UMA ABORDAGEM BASEADA EM YOLO PARA A DETECÇÃO E SEGMENTAÇÃO DE TUMORES ÓSSEOS EM IMAGENS RADIOGRÁFICAS

MARCELO DA SILVA DIAS[1]; MARINA DE SOUZA BRUM[2]
MARILTON SANCHOTENE DE AGUIAR[3]

[1] *Universidade Federal de Pelotas – marcelo.sdias@inf.ufpel.edu.br*
[2] *Universidade Federal de Pelotas – marina.sbrum@inf.ufpel.edu.br*
[3] *Universidade Federal de Pelotas – marilton@inf.ufpel.edu.br*

1. INTRODUÇÃO

Os ossos desempenham papel fundamental na sustentação e estruturação do corpo humano e podem ser afetados por diferentes condições patológicas, entre as quais se destacam os tumores ósseos (SAKDEJAYONT et al., 2025). Essas neoplasias constituem um grupo heterogêneo caracterizado pelo crescimento anormal e desordenado de células, resultando na formação de massas tumorais (HOSSEINI et al., 2025). Podem ser classificadas como primárias, quando se originam no próprio tecido ósseo e estão frequentemente associadas a alterações genéticas e moleculares, ou secundárias, quando decorrem de metástases provenientes de neoplasias localizadas em outras regiões do organismo (HOSSEINI et al., 2025).

Os tumores ósseos benignos frequentemente não apresentam sintomas evidentes, sendo muitas vezes identificados incidentalmente durante exames realizados por outras razões (SUBRAMANIAN; VISWANATHAN, 2026). Em contrapartida, os tumores malignos tendem a apresentar comportamento mais agressivo, podendo comprometer a integridade estrutural do tecido ósseo e afetar a qualidade de vida dos pacientes, causando dor, fraturas patológicas, fraqueza e limitações funcionais. Dados estatísticos indicam que os tumores benignos representam cerca de 79,3% das lesões diagnosticadas, com leve predominância no sexo feminino, enquanto os malignos correspondem a aproximadamente 20,7% e apresentam maior incidência no sexo masculino (SUBRAMANIAN; VISWANATHAN, 2026).

A avaliação de tumores ósseos em radiografias depende da análise de características morfológicas, como localização, forma, tamanho, margens, reação periosteal e mineralização (MATCUK et al., 2025). Nesse contexto, técnicas de aprendizado profundo podem auxiliar na identificação de padrões associados às lesões e ser empregadas na detecção e segmentação de tumores em imagens radiográficas, permitindo localizar e delimitar as regiões afetadas.

Diante desse cenário, este trabalho propõe a utilização de um modelo baseado na arquitetura *You Only Look Once* (YOLO) para a detecção e segmentação de tumores ósseos em imagens radiográficas. Para isso, é utilizado o conjunto de dados *Bone Tumor X-ray Radiograph Dataset* (BTXRD) (YAO et al., 2025), composto por 3.746 imagens de radiografias ósseas. O conjunto possui anotações por caixas delimitadoras e máscaras de segmentação, permitindo sua utilização nas duas tarefas. O modelo é treinado e avaliado quanto à capacidade de localizar e delimitar as regiões tumorais, buscando contribuir para o desenvolvimento de métodos computacionais aplicados à análise de imagens médicas.

2. METODOLOGIA

Este trabalho adotou uma metodologia estruturada nas seguintes etapas: (1) levantamento do referencial teórico e seleção do conjunto de dados; (2) preparação

do conjunto de dados; (3) remoção parcial do fundo preto; (4) treinamento do modelo; e (5) análise dos resultados obtidos.

Inicialmente, foi realizada uma revisão da literatura sobre detecção e segmentação de tumores ósseos em imagens radiográficas, abrangendo aprendizado profundo, redes neurais convolucionais, pré-processamento e técnicas de detecção e segmentação. A partir desse levantamento, identificou-se o uso recorrente do conjunto BTXRD em trabalhos recentes, que foi selecionado para este estudo por sua relevância na área e pela disponibilidade de anotações para detecção e segmentação. O conjunto é composto por 3.746 imagens, sendo 1.879 de ossos normais, 1.525 de tumores benignos e 342 de tumores malignos, com imagens tumorais anotadas por caixas delimitadoras e máscaras de segmentação.

Uma limitação do BTXRD é a ausência de um identificador individual de paciente na planilha de informações das imagens, o que dificulta a divisão dos dados e pode favorecer *data leakage* caso imagens relacionadas sejam distribuídas entre treinamento, validação e teste. Para minimizar esse risco, os registros foram analisados considerando centro de origem, sexo, idade, localização anatômica, ângulo de aquisição e presença de tumor, buscando identificar imagens potencialmente relacionadas. Esses registros foram agrupados e os grupos utilizados como unidade na divisão dos dados, mantendo imagens potencialmente relacionadas no mesmo subconjunto e reduzindo a sobreposição entre os conjuntos.

Para reduzir as regiões de fundo sem comprometer informações anatômicas, foi desenvolvido um procedimento de *cropping*. Os limites das regiões anatômicas são definidos pela combinação da energia de banda passante, obtida nas imagens original e processada com *Contrast Limited Adaptive Histogram Equalization* (CLAHE) (ZUIDERVELD, 1994), e da amplitude robusta de intensidade nas linhas e colunas. Esses limites são confrontados com as caixas delimitadoras das lesões e submetidos a verificações de segurança, rejeitando o recorte quando houver risco de perda anatômica. Após o recorte, as anotações de detecção e segmentação são remapeadas para as novas dimensões da imagem.

Por fim, foram avaliadas as variantes nano das arquiteturas YOLOv8, YOLOv9, YOLOv11 e YOLOv26, implementadas pela Ultralytics (JOCHER; QIU; CHAURASIA, 2023). A arquitetura de melhor desempenho foi selecionada para os experimentos subsequentes com *cropping* e aumento de dados. Com o objetivo de ampliar a variabilidade das amostras e favorecer a capacidade de generalização, foram empregadas técnicas como *Mosaic*, *Copy-Paste*, *MixUp* e variações de brilho. O desempenho nas tarefas de detecção e segmentação foi avaliado por meio das métricas *precision*, *recall*, F1-*score* e *mean Average Precision* (mAP).

3. RESULTADOS E DISCUSSÃO

Os resultados obtidos pelos modelos YOLOv8, YOLOv9, YOLOv11 e YOLOv26 são apresentados na Tabela 1, considerando as tarefas de detecção e segmentação no conjunto de teste do BTXRD. A partir desses resultados, são analisados o desempenho dos modelos e o impacto das estratégias de *cropping* e aumento de dados nas métricas avaliadas. Os valores destacados em negrito correspondem aos melhores resultados obtidos entre os experimentos apresentados.

Inicialmente, foram avaliadas as arquiteturas YOLOv8, YOLOv9, YOLOv11 e YOLOv26 em suas configurações padrão, com o objetivo de selecionar o modelo de

Tabela 1: Resultados dos modelos avaliados nas tarefas de detecção (Bbox) e segmentação (Mask) no conjunto de teste do BTXRD.

Modelos		***Precision***	***Recall***	**F1-*score***	**mAP@50**
YOLOv8	Bbox	0,3670	0,2880	0,3227	0,2730
	Mask	0,3720	0,2760	0,3169	0,2640
YOLOv9	Bbox	0,3240	0,2720	0,2957	0,2530
	Mask	0,3190	0,2680	0,2913	0,2470
YOLOv11	Bbox	0,3970	0,3340	0,3628	0,3130
	Mask	0,4020	0,2980	0,3423	0,3070
YOLOv26	Bbox	0,3850	0,3110	0,3441	0,3120
	Mask	0,3590	0,2910	0,3214	0,2650
YOLOv11 + Cropping	Bbox	0,4300	0,3590	0,3913	0,3450
	Mask	0,4710	0,3910	0,4273	0,3650
YOLOv11 + Cropping + aumento de dados	Bbox	**0,5440**	**0,3990**	**0,4604**	**0,3970**
	Mask	**0,5430**	**0,3900**	**0,4540**	**0,3870**

referência. A YOLOv11 apresentou o melhor desempenho geral, com F1-*score* de 0,3628 na detecção e 0,3423 na segmentação. Dessa forma, a arquitetura foi selecionada por apresentar o melhor equilíbrio entre *Precision* e *Recall*, sendo utilizada nos experimentos subsequentes com *cropping* e aumento de dados.

Como próximo experimento, foi aplicada a estratégia de *cropping* para reduzir as regiões de fundo preto e concentrar o processamento nas áreas de maior interesse, reduzindo informações pouco relevantes para a tarefa. Conforme apresentado na Tabela 1, a técnica proporcionou ganhos em ambas as tarefas. Na detecção, o mAP@50 aumentou de 0,3130 para 0,3450 e o F1-*score* de 0,3628 para 0,3913, enquanto, na segmentação, o mAP@50 passou de 0,3070 para 0,3650 e o F1-*score* de 0,3423 para 0,4273. Esses resultados sugerem que a redução do fundo favoreceu a localização e a delimitação das lesões.

No último experimento, foi utilizada uma estratégia de aumento de dados no modelo YOLOv11 com *cropping*, buscando ampliar a variabilidade das amostras e reduzir a dependência de padrões específicos do treinamento, favorecendo a generalização. Conforme apresentado na Tabela 1, a estratégia proporcionou os melhores resultados. Na detecção, o mAP@50 aumentou de 0,3450 para 0,3970 e o F1-*score* de 0,3913 para 0,4604. Na segmentação, o mAP@50 passou de 0,3650 para 0,3870 e o F1-*score* de 0,4273 para 0,4540. Esses ganhos indicam que o aumento da variabilidade contribuiu para uma detecção e segmentação mais robustas, com melhor equilíbrio entre *Precision* e *Recall*.

4. CONCLUSÕES

Este trabalho avaliou modelos da família YOLO para a detecção e segmentação de tumores ósseos em imagens radiográficas do conjunto BTXRD. A YOLOv11 apresentou o melhor desempenho entre os modelos avaliados em suas configurações padrão e foi utilizada nos experimentos subsequentes. Os resultados demonstraram que a aplicação de *cropping* contribuiu para o desempenho do modelo, enquanto a incorporação de técnicas de aumento de dados proporcionou ganhos adicionais, alcançando mAP@50 de 0,3970 na detecção e 0,3870 na segmentação, com F1-*score* de 0,4604 e 0,4540, respectivamente.

Os resultados obtidos evidenciam o potencial da abordagem proposta para

a localização e delimitação de tumores ósseos em imagens radiográficas, destacando a importância do tratamento adequado das imagens e da ampliação da variabilidade dos dados no treinamento. Dessa forma, o estudo contribui para a investigação de métodos automatizados de apoio à análise de tumores ósseos, especialmente em tarefas conjuntas de detecção e segmentação. Como trabalhos futuros, pretende-se explorar variantes de maior capacidade da YOLO, considerando que os experimentos foram realizados com as variantes nano, além de investigar arquiteturas alternativas, como U-Net e modelos baseados em *Transformers*, estratégias de balanceamento de classes e outros conjuntos de dados, visando ampliar a robustez e a capacidade de generalização dos modelos.

5. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

6. REFERÊNCIAS BIBLIOGRÁFICAS

HOSSEINI, H.; HEYDARI, S.; HUSHMANDI, K.; DANESHI, S.; RAESI, R. Bone tumors: a systematic review of prevalence, risk determinants, and survival patterns. **BMC Cancer**, London, England, v. 25, n. 1, p. 321, 2025.

JOCHER, G.; QIU, J.; CHAURASIA, A. **Ultralytics YOLO**. Ultralytics, 2023. Disponível em: <https://github.com/ultralytics/ultralytics>. Acesso em: 18 set. 2026.

MATCUK, G. R.; WALDMAN, L. E.; FIELDS, B. K. K.; COLANGELI, M.; PALMAS, M.; RIGHI, A.; FILONZI, G.; CROMBé, A.; SPINNATO, P. Conventional radiography for the assessment of focal bone lesions of the appendicular skeleton: fundamental concepts in the modern imaging era. **Skeletal Radiology**, v. 54, n. 7, p. 1391–1406, 2025.

SAKDEJAYONT, S.; CHOBPENTHAI, T.; SUKSIRIVECHARUK, P.; NINATKIATTIKUL, I.-F.; POOSIRIPINYO, T. A review on bone tumor management: Cutting-edge strategies in bone grafting, bone graft substitute, and growth factors for defect reconstruction. **Orthopedic Research and Reviews**, v. 17, p. 175–188, 2025.

SUBRAMANIAN, S.; VISWANATHAN, V. K. **Lytic Bone Lesions**. Treasure Island, Florida: StatPearls Publishing, 2026.

YAO, S.; HUANG, Y.; WANG, X.; ZHANG, Y.; PAIXAO, I. C.; WANG, Z.; CHAI, C. L.; WANG, H.; LU, D.; WEBB, G. I.; LI, S.; GUO, Y.; CHEN, Q.; SONG, J. A radiograph dataset for the classification, localization, and segmentation of primary bone tumors. **Scientific Data**, Springer Science and Business Media LLC, Berlin/Heidelberg, Germany, v. 12, n. 1, jan. 2025. ISSN 2052-4463.

ZUIDERVELD, K. Contrast limited adaptive histogram equalization. In: HECKBERT, P. S. (Ed.). **Graphics Gems IV**. San Diego, CA: Academic Press Professional, 1994. p. 474–485. ISBN 0-12-336155-9.