International Journal of Computer Assisted Radiology and Surgery (2024) 19:1469–1472
https://doi.org/10.1007/s11548-024-03186-z

SHORT COMMUNICATION

MRI-degad: toward accurate conversion of gadolinium-enhanced T1w MRIs to non-contrast-enhanced scans using CNNs

Feyi Ogunsanya[1,2] · Alaa Taha[1,3] · Greydon Gilmore[1,3] · Jason Kai[1,2] · Tristan Kuehn[1,3] · Arun Thurairajah[1,4] · Mauricio C. Tenorio[1,3] · Ali R. Khan[1,2,3] · Jonathan C. Lau[1,3,4]

Received: 22 March 2024 / Accepted: 10 May 2024 / Published online: 1 June 2024

Purpose

Acquiring magnetic resonance imaging (MRI) scans with peripherally injected gadolinium-based contrast agent (GBCA) is a crucial step to many neurosurgical interventions [1]. However, it presents challenges to common neuroimaging workflows focused on cortical surface reconstruction, segmentation and brain atlas registration [2], such as fMRIPrep and FreeSurfer [3, 4]. The potential solution of conducting two MRI sessions, with and without contrast, places a strain on healthcare resources. Machine learning (ML) models, particularly 3D convolutional neural networks (CNN), have been increasingly used for image synthesis and segmentation tasks, with a widely used architecture being the 'U-Net' [5]. A previous study by Bottani et al. explored the use of a 3D U-Net to translate contrast-enhanced scans into non-contrast-enhanced T1w brain MRIs [6]. Using segmentation tools derived from the Statistical Parametric Mapping (SPM) software, they found significantly smaller absolute volume differences in gray matter, white matter and cerebrospinal fluid between ground truth non-contrast scans and synthetic non-contrast scans, compared to the absolute differences between ground truth and contrast-enhanced images.

In this work, we employ the commonly used fMRIPrep/FreeSurfer neuroimaging pipeline to characterize biases in morphometric metrics when processing scans with contrast enhancement. We propose *MRI-degad*, a CNN model converting heterogeneous clinical T1w scans with gadolinium enhancement (MRI-gad) to counterpart scans with no contrast (MRI-nogad). We perform a robust validation strategy, leveraging voxel-based metrics and fMRIPrep outcomes.

Methods

We identified 63 patients who underwent stereoelectroencephalography and deep brain stimulation surgery and acquired 1.5-T MRI-gad and MRI-nogad scans. Preprocessing included bias correction, intra-subject rigid registration from MRI-nogad to MRI-gad space and isotropic resampling to 1 mm resolution. Subjects were randomly split into training ($n = 29$), validation ($n = 8$) and testing ($n = 17$) datasets. Using the Medical Open Network for AI (MONAI) framework [7], a 3D U-Net was implemented, herein termed the *MRI-degad* model, which was trained on subject-matched MRI-gad and MRI-nogad scans, with the former as the input and the latter as the target. A random search of the following hyperparameters was used to optimize our model:

- *Patch size* 3D patch sizes of length 16 mm and 32 mm were considered.
- *Batch size* Batches of size 32, 64 and 128 were considered.
- *Learning rate* 0.01, 0.001 and 0.005, 0.0001 were considered for the learning rate.
- *Initial number of filters* The number of channels in the first CNN layer was 16, 32 or 64.
- *Number of Convolutions* The number of convolutions in a single block was 2 or 3.

Our U-Net used three layers, batch normalization and 20% dropout. A mean absolute error (MAE) loss function was used. fMRIPrep was run with a 72-h time limit, extracting completion times, number of cortical surface holes and cortical thickness.

✉ Feyi Ogunsanya
feyi.ogunsanya@gmail.com

1 Robarts Research Institute, Western University, London, ON, Canada

2 Department of Medical Biophysics, Western University, London, ON, Canada

3 School of Biomedical Engineering, Western University, London, ON, Canada

4 Department of Neuroscience, Western University, London, ON, Canada

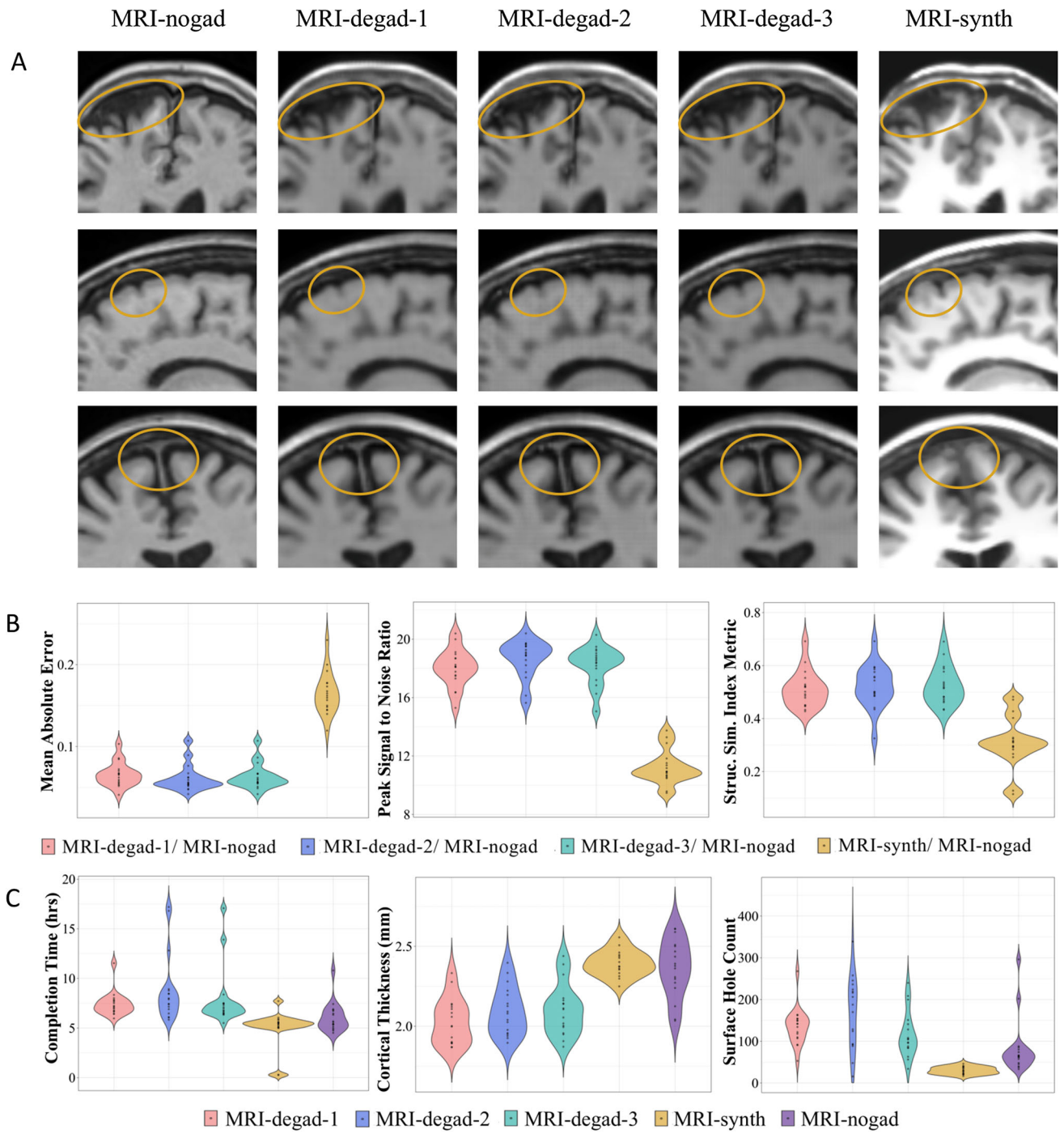

Fig. 1 Visual and quantitative comparisons between MRI-nogad, MRI-degad-1, MRI-degad-2, MRI-degad-3 and MRI-synth. **A** Patch-based inspection. **B** Violin plot of MAE, PSNR and SSIM ($n = 17$). **C** Violin plot of fMRIPrep completion times, cortical thickness and cortical surface holes ($n = 10$)

Table 1 Summary of mean fMRIPrep completion times, cortical surface holes and cortical thickness in our test dataset ($n = 10$)

	Mean fMRIPrep completion time (h)	Number of cortical surface holes	Cortical thickness (mm)
MRI-gad	22.644 ± 19.615	628.2 ± 231.0	2.51 ± 0.16
MRI-nogad	6.121 ± 0.959	57.6 ± 20.1	2.36 ± 0.18
MRI-degad-1	6.999 ± 0.641*	124.2 ± 33.4*/**	2.03 ± 0.14*/**
MRI-degad-2	8.141 ± 3.288*/**	132.4 ± 74.3*/**	2.09 ± 0.15*/**
MRI-degad-3	8.470 ± 3.821*/**	110.6 ± 67.7*/**	2.10 ± 0.16*/**
MRI-synth	5.654 ± 0.766*	30.4 ± 8.2*/**	2.39 ± 0.08*

*Significant difference between output and MRI-gad
**Significant difference between output and MRI-nogad

Test images were compared qualitatively through visual inspection by competent and expert neuroimaging technicians and quantitatively through MAE, structural similarity index metric (SSIM), peak signal to noise ratio (PSNR) and the fMRIPrep metrics. A summary of the preprocessing, model training and validation workflow is in Figure S1 of the supplementary materials. We compare model outputs to those of *SynthSR* (MRI-synth), an out-of-the-box model that resorts to a synthetic generator and segmentation CNN to create modality agnostic outputs [8].

Results

From the total dataset ($n = 63$), all MRI-nogad scans ran to completion, while 18 MRI-gad timed-out. Mean fMRIPrep workflow completion time for MRI-gad and MRI-nogad was 24.88 ± 16.68 h and 5.93 ± 0.75 h, respectively ($n = 45$). Following the cortical reconstruction step within fMRIPrep, the number of surface holes differed significantly between MRI-gad and MRI-nogad scans, with means of 642.16 ± 199.68 and 94.29 ± 48.81, respectively ($n = 45$). We provide a visual comparative evaluation between MRI-gad and MRI-nogad in the fMRIPrep workflow in Figure S2 of the supplementary materials.

We performed downstream analyses on the top three MRI-degads (MRI-degad-1, MRI-degad-2 and MRI-degad-3), whose details can be found in Table S1 of the supplementary materials, and on SynthSR outputs (MRI-synth). These models were selected from visual inspection based on their superior resemblance to ground truth MRI-nogad images, especially in areas with prior gadolinium enhancement. Upon a qualitative inspection of the MRI-degad images, as shown in Fig. 1A, our models are capable of extracting hyperintensities corresponding to the presence of contrast agent in vessels, while maintaining a resemblance of anatomical structures present in ground truth images. In the MRI-synth scan, significant artifacts that appeared to overestimate the boundaries of the sulci were identified and are circled in Fig. 1A.

A summary of mean fMRIPrep metrics for test dataset scans that successfully completed the workflow ($n = 10$) can be found in Table 1 and the violin plot representations can be seen in Fig. 1C. All MRI-degads successfully completed fMRIPrep processing, while two MRI-synth scans did not complete the fMRIPrep workflow, exiting at the registration step. All MRI-degads had significantly lower completion times, number of cortical surface holes and cortical thicknesses compared to MRI-gad. When compared to MRI-nogad, MRI-degad-1 had a completion time that was not statistically different from that of MRI-nogad. MRI-synth scans also showed no significant difference to MRI-nogad in terms of completion time and cortical thickness values.

Violin plots of mean MAE, SSIM and PSNR are shown in Fig. 1B. Against ground truth MRI-nogad, MRI-degads had a lower MAE, a higher PSNR and a higher SSIM compared to MRI-synth.

Conclusion

In this work, we noted that applying fMRIPrep to MRI-gad results in extended processing and elevated failure rates. Additionally, significant differences in cortical thickness and the number of cortical surface holes were observed when comparing MRI-gad to MRI-nogad. In the FreeSurfer pipeline, the projection of a scan onto a surface map involves binary closing operations, typically enhancing the inclusion of sulci in the mask, yielding a smoother output. However, the presence of gadolinium may disrupt this process, potentially affecting the accuracy of the hole-filling step, leading to a higher number of cortical surface holes. The higher fMRIPrep completion time in MRI-gad furthermore suggests an extended hole correction process and the significantly increased cortical thickness in MRI-gad indicates an overestimation of the cerebral cortical layer in the presence of gadolinium.

One of the MRI-degad models featured processing times with no statistical difference from MRI-nogad and overall, our workflow completion success rate was more akin to

MRI-nogad compared to MRI-synth. This can perhaps be attributed to its high image quality and the structural similarity between MRI-degad and MRI-nogad. Validation through visual inspection and voxel-based metrics, where MRI-degad outperformed MRI-synth, confirmed these findings. The two workflow failures in the MRI-synth group were likely caused by image distortions.

MRI-degad exhibited some limitations, including disparities in cortical thickness and cortical surface holes compared to MRI-nogad. To address these discrepancies, ongoing efforts will involve model optimization, dataset refinement, exploration of alternative models such as generative adversarial networks and stable diffusion. Additionally, incorporating more radiologists and neurosurgeons to assess image quality and evaluating the effect of MRI-degad scans on clinical diagnosis and surgical planning remain a critical next step. Our tool holds promise for clinicians and researchers seeking consistent neuroimaging analyses, while reducing MRI healthcare costs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11548-024-03186-z.

Author contributions Lau, Khan, Ogunsanya and Taha contributed to conception and design. Ogunsanya, Taha and Gilmore were involved in acquisition and preparation of data and manuscript. All authors contributed to critical review.

Funding Funding was provided by NSERC Discovery Grants (RGPIN-2023-05558 and RGPIN-2023-05562).

Availability of data and materials All raw and derivative data used in this study will be openly released upon publication of this work.

Declarations

Conflict of interest None.

Ethics approval Human Subject Research Ethics Board office at Western University for data collection (REB# 108952) and open release (REB# 109045).

Consent to participate Patients signed consent forms for undergoing clinical imaging and open release of de-identified study data.

Consent for publication The authors consent to publication of this work.

References

1. Lau JC, MacDougall KW, Arango MF, Peters TM, Parrent AG, Khan AR (2017) Ultra-high field template-assisted target selection for deep brain stimulation surgery. World Neurosurg 103:531–537. https://doi.org/10.1016/j.wneu.2017.04.043
2. Lie IA, Kerklingh E, Wesnes K, Nederpelt DRV, Brouwer I, Torkildsen Ø, Myhr K-M, Barkhof F, Bø L, Vrenken H (2022) The effect of gadolinium-based contrast-agents on automated brain atrophy measurements by FreeSurfer in patients with multiple sclerosis. Eur Radiol 32(5):3576–87. https://doi.org/10.1007/s00330-021-08405-8
3. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ (2019) fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods 16(1):111–116. https://doi.org/10.1038/s41592-018-0235-4
4. Fischl B (2012) Freesurfer. Neuroimage 62(2):774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021
5. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science, pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
6. Bottani S, Thibeau-Sutre E, Maire A, Ströer S, Dormont D et al (2022) Homogenization of brain MRI from a clinical data warehouse using contrast-enhanced to non-contrast-enhanced image translation with u-net derived models. In: SPIE medical imaging 2022: image processing, San Diego, United States, pp 576–582. https://doi.org/10.1117/12.2608565
7. Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y, Murrey B, et al (2022) Monai: an open-source framework for deep learning in healthcare. https://doi.org/10.48550/arXiv.2211.02701
8. Iglesias JE, Billot B, Balbastre Y, Magdamo C, Arnold SE, Das S, Edlow BL, Alexander DC, Golland P, Fischl B (2023) SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry. Sci Adv. https://doi.org/10.1126/sciadv.add3607