

Adaptação de uma rede U-net com duas MRIs de entrada para segmentação de lesões escleróticas

Isabella Medeiros de Sousa
Faculdade de Engenharia,
Universidade do Porto
Porto, Portugal
ORCID: 0000-0002-9577-7458

Marcela de Oliveira
Universidade Estadual Paulista “Júlio de
Mesquita Filho”
Bauru, Brasil
ORCID: 0000-0003-4144-7629

Jaime do Santos Cardoso
INESC TEC e Faculdade de Engenharia,
Universidade do Porto
Porto, Portugal
ORCID: 0000-0002-3760-2473

Resumo— A Esclerose Múltipla (EM) é uma doença autoimune que provoca lesões na substância branca no cérebro e na medula espinhal, em decorrência da desmielinização dos axônios. Essas lesões podem ser visualizadas através de imageamento por ressonância magnética (MRI do inglês *Magnetic Resonance Imaging*), auxiliando no diagnóstico e acompanhamento da evolução da doença. Este trabalho teve o objetivo de segmentar lesões escleróticas em MRI nas sequências FLAIR (do inglês *Fluid Attenuated Inversion Recovery*) e T1, através de Redes Neurais Convolucionais (CNN do inglês *Convolutional Neural Networks*). O modelo de segmentação foi testado em imagens de 20 pacientes diagnosticados com EM. O conjunto de segmentações foi avaliado estatisticamente, através das métricas coeficiente Dice (DSC do inglês *Dice Similarity Coefficient*), Valor Preditivo Positivo (VPP) e Diferença de Volume Absoluta (DVA). Os resultados indicaram que o método de segmentação automática apresenta bom desempenho, com grande potencial para ser utilizado para segmentar as lesões escleróticas em clínicas médicas no futuro.

Palavras-chave — Esclerose Múltipla, MRI, segmentação, CNN.

I. INTRODUÇÃO

A Esclerose Múltipla (EM) é uma doença crônica na qual células do sistema imunológico atacam a bainha de mielina dos neurônios do sistema nervoso central, resultando em um processo inflamatório e neurodegenerativo [1]. Esta doença afeta aproximadamente 2,8 milhões de pessoas ao redor do mundo, com grande número de jovens adultos afetados, além disso, foi observado que países afastados da linha equatorial apresentam altos números de doentes [2]. A EM é uma doença de origem complexa, que pode ter seu desenvolvimento associado a fatores genéticos e interações ambientais, como, por exemplo, baixa exposição solar (provocando baixo nível de vitamina), e outros fatores associados como a obesidade e infecção pelo vírus Espin-Barr [3]. A manifestação clínica da doença é bastante variável e depende do tamanho e localização das lesões; alguns possíveis sintomas presentes são: perda sensorial facial, parestias e neurite óptica unilateral [4].

O diagnóstico da EM é dado através da combinação da avaliação clínica, análise laboratorial e exames de imagem [5]. Os exames de imagem de ressonância magnética (MRI, do inglês *Magnetic Resonance Imaging*) apresentam os tecidos cerebrais em diferentes intensidades a depender da sequência de MRI utilizada, conforme Tabela 1, e podem ser utilizados para visualizar o processo de desmielinização [6]. Nas sequências T1, as lesões aparentam hipointensidades, enquanto na sequência T2 e FLAIR (do inglês *Fluid*

Attenuated Inversion Recovery) aparecem como hiperintensidades [7]. Entre estas modalidades, a imagem FLAIR é a que apresenta melhor diferença de sinal entre as lesões de EM e os tecidos saudáveis.

TABELA I. CARACTERÍSTICAS DE INTENSIDADE DE SINAL DAS LESÕES E TECIDOS CEREBRAIS NAS SEQUÊNCIAS DE MRI [6]

Tecido	Sequência de MRI		
	T1	T2	FLAIR
Líquido cefalorraquidiano	Preto	Branco	Preto
Substância branca	Cinza claro	Cinza escuro	Cinza escuro
Substância cinzenta	Cinza escuro	Cinza claro	Cinza claro
Desmielinização	Preto	Branco	Branco

Atualmente, a segmentação manual de lesões é considerada o padrão-ouro, embora consuma bastante tempo e esteja sujeita a variabilidade entre observadores [8]. Alternativamente, numerosos métodos de segmentação automática estão presentes na literatura [9]. Recentemente, a utilização de Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Networks*) ganhou destaque para segmentação de imagens médicas. Entretanto, a segmentação automática ainda é uma tarefa complicada de ser realizada, pela quantidade reduzida de imagens médicas disponíveis para a aplicação de modelos de *Deep Learning*, dessa forma a aplicação de *transfer learning* cresce como alternativa para superar essa dificuldade [10]. Além disso, o tamanho e localização das lesões escleróticas variam significativamente e suas bordas não são bem definidas [6]. Adicionalmente, é importante considerar os problemas de adaptação de domínio, quando os dados de treino e teste são obtidos por diferentes protocolos e equipamentos [11]. Outras dificuldades estão relacionadas com as imagens de MRI utilizadas, pois estas podem apresentar ruído, devido ao processo de aquisição e variações de intensidades de um mesmo tecido, causados pela não homogeneidade dos campos magnéticos do equipamento de realização do exame [12].

Em relação aos métodos de segmentação de lesões escleróticas encontrados na literatura, Roy *et al.* [13] propuseram uma arquitetura, na qual amostraram sequências T1 e FLAIR de MRI, sendo então processadas em paralelo por várias operações de convolução. Os resultados foram concatenados e passaram por outra série de operações de convolução até realizar a predição da segmentação. Valverde *et al.* [14] apresentaram uma rede de duas CNN em cascata, onde a primeira rede classificou possíveis lesões e a segunda

reduziu o número de falsos positivos. Além disso, em *Valverde et al.* [15], um método de adaptação de domínio das redes neurais por *transfer learning* foi implementado e obteve resultados satisfatórios. Em outro trabalho, utilizando 5 modalidades de MRI, *Ghosal et al.* [16] adaptaram a tradicional arquitetura Unet [17], reduzindo o número de operações, filtros e *skip connections* com a finalidade de simplificar o treinamento do modelo e obteve excelentes resultados. A identificação e quantificação das lesões são importantes para avaliar os efeitos de terapias [6]. Em *Oliveira et al.* [18], a quantificação de lesões escleróticas segmentadas por CNN apresentou bons resultados. Diante do exposto, este trabalho apresenta o resultado da implementação e segmentação do modelo cuja arquitetura foi adaptada de *Ghosal et al.* [16], utilizando apenas as sequências FLAIR e T1, dado que múltiplas aquisições de imagens podem causar desconforto para o paciente, além de aumentar a complexidade computacional dos algoritmos [19]. Hipotetiza-se que a utilização das sequências FLAIR e T1 seja suficiente para segmentar as lesões de EM no modelo adaptado, obtendo métricas comparativas satisfatórias.

II. MATERIAIS E MÉTODOS

A. Banco de dados

O banco de dados deste trabalho foi constituído por um banco de imagens para treinamento do modelo e outro para teste. As imagens utilizadas no treinamento foram selecionadas dos bancos de imagens do *International Symposium on Biomedical Imaging (ISBI) 2015* [20] e *MSSEG Challenge Proceedings: Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure* em *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) 2016* [21]. O banco de dados de ISBI 2015 foi composto de imagens de 5 sujeitos, com 4 ou 5 exames em diferentes momentos para cada paciente, e as imagens foram obtidas no equipamento Philips Medical Systems 3T. O banco de dados do MICCAI 2016 foi composto de imagens adquiridas em três diferentes domínios (equipamentos): Philips Ingenia 3T, Siemens Aera 1.5T e Siemens Verio 3T. Para o treinamento dos modelos, foram utilizadas as sequências FLAIR e T1.

O banco de dados para teste foi composto de imagens FLAIR e T1, obtidas através do sistema 3T Siemens Magnetom Trio, de 20 pacientes diagnosticados com EM, fornecido pela Universidade de Liubliana [22]. As segmentações *ground truth* fornecidas foram obtidas a partir de um consenso entre especialistas, como forma de diminuir a subjetividade de cada observador.

B. Pré-processamento das imagens

As imagens de ressonância magnética passaram por determinadas etapas de processamento através do *software Slicer 3D* [23]: extração das estruturas não-cerebrais através do módulo *SwissSkullStripper* [24], correção da não-homogeneidade das intensidades com o algoritmo *N4ITK* [25], reamostragem isotrópica e co-registro rígido pelo módulo do *BRAINSfit* [26] entre as sequências FLAIR e T1, com a finalidade de manter as estruturas biológicas correspondentes em ambas modalidades. Além dessas operações, as imagens foram redimensionadas para 256 x 256 pixels, os cortes sem informação foram descartados e as intensidades foram normalizadas.

C. Arquitetura da CNN

A rede neural convolucional implementada apresenta formato em U, constituído de um caminho de contração e um caminho de expansão [17], conforme observado na Fig 1. O primeiro caminho é formado por várias operações de convolução e *Max pooling*, para extrair informações das imagens analisadas [16]. As operações de *Max Pooling* extraem as características mais significantes das imagens e reduzem as suas dimensões, resultando em dados de baixa resolução. Essa primeira etapa permite que o modelo entenda o que está presente na imagem, entretanto perde informações relativas à localização. Dessa maneira, necessita-se do caminho de expansão, que é responsável por aumentar a dimensão e resolução dos dados, através de operações de convolução e *up-sampling* [16]. As duas etapas estão ligadas por *skip connections* alternadas, cuja finalidade é transferir informações espaciais do caminho de contração para o de expansão.

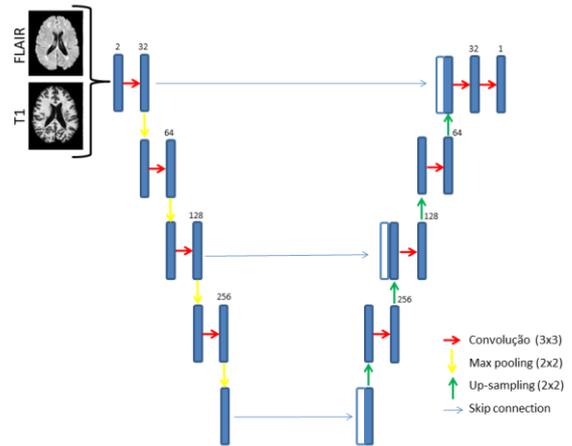


Fig. 1. Arquitetura *Light Weighted Unet* adaptada para duas sequências de MRI. As modalidades são concatenadas e servem de entrada à rede, passam por um caminho de contração e seguem para o caminho de expansão, por fim é realizada a previsão da segmentação.

D. Implementação

O modelo estudado foi implementado em Python com Keras e Tensorflow e o tempo de treinamento foi aproximadamente de duas horas utilizando GPU (NVIDIA Geforce GTX 1080). Os seguintes hiperparâmetros foram utilizados: *batch size* igual a 32, otimizador Adam com taxa de aprendizagem de 0,001, a função de perda escolhida foi a *binary crossentropy* e número de épocas foi igual a 60.

E. Pós-processamento das segmentações

O modelo implementado foi aplicado nas imagens do banco de teste, em seguida, estas imagens foram pós-processadas para aperfeiçoamento dos resultados. Nesta etapa, os pixels classificados como lesões de tamanho inferior a 3 mm² foram descartados da segmentação, por não ser indicativo de lesão de EM [27].

F. Métricas estatísticas

Métricas estatísticas foram utilizadas como meio de avaliar o desempenho da segmentação do modelo estudado. As métricas utilizadas foram o coeficiente de similaridade Dice (DSC), o valor preditivo positivo (VPP) e a diferença de volume absoluta (DVA), indicadas nas equações (1), (2) e (3), respectivamente [13]. Em que $|A|$ indica a segmentação automática e $|M|$ indica a segmentação manual.

$$DSC(A, M) = \frac{2 \cdot (|A \cap M|)}{|A| + |M|} \quad (1)$$

$$VPP(A, M) = \frac{|A \cap M|}{|A|} \quad (2)$$

$$DVA(A, M) = \frac{\text{abs}(|A| - |M|)}{|M|} \quad (3)$$

O DSC indica o grau de similaridade entre as duas imagens binárias através da sobreposição espacial, quanto mais próximo de uma unidade, melhor a segmentação. O VPP é dado pela razão entre os verdadeiros positivos e todos os pixels classificados como positivos, indicando quanto dos pixels classificados como lesão foram corretamente identificados. Por último, a DVA apresenta a diferença entre os volumes da segmentação automática e manual em relação ao volume dado pela segmentação realizada pelo especialista; quanto mais próximo de zero, melhor.

III. RESULTADOS

Nesta seção, os resultados obtidos da segmentação de imagens através do modelo treinado estão apresentados nas Figuras 2 e 3. As métricas estatísticas obtidas pelas análises das segmentações estão ilustradas nos gráficos *box-plot* na Fig. 2. Observa-se que o DSC apresentou mediana igual a 0,560, sendo este um valor razoável, porém o intervalo de valores obtidos sofreu uma grande variação, com valor mínimo igual a 0,067 e máximo de 0,705. Em relação ao VPP, os valores obtidos, no geral, foram excelentes e a mediana do conjunto foi igual a 0,882, ademais houve apenas alguns *outliers*. Por fim, o conjunto de dados da DVA apresentou mediana igual a 0,519 e houve grande variação entre os valores obtidos.

Métricas Comparativas

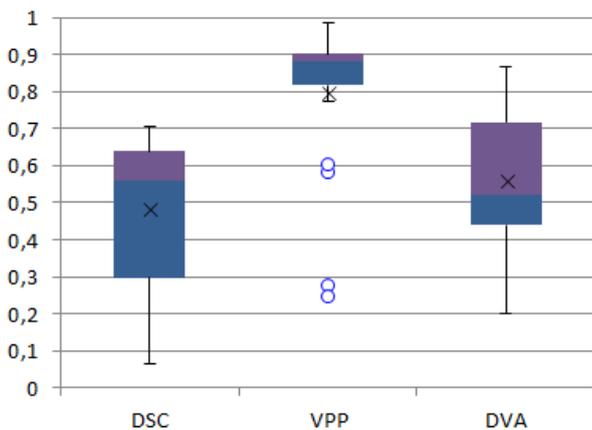


Fig 2. Métricas comparativas DSC, VPP e DVA, obtidas pelas segmentações das lesões do banco de teste.

Na Fig. 3, observam-se dois cortes (Fig. 3a e 3b) das imagens FLAIR e T1 de um mesmo paciente, e respectiva segmentação manual (*ground truth*) ilustrada em branco e segmentação automática em azul. Em ambos os cortes, o modelo identificou as lesões escleróticas de maneira apropriada, apresentando boa segmentação e apenas as bordas e pequenas lesões, em provável estágio inicial de desmielinização, não foram corretamente segmentadas.

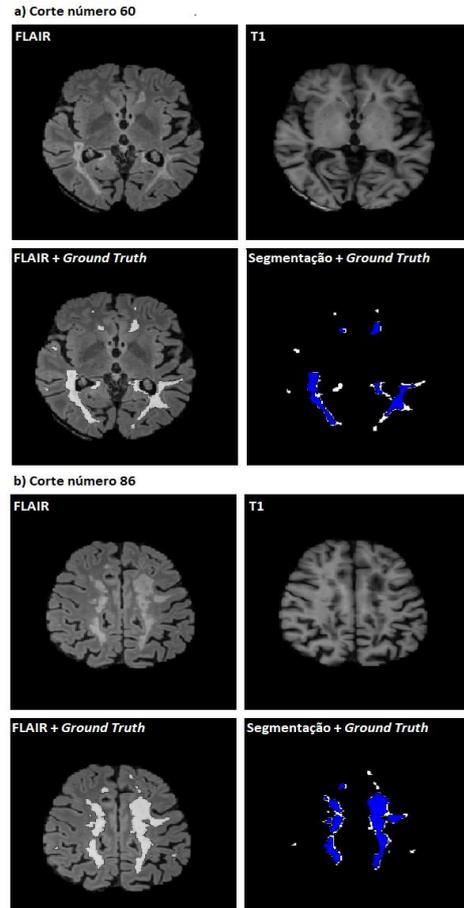


Fig.3. Diferentes cortes, a) corte de número 60 e b) 86, de MRI de um mesmo paciente. Na primeira linha, as sequências FLAIR e T1 estão ilustradas. Na segunda linha, observa-se a segmentações manual (em branco) sobreposta a imagem FLAIR e a segmentação automática (em azul) sobreposta à segmentação manual.

IV. DISCUSSÃO

Recentemente, vários algoritmos baseados em CNN foram propostos para segmentar as lesões de EM [13-16]. Neste estudo, o modelo apresentado foi uma versão adaptada do modelo proposto por Ghosal *et al.* [16], onde apenas duas sequências de MRI foram necessárias para treinar o modelo. Os resultados das segmentações foram avaliados através das métricas DSC, VPP e DVA em relação à segmentação manual. A partir destas análises, observaram-se, de maneira geral, que o grupo de dados de DSC apresentou bons valores, porém houve grande variação entre tais dados. De acordo com Roy *et al.* [13], os valores do DSC podem ser afetados pela grande diferença entre volumes das lesões e o volume total do cérebro, além disso espera-se que o valor médio seja próximo a 0,5 ou até menor, devido à variabilidade associada às segmentações manuais. Relativo ao VPP, o modelo apresentou excelentes resultados, indicando baixo índice de falsos positivos segmentados. Entretanto, a análise da DVA indicou subquantificação das lesões. É possível que o trabalho de Ghosal *et al.* [16] tenha apresentado métricas superiores, devido à utilização de 5 sequências de MRI. Ademais, as lesões escleróticas variam consideravelmente em localização, tamanho, forma e não apresentam bordas bem definidas, tornando a segmentação automática desafiante [13]. Uma possível solução para melhorar a identificação das lesões seria adicionar mais *skip connections* na arquitetura da rede. Além dessas observações, Valverde *et al.* [15] demonstrou que modelos utilizados para segmentar imagens de domínios diferentes em relação ao domínio de

treinamento apresentaram performance inferior aos modelos treinados considerando adaptação de domínio de imagem.

V. CONCLUSÃO

Métodos de *Deep Learning* têm se tornado popular para segmentação de imagens médicas. Neste estudo, a implementação de uma CNN em formato de U para segmentar lesões escleróticas na substância branca do cérebro com a utilização de apenas duas sequências de MRI foi apresentada. O tempo necessário para treinamento do modelo foi de aproximadamente duas horas e a segmentação das lesões foi realizada em poucos minutos. O trabalho apresentou ser possível segmentar as lesões escleróticas utilizando apenas as sequências FLAIR e T1 no modelo implementado de forma aceitável, tendo em vista os excelentes dados de VPP e plausíveis valores de DSC, apesar dos dados de DVA não serem tão próximos de zero. A partir desta pesquisa, acreditamos que os valores das métricas DSC e DVA, possam ser melhorados ao adicionar mais *skip connections* na arquitetura da CNN, para transferir melhor as informações sobre a localização das lesões do caminho de contração para o de expansão e assim ser possível detectar mais lesões. Além disso, os resultados podem ser aprimorados ao realizar um treinamento considerando adaptação de domínio, pois imagens de ressonância magnética adquiridas de diferentes equipamentos e protocolos variam em aparência.

REFERÊNCIAS

- [1] CW Stratton. “A review of multiple sclerosis as an infectious syndrome”. *Journal of Neurology & Neurophysiology*, vol. 7, pp. 400, 2016. doi: 10.4172/2155-9562.1000400.
- [2] Multiple Sclerosis International Federation, “Atlas of m.s.: mapping multiple sclerosis around the world”, 2020.
- [3] E. Waubant et al. “Environmental and genetic risk factors for ms: an integrated review”. *Annals of clinical and translational neurology*, vol. 6(9), pp. 1905–1922, 2019. doi:10.1002/acn3.50862.
- [4] H. Ford. “Clinical presentation and diagnosis of multiple sclerosis”, *Clinical Medicine*, vol. 20(4), pp. 380-383, 2020. doi: 10.7861/clinmed.2020-0292
- [5] W. J. Huang, W. W. Chen and X. Zhang. “Multiple sclerosis: pathology, diagnosis and treatments”. *Experimental and therapeutic medicine*, vol. 13(6), pp. 3163–3166, 2017. doi: 10.3892/etm.2017.4410.
- [6] Y. Ma et al. “Multiple sclerosis lesion analysis in brain magnetic resonance images: techniques and clinical applications”. *ArXiv Preprint: 2104.10029*, 2021.
- [7] S. Hunter. “Overview and diagnosis of multiple sclerosis”. *The American Journal of Managed Care*, vol. 22, pp.141–150, 2016.
- [8] E.M. Crowe, W. Alderson, J. Rossiter and C. Kent. “Expertise Affects Inter-Observer Agreement at Peripheral Locations within a Brain Tumor”, *Frontiers in Psychology*, vol. 8, pp.1628, 2017. doi: 10.3389/fpsyg.2017.01628.
- [9] C. Zeng, L. Gu, Z. Liu and S. Zhao. “Review of deep learning approaches for the segmentation of multiple sclerosis lesions on brain mri”. *Front. Neuroinform*, vol. 14, pp. 610967, 2020. doi: 10.3389/fninf.2020.610967.
- [10] Z. Akkus, A. Galimzianova, A. Hoogi, D.L. Rubin and B.J. Erickson. “Deep learning for brain mri segmentation: state of the art and future directions”. *Journal of digital imaging*, vol. 30(4), pp. 449–459, 2017. doi: 10.1007/s10278-017-9983-4.
- [11] E. Kondrateva, M. Pominova, E. Popova, M. Sharaev, A. Bernstein and E. Burnaev. “Domain shift in computer vision models for mri data analysis: an overview”. *ArXiv Preprint 2010.07222*, 2020.
- [12] I. Despotović, B. Goossens and W. Philips. “MRI segmentation of the human brain: challenges, methods, and applications”. *Computational and Mathematical Methods in Medicine*, 450341, 2015. doi:10.1155/2015/450341.
- [13] S. Roy, J. A. Butman, D. S. Reich, P. A. Calabresi and D. L. Pham. “Multiple sclerosis lesion segmentation from brain mri via fully convolutional neural networks”. *ArXiv Preprint: 1803.09172*, 2018.
- [14] S. Valverde et al. “Improving automated multiple sclerosis lesion segmentation with a cascaded 3d convolutional neural network approach”. *ArXiv Preprint: 1702.04869*, 2017.
- [15] S. Valverde et al. “One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks”. *ArXiv Preprint 1805.12415*, 2018.
- [16] P. Ghosal, P. K. C. Prasad, and D. Nandi. “A light weighted deep learning framework for multiple sclerosis lesion segmentation”. In *2019 Fifth International Conference on Image Information Processing (ICIIP)*, Shimla, India, 2019, pp. 526–531. doi: 10.1109/ICIIP47207.2019.8985674.
- [17] O. Ronneberger, P. Fischer and T. Brox. “U-net: convolutional networks for biomedical image segmentation”. In *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234-241. Springer, Cham. doi: 10.1007/978-3-319-24574-4_28.
- [18] M. d. Oliveira et al. "Quantification of brain lesions in multiple sclerosis patients using segmentation by convolutional neural networks," *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 2045-2048. doi: 10.1109/BIBM49941.2020.9313244.
- [19] R. Simões et al. “Automatic segmentation of cerebral white matter hyperintensities using only 3D FLAIR images”. *Magnetic Resonance Imaging*, vol. 31(7), pp. 1182-1189, 2013. doi: 10.1016/j.mri.2012.12.004.
- [20] A. Carass et al. “Longitudinal multiple sclerosis lesion segmentation: Resource and challenge”. *NeuroImage*, vol. 148, pp. 77–102, 2017. doi: 10.1016/j.neuroimage.2016.12.064.
- [21] O. Commowick, F. Cervenansky, and R. Ameli. “MSSEG challenge proceedings: multiple sclerosis lesions segmentation challenge using a data management and processing infrastructure”. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Athenes, Greece, 2016.
- [22] Ž. Lesjak, A. Galimzianova, A. Koren, M. Lukin, F. Pernuš, B. Likar, and Ž. Špiclin. “A novel public mr image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus”. *Neuroinformatics*, vol. 16(1), pp. 51–63, 2018. doi: 10.1007/s12021-017-9348-7.
- [23] A. Fedorov et al. “3d slicer as an image computing platform for the quantitative imaging network”. *Magnetic Resonance Imaging*, vol. 30(9), pp. 1323–41, 2012. PMID:

22770690. PMCID: PMC3466397. Available: <https://www.slicer.org/>.

[24] S. Bauer, T. Fejes, and M. Reyes. “A skull-stripping filter for itk”. *Insight Journal*, 2012.

[25] N. J. Tustison, B. B. Avants, P. A. Cook, A. Egan Y. Zheng, P. A. Yushkevich, and J. C. Gee. “N4ITK: improved n3 bias correction”. *IEEE Transactions on Medical Imaging*, vol. 29(6), pp. 1310–1320, 2010. doi: 10.1109/TMI.2010.2046908.

[26] H. Johnson, G. Harris, and K. Williams. “Brainsfit: mutual information registrations of whole-brain 3d images, using the insight toolkit”. *Insight Journal*, 2007.

[27] S. Grahl et al. “Evidence for a white matter lesion size threshold to support the diagnosis of relapsing remitting multiple sclerosis”. *Multiple Sclerosis and Related Disorders*, vol. 29, pp. 124–129, 2019. doi: 10.1016/j.msard.2019.01.042.