Industrialização, Digitalização, Desempenho

5º Simpósio Brasileiro de Tecnologia da Informação e Comunicação na Construção e 5º Workshop de Tecnologia de Processos e Sistemas Construtivos

FLORIANÓPOLIS-SC | 20 a 22 de agosto

[1]SEGMENTAÇÃO AUTOMATIZADA DE PAVIMENTOS EM IMAGENS DE AMBIENTES URBANOS USANDO YOLO V11

Automated Pavement Segmentation in Urban Environment Images Using YOLO V11

Gabriel de Souza Lima
Escola Politécnica da USP | São Paulo, SP | gabriel-souzalima@live.com

Jonathan Chefaly Mochon Zappile
Escola Politécnica da USP | São Paulo, SP | jonathanzappile@gmail.com

Marcio José Serra Paixão
Escola Politécnica da USP | São Paulo, SP | mjspaixao@gmail.com

Giovanni Bruno Molitor Schiffini
Escola Politécnica da USP | São Paulo, SP | giovannibruno658@gmail.com

Flavio Leal Maranhão, Prof. Dr.
Escola Politécnica da USP | São Paulo, SP | flavio.maranhao@usp.br

Angelo Sebastião Zanini, Prof. Dr.
Instituto Mauá de Tecnologia | São Paulo, SP | angelo.zanini@geovista.com.br

RESUMO

A segmentação de pavimentos em vias urbanas de forma automatizada traz para gestores públicos a possibilidade de planejamento antecipado para obras de manutenção, permitindo também a priorização de trabalhos em função do tipo do material presente em cada logradouro. Sendo assim, a pesquisa aqui apresentada tem como objetivo principal o treinamento de um modelo de inteligência artificial denominado Yolov11, sendo este capaz de analisar inúmeras imagens de forma rápida e acurada, o que possibilita a realização futura do inventário da malha viária de grandes metrópoles. Logo, para atingir tal objetivo, câmeras foram instaladas em uma frota de veículos para a obtenção de imagens. Como objetivo secundário e com um grande volume de imagens brutas disponíveis, um dataset anotado foi preparado, possibilitando o treinamento e teste de máquina. Resultados mostram que houve o aprendizado do modelo, onde foi obtido um mAP de 0.94.

Palavras-chave: segmentação do pavimento, visão computacional, monitoramento, yolov11, inteligência artificial.

ABSTRACT

The automated segmentation of pavements in urban roads provides public managers with the possibility of advance planning for maintenance works, also allowing the prioritization of work based on the type of material present in each street. Therefore, the research presented here has as its main objective the training of an artificial intelligence model called Yolov11, which can analyze numerous images quickly and accurately, which enables the future completion of an inventory of the road network of large metropolises. Therefore, to achieve this objective, cameras were installed in a fleet of vehicles to obtain images. As a secondary objective and with a large volume of raw images available, an annotated dataset was prepared, enabling machine training and testing. Results show that the model learned, where a mAP of 0.94 was obtained.

Keywords: *Pavement segmentation, Computer vision, monitoring, Yolov11, Artificial intelligence.*

1 INTRODUÇÃO

Em ambientes urbanos, conhecer a composição da malha viária é de suma importância, uma vez que o tipo de pavimento pode influenciar no conforto e segurança dos motoristas que ali trafegam (TIGHE, 2000). A geração de um inventário preciso pode ser uma ferramenta valiosa para gestores públicos, possibilitando assim uma administração eficiente baseada em dados.

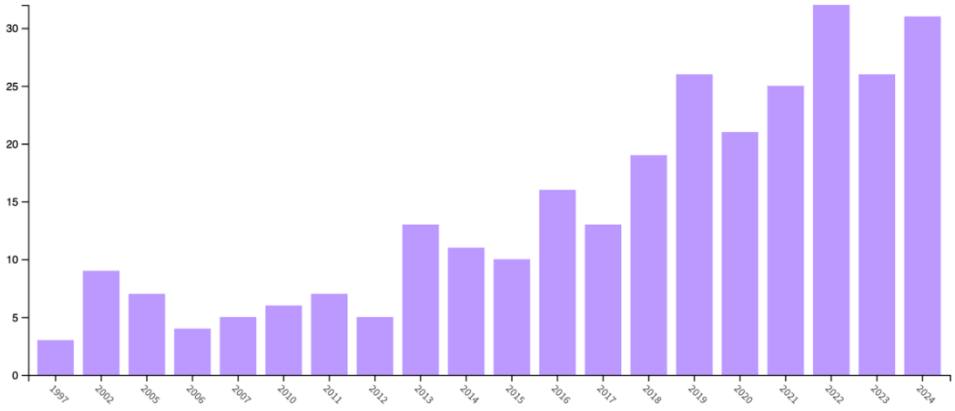
Atualmente, diversos métodos e tecnologias estão sendo empregados para avaliar pavimentos em tempo real, como a implantação de sensores inerciais em carros de passeios (PAIXÃO, 2020) (LIMA, 2024) para a verificação da qualidade superficial do pavimento, a utilização de veículos aéreos não tripulados para a

[1]LIMA, G.; ZAPPILE, J. C. M.; PAIXAO, M. J. S.; SCHIFFINI, G. B. M.; MARANHAO, F. L.; ZANINI, A. S. Segmentação Automatizada de Pavimentos em Imagens de Ambientes Urbanos usando YOLO V11. In: 5º SIMPÓSIO BRASILEIRO DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO NA CONSTRUÇÃO, 4., 2025, Florianópolis. **Anais [...].** Porto Alegre: ANTAC, 2025.

detecção de buracos nas vias (ALZAMZAMI, 2024), ou ainda a utilização de imagens de câmeras fixas para a detecção de outros objetos de interesse (HORITA, 2012).

Para problemas que demandam larga escala, como o levantamento e atualização do inventário das vias de uma cidade, o método crowdsensing torna-se uma opção, dado que múltiplos usuários poderão contribuir em tempo real com seus dados. Desse modo, pode-se observar um crescente aumento de trabalhos que embarcam câmeras em veículos para a análise de pavimentos, como mostra a Figura 1.

Figura 1: Gráfico de barras proveniente da pesquisa "cameras in vehicle pavement" na base de dados Web of Science



Fonte: Web Of Science (2025)

Por outro lado, a utilização da visão computacional para tais tarefas vem se mostrando promissora uma vez que, se aliado com o georreferenciamento das capturas pode-se obter não apenas uma evidencia visual para as detecções, mas também automatizar processos e abrir ordens de serviços para locais específicos. Porém, realizar a implementação desses sistemas em cenários dinâmicos é deveras desafiador, pois questões como iluminação, ângulo de visão e obstruções devem ser analisados no momento da concepção do projeto (LEE, 2024).

Além da visão computacional, a inteligência artificial vem desempenhando um papel importante para essa área, principalmente para aplicações de detecção, classificação e segmentação de objetos (FURTADO, 2024). Nesse contexto, podemos destacar o modelo de segmentação YOLOv11 (You Only Look Once) (ULTRALYTICS, 2024), sendo esta uma potente solução para inferências em grandes volumes de dados, possibilitando ainda, através do método transfer learning, o treinamento personalizado das classes de interesse (Nie, 2018).

Com sua arquitetura de estágio único, a YOLOv11 permite uma maior agilidade computacional no processamento dos frames comparado a outros modelos, garantindo também a acurácia necessária para as tarefas anteriormente mencionadas (JEGHAM, 2024).

Desse modo, o presente trabalho propõe, a partir da utilização de câmeras embarcadas em uma frota de veículos, a utilização do modelo de inteligência artificial YOLOv11 para a segmentação automatizada de pavimentos, sendo revestimento asfáltico, concreto, bloco intertravado e paralelepípedo as classes de interesse, conforme ilustrado na Figura 2.

Figura 2: Classes de Interesse. (a) revestimento asfáltico, (b) concreto, (c) bloco intertravado, (d) paralelepípedo.



Fonte: Autores (2025)

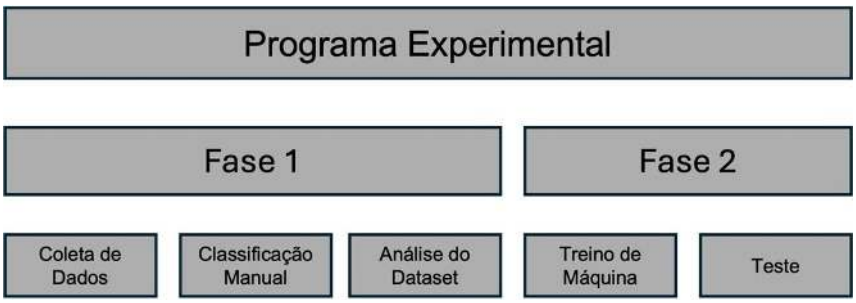
Como objetivo principal têm-se o treinamento do modelo, onde procura-se obter a capacidade de segmentar imagens de pavimentos em tempo real e com alta precisão, contribuindo assim para sistemas futuros de monitoramento da infraestrutura rodoviária, sendo que a inovação reside na aplicação de tal método para o problema prático mencionado.

Como objetivo secundário têm-se a geração de um dataset anotado, contendo imagens coletadas em campo de diversos tipos de pavimentos em diferentes contextos e cenários, possibilitando treinos de máquina não apenas para a atual pesquisa, mas também para trabalhos posteriores.

2 MATERIAIS E MÉTODOS

O programa experimental aqui proposto é dividido em duas fases, conforme ilustrado na Figura 3.

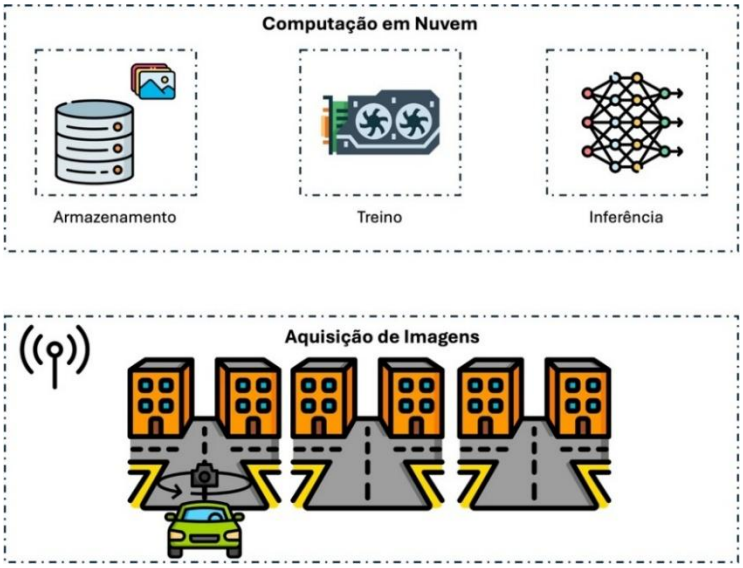
Figura 3: Fases do programa experimental.



Fonte: Autores (2025)

A primeira fase tem como objetivo gerar o dataset de treino, tendo como resultado esperado um conjunto plural de imagens que representam em sua totalidade os tipos de pavimento que se deseja detectar (revestimento asfáltico, concreto, bloco intertravado, paralelepípedo). Para isso, foi instalado um hardware equipado com câmera em uma frota descentralizada, onde imagens foram aquisitadas a 1 frame por segundo, gerando assim um conjunto de dados bruto para posterior análise. As imagens capturadas, junto com suas informações de geolocalização foram automaticamente enviadas para o servidor de armazenamento em nuvem através da conexão do celular do motorista a medida em que os carros aleatoriamente trafegavam pela cidade. A Figura 4 mostra em detalhes a arquitetura proposta para a coleta de dados.

Figura 4: Arquitetura proposta para a fase de coleta de dados.



Fonte: Autores (2025)

Em seguida, tem-se a etapa de classificação manual das imagens obtidas em campo, que se faz necessária pois o framework de treinamento utilizado se baseia no aprendizado supervisionado de máquina. Portanto, a geração de imagens anotadas com os elementos de interesse (ground truth) é o alvo desse processo.

A Figura 5 demonstra o sistema utilizado para a realização das classificações manuais, no qual cada amostra ao ser classificada é associada a um metadado contendo a máscara de pixels da classe observada.

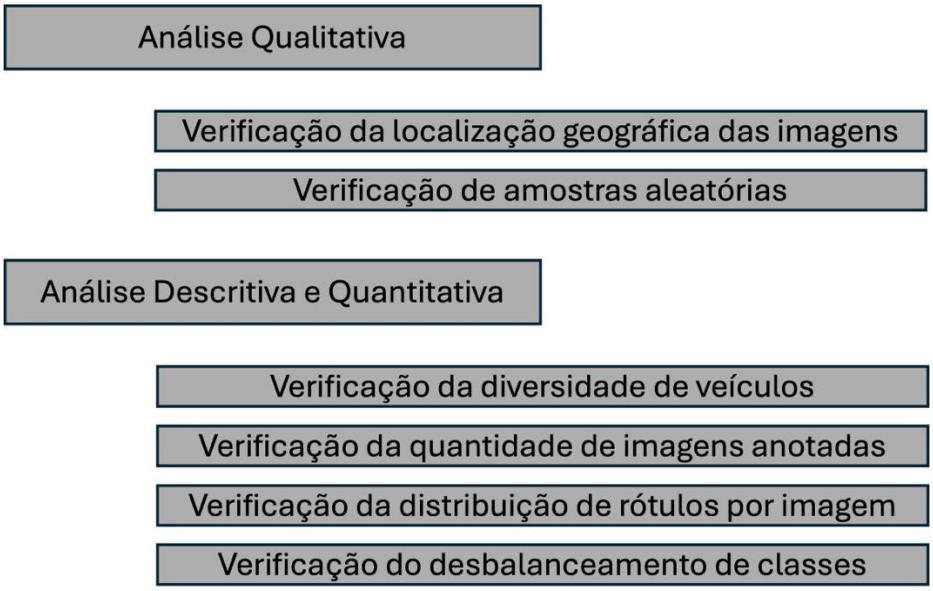
Figura 5: Sistema para segmentação manual de imagens.



Fonte: Autores (2025)

Por fim, para garantir a pluralidade do conjunto de dados, entender se há balanceamento entre as classes e eliminar possíveis inconsistências, foi realizada uma análise exploratória no dataset anotado, estando tal processo detalhado na Figura 6.

Figura 6: Método para exploração do dataset anotado.



Fonte: Autores (2025)

Na segunda fase é realizado o treino e teste de máquina, tendo como principal meta a generalização do modelo, desejando assim obter uma alta assertividade para amostras nunca observadas pelo sistema. Para esse propósito é dividido o dataset em três subconjuntos (validação, treino e teste), sendo validação e treino utilizados no momento de treinamento e teste utilizado para holdout.

As métricas que embasam as análises de assertividade do modelo são descritas pela (Equação 1), (Equação 2), (Equação 3), (Equação 4) e (Equação 5), sendo TP (número de verdadeiros positivos), TN (número de verdadeiros negativos), FP (número de falsos positivos) e FN (número de falsos negativos) valores obtidos no momento da inferência através da comparação do resultado obtido com o resultado esperado (Ground Truth).

$$Acurácia = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precisão(P) = \frac{TP}{TP+FP} \quad (2)$$

$$Sensibilidade(R) = \frac{TP}{TP+FN} \quad (3)$$

$$Average\ Precision = \sum_n (R_n - R_{n-1}) P_n \quad (4)$$

$$Mean\ Average\ Precision = \frac{1}{N} \sum_{i=1}^{n} AP_i \quad (5)$$

3 RESULTADOS E DISCUSSÕES

3.1 Fase 1

As imagens aqui utilizadas são oriundas de 52 veículos da frota contendo o hardware instalado, totalizando um montante de 43.381 imagens coletadas a uma resolução de 960 x 1280 pixels. Devido ao alto volume de dados, utilizou-se apenas uma janela de gravação de aproximadamente 30 minutos de cada automóvel. Foi escolhido o para-brisa dianteiro como local padrão para a instalação das câmeras, conforme mostra a Figura 7, uma vez que o local possui a visão completa do pavimento.

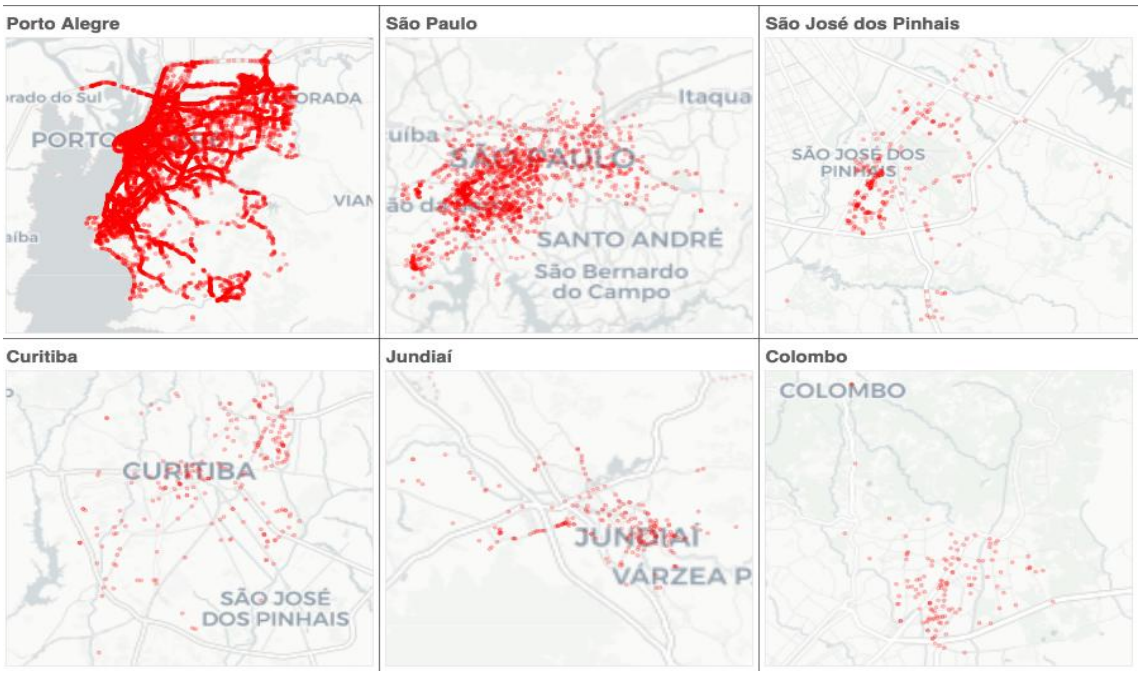
Figura 7: Posição de instalação da câmera.



Fonte: Autores (2025)

A Figura 8 detalha a distribuição espacial dos dados coletados, podendo ser notado a diversidade geográfica das capturas, aumentando assim a probabilidade de se aquisitar imagens de diversos tipos e condições de pavimentos.

Figura 8: Distribuição espacial dos dados coletados.



Fonte: Autores (2025)

As imagens foram integralmente disponibilizadas para três classificadores humanos devidamente treinados, sendo estes responsáveis por inserir as devidas marcações em cada amostra. A Figura 9 mostra um mosaico contendo uma amostra aleatória das imagens sem marcações.

Figura 9: Imagens brutas coletadas em campo.



Fonte: Autores (2025)

A etapa de classificação manual gerou 23.807 amostras anotadas com 24.022 classificações, representando 55% dos dados coletados, uma vez que imagens repetidas, com reflexo ou contendo outras anomalias foram descartadas no momento da verificação.

A Figura 10 mostra um mosaico contendo uma amostra aleatória das imagens com marcações.

Figura 10: Imagens anotadas coletadas em campo.



Fonte: Autores (2025)

No tocante a distribuição de rótulos, pode-se constatar que há majoritariamente uma classificação por imagem, porém, vias com mais de um tipo de pavimento acaba sendo comuns em casos de intervenções pontuais (recomposição asfáltica em ruas de paralelepípedo) ou locais com faixa exclusiva de ônibus. A Tabela 1 mostra a distribuição supracitada.

Tabela 1: Quantidade de detecções por imagem.

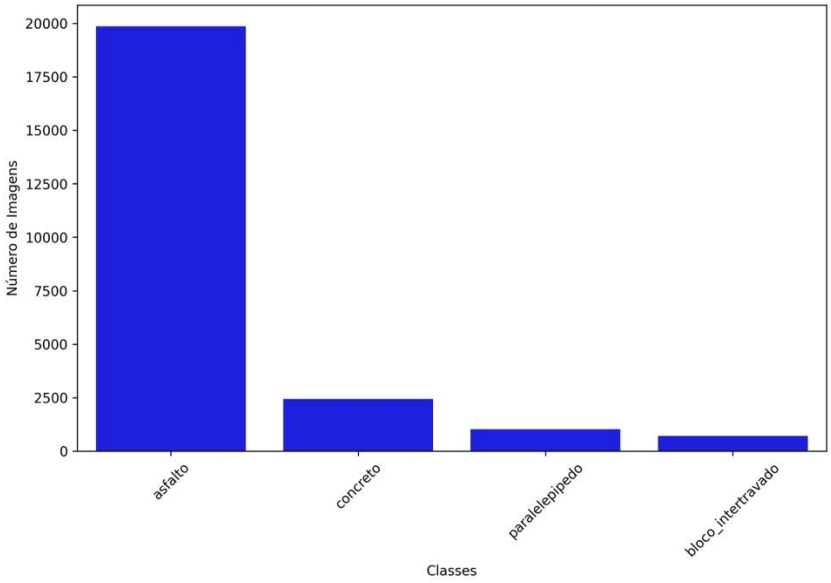
Número de Detecções por Imagem	Frequência
1	23.618
2	166
3	20
4	3

Fonte: Autores (2025)

Por fim, a Figura 11 apresenta a distribuição final de classes do dataset anotado, evidenciando um desbalanceamento entre as instâncias de interesse, o que é normal, uma vez que o pavimento de

revestimento asfáltico é majoritariamente utilizado em vias urbanas, sobrepondo a frequência de utilização de outros materiais nesses ambientes.

Figura 11: Distribuição de classes anotadas.

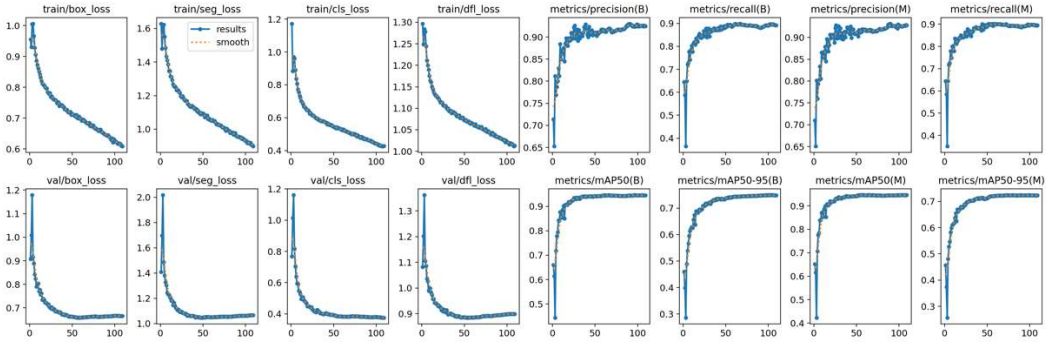


Fonte: Autores (2025)

3.2 Fase 2

O treino de máquina ocorreu com 109 épocas (early stop), conforme evidenciado na Figura 12, onde pode-se observar que houve a convergência, dado que as métricas analisadas nos conjuntos de validação e treino atingem valores assintóticos.

Figura 12: História do treinamento e validação de máquina.



Fonte: Autores (2025)

A Tabela 2 evidencia os hiper parâmetros utilizados nessa etapa.

Tabela 2: Valores dos hiperparâmetros de treino.

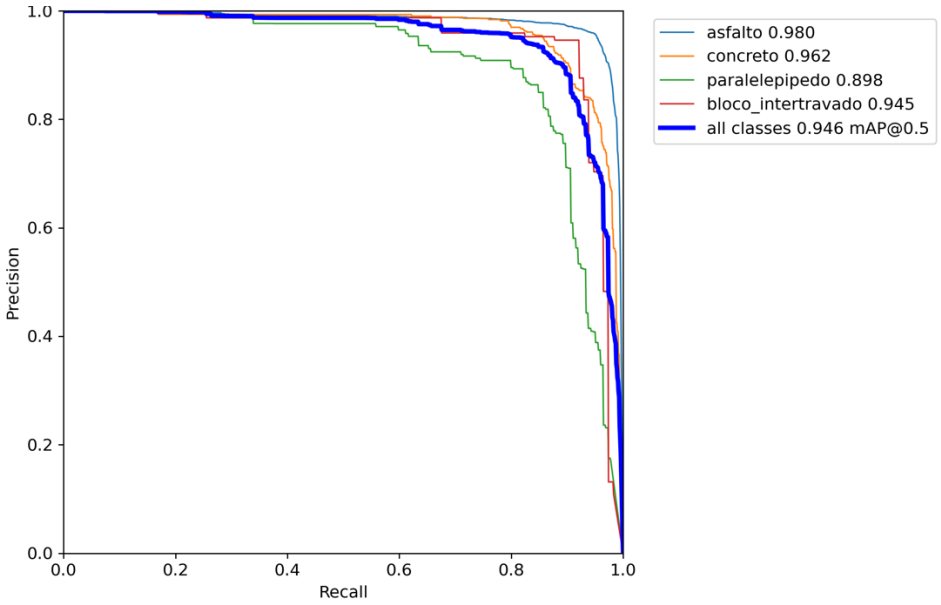
Hiperparâmetro	Valor
Batch	8
Épocas	200
Learning rate	0.01
Momentum	0.937
Weight decay	0.0005

Fonte: Autores (2025)

Outras curvas de treino foram analisadas para garantir que houve o aprendizado de máquina. A Figura 13 traz a precision recall curve, uma métrica importante para casos em que os dataset se encontra

desbalanceado, onde a área sob a curva (AUC) e o mAP (Mean Average Precision) de cada classe, próximos de 1, nos asseguram que o modelo esta performando bem.

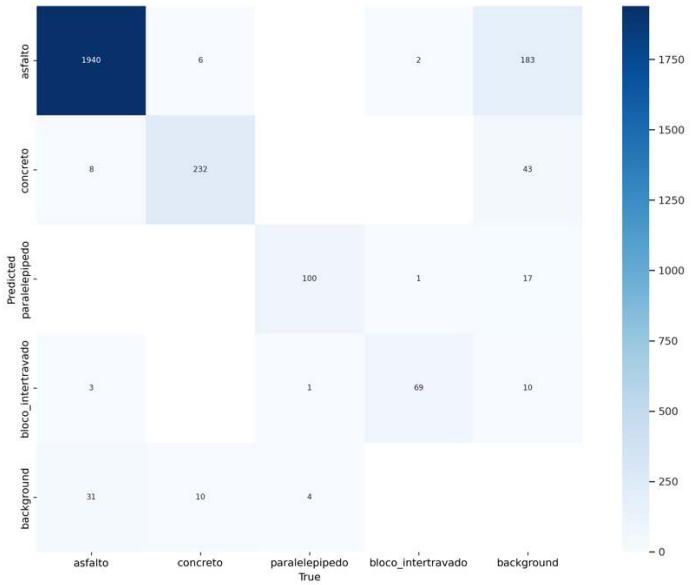
Figura 13: Precision Recall Curve.



Fonte: Autores (2025)

Por fim, com o modelo treinado foi realizado uma inferência para o conjunto de teste, buscando entender a assertividade real do sistema para amostras nunca analisadas. A Figura 14 mostra a matriz de confusão resultante desse processo.

Figura 14: Matriz de Confusão.



Fonte: Autores (2025)

4 CONSIDERAÇÕES FINAIS

O trabalho aqui realizado apresentou uma aplicação do modelo de inteligência artificial Yolov11 em um problema prático de segmentação do tipo de pavimento. A abordagem proposta se mostrou eficaz, uma vez que o modelo de inteligência artificial conseguiu “aprender” todos os tipos de pavimento pré-definidos a partir das imagens coletadas em campo. A criação de um dataset personalizado para pavimentos urbanos

possibilitou a generalização do modelo, mitigando assim os possíveis vieses, dado que a representatividade do conjunto de dados é de fato plural.

O modelo de criação de inventário de pavimentos proposto nesse trabalho apresenta custos reduzidos e alta periodicidade de coleta em relação aos métodos tradicionais, acarretando que a implementação desse modelo de inteligência artificial tem o potencial de impactar a forma como gestores públicos administram e monitoram suas vias e rodovias, possibilitando a geração de um banco de dados do pavimento urbano preciso e atualizado de forma automática.

Trabalhos futuros podem explorar a adaptação do modelo em regiões não previstas, como leitos naturais, assim como a implementação desse processo em larga escala em uma cidade como estudo de caso.

REFERÊNCIAS

ALZAMZAMI, O.; BABOUR, A.; BAALAWI, W.; AL KHUZAYEM, L. PDS-UAV: A Deep Learning-Based Pothole Detection System Using Unmanned Aerial Vehicle Images. Sustainability, v. 16, n. 21, 2024. Disponível em: https://doi.org/10.3390/su16219168. Acesso em: 30 mar. 2025.

FURTADO, Lara Sucupira; BESSA, Iuri Sidney; GURJÃO, Nayara de Oliveira; SOARES, Jorge Barbosa. Integrating smart city technologies for sustainable pavement infrastructure. Canadian Journal of Civil Engineering, 2024. Disponível em: https://doi.org/10.1139/cjce-2024-0217. Acesso em: 30 mar. 2025.

HORITA, Yuukou; KAWAI, Shohei; FURUKANE, Tatsuya; SHIBATA, Keiji. Efficient distinction of road surface conditions using surveillance camera images in night time. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, v. 19, 2012, p. 485–488. Disponível em: https://doi.org/10.1109/ICIP.2012.6466902. Acesso em: 30 mar.

JEGHAM, Nidhal; KOH, Chan Young; ABDELATTI, Marwan; HENDAWI, Abdeltawab. Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors. 2024. Pré-publicação. Disponível em: https://doi.org/10.48550/arXiv.2411.00201. Acesso em: 30 mar. 2025.

LEE, Jae Kwan; KIM, Bo Kyeong; CHOI, Hosik; CHANG, Seo Il. Road-pavement classification by artificial neural network model based on tire-pavement noise and road-surface image. Applied Acoustics, v. 225, p. 110-194, 2024. Disponível em: https://doi.org/10.1016/j.apacoust.2024.110194. Acesso em: 30 mar. 2025.

LIMA, Gabriel de Souza. Detecção de anomalias em pavimentos utilizando sensores inerciais. Dissertação (Mestrado) – Universidade de São Paulo, São Paulo, 2024. Disponível em: https://www.teses.usp.br/teses/disponiveis/3/3139/tde-26032024-111626/pt-br.php. Acesso em: 30 mar. 2025.

NIE, M.; WANG, K. Pavement Distress Detection Based on Transfer Learning. In: INTERNATIONAL CONFERENCE ON SYSTEMS AND INFORMATICS (ICSAI), v. 5, 2018, p. 435–439. Disponível em: https://doi.org/10.1109/ICSAI.2018.8599473. Acesso em: 30 mar. 2025.

PAIXÃO, Marcio; MARANHÃO, Flávio. Geovista – Gestão descentralizada das condições do pavimento em ambiente urbano / Geovista - Decentralized management of pavement conditions in an urban environment. Brazilian Journal of Development, v. 6, n. 7, 2020. Disponível em: https://doi.org/10.34117/bjdv6n7-825. Acesso em: 30 mar. 2025.

TIGHE, Susan; LI, Ningyuan; FALLS, Lynne Cowe; HAAS, Ralph. Incorporating Road Safety into Pavement Management. Transportation Research Record, v. 1699, n. 1, p. 1–10, 2000. Disponível em: https://doi.org/10.3141/1699-01. Acesso em: 30 mar. 2025.

ULTRALYTICS. YOLOv11 Documentation. 2024. Disponível em: https://docs.ultralytics.com/models/yolo11/. Acesso em: 30 mar. 2025.

WEB OF SCIENCE. Resumo de resultados de busca. Disponível em: https://www.webofscience.com/wos/woscc/summary/8bf0b322-3102-4e7a-b142-b5cc32eb7586-01567b8bc6/714968b0-4c62-461b-8db4-4320785e8e5a-01567b8bb8/relevance/1. Acesso em: 30 mar. 2025.