

POTENCIAL E LIMITES DA NNU-NET NA SEGMENTAÇÃO DE HEMORRAGIAS INTRACRANIANAS EM TOMOGRAFIA COMPUTADORIZADA

Luiz Felipe Torminn Rocha Lima¹
Gabriel de Paula Barros Botelho¹
Gabriel Fernandes Carvalho¹
Daniel Araujo Debastiani¹
Henrique Valle de Lima¹

RESUMO

Introdução: A segmentação de hemorragias intracranianas em tomografias computadorizadas é um desafio crítico na radiologia devido à elevada morbimortalidade e à complexidade da interpretação manual. O aprendizado profundo surge como alternativa promissora, destacando-se a arquitetura nnU-Net, reconhecida pela adaptação a diferentes domínios de imagens médicas. **Objetivo:** Avaliar a aplicabilidade da nnU-Net na segmentação automática de hemorragias intracranianas, analisando o desempenho e as limitações de suas variantes em cenários clínicos heterogêneos. **Método:** Foram utilizados 388 exames de TC dos datasets BHSD, MICCAI, Murtadha e CQ500, este último complementado por 14 máscaras manuais. O pré-processamento incluiu conversão volumétrica 3D e binarização das máscaras. O treinamento foi conduzido em 250 épocas com validação cruzada 5-fold, explorando as variantes default, medium, large e extra-large. A avaliação utilizou Dice, IoU, precisão, sensibilidade e especificidade. **Resultados:** As variantes apresentaram desempenho próximo. A configuração Large obteve os melhores índices de sobreposição (Dice = 0,3940; IoU = 0,3189) e maior precisão (0,601). Já a versão Default apresentou a maior sensibilidade (0,711), favorecendo a detecção de lesões em detrimento da precisão (0,544). Os achados evidenciam o equilíbrio entre robustez arquitetural e custo computacional. **Conclusões:** Apesar dos resultados ficarem aquém de benchmarks internacionais (Dice > 0,90), devido ao tamanho limitado da base e à ausência de técnicas avançadas, o estudo confirma a viabilidade da nnU-Net em cenários heterogêneos. O trabalho reforça seu potencial como ferramenta de apoio radiológico e aponta para pesquisas futuras com datasets ampliados, mecanismos de atenção e arquiteturas híbridas.

Palavras-chave: Hemorragia Intracraniana; Tomografia Computadorizada; Aprendizado Profundo; Inteligência Artificial.

INTRODUÇÃO

A Radiologia, originada do latim “radius” (raio) e “logia” (estudo ou ciência), constitui um campo essencial da medicina voltado à aplicação de radiações ionizantes e não ionizantes para fins diagnósticos, terapêuticos e de pesquisa. Nesse contexto, a Tomografia Computadorizada (TC) é especialmente útil na avaliação craniana devido à alta resolução, detalhamento e ampla disponibilidade, tornando-se exame inicial em várias condições (MORAES et al., 2023). No âmbito dos acidentes vasculares encefálicos (AVEs), as hemorragias intracranianas destacam-se como importantes causas de morbidade e mortalidade, afetando cerca de 15 milhões de

¹ Universidade Evangélica de Goiás – UniEVANGÉLICA

peças anualmente e contribuindo para aproximadamente 5,5 milhões de óbitos relacionados a doenças cardiovasculares (MARGARIDO et al., 2021). Contudo, a interpretação manual dessas imagens é trabalhosa e suscetível a falhas, sobretudo em casos complexos.

Nesse cenário, o Aprendizado de Máquina (AM) surge como ferramenta promissora, capaz de reconhecer padrões em grandes volumes de dados e desenvolver modelos automatizados para identificar hemorragias em exames de TC. A interação entre AM e Medicina configura-se como campo em expansão, com aplicações que apoiam diagnósticos complexos e impulsionam soluções inovadoras (NAJJAR; BREESAM, 2023). O presente estudo busca contribuir para essa integração, com ênfase na avaliação do nnU-Net (ISENSEE et al., 2024), modelo derivado da U-Net que incorpora convoluções tridimensionais para análise volumétrica e uma pipeline automatizada de ajuste de parâmetros, partindo da premissa de que configurações adequadas são mais determinantes para o desempenho do que modificações arquiteturais (KOK et al., 2022).

Para a condução do estudo, foram empregados datasets open-source com anotações validadas, incluindo MICCAI, BHSD e Murtadha (WU et al., 2023; GOLDBERGER et al., 2020; LI et al., 2022). Adicionalmente, integrou-se o CQ500 (CHILAMKURTHY et al., 2018), originalmente sem máscaras de referência, para o qual foram confeccionadas anotações manuais por médicos em formação sob supervisão. Essa complementação ampliou a base de dados e sua representatividade.

MATERIAIS E MÉTODOS

O fluxo metodológico teve como propósito explorar comparativamente o desempenho de diferentes arquiteturas de redes neurais aplicadas à segmentação de hemorragias intracranianas em tomografias computadorizadas (TC). Para tanto, realizou-se a busca e seleção de bases de dados em repositórios públicos de ampla relevância, além de conjuntos provenientes de competições internacionais. Entre os datasets selecionados, o CQ500, disponibilizado originalmente sem máscaras de anotação, foi complementado com 14 máscaras manuais confeccionadas por alunos de Medicina sob supervisão especializada, ampliando a representatividade do conjunto. Dessa forma, compuseram a base final de análise os datasets MICCAI (100

exames de TC), BHSD (192 exames de TC), Murtadha (82 exames de TC) e CQ500 (14 exames de TC – confeccionados), totalizando 388 exames.

Após a consolidação, foi realizado o pré-processamento das imagens para padronização dos dados. Os volumes foram convertidos para o formato tridimensional, assegurando compatibilidade com a rede da nnUNet. Em seguida, aplicou-se a binarização das máscaras, uniformizando anotações. Esse processo conferiu maior homogeneidade ao dataset, maximizando a confiabilidade dos resultados.

A implementação foi realizada com a pipeline da nnU-Net, reconhecida pela capacidade de adaptação automática a diferentes bases de imagens médicas. Essa escolha se justificou pela robustez já demonstrada em benchmarks internacionais e sua aceitação como referência na literatura em segmentação volumétrica. O treinamento foi conduzido por 250 épocas, com os demais hiperparâmetros mantidos nos valores padrão da nnU-Net, assegurando comparabilidade com estudos prévios. A divisão dos dados seguiu a estratégia de validação cruzada 5-fold, utilizando 294 exames para treinamento, 74 para validação e 20 exames independentes para teste, garantindo avaliação final em dados não vistos.

Na etapa de validação, buscou-se analisar o comportamento do modelo frente a dados externos ao treinamento. Foram conduzidos experimentos comparativos entre as variantes *default*, *medium* e *large* da nnU-Net, avaliando o impacto da complexidade arquitetural sobre a acurácia da segmentação e os custos computacionais. A avaliação quantitativa utilizou métricas consolidadas na literatura de segmentação médica, incluindo Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precisão (Precision), Recall (Sensibilidade) e Especificidade, assegurando uma análise abrangente da sobreposição espacial e da capacidade de detecção dos modelos.

RESULTADOS

Os experimentos conduzidos com as diferentes variantes da arquitetura nnU-Net revelaram desempenhos distintos em relação às métricas de segmentação avaliadas. A configuração *Large* apresentou o melhor equilíbrio global, com Dice = 0,3940 e IoU = 0,3189, superando as variantes *Default* (Dice = 0,3817; IoU = 0,3023) e *Medium* (Dice = 0,3912; IoU = 0,3136). No que se refere à precisão e sensibilidade, observou-se comportamento complementar: o modelo *Default* atingiu a maior

sensibilidade (Recall = 0,711) com menor precisão (0,544), enquanto a *Large* alcançou a melhor precisão (0,601) com sensibilidade levemente inferior (0,704). A versão *Medium* posicionou-se de forma intermediária, com precisão de 0,583 e sensibilidade de 0,654.

De modo geral, os resultados indicam que o aumento da complexidade da rede — de *Default* para *Large* — resultou em ganhos modestos nas métricas de Dice e IoU, sem alterações expressivas no desempenho geral. Ainda assim, a maior sensibilidade da configuração *Default* sugere maior propensão à detecção de lesões, aspecto relevante em cenários clínicos que priorizam a identificação de pequenas hemorragias.

CONCLUSÃO

Os resultados demonstraram que a arquitetura nnU-Net apresentou desempenho consistente na segmentação de hemorragias intracranianas em tomografias computadorizadas, embora com métricas inferiores ao estado da arte. A variante *Large* alcançou o melhor equilíbrio (Dice = 0,3940; IoU = 0,3189; Precisão = 0,601), enquanto a *Default* destacou-se pela maior sensibilidade (Recall = 0,711), indicando maior capacidade de detecção de lesões, ainda que com menor precisão (0,544). Esses achados reforçam que a escolha da variante deve considerar o contexto clínico em que o modelo será utilizado.

Em comparação com estudos recentes, como Isensee et al. (2024) e Li et al. (2025), que reportaram métricas de Dice superiores a 0,90, os valores obtidos aqui foram mais modestos. Essa diferença pode ser atribuída ao tamanho reduzido da base de dados, à variabilidade entre os conjuntos utilizados e à ausência de técnicas avançadas de atenção ou balanceamento de classes. Ainda assim, a integração de diferentes datasets e a confecção manual de máscaras no CQ500 demonstraram a viabilidade da aplicação da nnU-Net em cenários heterogêneos e realistas, contribuindo para a compreensão de suas potencialidades e limitações no contexto nacional.

A principal contribuição deste estudo é a avaliação comparativa das variantes da nnU-Net, destacando o impacto da sensibilidade e da precisão na prática clínica. Entre as limitações, ressalta-se a base restrita e a ausência de validação externa. Futuramente, recomenda-se ampliar os datasets, empregar mecanismos de atenção

e integrar estratégias multimodais para elevar o desempenho e aproximar esses modelos da prática clínica, consolidando o aprendizado profundo como ferramenta de apoio à radiologia.

AGRADECIMENTOS

Agradecemos ao CNPq pelo apoio financeiro e à UniEVANGÉLICA pelo suporte institucional e pela viabilização das atividades que possibilitaram a realização deste projeto.

REFERÊNCIAS BIBLIOGRÁFICAS

¹CHILAMKURTHY, Sasank et al. **Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans**. [S.l.: s.n.], 2018. arXiv: 1803.05854 [cs.CV]. Disponível em: <<https://arxiv.org/abs/1803.05854>>.

²GOLDBERGER, A. L. et al. **PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals**. v. 101. [S.l.: s.n.], 2020. e215–e220. Circulation [Online]. RRID:SCR_007345. Disponível em: <<https://physionet.org/content/ct-ich/1.3.1/>>.

³ISENSEE, Fabian et al. nnU-Net Revisited: A Call for Rigorous Validation in 3D Medical Image Segmentation, 2024. arXiv: 2404.09556 [cs.CV]. Disponível em: <<https://arxiv.org/abs/2404.09556>>.

⁴KOK, Yong En et al. Semantic Segmentation of Spontaneous Intracerebral Hemorrhage, Intraventricular Hemorrhage, and Associated Edema on CT Images Using Deep Learning. **Radiology: Artificial Intelligence**, v. 4, n. 6, e220096, 2022. PMID: 36523645. DOI: 10.1148/ryai.220096. eprint: <https://doi.org/10.1148/ryai.220096>. Disponível em: <<https://doi.org/10.1148/ryai.220096>>.

⁵LI, Xiangyu et al. **The 2022 Intracranial Hemorrhage Segmentation Challenge on Non-Contrast head CT (NCCT)**. [S.l.]: Zenodo, mar. 2022. DOI: 10.5281/zenodo.6362221. Disponível em: <<https://doi.org/10.5281/zenodo.6362221>>.

⁶LIN, Xinxin et al. Advanced Multi-label Brain Hemorrhage Segmentation Using an Attention-Based Residual U-Net Model. **BMC Medical Informatics and Decision Making**, v. 25, n. 1, p. 286, 2025. ISSN 1472-6947. DOI: 10.1186/s12911-025-03131-3. Disponível em: <<https://doi.org/10.1186/s12911-025-03131-3>>.

⁷MARGARIDO, Adriano Júnior Lucarelli et al. Epidemiologia do Acidente Vascular Encefálico no Brasil. **Revista Eletrônica Acervo Científico**, v. 39, e8859–e8859, 2021.

⁸MORAES, Helisson Alexander et al. Tomografia computadorizada em acidente vascular encefálico. **Research, Society and Development**, v. 12, n. 6, e2712641858–e2712641858, 2023.

⁹NAJJAR, Esraa; BREESAM, Aqeel Majeed. Supervised Machine Learning a Brief Survey of Approaches. **Al-Iraqia Journal for Scientific Engineering Research**, v. 2, n. 4, p. 71–82, 2023.

¹⁰WU, Biao et al. **BHSD: A 3D Multi-Class Brain Hemorrhage Segmentation Dataset**. [S.l.: s.n.], 2023. arXiv: 2308.11298 [cs.CV]. Disponível em: <<https://arxiv.org/abs/2308.11298>>.