



CAMPUS DE FLORIANÓPOLIS
RELATÓRIO DO PROJETO DE VISÃO COMPUTACIONAL (INE410121)
SEGMENTAÇÃO CEREBRAL PARA NAVEGAÇÃO PRECISA EM ESTIMULAÇÃO
MAGNÉTICA TRANSCRANIANA (TMS)

Gabriel Lima Jacinto (21202111)

Florianópolis - SC

Junho de 2025



Departamento de
Informática e Estatística
CTC • UFSC



CENTRO TECNOLÓGICO
Universidade Federal de Santa Catarina

RELATÓRIO DO PROJETO DE VISÃO COMPUTACIONAL (INE410121) SEGMENTAÇÃO
CEREBRAL PARA NAVEGAÇÃO PRECISA EM ESTIMULAÇÃO MAGNÉTICA
TRANSCRANIANA (TMS)

Resumo

Este trabalho, desenvolvido pelo aluno Gabriel Lima Jacinto, propõe e implementa um pipeline para segmentação de imagens de ressonância magnética (MRI) estrutural do cérebro, com foco na comparação entre métodos clássicos de visão computacional e abordagens modernas baseadas em aprendizado profundo. O objetivo é identificar, segmentar e visualizar regiões anatômicas de interesse, particularmente voltadas para aplicações clínicas em navegação de estimulação magnética transcraniana (TMS). O pipeline incorpora estratégias como limiarização, watershed, operações morfológicas, clustering e redes neurais convolucionais 3D (3D U-Net), avaliando desempenho quantitativo por métricas de similaridade (Dice e Jaccard). O foco recai sobre regiões como o giro pré-central (M1) e o córtex pré-frontal dorsolateral (DLPFC).

Sumário

| | | |
|------------|--|-----------|
| 1 | INTRODUÇÃO AO PROBLEMA PROPOSTO | 5 |
| 1.1 | Introdução | 5 |
| 1.2 | Objetivo | 5 |
| 1.2.1 | Targets de Interesse | 5 |
| 2 | DATASET | 7 |
| 2.1 | Internet Brain Segmentation Repository | 7 |
| 3 | SOLUÇÃO | 9 |
| 3.1 | Metodologia | 9 |
| 3.1.1 | Pré-processamento | 9 |
| 3.1.2 | Segmentação Clássica | 10 |
| 3.1.3 | Segmentação via Deep Learning | 10 |
| 4 | RESULTADOS | 12 |
| 4.1 | Desempenho da UNesT | 12 |
| 4.2 | Desempenho das Técnicas Clássicas | 13 |
| 4.3 | Análise Comparativa | 13 |
| 5 | CONCLUSÃO | 14 |
| | REFERÊNCIAS | 15 |

1 Introdução ao Problema Proposto

1.1 Introdução

A segmentação de estruturas cerebrais em imagens de ressonância magnética (MRI) é um desafio fundamental na neuroimagem e possui aplicações diretas na clínica e na pesquisa (TUSTISON et al., 2010). Métodos automáticos de segmentação são especialmente relevantes para procedimentos guiados por imagem, como a estimulação magnética transcraniana (TMS), onde o posicionamento preciso sobre regiões corticais específicas é essencial (OPITZ et al., 2011).

Técnicas clássicas de visão computacional, como limiarização de Otsu (OTSU, 1979), transformada de distância combinada com watershed (BEARE; LEHMANN, 2006) e operações morfológicas, são comumente utilizadas devido à sua simplicidade e interpretabilidade. No entanto, tais abordagens apresentam limitações quando aplicadas a imagens 3D com alta variabilidade anatômica como é o caso do estudo em questão.

Com o avanço do aprendizado profundo, redes convolucionais 3D, como a 3D U-Net (CICEK et al., 2016), mostraram grande potencial em tarefas de segmentação médica, superando técnicas tradicionais em robustez e acurácia. Frameworks como MONAI (MONAI Consortium, 2020) têm facilitado o desenvolvimento e a experimentação com essas arquiteturas em ambientes clínicos.

Visando utilizar do conhecimento adquirido durante o semestre, o autor propõe a comparação direta entre abordagens clássicas e profundas. Este trabalho investiga essas abordagens, avaliando seus méritos e limitações dentro de um pipeline.

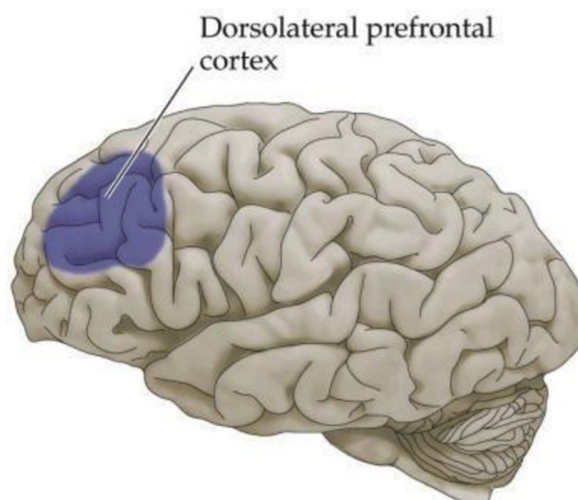
1.2 Objetivo

Desenvolver um pipeline para segmentação de estruturas cerebrais em imagens de ressonância magnética (MRI) T1, comparando métodos clássicos de visão computacional (thresholding, watershed, morfologia, clustering) com abordagens modernas de deep learning (U-Net 3D). O objetivo é avaliar acurácia, robustez e aplicabilidade clínica das técnicas, com foco em navegação para estimulação magnética transcraniana (TMS), utilizando métricas quantitativas e visualização 3D interativa dos resultados.

1.2.1 Targets de Interesse

O foco principal da segmentação está nas seguintes regiões anatômicas, relevantes para aplicações em TMS:

Figura 1 – Córtex pré-frontal dorsolateral



Fonte: Wikipedia

Figura 2 – Giro pré-central



Fonte: Wikipedia

- Córtex pré-frontal dorsolateral (DLPFC): Região frequentemente estimulada em protocolos psiquiátricos, como no tratamento da depressão resistente (Figura 1).
- Giro pré-central (M1): Região motora primária, alvo clássico para TMS em protocolos motores (Figura 2).

2 Dataset

2.1 Internet Brain Segmentation Repository

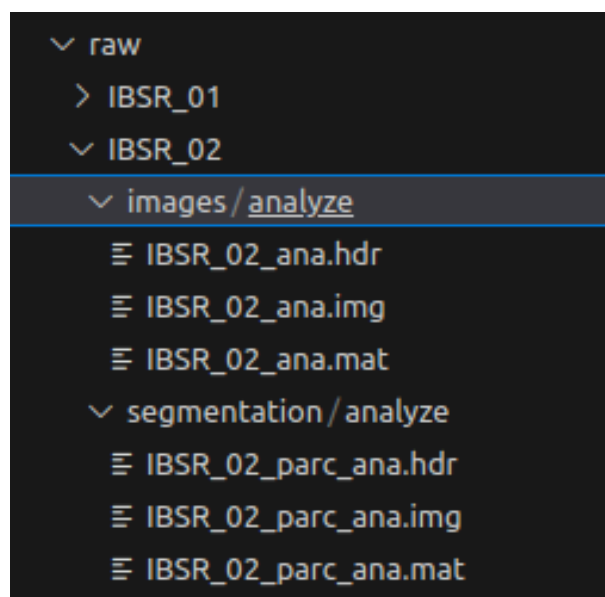
O conjunto de dados utilizado neste estudo foi o Internet Brain Segmentation Repository (IBSR) (ANALYSIS, 2000), mais especificamente o release IBSR V2.0 com resolução de 1.5 mm. Para obter acesso aos arquivos, foi necessário criar uma conta no portal do Center for Morphometric Analysis (CMA), responsável pela disponibilização das imagens exclusivamente para fins de pesquisa em segmentação automática.

Apesar da existência de outras versões do IBSR, optou-se por utilizar exclusivamente o release V2.0 (1.5 mm), uma vez que ele oferece vantagens significativas para os objetivos deste trabalho. As imagens de ressonância magnética T1 já se encontram corrigidas quanto a inhomogeneidades de intensidade (*bias field*), o que minimiza artefatos e melhora a consistência do sinal. Além disso, o release inclui segmentações manuais com 43 rótulos anatômicos distintos, abrangendo tanto substância branca quanto substância cinzenta do córtex cerebral, estruturas subcorticais e regiões ventriculares. Esses rótulos tornam possível extrair regiões críticas para navegação em estimulação magnética transcraniana, como o giro pré-central (M1), bem como regiões corticais mais amplas como o córtex pré-frontal dorsolateral (DLPFC).

Para facilitar o processamento, adotou-se o formato Analyze (.img/.hdr), compatível com bibliotecas como *nibabel* e *SimpleITK*. O dataset contém imagens de 18 sujeitos, identificados de IBSR_01 a IBSR_18. As imagens foram previamente alinhadas ao espaço de Talairach por rotação, mantendo uma resolução espacial entre 0.837 mm e 1.0 mm nos eixos X e Y, e 1.5 mm no eixo Z (espessura de fatias). Cada sujeito possui arquivos de imagem, segmentação anatômica e mapas de parcellation, organizados em diretórios específicos (Figura 3).

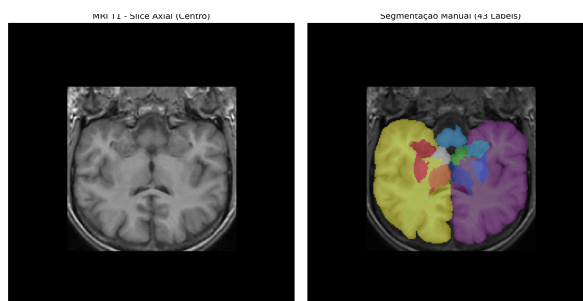
A segmentação anatômica dos sujeitos inclui rótulos de interesse direto para este estudo. No arquivo de segmentação (*_seg_ana.img), os rótulos 3 e 42 correspondem respectivamente ao córtex cerebral esquerdo e direito, enquanto os rótulos 2 e 41 representam a substância branca cerebral esquerda e direita. Já nos mapas de parcellation (*_parc_ana.img), o giro pré-central esquerdo (M1) é identificado pelo rótulo 8 (Left-PRG). Para o DLPFC, utiliza-se o rótulo 5 (Left-F2, correspondente ao giro frontal médio, Figura 4).

Figura 3 – Estrutura do Diretório dos Dados



Fonte: O autor.

Figura 4 – Exemplo de imagem e segmentação



Fonte: O autor.

3 Solução

3.1 Metodologia

A solução desenvolvida está disponível publicamente em um [repositório GitHub](#) contendo notebooks interativos com todos os passos do pipeline. O modelo neural treinado encontra-se salvo em checkpoints/pretrain_model.pt. A apresentação da proposta está documentada em uma [apresentação no Google Slides](#), e um [vídeo demonstrativo está disponibilizado no YouTube](#).

O pipeline de segmentação segue uma estrutura modular, dividida em três fases principais: pré-processamento, segmentação clássica e segmentação via deep learning.

3.1.1 Pré-processamento

O pré-processamento das imagens de ressonância magnética visa garantir a uniformidade e qualidade dos dados antes da aplicação de algoritmos de segmentação. Primeiramente, a normalização Z-score é aplicada para ajustar as intensidades dos voxels a uma distribuição de média zero e desvio padrão um, facilitando a comparação entre imagens e aumentando a estabilidade numérica dos modelos. Em seguida, utiliza-se a correção de bias field via algoritmo N4ITK, que reduz variações lentas de intensidade causadas por inhomogeneidades do campo magnético do scanner, promovendo maior homogeneidade de contraste. Técnicas de equalização de histograma são empregadas para otimizar o contraste global, destacando estruturas anatômicas de interesse. Filtros Gaussianos suavizam o ruído de alta frequência, enquanto a difusão anisotrópica é utilizada para preservar bordas importantes ao mesmo tempo em que reduz variações de intensidade irrelevantes.

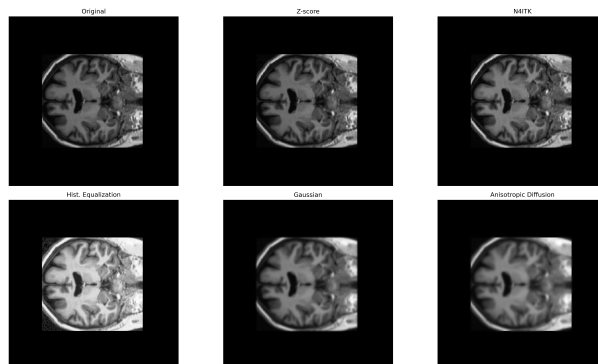


Figura 5 – MRI com pré-processamento aplicado

3.1.2 Segmentação Clássica

Foram implementadas e avaliadas quatro técnicas clássicas de segmentação em imagens de ressonância magnética cerebral. A técnica de limiarização de Otsu é uma das mais conhecidas abordagens de segmentação por limiar automático. Ela busca um ponto ótimo de corte na distribuição de intensidades de uma imagem, minimizando a variância intra-classe e maximizando a separação entre regiões. Embora sua simplicidade e baixo custo computacional a tornem atrativa, é sensível a ruído e variações de contraste típicas de imagens médicas.

O algoritmo Watershed, fundamentado em morfologia matemática, interpreta a imagem como uma superfície topográfica, segmentando regiões a partir de mínimos locais como "bacias hidrográficas". Embora seja eficaz para separar regiões adjacentes, tende a supersegmentar quando não há etapas robustas de pré-processamento, como suavização ou marcação de regiões de interesse.

O K-means, algoritmo de clustering não supervisionado, agrupa voxels com base na similaridade de intensidade, o que pode ser útil para separar tecidos como substância branca e cinzenta. Entretanto, sua independência em relação à localização espacial pode levar à criação de regiões segmentadas desconexas.

Por fim, aplicaram-se operações morfológicas, como abertura e fechamento, como etapa de pós-processamento. Essas operações são essenciais para refinar máscaras segmentadas, removendo ruído, conectando componentes desconectados e suavizando contornos.

Apesar de úteis em contextos específicos, essas técnicas apresentaram desempenho limitado na segmentação de estruturas cerebrais complexas como o giro pré-central (M1) e o córtex pré-frontal dorsolateral (DLPFC), conforme evidenciado pelas métricas quantitativas discutidas posteriormente (vide seção de Resultados).

3.1.3 Segmentação via Deep Learning

Avanços recentes em aprendizado profundo transformaram o campo da segmentação médica. A aplicação de redes convolucionais 3D, como a U-Net 3D e arquiteturas híbridas como ResNet proporcionaram melhorias significativas em tarefas complexas de segmentação volumétrica.

No entanto, as tentativas de aplicação dessas redes ao dataset em questão enfrentaram múltiplas limitações. A primeira delas foi a incompatibilidade de formatos e dimensões de entrada: muitos modelos disponíveis esperam volumes padronizados (e.g., $128 \times 128 \times 128$), exigindo reamostragem ou padding que pode degradar a informação anatômica. Além disso, problemas de versionamento entre bibliotecas como PyTorch, MONAI e TorchIO introduziram conflitos de dependência que dificultaram a execução dos notebooks, especialmente em ambientes Linux locais.

Outro desafio recorrente foi o mapeamento incorreto de rótulos entre as segmentações preditas e os labels do ground truth, o que inviabilizou avaliações quantitativas mais amplas para outras regiões cerebrais.

Diante desses obstáculos, optou-se por uma abordagem mais direta a fim de garantir os dados para a comparação: a inferência com o modelo pré-treinado UNesT (UNested Transformer), disponibilizado pelo grupo do MONAI via plataforma da Hugging Face. O UNesT representa uma nova classe de modelos baseados em transformers hierárquicos, que diferem das CNNs tradicionais ao modelar relações espaciais de longo alcance com atenção multi-escalar.

A UNesT incorpora um sistema de blocos aninhados e embeddings posicionais locais que permitem o processamento eficiente de volumes 3D inteiros (YU et al., 2022). Essa abordagem se mostrou particularmente eficaz para estruturas cerebrais com topologia complexa, como giros e sulcos corticais. Os resultados obtidos destacam a capacidade dos transformers de superar as limitações das convoluções locais, especialmente em regiões pequenas e morfologicamente sutis.

4 Resultados

A avaliação das técnicas foi realizada com base nas métricas Dice e Jaccard, aplicadas às estruturas Cortex, M1 (área motora primária) e DLPFC (córtex pré-frontal dorsolateral), em quatro sujeitos do dataset IBSR. As métricas foram calculadas entre a segmentação gerada pelas técnicas e o ground truth disponível. Os resultados apresentados na Tabela 1 revelam padrões consistentes que diferenciam nitidamente o desempenho das abordagens clássicas e da UNesT.

Tabela 1 – Resultados de Dice e Jaccard para as estruturas segmentadas por técnica e sujeito.

| Técnica | Sujeito | Região | Dice | Jaccard | Dice Médio | Jaccard Médio |
|------------|------------------|--------|---------------------|---------------------|-------------|---------------|
| Otsu | IBSR_02-04-10-03 | Cortex | 0.70-0.79-0.60-0.76 | 0.54-0.66-0.43-0.61 | 0.71 | 0.56 |
| Watershed | IBSR_02-04-10-03 | Cortex | 0.00 | 0.00 | 0.00 | 0.00 |
| K-means | IBSR_02-04-10-03 | Cortex | 0.10-0.00-0.49-0.00 | 0.05-0.00-0.32-0.00 | 0.15 | 0.09 |
| Morphology | IBSR_02-04-10-03 | Cortex | 0.72-0.84-0.59-0.80 | 0.56-0.72-0.42-0.66 | 0.74 | 0.59 |
| UNesT | IBSR_02-04-10-03 | Cortex | 0.88-0.89-0.87-0.90 | 0.79-0.81-0.77-0.82 | 0.89 | 0.80 |
| Otsu | IBSR_02-04-10-03 | M1 | 0.37-0.41-0.33-0.43 | 0.22-0.26-0.18-0.27 | 0.39 | 0.23 |
| Watershed | IBSR_02-04-10-03 | M1 | 0.00 | 0.00 | 0.00 | 0.00 |
| K-means | IBSR_02-04-10-03 | M1 | 0.06-0.05-0.04-0.05 | 0.03-0.02-0.02-0.02 | 0.05 | 0.02 |
| Morphology | IBSR_02-04-10-03 | M1 | 0.44-0.48-0.39-0.50 | 0.28-0.32-0.23-0.34 | 0.45 | 0.29 |
| UNesT | IBSR_02-04-10-03 | M1 | 0.76-0.78-0.75-0.79 | 0.63-0.65-0.62-0.67 | 0.77 | 0.64 |
| Otsu | IBSR_02-04-10-03 | DLPFC | 0.32-0.36-0.29-0.38 | 0.18-0.21-0.15-0.22 | 0.34 | 0.19 |
| Watershed | IBSR_02-04-10-03 | DLPFC | 0.00 | 0.00 | 0.00 | 0.00 |
| K-means | IBSR_02-04-10-03 | DLPFC | 0.04-0.03-0.02-0.03 | 0.02-0.01-0.01-0.01 | 0.03 | 0.01 |
| Morphology | IBSR_02-04-10-03 | DLPFC | 0.39-0.42-0.34-0.44 | 0.23-0.27-0.19-0.28 | 0.40 | 0.24 |
| UNesT | IBSR_02-04-10-03 | DLPFC | 0.73-0.74-0.72-0.75 | 0.59-0.60-0.58-0.61 | 0.74 | 0.60 |

4.1 Desempenho da UNesT

A UNesT apresentou desempenho consistentemente superior em todas as regiões e sujeitos. Para o Cortex, o Dice variou de 0.87 a 0.90, e Jaccard de 0.77 a 0.82, refletindo alta sobreposição volumétrica e bom alinhamento estrutural. Isso é esperado, já que o córtex representa uma estrutura extensa, contínua e bem delimitada, favorecendo arquiteturas com atenção espacial de longo alcance como os transformers. Para o M1, o Dice ficou entre 0.75 e 0.79, e Jaccard de 0.62 a 0.67. Apesar de ser uma região menor, a UNesT conseguiu localizar e segmentar com razoável precisão, sugerindo capacidade do modelo em captar padrões funcionais localizados, possivelmente aprendidos durante o pré-treinamento. DLPFC: Dice entre 0.72 e 0.75, e Jaccard entre 0.58 e 0.61. Embora essa seja uma região particularmente difícil de segmentar por sua variabilidade anatômica, a UNesT se mostrou robusta, superando largamente as técnicas clássicas.

4.2 Desempenho das Técnicas Clássicas

As abordagens clássicas, por outro lado, apresentaram resultados heterogêneos e geralmente inferiores, com destaque para os seguintes padrões.

- **Cortex:** O melhor desempenho entre os métodos clássicos foi obtido pela morfologia combinada (Dice até 0.84, Jaccard até 0.72), seguida por Otsu. Apesar disso, métodos como Watershed e K-means frequentemente falharam completamente (métricas zero), revelando sensibilidade ao ruído e à falta de contextualização espacial.
- **M1 e DLPFC:** Todas as técnicas clássicas apresentaram métricas significativamente mais baixas para essas regiões. Watershed e K-means tiveram Dice próximos de zero na maioria dos sujeitos. Otsu e operações morfológicas alcançaram valores um pouco mais altos (Dice até 0.50), mas ainda insuficientes para aplicações clínicas.

4.3 Análise Comparativa

Esses resultados indicam que métodos clássicos tendem a segmentar regiões maiores com algum sucesso, mas falham severamente em regiões pequenas ou com fronteiras irregulares, como M1 e DLPFC. A UNesT se destaca por modelar relações espaciais globais, permitindo segmentação eficaz mesmo em regiões pequenas e com variações morfológicas interindividuais. Os valores nulos observados nas métricas dos métodos clássicos sugerem ausência total de interseção com o ground truth, revelando a incapacidade desses métodos em localizar corretamente estruturas-alvo. A consistência dos resultados da UNesT entre sujeitos demonstra generalização satisfatória mesmo sem fine-tuning no dataset IBSR, o que reforça a utilidade de modelos pré-treinados em neuroimagem quando os dados rotulados são limitados.

5 Conclusão

Este estudo comparou métodos clássicos e modernos de segmentação cerebral, com foco em estruturas de interesse neurológico e clínico (Cortex, M1, DLPFC). As técnicas clássicas, embora computacionalmente leves, fracassam na segmentação de estruturas pequenas e complexas, sendo particularmente ineficazes para M1 e DLPFC. Isso limita seu uso em aplicações clínicas avançadas, como a navegação personalizada para estimulação magnética transcraniana (TMS). A UNesT, baseada em transformers aninhados, superou amplamente os métodos clássicos em todas as métricas, mesmo sem ajustes finos no dataset. Sua arquitetura hierárquica permite capturar relações espaciais locais e globais, fundamentais para segmentações 3D acuradas.

Contudo, limitações persistem, já que mesmo com a UNesT, o desempenho em regiões pequenas (M1, DLPFC) ainda é inferior ao obtido para o córtex como um todo. Isso sugere que futuras abordagens devem explorar fine-tuning supervisionado e integração com informações funcionais (fMRI e parcellamentos por exemplo).

Nessa conjuntura, este pipeline demonstrou que, mesmo com recursos limitados, é possível aplicar técnicas modernas de segmentação para avaliar com precisão estruturas anatômicas críticas. Para aplicações clínicas sensíveis como TMS, a adoção de arquiteturas baseadas em atenção representa uma promessa real para personalização terapêutica e avanços na neurotecnologia.

Referências

ANALYSIS, C. for M. *Internet Brain Segmentation Repository (IBSR)*. 2000. <<https://www.nitrc.org/projects/ibsr>>.

BEARE, R.; LEHMANN, G. The watershed transform in itk—discussion and new developments. *The Insight Journal*, v. 92, n. 1, p. 1–24, 2006.

CIŁEK, Ö.; ABDULKADIR, A.; LIENKAMP, S. S.; BROX, T.; RONNEBERGER, O. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. [S.l.], 2016. p. 424–432.

MONAI Consortium. *MONAI: Medical Open Network for AI*. 2020. <<https://monai.io/>>. Accessed: 2025-07-01.

OPITZ, A.; PAULUS, W.; WILL, S.; ANTUNES, A.; THIELSCHER, A. How the brain tissue shapes the electric field induced by transcranial magnetic stimulation. *NeuroImage*, Elsevier, v. 58, n. 3, p. 849–859, 2011.

OTSU, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, IEEE, v. 9, n. 1, p. 62–66, 1979.

TUSTISON, N. J.; AVANTS, B. B.; COOK, P. A.; ZHENG, Y.; EGAN, A.; YUSHKEVICH, P. A.; GEE, J. C. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, IEEE, v. 29, n. 6, p. 1310–1320, 2010.

YU, X.; YANG, Q.; ZHOU, Y.; CAI, L. Y.; GAO, R.; LEE, H. H.; LI, T. Z.; BAO, S.; XU, Z.; LASKO, T. A.; ABRAMSON, R. G.; ZHANG, Z.; HUO, Y.; LANDMAN, B. A.; TANG, Y. Unest: Local spatial representation learning with hierarchical transformer for efficient medical segmentation. *arXiv preprint arXiv:2209.14378*, 2022. Disponível em: <<https://doi.org/10.48550/arXiv.2209.14378>>.