, D. Using machine learning to predict the driving context whilst driving. In: *Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference*. New York, NY, USA: ACM, 2013. (SAICSIT '13), p. 47–55. Doi: 10.1145/2513456.2513472. Citado na página 54.

TUOMO, K.; HANNU, K.; JAKKE, M. Context-sensitive distraction warnings ? Effects on drivers? visual behavior and acceptance. *International Journal of Human-Computer Studies*, v. 90, p. 39–52, 2016. Citado na página 64.

VLAHOGIANNI, E. I.; BARMPOUNAKIS, E. N. Driving analytics using smartphones: Algorithms, comparisons and challenges. *Transportation Research Part C: Emerging Technologies*, v. 79, p. 196 – 206, 2017. Doi: 10.1016/j.trc.2017.03.014. Citado na página 131.

VORA, S.; RANGESH, A.; TRIVEDI, M. M. On generalizing driver gaze zone estimation using convolutional neural networks. In: *2017 IEEE Intelligent Vehicles Symposium (IV)*. [S.l.: s.n.], 2017. p. 849–854. Doi: 10.1109/IVS.2017.7995822. Citado 2 vezes nas páginas 59 e 112.

WANG, Y. et al. Determining Driver Phone Use by Exploiting Smartphone Integrated Sensors. *IEEE Transactions on Mobile Computing*, v. 15, n. 8, p. 1965–1981, ago. 2016. ISSN 1536-1233. Doi: 10.1109/TMC.2015.2483501. Citado na página 60.

WHO. *Mobile phone use: a growing problem of driver distraction*. [S.l.], 2011. Disponível em: <http://www.who.int/violence_injury_prevention/publications/road_traffic/en/index.html>. Citado 2 vezes nas páginas 30 e 51.

WOLLMER, M. et al. Online Driver Distraction Detection Using Long Short-Term Memory. *IEEE Transactions on Intelligent Transportation Systems*, v. 12, n. 2, p. 574–582, 2011. Citado na página 146.

XIAO, D.; FENG, C. Detection of drivers visual attention using smartphone. In: *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*. [S.l.: s.n.], 2016. p. 630–635. Doi: 10.1109/FSKD.2016.7603247. Citado 3 vezes nas páginas 54, 55 e 56.

YANNIS, G. et al. Simulation of texting impact on young drivers' behavior and safety on motorways. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 41, p. 10–18, 2016. Citado na página 85.

YE, M. et al. Detection of driver engagement in secondary tasks from observed naturalistic driving behavior. *Accident Analysis & Prevention*, v. 106, p. 385–391, 2017. Citado 3 vezes nas páginas 49, 54 e 56.

YEN, T.; HUO, C.; SUN, T. Adaptive lane departure warning system on android smartphone. *2014 IEEE International Conference on Consumer Electronics - Taiwan*, May. 2014. Doi: 10.1109/ICCE-TW.2014.6904104. Citado na página 28.

YOU, C.-W. et al. Carsafe app: Alerting drowsy and distracted drivers using dual cameras on smartphones. In: *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*. New York, NY, USA: ACM, 2013. (MobiSys '13), p. 13–26. Citado na página 54.

YUN, Y. et al. Video-based detection and analysis of driver distraction and inattention. In: *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*. [S.l.: s.n.], 2014. p. 190–195. Doi: 10.1109/SPIN.2014.6776946. Citado na página 54.

ZADEH, R.; GHATEE, M.; EFTEKHARI, H. Three-phases smartphone-based warning system to protect vulnerable road users under fuzzy conditions. *IEEE Transactions on Intelligent Transportation Systems*, v. 19, n. 7, October. 2017. Doi: 10.1109/TITS.2017.2743709. Citado na página 28.

ZHOU, Y.; CHELLAPPA, R. Computation of optical flow using a neural network. In: *IEEE 1988 International Conference on Neural Networks*. [S.l.: s.n.], 1988. Doi: 10.1109/ICNN.1988.23914. Citado na página 102.

ZHU, Y. et al. Hidden two-stream convolutional networks for action recognition. *CoRR*, abs/1704.00389, 2017. Disponível em: <http://arxiv.org/abs/1704.00389>. Citado na página 112.

ZYLIUS, G. Investigation of route-independent aggressive and safe driving features obtained from accelerometer signals. *IEEE Intelligent Transportation Systems Magazine*, v. 9, p. 103 – 113, 2017. Doi: 10.1109/MITS.2017.2666583. Citado na página 131.

APÊNDICE A . Artigos filtrados na revisão sistemática

Tabela 26. Parte 1: Lista de artigos filtrados.

Id	Título	Repositório
1	An observational study of secondary task engagement while driving on urban streets in Iranian Safe Communities	Science Direct
2	Tracking-Based Detection of Driving Distraction from Vehicular Interior Video	IEEE
3	Patterns of distracted driving behaviors among young adult drivers: Exploring relationships with personality variables	Science Direct
4	Older driver distraction: A naturalistic study of behaviour at intersections	Science Direct
5	Self-reported engagement in driver distraction: An application of the Theory of Planned Behaviour	Science Direct
6	Mobile phone use while driving in a sample of Spanish university workers	Science Direct
7	Managing in-vehicle distractions - evidence from the Psychological Refractory Period paradigm	ACM
8	Drivers? phone use at red traffic lights: A roadside observation study comparing calls and visual?manual interactions	Science Direct
9	The influence of roadway situation, other contextual factors, and driver characteristics on the prevalence of driver secondary behaviors	Science Direct
10	Driver Distraction Detection with a Camera Vision System	IEEE
11	Mobile phone use while driving: A hybrid modeling approach	Science Direct
12	Culture of distracted driving among intra-city commercial bus drivers in Ile-Ife, South-western Nigeria	Science Direct
13	Distraction and road user behavior: An observational pilot study across intersections in Washington, D.C.	Science Direct
14	Identifying the main factors contributing to driving errors and traffic violations ? Results from naturalistic driving data	Science Direct
15	Assessing characteristics related to the use of seatbelts and cell phones by 2 drivers: Application of a bivariate probit model	Science Direct
16	Exploring Boredom Proneness as a Predictor of Mobile Phone Use in the Car	ACM
17	Investigating Cell Phone Use While Driving in Qatar	Science Direct
18	CarSafe App: Alerting Drowsy and Distracted Drivers using Dual Cameras on Smartphones	ACM

Continued on next page

19	Distraction-induced driving error: An on-road examination of the errors made by distracted and undistracted drivers	Science Direct
20	A Pattern Recognition System for Detecting Use of Mobile Phones While Driving	IEEE
21	Image segmentation based on histogram of depth and an application in driver distraction detection	IEEE
22	Texting while driving: Is speech-based text entry less risky than handheld text entry?	Science Direct
23	Modeling safety risk perception due to mobile phone distraction among four wheeler drivers	Science Direct
24	Driving context influences drivers' decision to engage in visual?manual phone tasks: Evidence from a naturalistic driving study	Science Direct
25	Decisions and actions of distracted drivers at the onset of yellow lights	Science Direct
26	Analysis of vehicle-based lateral performance measures during distracted driving due to phone use	Science Direct
27	TEXIVE: Detecting Drivers Using Personal Smart Phones by Leveraging Inertial Sensors	ACM
28	Simulation of texting impact on young drivers? behavior and safety on motorways	Science Direct
29	Mutual interferences of driving and texting performance	Science Direct
30	Modelling driver distraction effects due to mobile phone use on reaction time	Science Direct
31	Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour	Science Direct
32	Effects of Smartphone Usage on Driver Safety Level Performance in Urban Road Conditions	IEEE
33	Drive Now, Text Later: Nonintrusive Texting-while-Driving Detection using Smartphones	IEEE
34	Determining Driver Phone Use by Exploiting Smartphone Integrated Sensors	IEEE
35	Detection of driver engagement in secondary tasks from observed naturalistic driving behavior	Science Direct
36	Detecting Driver?s Smartphone Usage via Non-intrusively Sensing Driving Dynamics	IEEE
37	Using Machine Learning to Predict the Driving Context whilst Driving	ACM
38	Toward Detection of Unsafe Driving with Inertial Head-Mounted Sensors	ACM
39	Examination of drivers' cell phone use behavior at intersections by using naturalistic driving data	Science Direct

Continued on next page

40	Using SHRP 2 naturalistic driving data to assess drivers' speed choice while being engaged in different secondary tasks	Science Direct
41	Risk prediction model for drivers? in-vehicle activities ? Application of task analysis and back-propagation neural network	Science Direct
42	Investigating the effect of area type and traffic conditions on distracted driving performance	Science Direct
43	Impact of mobile phone use on car-following behaviour of young drivers	Science Direct
44	Using an infotainment system while driving ? A continuous analysis of behavior adaptations	Science Direct
45	Real-world effects of using a phone while driving on lateral and longitudinal control of vehicles	Science Direct
46	Predicting Lane Keeping Behavior of Visually Distracted Drivers Using Inverse Suboptimal Control	IEEE
47	Model-Based Analysis and Classification of Driver Distraction Under Secondary Tasks	IEEE
48	Detecting Drivers? Mirror-Checking Actions and Its Application to Maneuver and Secondary Task Recognition	IEEE
49	Driver Distraction Detection By In-Vehicle Signal Processing	IEEE
50	Driver Distraction Detection for Vehicular Monitoring	IEEE
51	Variations in Driver?s Mechanical Admittance Facing Distracting Tasks	IEEE
52	Online Driver Distraction Detection Using Long Short-Term Memory	IEEE
53	Driver Identification and Driver's Emotion Verification using KDE and MLP Neural Networks	IEEE
54	Short Paper: Study of Driver Workload-based Smartphone Control Service	IEEE
55	Multiple Scale Faster-RCNN Approach to Driver?s Cell-phone Usage and Hands on Steering Wheel Detection	IEEE
56	Driver Distraction Recognition Based on Dual Compass Motion Sensoring	IEEE
57	Android OpenCV Based Effective Driver Fatigue and Distraction Monitoring System	IEEE
58	Driver Cell Phone Usage Detection From HOV/HOT NIR Images	IEEE
59	DriveSafe: an App for Alerting Inattentive Drivers and Scoring Driving Behaviors	IEEE
60	SafeWatch: A Wearable Hand Motion Tracking System for Improving Driving Safety	ACM
61	Estimating Gaze Direction of Vehicle Drivers using a Smartphone Camera	IEEE
62	What factors influence drivers? response time for evasive maneuvers in real traffic?	Science Direct

Continued on next page

63	Texting while driving using Google GlassTM: Promising but not distraction-free	Science Direct
64	Continuous Car Driving Intention Recognition with Syntactic Pattern Approach	IEEE
65	Context-Sensitive Distraction Warnings Effects on Drivers' Visual Behavior and Acceptance	Science Direct
66	Speech Feedback Reduces Driver Distraction Caused by In-vehicle Visual Interfaces	ACM
67	Analysis of Facial Features of Drivers Under Cognitive and Visual Distractions	IEEE
68	Driver Distraction Detection Using Semi-Supervised Machine Learning	IEEE
69	The Development of Driver Warning Algorithm for Safety Driving Support	IEEE
70	Estimating risk effects of driving distraction: A dynamic errorable car-following model	Science Direct
71	Detection of Distraction under Naturalistic Driving Using Galvanic Skin Responses	IEEE
72	A study of young adults examining phone dialing while driving using a touchscreen vs. a button style flip-phone	Science Direct
73	Automatic system for detecting driver use of mobile phones	Science Direct
74	A physiological measures-based method for detecting inattention in drivers using machine learning approach	Science Direct
75	distratto : Impaired Driving Detection Using Textile Sensors	IEEE
76	Detecting Driver Phone Calls in a Moving Vehicle Based on Voice Features on simulated driving performance and eye glance behaviour.	IEEE
77	Real-Time Detection System of Driver Distraction Using Machine Learning	IEEE
78	Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving	Science Direct
79	On Generalizing Driver Gaze Zone Estimation using Convolutional Neural Networks	IEEE
80	Autonomous Detection of Distracted Driving by Cell Phone	IEEE
81	Detection of Drivers Visual Attention Using Smartphone	IEEE
82	Young driver distraction by text messaging: A comparison of the effects of reading and typing text messages in Chinese versus English	Science Direct
83	Effects of phone type on driving and eye glance behaviour while text-messaging	Science Direct
84	Video-Based Detection and Analysis of Driver Distraction and Inattention	IEEE
85	The Study of Driver Distraction Characteristic Detection Technology	IEEE

Continued on next page

86	Detecting Driver Distraction Using Smartphones	IEEE

Tabela 27. Parte 2: Lista de artigos filtrados.

Id	**Técnica**	**Método de estudo**
1	chi-square test and Fisher's exact test	Macroscópico
2	KNN	Macroscópico
3	tabulação de dados	Macroscópico
4	Useful Field of View (UFOV), Trail Making Test B, Mini Mental Status Examination (MMSE)	Macroscópico
5	chi-square e tabulação de dados	Macroscópico
6	chi-square e tabulação de dados	Macroscópico
7	ANOVA and Stimulus onset asynchrony (SOA)	Efeito da distração
8	chi-square	Macroscópico
9	Binary logistic regression analyses	Macroscópico
10	rule?based and support-vector machine (SVM)	Macroscópico
11	integrated choice latent variables model (ICLV) and a mixed logit model (ML)	Macroscópico
12	tabulação de dados	Macroscópico
13	tabulação de dados	Macroscópico
14	tabulação de dados	Macroscópico
15	bivariate probit model	Macroscópico
16	hierarchical regression analysis	Macroscópico
17	chi-square and logistic regression	Macroscópico
18	SVM	Macroscópico
19	generalised estimating equations (GEE)	Microscópico
20	Support Vector Machine (SVM) with Polynomial kernel	Macroscópico
21	histogram of depth values	Macroscópico
22	tabulação de dados	Efeito da distração
23	maximum likelihood method	Macroscópico
24	chi-square e Wilcoxon signed rank test	Macroscópico
25	decision tree and logistic regression analyses	Efeitos da distração
26	ANOVA	Efeitos da distração
27	Hidden Markov Model (HMM) and naive Bayesian classifier	Efeitos da distração

Continued on next page

28	Lognormal regression methods and binary logistic methods	Efeitos da distração
29	Logistic regression	Efeitos da distração
30	parametric survival analysis	Efeitos da distração
31	linear regression models e binary logistic regression models	Efeitos da distração
32	tabulação de dados	Efeitos da distração
33	wavelet transform (WT) and Adaptive peak-thresholding	Efeitos da distração
34	threshold algorithm	Efeitos da distração
35	Artificial Neural Network (ANN)	Efeitos da distração
36	Hidden Markov Model (HMM) and naive Bayesian	Efeitos da distração
37	decision tree	Efeitos da distração
38	Support Vector Machine (SVM)	Efeitos da distração
39	ANOVA	Efeitos da distração
40	tabulação de dados	Efeitos da distração
41	Artificial Neural Network (ANN)	Efeitos da distração
42	descriptive analysis and generalized linear models	Efeitos da distração
43	ANOVA and Generalized Estimation Equation (GEE)	Efeitos da distração
44	ANOVA	Efeitos da distração
45	Linear mixed models (LMM)	Efeitos da distração
46	Markov Decision Processes (MDPs)	Efeitos da distração

Continued on next page

47	Support Vector Machine (SVM) e Artificial Neural Network (ANN)	Efeitos da distração
48	Support Vector Machine and Nearest Neighbor and Random Under-Sampling Boost	Efeitos da distração
49	Artificial Neural Network and Gaussian mixture model	Efeitos da distração
50	Gaussian Mixture model (GMM) and nonlinear extended 2-wheel vehicle dynamic model	Efeitos da distração
51	tabulação de dados	Efeitos da distração
52	long short-term memory (LSTM) recurrent neural networks	Efeitos da distração
53	Artificial Neural Network (ANN) and KDE	Efeitos da distração
54	Threshold Algorithm	Efeitos da distração
55	Region-based Convolutional Neural Network	Macroscópico
56	Threshold Algorithm	Microscópico
57	adaptive template matching and adaptive boosting	Macroscópica
58	Threshold Algorithm	Macroscópica
59	Threshold Algorithm	Efeitos da distração
60	tatistical hypothesis testing, Mahalanobis Distance	Microscópico
61	Suport Vector Machine (SVM)	Microscópico
62	ANOVA	Microscópico
63	tabução de dados	Macroscópica
64	Gramática livre de contexto	Microscópico
65	Threshold Algorithm	Microscópico
66	ANOVA	Microscópico
67	logistic regression analysis, KNN, SVM	Microscópico
68	Laplacian support vector machine and semi-supervised extreme learning machine	Microscópico
69	Fuzzy	Microscópico
70	dynamic time warping algorithm	Microscópico
71	kernel analysis and support vector machines (SVM)	Microscópico
72	Poisson and logistic regression	Microscópico
73	Threshold Algorithm	Microscópico
74	k-nearest neighbour analysis, linear discriminant analysis and quadratic discriminant analysis.	MIcroscópico

Continued on next page

75	Threshold Algorithm	Efeitos da distração
76	Threshold Algorithm	Microscópico
77	support vector machines (SVMs), static and dynamic neural networks (NNs) feedforward NNs (FFNNs) and layer-recurrent NNs (LRNNs), adaptive neuro- fuzzy inference systems (ANFISs)	Efeitos da distração
78	tabução de dados	Microscópico
79	Convolutional Neural Network (CNN)	Microscópico
80	Entropia	Microscópico
81	Suport Vector Machine (SVM)	Microscópico
82	ANOVA	Efeitos da distração
83	ANOVA	Microscópico
84	Suport Vector Machine (SVM)	Microscópico
85	Artificial Neural Network (ANN)	Macroscópica
86	Geometric Transformation Estimation	Microscópico

Tabela 28. Parte 3: Lista de artigos filtrados.

Id	**País**	**Tipo de estudo**	**Sensores**
1	Iran	Observacional Naturalístico	NÃO
2	Bangladesh	Naturalístico	Câmera
3	USA	Survey	NÃO
4	Austrália	Naturalístico	Câmera e IMU
5	Canadá	Survey	NÃO
6	Espanha	Survey	NÃO
7	Reino Unido	Observacional Simulação	Break Pedal
8	França	Observacional Naturalístico	NÃO
9	USA	Observacional Naturalístico	NÃO
10	Finlândia	Naturalístico	Câmera e CAN BUS
11	Colômbia	Survey	NÃO
12	Nigéria	Observacional Naturalístico	NÃO
13	USA	Observacional Naturalístico	NÃO
14	Alemanha	Observacional Naturalístico	Câmera
15	USA	Observacional Naturalístico	NÃO
16	Austrália	Survey	NÃO
17	Catar	Survey	NÃO
18	China	Naturalístico	Câmera e IMU
19	Austrália	Naturalístico	eye tracking,speed, brake pedal, acceleromoter, lane change, GPS

Continued on next page

20	Brasil	Naturalístico	Câmera
21	Vietnã	Naturalístico	Câmera
22	USA	Oservacional Simulação	brake response time
23	Índia	Survey	NÃO
24	Suécia	Observacional	Câmera e CAN BUS
25	Austrália	Simulação	lane position, speed, acceleration, and braking
26	Índia	Observacional Simulação	lane position, speed, acceleration, and braking
27	USA	Naturalístico	IMU
28	Grécia	Observacional Simulação	Speed
29	USA	Simulação	lane deviation recorded
30	Índia	Observacional Simulação	speed, lateral performance
31	Índia	Simulação e Survey	Speed
32	Tailândia	Observacioal Naturalístico	GPS e Acelerômetro
33	China	Naturalístico	GPS e IMU
34	USA	Naturalístico	IMU
35	USA	Naturalístico	GPS e Acelerômetro e throttle position
36	USA	Naturalístico	GPS e IMU
37	África do Sul	Naturalístico	GPS e IMU
38	USA	Naturalístico	IMU
39	USA	Observacional Naturalístico	GPS e Câmera
40	Alemanha	Observacional Naturalístico	Câmera e speed record
41	Taiwan	Simulação	câmera e speed record
42	Grécia	Observacional Simulação	Câmera e speed record
43	Austrália	Observacional Simulação	GPS e Acelerômetro
44	Alemanha	Observacional Simulação	GPS e Acelerômetro
45	Suécia	Observacional Naturalístico	Câmera e speed record
46	Alemanha	Simulação	Câmera e speed record
47	USA	Simulação	speed, accelerometer, throttle and brake pedal position
48	USA	Naturalístico	microphone, câmera e brake pedals
49	coreia do sul	Simulação	speed, accelerometer, Steering wheel
50	USA	Simulação	GPS
51	Japão	Observacional Simulação	steering wheel angle and the torque provided
52	Alemanha	Naturalístico	(CAN)-Bus e head tracking system
53	Paquistão	Naturalístico	Câmera e brake pedal signal
54	coreia do sul	Naturalístico	speed, accelerometer, lane change, turn curve
55	USA	Naturalístico	Câmera
56	China	Naturalístico	Câmera
57	Índia	Naturalístico	Câmera

Continued on next page

58	USA	Naturalístico	Câmera
59	Espanha	Naturalístico	Câmera, Microfone, IMU, GPS
60	USA	Naturalístico	IMU, GPS
61	USA	Naturalístico	Câmera
62	Suécia	Observacional Naturalístico	Câmera, speed, brake pedal position, steering wheel
63	USA	Observacilnal Simulação	speed, break pedal, steering wheel, lane change
64	Coréia do sul	Simulação	Eye Tracker
65	Finlândia	Naturalístico e Survey	GPS, Câmera
66	Suécia	Observacilnal Simulação	Eye Tracker
67	USA	Naturalístico	Câmera, Microfone, (CAN) Bus
68	singapura	Naturalístico	Eye Tracker
69	coréia do sul	Naturalístico	Câmera, bio-signal, CAN-Bus
70	USA	Naturalístico	Câmera
71	USA	Naturalístico	wearable sensor network, Galvanic Skin Response (GSR)
72	USA	Observacilnal Simulação	Eye Tracker, speed, break, steering wheel
73	Espanha	Naturalístico	power frequency
74	Índia	Simulação	electrocardiogram (ECG) and surface electromyogram (sEMG)
75	USA	Naturalístico	IMU, break pedal, gas pedal, GPS, arm rest, head rest
76	USA	Naturalístico	Microfone, On-Board Unit (OBU)
77	Itália	Simulação	Câmera, speed, brake pedal position, steering wheel, lane change
78	Suécia	Observacional Naturalístico	Câmera, CAN-Bus
79	USA	Naturalístico	Câmera
80	USA	Simulação	TouchScreen
81	China	Naturalístico	Eye Tracker e Câmera
82	China	Observacilnal Simulação	Lateral control, Longitudinal Control, Speed
83	Austrália	Observacilnal Simulação	Câmera, Speed, Lateral control
84	Suécia	Naturalístico	Eye Tracker, Câmera
85	China	Naturalístico	Câmera
86	USA	Naturalístico	Câmera