Magnetic Resonance Imaging 59 (2019) 143–152

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Technical note

Registration-based image enhancement improves multi-atlas segmentation of the thalamic nuclei and hippocampal subfields

Shunxing Bao[a,*], Camilo Bermudez[c], Yuankai Huo[b], Prasanna Parvathaneni[b], William Rodriguez[b], Susan M. Resnick[f], Pierre-François D'Haese[a,b,e], Maureen McHugo[d], Stephan Heckers[d], Benoit M. Dawant[a,b,c], Ilwoo Lyu[b], Bennett A. Landman[a,b,c,d]

[a] Computer Science, Vanderbilt University, Nashville, TN, United States of America
[b] Electrical Engineering, Vanderbilt University, Nashville, TN, United States of America
[c] Biomedical Engineering, Vanderbilt University, Nashville, TN, United States of America
[d] Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, United States of America
[e] Department of Neurological Surgery, Vanderbilt University Medical Center, Nashville, TN, United States of America
[f] Laboratory of Behavioral Neuroscience, National Institute on Aging, MD, United States of America

ARTICLE INFO

Keywords:
Non-rigid registration
Deep brain structure
Big data

ABSTRACT

Magnetic resonance imaging (MRI) is an important tool for analysis of deep brain grey matter structures. However, analysis of these structures is limited due to low intensity contrast typically found in whole brain imaging protocols. Herein, we propose a big data registration-enhancement (BDRE) technique to augment the contrast of deep brain structures using an efficient large-scale non-rigid registration strategy. Direct validation is problematic given a lack of ground truth data. Rather, we validate the usefulness and impact of BDRE for multi-atlas (MA) segmentation on two sets of structures of clinical interest: the thalamic nuclei and hippocampal subfields. The experimental design compares algorithms using T1-weighted 3 T MRI for both structures (and additional 7 T MRI for the thalamic nuclei) with an algorithm using BDRE. As baseline comparisons, a recent denoising (DN) technique and a super-resolution (SR) method are used to preprocess the original 3 T MRI. The performance of each MA segmentation is evaluated by the Dice similarity coefficient (DSC). BDRE significantly improves mean segmentation accuracy over all methods tested for both thalamic nuclei (3 T imaging: 9.1%; 7 T imaging: 15.6%; DN: 6.9%; SR: 16.2%) and hippocampal subfields (3 T T1 only: 8.7%; DN: 8.4%; SR: 8.6%). We also present DSC performance for each thalamic nucleus and hippocampal subfield and show that BDRE can help MA segmentation for individual thalamic nuclei and hippocampal subfields. This work will enable large-scale analysis of clinically relevant deep brain structures from commonly acquired T1 images.

1. Introduction

Brain MRI is an essential tool for analyzing deep brain structures. Visual inspection remains as the gold standard for identifying anatomical region boundaries in the brain. However, low contrast, spatial proximity of surrounding structures, and low resolution (LR) may inhibit clear visual assessment of deep brain structures [1]. Various post-acquisition image enhancement strategies have been proposed to address these issues. Traditional image enhancement employs a low-pass filter for denoising. Liu et al. propose an enhanced non-local mean using Gaussian filter to reduce the disturbance of the noise [2]. Manjón et al. use a non-local principal component analysis (PCA) thresholding within a rotationally invariant non-local mean filter, in which the local noise level is automatically determined in the image to make noise estimation and denoising [3]. A graph-based redundant wavelet transform approach is shown in [4] to represent MR images sparsely. Another image enhancement traditional research topic is to utilize low resolution image to reconstruct high resolution image. In [5], Rousseau proposes the use of anatomical inter-modality priors based on a physical model of image acquisition as a high resolution image reference to improve the resolution of the LR image. In [6], Bahrami et al., propose a hierarchical re-construction via group sparsity in a novel multi-level canonical correlation analysis (CCA) space, to improve the quality of 3 T MR image to be 7 T-like MRI. Recently, a trend has developed for deep learning for medical image enhancement area. Yang et al. use an artificial neural network to lower the bound of the variance/resolution

* Corresponding author.
E-mail address: shunxing.bao@vanderbilt.edu (S. Bao).

https://doi.org/10.1016/j.mri.2019.03.014
Received 24 August 2018; Received in revised form 9 March 2019; Accepted 13 March 2019

tradeoff and thus to decrease the size of the unachievable region in the variance/resolution plane of PET modality [7]. Oktay et al. incorporate anatomical prior knowledge into a convolutional neural network to enhance cardiac image [8]. Yet, none of the above works intrinsically focus on MRI deep brain enhancement.

Our inspiration comes from a recent work [9] that can perform averaging of thousands of images (pre-processed affinely into MNI space) in near real-time (< 3 min). The resulting image reveals enhanced contrast in some brain structures. When considering the generation of MNI average brain stereotaxic registration model [10], we posit that it is useful to use non-rigid big data averaging. However, such an approach has been computationally infeasible as a single reasonable non-rigid image registration would take 2–4 h; hence, registering a reference population would take about two months. Recently, ultra-fast image registration methods [11] have been developed that enable a 50–100-fold acceleration of the registration process. It is now feasible to routinely perform these computations. Herein, we propose big data registration-enhancement image (BDRE) to improve the intensity contrast of deep brain structures. Rather than parametric models of anatomy or specific sets of well-characterized atlases, we propose to use vast archives of unlabeled data (e.g., “big data”) to capture and enhance image features.

Briefly, our approach registers thousands of datasets to a single target image using non-rigid image registration [11]. We then create a new average image by deforming the registered data to reflect patterns of the target image. This approach reveals substantially increased detail in the deep brain structures. To the best of our knowledge, there is no research has presented a proposal to routinely apply registration of a large number of images (i.e., > 1000) per target image.

On standard 3 T T1 images, deep brain structures are visible, but substructures are not easily identifiable due to low contrast and low image resolution. The thalamic nuclei and hippocampal subfields are two sets of deep brain substructures that play significant roles in the pathophysiology and treatment of neurological and psychiatric disorders including Parkinson's disease, Alzheimer's disease, depression, and schizophrenia [12–14]. However, research studies or clinical practice do not commonly acquire delineation of these structures on MRI which typically requires high resolution 7 T or specialized 3 T images. In Figs. 1 and 2, we present examples of images specialized for identification of thalamic nuclei (Fig. 1: 7 T T1) and hippocampal subfields (Fig. 2: 3 T T2) along with reference segmentation. The appearance of these structures on standard 3 T T1-weighted images for both structures is also presented. As can be observed, the contrast of the thalamic nuclei and hippocampal subfields can be improved on standard 3 T T1 images by de-noising (DN), super resolution (SR), and BDRE. Because labeling of the thalamic nuclei and hippocampal subfields requires another imaging modality (i.e., 7 T MRI and 3 T T2 MRI), it is difficult to determine if contrast within BDREs are spatially well aligned with the original T1 target's structures since the generation of BDRE is based on a more common widely used 3 T scanner. In this work, rather than validate each boundary's correctness in BDRE due to difficulties of direct visualization of our target structures, we investigate the usefulness of BDRE for the multi-atlas (MA) segmentation on those deep brain structures and assume only T1 images are available.

2. Methods

The pipeline for BDRE image generation is as follows in Fig. 3. Briefly, a T1-weighted 3 T image is retrieved from an XNAT system [15] and affinely transformed into MNI space. All local atlases (a total of 5150 T1-weighted images pre-processed into MNI space, and we call them as 5 K atlas images for BDRE) undergo a linear transform to target image (in MNI space) followed by a non-rigid registration to the target image. Next, all registered atlases are averaged in target space with a uniform weight. Finally, the enhanced image is sent back to XNAT system for storage. Traditional registration tools, e.g., Advanced Normalization Tools (ANTS)'s Symmetric Normalization (SyN) algorithm [16], are time-consuming to perform at scale. With recent innovations in image registration efficiency, it is now becoming feasible to evaluate such scenarios. We evaluate non-parametric discrete registration with convex optimization. The tool, Dense Displacement Sampling - wbirLCC (deeds-wbirLCC), uses a fast implementation of a local cross-correlation metric evaluated densely over a discrete search region followed by Gaussian smoothing of the resulting fields, and the fields are coupled back into the similarity measure [11,17]. While the quality of the fast registration is generally similar to a traditional neuroimaging centric tool - ANTs SyN, the registration process is exceptionally speedy and sufficiently flexible to allow for high throughput registrations.

Specifically, we used NiftyReg [18] to perform the rigid, affine transformation to register the target image to MNI-305 template. We then used deeds' transformation tool – linearBCV – to operate linear transformation on 5 K atlases to the target image in MNI space for better compatibility with deeds-wbirLCC [11]. After that, we applied the big data registration tool deeds-wbirLCC to non-rigid registered transformed 5 K atlases to the target image in MNI template space. For averaging 5 K non-rigid registered images, we used ANTS AverageImages tool. All software packages were configured using default parameter settings, except for the deeds-WBIRLCC's regularization Gaussian sigma parameter that we empirically set as 1. We retrieved the 5 K atlas images for big data registration from normal healthy subjects collected by [19].

2.1. BDRE generation time

The proposed method takes < 7 h to perform 5150 non-rigid registrations to the target image in MNI space using ~200 CPU cores. For image averaging to generate one BDRE, we equally split 5150 registered atlases into 103 subgroups, average each group first and then average all intermediate results within 20 min.

In detail, the 5150 atlases images are all pre-rigid affine to MNI space since the target image would also be registered to MNI space. Above step is only a one time cost (based on our experimental cluster, it would take about 50 min). We did not do other preprocessing (e.g., N4 bias correction). If we did so, it would also be a one-time cost. The most time-consuming part is the large volume of non-linear registration, the testbed used for our validation involved 188 cores. At least 2GB RAM was available per CPU core. Each deeds-WBIRLCC takes four cores, thus in total 51 non-rigid registration can run simultaneously. Per each atlas, it would take about 4 min (due to CPU resource competition, the speed is slower than deeds-WBIRLCC on a single image with an empty cluster, which is 50× faster than ANTs around 2 min), and that is how we spent 7 h per one BDRE (5150/51*4 ≈ 404 min). The averaging time depended on the network setup of our cluster. Our experimental grid cluster used a shared network that was deployed with one-gigabit ethernet on each computation nodes and took 20 min to do averaging 5150 atlases. Moreover, if we had multiple BDRE generating, the image averaging can be applied slowly on a single core, the processing time of an on-going BDRE generation would overlap for the time of image averaging.

2.2. Image acquisition information

To explore the usefulness of BDRE, we focused on two deep brain structures: thalamic nuclei and hippocampal subfields. Under IRB approval, we collected MRI data from eight subjects with paired 7 T–3 T T1-weighted MRIs for full thalamus segmentation. The ground truth 46 labels (23 per each hemisphere) were manually delineated by visually fusing the information provided by the multiple image volumes in 7 T space [20]. Table 1 presents the detail of 23 thalamic nuclei structures.

We retrieved paired T1 and T2 MRI datasets for 29 subjects in deidentified form from a first episode psychosis hippocampal imaging

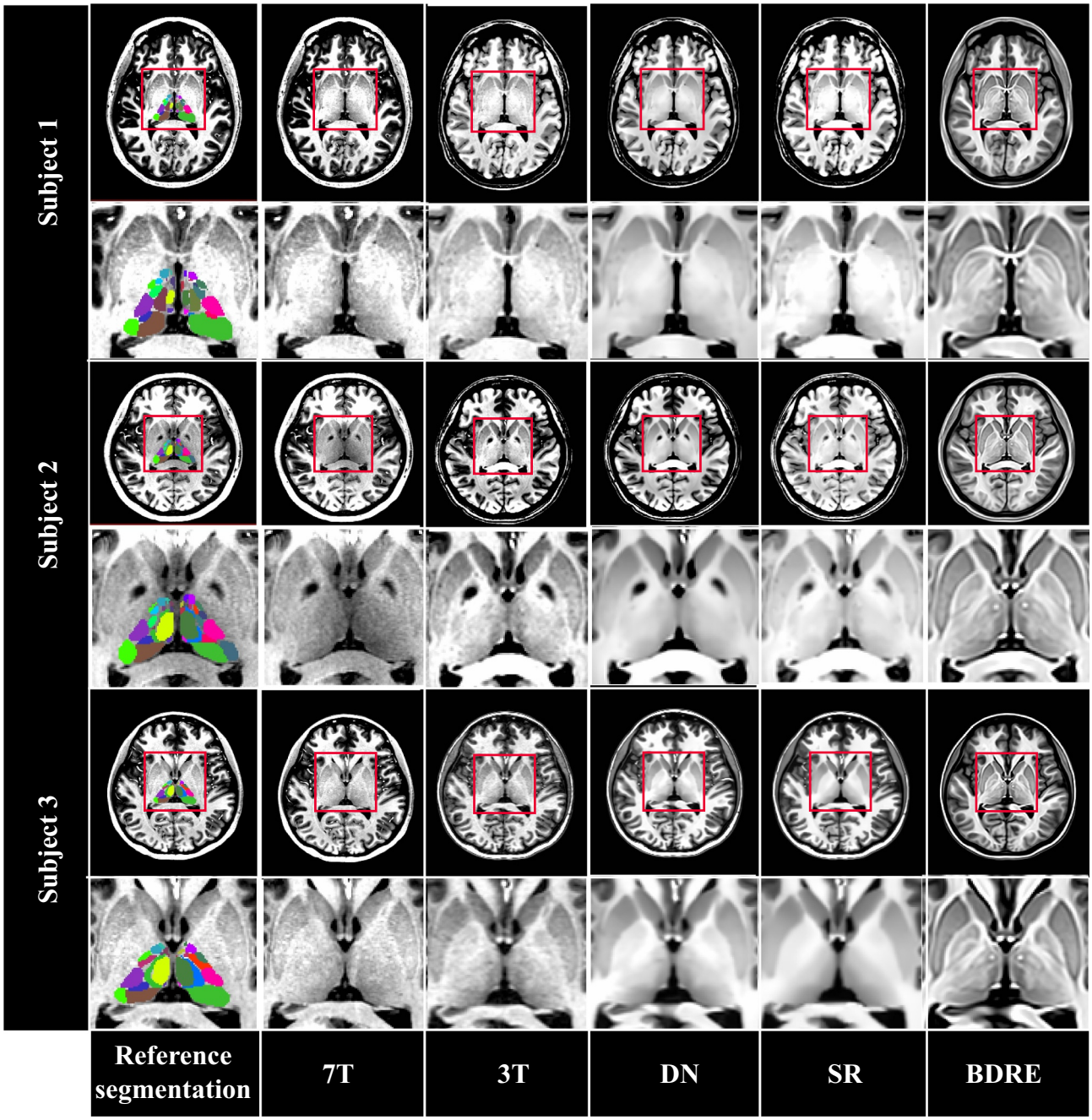

Fig. 1. Three samples of 7 T T1-weighted raw images with ground truth manual thalamic nuclei reference segmentation labels. Paired 3 T T1-weighted MRI images are shown. Each BDRE, DN and SR preprocessed images are generated by the corresponding 3 T image.

dataset. The reference standard hippocampal subfield labels were generated by automatic segmentation of hippocampal subfields (ASHS) pipeline on the T2 MRI (symmetric four labels for left/right side) as reference standards [21]: (1–2) CA1, CA2/3: cornu Ammonis. Here, we collapsed subfields CA2 and CA3 due to lack of distinguishing contrast in MRI and variability [22]; (3) DG: dentate gyrus; (4) SUB: subiculum.

2.3. Preprocessing

For each above 3 T T1-weighted image, we created one BDRE generated image. Meanwhile, we also created one denoising image and one super resolution image by software Baby Brain Toolkit (BTK) default setting. BTK is an Open-Source Toolkit intrinsically for fetal brain MR image processing, while the denoising technique can be applied for any 3D MRI, and super resolution can be applied to adults affected by pathologies (such as Parkinson's disease) leading to involuntary movements during image acquisition [23,24]. All images for MA segmentation pipeline were affinely transformed to the label space ($352 \times 352 \times 350$ voxels and 0.7 mm isotropic voxel size in the axial direction for thalamus; $480 \times 480 \times 26$ voxels with $0.48 \times 0.48 \times 2\,mm^3$ voxel size for Automatic Segmentation of hippocampal Subfields (ASHS) label space, for computation efficiency, we did not upsample atlases and targets.). Then, images were pre-processed by N4 bias field correction for correcting image intensity inhomogeneity. Next, we removed outliers by trimming each image and keep the intensity range within [0.15,99.85] percentiles. Finally, for image normalization, consider the original image I_0 with the mean intensity of μ and standard deviation of the intensity σ, the normalized image I_1, then

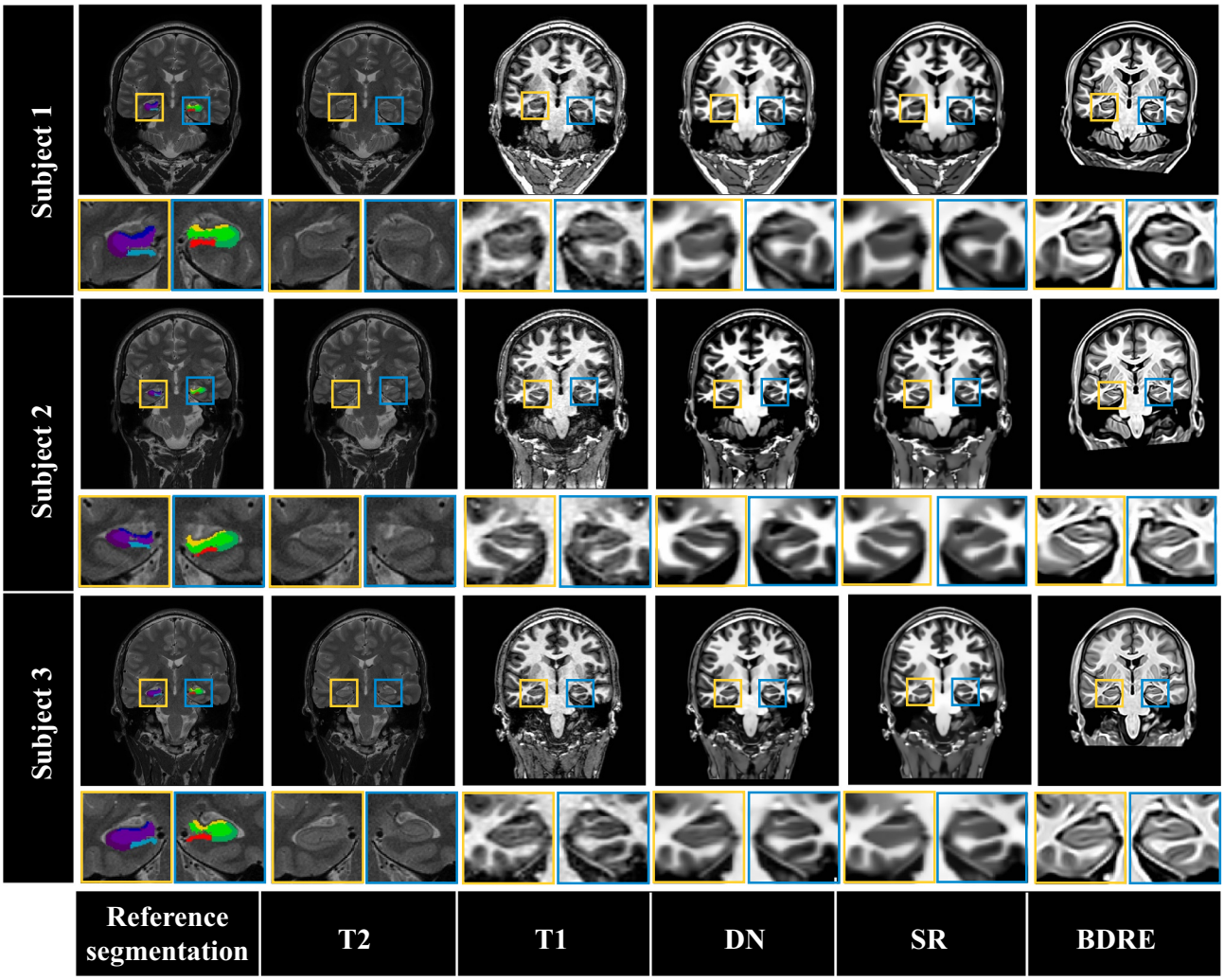

Fig. 2. Three samples of T2-weighted MRI with hippocampal subfields reference segmentation standard labels (generated by automatic segmentation of hippocampal subfields (ASHS) pipeline). Paired T1-weighted MRI images are shown. Each BDRE, DN and SR preprocessed images are generated by the corresponding T1 image.

a commonly used normalization based on z-scores can be described as follows.

$$I_1 = \frac{I_0 - \mu}{\sigma} \tag{1}$$

2.4. Experimental setup

We evaluate the usefulness of BDRE for MA segmentation on the above two deep brain structures. For the thalamic nuclei, we use three parallel MA pipelines as shown in Fig. 2. We use leave-one-out cross-validation, where we choose one subject as a target and the remaining seven are used as atlases. The process is repeated eight times using different subjects in the traverse. For the hippocampal subfields, based on input imaging modality, four parallel pipelines are used for verification. We first randomly chose 15 images from 29 datasets as atlases and remainder of the 14 images for testing. We then reverse the training and testing set for second fold validation.

For MA segmentation, we used NiftyReg's Reg Aladin algorithm for the efficient rigid, affine transformation. And we used ANTs SyN for non-rigid registration with the cross-correlation similarity metric (with radius 2) and a Gaussian regularizer with σ = 3. After registration, we warped reference segmentations from each of the atlases into the target image space. Finally, we used default joint label fusion [25] to make thalamus targets' label. For hippocampal subfield, we empirically set both patch radius for similarity measures and local search radius to 1x1x1 because of a limited number of slices in coronal view.

2.5. Validation

We quantified the performance of eight MAS approaches in terms of Dice Similarity Coefficient (DSC) to assess the contributions of BDRE in MA segmentation, which is the most commonly used accuracy metric in recent studies of image segmentation. Consider X as the segmentation result, Y the truth volume, and $|\cdot|$ the L1 norm operation, DSC is presented as Eq. (2).

$$DSC(X,Y) = \frac{2\,|X \cap Y|}{|X| + |Y|} \tag{2}$$

We also employed mean surface distance (MSD) error and Hausdorff distance (HD) error to characterize how far the surfaces of the segmentation to the truth as Eqs. (3) and (4), where A represents the vertices on the segmentation surface, B the vertices on the truth surface, sup the supremum, inf the infinum, *avg* the average, d the indicator of distance measure.

$$MSD(A,B) = \underset{a\in A}{\text{avg}}\ \underset{b\in B}{\text{inf}}\ d(A,B) \tag{3}$$

$$HD(A,B) = \underset{a\in A}{\text{sup}}\ \underset{b\in B}{\text{inf}}\ d(A,B) \tag{4}$$

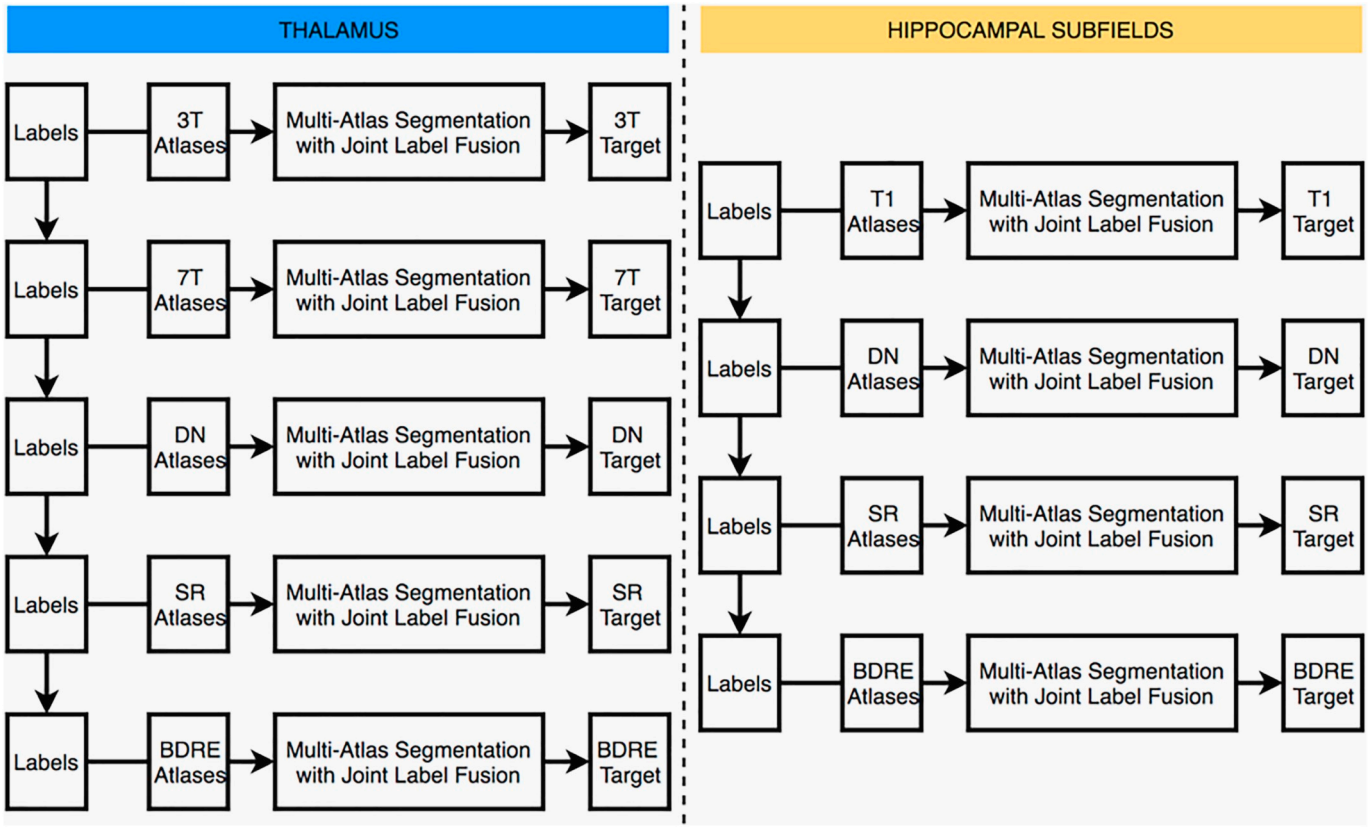

Fig. 3. Workflow for validation of BDRE usefulness for MA. For the thalamic nuclei segmentation, BDRE atlases to BDRE target approach is compared with using 3 T T1, 7 T T1. For the hippocampal subfields, 3 T T1 imaging is used. Each scenarios of BDRE is also compared with T1 images that are preprocessed by DN and SR.

3. Results

All assessments of significance below were performed with the Wilcoxon paired signed rank test. Between each BDRE (as reference) and other corresponding scenarios, we did one paired test is calculated between.

Fig. 4 illustrates qualitative results and Fig. 5 presents quantitative results for full thalamus MA segmentation of using different modalities. Three sample datasets are selected based on overall DSC performance: poor, average and good. Table 2 lists quantitative for overall DSC performance for each imaging modalities in detail. BDRE gets overall DSC of the mean (± standard deviation) of 0.488 (± 0.038), which significantly improves average thalamic nuclei segmentation compared with using 3 T only, 7 T only, DN only and SR only.

Specifically, BDRE MA performs significantly better with 10 out of a total of 46 thalamic nuclei structures when comparing BDRE and 3 T MRI MA segmentation; 11 out of a total of 46 thalamic nuclei structures are significantly better when comparing BDRE and 7 T MRI MA; 9 out of a total of 46 thalamic nuclei structures are significantly better when comparing BDRE and 3 T images preprocessed by DN; and 10 out of a total of 46 thalamic nuclei structures are significantly better when comparing BDRE and 3 T images preprocessed by SR. Meanwhile, we found 3 T MRI MA segmentation outperform than BDRE on the left lateral posterior nucleus (LP) and right anterior ventral dorsal (AVD), and when 3 T MRI preprocessed by DN, it performs better on left lateral pulvinar (PuL), right AVD and right pulvinar (Pu).

Similarly, Fig. 6 illustrates qualitative results and Fig. 7 presents quantitative results hippocampal subfields MA segmentation with three sample datasets. Table 3 lists quantitative for overall DSC performance in detail. BDRE performs overall DSC of the mean (± standard

Table 1
The detail of 23 thalamic nuclei substructures that were manually delineated.

Thalamic nuclei		
Anterior medial nucleus (AM)	Lateral posterior nucleus (LP)	Ventral anterior nucleus (VA)
Anterior ventral dorsal (AVD)	Mediodorsal nucleus (MD)	Ventral lateral anterior nucleus (Vla)
Central medial nucleus (CeM)	Mammillothalamic tract (mtt)	Lateral posterior nucleus (VLp)
Central lateral nucleus (CL)	Parafascicular (Pf)	Ventral medial nucleus (VM)
Centre median (CM)	Pulvinar (Pu)	Ventral posterior inferior nucleus (VPI)
Habenula (Hb)	Anterior pulvinar (PuA)	Ventral posterior lateral nucleus (VPL)
Lateral dorsal nucleus (LD)	Lateral pulvinar (PuL)	Ventral posteromedial nucleus (VPM)
Limitans (Li)	Paraventricular (PV)	

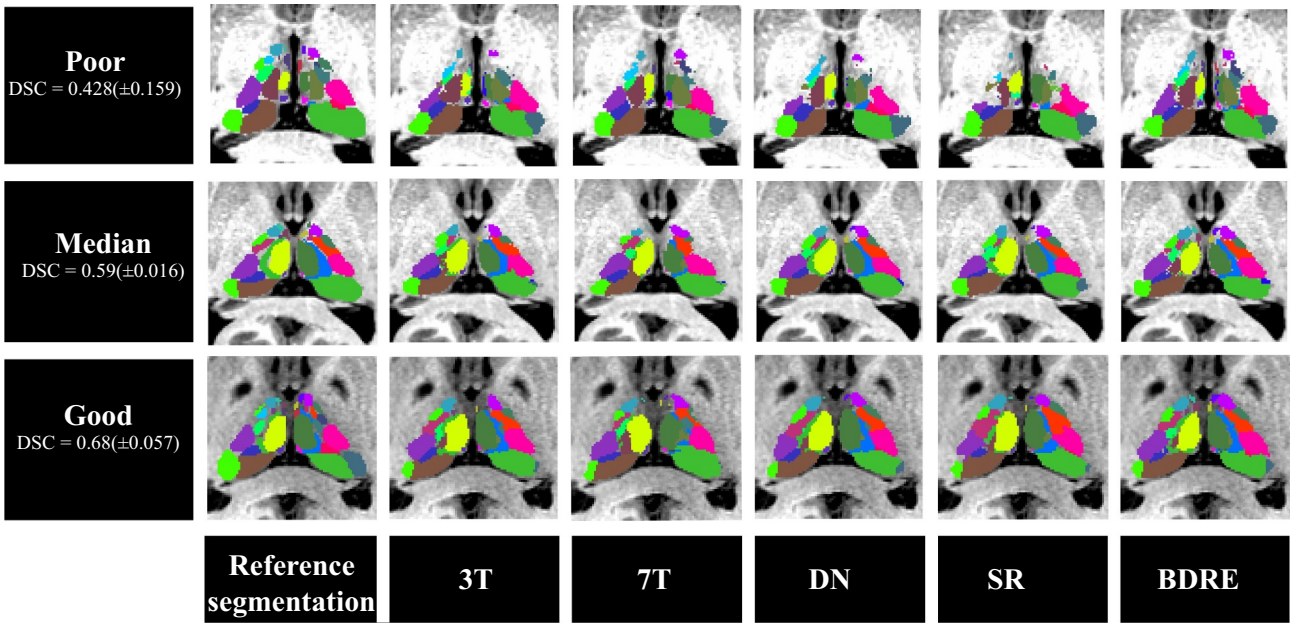

Fig. 4. Qualitative result for thalamic nuclei MA segmentation validation including ground truth reference segmentation, 3 T MRI only, 7 T MRI only, DN only, SR only and BDRE only. Three samples were selected by average overall DSC across each MA segmentation approach using different modalities, which were ordered by worst, median and best mean DSC similarity (± standard deviation).

deviation) of 0.66 (± 0.045) that significantly improves the average hippocampal subfield segmentation compared with using 3 T only, DN only and SR only.

Specifically, BDRE MA performs significantly better with all 8 hippocampal subfields structures when comparing BDRE and T1 MA segmentation; 5 out of total 8 hippocampal subfields structures when comparing BDRE and T1 images preprocessed by DN; and all 8 hippocampal subfields structures when comparing BDRE and T1 images preprocessed by SR.

MSD and HD performance also reveals BDRE MA outperforms the rest of baseline methods (details are provided in supplementary results), especially on both sides of the anterior medical nucleus (AM) and paraventricular (PV) for thalamus scenario, and both sides of CA1 for hippocampal subfields scenario.

4. Discussion & conclusion

Herein, we have presented a technique to use BDRE to enhance the image for deep brain structure MA segmentation. The full structure DSC performance results show that BDRE can significantly improve full thalamus and hippocampal subfields segmentation accuracy. For complete thalamus segmentation, BDRE significantly improves total thalamus segmentation accuracy with 9.1%, 15.6%, 6.9%, and 16.2% improvement over using 3 T imaging only, 7 T imaging only, DN only and SR only. The hippocampal subfields segmentation results in 8.7%, 8.4% and 8.6% improved DSC performance over utilizing T1 weighted modality only, DN only and SR only. We found that there were several structures when just using 3 T MRI or 3 T MRI preprocessed by DN outperformed BDRE. We also found that the overall performance when only using 7 T MRI as atlases and target was worse than the other four approaches. Although the ground truth thalamus nuclei are segmented in 7 T modality, the labels were not depicted by only using one 7 T image but multiple imaging sequences. So it can explain 7 T MA performs even worse.

Due to our limit number of 8 datasets, and the manual reproducibility of full thalamus is also very difficult, more datasets to validate should help us better understand the utilities of BDRE's enhanced thalamic nuclei. Furthermore, our work is not trying to design the best pipeline, (e.g., in [20], Liu et al. directly segment the thalamic nuclei in 3 T MRI using high resolution shape models which are manually delineated in high-field MR images. They start from entire thalamus shape and hierarchically fit joint models to capture the relationship between the thalamus and the internal nuclei and leads to a better result compared with our work). Instead, this works shows the value of the BDRE's contrast.

The BDRE approach has the potential to improve multi-atlas segmentation for different other deep brain structures and imaging protocols and offers enhanced contrast for other applications. We note that the intensity images produced by BDRE exaggerate contrast at boundaries and highlight intensity cues that might not be obvious in the original image. Some of the boundaries may be artefactual (i.e., a by-product of the registration method and/or data archive). Here, we sidestep issues of the validity of these boundaries and focus instead on the utility of the contrasts for driving image processing in situations where a true anatomical label is available via an alternative imaging modality (which is less routinely collected and more time consuming than typical 3 T T1 MRI). For a further understanding of BDRE and future validation, we plan to evaluate different numbers of atlases and/or utilize different time consuming non-rigid registration tools with a smaller number of atlases for computation efficiency in browsing if they impact the quality of deep brain structure contrast.

For this work, our primary purpose is to find the usefulness of BDRE

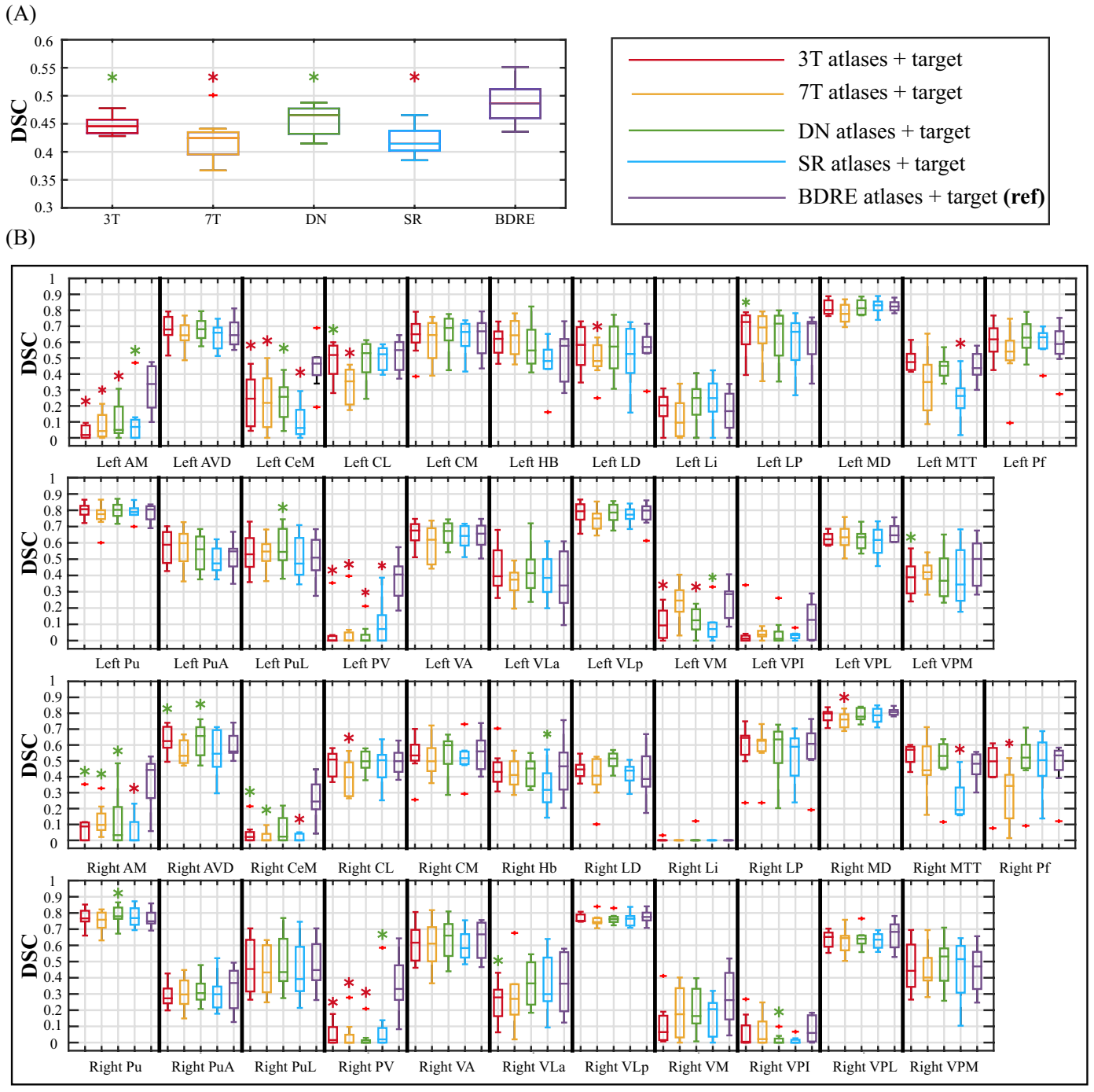

Fig. 5. Quantitative result for thalamic nuclei MA segmentation including 3 T MRI only, 7 T MRI only, DN only, SR only and BDRE only. (A) is the full thalamic nuclei MA segmentation DSC performance for each scenario; (B) is left/right thalamic nuclei MA segmentation DSC on each label based. Each Wilcox paired *t*-test is calculated between BDRE (as reference) and other corresponding modalities. Red "*" represents Wilcox paired *t*-test with $p < 0.01$, and green "*" represents $p < 0.05$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Average performance of thalamic nuclei MA segmentation approaches using different modalities as atlases and target. Average Dice similarity (± standard deviation (SD)) and median DSC are shown. Each Wilcox paired test is calculated between BDRE (as reference) and other corresponding scenarios, where "**" represents Wilcox paired *t*-test with $p < 0.01$, and "*" represents $p < 0.05$.

Atlases + target	3 T*	7 T**	DN *	SR**	BDRE (ref)
DSC(mean ± SD)	0.447 (± 0.017)	0.422 (± 0.04)	0.457 (± 0.028)	0.420 (± 0.028)	0.488 (± 0.038)
DSC(median)	0.446	0.425	0.466	0.414	0.477

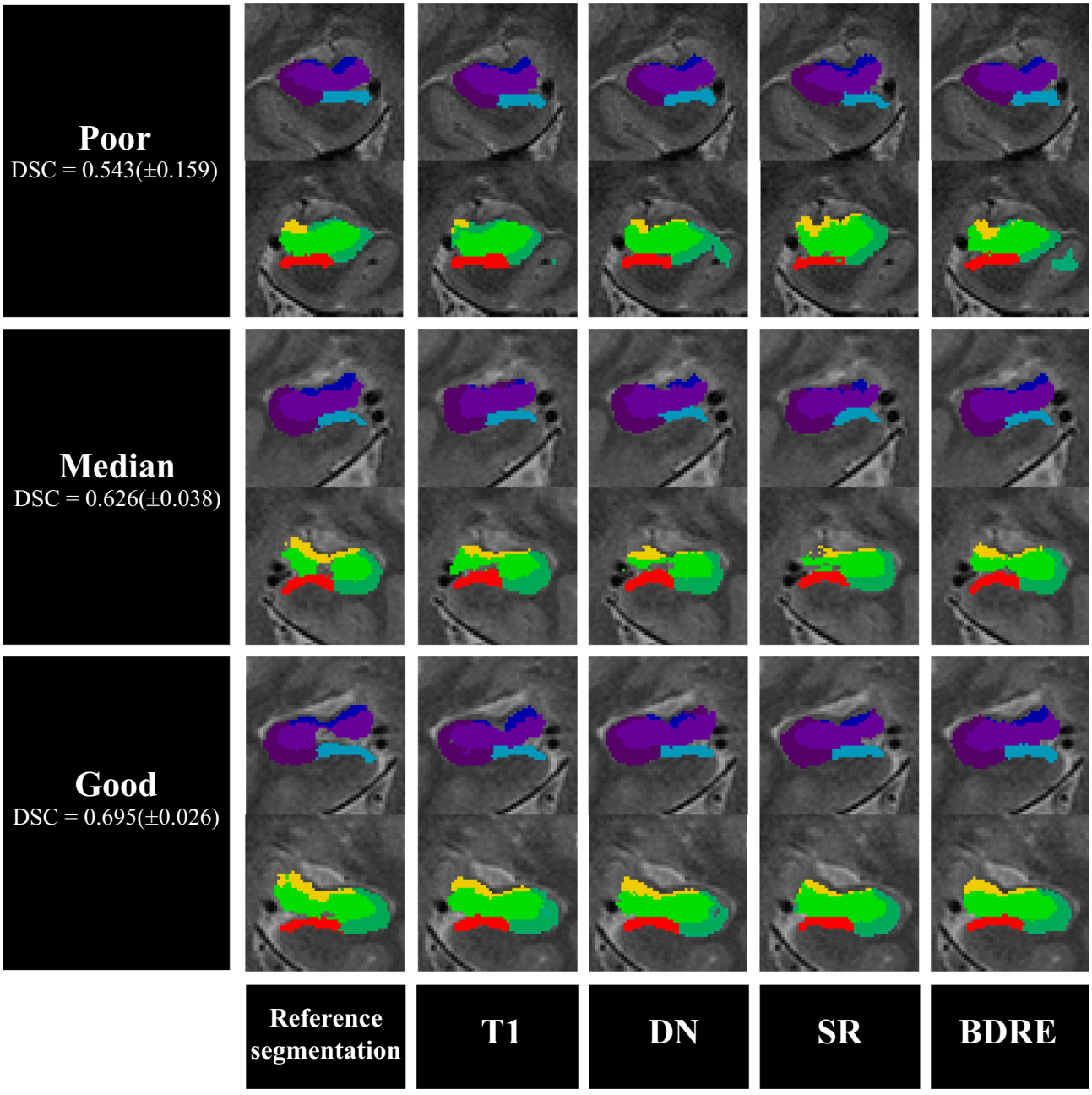

Fig. 6. Qualitative result for hippocampal subfields MA segmentation validation including reference segmentation standards from ASHS pipeline, T1 MRI only, DN only, SR only and BDRE only. Three samples were selected by average overall DSC across each MA segmentation approach using different modalities, which were ordered by worst, median and best mean DSC similarity (± standard deviation).

which is generated by large datasets. Using weighted averaging is a topic that has been highly studied, and tuning the weights is an essential future step to improve BDRE. We also tested on using multimodal approach for atlases (combined with original T1 MRI + BDRE) and the target was either T1 MRI or BDRE, the performance was worse than when only using BDRE as input atlases and target (results not shown).

For the potential "time consuming" issue of generating a BDRE image, 7 h for creating one image is long, but trying fast non-linear on large datasets (i.e. > 1000) is a new approach with interesting result (new contrast) which would take tremendous time when using traditional registration tool (i.e. ANTs SyN). Deep learning super resolution approaches can process an image in real time, but again recent approaches do not focused on deep brain structures. And in this work, we think it is irrelevant to compare BDRE with a deep learning approach currently since we focus on discovering an atlas with a new imaging

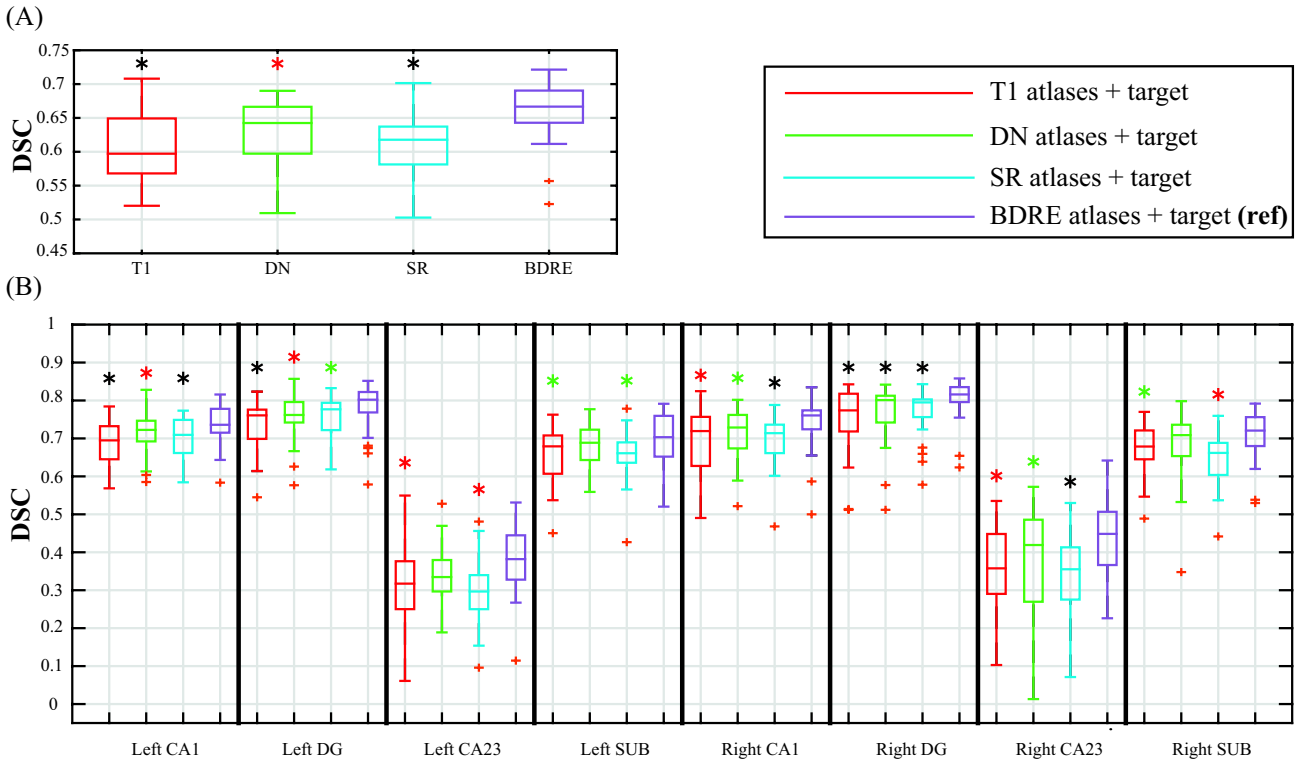

Fig. 7. Quantitative result for hippocampal subfields MA segmentation including T1 MRI only, DN only, SR only and BDRE only. (A) is the hippocampal subfields MA segmentation DSC performance for each scenario; (B) is left/right hippocampal subfields MA segmentation DSC on each label based. Each Wilcox paired *t*-test is calculated between BDRE (as reference) and other corresponding modalities. Black "*" represents Wilcox paired *t*-test with $p < 0.001$, Red "*" represents $p < 0.01$, and green "*" represents $p < 0.05$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Average performance of hippocampal subfields MA segmentation approaches using different modalities as atlases and target. Average Dice similarity (± standard deviation) are shown. Each Wilcox paired test is calculated between BDRE (as reference) and other corresponding scenarios, where "***" represents Wilcox paired t-test with $p < 0.001$, and "**" represents $p < 0.01$.

Atlases + target	3 T***	DN**	SR***	BDRE (ref)
DSC (mean ± SD)	0.607 (± 0.054)	0.628 (± 0.046)	0.608 (± 0.042)	0.660 (± 0.045)
DSC (median)	0.597	0.642	0.618	0.667

modality, which also has potential to train a deep network.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mri.2019.03.014.

Acknowledgements

This work was funded in part by NSF CAREER IIS 1452485, NIH 5R21EY024036, 1R21NS064534, R01NS095291, 2 UL1 TR000445-06, UL1 RR024975-01, support from Intramural Research Program, National Institute on Aging, NIH, and using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. We would like to thank all of the participants that made this work possible. Additionally, we'd like to thank and acknowledge the open access data platforms and data sources that were used for this work, including: XNAT, Neuroimaging Informatics Tool and Resources Clearinghouse (NITRC), 1000 Functional Connectomes Project (fcon-1000), Autism Brain Imaging Data Exchange (ABIDE I), Attention Deficit Hyperactivity Disorder 200 (ADHD-200), Baltimore Longitudinal Study of Aging (BLSA), Education and Brain Sciences Research Lab (EBRL), Information eXtraction from Images (IXI; http://brain-development.org/ixi-dataset), NIH-supported National Database for Autism Research (NDAR, NDAR is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in autism, this manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDAR; NIH MRI Study of Normal Brain Development, dataset identifier: #1151; http://www.bic.mni.mcgill.ca/nihpd/info/participating_centers.html), Nathan Kline Institute Rockland Sample (NKI Rockland), and Open Access Series of Imaging Studies (OASIS). The NIH MRI Study of Normal Brain Development is supported by the National Institute of Child Health and Human Development, the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke— Contract #s N01-HD02-3343, N01-MH9-0002, and N01-NS-9-2314, N01-NS-9-2315, N01-NS-9-2316, N01-NS-9-2317, N01-NS-9-2319 and N01-NS-9-2320. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH. Additional funding sources can be found at: http://fcon_1000.projects.nitrc.org/fcpClassic/FcpTable.html (1000 Functional ConnectomesProject); http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html (ABIDE); http://fcon_1000.projects.nitrc.org/indi/adhd200/ (ADHD-200); http://fcon_1000.projects.nitrc.org/indi/enhanced/ (NKI Rockland); http://www.oasis-brains.org/ (OASIS); NIH-supported National Database for Autism Research (NDAR), NDAR is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in autism. Dataset identifier(s): [NIMH Data Archive Collection ID(s) or NIMH Data Archive Digital Object Identifier (DOI)]. Any opinions, findings, and conclusions or recommendations expressed herein are those of the author(s) and do not necessarily reflect the views of the NIH or NSF.

References

[1] Cho Z-H, Min H-K, Oh H, Han J-Y, Park C-W, Chi J-G, et al. Direct visualization of deep brain stimulation targets in Parkinson disease with the use of 7-tesla magnetic resonance imaging. J Neurosurg 2010;113:639–47.

[2] Liu H, Yang C, Pan N, Song E, Green R. Denoising 3D MR images by the enhanced non-local means filter for Rician noise. Magn Reson Imaging 2010;28:1485–96.

[3] Manjón JV, Coupé P, Buades A. MRI noise estimation and denoising using non-local PCA. Med Image Anal 2015;22:35–47.

[4] Lai Z, Qu X, Liu Y, Guo D, Ye J, Zhan Z, et al. Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform. Med Image Anal 2016;27:93–104.

[5] Rousseau F. A non-local approach for image super-resolution using intermodality priors. Med Image Anal 2010;14:594–605.

[6] Bahrami K, Shi F, Zong X, Shin HW, An H, Shen D. Reconstruction of 7T-like images from 3T MRI. IEEE Trans Med Imaging 2016;35:2085–97.

[7] Yang B, Ying L, Tang J. Artificial neural network enhanced Bayesian PET image reconstruction. IEEE Trans Med Imaging 2018;37(6):1297–309.

[8] Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE Trans Med Imaging 2018;37:384–95.

[9] S. Bao, Y. Huo, P. Parvathaneni, A. J. Plassard, C. Bermudez, Y. Yao, *et al.*, "A data colocation grid framework for big data medical image processing: backend design," in Medical Imaging 2018: Imaging informatics for healthcare, research, and applications, 2018, p. 105790A.

[10] A. C. Evans, D. L. Collins, S. Mills, E. Brown, R. Kelly, and T. M. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in nuclear science symposium and medical imaging conference, 1993., 1993 IEEE Conference Record., 1993, pp. 1813–1817.

[11] Heinrich MP, Papież BW, Schnabel JA, Handels H. Non-parametric discrete registration with convex optimisation. International workshop on biomedical image registration. 2014. p. 51–61.

[12] Kringelbach ML, Jenkinson N, Owen SL, Aziz TZ. Translational principles of deep brain stimulation. Nat Rev Neurosci 2007;8:623.

[13] Pergola G, Selvaggi P, Trizio S, Bertolino A, Blasi G. The role of the thalamus in schizophrenia from a neuroimaging perspective. Neurosci Biobehav Rev 2015;54:57–75.

[14] Small SA, Schobel SA, Buxton RB, Witter MP, Barnes CA. A pathophysiological framework of hippocampal dysfunction in ageing and disease. Nat Rev Neurosci 2011;12:585.

[15] Harrigan RL, Yvernault BC, Boyd BD, Damon SM, Gibney KD, Conrad BN, et al. Vanderbilt University Institute of Imaging Science Center for Computational Imaging XNAT: a multimodal data archive and processing environment. NeuroImage 2016;124:1097–101.

[16] Avants BB, Tustison N, Song G. Advanced normalization tools (ANTS). Insight J 2009;2:1–35.

[17] Heinrich MP, Jenkinson M, Papież BW, Brady M, Schnabel JA. Towards realtime multimodal fusion for image-guided interventions using self-similarities. International conference on medical image computing and computer-assisted intervention. 2013. p. 187–94.

[18] Ourselin S, Roche A, Prima S, Ayache N. Block matching: A general framework to improve robustness of rigid registration of medical images. International conference on medical image computing and computer-assisted intervention. 2000. p. 557–66.

[19] Huo Y, Aboud K, Kang H, E L. Cutting, and B. A. Landman, "Mapping lifetime brain volumetry with covariate-adjusted restricted cubic spline regression from cross-sectional multi-site MRI,". International conference on medical image computing and computer-assisted intervention. 2016. p. 81–8.

[20] Y. Liu, P.-F. D'Haese, A. T. Newton, and B. M. Dawant, "Thalamic nuclei segmentation in clinical 3T T1-weighted images using high-resolution 7T shape models," in Medical imaging 2015: image-guided procedures, robotic interventions, and modeling, 2015, p. 94150E.

[21] Yushkevich PA, Pluta JB, Wang H, Xie L, Ding SL, Gertje EC, et al. Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical structures in mild cognitive impairment. Hum Brain Mapp 2015;36:258–87.

[22] Iglesias JE, Augustinack JC, Nguyen K, Player CM, Player A, Wright M, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. Neuroimage 2015;115:117–37.

[23] Rousseau F, Oubel E, Pontabry J, Schweitzer M, Studholme C, Koob M, et al. BTK: An open-source toolkit for fetal brain MR image processing. Comput Methods Programs Biomed 2013;109:65–73.

[24] Rousseau F, Kim K, Studholme C, Koob M, Dietemann J-L. On super-resolution for fetal brain MRI. International conference on medical image computing and computer-assisted intervention. 2010. p. 355–62.

[25] Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. Multi-atlas segmentation with joint label fusion. IEEE Trans Pattern Anal Mach Intell 2013;35:611–23.