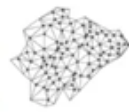




Modelagem Computacional no Sertão Mineiro



PPGMCS
Programa de Pós-Graduação em
Modelagem Computacional e Sistemas

MAPEAMENTO DE AGRUPAMENTOS NO TERRITÓRIO XAKRIABÁ: UMA ABORDAGEM DE DETECÇÃO DE OBJETOS COM VISÃO COMPUTACIONAL

Jhemerson Lincon Pereira da Silva¹ - jhemersonlincon00@gmail.com

Patrícia Takaki² - patricia.takaki@unimontes.br

¹Universidade Estadual de Montes Claros, UNIMONTES - Montes Claros, MG, Brazil

Abstract. A garantia dos Territórios Indígenas (TIs) é essencial para a preservação dos modos de vida das populações originárias. O povo Xakriabá, no norte de Minas Gerais, possui um profundo histórico de resistência territorial, tornando o monitoramento contínuo de suas aldeias fundamental para a gestão do território. Contudo, levantamentos convencionais frequentemente esbarram em barreiras logísticas. Nesse cenário, este trabalho propôs a aplicação de técnicas de visão computacional, utilizando a arquitetura YOLO e a biblioteca Ultralytics, para a detecção automatizada de aglomerações de residências na TI Xakriabá a partir de imagens de satélite. O modelo foi treinado durante 300 épocas, com um batch size de quatro imagens. Embora as métricas obtidas indiquem que ainda há limitações quanto à precisão das detecções, os resultados demonstram a viabilidade da abordagem para identificar padrões espaciais associados às localidades indígenas de forma não invasiva. Dessa forma, a integração entre aprendizado de máquina e sensoriamento remoto apresenta potencial para subsidiar ações de monitoramento territorial, fortalecendo a gestão das Terras Indígenas e apoiando a formulação de políticas públicas voltadas às comunidades indígenas.

Keywords: Território Xakriabá, Detecção de Objetos, Visão Computacional, Sensoriamento Remoto

1. INTRODUÇÃO

A população indígena Xakriabá mantém uma profunda relação histórica, cultural e espiritual com seu território, cuja demarcação foi resultado de décadas de mobilização e resistência em defesa de seus direitos territoriais (CIMI, 2024). Esse processo representou não apenas o reconhecimento legal de suas terras, mas também a reafirmação de sua identidade, a valorização de seus conhecimentos tradicionais e a continuidade de suas formas próprias de organização social.

Diante da importância desse território para a reprodução física, cultural e social do povo Xakriabá, a preservação de sua integridade e a garantia dos direitos conquistados dependem

do acompanhamento contínuo de sua dinâmica espacial. Nesse contexto, o monitoramento da ocupação das Terras Indígenas (TI) configura-se como uma atividade essencial, pois subsidia ações de planejamento territorial, gestão ambiental e proteção, permitindo compreender a distribuição das comunidades e acompanhar as transformações ocorridas ao longo do tempo (Brasil, 2012).

Entretanto, esse acompanhamento ainda depende, em grande parte, de levantamentos de campo e da interpretação manual, procedimentos que demandam tempo e recursos, dificultando sua aplicação em áreas extensas. É nesse cenário que o sensoriamento remoto, associado a técnicas computacionais, destaca-se por ampliar a capacidade de análise de imagens de satélite e subsidiar a tomada de decisão na gestão territorial (Silva et al., 2018).

Nos últimos anos, os avanços em inteligência artificial têm ampliado as possibilidades de automatização dessas análises (Zhu et al., 2017). Em especial, os modelos de visão computacional baseados em redes neurais convolucionais (CNNs) têm apresentado elevado desempenho na detecção de objetos em imagens de alta resolução. Entre esses modelos, a família YOLO (You Only Look Once) destaca-se pela capacidade de realizar inferências em tempo real, combinando elevada precisão e eficiência computacional (Redmon et al., 2016).

Diante desse cenário, este trabalho tem como objetivo desenvolver um modelo de visão computacional baseado na arquitetura YOLO para a detecção de aglomerações de residências nos territórios Xakriabá a partir de imagens de satélite, com o intuito de validar a capacidade do modelo em identificar agrupamentos em TI.

2. METODOLOGIA

Para a construção do conjunto de dados, foram utilizadas imagens de satélite da TI Xakriabá, localizada no município de São João das Missões, Minas Gerais (Martinez et al., 2002). Inicialmente, foi utilizado o software de código aberto QGIS como ambiente de Sistemas de Informação Geográfica (SIG) para aquisição das imagens, empregando o complemento QuickMapServices para acesso à camada Google Satellite, que fornece imagens de alta resolução (QGIS Development Team, 2026).

Como referencial espacial para a delimitação das áreas de interesse, utilizou-se a base vetorial de localidades indígenas fornecida pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Os pontos geográficos correspondentes a essas localidades atuaram como centróides para a exportação automatizada das imagens, processo conduzido por meio da ferramenta Atlas do QGIS. Para assegurar a padronização amostral e a mesma densidade de informações visuais em todas as extrações, fixou-se a escala cartográfica em 1:3.000. O procedimento culminou na consolidação de um conjunto de 43 imagens, representando os assentamentos mapeados no interior da Terra Indígena Xakriabá.

Após a aquisição das imagens, foi realizada a etapa de rotulagem manual utilizando a ferramenta LabelMe. As anotações consistiram na delimitação das aglomerações residenciais visíveis nas imagens de satélite, considerando como objeto de interesse conjuntos formados por, no mínimo, uma residência. Cada anotação foi realizada por meio de caixas delimitadoras (bounding boxes), gerando um conjunto de rótulos compatível com o treinamento do modelo de detecção de objetos. Ao término dessa etapa, o conjunto de dados foi composto por 43 imagens acompanhadas de seus respectivos arquivos de anotação.

Finalizada a etapa de demarcação, o conjunto de dados e suas respectivas anotações, foram redimensionado para a resolução de 1024×1024 pixels, para evitar perda de informações.

Além disso, para avaliar a capacidade de generalização da rede neural, o dataset foi dividido em dois subconjuntos, alocando-se 80% das amostras para o treinamento e os 20% restantes para a validação, resultando em 34 imagens para treinamento e 9 para validação.

Para a construção do modelo de detecção, utilizou-se a plataforma Google Colaboratory (Colab), que disponibiliza uma máquina virtual composta por uma CPU Intel Xeon com um núcleo e clock de 2 GHz, 13 GB de memória RAM, 78 GB de armazenamento e uma placa de vídeo Nvidia Tesla T4 com 15 GB de VRAM. Além disso, foi empregada a biblioteca Ultralytics (Jocher, 2023), amplamente reconhecida por suas implementações da família YOLO (You Only Look Once). A escolha específica da variante YOLOv8n justifica-se por sua arquitetura de estágio único (one-stage detector), que garante um excelente compromisso entre acurácia e custo computacional.

Em relação à etapa de treinamento, o modelo foi executado durante 300 épocas (epochs), utilizando um tamanho de lote (batch size) de 4 imagens, considerando a resolução de entrada de 1024×1024 pixels. Com o intuito de aumentar a variabilidade dos dados e contribuir para a capacidade de generalização do modelo, foram aplicadas técnicas de aumento de dados (data augmentation), por meio de rotações aleatórias entre -15° e $+15^\circ$ e translações horizontais e verticais de até 10%. Essas transformações foram utilizadas como estratégia para reduzir a possibilidade de sobreajuste (overfitting) durante o treinamento.

3. RESULTADOS E DISCUSSÕES

O modelo desenvolvido para a detecção de aglomerações residenciais apresentou resultados promissores na identificação automática de localidades indígenas em imagens de satélite, demonstrando a viabilidade da aplicação de modelos de detecção no monitoramento ocupacional. A Tabela 1 apresenta os valores obtidos para as métricas calculadas sobre o conjunto de validação.

Table 1: Desempenho do modelo YOLOv8 na base de validação.

Métricas	Valor
Precisão	55%
Recall	50%
mAP ₅₀	38%
mAP _{50–95}	12,5%

Como pode ser observado, a Precisão atingiu aproximadamente 55%. Esse valor indica que, na maioria das vezes em que o modelo gerou uma caixa delimitadora, ela efetivamente correspondia a uma aglomeração residencial. Embora demonstre uma capacidade de isolar áreas habitadas da vegetação, a métrica também revela uma incidência moderada de falsos positivos, com detecções ocorrendo em regiões sem a presença de aglomerações de residências.

O Recall alcançou cerca de 50%, evidenciando que metade das aglomerações presentes nas imagens não foram identificadas durante a inferência. Esse volume de falsos negativos sugere que o modelo apresenta dificuldades em generalizar a detecção para localidades com menor dimensão, baixa densidade construtiva ou configurações espaciais altamente esparsas, características comuns na organização territorial dessas áreas.

Em relação ao mAP₅₀, o modelo obteve aproximadamente 38%, demonstrando que apesar do baixo desempenho, o modelo conseguiu aprender padrões suficientes para diferenciar

possíveis aglomerações quando adotado um limiar de 50% do Intersection over Union (IoU). Esse resultado evidencia que, mesmo quando a localidade é identificada, a posição e o tamanho das caixas delimitadoras nem sempre coincidem adequadamente com as anotações de referência.

O mAP_{50-95} apresentou valor aproximado de 12,5%, resultado esperado devido ao caráter mais rigoroso dessa métrica, que considera limiares de IoU variando entre 50% e 95%. A diferença em relação ao mAP_{50} demonstra que o modelo possui maior dificuldade em produzir delimitações espacialmente precisas, especialmente em localidades com formatos irregulares ou maior concentração de residências.

A Figura 1 apresenta uma comparação entre as anotações originais do conjunto de dados e as predições produzidas pelo modelo. É possível observar que o modelo YOLOv8 foi capaz de identificar corretamente duas das quatro aglomerações residenciais presentes na imagem, atribuindo níveis de confiança superiores a 0,50 às detecções realizadas. Entretanto, localidades menores e agrupamentos mais extensos não foram detectados, evidenciando que o modelo apresenta maior dificuldade em representar regiões com elevada heterogeneidade espacial.

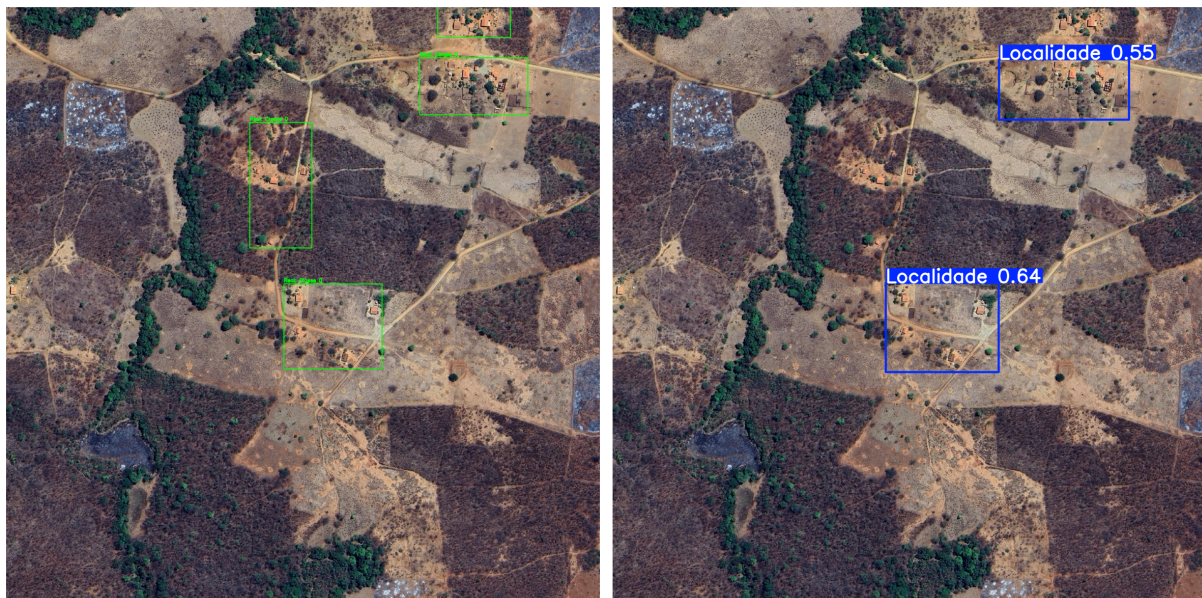


Figure 1: Comparação entre as anotações do conjunto de dados e as predições do modelo.

A análise conjunta dos resultados quantitativos e qualitativos indica que o modelo foi capaz de aprender características visuais relacionadas às aglomerações residenciais, mas ainda apresenta limitações na identificação de todos os objetos presentes e, principalmente, na delimitação espacial precisa dessas áreas. Dessa forma, os resultados obtidos devem ser interpretados como uma avaliação inicial da aplicabilidade da abordagem, e não como uma medida definitiva de sua capacidade de generalização.

Uma importante limitação do estudo está relacionada ao reduzido tamanho da base de dados, composta por 43 imagens. A quantidade limitada de amostras restringe a diversidade de características espaciais e visuais disponíveis para o aprendizado, o que pode dificultar a identificação de localidades com configurações distintas daquelas observadas durante o treinamento. Essa limitação é particularmente relevante considerando a heterogeneidade das aglomerações residenciais, que podem apresentar diferenças quanto ao número de edificações, distribuição espacial, densidade e extensão territorial.

4. CONCLUSÃO

Este trabalho apresentou uma abordagem baseada em visão computacional para a detecção automática de aglomerados de moradias no território indígena Xakriabá utilizando o modelo YOLOv8n. Os resultados obtidos demonstraram que, embora o desempenho quantitativo ainda seja limitado, o modelo foi capaz de aprender padrões visuais associados às localidades indígenas, identificando parte dos agrupamentos residenciais presentes nas imagens de satélite. As métricas obtidas evidenciam que ainda existem desafios relacionados à identificação de localidades com maior heterogeneidade espacial e à delimitação precisa das regiões detectadas. Esses resultados estão diretamente relacionados ao tamanho reduzido do conjunto de dados utilizado no treinamento, composto por apenas 43 imagens, o que restringe a capacidade de generalização do modelo. Apesar dessas limitações, o estudo demonstra a viabilidade da aplicação de técnicas de visão computacional como ferramenta de apoio ao monitoramento territorial de Terras Indígenas, reduzindo a dependência exclusiva da interpretação manual de imagens e contribuindo para processos de planejamento territorial e gestão ambiental. Como trabalhos futuros, pretende-se ampliar o conjunto de dados por meio da aquisição e anotação de novas imagens, empregar técnicas de aumento de dados (data augmentation) e investigar estratégias que permitam detectar moradias individuais, visando aumentar a precisão e a capacidade de generalização do modelo.

Acknowledgements

Os autores agradecem ao Programa Institucional de Bolsas de Iniciação Científica da Unimontes Ações Afirmativas (BIC-UNI/AF) e ao Programa de Pós-Graduação em Modelagem Computacional e Sistemas (PPGMCS) pela infraestrutura disponibilizada.

REFERENCES

- Brasil (2012), *Decreto nº 7.747, de 5 de junho de 2012. Política Nacional de Gestão Territorial e Ambiental de Terras Indígenas - PNGATI*, Diário Oficial da União, Brasília.
- Conselho Indigenista Missionário (2024), *Memória Xakriabá*, CIMI. Disponível em: <https://cimi.org.br/2024/02/memoria-xakriaba/>.
- QGIS Development Team (2026), *QGIS Geographic Information System*, Open Source Geospatial Foundation Project. Disponível em: <https://qgis.org>.
- Jocher, G. (2023), *Ultralytics YOLOv8* (Versão 8.0). Disponível em: <https://github.com/ultralytics/ultralytics>.
- Martinez, R., Vitali, L.H., Henriques, J.H.S., Machado, A.A., Albernaz, A., e Lima, A.A. (2002), *Inquérito soropidemiológico para infecções por fungos causadores de micoses sistêmicas na Reserva Indígena Xacriabá, Estado de Minas Gerais*, Revista da Sociedade Brasileira de Medicina Tropical, v. 35, n. 4, p. 347-350. DOI: 10.1590/S0037-86822002000400011.
- Redmon, J., Divvala, S., Girshick, R., e Farhadi, A. (2016), *You Only Look Once: Unified, Real-Time Object Detection*, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 779-788.
- Silva, T., Silva, M., Silva, D.T.A., Serrão, E., e Margalho, E. (2018), *Aplicações e Princípios do Sensoriamento Remoto*, Atena Editora, Ponta Grossa. ISBN: 978-85-85107-48-2.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., e Fraundorfer, F. (2017), *Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources*, IEEE Geoscience and Remote Sensing Magazine, v. 5, n. 4, p. 8-36.