

Detecção e Contagem de Ativos Verdes: Um Estudo Comparativo Usando YOLOv8, YOLOv11, RT-DETR e Faster R-CNN

Maxwell Pires Silva¹, Italo Francyles Santos da Silva¹, João Dallyson Sousa de Almeida¹,
Anderson Lopes Silva¹, Joana Kuelvia de Araujo Silva¹, João Pedro Gomes de Souza¹,
Anselmo Cardoso de Paiva¹, Laura Giovanna Freitas Castro², Luan Felipe Araujo Sousa²

¹ Núcleo de Computação Aplicada, Universidade Federal do Maranhão,
Av. dos Portugueses 1966 São Luís, Maranhão, Brasil

²Instituto de Ciencia e Tecnologia Grupo Equatorial, São Luís, Maranhão, Brazil

{maxwell.pires, francyles, jdallysson, anderson.silva}@nca.ufma.br,
{joana.silva, joao.souza, paiva}@nca.ufma.br,
laura.castro@equatorialenergia.com.br, luan.sousa@eqtlab.com.br

Abstract. *Urban green assets are relevant to climate resilience and well-being, but they may also pose operational risks when vegetation grows close to overhead power lines. This work evaluates the automatic detection and counting of green assets in street-level images through a comparative study involving YOLOv8, YOLOv11, RT-DETR, and Faster R-CNN. The experiments compare the original dataset and an extended version with additional annotations. The obtained results indicate that YOLOv11 achieved the best overall performance in both detection (mAP_{50} and Dice) and counting (MAE and RMSE), suggesting its suitability for urban vegetation monitoring applications.*

Resumo. *Ativos verdes urbanos são relevantes para a resiliência climática e o bem-estar, mas também podem gerar riscos operacionais quando a vegetação se aproxima de redes aéreas de energia. Este trabalho avalia a detecção e contagem automática de ativos verdes em imagens em nível de rua por meio de um estudo comparativo entre YOLOv8, YOLOv11, RT-DETR e Faster R-CNN. Os experimentos comparam o dataset original e uma versão ampliada com novas anotações. Os resultados obtidos indicam melhor desempenho global do YOLOv11 tanto em detecção (mAP_{50} e Dice) quanto em contagem (MAE e RMSE), sugerindo sua adequação para aplicações de monitoramento da vegetação urbana.*

1. Introdução

A aceleração do crescimento urbano intensifica desafios ambientais e sociais diretamente ligados à qualidade de vida, como ilhas de calor, poluição atmosférica, aumento do risco de inundações e maior exposição a eventos climáticos extremos. Nesse cenário, os chamados *ativos verdes*, como árvores de rua, parques e demais infraestruturas verdes, assumem papel estratégico por contribuírem para serviços ecossistêmicos relevantes ao bem-estar

e à saúde pública. Evidências compiladas pela OMS (Organização Mundial da Saúde) indicam que espaços verdes urbanos se associam a múltiplos benefícios em saúde e bem-estar [OMS 2017]. De modo complementar, o Painel Intergovernamental sobre Mudanças Climáticas destaca o potencial de soluções baseadas na natureza, como áreas verdes e infraestrutura verde-azul, para reduzir riscos urbanos e gerar co-benefícios sociais e ambientais [IPCC 2022].

Para que esses benefícios sejam sustentáveis ao longo do tempo, cidades precisam implementar rotinas de monitoramento e gestão da vegetação urbana baseadas em dados confiáveis, atualizados e espacialmente detalhados. Revisões sobre inventários arbóreos urbanos apontam que levantamentos de campo e medições diretas são fundamentais, mas podem demandar tempo, equipe especializada e custos recorrentes, especialmente quando há necessidade de atualização frequente e cobertura ampla [Nielsen et al. 2014]. Assim, cresce a demanda por abordagens digitais e escaláveis, alinhadas à perspectiva de *smart urban forestry*, que integra sensoriamento, dados e algoritmos para apoiar decisões de manejo e planejamento urbano [Nitoslawski et al. 2019].

Além dos benefícios ambientais, a vegetação urbana interage diretamente com infraestruturas críticas, em especial redes aéreas de distribuição e transmissão de energia elétrica. A proximidade de galhos e árvores aos condutores pode provocar desligamentos e interrupções no fornecimento, tornando o manejo de vegetação uma atividade recorrente e de alto impacto operacional para concessionárias. Estudos em redes de distribuição mostram que práticas mais rigorosas de poda e remoção em faixas próximas às linhas podem reduzir interrupções, tanto em análises estatísticas quanto em modelos de predição de falhas associadas a eventos climáticos [Taylor et al. 2022]. Outros trabalhos destacam que o risco relacionado à vegetação depende de variáveis locais, como estrutura das árvores, proximidade à rede e características do corredor elétrico [Wedagedara et al. 2023], havendo também evidências de conflito recorrente entre arborização urbana e rede elétrica no contexto brasileiro [Sant Ana et al. 2022].

Nesse contexto, abordagens baseadas em visão computacional e aprendizado profundo surgem como alternativas promissoras para automatizar a identificação de ativos verdes em imagens urbanas. O trabalho de [Jodas et al. 2022] discute que imagens em nível de rua constituem uma alternativa custo-efetiva para monitoramento urbano, pois permitem observar árvore, copa e tronco sob uma perspectiva próxima àquela utilizada em inspeções terrestres. Ele propõe a detecção de árvores e de seus componentes estruturais em imagens *street-view*, utilizando uma rede do tipo YOLO com *backbone* MobileNet, além de um procedimento para estimar altura a partir de um objeto de referência.

Apesar dos avanços existentes, ainda há espaço para avaliar como arquiteturas recentes e distintas de detecção de objetos se comportam no problema de detecção e contagem de ativos verdes urbanos. Em particular, a literatura ainda carece de comparações que considerem não apenas a localização dos objetos, mas também o impacto das detecções na contagem de árvores e componentes, tarefa relevante para inventário urbano e apoio à gestão da vegetação próxima à infraestrutura elétrica. Assim, este trabalho compara YOLOv8, YOLOv11, RT-DETR e Faster R-CNN em dois cenários, um com o dataset original de [Jodas et al. 2022] outro com a ampliação e correção desse dataset com novas marcações, buscando identificar o modelo com melhor equilíbrio entre qualidade de detecção e desempenho de contagem.

2. Metodologia

Este trabalho foi conduzido como um estudo comparativo entre arquiteturas de redes neurais para detecção de objetos aplicadas à identificação e contagem de ativos verdes em ambiente urbano. A estratégia experimental consistiu em preparar o conjunto de dados, ampliar as anotações do dataset base, treinar diferentes modelos sob um protocolo comum e comparar os resultados por métricas de detecção e contagem.

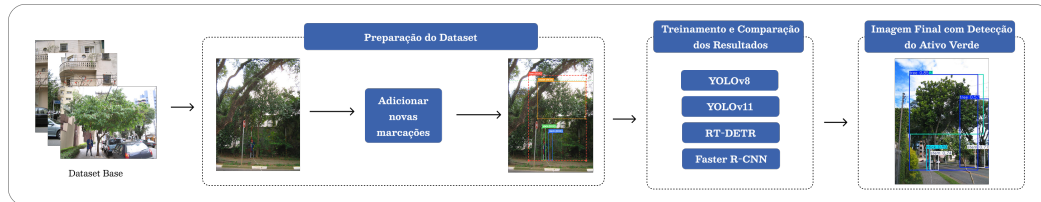


Figura 1. Visão geral do método proposto.

2.1. Base de dados

O conjunto de dados original [Jodas et al. 2022] é composto por imagens em nível de rua (*street-view*) contendo árvores urbanas e um objeto de referência (*stick*), com anotações manuais em caixas delimitadoras (*bounding boxes*). No trabalho base, foram anotadas quatro classes: *tree* (árvore), *crown* (copa), *stem* (tronco) e *stick* (régua de referência), totalizando 1434 imagens rotuladas. As anotações foram realizadas com a ferramenta LabellImg [Tzutalin 2015] e exportadas no formato Pascal.

Além das marcações originais, foi construída uma versão ampliada do dataset por meio da inclusão de novas anotações e do refinamento de caixas previamente existentes. Essa etapa teve como objetivo reduzir lacunas de anotação observadas no conjunto original, especialmente em imagens com instâncias visíveis de árvore, copa ou tronco sem a marcação correspondente, envolvendo assim a inclusão desses objetos não anotados. Para manter a compatibilidade com o dataset base, as novas anotações seguiram os mesmos critérios de marcação e os mesmos critérios semânticos das classes originais.

2.2. Arquiteturas avaliadas e protocolo experimental

Foram avaliadas quatro arquiteturas de detecção de objetos: YOLOv8, YOLOv11, RT-DETR e Faster R-CNN. A seleção buscou contemplar diferentes paradigmas de detecção, incluindo modelos *one-stage*, voltados à inferência eficiente, modelos baseados em Transformers e modelos *two-stage*, tradicionalmente associados ao refinamento de propostas de regiões. A inclusão simultânea de YOLOv8 e YOLOv11 permite verificar se a evolução recente da família YOLO resulta em ganhos práticos no problema analisado, enquanto RT-DETR e Faster R-CNN ampliam a comparação para arquiteturas com princípios distintos.

Todos os modelos foram treinados com 25 épocas e utilizando validação cruzada com 5 *Folds*. Esse valor foi adotado como equilíbrio entre custo computacional e padronização experimental, permitindo comparar as arquiteturas sob um mesmo padrão de treinamento. Para os modelos YOLOv8, YOLOv11 e RT-DETR, foram utilizadas as configurações padrão do Ultralytics, incluindo `imgsz=640` e `batch=16`. Para Faster R-CNN, utilizou-se a implementação do TorchVision com *backbone* ResNet-50-FPN.

Para avaliação da detecção, foram utilizadas Precisão, Recall, Dice e mAP_{50} . A Precisão mede a proporção de detecções corretas entre todas as predições, enquanto o Recall mede a proporção de objetos reais corretamente detectados. O coeficiente Dice avalia a consistência espacial entre predições e anotações de referência, sendo maior quando há melhor sobreposição entre as regiões. A métrica mAP_{50} corresponde à média da precisão média considerando IoU igual a 0,50, sendo amplamente utilizada para avaliar conjuntamente localização e classificação em detecção de objetos.

As métricas de contagem utilizadas foram MAE (*Mean Absolute Error*), RMSE (*Root Mean Square Error*), MBE (*Mean Bias Error*) e *Exact-count*. O MAE indica o erro médio absoluto entre a contagem estimada e a real. O RMSE também mede erro de contagem, mas penaliza mais fortemente erros grandes. O MBE indica tendência sistemática de superestimação ou subestimação, e o *Exact-count* [Wang 2025] representa o percentual de imagens em que o modelo acertou exatamente a contagem real.

3. Resultados e Discussão

Esta seção apresenta os resultados quantitativos obtidos nos datasets original e dataset ampliado com novas marcações. A análise considera métricas globais de detecção e contagem, seguida de uma comparação com o desempenho reportado no trabalho base.

Tabela 1. Resultados globais de detecção nos datasets original e com novas marcações.

Dataset	Modelo	Precisão	Recall	Dice	mAP_{50}
Original	YOLOv8	0.8278 ± 0.0087	0.7962 ± 0.0126	0.9243 ± 0.0078	0.8593 ± 0.0135
Original	YOLOv11	0.8334 ± 0.0255	0.8082 ± 0.0053	0.9297 ± 0.0065	0.8687 ± 0.0114
Original	RT-DETR	0.8353 ± 0.0198	0.7895 ± 0.0114	0.9221 ± 0.0093	0.8556 ± 0.0159
Original	Faster R-CNN	0.7903 ± 0.0314	0.8348 ± 0.0167	0.8106 ± 0.0174	0.8230 ± 0.0196
Ampliado	YOLOv8	0.8513 ± 0.0203	0.7983 ± 0.0194	0.9285 ± 0.0052	0.8667 ± 0.0091
Ampliado	YOLOv11	0.8546 ± 0.0115	0.8034 ± 0.0072	0.9318 ± 0.0052	0.8754 ± 0.0091
Ampliado	RT-DETR	0.8384 ± 0.0208	0.8022 ± 0.0181	0.9256 ± 0.0080	0.8616 ± 0.0139
Ampliado	Faster R-CNN	0.8015 ± 0.0181	0.8356 ± 0.0112	0.8175 ± 0.0114	0.8239 ± 0.0141

Tabela 2. Resultados globais de contagem nos datasets original e com novas marcações.

Dataset	Modelo	MAE	RMSE	MBE	<i>Exact-count</i> (%)
Original	YOLOv8	1.301	1.872	0.751	31.5
Original	YOLOv11	1.288	1.872	0.720	35.0
Original	RT-DETR	2.069	2.981	1.805	25.3
Original	Faster R-CNN	1.304	1.888	0.756	33.5
Ampliado	YOLOv8	1.200	1.708	0.692	32.7
Ampliado	YOLOv11	1.172	1.676	0.616	33.1
Ampliado	RT-DETR	2.644	3.734	2.512	21.6
Ampliado	Faster R-CNN	1.208	1.736	0.724	34.1

Nos resultados globais de detecção (Tabela 1), o YOLOv11 apresentou o melhor desempenho geral nos dois cenários avaliados, alcançando os maiores valores de Dice e mAP_{50} tanto no dataset original quanto no dataset com novas marcações. Esse comportamento indica maior equilíbrio entre localização e classificação dos ativos verdes. O Faster R-CNN obteve os maiores valores de Recall, sugerindo maior sensibilidade na recuperação de objetos, mas apresentou desempenho inferior em Dice e mAP_{50} , indicando

menor consistência espacial das detecções. O RT-DETR apresentou métricas competitivas de detecção, porém não superou o YOLOv11 nas métricas globais principais.

Em relação à contagem (Tabela 2), o YOLOv11 obteve os menores valores globais de MAE, RMSE e MBE no dataset com novas marcações, indicando contagens mais próximas dos valores reais e menor tendência de superestimação. A ampliação das anotações contribuiu para reduzir os erros de contagem em YOLOv8, YOLOv11 e Faster R-CNN, sugerindo que lacunas de anotação no dataset original impactavam negativamente a etapa de contagem. Apesar disso, a métrica *Exact-count* permaneceu relativamente baixa, o que evidencia a dificuldade da tarefa: pequenas variações na detecção de componentes, como copa e tronco, já são suficientes para tornar a contagem inexata.

3.1. Comparação com o trabalho base

Para contextualizar os resultados obtidos, o trabalho base reporta um desempenho médio em torno de $mAP_{50} \approx 0.81$ para configurações com YOLOv3, com e sem SPP (*Spatial Pyramid Pooling*), em validação cruzada com três *folds* [Jodas et al. 2022]. Comparando esses valores com os melhores resultados desta pesquisa, especialmente o YOLOv11 no dataset ampliado, observa-se aumento no mAP_{50} , sugerindo que arquiteturas mais recentes e a ampliação das marcações podem contribuir para maior qualidade de detecção em cenas urbanas.

Tabela 3. Desempenho reportado por [Jodas et al. 2022].

Modelo	mAP_{50}
YOLOv3 + MobileNet	0.81 ± 0.074
YOLOv3 + SPP + MobileNet	0.81 ± 0.084

De forma geral, os resultados reforçam que a escolha da arquitetura e a qualidade das anotações influenciam diretamente o desempenho da detecção e da contagem. O YOLOv11 apresentou o melhor resultado global, enquanto o Faster R-CNN mostrou maior Recall, mas menor consistência espacial. Já o RT-DETR, embora competitivo em detecção, apresentou maior erro de contagem, especialmente no dataset ampliado, indicando tendência de supercontagem em algumas condições.

Foram realizadas mais avaliações e os resultados completos, incluindo os por classe, um estudo de caso com imagens do *Street View* e o dataset ampliado com as novas marcações produzidas, foram disponibilizados em um repositório público.¹

4. Conclusão

Este trabalho apresentou um estudo comparativo de arquiteturas de detecção de objetos para detecção e contagem de ativos verdes em imagens urbanas em nível de rua. Foram avaliados YOLOv8, YOLOv11, RT-DETR e Faster R-CNN em dois cenários: dataset original e dataset ampliado com novas marcações. Os resultados indicaram melhor desempenho global do YOLOv11, especialmente nas métricas Dice e mAP_{50} , além de menores erros de contagem no dataset ampliado.

A ampliação das anotações contribuiu para melhorar a contagem e reduzir vieses em parte dos modelos, reforçando a importância da qualidade do dataset em tarefas de

¹<https://github.com/Maxw-11/Deteccao-e-Contagem-de-Ativos-Verdes>

detecção e contagem de componentes de árvores. Como trabalhos futuros, pretende-se ampliar a avaliação em novos conjuntos de imagens urbanas, investigar ajustes de limiares por classe e integrar os resultados a estratégias de priorização de risco em áreas próximas à infraestrutura elétrica.

5. Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil — Código de Financiamento 001; ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brasil; e à Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA), Brasil. Este trabalho foi apoiado pelo Instituto de Ciência e Tecnologia Grupo Equatorial e pelo Grupo Equatorial, por meio do programa de PDI ANEEL, sob o projeto nº 00037-0046/2022.

Referências

- IPCC (2022). Climate change 2022: Impacts, adaptation and vulnerability — chapter 6: Cities, settlements and key infrastructure. Technical report, Intergovernmental Panel on Climate Change (IPCC). Acesso em: 2026-02-02.
- Jodas, D. S., Yojo, T., Brazolin, S., Velasco, G. D. N., and Papa, J. P. (2022). Detection of trees on street-view images using a convolutional neural network. *International Journal of Neural Systems*, 32(1):2150042.
- Nielsen, A. B., Östberg, J., and Delshammar, T. (2014). Review of urban tree inventory methods used to collect data at single-tree level. *Arboriculture & Urban Forestry*, 40(2):96–111.
- Nitoslawski, S. A., Galle, N. J., Van Den Bosch, C. K., and Steenberg, J. W. (2019). Smarter ecosystems for smarter cities? a review of trends, technologies, and turning points for smart urban forestry. *Sustainable Cities and Society*, 51:101770.
- OMS (2017). Urban green space interventions and health: A review of impacts and effectiveness. Technical report, Organização Mundial da Saúde, Regional Office for Europe. Acesso em: 2026-02-02.
- Sant Ana, W. d. O., Faria, J. d., Zavarize, M. d. S., Pezente, D. P., Barbosa, V. d. C., Spacek, A. D., Vianna, M. P., and Junior, O. H. A. (2022). Methodology for defining priority locations for carrying out a forest inventory in points with conflict between urban tree planting and the electricity grid. *Energies*, 15(3):684.
- Taylor, W. O., Watson, P. L., Cerrai, D., and Anagnostou, E. (2022). A statistical framework for evaluating the effectiveness of vegetation management in reducing power outages caused during storms in distribution networks. *Sustainability*, 14(2):904.
- Tzutalin (2015). Labelimg. <https://github.com/tzutalin/labelImg>. Acesso em: 2026-02-08.
- Wang, S. (2025). Effectiveness of traditional augmentation methods for rebar counting using uav imagery with faster r-cnn and yolov10-based transformer architectures. *Scientific Reports*, 15(1):33702.
- Wedagedara, H., Witharana, C., Fahey, R., Cerrai, D., Joshi, D., and Parent, J. (2023). Modeling the impact of local environmental variables on tree-related power outages along distribution powerlines. *Electric Power Systems Research*, 221:109486.