

Segmentação automática da próstata em imagens médicas com deep learning por meio da técnica de federated learning

Luciana Silva Albuquerque de Melo (aluna)¹, Mário Olímpio de Menezes (orientador)¹

¹ Instituto de Pesquisas Energéticas e Nucleares – IPEN-CNEN, São Paulo, SP, Brasil.

lucianasmfl@gmail.com, mo.menezes@gmail.com

Resumo. A inteligência artificial (IA) e a ciência de dados têm crescido significativamente, especialmente na área da saúde, moldando o futuro da medicina e da assistência médica. No entanto, a aplicação de IA na saúde enfrenta desafios, especialmente relacionados à privacidade dos dados e à necessidade de grandes volumes de dados para treinar modelos eficazes. Este estudo visa desenvolver um método automático de segmentação da próstata utilizando IA e a técnica de federated learning (FL), que permite a colaboração entre instituições mantendo os dados localmente, sem compartilhamento direto. O método foi avaliado em dois cenários: centralizado e federado, com o cenário federado apresentando desempenho ligeiramente superior. Ainda assim, como o trabalho é preliminar, são necessários mais estudos para validar os resultados e explorar técnicas adicionais de proteção de privacidade no contexto do aprendizado federado.

Abstract. Artificial intelligence (AI) and data science have grown significantly, especially in the healthcare sector, shaping the future of medicine and medical care. However, the application of AI in healthcare faces challenges, particularly related to data privacy and the need for large volumes of data to train effective models. This study aims to develop an automatic method for prostate segmentation using AI and the technique of federated learning (FL), which allows collaboration between institutions while keeping data local, without direct sharing. The method was evaluated in two scenarios: centralized and federated, with the federated scenario showing slightly better performance. Nevertheless, as this is a preliminary work, further studies are needed to validate the results and explore additional techniques for privacy protection in the context of federated learning.

Palavras-chave: Inteligência artificial na saúde; Segmentação automática da próstata; Aprendizado federado.

1. CONTEXTO E MOTIVAÇÃO

A IA está revolucionando a assistência médica. Pesquisas indicam que ela pode desempenhar um papel tão competente quanto, ou até superior, ao dos seres humanos em tarefas relacionadas à saúde, como o diagnóstico de doenças (1).

O uso de exames de imagem para prevenção, diagnóstico e tratamento de tumores prostáticos é crescente, superando os métodos mais convencionais, como os exames de toque retal e o teste de antígeno prostático específico (PSA). A tarefa de segmentação da próstata que envolve a separação do volume prostático de outras estruturas presentes na imagem, é geralmente realizada manualmente, no entanto, é moroso e propenso a diferentes interpretações entre especialistas (2). Uma alternativa eficaz ao método manual é a exploração de métodos automáticos para a segmentação de imagens digitais (3). Essa abordagem pode não apenas acelerar o processo, bem como contribuir para resultados mais consistentes na identificação e delimitação da próstata em imagens médicas.

Contudo, o emprego da IA em saúde, especialmente na aplicação de algoritmos de machine learning (ML), enfrenta grandes desafios, destacando-se a preocupação com a privacidade dos dados, pois de acordo com a Lei Geral de Proteção a dados (Lei nº 13.709/2018), os dados médicos são considerados dados pessoais sensíveis. Além disso, a necessidade de treinamento com grandes volumes de dados agrava a complexidade deste cenário. Uma alternativa para superar esses

desafios é uma abordagem inovadora em ML, o aprendizado federado ou federated learning (FL). Essa técnica permite a colaboração no aprendizado mantendo os dados em suas respectivas fontes, eliminando a necessidade de compartilhamento e, assim, preservando a privacidade dos dados, além de melhorar a qualidade dos modelos (4).

1.1 Objetivo

Este estudo tem como objetivo desenvolver um modelo generalizável de ML utilizando a técnica de FL, capaz de realizar a segmentação automática de imagens de ressonância magnética da região da próstata. Ao integrar essa abordagem, pretende-se contribuir significativamente para o desenvolvimento de estratégias de tratamento mais precisas, como a radioterapia, que preservem os órgãos adjacentes, além de fornecer uma base para futuras pesquisas e contribuir para o avanço contínuo na segmentação de imagens médicas e aprendizado de máquina.

2. ATIVIDADES PRINCIPAIS

As principais atividades desenvolvidas foram:

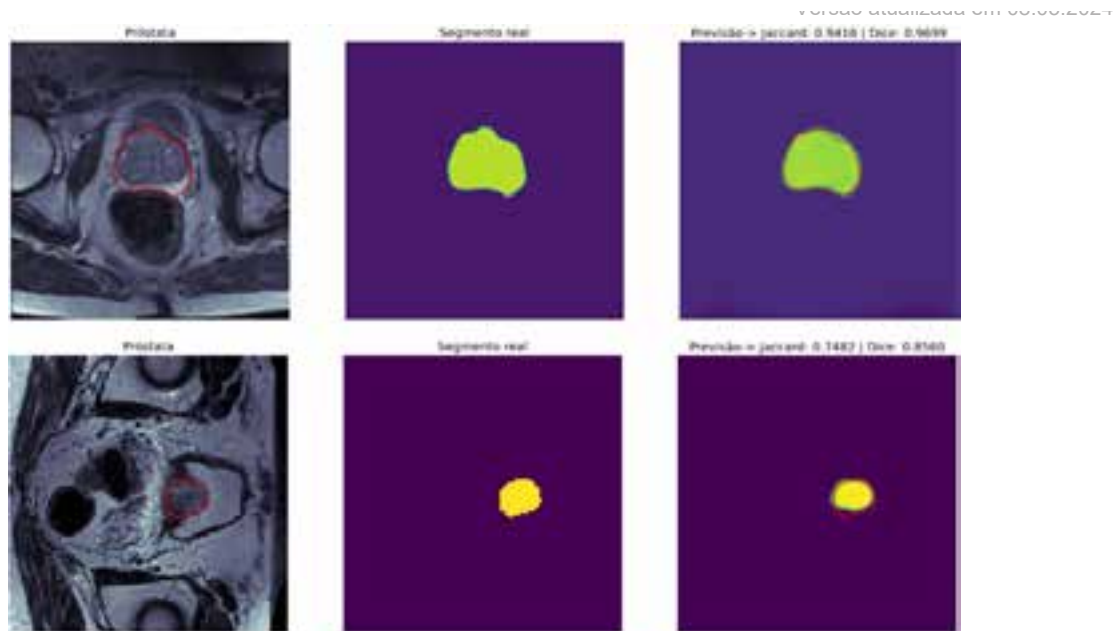
- Modelo de deep learning (DL) em ambiente centralizado: O treinamento inicial do modelo foi realizado utilizando dados centralizados, o que significa que todas as imagens utilizadas foram coletadas e armazenadas em um único local, seguindo o método tradicional de ML.
- Migração para um ambiente descentralizado: O fluxo de trabalho foi adaptado para um cenário descentralizado, utilizando o framework Flower para simular o FL.

Os próximos passos incluem a adição de novos datasets para aumentar a robustez do modelo, a integração de técnicas adicionais de proteção de privacidade, testar diferentes configurações de clientes e disponibilizar o modelo para uso com novas imagens.

3. DESENVOLVIMENTO DO TRABALHO (METODOLOGIA E RESULTADOS)

Este trabalho segue o ciclo de vida de um projeto de ciência de dados: 1- Definição do objetivo; 2- Coleta de 1.979 imagens reais de fontes públicas de dois datasets amplamente reconhecidos na área de segmentação de próstata, o PROMISE12 e o Task 05: Prostate. Destas, 1.380 imagens foram selecionadas por serem relevantes para o treinamento do modelo. 3- Foram desenvolvidos modelos de ML com a arquitetura U-Net, inicialmente treinados em um ambiente centralizado e depois migrados para um cenário descentralizado, simulando o FL com o framework Flower, envolvendo dois clientes. As configurações de clientes e treinamento estão sendo ajustadas para otimizar o desempenho e garantir a escalabilidade do modelo, visando expandi-lo para mais clientes e dados. 4- Os modelos foram avaliados utilizando as métricas Coeficiente de Jaccard e Dice Score. Os resultados preliminares mostraram índices médios de 0.7763 e 0.8613 no cenário centralizado, enquanto no federado os índices foram de 0.7872 e 0.8691, respectivamente. A Figura 1 demonstra a performance do modelo federado, evidenciando uma melhoria na precisão em segmentos maiores, destacando o potencial do FL para aprimorar a segmentação de imagens médicas. 5- Todas as etapas e resultados do projeto estão sendo devidamente documentados para garantir a replicabilidade e transparência dos processos adotados. 6- O modelo final será disponibilizado por meio de uma aplicação web, facilitando seu acesso para uso clínico e em pesquisas futuras.

Figura 1 – Exemplos de segmentação automática da próstata utilizando aprendizado federado



Fonte: O autor (2024)

4. DESAFIOS E APRENDIZADOS

Os principais desafios do projeto incluíram a necessidade de adquirir rapidamente conhecimentos técnicos sobre algoritmos de ML e FL, superada por meio de estudos intensivos, e a dificuldade em encontrar bases de dados públicas, parcialmente mitigada com o uso dos datasets PROMISE12 e Task 05: Prostate. Além disso, a integração de ferramentas e tecnologias enfrentou problemas de dependência entre bibliotecas, que foram solucionados pela migração para o Google Colaboratory.

5. REFERÊNCIAS

1. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616181/>. Acessado em 19/07/2024.
2. Ferreira JL, da Silva GLF, Reis ABS, Cavalcante AB, Silva AC, de Paiva AC. Segmentação Automática da Próstata em Imagens de Ressonância Magnética utilizando Redes Neurais Convolucionais e Mapa Probabilístico. In: *Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS), 18; 2018; Natal. Porto Alegre: Sociedade Brasileira de Computação; 2018. ISSN 2763-8952. doi: 10.5753/sbcas.2018.3683. Disponível em: https://sol.sbc.org.br/index.php/sbcas/article/view/3683. Acessado em: 19/07/2024.*
3. Mahapatra D, Buhmann JM. Prostate MRI segmentation using learned semantic knowledge and graph cuts. *IEEE Trans Biomed Eng.* 2014 Mar;61(3):756-64. doi: 10.1109/TBME.2013.2289306. Epub 2013 Nov 6. PMID: 24235297. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6656821>. Acessado em: 19/07/2024.
4. Sheller MJ, Edwards B, Reina GA, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Sci Rep.* 2020;10:12598. Disponível em: <https://www.nature.com/articles/s41598-020-69250-1>. Acessado em 19/07/2024.