

UNIVERSIDADE FEDERAL DO PARANÁ

ANDRÉ ROBERTO ORTONCELLI

SUMARIZAÇÃO E ANÁLISE DE ATIVIDADES PRÁTICAS DE DIREÇÃO PARA
DETECÇÃO E AVALIAÇÃO DE ESTRESSE E ANSIEDADE DE MOTORISTAS

CURITIBA PR

2021

ANDRÉ ROBERTO ORTONCELLI

SUMARIZAÇÃO E ANÁLISE DE ATIVIDADES PRÁTICAS DE DIREÇÃO PARA
DETECÇÃO E AVALIAÇÃO DE ESTRESSE E ANSIEDADE DE MOTORISTAS

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciência da Computação no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Olga Regina Pereira Bellon.

CURITIBA PR

2021

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

078s Ortoncelli, André Roberto
Sumarização e análise de atividades práticas de direção para
detecção e avaliação de estresse e ansiedade de motoristas
[recurso eletrônico] / André Roberto Ortoncelli – Curitiba, 2021.

Tese - Universidade Federal do Paraná, Setor de Ciências
Exatas, Programa de Pós-graduação em Informática.

Orientadora: Olga Regina Pereira Bellon

1. Visão por computador. 2. Motorista (Comportamento). I.
Universidade Federal do Paraná. II. Bellon, Olga Regina Pereira.
III. Título.

CDD: 006.37

Bibliotecária: Roseny Rivelini Morciani CRB-9/1585



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA -
40001016034P5

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **ANDRÉ ROBERTO ORTONCELLI** intitulada: **Sumarização e Análise de Atividades Práticas de Direção para Detecção e Avaliação de Estresse e Ansiedade de Motoristas**, sob orientação do Prof. Dr. OLGA REGINA PEREIRA BELLON, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 02 de Julho de 2021.

Assinatura Eletrônica

16/07/2021 15:33:21.0

OLGA REGINA PEREIRA BELLON
Presidente da Banca Examinadora

Assinatura Eletrônica

05/07/2021 12:59:55.0

LUCIANO SILVA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

13/07/2021 10:31:33.0

RODRIGO MINETTO
Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

Assinatura Eletrônica

05/07/2021 10:47:29.0

ROBERTO GREGORIO DA SILVA JUNIOR
Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

Rua Cel. Francisco H. dos Santos, 100 - Centro Politécnico da UFPR - CURITIBA - Paraná - Brasil

CEP 81531-980 - Tel: (41) 3361-3101 - E-mail: ppginf@inf.ufpr.br

Documento assinado eletronicamente de acordo com o disposto na legislação federal Decreto 8539 de 08 de outubro de 2015.

Gerado e autenticado pelo SIGA-UFPR, com a seguinte identificação única: 99181

Para autenticar este documento/assinatura, acesse <https://www.prppg.ufpr.br/siga/visitante/autenticacaoassinaturas.jsp> e insira o código 99181

*À minha família, dedico o resultado
do esforço realizado ao longo deste
percurso.*

AGRADECIMENTOS

A Deus, que me fortaleceu para chegar ao fim desse desafio.

A minha família, que sempre me incentivou e garantiu que eu não desistisse nunca.

A minha orientadora e aos membros da banca de avaliação, que contribuíram com dicas e sugestões.

Aos meus colegas, professores da Universidade Tecnológica Federal do Paraná campus Dois Vizinhos, que contribuíram para que eu pudesse dedicar um tempo maior para conclusão dessa tese.

A Autoescola Milênio da cidade de Fazenda Rio Grande-PR, cuja parceria foi fundamental para conclusão do trabalho.

Aos colegas do grupo de pesquisa IMAGO, pela sua ajuda preciosa dada ao desenvolvimento do meu trabalho.

Aos demais amigos, que forma direta ou indiretamente contribuíram para conclusão desse trabalho.

RESUMO

Esta tese de doutorado apresenta uma abordagem para detecção e análise de estados emocionais de estresse/ansiedade em motoristas a partir de dados coletados durante atividades práticas de direção. O método proposto tem o objetivo de identificar esses estados emocionais e agilizar o suporte necessário a ser prestado aos motoristas. Para isso, o método combina técnicas de priorização e sumarização de vídeos gravados durante atividades práticas de direção. Com os sumários produzidos, profissionais da área da psicologia podem identificar e compreender o comportamento do motorista com estresse/ansiedade de forma mais rápida, sem precisar assistir aos vídeos na sua integralidade, muito menos estar presente durante as atividades. Isso porque, com a técnica de priorização, é gerada uma lista de atividades práticas de direção, ordenadas pelo nível de estresse/ansiedade do motorista, possibilitando que sejam propostas medidas individuais para auxiliar motoristas a lidarem ou corrigirem esses estados. A técnica de priorização contribui com a abordagem de sumarização, pois agiliza o atendimento aos motoristas que mais precisam desse suporte. Ambas as técnicas propostas utilizam dados de Expressões Faciais (EF's), frequência cardíaca e movimentos do motorista — essas características foram selecionadas por já terem sido exploradas com sucesso em trabalhos relacionados. A maioria dos trabalhos correlatos focam na classificação automática de situações de estresse ou ansiedade em motoristas, portanto, a abordagem proposta nesta tese se diferencia por apresentar técnicas para sumarizar e priorizar atividades práticas de direção. A técnica de sumarização proposta utiliza também coordenadas de GPS, permitindo identificar o local no qual os comportamentos foram observados e, com isso, verificar gatilhos para estresse/ansiedade em cada motorista. A frequência cardíaca foi mensurada em termos de batimentos do coração por minuto. Quanto às EF's, um conjunto de 18 *Action Units* foi reconhecido com a ferramenta Open Face 2.0. Como movimentos, foram considerados hábitos do motorista que podem ocorrer em situações de estresse/ansiedade: pressionar, morder e lambear os lábios, roer as unhas e esfregar o rosto e/ou o cabelo. A técnica de priorização proposta combina um algoritmo clássico de ordenação com uma função comparativa baseada em uma Rede Neural Artificial. A função comparativa recebe uma lista de parâmetros extraídos de duas atividades de direção e identifica em qual delas o motorista apresentou maior nível de estresse/ansiedade. Quanto à sumarização, a abordagem proposta é apoiada por uma ferramenta desenvolvida para permitir a visualização de cada uma das atividades de direção em três perspectivas diferentes, exploradas para facilitar e organizar a análise dos dados: i) vídeos; ii) relatórios; e iii) sumários. Além das técnicas de priorização e sumarização, essa tese também apresenta um método para detecção automática dos movimentos das mãos dos motoristas, associados ao estresse/ansiedade. O método proposto para detecção desses gestos combina morfologia matemática com a detecção automática do rosto do motorista. Para avaliar a abordagem proposta, foram coletados dados em 60 atividades práticas de direção de alunos de uma autoescola. Os eventos de interesse foram rotulados manualmente na base de dados, que também foi priorizada, em relação ao nível de estresse/ansiedade dos motoristas, por uma psicóloga especializada no atendimento a motoristas. O resultado do método de priorização foi avaliado quanto à coesão e similaridade com o conjunto de referência que foi priorizado manualmente pela especialista. A ferramenta de suporte à sumarização de vídeo foi avaliada qualitativamente por profissionais ligados à autoescola na qual a base de dados foi produzida. Os resultados do método proposto para detecção automática de movimentos foram avaliados quanto à acurácia, revocação, precisão e F1-score, indicando que o suporte a motoristas pode ser agilizado, especialmente a aqueles com alto nível de estresse/ansiedade. Deve-se destacar que a

abordagem proposta é particularmente relevante no contexto de autoescolas com grande fluxo de alunos, nas quais é difícil que um especialista em Psicologia analise, em tempo hábil, todos os dados das atividades de direção realizadas. Nesse contexto, essa tese contribui diretamente para melhorar a formação de motoristas e conseqüentemente para a segurança no trânsito.

Palavras-chave: Visão Computacional. Comportamento de motoristas. Resumo de vídeo. Priorização de atividades.

ABSTRACT

This doctoral thesis presents an approach for detecting and analyzing the emotional states of drivers' stress/anxiety from data collected during practical driving activities. The proposed method aims to identify these emotional states and streamline the necessary support to be provided to drivers. For this, the method combines techniques of prioritization and summarization of videos recorded during practical activities of direction. With the summaries produced, professionals in the field of psychology can identify and understand the behavior of the driver with stress/anxiety more quickly, without having to watch the videos in their entirety, let alone being present during the activities. This is because, with the prioritization technique, a list of practical driving activities is generated, ordered by the driver's stress/anxiety level, enabling individual measures to be proposed to help drivers deal with or correct these states. The prioritization technique contributes to the summarization approach, as it speeds up the service to drivers who need this support the most. Both of the proposed techniques use data from Facial Expressions (EF's), heart rate and driver movements — these characteristics have been selected because they have already been successfully explored in related works. Most related works focus on the automatic classification of stress or anxiety situations in drivers, therefore, the approach proposed in this thesis differs by presenting techniques to summarize and prioritize practical driving activities. The proposed summarization technique also uses GPS coordinates, allowing to identify the place where the behaviors were observed and, with this, to verify triggers for stress/anxiety in each driver. Heart rate was measured in terms of heartbeats per minute. As for EF's, a set of 18 Action Units recognized with the Open Face 2.0 tool. As movements, it was considered driver's habits that can occur in stress/anxiety situations: pressing, biting and licking his lips, biting his nails, and rubbing his face and/or hair. The proposed prioritization technique combines a classic ordering algorithm with a comparative function based on an Artificial Neural Network. The comparative function receives a list of parameters extracted from two driving activities and identifies in which one the driver had the highest level of stress/anxiety. As for summarization, the proposed approach is supported by a tool developed to allow the visualization of each of the management activities in three different perspectives, explored to facilitate and organize the analysis of the data: i) videos; ii) reports; and iii) summaries. In addition to the prioritization and summarization techniques, this thesis also presents a method for automatic detection of drivers' hand movements, associated with stress/anxiety. The proposed method for detecting these gestures combines mathematical morphology with automatic detection of the driver's face. To assess the proposed approach, data were collected in 60 practical driving activities for students at a driving school. The events of interest were manually labeled in the database, which was also prioritized, in relation to the level of stress/anxiety of the drivers, by a psychologist specialized in assisting drivers. The result of the prioritization method was assessed for cohesion and similarity with the reference set that was prioritized manually by the specialist. The video summarization support tool was qualitatively evaluated by professionals linked to the driving school in which the database was produced. The results of the proposed method for automatic movement detection were evaluated for accuracy, recall, precision, and F1-score, indicating that support for drivers can be streamlined, especially for those with a high level of stress/anxiety. It should be noted that the proposed approach is particularly relevant in the context of driving schools with a large flow of students, in which it is difficult for a specialist in Psychology to analyze, in a timely manner, all the data of the direction activities performed. In

this context, this thesis contributes directly to improving the training of drivers and, consequently, to traffic safety.

Keywords: Computer Vision. Driver Behavior. Video Abstract. Prioritization of activities.

LISTA DE FIGURAS

1.1	Etapas da abordagem proposta.	23
2.1	<i>Framework</i> conceitual para sumarização de vídeo com quadros chaves (Ferreira et al., 2016)..	29
2.2	Atributos dos métodos de SV baseados em quadros chaves (Truong e Venkatesh, 2007)	30
2.3	Atributos dos métodos de SV baseados em cenas do vídeo (Truong e Venkatesh, 2007)	30
2.4	Geração de sumários baseados em curva de <i>skimming</i> (Truong e Venkatesh, 2007)	33
2.5	AU's utilizadas pela técnica de priorização - baseado em Farnsworth (2016).. . .	37
2.6	Exemplos de movimentos das mãos relacionados ao estresse.	38
3.1	Ano de publicação dos trabalhos selecionados no primeiro MSL	42
3.2	Ferramenta de visualização de vídeo proposta por Zhao et al. (2013)	48
3.3	Ano de publicação dos trabalhos selecionados no segundo MSL	52
3.4	Ferramenta de monitoramento da atividade de direção: (Kim et al., 2018).. . . .	59
3.5	Abordagem de visualização: mapa de estresse (Rodrigues et al., 2015)..	59
3.6	Dispositivo de visualização do nível de estresse: (MacLean et al., 2013)..	60
4.1	Câmeras utilizadas na etapa de coleta de dados	64
5.1	Processo de detecção de mãos e face nas imagens com o método 1	70
5.2	Exemplo de intersecção entre as regiões que representam a mão e a face do motorista	70
5.3	Exemplo de falso positivo na detecção de MMREA	71
5.4	Etapa de pós-processamento do método 1 de detecção de MMREA.	73
5.5	Principais problema do método 1 para detectar MMREA	73
5.6	Macro arquitetura da RNC VGG16 (Simonyan e Zisserman, 2014) - baseado em (Masood et al., 2020)	75
5.7	Macro arquitetura da RNC VGG19 (Simonyan e Zisserman, 2014) - baseado em (Masood et al., 2020)	75
5.8	Esquema de validação cruzado utilizado	78
6.1	Processo de produção do conjunto de treinamento	86
6.2	Gráficos de coesão e similaridade	89
6.3	Gráfico da acurácia da função comparativa.	91
7.1	Exemplo de conversão de valores do espaço discreto para o espaço contínuo . . .	95

7.2	Ferramenta de Visualização: Perspectiva de vídeos	96
7.3	Ferramenta de Visualização: Perspectiva de relatórios.	97
7.4	Exemplo de seleção de quadros para compor uma cena que fará parte de um sumário associado ao estresse/ansiedade	98
7.5	Ferramenta de Visualização: Perspectiva de sumários	99

LISTA DE TABELAS

3.1	Emoções/estados afetivos explorados pelos métodos de SV selecionados no MSL 1	43
3.2	Aplicações dos métodos de SV selecionados no MSL 1	44
3.3	Origem dos dados utilizados nos métodos de SV selecionados no MSL 1	45
3.4	Tipos de dados utilizados pelos métodos de SV selecionados no MSL 1	47
3.5	Aplicação e ambiente experimental dos trabalhos selecionados no MSL 2 - parte 1	53
3.6	Aplicação e ambiente experimental dos trabalhos selecionados no MSL 2 - parte 2	54
3.7	Dados utilizados pelos trabalhos selecionados no MSL 2 - parte 1	55
3.8	Dados utilizados pelos trabalhos selecionados no MSL 2 - parte 1	56
3.9	Técnica computacional utilizada nos trabalhos selecionados no MSL 2 - parte 1 .	57
3.10	Técnica computacional utilizada nos trabalhos selecionados no MSL 2 - parte 2 .	58
4.1	Cronograma referente a aprovação do projeto pelo CEP e coleta de dados	62
5.1	Resultados experimentais para 100% da base de dados	80
5.2	Resultados experimentais para 75% da base de dados*	80
5.3	Resultados experimentais para 100% da base de dados: não penalizando erros nos três quadros que precedem ou seguem o início ou fim de cada MMREA . . .	80
5.4	Resultados experimentais para 75% da base de dados *: não penalizando erros nos três quadros que precedem ou seguem o início ou fim de cada MMREA . . .	81
5.5	Matriz de confusão dos resultados do método 2 com VGG16 - 100% da base de dados	81
5.6	Matriz de confusão dos resultados do método 2 com VGG19 - 100% da base de dados	81
5.7	Matrizes de confusão da classificação binária dos resultados para 100% da base de dados	81
5.8	Matrizes de confusão da classificação binária dos resultados para 75% da base de dados *	82
6.1	Características utilizadas pelo método de priorização	85
6.2	Resultados de priorização obtidos	88
6.3	Matriz de Confusão: Função Comparativa com $\sqrt{ DA_{training} } = 50$ e Tempo = 25min	90
6.4	Acurácia da função comparativa	90

LISTA DE ACRÔNIMOS

ADEB	Algoritmo de Detecção de Estresse Biométrico
AENM	Atividade Elétrica de Nervos e Músculos
API	<i>Application Programming Interface</i>
AU	<i>Action Unit</i>
BPM	Batimentos por Minuto
CBC	<i>Connectivity-Based Clustering</i>
CE	Critério de Exclusão
CEPSH	Comitê de Ética em Pesquisa em Seres Humanos
CI	Critério de Inclusão
DINF	Departamento de Informática
EDA	<i>Electrodermal Activity</i>
EE	Elemento Estruturante
EEG	EletroEncefaloGrafia
EF	Expressão Facial
ELM	<i>Extreme Learning Machine</i>
FACS	<i>Facial Action Coding System</i>
FCM	<i>Fuzzy C-means</i>
GPS	<i>Global Positioning System</i>
HoG	<i>Histogram of Oriented Gradient</i>
IPCA	<i>Interaction Principal Component Analysis</i>
KBCS	<i>Kernel-Based Class Separability</i>
KMC	<i>K-means clustering</i>
KNN	<i>K-Nearest Neighbor</i>
LDA	<i>Linear Discriminant Analysis</i>
LBP	<i>Local Binary Patterns</i>
MMG	Modelos de Mistura Gaussiana
MMREA	Movimentos das Mãos Relacionados ao Estresse e a Ansiedade
MS	<i>Merge Sort</i>
MSL	Mapeamento Sistemático de Literatura
NFN	Número de Falsos Negativos
NFP	Número de Falsos Positivos
NMC	<i>Nearest Mean Classifier</i>
NVN	Número de Verdadeiros Negativos
NVP	Número de Verdadeiros Positivos
QP	Questão de Pesquisa

VE	Valência e Excitação
RL	Regressão Logística
RNA	Rede Neural Artificial
RNA-RN	Rede Neural Artificial Baseada em Regressão Numérica
RNC	Rede Neural Convolucional
RQMDS	Raiz Quadrada Média de Diferenças Sucessivas
SBL	<i>Sparse Bayesian Learning</i>
SFFS	<i>Sequential Forward Floating Selection</i>
SV	Sumarização de Vídeo
SVM	<i>Support Vector Machine</i>
UFPR	Universidade Federal do Paraná

SUMÁRIO

1	INTRODUÇÃO	16
1.1	OBJETIVOS	21
1.2	METODOLOGIA	22
1.3	CONTRIBUIÇÕES	24
1.4	ORGANIZAÇÃO DO TEXTO	25
2	FUNDAMENTAÇÃO TEÓRICA	27
2.1	DEFINIÇÕES BÁSICAS	27
2.2	ALGORITMOS DE PRIORIZAÇÃO	28
2.3	SUMARIZAÇÃO DE VÍDEO	28
2.4	EXPRESSÕES FACIAIS	34
2.5	BATIMENTO CARDÍACO NA ANÁLISE DE ESTRESSE	36
2.6	MOVIMENTOS DAS MÃOS RELACIONADOS AO ESTRESSE/ANSIEDADE	38
3	MAPEAMENTO SISTEMÁTICO DE LITERATURA	40
3.1	MSL 1: MÉTODOS DE SUMARIZAÇÃO DE VÍDEO EM INFORMAÇÕES AFETIVAS	40
3.1.1	Resultados do Primeiro Mapeamento Sistemático de Literatura	41
3.1.2	Análise dos Resultados do Primeiro Mapeamento Sistemático de Literatura	48
3.2	MSL 2: TÉCNICAS COMPUTACIONAIS APLICADAS NA CLASSIFICAÇÃO, ANÁLISE OU SUPORTE AO ESTRESSE/ANSIEDADE DE MOTORISTAS	49
3.2.1	Resultados do Segundo Mapeamento Sistemático de Literatura	52
3.2.2	Análise dos Resultados do Segundo Mapeamento Sistemático de Literatura	60
4	BASE DE DADOS EXPERIMENTAL	62
4.1	CRONOGRAMA	62
4.2	PROCESSO DE COLETA DE DADOS	63
4.3	ROTULAÇÃO DOS VÍDEOS	64
4.4	PRIORIZAÇÃO MANUAL DE ATIVIDADES DE DIREÇÃO	64
4.5	BASE DE DADOS DISPONIBILIZADA	65
5	DETECÇÃO DE MOVIMENTOS DAS MÃOS RELACIONADOS AO ESTRESSE E À ANSIEDADE	66
5.1	MÉTODO 1: BASEADO EM DETECÇÃO DA FACE E MORFOLOGIA MATEMÁTICA	67
5.1.1	Segmentação da Pele	67
5.1.2	Identificação de Objetos na Imagem	68
5.1.3	Detecção de Movimentos das Mãos Relacionados ao Estresse e a Ansiedade	69
5.1.4	Pós-processamento	71

5.1.5	Principais problemas	72
5.2	MÉTODO 2: BASEADO EM REDE NEURAL CONVOLUCIONAL	74
5.2.1	Modelo VGG16	74
5.2.2	Modelo VGG19	74
5.2.3	Criação dos conjuntos de teste e treinamento.	75
5.2.4	Validação cruzada.	77
5.2.5	Pós-processamento	77
5.3	EXPERIMENTOS	78
5.3.1	Métricas de Avaliação.	79
5.3.2	Resultados Experimentais.	79
5.3.3	Análise dos Resultados	82
5.4	CONSIDERAÇÕES FINAIS	83
6	PRIORIZAÇÃO DE ATIVIDADES PRÁTICAS DE DIREÇÃO	84
6.1	EXTRAÇÃO DE CARACTERÍSTICAS	84
6.2	FUNÇÃO COMPARATIVA	85
6.3	EXPERIMENTOS	86
6.3.1	Métricas de Avaliação.	86
6.3.2	Configuração Experimental.	87
6.3.3	Resultados Experimentais.	87
6.3.4	Análise dos Resultados	90
6.4	CONSIDERAÇÕES FINAIS	91
7	SUMARIZAÇÃO BASEADA NO COMPORTAMENTO DO MOTORISTA	93
7.1	TRANSFORMAÇÃO DOS DADOS	94
7.2	FERRAMENTA DE SUPORTE	95
7.2.1	Perspectiva de Visualização de Vídeos	95
7.2.2	Perspectiva de Visualização de Relatórios	97
7.2.3	Perspectiva de Visualização de Sumários.	97
7.3	EXPERIMENTOS	99
7.4	CONSIDERAÇÕES FINAIS	99
8	CONCLUSÃO	101
8.1	SUMÁRIO DAS CONTRIBUIÇÕES	101
8.2	TRABALHOS FUTUROS	102
8.3	CONSIDERAÇÕES FINAIS	103
	REFERÊNCIAS	104

1 INTRODUÇÃO

A palavra estresse é definida no dicionário Michaelis como “um estado físico e psicológico provocado por agressões que excitam e perturbam emocionalmente o indivíduo, levando o organismo a um nível de tensão e desequilíbrio, em consequência do aumento da secreção de adrenalina”. Já a ansiedade é definida como “um estado emocional frente a um futuro incerto e perigoso no qual um indivíduo se sente impotente e indefeso; sentimento e sensação de inquietude, medo ou receio” (Weiszflog, 2009).

Apesar da definição formal encontrada no dicionário, diferentes classificações são utilizadas para o estresse, tais como, o estresse agudo ou crônico. O estresse agudo é uma resposta de curto prazo a um determinado desafio — um estressor de curto prazo (Lazarus, 1966). O estresse crônico é uma resposta de longo prazo que pode surgir na vivência de experiências intensas de vida (Rabkin e Struening, 1976). Em ambas as definições, o estresse é sempre caracterizado como um estado complexo desencadeado por desequilíbrio entre as demandas colocadas sobre a pessoa e a sua capacidade de atender a essas demandas (Harrigan e O’Connell, 1996; Selye, 1956).

Já as definições de ansiedade têm em comum o fato de estarem relacionadas a um humor desagradável caracterizado por pensamentos de preocupação e medo, às vezes na ausência de real ameaça (Shin e Liberzon, 2010; Harrigan e O’Connell, 1996). Quando a ansiedade ocorre com frequência e em níveis desproporcionais ao nível de ameaça real, ela pode evoluir para uma ampla gama de distúrbios (Geddes e Andreasen, 2020).

Os termos estresse e ansiedade são frequentemente utilizados indistintamente, pois são dois estados emocionais complementares, relacionadas à resposta do organismo a demandas internas e externas (Selye, 1965), com características similares e apresentam um papel crucial na qualidade de vida das pessoas (Schneiderman et al., 2005). Dessa forma estresse e ansiedade podem impactar negativamente na saúde humana (Giannakakis et al., 2017; Padiaditis et al., 2015), em especial no comportamento de direção dos motoristas (Barnard e Chapman, 2018; Paschalidis et al., 2018).

A principal diferença entre esses estados afetivos é que a ansiedade geralmente não está diretamente ligada a pistas externas ou ameaças objetivas. Já o estresse é a resposta imediata às demandas diárias, sendo considerado mais adaptativo do que a ansiedade. No entanto, ambos envolvem sensações físicas semelhantes, tais como, aumento da frequência cardíaca, suor nas palmas das mãos e estômago embrulhado (Helgoe et al., 2005), que são comportamentos desencadeados por circuitos neuronais, quando o cérebro não consegue distinguir a diferença entre uma ameaça percebida e uma ameaça real. Essas semelhanças se estendem às expressões faciais associadas a cada um desses estados afetivos. Por essas semelhanças, nesse trabalho a ocorrência de ansiedade ou estresse é considerada como um único estado afetivo.

Destaca-se que a estratégia de considerar estresse e ansiedade como um único estado afetivo já foi explorada na literatura (Giannakakis et al., 2017; Padiaditis et al., 2015). Nesses trabalhos, a existência de estresse/ansiedade é reconhecida em filmagens de pessoas assistindo a vídeos, ou executando tarefas com o objetivo provocar os sentimentos de ansiedade/estresse ou relaxamento. Foram extraídas características do rosto (movimento da cabeça, dos olhos e da boca) e da Frequência Cardíaca (FC) dos participantes. Essas características foram utilizadas para reconhecer estresse/ansiedade automaticamente com algoritmos de aprendizagem de máquina.

Apesar de poucos trabalhos na literatura explorarem o reconhecimento em conjunto de estresse/ansiedade, muitos estudos focam na classificação desses estados afetivos de forma

individual, em diferentes situações, tais como, pessoas trabalhando (Sriramprakash et al., 2017), entrevistas de emprego (Sandulescu et al., 2015) e durante atividades práticas de direção (Rastgoo et al., 2019; Lee et al., 2020).

Com relação a condutores de carros, o estresse impacta no estado afetivo e no comportamento de direção por meio da percepção e atitude de risco (Hu et al., 2013), podendo afetar o processo de tomada de decisão, fazendo com que o condutor possa tomar decisões ruins que ele não faria em situações normais, prejudicando a sua segurança e de outras pessoas (Paschalidis et al., 2018; Starcke e Brand, 2012).

Quanto à ansiedade, especialistas apontam que as pessoas podem senti-la apenas em situações ou circunstâncias específicas de direção (tais como, condições meteorológicas e de trânsito), ou sua ansiedade pode fazer parte de um padrão mais amplo de evitação agorafóbica, ou relacionada ao trauma. Alguns motoristas descrevem leve ansiedade ou relutância em dirigir, enquanto outros relatam evitar completamente dirigir e/ou andar de veículo, e muitos estão entre esses extremos. A causa da ansiedade pode ser muito específica para alguns, por exemplo, medo de acidentes e críticas de outras pessoas, mas outros podem apresentar uma mistura dessas condições, que comumente impacta em diferentes tipos de ansiedade (Ehlers et al., 1994; Taylor et al., 2000, 2007). Pode-se destacar também que alunos que falham na prova prática de direção, tendem a apresentar uma ansiedade maior (Fairclough et al., 2006).

Além do estresse e da ansiedade, outras emoções positivas e negativas afetam o comportamento ao dirigir, de forma direta, por meio de uma direção agressiva do motorista, ou indireta, alterando os níveis de atenção (Steinhauser et al., 2018). Nesse contexto, pode-se afirmar que as emoções contribuem para acidentes de trânsito fatais, que podem levar a morte do motorista e dos demais envolvidos (Beanland et al., 2013; Brown et al., 2016).

Melhorar a segurança no trânsito é uma questão urgente para governos e indústrias no mundo todo (Rastgoo et al., 2019). Nesse contexto, se mostra importante o desenvolvimento de sistemas para automatizar a detecção e auxiliar no tratamento de estresse/ansiedade dos condutores. Destaca-se que é importante desenvolver sistemas para dar suporte a motoristas experientes, mas é particularmente relevante focar esforços em novos motoristas, que estão adquirindo habilidades de direção.

No Brasil, todo motorista deve aprender a dirigir em um Centro de Formação de Condutores (uma autoescola). Este é um momento importante na vida do aluno, pois é a primeira vez que ele irá dirigir um carro, o que pode ser uma atividade agradável ou estressante, envolvendo altos níveis de ansiedade ou nervosismo. Além disso, muitos dos pleiteantes a carteira de motorista são jovens, portanto, estão na faixa etária com maior risco de acidente de trânsito, devido à habilidade limitada de regulação emocional e falta de experiência de direção (Lee e Winston, 2016; Kaber et al., 2012).

Um aluno com estresse ou ansiedade ao dirigir pode passar nas provas práticas de direção, pois mesmo com a carga emocional envolvida, pode desempenhar as atividades técnicas necessárias para sua aprovação. Para esse motorista a atividade de direção pode ser uma atividade estressante para toda a sua vida, sem que ele nunca procure o suporte de um especialista.

Como estados afetivos, especificamente o estresse e ansiedade, podem impactar na qualidade da atividade de direção, esse perfil de motorista estará sempre correndo um risco maior ao dirigir, em relação aos outros motoristas. Nesse contexto, é importante se preocupar com o suporte prestado aos motoristas desde a sua formação, tentando, dessa forma, mitigar os problemas que esse estado afetivo pode provocar no trânsito.

Motoristas com maior nível de ansiedade ao dirigir comumente apresentam mais raiva no trânsito, menores sentimentos de segurança ao dirigir, cognições negativas e tendem a evitar dirigir. Apesar desses problemas, poucos motoristas que sofrem de ansiedade procuram ajuda, o

que indica a necessidade de desenvolver mecanismos para dar suporte a essas pessoas (Taylor, 2018).

Nesse contexto, essa tese propõe uma abordagem para detecção e análise dos estados emocionais de estresse/ansiedade de condutores de carros, através da sumarização de atividades práticas de direção. Nessa tese será considerado apenas um grupo de condutores de carros, formado por alunos de autoescolas, que serão chamados de motoristas. O método proposto tem o objetivo de viabilizar a análise de dados de um grande volume de alunos de direção, em um curto período de tempo, agilizando o suporte prestado de acordo com o nível de estresse/ansiedade. Como contribuições diretas da abordagem proposta, espera-se melhorar a segurança no trânsito e a formação de motoristas. Para tal, a abordagem combina uma técnica de priorização de atividades práticas de direção com base no nível de estresse/ansiedade e um método para produção de sumários de vídeos e relatórios, que é suportada por ferramenta gráfica para visualização desses resultados.

Como principais trabalhos relacionados, destaca-se métodos para sumarização (Cruz e Rinaldi, 2017) e visualização do estresse de motoristas (Kim et al., 2018; Rodrigues et al., 2015). Cruz e Rinaldi (2017) apresentam uma técnica de sumarização para viabilizar a análise automática da valência do motorista com base em Expressões Faciais (EF's). Nesse trabalho são selecionados quadros chaves de vídeos frontais da face de motoristas, que são utilizados para reduzir o custo do reconhecimento de emoções, partindo do pressuposto que não é necessário analisar todos os quadros para reconhecer uma emoção, mas sim, apenas os mais representativos. A diferença para a metodologia proposta nessa tese é que os sumários produzidos foram utilizados somente para melhorar o desempenho da classificação automática de EF's.

Kim et al. (2018) apresentam uma ferramenta para visualizar dados coletados de atividade prática de direção, com foco na análise do estresse do motorista em tempo real. Porém, não é apresentada nenhuma abordagem de sumarização. Já Rodrigues et al. (2015) mostram uma abordagem para visualização de um mapa de estresse dos motoristas de ônibus de uma cidade, com o uso da *Application Programming Interface* (API) do Google Maps.

A técnica de sumarização proposta nessa tese se diferencia desses trabalhos pois, além de produzir sumários de atividades práticas de direção, também apresenta uma ferramenta para visualização dos sumários, relatórios e vídeos coletados em diferentes perspectivas, explorando informações afetivas. Além disso, a abordagem também inclui uma técnica original de priorização de atividades, o que contribui para aumentar a agilidade do processo de análise desses dados.

Não foi possível identificar na literatura trabalhos focados na priorização de atividades práticas de direção com base no nível de estresse ou ansiedade de motoristas, porém, foram identificados trabalhos que apresentam métodos para o reconhecimento de níveis de estresse/ansiedade de motoristas, que são discutidos no Capítulo 3. Tais técnicas de classificação são relevantes e úteis, mas podem não ser suficientes, caso o motorista precise da ajuda de um profissional para controlar seu estresse/ansiedade. Nesse contexto, o objetivo desse trabalho é contribuir para aumentar a eficiência e eficácia no suporte prestado a esse grupo de motoristas.

Na abordagem proposta nessa tese, os parâmetros utilizados pelas técnicas de sumarização e priorização são relacionados à frequência cardíaca do motorista (em termos de Batimentos Por Minuto — BPM), EF's e movimentos referentes a comportamentos dos motoristas que podem ocorrer em situações de estresse/nervosismo. Cinco movimentos foram explorados, sendo que os quatro primeiros podem ser classificados como hábitos orais (Silva et al., 2019) e os dois últimos são gestos das mãos dos motoristas: i) pressionar os lábios; ii) morder os lábios; iii) lambe os lábios; iv) roer as unhas; e v) esfregar o rosto e/ou o cabelo.

As EF's foram coletas em termos de 18 *Action Units* (AU's) reconhecidas com a ferramenta Open Face 2.0 (Baltrusaitis et al., 2018). O BPM foi coletado com um relógio do

modelo Mi Band 5 da marca Xiaomi. Um método computacional foi proposto para reconhecer os gestos das mãos dos motoristas, que nessa tese são chamados de Movimentos das Mãos Relacionados ao Estresse e a Ansiedade - MMREA's. Os demais movimentos foram rotulados manualmente. Na literatura existem métodos baseados em Visão Computacional, para detecção de língua, boca e lábios (Nandini e Bhajantri, 2020; Tang et al., 2020; Ji et al., 2018; Vikram e Padmavathi, 2017) que podem ser adaptados para reconhecer hábitos orais do motorista.

As AU's e o BPM foram selecionados por já terem sido explorados com sucesso para análise de estresse ou ansiedade (Giannakakis et al., 2020; Rony e Ahmed, 2019; Rastgoo et al., 2018). Os demais movimentos utilizados podem ocorrer em situações de estresse/ansiedade (Silva et al., 2019; Lazari e Codinhoto, 2017; ?), por isso foram selecionados. Além disso, o conjunto de parâmetros foi selecionado pois é possível coletá-los com o uso de poucos equipamentos de baixo custo.

Além das técnicas de priorização e sumarização essa tese também apresenta experimentos conduzidos com dois métodos para detecção automática dos MMREA's. Um dos métodos é proposto nessa tese e explora o contato da região das mãos com a região do rosto do motorista, utilizando Morfologia Matemática (Serra e Soille, 2012) e detecção automática de faces (Zadeh et al., 2017). O segundo método é uma adaptação do trabalho de Masood e Alghamdi (2019), que utiliza Redes Neurais Convolucionais (RNC) para reconhecer gestos - essa técnica foi escolhida, pois não se encontrou na literatura métodos para classificar o mesmo grupo de MMREA's, e o trabalho desses autores está no estado na arte no reconhecimento de gestos das mãos dos motoristas associados à distração.

Um fator importante que foi considerado no desenvolvimento da abordagem proposta, é o fato de que ela não deve impactar no tempo das aulas práticas de direção, definido pela legislação nacional, nem gerar grandes despesas adicionais para os alunos ou para autoescolas, o que diminuiria o interesse das partes envolvidas. Além disso, é importante que a abordagem permita que especialistas analisem rapidamente os dados coletados, para que o suporte seja prestado aos alunos em tempo hábil, ou seja, enquanto o aluno está matriculado em aulas na autoescola.

Como os alunos da autoescola devem aceitar participar dos experimentos por livre e espontânea vontade, um fator considerado no planejamento da abordagem foi utilizar poucos sensores no corpo do motorista e câmeras em posições que não obstruíssem a visão dos alunos. Nesse contexto, a configuração experimental foi planejada para não gerar resistência por parte dos alunos e também para não prejudicá-los ou incomodá-los durante a atividade prática de direção.

Uma base de dados experimental foi coletada em 60 aulas práticas de direção de alunos em condições reais. Para validar adequadamente se a abordagem proposta pode ser aplicada sem impactar no tempo e escala de horários pré-definidos pela autoescola, todos os dados foram coletados em horários de aulas agendados pelo aluno, sem interferência nenhuma dos pesquisadores envolvidos nesse projeto.

Durante os experimentos, três tipos de equipamentos foram utilizados para coletar os dados: três câmeras, um relógio (para coletar dados de BPM) e um *smartphone*. As câmeras foram fixadas no carro antes do início de cada aula prática com uso de suportes com ventosas (em posições que possibilitaram registrar a face do motorista, a interação do motorista com o instrutor e o trajeto percorrido). O relógio é rapidamente colocado pelo motorista, já estando previamente sincronizado com o *smartphone*, que registra informações de BPM e GPS. Dessa forma, o processo de fixação/uso dos equipamentos e coleta de dados ocorre impactando muito pouco no tempo da atividade prática de direção.

Os dados de BPM foram utilizados pelo fato de já terem sido explorados com sucesso em trabalhos relacionados (Rastgoo et al., 2019; Sihem et al., 2019; Abayomi et al., 2017). As câmeras coletaram vídeos, que em conjunto com as coordenadas GPS, são importantes para que um especialista consiga analisar todo o contexto da atividade. Considerando que o ser humano responde de maneira estereotipada a situações de estresse (Li et al., 2020), os vídeos também foram importantes para identificar padrões de comportamentos dos motoristas, que foram rotulados e utilizados nas etapas de priorização e sumarização.

A maioria dos trabalhos relacionados à classificação de estresse/ansiedade focam em eventos estressores (que muitas vezes são adicionados de forma controlada com o uso de simuladores), criando um conjunto de referência em relação ao comportamento do motorista em situações com estresse, que permite treinar métodos para identificar esse padrão de comportamento do motorista com base em intervalos de tempo predefinidos (Li et al., 2020; Bianco et al., 2019). Outra abordagem que é utilizada na literatura é definir um trajeto padrão que será percorrido pelo motorista, com condições de estresse relacionadas ao volume de tráfego (Sihem et al., 2019; Chen et al., 2017). Também pode-se destacar trabalhos que analisam situações induzidas para provocar ansiedade em participantes que estão assistindo vídeos de atividades de direção (Lee et al., 2020).

Pensando em criar uma abordagem que possa ser aplicada sem interferir no processo que já é realizado nas autoescolas, não foi possível definir um trajeto específico com eventos estressores predefinidos nem simular condições relacionadas a esses eventos, pois em condições reais o instrutor da autoescola guia o aluno ao percurso adequado ao seu nível de habilidade, considerando condições de iluminação, clima e tráfego de acordo com o momento em que a atividade está ocorrendo.

A maioria dos trabalhos identificados na literatura analisam a presença de estresse em intervalos de tempo pré-estabelecidos, no qual o motorista dirige no mesmo tipo de percurso ou com o mesmo conjunto de estressores (Bianco et al., 2019; Abayomi et al., 2017; Chen et al., 2017). Como nessa tese, não foi possível fixar o trajeto, optou-se por analisar em conjunto os dados coletados durante toda a atividade prática de direção. Nesse contexto, a técnica de priorização proposta se diferencia da maioria dos trabalhos relacionados à detecção de estresse de motoristas. Com a abordagem utilizada nesse trabalho, é possível analisar o quanto o ato de dirigir é um evento estressante ao aluno e não apenas identificar variações no comportamento do motorista durante os principais eventos estressantes.

Além das contribuições já citadas, essa tese também apresenta Mapeamentos Sistemáticos de Literatura (MSL's) que contextualizam a abordagem proposta com o estado da arte. Com base nas contribuições apresentadas, pode-se destacar quatro fatores que diferenciam a abordagem proposta nessa tese do restante da literatura:

- Apresentar uma abordagem focada em sumarização com o objetivo de agilizar o suporte prestado a motoristas com diversos níveis de estresse. Na revisão de literatura realizada, não foram encontradas abordagens baseadas em priorização e sumarização relacionadas ao estresse de motoristas.
- O foco dado no suporte a alunos de autoescolas. Esse é um momento crítico, pois é o primeiro contato do aluno com automóveis e poucos trabalhos exploram essa problemática.
- O processo utilizado para extração de características, que considera dados coletados durante toda a atividade prática de direção (ou em grande parte dela). Nesse contexto esse trabalho se diferencia da grande maioria dos trabalhos relacionados à detecção de

estresse que têm foco em intervalos de tempo menores, ou seja, focam em identificar variações do comportamento do motorista durante um determinado evento estressor.

- Os métodos que focam na sumarização de vídeo com base em informações afetivas não apresentam ferramentas de visualização dos sumários que forneçam destaque visual às informações semânticas utilizadas. Nesse contexto a ferramenta de suporte à visualização dos sumários proposta nessa tese contribui com o estado da arte.

Detalhes sobre os objetivos dessa tese são apresentados na Seção 1.1. A metodologia utilizada para o desenvolvimento e validação da abordagem proposta é discutida na Seção 1.2. As contribuições do trabalho são ressaltadas na Seção 1.3. A Seção 1.4 apresenta a organização do restante dessa tese.

1.1 OBJETIVOS

O objetivo principal dessa tese é apresentar uma abordagem para detecção e análise de estados emocionais de estresse/ansiedade de motoristas através da sumarização de atividades práticas de direção. Com os métodos propostos, espera-se agilizar o suporte prestado aos motoristas por especialistas, tal como, profissionais da área da psicologia.

A partir do objetivo geral, foram definidos os objetivos específicos do presente trabalho, que são listados a seguir:

- analisar o comportamento dos motoristas iniciantes durante atividades práticas de direção para identificar características relacionadas a hábitos gestuais, expressões faciais e frequência cardíaca, que possam ser exploradas para o desenvolvimento da abordagem proposta.
- desenvolver técnicas que explorem informações fisiológicas e de comportamento dos motoristas para sumarizar atividades práticas de direção visando agilizar a análise do estresse/ansiedade dos motoristas.
- coletar e disponibilizar uma base de dados pública (detalhes no Capítulo 4), com informações referentes ao BPM dos motoristas e também sobre os eventos rotulados nos vídeos. A base de dados também contém a lista de atividades de direção priorizadas de acordo com o nível de estresse/ansiedade dos motoristas.
- propor uma abordagem de sumarização que combine uma técnica de priorização de atividades de direção com base no nível de estresse/ansiedade e uma técnica para produção de sumários com base no comportamento e BPM do motorista.
- desenvolver uma ferramenta computacional para dar suporte à abordagem de sumarização proposta, com a qual seja possível visualizar os vídeos, relatórios e sumários dando destaque às informações semânticas relacionadas.
- devido ao fato de não terem sido encontrados na literatura, técnicas para detecção automática dos MMREA rotulados na base de dados, têm-se como objetivo propor métodos computacionais para detectar esse grupo de movimentos com base nos vídeos coletados durante os experimentos.
- conduzir experimentos para avaliar cada um dos métodos e ferramentas que compõem a abordagem proposta.

1.2 METODOLOGIA

A abordagem proposta nesse trabalho é composta por cinco atividades principais: i) coleta de dados; ii) extração de características; iii) priorização de atividades de direção com base no nível de estresse/ansiedade dos motoristas; iv) sumarização da atividade de direção com base no comportamento e BPM do motorista; e v) análise eficiente do estresse/ansiedade dos motoristas por especialistas.

Destaca-se que todos os procedimentos executados para o desenvolvimento dessa tese foram previamente aprovados pelo Comitê de Ética em Pesquisa em Seres Humanos (CEPSH) do setor de Ciências da Saúde da Universidade Federal do Paraná (UFPR).

A Figura 1.1 ilustra a ordem sequencial de execução de cada uma das principais etapas da abordagem proposta. Ao longo do texto essas atividades serão detalhadas. A seguir, é apresentado de forma resumida a metodologia utilizada para execução de cada uma das atividades:

- **Coleta de dados:** essa atividade foi realizada em parceria com uma autoescola. Foram coletadas: coordenadas GPS, vídeos e BPM de alunos durante atividades práticas de direção. Para garantir a segurança dos alunos, durante os experimentos, todos os alunos foram acompanhados por um instrutor de direção e dirigiram veículos equipados com duplo comando dos pedais.
- **Extração de características:** três conjuntos de informações são utilizados para resumir e priorizar as atividades práticas de direção: i) ocorrência de AU's; ii) informações estatísticas sobre BPM; e iii) ocorrência de movimentos do motorista relacionados a cinco hábitos que podem ser associados com estresse/ansiedade: i) pressionar os lábios; ii) morder os lábios; iii) lamber os lábios; iv) roer as unhas; e v) esfregar o rosto e/ou o cabelo. As AU's e os cinco comportamentos do motorista foram rotuladas quadro a quadro. Um conjunto de 18 AU's foi reconhecido com a ferramenta Open Face 2.0 (Baltrusaitis et al., 2018). O BPM foi coletado com um relógio do modelo Mi Band 5 da marca Xiaomi. São apresentados dois métodos para detecção automática de gestos do motorista que envolvem o movimento das mãos (esfregar o cabelo/rosto e roer as unhas) – tais gestos são identificados pela sigla MMREA.

Para o método de sumarização foi utilizado dados sobre a ocorrência dos cinco hábitos rotulados e também sobre BPM, ambos sincronizados com as coordenadas GPS da localização do veículo. Para o método de priorização se utilizou as AU's, os MMREA'S e o BPM. Em relação aos dois métodos propostos para detecção automática dos MMREA's, um deles é baseado em uma abordagem proposta nessa tese, que utiliza Morfologia Matemática (Serra e Soille, 2012) e detecção automática de faces (Zadeh et al., 2017). Como não foram encontrados na literatura métodos para classificar o mesmo grupo de movimentos, à critério de comparação, também foram conduzidos experimento com um método recente que utiliza RNC para detectar gestos similares aos MMREA's, relacionado à distração de motoristas (Masood e Alghamdi, 2019).

- **Priorização de atividades de direção com base no nível de estresse do motorista:** Uma psicóloga especializada em atender motoristas priorizou manualmente o conjunto de atividades de direção de acordo com o nível de estresse dos motoristas, para servir como padrão. Características extraídas de cada atividade foram utilizadas para treinar uma função comparativa para identificar a atividade de direção na qual o motorista apresentou maior nível de estresse, de acordo com a classificação da especialista. O método de priorização proposto combina a função comparativa baseada em uma Rede

Neural Artificial com algoritmos clássicos de ordenação, para produzir um conjunto ordenado de atividades de acordo com a sua prioridade.

A função comparativa recebe como parâmetro duas atividades práticas de direção e extrai características dessas atividades. Tais características são referentes a dados estatísticos sobre o BPM e a porcentagem de tempo de ocorrência dos dois tipos de eventos: AU's e MMREA. Essas características são passadas como parâmetro para um modelo treinado com base em um algoritmo de aprendizagem de máquina. Com esse modelo a função retorna um valor que indica qual das atividades tem maior prioridade.

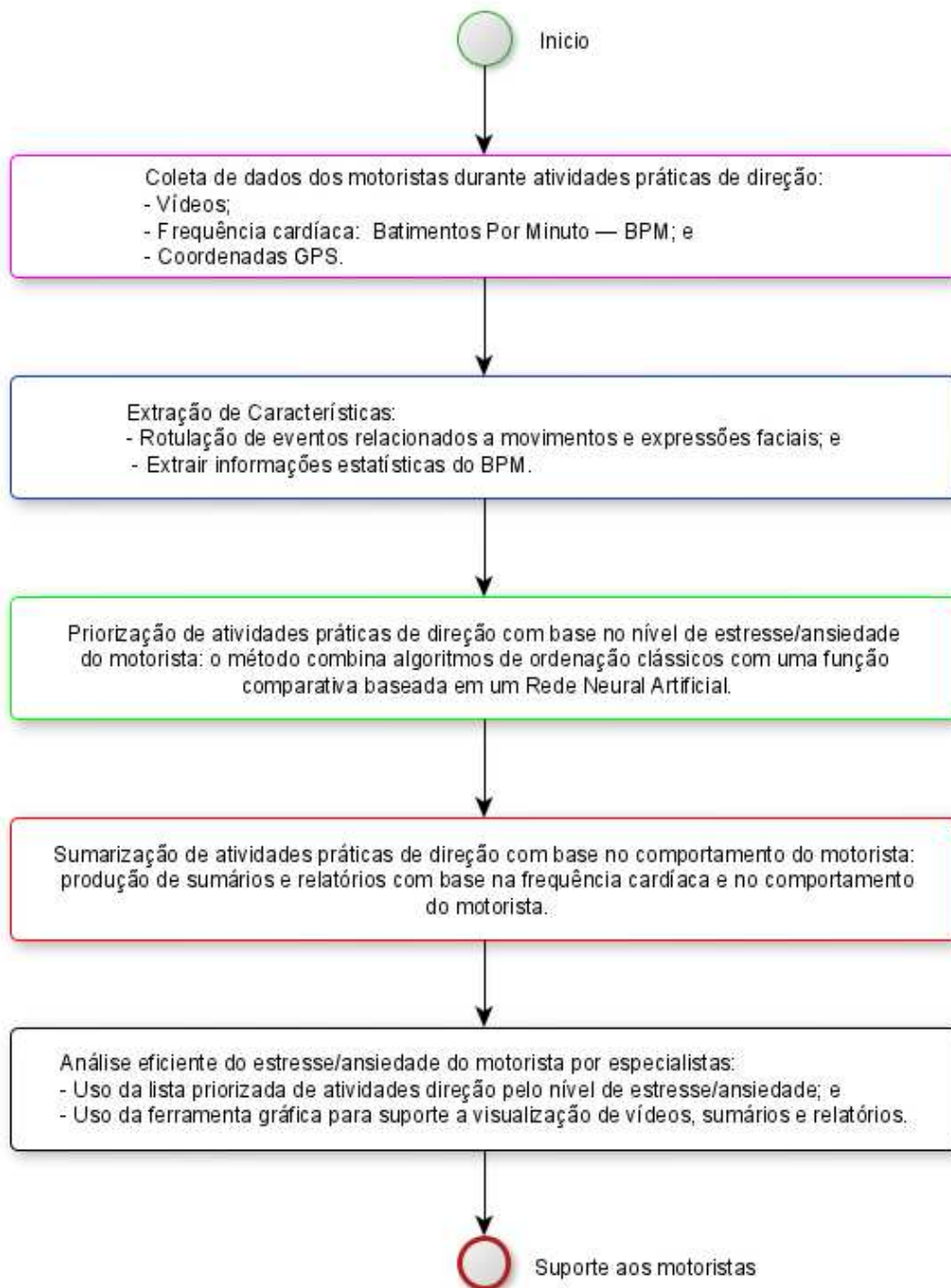


Figura 1.1: Etapas da abordagem proposta.

- **Sumarização de atividades práticas de direção com base no comportamento e BPM do motorista:** a técnica de sumarização proposta, identifica cenas dos vídeos de acordo com comportamento e BPM do motorista. Uma ferramenta foi proposta para visualizar os sumários produzidos. Essa ferramenta permite visualizar informações sobre a atividade prática de direção em três perspectivas diferentes: i) vídeos; ii) relatórios; e iii) sumários. As perspectivas são utilizadas para facilitar e organizar a análise dos dados (Ortoncelli et al., 2020).
- **Análise eficiente do estresse/ansiedade do motorista por especialistas:** a abordagem proposta pode ser utilizada por especialistas, tal como, psicólogos, que poderão compreender com mais facilidade as causas e o nível de estresse/ansiedade do motorista. A abordagem é particularmente relevante para o caso de autoescolas, onde no geral há um grande número de alunos, e seria difícil ou mesmo impraticável para um especialista analisar, em tempo hábil, os dados brutos sobre todas as atividades práticas de direção realizadas. Dessa forma, a metodologia desenvolvida viabiliza que tal análise seja executada de forma mais eficiente, em termos de demanda de tempo.

Para fundamentar a abordagem proposta em relação ao estado da arte foram conduzidos dois Mapeamentos Sistemáticos de Literatura:

- **MSL 1:** visa identificar tendências e desafios de pesquisas em relação aos métodos de Sumarização de Vídeo (SV) baseados em estados afetivos (que compreende o estado de ânimo ou humor, as emoções e os sentimentos de uma pessoa), respondendo a quatro questões de pesquisa: i) quais informações sobre estados afetivos têm contribuído positivamente com as técnicas de SV?; ii) quais as aplicações dos métodos de SV baseados em informações afetivas?; iii) como informações afetivas são adquiridas para auxiliar a SV?; e iv) como os resultados das técnicas de SV baseadas em estados afetivos são apresentadas ao usuário final?
- **MSL 2:** apresenta como técnicas computacionais são utilizadas na detecção, análise e suporte ao estresse/ansiedade de motoristas, respondendo a três questões de pesquisa: i) quais conjuntos de dados são utilizados para detecção, análise e suporte ao estresse/ansiedade de motoristas?; ii) quais as técnicas utilizadas para detecção, análise e suporte ao estresse/ansiedade de motoristas?; iii) quais técnicas de visualização e sumarização são exploradas nos trabalhos selecionados?

1.3 CONTRIBUIÇÕES

Esse trabalho contribui diretamente para melhorar a formação de motoristas, e consequentemente, para melhorar a segurança no trânsito. Com base nos objetivos apresentados na Seção 1.1, as principais contribuições dessa tese são resumidas a seguir:

- **Priorização automática de atividades práticas de direção com base no nível de estresse/ansiedade do motorista:** apresenta um método proposto para priorização de uma lista de atividades práticas de direção, que combina algoritmos de ordenação clássicos com uma função comparativa baseada em um Rede Neural Artificial para ordenar atividades de direção de acordo com o nível de estresse/ansiedade do motorista. Esse método é avaliado em relação à coesão e à similaridade com o conjunto de referência, priorizado manualmente por uma especialista. Com base na lista ordenada produzida pelo método, o suporte será prestado individualmente aos alunos de acordo

com o nível de estresse/ansiedade, aumentando a probabilidade de que os alunos com nível mais alto de estresse/ansiedade sejam atendidos de forma mais adequada.

- **Sumarização de atividades práticas de direção com base no comportamento do motorista:** é proposta uma técnica para sumarizar atividades práticas de direção com base nas informações rotuladas na base de dados experimental e no BPM do motorista. Uma ferramenta é proposta para dar suporte a visualização e análise dos sumários. Para cada fatia de informação analisada, os resultados podem ser exibidos em três perspectivas diferentes: i) vídeos, relatórios e sumários. Os resultados obtidos foram avaliados qualitativamente na opinião de profissionais vinculados à autoescola na qual a base de dados foi coletada. Com os sumários produzidos, os profissionais que irão prestar suporte ao aluno conseguem entender rapidamente o comportamento e problemas enfrentados pelo motorista, contribuindo para agilizar o suporte a ser prestado. Essa técnica pode ser combinada com o método de priorização proposto para aumentar a eficiência e eficácia do processo. Nos experimentos, o método de sumarização proposto foi validado com comportamentos relacionados a estresse e ansiedade, mas acredita-se que ele pode ser aplicado na análise de outras emoções e estados afetivos.
- **Detecção de MMREA:** são conduzidos experimentos com dois métodos de Visão Computacional para detecção de MMREA. Um desses métodos que é baseado em uma metodologia proposta nessa tese, analisa o contato da região das mãos com a região do rosto do motorista, explorando as seguintes tecnologias: Morfologia Matemática (Serra e Soille, 2012) e detecção automática de faces (Zadeh et al., 2017). O segundo método consiste em uma adaptação de um método do estado da arte, que utiliza RNC para detectar a movimentos das mãos de motoristas relacionados a distração (Masood e Alghamdi, 2019), que são similares aos MMREA's explorados nessa tese. Experimentos foram conduzidos para comparar os métodos em termos de acurácia, revocação, precisão e *F1-score*. Esses métodos contribuem para automatizar as técnicas de priorização e sumarização propostas. O método baseado na abordagem proposta nessa tese apresentou os melhores resultados em termos de revocação, acurácia e *F1-score*, porém, o tipo de vestuário ou penteado dos motoristas, podem afetar negativamente seus resultados e dificultar sua aplicação. O método baseado em RNC teve resultados melhores em termos de precisão e não sofreu interferências relacionadas ao tipo de roupa e penteado.

1.4 ORGANIZAÇÃO DO TEXTO

O restante desse texto encontra-se organizado da seguinte forma:

- **Capítulo 2 — Fundamentação Teórica:** apresenta conceitos e conhecimentos necessários para o entendimento da abordagem proposta nessa tese.
- **Capítulo 3 — Mapeamento Sistemático de Literatura:** apresenta a revisão da literatura conduzida para identificar o estado da arte relacionado a detecção, análise e suporte ao estresse de motoristas, e o estado da arte relacionado à sumarização de vídeo baseada em informações afetivas.
- **Capítulo 4 — Base de Dados Experimental:** apresenta detalhes da metodologia utilizada para produção da base de dados experimental: equipamentos utilizados, configuração do ambiente, procedimentos utilizados e também sobre a atividade de rotulagem dos dados.

- **Capítulo 5 — Detecção das Mãos Movimentos Relacionados ao Estresse/Ansiedade:** apresenta detalhes do método proposto para detecção de movimentos das mãos dos motoristas que podem ser associados a situações de estresse/ansiedade. São apresentados os experimentos conduzidos para avaliar esses métodos.
- **Capítulo 6 — Priorização de Atividades Práticas de Direção:** apresenta o método proposto para priorização de atividades práticas de direção com base no nível de estresse/ansiedade do motorista. Os resultados experimentais referentes à priorização de atividades também são apresentados no Capítulo 6.
- **Capítulo 7 — Sumarização Baseada no Comportamento do Motorista:** apresenta a técnica proposta para sumarizar uma atividade prática de direção com relação ao BPM ou informações relacionadas ao comportamento do motorista. Também são apresentados detalhes da ferramenta computacional que dá suporte à visualização dos sumários/relatórios e os resultados de uma avaliação qualitativa da ferramenta.
- **Capítulo 8 - Conclusão:** a conclusão do trabalho, resume as principais contribuições, descreve trabalhos futuros e apresenta as considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA

Esse Capítulo apresenta conceitos necessários para o entendimento da abordagem proposta nessa tese. Na Seção 2.1 é definida a nomenclatura base utilizada no texto. A Seção 2.2 contextualiza o conceito de priorização de atividades. Conceitos relacionados a técnicas de SV são apresentados na Seção 2.3. A Seção 2.4 define o conceito de *Action Units* e também apresenta as EF's e AU's utilizadas pela abordagem proposta. Trabalhos relacionados ao uso de BPM para análise de estresse e ansiedade são apresentados na Seção 2.5. A Seção 2.6 descreve os MMREA utilizados nesse trabalho e contextualiza as técnicas exploradas para reconhecer esse tipo de movimento.

2.1 DEFINIÇÕES BÁSICAS

Um conjunto de 60 atividades práticas de direção foi utilizado para conduzir os experimentos dessa tese. A base de dados experimental é denotada como: $DA = \{da_1, da_2, \dots, da_{60}\}$. Cada atividade $da_y \in DA$ é composta por três vídeos (V_a^y, V_b^y e V_c^y), dois vetores (B^y e G^y) e um valor inteiro (p^y).

Cada um dos vídeos é composto por n quadros: $V_a^y = \{fa_1, fa_2, \dots, fa_n\}$, $V_b^y = \{fb_1, fb_2, \dots, fb_n\}$ e $V_c^y = \{fc_1, fc_2, \dots, fc_n\}$. Três câmeras são utilizadas para gravar os vídeos. Cada uma das câmeras é identificada por uma letra: A, B e C. O vídeo V_a^y é gravado com a câmera A. Da mesma forma as câmeras B e C são utilizadas para gravar os vídeos V_b^y e V_c^y , respectivamente.

Após a coleta, os vídeos são manualmente sincronizados utilizando um evento chave como ponto de referência, por exemplo, o momento em que se inicia um som marcante, tal como o som de uma buzina. Com a sincronização, os quadros $fa_x \in V_a^y$, $fb_x \in V_b^y$ e $fc_x \in V_c^y$ representam o mesmo instante de tempo. Pode ocorrer variações de milissegundos na sincronização dos quadros devido a características das câmeras. Essa variação de tempo é constante, ou seja, não é cumulativa ao longo dos vídeos. Cabe destacar, que como a diferença de sincronização é muito pequena, ela não impacta no resultado das técnicas propostas nessa tese.

O vetor $B^y = \{b_1, b_2, \dots, b_z\}$ contém informações sobre o BPM. De forma análoga, o vetor $G^y = \{g_1, g_2, \dots, g_w\}$ contém informações sobre as coordenadas de GPS. Cada item $b_x \in B^y$ é uma tupla composta por dois valores: $b_x = [bpm_x, sec_x]$. O valor bpm_x representa o BPM coletado no instante de tempo sec_x . O valor sec_x representa um segundo no intervalo entre 0 e vt . Sendo vt a quantidade total de segundos do vídeo V_a^y . Por exemplo, se V_a^y tiver a duração de 10 minutos, o valor de vt será igual a 600, pois $10 \cdot 60 = 600$. Com os valores de cada item do vetor B^y é possível estimar o valor de BPM no momento em que cada um dos quadros dos vídeos foram gravados.

Com relação ao vetor G^y , cada item $g_x \in G^y$ é uma tupla composta por três itens: $g_x = [lat_x, lon_x, sec_x]$. Os valores lat_x, lon_x representam as coordenadas de latitude e longitude coletadas no instante de tempo sec_x . De forma análoga à informação de tempo utilizada na tupla b_x , na tupla g_x o valor de sec_x também está entre 0 e vt , permitindo estimar a coordenada de GPS na qual cada quadro dos vídeos foi gravado.

O valor p^y representa a prioridade de uma atividade de direção do conjunto DA em relação ao nível de estresse do motorista. O valor de p^y é um número inteiro entre 1 e z , sendo $z \leq |DA|$. Quanto maior o valor de p^y , maior é o nível de estresse do motorista.

A técnica de priorização de atividades proposta nessa tese recebe como parâmetro o conjunto DA sem o atributo p^y , que é estimado de forma automática pelo algoritmo de ordenação através da comparação de pares de atividades com uma função comparativa.

O método de sumarização apresentado nesse texto seleciona conjuntos de quadros adjacentes nos vídeos V_a^y , V_b^y e V_c^y . Cada conjunto de quadros selecionado representa uma cena do vídeo. O conjunto de cenas selecionados é utilizado como sumário da atividade de prática de direção.

2.2 ALGORITMOS DE PRIORIZAÇÃO

Um algoritmo de ordenação recebe como parâmetro uma lista de itens e coloca os seus elementos em uma certa ordem, geralmente numérica ou lexicográfica. Um algoritmo de priorização (ou ranqueamento) é basicamente um algoritmo de ordenação usado para ranquear uma lista de itens por sua prioridade.

Métodos de priorização são explorados na literatura para resolver problemas não triviais, tais como, listar pacientes com base na gravidade da doença (Mylavarapu. e Thomas, 2017), priorizar alunos com risco de não se formarem (Aguiar et al., 2015) e também em relação a problemas de Engenharia de Software: priorização de casos de teste (Lachmann et al., 2016; Bajaj e Sangwan, 2019), requisitos (Perini et al., 2012; Tonella et al., 2013) e atividades de desenvolvimento de software (Silveira et al., 2020).

Alguns desses métodos calculam individualmente a prioridade de cada item com técnicas de aprendizagem de máquina (Mylavarapu. e Thomas, 2017; Aguiar et al., 2015). Existem também algoritmos que não geram um índice com base em apenas um item. Tais métodos recebem parâmetros de um par de itens e definem o item com maior prioridade (Silveira et al., 2020; Lachmann et al., 2016; Perini et al., 2012). Também existem métodos de priorização baseados em algoritmos genéticos (Tonella et al., 2013; Bajaj e Sangwan, 2019).

A função comparativa aplicada nessa tese é baseada nas técnicas utilizadas em (Silveira et al., 2020; Lachmann et al., 2016; Perini et al., 2012). Na abordagem proposta também é utilizada uma função para comparar pares de itens com uma Rede Neural Artificial. A função comparativa é treinada com atributos relacionados ao comportamento do motorista e frequência cardíaca. Nos experimentos, essa função foi avaliada em conjunto com quatro algoritmos clássicos de ordenação: Bubble Sort, Insertion Sort, Merge Sort, e Quick Sort. Detalhes sobre esses algoritmos podem ser encontrados em (Cormen et al., 2009). Os melhores resultados foram obtidos com o algoritmo Merge Sort (MS).

2.3 SUMARIZAÇÃO DE VÍDEO

Métodos de SV têm o objetivo de criar um resumo do vídeo original, de modo que seja possível entender seu conteúdo em um curto período, devendo manter a mesma diversidade semântica do vídeo original (Fei et al., 2017), além ser atraente para a percepção humana (Yuan et al., 2011). Como o volume de vídeos gerados têm crescido exponencialmente ao longo dos anos, técnicas para resumir o conteúdo do vídeo se mostram extremamente importantes para facilitar o armazenamento eficiente, navegação rápida, indexação, recuperação rápida e compartilhamento rápido desse conteúdo. Esse contexto têm motivado pesquisas recentes relacionadas à SV (Tiwari e Bhatnagar, 2021; Hussain et al., 2021).

O método mais simples e computacionalmente eficiente para SV é a amostragem uniforme, porém tal abordagem pode não produzir sumários adequados, não selecionando segmentos importantes do vídeo original e/ou selecionando conteúdo repetido (Truong e

Venkatesh, 2007). Visando resolver tais problemas, diferentes métodos são explorados na literatura, podendo-se destacar duas abordagens que são comumente utilizadas para produção de sumários: quadros chaves (um conjunto de quadros representativos do vídeo original) e/ou cenas do vídeo (um conjunto de pequenos e relevantes pedaços de vídeo, que compõe o vídeo original).

No método proposto nessa tese, os sumários são baseados em cenas do vídeo. Uma vantagem do uso dessa abordagem, em relação ao uso de quadros chaves, é a possibilidade de incluir elementos de áudio e movimento que potencialmente aumentam tanto a expressividade quanto a informação do sumário produzido.

Essa Seção visa apresentar características dos diferentes tipos de métodos de SV, para tal serão apresentados os conceitos descritos nos trabalhos de Ferreira et al. (2016) e Truong e Venkatesh (2007). No trabalho de Ferreira et al. (2016) é apresentado um *framework* conceitual para sumarização de vídeos com base em quadros chaves. Em Truong e Venkatesh (2007), são descritos os atributos que são comumente explorados pelos métodos de SV.

A Figura 2.1 apresenta o *framework* proposto em Ferreira et al. (2016). Na primeira atividade desse *framework*, o vídeo original é dividido em cenas, utilizando um método de segmentação de vídeo. Cenas do vídeo consistem em uma coleção de segmentos de vídeo (e áudio correspondente) extraídos do vídeo original. A partir das cenas identificadas na primeira etapa, pode-se selecionar diretamente as cenas mais representativas para criar um sumário ou seguir para a segunda etapa, identificando os quadros chaves mais representativos para cada cena.

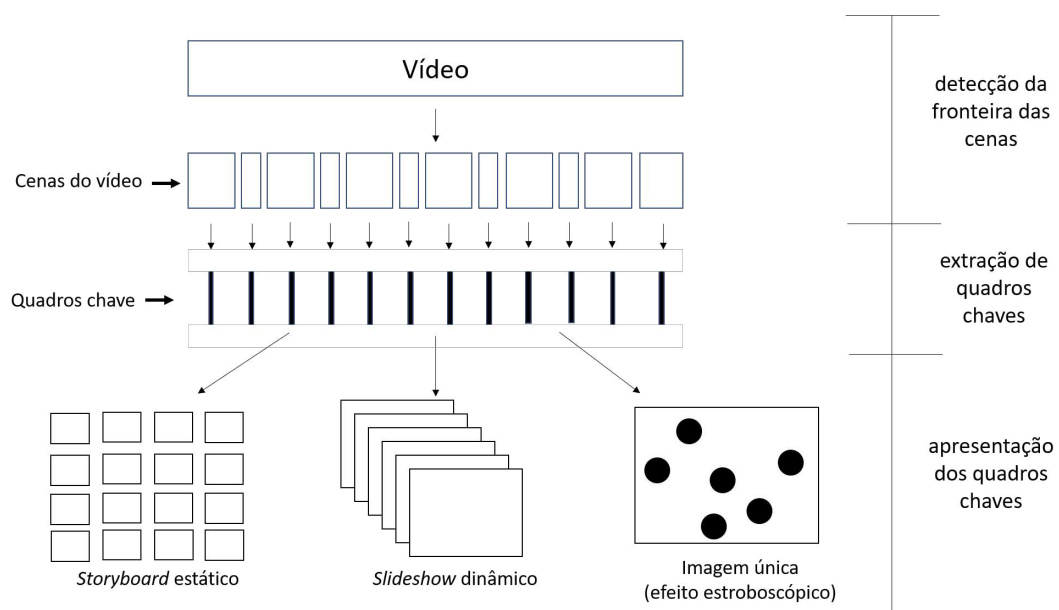


Figura 2.1: *Framework* conceitual para sumarização de vídeo com quadros chaves (Ferreira et al., 2016).

Por fim na última etapa, deve-se selecionar uma forma de exibição dos sumários produzidos. Quando se trabalha com quadros chaves, as principais técnicas são *storyboard* estático (sequência estática de quadros em ordem cronológica), *slideshow* dinâmico (apresenta os quadros selecionados, um por um, na tela) e imagem estroboscópica (elementos do fundo e do primeiro plano do vídeo são agregados em uma única imagem). Já os sumários no formato de cenas do vídeo comumente são apresentados na forma de um vídeo referente à concatenação das cenas selecionadas como mais representativas do vídeo original.

Além das principais etapas apresentadas no *framework* proposto Ferreira et al. (2016), diferentes parâmetros devem ser considerados quando se trabalha com quadros chaves ou com cenas do vídeo (Truong e Venkatesh, 2007). Tais parâmetros são apresentados respectivamente

nas Figuras 2.2 e 2.3, na forma de um diagrama. Detalhes sobre os atributos são apresentados a seguir.

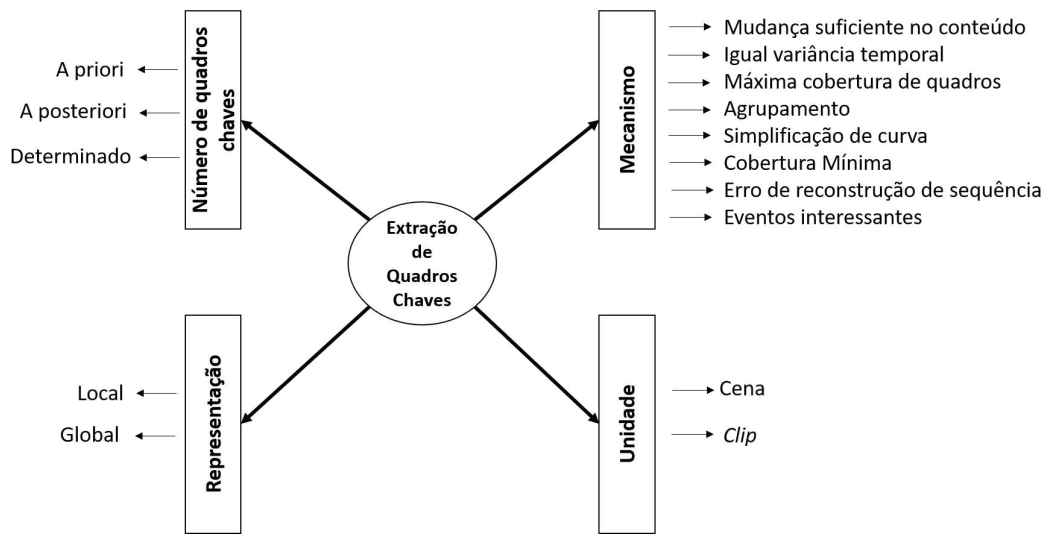


Figura 2.2: Atributos dos métodos de SV baseados em quadros chaves (Truong e Venkatesh, 2007)

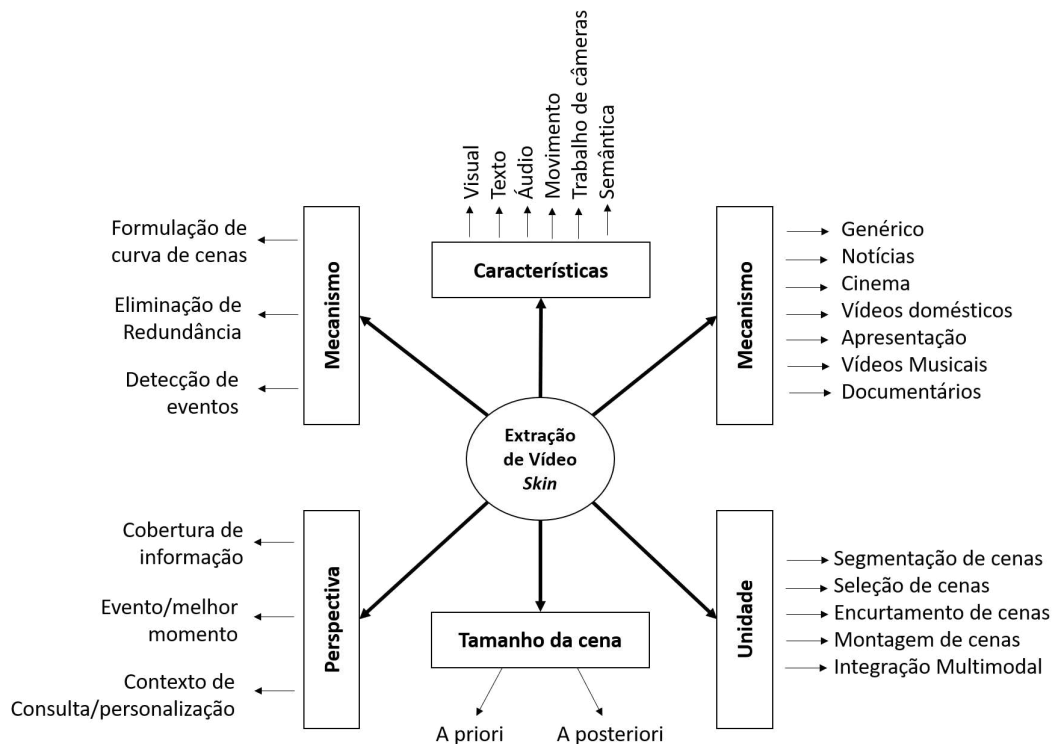


Figura 2.3: Atributos dos métodos de SV baseados em cenas do vídeo (Truong e Venkatesh, 2007)

Atributos dos métodos de SV baseados em quadros chaves:

- **Número de quadros chave:** o número de quadros chaves, seleccionados por um método de sumarização pode ser definido de três formas: i) a priori, quando o número de quadros chaves a ser produzido é definido e passado como uma restrição ao método de

SV; ii) a posteriori, quando o número de quadros chaves é desconhecido até a aplicação do método de SV terminar; e iii) determinado, que é na sua essência uma abordagem a posteriori, no entanto, o número apropriado de quadros-chave é definido antes que a extração completa seja executada;

- **Representação:** refere-se ao segmento do vídeo que é representado pelos quadros chaves, que podem representar segmentos específicos de vizinhança local de um vídeo ou representar o vídeo de forma global. A abordagem baseada na representação local preserva a informação temporal, enquanto a abordagem global pode produzir um conjunto menor de quadros chaves;
- **Unidade:** representa a unidade temporal básica que os quadros chaves devem representar. Nas técnicas de SV que representam cenas, cada segmento de vídeo é tratado de forma individual, a abordagem mais simples é selecionar o primeiro quadro, o último quadro ou o quadro do meio de uma cena. No entanto, a maioria dos métodos são adaptativos à dinâmica visual da cena. Métodos baseados em cenas podem produzir quadros chaves similares, devido à presença de cenas similares no vídeo, enquanto os métodos utilizados para representar um vídeo inteiro, podem criar um conjunto mais compacto de quadros chaves; e
- **Mecanismos computacionais:** Os métodos computacionais utilizados para produção de quadros chaves podem ser divididos por categorias. Em Truong e Venkatesh (2007), foram definidas oito categorias de mecanismos computacionais:
 1. Mudança suficiente no conteúdo: seleciona um quadro chave apenas quando o seu conteúdo divergir significativamente dos quadros anteriores ou dos quadros chaves previamente extraídos. Diferentes métricas podem ser utilizadas;
 2. Igual variância temporal: uma variante do método de mudança suficiente do conteúdo, que especifica previamente o tamanho do conjunto de quadros chaves a ser selecionado, assumindo que um bom conjunto de quadros chaves deve manter a variância temporal;
 3. Máxima cobertura de quadros: também referido como a abordagem baseada na fidelidade, este método procura selecionar quadros chaves com a maior cobertura de representação, ou seja, seleciona quadros chaves que representem a maior lista de imagens da sequência;
 4. Erro de reconstrução da sequência: também chamada de nível de reconstrução de cena, essa abordagem é baseada em mensurar a capacidade de um conjunto de quadros chaves reconstruírem o vídeo original, sendo útil em situações nas quais o número de quadros chaves é definido a priori e a progressão temporal é crucial para a aplicação;
 5. Simplificação de curva: De modo similar aos métodos de agrupamento, os métodos baseados em simplificação de curva também consideram cada quadro de um vídeo como um ponto em um espaço multidimensional. Porém, ao invés de trabalhar com agrupamentos, essa abordagem procura por um conjunto de pontos, de modo que a remoção dos pontos restantes modifica, pelo menos, a forma da curva que conecta todos os pontos por meio de sua ordenação temporal. Nesse sentido, a abordagem é similar aos métodos que utilizam Erro de reconstrução da sequência, se diferenciando pelo fato de que os pontos na curva não são necessariamente

igualmente espaçados pelo índice do quadro, sendo assim, uma modelagem explícita do erro entre a curva final e a trajetória original do quadro não é necessária;

6. Agrupamento: trata os quadros de um vídeo como pontos no espaço de recurso e utiliza um algoritmo de agrupamento, para identificar os quadros chaves. Nessa abordagem quatro etapas são comumente implementadas: i) pré processamento dos dados de entrada, que visa melhorar a eficácia e eficiência do processo de agrupamento; ii) agrupamento dos dados; iii) filtrar agrupamentos; e iv) extração dos pontos representativos de cada agrupamento;
7. Menor correlação entre quadros chave: tem o objetivo de selecionar quadros que são dissimilares entre si. Métodos que seguem essa abordagem, são baseados na premissa de que os quadros chaves devem ter a correlação mínima entre seus elementos, sendo que, em alguns casos é considerada apenas a correlação entre os elementos sucessivos; e
8. Eventos de interesse: diferentemente das abordagens anteriores, que se concentram na extensão e o equilíbrio da cobertura visual do conjunto de quadros chave, os métodos baseados em eventos de interesse focam em identificar quadros que são semanticamente importantes.

Atributos dos métodos de SV baseados em cenas do vídeo:

- **Tamanho da cena:** o tempo de duração do sumário produzido pode ser definido de duas formas: a priori (definido antes da execução do algoritmo de sumarização) ou a posteriori (definido de acordo com as características do vídeo);
- **Domínio do vídeo sumarizado:** existem técnicas de SV de vídeo que podem ser aplicadas em todos os tipos de vídeos (Dhagdi e Deshmukh, 2012; Varghese e Nair, 2015), porém grande parte dos métodos de sumarização de vídeo têm foco em um domínio específico de vídeos, tais como, esportes (Tejero-de Pablos et al., 2018), notícias (Ide et al., 2017), vídeos médicos (Muhammad et al., 2020), vigilância (Varghese e Nair, 2018); vídeos domésticos (Peng et al., 2011) e vídeos de atividades práticas de direção (Ortoncelli et al., 2020), que são o foco dese trabalho.
- **Processo:** o processo de geração de seleção de cenas do vídeo apresentado em (Truong e Venkatesh, 2007) conta com cinco etapas genéricas: segmentação de cenas, seleção de cenas, encurtamento de cenas, integração multimodal e montagem de trechos. Na prática, uma técnica de SV baseada em cenas do vídeo pode pular algumas dessas etapas ou combiná-las em diferentes variações, porém o princípio e a essência permanecem os mesmos;
- **Perspectiva preservada:** comumente para produção de sumários baseados em cenas do vídeo decide-se por preservar uma determinada perspectiva do vídeo. Diferentes perspectivas podem criar sumários diferentes, adequados para diferentes domínios e aplicações. Em (Truong e Venkatesh, 2007) são apresentadas três perspectivas que podem ser utilizadas: i) cobertura de informação, que consiste em eliminar a redundância e encurtar segmentos do vídeo, sem prejudicar o entendimento do telespectador sobre o conteúdo do vídeo; ii) eventos importantes/interessantes, consiste em preservar os eventos de interesse no vídeo; e iii) contexto de pesquisa e personalização, no qual os sumários são gerados de acordo com a especificação do usuário, na forma de uma consulta.

- **Mecanismos computacionais:** Os métodos computacionais utilizados para produção de sumários baseados em cenas do vídeo podem ser divididos em três grupos (Truong e Venkatesh, 2007):

1. **Eliminação de Redundância:** considerando que redundância são partes do vídeo que contém informações que já são conhecidas, o encurtamento de cenas é uma parte essencial da eliminação de redundância. O processo de eliminação de redundância pode ocorrer de diferentes formas, tais como, selecionando apenas quadros contínuos que maximizam a semelhança média com toda a sequência (Cooper e Foote, 2002), com processo de seleção de cenas via agrupamento (Ngo et al., 2003) ou com a seleção de partes de interesse de uma cena (Ma et al., 2002).
2. **Deteção de eventos e melhores momentos:** Consiste em identificar eventos importantes e os limites de ocorrência dos eventos identificados.
3. **Formulação de curva *skimming*:** Produz sumários com base em um valor limite aplicado à curva de perspectiva, tal como é ilustrado na Figura 2.4. Destaca-se que outros métodos podem ser aplicados para SV com base na curva produzida, tal como, a busca por máximos locais, se o elemento de interesse é tratado como um conceito temporalmente relativo (Ma et al., 2002).

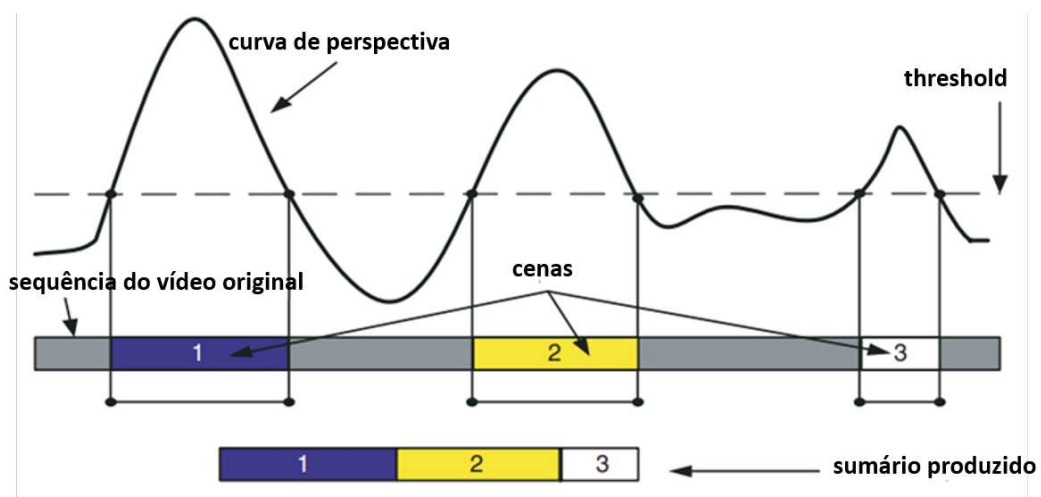


Figura 2.4: Geração de sumários baseados em curva de *skimming* (Truong e Venkatesh, 2007)

- **Características:** diferentes características podem ser exploradas para produção dos sumários, tais como, características visuais (cores, bordas e texturas), texto, áudio, dinâmica visual, movimento de câmera e informações semânticas.

As características exploradas pelos métodos de SV podem ser categorizadas em três grupos (Money e Harry, 2008): i) interna, informações obtidas diretamente do vídeo; ii) externa, que analisa informações que não estão no vídeo (no contexto de emoções, normalmente são utilizadas informações referentes aos telespectadores); e iii) híbridas, que combinam informações externas e internas.

A abordagem proposta nesse trabalho produz sumários de vídeos de atividades práticas de direção desenvolvidos tendo com perspectiva preservar eventos importantes. O tamanho das cenas dos sumários produzidos é definido a posteriori. Quanto ao mecanismo computacional

a abordagem proposta é baseada na formulação de curva *skimming*. Quanto às características exploradas, foram utilizados dados extraídos do vídeo, referentes a presença de AU's e MMREA's. Também foram exploradas características externas ao vídeo, referentes ao BPM do motorista.

2.4 EXPRESSÕES FACIAIS

Além da voz, diferentes formas de comunicação são exploradas pelos humanos, entre elas se destacam as EF's, movimentos das mãos e postura corporal. Em diferentes contextos, EF's podem ter funções de comunicação distintas, como sinalizar a tomada de turnos, transmitir informações biométricas, expressar a intensidade do esforço mental e indicar emoções (Corneanu et al., 2016). Pode-se destacar, trabalhos que apontam que EF's são um importante biomarcador relacionado a estresse/ansiedade (Faghel-Soubeyrand et al., 2020; Mayo e Heilig, 2019).

Duas abordagens são as mais utilizadas na literatura para definir a emoção humana representada por uma EF: dimensional (ou contínuo) e categórica (ou discreta) (Corneanu et al., 2016). Trabalhos relacionados a estados emocionais dimensionais ou contínuos utilizam respostas fisiológicas de pessoas para determinar o estado afetivo (Scheirer et al., 2002), considerando principalmente duas dimensões: Valência e Excitação (VE). Valência é o nível de aversão ou atração que um usuário ou telespectador sente em relação a um determinado estímulo. Excitação é o nível de intensidade de uma emoção decorrente de um estímulo específico (Money e Agius, 2013). Além de VE, outras dimensões contínuas também são exploradas, tais como controle (Hanjalic, 2005) e prazer (Arifin e Cheung, 2007).

A abordagem discreta é baseada em um conjunto finito de emoções, que na maioria dos trabalhos é composto pelas emoções primárias, que foram popularizadas por Paul Ekman (Ekman, 1994), que considera emoções como estados discretos que estão associados com expressões faciais, considerando a existência de seis emoções primárias: i) raiva; ii) tristeza; iii) nojo; iv) felicidade; v) medo; e vi) surpresa. Além dos conjuntos primários de emoções propostos por Ekman, outros grupos de emoções também são explorados na literatura (Ortony e Turner, 1990), além do uso de emoções secundárias e terciárias (Shaver et al., 1987).

Uma das principais abordagens utilizadas na literatura para o reconhecimento de EF's é o Sistema de Codificação de Ação Facial (tradução do termo em inglês *Facial Action Coding System* - FACS) (Ekman et al., 2002). De acordo com o FACS uma EF pode ser definida por um conjunto de ações individuais dos músculos faciais, cada um desses movimentos é uma unidade de ação (em inglês: *Action Unit* – AU).

O conceito de AU foi criado por Carl-Herman Hjortsjö, que definiu um conjunto de 23 AU's em 1970. No ano de 2002, Paul Ekman e Wallace Friesen, expandiram esse trabalho propondo um conjunto de 44 AU's (Ekman et al., 2002).

Desde então, o reconhecimento automático de FACS e AU's é um desafio que tem motivado vários pesquisadores, incluindo trabalhos recentes que contribuiram para melhorar a classificação automática de AU's (Batista et al., 2017; Martinez et al., 2017; Albiero et al., 2018; Liu et al., 2019; Baltrusaitis et al., 2018). Isso é positivo para a abordagem proposta neste trabalho, pois, quanto mais preciso for o método de reconhecimento da AU, mais precisos serão os resultados da abordagem proposta nessa tese.

Deve-se destacar que já foram observadas diferenças na frequência das AU's realizadas em cenários de direção de veículos e cenários de não direção (Li et al., 2021). Por esse motivo é importante desenvolver métodos para análise de estresse/ansiedade de motoristas focados nas EF's faciais dos condutores.

Além de (Li et al., 2021), outros autores analisaram AU's e FACS de motoristas (Ihme et al., 2018; Yüce et al., 2016; Moriyama et al., 2012). Em Ihme et al. (2018) é investigada a

ocorrência de FACS em estudo simulado relacionado à frustração no trânsito. Os resultados indicam que motoristas mostram mais atividade na área da boca quando estão frustrados em comparação com quando eles não são. Em Yüce et al. (2016) são analisadas AU's produzidas por motoristas em simuladores de direção, comparando condições normais de direção com condições de alta carga cognitiva, induzida por uma tarefa secundária a direção. Com base em características de 14 AU's um modelo foi treinado e alcançou um f1-score de 95% na detecção de atividades de alta carga cognitiva. Já o trabalho de Moriyama et al. (2012) investiga humores agressivos, como irritação e tensão, em termos de AU's, indicando que à medida que o nível de irritação dos sujeitos aumenta, se observa de forma mais evidente as AU's 4, 9, 14, 24 e 28.

Também pode-se destacar técnicas recentes que analisam o impacto das AU's em situações de estresse não relacionadas ao trânsito (Zhang et al., 2020; Lombardi e Marcolin, 2020; Viegas et al., 2018; Jáuregui et al., 2017). Zhang et al. (2020), propõe um modelo que analisa separadamente AU's e movimentos do corpo e então une esses resultados para determinar o nível de estresse de uma pessoa. Lombardi e Marcolin (2020) classifica automaticamente o estresse induzido em pessoas (por jogos, experiências assustadoras e entrevistas) com modelos de *Deep Learning* e algoritmos de aprendizagem de máquina, que são utilizados para identificar micro movimentos das EF's, que são registradas em 2D e 3D.

Em (Viegas et al., 2018) são propostos modelos treinados para reconhecer atividades estressantes realizadas por pessoas enquanto digitam textos. Os modelos utilizam características de 19 AU's, com uma precisão de classificação variando de 29% a 91%. Já Jáuregui et al. (2017) classifica situações de estresse de alunos de doutorado com base na FC, AU's e movimentos da cabeça e do corpo. Os dados foram coletados dos alunos durante atividades de orientação, apresentação, questionamentos e preparação da apresentação. Foram utilizadas imagens 2D e 3D para reconhecer as EF's e os demais movimentos. As características coletadas se mostraram importantes para estabelecer uma correlação com o estresse, permitindo alcançar uma acurácia de 85,42%.

Cabe destacar que os trabalhos de Zhang et al. (2020) e Jáuregui et al. (2017) combinam movimento corporal com AU's para classificar o estresse. As características relacionadas a gestos e movimentos serão descritas na Subseção 2.6.

Além desses trabalhos, também existem trabalhos que utilizam AU's para reconhecer o estresse e a ansiedade como um único estado afetivo (Giannakakis et al., 2020; Gavrilescu e Vizireanu, 2019). Em Giannakakis et al. (2020), a intensidade de diferentes AU's foi explorada para identificação automática de quatro estados de tensão. Quatro estados de estresse/ansiedade foram induzidos: i) exposição social; ii) lembrança emocional; iii) tensão mental; e iv) estímulos de vídeos estressantes. O trabalho concluiu que AU's apresentam padrões diferentes quando as pessoas estão em um estado de estresse em relação a um estado neutro. Na maioria dos casos, a intensidade das AU's foi significativamente maior durante as condições de estresse, indicando um rosto mais expressivo.

Em Gavrilescu e Vizireanu (2019) foram conduzidos experimentos com 128 pessoas escolhidas de modo a analisar diferentes gravidades de estresse, depressão e ansiedade. Os participantes assistiram a vídeos para induzir emoções. Um conjunto de AU's foi coletado desses participantes. Uma Rede Neural do tipo *Optimal Feed Forward* foi utilizada para analisar uma matriz dos níveis de intensidade da AU. Resultados significativos foram obtidos em relação à previsão de depressão, ansiedade e estresse.

Para validar o método de priorização apresentado nessa tese foram utilizadas 18 AU's reconhecidas pela ferramenta Open Face 2.0 (Baltrusaitis et al., 2018): AU's 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 25, 26, 28 e 45. AU's foram utilizadas pois existe na literatura técnicas automatizadas para reconhece-las e também por existirem evidências de que esses movimentos

contribuem para a detecção do estresse e da ansiedade (Giannakakis et al., 2020; Zhang et al., 2020; Jáuregui et al., 2017).

A Figura 2.5 apresenta exemplos das 18 AU's utilizadas. O conteúdo dessa figura é baseado em Farnsworth (2016), que também apresenta gifs animados relacionados a cada uma das AU's.

2.5 BATIMENTO CARDÍACO NA ANÁLISE DE ESTRESSE

A frequência cardíaca de uma pessoa consiste no número de pulsações do coração por uma unidade de tempo. Nesse trabalho a FC é mensurada pelo número de Batimentos Por Minuto (BPM).

Na literatura, a FC é um parâmetro explorado como um indicador de estresse/ansiedade (Rastgoo et al., 2018; Goessl et al., 2017; Soman et al., 2013). Diferentes características podem ser apontadas para justificar essa escolha:

- Em situações de estresse, as glândulas suprarrenais do corpo humano liberam o hormônio adrenalina no organismo. Existem estudos que indicam que esse hormônio potencializa a variação da FC (Rony e Ahmed, 2019).
- Vários trabalhos exploram a reação cardíaca, como um sintoma fisiológico de ansiedade (Wen et al., 2018; Thurston et al., 2013);
- A FC tem uma relação muito próxima com a atividade neural, podendo se alterar de acordo com o estado emocional (Kubota et al., 2001; McCraty et al., 1995); e
- Uma FC mais baixa geralmente está ligada a relaxamento, porém ela aumenta com a presença de estímulos emocionais ou esforço mental (Katsis et al., 2011).

Por esses motivos, optou-se por utilizar dados relacionados à frequência cardíaca na abordagem proposta nessa tese. Nos experimentos, essa informação foi coletada com uso de um relógio do modelo Mi Band 5 da marca Xiaomi. Esse equipamento, coleta dados de BPM com uma técnica de medição óptica chamada de fotopletismografia. Essa técnica é baseada em dois elementos básicos: uma fonte de luz infravermelha e um foto receptor que detecta variações na intensidade da luz refletida — que ocorrem por causa de variações volumétricas do sangue. Essa técnica é amplamente utilizada por equipamentos chamados de oxímetros de pulso (Shelley et al., 2001), como é o caso do sensor do relógio utilizado nos experimentos.

Com o uso desse relógio é possível coletar o BPM do motorista de forma não invasiva, sem nenhuma infiltração ou corte no motorista. Destaca-se que o uso de sensores vestíveis se mostra uma tendência em pesquisas relacionadas ao gerenciamento de estresse (Carneiro et al., 2017).

Atualmente se pode observar um aumento na quantidade de trabalhos que relacionam a frequência cardíaca com atividades cognitivas (Peng e Nagao, 2018). Correlações do BPM com o nível de estresse também vêm sendo exploradas em trabalhos que analisam variações de frequência cardíaca em situações que visam induzir estresse nos participantes dos experimentos (Pereira et al., 2017; Masood e Alghamdi, 2019).

Além dos trabalhos apresentados nessa Seção, também existem na literatura, técnicas que exploram informações relacionadas à frequência cardíaca para analisar ou detectar estresse/ansiedade em motoristas. Tais trabalhos são apresentados no Capítulo 3.

AU 01 - Levantador de sobrancelha Interna Músculos Faciais: <i>Frontalis, pars medialis</i>	
AU 02 - Levantador de sobrancelha Interna Músculos Faciais: <i>Frontalis, pars lateralis</i>	
AU 04 – Abaixador de sobrancelha Músculos Faciais: <i>Depressor glabellae, Depressor supercilii, Corrugator supercilii</i>	
AU 05 – Levantador de pálpebra superior Músculos Faciais: <i>Levator palpebrae superioris</i>	
AU 06 – Levantador de bochechas Músculos Faciais: <i>Orbicularis oculi, pars orbitalis</i>	
AU 07 – Apertador de pálpebra Músculos Faciais: <i>Orbicularis oculi, pars palpebralis</i>	
AU 09 – Enrugador de nariz Músculos Faciais: <i>Levator labii superioris alaquae nasi</i>	
AU 10 – Levantador de lábio superior Músculos Faciais: <i>Levator Labii Superioris, Caput infraorbitalis</i>	
AU 12 – Puxador do canto do lábio Músculos Faciais: <i>Zygomatic Major</i>	
AU 14 – Fazedor de covinhas Músculos Faciais: <i>Buccinator</i>	
AU 15 – Depressor de canto de lábio Músculos Faciais: <i>Depressor anguli oris (Triangularis)</i>	
AU 17 – Levantador de queixo Músculos Faciais: <i>Mentalis</i>	
AU 20 – Esticador de lábio Músculos Faciais: <i>Risorius</i>	
AU 23 – Endurecedor de lábio Músculos Faciais: <i>Orbicularis oris</i>	
AU 25 – Separador de lábios Músculos Faciais: <i>Depressor Labii, Relaxation of Mentalis (AU17), Orbicularis Oris</i>	
AU 26 – Queda de mandíbula Músculos Faciais: <i>Maseter; Temporal and Internal Pterygoid relaxed</i>	
AU 28 – Sucção de lábios Músculos Faciais: <i>Orbicularis oris</i>	
AU 45 – Piscar Músculos Faciais: <i>Relaxation of Levator Palpebrae and Contraction of Orbicularis Oculi, Pars Palpebralis</i>	

Figura 2.5: AU's utilizadas pela técnica de priorização - baseado em Farnsworth (2016).

Por fim, cabe destacar também que além da técnica utilizada nesse trabalho, existem outras abordagens para detecção da frequência cardíaca, tal como métodos baseados em vídeos (Hassan et al., 2017). Esse tipo de técnica não foi explorado nessa tese, pois, sua precisão é inferior a dos métodos baseados em fotopletismografia.

2.6 MOVIMENTOS DAS MÃOS RELACIONADOS AO ESTRESSE/ANSIEDADE

As emoções humanas podem ser associadas com diferentes movimentos do corpo. Diferentes características desses movimentos podem ser utilizadas para inferir emoções, tais como: amplitude, velocidade e fluidez do movimento, além de se tentar reconhecer gestos que representam emoções específicas (Li et al., 2016; Castellano et al., 2007).

Existem situações de estresse, ansiedade e nervosismo, nas quais o corpo humano pode responder com hábitos não normais de movimento (Azrin e Nunn, 1973). Nessa tese foi rotulado um conjunto de dois movimentos das mãos dos motoristas que podem ser associados ao estresse: i) esfregar o rosto e/ou cabelo; e ii) roer as unhas. A Figura 2.6 apresenta exemplos desses movimentos.



Figura 2.6: Exemplos de movimentos das mãos relacionados ao estresse.

Movimentos das mãos já foram explorados na literatura para classificar automaticamente o estresse de motoristas, porém nesses trabalhos, os movimentos foram detectados com o uso de sensores vestíveis na forma de luvas (Lee et al., 2016; Lee e Chung, 2016) ou embutidos no volante (Paredes et al., 2018). Em (Ortoncelli et al., 2020) esses movimentos foram explorados com base no vídeo, mas foram apenas rotulados manualmente.

Além de trabalhos com foco em motoristas, movimentos corporais também já foram explorados para classificar estresse em outros contextos (Zhang et al., 2020; Jáuregui et al., 2017). Em Zhang et al. (2020) são analisadas AU's e movimentos do corpo para reconhecer o nível de estresse de pessoas assistindo a vídeos, que foram selecionados para induzir esse estado afetivo. Apesar desse trabalho não deixar claro todos os movimentos, cita-se no texto os gestos de tocar a orelha e esfregar o cabelo. Além disso, o texto deixa claro que combinar AU's com movimentos do corpo contribui para a detecção do estresse.

No trabalho de Jáuregui et al. (2017) é analisado o estresse de alunos de doutorado durante atividades de orientação, apresentação, questionamentos e preparação da apresentação. O estresse é reconhecido com base na FC, AU's e movimentos da cabeça e do corpo. Nesse trabalho, os seguintes movimentos são reconhecidos em imagens 2D e 3D: movimentos da face e orientação da cabeça, o nível de contração do corpo, quantidade de movimento do corpo e variabilidade do centro de gravidade.

Apesar de não terem sido encontrados métodos que explorem Visão Computacional para reconhecer automaticamente os MMREA's explorados nessa tese, existem trabalhos que utilizam imagens para detectar movimentos das mãos dos motoristas relacionados a distração, tais como, o uso de celular, fazer a maquiagem, alcançar objetos, beber, operar o rádio e conversar com o passageiro (Masood et al., 2020; Craye e Karray, 2015). Em (Craye e Karray, 2015) são utilizadas imagens RGB-D, já em (Masood et al., 2020) são utilizadas imagens 2D classificadas com um RNC.

Cabe destacar que existem poucos trabalhos relacionados à detecção de movimentos das mãos para a classificar eventos de distração, porém a atenção de motoristas é um tema explorado em vários trabalhos, comumente envolvendo os movimentos dos olhos e a posição da cabeça (Le et al., 2020; Masood e Alghamdi, 2019).

Não foi encontrado na literatura nenhum método com foco nos MMREA's que são explorados nessa tese, porém, entre as técnicas existentes na literatura, o trabalho de Masood et al. (2020) se mostrou relacionado por explorar movimentos com características semelhantes (movimentos nos quais o motorista leva a mão próxima ao rosto). Por isso, além de propor um método para detecção de MMREA's, nos experimentos também é avaliada uma adaptação do método proposto por Masood et al. (2020).

O método proposto para detecção de MMREA é baseado em Visão Computacional, explorando detecção de face (Baltrusaitis et al., 2018; Zadeh et al., 2017) e Morfologia Matemática (Serra e Soille, 2012). O trabalho de (Masood e Alghamdi, 2019), utiliza Redes Neurais Convolucionais para detectar os gestos. Experimentos foram conduzidos para validar ambos os métodos com os vídeos da base de dados experimental. Detalhes sobre ambos os métodos e os experimentos realizados são apresentados no Capítulo 5.

Cabe destacar que nos experimentos realizados não foi considerada a amplitude ou intensidade dos movimentos. Como o motorista está constantemente com as mãos no volante e os movimentos ocorrem em um espaço restrito, eles comumente não se diferenciam em termos de intensidade/amplitude.

O método baseado em Morfologia Matemática apresentou os melhores resultados de revocação (0,806), acurácia (0,997) e *F1-score* (0,806). Quanto a precisão, os melhores resultados foram obtidos com o método baseado em RNC, que apresentou uma precisão de 0,992. Destaca-se que a acurácia dos dois métodos foi similar.

O método que utilizou Morfologia Matemática se mostrou sensível a determinados tipos de vestuário e penteados, que são descritos na Subseção 5.1.5. O segundo método não teve problemas com os mesmos tipos de roupas e penteados, apresentando resultados similares para todos os vídeos analisados. Como é difícil controlar essas variáveis em condições reais, recomenda-se o uso do método baseado em RNC.

3 MAPEAMENTO SISTEMÁTICO DE LITERATURA

Um Mapeamento Sistemático de Literatura (MSL) é método de revisão de literatura criado com o objetivo de identificar, analisar e interpretar as evidências disponíveis relacionadas a um tópico de pesquisa particular ou com um determinado fenômeno de interesse. Sendo um dos principais meios para sumarizar evidências de pesquisa (Felizardo et al., 2017; Kitchenham et al., 2004).

Nesse Capítulo são apresentados dois MSL's, que foram conduzidos de acordo com os cinco passos propostos por (Petersen et al., 2008): i) definição das Questões de Pesquisa (QP's); ii) pesquisa pelos estudos primários relevantes; iii) seleção dos estudos primários; iv) categorização dos artigos; e v) mapeamento e extração de informações que respondam às QP's estabelecidas.

O primeiro mapeamento (Seção 3.1) visa determinar quais as tendências e desafios de pesquisa em relação aos métodos de SV baseados em informações afetivas. O segundo mapeamento (Seção 3.2) tem como objetivo identificar como técnicas computacionais são aplicadas para a detecção, análise ou suporte ao estresse/ansiedade de motoristas.

Existem trabalhos similares ao segundo mapeamento (Zepf et al., 2020; Rastgoo et al., 2018). Em (Zepf et al., 2020) o foco foi dado na literatura relacionada ao reconhecimento de emoções de motoristas. Já em Rastgoo et al. (2018) é apresentada uma análise crítica de métodos para detecção de estresse de motoristas com base em características multimodais. Pode-se destacar também o trabalho de Arif et al. (2020), que apresenta uma revisão de técnicas baseadas em aprendizagem de máquina para classificação de ansiedade. Devido à similaridade dos trabalhos, além de aplicar os critérios de inclusão e exclusão nos trabalhos obtidos com a *string* de busca, esses critérios também foram aplicados aos artigos referenciados por Zepf et al. (2020); Arif et al. (2020) e Rastgoo et al. (2018).

3.1 MSL 1: MÉTODOS DE SUMARIZAÇÃO DE VÍDEO EM INFORMAÇÕES AFETIVAS

Para identificar as tendências e desafios de pesquisa em relação aos métodos de SV baseados em informações afetivas, foram criadas quatro QP's:

- QP_1 : quais informações sobre estados afetivos têm contribuído positivamente com as técnicas de SV?;
- QP_2 : quais as aplicações dos métodos de SV baseados em informações afetivas?;
- QP_3 : como informações afetivas são adquiridas para auxiliar a SV?; e
- QP_4 : como os resultados das técnicas de SV baseadas em estados afetivos são apresentadas ao usuário final?

Com base nas QP's foram definidos os Critérios de Inclusão (CI's) e os Critérios de Exclusão (CE's), com os quais é possível identificar quais estudos contribuem na resposta das questões de pesquisa. Foram definidos os seguintes CI's:

- CI_1 : Se vários artigos do mesmo autor principal apresentam estudos similares, somente o mais recente é incluído;

- CI_2 : Se houver versões do mesmo artigo, uma resumida e outra completa, apenas a completa deve ser incluída; e
- CI_3 : Artigos que descrevem o uso de técnicas da SV baseadas em informações afetivas devem ser incluídos.

Também com base nas QP's, foram definidos os seguintes CE's:

- CE_1 : Estudos que não estejam na língua inglesa serão excluídos;
- CE_2 : Relatórios técnicos, documentos que estão disponíveis na forma de resumos ou apresentações e também revisões de literatura secundárias (ou seja, revisões e mapeamentos sistemáticos da literatura) serão excluídos;
- CE_3 : Trabalhos que realizem a classificação de emoções ou estados afetivos apenas após a sumarização ou segmentação do vídeo serão excluídos;
- CE_4 : Trabalhos que não apresentem técnicas de sumarização de vídeo baseadas em estados afetivos serão excluídos.

Para responder às questões de pesquisa a seguinte *string* de busca foi utilizada: ((“*video summarization*” OR “*movie summarization*” OR “*video abstraction*” OR “*movie abstraction*” OR “*video highlights*” OR “*movie highlights*” OR “*video indexing*” OR “*movie indexing*”) AND (“*facial expression**” OR “*emotion**” OR “*affect**”)).¹

Os termos “*video indexing*”, “*movie indexing*”, apesar de não serem essencialmente técnicas de SV, foram adicionados a *string* de busca, devido ao fato de serem conceitos relacionados. Métodos de indexação visam adicionar índices ou rótulos a um determinado conteúdo, visando facilitar o acesso e a visualização. A indexação de vídeo pode ser uma etapa que antecede a SV.

Destaca-se que a *string* de busca apresentada foi adaptada de acordo com as características de cada uma das bases pesquisadas. As buscas foram conduzidas em bases de dados eletrônicas de grande importância para a Ciência da Computação (Dyba et al., 2007): ACM Digital Library, Science Direct, Engineering Village, IEEEExplore, Scopus, Springer e Web of Science.

Os resultados obtidos com esse MSL são apresentados na Subseção 3.1.1. Uma análise desses resultados é apresentada na Subseção 3.1.2.

3.1.1 Resultados do Primeiro Mapeamento Sistemático de Literatura

Após a busca por trabalhos através da *string* de busca, foram aplicados os CI's e CE's, resultando na seleção de 29 trabalhos, cujas identificações podem ser encontradas na Tabela 3.1. A distribuição desses trabalhos de acordo com o ano de publicação é apresentada em um gráfico na Figura 3.1. As Subseções 3.1.1.1 a 3.1.1.4, apresentam as respostas de cada uma das QP's.

Pode-se observar que apenas dois trabalhos relacionados foram publicados antes de 2010, e que, partir desse ano, as pesquisas na área começaram a se intensificar. Pode-se observar um pico no volume de trabalhos publicados no ano de 2016. Pode-se interpretar que o fato das pesquisas relacionadas se intensificarem na última década, se deve ao crescimento da internet, redes sociais, sistemas de vigilância, entre outras aplicações relacionadas a um grande volume de vídeos.

¹O '*' pode ser substituído por qualquer combinação de caracteres. Por exemplo, com o termo “*emotion**”, a *string* de busca irá aceitar as palavras *emotion*, *emotions* ou *emotinal*, entre outras possíveis combinações

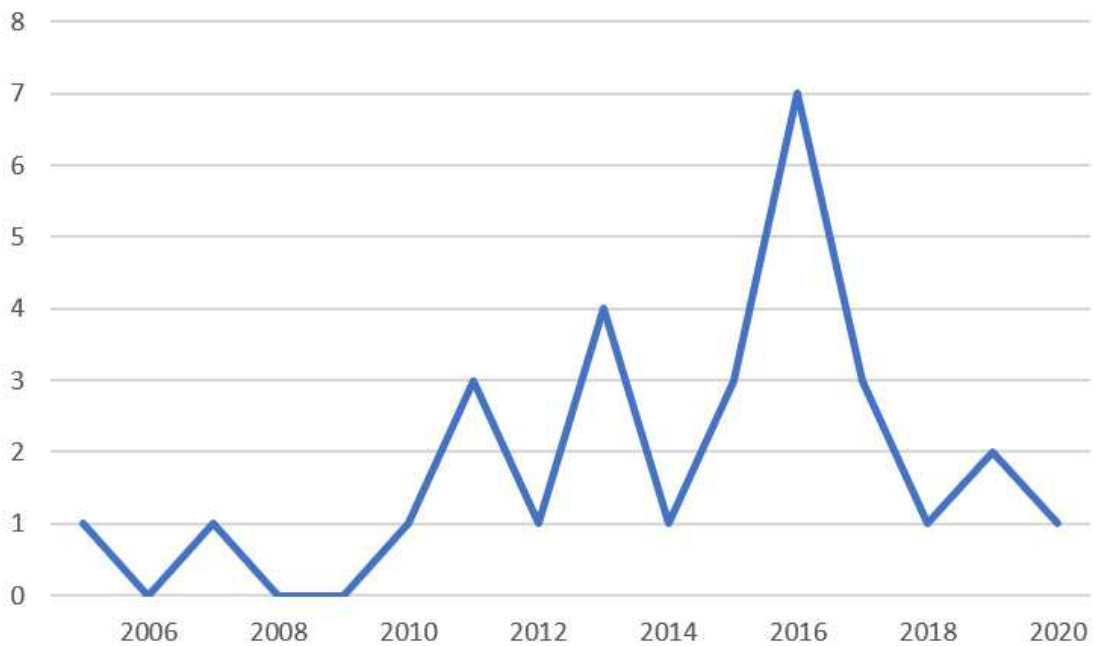


Figura 3.1: Ano de publicação dos trabalhos selecionados no primeiro MSL

3.1.1.1 *Quais informações sobre estados afetivos têm contribuído positivamente com as técnicas de SV?*

Todos os trabalhos selecionados relatam que o uso de informações afetivas contribuiu positivamente para SV. O objetivo dessa seção é apresentar todas as emoções ou estados afetivos utilizados pelos métodos de SV selecionados, que podem ser divididas em dois grupos: i) categórico ou discreto; e ii) dimensional ou contínuo. A Seção 2.4 apresenta detalhes sobre essas classificações.

Com relação ao conjunto categórico ou discreto, comumente são explorados o conjunto de emoções primárias, definido por (Ekman, 1994): raiva, tristeza, nojo, felicidade, medo e surpresa. Em alguns trabalhos essas emoções são divididas em três categorias: negativas, positivas e neutras, pois tal classificação, permite, em alguns casos, obter melhor desempenho preditivo (Malik e Hussain, 2017). Deve-se destacar, que as emoções que compõe cada uma das categorias não seguem um padrão, podendo ser diferentes de acordo com o trabalho analisado, por exemplo, em (Chakraborty et al., 2016), a categoria das emoções positivas é composta por felicidade e surpresa, porém (Peng et al., 2011) classificou apenas felicidade como uma emoção positiva.

Na Tabela 3.1 são apresentadas os afetivos utilizados em cada um dos trabalhos selecionados. A primeira coluna da identifica o trabalho, e na segunda são descritas os estados afetivos explorados.

3.1.1.2 *Quais as aplicações dos métodos de SV baseados em informações afetivas?*

Grande parte dos trabalhos selecionados são de aplicação genérica, ou seja, podem ser aplicados em qualquer tipo de vídeo, porém existem trabalhos com foco em tipos específicos de vídeos, tais como, vídeos domésticos (sem edição profissional), vídeos de esportes e também filmes com foco em gêneros específicos. A Tabela 3.2 apresenta a aplicação de cada um dos métodos de SV selecionados.

Tabela 3.1: Emoções/estados afetivos explorados pelos métodos de SV selecionados no MSL 1

Trabalho	Emoção/Estado afetivo
(Gao et al., 2020)	Valência e Excitação
(Ul Haq et al., 2019)	Alegria, tristeza, medo, nojo, surpresa, raiva e neutra
(Qayyum et al., 2019)	Alegria, tristeza, raiva, surpresa, amor e neutra
(Singhal et al., 2018)	Alegria, tristeza e neutro
(Salehin e Paul, 2017)	Valência e excitação
(Smith et al., 2017)	Medo
(Romão et al., 2017)	Não define a(s) emoção(ões) utilizada(s)
(Lan et al., 2016)	Não define a(s) emoção(ões) utilizada(s)
(Ashokkumar et al., 2016)	Neutra e positiva
(Chakraborty et al., 2016)	Valência, excitação e emoções neutras, positivas e negativas
(Raiturkar et al., 2016)	Valência e excitação
(Vermeulen et al., 2016)	Valência e excitação
(Mehmood et al., 2016)	Valência e excitação
(Xu et al., 2016)	Não define a(s) emoção(ões) utilizada(s)
(Qu et al., 2015)	Alegria, tristeza, raiva e neutra
(Dammak et al., 2015)	Alegria, tristeza, medo e neutra
(Xian et al., 2015)	Não define a(s) emoção(ões) utilizada(s)
(Wang et al., 2014)	Valência e excitação
(Lin et al., 2013)	Não define a(s) emoção(ões) utilizada(s)
(Hasan et al., 2013)	Valência e excitação
(Money e Agius, 2013)	Valência e excitação
(Zhao et al., 2013)	Alegria, tristeza, medo, nojo, surpresa e neutra
(Dhall e Goecke, 2012)	Alegria e neutra
(Peng et al., 2011)	Positiva e neutra
(Katti et al., 2011)	Valência e excitação
(Joho et al., 2011)	Alegria, tristeza, medo, nojo, surpresa e neutra
(Ong e Kameyama, 2010)	Valência e excitação
(Xu et al., 2007)	Medo
(Hanjalic, 2005)	Valência e excitação

Tabela 3.2: Aplicações dos métodos de SV selecionados no MSL 1

Trabalho	Aplicação
(Gao et al., 2020)	Vídeos de esportes
(Ul-Haq et al., 2019)	Qualquer tipo de vídeo
(Qayyum et al., 2019)	Qualquer tipo de vídeo
(Singhal et al., 2018)	Qualquer tipo de vídeo
(Salehin e Paul, 2017)	Qualquer tipo de vídeo
(Smith et al., 2017)	Filmes do gênero Terror
(Romão et al., 2017)	Vídeos de esportes
(Lan et al., 2016)	Qualquer tipo de vídeo
(Ashokkumar et al., 2016)	Qualquer tipo de vídeo
(Chakraborty et al., 2016)	Vídeos de esportes
(Raiturkar et al., 2016)	Qualquer tipo de vídeo
(Vermeulen et al., 2016)	Qualquer tipo de vídeo
(Mehmood et al., 2016)	Qualquer tipo de vídeo
(Xu et al., 2016)	Vídeos domésticos (sem edição profissional)
(Qu et al., 2015)	Qualquer tipo de vídeo
(Dammak et al., 2015)	Qualquer tipo de vídeo
(Xian et al., 2015)	Qualquer tipo de vídeo
(Wang et al., 2014)	Vídeos de esportes
(Lin et al., 2013)	Filmes do gênero Drama
(Hasan et al., 2013)	Vídeos de esportes
(Money e Agius, 2013)	Qualquer tipo de vídeo
(Zhao et al., 2013)	Qualquer tipo de vídeo
(Dhall e Goecke, 2012)	Qualquer tipo de vídeo
(Peng et al., 2011)	Vídeos domésticos (sem edição profissional)
(Katti et al., 2011)	Qualquer tipo de vídeo
(Joho et al., 2011)	Qualquer tipo de vídeo
(Ong e Kameyama, 2010)	Qualquer tipo de vídeo
(Xu et al., 2007)	Filmes do gênero Terror
(Hanjalic, 2005)	Vídeos de esportes

3.1.1.3 Como informações afetivas são adquiridas para auxiliar a SV?

Quanto às formas de aquisição dos estados afetivos utilizados para criar os SV nos trabalhos selecionados, dois critérios podem ser utilizados para classificar os trabalhos: i) origem dos dados e ii) tipo de dados utilizados. Em relação à origem dos dados, três tipos de classificações podem ser utilizadas para os métodos de SV (Money e Harry, 2008): i) interna, que analisa dados obtidos diretamente do vídeo; ii) externa, que analisa dados que não estão no vídeo (no contexto de emoções, normalmente são utilizadas informações referentes ao(s) telespectador(es)); e iii) híbridas, que combinam dados externos e internos. A classificação dos trabalhos selecionados quanto a origem dos dados é apresentada na Tabela 3.3.

Tabela 3.3: Origem dos dados utilizados nos métodos de SV selecionados no MSL 1

Trabalho	Interna	Externa	Híbrida
(Gao et al., 2020)	X		
(Ul Haq et al., 2019)	X		
(Qayyum et al., 2019)		X	
(Singhal et al., 2018)		X	
(Salehin e Paul, 2017)		X	
(Smith et al., 2017)	X		
(Romão et al., 2017)			X
(Lan et al., 2016)	X		
(Ashokkumar et al., 2016)		X	
(Chakraborty et al., 2016)		X	
(Raiturkar et al., 2016)		X	
(Vermeulen et al., 2016)		X	
(Mehmood et al., 2016)		X	
(Xu et al., 2016)	X		
(Qu et al., 2015)	X		
(Dammak et al., 2015)		X	
(Xian et al., 2015)		X	
(Wang et al., 2014)	X		
(Lin et al., 2013)	X		
(Hasan et al., 2013)	X		
(Money e Agius, 2013)		X	
(Zhao et al., 2013)	X		
(Dhall e Goecke, 2012)	X		
(Peng et al., 2011)		X	
(Katti et al., 2011)		X	
(Joho et al., 2011)		X	
(Ong e Kameyama, 2010)			X
(Xu et al., 2007)	X		
(Hanjalic, 2005)	X		

Quanto aos tipos de dados utilizados, podem ser provenientes de sensores ou do próprio vídeo. Os dados utilizados em cada um dos trabalhos selecionados são apresentados na Tabela 3.4, na qual a primeira coluna é utilizada para identificar o trabalho e a segunda coluna apresenta

a origem de dados. Todas as origens de dados dos trabalhos selecionados são apresentadas a seguir:

- Anotações Manuais: podem ser inseridas aos vídeos pelos telespectadores do vídeo, sendo possível analisar os estados afetivos com base no texto das anotações (Xian et al., 2015).
- Atividade Eletro-dérmica (do inglês, *Electrodermal Activity* - EDA): mede a condutividade elétrica da pele, estando diretamente ligada com a excitação (Gomez e Danuser, 2004). Na literatura, esse tipo de técnica também é chamado de Resposta Potencial da Pele;
- Áudio: comumente se utiliza características referentes ao tom e frequência para identificar emoções no áudio (Ooi et al., 2014). Tais características podem ser utilizadas para classificar a excitação no áudio (Gao et al., 2020). Também pode-se transcrever o texto de um áudio e analisar as informações com base no texto (Chatterjee et al., 2019).
- Eletroencefalografia (EEG): é um método utilizado para registrar a atividade elétrica do cérebro. Esse tipo de dado, pode ser utilizado para identificar o estado de excitação, curiosidade e pânico do telespectador de um vídeo (Salehin e Paul, 2017);
- Expressões Faciais (EF's): podem ter funções de comunicação distintas, como sinalizar a tomada de turnos, transmitir informações biométricas, expressar a intensidade do esforço mental e indicar emoções (Corneanu et al., 2016).
- Frequência Cardíaca (FC): a aceleração e desaceleração do batimento cardíaco também pode ser indicador de valência (Cacioppo et al., 1997);
- Vídeo: características dos vídeos, tal como, cor, forma, movimento e texto, podem ser utilizados para detecção da emoção contida no vídeo;
- Movimentos Corporais: emoções são frequentemente associadas com diferentes movimentos do corpo, podendo-se utilizar diferentes características do movimento para inferir emoções, tais como: amplitude, velocidade e fluidez do movimento, além de se tentar reconhecer formas e gestos que representam emoções específicas (Castellano et al., 2007);
- Olhos: a fixação e o movimentos dos olhos são uma importante informação pois podem indicar o nível de atenção de um telespectador (Peng et al., 2011). Também pode-se analisar a dilatação da pupila para mensurar o nível de excitação de uma pessoa (Raiturkar et al., 2016); e
- Respiração: a frequência respiratória é medida pelo número de ciclos respiratórios em um determinado intervalo de tempo. com o subsequente movimento expiratório, ou expiração. A taxa e a amplitude de respiração podem ser utilizadas para indicar níveis de excitação e valência (Philippot et al., 2002).

Tabela 3.4: Tipos de dados utilizados pelos métodos de SV selecionados no MSL 1

Trabalho	Tipos de dados utilizados
(Gao et al., 2020)	Áudio (tom de voz do locutor)
(Ul Haq et al., 2019)	EF
(Qayyum et al., 2019)	EEG
(Singhal et al., 2018)	EEG
(Salehin e Paul, 2017)	EEG
(Smith et al., 2017)	Áudio e Vídeo
(Romão et al., 2017)	Vídeo, áudio e anotações manuais
(Lan et al., 2016)	Video
(Ashokkumar et al., 2016)	EF
(Chakraborty et al., 2016)	EF e FC
(Raiturkar et al., 2016)	Olhos (dilatação da pupila)
(Vermeulen et al., 2016)	FC
(Mehmood et al., 2016)	EEG
(Xu et al., 2016)	Áudio e vídeo
(Qu et al., 2015)	Áudio e vídeo
(Dammak et al., 2015)	EF e movimento corporal
(Xian et al., 2015)	Anotações manuais (cementários dos telespectadores)
(Wang et al., 2014)	Áudio e vídeo
(Lin et al., 2013)	Áudio e faces*
(Hasan et al., 2013)	Áudio e vídeo
(Money e Agius, 2013)	EDA, FC e respiração
(Zhao et al., 2013)	Vídeo
(Dhall e Goecke, 2012)	EF
(Peng et al., 2011)	EF, áudio
(Katti et al., 2011)	Olhos (dilatação da pupila)
(Joho et al., 2011)	EF
(Ong e Kameyama, 2010)	FC, olhos (dilatação da pupila) e vídeo
(Xu et al., 2007)	Vídeo
(Hanjalic, 2005)	Áudio e vídeo

* Esse trabalho identificou a face dos atores, mas não utilizou a EF, apenas considerou a quantidade e tamanho das faces em cada quadro.

3.1.1.4 Como os resultados das técnicas de SV baseadas em estados afetivos são apresentadas ao usuário final?

Pode-se observar que os trabalhos selecionados se preocupam em produzir sumários que apresentem uma boa abstração do conteúdo do vídeo original, porém quanto a forma de exibição resultados de sumarização, grande parte dos trabalhos apenas apresentam os sumários gerados da forma original que foram criados: quadros chaves ou cenas do vídeo. Sumários baseados em quadros chaves são mais eficientes para indexação e recuperação de conteúdo, porém cenas do vídeo são mais vívidas e informativas no ponto de vista do usuário (Qu et al., 2015).

Dois dos trabalhos selecionados permitiram que os usuários selecionassem parâmetros para criar sumários de acordo com o seu interesse (Ul Haq et al., 2019; Romão et al., 2017). Tais trabalhos apresentaram uma ferramenta na qual o usuário pode escolher na perspectiva de qual emoção, ele gostaria de visualizar no sumário. Romão et al. (2017), também permite que o usuário crie sumários considerando mais de uma ou de todas as emoções. Apesar de as escolhas dos usuários interferirem nos sumários produzidos, esses trabalhos não exploram nenhuma interface que agregue informações durante a visualização de sumários.

Uma interface para visualização de sumários foi proposta em apenas um dos trabalhos selecionados no mapeamento (Zhao et al., 2013): que utilizou cores para representar cada emoção e também a sua intensidade, permitindo que o usuário navegue pelo conteúdo, sendo direcionado para as cenas nas quais possui interesse, tal como pode ser observado na Figura 3.2.

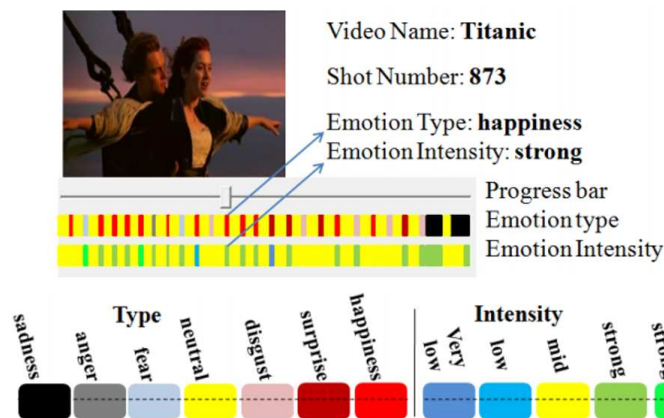


Figura 3.2: Ferramenta de visualização de vídeo proposta por Zhao et al. (2013)

3.1.2 Análise dos Resultados do Primeiro Mapeamento Sistemático de Literatura

Essa Seção apresentou um Mapeamento Sistemático da Literatura conduzido para identificar quais as tendências e desafios de pesquisa em relação aos métodos de SV baseados em informações afetivas. Existem ameaças à validade do estudo, já que alguns artigos relevantes podem não ter sido incluídos devido ao uso de um conjunto limitado de bibliotecas digitais (mesmo sendo um conjunto de bases de dados importantes) e também o sistema de classificação e as categorias elaboradas pode representar uma ameaça à validade (Pretorius e Budgen, 2008). Porém, mesmo com as possíveis ameaças, foi analisada uma parcela significativa dos trabalhos existentes na literatura, o que permite chegar a conclusões relevantes em relação às questões de pesquisa exploradas, identificando tendências e desafios de pesquisa.

Quanto a QP_1 (Quais informações sobre estados afetivos têm contribuído positivamente com as técnicas de SV?), os trabalhos selecionados reportaram resultados positivos com o uso

de informações afetivas. Vale ressaltar que alguns dos trabalhos compararam os resultados obtidos com métodos de SV que não utilizam informações emocionais, conseguindo resultados melhores com o uso de emoções, sendo esse, mais um subsídio para confirmar a hipótese de que informações afetivas têm um impacto positivo nos métodos de SV.

Com a QP_2 (Quais as aplicações dos métodos de SV baseados em informações afetivas?) pode-se observar que a maioria dos trabalhos propõe métodos genéricos, que podem ser aplicados em qualquer tipo de vídeo, permitindo concluir que o uso de informações afetivas tem potencial para contribuir com a sumarização de qualquer tipo de vídeo.

Já quanto a QP_3 (Como informações afetivas são adquiridas para auxiliar a SV?), a maioria dos trabalhos utiliza os afetivos dos telespectadores. Alguns desses trabalhos utilizam apenas dados coletados do próprio vídeo, enquanto outros trabalhos envolvem o uso de sensores para coletar informações fisiológicas, tais como, FC, EEG e EDA. Os resultados obtidos apontam que as duas abordagens são promissoras, podendo ser favorável o seu uso combinado.

Quanto a QP_4 (Como os resultados das técnicas de SV baseadas e, estados afetivos são apresentadas ao usuário final?), foi possível observar que os trabalhos selecionados têm uma grande preocupação em utilizar informações emocionais para superar problemas semânticos dos métodos de SV, porém quanto a exibição dos resultados, os trabalhos apenas apresentam o conteúdo produzido na forma de quadros chaves ou cenas do vídeo, que podem não ser as formas mais eficazes e eficientes de transmitir o conteúdo para usuário, principalmente quando o usuário tem a necessidade de analisar um conjunto grande de vídeos e não apenas um vídeo isolado, ou quando o mesmo não possui o conhecimento contextual adequado (Martins et al., 2011).

Por fim, pode-se concluir, que a possibilidade de criar sumários mais naturais aos espectadores e também de ampliar as aplicações de técnicas de SV (Arifin e Cheung, 2007) são fatores que motivaram a condução do projeto de pesquisa apresentado nessa tese.

3.2 MSL 2: TÉCNICAS COMPUTACIONAIS APLICADAS NA CLASSIFICAÇÃO, ANÁLISE OU SUPORTE AO ESTRESSE/ANSIEDADE DE MOTORISTAS

A revisão de literatura apresentada nessa Seção foi conduzida com o objetivo de identificar como técnicas computacionais são aplicadas para classificação, análise ou suporte ao estresse/ansiedade de motoristas. Com esse objetivo foram definidas três QP's²:

- QP_A : Quais as aplicações das técnicas computacionais relacionadas ao estresse/ansiedade de motoristas?
- QP_B : Quais os algoritmos e características exploradas nas técnicas computacionais relacionadas ao estresse/ansiedade de motoristas?;
- QP_C : Quais técnicas de visualização e sumarização são exploradas trabalhos selecionados?

Com base nas QP's, foram definidos os CI's e os CE's. A seguir são apresentados os CI's:

- CI_A : Se vários artigos do mesmo autor principal mostram estudos similares, somente o mais recente é incluído;

²Para evitar confusão durante a leitura do texto, as QP's foram identificadas com números no primeiro MSL e com letras no MSL apresentado nessa Seção - a mesma regra se aplica a CI's e CE's.

- CI_B : Se houver versões do mesmo artigo, uma resumida e outra completa, apenas a completa deve ser incluída; e
- CI_C : Trabalhos que apresentem a aplicação de técnicas computacionais para classificar, analisar ou dar suporte ao estresse/ansiedade de motorista, devem ser incluídos;

Também com base nas QP's, foram definidos os seguintes CE's:

- CE_A : Estudos que não estejam na língua inglesa serão excluídos;
- CE_B : Relatórios técnicos, documentos que estão disponíveis na forma de resumos, apresentações ou revisões de literatura secundárias (ou seja, revisões e mapeamentos sistemáticos da literatura) serão excluídos;
- CE_C : Trabalhos que apresentem o uso de recursos computacionais apenas para a coleta dos dados, serão excluídos. Tais trabalhos, depois de obter os dados não utilizam mais nenhuma técnica computacional. Esses trabalhos são excluídos, pois, as abordagens de coleta de dados apresentadas podem ser utilizadas para classificar diferentes estados afetivos e não exclusivamente estresse.
- CE_D : Trabalhos que explorem a identificação do estresse de motorista, apenas com o uso de equações matemáticas ou análises estatísticas serão excluídos, se não apresentarem nenhuma ferramenta computacional para classificar, analisar ou dar suporte ao estresse/ansiedade.

Um exemplo de trabalho que foi excluído com a CE_C é Affanni et al. (2019), que apresenta um sensor para coletar atividade eletro dérmica do motorista. Apesar da relevância desse trabalho em relação ao estresse/ansiedade de motoristas, o estudo somente avalia o quão corretos são os dados coletados, sem de fato apresentar uma aplicação desses dados na classificação/análise de estresse/ansiedade. Por esse motivo, esse estudo foi excluído com base no CE_C .

Com relação a CE_D , seguem exemplos de trabalhos que foram incluídos e excluídos:

- O trabalho de Ortoncelli et al. (2020), não aplica nenhuma técnica computacional para classificar níveis ou intensidade de estresses, porém, apresenta uma interface gráfica para dar suporte a sumarização e análise do estresse dos motoristas, por isso esse trabalho foi selecionado.
- O trabalho de (Kim et al., 2018) apresenta um arcabouço teórico para avaliar dados fisiológicos e operacionais em relação ao estresse de motoristas, mas todo esse trabalho não foi feito com uma técnica computacional, então o artigo não foi selecionado por essa contribuição. Porém, esse artigo apresenta uma ferramenta para visualização em tempo real de dados da atividade prática de direção, que podem ser utilizados para análise do estresse do motorista em tempo real. Por conta dessa ferramenta o trabalho de Kim et al. (2018) foi selecionado.
- Os trabalhos de Joosen et al. (2017); Krishnan (2017) calculam o estresse apenas com base em uma análise estatística, portanto, não foram selecionados.

Para responder às questões de pesquisa, foi criada uma *string* de busca, com termos relacionados a:

- Ansiedade do motorista: foram adicionados três termos sinônimos: “*driv* anxiety*”, “*motorist* anxiety*” e “*traffic anxiety*”;
- Estresse do motorista: foram adicionados três termos sinônimos: “*driv* stress*”, “*motorist* stress*” e “*traffic stress*”;
- Sensores: foram adicionados os termos relacionados utilizados na Revisão Sistemática de Literatura publicada por Seppälä et al. (2019): “*monitor*”, “*sensing*” e “*sensor*”;
- Visão Computacional e Inteligência Artificial: foram adicionados os termos relacionados utilizados na Revisão Sistemática de Literatura publicada por Patrício e Rieder (2018): “*computer vision*”, “*machine vision*”, “*artificial intelligence*”, “*machine learning*” e “*deep learning*”; e
- Tipos de dados que podem ser utilizados para detectar emoções e estados afetivos: foram utilizados termos relacionados aos dados explorados pelos trabalhos selecionados no primeiro MSL (Seção 3.1: “*facial expression*”, “*action units*”, “*Facial Action Coding System*”, “*body moviment*”, “*body motion*”, “*human moviment*”, “*human motion*”, “*hand* moviment*”, “*hand* motion*”, “*pupil* dilation*”, “*Beats Per Minute*”, “*BPM*”, “*heart rate*”, “*HR*”, “*electrodermal*”, “*EDA*”, “*respiration rate*”, “*respiration amplitude*”, “*Electroencephalogram*” e “*EEG*”).

Com os termos apresentados, foi criada a seguinte *string* de busca:

((“*driv* anxiety*” OR “*motorist* anxiety*” OR “*traffic anxiety*” OR “*driv* stress*” OR “*motorist* stress*” OR “*traffic stress*”) AND (“*monitor*” OR “*sensing*” OR “*sensor*” OR “*computer vision*” OR “*machine vision*” OR “*artificial intelligence*” OR “*machine learning*” OR “*deep learning*” OR “*facial expression*” OR “*action units*” OR “*Facial Action Coding System*” OR “*body moviment*” OR “*body motion*” OR “*human moviment*” OR “*human motion*” OR “*hand* moviment*” OR “*hand* motion*” OR “*pupil* dilation*” OR “*Beats Per Minute*” OR “*BPM*” OR “*heart rate*” OR “*HR*” OR “*electrodermal*” OR “*EDA*” OR “*respiration rate*” OR “*respiration amplitude*” OR “*Electroencephalogram*” OR “*EEG*”).

Essa *string* de busca apresentada foi adaptada de acordo com as características de cada uma das bases pesquisadas. As buscas foram conduzidas em bases de dados eletrônicas de grande importância para a Ciência da Computação (Dyba et al., 2007): ACM Digital Library, Elsevier (via Science Direct), Engineering Village, IEEEExplore, Scopus, Springer e Web of Science.

Além de analisar os trabalhos obtidos com a *string* de busca, também foram aplicados os CI’s e CE’s nos trabalhos referenciados por três revisões de literatura recentes, que apresentam trabalhos relacionados a análise de emoções de motoristas (Zepf et al., 2020; Rastgoo et al., 2018) e a classificação de ansiedade (Arif et al., 2020).

Os resultados obtidos com o segundo mapeamento são apresentados na Subseção 3.2.1. A análise desses resultados é apresenta na Subseção 3.2.2.

3.2.1 Resultados do Segundo Mapeamento Sistemático de Literatura

Após a busca por trabalhos através da *string* de busca, foram aplicados os CI's e CE's, resultando na seleção de 45 trabalhos. A Tabela 3.3 identifica os trabalhos selecionados. A distribuição desses trabalhos de acordo com o ano de publicação é apresentada na Figura 3.3. As Subseções 3.2.1.1 a 3.2.1.3, apresentam as respostas das QP's A, B e C.

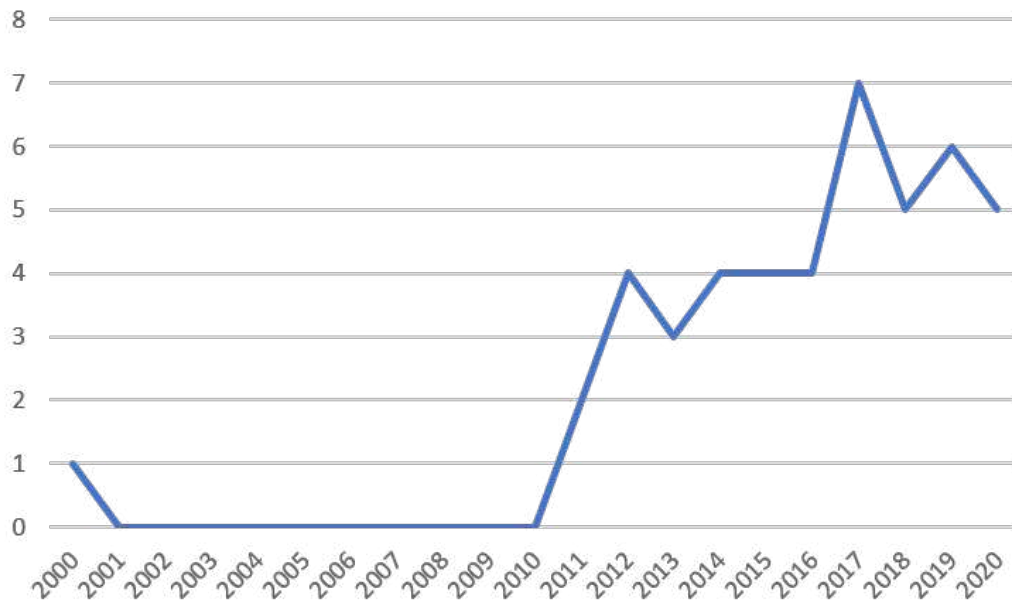


Figura 3.3: Ano de publicação dos trabalhos selecionados no segundo MSL

3.2.1.1 *Quais as aplicações das técnicas computacionais relacionadas ao estresse/ansiedade de motoristas?*

Para responder a essa *QP*, inicialmente foi identificado o tipo de problema resolvido para cada um dos trabalhos selecionados. Deve-se destacar que apenas um deles relatou trabalhar com a ansiedade de motoristas (Lee et al., 2020). Isso se justifica, pois, de acordo com (Lee et al., 2020), embora a ansiedade ao dirigir seja uma das emoções mais influentes para a segurança do motorista, poucos estudos mediram as respostas fisiológicas (Barnard e Chapman, 2016) e neurais (Lee et al., 2018) da ansiedade em comparação com outras respostas emoções negativas. Nesse contexto, o uso de técnicas computacionais é ainda mais restrito, como pode-se observar no levantamento da literatura.

Mesmo com o pequeno volume de trabalhos relacionados à ansiedade de motoristas, como ansiedade e estresse são estados afetivos diretamente relacionados (Giannakakis et al., 2017; Padiaditis et al., 2015), os trabalhos analisados permitem compreender o estado da arte no contexto desta tese..

Quanto ao estresse de motoristas, pode-se observar duas aplicações principais: i) classificação do nível ou probabilidade de estresse (40 trabalhos); e ii) sumarização ou visualização ou sumarização de informações relacionadas a estresse (4 trabalhos).

Na maioria desses trabalhos o estresse é classificado de forma binária, ou seja, apenas se identifica dois níveis de estresse: com estresse ou sem estresse (ou com ansiedade e sem ansiedade). Detalhes sobre os algoritmos e características utilizadas para classificar o estresse nesses artigos são apresentados na Subseção 3.2.1.2.

Os quatro trabalhos relacionados a sumarização ou visualização (Ortoncelli et al., 2020; Kim et al., 2018; Cruz e Rinaldi, 2017; Rodrigues et al., 2015) apresentam ferramentas visuais para análise de dados coletados durante as atividades práticas de direção. Em Ortoncelli et al. (2020) é apresentada uma técnica de sumarização e visualização de relatórios de atividades práticas de direção com foco na análise de estresse, que faz parte da abordagem proposta nessa tese, sendo detalhado no Capítulo 7. Em Kim et al. (2018) é apresentada uma ferramenta que permite visualizar em tempo real os dados coletados de atividade prática de direção, que podem ser utilizados para análise do estresse do motorista em tempo real. Já em Rodrigues et al. (2015) é apresentada uma abordagem para visualização de um mapa de estresses dos motoristas de ônibus de uma cidade, com uso da API Google Maps.

Um método de sumarização foi explorado em (Cruz e Rinaldi, 2017), que resume vídeos de faces de motoristas com o objetivo de reduzir os requisitos computacionais necessários para implementar um sistema de reconhecimento de estresse. Considerando que para emoções não é necessário analisar todos os quadros de um vídeo, o algoritmo proposto nesse trabalho seleciona um conjunto de quadros chaves representativos. Com base nesses quadros e em características do áudio do vídeo, é utilizada uma RNA para classificar a valência do motorista, que pode ser utilizada para descrever a positividade ou negatividade dos sentimentos, que é uma informação útil para extrapolar o nível de estresse. Os resultados apontam que a técnica contribuiu para melhorar a taxa de reconhecimento de emoções de motoristas, porém, não foi realizada uma avaliação da representatividade dos quadros, ou seja, não foi analisado se os quadros chaves representam momentos nos quais as emoções são mais proeminentes. Também não é apresentada nenhuma técnica para visualização dos sumários.

Além da aplicação dos artigos selecionados, também foi classificado em qual ambiente foram executados os experimentos. Dois ambientes foram utilizados nos trabalhos: simuladores ou ambientes reais (com um veículo real em condições reais de direção). As Tabelas 3.5 e 3.6 apresentam as aplicações e os ambientes experimentais considerados em cada um dos trabalhos selecionados.

Tabela 3.5: Aplicação e ambiente experimental dos trabalhos selecionados no MSL 2 - parte 1

Trabalho	Aplicação	Ambiente
(Ortoncelli et al., 2020)*	Sumarização de vídeo	Ambiente real
(Elgendi e Menon, 2020)	Classificação: 3 níveis de estresse	Ambiente real
(Lee et al., 2020)	Classificação: 2 níveis de ansiedade	Simulador
(Li et al., 2020)	Classificação: 2 níveis de estresse	Ambiente real
(Wang e Guo, 2020)	Classificação: 3 níveis de estresse	Ambiente real
(Bianco et al., 2019)	Classificação: 2 níveis de estresse	Simulador
(El Haouij et al., 2019)	Classificação: 3 níveis de estresse	Ambiente real
(Rastgoo et al., 2019)	Classificação: 3 níveis de estresse	Simulador
(Sihem et al., 2019)	Classificação: 3 níveis de estresse	Ambiente real
(Verma et al., 2019)	Classificação: 3 níveis de estresse	Ambiente real
(Affanni et al., 2018)	Classificação: 2 níveis de estresse	Simulador
(Baltodano et al., 2018)	Classificação: 2 níveis de estresse	Ambiente real

* Esse trabalho faz parte da abordagem proposta nessa tese e é detalhado no Capítulo 7.

Tabela 3.6: Aplicação e ambiente experimental dos trabalhos selecionados no MSL 2 - parte 2

Trabalho	Aplicação	Ambiente
(Kim et al., 2018)	Visualização de dados**	Ambiente real
(Napoletano e Rossi, 2018)	Classificação: 2 níveis de estresse	Simulador
(Zontone et al., 2018)	Classificação: 2 níveis de estresse	Simulador
(Abayomi et al., 2017)	Classificação: 3 níveis de estresse	Ambiente real
(Chen et al., 2017)	Classificação: 3 níveis de estresse	Ambiente real
(Cruz e Rinaldi, 2017)	Sumarização de vídeo	Ambiente real
(Magana e Munoz-Organero, 2017)	Classificação: 2 níveis de estresse	Ambiente real
(Mühlbacher-Karrer et al., 2017)	Classificação: 2 níveis de estresse	Simulador
(Saeed e Trajanovski, 2017)	Classificação: 2 níveis de estresse	Simulador
(Urbano et al., 2017)	Classificação: 2 níveis de estresse	Ambiente real
(Lee et al., 2016)	Classificação: 2 níveis de estresse	Simulador
(Lee e Chung, 2016)	Classificação: 2 níveis de estresse	Simulador
(Ollander et al., 2016)	Classificação: 2 níveis de estresse	Ambiente real
(Ooi et al., 2016)	Classifica: estresse e raiva	Simulador
(Gutmann et al., 2015)	Classificação: 2 níveis de estresse	Simulador
(Keshan et al., 2015)	Classificação: 3 níveis de estresse	Ambiente real
(Munla et al., 2015)	Classificação: 2 níveis de estresse	Ambiente real
(Rodrigues et al., 2015)	Visualização de mapa de estresse	Ambiente real
(Gao et al., 2014)	Classificação: 2 níveis de estresse	Simulador
(Lanatà et al., 2014)	Classificação: 2 níveis de estresse	Simulador
(Vhaduri et al., 2014)	Classificação: uma escala de 0 a 1	Ambiente real
(Soman et al., 2014)	Classificação: 2 níveis de estresse	Ambiente real
(Singh et al., 2013)	Classificação: 3 níveis de estresse	Ambiente real
(Wang et al., 2013)	Classificação: 2 níveis de estresse	Ambiente real
(Anzengruber e Riener, 2012)	Classificação: 2 níveis de estresse	Simulador
(Begum et al., 2012)	Classificação: 2 níveis de estresse	Simulador
(Conjeti et al., 2012)	Classificação: 3 níveis de estresse	Ambiente real
(Bořil et al., 2012)	Classificação: 2 níveis de estresse	Ambiente real
(Katsis et al., 2011)	Classificação: 2 níveis de estresse + desapontamento e euforia	Simulador
(Rigas et al., 2011)	Classificação: 2 níveis de estresse	Ambiente real
(Healey e Picard, 2000)	Classificação: 4 níveis de estresse	Simulador

** Apresenta uma ferramenta para visualização de dados da atividade prática de direção em tempo real.

3.2.1.2 *Quais os algoritmos e características exploradas nas técnicas computacionais relacionadas ao estresse de motoristas?*

Esse Subseção apresenta as características exploradas pelos trabalhos selecionados (Tabelas 3.7 e 3.8) e também os algoritmos que foram utilizados (Tabelas 3.9 e 3.10).

Grande parte das características exploradas nos trabalhos selecionados para o segundo MSL, são extraídas de fontes de dados já foram descritas na Subseção 3.1.1.3. Os tipos de informações que não foram descritos previamente são apresentados a seguir:

- **Atividade Elétrica de Nervos e Músculos (AENM):** A atividade dos músculos aumenta em situações de estresse, pois pessoas podem inconscientemente contrair os músculos de acordo com estado mental, mesmo quando nenhuma atividade física é necessária. Diferentes exames podem ser utilizados para medir essa atividade, tais como o Eletromiograma (Lundberg et al., 1994) e a análise de sinais faciais eletromiográficos (Katsis et al., 2008).
- **Dados do veículo:** o estresse pode afetar a forma com a qual o motorista interage com o veículo. Essas reações físicas dos motoristas podem ser monitoradas usando dados do veículo, como volante, aceleração e desaceleração do veículo (Rastgoo et al., 2019).
- **Dados do ambiente:** podem contribuir para o entendimento dos fatores que afetam o nível de estresse do motorista (Hill e Boyle, 2007). São exemplos de dados do ambiente: temperatura, horário, clima, intensidade de tráfego e distância para outros veículos.
- **GPS:** A localização e a velocidade do veículo podem ser obtidas a partir do GPS, que são informações que podem contribuir para identificar o estresse do motorista. Com base no horário e nas coordenadas GPS, também se pode identificar eventos estressores (Rodrigues et al., 2015).
- **Temperatura corporal:** existem trabalhos que apontam que a temperatura da pele pode estar diretamente ligada com o nível de estresse. Em (Anzengruber e Riener, 2012), é apresentado um estudo que indica que altas temperaturas do rosto estão ligadas com um maior nível de estresse, enquanto temperaturas mais baixas estão relacionadas a um estado mental mais relaxado e livre de estresse.

Com relação ao movimento corporal, existe uma diferença entre os trabalhos selecionados nas duas revisões apresentadas nessa tese. No primeiro MSL os movimentos foram classificados com uso de técnicas de visão computacional. Já no segundo mapeamento, os movimentos foram identificados com o uso de sensores vestíveis (Lee et al., 2016; Lee e Chung, 2016), ou somente com rotulação manual (Ortoncelli et al., 2020).

Tabela 3.7: Dados utilizados pelos trabalhos selecionados no MSL 2 - parte 1

Trabalho	Tipo de Dados
(Ortoncelli et al., 2020)*	EF, FC, GPS e movimentos corporais
(Elgendy e Menon, 2020)	FC, AENM, EDA e respiração
(Lee et al., 2020)	FC, EEG, EDA e dilatação da pupila
(Li et al., 2020)	FC
(Wang e Guo, 2020)	EDA
(Bianco et al., 2019)	EDA, FC e respiração
(El Haouij et al., 2019)	FC, AENM, EDA e respiração
(Noh et al., 2019)	FC, EEG, dados do veículo
(Rastgoo et al., 2019)	FC, dados do veículo e dados do ambiente
(Sihem et al., 2019)	FC
(Verma et al., 2019)	GPS e dados do veículo
(Affanni et al., 2018)	EDA
(Baltodano et al., 2018)	FC, EDA e temperatura corporal

* Esse trabalho faz parte da abordagem proposta nessa tese e é detalhado no Capítulo 7.

Tabela 3.8: Dados utilizados pelos trabalhos selecionados no MSL 2 - parte 1

Trabalho	Tipo de Dados
(Kim et al., 2018)	EEG, dados do veículo, olhos (direção do olhar) e vídeo
(Napoletano e Rossi, 2018)	FC
(Zontone et al., 2018)	EDA
(Abayomi et al., 2017)	FC, AENM, EDA e respiração
(Chen et al., 2017)	FC, AENM, EDA e respiração
(Cruz e Rinaldi, 2017)	EF e áudio
(Magana e Munoz-Organero, 2017)	FC e dados do veiculo
(Mühlbacher-Karrer et al., 2017)	FC, EDA, EEG
(Saeed e Trajanovski, 2017)	FC e EDA
(Urbano et al., 2017)	FC e EDA
(Lee et al., 2016)	FC e movimento corporal
(Lee e Chung, 2016)	Movimento corporal
(Ollander et al., 2016)	FC, AENM, EDA e respiração
(Ooi et al., 2016)	FC
(Gutmann et al., 2015)	FC, EDA, EEG, respiração e dados do veículo
(Keshan et al., 2015)	FC
(Munla et al., 2015)	FC
(Rodrigues et al., 2015)	FC, GPS e anotações manuais**
(Gao et al., 2014)	EF
(Lanata et al., 2014)	FC, EDA e respiração
(Soman et al., 2014)	FC E AENM
(Vhaduri et al., 2014)	GPS
(MacLean et al., 2013)	FC e EDA
(Singh et al., 2013)	FC e EDA
(Wang et al., 2013)	FC
(Anzengruber e Riener, 2012)	Temperatura corporal
(Begum et al., 2012)	FC
(Bořil et al., 2012)	Áudio
(Conjeti et al., 2012)	FC, EDA
(Katsis et al., 2011)	FC, AENM, EDA e respiração
(Rigas et al., 2011)	FC, EDA e respiração
(Healey e Picard, 2000)	FC, EDA e respiração

** Anotações manuais são feitas pelos próprios motoristas apertando um botão para indicar eventos estressores.

Quanto às técnicas computacionais utilizadas, estas podem ser divididas em quatro aplicações principais: extração de características, classificação, redução do espaço de características, e visualização de informações.

Para classificação foram utilizados os seguintes algoritmos: AdaBoost (Wang e Guo, 2020), Algoritmo de Detecção de Estresse Biométrico - ADEB (Garcia-Mancilla e Gonzalez, 2015), *Bayesian networks* (Jensen et al., 1996), Rede Neural Celular (Li et al., 2011), Rede Neural Convolucional - RNC (Albawi et al., 2017), Regressão Logística - RL (Speelman, 2014), *Connectivity-Based Clustering* - CBC (Elgendi e Menon, 2020). Deep learning (LeCun et al.,

2015), *Fuzzy C-Means* - FCM (Nayak et al., 2015), *Generalized Linear Mixed Model* - GLMM (Wolfinger e O'connell, 1993), *Extreme Learning Machine* ELM (Huang et al., 2011), *k-Nearest Neighbor* - KNN (Peterson, 2009), *K-means clustering* - KMC (Elgendi e Menon, 2020) Modelos de Mistura Gaussiana - MMG (Reynolds, 2009), Naive Bayes (Rish et al., 2001), *Nearest Mean Classifier* - NMC (Lanata et al., 2014), *Random Forest* (Breiman, 2001), *Random Trees* (LaValle e Kuffner, 2001), Rede Neural Artificial - RNA (Mehrotra et al., 1997), *Support Vector Machine* - SVM (Hearst et al., 1998) e Zero-R (Inamdar et al., 2011).

Para extração de características foram utilizados os algoritmos de processamento digital de imagens *Histogram of Oriented Gradient* - HoG (Dalal e Triggs, 2005), *Local Binary Patterns* - LBP (Pietikäinen et al., 2011), e o cálculo da diferença entre quadros. Também foram utilizados dados estatísticos e equações matemáticas no processo de extração características, porém essas técnicas não foram analisadas nesse MSL. Também foi utilizado o algoritmo de Viola e Jones (2004) para detectar rostos.

Para reduzir a dimensionalidade do espaço de características e selecionar conjuntos de características ideais foram utilizados os seguintes algoritmos: *Kernel-Based Class Separability* - KBCS (Wang, 2008) *Sparse Bayesian learning* - SBL, (Wipf e Rao, 2004), *Principal Component Analysis* - PCA (Wold et al., 1987), *Interaction Principal Component Analysis* - IPCA (Elgendi, 2019) e *Linear Discriminant Analysis* - LDA. (Balakrishnama e Ganapathiraju, 1998). Também foi utilizado o algoritmo *Sequential Forward Floating Selection* - SFFS, que foi aplicado para encontrar o conjunto ideal de características para reconhecer padrões de estresse.

Em relação à visualização de informações, pode-se destacar interfaces gráficas para visualização de sumários e relatórios (Ortoncelli et al., 2020) e mapas de estresse (Rodrigues et al., 2015). Também foi utilizado um hardware no formato de uma borboleta vestível que apresenta o nível de estresse para o motorista com base na intensidade do bater das asas (MacLean et al., 2013).

Tabela 3.9: Técnica computacional utilizada nos trabalhos selecionados no MSL 2 - parte 1

Trabalho	Técnica Computacional
(Ortoncelli et al., 2020)*	Ferramenta para visualização
(Elgendi e Menon, 2020)	IPCA, CBC, KMC
(Lee et al., 2020)	RL
(Li et al., 2020)	SVM, KNN e <i>Random Forest</i>
(Wang e Guo, 2020)	AdaBoost e RNA
(Bianco et al., 2019)	SVM, KNN e RNA
(El Haouij et al., 2019)	Random Forests
(Noh et al., 2019)	FCM
(Rastgoo et al., 2019)	RNC
(Verma et al., 2019)	RNA
(Affanni et al., 2018)	Filtros de passa-baixa e Operador de Energia Não Linear Suavizado (Calvagno et al., 2000)
(Sihem et al., 2019)	<i>Random Forest</i>
(Baltodano et al., 2018)	ADEB e API iMotions (Kulke et al., 2020)
(Kim et al., 2018)	Ferramenta para visualização
(Napoletano e Rossi, 2018)	SMV

* Esse trabalho faz parte da abordagem proposta nessa tese e é detalhado no Capítulo 7.

Tabela 3.10: Técnica computacional utilizada nos trabalhos selecionados no MSL 2 - parte 2

Trabalho	Técnica Computacional
(Zontone et al., 2018)	SVM
(Abayomi et al., 2017)	RNA, HOG, LBP
(Chen et al., 2017)	SBL, PCA, SVM e ELM
(Cruz e Rinaldi, 2017)	RNA
(Magana e Munoz-Organero, 2017)	Deep learning
(Mühlbacher-Karrer et al., 2017)	Rede Neural Celular
(Saeed e Trajanovski, 2017)	RNA
(Urbano et al., 2017)	LDA e SVM
(Lee et al., 2016)	SVM
(Lee e Chung, 2016)	SVM
(Ollander et al., 2016)	Naive Bayes
(Ooi et al., 2016)	SVM
(Gutmann et al., 2015)	RNC
(Keshan et al., 2015)	RNA, SVM, Naive Bayes, RL, Zero-R, Árvore de decisão, <i>Random Trees</i> e <i>Random Forest</i>
(Munla et al., 2015)	SVM, KNN e BRF
(Rodrigues et al., 2015)	Ferramenta para visualização
(Gao et al., 2014)	SVM e o algoritmo de Viola e Jones (2004)
(Lanata et al., 2014)	KNN e NMC e RNA
(Soman et al., 2014)	SVM
(Vhaduri et al., 2014)	GLMM
(Ortoncelli et al., 2020)*	Ferramenta para visualização
(MacLean et al., 2013)	Hardware para visualização do nível de estresse
(Singh e Queyam, 2013)	RNA
(Wang et al., 2013)	KBCS, PCA, LDA, KNN
(Anzengruber e Riener, 2012)	MMG e Visão Computacional
(Begum et al., 2012)	KNN
(Bořil et al., 2012)	MMG
(Conjeti et al., 2012)	RNA
(Katsis et al., 2011)	Árvore de decisão
(Rigas et al., 2011)	Bayesian <i>Networks</i>
(Healey e Picard, 2000)	SFFS e KNN

3.2.1.3 Quais técnicas de visualização e sumarização são exploradas nos trabalhos selecionados?

A maioria dos trabalhos selecionados com esse mapeamento se dedicou exclusivamente ao desenvolvimento de técnicas para a classificação automática do estresse. Apenas quatro dos trabalhos selecionados apresentaram abordagens para sumarização ou visualizações de informações (Ortoncelli et al., 2020; Kim et al., 2018; Cruz e Rinaldi, 2017; Rodrigues et al., 2015). Destaca-se também o trabalho de MacLean et al. (2013), que visualiza o nível de estresse com base no bater de asas de uma borboleta eletrônica.

A ferramenta apresentada em Ortoncelli et al. (2020), faz parte da abordagem de sumarização proposta nessa tese, sendo descrita com detalhes no Capítulo 7. No trabalho de Kim et al. (2018) é apresentada uma interface para visualização de dados de atividades de práticas de direção em tempo real (Figura 3.4).

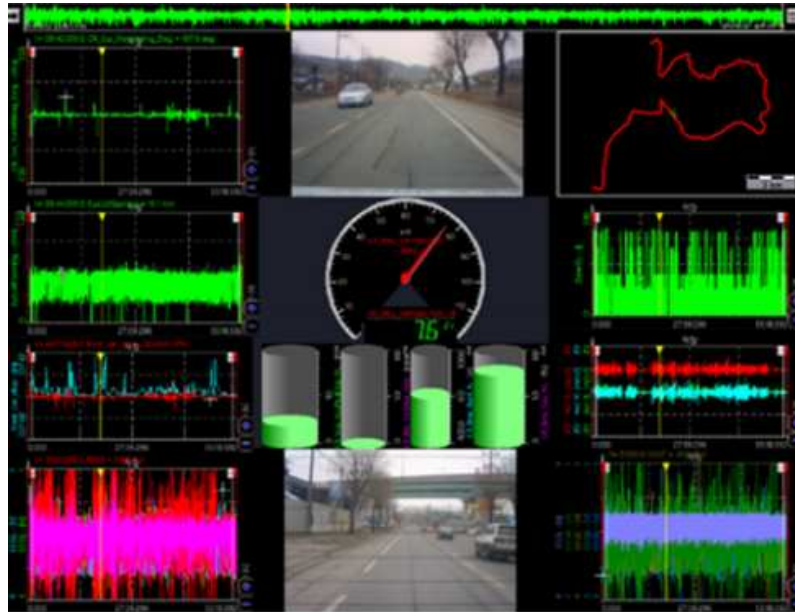


Figura 3.4: Ferramenta de monitoramento da atividade de direção: (Kim et al., 2018).

Uma técnica de sumarização é proposta em Cruz e Rinaldi (2017), que seleciona quadros chaves para otimizar o reconhecimento de emoções em vídeos frontais de faces de motoristas. Apesar de selecionar quadros chaves, o artigo não apresenta nenhum estudo da representatividade dos sumários produzidos quanto a capacidade de resumir o vídeo para a interpretação de usuários humanos, focando apenas na melhoria da classificação de emoções.

Em Rodrigues et al. (2015), é explorado um mapa de visualização de estresse utilizando a API Google Maps. Uma escala de cores e tamanho de regiões é utilizada para destacar eventos estressores para motoristas de ônibus de uma cidade (Figura 3.5).

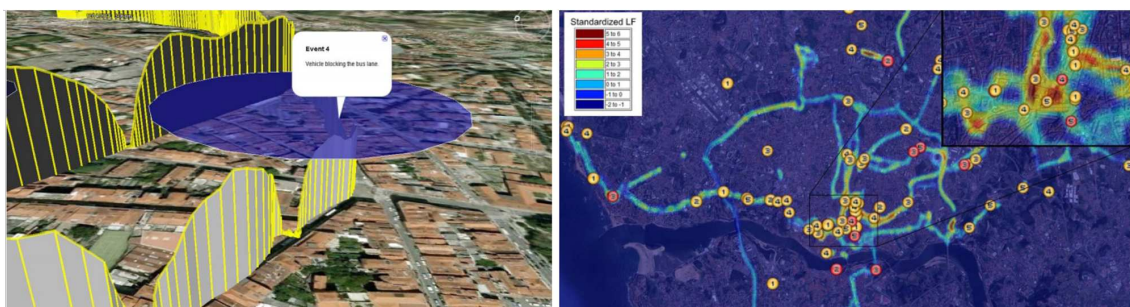


Figura 3.5: Abordagem de visualização: mapa de estresse (Rodrigues et al., 2015).

De todos os trabalhos selecionados, o único que apresentou uma abordagem de visualização de nível de estresse não baseada em uma interface gráfica foi (MacLean et al., 2013). Esse trabalho apresenta o nível de estresse para o motorista com base nos batimentos das asas de uma borboleta vestível (Figura 3.6). Os resultados experimentais desse trabalho foram negativos, pois a borboleta funcionou como um estressor. Os usuários comentaram que o movimento da borboleta os estressou porque viram como confirmação de seu estado de estresse.



Figura 3.6: Dispositivo de visualização do nível de estresse: (MacLean et al., 2013).

3.2.2 Análise dos Resultados do Segundo Mapeamento Sistemático de Literatura

Essa Seção apresentou uma revisão de literatura conduzida com o objetivo de identificar como técnicas computacionais são aplicadas para o classificar, analisar ou dar suporte ao estresse/ansiedade de motoristas. Existem ameaças a validade do estudo pois, alguns artigos relevantes podem não ter sido incluídos devido ao uso de um conjunto limitado de bibliotecas digitais (mesmo sendo um conjunto de bases de dados importantes) e também o sistema de classificação e as categorias elaboradas pode representar uma ameaça à validade (Pretorius e Budgen, 2008). Mesmo com essas ameaças, foi analisada uma parcela significativa da literatura, o que permite chegar a conclusões relevantes em relação às questões de pesquisa exploradas.

Com relação a QP_A , a grande maioria dos trabalhos teve foco no reconhecimento automático do estresse de motoristas, sendo que apenas um dos artigos foi direcionado à ansiedade de motoristas. Isso se justifica pois, apesar, da ansiedade ao dirigir ser uma das emoções que mais impactam na segurança no trânsito, poucos estudos mediram as respostas fisiológicas e neurais da ansiedade em comparação com outras respostas emoções negativas (Lee et al., 2020).

Quanto ao estresse, os trabalhos selecionados tiveram aplicação em duas áreas principais: i) classificação de estresse; e ii) sumarização ou visualização de informações. Todas essas abordagens são relevantes e úteis. Há espaço na literatura para contribuições que melhorem o estado da arte dessas técnicas, principalmente relacionadas à sumarização de atividades direção, devido ao menor volume de artigos publicados nessa área.

Em relação as respostas da QP_B , um grande conjunto de características e algoritmos são explorados na literatura para reconhecer o nível de estresse/ansiedade dos motoristas. Para algumas características exploradas há espaço para melhorias no processo de coleta de dados, tal como, a detecção de movimentos das mãos, que nos trabalhos selecionados é feita com o uso de sensores vestíveis. Nesse contexto, podem ser exploradas técnicas de visão computacional que detectem tais movimentos de forma mais natural. O Capítulo 5 apresenta métodos para a detecção desse tipo de movimento com base no vídeo.

Com relação a QP_C , apenas quatro dos trabalhos selecionados apresentaram técnicas de visualização de informações relacionadas a estresse/ansiedade do motorista (conforme foi apresentado na Subseção 3.2.1.3). Desses quatro trabalhos, apenas um teve foco em uma técnica de sumarização (Ortoncelli et al., 2020), apresentando uma abordagem que faz parte dessa tese, sendo detalhado no Capítulo 7.

Técnicas de sumarização são relevantes para agilizar o suporte a motoristas com alto nível de estresse. Devido à pequena quantidade de trabalhos relacionados na literatura, essa é uma área

que motiva o desenvolvimento de novas pesquisas. Também há espaço para o desenvolvimento de técnicas que possam ser utilizadas em conjunto com abordagens de sumarização de atividades práticas de direção, como é o caso da técnica de priorização apresentada no Capítulo 6.

4 BASE DE DADOS EXPERIMENTAL

Para produzir a base de dados experimental, todos os dados foram coletados em parceria com uma autoescola e seus alunos. Os dados foram coletados durante aulas prática de direção.

No Brasil, cada aula prática de direção tem a duração de 50 minutos, envolvendo cinco atividades: i) identificação biométrica inicial; ii) instruções iniciais; iii) atividade de direção; iv) *feedback* final; e v) identificação biométrica final. Para a produção da base de dados experimental foram considerados apenas dados coletados durante a atividade de direção.

A base de dados experimental é composta por dados coletados em 60 atividades práticas de direção com 60 motoristas diferentes — 39 mulheres, com média de idade de 29,5 anos e 21 homens, com a idade média de 21,9 anos. O tempo médio de duração das atividades de direção nas quais os dados foram coletados é de 38 minutos e 51 segundos, com um desvio padrão de 4 minutos e 58 segundos. Um grupo de 10 instrutores da autoescola participaram dos experimentos. Cada aluno foi acompanhado por apenas um instrutor.

Todas as etapas realizadas para conclusão do projeto proposto nessa tese foram previamente aprovadas pelo Comitê de Ética em Pesquisa (CEP) do setor de Ciências da Saúde da UFPR. As atividades referentes à aprovação do projeto, coleta de dados e produção da base de dados são apresentados nesse Capítulo. Também são apresentados detalhes da base de dados, que conta com informações sobre BPM, eventos rotulados e a prioridade das atividades práticas de direção.

O cronograma referente à aprovação do projeto e coleta de dados é mostrado na Seção 4.1. Os procedimentos e equipamentos utilizados para coleta de dados são apresentados na Seção 4.2. A atividade de rotulação dos vídeos é discutida na Seção 4.3. O processo de priorização manual das atividades de direção é descrito na Seção 4.4. As características da base de dados disponibilizada são apresentadas na Seção 4.5.

4.1 CRONOGRAMA

A Tabela 4.1 apresenta o cronograma de execução das atividades relacionadas à aprovação do projeto pelo CEP e também relacionadas a coleta de dados. Destaca-se, que devido a pandemia mundial da doença COVID-19 (Dong et al., 2020), o período de coleta de dados planejado para o projeto foi prejudicado.

No Brasil, as atividades das autoescolas foram inicialmente suspensas e em seguida retornaram com um protocolo de segurança: durante as aulas práticas e teóricas os alunos e instrutores eram obrigados a usar máscaras, impossibilitando a análise das AU's selecionadas nesse trabalho.

Tabela 4.1: Cronograma referente a aprovação do projeto pelo CEP e coleta de dados

Data	Evento
25/10/2019	O projeto recebe parecer favorável em reunião de colegiado do Programa de Pós-Graduação em Informática (PPGINF) da UFPR.
18/11/2019	O projeto é aprovado pelo CEP do setor de Ciências da Saúde da UFPR.
07/02/2020	Início da etapa de coleta de dados (Seção 4.2).
30/03/2020	Término precoce da etapa de coleta de dados devido a pandemia de COVID-19.

Com o impacto dessa pandemia, foi possível coletar dados de 60 atividades práticas de direção. Esperava-se obter um volume maior de dados. Também se planejava avaliar o impacto na abordagem proposta no suporte a motorista com alto nível de estresse — analisando os motoristas antes e depois do suporte prestado por uma psicóloga. Mesmo com essa limitação foi possível propor e avaliar uma abordagem para sumarização de atividades práticas de direção.

4.2 PROCESSO DE COLETA DE DADOS

O processo de coleta de dados contou com três atividades principais: i) convite e recrutamento dos participantes; ii) coleta de dados durante uma atividade prática de direção; e iii) entrevista final. Detalhes sobre essas atividades são apresentados a seguir:

- **Convite e recrutamento dos participantes:** alunos e instrutores da autoescola foram convidados a participarem do projeto. Os alunos foram convidados com uma carta convite entregue por um funcionário da autoescola. Instrutores foram convidados em reunião na qual o projeto foi apresentado. Os alunos e instrutores que aceitaram participar do projeto assinaram um termo de consentimento livre e esclarecido. Só foram coletados dados em atividades práticas nas quais tanto o aluno quanto o instrutor aceitaram participar do projeto.
- **Coleta de dados durante uma atividade prática de direção:** Os dados foram coletados durante uma aula prática de direção. Para garantir a segurança dos participantes, todos os experimentos foram realizados em um veículo da autoescola equipado com duplo mecanismo de pedais, com o acompanhamento de um instrutor de autoescola devidamente credenciado.

Foram coletados vídeos, BPM do motorista participante e coordenadas GPS. Tais dados foram adquiridos com os equipamentos e configuração descrita a seguir:

- três câmeras *full* HD (A Figura 4.1 ilustra a posição das câmeras):
 - * Câmera A - posicionada em uma visão frontal do rosto do motorista;
 - * Câmera B - posicionada diagonalmente no painel do veículo. Com essa imagem é possível analisar a interação entre o motorista e o instrutor, que pode ser uma das causas do estresse ao dirigir; e
 - * Câmera C - virada para o lado de fora do carro. Essa imagem permite identificar estressores externos entre outras informações contextuais sobre a atividade de direção.
- um relógio *smartwatch* coletou a frequência cardíaca dos motoristas; e
- um *smartphone* sincronizado por *Bluetooth* com o *smartwatch* foi utilizado para armazenar as informações de BPM e as coordenadas de GPS. O GPS em conjunto com a câmera C contribui para a análise de informações contextuais.

- **Entrevista Final:** Ao término da atividade, foi perguntado aos alunos se eles sentiram algum desconforto em relação ao experimento realizado durante a atividade prática de direção. Um dos participantes disse que se sentiu mais seguro com o uso dos equipamentos, os demais responderam que se sentiram indiferentes ao uso dos equipamentos.



Figura 4.1: Câmeras utilizadas na etapa de coleta de dados

4.3 ROTULAÇÃO DOS VÍDEOS

Apesar de três câmeras terem sido utilizadas para gravar os vídeos durante a coleta de dados, apenas os vídeos gravados com a câmera A foram utilizados na etapa de rotulação dos dados. Os vídeos obtidos com as demais câmeras foram utilizados para que especialistas compreendessem todo o contexto (estressores e eventos externos) de cada atividade prática de direção, dando suporte ao processo de priorização manual apresentado (Seção 4.4) e também sendo visualizados na ferramenta de suporte à sumarização (Capítulo 6).

Na etapa de rotulação foi utilizada uma ferramenta computacional para rotular manualmente quadro a quadro do vídeo, permitindo identificar o quadro inicial e o quadro final de cada um dos cinco hábitos do motorista que foram rotulados: morder, pressionar e lamber os lábios, roer as unhas e esfregar o rosto e/ou cabelo. Em relação aos movimentos das mãos do motorista, o Capítulo 5 apresenta detalhes sobre esses movimentos e também métodos computacionais propostos para os detectar. Também foi utilizada a ferramenta Open Face 2.0 (Baltrusaitis et al., 2018) para reconhecer automaticamente um conjunto de 18 AU's. Detalhes sobre as AU's foram apresentados na Seção 2.4.

4.4 PRIORIZAÇÃO MANUAL DE ATIVIDADES DE DIREÇÃO

Para produzir um conjunto de referência, com atividades de direção priorizadas pelo nível de estresse/ansiedade do motorista, a priorização manual da base de dados foi feita por uma psicóloga especializada em atender motoristas.

A psicóloga assistiu a todos os vídeos com a ferramenta de visualização apresentada na Seção 6. A ferramenta permite visualizar diferentes conjuntos de informações sincronizadas sobre a atividade de direção em três perspectivas: i) vídeo; ii) relatório; e iii) resumo. No processo de priorização manual, a ferramenta foi utilizada para analisar dados sobre BPM e os movimentos rotulados (Subseção 4.3).

Os sumários e relatórios visualizados possibilitaram que a psicóloga compreendesse de forma mais rápida o nível de estresse dos motoristas, contribuindo para agilizar o processo de priorização. Porém, mesmo com o uso da ferramenta é difícil priorizar todas as atividades em uma única iteração. Por essa razão, foi utilizado uma abordagem baseada no paradigma da divisão e conquista. Inicialmente, a psicóloga priorizou separadamente conjuntos de cinco atividades. Iterativamente os conjuntos de atividades foram mesclados recursivamente, usando a

mesma lógica do algoritmo *Merge Sort* (MS) — um método clássico de ordenação (Cormen et al., 2009).

Após a realização do processo de priorização, as 60 atividades de práticas de direção foram rotuladas com 41 níveis distintos de prioridade. Dessa forma, 19 atividades da base de dados têm a mesma prioridade de uma outra atividade, porém, não há mais do que três atividades com o mesmo rótulo (nível de prioridade).

4.5 BASE DE DADOS DISPONIBILIZADA

Os participantes dos experimentos não autorizaram a divulgação da sua imagem e dados pessoais, por isso os vídeos e informações pessoais dos participantes não fazem parte da base de dados que será disponibilizada mediante a solicitação e envio de termo de compromisso, de acordo com o processo descrito em <http://www.imago.ufpr.br/drivingstressdataset>.

Na base de dados, cada participante é identificado por um número inteiro entre 1 e 60. Para cada motorista, são disponibilizados dois arquivos de texto: i) um arquivo com o BPM e o instante de tempo no qual o BPM foi coletado; ii) um arquivo com as características extraídas de cada vídeos.

Também foi disponibilizada uma lista com a prioridade de cada atividade prática de direção, de acordo com o processo de priorização manual descrito na Seção 4.4.

A base de dados descrita nesse Capítulo foi disponibilizada com a licença *Creative Commons* CC BY¹, que autoriza os licenciados a copiar, distribuir, exibir e executar a obra e fazer trabalhos derivados dela, conquanto que deem créditos devidos ao autor. Todas às cópias derivadas do conteúdo poderão apenas ser licenciadas, com uma licença idêntica à obra original.

¹<https://creativecommons.org/licenses/by/4.0/>

5 DETECÇÃO DE MOVIMENTOS DAS MÃOS RELACIONADOS AO ESTRESSE E À ANSIEDADE

Vários estudos estabelecem uma correlação entre o nível de estresse dos motoristas com uma combinação de diferentes características, tais como, parâmetros fisiológicos, informações do veículo e informações contextuais. Apesar do volume de trabalho relacionado, poucos estudos utilizam o movimento das mãos do motorista como sinal de estresse/ansiedade em motoristas. Essas abordagens, em geral, utilizam dispositivos vestíveis com sensores de movimento para reconhecer esses movimentos (Lee et al., 2016; Lee e Chung, 2016; Choi et al., 2017).

Uma parte da abordagem proposta nessa tese que explora MMREA's já foi publicada em (Ortoncelli et al., 2020) - apesar de explorar MMREA's com base em vídeo, esse artigo não apresentou nenhuma abordagem para reconhecer automaticamente esses eventos, que foram rotulados manualmente.

Apesar de não terem sido encontrados algoritmos baseados em Visão Computacional para detecção automática de MMREA's, existem trabalhos relacionados que detectam movimentos das mãos, associados à distração do motorista (Masood et al., 2020; Craye e Karray, 2015) e também movimentos que podem ser utilizados para interação do usuário com veículos inteligentes (Manganaro et al., 2019).

Esse Capítulo apresenta os resultados obtidos com dois métodos automáticos para a detecção de dois tipos de MMREA's: i) esfregar o rosto ou o cabelo; e ii) roer as unhas. A Figura 2.6 mostrou exemplos desses movimentos. Um dos métodos é baseado em uma abordagem proposta nessa tese. O segundo método é uma adaptação do trabalho de Masood et al. (2020). As técnicas apresentadas são particularmente relevantes, pois não se encontrou abordagens para detecção do mesmo conjunto de movimentos. Detalhes sobre os dois métodos são apresentados a seguir:

- Método 1: explora a intersecção entre a região das mãos e do rosto, com base na detecção da região rosto (Baltrusaitis et al., 2018; Zadeh et al., 2017) e em técnicas de Morfologia Matemática (Serra e Soille, 2012), utilizadas para reconhecer a área das mãos dos motoristas.
- Método 2: é uma adaptação do método proposto em Masood et al. (2020) para detecção de movimentos das mãos dos motoristas relacionados a estresse/ansiedade. Utiliza uma Rede Neural Convolutiva (RNC) para detectar os movimentos. Essa técnica foi escolhida por estar no estado arte e explorar um conjunto de gestos similar aos MMREA's.

Todos os métodos apresentados classificam os MMREA's de forma binária, identificando a ocorrência ou não desses movimentos. Não foi considerada a amplitude nem a intensidade dos movimentos.

Experimentos foram realizados para comparar e avaliar os métodos apresentados para detecção de MMREA com vídeos 2D. Ambos os métodos foram comparados em termos de acurácia, revocação, precisão e *F1-score*.

O restante dessa Seção encontra-se organizado da seguinte forma. Os métodos para detecção de MMREA são detalhados nas Seções 5.1, e 5.2. Os experimentos são apresentados na Seção 5.3. Por fim, considerações finais sobre esse capítulo estão na Seção 5.4.

5.1 MÉTODO 1: BASEADO EM DETECÇÃO DA FACE E MORFOLOGIA MATEMÁTICA

O método apresentado nessa Seção explora uma característica comum aos três MMREA's analisados: o contato da mão do motorista com o rosto e/ou região dos cabelos. Com base nessa característica, foi desenvolvido um método baseado em Morfologia Matemática e reconhecimento facial para detectar MMREA's.

Inicialmente é segmentada a pele do motorista com base nos espaços de cores: RGB normalizado, HSV e YCbCr (Rahmat et al., 2016). Em seguida são aplicadas técnicas de Morfologia Matemática (Serra e Soille, 2012) para remover pequenos orifícios e conectar componentes. A área do rosto é detectada com base em pontos de referência faciais (*landmarks*) (Zadeh et al., 2017; Baltrusaitis et al., 2018). Com base na intersecção da área da face com as regiões segmentadas que representam as mãos do motorista, os MMREA's são detectados, de modo a criar um vetor V_{events} . Considerando um vídeo v que possui n quadros, $|V_{events}| = n$. Cada posição desse vetor representa a presença ou ausência de MMREA's em um quadro do vídeo. Depois de calcular V_{events} , são aplicadas operações morfológicas de abertura e fechamento neste vetor para remover pequenas falhas em um estágio de pós-processamento.

O restante dessa Seção é organizado da seguinte forma. A Subseção 5.1.1 descreve o método utilizado para segmentação da pele 5.1.1. A abordagem utilizada para segmentar as regiões que representam as mãos e o rosto do motorista é apresentada na Subseção 5.1.2. A metodologia aplicada para identificar os MMREA's é descrita na Subseção 5.1.3. Detalhes sobre a etapa de pós-processamento são apresentadas a Subseção 5.1.4 Por fim, a Subseção 5.1.5, aponta os principais problemas relacionados ao método de detecção descrito nessa Seção.

5.1.1 Segmentação da Pele

Para segmentar a pele do motorista foi utilizado o método proposto em (Rahmat et al., 2016). Esse método é baseado na segmentação combinada em todos os canais de crominância de três espaços de cores: HSV, YCbCr e RGB Normalizado.

O espaço de cor RGB consiste em três canais: vermelho (*Red* - R), verde (*Green* - G) e azul (*Blue* - B). Com base no RGB, o espaço de cores RGB Normalizado é obtido usando as Equações 5.1, 5.2 e 5.3.

$$R' = \frac{R}{R + G + B} \quad (5.1)$$

$$G' = \frac{G}{R + G + B} \quad (5.2)$$

$$B' = \frac{B}{R + G + B} \quad (5.3)$$

Tal como o RGB, o espaço de cores HSV é composto por três canais. O canal H que identifica uma cor. O canal S que representa a crominância e o canal V que representa a luminância.

O espaço de cores YCbCr possui dois elementos: luminância e a crominância. O canal Y representa luminância, que é a intensidade de luz. Os canais Cb e Cr representam a crominância, pela diferença dos componentes cromáticos azul e vermelho.

A técnica de segmentação proposta em (Rahmat et al., 2016) é baseada na união dos resultados de cinco regras de segmentação baseadas nesses espaços de cores: Equações 5.4, 5.5, 5.6, 5.7 e 5.8.

$$\frac{R'}{B'} > 0.185 \quad (5.4)$$

$$0.2 < S < 0.6 \quad (5.5)$$

$$(0 < H < 25) \cup (335 < H < 360) \quad (5.6)$$

$$77 \leq Cb \leq 127 \quad (5.7)$$

$$133 \leq Cr \leq 173 \quad (5.8)$$

Com base nessas equações, a segmentação de pele ocorre com a seguinte condição:
Equação 5.4 \cup Equação 5.5 \cup Equação 5.6 \cup Equação 5.7 \cup Equação 5.8.

5.1.2 Identificação de Objetos na Imagem

Após a etapa de segmentação da pele do motorista, para identificar as regiões segmentadas que representam as mãos e o rosto do motorista, são executadas seis etapas que são apresentadas nessa Subseção. Essas etapas são realizadas pela função `segmentImage(f_x)`, que recebe como parâmetro um quadro f_x do vídeo ($f_x \in V$). Essa função será utilizada pelo algoritmo apresentado na Subseção 5.1.3. A Figura 5.1 ilustra todas as etapas do processo de identificação das regiões em uma imagem, que são descritas a seguir:

- **Etapa 1** (Figura 5.1b): é produzida uma imagem $\mathbf{f}_x.\mathbf{skin}$ com a pele do motorista segmentada com o método apresentado na Subseção 5.1.1.
Após segmentar a pele, são retirados da imagem regiões com área inferior a α_1 e são aplicadas sequencialmente as operações morfológicas de fechamento e dilatação com um Elemento Estrutural (EE): quadrado com o tamanho $\alpha_2 \times \alpha_2$. Nos experimentos foi utilizado α_1 igual a 0,07% da área total da imagem e $\alpha_2 = 4$. Os valores de α_1 e α_2 valores foram definidos de forma empírica.
- **Etapa 2** (zoom na Figura 5.1c): é computado um vetor com 67 *landmarks* da face, denotado por: $\mathbf{f}_x.\mathbf{landmarks}$. Os pontos de referência são detectados com a ferramenta OpenFace 2.0 (Baltrusaitis et al., 2018) com o algoritmo *Convolutional Experts Constrained Local Model* (Zadeh et al., 2017). Essa ferramenta foi escolhida por apresentar bons resultados em situações de oclusão parcial e rotação facial.
- **Etapa 3** (zoom na Figura 5.1d): dois *landmarks* adicionais são computados com base no conjunto $f_x.\mathbf{landmarks}$. Esses dois pontos adicionais são adicionados ao conjunto $f_x.\mathbf{landmarks}$. Seja DL a distância entre os dois últimos pontos do lado esquerdo da face, é calculada uma linha reta que intercepta esses dois pontos e em seguida é adicionado um ponto que intercepta essa linha em uma distância $3DL$ do último ponto do lado esquerdo da face. Este procedimento é repetido com os dois últimos pontos do lado direito do rosto, gerando mais um *landmark*. Para computar os *landmarks* adicionais,

são considerados apenas os pontos da borda externa da face. o conjunto de *landmarks* passa a envolver também a área da testa do motorista.

É importante criar dois pontos de referência adicionais, pois esses pontos possibilitam uma análise mais precisa dos movimentos das mãos na área da testa e na parte superior do cabelo.

- **Etapa 4** (Figura 5.1e): é criada a imagem $f_x.\mathbf{face}$ preenchendo a região que está entre os *landmarks* computados nas etapas 2 e 3. Após preencher a área, é aplicada uma dilatação com um EE de formato quadro com tamanho $\alpha_2 \times \alpha_2$. Nos experimentos, foi utilizado $\alpha_2 = 4$, o mesmo parâmetro utilizado na etapa 1.
- **Etapa 5** (Figura 5.1f): é criada a imagem $f_x.\mathbf{aux}$ com as regiões da imagem $f_x.\mathbf{skin}$ que têm alguma intersecção com as regiões da imagem $f_x.\mathbf{face}$.
- **Etapa 6** (Figura 5.1g): é criada a imagem $f_x.\mathbf{blobHands}$ com as regiões da imagem que representam a segmentação das mãos do motorista. Essa imagem é criada com a seguinte equação: $f_x.\mathbf{blobHands} = f_x.\mathbf{skin} - f_x.\mathbf{aux}$.
- **Etapa 7** (Figura 5.1h): é criada a imagem $f_x.\mathbf{blobface}$ com uma única região que representa a face do motorista. Essa imagem é criada com a seguinte expressão: $f_x.\mathbf{blobface} = f_x.\mathbf{aux} \cup f_x.\mathbf{face}$.

5.1.3 Detecção de Movimentos das Mãos Relacionados ao Estresse e a Ansiedade

Para detectar MMREA, para cada par de quadros adjacentes $f_x, f_{x+1} \in V$, é aplicada a função $\text{computeIntersec}(f_{x+1}.\mathbf{blobFace}, f_x.\mathbf{blobHands}, f_x.\mathbf{landmarks})$. Essa função retorna o número de objetos da imagem $f_x.\mathbf{blobHands}$ que tem alguma intersecção com o objeto da imagem $f_{x+1}.\mathbf{blobFace}$. O parâmetro $f_x.\mathbf{landmarks}$ é usado para prevenir falsos positivos em relação à interação das áreas que deveriam representar a mão a face do motorista. A Figura 5.2 apresenta um exemplo de MMREA detectado.

Se uma região reg_x de $f_x.\mathbf{blobHands}$ tem alguma intersecção com a área da face na imagem $f_{x+1}.\mathbf{blobFace}$, esta intersecção só é considerada pelo método se a distância euclidiana entre um dos pontos de referência faciais em $f_x.\mathbf{landmarks}$ e pelo menos um *pixel* da região reg_x for menor ou igual a α_3 .

O limite de distância α_3 é importante porque ajuda a prevenir falsos positivos que podem prejudicar a precisão do método, como falhas de segmentação e situações em que o braço é dividido em duas regiões devido ao uso de relógios/pulseiras. Destaca-se que valores muito pequenos para α_3 podem aumentar o número de falsos negativos e valores muito grande não surtem efeito nenhum no método proposto. Nos experimentos foi utilizado $\alpha_3 = 180$, esse valor foi definido empiricamente.

A Figura 5.3 apresenta um exemplo um falso positivo relacionado a intersecção da região do rosto com a região que deve representar a(s) mão(s). Esse problema é evitado com o parâmetro α_3 . Observe que já existe uma interação ativa entre a mão e o rosto do motorista quando é detectada a nova intersecção apresentada nessa Figura.

Além de calcular o momento em que ocorre uma intersecção, o método proposto também é usado para calcular o momento em que uma intersecção acaba. A Figura 5.2 mostra um exemplo com o início e o final de uma intersecção relacionada a um MMREA.

Considerando $f_x, f_{x+1} \in V$, para calcular o momento em que se inicia uma intersecção são utilizados os seguintes parâmetros para a função $\text{computeIntersec}(f_{x+1}.\mathbf{face}, f_x.\mathbf{hands},$

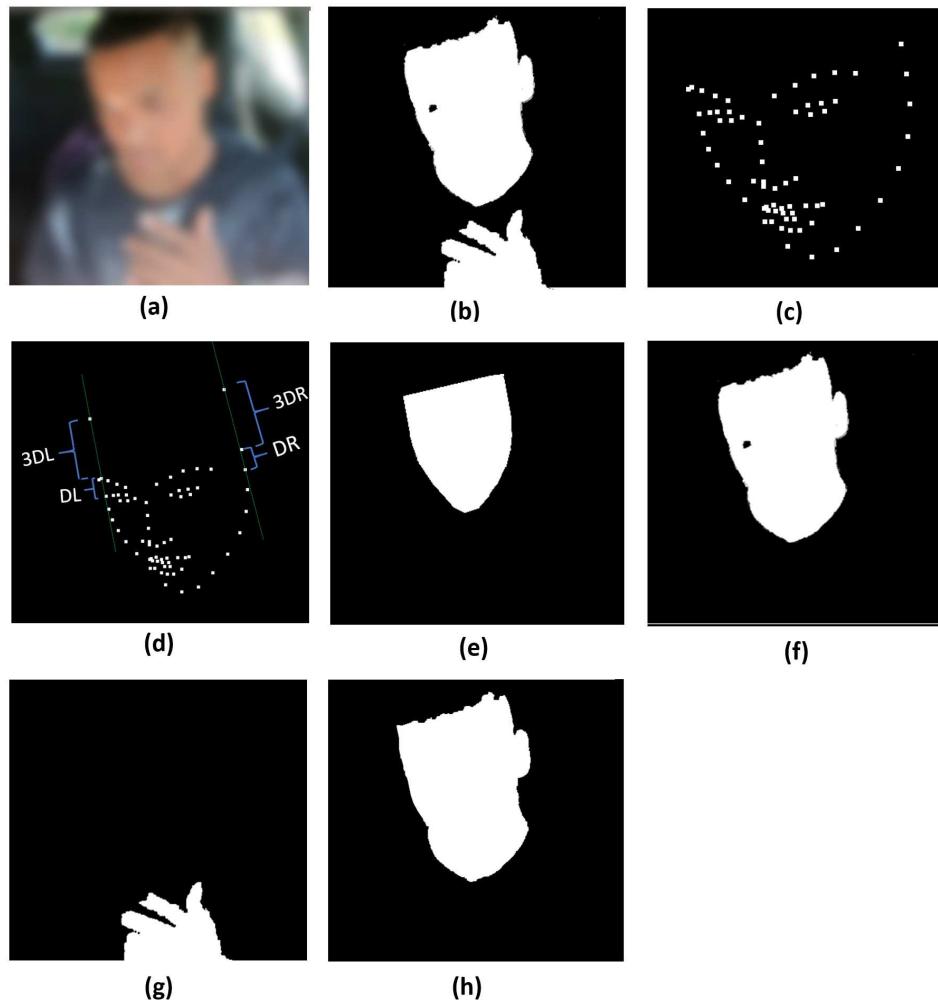


Figura 5.1: Processo de detecção de mãos e face nas imagens com o método 1



Figura 5.2: Exemplo de intersecção entre as regiões que representam a mão e a face do motorista

$f_x.landmarks$). Para computar o momento em que as intersecções se encerram são utilizados os seguintes parâmetros $computeIntersec(f_x.face, f_{x+1}.hands, f_{x+1}.landmarks)$.

O Algoritmo 1 apresenta o pseudocódigo do método proposto nessa Seção para detectar MMREA. O algoritmo recebe como parâmetro um vídeo V e retorna um vetor V_{events} . Para detectar os movimentos o algoritmo faz uso de duas funções: i) $computeIntersec$, que já foi

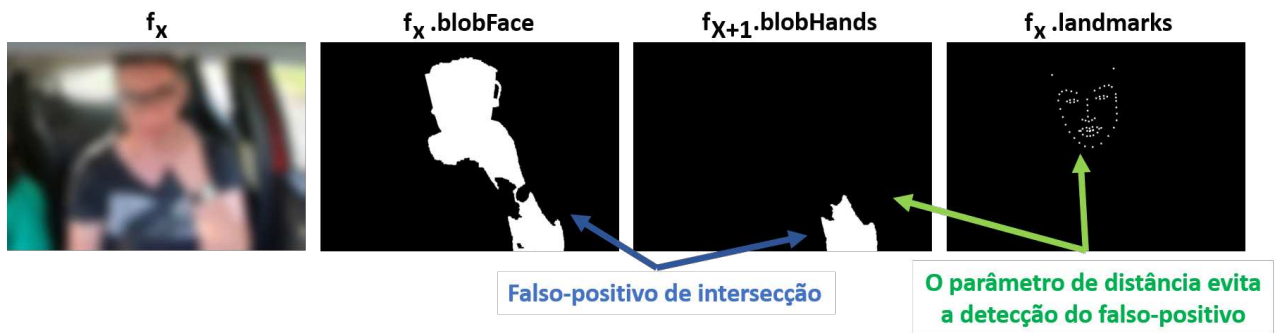


Figura 5.3: Exemplo de falso positivo na detecção de MMREA

explicada nesta Subsecção; e ii) `segmentImage` (f_x), que recebe uma imagem f_x e retorna um vetor de *landmarks* faciais e as imagens $f_x.hands$ e $f_x.face$, computadas com o processo apresentado na Subsecção 5.1.2.

No Algoritmo 1, é usada uma variável chamada “aux” para armazenar o número de interações ativas em um determinado quadro do vídeo. A variável “aux” recebe o valor 0 se não há nenhuma intersecção ativa, 1 se existe a intersecção de uma das mãos com o rosto, 2 se existe a intersecção de duas das mãos com o rosto. A variável pode receber valores maiores que 2, caso existam erros na detecção de MMREA. As principais causas de erros no algoritmo proposto nessa Seção são apresentadas na Subsecção 5.1.5.

Quatro medidas preventivas são tomadas para mitigar o impacto dos erros nos resultados do Algoritmo 1: i) a utilização do parâmetro de distância α_3 ; ii) não permitir que a variável “aux” receba um valor negativo; iii) uma etapa de pós-processamento descrita na Subsecção 5.1.4; e iv) estabelecer um tamanho máximo para os eventos detectados, conforme descrito a seguir.

Para limitar o tamanho dos eventos detectados, é utilizada uma variável chamada “count” que mantém o número de quadros consecutivos nos quais uma intersecção está ativa. Se o valor da variável “count” for igual a 1000 é interrompida a intersecção ativa, ou seja, é atribuído zero como o valor das variáveis “aux” e “count”.

O tamanho máximo para os eventos é utilizado para evitar problemas relacionados a não detecção do quadro no qual acaba a intersecção da mão do motorista com o rosto. Na base de dados experimental, não foi encontrado nenhum MMREA com duração superior a 1000 quadros, por isso esse valor foi selecionado como valor máximo possível para a variável “count”.

5.1.4 Pós-processamento

A etapa de pós-processamento é mais um cuidado tomado para mitigar problemas relacionados à detecção de MMREA's. Após computar o vetor V_{events} com o Algoritmo 1, a etapa de pós-processamento é aplicada para remover pequenas e grandes regiões adjacentes.

Seja V_{events} um vetor unidirecional composto de valores binários que indicam a presença (valor 1) ou não (valor 0) de um MMREA. Na etapa de pós-processamento, são executadas sequencialmente três operações em V_{events} :

- **Operação 1:** Atribuir o valor 1 aos componentes conexos do vetor V_{events} com valor 0 e tamanho menor ou igual a 10. Equivale à aplicação de uma operação de fechamento morfológico com um EE no formato de linha com tamanho 11.

Algoritmo 1 Detecção de movimentos das mãos dos motoristas relacionados a estresse/ansiedade

Require: *video*

```

1:  $V_{events} = []$ 
2:  $L1, F1, H1 = \text{segmentImage}(\text{video}[0])$ 
3:  $aux = 0$ 
4:  $count = 0$ 
5: for  $i = 1$  to  $\text{len}(\text{video})$  do
6:    $L2, F2, H2 = \text{segmentImage}(\text{video}[i])$ 
7:    $aux += \text{computeIntersec}(F2, H1, L1)$ 
8:    $aux -= \text{computeIntersec}(F1, H2, L2)$ 
9:   if  $(aux < 0)$  or  $(count == 1000)$  then
10:      $aux = 0$ 
11:      $count = 0$ 
12:   end if
13:   if  $aux > 0$  then
14:      $V_{events}.append(1)$   $count += 1$ 
15:   else
16:      $V_{events}.append(0)$ 
17:      $count = 0$ 
18:   end if
19:    $L1, F1, H1 = L2, F2, H2$ 
20: end for
21: return  $V_{events}$ 

```

- **Operação 2:** Atribuir o valor 0 aos componentes conexos do vetor V_{events} com valor 1 e tamanho menor ou igual a 3. É equivalente a aplicar uma operação de abertura morfológica com um EE no formato de linha com tamanho 4.
- **Operação 3:** Atribuir o valor 0 aos componentes conectados com o valor 1 e tamanho maior ou igual a 1000.

A Figura 5.4 apresenta um exemplo das operações realizadas na etapa de pós-processamento. Observe que para a terceira operação não é apresentado um exemplo de remoção de componentes conexos, pois não foi possível apresentar de forma legível um vetor de 1000 ou mais posições.

5.1.5 Principais problemas

Algumas situações podem causar erros nos resultados do método proposto. Pode-se destacar as quatro principais causas de erros que foram observadas durante os experimentos:

1. **Tipos de roupas:** com tipos de roupas específicos, uma região maior de pele pode ser segmentada junto com a área do rosto ou das mãos do motorista. Esse problema pode ser amenizado com o controle das roupas utilizadas. Nos experimentos, foram três tipos de roupas que causam esse problema: i) decote largo; ii) camiseta regata; e iii) roupa da cor da pele.
2. **Tipos de cabelos:** quando o motorista tem os cabelos longos na cor de pele, o método de segmentação pode segmentá-los em conjunto com área da face, o que pode impactar em

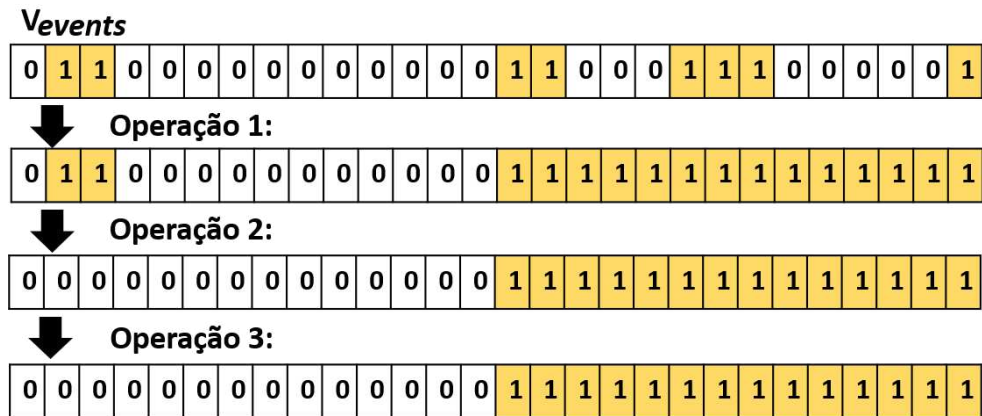


Figura 5.4: Etapa de pós-processamento do método 1 de detecção de MMREA

erros no cálculo da intersecção entre a área das mãos e a área do rosto. Esse problema pode ser controlado se o cabelo for preso na parte de trás da cabeça.

3. **Condições de iluminação:** a exposição à luz solar pode causar falhas na segmentação da pele. Apesar dessas falhas, a condição de iluminação afetou pouco os resultados experimentais obtidos com a abordagem proposta para detecção de MMREA's.
4. **Relógios ou braceletes:** Com esses itens, a região relacionada à mão/braço do motorista pode ser segmentada em duas regiões. Para minimizar o número de falsos positivos e falsos negativos em virtude dessa super segmentação, é utilizado o critério de distância (α_3) no Algoritmo 1. Um exemplo de um problema relacionado a relógios/pulseiras é apresentado na Figura 5.3.

O casos relacionados ao tipo de roupa e penteado causaram problemas mais sérios nos resultados experimentais. A Figura 5.5 apresenta exemplos dessas situações. Destaca-se, que os métodos 2 e 3 foram menos sucessíveis a esse grupo de problemas, de acordo com os resultados experimentais.



Figura 5.5: Principais problema do método 1 para detectar MMREA

5.2 MÉTODO 2: BASEADO EM REDE NEURAL CONVOLUCIONAL

Essa Seção descreve o método proposto em (Masood et al., 2020) para reconhecer movimentos de distração de motoristas, que é baseado em *Deep Learning*. Essa técnica foi implementada para permitir a comparação dos métodos de detecção de MMREA's propostos nessa tese com uma técnica no estado da arte.

Pelo fato de não ter sido encontrado literatura trabalhos que reconheçam os mesmos MMREA's utilizados nessa tese, foi selecionado o trabalho de Masood et al. (2020), que reconhece gestos com características semelhantes (as mãos dos motoristas se aproximarem do seu rosto), com uma acurácia superior a 98% - um conjunto de dez ações foi reconhecido originalmente nesse trabalho: i) direção segura; ii) digitar no celular com a mão direita; iii) digitar no celular com a mão esquerda; iv) conversar no celular usando a mão direita; v) conversar no celular usando a mão esquerda; vi) operar o rádio do carro; vii) beber em um copo; viii) alcançar algo no banco de trás; ix) arrumar o cabelo ou se maquiar; e x) interagir com o passageiro.

Com a técnica apresentada nessa Seção uma imagem RGB é passada como parâmetro para uma RNC que reconhece qual tipo de movimento está ocorrendo. Duas arquiteturas de RNC são utilizadas: VGG16 e VGG19 - dois modelos que foram propostos originalmente por Simonyan e Zisserman (2014). Para comparar os resultados obtidos com esse método em iguais condições com o método 1 (Seção 5.1), a etapa de pós processamento também foi aplicada. Como nesse método são utilizados quatro rótulos e não uma classificação binária dos eventos, uma adaptação foi realizada na atividade de pós processamento.

O restante dessa seção apresenta detalhes desse método. Os modelos VGG16 e VGG19 são detalhados nas Subseções 5.2.1 e 5.2.1, respectivamente. A Subseção 5.2.3, mostra as atividades realizadas para criar os conjuntos de teste e treinamento. A abordagem de validação cruzada realizada nos experimentos é descrita na Subseção 5.2.4. Por fim, o pós-processamento realizado é apresentado na Subseção 5.2.5.

5.2.1 Modelo VGG16

O VGG16 é uma RNC que recebe como entrada uma imagem RGB com dimensões de 224×224 pixels. Essa imagem passa por uma pilha de camadas convolucionais, com filtros com um campo receptivo de tamanho 3×3 .

A pilha de camadas convolucionais é seguida por três camadas totalmente conectadas: as duas primeiras camadas têm um total de 4096 canais cada, e a terceira camada executa uma classificação ILSVRC de 1000 vias. A camada final é a camada SOFTMAX. Todas as camadas ocultas são equipadas com retificação não linear (Hahnloser et al., 2000).

A macro arquitetura do modelo VGG16 é apresentado na Figura 5.6.

5.2.2 Modelo VGG19

O modelo VGG19 é muito similar com o modelo VGG16, mas como seu próprio nome sugere, é uma rede composta por 19 camadas. A macro arquitetura do modelo VGG19 é apresentado na Figura 5.7.

Na VGG19, três camadas convolucionais adicionais são adicionadas à pilha. Das três camadas RNC apresentadas, uma tem profundidade de 256, enquanto as outras duas camadas têm profundidade de 512 cada. Não há mudança nas camadas totalmente conectadas restantes da rede, em relação a VGG16.

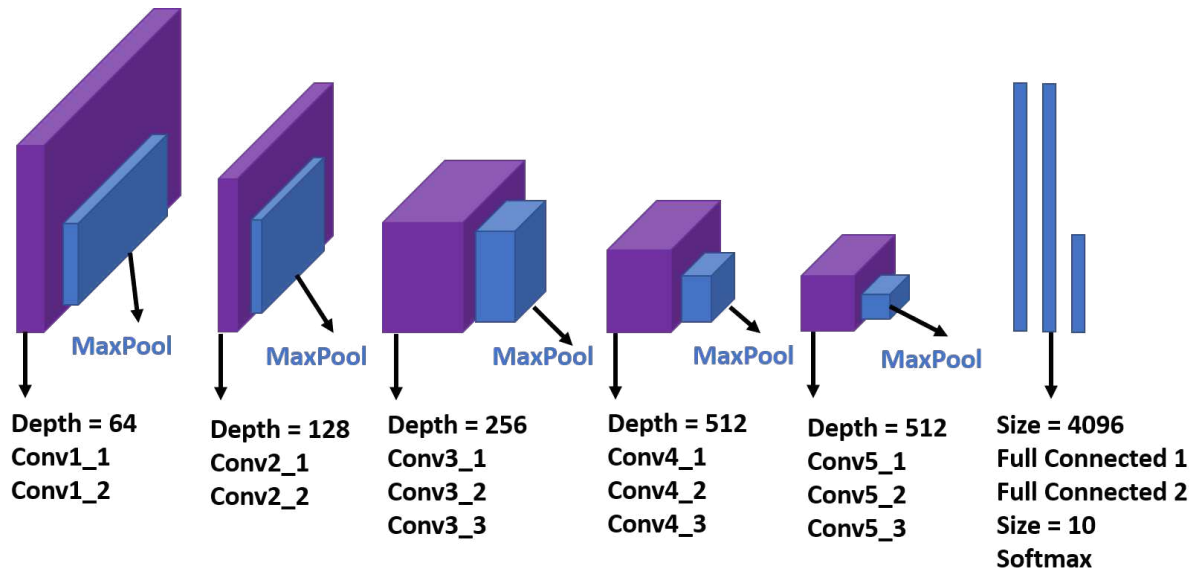


Figura 5.6: Macro arquitetura da RNC VGG16 (Simonyan e Zisserman, 2014) - baseado em (Masood et al., 2020)

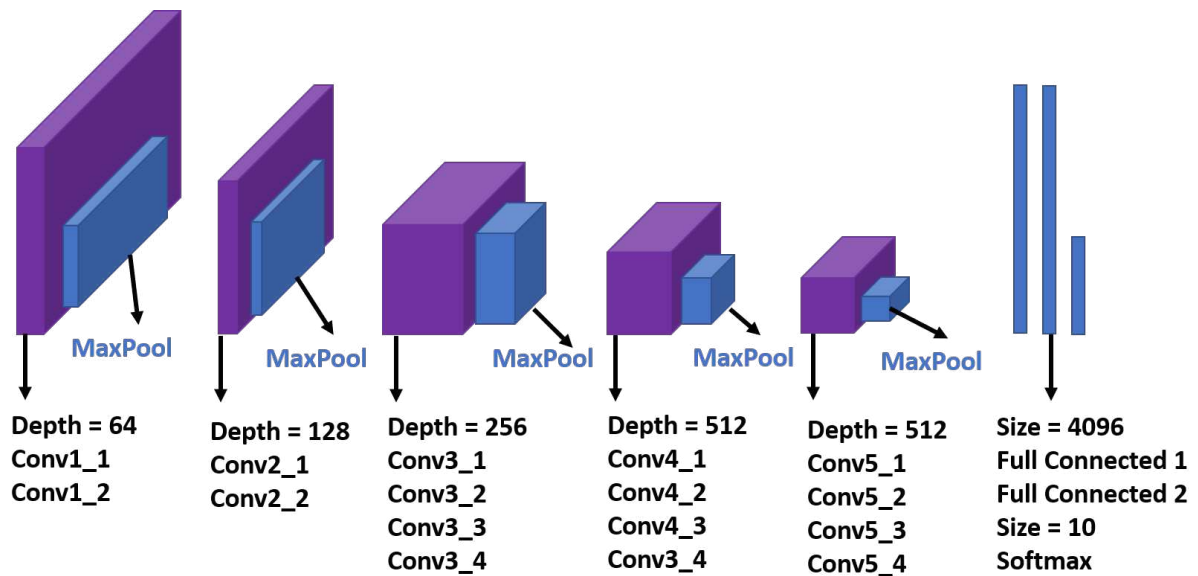


Figura 5.7: Macro arquitetura da RNC VGG19 (Simonyan e Zisserman, 2014) - baseado em (Masood et al., 2020)

5.2.3 Criação dos conjuntos de teste e treinamento

Para criar a base de dados utilizada para treinar os modelos VGG16 e VGG19, em um primeiro momento foram extraídos dos vídeos todos os quadros referentes a instantes de tempo nos quais ocorreram MMREA's. Essas imagens foram rotuladas com três rótulos distintos: i) MMREA's realizados com a mão direita; ii) MMREA's realizados com a mão esquerda; e iii) MMREA's realizados com às duas mãos. Além disso, também se utilizou um quarto rótulo, relacionado a ausência de MMREA's, com quadros que representam momentos nos quais não ocorreram MMREA's.

A abordagem de dividir movimentos de acordo com a mão que realiza a ação foi explorada em Masood et al. (2020). Devido ao fato dos MMREA's explorados nesse trabalho terem

características semelhantes, em relação ao tipo e amplitude dos movimentos das mãos/braços, optou-se por seguir essa estratégia.

Para cada vídeo da base de dados experimental, foi marcada manualmente a região na qual o corpo do motorista está centralizado. Um retângulo representa a região selecionada, cuja largura é equivalente a 1,3 vezes o comprimento da sua altura. Essas são as mesmas dimensões da base de dados experimental utilizada em Masood et al. (2020). Como para cada veículo a câmera é posicionada em uma posição fixa, para cada carro é necessário realizar essa marcação apenas uma única vez.

Após definir a posição da região selecionada nos vídeos, essa área foi recortada de cada um dos quadros dos vídeos, sendo redimensionada para o tamanho de 224 por 224 *pixels* - tamanho das imagens de entrada nos modelos VGG16 e VGG19.

Para cada vídeo foi selecionada uma quantidade q de quadros referentes a MMREA's, sendo q igual a quantidade de quadro com MMREA's. Esses quadros foram divididos em três subconjuntos $|Q1|$, $|Q2|$ e $|Q3|$, de acordo com o tipo de movimento realizado. Um conjunto $|Q4|$ foi criado com quadros que não apresentam MMREA's. Para cada vídeo, os quadros que compõe Q4 foram selecionados aleatoriamente, sendo $|Q4|$ igual ao maior valor entre $|Q1|$, $|Q2|$ e $|Q3|$.

Quanto aos conjuntos utilizados para testes, não foi aplicado nenhum filtro em relação à quantidade de quadros selecionados. Todos os quadros dos vídeos que não fizeram parte do conjunto de treinamento foram utilizados para testar o modelo treinado.

Diferentes instâncias dos conjuntos de teste e treinamento foram criadas, de acordo com as regras de validação cruzada, descritas na Seção 5.2.4.

Como a base de dados experimental possui mais de 2200 horas de vídeos gravados a uma taxa de 60 quadro por segundo, essa base conta com mais de 8 milhões de imagens. Apesar da quantidade de imagens, apenas uma pequena parte delas representa gestos das mãos relacionados a situações ao estresse/ansiedade. Além disso, devido ao fato de os vídeos terem sido gravados a uma taxa de 60 quadros por segundo, quadros adjacentes contém imagens similares. Para evitar problemas relacionados à quantidade de imagens com MMREA's, foram aplicadas técnicas para aumentar os dados. O processo de aumento de dados é descrito na Subseção 5.2.3.1.

Destaca-se que uma etapa de pré-processamento foi aplicada as imagens antes de aplicar o treinamento e reconhecimento de MMREA's com as RNC's. Detalhes sobre o pré-processamento são apresentados na Subseção 5.2.3.2.

5.2.3.1 Aumento de Dados

Redes neurais profundas requerem um conjunto de dados grande e diversificado para treinar o modelo de forma eficaz. A etapa de Aumento de Dados (do inglês, *Data Augmentation*), como o próprio nome sugere, consiste em aumentar o número de dados de treinamento de um modelo, para que se possa treiná-lo ainda mais e evitar o problema de sobreajuste (*overfitting* ou especialização de um modelo) (Shorten e Khoshgoftaar, 2019).

O *overfitting* ocorre quando o modelo está muito adaptado às imagens com as quais foi treinado, porém não é capaz de generalizar bem para novos dados. Em outras palavras, ocorre quando o modelo “decorou” o conjunto de dados de treino, mas não é capaz de diferenciar adequadamente os dados quando precisa classificar um novo conjunto de imagens.

Nessa tese foi utilizada a mesma abordagem de aumento de dados aplicada em (Masood et al., 2020). Em cada época de treinamento dos modelos, um novo conjunto de imagens foi gerado com diversas variações: intervalos máximos para graus de cisalhamento, zoom, deslocamento horizontal e vertical foram especificados. Valores aleatórios foram aplicados para cada um dos parâmetros mencionados, gerando novas imagens, mas conservando a classe/categoria da imagem.

5.2.3.2 Pré Processamento das Imagens

A etapa de treinamento envolve a multiplicação de pesos e a adição de vieses às entradas iniciais, que são então retro propagadas (Rumelhart et al., 1986) com gradientes para treinar os modelos. Para evitar que os gradientes fiquem fora de intervalos específicos é importante que cada recurso tenha valores semelhantes. Além disso, RNC's envolvem o compartilhamento de parâmetros e se as entradas não forem dimensionadas para que tenham valores de faixa semelhante, o compartilhamento não acontecerá facilmente. Para mitigar esse problema, o valor médio RGB de todos os *pixels* de todas as imagens da base de dados é subtraído de cada um dos *pixels* de cada uma das imagens, antes de realizar as atividades de treinamento e classificação.

5.2.4 Validação cruzada

A validação cruzada é uma técnica aplicada para avaliar a capacidade de generalização de um modelo (Kohavi et al., 1995). Consiste em treinar e testar modelos com diferentes subconjuntos de um mesmo conjunto original, com o objetivo de obter uma medida mais confiável sobre a capacidade do modelo.

Nesse trabalho foi explorada uma técnica baseada na validação cruzada *K-fold* (Blum et al., 1999), que consiste em dividir os dados em k subconjuntos mutuamente exclusivos de igual tamanho. A partir dessa divisão, são realizadas k instâncias do processo de treinamento e teste, sendo que, em cada uma delas, um conjunto k distinto é utilizado para teste e os $k-1$ conjuntos restantes são aplicados para treinar o modelo.

Nos experimentos realizados nesse trabalho, a base de dados foi dividida em $k = 6$ subconjuntos. Como a base conta com 60 vídeos, cada um dos k conjuntos é referente aos quadros extraídos de 10 vídeos. Dessa forma, cada um dos conjuntos representa a mesma quantidade de vídeos, mas não necessariamente tem a mesma quantidade de imagens, pois cada motorista realizou uma quantidade distinta de MMREA's. A Figura 5.8 apresenta o esquema de validação cruzada realizado.

5.2.5 Pós-processamento

Para validar os resultados obtidos com o método 3 em igualdade condições com os resultados dos métodos 1, a etapa de pós-processamento apresentado na Subseção 5.1.4, também foi aplicada.

Considere-se um vetor unidirecional no qual o valor de cada posição v_i representa o evento rotulado no quadro v_i do vídeo com o modelo treinado. Cada posição desse vetor pode ter três rótulos: 1) referente à presença de MMREA's realizados com a mão direita; 2) referente à presença de MMREA's realizados com a mão esquerda; 3) referente à presença de MMREA's realizados com às duas mãos; e 4) referente à ausência de MMREA's;

Esse vetor de resultados é binarizado iterativamente três vezes. Na primeira iteração, posições com o valor um (1) se mantém com o valor 1 e as demais recebem o valor zero (0). Na segunda iteração, posições com o valor dois (2) recebem o valor 1 e as demais recebem o valor 0. Na última iteração, posições com o valor 2 recebem o valor três (3) e as demais recebem o valor 0.

Em cada uma dessas iterações, o método de pós-processamento é aplicado no vetor. As métricas utilizadas para mensurar os resultados experimentais são aplicadas após a etapa de pós-processamento.

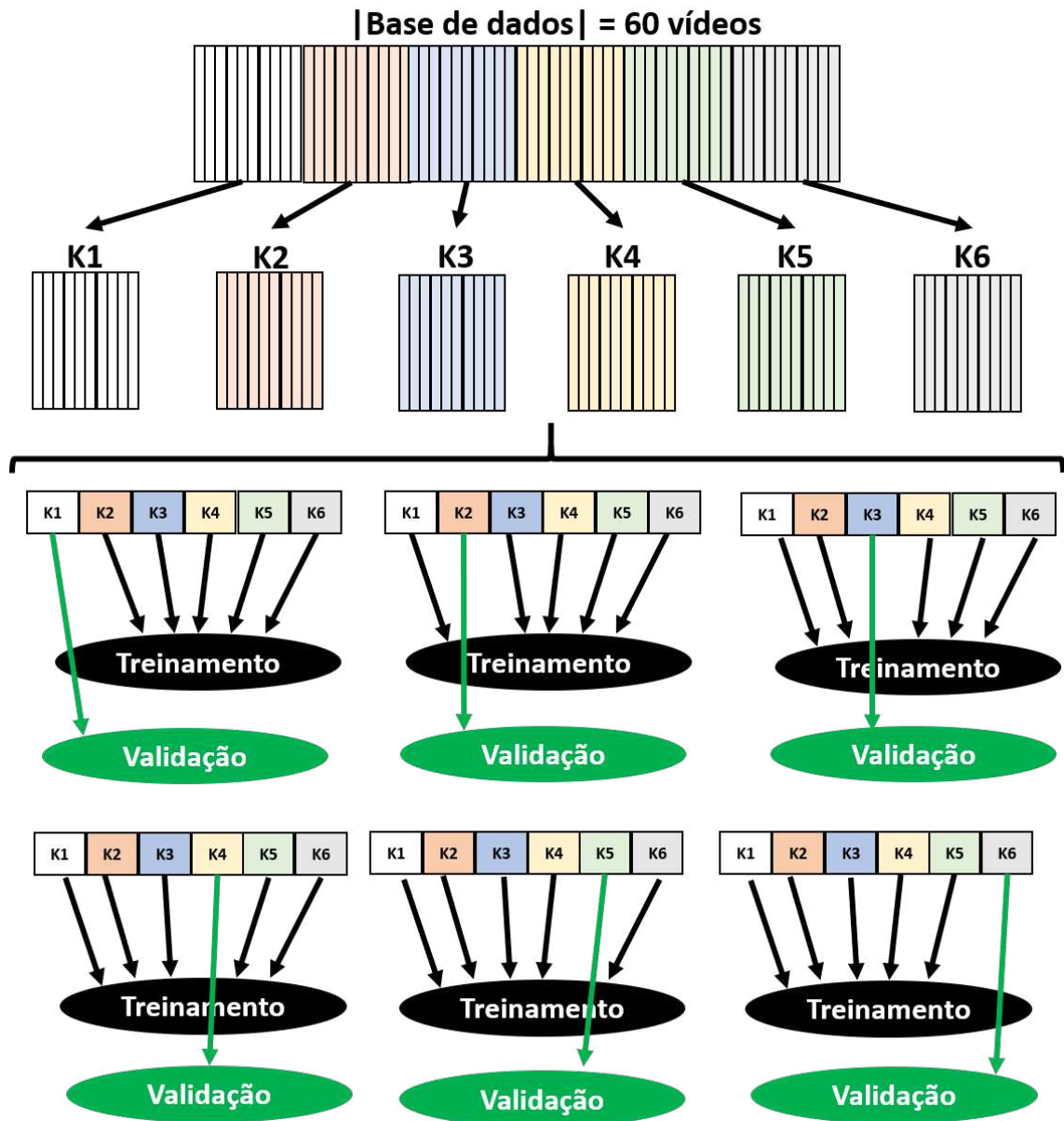


Figura 5.8: Esquema de validação cruzado utilizado

5.3 EXPERIMENTOS

Para conduzir os experimentos os métodos 1 e 2 foram implementados utilizando a linguagem de programação Python e as bibliotecas Numpy (Oliphant, 2006) e OpenCV (Howse, 2013). Quanto ao método 3, as duas arquiteturas utilizadas foram implementadas com as API's Keras (Gulli e Pal, 2017) e TensorFlow (Abadi et al., 2016). Destaca-se que todos os experimentos apresentados nesse Capítulo foram realizados com os vídeos coletados com a Câmera A (Figura 4.1).

Detalhes e resultados experimentais são apresentados nessa Seção. A Subseção 5.3.1 descreve as métricas utilizadas para avaliar os métodos propostos. Os resultados experimentais são apresentados na Subseção 5.3.2. A Subseção 5.3.3 apresenta uma análise dos resultados.

5.3.1 Métricas de Avaliação

Foram utilizadas quatro métricas para avaliar os resultados experimentais: i) precisão (a fração de instâncias relevantes entre as instâncias recuperadas); ii) revocação (a fração do total de instâncias relevantes que foram recuperadas); iii) acurácia (uma proporção de itens previstos corretamente); e iv) *F1-Score* (ou *f-measure*, é uma média harmônica entre precisão e revocação). Quatro parâmetros foram usados para calcular essas medidas:

- Número de Verdadeiros Positivos (NVP): número de imagens com MMREA nas quais o MMREA foi detectado;
- Número de Falsos Positivos (NFP): número de imagens com MMREA nas quais MMREA não foi detectado;
- Número de Verdadeiros Negativos (NVN): número de imagens sem MMREA nas quais não foi detectado MMREA;
- Número de Falsos negativos (NFN): número de imagens sem MMREA para as quais foi detectado MMREA;

A precisão, a revocação, a acurácia, e o *F1-Score* são computados com as equações 5.9, 5.10, 5.11 e 5.12.

$$precisão = \frac{NVP}{NVP + NFP} \quad (5.9)$$

$$revocação = \frac{NVP}{NVP + NFN} \quad (5.10)$$

$$acurácia = \frac{NVP + NVN}{NVP + NFP + NFN + NVN} \quad (5.11)$$

$$F1-Score = 2 * \frac{precisão * revocação}{precisão + revocação} \quad (5.12)$$

5.3.2 Resultados Experimentais

Essa Seção apresenta os resultados experimentais obtidos com os três métodos descritos para detecção de MMREA. Os experimentos foram conduzidos com os vídeos da base de dados experimental (Capítulo 4) obtidos com a câmera A. Os resultados de cada instância experimental foram mensurados com as métricas apresentadas na Subseção 5.3.1.

O métodos 1 classifica apenas a ocorrência ou não de MMREA's, já o método 2 utiliza *Deep Learning* para identificar um gesto entre quatro possíveis ações, sendo as três primeiras relacionadas à presença de MMREA's e a quarta relacionada à ausência de MMREA's. Como nessa tese o problema de sumarização e priorização de atividades de direção considera apenas a presença de MMREA's, independentemente de seu tipo, nos resultados apresentados nas Tabelas 5.1 a 5.4, os três rótulos relacionados a ocorrência de MMREA's (do método 2) foram considerados como um único rótulo.

A Tabela 5.1 apresenta os resultados obtidos com cada um dos métodos, considerando 100% da base de dados.

Tabela 5.1: Resultados experimentais para 100% da base de dados

Método	Precisão	Revocação	Acurácia	<i>F1-Score</i>
Método 1	0,642	0,301	0,986	0,410
Método 2 + VGG16	0,978	0,115	0,942	0,206
Método 2 + VGG19	0,989	0,249	0,977	0,398

Pode-se observar que duas características prejudicaram os resultados do método 1: o tipo de roupa e o penteado. Detalhes sobre esses problemas estão na Subseção 5.1.5. Devido a essas condições, dois conjuntos de vídeos foram utilizados nos experimentos: i) C_{100} , que contém todos os vídeos da base de dados; e C_{75} , que é composto pelos 75% dos vídeos para os quais não se identificou problemas relacionados a roupas e penteados, ou seja, vídeos nos quais aparência do motorista não se encaixa nos dois primeiros problemas descritos na Subseção 5.1.5.

A Tabela 5.2 apresenta os resultados obtidos com cada um dos métodos considerando apenas testes realizados com o conjunto C_{75} : o método 1 foi testado apenas com os vídeos do conjunto C_{75} e o método 2 foi treinado com os vídeos do conjunto e C_{100} e testado apenas com os vídeos contidos em C_{75} .

Tabela 5.2: Resultados experimentais para 75% da base de dados*

Método	Precisão	Revocação	Acurácia	<i>F1-Score</i>
Método 1	0,793	0,790	0,997	0,792
Método 2 + VGG16	0,979	0,115	0,941	0,206
Método 2 + VGG19	0,991	0,246	0,976	0,394

* Não considerando testes com vídeos com roupas e penteados descritos na Subseção 5.1.5

Foi observado também que muitos erros de detecção ocorrem nos três quadros que precedem ou seguem o início e o fim de cada MMREA. Este tipo de erro causa pouco impacto na visualização dos eventos detectados em vídeos gravados a 60 quadros por segundo. Por esse motivo, foram realizados experimentos considerando e sem considerar erros nos três quadros iniciais e finais de cada MMREA.

As Tabelas 5.3 e 5.4 apresentam os resultados experimentais sem penalizar erros nos três quadros iniciais e finais de cada MMREA. A Tabela 5.3 apresenta os resultados para 100% da base de dados. A Tabela 5.4 apresenta os resultados obtidos com os testes conduzidos com o conjunto de 75% dos vídeos que não apresentaram situações problemáticas para o método 1.

Tabela 5.3: Resultados experimentais para 100% da base de dados: não penalizando erros nos três quadros que precedem ou seguem o início ou fim de cada MMREA

Método	Precisão	Revocação	Acurácia	<i>F1-Score</i>
Método 1	0,650	0,307	0,986	0,418
Método 2 + VGG16	0,982	0,122	0,943	0,216
Método 2 + VGG19	0,992	0,255	0,977	0,406

Tabela 5.4: Resultados experimentais para 75% da base de dados *: não penalizando erros nos três quadros que precedem ou seguem o início ou fim de cada MMREA

Método	Precisão	Revocação	Acurácia	F1-Score
Método 1	0,806	0,806	0,997	0,806
Método 2 + VGG16	0,982	0,122	0,941	0,216
Método 2 + VGG19	0,992	0,257	0,977	0,408

* Não considerando testes com vídeos com roupas e penteados descritos na Subseção 5.1.5

Para possibilitar uma compreensão adequada do comportamento dos modelos treinados com o método 2, matrizes de confusão apresentam os resultados obtidos para cada rótulo de forma individual. As Tabelas 5.5 e 5.6 apresentam as matrizes de confusão dos resultados obtidos com modelos VGG16 e VGG19 considerando 100% da base de dados.

Tabela 5.5: Matriz de confusão dos resultados do método 2 com VGG16 - 100% da base de dados

	Rótulo 1	Rótulo 2	Rótulo 3	Rótulo 4
Rótulo 1: MMREA's - mão esquerda	23147	816	1428	259417
Rótulo 2: MMREA's - mão direita	1917	31916	1046	213665
Rótulo 3: MMREA's - duas mãos	35	78	1671	4416
Rótulo 4: sem MMREA's	898	187	251	7850369

Tabela 5.6: Matriz de confusão dos resultados do método 2 com VGG19 - 100% da base de dados

	Rótulo 1	Rótulo 2	Rótulo 3	Rótulo 4
Rótulo 1: MMREA's - mão esquerda	24869	792	1457	97137
Rótulo 2: MMREA's - mão direita	614	31945	959	90947
Rótulo 3: MMREA's - duas mãos	32	105	1895	1105
Rótulo 4: sem MMREA's	462	155	105	8138698

No método 2, os modelos VGG16 e VGG19 classificaram a presença de quatro grupos de movimentos (quatro rótulos), porém, o método 1 reconheceu apenas de forma binária a presença de MMREA's. Para comparar a matriz de confusão do método 1 com a matriz de confusão do método 2, os rótulos 1, 2 e 3 classificados pelo método 3 foram computados como uma única classificação relacionada à presença de MMREA's.

As tabelas 5.7 e 5.8 apresentam as matrizes de confusão dos dois métodos apresentados nessa Seção para classificação de MMREA's. A Tabela 5.7 apresenta os resultados considerando 100% da base de dados. Já a Tabela 5.8 apresenta os resultados dos testes realizados com 75% da base de dados - não considerando os vídeos nos quais os motoristas estão com os tipos de roupas e penteados descritos na Subseção 5.1.5.

Tabela 5.7: Matrizes de confusão da classificação binária dos resultados para 100% da base de dados

Método	Rótulo	MMREA	sem MMREA
Método 1	MMREA	40700	94626
	sem MMREA	22690	8233261
Método 2 + VGG16	MMREA	62054	477488
	sem MMREA	1366	7850369
Método 2 + VGG19	MMREA	62668	189189
	sem MMREA	722	8138698

Tabela 5.8: Matrizes de confusão da classificação binária dos resultados para 75% da base de dados *

Método	Rótulo	MMREA	sem MMREA
Método 1	MMREA	38681	10278
	sem MMREA	10069	6181384
Método 2 + VGG16	MMREA	47702	365765
	sem MMREA	1048	5825897
Método 2 + VGG19	MMREA	48289	147840
	sem MMREA	461	6043822

* Não considerando testes com vídeos com roupas e penteados descritos na Subseção 5.1.5

5.3.3 Análise dos Resultados

Em relação à acurácia, todos os métodos analisados apresentaram resultados acima de 0,941. Destaca-se que os valores de acurácia são consideravelmente superiores aos computados para a precisão. Isso indica que, embora as técnicas apresentem falhas na detecção de alguns dos eventos (ou seja, falsos positivos), elas têm uma capacidade maior de reconhecer momentos em que não ocorreram MMREA's (verdadeiros negativos).

Quanto à revocação, acurácia e *F1-score*, quando se utilizou 100% da base de dados, os melhores resultados foram obtidos com o método 1, que alcançou resultados de revocação de até 0,307, de acurácia igual a 0,986, e *F1-score* de até 0,418. Destaca-se que os resultados obtidos com o método 2 utilizando o modelo VGG19 foram próximos aos melhores resultados obtidos. Quando não foram considerados testes com vídeos com condições problemáticas para o método 1 (relacionadas ao tipo de roupa ou penteado), os resultados obtidos com o método 1 melhoram para uma revocação de 0,806, acurácia de 0,997 e *F1-score* de 0,806.

Com relação à precisão, os melhores resultados foram obtidos com o método 2 (sendo sempre superiores a 0,978), isso significa que os modelos treinados com Redes Neuras Convolucionais apresentaram a melhor capacidade em relação a não detectar falsos positivos.

Nesse contexto, para a maioria dos vídeos da base de dados experimental, os melhores resultados, em termos de revocação, acurácia e *F1-score*, foram obtidos com o método 1. Para mitigar o número de erros para esse método, deve-se ter cuidado com o tipo de roupa e penteado usados pelos motoristas. Caso não seja possível controlar essas variáveis, o usuário deve optar por utilizar o método 2.

Apesar do método 2 apresentar os melhores em relação à precisão, ele apresentou os piores resultados para as outras métricas avaliadas. Tal como pode ser observado nas matrizes de confusão, os resultados ruins se devem principalmente ao número alto de falsos positivos. Acredita-se que esses resultados podem ser melhorados em trabalhos futuros, nos quais se pode utilizar um novo conjunto de imagens, para aumentar a representatividade da base de dados utilizada para o treinamento dos modelos. Além disso, pode-se explorar outras técnicas de *Deep Learning*.

Os resultados de todos os métodos melhoram quando não são considerados erros nos três quadros que precedem ou seguem o início e o final de cada um dos MMREA rotulados na base de dados. Isso é positivo porque esse erro tem pouco impacto na interpretação dos resultados, considerando vídeos gravados a 60 quadros por segundo, que é o caso da base de dados experimental.

Muitos dos erros ocorreram no início ou no final de cada evento rotulado. Isso se deve ao fato de que os métodos trabalham com o conceito de intersecção entre a mão e a área do rosto em imagens 2D. Essa intersecção não começa necessariamente quando a mão aparece à frente do

rosto, por uma questão de perspectiva e profundidade. Em trabalhos futuros, este problema pode ser superado com imagem de profundidade.

Existem gestos com características similares aos MMREA's que são falso positivos, por exemplo, quando o motorista passa a mão na frente do rosto para apontar para algo. Embora esse tipo de falso positivo exista na base de dados, apenas uma pequena porcentagem dos erros dos experimentos se deve a esse tipo de problema. A maioria dos problemas está relacionada a falhas de segmentação, erros na detecção das mãos ou *landmarks* faciais e as situações apresentadas na Subseção 5.1.5. Trabalhos futuros podem melhorar os resultados, identificando esse tipo de falsos positivos.

5.4 CONSIDERAÇÕES FINAIS

Essa seção apresentou os métodos propostos para detectar automaticamente MMREA's em vídeos 2D. Dois gestos são identificados com os métodos propostos: i) esfregar o rosto ou o cabelo; e ii) roer as unhas.

Nos experimentos realizados, para a maioria dos vídeos, os melhores resultados em termos de revocação, acurácia e *F1-score*, foram obtidos com o método 1, baseado em detecção de face e Morfologia Matemática. Quanto à precisão, os melhores resultados foram obtidos com o método 2.

Apesar do método 1 apresentar os melhores resultados, para 25% dos vídeos ele teve problemas relacionados ao tipo de roupa e cabelo utilizado pelo motorista. Esse tipo de vestimenta pode-se ser controlado para mitigar os erros do método. Em situações reais, pode ser difícil controlar essas variáveis.

Quanto ao método 2, acredita-se que pode-se obter resultados melhores explorando um conjunto maior de imagens, que seja mais representativo. Além disso, pode-se explorar outras arquiteturas de RNC. Da mesma forma se pode melhorar os resultados do método 2, modificando a base de treinamento dos detectores de mãos e rosto, e também explorando outros modelos para essa detecção. Esses são tópicos de pesquisa importantes, que podem ser explorados em trabalhos futuros.

Informações relacionadas a MMREA são pouco exploradas na literatura. Não foi encontrado nenhum método publicado com foco na detecção automática desse tipo de movimento. Nesse contexto, espera-se que os métodos apresentados nesse Capítulo contribuam para aumentar a eficiência da análise desse tipo de movimento em trabalhos relacionados ao estresse/ansiedade de motoristas.

O método 1 é capaz apenas de identificar os três MMREA's utilizados de forma binária, ou seja, com um único rótulo. Em trabalhos futuros, se pode dedicar esforços para detectar individualmente cada um desses movimentos e também outros gestos do motorista que possam ser associados ao estresse/ansiedade, tais como, esfregar a roupa ou outras partes do corpo.

Com base nos experimentos, pode-se concluir que apesar do método 1 ter obtido os melhores resultados para a maioria dos vídeos analisados, em um ambiente real, as limitações que devem ser impostas ao tipo de vestimenta e penteado dos motoristas podem inviabilizar seu uso. Nesse contexto recomenda-se o método 2.

Em trabalhos futuros há espaço para melhorar os resultados explorando diferentes métodos de segmentação da pele e/ou detecção de mãos e rosto. Podem ser explorados outros métodos de reconhecimento de movimento e outros tipos de imagens, tal como imagens de profundidade. Também sugere-se o desenvolvimento de uma abordagem híbrida, que combine características dos dois métodos apresentados nesse Capítulo, para melhorar os resultados experimentais obtidos.

6 PRIORIZAÇÃO DE ATIVIDADES PRÁTICAS DE DIREÇÃO

Esse Capítulo apresenta o método proposto para priorizar atividades práticas de direção de acordo com o nível de estresse/ansiedade do motorista. O método proposto combina o algoritmo de ordenação Merge Sort (MS) com uma função comparativa baseada em uma Rede Neural Artificial baseada em Regressão Numérica (RNA-RN). O algoritmo de ordenação recebe uma lista de atividades não priorizadas e as ordena comparando pares de atividades com a função comparativa.

A função comparativa recebe como parâmetro uma lista de atributos relacionados à cada atividade prática de direção e com um algoritmo de aprendizagem de máquina determina qual das atividades possui a maior prioridade. Os atributos utilizados são extraídos de três conjuntos de dados: i) frequência cardíaca; ii) AU's 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 25, 26, 28 e 45 (Subseção 2.4); e iii) MMREA's rotulados (Seção 2.6).

A abordagem proposta foi validada com o conjunto de 60 atividades práticas de direção que compõem a base de dados experimental (Seção 4). Para treinar o modelo utilizado pela função comparativa foram conduzidos experimentos com uma Rede Neural Artificial Baseada em Regressão Numérica (RNA-RN).

Os resultados experimentais foram avaliados de acordo com a coesão e similaridade do conjunto priorizado automaticamente com o conjunto de referência priorizado por uma psicóloga especializada no atendimento a motoristas. Para calcular tais métricas foram utilizados o Índice de Jaccard (Hamers et al., 1989) e a Similaridade de Cosseno (Li e Han, 2013). O Índice de Jaccard é uma medida estatística usada para medir a similaridade e diversidade dos conjuntos de amostras, já a similaridade de cosseno é uma métrica que identifica o quão semelhantes são dois vetores medindo o cosseno do ângulo entre dois eles, projetado em um espaço multidimensional.

Com objetivo de apresentar o método de priorização e os resultados experimentais, o restante dessa Seção encontra-se organizado da seguinte forma. A Seção 6.1 descreve o conjunto de características utilizadas. A função comparativa é detalhada na Seção 6.2, que também apresenta o processo de criação dos conjuntos de teste e treinamento. Os experimentos realizados são descritos na Seção 6.3. Por fim, considerações finais sobre o método de priorização são apresentadas na Seção 6.4.

6.1 EXTRAÇÃO DE CARACTERÍSTICAS

Para cada atividade prática de direção da base de dados experimental $da_x \in DA$, foram extraídas 21 características de cada um dos seguintes tipos de dados: frequência cardíaca e comportamentos rotulados nos vídeos (AU's e MMREA's).

Com relação aos MMREA's foi extraída uma característica que denota a porcentagem de tempo de ocorrência de MMREA's, que equivale a divisão do número total de quadros no qual esse evento ocorre, pelo total de quadros do vídeo. Quanto as AU's foram extraídas 18 características, que são referentes a porcentagem de tempo de ocorrência de cada uma das AU's.

Quanto à frequência cardíaca do motorista, foram extraídas duas características que representam, respectivamente, a média dos valores de BPM e a Raiz Quadrada Média de Diferenças Sucessivas (RQMDS) do BPM. A RQMDS provou ser uma métrica importante para compreender as diferenças de variação da frequência cardíaca durante situações de estresse

(Pereira et al., 2017; Wang et al., 2009). Considerando um conjunto de χ valores de BPM $\{bpm_1, bpm_2, \dots, bpm_\chi\}$, a RQMDS é computada com a Equação 6.1.

$$RQMDS = \sqrt[2]{\frac{1}{\chi} \sum_{i=1}^{\chi} (|bpm_{i-1} - bpm_i|)^2} \quad (6.1)$$

A Tabela 6.1 apresenta o identificador e a descrição de cada uma das 21 características extraídas.

Tabela 6.1: Características utilizadas pelo método de priorização

ID	Descrição
$da_x[fc1]$	média do BPM.
$da_x[fc2]$	RQMDS do BPM.
$da_x[MMREA]$	Porcentagem de tempo do vídeo no qual ocorreram os MMREA's.
$da_x[au01]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 01.
$da_x[au02]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 02.
$da_x[au04]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 04.
$da_x[au05]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 05.
$da_x[au06]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 06.
$da_x[au07]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 07.
$da_x[au09]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 09.
$da_x[au10]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 10.
$da_x[au12]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 12.
$da_x[au14]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 14.
$da_x[au15]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 15.
$da_x[au17]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 16.
$da_x[au20]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 20.
$da_x[au23]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 23.
$da_x[au25]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 25.
$da_x[au26]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 26.
$da_x[au28]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 28.
$da_x[au45]$	Porcentagem de tempo do vídeo no qual ocorreu a AU 45.

6.2 FUNÇÃO COMPARATIVA

A função utilizada para comparar pares de atividades de direção em relação ao nível de estresse do motorista é denotada como $comp(da_x, da_y)$. Essa função recebe como parâmetros duas atividades práticas de direção e extrai características dessas atividades (Subseção 6.1). Essas características são passadas como parâmetros para um modelo treinado com base em uma Rede Neural Artificial. Após realizar essas etapas, a função comparativa retorna um valor λ que indica qual das atividades tem maior prioridade.

Considerando $\lambda = comp(da_x, da_y)$, se $\lambda > 0$, a prioridade de da_x é maior que a prioridade de da_y . Caso $\lambda < 0$, da_y tem a prioridade maior que da_x , e se, $\lambda = 0$, ambas as atividades têm a mesma prioridade.

Um conjunto de treinamento intitulado $DA_{training}$ é utilizado para treinar o modelo aplicado pela função comparativa. Cada item do conjunto $DA_{training}$ é composto pelas características extraídas de duas atividades práticas de direção e por um rótulo que indica qual atividade

tem a maior prioridade. Para cada par de atividades (da_x, da_y) utilizado para criar $DA_{training}$, o rótulo é computado pela função $priori_{da_x, da_y}$. O valor de $priori_{x,y}$ é definido pela Equação 6.2.

$$priori_{da_x, da_y} = \begin{cases} 1, & \text{se } da_x \text{ tem a maior prioridade} \\ 0, & \text{se } da_x \text{ e } da_y \text{ tem a mesma prioridade} \\ -1, & \text{se } da_y \text{ tem a maior prioridade} \end{cases} \quad (6.2)$$

Para criar o conjunto $DA_{training}$, é selecionado um subconjunto de τ atividades do conjunto DA . Sendo $|DA| = n$, então $\tau < n$. O conjunto $DA_{training}$ é composto por todas as possíveis combinações entre os τ itens extraídos de DA , portanto, $|DA_{training}| = \tau^2$. A Figura 6.1 ilustra o processo de criação de $DA_{training}$ considerando $\tau = 2$.

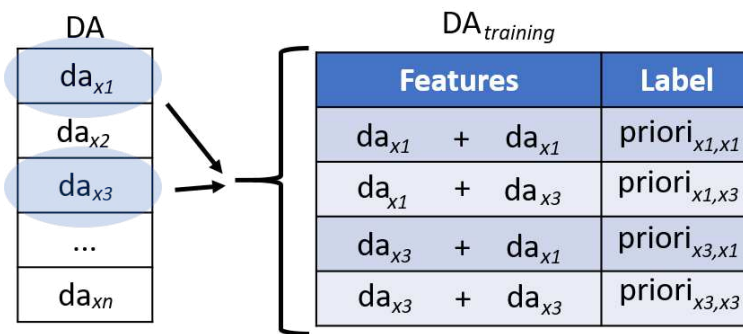


Figura 6.1: Processo de produção do conjunto de treinamento

Cada atividade $da_x \in DA$ que não faz parte do conjunto de treinamento ($da_x \notin DA_{training}$) é utilizada para compor o conjunto utilizado para testar a abordagem proposta. O conjunto de teste é intitulado DA_{test} .

6.3 EXPERIMENTOS

Experimentos foram realizados para avaliar a similaridade e coesão dos resultados obtidos com a abordagem proposta. Para esta avaliação, duas métricas foram selecionadas: Índice de Jaccard e Similaridade de Cosseno (Subseção 6.3.1). A precisão da função comparativa foi avaliada individualmente com relação à acurácia. Detalhes da configuração experimental e resultados estão nas Subseções 6.3.2 e 6.3.3, respectivamente. Finalmente, uma análise dos resultados é apresentada na Subseção 6.3.4.

6.3.1 Métricas de Avaliação

Para avaliar os resultados quantitativamente, foram utilizadas duas métricas: Índice de Jaccard e Similaridade de Cosseno. Ambas as medidas retornam valores de similaridade entre 0 e 1, sendo que, quanto mais próximo de 1 for o resultado, maior será a similaridade. Considerando dois conjuntos B e C , as Equações 6.3 e 6.4 representam como o Índice de Jaccard e a Similaridade de Cosseno são calculados, respectivamente.

O Índice Jaccard, também conhecido como “Intersecção sobre União”, mede a similaridade entre dois conjuntos de dados, levando em consideração a razão entre os tamanhos do conjunto de interseção e união.

$$J(B, C) = \frac{|B \cap C|}{|B \cup C|} \quad (6.3)$$

A Similaridade de Cosseno avalia o nível de coesão entre dois conjuntos a partir da distância angular entre eles no espaço. Essa métrica é particularmente importante para este trabalho, pois, permite avaliar se uma priorização, mesmo com alguns erros, estará próxima da ordem ideal esperada.

$$cos_{sim} = cos(\theta) = \frac{B \cdot C}{\|B\| \cdot \|C\|} \quad (6.4)$$

O banco de dados experimental possui diferentes atividades de direção com a mesma prioridade. Se no conjunto priorizado, duas atividades com a mesma prioridade estiverem em ordem diferente em relação ao conjunto de referência, essa troca não é penalizada como erro no cálculo das métricas de avaliação.

6.3.2 Configuração Experimental

Para realizar os experimentos, nosso banco de dados de 60 atividades de direção foi dividido em dois conjuntos: um número τ de atividades foi utilizado para criar o conjunto de treinamento ($DA_{training}$) e $60 - \tau$ atividades foram utilizadas para criar o conjunto de teste (DA_{teste}). Seis diferentes valores de τ foram explorados: 25, 30, 35, 40, 45 e 50. Além dos diferentes valores de τ , também foram realizados testes considerando sete diferentes intervalos (período de tempo utilizado para extrair as características de cada vídeo) de cada atividade de condução: i) primeiros 5 minutos; ii) primeiros 10 minutos; iii) primeiros 15 minutos; iv) primeiros 20 minutos; v) primeiros 25 minutos; vi) primeiros 30 minutos; e vii) tempo integral da atividade de direção. No caso de alguma das atividades práticas de direção ter duração menor do que o intervalo de tempo analisado, as características são extraídas, considerando o tempo integral desta atividade.

Nos experimentos, para cada combinação entre um intervalo de tempo e um valor de τ , 100 instâncias aleatórias de DA_{teste} e $DA_{training}$ foram criadas e usadas para validar o método de priorização proposto.

6.3.3 Resultados Experimentais

A Tabela 6.2 apresenta a média do Índice de Jaccard e da Similaridade de Cosseno, considerando os resultados obtidos para cada uma das configurações experimentais. A primeira coluna dessa tabela é referente ao tempo utilizado para extração de características. A segunda coluna denota o tamanho do conjunto de treinamento utilizado. As últimas duas colunas são referente à média do Índice de Jaccard e à média da Similaridade de Cosseno.

Os resultados da Tabela 6.2 também são apresentados na forma de gráficos na Figura 6.2: o primeiro gráfico é referente ao Índice de Jaccard e o segundo gráfico denota a Similaridade de Cosseno. O eixo vertical dos gráficos é referente ao valor médio computado para cada métrica. O eixo horizontal denota o intervalo de tempo considerado para extração de características. As linhas dos gráficos são referentes aos resultados obtidos com cada um dos tamanhos definido para os conjuntos de treinamento.

Destaca-se que todos os resultados apresentados nesse Capítulo são referentes à média das métricas calculadas para os 100 conjuntos aleatórios criados para cada tamanho de conjunto de treinamento utilizado.

Tabela 6.2: Resultados de priorização obtidos

Tempo	$\tau = \sqrt{ DA_{\text{training}} }$	Índice de Jaccard	Similaridade de Cosseno
5 min	25	0,07	0,75
	30	0,08	0,75
	35	0,10	0,75
	40	0,12	0,75
	45	0,17	0,76
	50	0,23	0,78
10 min	25	0,08	0,75
	30	0,08	0,75
	35	0,09	0,75
	40	0,13	0,76
	45	0,15	0,76
	50	0,23	0,78
15 min	25	0,08	0,75
	30	0,10	0,75
	35	0,10	0,75
	40	0,16	0,76
	45	0,21	0,77
	50	0,23	0,78
20 min	25	0,09	0,75
	30	0,11	0,76
	35	0,13	0,75
	40	0,16	0,77
	45	0,21	0,77
	50	0,31	0,80
25 min	25	0,08	0,75
	30	0,10	0,76
	35	0,12	0,77
	40	0,15	0,79
	45	0,23	0,81
	50	0,29	0,82
30 min	25	0,08	0,76
	30	0,09	0,77
	35	0,11	0,77
	40	0,13	0,77
	45	0,18	0,79
	50	0,28	0,81
Total	25	0,08	0,75
	30	0,10	0,77
	35	0,10	0,78
	40	0,13	0,77
	45	0,19	0,77
	50	0,24	0,80

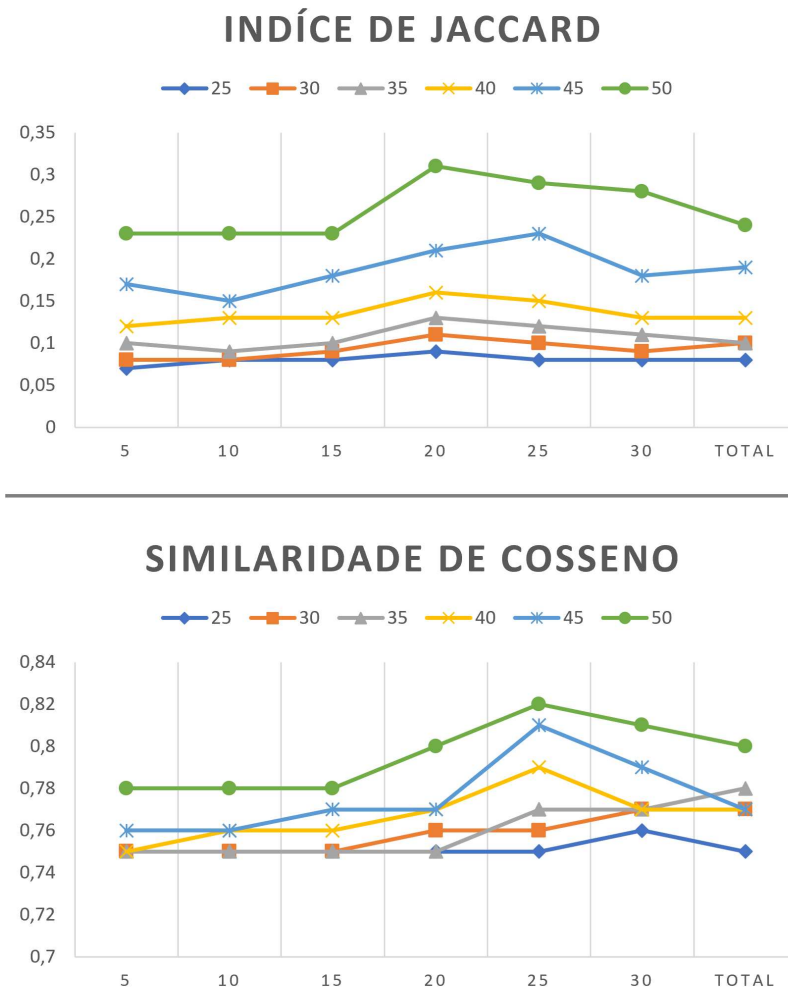


Figura 6.2: Gráficos de coesão e similaridade

Como a abordagem de priorização proposta combina uma função comparativa e um método de classificação, além de avaliar a similaridade e coesão da abordagem proposta, também é importante avaliar a acurácia da função comparativa. Se foram selecionadas τ atividades do conjunto DA para criar o conjunto $DA_{training}$, então $|DA_{training}| = \tau^2$ e $|DA_{test}| = (60 - \tau)^2$ (sendo $|DA| = 60$). Nos experimentos, para cada valor de ν , foram criadas 100 instâncias aleatórias de $DA_{training}$ e DA_{test} , dessa forma, para cada valor de τ é possível avaliar a acurácia da função comparativa com $100 * (60 - \tau)^2$ pares de atividades.

Para cada par de atividades $[A,B] \in DA_{test}$, três rótulos são os possíveis resultados da função comparativa: i) $A > B$, se A tem prioridade maior que B; ii) $B < A$, se B tem prioridade maior que A; e iii) $A = B$, se A e B tem a mesma prioridade. Considerando esses rótulos, a Tabela 6.3 apresenta uma matriz de confusão dos resultados da configuração experimental que apresentou os melhores resultados em termos de coesão: $\sqrt{|DA_{training}|} = \tau = 50$ com características extraídas dos primeiros 25 minutos da atividade.

Com a matriz de confusão apresentada na Tabela 6.3 é possível visualizar que só houve acertos em relação a dois rótulos: “ $A < B$ ” e “ $A > B$ ”. Quanto ao rótulo “ $A = B$ ”, não houve acertos. Erros com relação ao rótulo “ $A = B$ ” não são um problema para a abordagem de priorização porque a troca de posições de atividades com a mesma prioridade não afeta o resultado. Os erros relacionados ao rótulo “ $A = B$ ” podem ser justificados. Na abordagem proposta, apenas se

Tabela 6.3: Matriz de Confusão: Função Comparativa com $\sqrt{|DA_{training}|} = 50$ e Tempo = 25min

	A > B	A < B	A = B
A > B	3349	1085	0
A < B	1062	3372	0
A = B	593	539	0

considera que duas atividades têm a mesma prioridade se a função comparativa retornar o valor zero, porém, com algoritmo baseado em regressão numérica, é muito baixa a probabilidade de se obter exatamente o valor zero.

A acurácia da função comparativa pode ser quantificada, dividindo o número de resultados corretos pelo número total de pares de atividades analisadas. Com base nos valores da Tabela 6.3, se não for considerando o rótulo “A = B” a acurácia obtida é 0,76. Se considerarmos todos os resultados para “A = B” como corretos, a acurácia é 0,79.

Como erros em relação ao rótulo “A = B” não impactam nos resultados de priorização, a acurácia de todas as instâncias experimentais foi calculada considerando como corretos, todos os resultados obtidos para esse rótulo. A acurácia da função comparativa para cada configuração experimental é apresentada na Tabela 6.4. A Figura 6.3 apresenta um gráfico que compara os resultados de acurácia obtidos para cada intervalo de tempo considerado para extração de características.

Tabela 6.4: Acurácia da função comparativa

$\tau = \sqrt{ DA_{training} }$	Intervalo de tempo considerado						Total
	5	10	15	20	25	30	
25	0,68	0,69	0,70	0,73	0,73	0,72	0,70
30	0,68	0,69	0,70	0,74	0,73	0,72	0,71
35	0,70	0,70	0,71	0,75	0,75	0,74	0,72
40	0,71	0,71	0,72	0,75	0,75	0,74	0,72
45	0,73	0,73	0,73	0,77	0,77	0,75	0,74
50	0,74	0,74	0,76	0,79	0,79	0,77	0,76

6.3.4 Análise dos Resultados

Os melhores resultados experimentais foram obtidos considerando as características extraídas dos primeiros 25 minutos de cada atividade prática de direção, com $|DA_{training}| = 50$. Esses resultados indicam que a abordagem proposta pode ser utilizada para agilizar o suporte a motoristas com alto nível de estresse.

Mesmo que para algumas instâncias experimentais a similaridade entre os conjuntos de referência e o conjunto priorizado tenha sido baixa (de acordo com o Índice de Jaccard), a coesão entre os dois conjuntos se manteve sempre acima de 0,74 (em termos de Similaridade de Cosseno). Isso indica que mesmo com erros de priorização, atividades mal ordenadas estavam em uma posição próxima à correta.

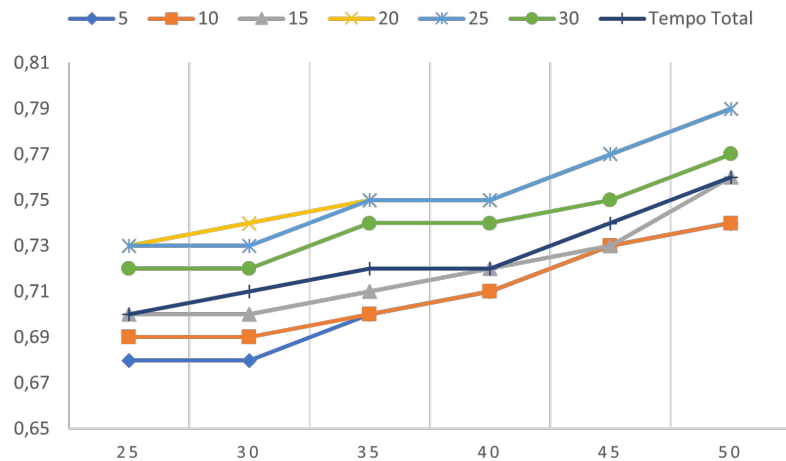


Figura 6.3: Gráfico da acurácia da função comparativa

Uma análise mais detalhada dos resultados da abordagem de priorização apresentada nessa Seção pode ser feita em relação a dois aspectos:

1. **$|\mathbf{DA}_{\text{training}}|$** : Quanto maior foi o tamanho do conjunto de treinamento utilizado nos experimentos, melhores foram os resultados obtidos. Este é um comportamento comum que pode ser observado na maioria dos trabalhos disponíveis na literatura, relacionados a aprendizagem de máquina.
2. **Intervalo de tempo analisado**: Com intervalos de tempo entre 5 e 25 minutos, quanto maior o período de tempo utilizado para extração de características, melhores são os resultados. Após 25 minutos, os resultados da priorização tenderam a piorar, porém os melhores resultados foram obtidos no intervalo entre 20 minutos e o tempo total da atividade prática de direção.

6.4 CONSIDERAÇÕES FINAIS

Essa Seção apresentou um método para priorizar as atividades de direção com base no nível de estresse/ansiedade do motorista. A técnica proposta combinou o algoritmo de ordenação Merge Sort com uma função comparativa baseada em uma Rede Neural Artificial, que recebe uma lista de características referentes a duas atividades práticas de direção, sendo capaz de identificar em qual delas o motorista apresentou maior nível de estresse.

A abordagem proposta é relevante pois, de acordo com o número de motoristas, é muito difícil para um especialista analisar os registros de todas as atividades práticas de direção. Com uma lista ordenada das atividades de direção, o especialista pode inicialmente analisar os casos em que os motoristas apresentaram maior nível de estresse, atendendo a esses motoristas com maior agilidade. Essa abordagem é particularmente relevante no contexto de auto escolas com um fluxo alto e grande de alunos.

Os resultados experimentais indicam que a função comparativa alcançou uma acurácia de até 0,79 com uma coesão igual ou maior a 0,74 em todas as instâncias experimentais. Isso indica que, mesmo quando ocorreram erros de priorização, as atividades mal ordenadas estavam em uma posição próxima da correta.

Algumas limitações podem ser apontadas em relação ao estudo de caso realizado, como, por exemplo, a metodologia utilizada para priorizar manualmente as atividades de direção, o

número limitado de atividades de direção analisadas e o conjunto de características extraídas. Apesar das limitações, os resultados indicam que a abordagem proposta pode ser utilizada para agilizar o apoio a alunos de autoescola com estresse/ansiedade.

O método proposto é particularmente relevante para autoescolas que possuem um grande fluxo de alunos. Se o suporte adequado para esses alunos, pode-se contribuir para a aprendizagem dos alunos e para aumentar a segurança no trânsito.

Devido ao fato de que a taxa de problemas relacionados ao estresse crônico é maior em motoristas que trabalham no transporte público ou privado (Rastgoo et al., 2018), acredita-se que a técnica de priorização apresentada também pode ser aplicada em outros contextos, tais como, empresas de logística e transporte. No Brasil, estudos apontam que entre 27,5% e 54,8% dos motoristas de transporte público apresentam manifestações de estresse (de Matos et al., 2015; Martins et al., 2014; Bigattão, 2005).

Mesmo que a abordagem proposta não tenha ordenado o conjunto de atividades de direção com 100% de similaridade e coesão em relação ao conjunto de referência, os resultados obtidos se mostraram adequados para contribuir na agilização do suporte prestado a motoristas com estresse/ansiedade. Em trabalho futuros, pode-se explorar outras técnicas e outros conjuntos de características para melhorar os resultados de priorização.

Por fim, destaca-se que para agilizar ainda mais o suporte aos motoristas, a técnica de priorização proposta pode ser combinada com a abordagem de sumarização apresentada no próximo Capítulo.

7 SUMARIZAÇÃO BASEADA NO COMPORTAMENTO DO MOTORISTA

O comportamento ao dirigir é um processo complexo que pode ser afetado por vários fatores, que envolvem condições de tráfego, características do veículo e do motorista - tais como idade, experiência, agressividade e estado emocional. Os efeitos dos estados afetivos têm recebido atenção considerável em pesquisas relacionadas à segurança no trânsito. Tais estudos consideram que o comportamento ao dirigir é afetado pelo estado mental/estresse dos motoristas, carga de trabalho cognitiva e distração. No entanto, esses estudos comumente são baseados em questionário e auto relatos que podem estar sujeitos a erros. Estudos apontam que apenas indivíduos altamente motivados respondem adequadamente as questões de pesquisa sobre estresse (Singh e Queyam, 2013).

Cabe destacar que os trabalhos que exploram técnicas computacionais relacionadas ao estresse/ansiedade de motoristas, comumente se concentraram na detecção desses estados afetivos, em vez de quantificar ou modelar seus efeitos nas decisões de condução (Paschalidis et al., 2018). Tal como foi apresentado no segundo MSL (Seção 3.3), a maioria dos trabalhos relacionados ao uso de técnicas computacionais para detecção, análise e suporte ao estresse/ansiedade de motoristas, têm foco na classificação automática desses eventos. Esses trabalhos focam em estabelecer uma correlação entre o nível de estresse/ansiedade com características do motorista, dados do veículo ou do ambiente.

Tais abordagens são relevantes e úteis, porém, se o nível de estresse/ansiedade for alto, apenas detectar os momentos estressantes pode não ser suficiente, pois o motorista pode precisar da ajuda de um especialista para aprender a lidar com esses sentimentos ao dirigir, que podem estar relacionados à vários fatores. Portanto, nem sempre é fácil identificar a causa e obter alívio do estresse/ansiedade relacionado à condução de veículos (Manseer e Riener, 2014).

Nesse contexto, o método para sumarização de atividades práticas de direção apresentado nesse Capítulo pode contribuir, pois a análise do especialista não será baseada apenas no auto relato do motorista, mas também em sumários/relatórios dos vídeos e dados fisiológicos. Além disso, por se tratar de uma abordagem de sumarização, ela também irá contribuir para agilizar e viabilizar a análise de dados coletados durante atividades práticas de direção. Isso é importante, pois, é impraticável para um psicólogo assistir a vídeos de atividades de práticas de direção de um grande conjunto de motoristas em um curto período de tempo (sendo essa a realidade de autoescolas com grande volume de alunos), pois cada atividade pode durar minutos ou horas.

Para facilitar o trabalho de especialistas e a auto compreensão do próprio motorista, a abordagem de sumarização apresentada neste Capítulo produz relatórios e resumos relacionados a uma atividade prática de direção, tendo como objetivo agilizar a análise do estresse/ansiedade de motoristas.

Além de apresentar a abordagem proposta, este Capítulo também descreve uma ferramenta desenvolvida para dar suporte à visualização das informações. A ferramenta permite visualizar uma atividade prática de direção em três perspectivas: i) visualização de vídeos; ii) relatórios; e iii) sumários. Em cada uma das perspectivas é possível visualizar diferentes fatias de informações, tais como, comportamentos relacionados a estresse/ansiedade e frequência cardíaca. A ferramenta apresenta o conteúdo dividido em perspectivas e fatias de informações, para organizar a visualização do conteúdo.

Também com o objetivo de facilitar a visualização e o processo de produção dos sumários, o conjunto discreto de dados rotulados nos vídeos, que indicam a presença ou não de um comportamento, é convertido para um espaço contínuo. Com os valores contínuos, um

vídeo pode ser facilmente resumido, selecionando uma sequência de quadros com valores mais altos que determinado limiar. Uma função é aplicada para realizar essa conversão do espaço discreto para o contínuo. Quanto maior o valor obtido após a conversão, maior é a intensidade ou frequência de um determinado comportamento.

O restante deste Capítulo encontra-se organizado da seguinte forma. A função utilizada para conversão de dados do espaço discreto para o espaço contínuo é descrita na Seção 7.1. A abordagem de sumarização e a ferramenta de suporte são apresentadas na Seção 7.2. A Seção 7.3, detalha os experimentos realizados para validar os resultados apresentados com a ferramenta proposta. Por fim, considerações finais sobre a abordagem de sumarização proposta são apresentadas na Seção 7.4.

7.1 TRANSFORMAÇÃO DOS DADOS

Considerando um vetor D de n posições, que representa os rótulos discretos associados a cada um dos n quadros de um vídeo, a posição x do vetor D é denotada por d_x . Cada posição $d_x \in D$ pode receber um valor binário (0 ou 1), que indica se um evento específico ocorre ou não no quadro x do vídeo. Se $d_x = 1$, o evento ocorreu no quadro x do vídeo, caso contrário não.

O vetor D é utilizado para criar o vetor C . Cada posição $c_x \in C$ representa um valor contínuo associado ao quadro f_x do vídeo V . O vetor C pode conter valores entre 0 e α . Seja α um parâmetro que indica o número de posições adjacentes do vetor D que são utilizadas para calcular o valor de cada posição do vetor C .

Os valores atribuídos para cada posição $c_x \in C$ são computados com a função $convert(x, m, D)$, que é apresentada na Equação 7.1.

$$c_x = convert(x, \alpha, D) = \sum_{y=-\alpha}^{\alpha} d_{x+y} \cdot \frac{\alpha - |y|}{\alpha} \quad (7.1)$$

Os valores atribuídos com a Equação 7.1 variam entre 0 e α . O valor 0 indica que todas as posições adjacentes a d_x em uma distância α , tem o valor 0. O valor de $convert(x, \alpha, D)$ é α se todas as posições adjacentes a x no vetor D tem o valor 1, também considerando uma adjacência de α posições. A Equação 7.1 também pode retornar valores entre 0 e α dependendo da combinação de valores discretos no vetor D .

Após aplicar a Equação 7.1 para todas as posições do vetor C , valores maiores indicam regiões com mais ocorrências de um determinado comportamento ou ocorrências com maior duração. Quanto maior o valor de uma posição $c_x \in C$ em relação à sua vizinhança, mais próxima será a posição c_x do centro de uma região em que um evento aconteceu.

A Figura 7.1 ilustra dois vetores correspondentes: D e C . O vetor C foi calculado com a Equação 7.1 recebendo como parâmetro $\alpha = 600$. Nos experimentos apresentados nessa tese, os vídeos foram gravados a uma taxa de 60 quadros por segundo, portanto, $\alpha = 600$ é equivalente a um intervalo de 10 segundos. Valores pequenos para α não são suficientes para transformar valores discretos em contínuos adequadamente. Valores altos para α também não são adequados, pois tendem a suavizar o gráfico que representa o vetor denotado pela letra C . Nesse contexto, $\alpha = 600$ se mostrou ser um valor adequado. Esse valor de α foi selecionado por uma análise empírica.

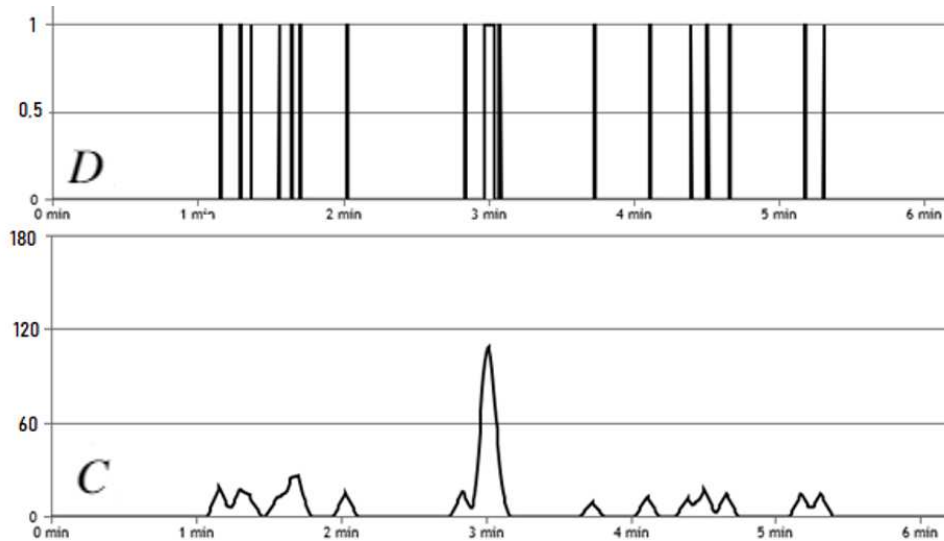


Figura 7.1: Exemplo de conversão de valores do espaço discreto para o espaço contínuo

7.2 FERRAMENTA DE SUPORTE

Uma ferramenta foi desenvolvida para apoiar a abordagem de sumarização proposta. A ferramenta permite analisar as informações sobre o estresse do motorista em três perspectivas distintas: i) vídeo; ii) relatório; e iii) resumo. Essas perspectivas são exploradas para organizar e facilitar a visualização do conteúdo.

Todas as informações da base de dados experimental (BPM e sinais comportamentais de estresse/ansiedade) podem ser visualizadas em cada uma das três perspectivas disponíveis na ferramenta. Os sinais comportamentais de estresse/ansiedade, que são originalmente dados de forma discreta, foram convertidos para o espaço contínuo, com a Equação 7.1.

Os botões na parte superior da ferramenta são usados para selecionar a informação e uma perspectiva de visualização. Exemplos de cada uma dessas perspectivas são apresentados na Figuras 7.2, 7.3 e 7.5. Nessas Figuras os rostos foram borrados para preservar a identidade dos participantes.

Além de informações sobre FC e sinais comportamentais de estresse/ansiedade diferentes conjuntos de dados podem ser sumarizados e apresentados com a mesma ferramenta. Em Ortoncelli et al. (2020), experimentos também foram conduzidos com dados rotulados sobre a interação do motorista com o instrutor e alegria, com base na expressão facial de felicidade.

As Subseções 7.2.1, 7.2.2 e 7.2.3 apresentam detalhes de cada uma das perspectivas de visualização da ferramenta proposta.

7.2.1 Perspectiva de Visualização de Vídeos

Nessa perspectiva a ferramenta apresenta um *player* de vídeo com as funções *play*, *pause*, *stop* e com uma linha do tempo que pode ser utilizada para selecionar um determinado momento nos vídeos. O usuário pode assistir de forma sincronizada o conteúdo dos três vídeos gravados e também visualizar de forma sincronizada a posição exata do mapa em que cada instante do vídeo foi gravado, já que um marcador no mapa se move em sincronia com os vídeos.

Nos mapas, duas abordagens foram utilizadas para apresentar informações com base em uma escala de cores. Para o BPM, as cores são exibidas na linha que representa o caminho percorrido pelo motorista. Para informações sobre o comportamento dos motoristas foram utilizadas marcações na forma de círculos posicionados no local onde ocorreu um determinado

evento - observou-se que o uso de linhas não se mostrou uma abordagem eficaz, pois vários dos movimentos ocorreram em poucos segundos (ou em menos de um segundo), dificultando a sua visualização em um pequeno espaço da linha. As duas abordagens usadas para representar informações no mapa são mostradas nas Figuras 7.2.1 e 7.2.2.

A Figura 7.2 apresenta um exemplo da perspectiva de visualização da ferramenta com informações sobre a FC do motorista. Oito componentes dessa perspectiva foram destacadas com círculos vermelhos e um identificador. Cada um desses componentes é detalhado a seguir:

- **1. Seleção do tipo de informação:** Um botão foi adicionado para a seleção de cada tipo de informação. O botão referente à informação selecionada fica destacado em azul;
- **2. Seleção da perspectiva de visualização:** Foram adicionados três botões, um para cada perspectiva de visualização. Tal como na seleção do tipo de informação, o botão referente à perspectiva de visualização selecionada também fica destacado em azul;
- **3. Legenda de cores:** apresenta uma escala de cores definida para cada tipo de dado. Essas cores são utilizadas para destacar informações no mapa;
- **4. Mapa:** utiliza a API Google Maps para destacar no mapa o local onde as informações ocorreram;
- **5. Vídeo da câmera C:** apresenta o vídeo gravado com a câmera C;
- **6. Vídeo da câmera B:** apresenta o vídeo gravado com a câmera B;
- **7. Player de vídeo:** com os botões *play*, *pause*, *stop* é possível visualizar de forma sincronizada os vídeos e a posição do cursor no mapa;
- **8. Vídeo da câmera A:** apresenta o vídeo gravado com a câmera A.

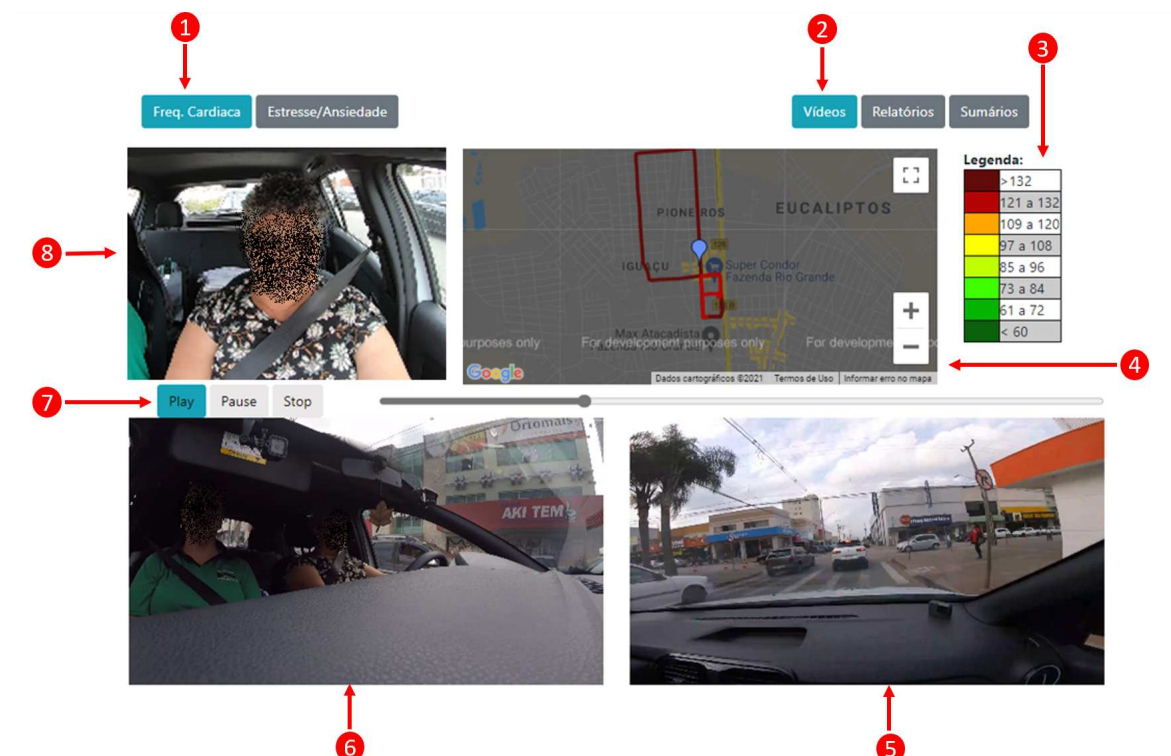


Figura 7.2: Ferramenta de Visualização: Perspectiva de vídeos

7.2.2 Perspectiva de Visualização de Relatórios

Na perspectiva de visualização de relatório a ferramenta apresenta gráficos e um mapa - o mesmo mostrado na perspectiva de visualização de vídeos. Dois tipos de gráficos são usados: linha e pizza. A Figura 7.3 ilustra a ferramenta proposta na perspectiva de visualização de relatórios sobre dados referentes aos sinais de estresse/ansiedade rotulados na base de dados. Seis componentes dessa perspectiva foram destacadas com círculos vermelhos e um identificador. Cada um desses componentes é detalhado a seguir:

- **1, 2 e 3:** esses itens são os mesmos da perspectiva de visualização de vídeos;
- **4. Gráfico de linhas:** os eixos horizontal e vertical exibem respectivamente a linha do tempo da atividade prática de direção e o BPM ou o valor rotulado para cada tipo de informação;
- **5. Gráfico de pizza:** mostra a porcentagem de tempo de ocorrência de cada tipo de informação apresentada na legenda (item 3); e
- **6. Mapa:** o mapa apresentado é o mesmo utilizado na perspectiva de visualização de vídeos. A diferença é que nessa perspectiva o mapa não é integrado com um *player* de vídeo.

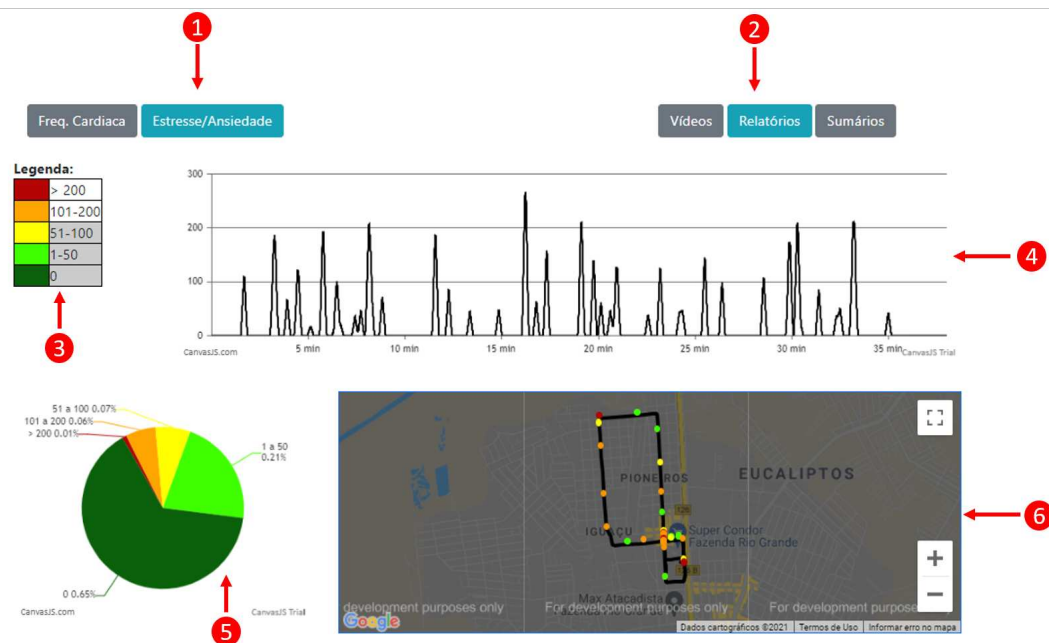


Figura 7.3: Ferramenta de Visualização: Perspectiva de relatórios

7.2.3 Perspectiva de Visualização de Sumários

Essa perspectiva apresenta sumários dos vídeos de acordo com uma perspectiva selecionada. Partes dos três vídeos são segmentadas e selecionadas para compor o sumário de acordo com um critério semântico.

Para cada momento selecionado para compor o sumário, cenas dos três vídeos que representam o mesmo evento são apresentadas lado a lado, sendo reproduzidas de forma síncrona quando o usuário interage com um *player* de vídeo. Todas as cenas são exibidas sequencialmente

em ordem temporal - uma abaixo da outra. Para complementar a exibição, é utilizada uma cor de fundo referente a uma escala de cores de acordo com a legenda apresentada na tela. Acima de cada cena são exibidos o tempo de duração e um valor referente à intensidade da fatia de informação representada na cena.

Quatro técnicas são utilizadas para selecionar as cenas dos vídeos que irão compor os sumários: i) maiores cinco: selecione os cinco momentos com maior valor; ii) maiores dez : selecione os dez momentos com maior valor; iii) valor maior que β_1 : selecione os momentos com algum valor maior que β_1 ; e iv) valor maior que β_2 : selecione os momentos com algum valor maior que β_2). O usuário pode escolher qual técnica de sumarização será utilizada com as opções apresentadas na parte superior da tela. Os valores de β_1 e β_2 foram definidos com base em uma análise empírica dos resultados. Para produzir sumários relacionados a frequência cardíaca, apenas às duas primeiras técnicas são utilizadas. Para os sumários relacionados ao estresse/ansiedade, às quatro técnicas foram utilizadas.

Para compor cada cena do sumário relacionado a frequência cardíaca, foram selecionados os quadros dos vídeos nos instantes de tempo que o motorista apresentou os maiores valores de BPM. Para os sumários associados aos sinais de estresse/ansiedade, para cada região conexa com valores diferentes de zero, apenas o maior valor dessa região é utilizado como critério de seleção da cena, sendo que, quando um quadro é selecionado, toda a região conexa a ele com valores maiores que zero também é selecionada para compor a cena. A Figura 7.4 apresenta um exemplo de seleção de quadros para compor uma cena que fará parte do sumário associado ao estresse/ansiedade.

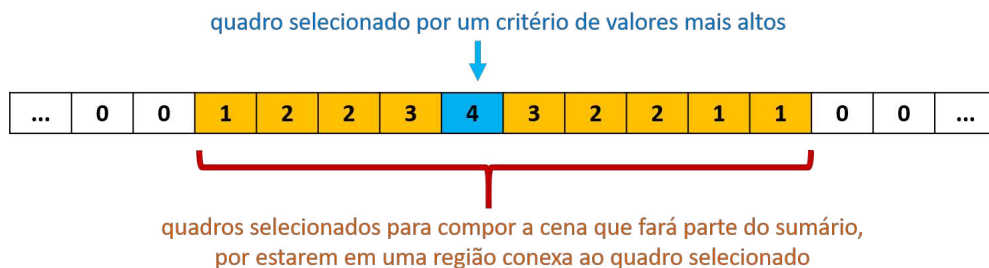


Figura 7.4: Exemplo de seleção de quadros para compor uma cena que fará parte de um sumário associado ao estresse/ansiedade

A Figura 7.5 apresenta a ferramenta proposta na perspectiva de visualização de sumários com dados referentes aos sinais de estresse/ansiedade rotulados na base de dados. Seis componentes dessa perspectiva foram destacadas com círculos vermelhos e um identificador. Cada um desses componentes é detalhado a seguir:

- **1, 2 e 3:** esses itens são os mesmos das outras duas perspectivas de visualização;
- **4. Sumário:** As partes dos vídeos selecionadas para compor os sumários são apresentadas lado a lado em ordem temporal. Uma cor de fundo é adicionada de acordo com a legenda. O maior valor associado a um dos quadros da cena, juntamente com o tempo de início e fim de cada evento são apresentados em cima de cada cena.
- **5. Seleção da técnica de sumarização:** O usuário pode escolher uma das quatro técnicas utilizadas para selecionar as cenas dos vídeos que irão compor os sumários.
- **6. Player de vídeo:** existe um *player* específico para cada cena selecionada para compor o sumário.

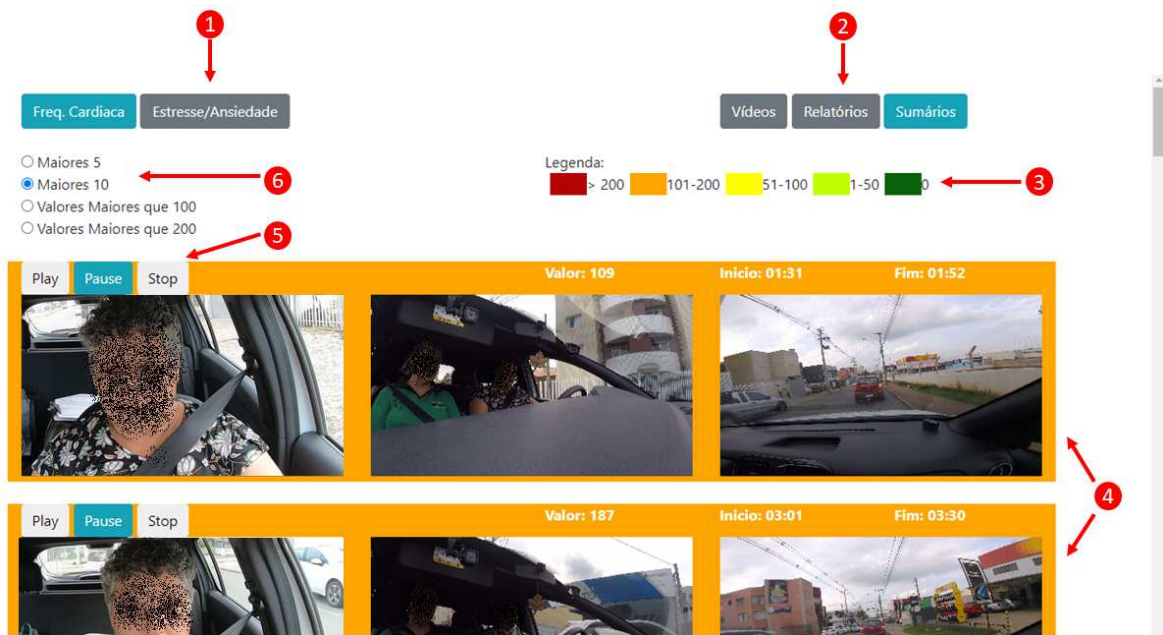


Figura 7.5: Ferramenta de Visualização: Perspectiva de sumários

7.3 EXPERIMENTOS

Experimentos foram realizados para avaliar de forma qualitativa os resultados da abordagem de sumarização e produção de relatórios, de acordo com a opinião de profissionais da autoescola na qual foi coletada a base de dados.

Um estudo foi publicado em Ortoncelli et al. (2020), no qual a ferramenta foi utilizada para visualizar informação sobre quatro tipos de informação: i) BPM; ii) comportamentos associados a nervosismos/estresse; iii) interação com o instrutor: eventos relacionados a comunicação verbal e não verbal do aluno com o instrutor foram rotulados; e iv) felicidade: a expressão facial de alegria foi rotulada. Nesse trabalho, um conjunto de 21 atividades práticas de direção foi utilizado para avaliar a ferramenta de suporte na perspectiva de dois profissionais: o instrutor de direção que acompanhou a atividade de dirigir e uma psicóloga especializada em atender motoristas.

A psicóloga inicialmente analisou os relatórios e os sumários e depois assistiu aos vídeos na íntegra. Os instrutores tiveram a liberdade de utilizar a ferramenta como desejarem, pois já haviam acompanhado o motorista na atividade. De acordo com a opinião dos instrutores de direção, a ferramenta resumia adequadamente as informações sobre a atividade de direção. A psicóloga também considerou pertinentes as informações apresentadas, permitindo-lhe compreender as características do motorista.

A ferramenta apresentada nessa Seção também foi utilizada durante o processo de priorização manual de atividades práticas de direção de acordo com o nível de estresse/ansiedade dos motoristas (Seção 4.4). Todo o processo de priorização manual foi feito por psicóloga especializada em atender motoristas, que considerou o conteúdo apresentado adequado para entender o nível de estresse/ansiedade do motorista.

7.4 CONSIDERAÇÕES FINAIS

A abordagem proposta visa auxiliar no suporte prestado a motoristas com alto nível de estresse que precisam da ajuda de um especialista, tal como, um psicólogo. Com a ferramenta

que apoia o método de sumarização, o especialista poderá analisar e compreender rapidamente informações sobre uma atividade de prática de direção em três diferentes perspectivas. Essas perspectivas contribuem para organizar os conteúdos analisados.

A ferramenta apresentada foi avaliada qualitativamente por instrutores de condução e uma psicóloga especializada no atendimento de motoristas. Pode-se concluir que os sumários resumem de forma adequada a atividade de direção e permitem compreender rapidamente o comportamento do condutor.

Tal como a técnica de priorização apresentada no Capítulo 6, o método de sumarização proposto também é particularmente relevante no contexto de auto escolas com uma grande quantidade de alunos. Esses dois métodos (sumarização e priorização) podem ser combinados para tornar o processo de análise de estresse de motoristas mais eficiente.

Além de aplicações para alunos de autoescolas, acredita-se que o método proposto pode ser aplicado em outros contextos, tal como empresas de transporte e logísticas, que desejam monitorar estados afetivos de seus motoristas. Como a taxa de problemas de saúde causados por estresse crônico são maiores em motoristas que trabalham no transporte público ou privado (Rastgoo et al., 2018), é importante explorar esse tipo de aplicação em trabalhos futuros. No Brasil, estudos apontam que taxas entre 27,5% e 54,8% dos motoristas de transporte público apresentam manifestações de estresse (de Matos et al., 2015; Martins et al., 2014; Bigattão, 2005).

Pesquisas futuras também podem focar no desenvolvimento de técnicas para automatizar o processo de sumarização, de modo que a análise dos dados possa ocorrer em tempo real. Além disso, pode-se avaliar a ferramenta com diferentes tipos de informações, tais como dados de EEG, EDA. Também pode-se aplicar a técnica de sumarização com foco em estados afetivos diferentes, tais como a felicidade/prazer de dirigir do motorista, que, por exemplo, pode ser utilizada como base para a implementação de um preditor da probabilidade de venda um de veículo após a realização de um *test drive*.

8 CONCLUSÃO

Essa tese apresentou uma abordagem para sumarização de atividades práticas de direção. Os resultados experimentais indicam que a abordagem pode ser aplicada para agilizar o suporte prestado a alunos com alto nível de estresse.

A abordagem proposta é particularmente relevante no contexto de autoescolas, devido a grande e rápida circulação de alunos. Nesse ambiente, é difícil que um especialista consiga analisar todos os dados coletados em tempo hábil para prestar o devido suporte aos alunos.

Detalhes sobre cada uma das contribuições e resultados dessa tese são apresentados na Seção 8.1. A Seção 8.2, apresenta uma perspectiva dos trabalhos futuros que podem ser realizados. Por fim, as últimas considerações sobre o trabalho são apresentadas na Seção 8.3.

8.1 SUMÁRIO DAS CONTRIBUIÇÕES

Esta Seção apresenta de forma resumida cada uma das contribuições dessa tese:

- **Base de dados:** Foi produzida e disponibilizada uma base de dados com todas as informações rotuladas nos vídeos e também o BPM dos motoristas. Os vídeos coletados não foram disponibilizados, pois, os participantes dos experimentos não autorizaram o uso da sua imagem.

Além de dados sobre o BPM e os eventos rotulados, a base de dados também disponibiliza a prioridade de cada uma das atividades práticas de direção, em termo dos rótulos atribuídos pelo especialista. Espera-se que os dados disponibilizados possam servir de base para o desenvolvimento de trabalhos futuros.

- **Deteção de MMREA:** Foram conduzidos experimentos com dois métodos para deteção automática dos MMREA's rotulados na base de dados. Os melhores resultados de precisão foram obtidos com o método 2, que é baseado em *Deep Learning*. Quanto as demais métricas (revocação, acurácia e *F1-score*), os melhores resultados foram obtidos com o método 1, baseado em deteção do rosto e morfologia matemática.

O método 1 se mostrou suscetível a problemas com determinados tipos de roupas/cabelos. Tais condições podem ser controladas para mitigar os erros. Como é difícil controlar essas variáveis em um ambiente real, isso pode inviabilizar o uso do método 1. Nesse contexto, recomenda-se o método 2, que apresentou a melhor precisão e taxas de acurácia e *F1-score* próximas às do método 1.

Destaca-se que não foram encontradas na literatura outras técnicas aplicadas para detectar automaticamente o mesmo conjunto de MMREA's. Nesse contexto, se reforça a relevância dos métodos apresentados.

- **Priorização de atividades práticas de direção:** Foi proposto um método para priorização de atividades práticas de direção que combina o algoritmo de ordenação Merge Sort com uma função comparativa baseada em uma Rede Neural Artificial.

Com os experimentos realizados se obteve um nível de similaridade de até 0,31 (computado com o Índice de Jaccard) e uma coesão de até 0,82 (computada com a métrica de Similaridade de Cosseno). Esses resultados podem ser considerados

adequados, pois mesmo quando ocorreram erros de priorização, as atividades mal ordenadas estavam em uma posição próxima da correta.

- **Sumarização de atividades práticas de direção:** Foi apresentada uma abordagem para sumarizar e produzir relatórios de atividades práticas de direção com base no BPM e no comportamento do motorista. Uma ferramenta foi apresentada para dar suporte a abordagem de sumarização, permitindo visualizar vídeos, relatórios para cada uma das fatias de informações analisadas.

Os resultados obtidos com essa técnica foram avaliados qualitativamente por profissionais da autoescola na qual os dados foram coletados. Os sumários se mostraram representativos e a ferramenta de visualização se mostrou adequada de acordo com a opinião de profissionais da auto escola na qual a base de dados foi coletada.

Os métodos para detecção automática de MMREA podem ser utilizados para automatizar a abordagem proposta nessa tese.

As técnicas propostas para priorização e sumarização podem combinadas para tornar mais eficiente o suporte a motoristas com alto nível de estresse. Também pode-se destacar que a abordagem proposta pode ser aprimorada em trabalhos futuros, tal como é descrito na próxima Seção.

8.2 TRABALHOS FUTUROS

Existem várias perspectivas de trabalhos futuros em relação ao tema dessa tese, entre elas pode-se destacar:

- os métodos propostos para detecção de movimentos e priorização de atividades não alcançaram os melhores resultados possíveis para acurácia, revocação, *F1-score*, similaridade e coesão, portanto, existe espaço para melhoria nesses métodos. Tais melhorias podem envolver, o treinamento de outras arquiteturas e/ou algoritmos de aprendizagem de máquina. E também o uso de base de dados diferentes, que podem envolver informações mais significativas para o treinamento dos modelos.
- podem ser conduzidos estudos para automatização da abordagem proposta com foco em uma análise em tempo real dos dados.
- além do conjunto de informações exploradas, outras fontes de dados também podem ser exploradas, tais como dados de EEG e EDA. Também pode-se explorar comportamentos relacionados a outros estados afetivos do motorista, tal como, felicidade/prazer de dirigir do motorista.
- nessa tese, os resultados experimentais foram avaliados de acordo com a opinião de uma única psicóloga e de um grupo pequeno de instrutores de autoescola. Para aumentar a confiabilidade dos resultados, pode-se trabalhar um grupo maior de profissionais, tanto na atividade de priorização manual de atividades práticas de direção, quanto para avaliação dos sumários e relatórios produzidos.
- outras aplicações para a abordagem proposta também podem ser exploradas, tal como, focar em empresas de transporte e logística que contam um grande número de motoristas. Trabalhos nessas áreas são importantes, pois, a taxa de problemas de saúde causados por estresse crônico são maiores em motoristas que trabalham no transporte público ou

privado (Rastgoo et al., 2018). Estudos conduzidos com motoristas de transporte público no Brasil, indicam que taxas entre 27,5% e 54,8% desses profissionais apresentam manifestações de estresse (de Matos et al., 2015; Martins et al., 2014; Bigattão, 2005). Nesse contexto, técnicas de detecção, sumarização e priorização podem ajudar a identificar e agilizar o suporte prestado a motoristas com estresse/ansiedade.

- também é importante analisar do impacto da abordagem proposta, no suporte prestado a motoristas com alto nível de estresse. Para essa análise, pode-se mensurar o nível de estresse do motorista antes e depois do suporte prestado com base nos sumários produzidos.

8.3 CONSIDERAÇÕES FINAIS

Com base nos resultados experimentais, se conclui que a abordagem proposta pode ser aplicada para agilizar o suporte prestado a alunos de autoescolas com estresse e ansiedade, contribuindo para melhorar a formação dos motoristas e a segurança no trânsito.

Mesmo com os resultados apresentados, o trabalho proposto nessa tese é apenas a ponta do *iceberg* em relação a trabalhos relacionados que podem ser desenvolvidos. Devido à relevância do tema, espera-se que o estudo instigue e inspire outras pesquisas relacionadas a sumarização de atividade práticas de direção com foco no estresse dos motoristas. Esse é um campo de pesquisa multidisciplinar, desafiador e, ao mesmo tempo, gratificante.

REFERÊNCIAS

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.
- Abayomi, A., Olugbara, O. O., Heukelman, D. e Adetiba, E. (2017). Physiological signals based automobile drivers' stress levels detection using shape and texture feature descriptors: An experimental study. Em *International Conference on Real Time Intelligent Systems*, páginas 436–447. Springer.
- Affanni, A., Bernardini, R., Piras, A., Rinaldo, R. e Zontone, P. (2018). Driver's stress detection using skin potential response signals. *Measurement*, 122:264–274.
- Affanni, A., Piras, A., Rinaldo, R. e Zontone, P. (2019). Dual channel electrodermal activity sensor for motion artifact removal in car drivers' stress detection. Em *IEEE Sensors Applications Symposium*, páginas 1–6. IEEE.
- Aguiar, E., Lakkaraju, H., Bhanpuri, N., Miller, D., Yuhas, B. e Addison, K. L. (2015). Who, when, and why: A machine learning approach to prioritizing students at risk of not graduating high school on time. Em *International Conference on Learning Analytics And Knowledge*, páginas 93–102.
- Albawi, S., Mohammed, T. A. e Al-Zawi, S. (2017). Understanding of a convolutional neural network. Em *International Conference on Engineering and Technology*, páginas 1–6. IEEE.
- Albiero, V., Bellon, O. R. e Silva, L. (2018). Multi-label action unit detection on multiple head poses with dynamic region learning. Em *IEEE International Conference on Image Processing*, páginas 2037–2041.
- Anzengruber, B. e Riener, A. (2012). "Facelight" potentials and drawbacks of thermal imaging to infer driver stress. Em *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, páginas 209–216.
- Arif, M., Basri, A., Melibari, G. et al. (2020). Classification of anxiety disorders using machine learning methods: A literature review. *Insights of Biomedical Research*, 4(1):95–110.
- Arifin, S. e Cheung, P. Y. (2007). A computation method for video segmentation utilizing the pleasure-arousal-dominance emotional information. Em *ACM International Conference on Multimedia*, páginas 68–77.
- Ashokkumar, S., Suresh, A., Priya, S. e Dhanasekaran, R. (2016). An optimized approach for multi-user facial expression-based video summarization. *Biomedical Research*.
- Azrin, N. e Nunn, R. (1973). Habit reversal: A method of eliminating nervous habits and tics. *Behaviour Research and Therapy*, 11(4):619–628.
- Bajaj, A. e Sangwan, O. P. (2019). A systematic literature review of test case prioritization using genetic algorithms. *IEEE Access*, 7:126355–126375.

- Balakrishnama, S. e Ganapathiraju, A. (1998). Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, 18(1998):1–8.
- Baltodano, S., Garcia-Mancilla, J. e Ju, W. (2018). Eliciting driver stress using naturalistic driving scenarios on real roads. Em *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, páginas 298–309.
- Baltrusaitis, T., Zadeh, A., Lim, Y. C. e Morency, L.-P. (2018). Openface 2.0: Facial behavior analysis toolkit. Em *IEEE International Conference on Automatic Face & Gesture Recognition*, páginas 59–66.
- Barnard, M. P. e Chapman, P. (2016). Are anxiety and fear separable emotions in driving? a laboratory study of behavioural and physiological responses to different driving environments. *Accident Analysis & Prevention*, 86:99–107.
- Barnard, M. P. e Chapman, P. (2018). The effects of instruction and environmental demand on state anxiety, driving performance and autonomic activity: Are ego-threatening manipulations effective? *Transportation Research part F: Traffic Psychology and Behaviour*, 55:123–135.
- Batista, J. C., Albiero, V., Bellon, O. R. e Silva, L. (2017). Aumpnet: simultaneous action units detection and intensity estimation on multipose facial images using a single convolutional neural network. Em *IEEE International Conference on Automatic Face & Gesture Recognition*, páginas 866–871.
- Beanland, V., Fitzharris, M., Young, K. L. e Lenné, M. G. (2013). Driver inattention and driver distraction in serious casualty crashes: Data from the australian national crash in-depth study. *Accident Analysis & Prevention*, 54:99–107.
- Begum, S., Ahmed, M. U., Funk, P. e Filla, R. (2012). Mental state monitoring system for the professional drivers based on heart rate variability analysis and case-based reasoning. Em *Federated Conference on Computer Science and Information Systems*, páginas 35–42. IEEE.
- Bianco, S., Napoletano, P. e Schettini, R. (2019). Multimodal car driver stress recognition. Em *International Conference on Pervasive Computing Technologies for Healthcare*, páginas 302–307.
- Bigattão, M. A. (2005). O estresse em motoristas no transporte coletivo de Ônibus em campo grande. 126 fl. Mestrado em psicologia, Universidade Católica Dom Bosco, Campo Grande.
- Blum, A., Kalai, A. e Langford, J. (1999). Beating the hold-out: Bounds for k-fold and progressive cross-validation. Em *Annual Conference on Computational Learning Theory*, páginas 203–208.
- Bořil, H., Boyraz, P. e Hansen, J. H. (2012). Towards multimodal driver's stress detection. Em *Digital Signal Processing for In-vehicle Systems and Safety*, páginas 3–19. Springer.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.
- Brown, T. G., Ouimet, M. C., Eldeb, M., Tremblay, J., Vingilis, E., Nadeau, L., Pruessner, J. e Bechara, A. (2016). Personality, executive control, and neurobiological characteristics associated with different forms of risky driving. *PLoS one*, 11(2):1–18.

- Cacioppo, J. T., Berntson, G. G., Klein, D. J. e Poehlmann, K. M. (1997). Psychophysiology of emotion across the life span. *Annual Review of Gerontology and Geriatrics*, 17:27–74.
- Calvagno, G., Ermani, M., Rinaldo, R. e Sartoretto, F. (2000). A multiresolution approach to spike detection in eeg. Em *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, páginas 3582–3585. IEEE.
- Carneiro, D., Novais, P., Augusto, J. C. e Payne, N. (2017). New methods for stress assessment and monitoring at the workplace. *IEEE Transactions on Affective Computing*, 10(2):237–254.
- Castellano, G., Villalba, S. D. e Camurri, A. (2007). Recognising human emotions from body movement and gesture dynamics. Em *International Conference on Affective Computing and Intelligent Interaction*, páginas 71–82. Springer.
- Chakraborty, P. R., Tjondronegoro, D., Zhang, L. e Chandran, V. (2016). Automatic identification of sports video highlights using viewer interest features. Em *ACM on International Conference on Multimedia Retrieval*, páginas 55–62.
- Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M. e Agrawal, P. (2019). Understanding emotions in text using deep learning and big data. *Computers in Human Behavior*, 93:309–317.
- Chen, L.-l., Zhao, Y., Ye, P.-f., Zhang, J. e Zou, J.-z. (2017). Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Systems with Applications*, 85:279–291.
- Choi, M., Koo, G., Seo, M. e Kim, S. W. (2017). Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness. *IEEE Transactions on Instrumentation and Measurement*, 67(3):634–645.
- Conjeti, S., Singh, R. R. e Banerjee, R. (2012). Bio-inspired wearable computing architecture and physiological signal processing for on-road stress monitoring. Em *IEEE-EMBS International Conference on Biomedical and Health Informatics*, páginas 479–482. IEEE.
- Cooper, M. e Foote, J. (2002). Summarizing video using non-negative similarity matrix factorization. Em *IEEE Workshop on Multimedia Signal Processing.*, páginas 25–28. IEEE.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L. e Stein, C. (2009). *Introduction to algorithms*. MIT press.
- Corneanu, C. A., Simón, M. O., Cohn, J. F. e Guerrero, S. E. (2016). Survey on RGB, 3D, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(8):1548–1568.
- Craye, C. e Karray, F. (2015). Driver distraction detection and recognition using rgb-d sensor. *arXiv preprint arXiv:1502.00250*.
- Cruz, A. C. e Rinaldi, A. (2017). Video summarization for expression analysis of motor vehicle operators. Em *International Conference on Universal Access in Human-Computer Interaction*, páginas 313–323. Springer.

- Dalal, N. e Triggs, B. (2005). Histograms of oriented gradients for human detection. Em *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, páginas 886–893. IEEE.
- Dammak, M., Wali, A. e Alimi, A. M. (2015). Viewer’s affective feedback for video summarization. *Journal of Information Processing Systems*, 11(1).
- de Matos, M. G., de Moraes, L. F. R. e Pereira, L. Z. (2015). Análise do estresse ocupacional em motoristas de coletivo urbano na cidade de belo horizonte. *Revista Gestão & Tecnologia*, 15(1):256–275.
- Dhagdi, M. S. T. e Deshmukh, P. (2012). Keyframe based video summarization using automatic threshold & edge matching rate. *International Journal of Scientific and Research Publications*, 2(7):1–12.
- Dhall, A. e Goecke, R. (2012). Group expression intensity estimation in videos via gaussian processes. Em *International Conference on Pattern Recognition*, páginas 3525–3528. IEEE.
- Dong, E., Du, H. e Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20(5):533–534.
- Dyba, T., Dingsoyr, T. e Hanssen, G. K. (2007). Applying systematic reviews to diverse study types: An experience report. Em *International Symposium on Empirical Software Engineering and Measurement*, páginas 225–234. IEEE.
- Ehlers, A., Hofmann, S. G., Herda, C. A. e Roth, W. T. (1994). Clinical characteristics of driving phobia. *Journal of Anxiety Disorders*, 8(4):323–339.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: a reply to russell’s mistaken critique. *American Psychological Association*.
- Ekman, P., Friesen, W. V. e Hager, J. C. (2002). Facial action coding system: The manual on cd rom. *A Human Face, Salt Lake City*.
- El Haouij, N., Poggi, J.-M., Ghozi, R., Sevestre-Ghalila, S. e Jaïdane, M. (2019). Random forest-based approach for physiological functional variable selection for driver’s stress level classification. *Statistical Methods & Applications*, 28(1):157–185.
- Elgendi, M. (2019). Characteristics of a highly cited article: A machine learning perspective. *IEEE Access*, 7:87977–87986.
- Elgendi, M. e Menon, C. (2020). Machine learning ranks ecg as an optimal wearable biosignal for assessing driving stress. *IEEE Access*, 8:34362–34374.
- Faghel-Soubeyrand, S., Lecomte, T., Bravo, M. A., Lepage, M., Potvin, S., Abdel-Baki, A., Villeneuve, M. e Gosselin, F. (2020). Abnormal visual representations associated with confusion of perceived facial expression in schizophrenia with social anxiety disorder. *NPJ schizophrenia*, 6(1):1–9.
- Fairclough, S. H., Tattersall, A. J. e Houston, K. (2006). Anxiety and performance in the british driving test. *Transportation Research part F: Traffic Psychology and Behaviour*, 9(1):43–52.
- Farnsworth, B. (2016). Facial action coding system (facs)—a visual guidebook. <https://imotions.com/blog/facial-action-coding-system/>.

- Fei, M., Jiang, W. e Mao, W. (2017). Memorable and rich video summarization. *Journal of Visual Communication and Image Representation*, 42:207–217.
- Felizardo, K. R., Nakagawa, E. Y., F., S. C. P. e Ferrari, F. C. (2017). *Revisão sistemática da literatura em engenharia de software: teoria e prática*. Elsevier Brasil.
- Ferreira, L., da Silva Cruz, L. e Assuncao, P. (2016). Towards key-frame extraction methods for 3D video: a review. *EURASIP Journal on Image and Video Processing*, 2016(1):28.
- Gao, H., Yüce, A. e Thiran, J.-P. (2014). Detecting emotional stress from facial expressions for driving safety. Em *IEEE International Conference on Image Processing*, páginas 5961–5965. IEEE.
- Gao, Y., Liu, C., Shen, Z., Han, Y., Gao, Z., Wang, P., Zhang, C., Ren, P. e Xie, X. (2020). Automatic key moment extraction and highlights generation based on comprehensive soccer video understanding. *International Conference on Multimedia & Expo Workshops*, 2(23):1041–1043.
- Garcia-Mancilla, J. e Gonzalez, V. M. (2015). Stress quantification using a wearable device for daily feedback to improve stress management. Em *Smart Health*, páginas 204–209. Springer.
- Gavrilescu, M. e Vizireanu, N. (2019). Predicting depression, anxiety, and stress levels from videos using the facial action coding system. *Sensors*, 19(17):3693.
- Geddes, J. R. e Andreasen, N. C. (2020). *New Oxford textbook of psychiatry*. Oxford University Press, USA.
- Giannakakis, G., Koujan, M. R., Roussos, A. e Marias, K. (2020). Automatic stress detection evaluating models of facial action units. Em *IEEE International Conference on Automatic Face & Gesture Recognition*.
- Giannakakis, G., Padiaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., Marias, K. e Tsiknakis, M. (2017). Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31:89–101.
- Goessl, V. C., Curtiss, J. E. e Hofmann, S. G. (2017). The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis. *Psychological Medicine*, 47(15):2578.
- Gomez, P. e Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of psychophysiology*, 53(2):91–103.
- Gulli, A. e Pal, S. (2017). *Deep learning with Keras*. Packt Publishing Ltd.
- Gutmann, M., Grausberg, P. e Kyamakya, K. (2015). Detecting human driver's physiological stress and emotions using sophisticated one-person cockpit vehicle simulator. Em *Information Technologies in Innovation Business Conference*, páginas 15–18. IEEE.
- Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J. e Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, 405(6789):947–951.
- Hamers, L. et al. (1989). Similarity measures in scientometric research: The jaccard index versus salton's cosine formula. *Information Processing and Management*, 25(3):315–18.

- Hanjalic, A. (2005). Adaptive extraction of highlights from a sport video based on excitement modeling. *IEEE Transactions on Multimedia*, 7(6):1114–1122.
- Harrigan, J. A. e O’Connell, D. M. (1996). How do you look when feeling anxious? facial displays of anxiety. *Personality and individual differences*, 21(2):205–212.
- Hasan, T., Bořil, H., Sangwan, A. e Hansen, J. H. (2013). Multi-modal highlight generation for sports videos using an information-theoretic excitability measure. *EURASIP Journal on Advances in Signal Processing*, 2013(1):173.
- Hassan, M. A., Malik, A. S., Fofi, D., Saad, N., Karasfi, B., Ali, Y. S. e Meriaudeau, F. (2017). Heart rate estimation using facial video: A review. *Biomedical Signal Processing and Control*, 38:346–360.
- Healey, J. e Picard, R. (2000). Smartcar: detecting driver stress. Em *International Conference on Pattern Recognition*, volume 4, páginas 218–221. IEEE.
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J. e Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, 13(4):18–28.
- Helgoe, L. A., Wilhelm, L. R. e Kommor, M. J. (2005). *The anxiety answer book*. Sourcebooks, Inc.
- Hill, J. D. e Boyle, L. N. (2007). Driver stress as influenced by driving maneuvers and roadway conditions. *Transportation Research part F: Traffic Psychology and Behaviour*, 10(3):177–186.
- Howse, J. (2013). *OpenCV computer vision with python*. Packt Publishing Ltd.
- Hu, T., Xie, X. e Li, J. (2013). Negative or positive? the effect of emotion and mood on risky driving. *Transportation Research part F: Traffic Psychology and Behaviour*, 16:29–40.
- Huang, G.-B., Zhou, H., Ding, X. e Zhang, R. (2011). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(2):513–529.
- Hussain, T., Muhammad, K., Ding, W., Lloret, J., Baik, S. W. e de Albuquerque, V. H. C. (2021). A comprehensive survey of multi-view video summarization. *Pattern Recognition*, 109:107567.
- Ide, I., Zhang, Y., Tanishige, R., Doman, K., Kawanishi, Y., Deguchi, D. e Murase, H. (2017). Summarization of news videos considering the consistency of auditory and visual contents. Em *International Symposium on Multimedia*.
- Ihme, K., Dömeland, C., Freese, M. e Jipp, M. (2018). Frustration in the face of the driver: A simulator study on facial muscle activity during frustrated driving. *Interaction Studies*, 19(3):487–498.
- Inamdar, S., Narangale, S. e Shinde, G. (2011). Preprocessor agent approach to knowledge discovery using zero-r algorithm. *International Journal*, 2:2011.
- Jáuregui, D. A. G., Castanier, C., Chang, B., Val, M., Cottin, F., Le Scanff, C. e Martin, J. C. (2017). Toward automatic detection of acute stress: Relevant nonverbal behaviors and impact of personality traits. Em *International Conference on Affective Computing and Intelligent Interaction*, páginas 354–361. IEEE.

- Jensen, F. V. et al. (1996). *An introduction to Bayesian networks*, volume 210. UCL press London.
- Ji, Y., Wang, S., Lu, Y., Wei, J. e Zhao, Y. (2018). Eye and mouth state detection algorithm based on contour feature extraction. *Journal of Electronic Imaging*, 27(5):051205.
- Joho, H., Staiano, J., Sebe, N. e Jose, J. M. (2011). Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents. *Multimedia Tools and Applications*, 51(2):505–523.
- Joosen, P., Exadaktylos, V., Taelman, J. e Berckmans, D. (2017). The effect of individual stress zones on car-racing performance. Em *International Conference on Wearable and Implantable Body Sensor Networks*, páginas 79–82. IEEE.
- Kaber, D., Zhang, Y., Jin, S., Mosaly, P. e Garner, M. (2012). Effects of hazard exposure and roadway complexity on young and older driver situation awareness and performance. *Transportation Research part F: Traffic Psychology and Behaviour*, 15(5):600–611.
- Katsis, C., Goletsis, Y., Rigas, G. e Fotiadis, D. (2011). A wearable system for the affective monitoring of car racing drivers during simulated conditions. *Transportation Research part C: Emerging Technologies*, 19(3):541–551.
- Katsis, C. D., Katertsidis, N., Ganiatsas, G. e Fotiadis, D. I. (2008). Toward emotion recognition in car-racing drivers: A biosignal processing approach. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(3):502–512.
- Katti, H., Yadati, K., Kankanhalli, M. e Tat-Seng, C. (2011). Affective video summarization and story board generation using pupillary dilation and eye gaze. Em *IEEE International Symposium on Multimedia*, páginas 319–326. IEEE.
- Keshan, N., Parimi, P. e Bichindaritz, I. (2015). Machine learning for stress detection from ecg signals in automobile drivers. Em *IEEE International Conference on Big Data*, páginas 2661–2669. IEEE.
- Kim, S., Rhee, W., Choi, D., Jang, Y. J. e Yoon, Y. (2018). Characterizing driver stress using physiological and operational data from real-world electric vehicle driving experiment. *International Journal of Automotive Technology*, 19(5):895–906.
- Kitchenham, B. A., Dyba, T. e Jorgensen, M. (2004). Evidence-based software engineering. Em *International Conference on Software Engineering*, páginas 273–281. IEEE.
- Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Em *International joint Conference on Artificial Intelligence*, volume 14, páginas 1137–1145. Montreal, Canada.
- Krishnan, M. J. (2017). Accident prone area detection using driver stress data base. Em *IEEE International Conference on Circuits and Systems*, páginas 220–223. IEEE.
- Kubota, Y., Sato, W., Toichi, M., Murai, T., Okada, T., Hayashi, A. e Sengoku, A. (2001). Frontal midline theta rhythm is correlated with cardiac autonomic activities during the performance of an attention demanding meditation procedure. *Cognitive Brain Research*, 11(2):281–287.

- Kulke, L., Feyerabend, D. e Schacht, A. (2020). A comparison of the affectiva imotions facial expression analysis software with emg for identifying facial expressions of emotion. *Frontiers in Psychology*, 11:329.
- Lachmann, R., Schulze, S., Nieke, M., Seidl, C. e Schaefer, I. (2016). System-level test case prioritization using machine learning. Em *IEEE International Conference on Machine Learning and Applications*, páginas 361–368. IEEE.
- Lan, Y., Wei, S., Liu, R. e Zhao, Y. (2016). Creating video summarization from emotion perspective. Em *IEEE International Conference on Signal Processing*, páginas 1112–1117. IEEE.
- Lanata, A., Valenza, G., Greco, A., Gentili, C., Bartolozzi, R., Bucchi, F., Frendo, F. e Scilingo, E. P. (2014). How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving. *IEEE Transactions on Intelligent Transportation Systems*, 16(3):1505–1517.
- LaValle, S. M. e Kuffner, J. J. (2001). Rapidly-exploring random trees: Progress and prospects. *Algorithmic and Computational Robotics: new Directions*, (5):293–308.
- Lazari, E. C. C. e Codinhoto, E. (2017). O estresse de professores em sala de aula. *Revista FAROL*, 2(2):98–112.
- Lazarus, R. S. (1966). *Psychological Stress and the Coping Process*. New York: McGraw-Hill.
- Le, A. S., Suzuki, T. e Aoki, H. (2020). Evaluating driver cognitive distraction by eye tracking: From simulator to driving. *Transportation Research Interdisciplinary perspectives*, 4:100087.
- LeCun, Y., Bengio, Y. e Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.
- Lee, B.-G. e Chung, W.-Y. (2016). Wearable glove-type driver stress detection using a motion sensor. *IEEE Transactions on Intelligent Transportation Systems*, 18(7):1835–1844.
- Lee, D. S., Chong, T. W. e Lee, B. G. (2016). Stress events detection of driver by wearable glove system. *IEEE Sensors Journal*, 17(1):194–204.
- Lee, S., Lee, T., Yang, T., Seomoon, E., Yoon, C. e Kim, S.-P. (2018). Neural correlates of anxiety induced by environmental events during driving. Em *IEEE Region 10 Conference*, páginas 0446–0450. IEEE.
- Lee, S., Lee, T., Yang, T., Yoon, C. e Kim, S.-P. (2020). Detection of drivers' anxiety invoked by driving situations using multimodal biosignals. *Processes*, 8(2):155.
- Lee, Y.-C. e Winston, F. K. (2016). Stress induction techniques in a driving simulator and reactions from newly licensed drivers. *Transportation Research part F: Traffic Psychology and Behaviour*, 42:44–55.
- Li, B. e Han, L. (2013). Distance weighted cosine similarity measure for text classification. Em *International Conference on Intelligent Data Engineering and Automated Learning*, páginas 611–618. Springer.
- Li, H., Liao, X., Li, C., Huang, H. e Li, C. (2011). Edge detection of noisy images based on cellular neural networks. *Communications in Nonlinear Science and Numerical Simulation*, 16(9):3746–3759.

- Li, J., Lv, J., Oh, B.-S., Lin, Z. e Yu, Y. J. (2020). Identification of stress state for drivers under different gps navigation modes. *IEEE Access*.
- Li, S., Cui, L., Zhu, C., Li, B., Zhao, N. e Zhu, T. (2016). Emotion recognition using kinect motion capture data of human gaits. *PeerJ*, 4:e2364.
- Li, W., Cui, Y., Ma, Y., Chen, X., Li, G., Zeng, G., Guo, G. e Cao, D. (2021). A spontaneous driver emotion facial expression (defe) dataset for intelligent vehicles: Emotions triggered by video-audio clips in driving scenarios. *IEEE Transactions on Affective Computing*.
- Lin, K.-S., Lee, A., Yang, Y.-H., Lee, C.-T. e Chen, H. H. (2013). Automatic highlights extraction for drama video using music emotion and human face features. *Neurocomputing*, 119:111–117.
- Liu, Z., Song, G., Cai, J., Cham, T.-J. e Zhang, J. (2019). Conditional adversarial synthesis of 3D facial action units. *Neurocomputing*, 355:200–208.
- Lombardi, L. e Marcolin, F. (2020). Psychological stress detection by 2d and 3d facial image processing. Em *Progresses in Artificial Intelligence and Neural Systems*, páginas 163–173. Springer.
- Lundberg, U., Kadefors, R., Melin, B., Palmerud, G., Hassmén, P., Engström, M. e Dohns, I. E. (1994). Psychophysiological stress and emg activity of the trapezius muscle. *International Journal of Behavioral medicine*, 1(4):354–370.
- Ma, Y.-F., Lu, L., Zhang, H.-J. e Li, M. (2002). A user attention model for video summarization. Em *ACM International Conference on Multimedia*, páginas 533–542.
- MacLean, D., Roseway, A. e Czerwinski, M. (2013). Moodwings: a wearable biofeedback device for real-time stress intervention. Em *International Conference on Pervasive Technologies Related to Assistive Environments*, páginas 1–8.
- Magana, V. C. e Munoz-Organero, M. (2017). Toward safer highways: predicting driver stress in varying conditions on habitual routes. *IEEE Vehicular Technology Magazine*, 12(4):69–76.
- Malik, M. e Hussain, A. (2017). Helpfulness of product reviews as a function of discrete positive and negative emotions. *Computers in Human Behavior*, 73:290–302.
- Manganaro, F., Pini, S., Borghi, G., Vezzani, R. e Cucchiara, R. (2019). Hand gestures for the human-car interaction: the briareo dataset. Em *International Conference on Image Analysis and Processing*, páginas 560–571. Springer.
- Manseer, M. e Riener, A. (2014). Evaluation of driver stress while transiting road tunnels. Em *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, páginas 1–6.
- Martinez, B., Valstar, M. F., Jiang, B. e Pantic, M. (2017). Automatic analysis of facial actions: A survey. *IEEE Transactions on Affective Computing*.
- Martins, F. d. F., Lopes, R. M. F. e Farina, M. (2014). Nível de estresse e principais estressores do motorista de transporte coletivo. *Boletim-Academia Paulista de Psicologia*, 34(87):523–536.
- Martins, P., Langlois, T. e Chambel, T. (2011). Movieclouds: content-based overviews and exploratory browsing of movies. Em *International Academic MindTrek Conference: Envisioning Future Media Environments*, páginas 133–140.

- Masood, K. e Alghamdi, M. A. (2019). Modeling mental stress using a deep learning framework. *IEEE Access*, 7:68446–68454.
- Masood, S., Rai, A., Aggarwal, A., Doja, M. N. e Ahmad, M. (2020). Detecting distraction of drivers using convolutional neural network. *Pattern Recognition Letters*, 139.
- Mayo, L. M. e Heilig, M. (2019). In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of Stress*, 10:100166.
- McCraty, R., Atkinson, M., Tiller, W. A., Rein, G. e Watkins, A. D. (1995). The effects of emotions on short-term power spectrum analysis of heart rate variability. *American Journal of cardiology*, 76(14):1089–1093.
- Mehmood, I., Sajjad, M., Rho, S. e Baik, S. W. (2016). Divide-and-conquer based summarization framework for extracting affective video content. *Neurocomputing*, 174:393–403.
- Mehrotra, K., Mohan, C. K. e Ranka, S. (1997). *Elements of artificial neural networks*. MIT press.
- Money, A. G. e Agius, H. (2013). ‘Mind the gap’: Evaluating user physiological response for multi-genre video summarisation. Em *International BCS Human Computer Interaction Conference*, páginas 1–6.
- Money, A. G. e Harry, A. (2008). Video summarisation: A conceptual framework and survey of the state of the art. *Journal of visual communication and image representation*, 2(19):121–143.
- Moriyama, T., Abdelaziz, K. e Shimomura, N. (2012). Face analysis of aggressive moods in automobile driving using mutual subspace method. Em *International Conference on Pattern Recognition*, páginas 2898–2901. IEEE.
- Muhammad, K., Hussain, T. e Baik, S. W. (2020). Efficient cnn based summarization of surveillance videos for resource-constrained devices. *Pattern Recognition Letters*, 130:370–375.
- Mühlbacher-Karrer, S., Mosa, A. H., Faller, L.-M., Ali, M., Hamid, R., Zangl, H. e Kyamakya, K. (2017). A driver state detection system—combining a capacitive hand detection sensor with physiological sensors. *IEEE Transactions on Instrumentation and Measurement*, 66(4):624–636.
- Munla, N., Khalil, M., Shahin, A. e Mourad, A. (2015). Driver stress level detection using HRV analysis. Em *International Conference on Advances in Biomedical Engineering*, páginas 61–64. IEEE.
- Mylavarapu., G. e Thomas, J. P. (2017). A multi-task machine learning approach for comorbid patient prioritization. Em *IEEE International Conference on Big Data*, páginas 3877–3881.
- Nandini, M. e Bhajantri, N. U. (2020). A comprehension of contemporary effort for tracking of lip. *International Journal of Bioinformatics Research and Applications*, 16(1):85–100.
- Napoletano, P. e Rossi, S. (2018). Combining heart and breathing rate for car driver stress recognition. Em *IEEE International Conference on Consumer Electronics*, páginas 1–5. IEEE.

- Nayak, J., Naik, B. e Behera, H. (2015). Fuzzy c-means (fcm) clustering algorithm: a decade review from 2000 to 2014. Em *Computational Intelligence in Data Mining-volume 2*, páginas 133–149. Springer.
- Ngo, C.-W., Ma, Y.-F. e Zhang, H.-J. (2003). Automatic video summarization by graph modeling. Em *IEEE International Conference on Computer Vision*, páginas 104–109. IEEE.
- Noh, Y., Kim, S. e Yoon, Y. (2019). Evaluation on diversity of drivers' cognitive stress response using eeg and ecg signals during real-traffic experiment with an electric vehicle. Em *IEEE Intelligent Transportation Systems Conference*, páginas 3987–3992. IEEE.
- Oliphant, T. E. (2006). *A guide to NumPy*, volume 1. Trelgol Publishing USA.
- Ollander, S., Godin, C., Charbonnier, S. e Campagne, A. (2016). Feature and sensor selection for detection of driver stress. Em *PhyCS*, páginas 115–122.
- Ong, K.-M. e Kameyama, W. (2010). Video summary based on human bio-signals-evaluation of its applicability to first-time viewers. Em *International Symposium on Aware Computing*, páginas 251–256. IEEE.
- Ooi, C. S., Seng, K. P., Ang, L.-M. e Chew, L. W. (2014). A new approach of audio emotion recognition. *Expert systems with applications*, 41(13):5858–5869.
- Ooi, J. S. K., Ahmad, S. A., Chong, Y. Z., Ali, S. H. M., Ai, G. e Wagatsuma, H. (2016). Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. Em *IEEE EMBS Conference on Biomedical Engineering and Sciences*, páginas 365–369. IEEE.
- Ortoncelli, A. R., Silva, L., Bellon, O. R. P., Oliveria, T. M. d. e Daga, J. (2020). Summarizing driving behavior to support driver stress analysis. *IEEE International Conference and Workshops on Face & Gesture Recognition*.
- Ortony, A. e Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3):315.
- Paredes, P. E., Ordonez, F., Ju, W. e Landay, J. A. (2018). Fast & furious: detecting stress with a car steering wheel. Em *Conference on Human Factors in Computing Systems*, páginas 1–12.
- Paschalidis, E., Choudhury, C. F. e Hess, S. (2018). Modelling the effects of stress on gap-acceptance decisions combining data from driving simulator and physiological sensors. *Transportation Research part F: Traffic Psychology and Behaviour*, 59:418–435.
- Patrício, D. I. e Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153:69–81.
- Pediaditis, M., Giannakakis, G., Chiarugi, F., Manousos, D., Pampouchidou, A., Christinaki, E., Iatraki, G., Kazantzaki, E., Simos, P. G., Marias, K. et al. (2015). Extraction of facial features as indicators of stress and anxiety. Em *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, páginas 3711–3714. IEEE.
- Peng, S. e Nagao, K. (2018). Automatic evaluation of presenters' discussion performance based on their heart rate. Em *International Conference on Computer Supported Education*, páginas 27–34.

- Peng, W.-T., Chu, W.-T., Chang, C.-H., Chou, C.-N., Huang, W.-J., Chang, W.-Y. e Hung, Y.-P. (2011). Editing by Viewing: automatic home video summarization by viewing behavior analysis. *IEEE Transactions on Multimedia*, 13(3):539–550.
- Pereira, T., Almeida, P. R., Cunha, J. P. e Aguiar, A. (2017). Heart rate variability metrics for fine-grained stress level assessment. *Computer methods and Programs in Biomedicine*, 148:71–80.
- Perini, A., Susi, A. e Avesani, P. (2012). A machine learning approach to software requirements prioritization. *IEEE Transactions on Software Engineering*, 39(4):445–461.
- Petersen, K., Feldt, R., Mujtaba, S. e Mattsson, M. (2008). Systematic mapping studies in software engineering. Em *International Conference on Evaluation and Assessment in Software Engineering*, páginas 1–10.
- Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2):1883.
- Philippot, P., Chappelle, G. e Blairy, S. (2002). Respiratory feedback in the generation of emotion. *Cognition & Emotion*, 16(5):605–627.
- Pietikäinen, M., Hadid, A., Zhao, G. e Ahonen, T. (2011). *Computer vision using local binary patterns*, volume 40. Springer Science & Business Media.
- Pretorius, R. e Budgen, D. (2008). A mapping study on empirical evidence related to the models and forms used in the UML. Em *ACM-IEEE international symposium on Empirical software engineering and measurement*, páginas 342–344.
- Qayyum, H., Majid, M. and ul Haq, E. e Anwar, S. M. (2019). Generation of personalized video summaries by detecting viewer’s emotion using Electroencephalograph. Em *Journal of Visual Communication and Image Representation*.
- Qu, W., Zhang, Y., Wang, D., Feng, S. e Yu, G. (2015). Semantic movie summarization based on string of IE-ROLENETS. *Computational Visual Media*, 1(2):129–141.
- Rabkin, J. G. e Struening, E. L. (1976). Life events, stress, and illness. *Science*.
- Rahmat, R. F., Chairunnisa, T., Gunawan, D. e Sitompul, O. S. (2016). Skin color segmentation using multi-color space threshold. Em *International Conference on Computer and Information Sciences*, páginas 391–396. IEEE.
- Raiturkar, P., Kleinsmith, A., Keil, A., Banerjee, A. e Jain, E. (2016). Decoupling light reflex from pupillary dilation to measure emotional arousal in videos. Em *ACM Symposium on Applied Perception*, páginas 89–96.
- Rastgoo, M. N., Nakisa, B., Maire, F., Rakotonirainy, A. e Chandran, V. (2019). Automatic driver stress level classification using multimodal Deep Learning. *Expert Systems with Applications*, 138:1–11.
- Rastgoo, M. N., Nakisa, B., Rakotonirainy, A., Chandran, V. e Tjondronegoro, D. (2018). A critical review of proactive detection of driver stress levels based on multimodal measurements. *ACM Computing Surveys*, 51(5):1–35.
- Reynolds, D. A. (2009). Gaussian Mixture Models. *Encyclopedia of Biometrics*, 741.

- Rigas, G., Goletsis, Y. e Fotiadis, D. I. (2011). Real-time driver's stress event detection. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):221–234.
- Rish, I. et al. (2001). An empirical study of the naive bayes classifier. Em *Workshop on empirical methods in artificial intelligence*, volume 3, páginas 41–46.
- Rodrigues, J. G., Kaiseler, M., Aguiar, A., Cunha, J. P. S. e Barros, J. (2015). A mobile sensing approach to stress detection and memory activation for public bus drivers. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3294–3303.
- Romão, T., Centieiro, P., Fião, G., Correia, N. e Dias, A. E. (2017). Designing a system for the automatic generation of sport video summaries. Em *ACM Symposium on Engineering Interactive Computing Systems*, páginas 69–74.
- Rony, R. J. e Ahmed, N. (2019). Monitoring driving stress using hrv. Em *International Conference on Communication Systems & Networks*, páginas 417–419. IEEE.
- Rumelhart, D. E., Hinton, G. E. e Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.
- Saeed, A. e Trajanovski, S. (2017). Personalized driver stress detection with multi-task neural networks using physiological signals. *arXiv preprint arXiv:1711.06116*.
- Salehin, M. M. e Paul, M. (2017). Affective video events summarization using EMD decomposed EEG signals (EDES). Em *International Conference on Digital Image Computing: Techniques and Applications*, páginas 1–6. IEEE.
- Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N. e Mozos, O. M. (2015). Stress detection using wearable physiological sensors. Em *International Work-Conference on the Interplay Between Natural and Artificial Computation*, páginas 526–532. Springer.
- Scheirer, J., Fernandez, R., Klein, J. e Picard, R. W. (2002). Frustrating the user on purpose: a step toward building an affective computer. *Interacting with Computers*, 14(2):93–118.
- Schneiderman, N., Ironson, G. e Siegel, S. D. (2005). Stress and health: psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology*, 1:607–628.
- Selye, H. (1956). *The stress of life*, McGraw-Hill, New York, NY, US, 1.
- Selye, H. (1965). The stress syndrome. *The American Journal of Nursing*, páginas 97–99.
- Seppälä, J., De Vita, I., Jämsä, T., Miettunen, J., Isohanni, M., Rubinstein, K., Feldman, Y., Grasa, E., Corripio, I., Berdun, J. et al. (2019). Mobile phone and wearable sensor-based mhealth approaches for psychiatric disorders and symptoms: systematic review. *JMIR Mental Health*, 6(2):e9819.
- Serra, J. e Soille, P. (2012). *Mathematical morphology and its applications to image processing*, volume 2. Springer Science & Business Media.
- Shaver, P., Schwartz, J., Kirson, D. e O'connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52(6):1061.
- Shelley, K., Shelley, S. e Lake, C. (2001). Pulse oximeter waveform: photoelectric plethysmography. *Clinical monitoring*, páginas 420–428.

- Shin, L. M. e Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology*, 35(1):169–191.
- Shorten, C. e Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):1–48.
- Sihem, N., Bitam, S. e Mellouk, A. (2019). A body area network for ubiquitous driver stress monitoring based on ECG signal. Em *International Conference on Wireless and Mobile Computing, Networking and Communications*, páginas 1–6. IEEE.
- Silva, L. C. d., Vedovello, S. A. S., Vedovello Filho, M., Meneghin, M. d. C., Ambrosano Bovi, G. M. e Degan, V. V. (2019). Anxiety and oral habits as factors associated with malocclusion. *CRANIO®*, páginas 1–5.
- Silveira, M. J., Marasca, A. L., Marcon, M. e Ortoncelli, A. R. (2020). Priorização de atividades de desenvolvimento de software com uma rede neural artificial. *Computer on the Beach*.
- Simonyan, K. e Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, M. e Queyam, A. B. (2013). A novel method of stress detection using physiological measurements of automobile drivers. *International Journal of Electronics Engineering*, 5(2):13–20.
- Singh, R. R., Conjeti, S. e Banerjee, R. (2013). A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals. *Biomedical Signal Processing and Control*, 8(6):740–754.
- Singhal, A., Kumar, P., Saini, R., Roy, P. P., Dogra, D. P. e Kim, B.-G. (2018). Summarization of videos by analyzing affective state of the user through crowdsourcing. *Cognitive Systems Research*, 52:917–930.
- Smith, J. R., Joshi, D., Huet, B., Hsu, W. e Cota, J. (2017). Harnessing ai for augmenting creativity: Application to movie trailer creation. Em *ACM International Conference on Multimedia*, páginas 1799–1808.
- Soman, K., Alex, V. e Srinivas, C. (2013). Analysis of physiological signals in response to stress using ecg and respiratory signals of automobile drivers. Em *International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing*, páginas 574–579. IEEE.
- Soman, K., Sathiyaa, A. e Suganthi, N. (2014). Classification of stress of automobile drivers using radial basis function kernel support vector machine. Em *International Conference on Information Communication and Embedded Systems*, páginas 1–5. IEEE.
- Speelman, D. (2014). Logistic regression. *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 43:487–533.
- Sriramprakash, S., Prasanna, V. D. e Murthy, O. R. (2017). Stress detection in working people. *Procedia Computer Science*, 115:359–366.
- Starcke, K. e Brand, M. (2012). Decision making under stress: a selective review. *Neuroscience & Biobehavioral Reviews*, 36(4):1228–1248.

- Steinhauser, K., Leist, F., Maier, K., Michel, V., Pärsch, N., Rigley, P., Wurm, F. e Steinhauser, M. (2018). Effects of emotions on driving behavior. *Transportation Research part F: Traffic Psychology and Behaviour*, 59:150–163.
- Tang, W., Gao, Y., Liu, L., Xia, T., He, L., Zhang, S., Guo, J., Li, W. e Xu, Q. (2020). An automatic recognition of tooth-marked tongue based on tongue region detection and tongue landmark detection via deep learning. *IEEE Access*, 8:153470–153478.
- Taylor, J. E. (2018). The extent and characteristics of driving anxiety. *Transportation Research part F: Traffic Psychology and Behaviour*, 58:70–79.
- Taylor, J. E., Deane, F. P. e Podd, J. (2007). Diagnostic features, symptom severity, and help-seeking in a media-recruited sample of women with driving fear. *Journal of Psychopathology and Behavioral Assessment*, 29(2):81–91.
- Taylor, J. E., Deane, F. P. e Podd, J. V. (2000). Determining the focus of driving fears. *Journal of Anxiety Disorders*, 14(5):453–470.
- Tejero-de Pablos, A., Nakashima, Y., Sato, T., Yokoya, N., Linna, M. e Rahtu, E. (2018). Summarization of user-generated sports video by using deep action recognition features. *IEEE Transactions on Multimedia*, 20(8):2000–2011.
- Thurston, R. C., Rewak, M. e Kubzansky, L. D. (2013). An anxious heart: anxiety and the onset of cardiovascular diseases. *Progress in Cardiovascular Diseases*, 55(6):524–537.
- Tiwari, V. e Bhatnagar, C. (2021). A survey of recent work on video summarization: approaches and techniques. *Multimedia Tools and Applications*, páginas 1–35.
- Tonella, P., Susi, A. e Palma, F. (2013). Interactive requirements prioritization using a genetic algorithm. *Information and Software Technology*, 55(1):173–187.
- Truong, B. T. e Venkatesh, S. (2007). Video abstraction: A systematic review and classification. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 3(1):3–es.
- Ul Haq, I., Ullah, A., Muhammad, K., Lee, M. Y. e Baik, S. W. (2019). Personalized movie summarization using deep cnn-assisted facial expression recognition. Em *Complexity*.
- Urbano, M., Alam, M., Ferreira, J., Fonseca, J. e Simíões, P. (2017). Cooperative driver stress sensing integration with ecall system for improved road safety. Em *International Conference on Smart Technologies*, páginas 883–888. IEEE.
- Varghese, J. e Nair, K. R. (2015). An algorithmic approach for general video summarization. Em *International Conference on Advances in Computing and Communications*, páginas 7–11. IEEE.
- Varghese, J. e Nair, K. R. (2018). Applying video summarization to aerial surveillance. Em *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR X*. International Society for Optics and Photonics.
- Verma, R., Mitra, B. e Chakraborty, S. (2019). Avoiding stress driving: Online trip recommendation from driving behavior prediction. Em *IEEE International Conference on Pervasive Computing and Communications*, páginas 1–10. IEEE.

- Vermeulen, J., MacDonald, L., Schöning, J., Beale, R. e Carpendale, S. (2016). Heartefacts: Augmenting mobile video sharing using wrist-worn heart rate sensors. Em *ACM Conference on Designing Interactive Systems*, páginas 712–723.
- Vhaduri, S., Ali, A., Sharmin, M., Hovsepian, K. e Kumar, S. (2014). Estimating drivers' stress from GPS traces. Em *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, páginas 1–8.
- Viegas, C., Lau, S.-H., Maxon, R. e Hauptmann, A. (2018). Towards independent stress detection: A dependent model using facial action units. Em *International Conference on Content-Based Multimedia Indexing*, páginas 1–6. IEEE.
- Vikram, K. e Padmavathi, S. (2017). Facial parts detection using viola jones algorithm. Em *International Conference on Advanced Computing and Communication Systems*, páginas 1–4. IEEE.
- Viola, P. e Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.
- Wang, J.-S., Lin, C.-W. e Yang, Y.-T. C. (2013). A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition. *Neurocomputing*, 116:136–143.
- Wang, K. e Guo, P. (2020). An ensemble classification model with unsupervised representation learning for driving stress recognition using physiological signals. *IEEE Transactions on Intelligent Transportation Systems*.
- Wang, L. (2008). Feature selection with kernel class separability. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9):1534–1546.
- Wang, X., Ding, X., Su, S., Li, Z., Riese, H., Thayer, J. F., Treiber, F. e Snieder, H. (2009). Genetic influences on heart rate variability at rest and during stress. *Psychophysiology*, 46(3):458–465.
- Wang, Z., Yu, J., He, Y. e Guan, T. (2014). Affection arousal based highlight extraction for soccer video. *Multimedia Tools and Applications*, 73(1):519–546.
- Weiszflog, W. (2009). Michaelis, dicionário da língua portuguesa. *Editora Melhoramentos*.
- Wen, W.-H., Liu, G.-Y., Mao, Z.-H., Huang, W.-J., Zhang, X., Hu, H., Yang, J. e Jia, W. (2018). Toward constructing a real-time social anxiety evaluation system: Exploring effective heart rate features. *IEEE Transactions on Affective Computing*.
- Wipf, D. P. e Rao, B. D. (2004). Sparse bayesian learning for basis selection. *IEEE Transactions on Signal processing*, 52(8):2153–2164.
- Wold, S., Esbensen, K. e Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.
- Wolfinger, R. e O'connell, M. (1993). Generalized linear mixed models a pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*, 48(3-4):233–243.
- Xian, Y., Li, J., Zhang, C. e Liao, Z. (2015). Video highlight shot extraction with time-sync comment. Em *International Workshop on Hot Topics in Planet-scale mObile computing and online Social neTworking*, páginas 31–36.

- Xu, B., Wang, X. e Jiang, Y.-G. (2016). Fast summarization of user-generated videos: Exploiting semantic, emotional, and quality clues. *IEEE MultiMedia*, 23(3):23–33.
- Xu, M., Jin, J. S., Luo, S. e Huang, Q. (2007). Using timing to detect horror shots in horror movies. Em *International Workshop on Advanced Pattern Recognition*, páginas 225–231.
- Yuan, Z., Lu, T., Wu, D., Huang, Y. e Yu, H. (2011). Video summarization with semantic concept preservation. Em *International Conference on Mobile and Ubiquitous Multimedia*, páginas 109–112.
- Yüce, A., Gao, H., Cuendet, G. L. e Thiran, J.-P. (2016). Action units and their cross-correlations for prediction of cognitive load during driving. *IEEE Transactions on Affective Computing*, 8(2):161–175.
- Zadeh, A., Chong Lim, Y., Baltrusaitis, T. e Morency, L.-P. (2017). Convolutional experts constrained local model for 3D facial landmark detection. Em *International Conference on Computer Vision Workshops*, páginas 2519–2528.
- Zepf, S., Hernandez, J., Schmitt, A., Minker, W. e Picard, R. W. (2020). Driver emotion recognition for intelligent vehicles: A survey. *ACM Computing Surveys*, 53(3):1–30.
- Zhang, H., Feng, L., Li, N., Jin, Z. e Cao, L. (2020). Video-based stress detection through deep learning. *Sensors*, 20(19):5552.
- Zhao, S., Yao, H., Sun, X., Jiang, X. e Xu, P. (2013). Flexible presentation of videos based on affective content analysis. Em *International Conference on Multimedia Modeling*, páginas 368–379. Springer.
- Zontone, P., Affanni, A., Bernardini, R., Piras, A. e Rinaldo, R. (2018). Low-complexity classification algorithm to identify drivers' stress using electrodermal activity (EDA) measurements. Em *The World Thematic Conference-Biomedical Engineering and Computational Intelligence*, páginas 25–33. Springer.