GAMER MRI : a deep dive in multiple sclerosis pathology and clinical disability

Inaugural dissertation

to
be awarded the degree of Dr. sc. med.

presented at
the Faculty of Medicine
of the University of Basel

by
Po-Jui Lu
from Taipei, Taiwan

Basel, 2022

Approved by the Faculty of Medicine
on application of

Prof. Dr. Cristina Granziera, University of Basel, *Primary advisor*
Prof. Dr. Philippe C. Cattin, University of Basel, *Secondary advisor*
Prof. Dr. Ender Konukoglu, ETH Zurich, *External expert*

Prof. Dr. Primo Leo Schär
Dean

Contents

Acknowledgment

I cannot begin to express my deepest gratitude to my supervisor, Prof. Dr. Cristina Granziera, for her invaluable advice and unwavering support during the whole PhD journey. I really appreciate all the constructive and in-depth discussions with her and her unparalleled expertise in the field, which were essential to the successful completion of the project. Moreover, I am thankful for her profound belief in and enthusiasm for my ideas.

I would also like to extend my sincere gratitude to Prof. Dr. Philippe Cattin, my second supervisor, for his helpful advice and support. I would like to thank Prof. Dr. Ender Konukoglu from ETH Zurich for kindly being the external reviewer evaluating the thesis. I would like to extend my thanks to previous and current colleagues in the ThINk group for the close collaboration and the fruitful discussion (in the alphabetical order): Dr. Muhamed Barakovic, Dr. Alessandro Cagol, Dr. Xinjie Chen, Dr. Riccardo Galbusera, Dr. Lester Melie-Garcia, Dr. Mario Ocampo-Pineda, Dr. Reza Rahmanzadeh, Dr. Ramona-Alexandra Todea and Dr. Matthias Weigel.

I very much appreciate the inspirational and comprehensive discussions with Dr. Benjamin Odry from Covera Health. Special thanks should also go to Dr. Robin Sandkühler from the Center for medical Image Analysis & Navigation (CIAN) group and Dr. Youngjin Yoo from Simens Healthineeers for the productive discussions. I would also like to thank Dr. Beat Fasel for the indispensable IT support in the Department of Biomedical Engineering and sciCORE for the computing resources.

In addition, I would like to acknowledge all the patients and healthy controls in the project. Without their willingness to share their time and their brain images, this project wouldn't be possible.

Finally, a heartful thanks to my beloved fiancée, Chu-Ying Lin, who is always there for me. I am extremely thankful to my family for their support and encouragement.

Summary

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system. Typical characteristics are multifocal inflammatory infiltration, demyelination, remyelination, and axonal loss in the microenvironment of brain tissue. The advanced MRI (aMRI) sequences and the quantitative measures derived from them can provide surrogate measurements on these microstructural changes. The information provided by aMRI is abundant, but partially redundant among them. There is a need to assess which aMRI or quantitative measures are important to a given task and explore the benefit of considering them jointly in studying MS axonal/myelin damage and repair.

We proposed and validated a novel method based on the convolutional neural network and gated attention mechanism (GAMER-MRI) in the application of well-understood stroke-related and multiple sclerosis-related lesion classification. The method gave an attention weight-based importance order of MR contrasts in line with clinical understanding. Next, we extended the method to tackle highly intercorrelated diffusion measures based on diffusion MRI in the classification of MS lesion and perilesional tissue. The correlation of selected measures with patient-level measures including the clinical scale of movement disability and the biological measure on the degraded axons were statistically significant, and the combinations of them had a stronger correlation. Last, we demonstrated the improvement of the method on the patient-level classification and a proposed approach to identify the brain regions contributing towards the importance of the images through the combination of the relevance maps and the corresponding attention weights.

Along with these developments, we demonstrated that GAMER-MRI was able to give us the importance of MR images from the local lesion-level analysis to the global patient-level analysis and be a new means to jointly combine the abundant information in different kinds of MRI images for a more comprehensive analysis in the future.

Chapter 1. Introduction

1.1 Multiple Sclerosis

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system, which affects around 2.8 million people worldwide. The mean age of diagnosis is 32 years, and female is two- to three-fold more likely affected (Dilokthornsakul et al., 2016; Walton et al., 2020). The MS hallmark includes multifocal inflammatory infiltration, demyelination, remyelination, and axonal loss. The compounding effect of these characteristics leads to various symptoms, including muscle weakness, numbness, tremor, blurry vision, fatigue, problems with bowel and bladder, speech difficulties, and movement disability. Some patients also suffer from mental health issues. MS patients are clinically classified as relapsing-remitting (RR), secondary-progressive (SP), or primary progressive (PP). Based on clinical metrics and the representation on MRI of disease activity, patients can be further categorized as active or non-active (Lublin et al., 2014). The two patient groups are particularly of interest because they exhibit quite opposite patterns in terms of axonal damage and repair. In active RRMS, the remyelination and the acute axonal damage, including demyelination, axonal transport disturbances, and axonal transections, have been shown increased (Albert et al., 2007; Barkhof et al., 2003a; Dziedzic et al., 2010a; Kuhlmann et al., 2002). On the other hand, in non-active progressive (naPMS), less acute axonal damage and remyelination, but dominated axonal loss have been found (Albert et al., 2007; Barkhof et al., 2003a; Dziedzic et al., 2010a; Kuhlmann et al., 2002). Therefore, these two distinctive groups of MS patients are suitable for an in-vivo study to disentangle the interplay between axonal damage and repair, such as remyelination and axonal reorganization.

1.2 MS Pathology

MS lesions are heterogeneous depending on the position, the tissue structure, the cellular composition, the degree of damage, and the level of repairment (Lucchinetti et al., 2000; Metz et al., 2014). White matter lesions and grey matter lesions are categorized in the position and have different pathologies. In white matter lesions, Blood-Brain-Barrier damage, significant inflammation, and gliosis can be observed, but not in grey matter lesions.

Neuronal, glial, and synaptic loss can be found in grey matter lesions (Geurts and Barkhof, 2008). In addition, Juxtacortical and periventricular lesions (JCL and PVL), for example, have distinctive remyelinating abilities and distinct extent of axonal and myelin damage in postmortem histopathological (Goldschmidt et al., 2009) and in-vivo PET (Poiron, 2018) studies.

MS lesions can be histopathologically categorized into four categories. Active lesions contain a large number of immune cells, including monocytes, macrophages, and microglia. Chronic active lesions have a rim of activated microglia/macrophages at the lesion boundary and demyelination or axonal loss in the lesion center. Inactive lesions are hypocellular and lack of oligodendrocytes, macrophages, and microglia. Remyelinating/remyelinated lesions present shorter and thinner myelin sheaths and oligodendrocyte progenitor cells, which generate remyelinating oligodendrocytes (Duncan et al., 2018; Kuhlmann et al., 2017; Patrikios et al., 2006). The surrounding tissue of lesions can also have some degrees of axon/myelin damage (Lieury et al., 2014; Mustafi et al., 2019).

Normal appearing white matter and grey matter (NAWM and NAGM) can also show signs of histopathologically identifiable damage, but look normal on conventional magnetic resonance imaging sequences (Geurts and Barkhof, 2008; Moll et al., 2011). The characteristics include a lesser degree of diffuse myelin and axonal damage and microglia clusters, compared with lesions (Granberg et al., 2017; Kutzelnigg et al., 2005; Lassmann, 2018). Depending on the closeness to white matter lesions, NAWM has different

pathologies (Moll et al., 2011). These heterogeneous MS lesions and normal-appearing brain tissue can be probed by conventional and advanced magnetic resonance imaging sequences (cMRI and aMRI) (Granziera et al., 2021; Rovira et al., 2015; Wattjes et al., 2015).

1.3 MRI

Magnetic Resonance Imaging (MRI) plays a fundamental role in MS diagnosis, prognosis, and disease monitoring on and off therapy (Wattjes et al., 2021). Among cMRI, white matter lesions appear hyperintense on the image of the Fluid-Attenuated Inversion Recovery (FLAIR) sequence, and grey matter lesions present as black holes on the image of the Magnetization Prepared - RApid Gradient Echo (MPRAGE).

aMRI can further provide metrics sensitive to various axonal components and structures (Granziera et al., 2021). **Myelin Water Fraction (MWF)** derived from **Myelin Water Imaging** quantifies the water locating within myelin layers and gives a surrogate measure of integrity of myelin. If there is demyelination, the MWF decreases. MWF is obtained by fitting multiple water components to the multi-echo T2 relaxometry as the T2 of the compartment for myelin water is around 10-20 ms and the one for the intra/extracellular water 50~200ms. The range of fraction is from 0 to 0.3 within brain (Laule et al., 2006; Nguyen et al., 2016). **Multi-shell diffusion MRI** measures the signal decay caused by water diffusion within a tissue in various directions subjective to the strength of diffusion gradients. This allows estimating different water compartments in fractions (e.g., intracellular volume fraction and isotropic volume fraction) by applying biophysical microstructural models based on different assumptions on the restricted diffusion within tissue. The models are, for example, the **Neurite Orientation Dispersion and Density Imaging (NODDI)** (Zhang et al., 2012), **NODDI with the Spherical Mean Technique (SMT-NODDI)** (Cabeen et al., 2019), and **Microstructure Bayesian approach** (Reisert et al., 2017). These diffusion measures are between 0 and 1. The intracellular volume fraction reflects the integrity of the neurites and the isotropic volume fraction indicates the movement of the free water. Therefore, in a white matter lesion, the

intracellular volume fraction is lower than surrounding normal appearing white matter (Novikov et al., 2019). **Quantitative Susceptibility Mapping (QSM)** estimates the magnetic susceptibility profile of tissue from the phase of the MR signal. The interesting part of the phase is induced by the local magnetic fields from the susceptibility of different molecules, such as diamagnetic myelin and paramagnetic iron content. Through the field inversion in addition to other post-processing steps, QSM can differentiate myelin and iron content and indicate various phases of myelin degradation. The value range within brain is within ±250 ppb. Chronic active MS lesions of a surrounding rim due to iron-laden macrophages can be identified on QSM (Deh et al., 2018; Spincemaille et al., 2020; Wang and Liu, 2015a). **Quantitative T1 relaxometry (qT1)** measures the T1 relaxation time from the excitation state back to thermal equilibrium, which depends on the spin-lattice relaxation subject to the tissue composition. Increasing water content and/or the less structural microenvironment increases qT1, such as the qT1 of white matter around 840 ms and the one cerebral spinal fluid is around 4000 ms (Bojorquez et al., 2017). It is, thus, sensitive to myelin, axonal diameter, and brain tissue organization. (Bonnier et al., 2014; Canty et al., 2013; Harkins et al., 2016; Kober et al., 2012; Stüber et al., 2014). **Magnetization Prepared 2 Rapid Gradient Echoes (MP2RAGE)** is an extension of MPRAGE. MP2RAGE exploits two inversion recoveries and the intrinsic bias field correction of the reception B_1 field (B_1^-) and the first order bias-field correction of the transmit B_1 field (B_1^+) in order to provide T1 weighted images with an optimal contrast between gray matter and white matter. qT1 can be estimated through the two inversion recovery images (Marques et al., 2010).

These aMRI sequences may jointly provide enlightening in-vivo descriptors of axonal damage and repair in MS patients and offer surrogate quantifiable biomarkers (Figure 1). However, most in-vivo studies of axonal damage and repair before this project utilized either one of mentioned aMRI sequences or a combination of two of them. How to jointly and selectively consider the information provided by aMRI and the derived quantitative measures to study the interplay between myelin and axonal damage/repair motivates this project.

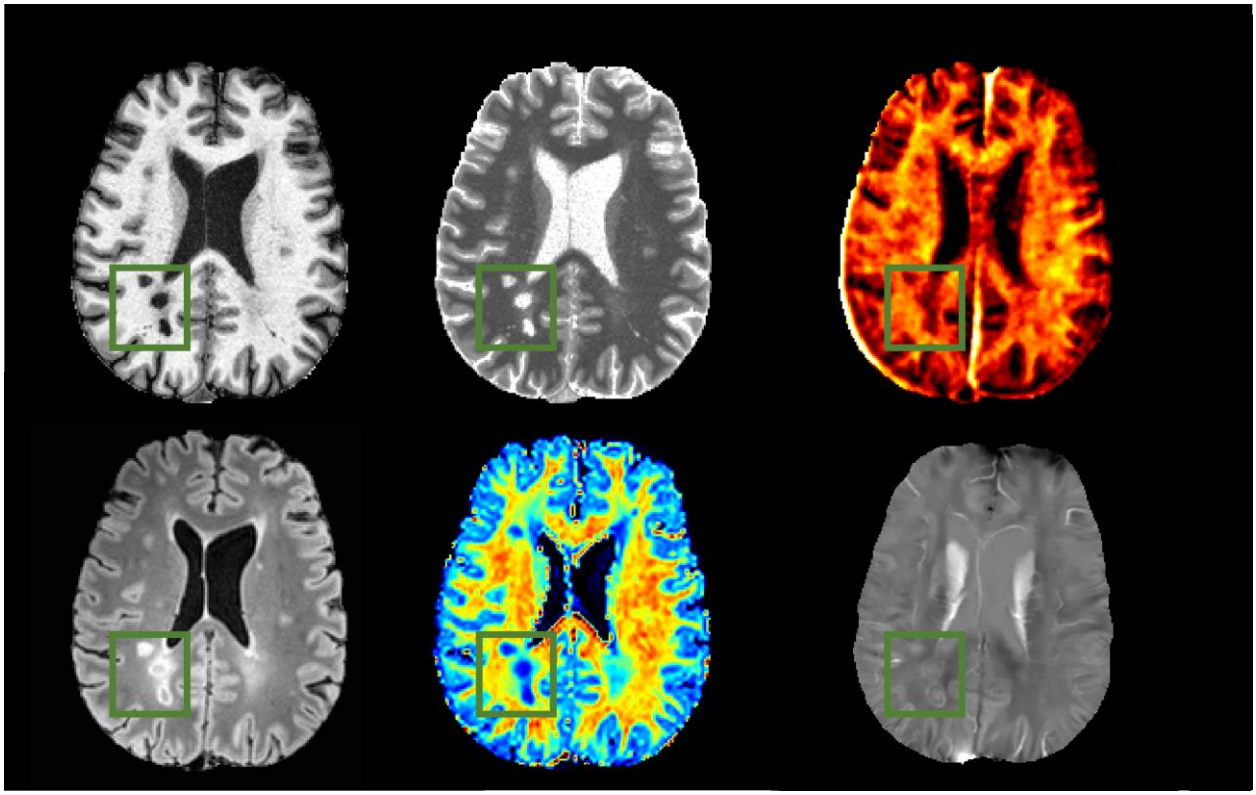


Figure 1:The white matter lesion representations of an exemplar MS patient on qualitative and quantitative MRI. Upper row from left to right: MP2RAGE, quantitative T1 relaxometry, Myelin Water Fraction. Lower row from left to right: FLAIR, NDI from NODDI and QSM. Lesions are indicated by the green squares. On MP2RAGE, lesions are hypointense and their centers can evolve like black holes. On quantitative T1 relaxometry, lesions are hyperintense because the demyelination and axonal damage increase the T1 relaxation time. On MWF, demyelination decreases the water in the myelin sheath, so lesions are hypointense. On FLAIR, the same process increases the T2 relaxation time and leads to hyperintense lesions. On NDI, the axonal damage and loss reduce the neurite density and thus lesions are hypointense. On QSM, lesions have rims formed by the iron-laden macrophages at the lesion boundaries, which are paramagnetic.

1.4 Deep Learning

The advancement of deep learning (DL) rendered possible the classification of disease patterns in MS patients (Barquero et al., 2020; Eitel et al., 2019; Shoeibi et al., 2021; Ye et al., 2020; Yoo et al., 2018; Zhang et al., 2018), prediction of mobile impairment (Marzullo et al., 2019; Tousignant et al., 2019) and accurate segmentation of MS lesions (Andermatt et al., 2018; Brosch et al., 2016; La Rosa et al., 2020) using MRI. The hierarchical structure and nonlinear activation operation of deep learning enable it to learn and extract meaningful patterns from raw or minimally pre-processed data (LeCun et al., 2015). Among various structures, the convolutional neural network (CNN) is suitable for MR images (Lundervold and Lundervold, 2019). The core components are the convolutional layer of different stationary kernels and the pooling layer in addition to the nonlinear activation. The convolutional layer learns the optimal values for the kernels to extract meaningful features (e.g., edges and textures) and its weight sharing nature enables it to detect the same kind of edge everywhere in the image. The pooling layer applies a summary statistic locally, such as maximum and mean operations, and the learned features are therefore invariant to small translations and condensed to lower dimensional features (Goodfellow et al., 2016). Through iteratively minimizing the error between the prediction and the ground truth of a given task given regularization constraints, CNN learns itself the features it deems relevant to the task.

Attention mechanism, which originated in the field of natural language processing, can instruct the NN to attend to useful correlated elements in the text. It was modeled by a feedforward NN inside the main part of the NN for translation. The feedforward NN considered a sequence of the encoded elements in the input sentence and the hidden state of the main NN and generated the probability of each element to be combined with the element to form an expected context vector for further decoding into translation (Bahdanau et al., 2015). Since then, various variants of the attention mechanism were developed (Lu et al., 2016; Luong et al., 2015; Niu et al., 2021; Vaswani et al., 2017; Voita et al., 2020). One of the variants, the gated attention mechanism, was

extended to images and found to successfully assign attention weights to non-overlapping patches from histopathological images containing malignant cancer cells (Ilse et al., 2018; Tomczak et al., 2018).

This data-driven characteristic can complement the use of summary statistics in clinical research. Often, the region of interest (ROI) was chosen based on human assumptions on the biological process. If the pathological conditions in ROI are complex, a summary statistic on certain aspects of the assumed process is calculated, while losing some information and considering a partial aspect of the process. In order to gain new insights into the damage and repair in MS, it would be imperative to have a deep learning tool providing a quantified measurement on the importance of the derived quantitative measures, which is relevant to a given task and based on all available information from the data. The additional benefits of quantified importance might establish the ground for forming a more informative pattern of the pathology.

1.5 Explainability Methods

The complex hierarchical NN has cast a limitation on how to understand its decision process. A plethora of methods has been developed to attempt to explain the decision of the NN. They include occlusion (Zeiler and Fergus, 2014), the saliency map(Simonyan et al., 2014), integrated gradients (Sundararajan et al., 2017), SHAP (Lundberg and Lee, 2017), GradCAM (Selvaraju et al., 2016) and Layer-wise Relevance Propagation (LRP) (Bach et al., 2015). Each of them is based on different rationales. The occlusion evaluates how much the performance changes when a patch of the image is whitened or zeroed. The saliency map considers the gradient of the output with respect to the input, i.e., how the small changes in the input image are going to affect the output. Integrated gradients is approximated by the sum of the gradients with respect to the input image along the straight-line path from a given baseline image to the input image. SHAP approximated the Shapley value in game theory with various approaches, including kernel-based and linear formulations. It gives a SHAP value to each feature by

summing the prediction differences between including and excluding that feature in all combinations of the subset of features in the dataset. GradCAM uses the gradients of prediction with respect to a convolutional layer to create a map and then upsamples the map to the resolution of the input image to reveal important regions. Among them, LRP has been shown effective in taking structures of NN into generation of the relevance map to explain the decision (Samek et al., 2021, 2017). LRP decomposes and redistributes the output score from the NN to the input image layer by layer based on the relative contributions of the neurons in two adjacent layers. The layer-by-layer approach can adapt different redistribution rules according to the kinds of the layers.

1.6 Contribution

In this thesis, we present a CNN-based method with gated attention mechanism (GAMER-MRI) to address the issue of joint and selective consideration of derived quantitative measures to decode the pattern of myelin and axonal damage and repair. The method provides attention weights as proxies of importance of input images of quantitative measures in a given task.

In the first work, the validity of attention weights was founded by giving attention weight-based important orders of qualitative MR contrasts and quantitative measures in line with the clinical understanding in three experiments. The first experiment was the classification of acute/subacute ischemic stroke patients vs healthy controls and patients with other brain pathology, where a pre-trained model was fine-tuned. The second experiment was the patch-based classification of acute ischemic stroke lesions vs healthy tissue. The last experiment was the classification of JCL and PVL MS lesions, supported by the passed statistical test result of the obtained attention weights.

In the second publication, we provided the solution when GAMER-MRI was applied to the highly correlated inputs. The experiment was conducted on the patch-level classification of lesion vs perilesional tissues using diffusion measures from eight biophysical microstructural models. The differences between lesions and perilesional

tissues, which also had various degrees of damage and repair, are smaller than between the lesions and healthy tissues. If GAMER-MRI could classify them with good performance, the diffusion measures being selected were highly sensitive. Due to their intercorrelated nature, their attention weights were close to each other. We proposed a simple multiplication within the network to enhance the differences in attention weights and a selection process based on attention weights. The selected measures averaged within and across lesions were shown to be correlated with the patient-level clinical and biological measures. The combinations of the selected measures had an even stronger correlation. This hinted at the possibility to combine quantitative measures for a more informative representation.

The third work illustrated two important aspects of GAMER-MRI. The first aspect was the applicability of GAMER-MRI to the patient-level classification using volumetric images of quantitative measures without a pre-trained model given a limited number of patients. The classification was performed on MS patients having severe movement disability vs mild disability. Second, it proposed an improvement of LRP in conjunction with the attention weights to uncover important regions for the classification. The experiments were the affected degree of classification performance because of the perturbation on the important regions of the quantitative measure maps and the correlation with the clinical measure of mobile impairment. The results showed that an informative map, which was obtained by linearly combining the attention weights and the relevance maps from LRP, unveiled the most important regions in both experiments.

By considering all these works, we demonstrated the suitability of GAMER-MRI in selective and joint consideration of quantitative measures in studying axonal and myelin damage/repair from the lesion-level to the patient-level analyses.

1.7 Outline

Based on the concepts in the introduction section, the details of the three aforementioned works are presented in Chapter 2 to 4. In Chapter 5, we discuss our work, limitations, and possible future work for further improvements and developments.

Chapter 2. GAMER-MRI: Gated-Attention MEchanism Ranking of multi-contrast MRI in brain pathology

In the publication of the following section, we presented a deep learning-based method, Gated-Attention MEchanism Ranking of multi-contrast Magnetic Resonance Imaging in brain pathology (GAMER MRI), to address the issue of selection of the most relevant magnetic resonance imaging contrasts for a given pathology. A multitude of novel quantitative and semiquantitative MRI techniques have provided new and complimentary information about the pathophysiology of neurological diseases. However, clinical research and practice are still limited by the time required to acquire multiple MR contrasts. The results showed that GAMER-MRI was able to rank the relative importance of MR measures in the classification of well-understood ischemic stroke lesions and in the classification of multiple sclerosis lesions where the relative importance of MR measures was less understood. In addition, the quantified importance may in fact help to choose the best combination of MR contrasts for a specific classification problem.

Publication.

The paper was published in the journal NeuroImage: Clinical on the 3rd of December 2020.

https://www.sciencedirect.com/science/article/pii/S2213158220303594

GAMER MRI: Gated-Attention MEchanism Ranking of multi-contrast MRI in brain pathology

Po-Jui Lu[1,2], Youngjin Yoo[3], Reza Rahmanzadeh[1,2], Riccardo Galbusera[1,2], Matthias Weigel[1,2,4], Pascal Ceccaldi[3], Thanh D. Nguyen[5], Pascal Spincemaille[5], Yi Wang[5], Alessandro Daducci[6], Francesco La Rosa[7,8,9], Meritxell Bach Cuadra[7,8,9], Robin Sandkuehler (Robin Sandkühler)[10], Kambiz Nael[11,12], Amish Doshi[12], Zahi A. Fayad[12,13], Jens Kuhle[2], Ludwig Kappos[2], Benjamin Odry[14], Philippe Cattin[10], Eli Gibson[3], Cristina Granziera[1,2]

1. Translational Imaging in Neurology (ThINk) Basel, Department of Medicine and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland.

2. Neurologic Clinic and Policlinic, Departments of Medicine, Clinical Research and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland.

3. Digital Technology and Innovation, Siemens Healthineers, Princeton, USA

4. Division of Radiological Physics, Department of Radiology, University Hospital Basel, Basel, Switzerland

5. Department of Radiology, Weill Cornell Medical College, New York, USA

6. Computer Science Department, University of Verona, Verona, Italy

7. Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne, Switzerland

8. Medical Image Analysis Laboratory, Center for Biomedical Imaging (CIBM), University of Lausanne, Switzerland

9. Department of Radiology, Lausanne University Hospital and University of Lausanne, Switzerland

10. Center for medical Image Analysis & Navigation, Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland.

11. Department of Radiological Sciences, David Geffen School of Medicine at University of California Los Angeles

12. Department of Diagnostic, Molecular and Interventional Radiology, Icahn School of Medicine at Mount Sinai

13. BioMedical Engineering and Imaging Institute, Department of Radiology, Icahn School of Medicine at Mount Sinai

14. AI for Clinical Analytics, Covera Health, New York, USA

Abstract

Introduction: During the last decade, a multitude of novel quantitative and semiquantitative MRI techniques have provided new information about the pathophysiology of neurological diseases. Yet, selection of the most relevant contrasts for a given pathology remains challenging. In this work, we developed and validated a method, Gated-Attention MEchanism Ranking of multi-contrast MRI in brain pathology (GAMER MRI), to rank the relative importance of MR measures in the classification of well-understood ischemic stroke lesions. Subsequently, we applied this method to the classification of multiple sclerosis (MS) lesions, where the relative importance of MR measures is less understood.

Methods: GAMER MRI was developed based on the gated attention mechanism, which computes attention weights (AWs) as proxies of importance of hidden features in the classification. In the first two experiments, we used Trace-weighted (Trace), apparent diffusion coefficient (ADC), Fluid-Attenuated Inversion Recovery (FLAIR), and T1-weighted (T1w) images acquired in 904 acute/subacute ischemic stroke patients and in 6,230 healthy controls and patients with other brain pathologies to assess if GAMER MRI could produce clinically meaningful importance orders in two different classification scenarios. In the first experiment, GAMER MRI with a pretrained convolutional neural network (CNN) was used in conjunction with Trace, ADC, and FLAIR to distinguish patients with ischemic stroke from those with other pathologies and healthy controls. In the second experiment, GAMER MRI with a patch-based CNN used Trace, ADC, and T1w to differentiate acute ischemic stroke lesions from healthy tissue. The last experiment explored the performance of patch-based CNN with GAMER MRI in ranking the importance of quantitative MRI measures to distinguish two groups of lesions with different pathological characteristics and unknown quantitative MR features. Specifically, GAMER MRI was applied to assess the relative importance of the myelin water fraction (MWF), quantitative susceptibility mapping (QSM), T1 relaxometry map (qT1), and neurite density index (NDI) in distinguishing 750 juxtacortical lesions from 242 periventricular lesions in 47 MS patients. Pair-wise permutation t-tests were used to evaluate the differences between the AWs obtained for each quantitative measure.

Results: In the first experiment, we achieved a mean test AUC of 0.881 and the obtained AWs of FLAIR and the sum of AWs of Trace and ADC were 0.11 and 0.89, respectively, as expected based on previous knowledge. In the second experiment, we achieved a mean test F1 score of 0.895 and a mean AW of Trace=0.49, of ADC=0.28, and of T1w=0.23, thereby confirming the findings of the first experiment. In the third experiment, MS lesion classification achieved test balanced accuracy=0.777, sensitivity=0.739, and specificity=0.814. The mean AWs of T1map, MWF, NDI, and QSM were 0.29, 0.26, 0.24, and 0.22 ($p < 0.001$), respectively.

Conclusions: This work demonstrates that the proposed GAMER MRI might be a useful method to assess the relative importance of MRI measures in neurological diseases with focal pathology. Moreover, the obtained AWs may in fact help to choose the best combination of MR contrasts for a specific classification problem.

Keywords (Maximum 6 words)

- Deep learning
- Attention mechanism
- Relative Importance order
- Stroke
- Multiple sclerosis
- Quantitative MRI

Highlights (3-5 bullet points, 85 characters including spaces per point):

- The attention mechanism can rank MR measures by relative importance.
- Proposed guideline for use of the attention mechanism with MR measures.
- Attention weights and quantitative MR measures can potentially form new patterns.

1. Introduction

Magnetic resonance imaging (MRI) has proven invaluable for the investigation of the pathophysiology of neurological diseases and guiding neurological diagnoses, prognoses, and evaluation of therapeutics. In fact, during the last decade, numerous fast MRI sequences and quantitative/semiquantitative MRI measures have been developed that provide complementary information to disentangle the pathological mechanisms and characteristics of brain diseases. In addition, specific biomarkers for diagnosis and response to therapy have been identified (Bozzali et al., 2016; González and Schwamm, 2016; Gupta et al., 2017). However, clinical research and practice are still limited by the time required to acquire multiple MR contrasts. It is imperative that these studies be conducted in a time frame compatible with patient tolerance, compliance, and in the case of clinical practice, the requirements dictated by the healthcare system. Therefore, the need to address the selection of the most informative MR contrasts is pivotal to avoid uncomfortably lengthy acquisitions, to lower the subsequent possibility of having motion artifacts, and to reduce the related cost.

Deep learning, especially convolutional neural networks (CNN), has proven promising in the segmentation of brain regions or lesions in MR images (Andermatt et al., 2018; Carass et al., 2017; Commowick et al., 2016; La Rosa et al., 2019; Wachinger et al., 2018), classification of brain diseases (Payan and Montana, 2015; Yoo et al., 2018), MR reconstruction (Akçakaya et al., 2019; Schlemper et al., 2018), and prediction of disease prognosis (Saha et al., 2020; Tousignant et al., 2019). The layer-wise neural network (NN) design can identify high-level hidden representations through iterative training, which are pivotal for a given classification task. Some of the deep learning designs specifically enhance the interpretability of the decision made by the NN, such as class activation maps (Selvaraju et al., 2016; Zhou et al., 2016) and Shapley Additive exPlanations (Lundberg and Lee, 2017). Nevertheless, these methods either give importance to the voxels in images or to post-hoc feature importance. On the contrary, the attention mechanism within a NN provides attention weights (AWs) representing the importance of specific features. The concept, which originated in the field of natural language processing, can instruct the NN to attend to useful correlated elements in the text (Bahdanau et al., 2015). One of its variants, the gated attention mechanism, was

extended to images and found to successfully assign AWs to non-overlapping patches from histopathological images in the classification of malignant cancer cells (Ilse et al., 2018; Tomczak et al., 2018).

In this work, we optimized the gated attention mechanism (Ilse et al., 2018) to develop a prototype of a Gated-Attention MEchanism Ranking of multi-contrast MRI in brain pathology (GAMER MRI). GAMER MRI specifically ranks the relative importance of global multi-contrast features, instead of the importance of local single-contrast patches, in the classification of focal lesions. This method was first validated for a clinical application where some MR-measure importance is known (e.g., ischemic stroke) and was then applied to the classification of specific subtypes of MS lesions, which are known to differ in the extent of myelin/axon damage and reparative capacity: this provided knew knowledge about which MRI measure – among those sensitive to axon and myelin integrity – is most suitable to distinguish lesions with different axon/myelin damage and repair in MS.

2. Materials and methods

2.1. MRI data

2.1.1. Stroke data

A total of 7,134 1.5T and 3T brain MRI studies obtained from a combination of inpatient and outpatient scanners at the Mount Sinai Hospital, New York, USA were randomly selected as the dataset. These imaging data were accumulated from the Mount Sinai BioMedical Engineering and Imaging Institute's HIPAA compliant Imaging Research Warehouse, including data from 10 scanners produced by two manufacturers (GE and Siemens Healthineers). The dataset consisted of various clinical acquisitions and

	TE (ms)	TR (ms)	FOV (mm^3)	SR (mm^3)	TI (ms)	b values (s/mm^2)
FLAIR	94	8000	230x230x160	0.72x0.72x5	2460	--
T1w	6.9	2876	179x220x160	0.69x0.69x5	840	--
DWI	113.8	7625	240x240x170	1.02x1.02x5	--	0,1000

Table 1. Acquisition parameters of each contrast in the stroke dataset. TE: echo time; TR: repetition time; TI: inversion time; FOV: field of view; SR: spatial resolution.

included healthy controls, patients with subacute and acute infarct stroke, and patients with subacute and acute hemorrhage and mass effect. Among these patients, 904 are subacute and acute infarct stroke patients (defined as group 1) and 6,230 are healthy controls and other patients (defined as group 2). The 2D axial protocol included conventional, isotropically weighted Diffusion Weighted Imaging (DWI), Fluid-Attenuated Inversion Recovery (FLAIR), and T1-weighted images (T1w) from the inversion recovery pulse sequence. The most important mean acquisition parameters are listed in Table 1. Trace-weighted contrast (Trace) and apparent-diffusion coefficient (ADC) were reconstructed on the scanner from DWI.

Acute infarct stroke has distinctive representations on the acquired contrasts (Fig. 1). In the acute phase, hyperintensity is seen on Trace while ADC appears hypointense (Allen et al., 2012). In the subacute infarct stroke phase, both contrasts develop towards pseudo-normality. The segmentation of acute stroke lesions was performed on Trace and ADC by an expert radiologist consulting for Siemens Healthineers.

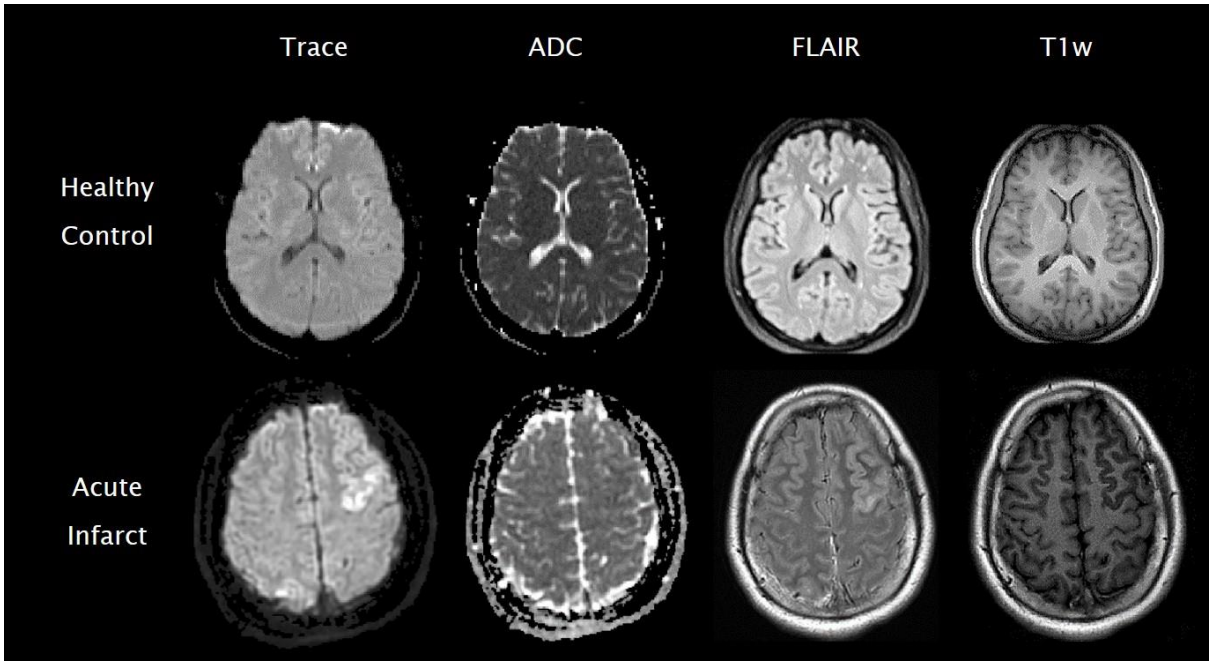


Figure 1. Examples of Trace, ADC, FLAIR and T1w images in the stroke dataset. The lesion is hyperintense on Trace but hypointense on ADC (Allen et al., 2012). On T1w, the lesion is more isointense than ADC and is faintly hyperintense on FLAIR.

2.1.2. Multiple Sclerosis data

Forty-seven MS patients (33 relapsing-remitting and 14 progressive, 31 females and 16 males, age range=43.6±14.4 years) were enrolled in the study approved by the local ethics committee of Basel University Hospital. Written consent was obtained prior to the MRI acquisition. Patients underwent a multi-parametric protocol on a 3T Siemens Healthineers MAGNETOM Prisma MRI system. The 3D protocol included SPACE-based FLAIR, Magnetization-Prepared 2 RApid Gradient Echoes (MP2RAGE) (Kober et al., 2012; Marques et al., 2010), Fast Acquisition with Spiral Trajectory and T2prep sequence (FAST-T2) (Nguyen et al., 2016), multi-shell Diffusion Weighted Imaging (mDWI), and Multi-Echo Gradient Recalled Echo MRI (ME-GRE) (Wang and Liu, 2015b). The most important acquisition parameters are shown in Table 2.

	TE (ms)	TR (ms)	FOV (mm^3)	SR (mm^3)	TI (ms)	Additional Parameters
FLAIR	386	5000	256x256x256	1x1x1	1800	--
MP2RAGE	3	5000	256x256x256	1x1x1	700, 2500	--
ME-GRE	6.7,10.8,14.8,18.9, 22.9,27,31.1,35.1, 39.2,43.2	49	195x240x180	0.75x0.75x3	--	--
FAST-T2	0.5	7.5	240x240x160	1.25x1.25x5	--	**T2prep times (ms)** 0 (T2prep turned off),7.5,17.5, 67.5,147.5,307.5
mDWI	75	4500	256x256x144	1.8x1.8x1.8	--	**b values (s/mm^2)** 0;700;1000; 2000;3000 /149 directions in total

Table 2: Acquisition parameters of each contrast in the MS dataset. TE: echo time; TR: repetition time; TI: inversion time; FOV: field of view; SR: spatial resolution.

From multi-parametric MRIs, quantitative MR maps (qMRs) were further reconstructed. Quantitative T1 relaxometry map (qT1) was reconstructed from MP2RAGE as in (Kober et al., 2012). Myelin water fraction map (MWF) was reconstructed

from FAST-T2 as in (Nguyen et al., 2016). Neurite density index (NDI) from the neurite orientation dispersion and density imaging model (Zhang et al., 2012) was reconstructed from DWI as in (Daducci et al., 2015). Quantitative Susceptibility Map (QSM) was reconstructed from ME-GRE as in (Wang and Liu, 2015b). Co-registration between images was performed using FMRIB Software Library (FSL) (Jenkinson et al., 2012) and FreeSurfer (Fischl et al., 2001), and the obtained transformation matrices were later used for finding the correspondence of MS lesions between different qMRs. qMRs were not resampled to the same resolution so that the effect of interpolation in the resampling would not confound the quantitative values reflective of physical characteristics. MS lesions in white matter (WM) show hyperintensities on FLAIR images and in grey matter (GM) are blackholes on MP2RAGE images of the uniform contrast in Fig. 2a. WM lesions were automatically segmented (La Rosa et al., 2019) and manually corrected by two expert raters. Juxtacortical lesions (JCLs) and periventricular lesions (PVLs) were defined as WM lesions located within 3 mm of the boundary between (i) WM and GM and (ii) WM and ventricles, respectively, in Fig. 2b. The aforementioned boundaries were obtained through FreeSurfer processing on MP2RAGE (Fujimoto et al., 2014). In the end, 750 JCLs and 242 PVLs were found with a class-imbalance ratio of 1:3.

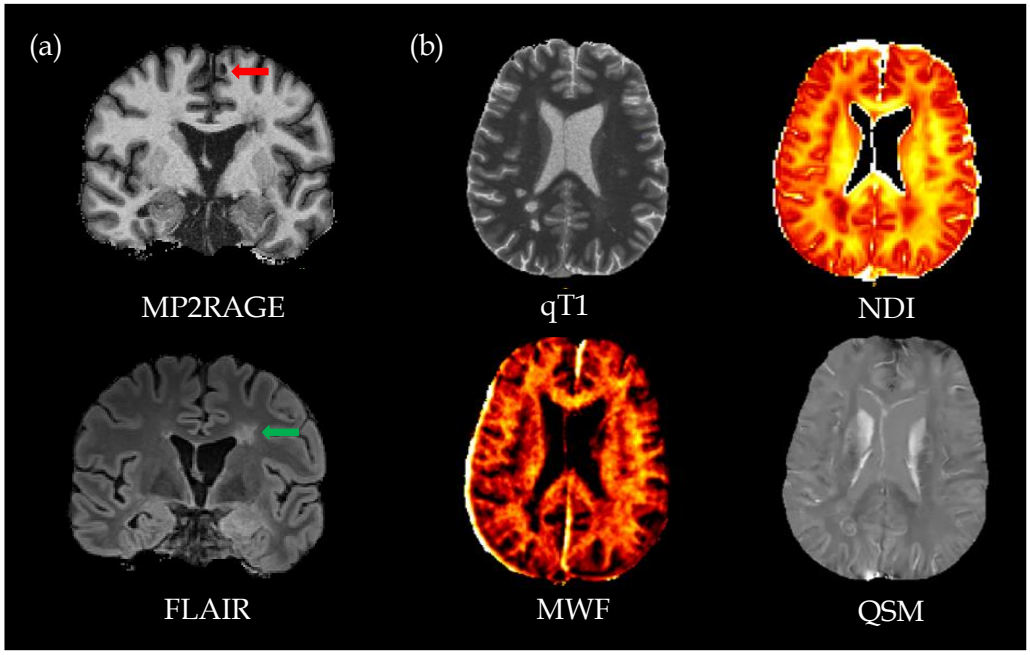


Figure 2. MS lesions and qMRs. In (a), on MP2RAGE, the MS lesion in GM is a black hole (red arrow) and on FLAIR, the MS lesion in WM is hyperintense (green arrow). In (b), qT1, NDI, MWF and QSM reflective of different aspects of the microenvironment illustrate various representations of lesions.

2.1.3. Study Summary

Fig. 3 summarizes the information about the two datasets and the training, validation, and test datasets in the following experiments.

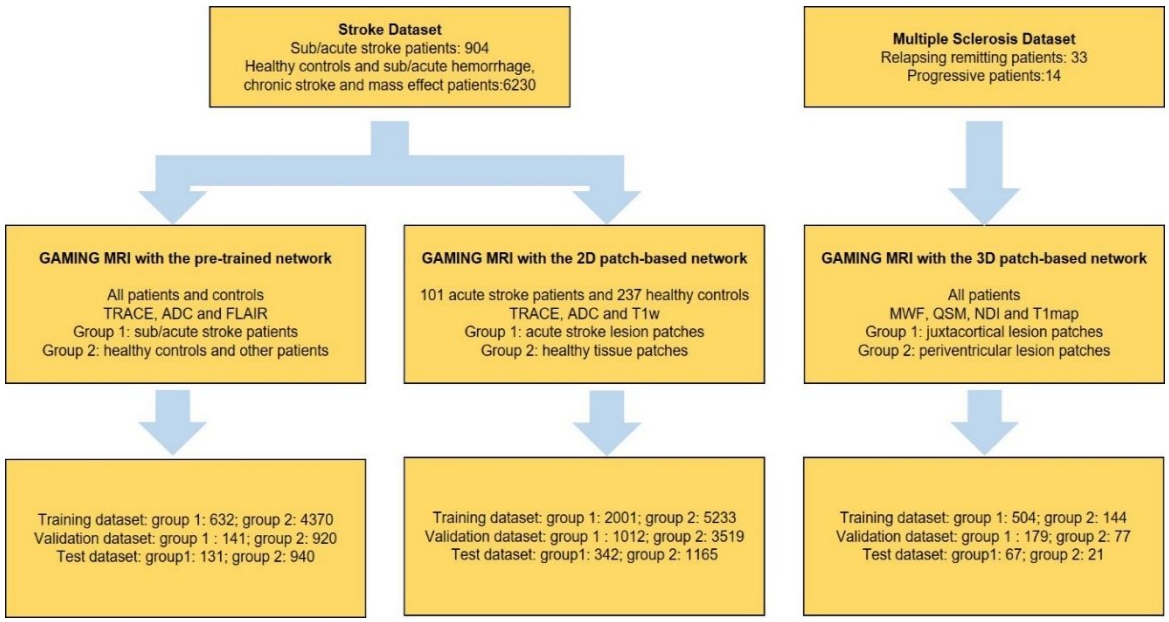


Figure 3. Study summary

2.2. GAMER MRI

The original gated attention mechanism proposed by Ilse et al. (Ilse et al., 2018) exploits the hidden representations of single-contrast patches to compute the corresponding AWs, which represents the relative importance among the hidden representations in the classification. The main theorem behind this rationale is the following (Zaheer et al., 2017):

Theorem 1. A prediction function $f(X)$ for a set of countable elements X is invariant to the permutation of the elements in X, if and only if, for suitable transformations g and h, $f(X)$ can be decomposed as:

$$f(X) = h(\sum_{x \in X} g(x)) \tag{1}$$

$g(.)$ and $h(.)$ were modeled by a NN. Based on (1), the gated attention mechanism is formulated as follows:

$$\boldsymbol{n} = \sum_{l=1}^{L} a_l \boldsymbol{m}_l = \sum_{x \in X} g(x) \tag{2}$$

$$a_l = softmax\left(\boldsymbol{w}^T\left(\tanh(\boldsymbol{U}\boldsymbol{m}_l) \odot \mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)\right)\right) \tag{3}$$

where $\boldsymbol{m}_l$ is the hidden representation of the l^{th} instance, a_l is the AW of the l^{th} instance, $\boldsymbol{U}$ and $\boldsymbol{V} \in R^{K\times M}$ are weights of the fully connected layers (FCs) following the hidden representations, sigm stands for the non-linear sigmoid function, $\odot$ is the element-wise multiplication operator, $\boldsymbol{w} \in R^{1\times K}$ is the weights of a FC, softmax stands for the softmax function.

Contrary to the original single-contrast approach to model $g(.)$, GAMER MRI adopted the multi-contrast multi-path approach on different MR contrasts and (2) becomes:

$$\sum_{l=1}^{L} a_l \boldsymbol{m}_l = \sum_{l=1}^{L} a_l q_l(x_l) = \sum_{x\in X} g(x) \tag{4}$$

where $q(x)$ is the encoding function of the NN and Equation (3) remains the same. It is a simple variant to extend the meaning of AWs to the assessment of the importance of the MR contrasts in studying diseases and the parallel encoding paths enable the MR contrasts to be ranked by AWs. The core implementation of the gated attention mechanism in the NN was the same as in (Ilse et al., 2018) and formed by an FC followed by the hyperbolic tangent function (the attention layer) and an FC followed by the sigmoid function (the gate layer). The outputs of the attention layer and the gate layer were element-wise multiplied and connected to a one-neuron FC and the softmax function to generate the normalized AWs. The number of neurons in the attention and gate layers depends on the experiment.

In order to validate our method and rank the importance of MRI features, the following three experiments were conducted: 1. volume-based classification of acute/subacute ischemic stroke vs other strokes and healthy controls; 2. patch-based classification of acute ischemic stroke lesions vs healthy tissue and 3. patch-based classification of JCLs vs PVLs in MS patients.

2.2.1. Pretrained network with GAMER MRI on stroke

To assess the performance of GAMER MRI as a ranking method, we combined GAMER MRI with the feature extracting compartment of an in-house pretrained NN from Siemens Healthineers (Princeton, NJ, USA), for the classification of acute/subacute

ischemic stroke vs other patients and healthy controls using volumetric Trace, ADC, and FLAIR. The pretrained NN was trained for the same classification and thus learned how to encode relevant hidden features from Trace, ADC, and FLAIR.

2.2.1.1. Inputs and preprocessing

Trace, ADC, and FLAIR images were considered in this experiment since these contrasts were used for training the pretrained network. Subacute and acute infarct stroke patients were categorized into group 1, while group 2 included other patients and healthy controls. There were 5,002 subjects (group 1: 632 and group 2: 4,370) in the training dataset. The validation dataset had 1,061 subjects (group 1: 141 and group 2: 920) and 1,071 subjects (group 1: 131 and group 2: 940) were in the test dataset.

2.2.1.2. Architecture

The combined NN was built with three main compartments, including the feature extracting compartment of the pretrained NN, GAMER MRI, and a classifier, as depicted in Fig. 4. The feature extracting compartment was, for each contrast, composed of two 3D convolutional blocks followed by two dense blocks based on the concept of DenseNet in (Huang et al., 2017). Each convolutional block consisted of a batch normalization layer, leaky ReLU units, and a 3D convolutional layer. In each dense block, there were two 3D convolutional blocks with the kernel size of 3x3x3 and 1x1x1. The number of initial features was 16 and the growth rate was 2. The hidden feature vectors from all contrasts were then concatenated as the input to the following GAMER MRI so that the hidden feature vector of each contrast was encoded independently prior to the computation of AWs. In the GAMER MRI, the number of neurons each in the attention layer and in the gate layer was 400. The classifier was one sigmoid neuron receiving the weighted sum of the hidden features and the AWs. The importance of each contrast is represented by the AW.

2.2.1.3. Training Strategy

The combined NN was trained with a cross-entropy loss function and mini-batches. The weighted sampler was used to account for the class imbalance during training. The network parameters, including the pretrained layers, were updated by the Adam optimizer

with decoupled weight decay (AdamW) (Loshchilov and Hutter, 2019). The evaluation metric was the area under the receiver operating characteristic curve (AUC), which was the same metric used in training the pretrained network. To avoid overfitting, data augmentation was independently performed for each contrast on-the-fly. Since there is inherent randomization in the initialization of network parameters and the split of mini-batches, the assessment of the effect of the random initialization is needed to properly describe the behavior of repeatability. The training, validation, and test datasets were kept the same during training, but the random seed changed in each repetition in the repeatability experiment. The leave-one-out (LOO) experiment on the selection of sequences was also conducted to characterize the method from a different perspective, namely by measuring the drop in the evaluation metrics reflecting the impact of the missing channel.

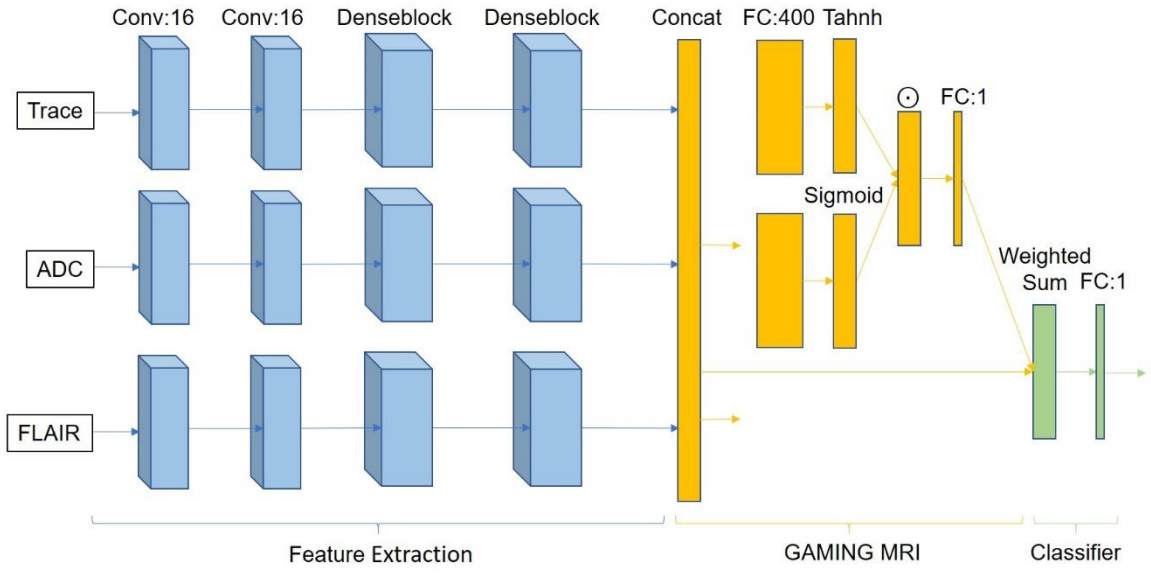


Figure 4: The network structure consists of the pretrained feature extraction, the GAMER MRI and the classifier. Conv stands for the 3D convolutional block. FC is the fully connected layer. Concat is the concatenating layer. ⊙ represents an element-wise multiplication.

2.2.2. Patch-based network with GAMER MRI on stroke

The second experiment was performed to assess the ability of the GAMER MRI in a neural network when it was trained from scratch on the stroke dataset. We hypothesize that if GAMER MRI can provide the weights reflective of the current clinical understanding

in the classification of acute infarct stroke lesions versus healthy tissues, it can be used in disease studies where the relative importance of MR contrasts is still unknown.

2.2.2.1. Inputs and preprocessing

In consideration of the limited number of existing acute infarct stroke lesion masks and in order to remove the effects of different scanners, Trace, ADC, and T1w from 101 acute infarct patients without other pathologies, like hemorrhage, and 237 healthy controls were selected from the stroke dataset for the patch-based experiment. T1w was registered to b0 of DWI because the right correspondence between contrasts is essential to patch sampling. Because acute infarct lesions are of varying sizes, care must be taken when choosing the sampled patch size. Too large of a patch size is detrimental to small lesions. On the other hand, too small of a patch size would under-represent large lesions. Thus, after inspecting a subset of acute infarct stroke images, 24x24 voxels was empirically chosen for 2D patches. For healthy controls, the patches were randomly upsampled three times within the brains so that the healthy brains would not be under-represented by a small number of patches. In the end, 3,355 lesion patches and 9,917 healthy patches were sampled. Patches were divided into training, validation, and test datasets according to the ratios: 0.6, 0.3, and 0.1. As a result, there were 7,234 patches (2,001 lesion patches and 5,233 healthy patches) in the training dataset; 4,531 patches (1,012 lesion patches and 3,519 healthy patches) in the validation dataset; 1,507 patches (342 lesion patches and 1,165 healthy patches) in the test dataset. The patches containing acute infarct stroke lesions were given the label=1.

2.2.2.2. Architecture

A patch-based multi-contrast CNN with GAMER MRI (NN2) could be decomposed into three compartments as the NN in 2.2.1.2 (Fig. 5). The feature extracting compartment included three convolutional blocks for each MR contrast. Each convolutional block included a convolutional layer of 128 filters, exponential leaky units, and a batch normalization layer. The number of filters was chosen based on the evaluation metrics without inspecting the AWs prior to the 100-time repetitions. The three connected convolutional blocks were followed by an FC of 128 neurons encoding the hidden feature

vector for each contrast. In the GAMER MRI, the number of neurons in the attention layer and in the gate layer were both 64. The classifier was the same as in 2.2.1.2.

2.2.2.3. Training strategy

The NN was trained with a weighted cross-entropy loss function to account for the effect of class imbalance. The mini-batch size was 128 for both training and evaluation. The optimizer was Adam (Kingma and Ba, 2015). The F1 score was chosen as the evaluation metric because the correct identification for positive cases, i.e., acute infarct stroke, was more important than healthy tissue. To avoid overfitting, data augmentation and early stopping were performed.

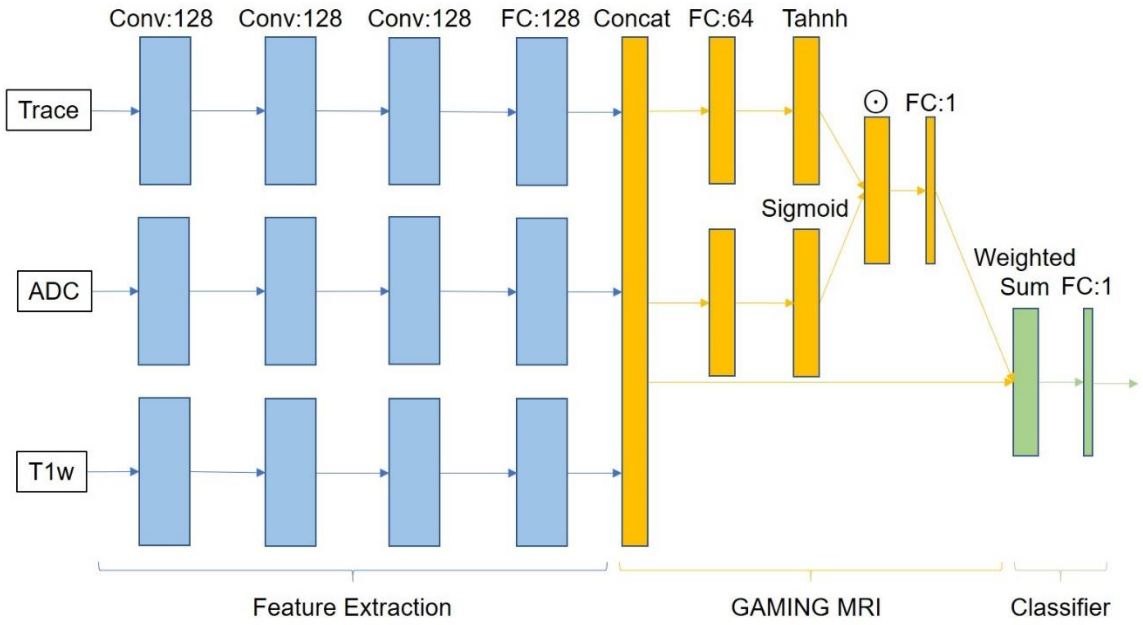


Figure 5: The network structure is composed of the feature extraction, the GAMER MRI and the classifier. Conv stands for a convolutional block of 2D convolutional filters. FC is the fully connected layer. Concat is the concatenating layer. ⊙ represents the element-wise multiplication.

To appropriately characterize the performance, in addition to the aforementioned training strategy on the different random initializations, the next level of assessment was to split training, validation, and test datasets differently in different repetitions to make sure the power of the method does not come from the split.

2.2.3. Patch-based network with GAMER MRI on MS lesions

The third experiment was the evaluation of the mechanism on the classification of JCLs and PVLs from the MS dataset using qMRs, where the relative importance is unknown in the clinic.

2.2.3.1. Inputs and preprocessing

3D patches close to 5x5x5 mm^3 were chosen as samples for training the neural network for the following three reasons: JCLs and PVLs are defined within 3 mm regions, the minimal slice thickness of qMRs is 5 mm, and various resolutions. This led to different patch sizes for each qMR to avoid confounding the quantitative values by the interpolation in the registration. Considering the defined JCLs and PVLs being in the WM, each qMR was masked by the WM mask. Lesion patches were divided into training, validation, and test datasets following the ratios: 0.6, 0.3, and 0.1. Therefore, there were 648 lesion patches (504 JCLs and 144 PVLs) in the training dataset, 256 lesion patches (179 JCLs and 77 PVLs) in the validation dataset, and 88 lesion patches (67 JCLs and 21 PVLs) in the test dataset.

2.2.3.2. Architecture

A patch-based multi-contrast CNN with a GAMER MRI similar to the NN in 2.2.2.2, was built (Fig.6). The feature extraction compartment included two convolutional blocks followed by an FC, as in 2.2.2.2, for each qMR. The convolutional layer in the convolutional block had 32 filters and the FC has 16 neurons encoding the hidden feature vector for each qMR. The criterion to choose the number of filters was the same as in 2.2.2.2. The hidden feature vectors from all qMRs were then concatenated as the input to the following GAMER MRI. In GAMER MRI, the number of neurons in the attention layer and gate layer were both 32. The classifier was the same as in 2.2.2.2.

2.2.3.3. Training strategy

The loss function and the mini-batch size, data augmentation, early stopping, and the learning-rate-reduce-plateau scheduler were the same as in 2.2.2.3. The optimizer and the evaluation metric were the same as in 2.2.1.3.

In addition to the characteristics evaluated in the previous two experiments, resampling patches prior to the split of datasets was performed. To avoid sampling bias in the patch-

based classification, randomly resampling patches is pivotal for reproducibility. The pair-wise one-sided 10,000 permutation t-tests were performed on the obtained orders of AW

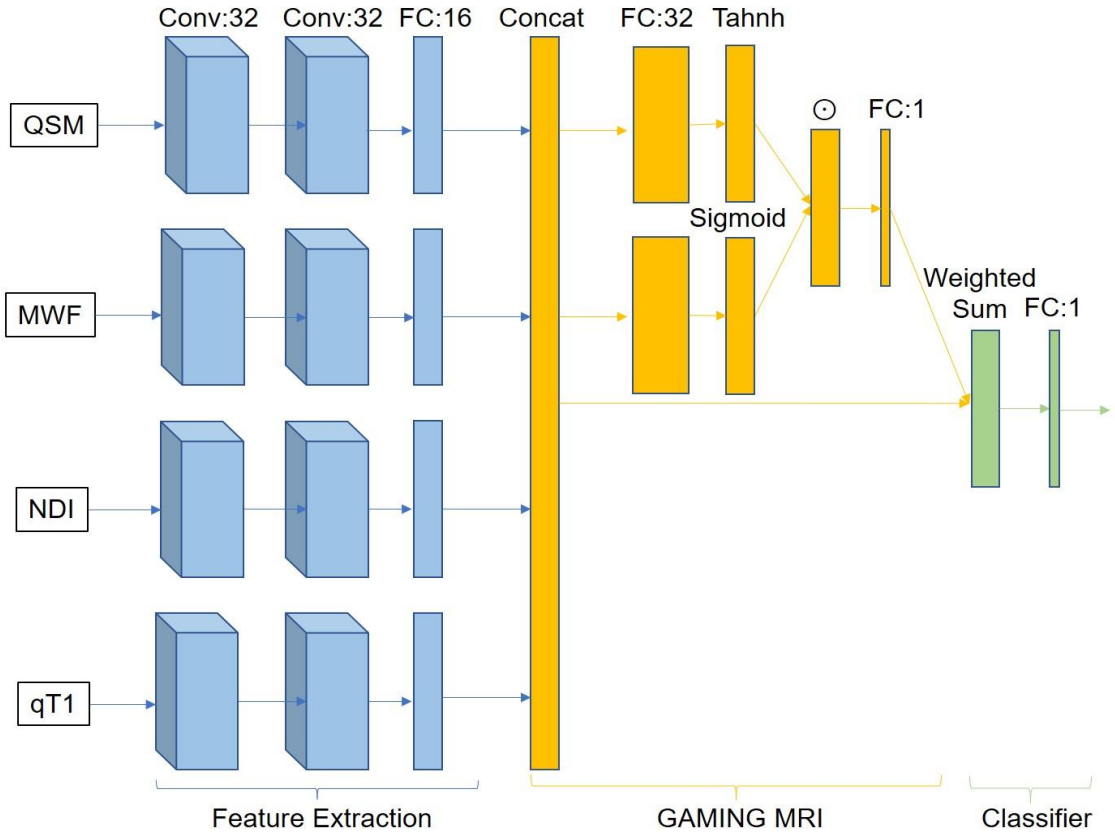


Figure 6: The network structure includes feature extraction, GAMER MRI and classifier. Conv stands for a convolutional block of 3D convolutional filters. FC is the fully connected layer. Concat is the concatenating layer. ⊙ represents the element-wise multiplication.

of all repetitions and the multiple comparison problem was tackled by Bonferroni correction.

2.3. Data and code availability statement

The datasets, provided by the Mount Sinai Hospital and Basel University Hospital, used in this study are not publicly available because the IRB of the study limits access to the data. The code used for training the models has dependencies on Siemens' internal tooling, infrastructure, and hardware, and its release is therefore not feasible. However, the architecture, layer details, and hyperparameters are described in sufficient details in the manuscript to support replication with non-proprietary libraries.

3. Results

3.1. Pretrained network with GAMER MRI on the stroke dataset

Validation and test results of the NN in 2.2.1 using three different random seeds for the random sampler, which led to different initializations and splits of mini-batches, are given in Table 3. In each repetition, the mean AW (mAW) was averaged over the AWs of the corrected predicted samples. The reported mean AWs (rmAWs) were the average of all mAW across repetitions.

Pretrained-network with GAMER MRI in stroke		
Dataset	**Validation**	**Test**
AUC	0.919	0.881
AWs	Trace+ADC, FLAIR (0.890, 0.110)	Trace+ADC, FLAIR (0.886, 0.114)

Table 3: Pretrained network with GAMER MRI on the stroke dataset: Mean validation and test results over three repetitions. The mean area under the curve (AUC) is reported to show the classification performance and the sum of reported mean attention weights (rmAWs) of Trace and ADC and the rmAW of FLAIR are shown to provide the importance ranking of the MRI.

The LOO experiment was conducted twice for each pair combination of Trace, ADC, and FLAIR. The drops in validation AUC were averaged across the repetitions and compared between combinations in Table 4.

Leave-one-out Experiment			
Input Contrasts	**ADC, FLAIR**	**Trace, FLAIR**	**Trace, ADC**
AUC	0.866	0.914	0.91
Drop in AUC	0.052	0.004	0.008
AWs	ADC, FLAIR (0.843, 0.157)	Trace, FLAIR (0.866, 0.134)	Trace, ADC (0.505*, 0.495*)

Table 4: The validation result of the leave-one-out experiment. The averaged AUC, the drop in performance, and the rmAWs are reported. For the experiment using only Trace and ADC, the rmAWs varied greatly in the repetitions. The rmAWs in the first repetition were (Trace, ADC)=(0.871, 0.129) and in the second repetition were (0.138, 0.861).

Patch-based network with GAMER MRI in stroke	**Validation**					
Training Strategy	Different Random Initialization			Different Split		
F1 score	0.917			0.895		
	Trace	ADC	T1w	Trace	ADC	T1w
AWs	0.5 ±0.043	0.293 ±0.036	0.207 ±0.028	0.49 ±0.049	0.281 ±0.039	0.229 ±0.030
	Trace, ADC and T1w	Trace, T1w and ADC	Other combinations	Trace, ADC and T1w	Trace, T1w and ADC	Other combinations
Order in Repetitions	98	2	---	82	18	---
Dataset	**Test**					
Training Strategy	Different Random Initialization			Different Split		
F1 score	0.885			0.895		
	Trace	ADC	T1w	Trace	ADC	T1w
AWs	0.507 ±0.043	0.284 ±0.035	0.208 ±0.028	0.494 ±0.050	0.278 ±0.040	0.228 ±0.029
	Trace, ADC and T1w	Trace, T1w and ADC	Other combinations	Trace, ADC and T1w	Trace, T1w and ADC	ADC, Trace and T1w	Other combinations
Order in Repetitions	94	6	---	80	19	1	---

Table 5: The validation and test result of the patch-based acute stroke classification. The averaged F1 score and the rmAWs and the standard deviation of mAWs across 100 repetitions are reported. The order in repetitions obtained by 100 repetitions shows the number of the corresponding order of attention.

3.2. Patch-based network with GAMER MRI on the stroke dataset

The NN in 2.2.2 was evaluated 100 times using the same training, validation and test datasets with different random initialization of the network and mini-batches.

Furthermore, the NN was evaluated 100 times using different splits of training, validation, and test datasets; respective validation results are reported in Table 5.

3.3. Patch-based network with GAMER MRI on MS lesions

The NN in 2.2.3 was trained 100 times on the resampled datasets to ensure reproducibility of the method for classification of MS lesions where the important order of sequences is unknown. The performance of the repetition experiment on the validation and test datasets is reported in Table 6. In Table 7, we report the mean and standard deviation of the mean AWs, which are defined as in 3.1. Furthermore, the results of the permutation t-test on the obtained order of AWs are reported.

Mean metrics (%)	Balanced Accuracy	Sensitivity	Specificity	F1 score
Validation dataset	78.34 ±3.09	74.21 ±7.86	82.49 ±5.12	64.66±4.00
Test dataset	77.65 ±5.49	73.90 ±10.10	81.41 ±8.21	65.09±8.93

Table 6: Performance of the patch-based network on MS lesions. The average mean and standard deviation of the metrics as a percentage of 100 repetitions are reported.

	Validation			
Contrast	**AW**	**Statistical test**	**P value**	**Significance**
qT1	0.285 ±0.027	**qT1 > MWF**	0.0001	***
MWF	0.256 ±0.015	**MWF > NDI**	0.0001	***
NDI	0.241 ±0.014	**NDI > QSM**	0.0001	***
QSM	0.218 ±0.022	---	---	---
	Test			
Contrast	**AW**	**Statistical test**	**P value**	**Significance**
qT1	0.284 ±0.030	**qT1 > MWF**	0.0001	***
MWF	0.256 ±0.016	**MWF > NDI**	0.0001	***
NDI	0.241 ±0.021	**NDI > QSM**	0.0001	***
QSM	0.218 ±0.023	---	---	---

Table 7: The rmAWs, the standard deviation of mAWs and the statistical test on the pair-wise comparison. The upper section shows the results of the validation dataset and in the lower section are the results of the test dataset. ***: corrected p<0.001

4. Discussion

We developed a Gated-Attention MEchanism Ranking of multi-contrast MRI in brain pathology (GAMER MRI) and demonstrated its ability to rank the relative importance of MRI contrasts / qMRs in the three different classification scenarios including the differentiation of well-studied infarct strokes and that of less understood MS lesions.

4.1. Pretrained network with GAMER MRI on stroke

To accomplish the classification task, the NN should be able to extract unique and common information from the input contrasts. We demonstrated in 3.1 that GAMER MRI could utilize the unique and common information from each contrast to provide the AW as a proxy of the importance of each contrast. The mean AUC in this experiment was comparable to the performance of the original pre-trained network in a similar classification task. In addition, the mean AUC of validation and test datasets (Table 3) indicated that the combination of a pretrained encoder and GAMER MRI well performed. Because the AWs of the correctly classified samples formed the correct pattern with the hidden features for the classifier to make the right decision, we then proceeded to average those AWs to obtain the mAW for each repetition: in fact, considering the AWs of the incorrectly classified samples would not have reflected their real importance in the identification of stroke lesions. The consistent ratio between the sum of rmAWs of Trace and ADC and the rmAW of FLAIR showed that FLAIR was less important compared to the other two contrasts in the given classification task. This is in line with the relative clinical importance of these contrasts for the diagnosis of acute and subacute infarct stroke (González and Schwamm, 2016).

We observed an inconsistent ratio between the rmAWs of Trace and that of ADC, which is probably due to the strong correlation between the contrasts. Because of the known evolvement of the infarct stroke representation from the acute to the subacute stage on Trace and ADC, the representations become pseudonormal and similar. This leads to a stronger correlation between the information brought by Trace and ADC in addition to the intrinsic physical correlation between these two contrasts.

In Table 4, the obvious drop in performance of the LOO experiment when Trace was excluded indicates that Trace provides more unique information than ADC and FLAIR. Indeed, in the two LOO experiments including FLAIR, the rmAWs of FLAIR were smaller

suggesting its relative lower importance in this classification, which echoes the result in Table 3. In the experiment performed without using FLAIR images, the inconsistent ratio of AWs between Trace and ADC bolsters the implication of confounds caused by the strong correlation between these two contrasts. One main assumption behind the previous interpretation is that the amount of extractable information is the same across the LOO experiments. The comparison of the obtained results with the current clinical understanding of stroke lesions indicates this assumption is valid.

Considering the results in 3.1, we obtained empirical evidence that the proposed variant of the attention mechanism can provide AWs representative of the importance of non-correlated contrasts.

4.2. Patch-based model with attention mechanism

We then aimed to assess whether GAMER MRI combined with a NN, which is trained from scratch for the classification of acute infarct stroke, is able to provide AWs representative of the known relative importance of MR contrasts in clinical practice.

The obtained F1 scores on both validation and test datasets (Table 5) indicated that the NN combined with GAMER MRI learned information necessary to classify the patch containing acute infarct stroke lesions regardless of being trained with different random initialization or different split of training, validation, and test datasets.

In both repetition experiments, the rmAWs of Trace, ADC, and T1w showed that Trace carried more characteristic information than ADC followed by T1w images, which is in accord with the clinical understanding that on Trace, acute infarct stroke is hyperintense and more obvious than on ADC, and that the light hypointense appearance of an acute stroke in T1w images is less evident than the drop in signal often observed in ADC maps (González and Schwamm, 2016). In addition, while the number of repetitions slightly varied across the different experimental conditions, the importance order of Trace, ADC, and T1w remained consistent (Table 5). This demonstrates that the NN was able to extract relevant information in the given samples. The order of importance of Trace, T1w, and ADC might be a result of the patches having similar information on Trace and ADC. As a result, the NN would learn less unique information from ADC leading to its lower mAW than the one of Trace.

The consistent results in 3.2 validated the assumption that the AWs obtained with GAMER MRI can be used to assess the relative importance of MRI contrasts without the restriction on an informative pre-trained NN. Also considering the results in 3.1, which demonstrated that GAMER MRI could obtain a clinically meaningful ranking of MRI contrasts, the method may be well applicable to neurological diseases that are less understood.

4.3. MS patch-based model with attention mechanism

In this last scenario, we aimed at assessing if the GAMER MRI could be applied to other MRI measures and diseases, where the relative importance of measures is less understood. Therefore, we studied whether the GAMER MRI could rank myelin/axonal sensitive measures such as qT1, MWF, NDI, and QSM to classify lesions that are known to have different myelin and axonal content, such as lesions located near the ventricles (PVL: lower myelin and axonal content) and next to the cortex (JCL: relatively higher myelin and axonal content) (Goldschmidt et al., 2009; Tonietto, 2018).

For both the validation and the test datasets, the network exhibited a moderate performance (Table 6): balanced accuracy was ca 78% - with a specificity that was slightly higher than the sensitivity (74% vs 82%), and the F1 score was ca 65%. In this experiment, different than in the previous one, we have assessed the network performance by using other summary measures than the F1 score: this is essential because the F1 score does not consider true negative results, hence it may not equally consider lesions, whose characteristics are not completely understood (i.e., JCLs and PVLs). The multiple statistical tests on pairwise rmAWs showed that the metric best discriminating PVL vs JCL microstructure is qT1 followed by MWF, NDI, and QSM. qT1 quantifies the overall microstructural tissue damage within MS lesions (Bonnier et al., 2014), whereas MWF and NDI provide specific information about myelin and axonal content (Nguyen et al., 2016; Zhang et al., 2012). The order of importance reflects the overall difference in myelin/axonal content revealed in pathological studies (Goldschmidt et al., 2009), which qT1 depicts with the highest sensitivity. Hence, through this experiment, we could establish the reliability of GAMER MRI in a context where the relative contribution of MR measures to the discrimination of focal pathology is not clear.

Compared to the results obtained on the stroke dataset, the smaller differences between rmAWs of different qMRs might be caused by the smaller size of MS lesion datasets and/or higher similarities between lesion groups. A much larger effect is expected if an increased number of samples in datasets is included. Another potential underlying cause of this difference is the fact that the applied qMRs have in part redundant information. Indeed, the microstructural environment measured by qT1 encompasses the myelin content and neuro-axonal integrity measured by MWF and NDI. On the other hand, QSM measures both iron deposition and myelin properties since it is sensitive to the susceptibility effect due to paramagnetic substances and to the orientation of myelin sheaths. Besides, it has to be considered that – different than the contrasts applied in stroke (e.g., Trace) – qMRs in the MS experiment could not sharply delineate the boundary of MS lesions, hereby reflecting the local variations surrounding the focal damage. Despite all this, however, GAMER-MRI still demonstrated a statistically significant difference between rmAWs of the qMRs.

4.4. Guideline on GAMER MRI

In consideration of the obtained results, we propose to use GAMER MRI as follows:

1. Train and evaluate the method multiple times to see if there is a strong or mild correlation between the resultant AWs of input measures. If there is a strong correlation, an ablation study should be performed to remove the correlated modality showing a smaller drop in performance. Train and evaluate the method on the remaining measures to obtain AWs.
2. If there is no strong correlation, the importance order based on the mean AWs across the repetitions is recommended.

4.5. Conclusion

Our work shows that GAMER MRI provides a clinically meaningful order of importance for MR-based features in the classification of infarct strokes. In addition, even though qMRs in the classification of JCLs and PVLs in MS had redundant information, GAMER MRI still managed to reveal a close but significant order of importance. Considering this importance order, it may be possible to reduce the number of input MRI measures while retaining most of the useful information.

The main disadvantage of this method is the need for multiple evaluations since the criteria for the so-called strong correlation is based on the AWs, not just on the input

contrasts, as shown in the experiments of NN2. Future work will be required to remove this constraint. Furthermore, future work should center on combining the proxy quantification of the importance of qMRs with the values of qMRs to form meaningfully combined patterns for further studies since qMRs characterize different physical processes and physiological environments.

Declaration of competing interest

Part of the work was performed while Po-Jui Lu was doing internship in Siemens Healthineers, Princeton, USA. Youngjin Yoo, Pascal Ceccaldi, and Eli Gibson are employed by Siemens Healthineers, Princeton, USA. Benjamin Odry is employed by Covera Health, New York, USA. Matthias Weigel has received research funding by Biogen for developing spinal cord MRI. Kambiz Nael has consulted for Olea Medical outside the scope of this work. Zahi Fayad has a research grant from Siemens Healthineers and is founder and board member of Trained Therapeutix Discovery.

CRediT authorship contribution statement

Po-Jui Lu: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **Reza Rahmanzadeh**: Data Curation, Writing – reviewing& editing. **Riccardo Galbusera**: Data Curation, Writing – reviewing& editing. **Matthias Weigel**: Resources, Data Curation, Writing – reviewing& editing. **Youngjin Yoo**: Methodology, Resources, Writing – reviewing& editing. **Pascal Ceccaldi**: Methodology, Resources, Writing – reviewing& editing. **Thanh D. Nguyen**: Resources, Writing – reviewing& editing. **Pascal Spincemaille**: Resources, Writing – reviewing& editing. **Yi Wang**: Resources, Writing – reviewing& editing. **Alessandro Daducci**: Resources, Writing – reviewing& editing. **Francesco La Rosa**: Data Curation, Writing – reviewing& editing. **Meritxell Bach Cuadra**: Resources, Writing – reviewing& editing. **Robin Sandkuehler**: Conceptualization, Writing – reviewing& editing. **Kambiz Nael**: Resources, Writing – reviewing& editing. **Amish Doshi**: Resources, Writing – reviewing& editing. **Zahi Fayad**: Resources, Writing – reviewing& editing. **Jens Kuhle**: Writing – reviewing& editing. **Ludwig Kappos**: Writing – reviewing& editing. **Benjamin Odry:** Methodology, Writing – reviewing& editing. **Philippe Cattin**: Supervision, Writing – reviewing& editing. **Eli Gibson**: Supervision, Methodology, Writing – reviewing& editing. **Cristina Granziera**: Supervision, Conceptualization, Funding Acquisition, Resources, Writing – reviewing& editing.

Acknowledgments and Disclaimer

We would like to acknowledge all the patients and healthy controls in this project. This project is supported by Swiss National Funds PZ00P3_154508, PZ00P3_131914, and PP00P3_176984. Francesco La Rosa and Meritxell Bach Cuadra are supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie project TRABIT (agreement No 765148) and by the Centre d'Imagerie BioMedicale (CIBM). We would also thank the Mount Sinai Hospital for providing the stroke dataset and Basel University Hospital for acquiring the MS dataset. The concepts and information presented in this paper are based on research results and GAMER MRI is not commercially available.

Chapter 3. GAMER-MRI in Multiple Sclerosis identifies the diffusion-based micro-structural measures that are most sensitive to focal damage: a deep-learning based-analysis and clinico-biological validation

In the following paper, we presented an improvement of GAMER-MRI to select the most discriminating, but intercorrelated diffusion-based microstructural measures in the classification of MS lesions and perilesional tissue. Diffusion measures of biophysical microstructure models derived from diffusion MRI can characterize the microstructural environment and probe the demyelination and the axonal injury of MS lesions. However, a microenvironment characteristic is measured differently by measures from different models due to the different biophysical assumptions. Therefore, how to select the most discriminating diffusion measures for a given neurological disorder and how to combine the complementary information they might provide are imperative issues to address. The results showed that the proposed extension of GAMER-MRI with a selection process can select sensitive diffusion measures, whose average over lesions of a patient correlated with the patient-level clinical measure of mobile impairment, the Expanded Disability Status Scale (EDSS), and the patient-level biological measure of neuron damage, the serum level of neurofilament light chain (sNfL). Furthermore, the combination of the selected measures had a stronger correlation with EDSS and sNfL than the individual measures. This work demonstrates that the proposed method might be useful to select the microstructural measures that are most discriminative of focal tissue damage, and that may also be combined to a unique contrast to achieve stronger correlations to clinical disability and neuroaxonal damage.

Publication. This work was published in the research topic Computational Neuroimage Analysis Tools for Brain (Diseases) Biomarkers of the section Brain Imaging Methods in the journal of Frontiers in Neuroscience on the 6th of April 2021.
https://www.frontiersin.org/articles/10.3389/fnins.2021.647535/full

GAMER-MRI in Multiple Sclerosis identifies the diffusion-based micro-structural measures that are most sensitive to focal damage: a deep-learning based-analysis and clinico-biological validation

Po-Jui Lu[1,2,3], Muhamed Barakovic[1,2,3], Matthias Weigel[1,2,3,4], Reza Rahmanzadeh[1,2,3], Riccardo Galbusera[1,2,3], Simona Schiavi[5], Alessandro Daducci[5], Francesco La Rosa[6,7,8], Meritxell Bach Cuadra[6,7,8], Robin Sandkühler[9], Jens Kuhle[2,3], Ludwig Kappos[2,3], Philippe Cattin[9], Cristina Granziera[1,2,3]

1. Translational Imaging in Neurology (ThINk) Basel, Department of Medicine and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland

2. Neurologic Clinic and Policlinic, Departments of Medicine, Clinical Research and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland

3. Research Center for Clinical Neuroimmunology and Neuroscience Basel, University Hospital Basel and University of Basel, Basel, Switzerland

4. Division of Radiological Physics, Department of Radiology, University Hospital Basel, Basel, Switzerland

5. Department of Computer Science, University of Verona, Verona, Italy

6. Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne, Switzerland

7. Medical Image Analysis Laboratory, Center for Biomedical Imaging (CIBM), University of Lausanne, Switzerland

8. Department of Radiology, Lausanne University Hospital and University of Lausanne, Switzerland

9. Center for medical Image Analysis & Navigation, Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland

** Correspondence: Cristina Granziera cristina.granziera@unibas.ch*

Keywords: Multiple sclerosis, Deep learning, Advanced Quantitative Diffusion MRI, Relative importance order, Clinically correlated measure selection.

Abstract

Conventional magnetic resonance imaging (cMRI) in multiple sclerosis (MS) patients provides measures of focal brain damage and activity, which are fundamental for disease diagnosis, prognosis, and the evaluation of response to therapy. However, cMRI is insensitive to the damage to the micro-environment of the brain tissue and the heterogeneity of MS lesions. In contrast, the damaged tissue can be characterized by mathematical models on multi-shell diffusion imaging data, which measure different compartmental water diffusion.

In this work, we obtained 12 diffusion measures from eight diffusion models and we applied a deep-learning attention-based convolutional neural network (CNN) (GAMER-MRI) to select the most discriminating measures in the classification of MS lesions and the perilesional tissue by attention weights. Furthermore, we provided clinical and biological validation of the chosen metrics - and of their most discriminative combinations - by correlating their respective mean values in MS patients with the corresponding Expanded Disability Status Scale (EDSS) and the serum level of neurofilament light chain (sNfL), which are measures of disability and neuroaxonal damage. Our results show that the neurite density index from Neurite Orientation and Dispersion Density Imaging (NODDI), the measures of the intra-axonal and isotropic compartments from the microstructural Bayesian approach and the measure of the intra-axonal compartment from the Spherical Mean Technique NODDI were the most discriminating (respective attention weights were 0.12, 0.12, 0.15 and 0.13). In addition, the combination of the neurite density index from NODDI and the measures for the intra-axonal and isotropic compartments from the microstructural Bayesian approach exhibited a stronger correlation with EDSS and sNfL than the individual measures.

This work demonstrates that the proposed method might be useful to select the microstructural measures that are most discriminative of focal tissue damage, and that may also be combined to a unique contrast to achieve stronger correlations to clinical disability and neuroaxonal damage.

1. Introduction

Conventional magnetic resonance imaging (cMRI) in multiple sclerosis (MS) plays a major role in MS diagnosis, prognosis, and the evaluation of patients' therapeutic response (Rovira et al., 2015; Wattjes et al., 2015). However, the heterogeneity of focal MS lesions, the pathology in normal-appearing white and grey matter (NAWM and NAGM), and the specific damage to myelin and axons are largely overlooked by cMRI. Multi-shell diffusion-weighted imaging (mDWI) provides a way to further probe tissue damage and repair in MS patients (Lakhani et al., 2020; Schneider et al., 2017). mDWI measures signal changes that are related to the diffusion of water molecules within central nervous system (CNS) tissue (Lakhani et al., 2020; Novikov et al., 2019), which is constrained by the local micro-environment (Novikov et al., 2019). This enables diffusion measures of biophysical microstructure models derived from mDWI to decode the information specific to different water compartments (e.g. intra-axonal and isotropic compartments) within the CNS tissue (Novikov et al., 2019). The intra-axonal compartment reflects the integrity of the neurites and the isotropic compartment indicates the movement of the free water (Novikov et al., 2019). These two compartments can describe the two pathological presentations of MS lesions, demyelination and axonal injury, and are commonly modeled by various biophysical microstructure models (Lakhani et al., 2020).

A micro-environment characteristic is measured differently by the measures from different mathematical models due to the different assumptions on the diffusion within the tissue. Yet, to our knowledge, the direct comparison of all considered diffusion measures on MS lesions and the possibility to combine them do not exist. Therefore, how to select the most discriminating diffusion measures for a given neurological disorder and how to combine the complementary information they might provide remain to be an open question and motivate this study.

Convolutional neural network (CNN) in deep learning has proven promising in various applications of MR images and is able to encode spatial patterns on the images into representative hidden features (Akçakaya et al., 2019; Andermatt et al., 2018; La Rosa et al., 2020; Saha et al., 2020; Yoo et al., 2018). In our previous work (Lu et al., 2021b), we used an attention-based CNN – GAMER MRI – to rank the importance of the

input quantitative MRIs in the classification of stroke and MS lesions. Here, we further developed the method to select discriminating inter-correlated diffusion measures in the classification of MS lesions and perilesional tissue. Compared to the conventional feature selection methods, this CNN-based method enables to utilize maximally available spatial information of the images and does not need to decide on how to find representative values for the samples of each contrast, such as the mean value only within a lesion neglecting the perilesional tissue. In addition, the method jointly considers all the contrasts, which is a limitation for most of the conventional feature selection methods. Furthermore, in this study, we have explored the relationship between the chosen measures, or their combinations, with the Expanded Disability Status Scale (EDSS) and the neurofilament light chain in the serum (sNfL), which are respectively (i) a clinical measure of disability in MS patients and (ii) a biological measure of neuroaxonal damage (Barro et al., 2018; Siller et al., 2019).

2. Material and methods

2.1. MRI data

One hundred twenty-three MS patients (84 relapsing-remitting and 39 progressive, 71 females and 52 males, age range=44.7±14.0, median EDSS=2.5, EDSS range 0.0-8.0) were enrolled in the study, which was approved by the local Ethics Committee of

	TE (ms)	TR (ms)	FOV (mm^3)	SR (mm^3)	TI (ms)	Additional Parameters
FLAIR	386	5000	256x256x256	1x1x1	1800	--
MP2RAGE	3	5000	256x256x256	1x1x1	700, 2500	--
mDWI	75	4500	256x256x144	1.8x1.8x1.8	--	b values (s/mm^2) 0/12 acquisitions and 12 reverse encoding acquisitions; 700;1000; 2000;3000 /137 directions in total

Table 1: Acquisition parameters of each contrast in the MS dataset. TE: echo time; TR: repetition time; TI: inversion time; FOV: field of view; SR: spatial resolution.

Basel University Hospital. All subjects gave written consent prior to the enrollment. MS patients underwent a multi-parametric protocol on a 3T whole-body MR system (Siemens MAGNETOM Prisma). The protocol included 3D SPACE-based FLAIR, 3D Magnetization-Prepared 2 RApid Gradient Echoes (MP2RAGE) (Marques et al., 2010), and mDWI (Table1).

mDWI was denoised by MRtrix (Cordero-Grande et al., 2019; Tournier et al., 2019). The correction of susceptibility-induced distortion with the reversed phase-encoding images, eddy currents, and movement was performed by FMRIB Software Library (FSL) (Andersson et al., 2003; Andersson and Sotiropoulos, 2016; Jenkinson et al., 2012; Smith et al., 2004). The quantitative diffusion measures for the isotropic and intra-axonal compartments were reconstructed from the eight open-source biophysical models, including Ball and Stick (Behrens et al., 2003), Neurite Orientation and Dispersion Density Imaging (NODDI)2(Zhang et al., 2012), NODDI with the Spherical Mean Technique (SMT-NODDI) (Cabeen et al., 2019), Microstructure Bayesian approach (MB) (Reisert et al., 2017), Multi-Compartment Microscopic Diffusion Imaging (MCMDI) (Kaden et al., 2016),

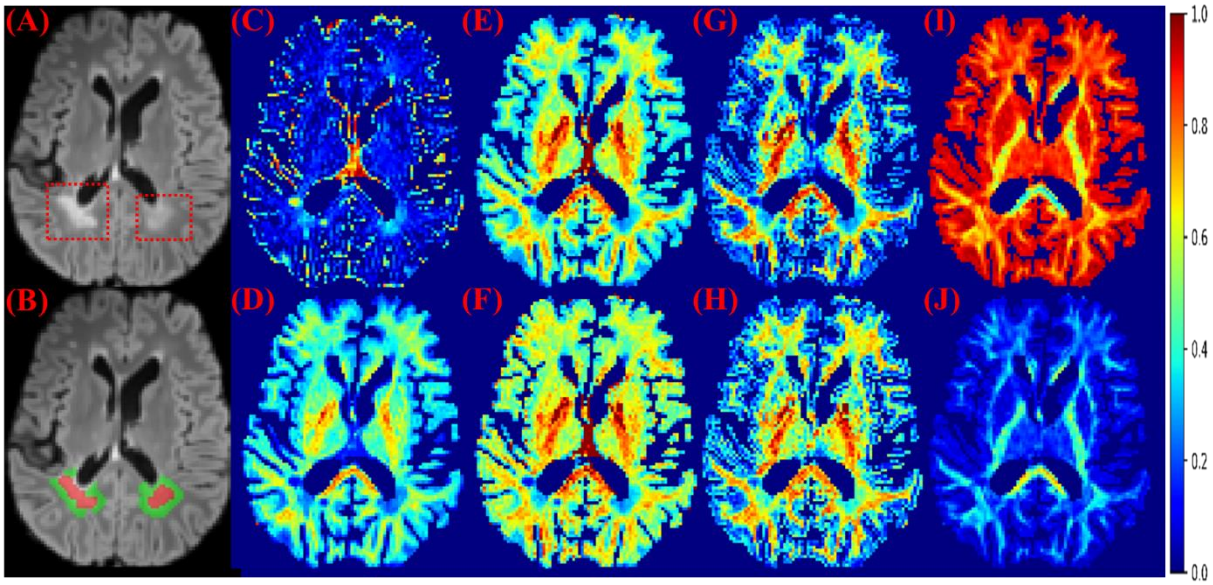


Figure 1: MS lesions on FLAIR and diffusion measures. (A) FLAIR: MS lesions are hyperintense and indicated by red dashed boxes. (B) Red: lesions; Green: perilesional white matter tissue. (C) The isotropic compartment from MB. (D) The intra-axonal compartment from MB. (E) The neurite density index from NODDI. (F) The intra-axonal compartment from SMT-NODDI. (G) The intra-axonal compartment from MCMDI (H) The intra-axonal compartment from NODDIDA (I) The isotropic compartment from Ball and Stick (J) The intra-axonal compartment from Ball and Stick. Other models in the analysis see in the Supplementary Figure 1.

Neurite Orientation Dispersion and Density Imaging with Diffusivities Assessment (NODDIDA) (Jelescu et al., 2015), DIstribution of 3-D Anisotropic MicrOstructural eNvironments in Diffusion-compartment imaging (DIAMOND) (Scherrer et al., 2016) and microstructure fingerprinting (Rensonnet et al., 2019). The exemplary diffusion measures and FLAIR are in Fig. 1.

The quantitative diffusion measures of each patient were masked by the brain mask to remove non-brain tissue including the ventricle. The brain mask was the binarized subcortical segmentation obtained from FreeSurfer (Fischl et al., 2001) on MP2RAGE (Fujimoto et al., 2014) and transformed by FSL to align with mDWI. The diffusion measures were then subject-wise normalized. Eighty-four patients were randomly selected to be used in a 5-fold cross-validation. The other 39/123 patients formed a pure test dataset. White matter lesions were automatically segmented using FLAIR and MP2RAGE (La Rosa et al., 2020) and manually corrected by two expert raters. The lesion segmentations were transformed by FSL to be aligned with mDWI. Lesions of size less than three voxels were excluded. The perilesional tissue was defined as white matter tissue locating within a three-voxel region around the lesions. Patches of 5x5x5 voxels were sampled on lesions and perilesional tissue considering the lesion sizes. To reduce the overlapping between lesion and perilesional patches due to their proximity, a constraint of at most 20% of a sampled patch being overlapped with another patch was applied. The numbers of patches being sampled on each lesion and perilesional tissue were proportional to the size of the lesion and the perilesional tissue, respectively. In the end, 3,007 lesion patches and 3,624 perilesional patches were sampled in the dataset for 5-fold cross-validation, and 1,402 lesion patches and 1,665 perilesional patches were sampled in the pure test dataset. The 5-fold cross-validation was based on the number of patients. Therefore, patches from a patient would not be present both in the training and in the validation datasets.

2.2. GAMER MRI

GAMER MRI was previously developed and validated as a method to obtain attention weights and the relative importance in a classification task of given input contrasts (Lu et al., 2021b). As we previously reported, the neural network consisted of three parts for feature extraction, gated attention mechanism (Ilse et al., 2018), and

classification (Lu et al., 2021b). The feature extraction part included three convolutional blocks for each contrast. Each convolutional block was composed of a layer of 16 convolutional filters and exponential leaky units followed by batch normalization. The kernel size of the convolutional filter was 3x3x3, and padding was applied correspondingly to maintain the patch size. After the last convolutional block, a 16-neuron fully connected layer (FCL) received the flattened vector of 125 elements and encoded the hidden feature of 16 elements. The gated attention mechanism was formed by an attention layer containing an eight-neuron FCL followed by the tanh function and a gate layer having an eight-neuron FCL followed by the sigmoid function. The outputs of tanh and sigmoid were element-wise multiplied. From the element-wise product, in the original implementation for not-highly-correlated input contrasts, the attention weights were obtained by a following one-neuron FCL and the softmax function (Lu et al., 2021b). However, this design was not effective for highly correlated inputs, i.e., diffusion measures in this work. The information content of measures is similar and thus, the difference in the obtained attention weights was small.

For the purpose of this study, we multiplied the outputs from the element-wise multiplication by 2. This enhanced the difference between the encoded features of the correlated diffusion measures during training because the exponential transformation in the softmax function could not properly reflect the difference of the small values and negative values. For example, 0.01 is 10 times larger than 0.001, but they become 1.01 and 1.001 after the exponential transformation. This leads to 0.502 and 0.498 as attention weights after the softmax function. The enhanced output was then connected to a one-neuron FCL followed by the softmax function to generate the normalized attention weights. The weighted sum of the hidden features and the corresponding attention weights formed

a combined hidden feature for the classifier. The classifier was one sigmoid neuron. The network structure is in Fig. 2.

The weighted sampler was used to account for the class imbalance, and the batch size was 256. The loss function was cross-entropy loss. The evaluation metric was the area under the receiver operating characteristic curve (AUC). The optimizer was AdamW (Loshchilov and Hutter, 2019) with the learning rate=5e-5 and the weight decay=1e-2. To avoid overfitting, data augmentation and a learning-rate scheduler were performed. On-the-fly data augmentation included random flipping in the left-right directions and

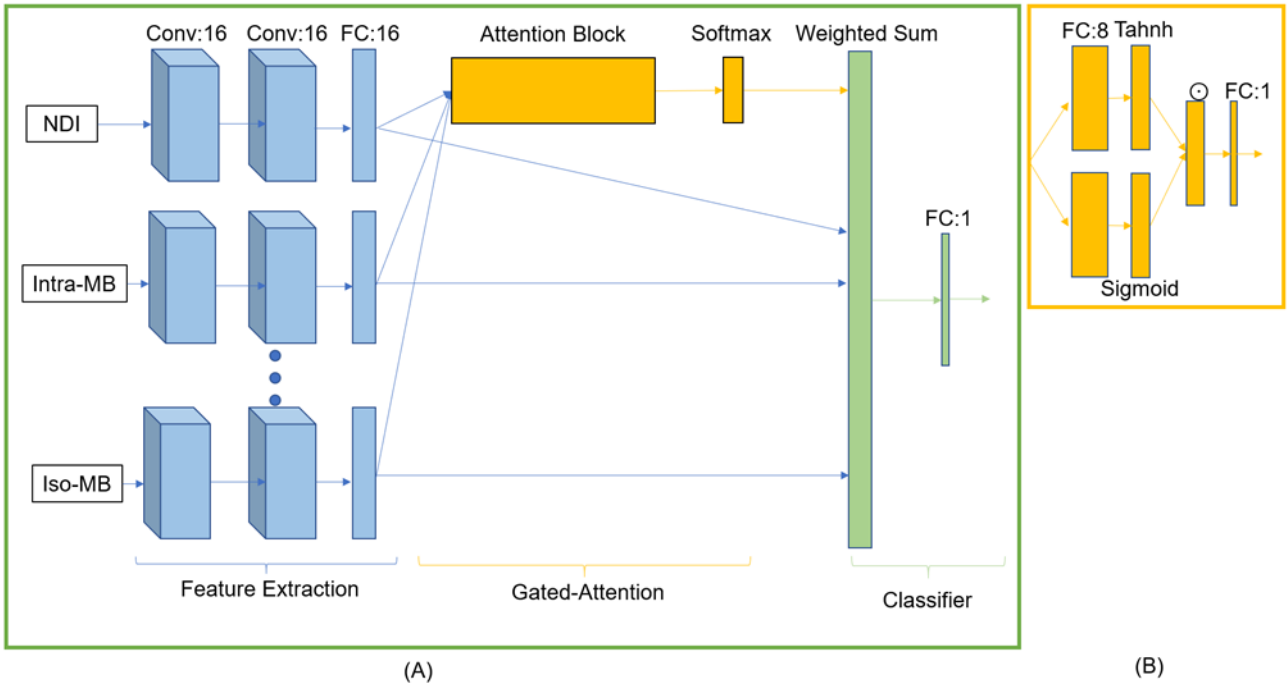


Figure. 2: GAMER-MRI. (A) The neural network. Conv stands for the convolutional block. FC is a fully connected layer. (B) Attention block. ⊙ represents an element-wise multiplication.

Gaussian noise with zero mean and unit standard deviation. The scheduler was the learning-rate-reduce-plateau scheduler with a patience of 15 epochs.

2.3. Selection of contrasts

An intrinsic strong correlation between the quantitative diffusion measures can lead to instability of the obtained attention weights and the ranked order, compared to the result in (Lu et al., 2021b). Therefore, to avoid determination solely based on the attention weights, the selection of discriminating measures was an iteration process. It started from the measure whose attention weight was dominant in the validation datasets in all the

cross-validation folds. If no measure was selected, the measures whose attention weights were ranked 1st or 2nd in all the folds were considered. If no measures were selected, the attention weights ranked 1st or 2nd and 3rd in all the folds were considered. The selection stopped when the sum of their attention weights was over 0.5, which meant that the selected measures were more important than 50% of the input diffusion measures in differentiating the lesion and perilesional tissue.

To assess which selected subject-wise normalized quantitative diffusion measures, or combination of those measures was best correlated with patients' EDSS as well as NfL in the pure test dataset, we first averaged the diffusion measures within each lesion and then over lesions within each patient. In 31/39 patients of the test dataset, we quantified sNfL. Then, we performed Spearman's correlation coefficient with two-sided 20,000 permutation tests. The Benjamin-Hochberg procedure(Benjamini and Hochberg, 1995) was performed to control the false discovery rate (FDR) with the threshold of 0.05. The flowchart is shown in Fig. 3.

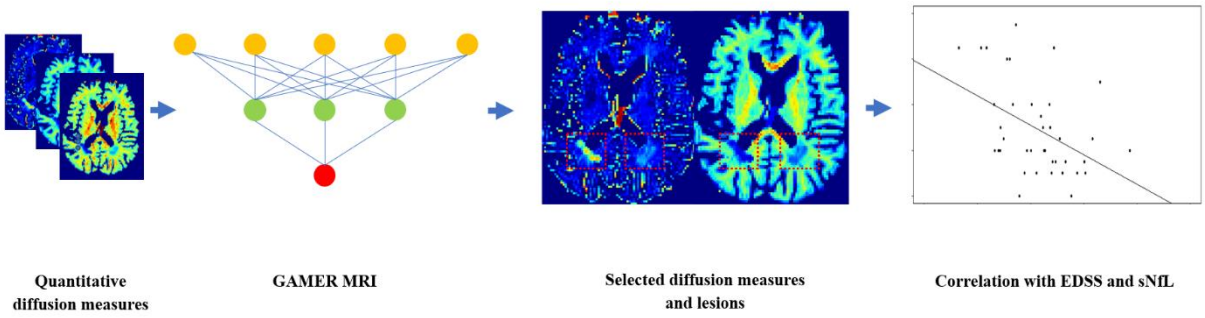


Figure. 3: Flowchart for using GAMER-MRI to select the most discriminating subject-wise normalized diffusion measures and correlating the combinations of selected diffusion measures with the Expanded Disability Status Scale and the serum level of neurofilament light chain.

3. Results

3.1 Lesion classification

In Table 2, we report the average performance of GAMER MRI using all the diffusion measures on the (i) validation dataset over five-fold cross-validation and (ii) on the pure test dataset.

The diffusion measures selected by using the validation datasets were the neurite density index (NDI) from NODDI, the intra-axonal and isotropic compartment from MB

(Intra-MB and Iso-MB), and the intra-axonal compartment from SMT-NODDI (Intra-SMT)

Mean metrics (%)	AUC	Balanced Accuracy	Sensitivity	Specificity	F1 score
Validation dataset	90.67±0.009	83.26 ±1.35	81.09 ±2.44	85.44 ±2.03	81.62±1.67
Test dataset	91.01±0.003	83.42 ±0.12	83.39 ±0.67	83.45 ±0.82	82.14±0.11
Selected Measures	NDI	Intra-MB	Iso-MB	Intra-SMT	
Attention Weights	0.121±0.014	0.117±0.014	0.145±0.007	0.131±0.015	

Table 2: Performance of the patch-based network on MS lesions and the selected diffusion measures on 5-fold cross-validation (first row, average mean and standard deviation are reported) and pure testing set (second row). Balanced Accuracy is defined as the average of sensitivity and specificity in each fold. F1 score is defined as the harmonic mean of precision and recall.

Lesion Load	ρ	P-value	Significance
Number of Lesions	0.13	0.41	--
Lesion Volume	0.25	0.12	--
Normalized Diffusion Measures			
NDI	-0.38	0.017	*
Intra-SMT	-0.31	0.057	--
Intra-MB	-0.40	0.013	*
Iso-MB	0.09	0.58	--
Intra-MB + Iso-MB	-0.39	0.014	*
Intra-MB + NDI	-0.43	0.007	*
Intra-SMT+NDI	-0.37	0.023	*
Intra-SMT + Intra-MB	-0.40	0.012	*
Intra-MB + Iso-MB + NDI	-0.45	0.004	*
Intra-MB + Iso-MB + Intra-SMT	-0.42	0.007	*
Intra-MB + Intra-SMT + NDI	-0.42	0.009	*
Intra-MB + Iso-MB + NDI + Intra-SMT	-0.41	0.009	*

Table 3: Spearman's correlation of selected normalized diffusion measures, or their combinations and EDSS. The significance is controlled by FDR with a threshold of 0.05. Only the combinations of significance are reported.

in Fig. 1. Their average attention weights of the corrected predicted samples are also reported in Table 2.

3.2 Spearman's correlation

3.2.1 Correlation with EDSS

The Spearman's correlation coefficients (ρ) and the corresponding original p-values of the selected normalized diffusion measures, or their statistically significant combinations and EDSS are reported in Table 3. The Spearman's correlation coefficients (ρ) of the conventional lesion load metrics are also reported. The number of potential combinations of four selected diffusion measures is 15, and there are two tests in the lesion load analysis. This led to in total 17 statistical tests. The significance controlled by FDR is indicated by *. The scatter plot of the combination having the strongest correlation is in Fig. 4A and an exemplary image of the combination is in Fig. 4B.

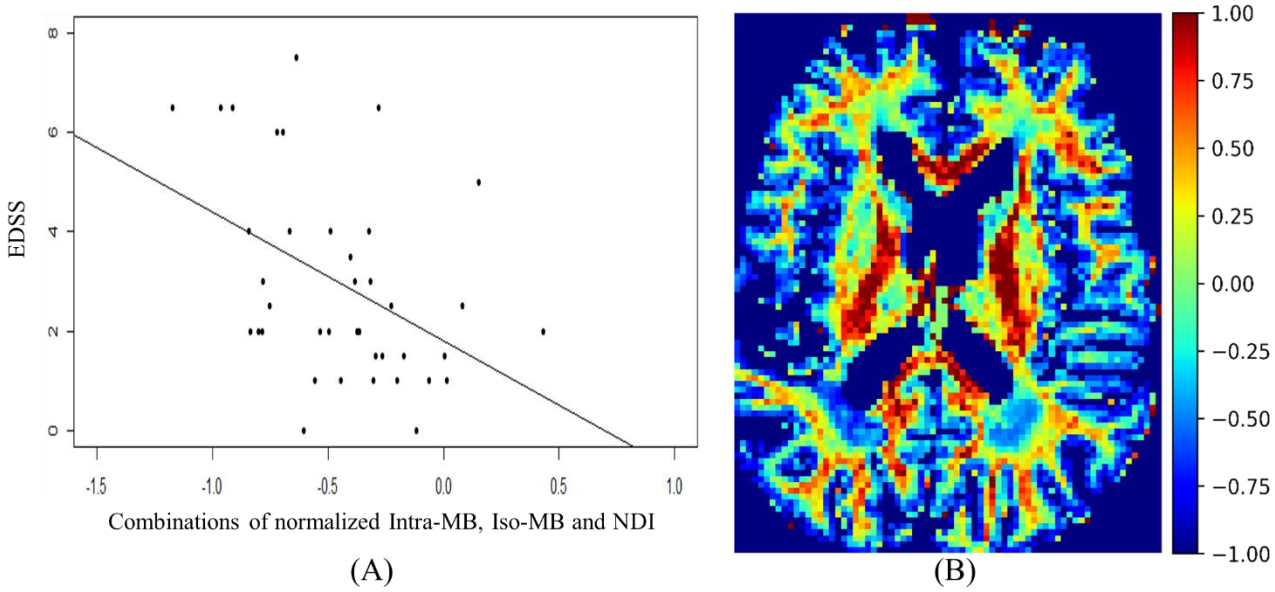


Figure 4: (A): Scatter plot and a regression line of EDSS and the combinations of normalized Intra-MB, Iso-MB and NDI, which has strongest correlation. (B): An exemplary image of the combined contrast.

3.2.2 Correlation with sNfL

The Spearman's correlation coefficients (ρ) and the corresponding original p-values are reported in Table 4. One patient had a relatively high sNfL level of 160 μg/mL, compared to the mean sNfL level of 8.9 μg/mL of the rest of the 30 patients. After this patient's data were excluded, the significance in Table 4 did not change, but the correlation was stronger. For illustration purpose, the scatter plot of the combination

having the strongest correlation (Fig. 5A) does not contain this outlier patient. An exemplary image of the combination is in Fig. 5B.

Lesion Load	ρ	P-value	Significance
Number of Lesions	0.48	0.006	*
Lesion Volume	0.45	0.01	*
Normalized Diffusion Measures			
NDI	-0.37	0.04	--
Intra-SMT	-0.27	0.14	--
Intra-MB	-0.42	0.02	*
Iso-MB	0.1	0.59	--
Intra-MB + Iso-MB	-0.51	0.004	*
Intra-MB + NDI	-0.43	0.02	*
Intra-MB + Iso-MB + NDI	-0.48	0.007	*
Intra-MB + Iso-MB + Intra-SMT	-0.45	0.01	*
Intra-MB + Iso-MB + NDI + Intra-SMT	-0.44	0.02	*

Table 4: Spearman's correlation of selected normalized diffusion measures, or their combinations and sNfL. The significance is controlled by FDR with a threshold of 0.05. Only the combinations of significance are reported.

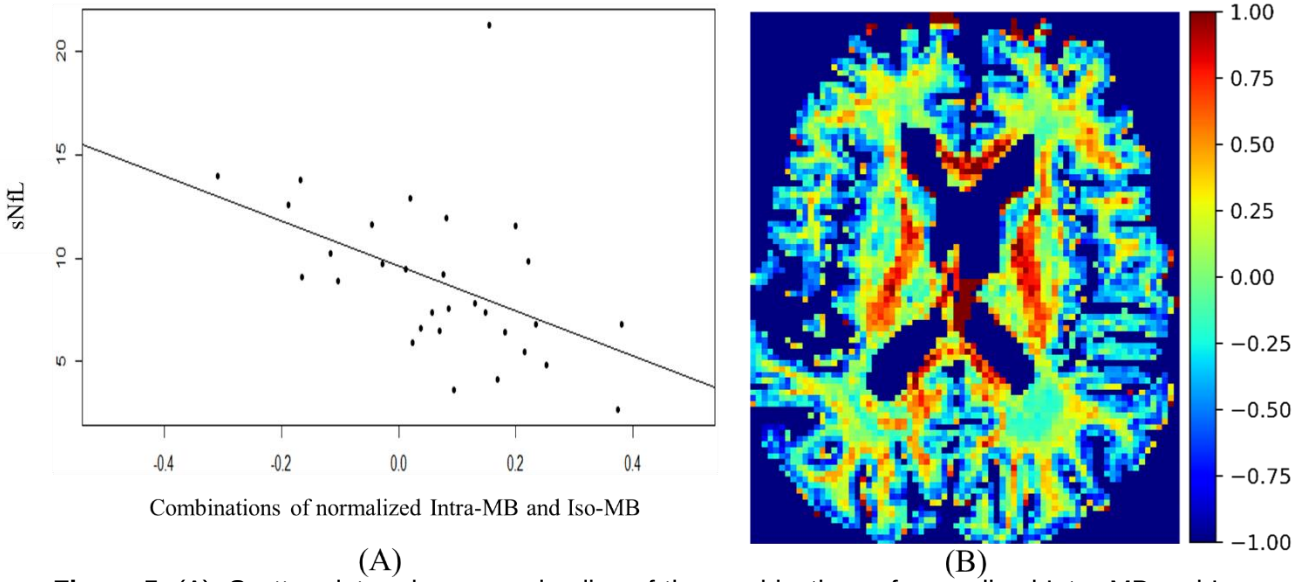


Figure 5: (A): Scatter plot and a regression line of the combinations of normalized Intra-MB and Iso-MB, which showed the strongest correlation with sNfL. (B): An exemplary image of the combined contrast.

4 Discussion

Our work provided evidence that a modified version of GAMER MRI, including a specific selection procedure for correlated measures, permits to identify the microstructural diffusion measures that are most discriminative of focal MS pathology among the ones obtained with eight open-source mathematical models of multi-shell diffusion data. Moreover, our data showed that some of the combinations of the selected normalized diffusion measures better correlated with patients' disability and neuroaxonal damage than the individual measures.

Diffusion-based microstructural measures quantify different compartments based on various assumptions. Nevertheless, the relative sensitivity of the different diffusion-based microstructural metrics to specific CNS pathologies is unclear. In this work, we have provided a methodological frame to discriminate the most sensitive diffusion microstructural measures to focal MS pathology in a large population of MS patients.

We first aimed at identifying which measure best discriminated MS lesions from the perilesional tissue because we judged that if the neural network was able to differentiate between lesions and the immediate surrounding tissue, the learned pattern would have been most sensitive to focal MS pathology than the one we would have derived by comparing lesions to distant normal-appearing tissue. The evaluation metrics in Table 2 indicated that the neural network was able to learn pivotal information for the target classification. As expected, because of the highly correlated nature of the studied diffusion-based measures, the difference among the obtained attention weights was small. The proposed selection process alleviated the fluctuating order of attention weights due to their small differences. The threshold of 0.5 in the selection process was empirically chosen considering the representativeness of selected diffusion measures and the multiple comparison problem.

The core idea of the attention mechanism is to enhance important features from the data themselves relevant to the specific application (Bahdanau et al., 2015). Therefore, in most of the applications in natural language processing and natural image classification, the attention weights were used to enhance the connection to the corresponding features based on their importance instead of quantifying the relative importance among the features (Hu et al., 2018; Maicas et al., 2017; Vaswani et al., 2017; Woo et al., 2018).

Using different designs of the attention mechanism, the attention weights provide also the relative importance among features as shown in a histopathological image classification and image captioning (Ilse et al., 2018; You et al., 2016). In GAMER MRI, attention weights were computed and validated on multi-contrast MRI measures in order to select their relative importance in a given neurological disease classification.

To our knowledge so far, only a few studies applied measures derived from microstructural models to study focal MS pathology (for a review see (Granziera et al., 2021)) and only one study used deep-learning to show the superior performance of diffusion basis spectrum imaging to segment voxel-wise different types of MS lesions compared to using diffusion tensor imaging (Ye et al., 2020). However, the joint comparison of multiple microstructural diffusion measures in MS lesions has not been explored yet. This work considered the potential interaction between the measures and tried to address this issue.

The four selected diffusion measures include three measures for the intra-axonal compartment from three models and one measure for the isotropic compartment from one of the three models. This means that most of the discriminating information of the damaged neurons was from the loss of axonal integrity. The additional information about the inflammatory processes might be reflected by the measure for the isotropic compartment to better characterize the distinction of lesions.

Besides, by combining the selected diffusion measures in the discrimination of focal pathology, it was possible to achieve a stronger correlation with patient disability than the one of those metrics alone or even conventional MRI metrics, such as the lesion number and volume. These results suggest that a comprehensive description of the tissue microstructure in regions of focal damage in MS patients may well help decrease the clinic-radiological paradox (Barkhof, 2002). Interestingly, the combined contrast achieving the best correlation with disability was the sum of measures quantifying intra-axonal and isotropic diffusion, which may be considered surrogate measures of the loss of integrity of axons and myelin as well as of inflammatory processes (i.e., increased cellularity and edema).

Most of the combinations that best correlated to EDSS were also highly related to the sNfL levels: remarkably, the correlation coefficients between sNfL and combinations

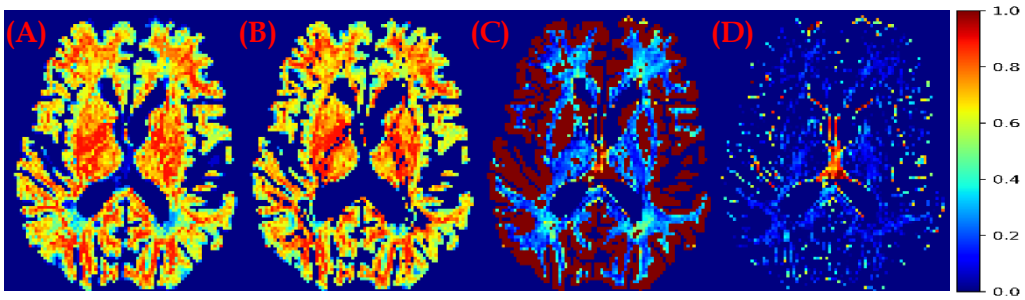
of diffusion-MRI metrics were even higher than the ones obtained between sNfL and the lesion load, which is known to be highly related to sNfL levels (Chitnis et al., 2018; Todea et al., 2020). The patient, who had an extremely high level of sNfL, had a relapse 2 months before the sNfL acquisition, which may have well influenced the strong increase in sNfL levels.

To perform the correlation analyses with EDSS and sNfL, we have used subject-wise normalized maps of diffusion-based microstructural measures, which were the ones encoded by GAMER MRI. We also trained the neural network on the original images, which - however - led to worse classification performance. Because subject-wise normalized maps were used, it is challenging to determine whether the network could learn the right pattern and to generate representative attention weights. Owing to the applied normalization procedure, the interpretation of the pathological meaning of the combined metrics is particularly difficult. Another limitation of this study was that we divided the cross-validation folds based on the number of patients instead of the number of patches: this led to different distributions of lesion and perilesional patches in the validation datasets of all cross-validation folds and to the fluctuation of the validation results. On the other hand, this also had the advantage of preventing the leak of information induced by the appearance of patches from one patient in both the training and validation dataset. Based on the obtained result (Table 2), the performance on the test dataset was stable, so the limitation was alleviated.

5 Conclusion

In summary, our work showed that the proposed attention-based neural network and the selection process based on the previous work can select important diffusion measures despite they are highly inter-correlated. Those measures have the potential to be combined to enhance the correlation with the clinical measures. Future work will be required to directly find the best combinations without using a statistical test and tackling the multiple comparison problem. Furthermore, the use of a combination of diffusion-based microstructural measures deserves further attention and development allowing better interpretability of its pathological meaning.

Supplementary Materials:



Supp. Figure 1: Other diffusion measures in the analysis. (A) The first intra-axonal compartment from microstructure fingerprinting (MF-f0). (B) The second intra-axonal compartment from microstructure fingerprinting (MF-f1). (C) The intra-axonal compartment from DIstribution of 3-D Anisotropic MicrOstructural eNvironments in Diffusion-compartment imaging (DIAMOND). (D) The isotropic compartment from Neurite Orientation and Dispersion Density Imaging (ISOVF)

Diffusion Measures	**NDI**	**Intra-MB**	**Iso-MB**	**Intra-SMT**	**ISOVF**	**Intra-MCMDI**
Attention Weights	0.121 ±0.014	0.117 ±0.014	0.145 ±0.007	0.131 ±0.015	0.038 ±0.014	0.098 ±0.010
Diffusion Measures	**Intra-Ball and Stick**	**Iso-Ball and Stick**	**Intra-DIAMOND**	**Intra-NODDIDA**	**MF-f0**	**MF-f1**
Attention Weights	0.049 ±0.012	0.055 ±0.005	0.03 0±0.012	0.049 ±0.007	0.090 ±0.001	0.077 ±0.004

Supp. Table 1: The attention weights of the diffusion measures in the validation dataset on 5-fold cross-validation (average mean and standard deviation are reported). The prefix, “intra-“, stands for the intra-axonal compartment of the following model. The prefix. “iso”, represents the isotropic compartment of the following model.

Footnote

1. https://github.com/AthenaEPI/dmipy
2. https://github.com/daducci/AMICO
3. https://bitbucket.org/reisert/baydiff/wiki/Home
4. https://github.com/robbert-harms/MDT
5. https://bitbucket.org/benoitscherrer/crldciestimate
6. The author needs to be contacted.
7. https://hub.docker.com/r/francescolr/ms_seg

Conflict of Interest

No conflicts of interest in relation to this work.

Author Contributions

Po-Jui Lu: Conceptualization, Data Curation, Methodology, Investigation, Formal analysis, Writing – original draft. **Muhamed Barakovic**: Data Curation, Methodology, Writing– reviewing& editing. **Matthias Weigel**: Resources, Data Curation, Writing – reviewing& editing. **Reza Rahmanzadeh**: Data Curation, Writing – reviewing& editing. **Riccardo Galbusera**: Data Curation, Writing – reviewing& editing. **Simona Schiavi**: Resources, Writing – reviewing& editing. **Alessandro Daducci**: Resources, Writing – reviewing& editing. **Francesco La Rosa**: Data Curation, Writing – reviewing& editing. **Meritxell Bach Cuadra**: Resources, Writing – reviewing& editing. **Robin Sandkühler**: Conceptualization, Writing – reviewing& editing. **Jens Kuhle**: Writing – reviewing& editing. **Ludwig Kappos**: Writing – reviewing& editing. **Philippe Cattin**: Supervision, Writing – reviewing& editing. **Cristina Granziera**: Supervision, Conceptualization, Funding Acquisition, Resources, Writing – reviewing& editing.

Funding

This project is supported by Swiss National Funds PZ00P3_154508, PZ00P3_131914, and PP00P3_176984. Francesco La Rosa and Meritxell Bach Cuadra are supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie project TRABIT (agreement No 765148) and by the Centre d'Imagerie BioMedicale (CIBM).

Acknowledgments

We would like to acknowledge all the patients and healthy controls in this project.

Chapter 4. GAMER-MRI identifies patterns of brain changes associated with disability in multiple sclerosis patients

In this work, we further extended GAMER-MRI from the previous lesion-level analysis to the patient-level analysis and presented an improvement of the Layer-wise Relevance Propagation (LRP) in combination with the attention weights from GAMER-MRI to address the issue of the unknown decision process of deep learning on MR images. Deep learning has a hierarchical layer structure containing nonlinear activation function to understand data on different scales at different layers. Nevertheless, the deep layer structure and non-linear activations cast a limitation on the explainability of the decision process. This becomes a major challenge for medical applications where reasoning of clinical decision is essential. LRP includes various rules to redistribute relevance to the input image and has been shown effective in increasing the explainability. Here, we proposed a new LRP rule for GAMER-MRI and used the attention weight as the relevance to redistribute. The results showed that the proposed improvement could reveal more regions deemed important by GAMER-MRI than the original LRP, and the quantitative MRI within the regions correlated with patients' EDSS.

GAMER-MRI identifies patterns of brain changes associated with disability in multiple sclerosis patients

Po-Jui Lu[1,2,3], Benjamin Odry[4], Muhamed Barakovic[1,2,3], Matthias Weigel[1,2,3,5], Robin Sandkühler[6], Reza Rahmanzadeh[1,2,3], Xinjie Chen[1,2,3], Mario Ocampo-Pineda[1,2,3], Jens Kuhle[2], Ludwig Kappos[2], Philippe Cattin[6], Cristina Granziera[1,2]

1. *Translational Imaging in Neurology (ThINk) Basel, Department of Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland*
2. *Department of Neurology, University Hospital Basel, Basel, Switzerland*
3. *Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Basel, Switzerland*
4. *AI for Clinical Analytics, Covera Health, New York, USA*
5. *Division of Radiological Physics, Department of Radiology, University Hospital Basel, Basel, Switzerland*
6. *Center for medical Image Analysis & Navigation, Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland.*

Abstract

Introduction: Deep neural networks are powerful but suffer from a black-box nature, as their exact decision process is elusive. Layer-wise relevance propagation (LRP) is a technique for increasing the explainability of the decision of deep neural networks. In this work, we have developed and validated a new LRP rule for GAMER-MRI – a neural network exploiting the gated attention mechanism – to explore brain regions relevant to the classification of multiple sclerosis patients having a severe or mild disability.

Methods: GAMER-MRI is based on the gated attention mechanism and was shown to provide attention weights as proxies of the importance of the quantitative magnetic resonance images (qMR images) in a given classification task. In this work, the encoder part of GAMER-MRI was based on DenseNet to tackle volumetric qMR images, including quantitative T1 relaxometry (qT1), neurite density index (NDI), and myelin water fraction (MWF). The proposed LRP considered the joint contribution nature of the gate and signal layers within the gated attention mechanism. We hypothesized that the attention weights could be more appropriate to relate to disability measures in patients with multiple sclerosis (MS) in LRP than the output score from the classifier. We also hypothesized that the relevance maps of the qMR images could be linearly combined with the qMR images to form a more informative combined map. The important regions revealed by the LRP were assessed by the impact on the area under the receiver operating characteristic curve (AUC) if voxel values within the regions were inverted. The Spearman's correlation with the Expanded Disability Status Scale (EDSS) – a measure of disability in MS patients - accompanied by the two-sided permutation test was also conducted for the assessment.

Results: The model performance of GAMER-MRI was AUC=0.864 on the validation folds on average in the three-fold cross-validation and 0.885 on the independent test dataset. qT1 was the most sensitive, followed by NDI and MWF. In the voxel inverting experiment, the most affected AUC was achieved by using the combined map based on the proposed rule and the attention weights. The importance mask, which was defined by the top 40th quantiles of the combined map, covered regions, where qT1 and NDI moderately correlated with EDSS (ρ = -0.371 and 0.440, respectively).

Conclusions: This work demonstrates that the proposed configurations for GAMER-MRI and the modified LRP rule based on attention weights might be useful to assess the relative importance of MRI measures for MS patients with a severe or mild movement disability and reveal potentially patient-specific regions to patients' impairment.

Keywords

Deep learning, Attention Mechanism, Multiple Sclerosis, quantitative MRI, Layer-wise Relevance Propagation

1. Introduction

Deep neural networks (DNNs) have achieved remarkable success in various applications using Magnetic Resonance Imaging (MRI), including brain lesion detection (Barquero et al., 2020; Koschmieder et al., 2022; Maggi et al., 2020), classification of brain disease (Nael et al., 2021; Wen et al., 2020; Yoo et al., 2018), segmentation of brain lesions (Andermatt et al., 2018; La Rosa et al., 2020; Zeineldin et al., 2020), and prediction of disease prognosis (Saha et al., 2020; Sivaranjini and Sujatha, 2020). DNNs have a hierarchical layer structure containing nonlinear activation functions. Thus, they have the capacity to understand the data on different scales at different layers (LeCun et al., 2015). This characteristic allows DNNs to gradually learn and condense raw or minimally pre-processed data into meaningful and lower-dimensional features, during the iterative training process for a given task. In addition, it may shed light on unexpected patterns in data for further exploration. However, the deep layer structure and nonlinear activations represent a limitation on the explainability of the decision process of DNNs (Samek et al., 2021). This becomes a major challenge for medical applications where the reasoning of clinical decisions is important. To tackle this issue, several methods have been developed to improve the explainability of the decision of DNNs, including saliency (Simonyan et al., 2013), integrated gradients (Sundararajan et al., 2017), and layer-wise relevance propagation (LRP) (Bach et al., 2015). LRP has been shown to be more effective than other methods (Samek et al., 2021, 2017). It leverages the DNN layer structure to generate a map, which is based on the output score from the classifier before the sigmoid/softmax function and designed rules, showing the relevance of individual

pixels to the given task (Arras et al., 2019; Bach et al., 2015; Montavon et al., 2019). It has been shown that the relevance maps were able to illustrate the disease-specific evidence (Böhle et al., 2019).

In our previous work, we developed a gated attention-based convolutional neural network (CNN) – GAMER-MRI – using qualitative MRI and quantitative MRI (qMRI) for patch-based classification of acute infarct stroke lesions and of perilesional and juxtacortical multiple sclerosis (MS) lesions (Lu et al., 2021b). The attention weight (AW) in GAMER-MRI was demonstrated to be a proxy quantification of the clinical importance of the individual MR contrasts. In (Lu et al., 2021a), GAMER-MRI was used for the classification of lesions and peri-lesion tissue using various diffusion measures of biophysical microstructure models derived from diffusion MRI. The AWs were able to select a subset of diffusion measures, which could be summed to enhance the statistically significant correlation with biological and clinical measures. Based on these findings, we hypothesized that an LRP relevance map based on AWs would be more relevant than a map based on the output score from the classifier, and the LRP relevance map could be linearly combined with qMR images to form an especially relevant relevance map. In addition, we proposed a new LRP rule to account for the gated attention mechanism in GAMER-MRI. To tackle the high number of voxels in 3D qMR images, the encoder in GAMER-MRI was based on the DenseNet architecture (Huang et al., 2017) because DenseNet encourages the reuse of features at different layers and greatly reduces the number of learnable parameters. DenseNet has been used in various neuroimaging studies (Ghesu et al., 2021; Ruiz et al., 2020; Zeineldin et al., 2020). We assessed our hypotheses and the effect of the proposed rule on an MS 3D qMR dataset aiming to decode the pattern of the movement disability of MS on qMR images. Some works deployed CNN and LRP in the classification of MS patients/healthy controls using volumetric conventional MRI and slices of susceptibility weighted images (Eitel et al., 2019; Lopatina et al., 2020) and classification of relapsing and remitting MS/progressive MS (RRMS/PMS) (Cruciani et al., 2021). Each of their studies used a single rule for different kinds of layers in the NNs and demonstrated the applicability of LRP on MS MRI datasets. In addition, they obtained the relevance maps based on the classification

probability instead of the output score before the sigmoid/softmax function as shown in (Bach et al., 2015; Montavon et al., 2019).

MS is a chronic inflammatory disease of the central nervous system, which affects ~ 2.3 million people worldwide (Walton et al., 2020). The representative characteristics include multifocal inflammatory infiltration, demyelination, remyelination, and axonal loss leading to the accumulation of disability in MS patients (Albert et al., 2007; Barkhof et al., 2003; Dziedzic et al., 2010; Kuhlmann et al., 2002). These microstructural properties of brain tissue can be better measured by qMRI (e.g. myelin – myelin water fraction (MWF) and quantitative T1 relaxometry (qT1); axon – qT1 and neurite density index (NDI)) (Granziera et al., 2021), compared to conventional MRI. Disability in MS patients is clinically assessed by the Expanded Disability Status Scale (EDSS), a nonlinear representation of clinical disability ranging from 0 to 10. Starting from 5, MS patients are considered to have a severe disability.

Therefore, we have two specific aims in this work: (i) the classification of MS patients of severe/moderate-mild disability using GAMER-MRI with volumetric qMR images and (ii) assessment of the relevance maps of qMR based on AWs and/or the proposed LRP rule in terms of changes in performance and correlation with EDSS.

2. Materials and methods

2.1 MRI data

One-hundred-sixty-six MS patients (100 RRMS and 66 PMS, 99 females and 67 males, age range=45.9±14.3 years, median EDSS=2.5) were enrolled in the study, which was approved by the Ethics Committee Northwest and Central Switzerland and the Basel University Hospital. Written consent was obtained prior to the MRI acquisition. Forty out of the 166 patients had EDSS ≥5 (severe disability group) and 126/166 had a mild disability (<5). Sixty-nine patients had a two-year follow-up acquisition. Patients underwent a multi-parametric protocol on a 3T Siemens Healthineers MAGNETOM Prisma MRI system. The volumetric protocol included 3D Magnetization-Prepared 2 RApid Gradient Echoes (MP2RAGE) (Kober et al., 2012; Marques et al., 2010), 3D turbo spin echo (SPACE) based Fluid Attenuated Inversion Recovery (FLAIR), 3D Fast

	TE (ms)	TR (ms)	FOV (mm^3)	SR (mm^3)	TI (ms)	Additional Parameters
MP2RAGE	3	5000	256x256x256	1x1x1	700, 2500	--
FLAIR	386	5000	256x256x256	1x1x1	1800	--
FAST-T2	0.5	7.5	240x240x160	1.25x1.25x5	--	**T2prep times (ms)** 0 (T2prep turned off),7.5,17.5, 67.5,147.5,307.5
Multi-shell DWI	75	4500	256x256x144	1.8x1.8x1.8	--	**b values (s/mm^2)** 0;700;1000; 2000;3000 /149 directions in total

Table 1. Acquisition parameters of each MR contrast in the MS dataset. TE: echo time; TR: repetition time; FOV: field of view; SR: spatial resolution; TI: inversion time.

Acquisition with Spiral Trajectory and T2prep sequence (FAST-T2) (Nguyen et al., 2016), and multi-shell Diffusion-Weighted Imaging (multi-shell DWI) with contiguous 2D slices. The most important acquisition parameters are shown in Table 1.

White matter lesions (WM lesions) were automatically segmented using FLAIR and MP2RAGE (La Rosa et al., 2019) and manually corrected by three expert raters. The WM lesion masks were used for lesion-filling on MP2RAGE images by FMRIB Software Library (FSL) (Jenkinson et al., 2012). FreeSurfer (Fischl et al., 2001) processing was

performed on lesion-filled MP2RAGE images to obtain brain segmentation. From multi-parametric MRIs, qMR images were further reconstructed. The qT1 was reconstructed from MP2RAGE as in (Kober et al., 2012; Marques et al., 2010). The MWF map was reconstructed from FAST-T2 as in (Nguyen et al., 2016). The neurite density index (NDI) from the neurite orientation dispersion and density imaging model (NODDI) (Zhang et al., 2012) was reconstructed from multi-shell DWI as in (Daducci et al., 2015). The qMR images were co-registered using FSL. The brain segmentation was transformed to the coordinates of each qMR image to remove non-brain tissue. qMR values within the brain were then normalized between zero and one. qT1 feasible values fall between 0 and 2500 ms excluding ventricles (Bojorquez et al., 2017; Bonnier et al., 2014). MWF feasible values within the brain are at most 30 % (MacKay and Laule, 2016). Normalized qMR images were then transformed to the coordinates of NDI. NDI was chosen as the reference coordinate for the trade-off between the lower degree of slice interpolation for MWF and the retainment of fine resolution of qT1. Cerebellum details were coarse under the resolution of NDI and it was removed to improve the performance of the subsequent training. The matrix size of each qMR image was (96, 96, 112) after the removal of empty space.

The dataset was divided into an independent test dataset and a dataset for cross-validation. In this case, the dataset for cross-validation was used for the optimization of the hyperparameters and the model selection. The independent test dataset was used to estimate the model performance on unforeseen data. The test dataset included 21/166 patients (16 in the mild disability group and 5 in the severe group) and 20 of them had the follow-up acquisition. The baseline and follow-up acquisitions were considered two samples in the same dataset. In light of the number of the remaining patients in the severe group being only 35, stratified 3-fold cross-validation was used for evaluating our method. As a result, the dataset for cross-validation was divided into three folds. In each turn of cross-validation, two folds formed the training dataset and the remaining fold was the validation dataset

2.2 GAMER-MRI

The core idea of GAMER-MRI was to use the gated attention mechanism (Ilse et al., 2018) and a parallel encoding structure to generate attention weights as proxies of relative importance among input MR images. The variant gated attention mechanism (Lu et al., 2021b) is formulated as follows:

$$\boldsymbol{n} = \sum_{l=1}^{L} a_l \boldsymbol{m}_l = \sum_{l=1}^{L} a_l q_l(x_l) = \sum_{x \in X} g(x), \quad (1)$$

$$a_l = softmax\left(\boldsymbol{w}^T\left(\tanh(\boldsymbol{U}\boldsymbol{m}_l) \odot \mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)\right)\right), \quad (2)$$

where $\boldsymbol{n}$ is the combined representation for the classifier, $\boldsymbol{m}_l$ is the hidden representation of the l^{th} instance, $q_l(x)$ is the encoding function and a_l is the AW of the l^{th} instance, $\boldsymbol{U}$ and $\boldsymbol{V} \in R^{K \times M}$ are weights of the fully connected layers (FCs) following the hidden representations, $tanh$ stands for the nonlinear hyperbolic tangent function, sigm stands for the nonlinear sigmoid function, $\odot$ is the element-wise multiplication operator, $\boldsymbol{w} \in R^{1 \times K}$ contains the weights of an FC, $softmax$ stands for the softmax function.

2.2.1 Architecture

The multi-path neural network included three main compartments, including the parallel feature extracting compartment, GAMER-MRI, and a classifier, as depicted in Fig. 1. The feature extracting compartment was based on the DenseNet architecture (Huang et al., 2017). DenseNet concatenated all feature maps generated within the same dense block to reuse feature maps and facilitate feature propagation. This reduced the number of trainable parameters substantially and alleviated the vanishing gradient issue. The feature extraction consisted of a convolutional block, four dense blocks, and three transition layers. The convolutional block was composed of an initial convolutional layer (Conv) of 16 filters with a kernel size of 3x3x3, a batch normalization layer (BN), rectified linear units (ReLU) and a 3D max pooling layer with a kernel size of 3, a stride of 2 and a padding of 1. Each dense block contained two dense layers, each of which was a collection of BN, ReLU, and Conv with a kernel size of 1x1x1 and replicate padding, and BN, ReLU, and Conv with a kernel size of 3x3x3. The transition layer was BN, ReLU, Conv (1x1x1), and average pooling with a kernel size of 2x2x2 and a stride of 2. The growth rate of the number of feature maps was four. The fourth dense block was followed by a BN, ReLU, and a fully connected layer (FC) of 32 neurons to form a hidden feature

vector for the input image. The main components of GAMER-MRI were the signal layer and the gate layer. The FCs in the signal layer and the gate layer had 16 neurons. The combined hidden feature vector formed by the AWs and the hidden feature vectors from all paths was concatenated by patients' age information. Patients' ages were divided by 100 prior to concatenation. The incorporation of age information accounted for the age effect (Zeydan and Kantarci, 2020). The classifier was an FC of a neuron outputting $f(\boldsymbol{x})$ that was subsequently transformed to classification probability by sigm.

2.2.2 Training Strategy

As the loss function, the binary cross-entropy was used. In consideration of the heterogeneity within each group and across groups measured by EDSS, the training loss of each sample was weighted by the closeness between EDSS=5 and patients' EDSS ($weight = 2 - \frac{|EDSS-5|}{5}$). The mini-batch size was 70 for training. The weighted sampler was used to account for the class imbalance during training, and the optimizer was the Adam optimizer with decoupled weight decay (AdamW) with the learning rate=5e-5 and the default weight decay=1e-2 (Loshchilov and Hutter, 2019). The evaluation metric was the area under the receiver operating characteristic curve (AUC). To alleviate overfitting, data augmentation included random flipping, random 90-degree rotation, random

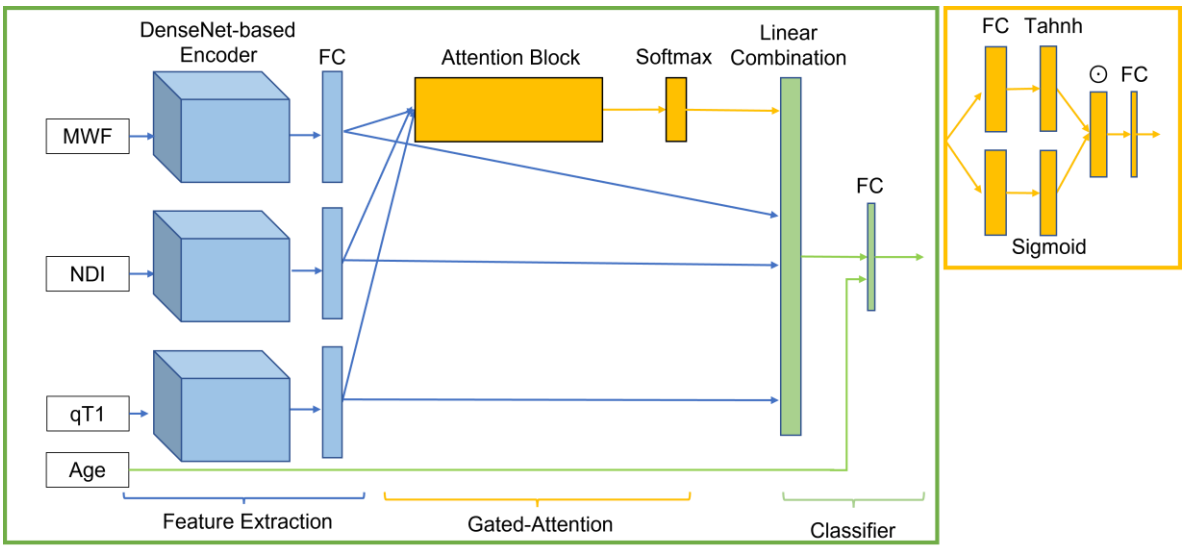


Figure 1: The network structure consists of the feature extraction, the GAMER-MRI and the classifier. FC is the fully connected layer. The linear combination forms the combined feature vector. ⊙ represents an element-wise multiplication.

Gaussian noise of zero mean and standard deviation equal to 0.2, and random affine transformation with maximum rotation ±30 degrees and ±10% scaling. We implemented our method in PyTorch v1.7 and used two A100 GPUs (Nvidia, Santa Clara, CA, USA) for training.

2.2.3 Layer-wise relevance propagation

Layer-wise relevance propagation (LRP) is a post-hoc explaining technique for neural networks and is based on Deep Taylor Decomposition (DTD) (Montavon et al., 2017). The main rationale is to redistribute the prediction backward through the layer structure of the NN down to the input data based on the percentage of contribution from the individual neuron in the forward pass. It is formulated as follows:

$$\boldsymbol{R}_x = D_j\left(D_{j+1}\left(\dots D_c\big(f(x)\big)\right)\right), \tag{3}$$

$$\sum_{v \in x} R_v = f(x), \tag{4}$$

where $\boldsymbol{R}_x$ is the relevance map of the input image x, R_v is the relevance of the voxel v, D_c is the redistribution rule function for the classifier and D_j is the function for the jth intermediate layer. For (4) to be valid, the bias parameters in the NN are considered in the forward and backward passes during training and prediction, but aren't considered in the LRP backward passes to get the relevance map (Samek et al., 2021). For the rest of the paper, all the rules do not consider the bias parameters.

The redistribution rules depend on the kinds of layers and are thus versatile. The common rules included the ε-rule, the αβ-rule, the z^β-rule and others (Arras et al., 2019; Montavon et al., 2019). In (Montavon et al., 2019), they recommended different rules for layers in different positions in the hierarchical structure and the kinds of layers, so we applied an orderly mixture of rules. The ε-rule was used in FCs for the relevance R_s of the neuron s in the jth layer:

$$R_s = D_j\big(\boldsymbol{R}_{j+1}\big) = \sum_{t \in (j+1)th\ layer} \frac{u_s \omega_{s\to t}}{\varepsilon + \sum_{s \in jth\ layer} u_s \omega_{s\to t}} R_t, \tag{5}$$

where u_s is the output from the neuron s in the jth layer, $\omega_{s\to t}$ is the weight connecting s and the neuron t in the $(j+1)$th layer, R_t is the relevance of t and ε is a small value to avoid zero division and increasing ε can also function as absorbing small relevance to

keep more salient features on the relevance map. In this work, ε was 1e-8. The αβ-rule was used for Conv and is formulated as follows:

$$R_s = \sum_{t \in (j+1)th\ layer} (\alpha \frac{(u_s \omega_{s \to t})^+}{\sum_{s \in jth\ layer} (u_s \omega_{s \to t})^+} - ß \frac{(u_s \omega_{s \to t})^-}{\sum_{s \in jth\ layer} (u_s \omega_{s \to t})^-}) R_t, \quad (6)$$

where $(u_s \omega_{s \to t})^+$ takes only positive values and $(u_s \omega_{s \to t})^-$ takes negative values, $\alpha - ß = 1$. In this study, $(\alpha, ß)$ were $(1, 0)$ because it can be interpreted by DTD. The z^β-rule was for the layer taking the input MR image and is given as:

$$R_v = \sum_{t \in (j+1)th\ layer} \frac{v\omega_{v \to t} - l{\omega_{v \to t}}^+ - h{\omega_{v \to t}}^-}{\sum_{v \in x} v\omega_{v \to t} - l{\omega_{v \to t}}^+ - h{\omega_{v \to t}}^-} R_t, \quad (7)$$

where v is the voxel value and l and h are the lowest and highest possible values on the input image, respectively.

To account for the element-wise multiplication $\tanh(\boldsymbol{U}\boldsymbol{m}_l) \odot \mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)$ in the gated attention mechanism, we proposed a new rule. In (Arras et al., 2019), they proposed the LRP-all rule for the element-wise multiplication of the cell input and the input gate within the long short-term memory network (LSTM) and showed that it was more suitable than the LRP-prop (Ding et al., 2017) and LRP-abs rules. The LRP-all only lets the cell input take all the relevance. However, for the gated attention mechanism, this rule neglects the nature of the structure that both $\tanh(\boldsymbol{U}\boldsymbol{m}_l)$ and $\mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)$ contribute towards the attention weights. The proposed modification is formulated as follows:

$$R_s = \frac{|\mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)|}{|\mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)| + |\tanh(\boldsymbol{U}\boldsymbol{m}_l)|} R_a, \quad (8)$$

$$R_g = \frac{|\tanh(\boldsymbol{U}\boldsymbol{m}_l)|}{|\mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)| + |\tanh(\boldsymbol{U}\boldsymbol{m}_l)|} R_a, \quad (9)$$

where R_s is the relevance for the branch of $\mathrm{sigm}(\mathbf{V}\boldsymbol{m}_l)$, R_g is the relevance for the branch of $\tanh(\boldsymbol{U}\boldsymbol{m}_l)$ and R_a is the relevance received by the attention weight. The modification is similar to LRP-abs, which in principle takes the absolute values of $\mathbf{V}\boldsymbol{m}_l$ and $\boldsymbol{U}\boldsymbol{m}_l$, if it is applied to the gated attention mechanism. The scales of $\mathbf{V}\boldsymbol{m}_l$ and $\boldsymbol{U}\boldsymbol{m}_l$ are not bounded, so the LRP-abs rule cannot properly reflect the individual contributions in the element-wise multiplication. In addition, there is also a similar element-wise structure in the linear combination of AWs and the hidden feature vectors to form the combined hidden feature vector. Applying the principle of the LRP-all rule, AWs would be considered as $\omega_{s \to t}$ and

elements of hidden feature vectors as u_s in (5) if relevance maps are based on $f(\boldsymbol{x})$. No relevance would be redistributed through the gated attention mechanism, but the hidden feature vectors. The proposed rule was also applied to redistribute the relevance to AWs and hidden feature vectors. For LRP to be applicable to the DenseNet, we followed suit in (Binder, 2020) to merge the collection of BN, ReLU, and Conv into equivalent Conv and the collection of BN and FC into equivalent FC during the LRP backward pass. Other layers in the NN were the same as in (Montavon et al., 2019). For each qMR image, a relevance map was generated.

Here, we proposed to start the LRP backward pass from AWs for the relevance map instead of the output value $f(\boldsymbol{x})$. The AW as the proxy for the importance of the input image should focus on more relevant features. In addition, the parallel encoding structure of GAMER-MRI and the corresponding AWs could potentially allow a linearly combined map incorporating the joint information of the input qMR images.

$$Combined\ map\ =\ R_{MWF} * MWF + R_{NDI} * NDI + R_{qT1} * qT1. \quad (10)$$

2.2.4 Assessment

To assess if the proposed modifications could reveal more important brain regions, a voxel value inverting experiment from the aspect of model performance and Spearman's correlation (ρ) with EDSS from the clinical aspect were conducted. The joint dataset from the test and the validation datasets was in use for both assessments to have a sufficiently large number of samples. The relevance maps and the combined maps were binarized from high to low quantiles of positive relevance within the brain. In the voxel inverting experiment, voxel values of the normalized qMR images were inverted, i.e., $qMR = 1 - qMR$, according to the quantile masks. If regions identified by a quantile mask are important, the AUC is affected by inverting voxel values in the regions. If a quantile mask based on a scenario can affect the AUC more than the quantile mask based on other scenarios, this means the mask identifies more important regions. There were three conditions to be evaluated, including (i) whether the relevance map was based on AWs or $f(\boldsymbol{x})$, (ii) if the proposed rule or the LRP-all rule was used and (iii) whether the individual relevance maps or the combined map were/was considered. Table 2 lists the eight

Scenarios				
	AW		$f(\boldsymbol{x})$	
Rule	**Proposal**	**LRP-all**	**Proposal**	**LRP-all**
Individual Relevance Map	1	2	5	6
Combined Map	3	4	7	8

Table 2: The eight scenarios for comparison.

Scenarios				
	AW		$f(\boldsymbol{x})$	
Rule	**Saliency**	**Integrated Gradient**	**Saliency**	**Integrated Gradient**
Individual Relevance Map	9	10	11	12

Table 3: Four additional scenarios for comparison using saliency and integrated gradients to obtain relevant maps.

possible scenarios. Furthermore, heatmaps based on the saliency and integrated gradient methods were also in comparison (Table 3). The scenario/scenarios that achieved the largest drop in AUC was/were assessed for the correlation with EDSS. The correlation was performed on the normalized qMR values, which were averaged within the quantile mask of the scenario. The two-sided permutation test with 20,000 permutations was used for testing the strongest correlation achieved by the quantile masks in the scenario. The statistics were performed in R.

The quantile mask of the strongest correlation of each patient was further transformed to the MNI152 template (Grabner et al., 2006) for exploring potential group effect areas on the heatmap. The transformation matrices were obtained by greedy (Yushkevich, n.d.; Yushkevich et al., 2016) in ITK-Snap (Yushkevich et al., 2006) using the normalized cross-correlation metric with a 2x2x2 patch radius in the nonlinear deformable registration of MP2RAGE to MNI152. The c3d_affine_tool of ITK-Snap was used to convert the previously obtained transformation matrix between MP2RAGE and multi-shell DWI to the itk format.

3. Results

3.1 Performance

The test and averaged cross-validation results in AUC, accuracy, specificity, sensitivity, and the mean AWs of qMRs are reported in Table 4 and Table 5. As in (Lu et al., 2021b, 2021a), the reported AWs were averaged across the correctly classified samples. qT1 was the most important, followed by NDI and MWF in terms of the AWs both in the test dataset and the cross-validation folds on average.

Cross-Validation dataset

Metrics	AUC	Accuracy	Specificity	Sensitivity
	0.864	0.809	0.839	0.718
	MWF	**NDI**		**qT1**
AWs	0.188	0.309		0.503

Table 4: The average cross-validation result.

Test dataset

Metrics	AUC	Accuracy	Specificity	Sensitivity
	0.885	0.854	0.844	0.889
	MWF	**NDI**		**qT1**
AWs	0.164	0.369		0.467

Table 5: The test dataset result.

3.2 Voxel inverting experiment

The AUCs in the scenarios 1,2,3,4 are given in Fig. 2. The scenarios 3 and 4, where the quantile masks were defined by the combined map based on AWs using, respectively, the proposed rule and the LRP-all rule, had similar AUC for most of the compared quantiles. From the top 40^{th} quantile, the scenario 3 had a lower AUC than other scenarios. The performance of the scenarios 5 to 8 is illustrated in Fig. 3 and of the

scenarios 9 to 12 in Fig. 4 in addition to the best scenario 3 are shown in Fig. 2. The scenario 3 achieved the largest decrease in AUC.

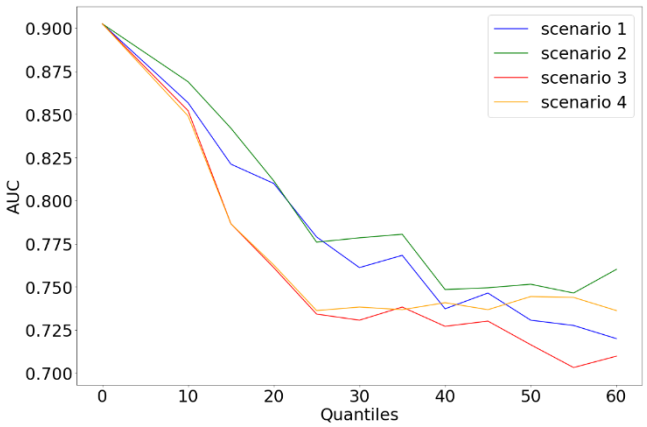


Figure 2: The AUC of different top Nth quantile masks in the voxel inverting experiment for the scenarios from 1 to 4. The scenarios used the relevance maps based on the attention weights and the combinations of different rules and the individual masks or the combined mask.

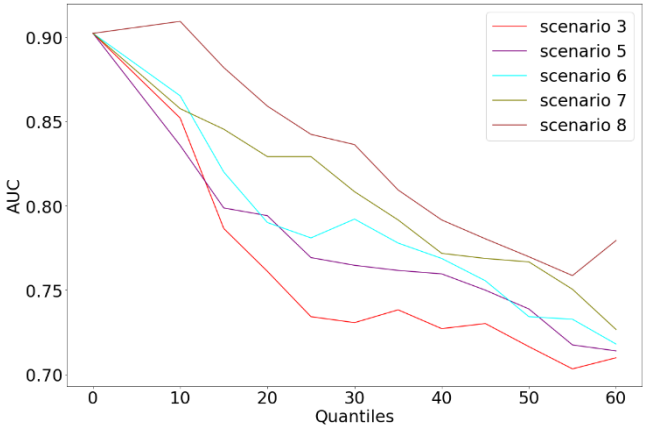


Figure 3: The AUC of different top Nth quantile masks in the voxel inverting experiment for the scenarios from 5 to 8 and the scenario 3. The scenarios used the relevance maps based on the output $f(\boldsymbol{x})$ and the combinations of different rules and the individual masks or the combined mask.

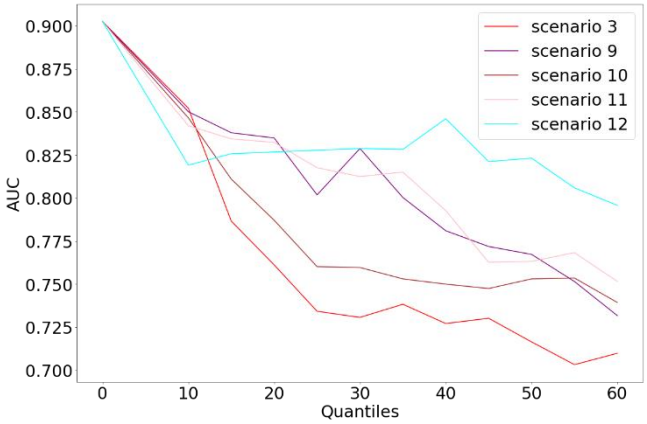


Figure 4: The AUC of different top Nth quantile masks in the voxel inverting experiment for the scenarios from 9 to 12 and the scenario 3. The scenarios used the heatmaps of the saliency method and the integrated gradient method based on the output $f(\boldsymbol{x})$ and the attention weights.

3.3 Correlation with EDSS

When qMR voxel values within the top 40th quantile mask, i.e., the 60th quantile, were averaged, ρ of qT1 and NDI were the largest (Table 6). Compared with ρ in the scenario 3 (Table 6), ρ of qT1 and NDI in the scenario 4 (Table 7) were smaller. Therefore, the permutation test was performed on the regions covered by the top 40th quantile mask in the scenario 3. The correlation is statistically significant, and the p-values are reported in Table 8. The NDI images of two exemplary patients in the two groups, as well as the top 40th quantile masks, are shown in Fig. 5. The mask of the patient in the severe disability group covered extensively the posterior limb of the internal capsule, where the corticospinal tract goes through and thus, the average NDI values are high in the relevant regions.

Spearman's Correlation Coefficient

Top Nth Quantile	qT1	NDI
10	-0.288	0.341
20	-0.344	0.414
30	-0.362	0.429
40	-0.371	0.440
50	-0.366	0.431
60	-0.346	0.411

Table 6: Spearman's correlation coefficients by different quantile masks of the scenario 3.

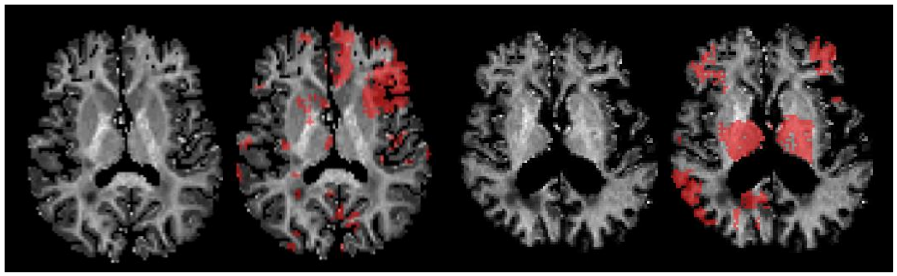
Spearman's Correlation Coefficient

Top Nth Quantile	qT1	NDI
10	-0.219	0.239
20	-0.248	0.314
30	-0.269	0.337
40	-0.281	0.348
50	-0.277	0.345
60	-0.261	0.343

Table 7: Spearman's correlation coefficients by different quantile masks of the scenario 4.

	qT1	NDI	MWF
Correlation	-0.371	0.440	0.298
P-value	<0.001	<0.0001	<0.01

Table 8: Spearman's correlation coefficient in the scenario 3 with a top 40th quantile mask and the corresponding p-values from the permutation test



(A) (B) (C) (D)

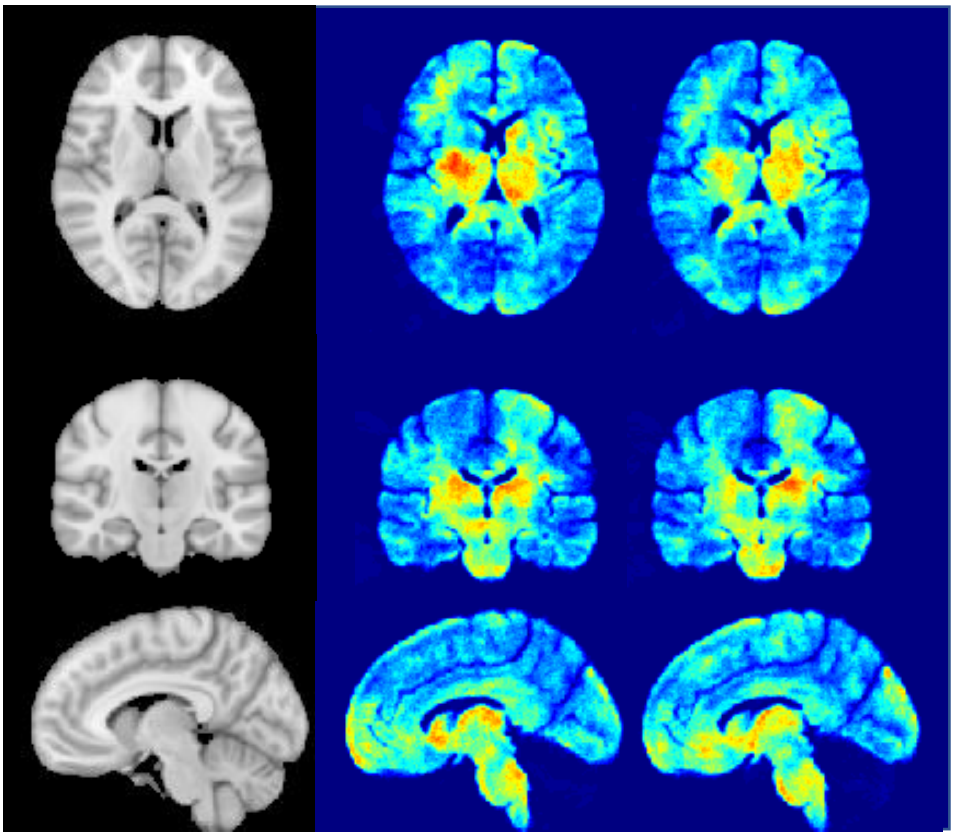
Figure 5: The top 40th quantile masks of the scenario 3 on two exemplary patients' NDI images. (A) the NDI image of a patient in the mild disability group; (B) the NDI image in (A) overlapped with the mask; (C) the NDI image of a patient in the severe disability; (D) The NDI image in (C) overlapped with the mask.

3.4 Heatmap on MNI152

Fig. 6 illustrates the heatmaps of the scenarios 3 and 4. Identified regions include the left thalamus, the left internal capsule, and part of the putamen. Scenario 3, which used the proposed rule and the combined mask based on AWs, identified more regions, including the left caudate, a larger part of the right putamen, and the right internal capsule.

4. Discussion

In this work, we extended the capability of GAMER-MRI trained from scratch using volumetric qMR images to classify MS patients with a severe or mild disability. Besides, we provided evidence that the proposed LRP rule for the gated attention mechanism in combination with AWs and the combined map could better identify brain regions whose alterations were most related to patients' disability than the LRP-all rule.



(A) (B) (C)

Figure 6: The heatmaps of the top 40th quantile masks of the scenarios 3 and 4 on the MNI 152 template. (A) MNI152 template (B) the heatmap of the scenario 3. (C) the heatmap of the scenario 4. The color scale is from 0 to 0.65.

The comparable NN performance on cross-validation and test datasets in Table 4 and 5 indicated that the NN learned a right representation to classify patients into severe and moderate/mild disability groups. This established the ground for further analyses related to LRP. According to the obtained attention weights, qT1 was the measure that best discriminated clinical severity in our cohort of MS patients. This might be due to different reasons: (i) qT1 provides a more comprehensive representation of tissue damage compared to MWF and NDI (i.e. global microstructural damage and iron accumulation vs myelin and axon-related damage) (Granziera et al., 2021) or (ii) the original higher spatial resolution and the higher white-grey matter contrast of qT1 reveal more details compared to MWF and NDI. Fig. 2, 3, and 4 demonstrate that AUC is affected most by inverting qMR voxel values within regions identified by the combined map based on AWs. These results meet our hypothesis on the advantage of the combined map, which is formed by the linear combination of AWs and qMR images. Furthermore,

the proposed rule and the relevance maps based on AWs affected more performance than the maps based on $f(\boldsymbol{x})$. This supports our hypothesis that AWs, used as a proxy for the importance of input qMR images, contain more relevant information than the output from the classifier. Scenario 3 shows a statistically significant correlation with EDSS, stronger than that with other scenarios listed in Table 2. From a clinical perspective, this supports our hypothesis that the identified regions were not merely important to the NN, but also meaningfully related to clinical measures of disability in MS patients. In Fig. 3, the impacted performance by the mask of scenarios 5 and 6 using the proposed rule was less than in scenarios 7 and 8 using the LRP-all rule. This may be due to the fact that the elements of the hidden feature vectors are unbounded in the element-wise multiplication contrary to AWs, which are between 0~1 and sum to 1. A potential way to tackle this issue is to apply an activation function to the hidden feature vectors prior to the gated attention mechanism so that the scale of elements of the hidden feature vectors is on par with AWs.

In the brain regions identified as relevant by the NN, the negative correlation of qT1 and the positive correlation of NDI with EDSS possibly indicate that the integrity or the structural composition of those brain regions – rather than their damage – is the driver of the classification. NDI provides specific information about axonal content (Zhang et al., 2012), and qT1 quantifies the overall microstructural tissue environment (Bonnier et al., 2014). The axonal damage and demyelination within lesions should increase the qT1 and decrease NDI. In Fig. 5, the relevant regions overlap less with the MS lesions, which are near the posterior horn of the lateral ventricles. For the patient in the severe disability group (Fig. 5), the relevant regions include a larger part of the posterior limb of the internal capsule and the thalamus, which contain the bundled corticospinal tract and collateral fibers to the tract (Tovar-Moll et al., 2015) and thus, moderate to high NDI values in comparison with other brain regions. Therefore, the opposite trend of the correlation of qT1 and NDI points us in the direction of assessing the structural integrity of brain tissue for patients' ability to maintain their function. This might also be partly reflected by the fact that the average cross-validation performance achieved specificity higher than sensitivity. To further validate the interpretation in this cross-sectional patient study, future works on the differentiation between healthy controls and patients and the assessment of the longitudinal changes in a larger group of patients having a severe disability are required.

On the MNI152 template, the regions identified by the heatmap of the top 40th quantile masks of all samples overlap with part of the corticospinal tract and regions of many collateral fibers (Fig. 6), which are strongly related to EDSS and disability (Tovar-Moll et al., 2015).

In (Eitel et al., 2019), healthy controls and MS patients were classified using 3D FLAIR and a CNN, and the ε-rule was applied to obtain a relevance map. In (Cruciani et al., 2021), the authors trained single-path CNNs to classify relapsing-remitting MS and progressive MS patients for each of 3D T1-weighted, diffusion MRI and microstructural maps derived from diffusion MRI within grey matter and the αß-rule with α=1, ß=0 was used. Different than in those previous works, in our study, we used the multi-path CNN-based NN to simultaneously take into account the information provided by multiple quantitative MRIs of tissue properties. Additionally, we used a collection of LRP, an approach that was shown to perform better than using single rules (Montavon et al., 2019). Besides, we have evaluated the relevance of the maps by using a quantile-based region-of-interest analysis as previously performed (Lopatina et al., 2020).

There were some limitations to this work. The first limitation was the size of the dataset. Even though the size of the dataset is large compared with other neuroimaging studies exploiting qMRI and DenseNet greatly reduces the required number of learnable parameters, the original DenseNet configuration with the parallel encoding structure of GAMER-MRI leads to a sizeable number of parameters. It is easily prone to overfitting using qMR images in this work. The minimal configuration of the DenseNet encoder was experimented with to have better performance. The richness of hidden representation is limited to certain degrees. Transfer learning can alleviate this issue and was demonstrated in (Eitel et al., 2019) using the Alzheimer's Disease Neuroimaging Initiative dataset. We have attempted to pre-train the model on the same patients' conventional MRI contrasts and fine-tuned it on the qMR images, but the model performance did not improve. It might be that the representation of conventional MRI contrasts learned by the pre-trained model was quite different from the target application using the qMR images and the model suffered from the negative transfer issue (Wang et al., 2019). The second limitation was the choice of LRP rules. Here we only utilized rules including the ε-rule, the αβ-rule, and the z^{β}-rule in addition to the proposed and LRP-all rules. There are more

LRP rules for different types and positions of layers and NN structures (Hui and Binder, 2019; Motzkus et al., 2022), and it might be beneficial for a more comprehensive assessment to experiment on the best combinations of rules. Another limitation was that the relevance maps and the combined map were used as different quantile masks instead of the values being used. The average number of voxels in a volumetric MR image is on the order of 10^6 and the AW is 0~1, so the relevance value received by a voxel was often tiny. Results in (Eitel et al., 2019) and (Cruciani et al., 2021) were also affected by the tiny values. The value itself can be affected differently across data samples depending on the choice of ε in the ε-rule and the numerical precision used for training the NN. In (Eitel et al., 2019), ε was 0.001, while it was 1 in (Lopatina et al., 2020). The use of the mask based on quantiles of the relevance values functioned as a workaround to this issue. Furthermore, only the positive relevance values were considered and the assessment of the negative relevance values will be a potential future work. The validity of (4) also imposed a limitation. The bias parameters were not considered during the LRP backward pass and this led to the fact that the obtained relevance map reflected only the interaction of the input images and the partial pattern learned by the NN. On the other hand, incorporation of the bias cannot enforce the sum of the relevance maps to be the same as the starting value, and the relevance values can be of different orders of magnitude.

5. Conclusion

In this work, we demonstrate that GAMER-MRI identifies brain regions that are potentially related to patients' disability in qMRI maps, which are sensitive to MS pathology. In addition, the combination of the proposed LRP rule for the gated attention mechanism, starting the LRP backward pass from the attention weights, and the combined importance map reveal potentially important regions to the integrity of the mobile function. Future work will aim at (i) investigating the hidden representations learned by GAMER-MRI and their pathological meaning; (ii) integrating other qMRI measures such as quantitative susceptibility mapping (QSM) and magnetization transfer saturation (MTsat) and (iii) increasing the richness and completeness of the representations by the incorporation of bias parameters during LRP backward pass. The method is suitable and easily adaptable to study different diseases.

Chapter 5. Discussion and Conclusion

The goal of this project was to develop a tool able to jointly and selectively consider the advanced MRI or derived quantitative measures in studying MS axonal and myelin damage and repair. Advanced MRI and the derived quantitative measures provide abundant information on the microstructural changes caused by MS inflammation, demyelination, remyelination, and axonal damage. How to maximize the utilization of all the available information, while selecting the most informative MR contrasts to reduce redundancy was an issue.

In the first work, we showed that GAMER MRI provided (through its parallel encoding structure) attention weights, whose order met the clinically meaningful order of importance for MR-based features, in the classification of infarct strokes. GAMER MRI, furthermore, proposed a potential order of importance for quantitative measures in the classification of Juxtacortical and periventricular lesions in MS. These results demonstrated its clinical potential in selecting the more informative MRI measures based on the corresponding attention weights. However, in this work, multiple evaluations were required to overcome the strong correlation of attention weights of the MR contrasts in the experiment of classification of acute/subacute infarct stroke patients vs other patients and healthy controls. The sources of the strong correlation were hypothesized to include the intrinsically physical correlation and the similar image representation of infarct lesions on TRACE and ADC at one of the phases of the stroke evolvement.

In the second publication, we tackled the issue of the aforementioned correlation among different diffusion measures derived from the same diffusion MRI but based on different biophysical assumptions in the classification of lesion and perilesional tissue patches. The results showed that by enhancing the differences prior to the conversion to attention weights by the softmax function and the proposed selection process, GAMER-MRI was able to select informative diffusion measures. In addition, the combination of the selected measures at the lesion level led to a stronger correlation with the patient-

level clinical measure of mobile impairment and the biological measure of axonal damage. However, the use of subject-wise normalization rendered the interpretation of the combined value difficult since it diminished the physical meaning of the measures. This suggested the use of a combination of diffusion-based microstructural measures deserved further attention for better interpretability of its pathological meaning.

The third work demonstrated the feasibility of GAMER-MRI to select important volumetric quantitative MR images while being trained from scratch for patient-level classification, despite the generality of and heterogeneity within EDSS - a measure of disability in MS patients. With the proposed relevance propagation rule, the new referent relevance to propagate, and the linearly combined map, we could identify brain regions whose alterations were mostly related to EDSS by GAMER-MRI. The average quantitative measures within the identified regions appeared to reflect the integrity of the mobile function considering the physical meanings and the respective correlation with EDSS of qT1 and NDI. This result of the linearly combined map addressed the joint consideration of quantitative measures. There are some limitations that need to be addressed or alleviated in the future. The first limitation was the large number of trainable parameters due to the parallel encoding structure of GAMER-MRI since each quantitative measure required an individual encoder. This constituted a challenge to avoid/alleviate overfitting given the number of the volumetric data and to include many more measures. The second limitation was the loss of information of the negative relevance. In the analysis, only the positive relevance within brain was considered. Future work will need to include negative relevance for more comprehensive analyses. In addition, the relevance maps of each quantitative measure and the combined map were used as a mask defined by the quantiles of the relevance values. It is of deep interest to further develop on how to use the relevance values directly instead of being used for the creation of masks. By addressing this, it might be possible to condense the combined relevance values into a single index indicating the extent of the integrity of brain function.

To conclude, the developed GAMER-MRI addressed the issue of selective consideration of advanced MRI and the derived quantitative measures in disentangling the interplay between axonal damage and repair, such as remyelination and axonal reorganization through the intensity of attention weights. The issue of joint consideration was addressed by the linearly combined map based on attention weights and the incorporation of the proposed relevance propagation rule with other layer-wise relevance propagation rules.

Publication

First-Authored Papers:

1. **Lu, P.-J.,** Barakovic, M., Weigel, M., Rahmanzadeh, R., Galbusera, R., Schiavi, S., Daducci, A., La Rosa, F., Bach Cuadra, M., Sandkühler, R., Kuhle, J., Kappos, L., Cattin, P., Granziera, C., 2021a. *GAMER-MRI in Multiple Sclerosis Identifies the Diffusion-Based Microstructural Measures That Are Most Sensitive to Focal Damage: A Deep-Learning-Based Analysis and Clinico-Biological Validation*. Front. Neurosci. 15. https://doi.org/10.3389/fnins.2021.647535

2. **Lu, P.-J.,** Yoo, Y., Rahmanzadeh, R., Galbusera, R., Weigel, M., Ceccaldi, P., Nguyen, T.D., Spincemaille, P., Wang, Y., Daducci, A., La Rosa, F., Bach Cuadra, M., Sandkühler, R., Nael, K., Doshi, A., Fayad, Z.A., Kuhle, J., Kappos, L., Odry, B., Cattin, P., Gibson, E., Granziera, C., 2021b. *GAMER MRI: Gated-attention mechanism ranking of multi-contrast MRI in brain pathology*. NeuroImage Clin. 29, 102522. https://doi.org/10.1016/j.nicl.2020.102522

3. Todea, R.A.*, **Lu, P.J.*,** Fartaria, M.J., Bonnier, G., Du Pasquier, R., Krueger, G., Bach Cuadra, M., Psychogios, M.N., Kappos, L., Kuhle, J., Granziera, C., 2020. Evolution of Cortical and White Matter Lesion Load in Early-Stage Multiple Sclerosis: Correlation With Neuroaxonal Damage and Clinical Changes. Front. Neurol. https://doi.org/10.3389/fneur.2020.00973
   * Equally contributed

Coauthored Papers:

1. Cagol, A., Schaedelin, S., Barakovic, M., Benkert, P., Todea, R.-A., Rahmanzadeh, R., Galbusera, R., **Lu, P.-J.,** Weigel, M., Melie-Garcia, L., Ruberte, E., Siebenborn, N., Battaglini, M., Radue, E.-W., Yaldizli, Ö., Oechtering, J., Sinnecker, T., Lorscheider, J., Fischer-Barnicol, B., Müller, S., Achtnichts, L., Vehoff, J., Disanto, G., Findling, O., Chan, A., Salmen, A., Pot, C., Bridel, C., Zecca, C., Derfuss, T., Lieb, J.M., Remonda, L., Wagner, F., Vargas, M.I., Du Pasquier, R., Lalive, P.H., Pravatà, E., Weber, J., Cattin, P.C., Gobbi, C., Leppert, D., Kappos, L., Kuhle, J., Granziera, C., 2022. *Association of Brain Atrophy With Disease Progression Independent of Relapse Activity in Patients With Relapsing Multiple Sclerosis.* JAMA Neurol. https://doi.org/10.1001/jamaneurol.2022.1025

2. Bosticardo, S., Schiavi, S., Schaedelin, S., **Lu, P.-J.,** Barakovic, M., Weigel, M., Kappos, L., Kuhle, J., Daducci, A., Granziera, C., 2022. *Microstructure-Weighted Connectomics in Multiple Sclerosis.* Brain Connect. 12, 6–17. https://doi.org/10.1089/brain.2021.0047

3. Schiavi, S., **Lu, P.-J.,** Weigel, M., Lutti, A., Jones, D.K., Kappos, L., Granziera, C., Daducci, A., 2022. *Bundle myelin fraction (BMF) mapping of different white matter connections using microstructure informed tractography*. Neuroimage 249, 118922. https://doi.org/10.1016/j.neuroimage.2022.118922

4. Maggi, P., Kuhle, J., Schädelin, S., van der Meer, F., Weigel, M., Galbusera, R., Mathias, A., **Lu, P.-J.,** Rahmanzadeh, R., Benkert, P., La Rosa, F., Bach Cuadra, M., Sati, P., Théaudin, M., Pot, C., van Pesch, V., Leppert, D., Stadelmann, C., Kappos, L., Du Pasquier, R., Reich, D.S., Absinta, M., Granziera, C., 2021. *Chronic White Matter Inflammation and Serum Neurofilament Levels in Multiple Sclerosis*. Neurology 97, e543–e553. https://doi.org/10.1212/WNL.0000000000012326

5. Rahmanzadeh, R., **Lu, P.-J.,** Barakovic, M., Weigel, M., Maggi, P., Nguyen, T.D., Schiavi, S., Daducci, A., La Rosa, F., Schaedelin, S., Absinta, M., Reich, D.S., Sati, P., Wang, Y., Bach Cuadra, M., Radue, E.-W., Kuhle, J., Kappos, L., Granziera, C., 2021. *Myelin and axon pathology in multiple sclerosis assessed by myelin water and multi-shell diffusion imaging*. Brain 144, 1684–1696. https://doi.org/10.1093/brain/awab088

6. Weigel, M., Dechent, P., Galbusera, R., Bahn, E., Nair, G., **Lu, P.-J.,** Kappos, L., Brück, W., Stadelmann, C., Granziera, C., 2021. *Imaging multiple sclerosis pathology at 160 µm isotropic resolution by human whole-brain ex vivo magnetic resonance imaging at 3 T*. Sci. Rep. 11, 15491. https://doi.org/10.1038/s41598-021-94891-1

7. Barquero, G., La Rosa, F., Kebiri, H., **Lu, P.-J.**, Rahmanzadeh, R., Weigel, M., Fartaria, M.J., Kober, T., Théaudin, M., Du Pasquier, R., Sati, P., Reich, D.S., Absinta, M., Granziera, C., Maggi, P., Bach Cuadra, M., 2020. *RimNet: A deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis*. NeuroImage Clin. 28, 102412. https://doi.org/10.1016/j.nicl.2020.102412

8. La Rosa, F., Abdulkadir, A., Fartaria, M.J., Rahmanzadeh, R., **Lu, P.-J.**, Galbusera, R., Barakovic, M., Thiran, J.-P.P., Granziera, C., Cuadra, M.B., 2020. *Multiple sclerosis cortical and WM lesion segmentation at 3T MRI: a deep learning method based on FLAIR and MP2RAGE*. NeuroImage Clin. 27, 102335. https://doi.org/10.1016/j.nicl.2020.102335

Bibliography

Akçakaya, M., Moeller, S., Weingärtner, S., Uğurbil, K., 2019. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. Magn. Reson. Med. https://doi.org/10.1002/mrm.27420

Albert, M., Antel, J., Brück, W., Stadelmann, C., 2007. Extensive cortical remyelination in patients with chronic multiple sclerosis. Brain Pathol. 17, 129–38. https://doi.org/10.1111/j.1750-3639.2006.00043.x

Allen, L.M., Hasso, A.N., Handwerker, J., Farid, H., 2012. Sequence-specific MR imaging findings that are useful in dating ischemic stroke. Radiographics. https://doi.org/10.1148/rg.325115760

Andermatt, S., Pezold, S., Cattin, P.C., 2018. Automated Segmentation of Multiple Sclerosis Lesions Using Multi-dimensional Gated Recurrent Units. pp. 31–42. https://doi.org/10.1007/978-3-319-75238-9_3

Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. Neuroimage. https://doi.org/10.1016/S1053-8119(03)00336-7

Andersson, J.L.R., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage. https://doi.org/10.1016/j.neuroimage.2015.10.019

Arras, L., Arjona-Medina, J.A., Widrich, M., Montavon, G., Gillhofer, M., Müller, K.-R., Hochreiter, S., Samek, W., 2019. Explaining and Interpreting LSTMs. https://doi.org/10.1007/978-3-030-28954-6_11

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.R., Samek, W., 2015. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS One 10, 1–46. https://doi.org/10.1371/journal.pone.0130140

Bahdanau, D., Cho, K.H., Bengio, Y., 2015. Neural machine translation by jointly learning to align and translate, in: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

Barkhof, F., 2002. The clinico-radiological paradox in multiple sclerosis revisited. Curr. Opin. Neurol. https://doi.org/10.1097/00019052-200206000-00003

Barkhof, F., Brück, W., De Groot, C.J., Begers, E., Hulshof, S., Geurts, J., Polman, C.H., 2003a. Remyelinated Lesions in Multiple Sclerosis. Arch. Neurol. 60, 1073–81.

Barkhof, F., Bruck, W., De Groot, C.J.A., Bergers, E., Hulshof, S., Geurts, J., Polman, C.H., van der Valk, P., 2003b. Remyelinated lesions in multiple sclerosis: magnetic resonance image appearance. Arch. Neurol. 60, 1073–81. https://doi.org/10.1001/archneur.60.8.1073

Barquero, G., La Rosa, F., Kebiri, H., Lu, P.-J., Rahmanzadeh, R., Weigel, M., Fartaria, M.J., Kober, T., Théaudin, M., Du Pasquier, R., Sati, P., Reich, D.S., Absinta, M., Granziera, C., Maggi, P., Bach Cuadra, M., 2020. RimNet: A deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis. NeuroImage Clin. 28, 102412. https://doi.org/10.1016/j.nicl.2020.102412

Barro, C., Benkert, P., Disanto, G., Tsagkas, C., Amann, M., Naegelin, Y., Leppert, D., Gobbi, C., Granziera, C., Yaldizli, Ö., Michalak, Z., Wuerfel, J., Kappos, L., Parmar, K., Kuhle, J., 2018. Serum neurofilament as a predictor of disease worsening and brain and spinal cord

atrophy in multiple sclerosis. Brain. https://doi.org/10.1093/brain/awy154
Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. Magn. Reson. Med. https://doi.org/10.1002/mrm.10609
Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
Binder, A., 2020. Notes on canonization for resnets and densenets [WWW Document]. URL https://github.com/AlexBinder/LRP_Pytorch_Resnets_Densenet/canonization doc.pdf
Böhle, M., Eitel, F., Weygandt, M., Ritter, K., 2019. Layer-Wise Relevance Propagation for Explaining Deep Neural Network Decisions in MRI-Based Alzheimer's Disease Classification. Front. Aging Neurosci. 11. https://doi.org/10.3389/fnagi.2019.00194
Bojorquez, J.Z., Bricq, S., Acquitter, C., Brunotte, F., Walker, P.M., Lalande, A., 2017. What are normal relaxation times of tissues at 3 T? Magn. Reson. Imaging 35, 69–80. https://doi.org/10.1016/j.mri.2016.08.021
Bonnier, G., Roche, A., Romascano, D., Simioni, S., Meskaldji, D., Rotzinger, D., Lin, Y.C., Menegaz, G., Schluep, M., Du Pasquier, R., Sumpf, T.J., Frahm, J., Thiran, J.P., Krueger, G., Granziera, C., 2014. Advanced MRI unravels the nature of tissue alterations in early multiple sclerosis. Ann. Clin. Transl. Neurol. 1, 423–432. https://doi.org/10.1002/acn3.68
Bozzali, M., Serra, L., Cercignani, M., 2016. Quantitative MRI to understand Alzheimer's disease pathophysiology. Curr. Opin. Neurol. https://doi.org/10.1097/WCO.0000000000000345
Brosch, T., Tang, L.Y.W., Yoo, Y., Li, D.K.B., Traboulsee, A., Tam, R., 2016. Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation. IEEE Trans. Med. Imaging 35, 1229–1239. https://doi.org/10.1109/TMI.2016.2528821
Cabeen, R.P., Sepehrband, F., Toga, A.W., 2019. Rapid and Accurate NODDI Parameter Estimation with the Spherical Mean Technique, in: ISMRM.
Canty, A.J., Huang, L., Jackson, J.S., Little, G.E., Knott, G., Maco, B., De Paola, V., 2013. In-vivo single neuron axotomy triggers axon regeneration to restore synaptic density in specific cortical circuits. Nat. Commun. https://doi.org/10.1038/ncomms3038
Carass, A., Roy, S., Jog, A., Cuzzocreo, J.L., Magrath, E., Gherman, A., Button, J., Nguyen, J., Prados, F., Sudre, C.H., Jorge Cardoso, M., Cawley, N., Ciccarelli, O., Wheeler-Kingshott, C.A.M., Ourselin, S., Catanese, L., Deshpande, H., Maurel, P., Commowick, O., Barillot, C., Tomas-Fernandez, X., Warfield, S.K., Vaidya, S., Chunduru, A., Muthuganapathy, R., Krishnamurthi, G., Jesson, A., Arbel, T., Maier, O., Handels, H., Iheme, L.O., Unay, D., Jain, S., Sima, D.M., Smeets, D., Ghafoorian, M., Platel, B., Birenbaum, A., Greenspan, H., Bazin, P.L., Calabresi, P.A., Crainiceanu, C.M., Ellingsen, L.M., Reich, D.S., Prince, J.L., Pham, D.L., 2017. Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. Neuroimage 148, 77–102. https://doi.org/10.1016/j.neuroimage.2016.12.064
Chitnis, T., Gonzalez, C., Healy, B.C., Saxena, S., Rosso, M., Barro, C., Michalak, Z., Paul, A., Kivisakk, P., Diaz-Cruz, C., Sattarnezhad, N., Pierre, I. V., Glanz, B.I., Tomic, D., Kropshofer, H., Häring, D., Leppert, D., Kappos, L., Bakshi, R., Weiner, H.L., Kuhle, J., 2018. Neurofilament light chain serum levels correlate with 10-year MRI outcomes in multiple sclerosis. Ann. Clin. Transl. Neurol. https://doi.org/10.1002/acn3.638
Commowick, O., Cervenansky, F., Ameli, R., 2016. MSSEG Challenge Proceedings: Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing

Infrastructure. Miccai.
Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A.N., Hajnal, J. V., 2019. Complex diffusion-weighted image estimation via matrix recovery under general noise models. Neuroimage. https://doi.org/10.1016/j.neuroimage.2019.06.039
Cruciani, F., Brusini, L., Zucchelli, M., Retuci Pinheiro, G., Setti, F., Boscolo Galazzo, I., Deriche, R., Rittner, L., Calabrese, M., Menegaz, G., 2021. Interpretable deep learning as a means for decrypting disease signature in multiple sclerosis. J. Neural Eng. 18, 0460a6. https://doi.org/10.1088/1741-2552/ac0f4b
Daducci, A., Canales-Rodríguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., Thiran, J.P., 2015. Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. Neuroimage. https://doi.org/10.1016/j.neuroimage.2014.10.026
Deh, K., Ponath, G.D., Molvi, Z., Parel, G.-C.T., Gillen, K.M., Zhang, S., Nguyen, T.D., Spincemaille, P., Ma, Y., Gupta, A., Gauthier, S.A., Pitt, D., Wang, Y., 2018. Magnetic susceptibility increases as diamagnetic molecules breakdown: Myelin digestion during multiple sclerosis lesion formation contributes to increase on QSM. J. Magn. Reson. Imaging 48, 1281–1287. https://doi.org/10.1002/jmri.25997
Dilokthornsakul, P., Valuck, R.J., Nair, K. V., Corboy, J.R., Allen, R.R., Campbell, J.D., 2016. Multiple sclerosis prevalence in the United States commercially insured population. Neurology 86, 1014–1021. https://doi.org/10.1212/WNL.0000000000002469
Ding, Y., Liu, Y., Luan, H., Sun, M., 2017. Visualizing and understanding neural machine translation. ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap. 1, 1150–1159. https://doi.org/10.18653/v1/P17-1106
Duncan, I.D., Radcliff, A.B., Heidari, M., Kidd, G., August, B.K., Wierenga, L.A., 2018. The adult oligodendrocyte can participate in remyelination. Proc. Natl. Acad. Sci. 115. https://doi.org/10.1073/pnas.1808064115
Dziedzic, T., Metz, I., Dallenga, T., König, F.B., Müller, S., Stadelmann, C., Brück, W., 2010a. Wallerian degeneration: A major component of early axonal pathology in multiple sclerosis. Brain Pathol. 20, 976–985. https://doi.org/10.1111/j.1750-3639.2010.00401.x
Dziedzic, T., Metz, I., Dallenga, T., König, F.B., Müller, S., Stadelmann, C., Brück, W., 2010b. Wallerian degeneration: a major component of early axonal pathology in multiple sclerosis. Brain Pathol. 20, 976–85. https://doi.org/10.1111/j.1750-3639.2010.00401.x
Eitel, F., Soehler, E., Bellmann-Strobl, J., Brandt, A.U., Ruprecht, K., Giess, R.M., Kuchling, J., Asseyer, S., Weygandt, M., Haynes, J.-D., Scheel, M., Paul, F., Ritter, K., 2019. Uncovering convolutional neural network decisions for diagnosing multiple sclerosis on conventional MRI using layer-wise relevance propagation. NeuroImage Clin. 24, 102003. https://doi.org/10.1016/j.nicl.2019.102003
Fischl, B., Liu, A., Dale, A.M., 2001. Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans. Med. Imaging. https://doi.org/10.1109/42.906426
Fujimoto, K., Polimeni, J.R., van der Kouwe, A.J.W., Reuter, M., Kober, T., Benner, T., Fischl, B., Wald, L.L., 2014. Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7T. Neuroimage 90, 60–73. https://doi.org/10.1016/j.neuroimage.2013.12.012
Geurts, J.J., Barkhof, F., 2008. Grey matter pathology in multiple sclerosis. Lancet Neurol. 7, 841–851. https://doi.org/10.1016/S1474-4422(08)70191-1
Ghesu, F.C., Georgescu, B., Mansoor, A., Yoo, Y., Gibson, E., Vishwanath, R.S., Balachandran, A., Balter, J.M., Cao, Y., Singh, R., Digumarthy, S.R., Kalra, M.K., Grbic, S., Comaniciu, D.,

2021. Quantifying and leveraging predictive uncertainty for medical image assessment. Med. Image Anal. 68, 101855. https://doi.org/10.1016/j.media.2020.101855
Goldschmidt, T., Antel, J., König, F.B., Brück, W., Kuhlmann, T., 2009. Remyelination capacity of the MS brain decreases with disease chronicity. Neurology 72, 1914–1921. https://doi.org/10.1212/WNL.0b013e3181a8260a
González, R.G., Schwamm, L.H., 2016. Imaging acute ischemic stroke. pp. 293–315. https://doi.org/10.1016/B978-0-444-53485-9.00016-7
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.
Grabner, G., Janke, A.L., Budge, M.M., Smith, D., Pruessner, J., Collins, D.L., 2006. Symmetric Atlasing and Model Based Segmentation: An Application to the Hippocampus in Older Adults. pp. 58–66. https://doi.org/10.1007/11866763_8
Granberg, T., Fan, Q., Treaba, C.A., Ouellette, R., Herranz, E., Mangeat, G., Louapre, C., Cohen-Adad, J., Klawiter, E.C., Sloane, J.A., Mainero, C., 2017. In vivo characterization of cortical and white matter neuroaxonal pathology in early multiple sclerosis. Brain 140, 2912–2926. https://doi.org/10.1093/brain/awx247
Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, Nicola, Enzinger, C., Evangelou, N., Filippi, M., Geurts, J.J.G., Reich, D.S., Rocca, M.A., Ropele, S., Rovira, À., Sati, P., Toosy, A.T., Vrenken, H., Gandini Wheeler-Kingshott, C.A.M., Kappos, L., Barkhof, F., de Stefano, N, Sastre-Garriga, J., Ciccarelli, O., Enzinger, C., Filippi, M., Gasperini, C., Kappos, L., Palace, J., Vrenken, H., Rovira, À., Rocca, M.A., Yousry, T., 2021. Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. Brain 144, 1296–1311. https://doi.org/10.1093/brain/awab029
Gupta, A., Al-Dasuqi, K., Xia, F., Askin, G., Zhao, Y., Delgado, D., Wang, Y., 2017. The use of noncontrast quantitative MRI to detect gadolinium-enhancing multiple sclerosis brain lesions: A systematic review and meta-analysis. Am. J. Neuroradiol. https://doi.org/10.3174/ajnr.A5209
Harkins, K.D., Xu, J., Dula, A.N., Li, K., Valentine, W.M., Gochberg, D.F., Gore, J.C., Does, M.D., 2016. The microstructural correlates of T1in white matter. Magn. Reson. Med. https://doi.org/10.1002/mrm.25709
Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-Excitation Networks. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 7132–7141. https://doi.org/10.1109/CVPR.2018.00745
Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. https://doi.org/10.1109/CVPR.2017.243
Hui, L.Y.W., Binder, A., 2019. BatchNorm Decomposition for Deep Neural Network Interpretation. pp. 280–291. https://doi.org/10.1007/978-3-030-20518-8_24
Ilse, M., Tomczak, J.M., Welling, M., 2018. Attention-based Deep Multiple Instance Learning.
Jelescu, I.O., Veraart, J., Adisetiyo, V., Milla, S.S., Novikov, D.S., Fieremans, E., 2015. One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI? Neuroimage. https://doi.org/10.1016/j.neuroimage.2014.12.009
Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. Review FSL. Neuroimage. https://doi.org/10.1016/j.neuroimage.2011.09.015
Kaden, E., Kelm, N.D., Carson, R.P., Does, M.D., Alexander, D.C., 2016. Multi-compartment microscopic diffusion imaging. Neuroimage. https://doi.org/10.1016/j.neuroimage.2016.06.002
Kingma, D.P., Ba, J.L., 2015. Adam: A method for stochastic optimization, in: 3rd International

Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.
Kober, T., Granziera, C., Ribes, D., Browaeys, P., Schluep, M., Meuli, R., Frackowiak, R., Gruetter, R., Krueger, G., 2012. MP2RAGE multiple sclerosis magnetic resonance imaging at 3 T. Invest. Radiol. https://doi.org/10.1097/RLI.0b013e31824600e9
Koschmieder, K., Paul, M.M., van den Heuvel, T.L.A., van der Eerden, A.W., van Ginneken, B., Manniesing, R., 2022. Automated detection of cerebral microbleeds via segmentation in susceptibility-weighted images of patients with traumatic brain injury. NeuroImage Clin. 35, 103027. https://doi.org/10.1016/j.nicl.2022.103027
Kuhlmann, T., Lingfeld, G., Bitsch, A., Schuchardt, J., Brück, W., 2002. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. Brain 125, 2202–12.
Kuhlmann, T., Ludwin, S., Prat, A., Antel, J., Brück, W., Lassmann, H., 2017. An updated histological classification system for multiple sclerosis lesions. Acta Neuropathol. 133, 13–24. https://doi.org/10.1007/s00401-016-1653-y
Kutzelnigg, A., Lucchinetti, C.F., Stadelmann, C., Brück, W., Rauschka, H., Bergmann, M., Schmidbauer, M., Parisi, J.E., Lassmann, H., 2005. Cortical demyelination and diffuse white matter injury in multiple sclerosis. Brain 128, 2705–2712. https://doi.org/10.1093/brain/awh641
La Rosa, F., Abdulkadir, A., Fartaria, M.J., Rahmanzadeh, R., Lu, P.-J.J., Galbusera, R., Barakovic, M., Thiran, J.-P.P., Granziera, C., Cuadra, M.B., 2020. Multiple sclerosis cortical and WM lesion segmentation at 3T MRI: a deep learning method based on FLAIR and MP2RAGE. NeuroImage Clin. 27, 102335. https://doi.org/10.1016/j.nicl.2020.102335
La Rosa, F., Fartaria, M.J., Kober, T., Richiardi, J., Granziera, C., Thiran, J.-P., Cuadra, M.B., 2019. Shallow vs Deep Learning Architectures for White Matter Lesion Segmentation in the Early Stages of Multiple Sclerosis. pp. 142–151. https://doi.org/10.1007/978-3-030-11723-8_14
Lakhani, D.A., Schilling, K.G., Xu, J., Bagnato, F., 2020. Advanced multicompartment diffusion MRI models and their application in multiple sclerosis. Am. J. Neuroradiol. https://doi.org/10.3174/AJNR.A6484
Lassmann, H., 2018. Multiple Sclerosis Pathology. Cold Spring Harb. Perspect. Med. 8, a028936. https://doi.org/10.1101/cshperspect.a028936
Laule, C., Leung, E., Lis, D.K.B., Traboulsee, A.L., Paty, D.W., MacKay, A.L., Moore, G.R.W., 2006. Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. Mult. Scler. 12, 747–53. https://doi.org/10.1177/1352458506070928
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539
Lieury, A., Chanal, M., Androdias, G., Reynolds, R., Cavagna, S., Giraudon, P., Confavreux, C., Nataf, S., 2014. Tissue remodeling in periplaque regions of multiple sclerosis spinal cord lesions. Glia 62, 1645–1658. https://doi.org/10.1002/glia.22705
Lopatina, A., Ropele, S., Sibgatulin, R., Reichenbach, J.R., Güllmar, D., 2020. Investigation of Deep-Learning-Driven Identification of Multiple Sclerosis Patients Based on Susceptibility-Weighted Images Using Relevance Analysis. Front. Neurosci. 14. https://doi.org/10.3389/fnins.2020.609468
Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization, in: 7th International Conference on Learning Representations, ICLR 2019.
Lu, J., Yang, J., Batra, D., Parikh, D., 2016. Hierarchical Question-Image Co-Attention for Visual Question Answering, in: Proceedings of the 30th International Conference on Neural

Information Processing Systems, NIPS'16. Curran Associates Inc., Red Hook, NY, USA, pp. 289–297.
Lu, P.-J., Barakovic, M., Weigel, M., Rahmanzadeh, R., Galbusera, R., Schiavi, S., Daducci, A., La Rosa, F., Bach Cuadra, M., Sandkühler, R., Kuhle, J., Kappos, L., Cattin, P., Granziera, C., 2021a. GAMER-MRI in Multiple Sclerosis Identifies the Diffusion-Based Microstructural Measures That Are Most Sensitive to Focal Damage: A Deep-Learning-Based Analysis and Clinico-Biological Validation. Front. Neurosci. 15. https://doi.org/10.3389/fnins.2021.647535
Lu, P.-J., Yoo, Y., Rahmanzadeh, R., Galbusera, R., Weigel, M., Ceccaldi, P., Nguyen, T.D., Spincemaille, P., Wang, Y., Daducci, A., La Rosa, F., Bach Cuadra, M., Sandkühler, R., Nael, K., Doshi, A., Fayad, Z.A., Kuhle, J., Kappos, L., Odry, B., Cattin, P., Gibson, E., Granziera, C., 2021b. GAMER MRI: Gated-attention mechanism ranking of multi-contrast MRI in brain pathology. NeuroImage Clin. 29, 102522. https://doi.org/10.1016/j.nicl.2020.102522
Lublin, F.D., Reingold, S.C., Cohen, J. a, Cutter, G.R., Thompson, A.J., Wolinsky, J.S., Fox, R.J., Freedman, M.S., Goodman, A.D., Lubetzki, C., 2014. Defining the clinical course of multiple sclerosis : The 2013 revisions Defining the clinical course of multiple sclerosis The 2013 revisions 1–10. https://doi.org/10.1212/WNL.0000000000000560
Lucchinetti, C., Br&#xfffd;ck, W., Parisi, J., Scheithauer, B., Rodriguez, M., Lassmann, H., 2000. Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. Ann. Neurol. 47, 707–717. https://doi.org/10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems.
Lundervold, A.S., Lundervold, A., 2019. An overview of deep learning in medical imaging focusing on MRI. Z. Med. Phys. 29, 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002
Luong, T., Pham, H., Manning, C.D., 2015. Effective Approaches to Attention-based Neural Machine Translation, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 1412–1421. https://doi.org/10.18653/v1/D15-1166
MacKay, A.L., Laule, C., 2016. Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin. Brain Plast. 2, 71–91. https://doi.org/10.3233/BPL-160033
Maggi, P., Fartaria, M.J., Jorge, J., La Rosa, F., Absinta, M., Sati, P., Meuli, R., Du Pasquier, R., Reich, D.S., Cuadra, M.B., Granziera, C., Richiardi, J., Kober, T., 2020. CVSnet: A machine learning approach for automated central vein sign assessment in multiple sclerosis. NMR Biomed. 33. https://doi.org/10.1002/nbm.4283
Maicas, G., Carneiro, G., Bradley, A.P., Nascimento, J.C., Reid, I., 2017. Deep reinforcement learning for active breast lesion detection from DCE-MRI, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-66179-7_76
Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. Neuroimage. https://doi.org/10.1016/j.neuroimage.2009.10.002
Marzullo, A., Kocevar, G., Stamile, C., Calimeri, F., Terracina, G., Durand-Dubief, F., Sappey-Marinier, D., 2019. Prediction of Multiple Sclerosis Patient Disability from Structural Connectivity using Convolutional Neural Networks, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 2087–2090. https://doi.org/10.1109/EMBC.2019.8856845

Metz, I., Weigand, S.D., Popescu, B.F.G., Frischer, J.M., Parisi, J.E., Guo, Y., Lassmann, H., Brück, W., Lucchinetti, C.F., 2014. Pathologic heterogeneity persists in early active multiple sclerosis lesions. Ann. Neurol. 75, 728–738. https://doi.org/10.1002/ana.24163

Moll, N.M., Rietsch, A.M., Thomas, S., Ransohoff, A.J., Lee, J.-C., Fox, R., Chang, A., Ransohoff, R.M., Fisher, E., 2011. Multiple sclerosis normal-appearing white matter: Pathology-imaging correlations. Ann. Neurol. 70, 764–773. https://doi.org/10.1002/ana.22521

Montavon, G., Binder, A., Lapuschkin, S., Samek, W., Müller, K.R., 2019. Layer-Wise Relevance Propagation: An Overview. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11700 LNCS, 193–209. https://doi.org/10.1007/978-3-030-28954-6_10

Montavon, G., Lapuschkin, S., Binder, A., Samek, W., Müller, K.-R., 2017. Explaining nonlinear classification decisions with deep Taylor decomposition. Pattern Recognit. 65, 211–222. https://doi.org/10.1016/j.patcog.2016.11.008

Motzkus, F., Weber, L., Lapuschkin, S., 2022. Measurably Stronger Explanation Reliability via Model Canonization.

Mustafi, S., Harezlak, J., Kodiweera, C., Randolph, J., Ford, J., Wishart, H., Wu, Y.-C., 2019. Detecting white matter alterations in multiple sclerosis using advanced diffusion magnetic resonance imaging. Neural Regen. Res. 14, 114. https://doi.org/10.4103/1673-5374.243716

Nael, K., Gibson, E., Yang, C., Ceccaldi, P., Yoo, Y., Das, J., Doshi, A., Georgescu, B., Janardhanan, N., Odry, B., Nadar, M., Bush, M., Re, T.J., Huwer, S., Josan, S., von Busch, H., Meyer, H., Mendelson, D., Drayer, B.P., Comaniciu, D., Fayad, Z.A., 2021. Automated detection of critical findings in multi-parametric brain MRI using a system of 3D neural networks. Sci. Rep. 11, 6876. https://doi.org/10.1038/s41598-021-86022-7

Nguyen, T.D., Deh, K., Monohan, E., Pandya, S., Spincemaille, P., Raj, A., Wang, Y., Gauthier, S.A., 2016. Feasibility and reproducibility of whole brain myelin water mapping in 4 minutes using fast acquisition with spiral trajectory and adiabatic T2prep (FAST-T2) at 3T. Magn. Reson. Med. 76, 456–465. https://doi.org/10.1002/mrm.25877

Niu, Z., Zhong, G., Yu, H., 2021. A review on the attention mechanism of deep learning. Neurocomputing 452, 48–62. https://doi.org/10.1016/j.neucom.2021.03.091

Novikov, D.S., Fieremans, E., Jespersen, S.N., Kiselev, V.G., 2019. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. NMR Biomed. https://doi.org/10.1002/nbm.3998

Patrikios, P., Stadelmann, C., Kutzelnigg, A., Rauschka, H., Schmidbauer, M., Laursen, H., Sorensen, P.S., Bruck, W., Lucchinetti, C., Lassmann, H., 2006. Remyelination is extensive in a subset of multiple sclerosis patients. Brain 129, 3165–3172. https://doi.org/10.1093/brain/awl217

Payan, A., Montana, G., 2015. Predicting Alzheimer's disease a neuroimaging study with 3D convolutional neural networks, in: ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods, Proceedings.

Poiron, E., 2018. Remyelination fails in the periventricular white matter in MS. ECTRIMS.

Reisert, M., Kellner, E., Dhital, B., Hennig, J., Kiselev, V.G., 2017. Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. Neuroimage. https://doi.org/10.1016/j.neuroimage.2016.09.058

Rensonnet, G., Scherrer, B., Girard, G., Jankovski, A., Warfield, S.K., Macq, B., Thiran, J.P., Taquet, M., 2019. Towards microstructure fingerprinting: Estimation of tissue properties from a dictionary of Monte Carlo diffusion MRI simulations. Neuroimage. https://doi.org/10.1016/j.neuroimage.2018.09.076

Rovira, Á., Wattjes, M.P., Tintoré, M., Tur, C., Yousry, T.A., Sormani, M.P., De Stefano, N., Filippi, M., Auger, C., Rocca, M.A., Barkhof, F., Fazekas, F., Kappos, L., Polman, C., Miller, D., Montalban, X., 2015. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis - Clinical implementation in the diagnostic process. Nat. Rev. Neurol. https://doi.org/10.1038/nrneurol.2015.106

Ruiz, J., Mahmud, M., Modasshir, M., Shamim Kaiser, M., Alzheimer's Disease Neuroimaging In, for the, 2020. 3D DenseNet Ensemble in 4-Way Classification of Alzheimer's Disease. pp. 85–96. https://doi.org/10.1007/978-3-030-59277-6_8

Saha, S., Pagnozzi, A., Bourgeat, P., George, J.M., Bradford, D.K., Colditz, P.B., Boyd, R.N., Rose, S.E., Fripp, J., Pannek, K., 2020. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model. Neuroimage. https://doi.org/10.1016/j.neuroimage.2020.116807

Samek, W., Binder, A., Montavon, G., Lapuschkin, S., Muller, K.-R., 2017. Evaluating the Visualization of What a Deep Neural Network Has Learned. IEEE Trans. Neural Networks Learn. Syst. 28, 2660–2673. https://doi.org/10.1109/TNNLS.2016.2599820

Samek, W., Montavon, G., Lapuschkin, S., Anders, C.J., Muller, K.-R., 2021. Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. Proc. IEEE 109, 247–278. https://doi.org/10.1109/JPROC.2021.3060483

Scherrer, B., Schwartzman, A., Taquet, M., Sahin, M., Prabhu, S.P., Warfield, S.K., 2016. Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND). Magn. Reson. Med. https://doi.org/10.1002/mrm.25912

Schlemper, J., Caballero, J., Hajnal, J. V., Price, A.N., Rueckert, D., 2018. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2017.2760978

Schneider, T., Brownlee, W., Zhang, H., Ciccarelli, O., Miller, D.H., Wheeler-Kingshott, C.G., 2017. Sensitivity of multi-shell NODDI to multiple sclerosis white matter changes: A pilot study. Funct. Neurol. https://doi.org/10.11138/FNeur/2017.32.2.097

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2016. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Int. J. Comput. Vis. 128, 336–359. https://doi.org/10.1007/s11263-019-01228-7

Shoeibi, A., Khodatars, M., Jafari, M., Moridian, P., Rezaei, M., Alizadehsani, R., Khozeimeh, F., Gorriz, J.M., Heras, J., Panahiazar, M., Nahavandi, S., Acharya, U.R., 2021. Applications of deep learning techniques for automated multiple sclerosis detection using magnetic resonance imaging: A review. Comput. Biol. Med. 136, 104697. https://doi.org/10.1016/j.compbiomed.2021.104697

Siller, N., Kuhle, J., Muthuraman, M., Barro, C., Uphaus, T., Groppa, S., Kappos, L., Zipp, F., Bittner, S., 2019. Serum neurofilament light chain is a biomarker of acute and chronic neuronal damage in early multiple sclerosis. Mult. Scler. J. https://doi.org/10.1177/1352458518765666

Simonyan, K., Vedaldi, A., Zisserman, A., 2014. Deep inside convolutional networks: Visualising image classification models and saliency maps. 2nd Int. Conf. Learn. Represent. ICLR 2014 - Work. Track Proc. 1–8.

Simonyan, K., Vedaldi, A., Zisserman, A., 2013. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.

Sivaranjini, S., Sujatha, C.M., 2020. Deep learning based diagnosis of Parkinson's disease using convolutional neural network. Multimed. Tools Appl. 79, 15467–15479.

https://doi.org/10.1007/s11042-019-7469-8
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL, in: NeuroImage. https://doi.org/10.1016/j.neuroimage.2004.07.051
Spincemaille, P., Anderson, J., Wu, G., Yang, B., Fung, M., Li, K., Li, S., Kovanlikaya, I., Gupta, A., Kelley, D., Benhamo, N., Wang, Y., 2020. Quantitative Susceptibility Mapping: MRI at 7T versus 3T. J. Neuroimaging 30, 65–75. https://doi.org/10.1111/jon.12669
Stüber, C., Morawski, M., Schäfer, A., Labadie, C., Wähnert, M., Leuze, C., Streicher, M., Barapatre, N., Reimann, K., Geyer, S., Spemann, D., Turner, R., 2014. Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. Neuroimage 93 Pt 1, 95–106. https://doi.org/10.1016/j.neuroimage.2014.02.026
Sundararajan, M., Taly, A., Yan, Q., 2017. Axiomatic attribution for deep networks. 34th Int. Conf. Mach. Learn. ICML 2017 7, 5109–5118.
Todea, R.A., Lu, P.J., Fartaria, M.J., Bonnier, G., Du Pasquier, R., Krueger, G., Bach Cuadra, M., Psychogios, M.N., Kappos, L., Kuhle, J., Granziera, C., 2020. Evolution of Cortical and White Matter Lesion Load in Early-Stage Multiple Sclerosis: Correlation With Neuroaxonal Damage and Clinical Changes. Front. Neurol. https://doi.org/10.3389/fneur.2020.00973
Tomczak, J.M., Ilse, M., Welling, M., Jansen, M., Coleman, H.G., Lucas, M., Laat, K. de, Bruin, M. de, Marquering, H., Wel, M.J. van der, Boer, O.J. de, Heijink, C.D.S., Meijer, S.L., 2018. Histopathological classification of precursor lesions of esophageal adenocarcinoma: A Deep Multiple Instance Learning Approach. Med. Imaging with Deep Learn. 3–5.
Tonietto, M., 2018. Periventricular remyelination is associated with grey matter atrophy in MS. ECTRIMS.
Tournier, J.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A., 2019. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. Neuroimage. https://doi.org/10.1016/j.neuroimage.2019.116137
Tousignant, A., Paul Lemaˆıtre Lemaˆıtre, M., Doina Precup, C., Arnold, D.L., 2019. Prediction of Disease Progression in Multiple Sclerosis Patients using Deep Learning Analysis of MRI Data Tal Arbel 3, Proceedings of Machine Learning Research.
Tovar-Moll, F., Evangelou, I.E., Chiu, A.W., Auh, S., Chen, C., Ehrmantraut, M., Ohayon, J.M., Richert, N., Bagnato, F., 2015. Diffuse and Focal Corticospinal Tract Disease and Its Impact on Patient Disability in Multiple Sclerosis. J. Neuroimaging 25, 200–206. https://doi.org/10.1111/jon.12171
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need.
Voita, E., Talbot, D., Moiseev, F., Sennrich, R., Titov, I., 2020. Analyzing multi-head self-attention: Specialized heads do the heavy lifting, the rest can be pruned. ACL 2019 - 57th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. 5797–5808. https://doi.org/10.18653/v1/p19-1580
Wachinger, C., Reuter, M., Klein, T., 2018. DeepNAT: Deep convolutional neural network for segmenting neuroanatomy. Neuroimage 170, 434–445. https://doi.org/10.1016/j.neuroimage.2017.02.035
Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R.A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N.,

Baneke, P., 2020. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. Mult. Scler. J. 26, 1816–1821. https://doi.org/10.1177/1352458520970841
Wang, Y., Liu, T., 2015a. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. 73, 82–101. https://doi.org/10.1002/mrm.25358
Wang, Y., Liu, T., 2015b. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. 73, 82–101. https://doi.org/10.1002/mrm.25358
Wang, Z., Dai, Z., Póczos, B., Carbonell, J., 2019. Characterizing and avoiding negative transfer, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11293–11302.
Wattjes, M.P., Ciccarelli, O., Reich, D.S., Banwell, B., de Stefano, N., Enzinger, C., Fazekas, F., Filippi, M., Frederiksen, J., Gasperini, C., Hacohen, Y., Kappos, L., Li, D.K.B., Mankad, K., Montalban, X., Newsome, S.D., Oh, J., Palace, J., Rocca, Maria A, Sastre-Garriga, J., Tintoré, M., Traboulsee, A., Vrenken, H., Yousry, T., Barkhof, F., Rovira, À., Wattjes, M.P., Ciccarelli, O., de Stefano, N., Enzinger, C., Fazekas, F., Filippi, M., Frederiksen, J., Gasperini, C., Hacohen, Y., Kappos, L., Mankad, K., Montalban, X., Palace, J., Rocca, María A, Sastre-Garriga, J., Tintore, M., Vrenken, H., Yousry, T., Barkhof, F., Rovira, A., Li, D.K.B., Traboulsee, A., Newsome, S.D., Banwell, B., Oh, J., Reich, D.S., Reich, D.S., Oh, J., 2021. 2021 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. Lancet Neurol. 20, 653–670. https://doi.org/10.1016/S1474-4422(21)00095-8
Wattjes, M.P., Rovira, À., Miller, D., Yousry, T.A., Sormani, M.P., De Stefano, N., Tintoré, M., Auger, C., Tur, C., Filippi, M., Rocca, M.A., Fazekas, F., Kappos, L., Polman, C., Barkhof, F., Montalban, X., 2015. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis - Establishing disease prognosis and monitoring patients. Nat. Rev. Neurol. https://doi.org/10.1038/nrneurol.2015.157
Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-González, J., Routier, A., Bottani, S., Dormont, D., Durrleman, S., Burgos, N., Colliot, O., 2020. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. Med. Image Anal. 63, 101694. https://doi.org/10.1016/j.media.2020.101694
Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: Convolutional Block Attention Module, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 3–19. https://doi.org/10.1007/978-3-030-01234-2_1
Ye, Z., George, A., Wu, A.T., Niu, X., Lin, J., Adusumilli, G., Naismith, R.T., Cross, A.H., Sun, P., Song, S.K., 2020. Deep learning with diffusion basis spectrum imaging for classification of multiple sclerosis lesions. Ann. Clin. Transl. Neurol. https://doi.org/10.1002/acn3.51037
Yoo, Y., Tang, L.Y.W., Brosch, T., Li, D.K.B., Kolind, S., Vavasour, I., Rauscher, A., MacKay, A.L., Traboulsee, A., Tam, R.C., 2018. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. NeuroImage Clin. 17, 169–178. https://doi.org/10.1016/j.nicl.2017.10.015
You, Q., Jin, H., Wang, Z., Fang, C., Luo, J., 2016. Image captioning with semantic attention, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2016.503
Yushkevich, P.. ., n.d. No Title [WWW Document]. URL

https://github.com/pyushkevich/greedy
Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage 31, 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015
Yushkevich, P.A., Pluta, J., Wang, H., Wisse, L.E.M., Das, S., Wolk, D., 2016. IC-P-174: Fast Automatic Segmentation of Hippocampal Subfields and Medial Temporal Lobe Subregions In 3 Tesla and 7 Tesla T2-Weighted MRI. Alzheimer's Dement. 12. https://doi.org/10.1016/j.jalz.2016.06.205
Zaheer, M., Kottur, S., Ravanbhakhsh, S., Póczos, B., Salakhutdinov, R., Smola, A.J., 2017. Deep sets, in: Advances in Neural Information Processing Systems.
Zeiler, M.D., Fergus, R., 2014. Visualizing and Understanding Convolutional Networks. pp. 818–833. https://doi.org/10.1007/978-3-319-10590-1_53
Zeineldin, R.A., Karar, M.E., Coburger, J., Wirtz, C.R., Burgert, O., 2020. DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. Int. J. Comput. Assist. Radiol. Surg. 15, 909–920. https://doi.org/10.1007/s11548-020-02186-z
Zeydan, B., Kantarci, O.H., 2020. Impact of Age on Multiple Sclerosis Disease Activity and Progression. Curr. Neurol. Neurosci. Rep. 20, 24. https://doi.org/10.1007/s11910-020-01046-2
Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 61, 1000–16. https://doi.org/10.1016/j.neuroimage.2012.03.072
Zhang, Y.-D., Pan, C., Sun, J., Tang, C., 2018. Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. J. Comput. Sci. 28, 1–10. https://doi.org/10.1016/j.jocs.2018.07.003
Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016. Learning Deep Features for Discriminative Localization, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 2921–2929. https://doi.org/10.1109/CVPR.2016.319