Proc. Annual Meeting of the Austrian Society for Biomedical Engineering 2021
DOI: 10.3217/978-3-85125-826-4-16

Relevance-guided Deep Learning for Feature Identification in R2* Maps in Alzheimer's Disease Classification

C. Tinauer[1], S. Heber[1], L. Pirpamer[1], A. Damulina[1], M. Soellradl[2], M. Sackl[1], E. Hofer[1], M. Koini[1], R. Schmidt[1], R. Stollberger[2], S. Ropele[1], C. Langkammer[1]

[1]Department of Neurology, Medical University of Graz, Austria

[2]Institute of Medical Engineering, Graz University of Technology, Austria

christian.tinauer@medunigraz.at

Abstract- *When using deep neural networks to separate Alzheimer's disease patients (n=119) from normal controls (n=131) by using MR images, heat mapping revealed that the image preprocessing is introducing misleading features used by the classifier. Therefore we systematically investigated the influence of registration and brain extraction on the learned features by heat mapping. Results were compared to a novel relevance-guided training method, focusing on brain tissue. The relevance-guided configurations yielded highest classification accuracies and also confirmed histopathologically relevant regional iron deposition.*

Keywords- *Deep convolutional neural networks, heat mapping, relevance guidance, Alzheimer's disease*

Introduction

Deep learning techniques are increasingly utilized in medical applications, including image reconstruction [1], segmentation [2], and classification [3,4]. However, despite the good performance those models are not easily interpretable by humans [5]. Especially medical applications require verification that the high accuracy of those models is not the result of exploiting artifacts in the data [6]. Our previous experiments on Alzheimer's disease (AD) classification showed that Deep Neural Networks such as Convolutional Neural Networks (CNNs) might learn from features introduced by the brain extraction algorithm [7]. Therefore, in this work we investigated how preprocessing steps including registration and brain extraction determine which features in the R2* maps are relevant for the separation of patients with AD from normal controls. MR-based R2* mapping enables the in vivo detection of iron. Brain iron accumulates during aging and has been associated with neurodegenerative disorders including AD.

Methods

Dataset: We retrospectively selected 252 MRI datasets from 119 patients with probable AD (mean age=72.4±9.0 years) from our outpatient clinic and 133 MRIs from 131 age-matched healthy controls (mean age=70.3±9.1 years) from a local community dwelling study. Patients and controls were scanned using a consistent MRI protocol performed at the same scanner at 3 Tesla (Siemens TimTrio) including a T1-weighted MPRAGE sequence (1mm isotropic resolution) and a spoiled FLASH sequence (0.9x0.9x2mm³, TR/TE_1=35/4.92ms, 6 echoes, 4.92ms echo spacing, 64 slices). The AD data was randomly split up into 178 training, 37 validation and 37 test scans and the normal control data was randomly split up into 95 training, 19 validation and 19 test scans, creating 1 partition.

Preprocessing: Binary brain masks from each subject were obtained using FSL-SIENAX [8] and subsequently used for brain extraction (BET) to isolate the brain tissue from the skull. R2* maps were calculated voxelwise using a numerical correction model [9] and nonlinearly registered to the 1mm MNI152 template using FSL fnirt [10].

Classifier network: We utilized a 3D classifier network, combining a single convolutional layer (kernel 8x8x8, 8 channels) with a down-convolutional layer (kernel 8x8x8, 8 channels, striding 2x2x2) as the main building block. The overall network stacks 4 of these main building blocks followed by two fully connected layers (16 and 2 units) (Figure 1). Each layer is followed by a Rectified Linear Unit (ReLU) nonlinearity, except for the output layer where a Softmax activation is applied.

CC BY

Published by Verlag der TU Graz
Graz University of Technology

DOI: 10.3217/978-3-85125-826-4-16

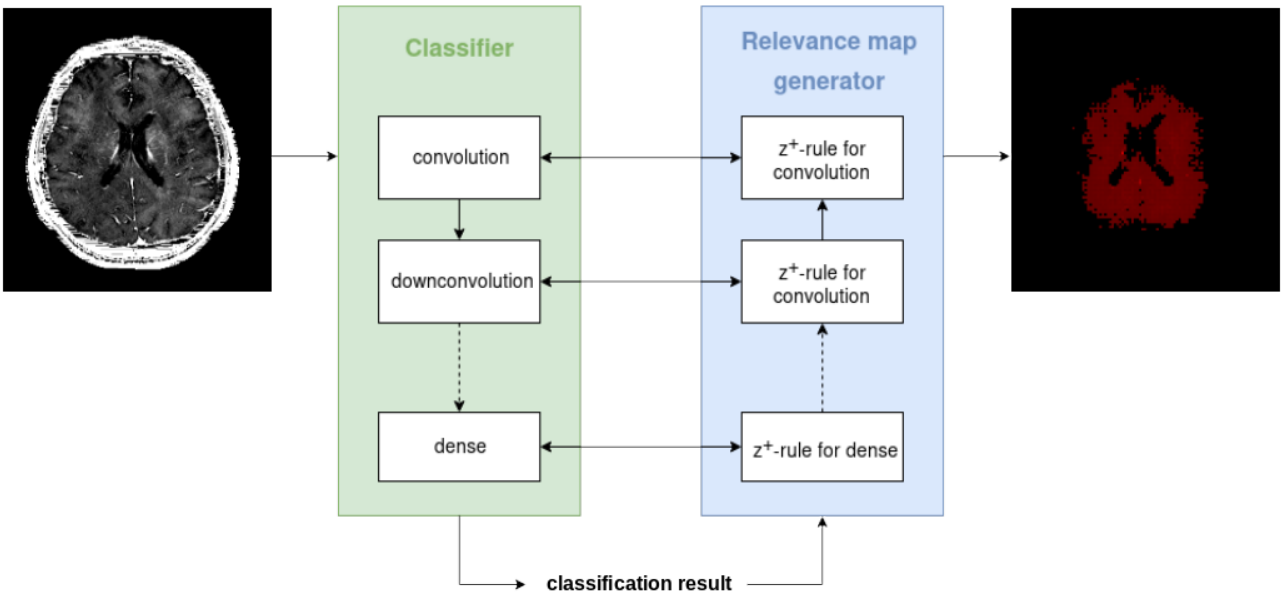

Figure 1: Overview of the used relevance-guided classifier network. We extend a default classifier network (green) with a relevance map generator (blue). For every layer of the classifier network a corresponding relevance redistribution layer is added to the generator network. The output relevance map of the generator has the same resolution as the R2* map used as input for the classifier and allows to guide the training of the classifier network by adding a term that sums relevance values outside a given brain mask to the categorical cross entropy loss.

Relevance-guided classifier network: To focus the network on "relevant features", we propose a relevance-guided network architecture that extends the given classifier network with a relevance map generator (Figure 1). To this end we implemented the deep Taylor decomposition (z^+-rule) [11] to generate the relevance maps of each input image depending on the classifier's current parameters.

Training: We trained models for two differently preprocessed types of R2*:

- R2* images in native subject space
- R2* images nonlinearly registered to the 1mm MNI152 template

For each type we compare the two standard classifier networks (unmasked and masked) with the outcome of our relevance-guided method. Each model was trained using Adam optimizer [12] for 60 epochs with a batch size of 8. The difference in the class sizes was accounted for using a class weighting in the loss function.

Heat map presentation: Besides qualitatively comparing individual heat maps, we compared average heat maps by accumulating the bin contents of each averaged heat map histogram from top until we reached 20 % of all relevance within the heat map overlaid on an MNI152 1mm template.

Results

The resulting balanced classification accuracy between normal controls and AD shows increased performance on the test set for the relevance-guided models (Table 1). However, the obtained relevance maps (Figure 2) show that using unmasked images or brain masking yield highly relevant features for AD/normal control classification at the respective outer boundaries (left and center column). In contrast, relevance-guided training identifies regions within brain tissue, with the highest feature density in the basal ganglia. The corresponding receiver operating characteristics (ROC) curves for all six models are shown in Figure 3.

CC BY
Published by Verlag der TU Graz
Graz University of Technology

DOI: 10.3217/978-3-85125-826-4-16

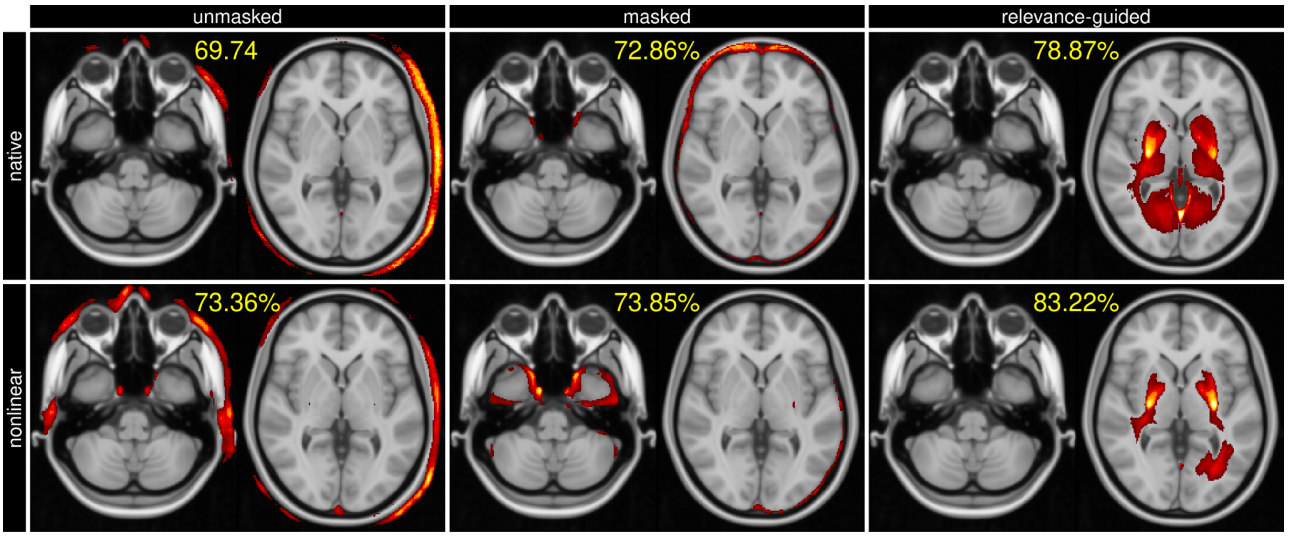

Figure 2: Mean relevance maps (highest relevances in yellow) overlaid on MNI152 template and balanced classification accuracy (percentage) obtained for all six models. Unmasked and masked MRI classifiers obtain relevant image features overwhelmingly from volumetric information (left and center columns). In contrast, the proposed relevance-guided method bases the classifier's decision on deep brain image features, virtually independently of the registration method (right column).

Table 1: Performance (in %) for the different models on the test set. Highest values per column are highlighted in bold.

Class., Classifier; BET, brain extraction; Reg., registration; Bal. acc.; balanced accuracy; Sens, sensitivity; Spec., specificity; AUC, area under the curve of the receiver operating characteristics; CNN, convolutional neural network; RG, relevance-guided

Class.	BET	Reg.	Bal. acc.	Sens.	Spec.	AUC
CNN	no	-	70%	89%	50%	0.70
CNN	yes	-	73%	89%	56%	0.79
CNN	no	nlin	73%	84%	63%	0.77
CNN	yes	nlin	74%	79%	69%	0.79
CNN+ RG	no	-	79%	**92%**	66%	0.80
CNN+ RG	no	nlin	**83%**	79%	**88%**	**0.85**

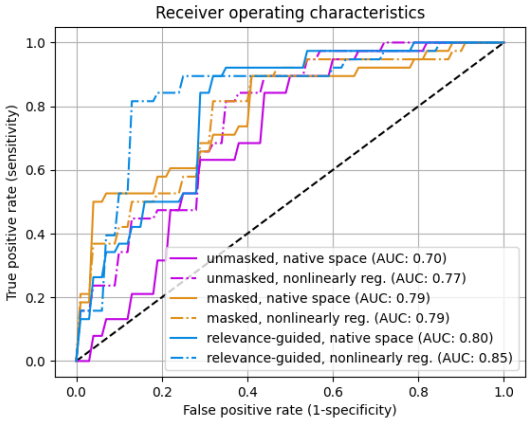

Figure 3: Comparison of receiver operating characteristics for all six configurations. The relevance-guided models (blue) show higher values for the area under the curve (AUC in legend) compared to unmasked (purple) and masked (orange) configurations.

Discussion and Conclusion

In this explorative study we demonstrate that the preprocessing of MR images is crucial for the feature identification by DNNs. While previous work has shown that skull stripping is necessary to avoid identification of features outside the brain, this introduces new edges by the brain mask, which are subsequently used by the DNN for classification. In this context, it was demonstrated that the outcome of brain extraction algorithms can be biased by the patient cohort [13]. In contrast, when using the proposed relevance-guided approach and independently of preprocessing, the regions of highest relevance were found in the basal ganglia. R2* is considered as a measure of iron content [14]. Histological and in-vivo studies [15], [16] have shown that brain iron concentration is higher in these regions in AD patients compared to normal controls. In conclusion, our results are in good agreement with findings from iron mapping studies and strongly support the hypothesis that the relevance-guided approach is minimizing the impact of preprocessing steps such as skull stripping and registration. Additionally,

CC BY
Published by Verlag der TU Graz
Graz University of Technology

DOI: 10.3217/978-3-85125-826-4-16

relevance-guiding forces the feature identification to focus on the parenchyma only and therefore provides more plausible results with higher classification accuracy.

Acknowledgements

This study was funded by the Austrian Science Fund (FWF grant numbers: KLI523, P30134). This research was supported by NVIDIA GPU hardware grants.

References

[1] K. Hammernik *et al.*, “Learning a variational network for reconstruction of accelerated MRI data,” *Magn Reson Med*, vol. 79, no. 6, pp. 3055–3071, Jun. 2018, doi: 10.1002/mrm.26977.

[2] J. Kleesiek *et al.*, “Deep MRI brain extraction: A 3D convolutional neural network for skull stripping,” *Neuroimage*, vol. 129, pp. 460–469, Apr. 2016, doi: 10.1016/j.neuroimage.2016.01.024.

[3] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, Art. no. 7639, Feb. 2017, doi: 10.1038/nature21056.

[4] E. Hosseini-Asl *et al.*, “Alzheimer’s disease diagnostics by a 3D deeply supervised adaptable convolutional network,” *Front Biosci (Landmark Ed)*, vol. 23, pp. 584–596, Jan. 2018, doi: 10.2741/4606.

[5] W. Samek, T. Wiegand, and K.-R. Müller, “Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models,” *ITU Journal: ICT Discoveries*, vol. 1, no. 1, pp. 39–48, 2018.

[6] C. Davatzikos, “Machine learning in neuroimaging: Progress and challenges,” *NeuroImage*, vol. 197, pp. 652–656, Aug. 2019, doi: 10.1016/j.neuroimage.2018.10.003.

[7] C. Tinauer *et al.*, “Relevance-guided Feature Extraction for Alzheimer’s Disease Classification,” presented at the ISMRM 27th Annual Meeting & Exhibition, Montréal, CANADA, May 2019.

[8] S. M. Smith *et al.*, “Accurate, robust, and automated longitudinal and cross-sectional brain change analysis,” *Neuroimage*, vol. 17, no. 1, pp. 479–489, Sep. 2002, doi: 10.1006/nimg.2002.1040.

[9] M. Soellradl *et al.*, “Assessment and correction of macroscopic field variations in 2D spoiled gradient-echo sequences,” *Magn Reson Med*, vol. 84, no. 2, pp. 620–633, Aug. 2020, doi: 10.1002/mrm.28139.

[10] S. M. Smith *et al.*, “Advances in functional and structural MR image analysis and implementation as FSL,” *Neuroimage*, vol. 23 Suppl 1, pp. S208-219, 2004, doi: 10.1016/j.neuroimage.2004.07.051.

[11] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, and K.-R. Müller, “Explaining nonlinear classification decisions with deep Taylor decomposition,” *Pattern Recognition*, vol. 65, pp. 211–222, May 2017, doi: 10.1016/j.patcog.2016.11.008.

[12] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *ICLR*, 2015.

[13] C. Fennema-Notestine *et al.*, “Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: effects of diagnosis, bias correction, and slice location,” *Hum Brain Mapp*, vol. 27, no. 2, pp. 99–113, Feb. 2006, doi: 10.1002/hbm.20161.

[14] S. Ropele and C. Langkammer, “Iron quantification with susceptibility,” *NMR Biomed*, vol. 30, no. 4, Apr. 2017, doi: 10.1002/nbm.3534.

[15] N. Schröder, L. S. Figueiredo, and M. N. M. de Lima, “Role of brain iron accumulation in cognitive dysfunction: evidence from animal models and human studies,” *J Alzheimers Dis*, vol. 34, no. 4, pp. 797–812, 2013, doi: 10.3233/JAD-121996.

[16] C. Ghadery *et al.*, “R2* mapping for brain iron: associations with cognition in normal aging,” *Neurobiol Aging*, vol. 36, no. 2, pp. 925–932, Feb. 2015, doi: 10.1016/j.neurobiolaging.2014.09.013.

CC BY

Published by Verlag der TU Graz
Graz University of Technology