

Washington Rodrigo Dias da Silva

**Classificação de embriaguez a partir de imagens
faciais usando redes neurais convolucionais:
Uma abordagem baseada em respostas
fisiológicas induzidas pelo álcool**

Sorocaba, SP

10 de Setembro de 2021

Washington Rodrigo Dias da Silva

Classificação de embriaguez a partir de imagens faciais usando redes neurais convolucionais: Uma abordagem baseada em respostas fisiológicas induzidas pelo álcool

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação (PPGCC-So) da Universidade Federal de São Carlos como parte dos requisitos exigidos para a obtenção do título de Mestre em Ciência da Computação. Linha de pesquisa: Computação Científica e Inteligência Computacional.

Universidade Federal de São Carlos – UFSCar

Centro de Ciências em Gestão e Tecnologia – CCGT

Programa de Pós-Graduação em Ciência da Computação – PPGCC-So

Orientador: Prof. Dr. Siovani Cintra Felipussi

Sorocaba, SP

10 de Setembro de 2021

Silva, Washington Rodrigo Dias da

Classificação de embriaguez a partir de imagens faciais usando redes neurais convolucionais: uma abordagem baseada em respostas fisiológicas induzidas pelo álcool / Washington Rodrigo Dias da Silva -- 2021. 93f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus Sorocaba, Sorocaba
Orientador (a): Siovani Cintra Felipussi
Banca Examinadora: Siovani Cintra Felipussi, Anderson de Rezende Rocha, Katti Faceli
Bibliografia

1. Classificação de embriaguez. 2. Redes neurais convolucionais. 3. Análise de imagens térmicas. I. Silva, Washington Rodrigo Dias da. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Maria Aparecida de Lourdes Mariano -
CRB/8 6979



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências em Gestão e Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Washington Rodrigo Dias da Silva, realizada em 24/09/2021.

Comissão Julgadora:

Prof. Dr. Siovani Cintra Felipussi (UFSCar)

Prof. Dr. Anderson de Rezende Rocha (UNICAMP)

Profa. Dra. Katti Faceli (UFSCar)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Computação.

Aos meus pais, Normeide e Benedito;
Ao meu orientador, Prof. Dr. Siovani Cintra Felipussi;
Àqueles cujas vidas foram afetadas por uma imprudência que poderia ser evitada.

Agradecimentos

Agradeço,

aos meus pais, Normeide e Benedito, pelo apoio e dedicação que possibilitaram a realização deste mestrado.

ao meu orientador, Prof. Dr. Siovani Cintra Felipussi, por aceitar me orientar neste primeiro caminho da carreira acadêmica; por seus valiosos ensinamentos; por acreditar no potencial desta pesquisa; pelo apoio e incentivo durante os momentos difíceis e por sua paciência durante todo o período de desenvolvimento desta dissertação.

aos meus amigos, que estiveram presentes em todos os momentos desta jornada e contribuíram para que eu chegasse ao final com minha sanidade preservada; ou quase.

aos meus companheiros de laboratório e professores, que de alguma maneira contribuíram para a realização desta dissertação.

à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), pelo apoio financeiro que possibilitou a realização desta pesquisa (Código de Financiamento 001).

“Uma jornada de mil milhas começa com um único passo”
(Lao Tzu)

Resumo

Devido ao crescimento populacional e ao aumento da frota de veículos motorizados em todo o mundo, a segurança viária ainda é um desafio global que se mantém a despeito das campanhas de conscientização e dos instrumentos legais de fiscalização. Em sua maioria, as causas de um acidente de trânsito são evitáveis e neste sentido a embriaguez ao volante requer uma atenção especial pois algumas doses de álcool podem ser suficientes para prejudicar funções motoras e sensoriais necessárias para se conduzir um veículo de maneira segura. Deste modo, um indivíduo nesta condição pode representar um risco à segurança viária ainda que apresente um índice de alcoolemia abaixo do limite legal. Considerando que a embriaguez ao volante é um problema recorrente, esta dissertação propôs um método complementar para a fiscalização de motoristas alcoolizados; o qual considera as respostas fisiológicas induzidas pelo álcool como indicativos de intoxicação. Neste contexto, o critério adotado para a classificação de embriaguez foi o comportamento da temperatura facial. Os sinais pertinentes a tal fenômeno foram identificados por meio da aplicação de uma Rede Neural Convolucional a um conjunto de imagens térmicas faciais provenientes de indivíduos sóbrios e ébrios. Esta abordagem possibilitou a extração automática de características invariantes à pose e a visualização dos atributos determinantes para a inferência do modelo, permitindo a interpretação das variações de temperatura nas regiões da testa e do nariz como um indicativo de embriaguez. Outro comportamento identificado pelo modelo foi a distribuição assimétrica de temperatura entre os lados esquerdo e direito do rosto em indivíduos embriagados, o qual não foi evidenciado em estudos anteriores. Deste modo, foi possível demonstrar a relação entre os padrões de distribuição da temperatura facial e o ciclo metabólico do álcool, evidenciando a viabilidade do uso de imagens para reconhecer os efeitos fisiológicos decorrentes do consumo de bebidas alcoólicas.

Palavras-chaves: Computação forense. Farmacocinética do álcool. Segurança viária. Classificação de embriaguez. Aprendizado profundo. Redes Neurais Convolucionais. Análise de imagens térmicas.

Abstract

Due to population growth and an increasing motorized vehicle fleet worldwide, road safety is still a global challenge that persists despite the awareness campaigns and the legal enforcement instruments. Most of the road safety risk factors are completely preventable, and in this sense the drunk-driving needs a special attention since a few doses of alcohol can be enough to impair motor and sensory functions required for safe driving. Hence, an individual in such condition may represent a road safety risk even presenting a blood alcohol concentration within the local legal limits. Since drunk-driving is a recurrent problem, in this master thesis we proposed a complementary method for drunk drivers enforcement considering the physiological responses induced by alcohol as intoxication indicatives. We addressed the drunkenness classification problem by identifying facial temperature distribution behaviors related to alcohol consumption. Such features were extracted from individuals facial thermal images in both drunk and sober states by a Convolutional Neural Network, which enabled the automatic pose invariant feature extraction and the model inference interpretation. This approach allowed us to identify temperature variations between the forehead and nose regions as a drunkenness indicative and also to identify an asymmetric facial temperature distribution in inebriated individuals that was not showed by previous studies. Hence, we were able to demonstrate a relation between facial temperature distribution patterns and the alcohol metabolic cycle; evidencing the viability of using images to recognize physiological effects caused by alcohol.

Key-words: Forensic Computing. Alcohol Pharmacokinetics. Road Safety. Drunkenness Classification. Deep Learning. Convolutional Neural Networks. Thermal Imaging.

Lista de ilustrações

Figura 1 – Algoritmo para estimativa de características geométricas da pupila. . .	31
Figura 2 – Impressão térmica dos vasos sanguíneos faciais. Em (a) e (b) são ilustradas as assinaturas térmicas da atividade vascular facial de um indivíduo em estado sóbrio e ébrio, respectivamente.	33
Figura 3 – Arquitetura base de uma Rede Neural Convolutacional.	34
Figura 4 – Representação gráfica do processo de <i>transfer learning</i>	35
Figura 5 – Subtração de termogramas. Em (a) e (b) são apresentados os quadros inicial e final de uma determinada sequência de capturas 30 minutos após a ingestão das doses de bebida alcoólica. Em (c) é ilustrado o resultado da subtração entre a última e a primeira captura.	44
Figura 6 – Representações abstraídas pelo modelo pré-treinado. O mapa de características ilustrado foi extraído da imagem térmica facial de um determinado indivíduo em estado de embriaguez (90 minutos após o consumo das doses de bebida alcoólica) na 2 ^o camada de convolução da rede VGG16.	49
Figura 7 – Representação gráfica do processo de validação cruzada (estratificada).	50
Figura 8 – Representação gráfica do processo de mapeamento de ativação de classes.	53
Figura 9 – Curva de aprendizado do modelo base.	59
Figura 10 – Curva de aprendizado do modelo base após o processo de <i>fine-tuning</i>	61
Figura 11 – Curva de aprendizado do modelo final.	64
Figura 12 – Curva de aprendizado do modelo final após o processo de <i>fine-tuning</i>	65
Figura 13 – Mapa de ativação de classes. Em (a) é ilustrada a assinatura térmica facial de um indivíduo em estado sóbrio. Em (b), (c) e (d) são apresentadas as características relacionadas à embriaguez consideradas determinantes pelo modelo em cada período de aquisição subsequente à ingestão de álcool. A relevância de cada atributo extraído pela RNC pode ser interpretada pela intensidade dos <i>pixels</i> destacados em amarelo.	67
Figura 14 – Indicativos de embriaguez apresentados por homens e mulheres. Nesta Figura é possível observar que a distribuição da temperatura facial induzida pelo álcool pode variar em extensão e intensidade ao longo do tempo para indivíduos do sexo oposto.	72
Figura 15 – Indicativos de embriaguez apresentados por indivíduos com peso e idade distintos. Em (b) e (c) é possível observar que a temperatura nas regiões do nariz e testa aumenta e se propaga para a região superior esquerda do rosto, enquanto em (f) e (g) a temperatura da testa e do lado direito do rosto diminui.	73

Figura 16 – Indicativos de embriaguez apresentados por indivíduos com o mesmo sexo, peso e idade. Nos três primeiros períodos de aquisição é possível observar padrões semelhantes de assinatura térmica, contudo, nota-se que em (d) a temperatura da testa e do lado esquerdo do rosto aumenta enquanto em (h) esta característica permanece semelhante ao período anterior. 74

Lista de tabelas

Tabela 1 – Evolução das arquiteturas de RNCs entre os anos de 2012 e 2020. . . .	46
Tabela 2 – Distribuições utilizadas no processo de busca randômica.	52
Tabela 3 – Resultados do processo de busca randômica na fase de transferência de aprendizado.	58
Tabela 4 – Resultados do processo de busca manual na fase de transferência de aprendizado.	58
Tabela 5 – Seleção do modelo base para a etapa de <i>fine-tuning</i>	59
Tabela 6 – Matriz de confusão do modelo base (etapa de transferência de aprendi- zado).	59
Tabela 7 – Resultados obtidos após a atualização dos pesos das camadas de convo- lução mais profundas.	60
Tabela 8 – Seleção do modelo base para a etapa final de treinamento.	60
Tabela 9 – Matriz de confusão do modelo base após a etapa de <i>fine-tuning</i>	60
Tabela 10 – Resultados do processo de busca randômica na fase de treinamento do modelo final (7 partições de validação cruzada).	61
Tabela 11 – Avaliação dos hiperparâmetros obtidos no processo de busca randômica (conjunto de teste).	62
Tabela 12 – Seleção do número de unidades ocultas para a expansão da topologia do modelo.	63
Tabela 13 – Resultados do processo de otimização bayesiana para busca de hiperpa- râmetros.	63
Tabela 14 – Seleção do modelo final para a etapa de <i>fine-tuning</i>	64
Tabela 15 – Matriz de confusão do modelo final.	64
Tabela 16 – Avaliação do modelo final após a etapa de <i>fine-tuning</i>	65
Tabela 17 – Matriz de confusão do modelo final após o processo de <i>fine-tuning</i> . . .	65
Tabela 18 – Especificação dos hiperparâmetros do modelo de classificação de embri- aguez.	66
Tabela 19 – Quadros-chave selecionados de acordo com o critério de máxima entropia.	66

Lista de abreviaturas e siglas

RNC	Rede Neural Convolutacional
CTB	Código Brasileiro de Trânsito
g	gramas
kg	kilogramas
h	hora
k	mililitros
ml	mililitros
mg	miligramas
L	Litro
VGG	<i>Visual Geometry Group</i>
Grad-CAM	<i>Gradient-weighted Class Activation Mapping</i>
AUC	<i>Area Under Curve</i>
ROC	<i>Receiver Operating Characteristic</i>
Relu	<i>Rectified linear unit</i>
RMSprop	<i>Root Mean Square propagation</i>
Adam	<i>Adaptive moment estimation</i>
N°	Número
CIFAR	<i>Canadian Institute For Advanced Research</i>
min.	minutos
ACT	Água Corporal Total
ADH	<i>Alcohol Dehydrogenase</i>
ALDH	<i>Aldehyde Dehydrogenase</i>
°C	graus Celsius

Lista de símbolos

f_l^k	Matriz de características para a l – <i>ésima</i> camada e k – <i>ésimo</i> neurônio
$f_l^k(p, q)$	Elemento da matriz de características, p e q representam os índices de linha e coluna, respectivamente
\sum_c	Soma dos produtos entre os <i>pixels</i> e o filtros de convolução no canal c
$\sum_{x,y}$	Soma dos produtos entre o <i>pixel</i> localizado na linha x , coluna y e o filtros de convolução
$i_c(x, y)$	Elemento (linha x , coluna y) do c – <i>ésimo</i> canal de uma imagem
$e_l^k(u, v)$	Elemento (linha u , coluna v) do k – <i>ésimo</i> <i>kernel</i> da l – <i>ésima</i> camada
E	Entropia
$\sum_{i=0}^{L-1}$	Soma dos produtos entre a probabilidade de ocorrência do i – <i>ésimo</i> nível de cinza relativo a um determinado <i>pixel</i> e seu logaritmo
$\sum_{j=0}^{L-1}$	Soma dos produtos entre a probabilidade de ocorrência do j – <i>ésimo</i> nível de cinza relativo a um determinado <i>pixel</i> e seu logaritmo
L	Níveis de tons de cinza de uma imagem
p_{ij}	Probabilidade de ocorrência do i – <i>ésimo</i> e j – <i>ésimo</i> níveis de cinza
$\log_2(p_{ij})$	Logaritmo da probabilidade de ocorrência do i – <i>ésimo</i> e j – <i>ésimo</i> níveis de cinza
ε	Constante de valor arbitrário positivo
w_n^c	gradiente (w) correspondente ao n – <i>ésimo</i> mapa de características da classe c

Sumário

	Introdução	25
1	ÁLCOOL E DIREÇÃO: UMA NOVA PERSPECTIVA SOBRE OS MÉTODOS DE FISCALIZAÇÃO	29
1.1	Identificação de alterações oculares e oculomotoras induzidas pelo álcool	31
1.2	Identificação de alterações na distribuição de temperatura facial induzidas pelo álcool	32
1.2.1	Extração de características faciais por Redes Neurais Convolucionais	33
1.3	Considerações sobre o capítulo	36
2	CONTEXTUALIZAÇÃO BIBLIOGRÁFICA	39
2.1	Considerações sobre o capítulo	40
3	METODOLOGIA	43
3.1	Descrição da base de dados	43
3.2	Protocolo experimental	44
3.3	Materiais	54
3.4	Considerações sobre o capítulo	54
4	RESULTADOS	57
4.1	Seleção dos hiperparâmetros de treinamento	57
4.2	Revisão das características extraídas pela RNC	66
4.3	Considerações sobre o capítulo	67
5	DISCUSSÃO	69
5.1	Influência do sexo nas respostas fisiológicas induzidas pelo álcool	71
5.2	Influência do peso e da idade nas respostas fisiológicas induzidas pelo álcool	71
5.3	Influência genética nas respostas fisiológicas induzidas pelo álcool	73
5.4	Influência do tipo de bebida alcoólica nas respostas fisiológicas induzidas pelo álcool	74
5.5	Influência do esvaziamento gástrico nas respostas fisiológicas induzidas pelo álcool	75
5.6	Influência do ciclo circadiano nas respostas fisiológicas induzidas pelo álcool	75

5.7	Influência de fatores ambientais nas respostas fisiológicas induzidas pelo álcool	76
5.8	Considerações sobre o capítulo	77
	Conclusão	81
	Referências	85

Introdução

Em Setembro de 2015, a Assembleia Geral das Nações Unidas estabeleceu um conjunto de metas para promover o avanço global da segurança viária. Esta ação colocou em evidência um problema de saúde pública internacional que sem as devidas diretivas e a atuação dos países membros poderia comprometer o progresso da Agenda de Metas de Desenvolvimento Sustentável para 2030 ([UNITED NATIONS, 2020](#)).

As causas de um acidente de trânsito podem ser associadas a diversos fatores, como o comportamento do condutor, a infraestrutura das vias, a segurança dos veículos ou até mesmo uma combinação destes elementos os quais em sua maioria podem ser prevenidos. Neste contexto, convém salientar que a ingestão de bebidas alcoólicas demanda uma atenção especial por se tratar de uma causa evitável e que implica sérios riscos à segurança viária. Anualmente, 1.35 milhão de pessoas em todo o mundo são vítimas de acidentes de trânsito fatais e cerca de 35% destes são causados por motoristas embriagados ([WORLD HEALTH ORGANIZATION, 2018b](#)).

Segundo [Fell e Voas \(2006\)](#), restringir o limite legal de alcoolemia a 0.05% poderia reduzir o número de óbitos ocasionados pela violência no trânsito em até 18%. Contudo, mesmo em concentrações abaixo deste índice, o álcool compromete funções motoras e sensoriais que reduzem as aptidões necessárias para se conduzir um veículo; tornando um indivíduo nesta condição até três vezes mais propenso a causar um acidente de trânsito em comparação a uma pessoa que não consumiu bebida alcoólica ([MARTIN et al., 2013](#); [COMPTON](#); [BERNING, 2015](#)).

Deste modo, acredita-se que o atual formato de fiscalização de embriaguez ao volante poderia se beneficiar de indicativos complementares à concentração de álcool no sangue ou no ar alveolar. Bebidas alcoólicas e outras substâncias psicoativas promovem sinais de intoxicação que não são observáveis visualmente mas importantes para a identificação de indivíduos nesta condição. Neste contexto, a variação de temperatura da pele representa um indício possivelmente determinante para a categorização do estado de embriaguez. O álcool afeta o processo de termorregulação do organismo humano induzindo respostas fisiológicas de perda de calor, como a vasodilatação. Este comportamento vascular implica na expansão dos vasos sanguíneos e no aumento do fluxo de sangue, permitindo a transferência de calor entre o corpo e o ambiente por meio da pele ([KALANT; LE, 1983](#); [FREUND; O'BRIEN; YOUNG, 1994](#); [DANEL; LIBERSA; TOUITOU, 2001](#); [YODA et al., 2005](#); [BRICK; ERICKSON, 2009](#)).

A difusão de calor entre os vasos sanguíneos e o tecido ao seu redor possibilita que as alterações de temperatura na pele sejam observadas por meio de imagens térmicas,

correlacionando a intensidade dos *pixels* e a atividade vascular local (BUDDHARAJU et al., 2007). Assim, a hipótese abordada nesta dissertação é de que a embriaguez pode ser identificada a partir dos efeitos fisiológicos decorrentes da intoxicação alcoólica¹ por meio da análise de características faciais extraídas da imagem térmica do rosto de um indivíduo.

A aplicação de termografia em estudos sobre biometria facial e análise de sentimentos tem apresentado resultados promissores (BUDDHARAJU et al., 2007; KHAN; WARD; INGLEBY, 2009; KOPACZKA et al., 2018). No contexto da classificação de embriaguez, alguns trabalhos abordaram a variação das distribuições de temperatura facial e ocular como indicativo de alcoolemia, demonstrando métodos estatísticos e baseados em aprendizado de máquina para associar tais comportamentos aos estados de sobriedade e ebriedade (KOUKIOU; ANASTASSOPOULOS, 2012; KOUKIOU; ANASTASSOPOULOS, 2015; KOUKIOU; ANASTASSOPOULOS, 2016a; KOUKIOU; ANASTASSOPOULOS, 2018a; HERMOSILLA et al., 2018).

Ainda que as abordagens propostas tenham apresentado resultados promissores, os métodos adotados para extração de características são dependentes de *pixels* específicos selecionados manualmente e sensíveis à variações de pose. Tais fragilidades podem comprometer a eficácia dos classificadores propostos e reduzir sua capacidade de generalização. Deste modo, esta dissertação propõe uma abordagem para classificação de embriaguez orientada a dados, a qual se baseia no uso de Redes Neurais Convolucionais (RNCs) para extração de características e técnicas de visualização para interpretação dos resultados.

Considerando que tais métodos possibilitam a correlação entre os efeitos toxicológicos do álcool e os indicativos de embriaguez reconhecidos pelo modelo de classificação, acredita-se que a abordagem proposta é passível de aplicação em outros contextos, como veículos inteligentes. Neste cenário, seria possível conceder ao automóvel a capacidade de analisar o estado do motorista e permitir sua condução apenas sob determinadas condições. Outro exemplo de aplicação a se considerar é o controle de acesso, visto a possibilidade de se restringir a entrada de indivíduos alcoolizados em ambientes nos quais sua presença poderia representar um risco a sua própria segurança e às demais pessoas, como aeronaves, hospitais ou linhas de produção, por exemplo.

Objetivos e contribuições

O objetivo desta dissertação é propor um método para classificar o estado de embriaguez por meio da identificação de respostas fisiológicas induzidas pelo álcool a partir de imagens térmicas faciais. Isto posto, a presente pesquisa contribui com a literatura expandindo a base de conhecimento nos seguintes aspectos:

¹ Neste estudo, a condição de intoxicação alcoólica foi considerada como a alteração de funções cognitivas, motoras, fisiológicas e/ou comportamentais decorrente do consumo de álcool; independentemente do nível de gravidade associado aos sintomas.

- Extração de características invariantes à pose de imagens térmicas faciais por meio de RNCs.
- Classificação de embriaguez a partir de imagens térmicas faciais.
- Interpretabilidade da inferência do modelo.
- Identificação de fatores físicos, fisiológicos e ambientais que podem influenciar o comportamento dos indicativos de embriaguez reconhecidos pelo modelo.
- Verificação da viabilidade do uso imagens para a identificação de sinais de intoxicação causados por outras substâncias psicoativas.

Organização

Esta dissertação está organizada da seguinte maneira: No Capítulo 1 são apresentados conceitos fundamentais sobre os efeitos do álcool no organismo humano e as técnicas computacionais que permitem detectá-los. No Capítulo 2 são apresentados os trabalhos que abordaram o problema de classificação de embriaguez e que fundamentam esta pesquisa. O Capítulo 3 detalha as propriedades do conjunto de dados e as técnicas empregadas para a elaboração do modelo de classificação. Os critérios considerados para avaliação da abordagem proposta e as métricas utilizadas para a análise de desempenho do classificador são apresentadas no Capítulo 4. No Capítulo 5 os resultados são analisados e a possibilidade do uso de imagens para o reconhecimento de intoxicação por substâncias psicoativas, como o álcool, é avaliada. Por fim, o Capítulo 6 apresenta as considerações finais desta dissertação e direcionamentos para pesquisas futuras.

1 Álcool e direção: Uma nova perspectiva sobre os métodos de fiscalização

O consumo de bebidas alcoólicas é uma prática que faz parte da rotina de inúmeras pessoas, seja em situações de socialização ou de forma recreativa ([WORLD HEALTH ORGANIZATION, 2018a](#)). Em doses moderadas o álcool proporciona sensações de relaxamento e euforia, motivo pelo qual o seu uso tenha sido aceito culturalmente e difundido em diversas sociedades. No entanto, seus efeitos psicoativos são dependentes do volume consumido e implicam em reações adversas à medida em que sua concentração no sangue aumenta ([OSCAR-BERMAN; MARINKOVIĆ, 2007](#); [JONES, 2019b](#)).

O consumo excessivo de álcool acarreta alterações fisiológicas que prejudicam funções psicomotoras e cognitivas, reduzindo o desempenho de tarefas complexas que exijam atenção, coordenação motora e raciocínio lógico; como a condução de um veículo ([OSCAR-BERMAN; MARINKOVIĆ, 2007](#); [BRICK; ERICKSON, 2009](#); [JONES, 2019b](#)). Contudo, [Martin et al. \(2013\)](#) evidencia que tais aptidões podem ser comprometidas em níveis menores de alcoolemia (0.02-0.03%), aumentando significativamente o risco de um acidente de trânsito.

No atual formato de fiscalização de embriaguez ao volante, a categorização de um indivíduo quanto a sua capacidade de condução é realizada por meio de um teste de alcoolemia (etilômetro) ou pela observação dos sinais de alteração psicomotora ([CONSELHO NACIONAL DE TRÂNSITO, 2013](#)). Ainda que este modelo seja amplamente utilizado como uma estratégia de prevenção a tal infração ([WORLD HEALTH ORGANIZATION, 2018a](#)), é possível que motoristas sob o efeito de álcool não sejam identificados em certas ocasiões.

De acordo com [Brick e Erickson \(2009\)](#), indivíduos nesta condição podem não exibir sinais visíveis de intoxicação mesmo com uma concentração de álcool no sangue que os torne inaptos para dirigir. Ademais, diversos trabalhos na literatura evidenciam que os efeitos adversos da intoxicação alcoólica tendem a se manter durante a fase descendente da curva de alcoolemia e afetam funções cognitivas e motoras mesmo após o teor de álcool no sangue retornar a zero ([OGDEN; MOSKOWITZ, 2004](#); [LIU; HO, 2010](#); [MARTIN et al., 2013](#); [STARKEY; CHARLTON, 2014](#); [VERSTER et al., 2014](#); [CHARLTON; STARKEY, 2015](#)).

Deste modo, acredita-se que o reconhecimento das respostas fisiológicas induzidas pelo álcool poderia contribuir com o protocolo adotado para a identificação de condutores alcoolizados; uma vez que o organismo humano pode manifestar sinais que representem o

comprometimento das aptidões necessárias para se conduzir um veículo de maneira segura (VONGHIA et al., 2008; BRICK; ERICKSON, 2009; WORLD HEALTH ORGANIZATION, 2018a). O comportamento pupilar é um exemplo de indicativo que pode ser relacionado à deterioração das funções visuais. Esta correlação é possível pois a dilatação da pupila implica no aumento da incidência de luz sobre a retina, afetando a discriminação visual, a sensibilidade ao contraste e a resposta acomodativa do globo ocular¹ (CASTRO et al., 2014; CASARES-LÓPEZ et al., 2020; CASARES-LOPEZ et al., 2021).

De acordo com Castro et al. (2014), Casares-López et al. (2020) e Casares-Lopez et al. (2021), tais distúrbios podem ocorrer em níveis de alcoolemia inferiores a 0.02% e contribuem significativamente para o detrimento das habilidades de condução. O estudo de Tyson et al. (2021) corrobora estes resultados, identificando o comprometimento de funções cognitivas e visuais a partir de concentrações de 0.015 gramas de álcool por 100 mililitros de sangue. Dentre os comportamentos observados, o autor destaca a redução do processamento de informações visuais como direção, velocidade e movimentação associada à ocorrência de movimentos involuntários do globo ocular.

Segundo o Conselho Nacional de Trânsito (2013), a concentração de álcool por litro de ar alveolar² correspondente a este índice de alcoolemia está abaixo do limite estabelecido como infração ao Código Brasileiro de Trânsito (CTB). No entanto, ao considerar os estudos supracitados é possível inferir que um indivíduo nesta condição ainda representa um risco para a segurança viária.

Em adição aos distúrbios visuais (CASTRO et al., 2014; CASARES-LÓPEZ et al., 2020; CASARES-LOPEZ et al., 2021) e oculomotores (TYSON et al., 2021), há outros sinais que podem indicar a influência do álcool no organismo, tais como: aumento de temperatura da pele (decorrente do efeito de vasodilatação), das taxas de batimentos cardíacos e respiração, rubor facial, hipotensão, hipotermia e alterações comportamentais (KALANT; LE, 1983; ERIKSSON, 2001; VONGHIA et al., 2008). As variações de temperatura relacionadas aos efeitos do álcool no sistema cardiovascular, bem como as alterações oculares que indicam o comprometimento do sistema visual, são características promissoras a se considerar visto sua correlação com os efeitos de pós-intoxicação alcoólica.

O comportamento ocular pode ser associado ao desempenho de funções cognitivas (TYSON et al., 2021), que de acordo com Lantman et al. (2017) podem permanecer comprometidas por até 18 horas após um episódio de consumo excessivo de bebida alcoólica. Segundo Reinberg et al. (2010), a influência do álcool no processo de termorregulação do organismo também representa um sinal de que funções cognitivas e psicomotoras

¹ A acomodação é uma função visual que possibilita a adaptação do foco para que a imagem produzida na retina se mantenha nítida em diferentes distâncias (CASARES-LOPEZ et al., 2021).

² A conversão entre as concentrações de álcool no sangue e no ar alveolar expirado foi calculada de acordo a proporção 2,300:1, a qual segundo Jones (2019b) reduz o viés da estimativa quando as amostras são coletadas durante a fase descendente da curva de alcoolemia.

relacionadas ao tempo de reação podem estar comprometidas. [Devaney, Graham e Greeley \(2003\)](#) evidenciam que a oscilação da temperatura corporal pode se estender por até 14 horas após o período de ingestão.

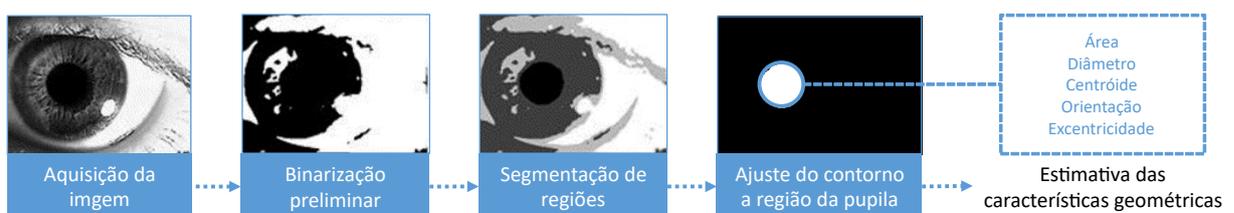
Deste modo, constata-se a necessidade de se considerar as respostas fisiológicas, induzidas pelo consumo de bebidas alcoólicas, como um parâmetro complementar às concentrações de álcool obtidas por meio de etilômetro ou exame de sangue. As técnicas computacionais que possibilitam a identificação destes sinais serão abordadas nas seções a seguir.

1.1 Identificação de alterações oculares e oculomotoras induzidas pelo álcool

Conforme mencionado na seção anterior, o consumo de bebidas alcoólicas, mesmo que moderado, compromete o desempenho de funções visuais e cognitivas de um indivíduo. As respostas fisiológicas associadas a tais efeitos podem ser relacionadas a dois sinais: a dilatação da pupila e a ocorrência de movimentos involuntários do globo ocular.

As oscilações do comportamento pupilar podem ser detectadas a partir de características geométricas, como o diâmetro. Tais propriedades podem ser obtidas com o emprego de técnicas de segmentação de imagens e a atribuição de um contorno elíptico à região da pupila para a estimativa dos parâmetros morfológicos associados com a sua geometria ([SANTIS; IACOVIELLO, 2006](#)). Para identificar os elementos da estrutura ocular a partir de uma imagem, [Santis e Iacoviello \(2006\)](#) propuseram um algoritmo de segmentação binária recursiva; em que as duas regiões obtidas após uma etapa de binarização são binarizadas sucessivamente. Segundo os autores, quatro níveis de segmentação são suficientes para se obter uma representação simplificada do olho humano e preservar as características relevantes para a estimativa dos parâmetros morfológicos da pupila (Figura 1).

Figura 1 – Algoritmo para estimativa de características geométricas da pupila.



Fonte: Adaptado de: [Santis e Iacoviello \(2006\)](#).

Este método também possibilita a identificação de dois atributos determinantes para a detecção de movimentos involuntários do globo ocular: as coordenadas dos eixos

horizontal e vertical do contorno elíptico. Tais propriedades possibilitam que a partir de um vídeo da região periocular se obtenha a posição do centro da pupila detectada. Assim, os parâmetros obtidos em cada quadro podem ser comparados com valores de referência para determinar a ocorrência de movimentos anômalos (ZHANG et al., 2021).

1.2 Identificação de alterações na distribuição de temperatura facial induzidas pelo álcool

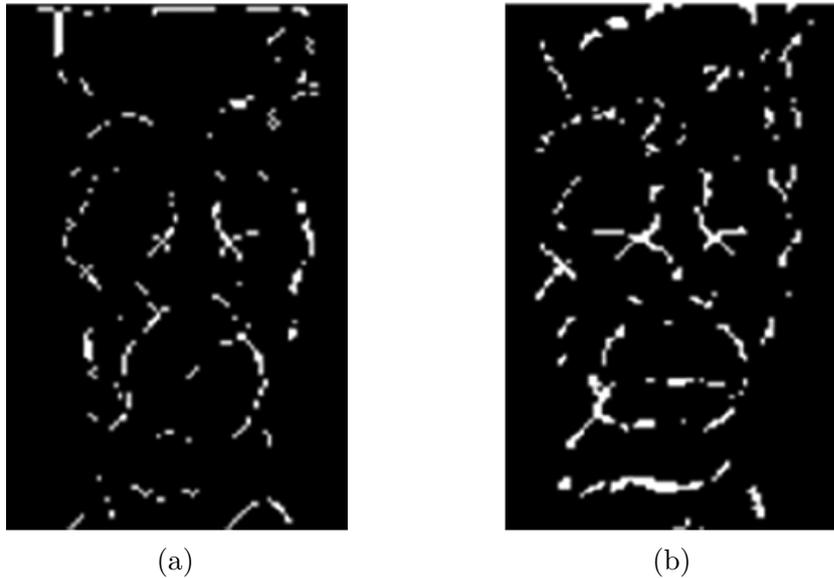
O efeito de vasodilatação induzido pelo álcool implica no aumento do fluxo sanguíneo em regiões periféricas do corpo, como o rosto, acarretando a elevação da temperatura cutânea nestes locais. Tal comportamento possibilita a identificação de respostas fisiológicas induzidas pelo álcool considerando o uso de imagens térmicas. Neste contexto, o estado de embriaguez pode ser caracterizado pela variação dos níveis de intensidade dos *pixels* e pela maneira como tais valores estão distribuídos pela área de interesse (KALANT; LE, 1983; WOLF; TÜZÜN; TÜZÜN, 1999; ERIKSSON, 2001; VONGHIA et al., 2008; KAWANO, 2010; FERNÁNDEZ-CUEVAS et al., 2015; BLANER et al., 2017).

Na literatura, diversos métodos foram propostos para a identificação de padrões de distribuição de temperatura associados à embriaguez, como redes neurais artificiais (KOUKIOU; ANASTASSOPOULOS, 2015), padrões locais binários (KOUKIOU; ANASTASSOPOULOS, 2018b) e cadeias markovianas (KOUKIOU, 2021). Tais modelos consideram a região da testa como o local mais promissor para a detecção de alterações na assinatura térmica, no entanto, Kubicek et al. (2018) e Kubicek et al. (2019) demonstraram por meio de técnicas de segmentação e análise de agrupamento que a região do nariz também apresenta variações de temperatura à medida em que a concentração de álcool no sangue aumenta.

Em adição ao comportamento dos *pixels*, a influência do álcool na atividade vascular facial também pode ser analisada a partir da impressão térmica dos vasos sanguíneos. Para tal, Koukiou e Anastassopoulos (2016b) propuseram o uso da técnica de difusão anisotrópica em conjunto com a transformação morfológica *Top-hat* para destacar os *pixels* com maior intensidade; possibilitando a comparação de suas características nos estados de embriaguez e sobriedade. Na Figura 2 (b) é possível observar que os vasos sanguíneos das regiões da testa, nariz, boca e olhos do indivíduo apresentam maior atividade após o consumo de bebida alcoólica.

Segundo Kang et al. (2007), Weber et al. (2013) e Toris, Tye e Pattabiraman (2019), a atividade vascular intraocular também pode variar de acordo com o volume de bebida alcoólica consumido, o que faz da região periocular um local promissor para extração de características assim como as regiões do nariz e da testa. Deste modo, a abordagem proposta nesta dissertação se destaca dos métodos apresentados por possibilitar a identificação de

Figura 2 – Impressão térmica dos vasos sanguíneos faciais. Em (a) e (b) são ilustradas as assinaturas térmicas da atividade vascular facial de um indivíduo em estado sóbrio e ébrio, respectivamente.



Fonte: Koukiou e Anastassopoulos (2016b).

alterações pertinentes a estes locais de maneira holística. Assim, a partir de uma imagem do rosto de um indivíduo é possível identificar padrões de distribuição de temperatura que não seriam reconhecidos ao considerar tais regiões isoladamente. O método de extração de características por RNCs será detalhado na subseção a seguir.

1.2.1 Extração de características faciais por Redes Neurais Convolucionais

De acordo com LeCun, Bengio e Hinton (2015), modelos computacionais compostos por múltiplas camadas de processamento permitem o aprendizado de características em diferentes níveis de abstração. Deste modo, a topologia de uma RNC é estruturada de maneira hierárquica representando diferentes etapas de aprendizado tais como convolução, processamento não linear e subamostragem.

Durante a fase de convolução, uma imagem é dividida em segmentos locais que são filtrados por um conjunto de *kernels* em mapas de características. De acordo com Khan et al. (2020), esta operação pode ser expressa pela equação 1.1:

$$f_l^k(p, q) = \sum_c \sum_{x, y} i_c(x, y) \cdot e_l^k(u, v) \quad (1.1)$$

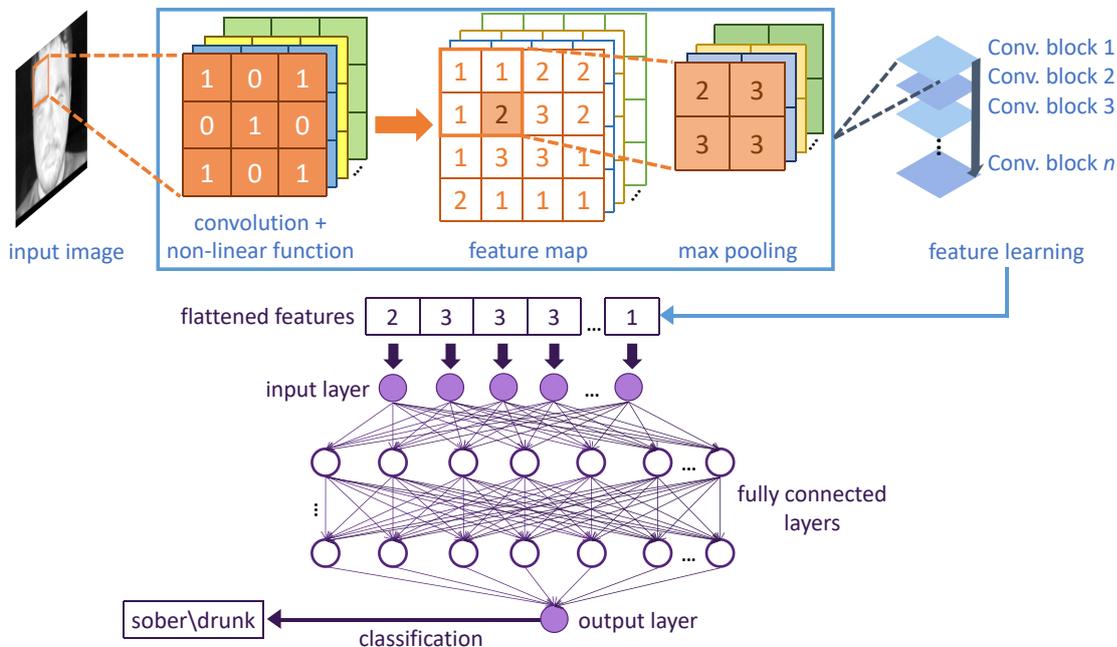
Sendo que (p, q) representa um elemento da matriz de características f obtida pelo *kernel* k na camada l . Cada atributo deste conjunto é constituído pela soma dos produtos obtidos com a multiplicação elemento a elemento dos *pixels* da imagem i (em que x e y

representam as linhas e colunas do canal c) por cada índice (u, v) do filtro de convolução e .

Deste modo, um mesmo padrão pode ser identificado em diferentes regiões de uma imagem, como a variação de intensidade dos *pixels* localizados nas áreas de interesse discutidas na seção anterior (olhos, nariz e testa). Ademais, as propriedades extraídas são invariantes à pose ou orientação, pois os mapas de características decorrentes da operação de convolução apresentarão o mesmo comportamento de sua origem (LECUN et al., 1998).

Em seguida, o espaço vetorial obtido é atribuído a uma função não linear que estabelece diferentes padrões de ativação em resposta aos estímulos da rede (dados de entrada). Na etapa de subamostragem, o mapa de características derivado da função de ativação é sumarizado semanticamente a partir da extração do valor máximo de cada segmento local. Este procedimento permite a redução de dimensionalidade das representações e proporciona invariância às distorções geométricas ao mitigar a influência da localização dos *pixels*. A organização destes processos em múltiplas camadas interconectadas (Figura 3) permite ao modelo abstrair representações de alto nível nas camadas mais profundas da rede ao combinar características de baixo nível (como contornos e bordas) obtidas nas camadas anteriores (LECUN et al., 1998; LECUN; BENGIO; HINTON, 2015; KHAN et al., 2020).

Figura 3 – Arquitetura base de uma Rede Neural Convolutacional.



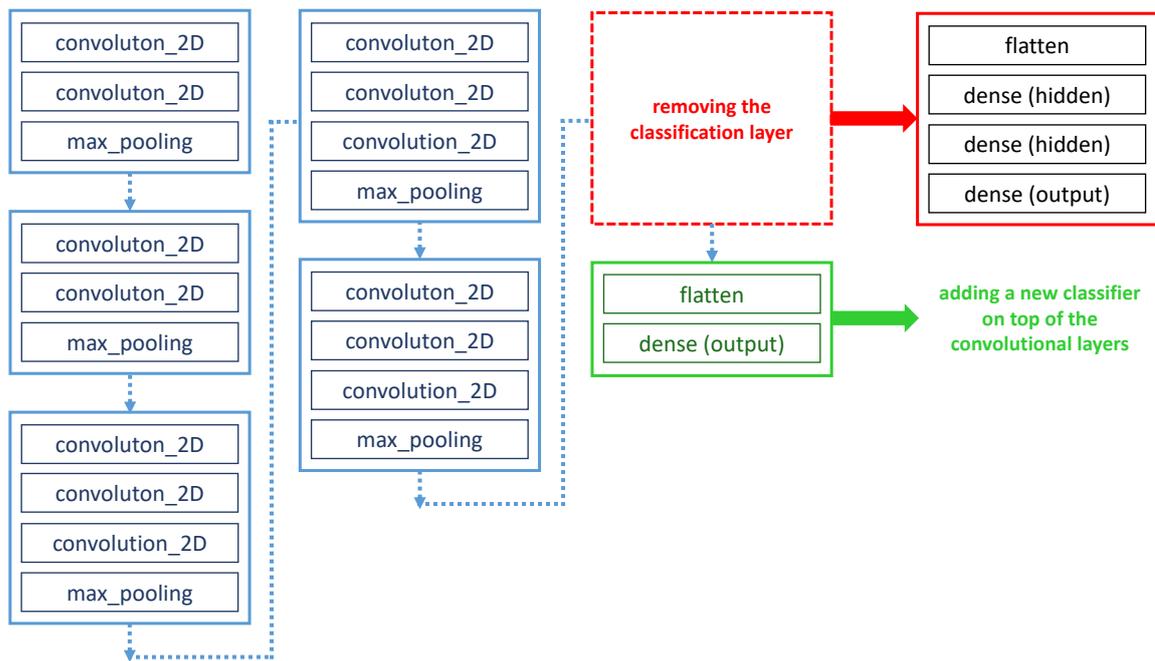
Fonte: Elaborado pelo autor.

Durante o processo de treinamento, os pesos dos filtros de convolução e das camadas totalmente conectadas são inicializados randomicamente e atualizados por um processo

de propagação reversa dos gradientes da rede (LECUN et al., 1998; LECUN; BENGIO; HINTON, 2015). Deste modo, a efetividade das aplicações baseadas em RNCs está diretamente relacionada à disponibilidade de dados para o treinamento dos modelos. Em cenários em que este recurso é restrito, a adoção do paradigma *transfer learning* representa uma solução pois possibilita a aplicação de filtros de convolução pré-treinados para a extração de características genéricas de imagens que pertencem a um domínio específico (Figura 4).

RNCs treinadas com alto volume de dados podem abstrair representações de baixo nível que são comuns a diversas classes de tarefas, assim, a substituição de sua camada totalmente conectada (destacada em vermelho na Figura 4) por um classificador específico da tarefa alvo (destacado em verde na Figura 4) induz o modelo a abstrair um novo conhecimento a partir de representações genéricas. (PAN; YANG, 2010; LECUN; BENGIO; HINTON, 2015; SHORTEN; KHOSHGOFTAAR, 2019).

Figura 4 – Representação gráfica do processo de *transfer learning*.



Fonte: Elaborado pelo autor.

Outra estratégia que pode ser adotada em cenários com disponibilidade restrita de dados é a *data-augmentation*. Este método abrange múltiplas técnicas que possibilitam incrementar o volume de um conjunto de dados, as quais são categorizadas em algoritmos que transformam imagens existentes (*data warping*) ou que geram instâncias sintéticas a partir das amostras disponíveis (*data oversampling*) (SHORTEN; KHOSHGOFTAAR,

2019).

Estratégias baseadas em *data warping* promovem transformações geométricas ou fotométricas que alteram as características estruturais de uma imagem, como rotação, orientação, translação, cor ou contraste, por exemplo. Os métodos baseados em *data oversampling*, por sua vez, envolvem o uso de técnicas como combinação de imagens, aumento do espaço de características e redes neurais generativas para a concepção de novas instâncias (SHORTEN; KHOSHGOFTAAR, 2019).

Contudo, é necessário considerar que tais abordagens alteram significativamente a estrutura das imagens base e podem comprometer a classe das amostras de acordo com o problema investigado. Neste contexto, convém salientar que os algoritmos baseados em *data oversampling* apresentam uma desvantagem em relação àqueles orientados a *data warping* pois comprometem a explicabilidade dos fatores que induzem a melhora de desempenho do modelo e a interpretação das representações aprendidas (SHORTEN; KHOSHGOFTAAR, 2019).

Deste modo, a aplicação do método de *data-augmentation* deve se basear em algoritmos que adicionem diversidade ao conjunto de dados sem descaracterizar a natureza das imagens. No contexto da classificação de embriaguez, a intensidade dos *pixels* está diretamente relacionada à assinatura térmica de cada classe, portanto, operações que alterem o contraste das amostras de treinamento devem ser evitadas pois podem induzir o modelo à classificações incorretas.

1.3 Considerações sobre o capítulo

Ao longo deste capítulo foram discutidos assuntos referentes ao consumo de bebidas alcoólicas e suas implicações para a segurança viária. Preliminarmente, discursou-se sobre a relação entre o volume de álcool ingerido e as alterações fisiológicas decorrentes de sua absorção pelo organismo. Neste contexto, foram apresentados trabalhos que evidenciam que mesmo doses moderadas desta substância podem comprometer as aptidões necessárias para se conduzir um veículo com segurança; ainda que as concentrações presentes na corrente sanguínea estejam abaixo dos limites legais. Assim, constatou-se que os atuais métodos de fiscalização podem não ser suficientes para determinar se um indivíduo está de fato sob influência de álcool. Como uma possível solução, foi sugerido um parâmetro complementar ao protocolo adotado para a identificação de condutores alcoolizados: o reconhecimento dos efeitos fisiológicos induzidos pelo consumo de bebidas alcoólicas.

Dentre os indicativos apresentados, os comportamentos pupilar e oculomotor e a assinatura térmica facial foram destacados por estarem diretamente associados à redução de desempenho de funções visuais, cognitivas e psicomotoras. Ademais, tais características apresentam alterações em níveis menores de alcoolemia e possibilitam a identificação

de condutores sob o efeito de álcool mesmo após o índice de alcoolemia ter retornado a zero. Nas seções 1.1 e 1.2 foram apresentadas técnicas computacionais que possibilitam a identificação destes sinais.

Devido a disponibilidade restrita de bases de dados públicas referentes ao problema investigado, o escopo desta dissertação foi delimitado à identificação de padrões de distribuição de temperatura facial associados ao álcool. O próximo capítulo apresentará os trabalhos que abordaram a classificação de embriaguez por meio de imagens térmicas e discutirá as contribuições da metodologia proposta em relação a estes estudos.

2 Contextualização bibliográfica

O uso de imagens térmicas em tarefas de visão computacional tem se mostrado promissor ao longo dos últimos anos, possibilitando soluções para problemas específicos ao domínio do espectro visível (como variação de luz ambiente em aplicações de reconhecimento facial) e contribuindo para o surgimento de novas linhas de investigação (como a detecção de sinais psicofisiológicos) (BUDDHARAJU et al., 2007; KOUKIOU; ANASTASSOPOULOS, 2012; KOPACZKA et al., 2018).

Considerando a influência do álcool no processo de termorregulação do corpo humano, Koukiou e Anastassopoulos (2012) investigaram a relação entre as respostas fisiológicas decorrentes da ingestão de bebidas alcoólicas e as intensidades dos *pixels* de uma imagem térmica facial. Os autores constataram que a temperatura de certas regiões do rosto tende a aumentar de acordo com a concentração de álcool no sangue possibilitando a identificação de indicativos de embriaguez a partir da distribuição da temperatura facial.

Com base nestes resultados, Koukiou e Anastassopoulos (2015) empregaram algoritmos de redes neurais artificiais para classificar o estado de embriaguez a partir de imagens térmicas. Os autores concluíram que as informações contidas no padrão de distribuição da temperatura facial são suficientes para categorizar a condição de um indivíduo, evitando a necessidade de comparação entre imagens referentes aos estados de sobriedade e embriaguez. No entanto, o modelo apresentado pode ser sensível a falsos positivos uma vez que as características consideradas determinantes para a classificação (assinatura térmica da região da testa) também podem ser relacionadas a outras causas. O estresse é um exemplo de condição fisiológica que, assim como a embriaguez, implica em aumento do fluxo sanguíneo na região da testa (PAVLIDIS et al., 2007).

Em um estudo posterior, Koukiou e Anastassopoulos (2016a) constataram que a distribuição da temperatura ocular também é influenciada pela condição fisiológica de um indivíduo. Os experimentos revelaram que no estado de embriaguez, a esclera apresenta um aumento de temperatura em relação a íris. No entanto, 5 dos 41 participantes envolvidos no estudo não apresentaram a variação de temperatura esperada. Os autores consideraram os hábitos de consumo de álcool desses participantes como uma possível causa para este comportamento, mas também destacaram que outros fatores como sexo, peso, idade e o estado psicológico dos participantes podem ter contribuído para o resultado observado.

Koukiou e Anastassopoulos (2018a) superaram as limitações dos indicativos de embriaguez estudados anteriormente adotando uma abordagem de fusão de características. O modelo apresentado é baseado em redes neurais artificiais e considera a combinação de atributos pertinentes às distribuições de temperatura facial (KOUKIOU; ANASTASSO-

POULOS, 2012) e ocular (KOUKIOU; ANASTASSOPOULOS, 2016a) para a classificação do estado de intoxicação alcoólica. Com este estudo, os autores evidenciam que a utilização de um conjunto de características possibilita a elaboração de um classificador com nível de confiança superior em relação àqueles que abordam assinaturas térmicas locais individualmente.

Contudo, ainda que este modelo tenha apresentado resultados promissores, sua robustez quanto ao reconhecimento dos indicativos de embriaguez pode ser questionada devido a sensibilidade dos métodos de extração de características à variações de pose. Tal fragilidade implica na possível seleção de *pixels* que não são determinantes para a discriminação de assinaturas térmicas associadas à embriaguez (XIE et al., 2016). Conforme mencionado na subseção 1.2.1, a aplicação de RNCs pode superar esta limitação pois possibilita a extração das principais características de uma imagem de maneira automática e é invariante à posição, orientação, escala e distorções locais (LECUN et al., 1998).

Considerando tais propriedades e os resultados alcançados por Koukiou e Anastassopoulos (2018a), Obi-Alago, Yanushkevich e Wetherley (2019) propuseram um método para classificação de embriaguez baseado na fusão de RNCs aplicadas às regiões da testa e dos olhos. Os autores utilizaram as mesmas bases de dados estudadas por Koukiou e Anastassopoulos (2015) e Koukiou e Anastassopoulos (2016a) e obtiveram resultados promissores, apresentando uma taxa de acerto de 96.92%. Contudo, o estudo não apresenta quais características contribuem de maneira mais efetiva para a classificação e como seus comportamentos variam ao longo dos períodos de aquisição das imagens. Neste contexto, acredita-se que o emprego de técnicas de visualização em adição aos métodos de fusão de características poderia contribuir substancialmente com a literatura por possibilitar a interpretação dos atributos extraídos pelas RNCs.

2.1 Considerações sobre o capítulo

Este capítulo apresentou alguns trabalhos que abordam o problema de classificação da embriaguez e que fundamentam a pesquisa proposta nesta dissertação. Apesar de suas limitações, os estudos supracitados apresentam resultados promissores e uma base de conhecimento substancial sobre o tema em questão. Neste contexto, identificou-se dois aspectos que justificam a necessidade de novas investigações: a) a relação entre as características extraídas e as respostas fisiológicas induzidas pelo álcool e b) o método de treinamento dos modelos.

Preliminarmente, verificou-se que o uso de atributos locais pode comprometer a confiabilidade dos métodos apresentados devido a possível ocorrência de alterações térmicas análogas à embriaguez por condições distintas (como o estresse). Ademais, foi observado que a sensibilidade de tais características à posição do rosto na imagem capturada pode

implicar na obtenção de valores incorretos pois sua seleção é baseada na localização de *pixels* em regiões específicas do rosto a partir de coordenadas fixas. Em seguida, discutiu-se sobre como o uso de técnicas de visualização poderia complementar os métodos de classificação apresentados na literatura, uma vez que tal abordagem possibilitaria a interpretação dos atributos extraídos e a verificação da correspondência entre os padrões identificados pelos modelos e os efeitos causados pela intoxicação alcoólica.

Deste modo, o método de classificação de embriaguez proposto nesta dissertação contribui com a base de conhecimento estabelecida pelos estudos supracitados ao correlacionar as respostas fisiológicas induzidas pelo álcool e os sinais que as caracterizam. O classificador em questão foi baseado no comportamento das propriedades intrínsecas dos dados, identificadas a partir da aplicação de uma RNC a imagens térmicas faciais. Esta estratégia possibilita que o modelo analise as características pertinentes à distribuição da temperatura facial de maneira holística, considerando informações que em conjunto podem minimizar a possibilidade de sua inferência ser influenciada por assinaturas térmicas locais sensíveis às variações de pose ou aos efeitos de outras condições fisiológicas. Ademais, o uso de RNC neste contexto também permite a visualização e a revisão das características consideradas pelo modelo para a categorização do estado de embriaguez. A metodologia de implementação da abordagem proposta será detalhada no próximo capítulo.

3 Metodologia

3.1 Descrição da base de dados

O modelo de classificação de embriaguez proposto nesta dissertação foi elaborado com o subconjunto facial da base de dados *Sober-Drunk Database*, desenvolvida por Koukiou e Anastassopoulos (2015)¹. Este conjunto de dados é composto por imagens térmicas de 41 indivíduos (31 homens e 10 mulheres), capturadas nos períodos anterior a ingestão de bebida alcoólica e em intervalos de 30, 60 e 90 minutos após o consumo. Os participantes envolvidos no experimento ingeriram 480 ml de vinho contendo 62,4 ml de álcool ao longo de uma hora. Em cada período de aquisição, uma sequência de 50 quadros com resolução de 128x160 *pixels* foi obtida com um intervalo de 100 milissegundos entre cada captura. Todos os participantes apresentaram uma concentração de álcool por litro de ar exalado superior a 0.2 mg (correspondente a aproximadamente 0.05 g por 100 ml de sangue); variando entre 0.22 e 0.37 mg/L para os homens e de 0.49 a 0.89 mg/L para as mulheres (KOUKIOU; ANASTASSOPOULOS, 2015).

Segundo Eriksson (2001), a variação de temperatura da pele, no contexto da embriaguez, está associada a um subproduto metabólico do etanol (acetaldeído) que é liberado na corrente sanguínea a uma taxa de g/kg/h (JONES, 2010). Assim, convém salientar que a despeito do comportamento variável das assinaturas térmicas faciais ao longo dos períodos de aquisição, o modelo de classificação proposto foi elaborado com o intuito de categorizar o estado dos participantes como *sóbrio* ou *embriagado*.

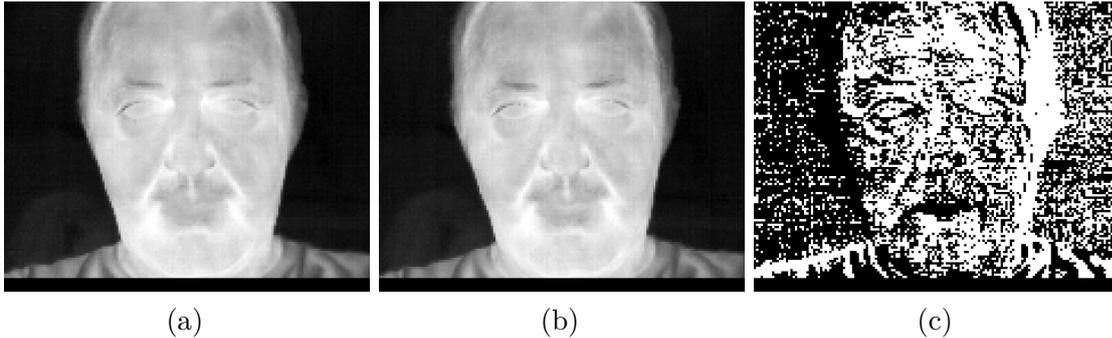
Diferentemente de câmeras sensíveis apenas ao espectro visível que captam a luz refletida em uma cena e a converte em sinais elétricos, dispositivos termográficos detectam a radiação emitida por um objeto por meio de uma matriz de sensores térmicos, de modo que a temperatura registrada por cada sensor é convertida em um *pixel*. Em vista disso, os valores atribuídos a cada *pixel* da imagem estão sujeitos a influência de flutuações térmicas decorrentes de variáveis ambientais ou do calor emitido pelos componentes eletrônicos do próprio dispositivo e tal fenômeno é referido como *drift error* (RIOU; BERREBI; BREMOND, 2004; USAMENTIAGA et al., 2014; TATTERSALL, 2016).

Câmeras termográficas de alto desempenho apresentam funções que minimizam a variação *interpixel* decorrente dos ruídos ocasionados pelas flutuações térmicas. Para verificar se o conjunto de dados utilizado neste estudo é afetado por este fenômeno foi realizada uma subtração elemento-a-elemento entre os *pixels* dos quadros final e inicial de cada sequência de capturas, de modo que na ausência de *drift error* espera-se que o valor

¹ <http://www.physics.upatras.gr/sober/>

de cada *pixel* seja 0 (TATTERSALL, 2016). Assim, é possível observar se houve variações entre as intensidades dos *pixels* ao longo do período de amostragem. A Figura 5 ilustra o resultado do experimento em questão para uma determinada sequência de capturas.

Figura 5 – Subtração de termogramas. Em (a) e (b) são apresentados os quadros inicial e final de uma determinada sequência de capturas 30 minutos após a ingestão das doses de bebida alcoólica. Em (c) é ilustrado o resultado da subtração entre a última e a primeira captura.



Fonte: Elaborado pelo autor.

Ao analisar as Figuras 5(a) e 5(b) constata-se que a ocorrência de tais distúrbios é praticamente imperceptível visualmente, no entanto, a Figura 5(c) demonstra que as variações decorrentes do *drift error* podem ocorrer em regiões de interesse determinantes para a classificação do estado de embriaguez. Deste modo, optou-se por manter cada quadro como uma amostra individual para que o modelo fosse capaz de identificar os sinais que caracterizam as respostas fisiológicas induzidas pelo álcool a despeito da influência do *drift error*. A próxima seção abordará o protocolo experimental empregado para a elaboração do classificador proposto.

3.2 Protocolo experimental

Ao analisar a base de dados descrita na seção anterior, foi observado que o participante n° 20 não possui imagens referentes ao último período de aquisição (90 minutos após o consumo das doses de bebida alcoólica). Deste modo, suas amostras não foram consideradas para o treinamento do modelo porque os demais integrantes do experimento não apresentam tal inconformidade. Assim, o conjunto final foi composto por 8.000 imagens térmicas faciais provenientes de 40 indivíduos (31 homens e 9 mulheres).

Conforme mencionado na subseção 1.2.1, a efetividade de modelos baseados em RNCs está associada ao volume de dados disponível para o seu treinamento e por conseguinte, optou-se por empregar a estratégia de *transfer learning* (descrita na subseção 1.2.1) utilizando RNCs pré-treinadas com imagens provenientes do *ImageNet*. Este banco de dados de larga escala contempla 14.197.122 de imagens de alta resolução manualmente

rotuladas e organizadas em 21.841 categorias de acordo com a hierarquia semântica do *WordNet*² (MILLER, 1995; RUSSAKOVSKY et al., 2015).

O *ImageNet* possui uma contribuição substancial para a literatura por possibilitar a popularização de técnicas de reconhecimento de objetos baseadas em RNCs e as decorrentes inovações deste campo de estudo ao longo dos últimos anos (LECUN; BENGIO; HINTON, 2015; RUSSAKOVSKY et al., 2015; KHAN et al., 2020). Khan et al. (2020) apresentaram uma revisão das principais arquiteturas de RNCs publicadas entre os anos de 2012 e 2020, classificando-as em sete categorias: exploração espacial, profundidade, largura, múltiplos caminhos, exploração de mapas de características, aumento de canais e atenção. Considerando a natureza do problema investigado nesta dissertação, verificou-se que as RNCs baseadas em exploração espacial podem ser mais adequadas para a abordagem proposta uma vez que consideram a correlação entre as informações distribuídas na vizinhança de um *pixel*.

Conforme citado na seção 1.2, a influência do álcool no aumento da temperatura facial pode ser caracterizada pela variação das intensidades dos *pixels* e pela distribuição destes valores em regiões de interesse do rosto. Neste contexto, dentre as arquiteturas apresentadas por Khan et al. (2020), verificou-se que a VGG³ representa a opção mais promissora para abordar o problema investigado devido a simplicidade e homogeneidade de sua topologia. Segundo Simonyan e Zisserman (2015), as configurações mais profundas desta arquitetura (16 e 19 camadas) com filtros de convolução 3x3 possuem o melhor desempenho para tarefas de classificação.

Considerando que a dimensionalidade dos *kernels* de uma RNC determina o nível de detalhes da informação extraída, é possível inferir que tal composição possibilitaria o encapsulamento de variações sutis nos níveis de intensidade dos *pixels* de uma imagem térmica. Outro fator que incentiva o uso da rede VGG é o fato de que esta se beneficia do conceito de profundidade para a abstração de representações de alto nível, mas não apresenta inconsistências que podem ocorrer em redes com topologias mais extensas, como a degradação da informação contida nos mapas de características durante sua propagação pelas camadas de convolução (KHAN et al., 2020). A Tabela 1 resume as potencialidades e limitações das arquiteturas revisadas por Khan et al. (2020).

A rede VGG foi treinada com um subconjunto do *ImageNet* que é disponibilizado periodicamente para o desafio anual de reconhecimento visual em larga escala. Tal conjunto compreende 1.281.167 imagens para treinamento, 50.000 para validação e 100.000

² O *WordNet* é uma base dados léxica da língua inglesa composta por substantivos, adjetivos, verbos e advérbios organizados em conjuntos de sinônimos (*synsets*) hierarquicamente estruturados, de modo que cada conjunto representa um conceito. Tais conceitos são relacionados por meio ligações semânticas entre os *synsets* (MILLER, 1995)

³ O significado da sigla VGG está associado ao nome do grupo de pesquisa responsável pela elaboração do modelo (Visual Geometry Group) (SIMONYAN; ZISSERMAN, 2015).

Tabela 1 – Evolução das arquiteturas de RNCs entre os anos de 2012 e 2020.

Arquitetura	Inovações	Limitações
LeNet	Primeira arquitetura a explorar o conceito de correlação espacial e aprender a extrair características automaticamente a partir dos <i>pixels</i> .	Limitada a tarefa de reconhecimento de dígitos manuscritos; baixa capacidade de adaptação a outras classes de problemas; filtros de convolução com alta dimensionalidade; proporciona apenas representações de baixo nível.
AlexNet	Extração de características de baixo, médio e alto nível utilizando filtros de convolução com múltiplas dimensionalidades; apresentou uma arquitetura com camadas de convolução mais profundas, alavancando sua capacidade de aprendizado e aplicabilidade; introduziu o uso de regularização em RNCs; promoveu o treinamento paralelo de modelos complexos em aceleradores gráficos.	<i>Kernels</i> inativos nas duas primeiras camadas; artefatos de <i>aliasing</i> nos mapas de características devido a alta dimensionalidade dos filtros de convolução.
ZfNet	Introduziu o conceito de ajuste de parâmetros a partir da visualização das características extraídas pelos <i>kernels</i> das camadas intermediárias; possibilitou o ajuste da topologia da rede AlexNet, aperfeiçoando seu desempenho em relação ao modelo anterior.	Requer o processamento de informações adicionais para a visualização dos mapas de características.
VGG	Propôs uma arquitetura modular com mais camadas de convolução (aperfeiçoando a capacidade representacional do modelo) e uma configuração de <i>kernels</i> homogênea, a qual se mostrou mais eficiente tanto em desempenho como em consumo de recursos computacionais.	Ainda que a composição de filtros de convolução com menor dimensionalidade tenha reduzido a complexidade computacional do processo de extração de características, o elevado número de parâmetros deste modelo restringe sua aplicação em sistemas com recursos computacionais limitados.
GoogleNet (Inception)	Apresentou um novo conceito de arquitetura denominado <i>Inception block</i> , no qual as camadas de convolução convencionais são substituídas por blocos menores que executam operações de convolução em múltiplas escalas, incorporando processos de divisão, transformação e combinação; propôs mudanças de arquitetura que reduzem do número de parâmetros; introduziu o uso de classificadores auxiliares para melhorar a taxa de convergência.	Complexa customização de parâmetros devido sua arquitetura heterogênea; possível perda de informação devido ao gargalo representacional entre as camadas.
Inception (V3)	Apresentou uma nova configuração de filtros de convolução assimétricos para reduzir o custo computacional da rede sem afetar sua capacidade de generalização.	Arquitetura complexa e sem homogeneidade, limitações semelhantes as da versão anterior.
Inception (V4)	Propôs uma atualização das arquiteturas anteriores, propiciando hierarquias de características mais profundas e múltiplos níveis de representação.	Longo período de treinamento.

Highway Networks	Introduziu uma nova abordagem para otimizar o treinamento de redes com topologias mais extensas, utilizando conexões auxiliares em adição às conexões diretas.	Alta complexidade de parametrização e implementação dos mecanismos de controle dos gradientes.
ResNet	Aperfeiçoou o conceito introduzido pelas <i>Highway Networks</i> reduzindo sua complexidade de parametrização; propôs uma arquitetura que possibilita o aprendizado de informações residuais; diminuiu a taxa de erro de redes muito profundas; reduziu o problema de desaparecimento de gradientes nas camadas inferiores.	Arquitetura complexa; degradação da informação encapsulada pelos mapas de características durante sua propagação pelas camadas da rede; requer a adaptação de seus hiperparâmetros para tarefas específicas; diversas camadas podem contribuir com pouca ou nenhuma informação; possível abstração de características redundantes.
Inception-ResNet	Combina as potencialidades da arquitetura de aprendizado residual e do bloco Inception.	Alto custo computacional.
DenseNet	Propôs mudanças de arquitetura que melhoraram o fluxo de informação entre as camadas, evitando a abstração de mapas de características redundantes e possibilitando o acesso das camadas totalmente conectadas às representações de baixo e alto nível.	Aumento expressivo no número de parâmetros devido a adição de mais mapas de características em cada camada.
Wide ResNet	Demonstrou a efetividade do uso de transformações paralelas ao expandir a largura da arquitetura ResNet e diminuir sua profundidade; propôs a utilização de <i>dropouts</i> entre as camadas de convolução ao invés de aplicá-los no interior dos blocos residuais, aprimorando a eficiência do modelo e mitigando o problema de reutilização de características redundantes.	Possível sensibilidade a <i>overfitting</i> ; elevado número de parâmetros em relação às arquiteturas mais estreitas.
Pyramidal Net	Introduziu o conceito de expansão gradual da largura da rede, evitando a perda de informação espacial entre os blocos residuais da arquitetura ResNet.	Alta complexidade temporal e espacial.
Xception	Propôs alterações na arquitetura Inception que otimizam sua capacidade de aprendizado e seu desempenho; introduziu o uso de cardinalidade nos blocos Xception e transformações convolucionais em profundidade para abstrair melhores representações.	Alto custo computacional.
ResNeXt	Aperfeiçoou a arquitetura Inception substituindo os filtros de convolução com dimensões distintas pela configuração homogênea introduzida pela rede VGG; demonstrou que o uso de cardinalidade no bloco Inception melhora o desempenho da classificação; introduziu o uso de transformações residuais agregadas para otimizar a convergência do modelo.	Alto custo computacional.

Residual Attention Network	Propôs a aplicação do conceito de atenção em redes residuais profundas, introduzindo um novo módulo à arquitetura ResNet que possibilita ao modelo "focar" em características inerentes ao contexto da classificação; aperfeiçoou a capacidade da rede de abstrair características em diferentes níveis; facilmente escalável devido a estrutura hierárquica dos blocos residuais.	Arquitetura complexa.
SE-Network	Propôs um novo bloco de operações para selecionar os mapas de características mais relevantes para a discriminação do objeto de classificação; introduziu uma arquitetura genérica, aplicável a qualquer modelo baseado em RNCs.	Quando aplicado à arquitetura ResNet considera apenas a informação residual, minimizando a influência do bloco SE e implicando na abstração de mapas de características redundantes.
Convolutional Block Attention Module	Propôs uma arquitetura simplificada do bloco de atenção, aplicável a qualquer modelo baseado em RNCs; considera tanto os mapas de características quanto a informação espacial para determinar <i>o quê</i> é relevante e <i>onde</i> tal objeto está localizado; demonstrou que o emprego das funções <i>global average pooling</i> e <i>max pooling</i> em conjunto podem alavancar o potencial representacional da rede.	Aumento da carga computacional.
CMPE-SE Network	Apresentou mudanças de arquitetura que otimizam o desempenho do bloco SE em redes residuais profundas; introduziu o conceito de competição entre os mapas de características obtidos pelos blocos residuais e pelas funções de identidade para determinar a relevância das representações.	Não suporta o mecanismo de atenção.

Fonte: Adaptado de: [Khan et al. \(2020\)](#), [Szegedy et al. \(2017\)](#).

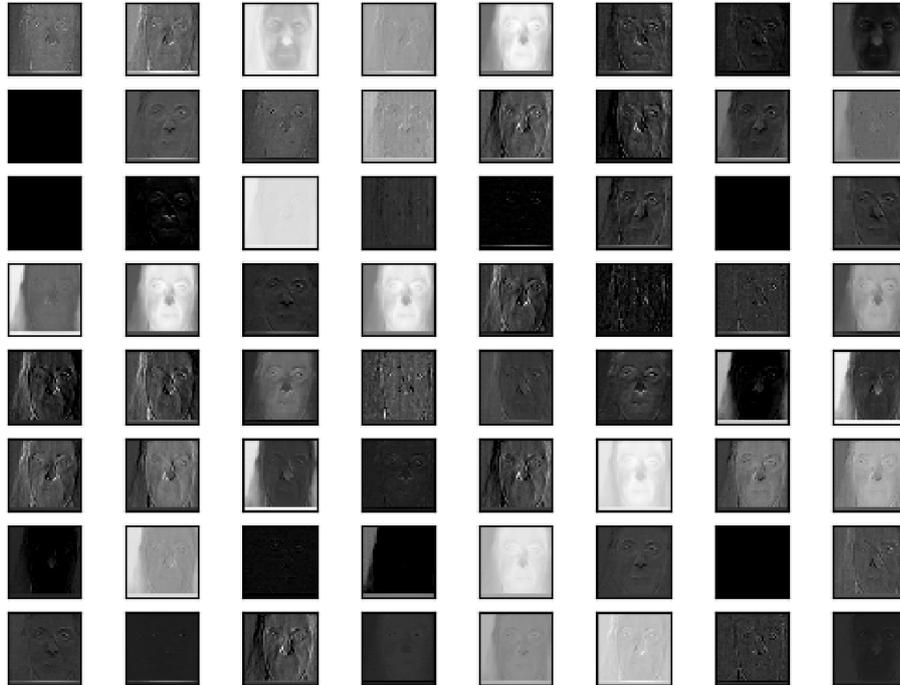
para testes, representando 1.000 classes de objetos ([SIMONYAN; ZISSERMAN, 2015](#); [RUSSAKOVSKY et al., 2015](#)). Ainda que o *ImageNet* não contemple imagens térmicas faciais de indivíduos sob o efeito de álcool, é possível observar pela Figura 6 que os filtros de convolução pré-treinados são capazes de abstrair representações de baixo nível que caracterizam a estrutura facial do rosto humano.

Deste modo, a estratégia de *transfer learning* foi empregada com o intuito que o modelo proposto assimilasse as variações de intensidade dos *pixels* que indicam o aumento de temperatura da pele decorrente das respostas fisiológicas induzidas pelo álcool. Para tal, o processo de treinamento do classificador foi dividido em duas etapas: transferência de aprendizado e composição do modelo final. Preliminarmente, a escala dos valores dos *pixels* foi normalizada para o intervalo $[0, 1]$ de acordo com a equação 3.1:

$$i' = (i - \min(i)) / (\max(i) - \min(i)) \quad (3.1)$$

Em que i' representa a imagem i reescalada a partir da razão entre a subtração

Figura 6 – Representações abstraídas pelo modelo pré-treinado. O mapa de características ilustrado foi extraído da imagem térmica facial de um determinado indivíduo em estado de embriaguez (90 minutos após o consumo das doses de bebida alcoólica) na 2ª camada de convolução da rede VGG16.



Elaborado pelo autor.

de cada elemento pelo escalar com valor mínimo ($\min(i)$) e a diferença entre os *pixels* com os valores mínimo ($\min(i)$) e máximo ($\max(i)$), respectivamente. Considerando que a rede VGG foi capaz de distinguir as propriedades dos planos de fundo e primário (Figura 6), extraíndo as características possivelmente determinantes para a classificação apenas da região do rosto, não foram empregadas etapas adicionais de pré-processamento, como a segmentação da região de interesse do plano de fundo da imagem.

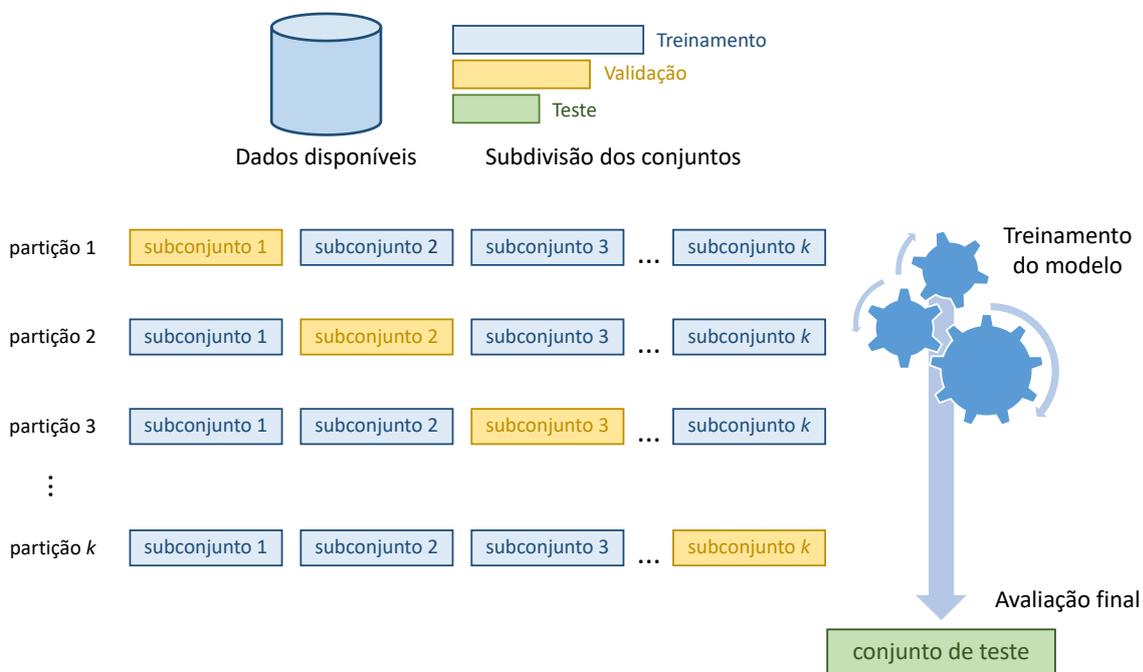
Em seguida, o conjunto de dados foi decomposto em três subconjuntos: treinamento, validação e teste. Os dois primeiros foram subdivididos a partir da aplicação do método *stratified k-fold cross validation*. Neste processo, as amostras são particionadas estratificadamente em k subconjuntos de modo que cada conjunto de validação contenha n/k elementos, em que n representa o volume de dados disponível e k o número de partições. Esta abordagem evita que uma mesma imagem seja aplicada aos procedimentos de treinamento e validação à medida em que também preserva a proporção das amostras de cada classe distribuídas nos subconjuntos, prevenindo assim a redução da capacidade de generalização dos modelos (ARLOT; CELISSE et al., 2010). A Figura 7 ilustra o procedimento descrito.

Convém salientar que a aplicação do método citado não garante a estratificação das amostras quanto ao sexo dos participantes, apenas em relação as classes que representam seus estados (sóbrio ou embriagado). Arlot, Celisse et al. (2010) sugerem que a parametrização de k considere valores entre 5 e 10, pois as iterações entre as partições correspondentes a tal intervalo se mantêm computacionalmente viáveis e valores acima de 10 não apresentam melhoras significativas de desempenho.

Considerando que k determina o volume de dados nos conjuntos de treinamento e validação, foram selecionados três valores para verificar a influência deste parâmetro no comportamento dos modelos, são eles: 5, 7 e 10. Tais medidas foram escolhidas pois propiciam a formação de subconjuntos com um volume amostras proporcional a cardinalidade das sequências de quadros obtidas de cada participante. Dado um conjunto com 3.500 imagens, por exemplo, $k=6$ resultaria em um subconjunto de validação com 583 amostras, o qual não contempla uma sequência de quadros em sua totalidade.

O conjunto de teste, por sua vez, foi subdividido previamente ao particionamento dos subconjuntos de treinamento e validação a partir da seleção das sequências de quadros dos 5 últimos participantes. As imagens pertinentes a este conjunto foram utilizadas apenas para a avaliação de desempenho dos modelos.

Figura 7 – Representação gráfica do processo de validação cruzada (estratificada).



Elaborado pelo autor.

Assim, na etapa de transferência de aprendizado (Figura 4) três modelos pré-treinados baseados na arquitetura VGG foram selecionados como classificadores base, sendo

eles: VGG16 (SIMONYAN; ZISSERMAN, 2015), VGG19 (SIMONYAN; ZISSERMAN, 2015) e VGGFace (PARKHI; VEDALDI; ZISSERMAN, 2015). Em cada experimento, os pesos dos filtros de convolução foram congelados para possibilitar a extração das características genéricas das imagens e prevenir a atualização destes parâmetros durante o processo de treinamento. Em seguida, as camadas totalmente conectadas foram substituídas por um classificador binário de camada única do tipo *perceptron* para induzir os modelos a abstrair um novo conhecimento a partir de tais representações.

Deste modo, foram consideradas imagens de apenas dois períodos de aquisição (antes da ingestão de bebida alcoólica e 90 minutos após o consumo) para que fosse possível observar a eficácia dos modelos ao distinguir as assinaturas térmicas faciais que caracterizam estas duas classes.

A capacidade de um algoritmo em encontrar padrões que caracterizem um determinado comportamento depende do ajuste de diversos hiperparâmetros, como as funções de ativação, otimização e custo, a taxa de aprendizagem, o número de *epochs*, o tamanho do *mini-batch*, dentre outros (LI et al., 2021). Para identificar o valor ótimo de cada hiperparâmetro empregou-se a técnica de busca randômica, uma vez que é computacionalmente mais eficiente e possibilita a obtenção de modelos com desempenho equiparável àqueles obtidos por meio dos métodos de busca em grade ou manual (BERGSTRA; BENGIO, 2012).

Considerando que nesta etapa utilizou-se a arquitetura e os pesos de um modelo pré-treinado, a busca de hiperparâmetros limitou-se a taxa de aprendizagem, tamanho do *mini-batch* e o número de *epochs*. Devido a natureza binária da tarefa de classificação investigada, a função de custo utilizada foi a *binary cross-entropy*. Em relação a função de otimização, empregou-se a estimativa adaptativa de momento (Adam) uma vez que tal método é reconhecido na literatura por apresentar resultados consistentes em diversas classes de problemas. (LI et al., 2021).

Nesta primeira etapa optou-se por manter a ordem das amostras contidas em cada *mini-batch* semelhante a organização original do banco de dados, uma vez que tal estratégia garante que em cada *epoch* de treinamento o modelo seja aplicado a imagens de indivíduos com características físicas distintas em ambos os estados de sobriedade e embriaguez.

A distribuição de valores utilizada no processo de busca randômica pode ser observada na Tabela 2. O número de iterações deste processo foi definido como 64 em conformidade aos resultados obtidos por Bergstra e Bengio (2012). Devido a natureza iterativa deste procedimento, que é aplicado em conjunto ao método de validação cruzada, os modelos foram avaliados de acordo com a média dos valores de acurácia (no conjunto de validação) obtidos em cada partição e seu respectivo desvio padrão, o qual sumariza a variância de desempenho da classificação.

Tabela 2 – Distribuições utilizadas no processo de busca randômica.

Hiperparâmetro	Valores
<i>learning rate</i>	[1e-3, 1e-4, 1e-5, 1e-6]
<i>mini-batch size</i>	[10, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]
<i>epochs</i>	[10, 20, 40, 60, 80, 100, 200, 300, 400, 500]

Fonte: Elaborado pelo autor.

Uma vez identificado os valores ótimos dos hiperparâmetros, o procedimento de validação cruzada foi executado novamente e os classificadores foram avaliados de acordo com seus respectivos desempenhos no conjunto de teste (de maneira similar ao processo descrito no parágrafo anterior). O modelo com a melhor acurácia média foi selecionado e aplicado ao processo de *fine-tuning*, o qual consiste em descongelar os pesos das camadas de convolução mais profundas e retreiná-los para que o modelo seja capaz de abstrair características de alto nível relacionadas a tarefa alvo da transferência de aprendizado.

Deste modo, o classificador obtido foi utilizado como base para a elaboração do modelo final em um procedimento semelhante ao descrito na etapa de transferência de aprendizado. Neste estágio, as imagens dos 4 períodos de aquisição foram utilizadas e em vista disso, a função de custo empregada no processo de treinamento foi ponderada para evitar que o desbalanceamento de classes afetasse a eficácia do classificador. Considerando que este contexto implica no aumento da complexidade da tarefa de classificação, a profundidade das camadas totalmente conectadas foi estendida para aperfeiçoar a capacidade representacional do modelo.

Segundo [Goodfellow, Bengio e Courville \(2016a\)](#), uma única camada oculta pode ser suficiente para abstrair as características não lineares dos dados e possui uma complexidade de otimização inferior quando comparada à estruturas mais profundas. Assim, optou-se pela adoção desta estratégia para aumentar a profundidade do modelo. O número de unidades ocultas desta camada foi definido seguindo o processo de busca em grade ([BERGSTRÄ; BENGIO, 2012](#)), incrementando a quantidade de neurônios em cada iteração por múltiplos de 2 até o valor máximo de 512.

Durante esta etapa as amostras contidas em cada *mini-batch* foram reordenadas de maneira aleatória⁴, ignorando a organização original do banco de dados. Considerando a eficácia do modelo base, esta estratégia introduz ao processo de treinamento do modelo final o princípio da incerteza; que por sua vez pode aperfeiçoar o processo de otimização do classificador ao longo dos *epochs* e alavancar sua capacidade de discriminação ([GOODFELLOW; BENGIO; COURVILLE, 2016d](#); [GOODFELLOW; BENGIO; COURVILLE,](#)

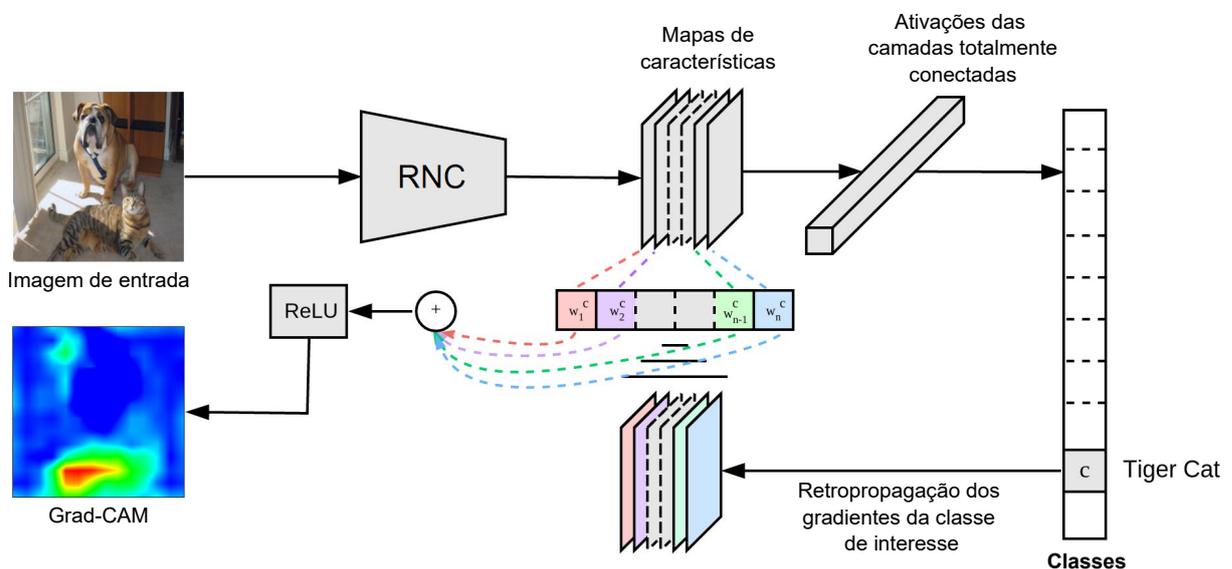
⁴ O banco de dados foi reordenado previamente a distribuição das amostras nos conjuntos de treinamento e validação, assim, no momento da subdivisão dos *mini-batches* os dados já apresentavam uma organização distinta à sequência original.

2016b).

O desempenho do modelo base foi analisado por meio da métrica acurácia, a qual representa a razão entre as classes corretamente classificadas e o número total de amostras avaliadas. Considerando o desbalanceamento de classes do conjunto de dados utilizado para o treinamento do modelo final, o desempenho do classificador em questão foi apreciado por meio da métrica F-Medida; a qual representa a média harmônica entre os valores de precisão (amostras corretamente classificadas como positivas dentre o número total predições referentes a classe positiva) e revocação (proporção de amostras positivas corretamente classificadas) (HOSSIN; SULAIMAN, 2015).

As características extraídas pelo modelo foram revisadas com o auxílio de uma técnica de visualização referida como *Gradient-weighted Class Activation Mapping* (Grad-CAM) que utiliza os gradientes da última camada de convolução para atribuir valores de importância às conexões neurais responsáveis pela classificação. Neste contexto, os gradientes de uma classe de interesse são retropropagados pela RNC e combinados aos seus respectivos mapas de características para destacar os atributos discriminantes da classe predita, conforme ilustra a Figura 8 (SELVARAJU et al., 2017).

Figura 8 – Representação gráfica do processo de mapeamento de ativação de classes.



Fonte: Adaptado de Selvaraju et al. (2017).

Para a execução do método supracitado, selecionou-se um quadro-chave de cada sequência de capturas correspondente aos indivíduos que representam o conjunto de teste. No contexto da teoria da informação, a variabilidade *interpíxel* pode ser relacionada com a informação visual de uma imagem por meio da entropia. Esta medida estatística possibilita a estimativa da informação média contida em cada *pixel* a partir de seus respectivos níveis

de intensidade (GONZALES; WOODS, 2008a). Deste modo, considerou-se como critério de seleção o valor de entropia máximo entre os quadros de cada sequência de capturas. Para tal, foi utilizada uma matriz de coocorrência de níveis de cinza para verificar a distribuição de intensidades entre as posições relativas dos *pixels*.

Neste contexto, a vizinhança considerada corresponde a 1 posição à esquerda e à direita em um ângulo horizontal (HARALICK; SHANMUGAM; DINSTEIN, 1973; GONZALES; WOODS, 2008b). O histograma decorrente desta operação foi aplicado à equação 3.2, em que a entropia E é obtida pela soma dos produtos entre a probabilidade p do i -ésimo e j -ésimo níveis de intensidade relativos a um determinado *pixel* ocorrerem e seu logaritmo, delimitada pelos níveis de tons de cinza ($L - 1$) da imagem. Uma constante de valor arbitrário positivo ε foi incluída na equação para evitar possíveis erros decorrentes do cálculo do logaritmo de probabilidades nulas ($p_{ij} = 0$) (HARALICK; SHANMUGAM; DINSTEIN, 1973; GONZALES; WOODS, 2008b).

$$E = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} \log_2(p_{ij} + \varepsilon) \quad (3.2)$$

3.3 Materiais

Os métodos descritos neste capítulo foram implementados com a linguagem de programação *Python* e os pacotes: *Tensorflow*, *Keras*, *Scikit-learn*, *Scikit-image*, *OpenCV*, *Pillow*, *Numpy*, *Scipy*, *Matplotlib* e *TF-Explain*. Quanto aos recursos computacionais, utilizou-se um ambiente de computação remoto provido pela plataforma *Kaggle* com 13 GB de memória RAM, 20 GB de armazenamento, 1 unidade de processamento central com 2 núcleos (modelo Intel Xeon) e uma unidade de processamento gráfico com 16 GB de VRAM (modelo Nvidia Tesla P-100).

3.4 Considerações sobre o capítulo

Este capítulo apresentou os métodos que compõem a abordagem de classificação de embriaguez proposta nesta dissertação. Na seção 3.1 foram descritas as características da base de dados utilizada para o treinamento do classificador.

A seção 3.2 apresentou o protocolo experimental adotado para a elaboração do modelo, descrevendo as estratégias empregadas durante as fases do processo de treinamento. Neste contexto, discursou-se sobre os critérios considerados para o emprego da técnica de transferência de aprendizado e a seleção da arquitetura de RNC. Ademais, também foram apresentados os métodos utilizados para a seleção dos hiperparâmetros de treinamento. Sobre este aspecto, convém salientar que em adição aos processos de busca randômica e busca em grade utilizou-se a busca manual com o intuito de aprimorar os valores obtidos pelos algoritmos de otimização automática, uma vez que os comportamentos de

determinados hiperparâmetros podem ser estimados com base na influência que exercem no processo de treinamento (GOODFELLOW; BENGIO; COURVILLE, 2016c; LI et al., 2021). Para validar o conjunto de configurações obtido, empregou-se o método de otimização bayesiana (BERGSTRA; BENGIO, 2012).

Em seguida, descreveu-se as métricas de avaliação utilizadas para a análise de desempenho do modelo e o método de visualização empregado para a interpretação de sua inferência. Por fim, a seção 3.3 apresentou a linguagem de programação, as bibliotecas e os recursos computacionais utilizados para a implementação dos métodos descritos ao longo deste capítulo.

4 Resultados

Para validar a hipótese de que a embriaguez pode ser identificada a partir da análise de características faciais extraídas da imagem do rosto de um indivíduo, os modelos de classificação descritos no capítulo anterior foram avaliados quantitativamente com relação a sua capacidade de categorização e qualitativamente quanto a interpretabilidade de sua inferência. Os resultados obtidos com a aplicação dos métodos de busca de hiperparâmetros, assim como as estatísticas de desempenho do classificador, serão apresentados na seção 4.1. Por fim, as amostras obtidas com o método de seleção de quadros-chave e as características consideradas determinantes para a classificação serão especificados na seção 4.2.

4.1 Seleção dos hiperparâmetros de treinamento

Conforme citado na seção 3.2, o desempenho de um modelo baseado em aprendizado de máquina está associado às especificações de um conjunto de propriedades referido como hiperparâmetros. A Tabela 3 apresenta os resultados obtidos após a aplicação do método de busca randômica na fase de transferência de aprendizado, na qual foram investigados três modelos pré-treinados baseados na arquitetura VGG: VGG16 (SIMONYAN; ZISSERMAN, 2015), VGG19 (SIMONYAN; ZISSERMAN, 2015) e VGGFace (PARKHI; VEDALDI; ZISSERMAN, 2015). Para não comprometer a leitura deste capítulo, a tabela supracitada apresentará apenas os melhores resultados de cada experimento, os *logs* completos podem ser conferidos neste repositório¹.

Conforme pode ser observado na Tabela 3, os experimentos preliminares não apresentaram resultados promissores. Para aperfeiçoá-los, analisou-se todos as possíveis configurações de hiperparâmetros obtidas e partir destas adotou-se o método de busca manual². Para tal, foram considerados os valores comuns às propriedades que proporcionaram os melhores resultados tais como: k igual a 5 ou 7; taxa de aprendizagem entre 1e-5 e 1e-6; tamanho do *mini-batch* e número de *epochs* iguais ou superiores a 200. Como critério de seleção, considerou-se o conjunto de hiperparâmetros que proporcionasse a maior acurácia média entre os modelos avaliados em cada partição do processo de validação cruzada e o menor desvio padrão entre acurácias individuais.

Neste contexto, a partir da experimentação empírica e da observação do comportamento dos modelos quando aplicados ao conjunto de teste, verificou-se que a rede VGG16 com os seguintes hiperparâmetros propiciou resultados mais promissores: k igual 5, taxa

¹ <<https://bit.ly/3dfNecZ>>

² Em conformidade com Bergstra e Bengio (2012)

Tabela 3 – Resultados do processo de busca randômica na fase de transferência de aprendizado.

Modelo pré-treinado	K*	Learning rate	Batch size	Epochs	Acurácia média	Desvio padrão
VGG16	5	1e-06	50	200	0.67	0.11
VGG16	7	1e-05	300	200	0.69	0.1
VGG16	10	1e-04	600	40	0.66	0.19
VGG19	5	1e-05	900	300	0.7	0.13
VGG19	7	1e-04	600	60	0.64	0.06
VGG19	10	1e-04	600	60	0.63	0.19
VGGFace	5	1e-06	10	300	0.66	0.17
VGGFace	7	1e-04	400	10	0.63	0.1
VGGFace	10	1e-03	500	20	0.63	0.19

Fonte: Elaborado pelo autor.

Nota: *Partições utilizadas no processo de validação cruzada (estratificada).

de aprendizagem igual a 1e-6, tamanho do *mini-batch* igual a 300 e número de *epochs* igual a 205. A Tabela 4 apresenta o desempenho dos modelos analisados.

Tabela 4 – Resultados do processo de busca manual na fase de transferência de aprendizado.

Modelo pré-treinado	K*	Learning rate	Batch size	Epochs	Acurácia média	Desvio padrão
VGG16	5	1e-06	300	205	0.86	0.1
VGG19	5	1e-05	300	240	0.79	0.05
VGGFace	5	1e-06	300	210	0.74	0.05

Fonte: Elaborado pelo autor.

Nota: *Partições utilizadas no processo de validação cruzada (estratificada).

Deste modo, para a execução da fase de *fine-tuning* do modelo base, selecionou-se o classificador que apresentou o desempenho mais equilibrado entre os conjuntos de treinamento, validação e teste dentre aqueles avaliados em cada partição do método de validação cruzada. A Tabela 5 descreve tais dados e destaca o modelo selecionado, a Tabela 6 demonstra a matriz de confusão resultante de sua aplicação ao conjunto de teste e por fim, a Figura 9 ilustra a curva de aprendizado do modelo base durante todo o processo de validação cruzada.

Para induzir o modelo base a abstrair características de alto nível pertinentes a classe do problema investigado, empregou-se o método de *fine-tuning*. A efetividade deste procedimento foi avaliada após a observação do comportamento do modelo ao retreinar os

Tabela 5 – Seleção do modelo base para a etapa de *fine-tuning*.

Partição	Acurácia no conjunto de treinamento	Acurácia no conjunto de validação	Acurácia no conjunto de teste
1	0.7126	0.4629	0.6679
2	0.6789	0.7643	0.9359
3	0.7389	0.7900	0.8960
4	0.7889	0.8271	0.9240
5	0.8300	0.8286	0.9

Fonte: Elaborado pelo autor.

Tabela 6 – Matriz de confusão do modelo base (etapa de transferência de aprendizado).

	Embriagado (condição real)	Sóbrio (condição real)
Embriagado (condição predita)	250	50
Sóbrio (condição predita)	0	200

Fonte: Elaborado pelo autor.

Figura 9 – Curva de aprendizado do modelo base.



Fonte: Elaborado pelo autor.

três últimos blocos de convolução. Conforme observado na Tabela 7, não houve melhorias de desempenho ao reajustar os pesos das camadas convolucionais superiores ao bloco 4, portanto, apenas os últimos dois blocos de convolução foram atualizados. Convém salientar que os valores pertinentes ao número de *epochs* e ao tamanho do *mini-batch* foram mantidos em relação a etapa anterior, contudo, foi necessário alterar a taxa de aprendizagem ($1e-8$)

para se obter os resultados apresentados na Tabela 7.

A seleção do modelo utilizado para a elaboração do classificador final foi realizada de maneira semelhante ao método descrito na etapa de transferência de aprendizado e será apresentada na Tabela 8. A Tabela 9 e a Figura 10 demonstram a matriz de confusão e a curva de aprendizado do modelo base obtidos após o processo de *fine-tuning*, respectivamente.

Tabela 7 – Resultados obtidos após a atualização dos pesos das camadas de convolução mais profundas.

Blocos de convolução retreinados	Acurácia média (conjunto de teste)	Desvio padrão
Bloco 5	0.9	0.004
Blocos 5 e 4	0.94	0.04
Blocos 5, 4 e 3	0.94	0.03

Fonte: Elaborado pelo autor.

Tabela 8 – Seleção do modelo base para a etapa final de treinamento.

Partição	Acurácia no conjunto de treinamento	Acurácia no conjunto de validação	Acurácia no conjunto de teste
1	0.9043	0.6829	0.8999
2	0.8989	0.9571	0.9259
3	0.9450	0.8986	0.9079
4	0.9368	1.0000	0.9919
5	0.9546	1.0000	0.9819

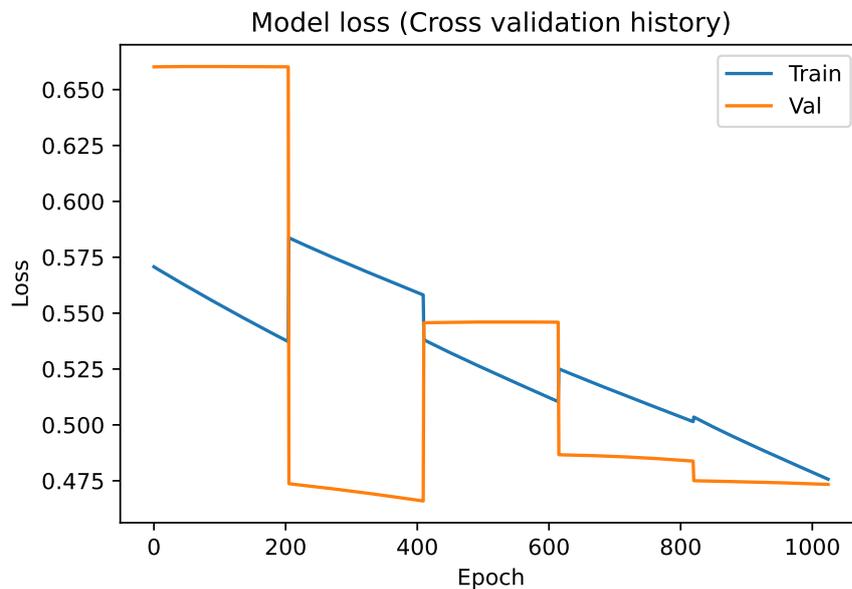
Fonte: Elaborado pelo autor.

Tabela 9 – Matriz de confusão do modelo base após a etapa de *fine-tuning*.

	Embriagado (condição real)	Sóbrio (condição real)
Embriagado (condição predita)	246	5
Sóbrio (condição predita)	4	245

Fonte: Elaborado pelo autor.

Uma vez identificada a capacidade de categorização do modelo, este foi utilizado como base para a elaboração do classificador final em um processo semelhante ao descrito na etapa de etapa de transferência de aprendizado. Nesta fase do processo de treinamento, o método de busca randômica retornou as mesmas composições de hiperparâmetros para k igual a 5 e 7, respectivamente. Contudo, a versão avaliada com 7 subdivisões apresentou um desempenho superior. Não foi possível realizar este experimento com 10 partições de

Figura 10 – Curva de aprendizado do modelo base após o processo de *fine-tuning*.

Fonte: Elaborado pelo autor.

validação cruzada pois o tempo de execução superou o limite estabelecido pela plataforma *Kaggle* e desta feita, foi cancelado antes de sua conclusão. Para não prejudicar a leitura deste capítulo, a Tabela 10 apresenta 5 dos resultados mais promissores obtidos com o processo de busca randômica. Os *logs* completos podem ser conferidos [neste repositório](#)³.

Tabela 10 – Resultados do processo de busca randômica na fase de treinamento do modelo final (7 partições de validação cruzada).

<i>Learning rate</i>	<i>Batch size</i>	<i>Epochs</i>	Acurácia média	Desvio padrão
1e-06	10	300	0.96	0.005
1e-06	50	200	0.91	0.018
1e-03	500	20	1.0	0.0
1e-05	400	300	0.96	0.007
1e-04	600	300	1.0	0.0

Fonte: Elaborado pelo autor.

Em contraste aos resultados observados na Tabela 3, a aplicação do método de busca randômica na etapa final do processo de treinamento propiciou diversas composições de hiperparâmetros promissoras, dentre as quais 5 foram apresentadas na Tabela 10. Deste modo, as configurações que proporcionaram uma acurácia média igual ou superior a 90% no conjunto de validação foram selecionadas e atribuídas ao modelo para avaliação no conjunto de teste. Para não comprometer a leitura deste capítulo, a Tabela 11 apresentará

³ <<https://bit.ly/3dfNecZ>>

os resultados obtidos para os conjuntos de hiperparâmetros descritos na Tabela 10. Contudo, convém salientar que tais instâncias ilustram o comportamento geral observado ao longo dos experimentos.

Tabela 11 – Avaliação dos hiperparâmetros obtidos no processo de busca randômica (conjunto de teste).

<i>Learning rate</i>	<i>Batch size</i>	<i>Epochs</i>	Acurácia média	Desvio padrão
1e-06	10	300	0.71	0.008
1e-06	50	200	0.70	0.037
1e-03	500	20	0.72	0.008
1e-05	400	300	0.71	0.029
1e-04	600	300	0.76	0.006

Fonte: Elaborado pelo autor.

Conforme observado na Tabela 11, os conjuntos de hiperparâmetros obtidos com o método de busca randômica não demonstraram os resultados promissores obtidos no conjunto de validação quando aplicados ao conjunto de teste. Deste modo, assim como na etapa de transferência de aprendizado, empregou-se o método de busca manual considerando os valores comuns às propriedades que proporcionaram os melhores desempenhos de classificação para a identificação da especificação ótima de cada hiperparâmetro. Em ato contínuo, para aperfeiçoar a capacidade representacional do modelo, aplicou-se o método de busca em grade para determinar a profundidade da camada totalmente conectada. Neste contexto, obteve-se a seguinte composição de hiperparâmetros: taxa de aprendizagem: 1e-06; tamanho do *mini-batch*: 500; *epochs*: 205. A Tabela 12 apresenta a seleção do número de unidades da camada oculta do classificador.

Considerando as especificações obtidas após os processos de busca manual e em grade, empregou-se o método de otimização bayesiana para verificar se os hiperparâmetros identificados de fato proporcionam o melhor desempenho de classificação. Conforme pode ser observado na Tabela 13, tal procedimento resultou em configurações que não promovem melhorias de performance. Deste modo, as configurações previamente citadas foram utilizadas para o treinamento do modelo final. A Tabela 14 e a Figura 11 apresentam o comportamento do classificador e sua curva de aprendizado ao longo do processo de validação cruzada, respectivamente. A Tabela 15 demonstra a matriz de confusão do modelo selecionado para a última etapa de *fine-tuning* (destacado na Tabela 14).

Para a execução do processo de *fine-tuning* do modelo final considerou-se a mesma estratégia empregada na etapa de transferência de aprendizado, em que se manteve a composição de hiperparâmetros utilizada na fase de treinamento com exceção da taxa de aprendizagem (a qual foi reduzida para 1e-08) e atualizou-se os pesos dos dois últimos blocos de camadas convolucionais. A Tabela 16 apresenta o comportamento do modelo

Tabela 12 – Seleção do número de unidades ocultas para a expansão da topologia do modelo.

Número de unidades ocultas	Número de camadas ocultas	Acurácia média no conjunto de teste	Desvio padrão
2	1	0.71	0.022
4	1	0.79	0.043
8	1	0.79	0.030
16	1	0.81	0.014
32	1	0.76	0.035
64	1	0.73	0.026
128	1	0.75	0.027
256	1	0.75	0.027
512	1	0.76	0.022

Fonte: Elaborado pelo autor.

ao longo do processo de validação cruzada. Considerando o desbalanceamento de classes do conjunto de dados utilizado para o treinamento do classificador, a avaliação final do desempenho de classificação foi baseada na métrica F-Medida conforme descrito na seção 3.2.

Tabela 13 – Resultados do processo de otimização bayesiana para busca de hiperparâmetros.

Iterações	Acurácia média (conjunto de teste)	Desvio padrão	<i>Learning rate</i>	Camadas ocultas	Unidades ocultas
16	0.76	0.04	2.6e-07	5	201
32	0.81	0.0	5.7e-03	6	353
48	0.79	0.03	8e-08	7	512

Fonte: Elaborado pelo autor.

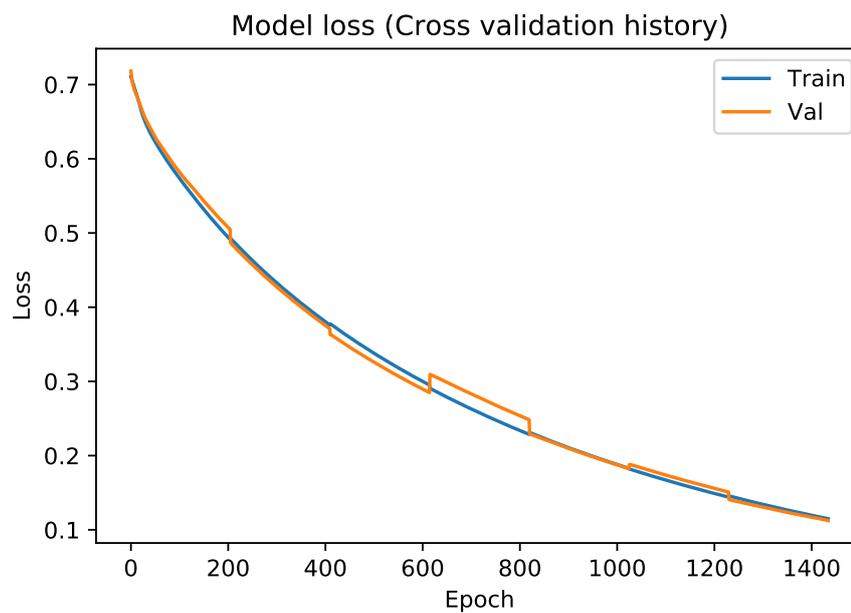
A curva de aprendizado obtida após a última etapa de *fine-tuning* é ilustrada na Figura 12, a Tabela 17 demonstra a matriz de confusão do modelo destacado na Tabela 16. Por fim, a Tabela 18 sumariza os hiperparâmetros utilizados para o treinamento dos classificadores descritos ao longo desta seção.

Tabela 14 – Seleção do modelo final para a etapa de *fine-tuning*.

Partição	Acurácia no conjunto de treinamento	Acurácia no conjunto de validação	Acurácia no conjunto de teste
1	0.9025	0.8970	0.8010
2	0.9320	0.9420	0.7890
3	0.9433	0.9560	0.8010
4	0.9502	0.9390	0.8140
5	0.9645	0.9620	0.8190
6	0.9808	0.9780	0.8289
7	0.9955	0.9950	0.8280

Fonte: Elaborado pelo autor.

Figura 11 – Curva de aprendizado do modelo final.



Fonte: Elaborado pelo autor.

Tabela 15 – Matriz de confusão do modelo final.

	Embriagado (condição real)	Sóbrio (condição real)
Embriagado (condição predita)	681	102
Sóbrio (condição predita)	69	148

Fonte: Elaborado pelo autor.

Tabela 16 – Avaliação do modelo final após a etapa de *fine-tuning*.

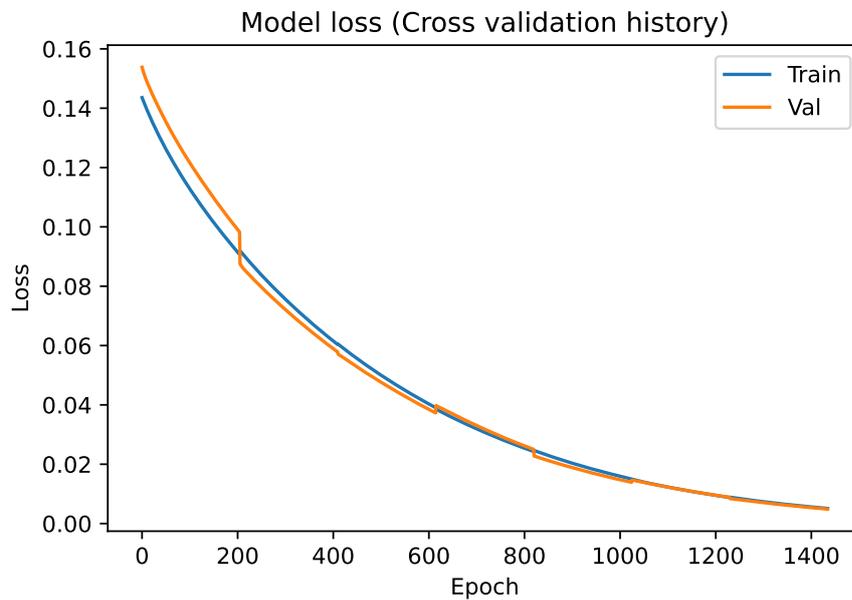
Partição	Acurácia no conjunto de treinamento	Acurácia no conjunto de validação	Acurácia no conjunto de teste
1	1.0	1.0	0.8420
2	1.0	1.0	0.8350
3	1.0	1.0	0.8360
4	1.0	1.0	0.8570
5	1.0	1.0	0.8840
6	1.0	1.0	0.8830
7	1.0	1.0	0.8750

Desvio padrão: 0.02

F-Medida média (conjunto de teste): 0.91

F-Medida individual (modelo destacado): 0.92

Fonte: Elaborado pelo autor.

Figura 12 – Curva de aprendizado do modelo final após o processo de *fine-tuning*

Fonte: Elaborado pelo autor.

Tabela 17 – Matriz de confusão do modelo final após o processo de *fine-tuning*.

	Embragado (condição real)	Sóbrio (condição real)
Embragado (condição predita)	696	62
Sóbrio (condição predita)	54	188

Fonte: Elaborado pelo autor.

Tabela 18 – Especificação dos hiperparâmetros do modelo de classificação de embriaguez.

Parâmetro de treinamento	Modelo base	Modelo final
Camadas ocultas	–	1
Unidades ocultas	–	16
Função de ativação das camadas ocultas	–	<i>Relu</i>
Função de ativação da camada de saída	<i>Sigmoid</i>	<i>Sigmoid</i>
Função de otimização	<i>Adam</i>	<i>Adam</i>
<i>Learning rate</i> (treinamento)	1e-06	1e-06
<i>Learning rate</i> (<i>fine-tuning</i>)	1e-08	1e-08
Função de custo	<i>Binary Cross Entropy</i>	<i>Binary Cross Entropy</i>
<i>Batch size</i> (treinamento)	300	500
<i>Batch size</i> (<i>fine-tuning</i>)	300	500
<i>Epochs</i> (treinamento)	205	205
<i>Epochs</i> (<i>fine-tuning</i>)	205	205

Fonte: Elaborado pelo autor.

4.2 Revisão das características extraídas pela RNC

Para que fosse possível visualizar as características extraídas pelo modelo e verificar sua associação com as respostas fisiológicas induzidas pelo metabolismo do álcool, foi necessário selecionar um quadro-chave que resumisse a informação contida nos quadros de cada sequência de capturas. Para tal, empregou-se um método de seleção baseado no critério de máxima entropia (descrito na seção 3.2). A Tabela 19 apresenta as amostras do conjunto de teste que foram selecionadas de acordo com este processo.

Tabela 19 – Quadros-chave selecionados de acordo com o critério de máxima entropia.

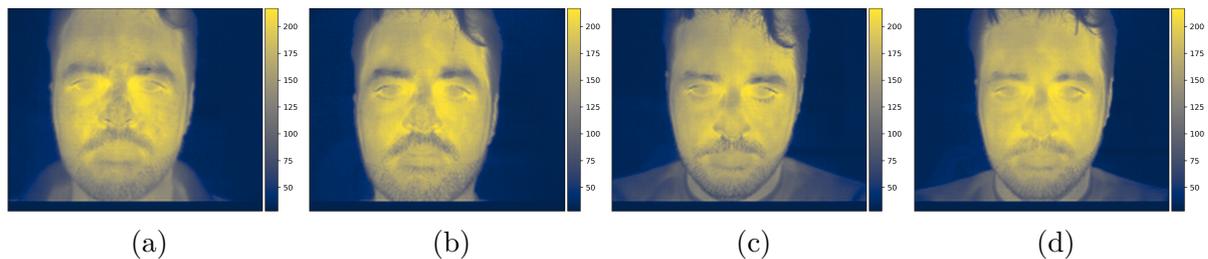
Participante	1° período de aquisição	2° período de aquisição	3° período de aquisição	4° período de aquisição
37	34° quadro	20° quadro	26° quadro	33° quadro
38	38° quadro	3° quadro	43° quadro	37° quadro
39	11° quadro	7° quadro	36° quadro	50° quadro
40	1° quadro	4° quadro	44° quadro	16° quadro
41	28° quadro	22° quadro	26° quadro	2° quadro

Fonte: Elaborado pelo autor.

Em seguida, empregou-se o método Grad-CAM para a concepção das representações visuais que indicam a relevância das características extraídas pelo modelo (destacado na Tabela 16) para a classificação do estado de embriaguez. No mapa de ativação de classes

ilustrado pela Figura 13 é possível observar que o classificador considerou dois comportamentos como indicativos de intoxicação alcoólica: as diferenças entre as intensidades dos *pixels* nos lados esquerdo e direito do rosto e nas regiões do nariz e da testa.

Figura 13 – Mapa de ativação de classes. Em (a) é ilustrada a assinatura térmica facial de um indivíduo em estado sóbrio. Em (b), (c) e (d) são apresentadas as características relacionadas à embriaguez consideradas determinantes pelo modelo em cada período de aquisição subsequente à ingestão de álcool. A relevância de cada atributo extraído pela RNC pode ser interpretada pela intensidade dos *pixels* destacados em amarelo.



Fonte: Elaborado pelo autor.

Ao analisar as características destacadas na Figura 13, nota-se que a RNC foi capaz de identificar variações nos valores dos *pixels* pertinentes às áreas de interesse descritas na seção 1.2. Em adição ao reconhecimento de tais características, foi observado que o modelo também abstraiu as mudanças de comportamento destas representações ao longo do período de amostragem.

4.3 Considerações sobre o capítulo

Neste capítulo foram apresentados os resultados obtidos com a aplicação dos métodos que compõem a abordagem proposta para classificação de embriaguez. Ao longo da seção 4.1 detalhou-se o processo de treinamento do classificador e a evolução de seu desempenho durante as etapas de transferência de aprendizado e composição do modelo final, considerando também os métodos de busca de hiperparâmetros utilizados em cada fase. Neste contexto, foi demonstrado que o método de classificação de embriaguez proposto nesta dissertação foi capaz de discriminar as assinaturas térmicas faciais correspondentes ao estado de embriaguez em 92% dos casos, identificando corretamente 696 dos 750 quadros do conjunto de teste que representam a classe em questão.

Em comparação ao estudo de [Obi-Alago, Yanushkevich e Wetherley \(2019\)](#), que propõe a fusão de RNCs aplicadas às regiões da testa e olhos para identificar os sinais de intoxicação alcoólica, o classificador proposto no presente trabalho apresentou um desempenho aproximadamente 5% inferior. Contudo, convém salientar que o mesmo possui um fator de explicabilidade superior. Conforme ilustrado na seção 4.2, o modelo descrito

ao longo deste capítulo foi capaz de identificar variações de comportamento em *pixels* de regiões do rosto que são reconhecidas na literatura por apresentar um aumento de temperatura de acordo com a concentração de álcool no sangue.

5 Discussão

Nesta dissertação, a classificação de embriaguez foi abordada sob a perspectiva do comportamento das propriedades intrínsecas dos dados e de sua relação com as respostas fisiológicas induzidas pelo álcool. Para tal, foi utilizado um conjunto de imagens térmicas faciais obtidas em quatro períodos distintos: antes da ingestão de bebida alcoólica e em três intervalos subsequentes de 30 minutos após o consumo, de modo que em cada etapa foram obtidos 50 quadros de cada participante com um intervalo de amostragem de 100 milissegundos (KOUKIOU; ANASTASSOPOULOS, 2015).

Neste contexto, considerou-se que as possíveis variações de temperatura entre cada captura podem estar relacionadas às flutuações térmicas registradas pelo sensor termográfico em decorrência do calor emitido pelos componentes eletrônicos do próprio dispositivo ou pela influência de variáveis ambientais (RIOU; BERREBI; BREMOND, 2004; USAMENTIAGA et al., 2014; TATTERSALL, 2016). Assim, considerando a possibilidade deste fenômeno ocorrer em cenários reais, optou-se por treinar o modelo de classificação com todos os quadros disponíveis para que o classificador fosse capaz de identificar os sinais que caracterizam as respostas fisiológicas induzidas pelo álcool a despeito do ruído registrado pelo sensor.

Em vista disso, adotou-se uma abordagem baseada em RNCs uma vez que tal método possibilita a extração automática das principais características de uma imagem, é invariante à pose e possibilita a interpretação das representações consideradas determinantes pelo classificador para a inferência da classe de uma amostra (LECUN et al., 1998; LECUN; BENGIO; HINTON, 2015; SELVARAJU et al., 2017). Considerando que a eficácia das RNCs depende do volume de dados disponível para o seu treinamento, optou-se por empregar a estratégia de *transfer learning* utilizando um modelo pré-treinado em uma base de dados de larga escala. Neste contexto, a rede VGG16 (SIMONYAN; ZISSERMAN, 2015) foi selecionada como base para a elaboração do classificador proposto nesta dissertação.

Deste modo, o processo de treinamento foi dividido em duas etapas: transferência de aprendizado e elaboração do modelo final. Durante a primeira fase foi observado que o subconjunto de dados utilizado na primeira partição do método de validação cruzada (Tabela 5) implicou em um desempenho consideravelmente inferior aos demais. Acredita-se que a ocorrência de tal fenômeno está relacionada ao fato de que uma parte das imagens que compõem este subconjunto está desfocada, o que compromete a qualidade das características extraídas pela rede. Contudo, a despeito deste fenômeno, o modelo base apresentou resultados promissores ao final da etapa de *fine-tuning* (Tabela 9), superando inclusive o desempenho de classificação do modelo final (Tabela 17).

Neste contexto, acredita-se que o *drift error* possa ter influenciado o resultado final, uma vez que as assinaturas térmicas das imagens pertinentes ao 4º período de aquisição apresentam uma mudança de comportamento distintivamente superior aos demais em relação as amostras correspondentes ao estado de sobriedade. Não obstante, conforme ilustrado pela Figura 13, o classificador reconheceu as variações de temperatura entre as regiões da testa e do nariz como um indicativo de embriaguez de forma análoga ao trabalho de Koukiou e Anastassopoulos (2012).

Ademais, a presente pesquisa também contribui com a literatura ao identificar uma característica do estado de embriaguez não evidenciada por estudos anteriores: a assimetria térmica entre os lados direito e esquerdo do rosto. Este resultado evidencia que o objetivo desta dissertação foi alcançado e confirma a premissa estabelecida na seção 1.2 de que a partir de uma imagem do rosto de um indivíduo é possível identificar padrões de distribuição de temperatura que não seriam reconhecidos ao se considerar regiões faciais específicas isoladamente; conforme abordado em estudos anteriores (KOUKIOU; ANASTASSOPOULOS, 2015; KOUKIOU; ANASTASSOPOULOS, 2016a; KOUKIOU; ANASTASSOPOULOS, 2016b; KOUKIOU; ANASTASSOPOULOS, 2018a; HERMO-SILLA et al., 2018; OBI-ALAGO; YANUSHKEVICH; WETHERLEY, 2019; KOUKIOU, 2021).

Convém salientar que ainda que o modelo proposto tenha identificado variações na assinatura térmica da região periocular, o mesmo não reconheceu as alterações de temperatura entre a íris e a esclera descritas por Koukiou e Anastassopoulos (2016a). Acredita-se que este resultado esteja relacionado à configuração dos filtros de convolução da rede VGG16, implicando em mapas de características incapazes de representar o comportamento dos *pixels* selecionados manualmente por Koukiou e Anastassopoulos (2016a). Assim, uma possível linha de investigação para trabalhos futuros seria a atualização da arquitetura das camadas de convolução para possibilitar o encapsulamento do contraste de temperatura entre a íris e a esclera.

Considerando que os indicativos de alcoolemia identificados pelo modelo proposto são baseados na difusão de calor entre os vasos sanguíneos e o tecido ao seu redor, convém salientar que mesmo que o comportamento da distribuição de temperatura seja semelhante, os padrões de assinatura térmica podem variar entre os indivíduos devido a espessura do tecido facial e a influencia da pigmentação da pele em sua condutância térmica (tonalidades mais escuras tendem a ser menos condutivas) (BUDDHARAJU et al., 2007).

Ademais, a representação dos sinais de embriaguez observados na Figura 13 também pode variar de acordo com a intensidade e a periodicidade das respostas fisiológicas induzidas pelo álcool, que por sua vez podem ser influenciadas pela tolerância de um indivíduo ao álcool e ao nível de alcoolemia alcançado após o consumo de bebidas alcoólicas. As variações destas condições dependem de diversos fatores, tais como: sexo, peso, idade,

o tipo da bebida alcoólica, a dose consumida, a presença de comida no estômago antes ou durante a ingestão do álcool, os hábitos de consumo do indivíduo, o período do dia em que a bebida alcoólica é consumida, a temperatura ambiente e o nível de estresse associado à execução de uma tarefa (como a condução de um veículo) (KAWANO, 2010; CEDERBAUM, 2012; PAVLIDIS et al., 2016; CARDONE; MERLA, 2017; JONES, 2019a; LOHANI; PAYNE; STRAYER, 2019). A influência e os efeitos destes fatores serão abordados nas próximas seções.

5.1 Influência do sexo nas respostas fisiológicas induzidas pelo álcool

Conforme descrito na seção 3.1, durante o procedimento de aquisição da base de dados *Sober-Drunk Database* a concentração de álcool por litro de ar exalado variou entre 0.22 e 0.37 mg/L para os homens e de 0.49 a 0.89 mg/L para as mulheres (KOUKIOU; ANASTASSOPOULOS, 2015). Essas leituras sugerem que de acordo com o sexo, as pessoas podem ser afetadas de diferentes maneiras consumindo o mesmo volume de álcool. Este comportamento tende a ocorrer pois a distribuição do álcool pelo organismo varia conforme a proporção entre gordura e água no corpo humano. Considerando que o percentual de gordura corporal é superior em mulheres, o volume de etanol distribuído aos tecidos pela corrente sanguínea é menor quando comparado aos homens e desta feita, ocasionando o elevado pico de alcoolemia (BARAONA et al., 2001; BRICK, 2006; CEDERBAUM, 2012).

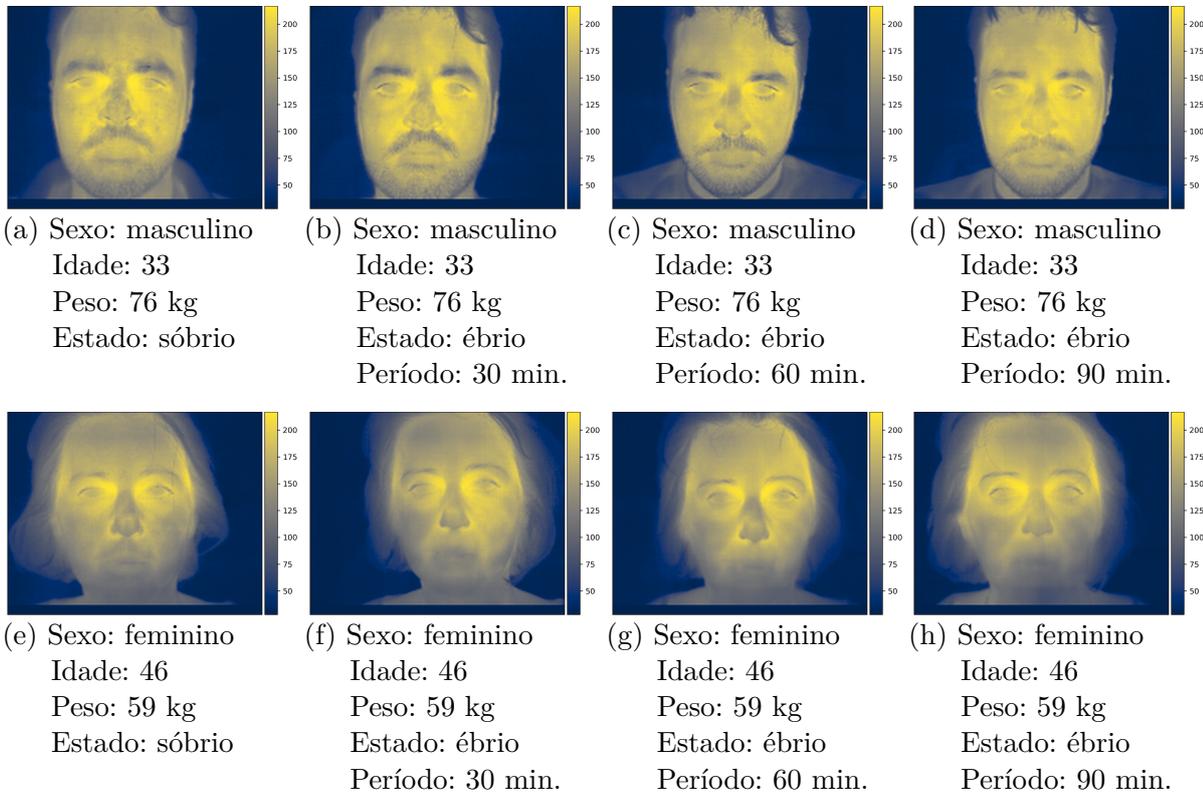
As diferenças físicas e fisiológicas entre homens e mulheres também podem levar a representações distintas de embriaguez, uma vez que as mulheres metabolizam o etanol a uma taxa mais elevada que os homens. Esta característica do metabolismo feminino implica na produção de acetaldeído em um período inferior, causando a manifestação precoce de sintomas de intoxicação (BARAONA et al., 2001). A Figura 14 ilustra o mapeamento de ativação de classes para as amostras de dois participantes de sexos distintos, na qual é possível observar o comportamento mencionado.

Na próxima seção será discutido como o peso e a idade também podem contribuir para a variação destas representações.

5.2 Influência do peso e da idade nas respostas fisiológicas induzidas pelo álcool

Na seção anterior foi mencionado que homens e mulheres podem apresentar sinais de intoxicação alcoólica com intensidades distintas, uma vez que as taxas de distribuição e eliminação do etanol variam de acordo com suas características físicas e fisiológicas.

Figura 14 – Indicativos de embriaguez apresentados por homens e mulheres. Nesta Figura é possível observar que a distribuição da temperatura facial induzida pelo álcool pode variar em extensão e intensidade ao longo do tempo para indivíduos do sexo oposto.

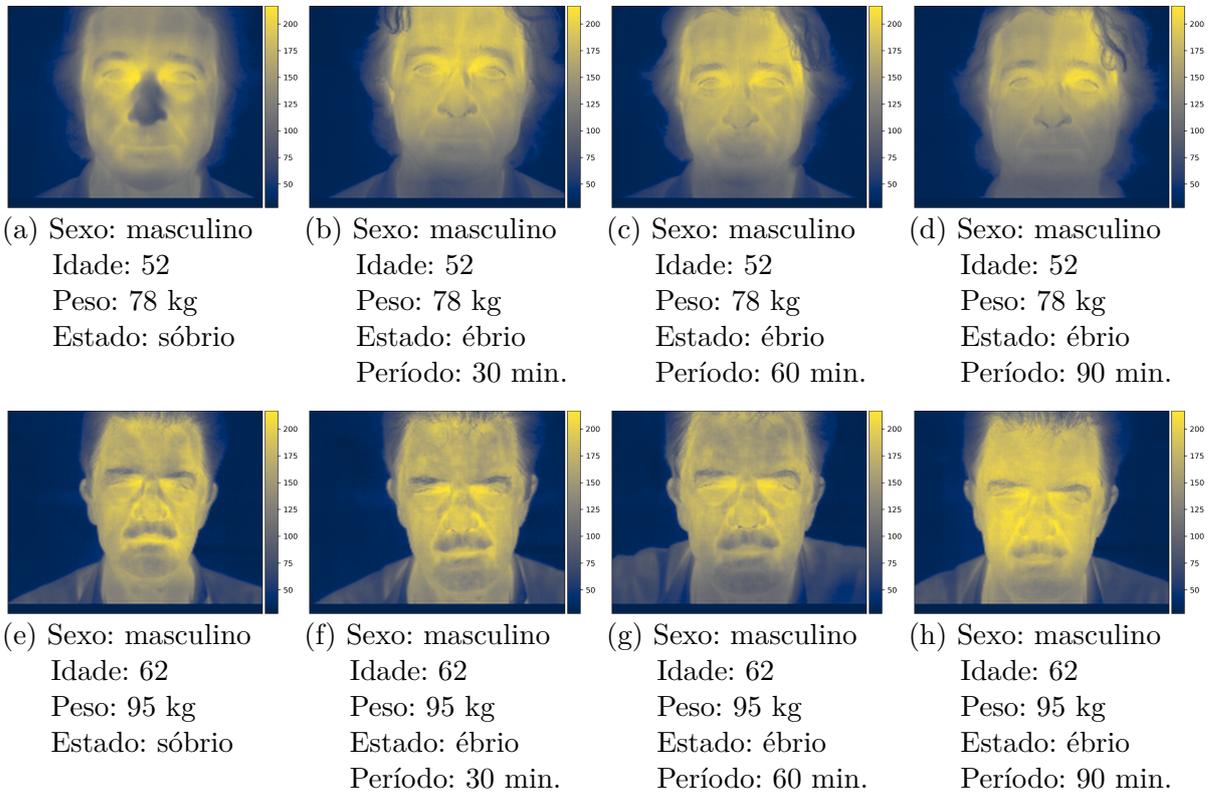


Não obstante, o peso e a idade são fatores que também podem contribuir para tal comportamento. Isto ocorre pois a medida em que um indivíduo envelhece seu índice de Água Corporal Total (ACT) diminui e desta feita, o volume de etanol distribuído aos tecidos e fluidos extravasculares é reduzido (NORBERG *et al.*, 2003).

Ademais, a proporção entre massa muscular e gordura corporal também pode influenciar significativamente o processo de distribuição do etanol pois a concentração de álcool em um tecido depende de seu conteúdo relativo de água. Deste modo, considerando que o envelhecimento e o sobrepeso reduzem a massa muscular e aumentam o tecido adiposo, indivíduos mais velhos e com um alto índice de gordura corporal tendem a apresentar concentrações elevadas de álcool no sangue por apresentarem menos ACT por kg de peso corporal para diluir o etanol (NORBERG *et al.*, 2003; BRICK, 2006; CEDERBAUM, 2012; JONES, 2019a), portanto, é esperado que indivíduos com características físicas distintas apresentem sinais de embriaguez com intensidade e periodicidade diferentes, mesmo consumindo a mesma dose de álcool. A Figura 15 ilustra este fenômeno.

Também foi observado que o peso e a idade individualmente podem não ser os únicos fatores que afetam a representação dos sinais de embriaguez. Conforme ilustrado pela Figura 16, pessoas com o mesmo sexo, peso e idade podem apresentar padrões de

Figura 15 – Indicativos de embriaguez apresentados por indivíduos com peso e idade distintos. Em (b) e (c) é possível observar que a temperatura nas regiões do nariz e testa aumenta e se propaga para a região superior esquerda do rosto, enquanto em (f) e (g) a temperatura da testa e do lado direito do rosto diminui.



assinatura térmica facial distintos.

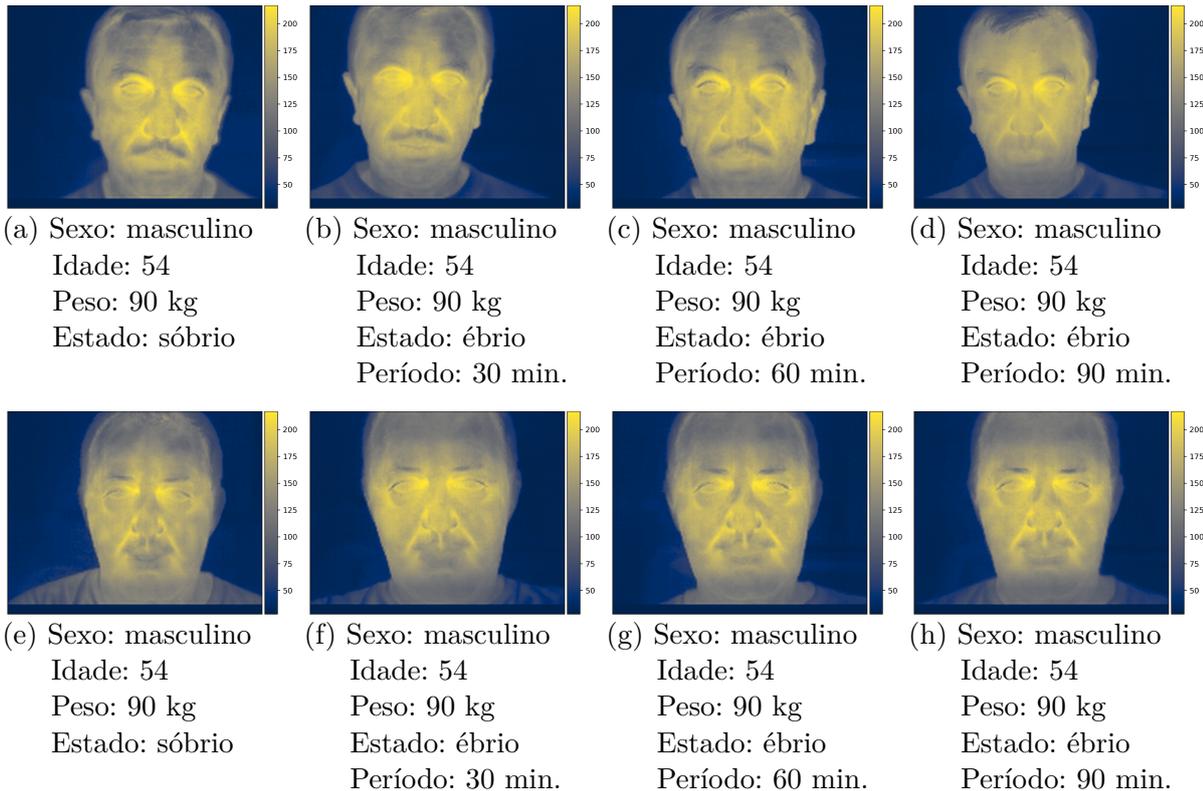
Na próxima seção será discutido como as características genéticas de um indivíduo podem influenciar a concentração de álcool no sangue e a intensidade dos sinais de intoxicação alcoólica.

5.3 Influência genética nas respostas fisiológicas induzidas pelo álcool

Em adição aos fatores previamente citados, convém salientar que diferentes grupos étnicos podem apresentar os indicativos de embriaguez identificados neste estudo de maneiras distintas devido a diversidade de suas características genéticas. Durante o processo metabólico do álcool, a enzima álcool desidrogenase (ADH) converte o etanol das bebidas alcoólicas em acetaldeído e a enzima aldeído desidrogenase (ALDH) o transforma em acetato. Um polimorfismo da enzima ALDH2 (ALDH2*2) compromete a oxidação do acetaldeído aumentando sua concentração na corrente sanguínea, assim, indivíduos com esta característica genética consumindo até mesmo um volume reduzido de álcool,

tendem a apresentar respostas fisiológicas altamente aversivas como rubor facial, náuseas, taquicardia e hipotensão, por exemplo (ERIKSSON, 2001; ZAKHARI, 2006; EDENBERG, 2007; KAWANO, 2010; CEDERBAUM, 2012).

Figura 16 – Indicativos de embriaguez apresentados por indivíduos com o mesmo sexo, peso e idade. Nos três primeiros períodos de aquisição é possível observar padrões semelhantes de assinatura térmica, contudo, nota-se que em (d) a temperatura da testa e do lado esquerdo do rosto aumenta enquanto em (h) esta característica permanece semelhante ao período anterior.



Como consequência, estes sintomas podem alterar os padrões de distribuição de temperatura facial reconhecidos pelo classificador proposto. Alguns estudos indicam que a variação genética $ALDH2^{*2}$ é mais comum em populações oriundas do Leste Asiático e em seus descendentes que em outros grupos étnicos (GOEDDE et al., 1992; EDENBERG, 2007; ENG; LUCZAK; WALL, 2007; ZHONG et al., 2018).

Na próxima seção será discutido como o tipo e o volume de bebida alcoólica consumida podem afetar o índice de alcoolemia de um indivíduo.

5.4 Influência do tipo de bebida alcoólica nas respostas fisiológicas induzidas pelo álcool

De acordo com Cederbaum (2012), o tipo de bebida alcoólica é um fator que não influencia significativamente o índice de alcoolemia quando o volume de álcool consumido é

o mesmo em cada cenário. Deste modo, se dois indivíduos consumirem 4 doses de cerveja e 4 doses de vinho, respectivamente, as respostas fisiológicas de seus organismos tendem a se manifestar em diferentes níveis de intensidade. Isto ocorre pois o vinho e a cerveja possuem teores alcoólicos distintos e desta feita, ainda que a dosagem das bebidas consumidas tenha sido a mesma, o volume de etanol metabolizado por cada indivíduo neste contexto será desproporcional.

Considerando este princípio, os participantes que contribuíram para a elaboração da base de dados *Sober-Drunk Database* (KOUKIOU; ANASTASSOPOULOS, 2015) teriam que consumir cerca de 1,56 L de cerveja (com teor alcoólico de 5%) para que os sinais de intoxicação identificados após o consumo de 480 ml de vinho (com teor alcoólico de 13%) fossem apresentados.

Na próxima seção será discutido como o volume do conteúdo estomacal pode influenciar o volume de álcool absorvido pelo organismo.

5.5 Influência do esvaziamento gástrico nas respostas fisiológicas induzidas pelo álcool

A ingestão de bebidas alcoólicas durante ou após uma refeição também pode influenciar o índice de alcoolemia, uma vez que parte do volume de álcool ingerido é metabolizado no estômago em um processo referido como metabolismo de primeira passagem. A presença de conteúdo estomacal não apenas reduz a absorção do etanol mas também acelera sua eliminação, implicando em uma concentração de álcool no sangue menor e que leva mais tempo para alcançar seu momento de pico. Assim, indivíduos que consomem bebidas alcoólicas durante ou após uma refeição tendem a apresentar respostas fisiológicas com intensidade e periodicidade distintas quando comparados a pessoas que consumiram a mesma dose de álcool em jejum (JONES, 2010; CEDERBAUM, 2012; JONES, 2019a).

Na próxima seção será discutido como o período do dia em que a bebida alcoólica é consumida pode afetar os efeitos fisiológicos decorrentes.

5.6 Influência do ciclo circadiano nas respostas fisiológicas induzidas pelo álcool

Em relação ao período do dia em que as bebidas alcoólicas são consumidas, Kawano (2010) evidencia que comportamentos vasculares distintos podem ocorrer de acordo com a interação do álcool e as fases dos ciclo circadiano. Diversos estudos indicam que o metabolismo do etanol induz respostas fisiológicas vasoconstritivas e vasodilativas nos

períodos noturno e diurno, respectivamente, implicando efeitos de hipertermia e hipotermia durante essas fases do dia (DANEL; LIBERSA; TOUITOU, 2001; WASIELEWSKI; HOLLOWAY, 2001; DEVANEY; GRAHAM; GREELEY, 2003; KAWANO, 2010; HASLER et al., 2012).

Deste modo, é possível inferir que os indicativos de embriaguez identificados pelo modelo proposto também podem variar de acordo com o ciclo circadiano. Considerando que o processo de aquisição das imagens utilizadas neste estudo ocorreu durante o ciclo noturno¹, recomenda-se que novas investigações sejam conduzidas para verificar a ocorrência de padrões de distribuição de temperatura facial característicos dos períodos diurno e noturno.

Na próxima seção serão apresentados outros fatores ambientais que podem influenciar o comportamento vascular e os padrões de distribuição da temperatura facial.

5.7 Influência de fatores ambientais nas respostas fisiológicas induzidas pelo álcool

Conforme descrito ao longo desta dissertação, a abordagem proposta para classificação de embriaguez se baseia nas variações de distribuição da temperatura facial induzidas pelo consumo de bebidas alcoólicas. Deste modo, deve-se considerar que o classificador em questão pode ser afetado por fatores que alterem o comportamento desta característica. Segundo Gordon (2018), indivíduos sob o efeito de álcool podem apresentar respostas termorreguladoras distintas de acordo com a temperatura ambiente. Yoda et al. (2005) evidenciaram que em 33°C a circulação sanguínea da pele e a produção de suor tendem a aumentar significativamente após o consumo de álcool, implicando na redução da temperatura corporal.

Em um estudo posterior, Yoda et al. (2008) utilizaram um protocolo experimental semelhante para verificar a influência do álcool no sistema termorregulador em ambientes com baixa temperatura. No experimento em questão, foi identificado que em condições climáticas abaixo de 20°C a temperatura média da pele e a corporal dos participantes envolvidos no experimento foram reduzidas em proporções semelhantes nos períodos anterior e posterior à ingestão de álcool. Contudo, foi observado que os indivíduos alcoolizados apresentaram uma taxa metabólica inferior àqueles em estado de sobriedade.

Estes resultados indicam que ainda que o álcool afete as respostas termorreguladoras do corpo humano, a maneira como tal influência é manifestada em ambientes frios e quentes podem se distinguir dos indicativos apresentados nesta dissertação. Isto posto, recomenda-se que novas investigações sejam conduzidas para verificar como as assinaturas térmicas faciais de indivíduos sob o efeito de álcool podem variar em temperaturas ambientes distintas. Os

¹ <http://www.physics.upatras.gr/sober/Sober_Drunk_DataBase/00_Readme_sober_drunk.txt>

autores da base de dados utilizada nesta dissertação (KOUKIOU; ANASTASSOPOULOS, 2015) não mencionam a condições climáticas do laboratório em que as imagens foram obtidas.

Outro fator que pode alterar o comportamento da assinatura térmica facial e afetar a eficácia do modelo proposto, é a influência de agentes estressores (CARDONE; MERLA, 2017). De acordo com Lohani, Payne e Strayer (2019), a intensidade da carga cognitiva a qual um indivíduo está submetido pode mudar o comportamento vascular de determinadas regiões do rosto, acarretando variações de temperatura em locais como a testa, o nariz, as bochechas e o queixo. Neste contexto, Pavlidis et al. (2016) demonstrou que estressores cognitivos, emocionais, sensório-motores e mistos estão associados com o aumento de transpiração na região do nariz. Ao analisar o desempenho de indivíduos influenciados por tais estímulos em um simulador de trânsito, os autores observaram que estressores sensório-motores e mistos induziram os participantes a comportamentos de risco (como saídas acentuadas de faixa).

Tais evidências sugerem que a termografia facial poderia ser empregada para a classificação de embriaguez em tempo real, ou seja, enquanto um indivíduo conduz o veículo. Contudo, no contexto da fiscalização policial deve-se considerar que agentes estressores podem afetar a assinatura térmica facial no momento da aquisição da imagem e desta feita, induzir o modelo a uma classificação incorreta.

De acordo com Koukiou e Anastassopoulos (2015), durante a aquisição da base de dados *Sober-Drunk Database* os participantes foram privados de estímulos estressores, portanto, recomenda-se que novas investigações sejam conduzidas para que se verifique como este fator pode influenciar os padrões de distribuição da temperatura facial em indivíduos alcoolizados.

5.8 Considerações sobre o capítulo

Neste capítulo foi apresentada uma discussão sobre o uso de imagens térmicas faciais para a identificação de respostas fisiológicas associadas à embriaguez. Preliminarmente, discursou-se sobre as características do conjunto de dados utilizado para a condução do estudo apresentado nesta dissertação. Em seguida, discutiu-se sobre o método de classificação proposto e as observações realizadas durante as etapas de treinamento do modelo; as quais podem estar relacionadas às imagens desfocadas presentes no conjunto de dados e às flutuações térmicas capturadas pelo sensor termográfico durante o processo de aquisição dos dados.

Ao analisar o comportamento do classificador, constatou-se que este foi capaz de reconhecer variações de temperatura em regiões de interesse previamente estudadas na literatura (seção 1.2), ademais, também foi observado que o modelo identificou um padrão

de assinatura térmica não evidenciado em trabalhos anteriores: a distribuição assimétrica de temperatura entre os lados esquerdo e direito do rosto em indivíduos alcoolizados. Isto se deve pois a abordagem adotada para a elaboração do classificador possibilita a análise das características pertinentes à distribuição da temperatura facial de maneira holística, evitando que a inferência do modelo seja orientada por assinaturas térmicas locais.

Considerando que o modelo proposto apresentou um desempenho promissor na classificação de indivíduos alcoolizados (capítulo 4) e que identificou corretamente as características associadas às respostas fisiológicas induzidas pelo álcool, constatou-se que o uso de imagens térmicas faciais para classificação de embriaguez é um método viável se suas limitações forem consideradas.

Deste modo, ao longo deste capítulo também foram apresentados diversos fatores que podem afetar o desempenho deste tipo de aplicação e como estes podem ser abordados em trabalhos futuros para o aprimoramento deste método. Preliminarmente, foi evidenciado que a intensidade e a periodicidade das respostas fisiológicas induzidas pelo álcool são influenciadas pela alcoolemia alcançada após o consumo de bebidas alcoólicas. Isto posto, deve-se considerar que a variação deste índice depende de fatores individuais, ambientais e circunstanciais que conseqüentemente irão afetar as representações das características extraídas pelo modelo.

Nas seções 5.1, 5.2 foi evidenciado como a diversidade de características físicas e fisiológicas entre as pessoas pode implicar em assinaturas térmicas faciais distintas. Neste contexto, a seção 5.3 evidenciou que a variação genética dos indivíduos pode influenciar o comportamento das enzimas responsáveis pelo metabolismo do etanol presente nas bebidas alcoólicas e desta feita, proporcionar respostas fisiológicas com intensidades distintas. Como consequência, este fator pode alterar os padrões de distribuição de temperatura facial reconhecidos pelo classificador proposto.

Nas seções 5.4 e 5.5 foi observado que fatores circunstanciais como o tipo e o volume da bebida alcoólica consumida e a presença de comida no estômago antes ou durante a ingestão do álcool podem influenciar a concentração máxima de etanol na corrente sanguínea e o período de ascensão da curva de alcoolemia. Considerando que tais propriedades estão diretamente relacionadas ao comportamento das respostas fisiológicas induzidas pelo álcool, convém salientar que suas implicações podem afetar o modelo de maneira semelhante às condições mencionadas no parágrafo anterior.

Por fim, nas seções 5.6 e 5.7 foi evidenciado que o período do dia em que a bebida alcoólica é consumida, a temperatura ambiente e a atuação de agentes estressores (cognitivos, emocionais, sensório-motores ou mistos) podem induzir respostas termorreguladoras ou simpáticas que alterem o padrão de distribuição de temperatura facial característico do estado de embriaguez.

Deste modo, acredita-se que se estes fatores forem abordados em pesquisas futuras, será possível obter um classificador com alto índice de confiança e passível de aplicação em cenários reais de fiscalização de condutores alcoolizados, veículos inteligentes ou no contexto de controle de acesso.

Conclusão

Esta dissertação apresentou uma revisão da literatura que evidencia que mesmo em baixas concentrações o álcool pode comprometer o desempenho de funções motoras e sensoriais fundamentais para se conduzir um veículo de maneira segura. Isto posto, constatou-se que um indivíduo pode representar um risco à segurança viária ainda que apresente um índice de alcoolemia abaixo do limite legal. Deste modo, esta pesquisa abordou o problema de classificação de embriaguez sob o contexto das respostas fisiológicas induzidas pelo álcool, identificando os sinais correspondentes aos efeitos da intoxicação alcoólica para categorizar o estado de um indivíduo.

Para alcançar este objetivo, elaborou-se um modelo de classificação baseado em RNCs a partir do subconjunto facial da base de dados *Sober-Drunk Database*, uma vez que esta é a única fonte publicamente disponível de imagens térmicas que correspondem ao problema investigado.

Esta metodologia possibilitou a extração de atributos invariantes à pose, evitando a seleção imprecisa de *pixels* locais ou a introdução de viés durante o processo de engenharia de características. Em comparação com estudos anteriores, a abordagem proposta foi orientada pelas propriedades intrínsecas das imagens em relação à variação da temperatura facial induzida pelo álcool, possibilitando a classificação do estado de embriaguez com uma taxa de acerto de 91% e validando a hipótese estabelecida no início desta dissertação: de que a embriaguez pode ser identificada a partir dos efeitos fisiológicos decorrentes da intoxicação alcoólica por meio da análise de características faciais extraídas da imagem térmica do rosto de um indivíduo.

Ademais, os procedimentos adotados para a elaboração deste modelo nos permitiu interpretar sua inferência ao correlacionar as características discriminantes da classe predita com o ciclo metabólico do álcool, estabelecendo uma relação entre os padrões de assinatura térmica e a variação do índice de alcoolemia. Deste modo, foi possível identificar uma distribuição assimétrica da temperatura facial em indivíduos embriagados que não foi evidenciada em estudos anteriores.

Convém salientar que o modelo apresentado ainda possui limitações, uma vez que pode ser influenciado por diversos fatores ambientais, individuais e circunstanciais (abordados no capítulo anterior). Ademais, por considerar apenas características relacionadas à variação térmica como indicativos de embriaguez, este modelo de classificação também poderia classificar padrões de distribuição de temperatura facial relacionados a outras condições fisiológicas como um sinal falso positivo de intoxicação alcoólica. As possíveis linhas de investigação que podem superar essas fragilidades serão abordadas na próxima

seção.

A despeito das limitações mencionadas, o modelo proposto foi capaz de identificar os mesmos indicativos de embriaguez em relação aos fatores sexo, peso e idade: a) a divergência de temperatura entre as regiões do nariz e da testa; b) a distribuição térmica assimétrica entre os lados esquerdo e direito do rosto. Assim, os resultados obtidos sugerem a viabilidade do uso de imagens para reconhecer os efeitos fisiológicos causados pelo álcool e incentivam a expansão desta pesquisa para identificar sinais de intoxicação induzidos por outras substâncias psicoativas.

Isto posto, acredita-se que as contribuições desta dissertação podem beneficiar o atual formato de fiscalização de embriaguez ao volante auxiliando a reduzir o número de indivíduos que dirigem sob o efeito do álcool, diminuindo assim os impactos econômicos e sociais decorrentes dos acidentes de trânsito.

Contudo, convém salientar que tais contribuições não são exclusivas ao contexto estudado nesta dissertação e podem beneficiar diversas aplicações. No domínio de veículos inteligentes, por exemplo, o conhecimento obtido com esta pesquisa poderia ser empregado para monitorar o estado do motorista e conceder ao automóvel a capacidade de adotar medidas preventivas de acordo a situação, como permitir sua condução apenas sob determinadas condições. Outro potencial domínio de aplicação é a segurança sob o contexto do controle de acesso. Neste cenário, um modelo de classificação com características análogas ao apresentado nesta dissertação poderia ser empregado para restringir a entrada de indivíduos alcoolizados em determinados ambientes, evitando situações que impliquem risco a sua própria segurança e às demais pessoas no local.

Direcionamentos para trabalhos futuros

Considerando os assuntos discutidos ao longo desta dissertação, acredita-se que a classificação de embriaguez é um problema que pode ser abordado em pesquisas futuras sob diferentes linhas de investigação, tais como: a) a elaboração de uma nova base de dados; b) a análise de indicativos de intoxicação alcoólica em imagens de diferentes faixas do espectro eletromagnético c) a classificação de embriaguez a partir da fusão de características multispectrais; d) a predição do índice de alcoolemia a partir de imagens.

Tendo em vista que fatores ambientais, individuais e circunstanciais podem afetar as representações dos indicativos de embriaguez, acredita-se que a elaboração de um conjunto de dados que contemple tais propriedades aumentaria a representatividade estatística da população analisada e contribuiria para a identificação de padrões que caracterizem tais fatores como sinais de intoxicação alcoólica.

Neste contexto, imagens de diferentes faixas do espectro eletromagnético também

poderiam ser estudadas com o intuito de se analisar características complementares à variação da distribuição de temperatura facial. [Arora et al. \(2012\)](#) demonstraram que a ingestão de álcool também afeta os processos de dilatação e contração das pupilas e que tais sinais podem ser identificados a partir de imagens em infravermelho próximo. Ademais, considerando que um dos primeiros sinais de embriaguez é a alteração comportamental ([DUBOWSKI, 1980](#)), o uso de imagens do espectro visível ou infravermelho próximo possibilitaria a análise das expressões faciais dos indivíduos.

Em relação ao espectro infravermelho longo (térmico), [Bhuyan, Bora e Koukiou \(2019\)](#) evidenciam que em adição às regiões do rosto e olhos, as mãos e o perfil lateral são locais com alto potencial para extração de características relacionadas ao consumo de álcool. Deste modo, acredita-se que métodos de classificação baseados nessas regiões poderiam complementar o conhecimento adquirido com a abordagem apresentada nesta dissertação.

Ademais, também é possível considerar respostas fisiológicas associadas à vasodilatação como atributos complementares a distribuição de temperatura facial no contexto da classificação de embriaguez. Dentre tais sintomas é possível citar as variações das taxas de respiração, batimentos cardíacos e ativação das glândulas sudoríparas ([KALANT; LE, 1983](#); [ERIKSSON, 2001](#); [CARDONE; PINTI; MERLA, 2015](#); [JONES, 2019a](#)).

Uma possível linha de investigação também relacionada ao uso de imagens térmicas e brevemente abordada nesta dissertação, refere-se a relação entre o *drift error* e a distribuição dos níveis de cinza entre os *pixels* que caracterizem o fenômeno a ser observado. Acredita-se que um estudo mais profundo sobre este assunto possa esclarecer como a metodologia de amostragem influencia as propriedades de uma imagem térmica e de que maneira tais características afetam a eficácia de aplicações relacionadas à detecção de respostas psicofisiológicas.

Não obstante, considerando que condições fisiológicas distintas podem se manifestar por meio de sinais análogos à intoxicação alcoólica ([PAVLIDIS et al., 2007](#); [TOMEOREYES; ROSS; CHANDRAN, 2016](#)), acredita-se que a utilização de imagens obtidas em diferentes faixas do espectro eletromagnético e a fusão de suas respectivas características possibilitaria a abstração de associações entre essas representações, contribuindo assim para a redução do nível de incerteza associado à natureza dos sinais identificados.

Em adição às sugestões previamente citadas, também destaca-se a possibilidade de predição do índice de alcoolemia por meio de imagens. Conforme mencionado ao longo desta dissertação, fatores como sexo, peso e idade podem influenciar a concentração de álcool no sangue e as decorrentes respostas fisiológicas de cada pessoa. No entanto, ao considerar o volume de álcool ingerido e o tempo decorrido desde o consumo da última dose, tais elementos podem ser utilizados para se estimar a alcoolemia de um indivíduo ([BRICK, 2006](#)). Deste modo, acredita-se que seja possível associar tal estimativa aos mapas

de ativação de classes para a predição da taxa de concentração de álcool no sangue após a classificação de embriaguez.

Referências

- ARLOT, S.; CELISSE, A. et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, v. 4, p. 40–79, 2010. doi:10.1214/09-SS054. Citado 2 vezes nas páginas 49 e 50.
- ARORA, S. S. et al. Iris recognition under alcohol influence: A preliminary study. In: *2012 5th IAPR International Conference on Biometrics (ICB)*. New Delhi, India: IEEE, 2012. p. 336–341. doi:10.1109/ICB.2012.6199829. Citado na página 83.
- BARAONA, E. et al. Gender differences in pharmacokinetics of alcohol. *Alcoholism: Clinical and Experimental Research*, Wiley Online Library, v. 25, n. 4, p. 502–507, 2001. doi:10.1111/j.1530-0277.2001.tb02242.x. Citado na página 71.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, v. 13, n. 10, p. 281–305, 2012. Disponível em: <<http://jmlr.org/papers/v13/bergstra12a.html>>. Citado 4 vezes nas páginas 51, 52, 55 e 57.
- BHUYAN, M.; BORA, K.; KOUKIOU, G. Detection of intoxicated person using thermal infrared images. In: *2019 IEEE 6th Asian Conference on Defence Technology (ACDT)*. Bali, Indonesia: IEEE, 2019. p. 59–64. doi:10.1109/ACDT47198.2019.9072741. Citado na página 83.
- BLANER, W. S. et al. Chronic alcohol consumption decreases brown adipose tissue mass and disrupts thermoregulation: a possible role for altered retinoid signaling. *Scientific Reports*, Nature Publishing Group, v. 7, n. 43474, p. 1–12, 2017. doi:10.1038/srep43474. Citado na página 32.
- BRICK, J. Standardization of alcohol calculations in research. *Alcoholism: Clinical and Experimental Research*, Wiley Online Library, v. 30, n. 8, p. 1276–1287, 2006. doi:10.1111/j.1530-0277.2006.00155.x. Citado 3 vezes nas páginas 71, 72 e 83.
- BRICK, J.; ERICKSON, C. K. Intoxication is not always visible: an unrecognized prevention challenge. *Alcoholism: Clinical and Experimental Research*, Blackwell Publishing Ltd Oxford, UK, v. 33, n. 9, p. 1489–1507, 2009. doi:10.1111/j.1530-0277.2009.00979.x. Citado 3 vezes nas páginas 25, 29 e 30.
- BUDDHARAJU, P. et al. Physiology-based face recognition in the thermal infrared spectrum. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 29, n. 4, p. 613–626, 2007. doi:10.1109/TPAMI.2007.1007. Citado 3 vezes nas páginas 26, 39 e 70.
- CARDONE, D.; MERLA, A. New frontiers for applications of thermal infrared imaging devices: Computational psychophysiology in the neurosciences. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 17, n. 5, p. 1–21, 2017. doi:10.3390/s17051042. Citado 2 vezes nas páginas 71 e 77.
- CARDONE, D.; PINTI, P.; MERLA, A. Thermal infrared imaging-based computational psychophysiology for psychometrics. *Computational and Mathematical Methods in*

Medicine, Hindawi, v. 2015, n. 984353, p. 1–8, 2015. doi:10.1155/2015/984353. Citado na página 83.

CASARES-LÓPEZ, M. et al. Contrast sensitivity and retinal straylight after alcohol consumption: effects on driving performance. *Scientific Reports*, Nature Publishing Group, v. 10, n. 13599, p. 1–12, 2020. doi:10.1038/s41598-020-70645-3. Citado na página 30.

CASARES-LOPEZ, M. et al. Changes in accommodation dynamics after alcohol consumption, for two different doses. *Graefe's Archive for Clinical and Experimental Ophthalmology*, Springer, v. 259, n. 4, p. 919–928, 2021. doi:10.1007/s00417-020-04978-0. Citado na página 30.

CASTRO, J. J. et al. Retinal-image quality and night-vision performance after alcohol consumption. *Journal of Ophthalmology*, Hindawi, v. 2014, p. 1–7, 2014. doi:10.1155/2014/704823. Citado na página 30.

CEDERBAUM, A. I. Alcohol metabolism. *Clinics in Liver Disease*, Elsevier, v. 16, n. 4, p. 667–685, 2012. doi:10.1016/j.cld.2012.08.002. Citado 4 vezes nas páginas 71, 72, 74 e 75.

CHARLTON, S. G.; STARKEY, N. J. Driving while drinking: performance impairments resulting from social drinking. *Accident Analysis & Prevention*, Elsevier, v. 74, p. 210–217, 2015. doi:10.1016/j.aap.2014.11.001. Citado na página 29.

COMPTON, R. P.; BERNING, A. *Drug and alcohol crash risk. Traffic Safety Facts Research Note*. Washington, DC, 2015. Citado na página 25.

CONSELHO NACIONAL DE TRÂNSITO. *Resolução n. 432 de 23 de janeiro de 2013*. 2013. Dispõe sobre os procedimentos a serem adotados pelas autoridades de trânsito e seus agentes na fiscalização do consumo de álcool ou de outra substância psicoativa que determine dependência, para aplicação do disposto nos arts. 165, 276, 277 e 306 da Lei n. 9.503, de 23 de setembro de 1997 - Código de Trânsito Brasileiro (CTB). Disponível em: <<https://www.gov.br/infraestrutura/pt-br/assuntos/transito/conteudo-contran/resolucoes/resolu-o-uo-432-2013c.pdf>>. Acesso em: 14 jul. 2021. Citado 2 vezes nas páginas 29 e 30.

DANEL, T.; LIBERSA, C.; TOUITOU, Y. The effect of alcohol consumption on the circadian control of human core body temperature is time dependent. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, American Physiological Society Bethesda, MD, v. 281, n. 1, p. R52–R55, 2001. doi:10.1152/ajpregu.2001.281.1.R52. Citado 2 vezes nas páginas 25 e 76.

DEVANEY, M.; GRAHAM, D.; GREELEY, J. Circadian variation of the acute and delayed response to alcohol: investigation of core body temperature variations in humans. *Pharmacology Biochemistry and Behavior*, Elsevier, v. 75, n. 4, p. 881–887, 2003. doi:10.1016/S0091-3057(03)00170-9. Citado 2 vezes nas páginas 31 e 76.

DUBOWSKI, K. M. Alcohol determination in the clinical laboratory. *American journal of clinical pathology*, Oxford University Press, v. 74, n. 5, p. 747–750, 1980. doi:10.1093/ajcp/74.5.747. Citado na página 83.

EDENBERG, H. J. The genetics of alcohol metabolism: Role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health*, National Institute on

- Alcohol Abuse and Alcoholism, v. 30, n. 1, p. 5–13, 2007. PMID: 17718394; PMCID: PMC3860432. Citado na página 74.
- ENG, M. Y.; LUCZAK, S. E.; WALL, T. L. Aldh2, adh1b, and adh1c genotypes in asians: a literature review. *Alcohol Research & Health*, National Institute on Alcohol Abuse and Alcoholism, v. 30, n. 1, p. 22–27, 2007. PMID: 17718397; PMCID: PMC3860439. Citado na página 74.
- ERIKSSON, C. P. The role of acetaldehyde in the actions of alcohol (update 2000). *Alcoholism: Clinical and Experimental Research*, Wiley Online Library, v. 25, p. 15S–32S, 2001. doi:10.1111/j.1530-0277.2001.tb02369.x. Citado 5 vezes nas páginas 30, 32, 43, 74 e 83.
- FELL, J. C.; VOAS, R. B. The effectiveness of reducing illegal blood alcohol concentration (bac) limits for driving: evidence for lowering the limit to .05 bac. *Journal of safety research*, Elsevier, v. 37, n. 3, p. 233–243, 2006. doi:10.1016/j.jsr.2005.07.006. Citado na página 25.
- FERNÁNDEZ-CUEVAS, I. et al. Classification of factors influencing the use of infrared thermography in humans: A review. *Infrared Physics & Technology*, Elsevier, v. 71, p. 28–55, 2015. doi:10.1016/j.infrared.2015.02.007. Citado na página 32.
- FREUND, B. J.; O'BRIEN, C.; YOUNG, A. J. Alcohol ingestion and temperature regulation during cold exposure. *Journal of Wilderness Medicine*, Elsevier, v. 5, n. 1, p. 88–98, 1994. doi:10.1580/0953-9859-5.1.88. Citado na página 25.
- GOEDDE, H. et al. Distribution of adh2 and ald2 genotypes in different populations. *Human Genetics*, Springer, v. 88, n. 3, p. 344–346, 1992. doi:10.1007/BF00197271. Citado na página 74.
- GONZALES, R. C.; WOODS, R. E. Image Compression. In: _____. *Digital Image Processing*. 3. ed. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2008. cap. 8, p. 525–626. Citado na página 54.
- GONZALES, R. C.; WOODS, R. E. Representation and Description. In: _____. *Digital Image Processing*. 3. ed. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2008. cap. 11, p. 795–860. Citado na página 54.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Feedforward Networks. In: _____. *Deep Learning*. MIT Press, 2016. cap. 6, p. 168–227. Disponível em: <<http://www.deeplearningbook.org>>. Citado na página 52.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Optimization for Training Deep Models. In: _____. *Deep Learning*. MIT Press, 2016. cap. 8, p. 274–329. Disponível em: <<http://www.deeplearningbook.org>>. Citado 2 vezes nas páginas 52 e 53.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Practical Methodology. In: _____. *Deep Learning*. MIT Press, 2016. cap. 11, p. 421–442. Disponível em: <<http://www.deeplearningbook.org>>. Citado na página 55.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Probability and Information Theory. In: _____. *Deep Learning*. MIT Press, 2016. cap. 3, p. 53–79. Disponível em: <<http://www.deeplearningbook.org>>. Citado 2 vezes nas páginas 52 e 53.

- GORDON, C. J. Response of the thermoregulatory system to toxic chemicals. In: SHRIVASTAVA, D. (Ed.). *Theory and Applications of Heat Transfer in Humans*. West Sussex, UK: John Wiley & Sons Ltd, 2018. v. 1, cap. 25, p. 529–552. doi:10.1002/9781119127420.ch25. Citado na página 76.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, IEEE, SMC-3, n. 6, p. 610–621, 1973. doi:10.1109/TSMC.1973.4309314. Citado na página 54.
- HASLER, B. P. et al. Circadian rhythms, sleep, and substance abuse. *Sleep Medicine Reviews*, Elsevier, v. 16, n. 1, p. 67–81, 2012. doi:10.1016/j.smr.2011.03.004. Citado na página 76.
- HERMOSILLA, G. et al. Face recognition and drunk classification using infrared face images. *Journal of Sensors*, Hindawi, v. 2018, 2018. doi:10.1155/2018/5813514. Citado 2 vezes nas páginas 26 e 70.
- HOSSIN, M.; SULAIMAN, M. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, Academy & Industry Research Collaboration Center (AIRCC), v. 5, n. 2, p. 1–11, 2015. doi:10.5121/ijdkp.2015.5201. Citado na página 53.
- JONES, A. W. Evidence-based survey of the elimination rates of ethanol from blood with applications in forensic casework. *Forensic Science International*, Elsevier, v. 200, n. 1-3, p. 1–20, 2010. doi:10.1016/j.forsciint.2010.02.021. Citado 2 vezes nas páginas 43 e 75.
- JONES, A. W. Alcohol, its absorption, distribution, metabolism, and excretion in the body and pharmacokinetic calculations. *Wiley Interdisciplinary Reviews: Forensic Science*, Wiley Online Library, v. 1, n. 5, p. e1340, 2019. doi:10.1002/wfs2.1340. Citado 4 vezes nas páginas 71, 72, 75 e 83.
- JONES, A. W. Alcohol, its analysis in blood and breath for forensic purposes, impairment effects, and acute toxicity. *Wiley Interdisciplinary Reviews: Forensic Science*, Wiley Online Library, v. 1, n. 6, p. e1353, 2019. doi:10.1002/wfs2.1353. Citado 2 vezes nas páginas 29 e 30.
- KALANT, H.; LE, A. Effects of ethanol on thermoregulation. *Pharmacology & Therapeutics*, Elsevier, v. 23, n. 3, p. 313–364, 1983. doi:10.1016/0163-7258(83)90018-9. Citado 4 vezes nas páginas 25, 30, 32 e 83.
- KANG, J. H. et al. Prospective study of alcohol consumption and the risk of primary open-angle glaucoma. *Ophthalmic Epidemiology*, Taylor & Francis, v. 14, n. 3, p. 141–147, 2007. doi:10.1080/09286580601187963. Citado na página 32.
- KAWANO, Y. Physio-pathological effects of alcohol on the cardiovascular system: its role in hypertension and cardiovascular disease. *Hypertension Research*, Nature Publishing Group, v. 33, n. 3, p. 181–191, 2010. doi:10.1038/hr.2009.226. Citado 5 vezes nas páginas 32, 71, 74, 75 e 76.
- KHAN, A. et al. A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, Springer, v. 53, n. 8, p. 5455–5516, 2020. doi:10.1007/s10462-020-09825-6. Citado 4 vezes nas páginas 33, 34, 45 e 48.

- KHAN, M. M.; WARD, R. D.; INGLEBY, M. Classifying pretended and evoked facial expressions of positive and negative affective states using infrared measurement of skin temperature. *ACM Transactions on Applied Perception*, ACM, v. 6, n. 1, p. 1–22, 2009. doi:10.1145/1462055.1462061. Citado na página 26.
- KOPACZKA, M. et al. A thermal infrared face database with facial landmarks and emotion labels. *IEEE Transactions on Instrumentation and Measurement*, IEEE, v. 68, n. 5, p. 1389–1401, 2018. doi:10.1109/TIM.2018.2884364. Citado 2 vezes nas páginas 26 e 39.
- KOUKIOU, G. Intoxicated person identification using markov chains and neural networks. *Neural Computing and Applications*, Springer, v. 33, n. 8, p. 3459–3467, 2021. doi:10.1007/s00521-020-05219-5. Citado 2 vezes nas páginas 32 e 70.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Drunk person identification using thermal infrared images. *International Journal of Electronic Security and Digital Forensics*, Inderscience Publishers, v. 4, n. 4, p. 229–243, 2012. doi:10.1504/IJESDF.2012.049747. Citado 4 vezes nas páginas 26, 39, 40 e 70.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Neural networks for identifying drunk persons using thermal infrared imagery. *Forensic Science International*, Elsevier, v. 252, p. 69–76, 2015. doi:10.1016/j.forsciint.2015.04.022. Citado 10 vezes nas páginas 26, 32, 39, 40, 43, 69, 70, 71, 75 e 77.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Drunk person screening using eye thermal signatures. *Journal of Forensic Sciences*, Wiley Online Library, v. 61, n. 1, p. 259–264, 2016. doi:10.1111/1556-4029.12989. Citado 4 vezes nas páginas 26, 39, 40 e 70.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Intoxicated person discrimination using infrared signature of facial blood vessels. *Australian Journal of Forensic Sciences*, Taylor & Francis, v. 48, n. 3, p. 326–338, 2016. doi:10.1080/00450618.2015.1060522. Citado 3 vezes nas páginas 32, 33 e 70.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Fusion using neural networks for intoxication identification. In: *2018 International Workshop on Biometrics and Forensics (IWBF)*. Sassari, Italy: IEEE, 2018. p. 1–5. doi:10.1109/IWBF.2018.8401556. Citado 4 vezes nas páginas 26, 39, 40 e 70.
- KOUKIOU, G.; ANASTASSOPOULOS, V. Local difference patterns for drunk person identification. *Multimedia Tools and Applications*, Springer, v. 77, n. 8, p. 9293–9305, 2018. doi:10.1007/s11042-017-4892-6. Citado na página 32.
- KUBICEK, J. et al. Detection and dynamical tracking of temperature facial distribution caused by alcohol intoxication with using of modified otsu regional segmentation. In: NGUYEN, N. et al. (Ed.). *Asian Conference on Intelligent Information and Database Systems. ACHIIDS 2018. Lecture Notes in Computer Science*. Dong Hoi City, Vietnam: Springer, 2018. v. 10752, p. 357–366. doi:10.1007/978-3-319-75420-8_34. Citado na página 32.
- KUBICEK, J. et al. Prediction model of alcohol intoxication from facial temperature dynamics based on k-means clustering driven by evolutionary computing. *Symmetry*, Multidisciplinary Digital Publishing Institute, v. 11, n. 8, p. 1–31, 2019. doi:10.3390/sym11080995. Citado na página 32.

- LANTMAN, M. van S. et al. Total sleep time, alcohol consumption, and the duration and severity of alcohol hangover. *Nature and Science of Sleep*, Dove Press, v. 9, p. 181–186, 2017. doi:10.2147/NSS.S136467. Citado na página 30.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. doi:10.1038/nature14539. Citado 5 vezes nas páginas 33, 34, 35, 45 e 69.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. doi:10.1109/5.726791. Citado 4 vezes nas páginas 34, 35, 40 e 69.
- LI, Z. et al. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, p. 1–21, 2021. doi:10.1109/TNNLS.2021.3084827. Citado 2 vezes nas páginas 51 e 55.
- LIU, Y.-C.; HO, C. H. Effects of different blood alcohol concentrations and post-alcohol impairment on driving behavior and task performance. *Traffic Injury Prevention*, Taylor & Francis, v. 11, n. 4, p. 334–341, 2010. doi:10.1080/15389581003747522. Citado na página 29.
- LOHANI, M.; PAYNE, B. R.; STRAYER, D. L. A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience*, Frontiers, v. 13, n. 57, p. 1–27, 2019. doi:10.3389/fnhum.2019.00057. Citado 2 vezes nas páginas 71 e 77.
- MARTIN, T. L. et al. A review of alcohol-impaired driving: The role of blood alcohol concentration and complexity of the driving task. *Journal of Forensic Sciences*, Wiley Online Library, v. 58, n. 5, p. 1238–1250, 2013. doi:10.1111/1556-4029.12227. Citado 2 vezes nas páginas 25 e 29.
- MILLER, G. A. Wordnet: A lexical database for english. *Communications of the ACM*, ACM, v. 38, n. 11, p. 39–41, 1995. doi:10.1145/219717.219748. Citado na página 45.
- NORBERG, Å. et al. Role of variability in explaining ethanol pharmacokinetics. *Clinical Pharmacokinetics*, Springer, v. 42, n. 1, p. 1–31, 2003. doi:10.2165/00003088-200342010-00001. Citado na página 72.
- OBI-ALAGO, O.; YANUSHKEVICH, S. N.; WETHERLEY, H. Detecting thermal face signature abnormalities. In: IEEE. *2019 Eighth International Conference on Emerging Security Technologies (EST)*. Colchester, UK, 2019. p. 1–6. doi:10.1109/EST.2019.8806217. Citado 3 vezes nas páginas 40, 67 e 70.
- OGDEN, E. J.; MOSKOWITZ, H. Effects of alcohol and other drugs on driver performance. *Traffic Injury Prevention*, Taylor & Francis, v. 5, n. 3, p. 185–198, 2004. doi:10.1080/15389580490465201. Citado na página 29.
- OSCAR-BERMAN, M.; MARINKOVIĆ, K. Alcohol: effects on neurobehavioral functions and the brain. *Neuropsychology Review*, Springer, v. 17, n. 3, p. 239–257, 2007. doi:10.1007/s11065-007-9038-6. Citado na página 29.
- PAN, S. J.; YANG, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, IEEE, v. 22, n. 10, p. 1345–1359, 2010. doi:10.1109/TKDE.2009.191. Citado na página 35.

PARKHI, O. M.; VEDALDI, A.; ZISSERMAN, A. Deep face recognition. In: XIE, X.; JONES, M. W.; TAM, G. K. L. (Ed.). *Proceedings of the British Machine Vision Conference (BMVC)*. Swansea, UK: BMVA Press, 2015. p. 41.1–41.12. doi:10.5244/C.29.41. Citado 2 vezes nas páginas 51 e 57.

PAVLIDIS, I. et al. Dissecting driver behaviors under cognitive, emotional, sensorimotor, and mixed stressors. *Scientific Reports*, Nature Publishing Group, v. 6, n. 25651, p. 1–12, 2016. doi:10.1038/srep25651. Citado 2 vezes nas páginas 71 e 77.

PAVLIDIS, I. et al. Interacting with human physiology. *Computer Vision and Image Understanding*, Elsevier, v. 108, n. 1-2, p. 150–170, 2007. doi:10.1016/j.cviu.2006.11.018. Citado 2 vezes nas páginas 39 e 83.

REINBERG, A. et al. Habitual moderate alcohol consumption desynchronizes circadian physiologic rhythms and affects reaction-time performance. *Chronobiology International*, Taylor & Francis, v. 27, n. 9-10, p. 1930–1942, 2010. doi:10.3109/07420528.2010.515763. Citado na página 30.

RIOU, O.; BERREBI, S.; BREMOND, P. Nonuniformity correction and thermal drift compensation of thermal infrared camera. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Thermosense XXVI*. Florida, United States: SPIE, 2004. v. 5405, p. 294–302. doi:10.1117/12.547807. Citado 2 vezes nas páginas 43 e 69.

RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, Springer, v. 115, n. 3, p. 211–252, 2015. doi:10.1007/s11263-015-0816-y. Citado 2 vezes nas páginas 45 e 48.

SANTIS, A. D.; IACOVIELLO, D. Optimal segmentation of pupillometric images for estimating pupil shape parameters. *Computer Methods and Programs in Biomedicine*, Elsevier, v. 84, n. 2-3, p. 174–187, 2006. doi:10.1016/j.cmpb.2006.07.005. Citado na página 31.

SELVARAJU, R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE international conference on computer vision (ICCV)*. Venice, Italy: IEEE, 2017. p. 618–626. doi:10.1109/ICCV.2017.74. Citado 2 vezes nas páginas 53 e 69.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, Springer, v. 6, n. 1, p. 1–48, 2019. doi:10.1186/s40537-019-0197-0. Citado 2 vezes nas páginas 35 e 36.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: BENGIO, Y.; LECUN, Y. (Ed.). *3rd International Conference on Learning Representations (ICLR)*. San Diego, CA, USA, 2015. p. 14. Disponível em: <<http://arxiv.org/abs/1409.1556>>. Citado 5 vezes nas páginas 45, 48, 51, 57 e 69.

STARKEY, N. J.; CHARLTON, S. G. The effects of moderate alcohol concentrations on driving and cognitive performance during ascending and descending blood alcohol concentrations. *Human Psychopharmacology: Clinical and Experimental*, Wiley Online Library, v. 29, n. 4, p. 370–383, 2014. doi:10.1002/hup.2415. Citado na página 29.

- SZEGEDY, C. et al. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco, USA: AAAI Press, 2017. p. 4278–4284. Disponível em: <<https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14806/14311>>. Citado na página 48.
- TATTERSALL, G. J. Infrared thermography: A non-invasive window into thermal physiology. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, Elsevier, v. 202, p. 78–98, 2016. doi:10.1016/j.cbpa.2016.02.022. Citado 3 vezes nas páginas 43, 44 e 69.
- TOMEIO-REYES, I.; ROSS, A.; CHANDRAN, V. Investigating the impact of drug induced pupil dilation on automated iris recognition. In: *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. Niagara Falls, NY, USA: IEEE, 2016. p. 1–8. doi:10.1109/BTAS.2016.7791178. Citado na página 83.
- TORIS, C. B.; TYE, G.; PATTABIRAMAN, P. Changes in parameters of aqueous humor dynamics throughout life. In: GUIDOBONI, G.; HARRIS, A.; SACCO, R. (Ed.). *Ocular Fluid Dynamics. Modeling and Simulation in Science, Engineering and Technology*. 1. ed. Cham, Switzerland: Springer, 2019. p. 161–190. doi:10.1007/978-3-030-25886-3_6. Citado na página 32.
- TYSON, T. L. et al. Dose-dependent sensorimotor impairment in human ocular tracking after acute low-dose alcohol administration. *The Journal of Physiology*, Wiley Online Library, v. 599, n. 4, p. 1225–1242, 2021. doi:10.1113/JP280395. Citado na página 30.
- UNITED NATIONS. *Resolution adopted by the General Assembly on 31 August 2020*. 2020. Improving global road safety. Disponível em: <<https://undocs.org/en/A/RES/74/299>>. Acesso em: 28 jul. 2021. Citado na página 25.
- USAMENTIAGA, R. et al. Infrared thermography for temperature measurement and non-destructive testing. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 14, n. 7, p. 12305–12348, 2014. doi:10.3390/s140712305. Citado 2 vezes nas páginas 43 e 69.
- VERSTER, J. C. et al. Effects of alcohol hangover on simulated highway driving performance. *Psychopharmacology*, Springer, v. 231, n. 15, p. 2999–3008, 2014. doi:10.1007/s00213-014-3474-9. Citado na página 29.
- VONGHIA, L. et al. Acute alcohol intoxication. *European Journal of Internal Medicine*, Elsevier, v. 19, n. 8, p. 561–567, 2008. doi:10.1016/j.ejim.2007.06.033. Citado 2 vezes nas páginas 30 e 32.
- WASIELEWSKI, J. A.; HOLLOWAY, F. A. Alcohol's interactions with circadian rhythms: A focus on body temperature. *Alcohol Research & Health*, v. 25, p. 94–100, 2001. PMID: 11584555; PMCID: PMC6707125. Citado na página 76.
- WEBER, A. et al. Retrobulbar blood flow and visual field alterations after acute ethanol ingestion. *Clinical Ophthalmology*, Dove Press, v. 7, p. 1641–1646, 2013. doi:10.2147/OPHTH.S45304. Citado na página 32.
- WOLF, R.; TÜZÜN, B.; TÜZÜN, Y. Alcohol ingestion and the cutaneous vasculature. *Clinics in Dermatology*, Elsevier, v. 17, n. 4, p. 395–403, 1999. doi:10.1016/S0738-081X(99)00023-1. Citado na página 32.

WORLD HEALTH ORGANIZATION. *Global status report on alcohol and health 2018*. Geneva: World Health Organization, 2018. Citado 2 vezes nas páginas 29 e 30.

WORLD HEALTH ORGANIZATION. *Global status report on road safety 2018*. Geneva: World Health Organization, 2018. Citado na página 25.

XIE, Z. et al. Drunk identification using far infrared imagery based on dct features in dwt domain. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Infrared Technology and Applications, and Robot Sensing and Advanced Control*. Beijing, China: SPIE, 2016. v. 10157, p. 345–350. doi:10.1117/12.2246469. Citado na página 40.

YODA, T. et al. Effects of alcohol on thermoregulation during mild heat exposure in humans. *Alcohol*, Elsevier, v. 36, n. 3, p. 195–200, 2005. doi:10.1016/j.alcohol.2005.09.002. Citado 2 vezes nas páginas 25 e 76.

YODA, T. et al. Effects of alcohol on autonomic responses and thermal sensation during cold exposure in humans. *Alcohol*, Elsevier, v. 42, n. 3, p. 207–212, 2008. doi:10.1016/j.alcohol.2008.01.006. Citado na página 76.

ZAKHARI, S. Overview: How is alcohol metabolized by the body? *Alcohol Research & Health*, National Institute on Alcohol Abuse and Alcoholism, v. 29, n. 4, p. 245–254, 2006. PMID: 17718403; PMCID: PMC6527027. Citado na página 74.

ZHANG, W. et al. Deep learning based torsional nystagmus detection for dizziness and vertigo diagnosis. *Biomedical Signal Processing and Control*, Elsevier, v. 68, n. 102616, 2021. doi:10.1016/j.bspc.2021.102616. Citado na página 32.

ZHONG, Z. et al. Genetic polymorphisms of the mitochondrial aldehyde dehydrogenase aldh2 gene in a large ethnic hakka population in southern china. *Medical science monitor: international medical journal of experimental and clinical research*, International Scientific Information, Inc., v. 24, p. 2038–2044, 2018. doi:10.12659/msm.906606. Citado na página 74.