Deep-PRL: a deep learning network for the identification of paramagnetic rim lesions in multiple sclerosis

Federico Spagnolo[1,2,3,4], Aarushi Bhardwaj[1,5], Pedro M. Gordaliza[6,7], Po-Jui Lu[1,2,3], Mario Ocampo-Pineda[1,2,3], Meritxell Bach Cuadra[6,7], Xinjie Chen[1,2,3], Batuhan Ayci[1,8], Alessandro Cagol[1,2,3,9], Vincent Andrearczyk[4], Adrien Depeursinge[4,10], Cristina Granziera[1,2,3]

[1]Translational Imaging in Neurology (ThINK) Basel, Department of Biomedical Engineering, Faculty of Medicine, 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]MedGIFT, Institute of Informatics, School of Management, HES-SO Valais-Wallis University of Applied Sciences and Arts Western Switzerland, Sierre, Switzerland; [5]University of Guelph, Guelph, ON, Canada; [6]CIBM Center for Biomedical Imaging, Lausanne, Switzerland; [7]Radiology Department, Lausanne University Hospital (CHUV) and University of Lausanne, Lausanne, Switzerland; [8]Istanbul University-Cerrahpasa, Cerrahpasa Medical School, Istanbul, Turkey; [9]Dipartimento di Scienze della Salute, Università degli Studi di Genova, Genova, Italy; [10]Nuclear Medicine and Molecular Imaging Department, Lausanne University Hospital (CHUV) and University of Lausanne, Lausanne, Switzerland.

federico.spagnolo@unibas.ch
aarushi@uoguelph.ca
pedro.maciasgordaliza@unil.ch
p.lu@unibas.ch
mario.ocampopineda@unibas.ch
meritxell.bachcuadra@unil.ch
xinjie.chen@unibas.ch
batuhanayci@gmail.com
alessandro.cagol@unibas.ch
vincent.andrearczyk@hevs.ch
adrien.depeursinge@hevs.ch
cristina.granziera@unibas.ch

Synopsis

Keywords: Machine Learning, Multiple Sclerosis, Paramagnetic Rim Lesion, Classification, Chronic Active Lesion

Motivation: PRLs are an important diagnostic biomarker in people with multiple sclerosis (pwMS). Their identification on MRI is time-consuming and subject to high inter-rater variability. However, the use of AI could support this identification process.

Goal(s): We leverage multi-contrast MRI to improve the identification of PRLs.

Approach: Deep-PRL is an attention-based CNN, fusing features of T1-w and unwrapped phase images from 185 pwMS. The approach follows a nested cross-validation with patient stratification.

Results: The test performance outperformed state-of-the-art methods, achieving a mean F1 score of 0.860 ± 0.048 and an AUC of 0.982 ± 0.007.

Impact: These results represent a significant step towards the integration of an AI tool to assist clinicians in the identification of PRLs, thereby improving the management of pwMS.

Body of the abstract

Introduction

Paramagnetic rim lesions (PRLs) are an emerging biomarker valuable for diagnosing multiple sclerosis (MS), and have potential for patient prognosis and stratification[1]. Identifying PRLs on magnetic resonance imaging (MRI) is time-consuming and prone to high inter-rater variability. Artificial intelligence (AI) has the potential to improve PRL identification, providing significant benefits for both clinicians and people with MS (pwMS). PRLs are a subset of white matter lesions (WMLs), which can be distinguished from non-PRLs on susceptibility-sensitive MRI contrasts, such as unwrapped phase (UP), quantitative susceptibility mapping (QSM), and susceptibility-weighted imaging. According to a recent consensus statement[1], each of these contrasts has distinct advantages and disadvantages for assessing this biomarker. In histopathology, PRLs findings are highly correlated with the presence of chronic active lesions[2] (CALs), characterized by a demyelinated core surrounded by a ferritin-bound iron rim. Recent AI-based methods exploit a combination of conventional MRI, and either UP or QSM to distinguish PRLs from other WMLs. $APRL_1$[3], RimNet[4,5] and $APRL_2$[6] used UP, while QSMRim-Net[7] and DeDA[8] adopted QSM, following deep learning and radiomics approaches.

Methods

Brain MRI acquisitions from 185 pwMS (age: 47 ± 14; sex: 111 females; 61 secondary progressive, 23 primary progressive, and 101 relapsing-remitting; 495 PRLs and 5415 non-PRLs) were collected at the University Hospital of Basel, Switzerland. A WML mask was obtained using a CNN[9] and corrected by a neurologist on FLAIR; non-confluent PRLs were independently annotated by an expert neurologist and a medical student, reaching a consensus on UP images in the T2* space. UP were derived with MEDI algorithm[10] from T2* segmented echo planar imaging (EPI). Magnetization prepared 2 rapid acquisition gradient echoes (MP2RAGE) images were skull-stripped using FreeSurfer[11,12] and HD-BET[13], and registered to the T2* space using FSL[14,15,16]. The acquisition MRI protocol is described in Figure 1.
We propose Deep-PRL, a patch-based convolutional neural network (CNN) that uses three inputs in the T2* space: MP2RAGE, UP, and the dilated WML mask. The architecture, illustrated in Figure 2, exploits the WML mask through an attention branch.
To train and test the network, pwMS were stratified into four groups based on PRL count (i.e. 0, 1-3, 4-7, >7), and a nested cross-validation technique was applied (k=5 outer loop, k=3 inner loop). Patches of around 28x28x28 voxels were extracted around each lesion's center of mass, normalizing intensities between 0 and 1. In the training set, patches of positive examples were augmented by shifting the center of mass by 5 voxels. Data augmentation was used for random Gaussian noise injection, intensity shifts, flips along axes, 90-degree rotations, and affine transformations. The network was trained for 100 epochs using a polynomial learning rate scheduler, Adam optimizer, a batch size of 32 and a focal loss[17] function (γ=2, α=0.2).
Models with the best F1 validation performance were selected for external tests, and the average F1, sensitivity, specificity, and area under the receiver operating characteristic curve (ROC AUC) were reported.

Results

Deep-PRL achieved a mean (± standard deviation) test F1 score of 0.860 ± 0.048, sensitivity of 0.874 ± 0.031, specificity of 0.986 ± 0.09, and a ROC AUC of 0.982 ± 0.007. A comparison to state-of-the-art methods is provided in Figure 3, with detailed results for each of our test sets shown in Figure 4.

Discussion

As summarized in Figure 3, our network performs better than state-of-the-art methods in classifying PRLs, with significantly higher F1 scores, sensitivity, and AUC. However, comparisons must be interpreted cautiously, as results from these methods are obtained with different premises, datasets, and MRI contrasts. For example, confluent lesions were manually split by one rater and included in RimNet, QSMRim-Net and DeDA, whereas $APRL_1$ and $APRL_2$ followed an automatic pipeline. Additionally, $APRL_1$, QSMRim-Net and DeDA report validation performances, while RimNet, $APRL_2$ and Deep-PRL adopted a holdout test approach.
To enhance the clinical applicability of our method, future work should address the following limitations: 1) the exclusion of confluent PRLs, which may result in neglecting a significant number of PRLs; 2) the need for deeper characterization of artifacts or pseudo-PRLs[1,18] in UP that may reduce annotation specificity, potentially impacting the network's ability to identify CALs; 3) in Deep-PRL a lesion mask or clinician's input is required to define the target patch, and to enhance performances through the attention branch; 4) expanding experiments from single to multi-centric data would boost the generalizability of results.

Conclusions

Deep-PRL represents a significant advancement in the automatic identification of PRLs, offering a promising approach to expedite manual assessment and potentially facilitate clinical use.

Acknowledgements

This work was supported by the Hasler Foundation with the project MSxplain number 21042.

Disclosures

Fedrico Spagnolo has nothing to disclose.

Aarushi Bhardwaj has nothing to disclose.

Pedro M. Gordaliza has nothing to disclose.
Po-Jui Lu has nothing to disclose.
Mario Ocampo-Pineda has nothing to disclose.
Meritxell Bach Cuadra has nothing to disclose.

Xinjie Chen has nothing to disclose.
Batuhan Ayci has nothing to disclose.
Alessandro Cagol is supported by EUROSTAR E!113682 HORIZON2020, and received speaker honoraria from Novartis and Roche.
Vincent Andrearczyk has nothing to disclose.
Adrien Depeursinge has nothing to disclose.
Cristina Granziera as an employee of the University Hospital Basel (USB) has received the following fees which were used exclusively for research support: (1) advisory board and

consultancy fees from Actelion, Novartis, Genzyme, and F. Hoffmann-La Roche Ltd; (2) speaker fees from Biogen and Genzyme-Sanofi; and (3) research support by F. Hoffmann-La Roche Ltd.

References

1. Francesca Bagnato, Pascal Sati, Christopher C Hemond, Colm Elliott, Susan A Gauthier, Daniel M Harrison, Caterina Mainero, Jiwon Oh, David Pitt, Russell T. Shinohara, Seth A. Smith, Bruce Trapp, Christina J. Azevedo, Peter A. Calabresi, Roland G. Henry, Cornelia Laule, Daniel Ontaneda, William D. Rooney, Nancy L Sicotte, Daniel S. Reich, Martina Absinta, on behalf of the NAIMS Cooperative, Imaging chronic active lesions in multiple sclerosis: a consensus statement, *Brain*, 147 (9), 2913–2933, DOI: 10.1093/brain/awae013 (2024).
2. Rahmanzadeh, R., Galbusera, R., Lu, P.-J., Bahn, E., Weigel, M., Barakovic, M., Franz, J., Nguyen, T.D., Spincemaille, P., Schiavi, S., Daducci, A., La Rosa, F., Absinta, M., Sati, P., Bach Cuadra, M., Radue, E.-W., Leppert, D., Kuhle, J., Kappos, L., Brück, W., Reich, D.S., Stadelmann, C., Wang, Y. and Granziera, C., A New Advanced MRI Biomarker for Remyelinated Lesions in Multiple Sclerosis. Ann Neurol, 92: 486-502. DOI: 10.1002/ana.26441 (2022).
3. Carolyn Lou, Pascal Sati, Martina Absinta, Kelly Clark, Jordan D. Dworkin, Alessandra M. Valcarcel, Matthew K. Schindler, Daniel S. Reich, Elizabeth M. Sweeney, Russell T. Shinohara, Fully automated detection of paramagnetic rims in multiple sclerosis lesions on 3T susceptibility-based MR imaging, NeuroImage: Clinical, 32, 102796, DOI: 10.1016/j.nicl.2021.102796 (2021).
4. Germán Barquero, Francesco La Rosa, Hamza Kebiri, Po-Jui Lu, Reza Rahmanzadeh, Matthias Weigel, Mário João Fartaria, Tobias Kober, Marie Théaudin, Renaud Du Pasquier, Pascal Sati, Daniel S. Reich, Martina Absinta, Cristina Granziera, Pietro Maggi, Meritxell Bach Cuadra, RimNet: A deep 3D multimodal MRI architecture for paramagnetic rim lesion assessment in multiple sclerosis, NeuroImage: Clinical, 28, 102412, DOI: 10.1016/j.nicl.2020.102412 (2020).
5. Najm, Joe, Gordaliza, Pedro M., Barquero, Germán, LaRosa, Francesco, Wynen, Maxence, and Bach Cuadra, Meritxell, Streamline RimNet: Tools for Automatic Classification of Paramagnetic Rim Lesions in MRI of Multiple Sclerosis, Zenodo. DOI: 10.5281/zenodo.7962482 (2023).
6. Chen L, Ren Z, Clark KA, Lou C, Liu F, Cao Q, et al. Multicenter validation of automated detection of paramagnetic rim lesions on brain MRI in multiple sclerosis. *J Neuroimaging*, 1–8, DOI: 10.1111/jon.13242 (2024).
7. Hang Zhang, Thanh D. Nguyen, Jinwei Zhang, Melanie Marcille, Pascal Spincemaille, Yi Wang, Susan A. Gauthier, Elizabeth M. Sweeney, QSMRim-Net: Imbalance-aware learning for identification of chronic active multiple sclerosis lesions on quantitative susceptibility maps, NeuroImage: Clinical, 34, 102979, DOI: 10.1016/j.nicl.2022.102979 (2022).
8. Zhang, H., Wang, R., Hu, R., Zhang, J., Li, J. (2023). DeDA: Deep Directed Accumulator. In: Greenspan, H., *et al.* Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. MICCAI 2023. Lecture Notes in Computer Science, vol 14221. Springer, Cham. DOI: 10.1007/978-3-031-43895-0_72 (2023).
9. La Rosa, F. et al. Multiple sclerosis cortical and WM lesion segmentation at 3T MRI: a deep learning method based on FLAIR and MP2RAGE. NeuroImage: Clin. 27, 102335, DOI: 10.1016/j.nicl.2020.102335 (2020).

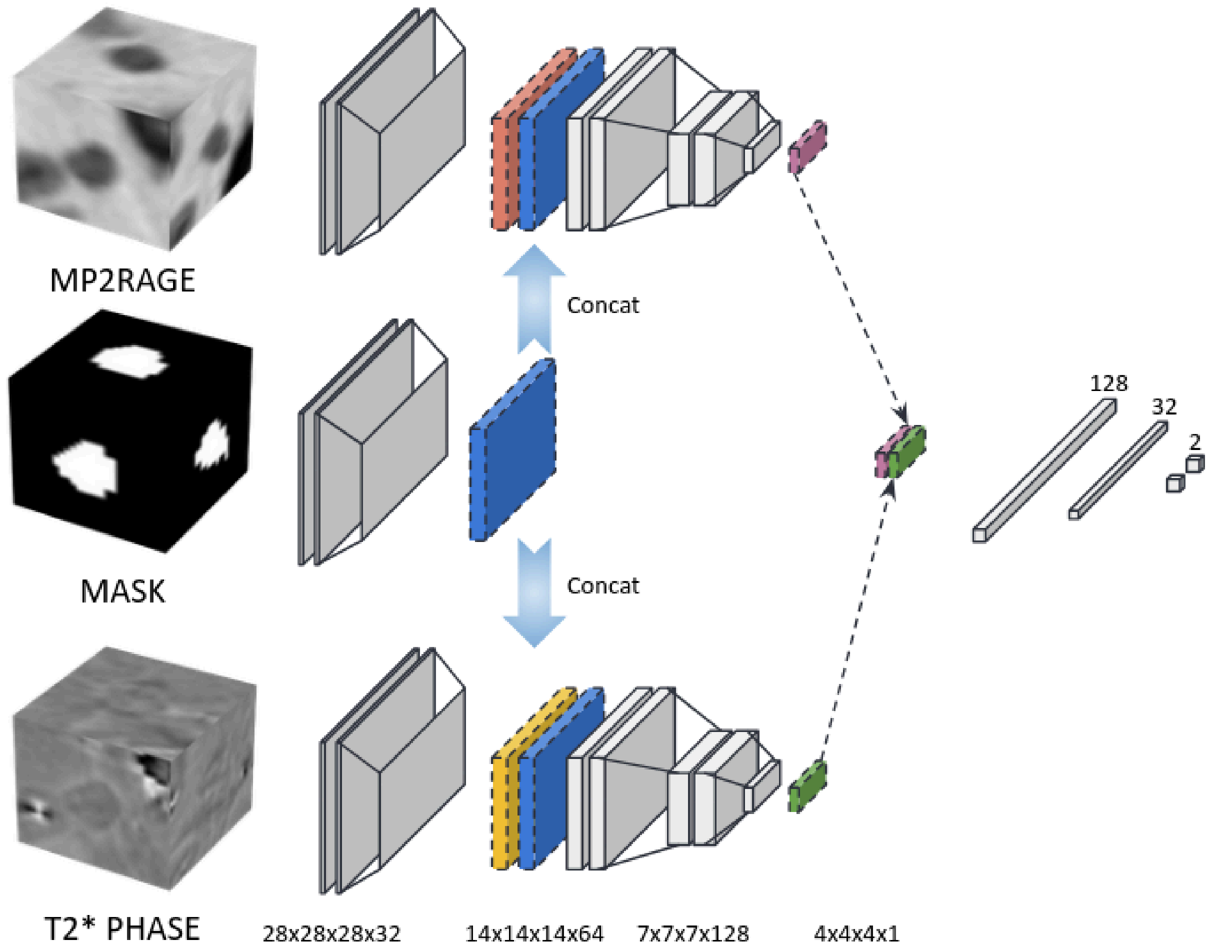
10. Liu T, Xu W, Spincemaille P, Avestimehr AS, Wang Y. Accuracy of the morphology enabled dipole inversion (MEDI) algorithm for quantitative susceptibility mapping in MRI. *IEEE Trans Med Imaging*. 2012; 31(3):816-824 (2012).
11. B. Fischl, D.H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A.M. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron, 33, 341-355, DOI: 10.1016/S0896-6273(02)00569-X (2002).
12. Kyoko Fujimoto, Jonathan R. Polimeni, André J.W. van der Kouwe, Martin Reuter, Tobias Kober, Thomas Benner, Bruce Fischl, Lawrence L. Wald, Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7T, NeuroImage, 90, 60-73, DOI: 10.1016/j.neuroimage.2013.12.012 (2014).
13. Isensee F, Schell M, Tursunova I, Brugnara G, Bonekamp D, Neuberger U, Wick A, Schlemmer HP, Heiland S, Wick W, Bendszus M, Maier-Hein KH, Kickingereder P. Automated brain extraction of multi-sequence MRI using artificial neural networks. Hum Brain Mapp., 1–13, DOI: 10.1002/hbm.24750 (2019).
14. Jenkinson, M., Bannister, P., Brady, J. M. and Smith, S. M. Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. NeuroImage, 17(2), 825-841 (2002).
15. Jenkinson, M. and Smith, S. M. A Global Optimisation Method for Robust Affine Registration of Brain Images. Medical Image Analysis, 5(2), 143-156 (2001).
16. Greve, D.N. and Fischl, B. Accurate and robust brain image alignment using boundary-based registration. NeuroImage, 48(1), 63-72 (2009).
17. T. -Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal Loss for Dense Object Detection," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, 2999-3007, DOI: 10.1109/ICCV.2017.324 (2017).
18. Huang W., Sweeney E. M., Kaunzner U. W., Wang Y., Gauthier S. A., Nguyen T. D. Quantitative susceptibility mapping versus phase imaging to identify multiple sclerosis iron rim lesions with demyelination. J Neuroimaging, 32(4):667-675. DOI: 10.1111/jon.12987 (2022).

Figure 1

	MP2RAGE	**FLAIR**	**T2* (3D EPI)**
TR (ms)	5000	5000	64
TE (ms)	2.98	386	35
TI (ms)	700; 2500	1800	/
FA (°)	4; 5	/	10
Resolution	1 mm iso	1 mm iso	0.67 mm iso

Description of the MRI protocol. Abbreviations: repetition time (TR), echo time (TE), inversion time (TI), flip angle (FA), fluid-attenuated inversion recovery (FLAIR), echo planar imaging (EPI), magnetization prepared 2 rapid acquisition gradient echoes (MP2RAGE).

Figure 2



Overview of the network's architecture. The attention branch combines low level features from the lesion mask with those of T2* phase and MP2RAGE. This encourages the network to extract high level features on meaningful regions of T2* phase and MP2RAGE.

Figure 3

	RimNet	**APRL$_1$**	**APRL$_2$**	**QSMRim-Net**	**DeDA**	**Deep-PRL**
TPs	/	24 (6)	70	120	/	86
FPs	/	31 (19)	482	44	/	15
FNs	/	8 (0)	45	57	/	13
TNs	/	135 (47)	1992	3942	/	1068
F1 score	0.623	0.552 (0.387)	0.210	0.704	0.750	**0.860**
specificity	0.951	0.813 (0.712)	0.805	0.989	**0.992**	0.986
sensitivity	0.758	0.750 (1.0)	0.609	0.678	0.712	**0.874**
PPV	0.528	0.436 (0.24)	0.127	0.732	0.792	**0.851**
ROC AUC	0.958	0.82 (0.88)	0.73	0.760	0.975	**0.982**

Performance comparison between Deep-PRL and state-of-the-art methods. In APRL$_1$ the numbers in parentheses are obtained excluding confluent lesions. Abbreviations: true positives (TPs), false positives (FPs), false negatives (FNs), true negatives (TNs), positive predictive value (PPV), area under the receiver operating characteristic curve (ROC AUC).

Figure 4

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average	STD
TPs	98	91	85	86	72	86.4	9.5
FPs	12	8	9	16	31	15.2	9.4
FNs	14	17	14	7	11	12.6	3.8
TNs	1149	1158	1064	961	1007	1067.8	86.4
F1 score	0.883	0.879	0.881	0.882	0.774	0.860	0.048
specificity	0.990	0.993	0.992	0.984	0.970	0.986	0.009
sensitivity	0.875	0.843	0.859	0.925	0.867	0.874	0.031
PPV	0.891	0.919	0.904	0.843	0.699	0.851	0.090
ROC AUC	0.978	0.990	0.976	0.989	0.975	0.982	0.007

Detailed performance in the separate test folders. Abbreviations: true positives (TPs), false positives (FPs), false negatives (FNs), true negatives (TNs), positive predictive value (PPV), area under the receiver operating characteristic curve (ROC AUC).

Thumbnail