

Arquitetura Baseada em Aprendizado de Máquina para Inferência de Fadiga em Condutores a partir de Sinais Fisiológicos Multimodais

Luísa C. de Souza^{*†} , Karolayne S. Azevedo^{*†} , Marianne Silva[‡] ,
Ivanovitch M. D. da Silva[§] , Marcelo A. C. Fernandes^{*†§¶} 

^{*}InovAI Lab, nPITI/IMD, Universidade Federal do Rio Grande do Norte (UFRN), Natal, RN, Brasil.

[†]Leading Advanced Technologies Center of Excellence (LANCE), nPITI/IMD, UFRN, Natal, RN, Brasil.

[‡]Universidade Federal de Alagoas (UFAL) – Unidade de Ensino Penedo, Penedo, AL, Brasil.

[§]Departamento de Engenharia da Computação e Automação (DCA), UFRN, Natal, RN, Brasil.

[¶]Centro Multiusuário de Bioinformática (BioME) – IMD, UFRN, Natal, RN, Brasil.

Emails: luisa.souza.103@ufrn.edu.br, karolayneazevedosantos@gmail.com, marianne.silva@penedo.ufal.br,
ivanovitch.silva@ufrn.br, mfernandes@dca.ufrn.br

Abstract—Fatores relacionados ao condutor continuam sendo uma das principais causas de acidentes viários, com a fadiga e a redução do estado de alerta representando riscos significativos, especialmente em cenários de condução automatizada. Este trabalho propõe uma metodologia para caracterização do comportamento de motoristas e inferência de estados de fadiga com base na análise de sinais fisiológicos multimodais coletados em simulações de direção. Os sinais foram transformados em representações visuais por meio da Transformada de *wavelet* Contínua e processados com uma rede neural convolucional pré-treinada (ResNet50) para extração de características. Técnicas de aprendizado não supervisionado e análise exploratória permitiram agrupar os dados e investigar padrões associados à fadiga, utilizando parâmetros veiculares e frequência de piscadas. Os resultados sugerem que a área dos gráficos de radar, aliada à redução da taxa de piscadas, pode indicar variações no estado de fadiga, embora a correlação não se confirme de forma universal. A abordagem mostrou-se promissora e destaca o potencial da análise multimodal e a integração de múltiplos parâmetros assistida por inteligência artificial para aplicações em segurança veicular. Como trabalho futuro, propõe-se a incorporação de métricas complementares, além da avaliação de algoritmos de agrupamento mais robustos.

Index Terms—Fadiga do condutor, sinais fisiológicos, aprendizado não supervisionado, processamento de sinais, condução automatizada, análise multimodal.

I. INTRODUCTION

Apesar dos avanços na tecnologia automotiva, os acidentes rodoviários continuam sendo uma questão global crítica. De acordo com a Organização Mundial da Saúde, os acidentes de trânsito resultam em aproximadamente 1,19 milhão de mortes e cerca de 20 a 50 milhões de feridos a cada ano, com um impacto econômico estimado em até 3% do Produto Interno Bruto (PIB) anual de um país [1]. Notavelmente, cerca de 90% desses acidentes são atribuídos a erro humano, com fatores contribuintes importantes, incluindo superconfiança do motorista, distração, fadiga, excesso de velocidade, estresse, pressão do tempo e comportamento imprudente [2]–[4].

Diante da crescente demanda por segurança e eficiência no trânsito, iniciativas de pesquisa voltadas à implementação e avaliação de condições para a navegação autônoma, bem como para a detecção de alterações nas condições psicofísicas dos motoristas, têm ganhado destaque [5]. Além disso, o desempenho de sistemas de condução autônoma pode ser aprimorado por meio do estudo do comportamento e do estado mental do condutor, os quais podem ser inferidos a partir de sinais fisiológicos, como frequência cardíaca e padrões respiratórios, com o apoio de técnicas de inteligência artificial [6].

Na literatura, métodos tradicionais baseados exclusivamente na análise de imagens e vídeos para detectar comportamentos anormais de motoristas têm se mostrado limitados na captação de padrões temporais complexos. Para superar essas limitações, estudos acadêmicos vêm explorando abordagens que integram a análise de expressões faciais, sinais biológicos e a dinâmica do veículo. Adicionalmente, sistemas baseados em inteligência artificial têm sido desenvolvidos recentemente com o objetivo de detectar e mitigar comprometimentos do condutor, contribuindo para o aumento da segurança viária [7], [8].

O corpo humano gera continuamente uma variedade de sinais biológicos que refletem seu estado fisiológico, sendo que alterações em seus padrões muitas vezes indicam mudanças nas condições de saúde [9], [10]. Devido à natureza intrinsecamente não estacionária, não linear e não gaussiana desses sinais, a análise fisiológica enfrenta desafios significativos, especialmente no processo de extração de características. Essa etapa desempenha um papel fundamental na análise de sinais fisiológicos, pois as características extraídas devem representar com precisão os padrões e dinâmicas subjacentes dos sinais [11]. Além disso, a informação transmitida pelos biosinais está intrinsecamente associada às suas condições iniciais, duração e variabilidade. Portanto, uma extração de características restrita a um único domínio, como tempo, espaço

ou frequência é insuficiente para representar plenamente as informações incorporadas [10].

Neste contexto, este trabalho propõe um estudo voltado à caracterização e detecção de comportamentos de motoristas por meio da análise de sinais fisiológicos multimodais, utilizando técnicas de processamento de sinais e aprendizado de máquina. Para isso, será utilizado um conjunto de dados disponibilizado pelo projeto chamado de *Holistic Approach for Driver Role Integration and Automation Allocation for European Mobility Needs* (HADRIAN) [12], o qual reúne sinais fisiológicos captados durante sessões de simulação de direção, além de informações sobre o desempenho dos motoristas e as condições dinâmicas da condução. A metodologia apresenta o desenvolvimento de uma técnica de representação baseada na Transformada de *wavelet* Contínua (TWC) para converter os sinais em imagens, facilitando a extração de padrões relevantes. Ademais, será implementada uma metodologia de análise exploratória combinada com aprendizado de máquina não supervisionado, voltada à inferência de fadiga e estados mentais de condutores.

II. TRABALHOS RELACIONADOS

A análise e anotação de sinais fisiológicos, inicialmente realizada de forma manual, evoluiu com a introdução de filtros de redução de ruído e técnicas de extração de características baseadas em processamento digital de sinais e algoritmos estatísticos de classificação. Essa evolução viabilizou o desenvolvimento de métodos automatizados voltados à orientação diagnóstica e apoio à decisão clínica [13]. Com o avanço da inteligência artificial, têm-se observado progressos significativos na análise de dados fisiológicos e na detecção de estados mentais de condutores.

Em [14], os autores propõem o uso da transformada de Radon para representar sinais de eletrocardiograma (ECG) em formato tomográfico bidimensional, possibilitando a identificação de padrões espectrais associados à doença isquêmica cardíaca precoce não visíveis na análise tradicional de Fourier. A detecção de emoções por meio de sinais fisiológicos também tem sido explorada. Em [15], sinais de atividade eletrodérmica (EDA) foram usados como entrada para redes neurais convolucionais unidimensionais (1D-CNN), alcançando F1-scores superiores a 91% em modelos híbridos com LSTM e BiLSTM. Em [16], foi proposta uma fusão temporal multimodal de sinais de eletroencefalograma (EEG) e pulso de volume sanguíneo (*Blood Volume Pulse* - BVP), utilizando uma arquitetura *end-to-end* com CNN e LSTM, atingindo acurácia superior a 71% na classificação de emoções.

O trabalho de [17] empregou técnicas de extração estatística e modelos como *Ensemble Bagged Tree* e KNN para classificar emoções em três bases de dados públicas com sinais de EEG, ECG entre outros. A abordagem apresentou elevadas taxas de acerto, com 100% de taxa de verdadeiros positivos no conjunto SWELL. Focando na detecção de fadiga em veículos automatizados, [18] utilizaram sinais de EEG processados com *Independent Component Analysis* e redes CNN-LSTM, atingindo F1-scores superiores a 97,4%, mesmo em condições

de ruído, prevenindo transições de fadiga com até 13,8 segundos de antecedência.

Abordagens baseadas em visão computacional também são utilizadas para detectar fadiga. Em [19], marcos faciais são detectados para calcular o *Eye Aspect Ratio* (EAR) e o *Mouth Aspect Ratio* (MAR), monitorando piscadas e bocejos. Caso os valores de EAR e MAR se mantiverem, respectivamente, abaixo e acima de limiares pré definidos, por um número consecutivo de (frames), alertas de sonolência são disparados. De forma semelhante, [20] propuseram um sistema de baixo custo baseado em vídeos e técnicas como Viola-Jones para detecção facial, seguido de classificação por SVM e Adaboost. O sistema identificou episódios de micro-sono com base em limiares de olhos fechados em quadros consecutivos.

Esses estudos demonstram a viabilidade de integrar sinais fisiológicos e técnicas de aprendizagem de máquina para inferência do estado mental do condutor, destacando o potencial do uso de abordagens híbridas e multimodais na detecção de fadiga e excitação reduzida.

III. METODOLOGIA

Esta seção detalha a metodologia proposta para a análise do comportamento e das condições de motoristas, empregando técnicas de aprendizagem de máquina e análise exploratória de dados. Para tal, a pesquisa se baseia em sinais fisiológicos e dados de simulação de direção, provenientes do Hadrian Project Dataset, permitindo a extração de inferências relevantes a partir de uma abordagem multimodal. Embora o conjunto de dados não forneça rótulos objetivos para estados emocionais complexos, ele oferece uma rica fonte de informações sobre o comportamento físico do motorista e seu perfil de direção. Serão abordados, sequencialmente, a descrição do dataset, os procedimentos para a representação de sinais fisiológicos, a detecção da direção manual e a inferência do estado mental do motorista relacionado à fadiga.

A. Conjunto de Dados do Projeto HADRIAN

O projeto HADRIAN tem como foco a integração homem-sistema na direção automatizada, considerando fatores humanos, ambientais e de projeto sistêmico para melhorar a segurança e a experiência do usuário [12], [21].

Foram conduzidos três experimentos principais: LAS (*Low Arousal and Sleepiness*), DD (*Driver Distraction*) e SH (*Stress and Hazard*). Este trabalho utiliza exclusivamente os dados do experimento LAS, que avalia o impacto da monotonia na direção sobre o estado de alerta e o comportamento do condutor.

No protocolo LAS, 21 indivíduos participaram do experimento, cada um identificado por um 'ID' distinto. Conforme ilustrado na Figura 1, os participantes conduziram um veículo simulado por três horas em uma rodovia sem tráfego, alternando entre os modos de direção manual e autônoma. Apenas os 90 minutos referentes à Direção Manual foram utilizados para as análises de fadiga, segmentados em 18 blocos de 5 minutos cada.

Foram utilizados quatro conjuntos de dados:

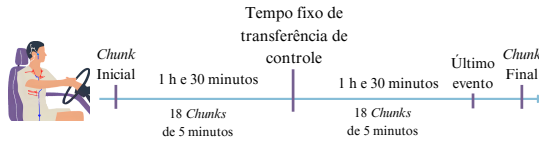


Fig. 1: Estrutura do experimento LAS, com duas sessões de 90 minutos separadas por uma transição de controle (take-over).

- **Sinais Fisiológicos** [22]: incluem EDA, ECG, BVP e respiração, amostrados a 400 Hz, uniformemente reamostrados após pré-processamento.
- **Índices do Simulador de Direção** [23]: contêm 45 variáveis relacionadas à dinâmica veicular, como aceleração, frenagem, velocidade, posição na pista, ângulo do volante e condições ambientais, segmentadas em 39 blocos temporais.
- **Movimentos Oculares** [24]: 187 parâmetros extraídos via sistema Smart Eye PRO a 120 Hz, incluindo posições e orientações da cabeça e dos olhos, direção do olhar, piscadas e métricas de qualidade de rastreamento.
- **Anotações Comportamentais** [25]: informações qualitativas sobre eventos como expressões faciais (bocejos, caretas), gestos manuais (coçar o rosto, mãos fora do volante), posturas corporais e ocorrências de colisões ou quase-colisões.

B. Representação dos Sinais Fisiológicos

A Figura 2 apresenta um minuto dos sinais brutos de Respiração, EDA, ECG e BVP obtidos do participante de ID 53 do experimento LAS.

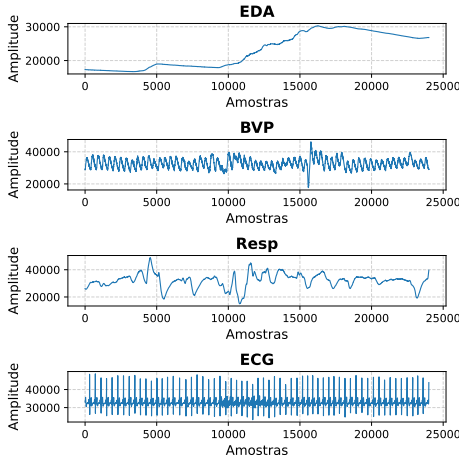


Fig. 2: Um minuto de cada sinal fisiológico bruto captado no experimento LAS.

A representação adequada de sinais fisiológicos é essencial para a extração de características discriminativas que possam ser utilizadas em tarefas de classificação e inferência [11]. Neste trabalho, os sinais unidimensionais, exibidos na Figura 2, foram convertidos em representações visuais bidimensionais

por meio da TWC, a qual é particularmente eficaz para sinais não estacionários, como os sinais fisiológicos [26], [27]. O fluxograma apresentado na Figura 3 ilustra o fluxo de atividades para a geração de uma representação de sinal utilizando a transformada *wavelet*.

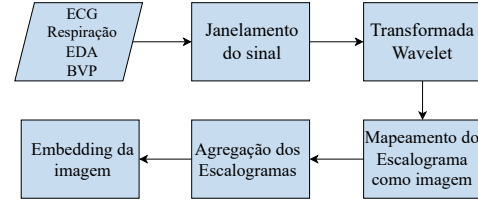


Fig. 3: Fluxograma da Aplicação da Transformada *wavelet* para Representação de Imagens de Sinais.

Cada sinal fisiológico é processado individualmente por meio do mesmo pipeline, exibido na Figura 3. Dessa forma, adotou-se a notação genérica $\mathbf{x} \in \{\mathbf{e}, \mathbf{c}, \mathbf{r}, \mathbf{b}\}$, em que: $\mathbf{e}$ representa o sinal EDA, $\mathbf{c}$ representa o ECG, $\mathbf{r}$ representa o sinal de respiração, e $\mathbf{b}$ representa o BVP. Dessa forma, $\mathbf{x}$ assume, em cada iteração, um dos quatro sinais fisiológicos disponíveis.

Seja $\mathbf{x}$ um sinal fisiológico capturado com frequência de amostragem constante $F_s = 400$ Hz. Esse sinal foi inicialmente segmentado em janelas temporais não sobrepostas de 60 segundos, originando um conjunto de subsequências denotadas por $\mathbf{x}^{(j)}$, onde $j \in \{1, 2, \dots, J\}$ representa o índice da janela.

Cada subsequência $\mathbf{x}^{(j)}$ possui comprimento fixo $N = F_s \times 60$, correspondente ao número de amostras em uma janela de 60 segundos. Assim, $\mathbf{x}^{(j)} = \{x^{(j)}[0], x^{(j)}[1], \dots, x^{(j)}[N-1]\}$, e constitui a entrada individual para a aplicação da TWC, que converte sinais unidimensionais em representações visuais no domínio tempo-frequência, evidenciando eventos transitórios e variações espectrais.

A TWC possibilita a análise de sinais discretizados ao substituir a integral contínua por uma somatória discreta, resultando em uma representação altamente redundante, porém informativa, no domínio tempo-frequência. Na prática, essa transformada é implementada por meio de uma convolução discreta entre o sinal de entrada e versões escaladas e transladadas da *wavelet* mãe.

Os dados de entrada para a aplicação computacional da TWC consistiram na janela de sinal $\mathbf{x}^{(j)}$, na frequência de amostragem F_s , e na escolha da *wavelet* mãe analítica. Neste estudo, adotou-se a *wavelet* Morse, configurada com parâmetro de simetria $\gamma = 3$ e produto tempo-largura igual a 60, reconhecida por sua elevada capacidade de localização espectral [28]. A TWC gera como saída uma matriz de coeficientes complexos $C \in \mathbb{C}^{S \times L}$, onde S representa o número de escalas (associadas à frequência F_s) e L é o número de amostras no tempo.

A visualização da matriz de coeficientes *wavelet* $|C|$ como uma imagem gera o que se denomina escalograma, que é capaz de representar a energia do sinal ao longo do tempo e das

frequências. A magnitude desses coeficientes, $|C|$, é mapeada como intensidade de cor para cada ponto da imagem, com os eixos representando, respectivamente, o tempo (em segundos) e a frequência (em Hz).

A redundância da representação gerada pela TWC pode implicar em elevado custo computacional para tarefas subsequentes de processamento e análise. Para mitigar esse problema, os coeficientes C foram redimensionados para uma matriz de resolução fixa 210×280 , por meio de interpolação bilinear. Essa etapa de redimensionamento foi fundamental para representar o escalograma como imagem de entrada para redes neurais convolucionais, mantendo a coerência espacial e realçando os padrões relevantes, ao mesmo tempo em que reduz a redundância visual e os custos computacionais associados.

A interpolação bilinear garante uma aproximação contínua e suave da matriz original, adaptando a dimensão temporal L e a dimensão de escala S ao novo formato da imagem. Como o eixo vertical (frequência) é exibido em escala logarítmica, o redimensionamento também considerou o espaçamento não linear das frequências para preservar informações importantes, especialmente aquelas concentradas nas faixas de baixa frequência. A Figura 4 apresenta os escalogramas obtidos para cada sinal fisiológico do experimento LAS.

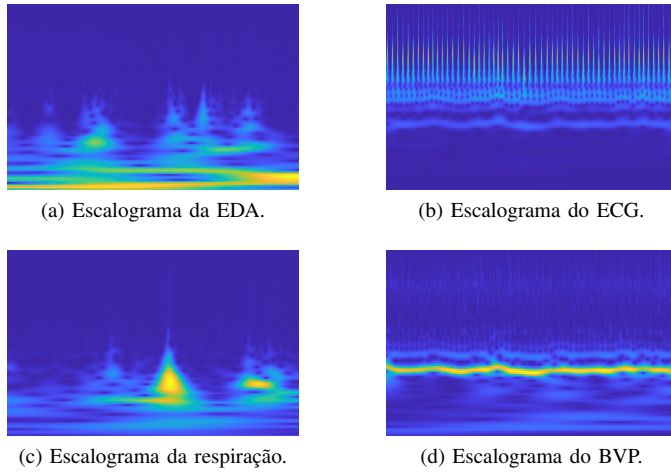


Fig. 4: Escalogramas resultantes da Transformada *wavelet* Contínua.

Após a obtenção dos escalogramas individuais, estes foram organizados em uma única imagem composta, na qual cada quadrante representa um dos sinais fisiológicos em diferentes escalas e localizações temporais, conforme processado pela transformada *wavelet* contínua. A Figura 5 ilustra essa representação agregada.

Como essas imagens foram utilizadas em técnicas de aprendizado não supervisionado, foi necessária uma etapa de *embedding* para redução de dimensionalidade. Essa transformação diminui ainda mais os custos computacionais e melhora a eficiência dos algoritmos de agrupamento, preservando as relações espaciais relevantes da imagem e permitindo uma

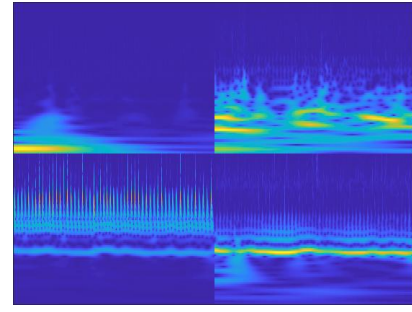


Fig. 5: Representação agregada dos escalogramas dos sinais fisiológicos para inferência de fadiga.

representação mais adequada das distribuições não lineares dos dados [29].

Para essa finalidade, utilizou-se a rede neural convolucional pré-treinada ResNet50, proposta por [30], devido ao seu desempenho satisfatório na extração de características robustas e generalizáveis em diferentes contextos visuais, incluindo imagens biomédicas [31], [32]. Neste estudo, a ResNet50 foi empregada sem ajuste fino, mantendo-se os pesos originais treinados no conjunto ImageNet. Essa abordagem simplifica o treinamento e preserva a capacidade da rede de extrair características genéricas, adequadas aos escalogramas analisados. Cada imagem resultou em um vetor de características com dimensão 1×2048 , totalizando 90 vetores por indivíduo durante a sessão de direção manual.

C. Inferência da Condição de Fadiga do Conductor

A Figura 6 apresenta o fluxo de atividades desenvolvido para a inferência do estado de fadiga do conductor.

Nesta análise, considerando que a manifestação da fadiga é subjetiva e que não há rótulos disponíveis para classificação supervisionada, foram adotadas técnicas de aprendizado não supervisionado com o objetivo de descrever e caracterizar os dados dos sinais fisiológicos por meio de agrupamento. Dessa forma, foram conduzidas análises exploratórias sobre os dados obtidos na simulação de direção, a fim de identificar padrões de comportamento e inferir possíveis estados de fadiga dos condutores.

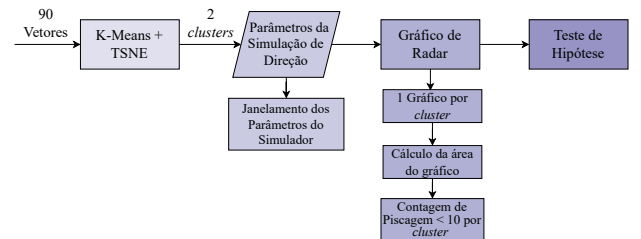


Fig. 6: Fluxograma ilustrando o fluxo de atividades para a inferência da condição de fadiga do condutor.

A Figura 6 apresenta o fluxo de atividades desenvolvido para a inferência da condição de fadiga do condutor, a partir

das representações extraídas dos sinais fisiológicos. Inicialmente, os 90 vetores gerados por indivíduo, provenientes da etapa de *embedding*, são submetidos ao algoritmo de agrupamento K-Means, com visualização em duas dimensões por meio do t-SNE. Os dados são agrupados em dois *clusters*, os quais são utilizados para resgatar os parâmetros da simulação de direção associados a cada janela temporal. Esses parâmetros são organizados e analisados individualmente por *cluster*, possibilitando a construção de gráficos de radar que representam o comportamento de direção médio de cada grupo. A seguir, é realizado o cálculo da área de cada gráfico, e, com base nos dados de rastreamento ocular, contabiliza-se a frequência de piscagens inferiores a 10 por minuto em cada grupo. Por fim, os resultados dessas análises são submetidos a um teste de hipótese, com o objetivo de investigar a relação entre o aumento da área do gráfico de radar e a maior incidência de indicadores fisiológicos de fadiga.

1) *Aprendizado de Máquina Não Supervisionado para Caracterização dos Dados*: A identificação do estado de fadiga do condutor é uma tarefa complexa, especialmente diante da ausência de rótulos explícitos que indiquem os momentos em que os participantes apresentaram fadiga ou baixa excitação fisiológica. Embora o experimento LAS forneça diversos parâmetros relevantes, não há anotações que caracterizem diretamente esses estados. Diante disso, adotaram-se técnicas de aprendizado de máquina não supervisionado, com o objetivo de explorar e caracterizar os dados fisiológicos a partir da identificação de padrões intrínsecos por meio de agrupamento baseado em similaridade, utilizando o algoritmo *k-means*.

Para a escolha do número ótimo de clusters k , foi adotado inicialmente o método do cotovelo, que avalia a soma dos erros quadráticos (*Sum of Squared Errors* – SSE) em função de diferentes valores de k . O ponto de inflexão da curva representa o valor ideal, onde aumentos adicionais em k resultam em ganhos marginais na redução da SSE [33]. A Figura 7 ilustra esse procedimento para os dados do ID 55.

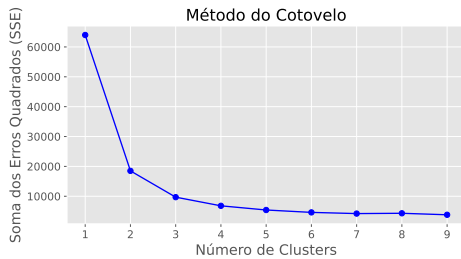


Fig. 7: Aplicação do método do cotovelo para o ID 55.

Com o objetivo de validar a escolha do número de grupos, utilizou-se o *silhouette score*, que mede o grau de separação entre clusters. Esse coeficiente varia de -1 a 1: valores próximos de 1 indicam boa separação, próximos de 0 indicam sobreposição entre grupos e negativos sugerem má atribuição. A média dos escores de silhueta foi usada para avaliar a qualidade global do agrupamento [33]. A distância adotada foi

a *cityblock* (Manhattan), com até 10 iterações por execução e 50 replicações independentes. A inicialização dos centróides foi realizada via método *k-means++*.

Para visualização dos agrupamentos, foi empregada a técnica t-SNE (*t-Distributed Stochastic Neighbor Embedding*), que reduz a dimensionalidade dos dados ao projetá-los em um espaço bidimensional. Essa técnica transforma as distâncias euclidianas em probabilidades condicionais, permitindo visualizar melhor as relações de proximidade e distância entre os dados [34].

2) *Análise Exploratória: Simulação de Direção e Movimento Ocular*: A literatura classifica os sistemas de detecção de fadiga em duas categorias principais: baseados no condutor e baseados no veículo [8]. Sistemas baseados no condutor utilizam sinais fisiológicos como EEG, ECG e EDA, além de características visuais, como movimentos oculares e expressões faciais. Por outro lado, sistemas baseados no veículo utilizam parâmetros como velocidade, aceleração, ângulo do volante e posição na pista, oriundos de simulações ou sensores embarcados.

Com base na literatura, foram selecionados cinco parâmetros do veículo para inferência de fadiga: desvio padrão da aceleração (STD A), da velocidade (STD V), do ângulo do volante (STD AV), das mudanças de faixa (MF) e da posição na pista (STD PP). Esses indicadores são reconhecidos por refletirem alterações no comportamento de direção sob fadiga [3], [35]–[37].

Seguindo a metodologia proposta por [38], os parâmetros de direção foram utilizados para construir gráficos de radar, cujas áreas representam o grau de variabilidade nos dados. Um aumento na área do polígono é interpretado como maior instabilidade comportamental associada à fadiga. Cada gráfico foi gerado a partir da média dos parâmetros em janelas de 1 minuto, sincronizadas com as janelas geradas para a representação dos sinais fisiológicos.

Os dados dos parâmetros foram normalizados no intervalo $[0, 1]$ e divididos segundo os clusters obtidos pelo algoritmo *k-means* [39]. Para cada cluster, foi construído um gráfico de radar, e a área total foi calculada pela soma das áreas dos cinco triângulos formados pelos eixos do gráfico, utilizando a lei dos senos [38].

No que se refere aos movimentos oculares, o número de piscadas por minuto (PPM) foi utilizado como indicador de fadiga, tendo em vista que a taxa de piscadas tende a diminuir com o aumento do cansaço [20]. Utilizou-se a coluna *Blink*, que representa um identificador crescente para cada piscada detectada. Para cada janela de 1 minuto, foi contada a quantidade de piscadas, e em seguida, calculou-se a porcentagem de janelas com $PPM < 10$ dentro de cada cluster, gerando o indicador $\%PPM < 10$.

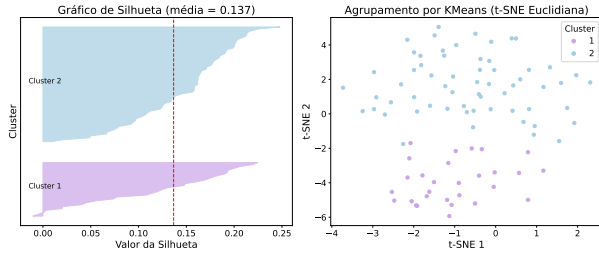
3) *Formulação de Hipótese*: Com base nas evidências comportamentais, propôs-se a hipótese nula H_0 : *quanto maior a área do gráfico de radar para um dado cluster, maior é o nível de fadiga do condutor, evidenciado por uma maior proporção de janelas com $PPM < 10$* . Para verificar essa hipótese, compararam-se os valores de área dos gráficos de

radar e os valores de $\%PPM < 10$ entre os dois clusters identificados para cada participante.

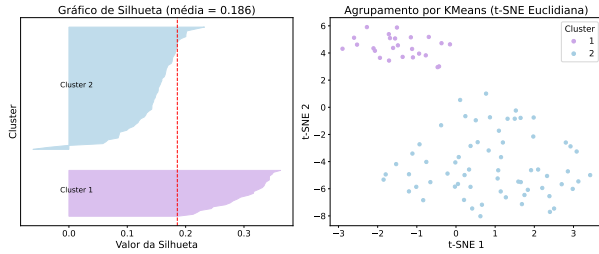
IV. RESULTADOS E DISCUSSÃO

Como em geral, o método de cotovelo apontou dois valores ideais para k , os resultados obtidos com a aplicação do k -means para estes, $k = 2$ e $k = 3$ indicaram melhor desempenho médio dos escores de silhueta para $k = 2$, com valores variando entre 0,1270 e 0,4380. Apesar de não elevados, tais valores demonstraram separação superior à obtida com três grupos, justificando a escolha de dois clusters para os experimentos subsequentes.

A Figura 8 apresenta a visualização dos agrupamentos com t-SNE e os respectivos gráficos de silhueta para os IDs 55 e 57.



(a) ID 55.



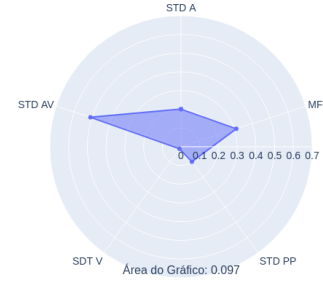
(b) ID 57.

Fig. 8: Visualização dos agrupamentos com t-SNE e gráficos de silhueta para $k = 2$.

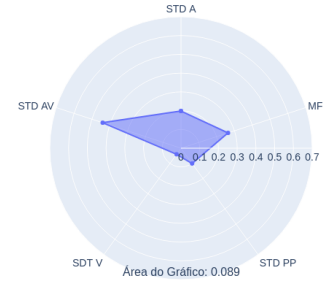
Após a definição dos dois clusters, cada vetor foi associado ao respectivo intervalo temporal de criação do escalograma, utilizando a marcação de tempo presente nos datasets (TIMESTAMP). Esses intervalos foram posteriormente utilizados para recuperar os dados correspondentes da simulação de direção e do conjunto de registros de movimento ocular, permitindo a análise comportamental de cada grupo em termos de variáveis externas relacionadas à fadiga.

As Figuras 9 e 10 apresentam os gráficos de radar construídos para os clusters dos IDs 55 e 57, respectivamente. Cada eixo representa um dos cinco parâmetros veiculares, e a distância ao centro indica sua magnitude normalizada.

A Tabela I apresenta os casos em que a hipótese H_0 foi confirmada, ou seja, o cluster com maior área de gráfico de radar também apresentou maior $\%PPM < 10$, indicando maior fadiga. Já a Tabela II exibe os casos em que essa relação não foi observada.

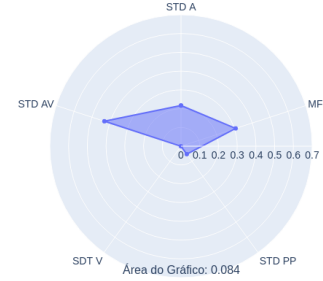


(a) Cluster 1.

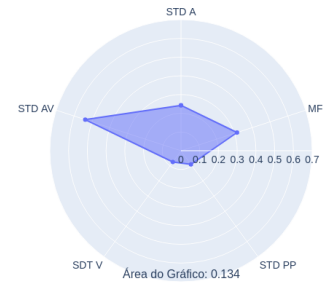


(b) Cluster 2.

Fig. 9: Gráficos de radar para o ID 55.



(a) Cluster 1.



(b) Cluster 2.

Fig. 10: Gráficos de radar para o ID 57.

TABLE I: Casos em que a Hipótese Nula (H_0) é Confirmada.

ID	Média da Silhouette	Cluster 1 Qnt (% PPM < 10)	Área do Gráfico do Cluster 1	Cluster 2 Qnt (% PPM < 10)	Área do Gráfico do Cluster 2
ID_53	0.2014	19 (31.5%)	0.1131	71 (14%)	0.0937
ID_55	0.2780	29 (96.5%)	0.0967	61 (65.5%)	0.0892
ID_57	0.3627	25 (36%)	0.0837	65 (44.6%)	0.1340
ID_62	0.1767	73 (63%)	0.1688	17 (76.4%)	0.2003
ID_63	0.4380	38 (7.9%)	0.1550	52 (75%)	0.1551
ID_64	0.2003	64 (6.2%)	0.1767	26 (19.2%)	0.2105
ID_66	0.1716	43 (34.9%)	0.1764	47 (17%)	0.1426
ID_68	0.1716	40 (55%)	0.1509	50 (48%)	0.1316
ID_69	0.2259	60 (81.6%)	0.1313	30 (86.6%)	0.1356
ID_82	0.2040	33 (12.1%)	0.0992	57 (17.5%)	0.1709
ID_83	0.2630	34 (5.8%)	0.1405	56 (3.6%)	0.1349
ID_91	0.2135	27 (96%)	0.2529	63 (79%)	0.2327

TABLE II: Casos em que a Hipótese Nula (H_0) é Rejeitada.

ID	Média da Silhouette	Cluster 1 Qnt (% PPM < 10)	Área do Gráfico do Cluster 1	Cluster 2 Qnt (% PPM < 10)	Área do Gráfico do Cluster 2
ID_61	0.1794	46 (21.1%)	0.2096	44 (6.8%)	0.2592
ID_72	0.2410	50 (12%)	0.1678	40 (30%)	0.1119
ID_73	0.1714	39 (25.6%)	0.1292	51 (27.4%)	0.1023
ID_85	0.3932	64 (1.5%)	0.1620	26 (42.3%)	0.1173

Para os IDs 74, 75, 76, 78 e 80, não foram observadas janelas com PPM < 10, inviabilizando a validação da hipótese.

Nos 12 indivíduos da Tabela I, foi observada uma correlação positiva entre o aumento da área do gráfico e maior ocorrência de baixa taxa de piscadas, sugerindo níveis mais altos de fadiga. Nos quatro indivíduos da Tabela II, essa relação não se confirmou, o que pode estar relacionado à limitações no parâmetro utilizado como verdade de base ou à subjetividade da fadiga.

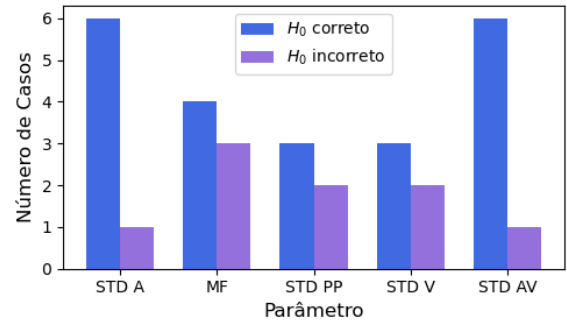
Pesquisas indicam que, embora o número de piscadas possa diminuir com a fadiga, a duração das piscadas tende a aumentar. Assim, é recomendável incluir indicadores complementares, como o PERCLOS (*Percent Eye Closure*) ou movimentos de cabeça e expressões faciais [20].

Por fim, foi realizada uma análise dos parâmetros veiculares que mais contribuíram para a distinção entre clusters. Para cada indivíduo, identificaram-se os dois parâmetros com maior diferença média entre os grupos. A Figura 11 apresenta o ranqueamento dos parâmetros mais influentes.

O gráfico indica que STD A e STD AV foram os parâmetros mais relevantes, com maior frequência de acerto na validação de H_0 . Os parâmetros STD V e STD PP apresentaram influência moderada, enquanto MF mostrou impacto maior nos casos de erro, sugerindo que os dois primeiros são mais confiáveis para inferência de fadiga com base em comportamento de direção.

V. CONCLUSÃO

Este trabalho propôs uma metodologia para a inferência de estados de fadiga em motoristas por meio da análise de sinais fisiológicos multimodais obtidos em simulações de direção. A abordagem envolveu a conversão dos sinais em representações visuais utilizando a Transformada de *wavelet*

Fig. 11: Ranqueamento dos parâmetros veiculares mais influentes na inferência da hipótese H_0 .

Contínua, seguidas por extração de características com redes neurais convolucionais e análise de agrupamento não supervisionado. Os resultados indicam que a combinação entre parâmetros veiculares, movimento ocular e análise de escalogramas é promissora para identificar variações comportamentais relacionadas à fadiga. A validação por meio de radar charts e taxas de piscagem revelou uma correlação significativa para a maioria dos participantes, reforçando o potencial da abordagem adotada.

Como trabalhos futuros, pretende-se explorar outras técnicas de agrupamento mais robustas, como DBSCAN e HDBSCAN, que podem capturar estruturas mais complexas nos dados. Além disso, busca-se aprimorar os indicadores de validação da fadiga com o uso de métricas complementares, como o PERCLOS, movimentações de cabeça e expressões faciais, possibilitando uma análise mais abrangente e personalizada do estado do condutor. A integração de múltiplos parâmetros e métodos, considerando abordagens personalizadas, permanece

um caminho promissor para o desenvolvimento de sistemas eficazes de monitoramento do motorista.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte e financiamento.

REFERENCES

- [1] World Health Organization (WHO), "Road traffic injuries," Dec. 2023, fact sheet. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>
- [2] N. M. Zanon and L. D. F. R. Brisotto, "Comportamento de risco e a contribuição da psicologia para a redução dos acidentes de trânsito: uma revisão da literatura," *Revista Destaques Acadêmicos*, vol. 12, no. 2, 2020.
- [3] W. Hailin, L. Hanhui, and S. Zhumei, "Fatigue driving detection system design based on driving behavior," in *2010 international conference on optoelectronics and image processing*, vol. 1. IEEE, 2010, pp. 549–552.
- [4] G. Fountas, S. S. Pantangi, K. F. Hulme, and P. C. Anastasopoulos, "The effects of driver fatigue, gender, and distracted driving on perceived and observed aggressive driving behavior: A correlated grouped random parameters bivariate probit approach," *Analytic methods in accident research*, vol. 22, p. 100091, 2019.
- [5] A. Giorgi, V. Ronca, A. Vozzi, P. Aricò, G. Borghini, R. Capotorto, L. Tamborra, I. Simonetti, S. Sportiello, M. Petrelli *et al.*, "Neurophysiological mental fatigue assessment for developing user-centered artificial intelligence as a solution for autonomous driving," *Frontiers in Neurobotics*, vol. 17, p. 1240933, 2023.
- [6] T. Arakawa, R. Hibi, and T.-a. Fujishiro, "Psychophysical assessment of a driver's mental state in autonomous vehicles," *Transportation research part A: policy and practice*, vol. 124, pp. 587–610, 2019.
- [7] S. A. El-Nabi, W. El-Shafai, E.-S. M. El-Rabaie, K. F. Ramadan, F. E. Abd El-Samie, and S. Mohsen, "Machine learning and deep learning techniques for driver fatigue and drowsiness detection: a review," *Multimedia Tools and Applications*, vol. 83, no. 3, pp. 9441–9477, 2024.
- [8] M. H. Alkinani, W. Z. Khan, and Q. Arshad, "Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges," *IEEE Access*, vol. 8, pp. 105 008–105 030, 2020.
- [9] J. Rahul, M. Sora, and L. D. Sharma, "An overview on biomedical signal analysis," *Int. J. Recent Technol. Eng.*, vol. 7, no. 5, pp. 206–209, 2019.
- [10] A. Gómez-Echavarría, J. P. Ugarte, and C. Tobón, "The fractional fourier transform as a biomedical signal and image processing tool: A review," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 1081–1093, 2020.
- [11] A. K. Singh and S. Krishnan, "Ecg signal feature extraction trends in methods and applications," *BioMedical Engineering OnLine*, vol. 22, no. 1, p. 22, 2023.
- [12] Hadrian, "Hadrian Project — Holistic Approach for Driver Role Integration and Automation Allocation for European Mobility Needs — hadrianproject.eu," <https://hadrianproject.eu/>, 2019, [Accessed 29-07-2024].
- [13] N. Ganapathy, R. Swaminathan, and T. M. Deserno, "Deep learning on 1-d biosignals: a taxonomy-based survey," *Yearbook of medical informatics*, vol. 27, no. 01, pp. 098–109, 2018.
- [14] Y. M. Belousov, N. Elkin, V. Man'ko, E. Popov, and S. Revenko, "Tomographic representation of electrocardiogram signals," *Journal of Russian Laser Research*, vol. 39, pp. 302–313, 2018.
- [15] R. Sánchez-Reolid, F. L. de la Rosa, M. T. Lopez, and A. Fernández-Caballero, "One-dimensional convolutional neural networks for low/high arousal classification from electrodermal activity," *Biomedical Signal Processing and Control*, vol. 71, p. 103203, 2022.
- [16] B. Nakisa, M. N. Rastgoo, A. Rakotonirainy, F. Maire, and V. Chandran, "Automatic emotion recognition using temporal multimodal deep learning," *IEEE Access*, vol. 8, pp. 225 463–225 474, 2020.
- [17] A. Alam, S. Urooj, and A. Q. Ansari, "Human emotion recognition models using machine learning techniques," in *2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON)*. IEEE, 2023, pp. 329–334.
- [18] S. Sheykhivand, T. Y. Rezaii, Z. Mousavi, S. Meshgini, S. Makouei, A. Farzamnia, S. Danishvar, and K. Teo Tze Kin, "Automatic detection of driver fatigue based on eeg signals using a developed deep neural network," *Electronics*, vol. 11, no. 14, p. 2169, 2022.
- [19] H. Garg, "Drowsiness detection of a driver using conventional computer vision application," in *2020 international conference on power electronics & IoT applications in renewable energy and its control (PARC)*. IEEE, 2020, pp. 50–53.
- [20] B. Fatima, A. R. Shahid, S. Ziauddin, A. A. Safi, and H. Ramzan, "Driver fatigue detection using viola jones and principal component analysis," *Applied Artificial Intelligence*, vol. 34, no. 6, pp. 456–483, 2020.
- [21] P. Moertl, "Human-systems integration for driving automation systems: Holistic approach for driver role integration and automation allocation for european mobility needs," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 78–88.
- [22] L. L. D. Stasi, F. Angioi, M. A. C. Fernandes, M. J. Caurcel-Cara, C. Prat, J. Sodnik, and C. Diaz-Piedra, "Physiological indices directory," 2 2024. [Online]. Available: https://plus.figshare.com/articles/dataset/Physiological_indices_directory/22277626
- [23] —, "Driving simulator indices directory," 2 2024. [Online]. Available: https://plus.figshare.com/articles/dataset/Driving_simulator_indices_directory/22293385
- [24] —, "Eye movement directory," 2 2024. [Online]. Available: https://plus.figshare.com/articles/dataset/Eye_movement_directory/22285084
- [25] —, "Annotations directory," 2 2024. [Online]. Available: https://plus.figshare.com/articles/dataset/Annotations_directory/22285732
- [26] J. M. Lilly, "Element analysis: A wavelet-based method for analysing time-localized events in noisy time series," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2200, p. 20160776, 2017.
- [27] A.-H. Najmi, J. Sadowsky *et al.*, "The continuous wavelet transform and variable resolution time-frequency analysis," *Johns Hopkins apl technical digest*, vol. 18, no. 1, pp. 134–140, 1997.
- [28] A. Shalhaf, S. Bagherzadeh, and A. Maghsoudi, "Transfer learning with deep convolutional neural network for automated detection of schizophrenia from eeg signals," *Physical and Engineering Sciences in Medicine*, vol. 43, pp. 1229–1239, 2020.
- [29] Z. Li, L. Yao, S. Wang, S. Kanhere, X. Li, and H. Zhang, "Adaptive two-dimensional embedded image clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, 2020, pp. 4796–4803.
- [30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [31] V. Vinaykumar, J. A. Babu, and J. Frnda, "Optimal guidance whale optimization algorithm and hybrid deep learning networks for land use land cover classification," *EURASIP Journal on Advances in Signal Processing*, vol. 2023, no. 1, p. 13, 2023.
- [32] A. Rath, B. S. P. Mishra, and D. K. Bagal, "Resnet50-based deep learning model for accurate brain tumor detection in mri scans," *Next Research*, vol. 2, no. 1, p. 100104, 2025.
- [33] C. Shi, B. Wei, S. Wei, W. Wang, H. Liu, and J. Liu, "A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm," *EURASIP journal on wireless communications and networking*, vol. 2021, pp. 1–16, 2021.
- [34] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. 11, 2008.
- [35] H. Du, X. Zhao, X. Zhang, Y. Zhang, and J. Rong, "Effects of fatigue on driving performance under different roadway geometries: a simulator study," *Traffic injury prevention*, vol. 16, no. 5, pp. 468–473, 2015.
- [36] J. Lowrie and H. Brownlow, "The impact of sleep deprivation and alcohol on driving: a comparative study," *BMC public health*, vol. 20, pp. 1–9, 2020.
- [37] L. N. Boyle, J. Tippin, A. Paul, and M. Rizzo, "Driver performance in the moments surrounding a microsleep," *Transportation research part F: traffic psychology and behaviour*, vol. 11, no. 2, pp. 126–136, 2008.
- [38] M. B. D. d. Silva, "Uma metodologia orientada a fluxo de dados para modelagem do comportamento de motoristas," Ph.D. dissertation, 2022.
- [39] B. Chong *et al.*, "K-means clustering algorithm: a brief review," *Academic Journal of Computing & Information Science*, vol. 4, no. 5, pp. 37–40, 2021.