

DETECÇÃO DE DISTRAÇÃO DE MOTORISTAS EM TEMPO REAL USANDO DEEP LEARNING EM DISPOSITIVOS MÓVEIS¹

Carlos Gil Martins da Silva², Geraldo P. Rocha Filho³

RESUMO

Este trabalho apresenta o desenvolvimento de um aplicativo móvel que utiliza visão computacional para detectar distrações em tempo real. O aplicativo emprega um modelo de *Deep Learning* baseado na arquitetura *You Only Look Once* (YOLO), treinado a partir de um conjunto de dados híbrido com imagens categorizadas em quatro comportamentos de risco: (i) uso de celular, (ii) operação do rádio, (iii) ingestão de bebidas; e (iv) alcançar objetos no banco traseiro. A avaliação no conjunto de teste demonstrou eficácia, alcançando 91,6% de precisão e 83,5% de revocação. O desempenho na detecção do risco “usando celular” foi superior, com 96,9% de revocação, em comparação com a classe “operando o rádio”, que alcançou 66,7%. O estudo conclui que a abordagem é viável, resultando em uma ferramenta funcional que traduz o desempenho do modelo em alertas para o condutor.

PALAVRAS-CHAVE: Aplicativo Móvel, Detecção de Objetos, Distração de Motoristas.

REAL-TIME DRIVER DISTRACTION DETECTION USING DEEP LEARNING ON MOBILE DEVICES

ABSTRACT

This paper presents the development of a mobile application that uses computer vision to detect distractions in real time. The application employs a Deep Learning model based on the You Only Look Once (YOLO) architecture, trained on a hybrid dataset with images categorized into four risk behaviors: (i) cell phone use; (ii) radio use; (iii) drinking; and (iv) reaching for objects in the backseat. Evaluation on the test set demonstrated effectiveness, achieving 91.6% precision and 83.5% recall. Performance in detecting the risk category "using cell phone" was superior, with 96.9% recall, compared to the category "operating radio," which achieved 66.7%. The study concludes that the

¹ FAPESB - Fundação de Amparo à Pesquisa do Estado da Bahia

² Discente do Curso de Ciência da Computação, Universidade Estadual do Sudoeste da Bahia (UESB).

³ Docente do Curso de Ciência da Computação, Universidade Estadual do Sudoeste da Bahia (UESB).

approach is viable, resulting in a functional tool that translates the model's performance into alerts for the driver.

KEYWORDS: Mobile Application, Object Detection, Driver Distraction.

INTRODUÇÃO

A segurança viária representa um grave desafio global, com estatísticas da Organização Mundial de Saúde apontando para 1,19 milhão de óbitos anuais em acidentes de trânsito (World Health Organization, 2023). A base deste problema reside no fator humano, responsável por mais de 96% das colisões, com a distração do condutor sendo elemento de risco crítico (ABOJARADEH, 2023). Este cenário reflete-se diretamente no Brasil, onde a falta de atenção é a principal causa de acidentes com vítimas em rodovias federais (ANDRADE; ANTUNES, 2020). Agravado pelo uso crescente de *smartphones*, o impacto da desatenção vai além dos óbitos, reforçando a urgência por tecnologias de monitoramento preventivo.

Para enfrentar esse desafio, o presente trabalho apresenta um aplicativo móvel que utiliza Inteligência Artificial para promover uma condução mais segura. O aplicativo emprega visão computacional e *Deep Learning*, consolidando a análise por imagem como uma das abordagens mais promissoras da área (BOUHSISSIN; SAEL; BENABBOU, 2023). Especificamente, utiliza-se a arquitetura de detecção de objetos em tempo real YOLO para analisar o fluxo de vídeo da câmera de um *smartphone* e identificar padrões de distração. Ao detectar um comportamento de risco, o aplicativo fornece um alerta imediato, apresentando-se como uma ferramenta acessível que opera em um *smartphone*, com o objetivo de mitigar as distrações e contribuir para a reeducação de hábitos dos condutores.

MATERIAIS E MÉTODOS

O desenvolvimento do modelo baseou-se em um conjunto de dados consolidado a partir de três fontes: o acervo público *State Farm Distracted Driver Detection* (KAGGLE, 2016), o *benchmark Drive&Act* (MARTIN-GUTIERREZ et al., 2019) e uma coleta de imagens em cenário real, totalizando 4.063 amostras. O dataset foi dividido em quatro classes de distração, sendo elas: 'C0 - Usando celular' com 1.971 imagens, 'C2 - Bebendo' com 1.003 imagens, 'C1 - Operando Rádio' com 601 imagens e 'C3 - Alcançando Atrás' com 488 imagens. O preparo destes dados incluiu a anotação manual com caixas delimitadoras (*bounding boxes*), a inclusão de imagens negativas para redução de falsos positivos, e a aplicação de um pipeline de *data augmentation* com transformações fotométricas e geométricas para aumentar a generalização do modelo.

A metodologia de detecção utiliza a arquitetura YOLOv11n (JOCHER; QIU, 2024), escolhida por seu balanço otimizado entre precisão e eficiência computacional, ideal para a execução em dispositivos de borda. O treinamento seguiu uma abordagem de aprendizado por transferência, iniciando com pesos pré-treinados e, subsequentemente, realizando um *fine-tuning* com o dataset de dados coletados em um cenário real e uma taxa de aprendizado reduzida. Após a validação, o modelo treinado foi convertido para o formato TensorFlow Lite e integrado ao aplicativo, desenvolvido com o framework Flutter. A inferência local no smartphone é executada por meio da biblioteca *ultralytics_yolo*, permitindo a detecção em tempo real.

RESULTADOS E DISCUSSÃO

A avaliação final do modelo, realizada no conjunto de teste com dados de cenário real, demonstrou um bom desempenho, conforme detalhado na Tabela 1. O modelo alcançou uma precisão geral de 91,6% e a revocação de 83,5% que demonstram que o aplicativo é confiável, com uma baixa taxa de falsos positivos, e abrangente, sendo capaz de identificar a grande maioria das distrações reais.

A análise por classe evidencia a capacidade do modelo em aprender ações distintas. A distração "C0 - Usando celular" atingiu 96,9% tanto em precisão quanto em revocação. Em contraste, a classe "C1 - Operando Rádio" mostrou-se mais desafiadora, com uma revocação de 66,7%. Essa diferença sugere que a variabilidade e a sutileza visual de interagir com o rádio são mais difíceis de generalizar em comparação com o gesto claro de segurar um celular. Ainda assim, os resultados gerais validam o aplicativo como uma ferramenta viável para o monitoramento de distrações ao volante.

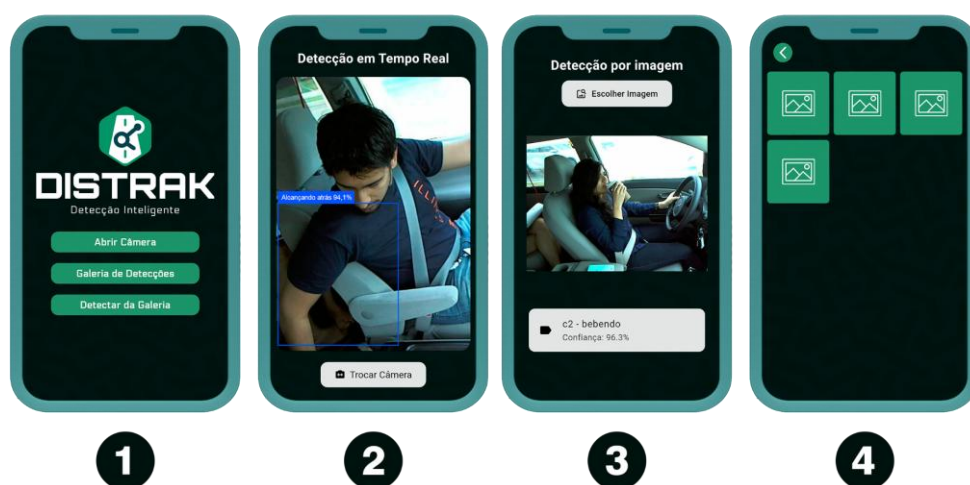
TABELA 1: Resultados detalhados por classe na validação

Classe	Precisão	Revocação
Todas	0.916	0.835
C0 - Usando celular	0.969	0.969
C1 - Operando Rádio	0.878	0.667
C2 - Bebendo	1.000	0.848
C3 - Alcançando Atrás	0.817	0.857

Fonte: Do autor.

A aplicação prática do modelo é demonstrada pela interface do aplicativo, ilustrada na Figura 1. A tela principal (Figura 1, Rótulo 1) serve como um menu para as funcionalidades principais: a detecção em tempo real (Figura 1, Rótulo 2), que sobrepõe na imagem da câmera a classe da distração e a confiança do modelo, e a análise de imagens estáticas selecionadas pelo usuário (Figura 1, Rótulo 3). Todas as detecções confirmadas são armazenadas em uma galeria interna (Figura 1, Rótulo 4), permitindo ao condutor consultar seu histórico. A interface traduz o desempenho do modelo, apresentado anteriormente, em uma ferramenta interativa e de fácil usabilidade.

FIGURA 1: Principais telas da interface do aplicativo



Fonte: Do autor.

CONCLUSÕES/CONSIDERAÇÕES

O presente trabalho desenvolveu um aplicativo de detecção de distrações de motoristas baseado na arquitetura YOLOv11n. Os resultados obtidos demonstraram a eficácia do modelo, que desempenhou na identificação de comportamentos de risco críticos. A principal contribuição do trabalho é a materialização de um aplicativo de baixo custo e acessível, capaz de funcionar em um *smartphone*, oferecendo um feedback ao condutor para a prevenção de acidentes e promoção da segurança viária.

Observou-se como limitação uma menor performance na detecção de gestos mais sutis, como "operando o rádio", indicando um desafio em ações de alta variabilidade visual. Como trabalhos futuros, pretende-se a expansão do conjunto de dados com foco nessas classes de menor desempenho para aprimorar a generalização. Além disso, o aplicativo pode ser evoluído para incorporar novas funcionalidades, como a detecção de sonolência, e submetido a estudos de usabilidade em cenários reais para avaliar seu impacto prático no comportamento dos motoristas, consolidando sua contribuição para um trânsito mais seguro.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] ABOJARADEH, M. Analysis of driver distraction behaviour causing risk of accidents in Jordan. European Transport/Trasporti Europei, p. 1–15, 03 2023.

[2] ANDRADE, F. R. d.; ANTUNES, J. L. F. Falta de atenção ao conduzir veículo automotor como causa de acidentes de trânsito nas rodovias federais brasileiras. Revista Brasileira de Epidemiologia, Associação Brasileira de Saúde Coletiva, v. 23, p. e200085, 2020. ISSN 1415-790X. Disponível em: <<https://doi.org/10.1590/1980-549720200085>>.

[3] BOUHSISSIN, S.; SAEL, N.; BENABBOU, F. Driver behavior classification: A systematic literature review. IEEE Access, v. 11, p. 14128–14153, 2023.

[4] JOCHER, G.; QIU, J. Ultralytics YOLO11. 2024. Disponível em: <<https://github.com/ultralytics/ultralytics>>.

[5] MARTIN, M. et al. Drive&act: A multi-modal dataset for fine-grained driver behavior recognition in autonomous vehicles. In: The IEEE International Conference on Computer Vision (ICCV). [S.l.: s.n.], 2019.

[6] MONTOYA, A. et al. State Farm Distracted Driver Detection. 2016. <<https://kaggle.com/competitions/state-farm-distracted-driver-detection>>. Kaggle.

[7] World Health Organization. Global status report on road safety 2023. Geneva, 2023. Disponível em: <<https://www.who.int/publications/i/item/9789240086517>>.