

ANÁLISE DE PATOLOGIA DE PAVIMENTOS ATRAVÉS DE FOTOS CAPTURADAS POR DRONE UTILIZANDO IA

Universidade Presbiteriana Mackenzie

Coordenadoria de Fomento à Pesquisa - Pró-reitoria de Pesquisa e
Pós-Graduação

Breno Panarelli Wanderley
Apoio: PIVIC Mackenzie

Patrícia Barboza da Silva

Unidade São Paulo

10239072@mackenzista.com.br
patricia.silva1@mackenzie.br

2025

RESUMO

Este trabalho investiga a aplicação de técnicas de *Machine Learning*, com ênfase no modelo YOLO (*You Only Look Once*), para a identificação de patologias em pavimentos a partir de imagens capturadas por drones ou outras fontes visuais. O objetivo é reduzir a necessidade de interdições, minimizar transtornos e otimizar custos na avaliação das condições das vias. O modelo foi desenvolvido em Python, utilizando a plataforma Google Colab, o que possibilitou a execução de treinamentos e experimentações com recursos de aceleração por GPU. A base de dados foi composta por imagens obtidas no Roboflow, garantindo diversidade e ampliando a capacidade de generalização do modelo. A rede neural convolucional implementada permitiu a detecção automática de defeitos, apoiando o monitoramento das condições do pavimento. Os testes indicaram desempenho satisfatório na identificação de patologias em imagens complexas, demonstrando potencial para aplicações em manutenção preditiva e gestão de infraestrutura viária. Contudo, melhorias adicionais são necessárias para aprimorar a precisão do reconhecimento em cenários variados.

Palavras-chave: Pavimentos. Machine Learning. Drones. Redes Neurais Convolucionais. YOLO. Patologias.

ABSTRACT

This study investigates the application of *Machine Learning* techniques, focusing on the YOLO (*You Only Look Once*) model, for detecting pavement distresses from images captured by drones or other visual sources. The aim is to reduce the need for road closures, minimize disruptions, and optimize costs in road condition assessment. The model was developed in Python using the Google Colab platform, enabling training and experimentation with GPU acceleration. The dataset was composed of images obtained from Roboflow, ensuring diversity and enhancing the model's generalization capabilities. The implemented convolutional neural network enabled the automatic detection of defects, supporting pavement condition monitoring. Test results showed satisfactory performance in identifying distresses in complex images, demonstrating potential for predictive maintenance and infrastructure management applications. However, further improvements are needed to increase detection accuracy in diverse scenarios.

Keywords: Pavement. Machine Learning. Drones. Convolutional Neural Networks. YOLO. Pathologies.

1 INTRODUÇÃO

De acordo com a Confederação Nacional dos Transportes (CNT, 2023), aproximadamente 67,5% das rodovias brasileiras apresentam condições regulares, ruins ou péssimas, enquanto apenas 32,5% estão classificadas como boas ou ótimas. A degradação dos pavimentos compromete a segurança e a eficiência do tráfego, tornando essencial a realização de inspeções periódicas para a identificação de patologias e a definição de estratégias de manutenção.

Atualmente, a avaliação das condições dos pavimentos é majoritariamente realizada de forma manual, o que demanda tempo, apresenta riscos à segurança dos inspetores e pode resultar em diagnósticos subjetivos. Além disso, a necessidade de interdições parciais ou totais das vias impacta negativamente a mobilidade urbana e o transporte de cargas, gerando custos adicionais e transtornos para a sociedade.

A utilização de novas tecnologias, como drones e inteligência artificial, surge como alternativa viável para otimizar esse processo. O emprego de drones permite a captura rápida e eficiente de imagens de pavimentos sem a necessidade de interrupções no tráfego, enquanto modelos de aprendizado de máquina possibilitam a análise automatizada dessas imagens, garantindo maior precisão na identificação de patologias.

Neste contexto, este estudo investiga a aplicação de redes neurais convolucionais para a identificação de defeitos em pavimentos a partir de imagens aéreas. A abordagem adotada busca aprimorar a manutenção preditiva, contribuindo para a gestão mais eficiente da infraestrutura viária. Apesar dos desafios técnicos e da necessidade de aperfeiçoamento dos modelos de reconhecimento, os resultados obtidos demonstram o potencial dessas tecnologias para transformar as práticas tradicionais de inspeção e monitoramento de rodovias.

2 REFERENCIAL TEÓRICO

2.1 Pavimentos

De acordo com Bernucci et al. (2022), a pavimentação tem como principal objetivo garantir conforto e segurança aos usuários das vias, distribuindo de maneira eficiente as cargas transmitidas pelos veículos. Os pavimentos podem ser classificados como flexíveis, semirrígidos ou rígidos. Pavimentos flexíveis são aqueles que apresentam deformação elástica sob cargas aplicadas, enquanto os rígidos são compostos por camadas que atuam principalmente sob tração, utilizando placas de concreto Portland. Já os semirrígidos possuem características intermediárias entre os dois tipos anteriores.

A correta concepção e manutenção do pavimento são fundamentais para sua durabilidade e funcionalidade. Erros de projeto, falhas na execução, tráfego acima do previsto e falta de conservação podem comprometer sua estrutura, reduzindo sua vida útil.

2.2 Patologias em Pavimentos

As patologias em pavimentos asfálticos representam formas de deterioração que podem ocorrer desde os primeiros anos de uso até fases avançadas de sua vida útil. Elas decorrem de fatores como fadiga estrutural, envelhecimento do ligante, sobrecarga de tráfego, variações térmicas, falhas construtivas e problemas de drenagem (BERNUCCI et al., 2022; CNT, 2018). A identificação correta desses defeitos é essencial para adoção de estratégias de manutenção eficientes e econômicas.

Entre as manifestações mais comuns estão as fissuras e trincas, que podem ser transversais, longitudinais, interligadas (couro de jacaré) ou em blocos, cada uma associada a causas específicas como retração térmica, endurecimento do ligante, fadiga estrutural e variações climáticas (CNT, 2018; RIBEIRO, 2017). A evolução dessas trincas, quando não tratadas, pode gerar cavidades e buracos.

As deformações incluem afundamentos — plásticos ou por consolidação —, ondulações e corrugações, geralmente provocadas por instabilidade da mistura, compactação insuficiente ou esforços tangenciais intensos (CNT, 2018; NÚCLEO DO CONHECIMENTO, 2020). Outras patologias relevantes são o escorregamento do revestimento, causado por falhas na ligação entre camadas (DNIT, 2003), a exsudação, resultante do excesso de ligante (CNT, 2018), e a desagregação superficial, associada à perda de agregados por falhas de adesividade ou envelhecimento do ligante (CNT, 2018; ALÉM DA INÉRCIA, 2018).

As panelas, ou buracos, representam estágios avançados de degradação, demonstrado na figura 1, geralmente originados pela evolução de trincas de fadiga e infiltração de água, agravados pela ausência de manutenção e por falhas construtivas (CNT, 2018; DNIT, 2003).



Figura 1 – Buraco em Pavimento.

Fonte: ONSV, 2023.

2.3 Inspeção de Pavimentos

A inspeção de pavimentos é uma etapa fundamental para garantir a durabilidade e a segurança das vias, permitindo a identificação precoce de defeitos e a adoção de medidas corretivas adequadas. Essa avaliação fornece subsídios técnicos para o planejamento de ações de manutenção preventiva e corretiva, auxiliando na gestão dos recursos disponíveis e na otimização do ciclo de vida do pavimento (BERNUCCI et al., 2022; OLIVEIRA et al., 2021).

De acordo com a norma DNIT 008/2003 – PRO, a inspeção da superfície de pavimentos flexíveis e semirrígidos pode ser realizada por meio de métodos visuais ou por tecnologias mais avançadas (DNIT, 2003). O levantamento visual, amplamente utilizado no Brasil, consiste na observação direta do pavimento por técnicos treinados, que registram os tipos, a severidade e a extensão das patologias presentes. Embora esse método seja de baixo custo e permita uma avaliação imediata, ele apresenta limitações como a subjetividade do avaliador, a suscetibilidade a erros humanos e o elevado tempo necessário para a cobertura de longos trechos rodoviários (CNT, 2018; ROCHA et al., 2020).

2.4 Redes Neurais Convolucionais na Identificação de Patologias

Redes Neurais Convolucionais (RNCs) são um modelo de inteligência artificial especializado na análise de imagens. Inspiradas no funcionamento do córtex visual humano, essas redes são compostas por camadas convolucionais que extraem características dos dados visuais, permitindo a detecção automática de padrões.

Estudos recentes demonstram que as RNCs são eficazes na identificação de patologias em pavimentos. No estudo de Rosa et al. (2020), um modelo de RNC foi capaz de identificar corretamente 76% das patologias analisadas. Já a pesquisa de Nóbrega (2023) aplicou RNCs na classificação de imagens dermatológicas, alcançando 92% de precisão.

A aplicação dessas redes na detecção de defeitos em pavimentos permite a automação do processo de inspeção, reduzindo a necessidade de intervenção humana e aumentando a precisão na identificação das patologias. Assim, o uso de RNCs aliado a imagens capturadas por drones representa um avanço significativo na manutenção preditiva e na gestão da infraestrutura viária.

2.5 Aplicação do Modelo YOLO na Identificação de Defeitos

O modelo YOLO (*You Only Look Once*) é uma arquitetura de detecção de objetos em tempo real que processa a imagem inteira em uma única passagem pela rede neural, diferindo de métodos tradicionais baseados em múltiplas etapas (ALVES, 2020). Essa abordagem possibilita alta velocidade e precisão, sendo capaz de detectar simultaneamente diversos alvos por meio de redes neurais convolucionais que dividem a imagem em grades e preveem caixas delimitadoras e classes (REDMON et al., 2016).

Na inspeção de pavimentos, o YOLO se destaca pela capacidade de identificar e localizar patologias como trincas e panelas com precisão, permitindo integração a sistemas de informação geográfica para planejamento de manutenção (NÓBREGA, 2023; ABBAS; ISMAEL, 2021). Seu desempenho está diretamente relacionado à qualidade e diversidade do

conjunto de dados de treinamento, sendo potencializado pelo uso de técnicas como *transfer learning*, que reduzem o tempo de treino e melhoram a generalização (TANWAR, 2024; BOCHKOVSKIY; WANG; LIAO, 2020).

Combinado a tecnologias de captura automatizada, como drones e veículos instrumentados, o YOLO representa um avanço significativo frente aos métodos tradicionais, oferecendo rapidez, escalabilidade e suporte estratégico à gestão da infraestrutura viária (ALVES, 2020; TANWAR, 2024).

2.6 Google Colab e a Implementação do Modelo

O Google Colaboratory, ou simplesmente Google Colab, é uma plataforma de desenvolvimento baseada em nuvem disponibilizada gratuitamente pelo Google, que permite a execução de códigos na linguagem Python sem a necessidade de instalação local de softwares ou configurações complexas. Essa característica torna a ferramenta amplamente utilizada por pesquisadores, estudantes e desenvolvedores para prototipagem e implementação de modelos de inteligência artificial (GOOGLE, 2024).

Entre suas principais vantagens destaca-se a integração nativa com unidades de processamento gráfico (GPUs) e unidades de processamento tensorial (TPUs), que proporcionam ganhos expressivos de desempenho no treinamento de redes neurais convolucionais e outros modelos de aprendizado profundo. Esses recursos de aceleração de hardware são disponibilizados sem custos adicionais e permitem a execução de tarefas de alta demanda computacional de forma mais rápida e eficiente, mesmo para usuários que não dispõem de infraestrutura física própria (GOOGLE, 2024; AWS, 2024).

Outro diferencial do Google Colab é a sua integração com o Google Drive, possibilitando o armazenamento e a manipulação de grandes volumes de dados diretamente na nuvem. Essa funcionalidade facilita a organização de bases de imagens, scripts e resultados. (ALURA, 2023).

2.7 Utilização do Roboflow para o treinamento de IA

O Roboflow é uma plataforma online voltada para o gerenciamento, processamento e anotação de bases de dados de imagens, amplamente utilizada em projetos de visão computacional para tarefas de classificação, detecção e segmentação de objetos. A ferramenta fornece um ambiente integrado que facilita desde a aquisição de conjuntos de dados até a exportação de datasets prontos para uso em diferentes arquiteturas de redes neurais, incluindo o modelo YOLO (ROBOFLOW, 2024).

Entre suas funcionalidades destacam-se a capacidade de realizar pré-processamento e aumento de dados (*data augmentation*), aplicando automaticamente transformações como rotações, ajustes de brilho e contraste, inserção de ruído, recortes e espelhamentos. Essas operações aumentam a diversidade e a robustez do conjunto de treinamento, contribuindo para melhorar a capacidade de generalização do modelo frente a variações de iluminação, ângulos de captura e condições ambientais (ROBOFLOW, 2024; SHORTEN; KOSHY, 2019).

No presente trabalho, o Roboflow foi utilizado tanto para a organização e curadoria da base de dados quanto para o refinamento do modelo de detecção de defeitos em pavimentos. A plataforma permitiu que milhares de imagens fossem processadas e anotadas de forma eficiente, reduzindo o tempo necessário para a preparação do *dataset*. Além disso, suas funcionalidades de integração direta com o Google Colab e outros ambientes de desenvolvimento possibilitaram que o treinamento fosse realizado de forma contínua, com ajustes nos parâmetros e testes rápidos de desempenho.

3 METODOLOGIA

A metodologia adotada neste estudo foi estruturada para permitir o desenvolvimento, treinamento, validação e análise comparativa de diferentes modelos de detecção de patologias em pavimentos utilizando a arquitetura YOLO. O processo contemplou desde a coleta e organização das imagens até a avaliação sistemática do desempenho, buscando garantir reprodutibilidade e consistência científica.

Inicialmente, foram empregadas plataformas integradas que viabilizaram tanto o gerenciamento de dados quanto o processamento computacional. O Roboflow foi utilizado como ambiente central para curadoria, anotação e treinamento preliminar dos modelos, possibilitando a aplicação de técnicas de *data augmentation* para aumentar a diversidade visual do conjunto de imagens. O Google Drive foi adotado como repositório para armazenamento e organização dos dados, enquanto o Google Colab foi empregado para análises complementares e ajustes mais específicos nos parâmetros de treinamento, aproveitando sua capacidade de processamento em GPU/TPU e integração direta com o Roboflow.

Ao longo da pesquisa, cada modelo foi submetido a um processo de validação utilizando imagens inéditas — não presentes na etapa de treinamento — para mensurar métricas como mAP@50, Precisão e Recall, além do F1-score estimado, permitindo comparar não apenas a acurácia, mas também o equilíbrio entre sensibilidade e seletividade. Essa abordagem metodológica possibilitou identificar o efeito da quantidade e qualidade de dados, bem como da estratégia de treinamento, sobre o desempenho final da rede.

3.1 Coleta de Dados

Para o desenvolvimento do modelo de detecção baseado na arquitetura YOLO, foram priorizadas imagens de patologias do tipo panela, devido à sua elevada frequência de ocorrência em pavimentos asfálticos e à ampla disponibilidade de registros em bases públicas. Inicialmente, foram coletadas aproximadamente 250 imagens provenientes de diferentes fontes, principalmente do acervo do Roboflow Universe, que reúne mais de 500 milhões de imagens gratuitas.

Para ampliar a diversidade e a robustez do treinamento, as imagens passaram por um processo de aumento de dados (*data augmentation*), incluindo operações de rotação, distorção, variação de escala, ajustes de brilho e contraste, recortes e espelhamento. Essas transformações

visaram expor o modelo a diferentes condições de iluminação, ângulos de captura e contextos visuais, aumentando sua capacidade de generalização. Apesar da aplicação dessas técnicas, a quantidade original de imagens foi mantida como referência, preservando o equilíbrio entre a base original e as variações geradas artificialmente, tornando o total próximo de 1000 imagens para o primeiro modelo.

3.2 Treinamento do Modelo de IA

As imagens coletadas foram utilizadas para o treinamento de uma Rede Neural Convolutacional (RNC) implementada na arquitetura YOLO, com o objetivo de identificar a localização e os contornos das patologias em pavimentos, com foco específico no tipo panela. O treinamento foi conduzido de forma supervisionada e manual, de modo que o modelo adquirisse uma base sólida para a análise das imagens de validação e teste.

O processo consistiu na apresentação de imagens ao algoritmo, acompanhadas de anotações poligonais que delimitavam a área exata da patologia, podendo-se observar esse processo na Figura 2. A partir da análise repetida de centenas dessas imagens anotadas, a rede aprendeu a associar padrões visuais presentes dentro das regiões marcadas à ocorrência de defeitos. Essa abordagem possibilitou que o modelo, ao ser exposto a novas imagens contendo múltiplas informações visuais, identificasse com precisão as regiões correspondentes às panelas.

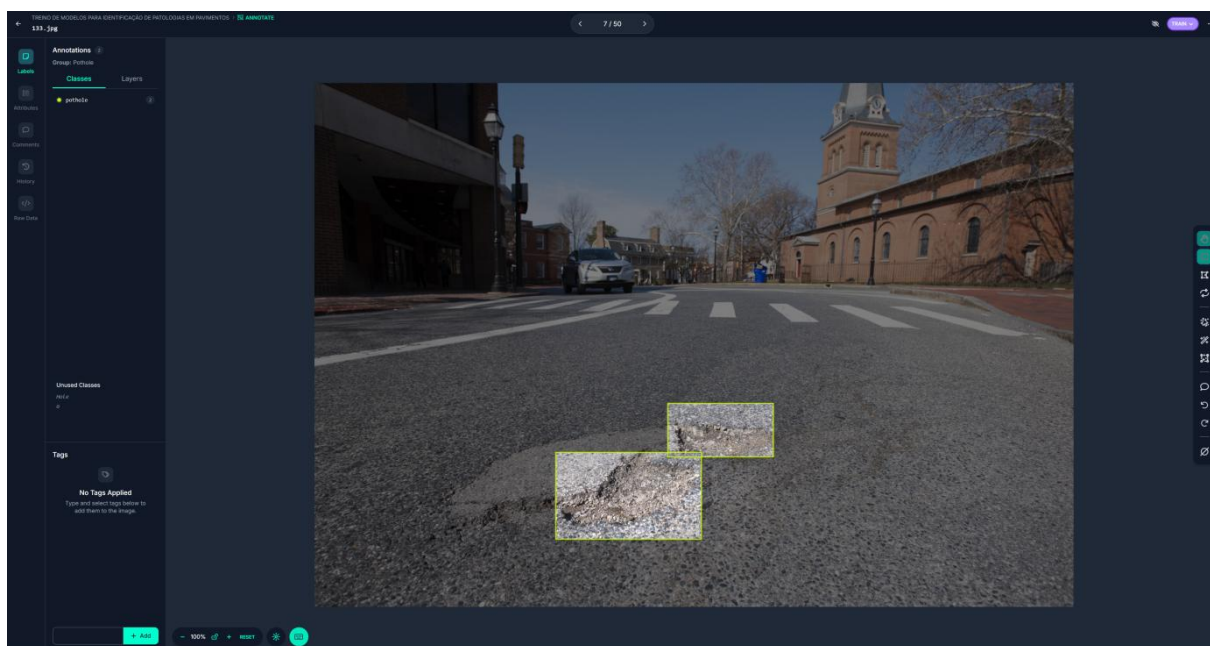


Figura 2 – Delimitação Manual das Patologias para reconhecimento de Padrão

Fonte: kaggle, 2020.

O primeiro modelo desenvolvido, denominado patologia-v1, foi criado em 15 de outubro de 2024, com 326 imagens destinadas ao treino, 93 para validação e 46 para teste. Ao longo da pesquisa, esse modelo passou por sucessivas atualizações com a inclusão de novos dados e refinamentos no processo de anotação manual, alcançando 909 imagens no conjunto de treino como mostrado na Figura 3, mantendo inalterados os conjuntos de validação e teste. Esse modelo apresentou o desempenho mais consistente entre todos, sendo capaz de identificar com precisão até mesmo múltiplos buracos em uma única imagem — característica que não foi replicada pelos modelos com bases menores e menor tempo de treinamento supervisionado.

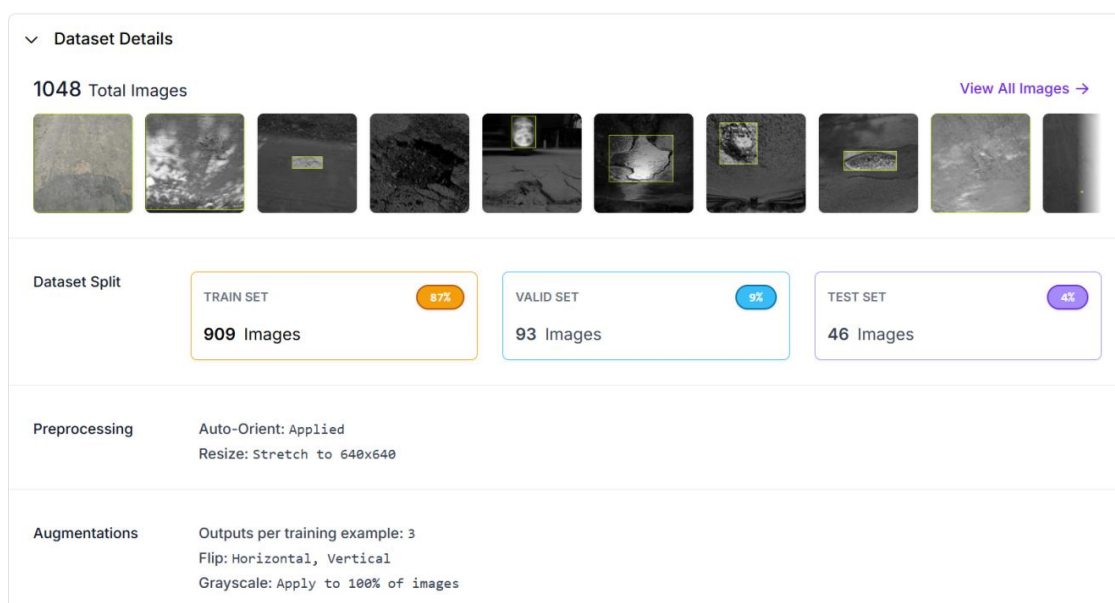


Figura 3 – Detalhes do primeiro modelo

Fonte: Autor, 2025.

Ao todo, foram criados cinco modelos, dos quais quatro foram utilizados de forma ativa. O segundo modelo relevante foi desenvolvido em 12 de novembro de 2024, com 752 imagens de treino, 214 de validação e 107 de teste. Já o terceiro modelo, criado em 29 de novembro de 2024, utilizou exatamente o mesmo conjunto de dados do segundo, sendo concebido com o objetivo de avaliar as possibilidades de diversificação de resultados a partir de um dataset e procedimento de treino idênticos. O último modelo foi gerado apenas para testes preliminares, incluindo um treinamento automatizado sem verificação manual, com base de dados 2,5 vezes maior que a do patologia-v1, utilizado para avaliar o impacto da ausência de supervisão e treinamento humano na qualidade do resultado.

3.3 Validação e Testes

A etapa de validação e testes teve como objetivo avaliar a precisão e a robustez dos modelos desenvolvidos. Para isso, foram utilizadas imagens inéditas, não presentes no conjunto

de treinamento, a fim de garantir que a avaliação refletisse a capacidade real de generalização da rede neural.

O processo consistiu na comparação direta entre as predições automáticas do modelo e as anotações realizadas manualmente. Essa análise permitiu verificar a taxa de acertos na detecção das patologias do tipo panela, bem como identificar eventuais falsos positivos e falsos negativos. Foram observados, ainda, casos de variação na precisão em função de fatores como iluminação, textura do pavimento e presença de elementos visuais que pudessem confundir a rede.

Os resultados obtidos foram utilizados para retroalimentar o processo de treinamento, permitindo ajustes nos hiperparâmetros e na base de dados. Dessa forma, a etapa de validação não se limitou apenas a mensurar o desempenho, mas também contribuiu para o refinamento contínuo do modelo, visando maior acurácia e confiabilidade nas aplicações práticas.

4 RESULTADOS E DISCUSSÕES

A avaliação de desempenho dos modelos desenvolvidos baseou-se em três métricas principais: mAP@50, Precisão (*Precision*) e Sensibilidade (*Recall*). Essas métricas foram amplamente utilizadas em trabalhos de detecção de objetos por redes neurais convolucionais, sendo consideradas padrão em competições e pesquisas na área (EVERINGHAM et al., 2010; REDMON et al., 2016).

4.1 Definição das Métricas Utilizadas

A mAP@50 (*Mean Average Precision at Intersection over Union = 50%*) mede a média da precisão considerando diferentes limiares de confiança, mas utilizando como critério mínimo de sobreposição (*Intersection over Union*, IoU) 50% entre a área detectada e o objeto real (EVERINGHAM et al., 2010). Essa métrica é particularmente útil para avaliar o equilíbrio entre a capacidade de localizar corretamente o objeto e a confiança na detecção. Valores mais altos indicam melhor desempenho global, conciliando precisão e sensibilidade.

A Precisão (*Precision*) representa a proporção de detecções corretas entre todas as detecções realizadas pelo modelo (SOKOLOVA; LAPALME, 2009). É calculada pela Equação (1):

$$(1): \quad \text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}}$$

Na prática, valores elevados de precisão indicam que o modelo comete poucos falsos positivos, ou seja, raramente detecta um defeito inexistente.

A Sensibilidade (*Recall*), também conhecida como taxa de acertos ou verdadeiros positivos, mede a proporção de objetos reais detectados corretamente pelo modelo (SOKOLOVA; LAPALME, 2009). É calculada pela Equação (2):

$$(2): \quad \text{Recall} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}}$$

Valores altos de *recall* indicam que o modelo é eficaz em identificar a maioria dos defeitos presentes nas imagens, reduzindo a ocorrência de falsos negativos.

Em conjunto, essas métricas fornecem uma visão abrangente do desempenho do sistema, permitindo identificar se o modelo prioriza mais a precisão ou a cobertura das detecções, e se há equilíbrio entre ambas.

4.2 Resultados Obtidos

4.2.1 Modelo *Patologia-v1*

O primeiro modelo (*patologia-v1*), treinado progressivamente com maior volume de dados e anotações manuais de alta qualidade, apresentou $\text{mAP@50} = 55,8\%$, Precisão = $60,6\%$ e Recall = $56,4\%$. Esse desempenho demonstra um equilíbrio consistente entre seletividade e abrangência: a precisão ligeiramente superior ao *recall* indica que o modelo tende a evitar falsos positivos, mesmo que isso implique deixar passar alguns defeitos mais sutis.

O F1-score estimado é $58,4\%$, confirmando sua capacidade de manter um balanço satisfatório entre detecções corretas e cobertura da detecção. O bom desempenho pode ser atribuído à expansão gradual da base de dados (de 326 para 909 imagens de treino) e à curadoria manual intensiva, que reduziu ambiguidades nas anotações e aprimorou a capacidade do modelo de lidar com imagens contendo múltiplos defeitos.

4.2.2 Segundo Modelo

O segundo modelo, treinado em 12/11/2024 com 752 imagens de treino, obteve $\text{mAP@50} = 55,3\%$, Precisão = $55,7\%$ e Recall = $51,3\%$. Em relação ao *patologia-v1*, observou-se queda nas três métricas: $-0,5$ p.p. (pontos percentuais) no mAP@50 , $-4,9$ p.p. em precisão e $-5,1$ p.p. em *recall*.

Essa redução indica que a menor quantidade de dados e a provável redução no detalhamento das anotações impactaram negativamente a generalização e aumentaram a ocorrência de falsos positivos e falsos negativos. O F1-score foi estimado em $53,4\%$, confirmando menor robustez em relação ao primeiro modelo.

4.2.3 Terceiro Modelo

O terceiro modelo, criado em 29/11/2024 com o mesmo dataset do segundo, apresentou $\text{mAP@50} = 53,9\%$, Precisão = $57,0\%$ e Recall = $55,1\%$. Comparado ao segundo modelo,

apresentou ganho de 3,8 p.p. em *recall* e aumento de 1,3 p.p. em precisão, mas com redução de 1,4 p.p. no mAP@50.

O F1-score estimado foi 56,0%, sugerindo que ajustes no processo de treinamento — mesmo mantendo a mesma base de dados — podem deslocar o modelo para um ponto de operação mais voltado à sensibilidade (maior detecção de defeitos), ainda que à custa de ligeira perda na consistência global medida pelo mAP@50.

4.2.4 Quarto Modelo

O quarto modelo foi gerado com caráter experimental, visando avaliar o impacto da ausência de supervisão e treinamento humano no desempenho. Para isso, utilizou-se uma base de dados 2,5 vezes maior que a do patologia-v1, porém sem verificação manual das anotações. O treinamento foi totalmente automatizado, dispensando a curadoria que caracterizou os modelos anteriores.

Os resultados obtidos foram significativamente inferiores: mAP@50 = 29,5%, Precisão = 34,6% e Recall = 37,9%, resultando em um F1-score de aproximadamente **36,2%**. Apesar do grande volume de dados, a ausência de supervisão manual comprometeu drasticamente a qualidade das anotações, levando a uma alta incidência de falsos positivos e falsos negativos.

Esse resultado reforça a premissa de que, para modelos de detecção de objetos, a qualidade das anotações é mais determinante que a quantidade bruta de dados. A inclusão de exemplos incorretamente rotulados durante o treino provavelmente gerou ruído nas representações aprendidas pela rede, prejudicando a capacidade de diferenciar corretamente padrões de defeitos e não-defeitos.

Foi realizado um teste utilizando uma imagem externa, não pertencente ao banco de dados de treinamento, com o objetivo de avaliar a confiabilidade e a capacidade de identificação de cada modelo. Para garantir a comparabilidade dos resultados e permitir uma análise fiel do desempenho, a mesma imagem foi utilizada em todos os testes mostrados nas figuras a seguir.

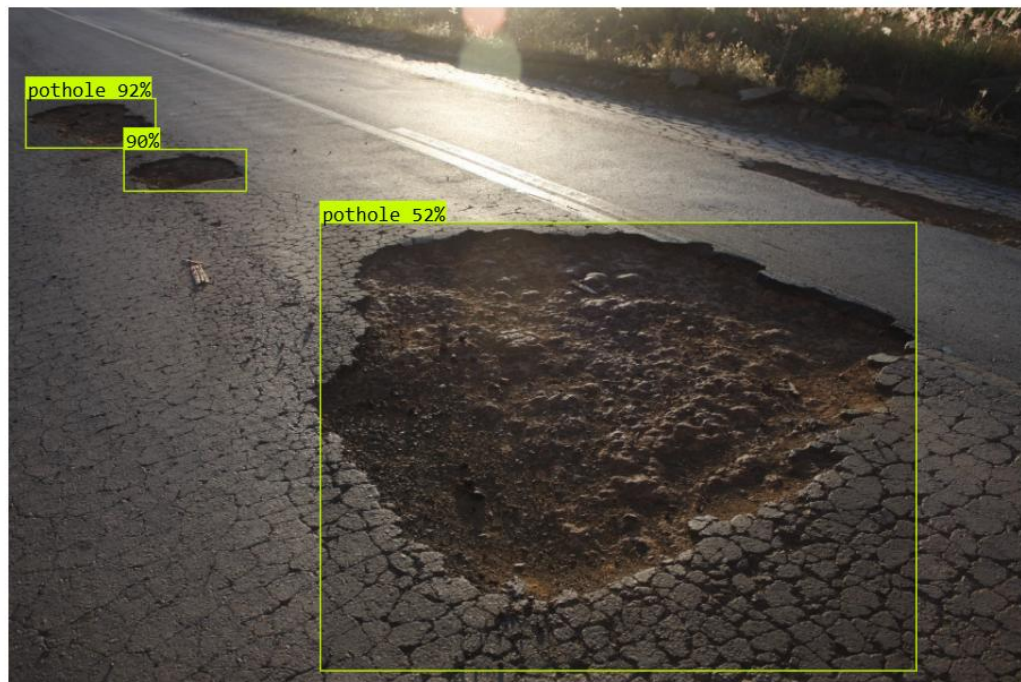


Figura 4 – Teste Real patologia-v1 (primeiro modelo)

Fonte: Redediocesana, 2024.



Figura 5 – Teste Real segundo modelo

Fonte: Redediocesana, 2024.



Figura 6 – Teste Real terceiro modelo

Fonte: Redediocesana, 2024.



Figura 7 – Teste Real quarto modelo

Fonte: Redediocesana, 2024.

Tabela 1 – Comparativa de métricas dos modelos desenvolvidos

Modelo	mAP@50 (%)	Precisão (%)	Recall (%)	F1-score (%)	Observações Principais
Patologia-v1	55,8	60,6	56,4	58,4	Maior equilíbrio entre precisão e recall; base expandida com curadoria manual intensiva.
Segundo Modelo	55,3	55,7	51,3	53,4	Base menor e menos curada; queda geral no desempenho.
Terceiro Modelo	53,9	57,0	55,1	56,0	Mesmo dataset do segundo; melhor recall, mas menor mAP@50.
Quarto Modelo	29,5	34,6	37,9	36,2	Treinamento totalmente automatizado, sem supervisão; base 2,5× maior, mas com queda drástica de desempenho.

Fonte: Autor, 2025.

5 CONSIDERAÇÕES FINAIS

O presente estudo demonstrou que a combinação de imagens aéreas obtidas por drones com técnicas de visão computacional baseadas no modelo YOLO representa uma solução eficiente para a inspeção de pavimentos, permitindo diagnósticos rápidos e menos invasivos, com potencial para redução de custos e aumento da eficiência operacional.

A análise comparativa dos quatro modelos desenvolvidos evidenciou que a qualidade das anotações tem influência mais determinante no desempenho do que a simples ampliação do volume de dados. O modelo patologia-v1, treinado com supervisão manual e expansão gradual da base, apresentou mAP@50 de 55,8%, precisão de 60,6% e recall de 56,4%, configurando-se como o mais equilibrado entre as métricas avaliadas, sendo capaz de reconhecer todas as patologias, conforme apresentado anteriormente em exemplo da Figura 24 com diferentes graus de precisão.

O terceiro modelo, embora inferior ao patologia-v1 em mAP@50 e precisão, apresentou recall próximo, porém não foi capaz de reconhecer uma das patologias da Figura 26 onde ambos o modelo 1 e o modelo 2 foram capazes.

O segundo modelo apresentou resultados inferiores ao patologia-v1 em todas as métricas, porém para patologias de maior área apresentou uma confiança superior ao primeiro modelo.

O quarto modelo, desenvolvido para testes preliminares com base de dados 2,5 vezes maior que a do patologia-v1, porém sem verificação manual, apresentou queda acentuada de desempenho, atingindo mAP@50 de 29,5%, precisão de 34,6% e recall de 37,9%. Esse

resultado confirma que anotações incorretas ou inconsistentes durante o treinamento comprometem a capacidade de generalização do modelo, mesmo em bases extensas, sendo incapaz de identificar qualquer patologia de forma correta na Figura 27.

Todos os modelos apresentaram dificuldades em condições adversas, como iluminação inadequada, presença de sombras intensas, semelhança visual entre defeitos e elementos não relacionados (como remendos e manchas de óleo) e detecção de defeitos pequenos ou distantes. Essas limitações apontam para a necessidade de aprimoramentos como aumento da diversidade nas técnicas de data *augmentation*, uso de treinamento multi-escala, ajustes dinâmicos nos parâmetros de confiança e IoU, aprimoramento das anotações poligonais e inclusão de cenários mais desafiadores no conjunto de validação.

Com base nos resultados obtidos, destacam-se três conclusões principais: (i) a qualidade das anotações influencia diretamente o mAP@50 e o equilíbrio entre precisão e recall; (ii) variações no processo de treinamento, mesmo com o mesmo conjunto de dados, podem alterar significativamente o comportamento do modelo; e (iii) modelos treinados com supervisão consistente e base de dados amadurecida tendem a apresentar melhor desempenho médio e maior previsibilidade em campo.

Referências

ALVES, João. Aplicações do modelo **YOLO** para detecção de objetos em tempo real. *Revista de Inteligência Artificial Aplicada*, v. 5, n. 2, p. 45–58, 2020.

ANDREWMVD. Pothole Detection. Kaggle, 202x. Disponível em: <https://www.kaggle.com/datasets/andrewmvd/pothole-detection>. Acesso em: 18 ago. 2025.

BERNUCCI, et al. Manifestações **patológicas** em estruturas asfálticas: estudo. *GETEC*, v. 10, n. 30, p. 54–70, 2021.

BRÍGIDO, Rodrigo. Diferenças básicas entre **pavimento** rígido e flexível. *Engenharia Conceitual*, 2018. Disponível em: <https://designbrigido.wixsite.com/eng-conceitual/single-post/2018/03/17/diferen%C3%A7as-b%C3%A1sicas-entre-pavimento-r%C3%ADgido-e-flex%C3%ADvel>. Acesso em: 15 ago. 2025.

CBBASFALTOS. **Desagregação** precoce do concreto asfáltico usinado a quente. *Blog CBB Asfaltos*, 2021. Disponível em: <https://blog.cbbasfaltos.com.br/desagregacao-precoce-do-concreto-asfaltico-usinado-a-quente/>. Acesso em: 03 jan. 2025.

CONFEDERAÇÃO NACIONAL DO TRANSPORTE (CNT). Conheça os 13 principais **defeitos** do **pavimento** das rodovias. 2018. Disponível em: <https://www.cnt.org.br/agencia-cnt/conheca-principais-defeitos-pavimento>. Acesso em: 21 ago. 2024.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES (DNIT). DNIT 005/2003 – TER: **Defeitos** nos **pavimentos** asfálticos – Terminologia. Brasília: DNIT, 2003. Disponível em: https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-normas/terminologia-ter/dnit_005_2003_ter-1.pdf. Acesso em: 15 set. 2024.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES (DNIT). **Manual de restauração** de **pavimentos** asfálticos. Brasília: DNIT/IPR, 2006. Disponível em: https://www.gov.br/dnit/pt-br/assuntos/planejamento-e-pesquisa/ipr/coletanea-de-manuais/vigentes/720_manual_restauracao_pavimentos_afalticos.pdf. Acesso em: 30 out. 2024.

DATASCIENTEST. **You Only Look Once (YOLO)**: o que é? 2023. Disponível em: <https://datascientest.com/en/you-only-look-once-yolo-what-is-it>. Acesso em: 27 mar. 2025.

GLOBO, O. **Buracos e ondulações** prejudicam motoristas e ciclistas na Zona Sul. *O Globo*, 2021. Disponível em: <https://oglobo.globo.com/eu-reporter/buracos-ondulacoes-prejudicam-motoristas-ciclistas-na-zona-sul-4041611>. Acesso em: 19 set. 2024.

GOOGLE. **Google Colaboratory**: recursos e especificações técnicas. 2024. Disponível em: <https://colab.research.google.com>. Acesso em: 05 jan. 2025.

INOVACIVIL. As **patologias** mais comuns nas estradas. 2022. Disponível em: <https://inovacivil.com.br/as-patologias-mais-comuns-nas-estradas/>. Acesso em: 08 dez. 2024.

IONOS. **Convolutional Neural Network (CNN)**. 2023. Disponível em: <https://www.ionos.com/pt-br/digitalguide/sites-de-internet/desenvolvimento-web/convolutional-neural-network/>. Acesso em: 14 fev. 2025.

LABORATÓRIO, et al. Comparative study of commercial electrodes for underwater wet welding. *ResearchGate*, 2010. Disponível em: https://www.researchgate.net/figure/Exemplo-de-trincas-transversais-em-secao-longitudinal-do-metal-de-solda-dos-eletrodos-B_fig3_262660359. Acesso em: 12 jun. 2025.

MARIANI, B.; BARROS, M. **Exsudação** aparente na via. *ResearchGate*, 2020. Disponível em: https://www.researchgate.net/figure/Figura-7-Exsudacao-aparente-na-via_fig2_343332220. Acesso em: 05 nov. 2024.

REDE DIOCESANA. Condições precárias da BR-158 entre Piranhas e Caiapônia causam onda de acidentes. Rede **Diocesana**, 2023. Disponível em: <https://redediocesana.com/caiaponia-e-regiao/condicoes-precarias-da-br-158-entre-piranhas-e-caiaponia-causam-onda-de-acidentes/>. Acesso em: 15 jun. 2025.

ONSV. Ocorrências de trânsito causadas por **buracos** na via. *Observatório Nacional de Segurança Viária*, 2022. Disponível em: <https://www.onsv.org.br/comunicacao/observadores-certificados/ocorrencias-de-transito-causadas-por-buracos-na-via>. Acesso em: 22 jan. 2025.

TANWAR, Rohit. **Pavement damage detection** using YOLO and deep learning techniques. *International Journal of Computer Vision and Artificial Intelligence*, v. 12, n. 1, p. 25–37, 2024.

THEAI SUMMER. **YOLO – You Only Look Once**. 2023. Disponível em: <https://theaisummer.com/YOLO/>. Acesso em: 10 jul. 2025.