

ALGORITMO DE DETECÇÃO DE RACHADURAS NO CONCRETO ASFÁLTICO BETUMINOSO POR MEIO DE VISÃO COMPUTACIONAL

Abdul Rahman Kassem Kassem
União Dinâmica de Faculdades Cataratas (UDC)
Foz do Iguaçu, Brasil
dinhorkk25@gmail.com

Dr. Miguel Diogenes Matrakas
União Dinâmica de Faculdades Cataratas (UDC)
Foz do Iguaçu, Brasil
mdmatrakas@gmail.com

Dra. Alessandra Bussador
União Dinâmica de Faculdades Cataratas (UDC)
Foz do Iguaçu, Brasil
alebussador@gmail.com

Abstract—This study explores the use of computer vision techniques to automate the detection of cracks in digital asphalt images, aiming to improve the maintenance of Brazilian highways. Poor road conditions not only increase operational costs for transportation companies but also compromise user safety, raising the risk of accidents. By implementing an automated system for early crack detection, it is possible to carry out preventive interventions, avoiding more severe damage to the pavement and reducing maintenance costs over time. The research proposes the development of a prototype using the YOLO algorithm for crack detection, laying the groundwork for future implementations. Although the creation of a diversified image database and the practical implementation of the prototype have yet to be accomplished, the presented methodology provides a solid foundation for future work. The next steps will involve building the image database, training the model, and evaluating its effectiveness.

Keywords—Computer Vision; YOLO; Crack Detection.

Resumo—Este estudo explora o uso de técnicas de visão computacional para automatizar a detecção de rachaduras em imagens digitais de asfalto, com o objetivo de melhorar a manutenção das rodovias brasileiras. A má conservação das estradas não apenas aumenta os custos operacionais para as empresas de transporte, mas também compromete a segurança dos usuários, elevando o risco de acidentes. Com a implementação de um sistema automatizado para a detecção precoce de rachaduras, é possível realizar intervenções preventivas, evitando danos mais graves ao pavimento e reduzindo os custos de manutenção ao longo do tempo. A pesquisa propõe o desenvolvimento de um protótipo utilizando o algoritmo YOLO para a detecção de rachaduras, estabelecendo as bases para futuras implementações. Embora a criação de um banco de imagens diversificado e a implementação prática do protótipo ainda estejam por ser realizadas, a metodologia apresentada fornece uma base sólida para trabalhos futuros.

Os próximos passos envolverão a construção do banco de imagens, o treinamento do modelo e a avaliação de sua eficácia.

Palavras-chave—Visão Computacional; YOLO; Detecção de Rachaduras.

I. INTRODUÇÃO

O Brasil possui a maior concentração rodoviária de transporte de cargas e passageiros entre as principais economias mundiais, sendo que 58% do transporte total é realizado por rodovias. A manutenção dessas rodovias é crucial, pois afeta diretamente a economia, a competitividade e a segurança viária do país. No entanto, uma pesquisa realizada pela Confederação Nacional do Transporte (CNT) revela que 67,5% das rodovias brasileiras apresentam problemas, classificadas como regulares, ruins ou péssimas. Esses problemas incluem irregularidades como rachaduras, buracos e desníveis, que aumentam os riscos de acidentes e comprometem a segurança dos motoristas [1] [2].

A pesquisa busca explorar o uso de técnicas de visão computacional para automatizar a detecção de rachaduras em imagens digitais de asfalto, com o objetivo de melhorar a manutenção das rodovias brasileiras.

A má conservação das estradas, além de aumentar os custos operacionais para as empresas de transporte, compromete a segurança dos usuários, elevando o risco de acidentes. A adoção de um sistema automatizado para a detecção precoce dessas rachaduras permite a realização de intervenções preventivas,

evitando danos mais graves ao pavimento e, conseqüentemente, reduzindo os custos de manutenção ao longo do tempo.

O objetivo principal deste estudo é desenvolver um protótipo que utilize técnicas de visão computacional para identificar rachaduras em imagens digitais de asfalto, contribuindo para a manutenção preventiva das rodovias. Para isso, será criado um banco diversificado de imagens de asfalto com rachaduras, que servirá para treinar o algoritmo responsável pela detecção. A eficácia do modelo será testada em diferentes condições viárias, a fim de garantir sua precisão e aplicabilidade em cenários reais.

II. FUNDAMENTAÇÃO TEÓRICA

A. Tipos de Rachaduras

As rachaduras podem assumir diferentes formas e direções, cada uma com suas características e implicações. A seguir, são descritos os dois tipos mais comuns de rachaduras encontradas em pavimentos [3]:

- **Rachadura Transversal:** Rachadura isolada que se forma perpendicularmente ao eixo da via. Com extensão de até 100 cm, é classificada como transversal curta, e acima de 100 cm, transversal longa.
- **Rachadura Longitudinal:** Trinca isolada paralela ao eixo da via. Classificada como longitudinal curta se tiver até 100 cm e longitudinal longa se a extensão ultrapassar 100 cm.

Os exemplos desses tipos de rachaduras estão na Figura 4.



Fig. 1. Rachadura Longitudinal e Transversal [3]

B. Visão Computacional

A visão computacional, um subcampo da inteligência artificial e do *machine learning*, visa capacitar computadores a entenderem dados visuais de forma detalhada, imitando a visão humana. Essa habilidade permite que máquinas interpretem imagens e vídeos, utilizando essas informações para tomar decisões automatizadas em diversas aplicações, sem intervenção humana [4].

Embora o campo seja ativo atualmente, suas raízes datam da década de 1950, com um desenvolvimento significativo nos anos 1970, quando começaram os esforços para combinar

visão computacional com inteligência artificial. No entanto, a complexidade da visão humana tornou claro que a replicação dessa capacidade em máquinas seria um desafio maior do que o inicialmente previsto. As aplicações da visão computacional são vastas, abrangendo setores como indústria, medicina, segurança e veículos autônomos [5].

C. Processamento Digital de Imagem

Processamento digital de imagem (PDI) é uma disciplina que se concentra na análise e manipulação de imagens digitais por meio de computadores. Um aspecto essencial do processamento de imagem é entender que uma imagem digital é formada por um número finito de elementos, conhecidos como píxeis, cada um com uma localização e valor específicos [6].

O processamento pode ser organizada em três níveis: baixo, médio e alto. Os processos de baixo nível incluem operações básicas, como a redução de ruído, melhoria de contraste e nitidez da imagem. Os processos de médio nível envolvem tarefas mais complexas, como a segmentação, descrição e classificação de objetos dentro de uma imagem. Já os processos de alto nível focam na compreensão e interpretação de conjuntos de objetos reconhecidos, especialmente em contextos de análise de imagem e visão computacional [6].

D. Redes Neurais Convolucionais

O aprendizado profundo, também conhecido como *deep learning*, é uma subcategoria do aprendizado de máquina que capacita modelos computacionais a desenvolverem múltiplas camadas de processamento para aprender representações de dados com diversos níveis de abstração. Um dos tipos mais utilizados de *deep Learning* são as redes neurais convolucionais [7].

As redes neurais convolucionais (CNN), representam uma arquitetura treinável inspirada biologicamente, capaz de aprender características invariantes. Elas consistem em uma série de camadas que extraem características de imagens de entrada por meio de convoluções sucessivas e redimensionamento, culminando na identificação da classe à qual a imagem pertence. A CNN é uma rede neural que implementa várias camadas distintas, sendo que as principais são: camada de convolução, de Pooling e totalmente conectadas ou Fully-connected. A Figura 1 ilustra as três principais camadas da CNN [7].

- **Camada Convulucional:** A camada convolucional tem por funcionalidade extrair atributos dos volumes de entradas.
- **Camada de Pooling:** A camada de Pooling é responsável por reduzir a dimensionalidade do volume resultante após as camadas convolucionais, ajudando a tornar a representação invariante a pequenas translações na entrada.

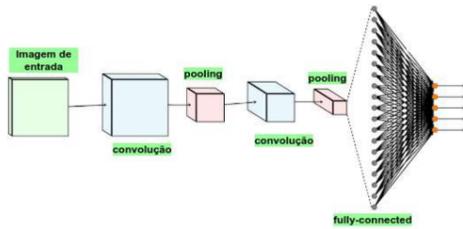


Fig. 2. Arquitetura CNN [7]

- **Camada Totalmente Conectada:** a camada totalmente conectada é responsável pela propagação do sinal por meio da multiplicação ponto a ponto e uso de uma função de ativação.

E. Detecção de Objetos em Imagens

A detecção de objetos é uma área desafiadora dentro da visão computacional. Utilizando técnicas de aprendizado profundo, que se tornaram populares no início dos anos 2000, a detecção de objetos procura identificar e localizar precisamente objetos em imagens ou vídeos, atribuindo-lhes rótulos de classe apropriados. Essa técnica tem várias aplicações, como previsão de valores de ações, reconhecimento de fala e detecção de intrusos [8].

Segundo [8], os principais passos de implementação para detecção de objetos são:

- **Coleta e anotação de dados:** Recolher imagens ou vídeos com objetos rotulados, incluindo classe e coordenadas das caixas delimitadoras.
- **Pré-processamento de dados:** Preparar os dados para treinamento realizando tarefas nos dados.
- **Seleção de um modelo:** Escolher um modelo de detecção adequado com base em precisão, velocidade e recursos computacionais.
- **Treinamento do modelo:** Treinar o modelo no conjunto de dados rotulado, ajustando parâmetros para minimizar a função de perda.
- **Validação e teste:** Validar o modelo em dados separados e ajustar hiperparâmetros. Testar o modelo final para avaliar sua generalização.
- **Implantação:** Implantar o modelo treinado em produção, otimizando para velocidade de inferência e compatibilidade com a plataforma de hardware e software.

Além da detecção de objetos, os modelos realizam tarefas adicionais, como classificação, localização e segmentação. A classificação envolve atribuir um rótulo de classe a uma imagem ou a uma região específica de interesse. A localização identifica a posição do objeto desenhando uma caixa delimitadora ao seu redor. A segmentação de instâncias identifica os pixels que

pertencem a um objeto, criando uma máscara de nível de pixel que delinea sua forma. Como mostrado na Figura 2 [8].

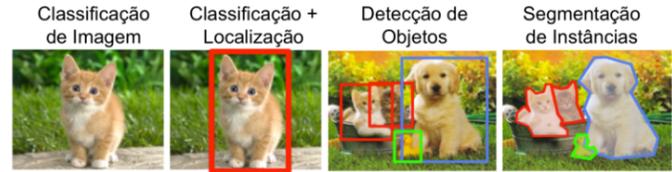


Fig. 3. Classificação, Localização e Segmentação [9]

F. YOLO

O *You Only Look Once* (YOLO) é uma abordagem na detecção de objetos que utiliza uma rede neural end-to-end para prever caixas delimitadoras e probabilidades de classe simultaneamente.

Como o YOLO funciona?

- 1) O YOLO começa dividindo a imagem em uma grade de células. Nas primeiras versões, essa grade era de 13x13, totalizando 169 células. Cada célula prevê até 5 caixas delimitadoras, resultando em um total de 845 caixas (13x13x5), cada uma analisando uma parte da imagem [10].
- 2) Cada caixa delimitadora é um retângulo que contém a localização de um objeto e pode variar em tamanho. Durante a detecção, o YOLO atribui uma pontuação de confiança a cada caixa, indicando a certeza de que a caixa contém um objeto, mas sem identificar o tipo de objeto. A espessura das bordas das caixas destaca a pontuação de confiança atribuída a cada uma: quanto maior a confiança de que um objeto está contido, mais espessa é a borda da caixa [10].
- 3) Além da detecção, o YOLO classifica os objetos em cada caixa, atribuindo uma probabilidade para cada classe treinada. A pontuação final combina a confiança da caixa com a probabilidade da classe. Para melhorar a precisão, pode-se definir um limiar de confiança (por exemplo, 70%) para filtrar caixas com baixa confiança, mantendo apenas as mais confiáveis, como ilustrado na Figura 3 [10].

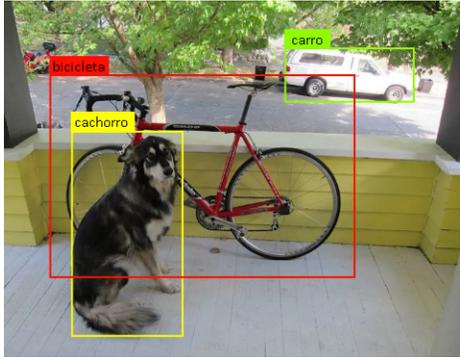


Fig. 4. Resultado da Detecção [10]

G. Matriz de Confusão

Segundo [11], a matriz de confusão é uma tabela que mostra os acertos e erros do modelo em comparação com os resultados esperados.

H. Aumento de Dados

O aumento de dados, é uma técnica que melhora o desempenho de modelos de aprendizado de máquina com dados limitados, criando versões adicionais do conjunto original para aumentar a eficiência dos algoritmos [12].

III. METODOLOGIA

A pesquisa envolve a coleta de imagens digitais de asfalto para o desenvolvimento de um sistema de detecção de rachaduras. As imagens serão obtidas por meio de captura de vídeo com uma câmera de celular de 48 megapixels instalada em um veículo. Os vídeos capturados serão transformados em imagens no formato .jpg para análise.

Para melhorar a robustez do modelo, serão aplicadas técnicas de aumento de dados usando a biblioteca OpenCV. Isso incluirá transformações nas imagens existentes para criar uma diversidade que aumente a capacidade do modelo de generalizar para novas imagens de rachaduras. Após o aumento de dados, o conjunto de 2000 imagens será dividido em 80% para treinamento e 20% para teste.

O treinamento do modelo será realizado utilizando o algoritmo YOLO, conhecido por sua eficiência na detecção de objetos. O processo incluirá a rotulagem manual das imagens com o software LabelMe para destacar as rachaduras, configuração do modelo YOLO com ajustes específicos para a detecção de rachaduras, e o treinamento do modelo com essas imagens rotuladas.

Após o treinamento, o modelo será avaliado usando o conjunto de imagens de teste, com a análise de métricas de desempenho e a geração de uma matriz de confusão para medir a eficácia na detecção de rachaduras.

IV. CONCLUSÃO

A pesquisa realizada destacou a importância da manutenção eficiente das rodovias brasileiras, considerando os impactos econômicos e de segurança associados à deterioração das vias. O estudo explora o potencial das técnicas de visão computacional, como o algoritmo YOLO, para automatizar a detecção de rachaduras em imagens digitais de asfalto, com o objetivo de melhorar a manutenção preventiva das rodovias. A criação de um banco de imagens diversificado e a implementação prática do protótipo estão em desenvolvimento, com a metodologia proposta estabelecendo uma base sólida para futuras etapas do projeto.

Como trabalho futuro, será necessário concluir o treinamento do modelo YOLO com as imagens rotuladas e realizar a validação com os dados de teste para avaliar seu desempenho na detecção de rachaduras. Além disso, será importante otimizar o modelo por meio de ajustes de hiperparâmetros. A continuidade deste trabalho é crucial para garantir que as técnicas desenvolvidas possam ser aplicadas de forma eficaz.

REFERÊNCIAS

- [1] G1, “Por que o Brasil depende tanto do transporte rodoviário?” 5 2018. [Online]. Available: <https://g1.globo.com/economia/noticia/por-que-o-brasil-depende-tanto-do-transporte-rodoviario.ghtml>
- [2] —, “Veja ranking das piores e melhores rodovias do país, segundo confederação nacional do transporte,” 11 2023. [Online]. Available: <https://11nk.dev/FiveL>
- [3] CNT, “Conheça os 13 principais defeitos do pavimento das rodovias,” 2 2018. [Online]. Available: <https://www.cnt.org.br/agencia-cnt/conheca-principais-defeitos-pavimento>
- [4] M. Gonçalves, T. Marques, P. D. Gaspar, V. N. Soares, and J. M. Caldeira, “Road pavement damage detection using computer vision techniques: Approaches, challenges and opportunities,” *Revista de Informatica Teorica e Aplicada*, vol. 30, pp. 22–35, 10 2023.
- [5] D. Milano and L. Honorato, “Visão computacional,” p. 7, 2010.
- [6] R. C. Gonzalez and R. E. R. E. Woods, *Digital image processing*. Pearson, 12 2018.
- [7] M. V. Ferreira, K. M. D. S. Oliveira, A. O. C. Filho, and A. Dalília, “Deep learning uma introdução às redes neurais,” p. 21, 2017.
- [8] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, “Statistical analysis of design aspects of various yolo-based deep learning models for object detection,” *International Journal of Computational Intelligence Systems*, vol. 16, 12 2023.
- [9] DSA, “Segmentação de imagens médicas com deep learning,” 3 2024. [Online]. Available: <https://blog.dsacademy.com.br/segmentacao-de-imagens-medicas-com-deep-learning/>
- [10] Piemontez, “Yolo para detecção de objetos – visão geral,” 9 2022. [Online]. Available: <https://visaocomputacional.com.br/yolo-para-deteccao-de-objetos-visao-geral/>
- [11] V. Rodrigues, “Métricas de avaliação: acurácia, precisão, recall... quais as diferenças?” 4 2019. [Online]. Available: <https://vitorborbarodrigues.medium.com/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-acur%C3%A1cia-precis%C3%A3o-recall-quais-as-diferen%C3%A7as-c8f05e0a513c>
- [12] P. Ostwal, “Data augmentation for computer vision,” 2 2023. [Online]. Available: <https://pranjal-ostwal.medium.com/data-augmentation-for-computer-vision-b88b818b6010>