

Visão Computacional para Detecção de Fadiga de Motoristas em Áreas de Pesquisa Mineral

1st Wherbert Silva

Programa de Pós-graduação em
Computação Aplicada
Universidade Federal do Pará
Parauapebas-PA
wherbert.silva@tucurui.ufpa.br

2nd Daniel Pinheiro

Programa de Pós-graduação em
Computação Aplicada
Universidade Federal do Pará
Tucurú-PA
dpinheiro@ufpa.br

3rd Otávio Teixeira

Programa de Pós-graduação em
Computação Aplicada
Universidade Federal do Pará
Tucurú-PA
otaviont@ufpa.br

4th Vigner Santos

Programa de Pós-graduação em
Computação Aplicada
Universidade Federal do Pará
Tucurú-PA
vigner.santos@tucurui.ufpa.br

Resumo—Este trabalho apresenta um sistema inteligente para detecção de fadiga em motoristas com base em técnicas de Visão Computacional e aprendizado de máquina, aplicado em operações de pesquisa mineral da Vale S.A. O sistema utiliza modelos YOLO para detecção facial e técnicas baseadas em pontos faciais (MediaPipe e dlib) para extração de indicadores fisiológicos, como EAR, MAR e inclinação da cabeça. O treinamento foi realizado com imagens reais coletadas em operações de sondagem, permitindo maior adaptação ao ambiente. Os resultados mostram que a combinação de pré-filtragem facial com YOLO e extração de pontos faciais aumenta a precisão da detecção de sinais visuais de fadiga, como olhos semicerrados, bocejos e fechamento ocular.

Palavras-chave—Fadiga de motorista, pesquisa mineral, Visão Computacional

I. INTRODUÇÃO

A fadiga de motoristas é um dos principais fatores de risco em rodovias e ambientes industriais, contribuindo para a ocorrência de acidentes [1], [6]. No contexto da mineração, este risco é potencializado por longas jornadas de trabalho, terrenos irregulares e tarefas repetitivas. A aplicação de técnicas de Visão Computacional para detectar sinais faciais de fadiga, como fechamento ocular prolongado, bocejos e inclinação da cabeça, vem ganhando destaque nos últimos anos [4], [13].

Entretanto, métodos convencionais treinados com bases genéricas apresentam limitações quando aplicados a cenários operacionais reais, sujeitos a variações de iluminação, posicionamento do condutor e reflexos. Este estudo apresenta um sistema desenvolvido e testado com dados reais coletados em operações de sondagem da Vale S.A., abordando estratégias para aumentar a precisão e robustez da detecção em tempo real.

Este trabalho contou com o apoio da Diretoria de Geociências e Sondagem da Vale S.A., no âmbito das atividades de Pesquisa e Desenvolvimento realizadas no contexto do projeto.

II. TRABALHOS RELACIONADOS

Estudos recentes têm utilizado modelos YOLO (You Only Look Once), uma das arquiteturas de Visão Computacional mais conhecidas e eficientes para detecção de padrões em imagens e vídeos em tempo real, aplicando-a à detecção facial e à análise de sinais oculares com elevado desempenho [2], [4], [7].

O YOLO, é construído sobre avanços de ponta em aprendizado profundo e visão computacional, oferecendo

desempenho incomparável em termos de velocidade e precisão[11].

Na Tabela I, compararmos 6 trabalhos, onde há predominância de redes neurais convolucionais, especialmente YOLOv5, v8 e v10, para detecção de fadiga em tempo real.

TABELA I - COMPARATIVO ENTRE ESTUDOS

Autor	Trabalho	Técnica Principal	Indicadores	Acurácia
Xie et al. (2023)	Revolutionizing Road Safety: YOLOv8-Powered Driver Fatigue Detection	YOLOv8 + PERCLOS + POM	Abertura ocular, bocejo	96,9% (tempo real, 11,4 ms/frame)
Guo et al. (2022)	Monitoring and Detection of Driver Fatigue from Monocular Cameras Based on YOLOv5	YOLOv5	Olhos e boca	42 FPS (GTX1650), alta precisão
Dijaya et al. (2023)	Facial Fatigue Detection in High-Risk Occupational Environments: Leveraging YOLOv4 for Enhanced Worker Safety	YOLOv4	Olhos cansados, bocejo	mAP 98,3%
Tang & Guo (2024)	Fatigue Driving Detection Methods Based on Drivers Wearing Sunglasses	YOLOv8 + Transfer Learning	Olhos fechados, bocejo (com óculos escuros)	>98%
Li et al. (2022)	Driver Fatigue Detection Method Based on Human Pose Information Entropy	Pose + Entropia da Informação	Entropia da pose	Não especificado
Liu et al. (2024)	A Review of Driver Fatigue Detection: Progress and Prospect	Revisão	Vários	-

Fonte: Criação própria, tal como apresento tabela acima.

Apesar dos avanços observados nos estudos recentes, as abordagens propostas apresentam limitações quanto à generalização e aplicabilidade em ambientes operacionais no contexto da pesquisa mineral. Xie et al. (2023) obtiveram alta acurácia (96,9%) ao integrar YOLOv8, PERCLOS e POM, contudo, o modelo foi treinado em bases públicas com

condições controladas de iluminação e ângulo, não refletindo as variações inerentes a contextos industriais. Guo et al. (2022) aprimoraram a velocidade de processamento com YOLOv5 (42 FPS), porém restringiram-se à detecção de olhos e boca, sem tratar interferências externas como vibração ou oclusões. O método de Dijaya et al. (2023), baseado em YOLOv4, alcançou elevado desempenho em ambientes laboratoriais (mAP 98,3%), mas com conjunto de dados reduzido (500 imagens) e sem validação em campo. Já Tang e Guo (2024) introduziram um modelo adaptado para condutores utilizando óculos escuros, demonstrando robustez com YOLOv8 e transfer learning, ainda assim, o enfoque restrito à detecção facial limita sua integração com dados de telemetria ou variáveis de comportamento veicular. Li et al. (2022) propuseram uma abordagem baseada na entropia de pose humana, mitigando o problema das oclusões faciais, mas apresentando maior complexidade computacional e carência de análise em tempo real. Por fim, Liu et al. (2024), em sua revisão sistemática, destacou que os sistemas de detecção atuais carecem de modelos híbridos capazes de combinar múltiplas fontes de informação, como Visão Computacional, sinais fisiológicos e dados de condução, de modo integrado e adaptável a condições adversas.

III. METODOLOGIA

A. Modelos baseados em YOLO

Foram conduzidos experimentos com versões YOLOv10 (Nano, Small, Large e X) e YOLOv11-S. O primeiro utilizou imagens de repositórios abertos, como: Images encontradas em sites de busca e imagens simuladas.

Após a aquisição das imagens, foi montado o dataset e anotações das imagens para identificação de olhos abertos, olhos sonolento, olhos fechados, bocejos, atento e cabeça baixa.



Fig. 1. Ilustração dos delimitadores adicionados nas imagens do dataset
Fonte: Criação própria, tal como apresento na figura acima.

A variante YOLOv10-S apresentou melhor desempenho geral (96,1%) com boa separação entre classes visuais, vide tabela II.

TABELA II - COMPARAÇÃO DE ACURÁCIA DOS MODELOS YOLOV10

M. YOLO V10	Ac Geral %	Atento %	Bocej o %	Cabeç a Baixa %	Olkos Abertos %	Olkos Fechados %	Sonolent o %
L	95.8	99	100	98	92	97	91
N	95.6	100	100	97	95	95	92
S	96.1	99	100	98	97	97	90
X	96.0	99	99	95	98	96	86

Fonte: Criação própria, tal como apresento tabela acima.

Em seguida, um modelo usando YOLO V10 “S” foi treinado com 1.646 imagens coletadas do sistema de monitoramento veicular da Vale, representando melhor o contexto real.

A eficiência do modelo pode ser avaliada matriz de confusão abaixo.

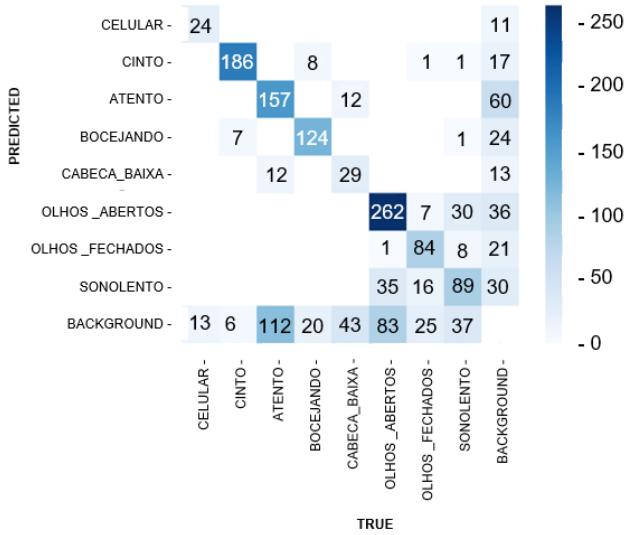


Fig. 2. Matriz de confusão, modelo YOLO com imagens do sistema de monitoramento de fadiga da Vale

Fonte: Criação própria, tal como apresento na figura acima.

Houve melhora na detecção de estados visuais, embora persistisse confusão entre “sonolento” e “olhos fechados”, relacionada à resolução temporal e variações de iluminação. Pensando em projetos futuros foram anotados no modelo, classes para detecção de cinto de segurança e celular.

B. Extração de pontos faciais com MediaPipe e dlib

Para aprimorar a mensuração dos estados de identificação, olhos abertos, olhos sonolentos, olhos fechados, bocejos, atenção e cabeça baixa, foram utilizadas técnicas baseadas em proporções geométricas extraídas de pontos faciais. Destacam-se o EAR (Eye Aspect Ratio), amplamente empregado para detecção de piscadas e fechamento prolongado dos olhos [10]; o MAR (Mouth Aspect Ratio), utilizado para identificação de bocejos; e a inclinação vertical da cabeça, variável que complementa os sinais oculares e tem sido associada a estados de sonolência e fadiga postural [13].

- MediaPipe Face Mesh: estima até 468 pontos faciais 3D [9], o que possibilitou fazermos os cálculos para definição de EAR (Proporção de Aspecto dos Olhos), MAR (Proporção de Aspecto da Boca) e inclinação vertical da cabeça. Em análise empírica, o desempenho foi consistente usando as imagens do sistema da Vale.
- dlib: Apresenta boa precisão em condições ideais, permitindo medições com base em pontos da testa, nariz, queixo e lábios [8]. Entretanto, no teste de desempenho realizado, mostrou menor estabilidade em vídeos reais, além de custo computacional mais alto.

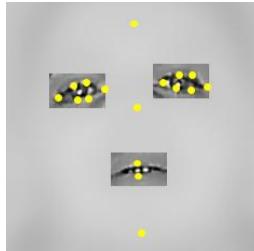


Fig. 3. Resultado da detecção com a técnica Face Mesh [aplicado tratamento na imagem para preservar a identidade do motorista].

Fonte: Criação própria, tal como apresento na figura acima.

Durante os testes com vídeos reais provenientes do sistema de monitoramento veicular da Vale, observou-se que, em determinadas amostras, a detecção dos pontos faciais apresentava interrupções momentâneas ou, em alguns casos, falhas completas na identificação, resultando em quadros sem pontos faciais válidos.

C. Pré-filtragem Facial com YOLO

Para contornar falhas na detecção de pontos faciais em vídeos reais, foi adotada uma etapa de pré-filtragem facial com YOLO V11 “S” personalizado, treinado com 4.884 imagens, desempenho na matriz de confusão abaixo.

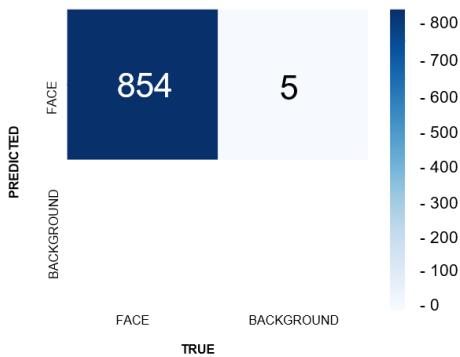


Fig. 4. Matriz de confusão, modelo YOLO para pré-filtragem

Fonte: Criação própria, tal como apresento na figura acima.

Uma vez estabelecido um processo de pré-filtragem, que consiste em delimitar a região de interesse (ROI) da face com alta precisão, enviando somente essa região para o MediaPipe.

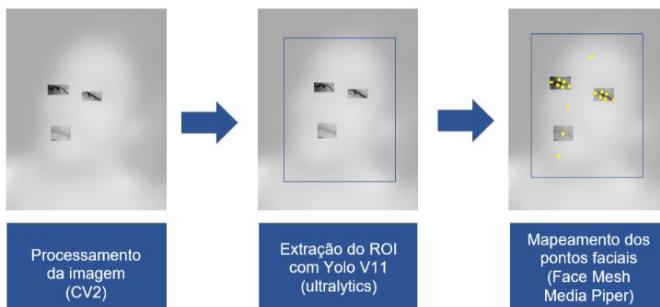


Fig. 5. Processamento das imagens usando a filtragem do ROI com YOLO [aplicado tratamento na imagem para preservar a identidade do motorista].

Fonte: Criação própria, tal como apresento na figura acima.

IV. RESULTADOS

A avaliação quantitativa contemplou métricas de acurácia e análises empíricas de desempenho, permitindo verificar a eficácia dos métodos em cenários operacionais reais. A Tabela III sintetiza os principais resultados obtidos.

TABELA III - COMPARATIVO DE DESEMPENHO ENTRE TÉCNICAS DE DETECCÃO E RASTREAMENTO FACIAL EM CENÁRIOS OPERACIONAIS

Técnica	Acurácia Geral	Pontos Fortes	Limitações Principais
YOLOv10 - Base Genérica	96,1%	Alta acurácia em classes visuais distintas (ex: olhos abertos/fechados)	Baixa generalização em ambiente real; confusão entre classes visuais
YOLOv10 - Imagens da Vale	95,8%	Melhor adaptação ao cenário real; maior precisão em classe "atento"	Persistente confusão entre "sonolento" e "olhos fechados"
MediaPipe Face Mesh	-	Detecção estável de pontos faciais; ideal para extração de EAR, MAR e PERCLOS	Sensível à iluminação e à centralização do rosto
dlib	-	Alta precisão em condições ideais (luz frontal, rosto centralizado)	Elevado custo computacional; falhas recorrentes em vídeos reais
YOLO - Detecção facial (ROI)	>99% (854/859)	Redução de falsos negativos na localização da face	-

Fonte: Criação própria, tal como apresento tabela acima.

Os modelos YOLOv10-S e YOLOv11-S treinados com imagens reais apresentaram desempenho superior aos modelos baseados em bases genéricas. A filtragem facial reduziu falhas na extração de pontos faciais, aumentando a estabilidade do sistema em condições adversas. O uso combinado de YOLO + MediaPipe mostrou-se mais robusto que abordagens unimodais, especialmente para sinais sutis como olhos semicerrados.

V. CONCLUSÃO

Modelos de detecção de sinais de fadiga treinados com dados representativos do ambiente operacional, apresentaram desempenho superior na identificação de sinais visuais de fadiga. A integração entre a filtragem facial baseada em YOLO e a extração de pontos faciais via MediaPipe Face Mesh, permitiu, mesmo sob condições adversas típicas da mineração, detectar de forma aceitável, indicadores faciais em amostras de vídeos onde os condutores fecharam os olhos por fadiga crítica.

Para trabalhos futuros, indica-se a exploração de técnicas de suavização temporal para aumentar a estabilidade das inferências e o desenvolvimento de modelos de inferência capazes de quantificar gradualmente os níveis de fadiga dos motoristas, ampliando a aplicabilidade do sistema em cenários industriais complexos.

REFERÊNCIAS

- [1] T. Bauerle, S. M. Smith, and M. R. Ross, “Mineworker fatigue: A review of what we know and future decisions,” *Professional Safety*, vol. 70, no. 3, pp. 33–40, 2018.
- [2] S. Xie, Y. Li, and C. Zhang, “Revolutionizing road safety: YOLOv8-powered driver fatigue detection,” in *Proc. IEEE i-COSTE*, 2023.
- [3] R. Dijaya, A. R. Fajaresta, I. Anshory, and R. A. Sukmono, “Facial fatigue detection in high-risk occupational environments: Leveraging YOLOv4 for enhanced worker safety,” in *Proc. ICONNIC*, 2023.
- [4] Z. Guo, H. Wang, and L. Zhao, “Monitoring and detection of driver fatigue from monocular cameras based on YOLOv5,” in *Proc. IEEE CVCI*, pp. 428–433, 2022.
- [5] Ultralytics, “Train custom YOLO models,” 2024. [Online]. Disponível: <https://docs.ultralytics.com>
- [6] A. B. S. Silva, M. R. Oliveira, and P. C. Ferreira, “Fadiga e seu impacto na morbimortalidade de motoristas profissionais,” *Saúde, Ética & Justiça*, vol. 29, no. 2, 2024.
- [7] X.-X. Tang and P.-Y. Guo, “Fatigue driving detection methods based on drivers wearing sunglasses,” *IEEE Access*, vol. 12, pp. 70946–70959, 2024. DOI: 10.1109/ACCESS.2024.3394218.
- [8] D. E. King, “Dlib-ml: A machine learning toolkit,” *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, 2009.
- [9] Google, “MediaPipe Face Mesh,” 2024. [Online]. Disponível: <https://mediapipe.readthedocs.io>
- [10] T. Soukupová and J. Čech, “Real-time eye blink detection using facial landmarks,” *Proc. CVWW*, 2016.
- [11] Ultralytics, “YOLOv10 – Real-Time Object Detection,” 2024. [Online]. Disponível: <https://docs.ultralytics.com>
- [12] Trimble Transportation, “Fadiga no trânsito: entenda os riscos e como prevenir acidentes,” 2024.
- [13] W. Zhao, H. Liu, and Q. Chen, “Multimodal driver fatigue detection based on attention fusion,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 8, pp. 9212–9225, 2023.
- [14] Roboflow, “Getting started with Roboflow,” 2024. [Online]. Disponível: <https://roboflow.com>
- [15] T. Li, T. Zhang, Y. Zhang, and L. Yang, “Driver fatigue detection method based on human pose information entropy,” *Journal of Advanced Transportation*, vol. 2022, Article ID 7213841, pp. 1–15, May 2022. DOI: 10.1155/2022/7213841.
- [16] F. Liu, X. Li, T. Lv, and F. Xu, “A review of driver fatigue detection: Progress and prospect,” *IEEE Access*, vol. 12, pp. 70960–70982, 2024. DOI: 10.1109/ACCESS.2024.3394219.