Inferring Maps of Cellular Structures from MRI Signals using Deep Learning

Zifei Liang[1], Choong Heon Lee[1], Tanzil M. Arefin[1], Zijun, Dong[1], Piotr Walczak[2], Song-Hai Shi[3], Florian Knoll[1], Yulin Ge[1], Leslie Ying[4], Jiangyang Zhang[1]

[1] Bernard and Irene Schwartz Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, New York, NY 10016, USA

[2] Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland, Baltimore, MD

[3] Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA;

[4] Departments of Biomedical Engineering, Electrical Engineering, University at Buffalo, the State University of New York, Buffalo, NY, United States

Running title: Connect histology and MRI using deep learning

Corresponding author: Jiangyang Zhang, Ph.D.
Bernard and Irene Schwartz Center for Biomedical Imaging,
Department of Radiology,
New York University School of Medicine,
660 First Ave.
New York, NY 10016, USA
Jiangyang.zhang@nyulangone.org

Summary (150 words):

^{1}H MRI maps brain anatomy and pathology non-invasively through contrasts generated by exploiting inhomogeneities in tissue micro-environments. Inferring histopathological information from MRI findings, however, remains challenging due to the absence of direct links between MRI signals and specific tissue compartments. Here, we show that convolutional neural networks, developed using co-registered multi-contrast MRI and histological data of the mouse brain, can generate virtual histology from MRI results. Our networks provide maps that mirror histological stains for axons and myelin with enhanced specificity compared to existing MRI markers. Furthermore, by introducing random perturbations to the inputs, the relative contribution of each MRI contrast within the networks can be estimated and guide the optimization of MRI acquisition. We anticipate our method to be a starting point for translation of MRI results into easy-to-understand virtual histology for neurobiologists and provide resources for developing novel MRI contrasts.

Introduction:

Magnetic resonance imaging (MRI) is one of a few techniques that can image the brain non-invasively and without ionizing radiation. This advantage is further augmented by a large and still increasing array of versatile tissue contrasts (e.g., diffusion[1], magnetization transfer[2], and manganese enhanced MRI[3]). While the rich tissue contrasts provide unparalleled insights into brain structure and function at the macroscopic level[4], inferring the spatial organization of microscopic structures (e.g. axons and myelin) and their integrity in the brain based on MRI findings remains an ill-posed inverse problem. Aside from the resolution differences, most MRI contrasts are not tied to specific cellular structures and therefore susceptible to confounding factors, especially under pathological conditions. As a result, even though MRI has been widely used to detect certain neuropathology (e.g. ischemic stroke and demyelination), uncertainty arises when the exact pathological events and their severities need to be determined[5-7]. The lack of specificity hinders the direct translation of MRI findings into histopathology and limits its diagnostic value.

Tremendous efforts have been devoted to discover new mechanisms to amplify the affinity of MRI signals to unique physical and chemical properties of target cellular structures and, by doing so, generate new contrasts with improved sensitivity and specificity. Recent progress in multi-modal MRI promises enhanced specificity by integrating multiple MRI contrasts that target distinct aspects of a particular

cellular structure[8]. For example, magnetization transfer (MT), T_2, and diffusion MRI are sensitive to the chemical, magnetic susceptibility, and micro-structural properties of myelin, and combining them together may lead to more specific myelin measurements than individual contrast[9]. Progress in this front, however, has been hindered by the lack of realistic tissue models for inference and ground truth histological data for validation.

The objective of this study is to test whether deep learning networks developed based on co-registered histology and MRI data can bypass the above-mentioned obstacles and enhance our ability to infer the state of key cellular structures from MRI signals. With its capability to bridge data acquired with different modalities, the deep learning framework[10] has clear advantages over existing modeling approaches, as it is data-driven and not limited by underlying assumptions associated with models. As MRI signals are the ensemble average of spins within each pixel, a typical set of three-dimensional (3D) MRI data with millions of pixels thus provides a large number of instances to train the deep learning networks. Through training, the network can learn the formation of MRI signals from cellular structures embedded in co-registered histology. Our results demonstrate that deep convolutional neuronal networks (CNNs) can translate multi-contrast MRI data into maps that mimic histology and offer enhanced specificity for detecting axons and myelin compared to existing MRI-based markers. Furthermore, careful examinations of the networks allow us to measure the relative contribution of individual MRI contrast, which can be used to optimize multi-contrast MRI strategy and evaluate novel imaging contrasts.

Results:

Prediction of auto-fluorescence (AF) images of the mouse brain from MR images using deep learning

We first demonstrated our method using co-registered three-dimensional (3D) MRI and AF data. The MRI data consisted of 67 3D T_2, MT, and diffusion MR images of *ex vivo* C57BL/6 mouse brains (P60, n=6), which were spatially normalized to the Allen Reference Atlas (ARA)[11] (**Fig. 1A**). We then selected 100 AF datasets from the Allen Mouse Brain Connectivity Atlas (AMBCA)[12] with a minimal amount of tracer signals in the forebrain. The contrast in the AF data is not specific to a particular structure, but a majority of hypointense regions colocalize with myelinated white matter tracts[13]. These 3D AF data had already been normalized to the ARA and were down sampled to the resolution of the MRI data (0.06 mm isotropic). The estimated mismatches between the MRI and AF images were mostly within one pixel (**Supplementary Fig. 1A**).

A deep CNN, called MRH-AF, was trained used multiple 3x3 patches from each MRI data (n=6) within the forebrain region (~40,000 patches) as inputs and corresponding patches in the co-registered AF data as targets (Fig. 1B) (Details on the network and training can be found in the Methods section). In order to determine the amount of training data to capture the relationship between the two modalities, the target AF data ranged from randomly selected 60 subjects (i), 6 subjects (ii), single subject data (iii), and 5,000 and 1,000 randomly selected 3x3 patches within a single subject data (iv & v) (**Fig. 1C**). We chose small patches instead of the entire image for training because MRI signals reflect the ensemble-average within a pixel so that the relationship between MRI and histological signals should be local. The 3x3 patch size was shown to accommodate mismatches between MRI and AF data (**Supplementary Fig. 1**).

The performance of MRH-AF was evaluated using the average 3D AF data in the ARA (CCF version 3, average of 1,675 mouse brains)[12] as the reference and MRI data from a separate group of mice (P60, n=4) as inputs. The MRH-AF results with 60-subject AF data as training targets (i) showed good agreement with the reference (**Fig. 1D**) and strong voxel-wise signal correlation (R^2=0.73, $p<0.001$, **Supplementary Fig. 5A**). The agreement was maintained for (ii) and (iii) both visually and quantitatively, as measured by the root mean square errors (RMSE) and structural similarity index SSIM (**Fig. 1D-E**). The specificity to hypo-intense regions in the co-registered reference defined by optimal thresholding was evaluated using receiver operating characteristic (ROC) analysis, and the results (**Fig. 1F)** showed slightly decreased areas under the curve (AUCs) in (iii) compared to (i) (0.937 vs. 0.940) (**Fig. 1F**). The variation in the ROC curves in (iii), caused by the inter-subject variations in AF signals, was relatively small. Further reducing the size of training data (e.g., iv and v) resulted in declined performances (**Fig. 1D-F)**, emphasizing the need for sufficient training data.

The way that MRH-AF in (iii) translated individual 3x3 MR patches into AF signals was visualized in the 2D feature space derived by *t*-Distributed Stochastic Neighbor Embedding (*t*-SNE) analysis[14] (**Fig. 1G**). MR patches that were assigned with hypo-intense AF signals (orange) mostly clustered at the lower right corner, well separated from patches that were assigned with hyper-intense AF signals (turquoise) or near the brain surface (gray). Representative patches from the first two categories showed distinctive signal profiles (**Fig. 1H**). Overall, the result demonstrates that the ability of MRH-AF to translate multi-contrast MRI data into maps the mimic the tissue AF contrast in the AMBCA.

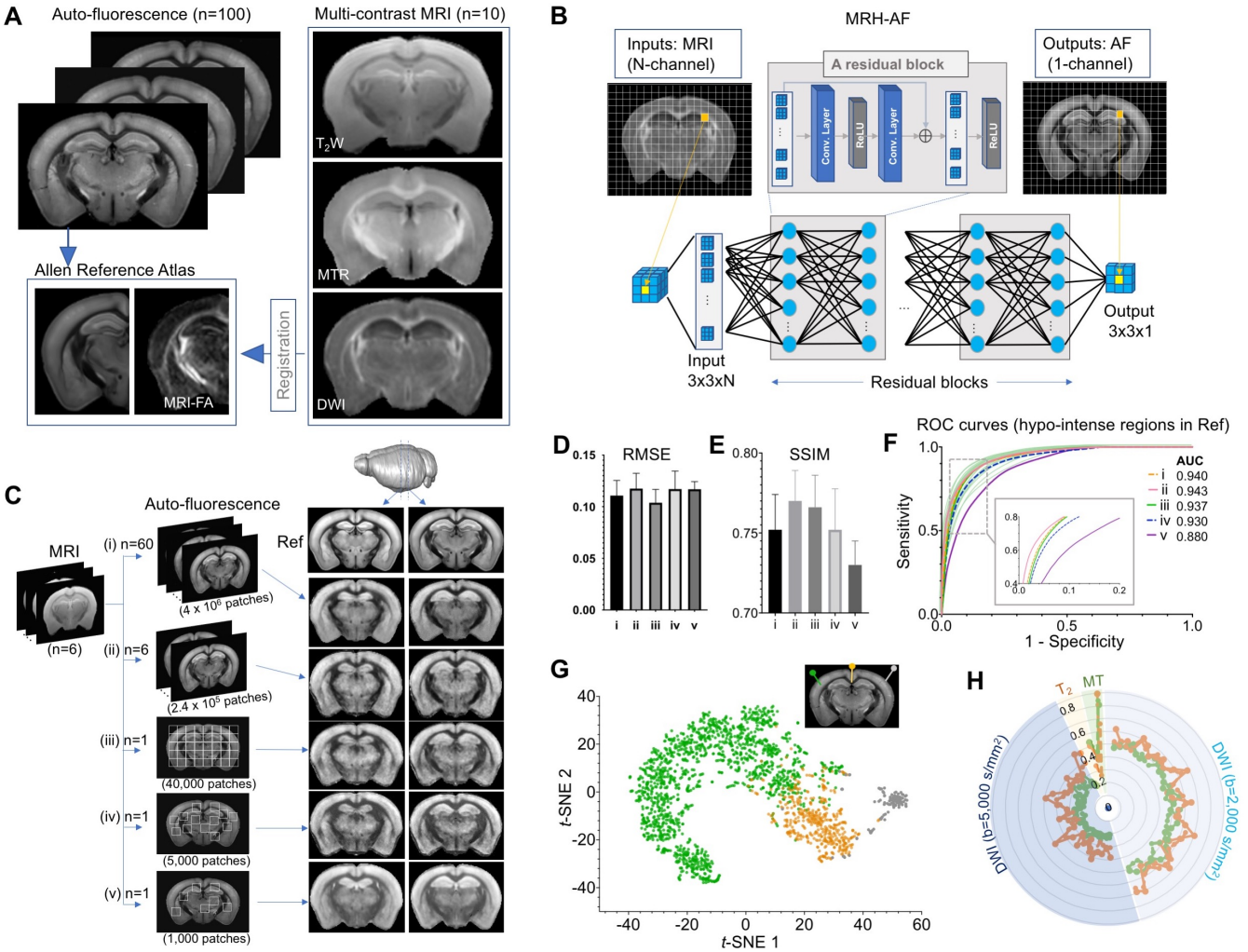

Fig. 1: Connect multi-contrast MRI and auto-fluorescence (AF) data of the mouse brain using deep learning. **A:** T_2-weighted (T_2W), magnetization transfer ratio (MTR), and diffusion weighted images (DWI) were registered to the ARA, from which 100 already registered AF data were selected. **B:** The deep CNN contained 64 layers and was trained using multiple 3x3 MRI patches as inputs and corresponding 3x3 patches from histology as targets. **C:** The CNN was trained using 6-subject MRI data and a wide range of randomly selected AF data (i – v), and the results generated by applying the CNN to a separate set of 4-subject MRI data were shown on the right for visual comparison with the reference (Ref: average AF data from 1,700 subject). **D-E:** Quantitative evaluation of the results in **C** with respect to the reference using RMSE and SSIM. The error bars indicate the standard deviations due to random selections of AF data used to train the network. **F:** The receiver operating characteristic curves of the results in **C** in identifying hypo-intense structures in the reference and their areas under the curve (AUCs). The ROC curves from 25 separate experiments in (iii) (light green) show the variability with respect to the mean ROC curve (dark green) due to inter-subject variations in AF intensity. **G:** The distribution of randomly selected 3x3 MRI patches in the network's 2D feature space, defined using *t*-SNE analysis based on case (iii) in **C**, shows three clusters of patches based on the intensity of their corresponding patches in the reference AF data (turquoise: hyper-intense, orange: hypo-intense; gray: brain surfaces). **H:** MRI signals from two representative patches with hyperintense AF signals (turquoise) and two patches with hypointense AF signals (orange). The orange profiles

show higher diffusion-weighted imaging (DWI) signals and larger oscillation among them than the turquoise profiles (both at b=2,000 s/mm^2 and 5,000 s/mm^2).

Estimation of the relative contribution of individual MRI contrast

Based on the local ensemble average property of MRI signals, we probed the inner-working of MRH-AF by adding random noises to each of the 67 MR images, one at a time, as perturbations to the network[15]. We then measured the effect on network outcomes with respect to noise-free cases (**Fig. 2A**), and this measurement reflected how each MR image influenced the outcome of MRH-AF or its relative contribution in the network. Similar information can also be obtained by training the networks with different subsets of the MRI contrasts and comparing the network predictions, but the perturbation method allows us to probe the network with all the inputs present. We found that adding noises to a few images (e.g., T_2 and MT images) produced noticeably larger effects, in terms of output image quality and the ability of the network to separate different tissue types, than adding a comparable level of noises to other images (**Supplementary Figs. 6-7**). We can accelerate MRI acquisition by prioritizing the acquisition of images or contrasts with higher relative contributions. The top 4, 17, and 38 images based on their contributions accounted for 28%, 50%, and 75% of the total contribution to the final result, respectively (**Fig. 2B**). Results from training the network with the top 38 MR images as inputs showed comparable visual quality (**Fig. 2B-C**) and diagnostic power (**Fig. 2D**) as the results based on the full dataset, and only required 57% the imaging time.

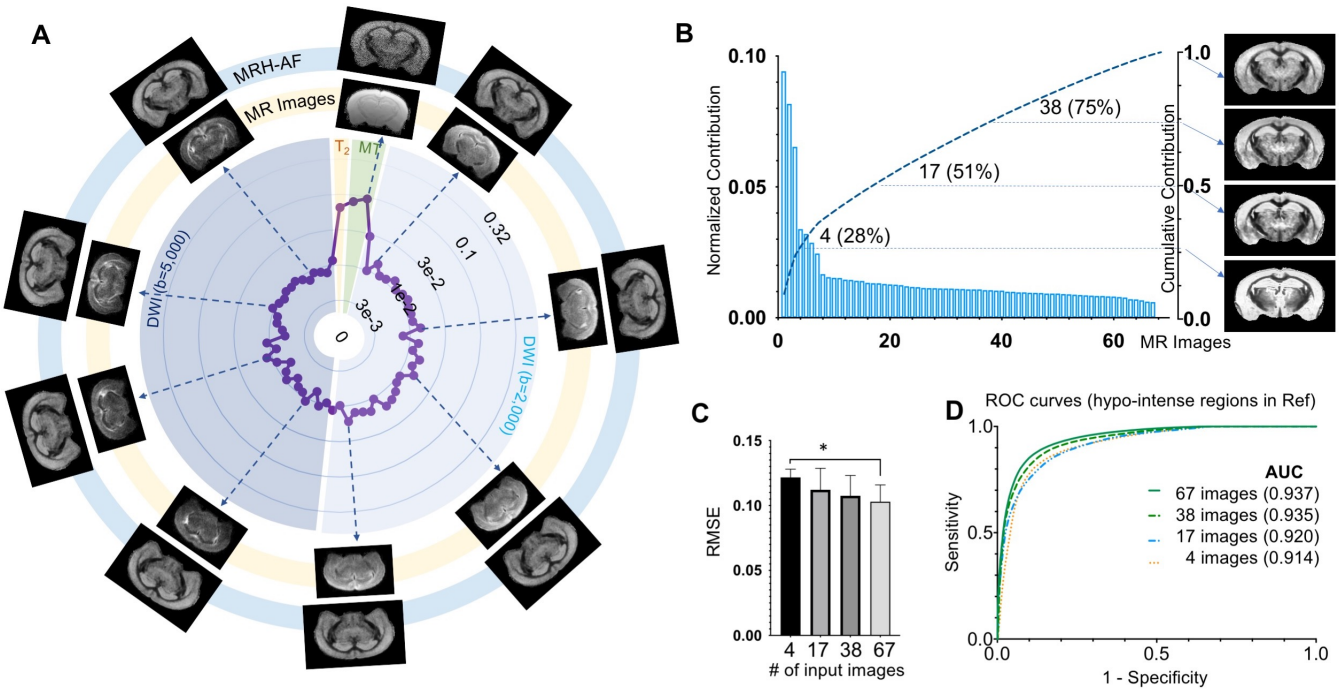

Fig. 2: Estimating how each MR image influence the final results. **A:** Plots of the relative contribution of individual MRI images, normalized by the total contribution of all MR images. Images displayed on the outer ring (light blue, MRH-AF) show the network outcomes after adding 10% random noises to a specific MR image on the inner ring (light yellow). **B:** The relative contributions of all 67 MR images in a descending order and their cumulative contribution. The images on the right show the MRH-AF results with the network trained using the top 4, 17, 38, and all images as inputs. **C:** RMSE measurements of images in **B** (n=4) with respect to the reference AF data. * indicates statistically significant difference ($p<0.05$). **D:** ROC curves of MRH-AF results in **B** and the AUC values.

Use deep learning to predict maps of axon and myelin and enhance specificity

Next, we trained our network using serial histological sections from the Allen mouse brain atlas, immuno-stained for neurofilament (NF) and myelin basic protein (MBP), two commonly used markers for axons and myelin. These images were down sampled, normalized to the ARA (**Supplementary Fig. 2**), and used as both training target and reference here. The MRH results (**Fig. 3A**) showed closer visual congruence with the histological references than commonly used MRI-based markers for axons (fractional anisotropy, FA) and myelin (magnetization transfer ratio, MTR, and radial diffusivity[16], D_R). Even though MRH was trained using coronal sections, it can generate maps along other axes when applied to 3D MRI data **(Fig. 3A)**. The MRH-NF/MBP results also showed strong signal correlations with the reference data ($R^2 = 0.61/0.73$, respectively, **Supplementary Fig. 5B-C**). ROC analyses (**Fig. 3B**) on detecting axon and myelin rich structures demonstrate improved specificity, and *t*-SNE analysis visualize how the two networks separate the patches in MRI data that corresponded to NF and MBP rich structures from the rest (**Fig. 3C**). Analysis on image relative contributions again showed high contributions from T_2W and MT images (**Supplementary Fig. 8**). Applying the MRH-MBP network to MRI data that were collected from dysmyelinated *shiverer* and control mouse brains (n=5/5) generated maps that resembled the MBP-stained histology (**Fig. 3D**). In the corpus callosum, the MRH-MBP results showed similar contrasts between *shiverer* and control moue brains as MTR (**Fig. 3E**).

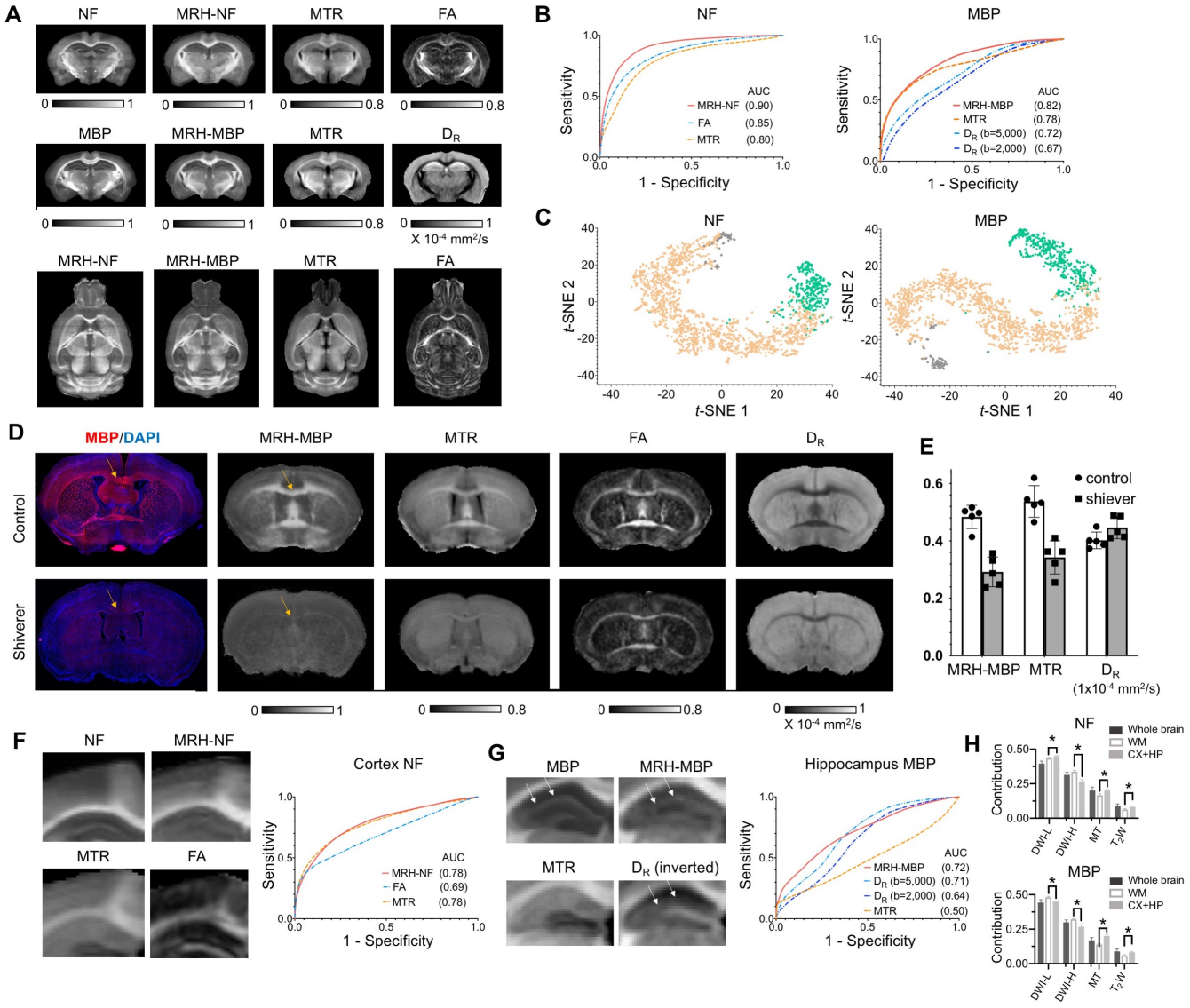

Fig. 3: Inferring maps of neurofilament (NF) and myelin basic protein (MBP) from multi-contrast MRI data. **A:** Comparisons of MRH-NF/MBP results with reference histology and MRI-based markers that are commonly used to characterize axon and myelin in the brain (MTR: magnetization transfer ratio; FA: fractional anisotropy; D_R: radial diffusivity). In the bottom row, horizontal sections of MRH-NF/MBP, MTR and FA maps. **B:** ROC curves of MRH-NF and MRH-MBP show enhanced specificity to their target structures defined in the reference data than MTR, FA, and D_R. Here, D_R values from DWIs with b-values of 2,000 and 5,000 s/mm^2 are examined separately. **C:** The distribution of randomly selected 3x3 MRI patches in the network's 2D feature spaces of MRH-NF and MRH-MBP defined using *t*-SNE analyses. **D:** Representative MRH-MBP results from dysmyelinated *shiverer* and control mouse brains compared with histology and maps of MTR, FA, and D_R. **E:** Differences in MRH-MBP, MTR, and D_R values of the corpus callosum. **F-G:** Enlarged maps of the cortical and hippocampal regions and corresponding ROC curves for detecting NF and MBP-rich structures within each region. In **(G)**, white arrows point to a layer structure in the hippocampus. **H:** Relative contributions of T_2W, MT, diffusion MRI (DWI-L: b=2,000 s/mm^2; DWI-H: b=5,000 s/mm^2) for the whole brain, white matter, and cortex/hippocampus. *: p<0.005.

The MRH results, in combination with the rich structural labels in ARA, provided insights into how the networks balanced multiple MRI contrasts to detect axons and myelin in brain regions with distinct microstructural compositions. In the cortex, MRH-NF and MTR showed similar contrasts and comparable specificities to axons (**Fig. 3D**), whereas, in the whole brain, MTR had lower specificity than MRH-NF and FA (**Fig. 3B**). In the hippocampus, the MRH-MBP results revealed a layer structure in the hippocampus, which was not obvious in the MTR map but visible in the radial diffusivity (D_R) map at b=5,000 s/mm^2 (**Fig. 3E**). In ROC analysis for pixels within the hippocampus, the curve of MRH-MBP closely followed the curve of D_R at b=5,000 s/mm^2, in a departure from the whole brain result (**Fig. 3B**). Relative contributions of T_2W and MT signals were significantly higher in the cortex and hippocampus than in white matter regions for both MRH-NF and MRH-MBP (**Fig. 3F**).

Use deep learning to generate maps that mimic Nissl staining:

We further used MRH to test whether cellularity in the mouse brain can be inferred from diffusion MRI signals, as our previous studies suggest that oscillating gradient spin echo (OGSE)[17] diffusion MRI can generate a contrast similar to Nissl-stained sections in both normal and injured mouse brains[18,19]. We separated the down sampled single subject 3D Nissl data from ARA into two parts. One was used as the training target, and the rest was used as the reference for testing (**Fig. 4A**). The inputs to the so-called MRH-Nissl network included conventional pulsed gradient spin echo (PGSE) and recently developed OGSE diffusion MRI data. In the testing regions, the network that utilized all OGSE and PGSE data as inputs generated maps with good agreement with the ground truth Nissl data (**Fig. 4A**), and showed higher sensitivity and specificity than PGSE (**Fig. 4B**). In the 2D feature space from *t*-SNE analysis (**Fig. 4C**), the patches that correspond to regions with low Nissl signals were separated from patches that correspond to regions with strong Nissl signals. Representative signal profiles from the three categories (**Fig. 4D**) revealed that signals in the high Nissl signal patches decreased as the oscillating frequency increased, whereas the other two types of patches showed no such pattern. Detailed analysis of contrast contribution showed that PGSE and OGSE data contribute equally (**Fig. 4E**). The MRH-Nissl map of the *sas4*$^{-/-}$*p53*$^{-/-}$ mouse brains, which contain a band of heterotopia consists of undifferentiated neurons[20], produced image contrasts that matched Nissl-stained histology (**Fig. 4F**).

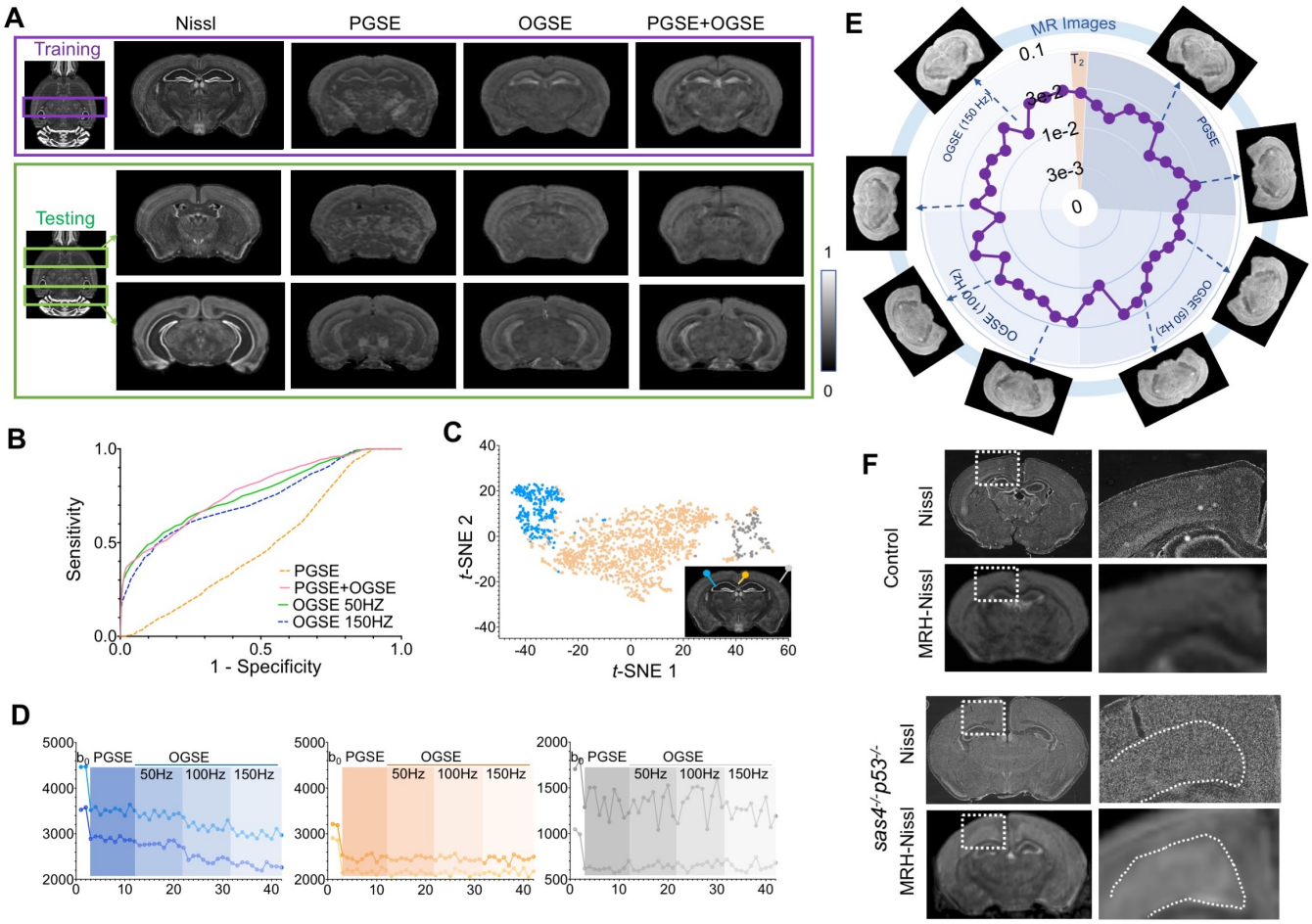

Fig. 4: Generating maps that mimic Nissl stained histology from multi-contrast MRI data. **A:** Comparisons of reference Nissl histology and MRH-Nissl results with PGSE, OGSE, combined PGSE and OGSE diffusion MRI data in both training and testing datasets. The entire datasets consist of PGSE and OGSE data acquired with oscillating frequencies of 50, 100, and 150 Hz, a total of 42 images. **B:** ROC curves of MRH-Nissl show enhanced specificity for structures with high cellularity (strong Nissl staining) when both PGSE and OGSE data were included in the inputs than PGSE only. **C:** The distribution of randomly selected 3x3 MRI patches in the network's 2D feature spaces of MRH-Nissl defined using *t*-SNE analyses. Green and orange dots correspond to regions with high and low cellularity, respectively, and gray dots represent patches on the brain surface. **D:** Representative signal profiles from different groups in **C**. **E:** Relative contributions of PGSE and three OGSE diffusion MRI datasets **F:** Representative MRH-Nissl results from *sas4*$^{-/-}$*p53*$^{-/-}$ and control mouse brains compared with Nissl stained sections. The location of the cortical heterotopia, consists of undifferentiated neurons, is indicated by the dashed lines in the mutant mouse brain image.

Discussions:

The present study focused on inferring maps of key cellular structures in the mouse brain from MRI signals. Previous works on this problem include: new MRI contrasts that capture specific aspects of cellular structures of interest[21-24]; carefully constructed tissue models for MR signals[25-28]; statistical

methods to extract relevant information from multi-contrast MRI[8]; and techniques to register histology and MRI data[29-31] to produce ground truth for validation[32-34]. Here, we built on these efforts by demonstrating that deep learning networks trained by co-registered histological and MRI data can improve our ability to detect target cellular structures.

Previous studies on the relationship between histology and MRI signals focused on correlating histological and MRI markers as co-registered MRI and histological data as well as realistic tissue models are scarce[28,35]. Adopting similar approaches described by recent reports[36-38] on using deep learning to generate histological labels from unlabeled images, we demonstrate a proof of concept of using deep learning to solve the inverse problem of inferring histological information from MRI signals. Even though the resolution of the inferred maps of cellular structures is inevitably limited by the resolution of the input MRI data (~100 μm/pixel) (**Supplementary Fig. 4**), the presented approach has many potential applications in biomedical research involving MRI. It can enhance our ability to accurately map selected cellular structures and their pathology in mouse models of diseases using non-invasive MRI, with contrasts familiar to neurobiologists. In addition, the co-registered histology and MRI dataset provide a testbed for developing new MRI strategies. As it is relatively straightforward to normalize any new MRI data to our 3D multi-contrast MRI data and co-registered histology, the sensitivity and specificity of a new MRI contrast to target cellular structures can be evaluated. With quantitative information on the contributions of different MR contrasts, it is now straightforward to design accurate and efficient multi-contrast MRI strategy.

Perfect co-registration between MRI and histology is highly challenging, as conventional tissue preparations inevitably introduce tissue deformation and damages. In addition, differences in tissue contrasts between histology and MRI also limit the accuracy of registration. Serial two-photon tomography[39] used by the Allen Institute and similar methods allow 3D uniform sampling of the entire brain, which facilitate registration using established registration pipelines[29,40]. We expect recent advances in tissue clearing techniques[41,42] can assist in this aspect once issues such as tissue shrinkage and penetration of antibodies for more target cellular structures are resolved. Remaining mismatches can be accommodated by choosing the appropriate patch size in the network as shown in our results and earlier studies[38].

There are several directions to further improve our work. First, it is important to curate a training dataset that covers a broad spectrum of conditions and, ideally, with MRI and histology data acquired from the same animal. The data included in this study were adult mouse brain and most white matter structures are

myelinated. As a result, the network to predict axons place a substantial weight on MRI contrasts that reflect myelin content (e.g., MT). With the inclusion of unmyelinated embryonic or neonatal mouse brains, we anticipate that the contribution of myelin will be reduced. Inclusion of pathological examples, such as *shiverer* mouse brain data, for training will likely improve our ability to characterize pathological conditions. Secondly, the CNNs constructed in this study involved several common building blocks of deep learning, and new strategies on network architecture design (e.g. [43]) could further enhance the performance. Third, developing similar networks for *in vivo* MRI data and potential clinical application will require additional effort. The MRI data used in this study were collected from post-mortem mouse brain specimens, which are different from *in vivo* mouse brains due to death and chemical fixation. Differences in tissue properties between human and mouse brains also require additional steps. Finally, deep learning cannot replace the good understanding of the physics involved in MRI contrasts and the development of new MRI contrast that targets specific cellular structures.

Methods:

Animals and *ex vivo* MRI. All animal experiments have been approved by the Institute Animal Care and Use Committee at New York University, Memorial Sloan Kettering Cancer Center, and Johns Hopkins University. Adult C57BL/6 mice (P60, n=10, Charles River, Wilmington, MA, USA), *sas4*$^{-/-}$*p53*$^{-/-}$ [20] and littermate controls (n=4/4, P28), *rag2*$^{-/-}$ *shiverer* and littermate controls (n=5/5, P50) were perfusion fixed with 4% paraformaldehyde (PFA) in PBS. The samples were preserved in 4% PFA for 24 hours before transferring to PBS. *Ex vivo* MRI of mouse brain specimens was performed on a horizontal 7 Tesla MR scanner (Bruker Biospin, Billerica, MA, USA) with a triple-axis gradient system. Images were acquired using a quadrature volume excitation coil (72 mm inner diameter) and a receive-only 4-channel phased array cryogenic coil. The specimens were imaged with the skull intact and placed in a syringe filled with Fomblin (perfluorinated polyether, Solvay Specialty Polymers USA, LLC, Alpharetta, GA, USA) to prevent tissue dehydration. Three-dimensional diffusion MRI data were acquired using a modified 3D diffusion-weighted gradient- and spin-echo (DW-GRASE) sequence[44] with the following parameters: echo time (TE)/repetition time (TR) = 30/400ms; two signal averages; field of view (FOV) = 12.8 mm x 10 mm x 18 mm, resolution = 0.1 mm x 0.1 mm x 0.1 mm; two non-diffusion weighted images (b_0s); 30 diffusion encoding directions; and b = 2,000 and 5,000 s/mm^2, total 60 diffusion weighted images (DWIs). Co-registered T_2-weighted and magnetization transfer (MT) MRI data were acquired using a rapid acquisition with relaxation enhancement (RARE) sequence with the same FOV, resolution, and signal

averages as the diffusion MRI acquisition and the following parameters: T_2: TE/TR=50/3000 ms; MT: TE/TR=8/800ms, one baseline image (M_0) and one MT-weighted (M_t) images with offset frequency/power = -3 KHz/20 μT were acquired. The total imaging time was approximately 12 hours for each specimen. For the *sas4*$^{-/-}$*p53*$^{-/-}$ and littermate controls, PGSE and OGSE diffusion MRI data were acquired with the protocol described in [19] and a spatial resolution of 0.1 mm x 0.1 mm x 0.1 mm. All 3D MRI data were interpolated to a numerical resolution of 0.06 mm x 0.06 mm x 0.06 mm to match the resolution of our MRI-based atlas[45].

Magnetization transfer ratio (MTR) images were generated as MTR=(M_0-M_t)/M_0. From the diffusion MRI data, diffusion tensors were calculated using the log-linear fitting method implemented in MRtrix (http://www.mrtrix.org) at each pixel, and maps of mean and radial diffusivities and fractional anisotropy were generated, The mouse brain images were spatial normalized to an ex vivo MRI template[45] using the Large Deformation Diffeomorphic Metric Mapping (LDDMM) method[46] implemented in the DiffeoMap software (www.mristudio.org). The template images had been normalized to the Allen reference atlas using landmark-based image mapping and LDDMM.

Histological data. From the Allen moue brain atlas, single subject 3D Nissl data and 3D AF data (n=100), which were already registered to the ARA space, were down sampled to 0.06 mm isotropic resolution. The *sas4*$^{-/-}$*p53*$^{-/-}$, *rag2*$^{-/-}$ *shiverer* and control mouse brains were cryopreserved and cut into 30 μm coronal sections and processed for Nissl and immunofluorescence. For immunofluorescence, sections were first washed with PBS, blocked with 5% bovine serum albumin, and incubated overnight at 4 °C with primary antibodies: anti-myelin basic protein (AbD Serotec, MCA4095). Sections were rinsed with PBS and incubated with Alexa fluor secondary antibodies (Invitrogen) cover-slipped with anti-fade mounting medium containing DAPI (Vectrolabs, H-1200). Images were obtained and tile-stitched using an inverted microscope (Zeiss, Axio Observer.Z1) equipped with a motorized table.

Registration of MRI and histological data. Group average 3D MRI data in our previously published mouse brain atlas[45] were first spatially normalized to the ARA space. Briefly, fourteen major brain structures (e.g., cortex, hippocampus, striatum) in the atlas MRI data were manually segmented following the structural delineations in the ARA. Voxels that belong to these structures in the MRI and average 3D AF data in the ARA (down sampled to 0.06 mm isotropic resolution) were assigned distinct intensity values, and a diffeomorphic mapping between the discretized atlas MRI and ARA AF data was computed using LDDMM. The mapping was then applied to the original atlas MRI data to generate an MRI template registered to the ARA space. Using dual-channel LDDMM[46] based on tissue contrasts in the average DWI

and FA images and the MRI template, the 3D MRI data acquired in this study were accurately normalized to the ARA space.

NF and MBP stained images of the C57BL/6 mouse brain were downloaded from the ARA reference dataset. Images with major artifacts or tissue loss were excluded. Small tissue tearing and staining artifacts were removed using the Inpainting feature implemented in the photoshop heading brush tool (www.adobe.com), and dark pixels in the ventricles were replaced by the average intensity values of the cortex to match MRI data (Supplementary Fig. 4A). The repaired images were down-sampled to an in-plane resolution of 0.06 mm/pixel. For each 2D histological image, the best-matching MRI section in the MRI template was identified, and a coarse-to-fine alignment from histology to MRI using affine transform and landmark-based image warping tool in ImageJ (https://imageJ.nete/BUnwarpJ). The aligned 2D sections were then assembled into a 3D volume and mapped to the MRI template using LDDMM (between NF/MBP and FA) to further improve the quality of registration.

Evaluation of image resolution. The resolution of MRI, histological images, were evaluated using a parameter-free decorrelation analysis method[47], without the initial edge apodization step.

Design and training of the MRH networks. MRH-networks are constructed using a CNN model with convolutions from the MRI to histology space. The model has previously been reported to perform well on MRI[48] or histological[38] data. The MRH networks were implemented using the deep learning toolbox in Matlab (www.mathworks.com) using the directed acyclic graph architecture. To accommodate residual mismatches between MRI and histological data, the network consistently applied convolutional layers until the end layer that computed the distance loss to the target histology. The number of layers and neurons in each layer were determined empirically to balance performance and the risk of overfitting. We chose 64 hidden layers, each with 64 neurons, which applied a filter size of 3x3 and included a rectified linear unit (ReLu). Most of the hidden layers utilized skip connections to jump over layers to avoid vanishing gradients[49]. The network was initialized with orthogonal random weights and was trained using a stochastic gradient descent optimizer (**Supplementary Fig. 3**) to minimize pixel-level absolute distance loss. Several choices of minibatch sizes were tested and the size was set to 128 to attain a balance between training speed and performance. Stochastic gradient descent with a momentum beta of 0.9 was used for stochastic optimization, with an initial learning rate of 0.1 and a learning rate factor of 0.1. During training, the learning rate was reduced by a factor of 0.1 every 10 epochs. Maximum epoch number was set at 60, but early stopping was employed if the validation set loss did not decrease in 5 epochs. During hyper-parameter tuning, 1000 3x3 patches were randomly held out to as the validation dataset and isolated from

the training dataset. The weights from the epoch with the lowest validation loss were selected for final testing.

***t*-Distributed Stochastic Neighbor Embedding(t-SNE) analysis.** The *t*-SNE cluster was performed using the Matlab t-SNE analysis function on the network prediction based on values in 2,000 randomly selected 3x3 patches in the mouse brain MRI data.

Contribution analysis. Following the perturbation method described by Olden et al.[15], Rician noises were added to one input MR image to the pre-trained MRH networks, and RMSE between the noise contaminated outputs and the original output without noise was recorded. By repeating this procedure for all MR images, the sensitivities of MRH to each of the 67 input MR image or their contributions were obtained.

Evaluate the effect of pixel mismatches. In the experiment that used the DWI and FA data of the mouse brains to train an MRH network, simulated pixel displacements were used to deform the FA data (target), which were perfectly co-registered to the DWI data (inputs), to test the effect of pixel mismatches between input and target data on network prediction. Gaussian random displacement fields were generated for pixels on a 1 mm by 1 mm grid in the coronal plane and propagated to other pixels by B-spline interpolations. The displacement fields followed a Chi distribution with 2 degrees of freedom and were adjusted to match the level of pixel mismatches observed between MRI and histological data.

References

1 Le Bihan, D. Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience* **4**, 469-480, doi:Doi 10.1038/Nrn1119 (2003).

2 Wolff, S. D. & Balaban, R. S. Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. *Magn Reson Med* **10**, 135-144, doi:10.1002/mrm.1910100113 (1989).

3 Koretsky, A. P. & Silva, A. C. Manganese-enhanced magnetic resonance imaging (MEMRI). *NMR in biomedicine* **17**, 527-531, doi:10.1002/nbm.940 (2004).

4 Lerch, J. P. *et al.* Studying neuroanatomy using MRI. *Nature Neuroscience* **20**, 314-326, doi:10.1038/nn.4501 (2017).

5 Purushotham, A. *et al.* Apparent diffusion coefficient threshold for delineation of ischemic core. *Int J Stroke* **10**, 348-353, doi:10.1111/ijs.12068 (2015).

6 Nael, K. *et al.* Multiparametric Magnetic Resonance Imaging for Prediction of Parenchymal Hemorrhage in Acute Ischemic Stroke After Reperfusion Therapy. *Stroke* **48**, 664-670, doi:10.1161/STROKEAHA.116.014343 (2017).

7 Jelescu, I. O. *et al.* In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. *Neuroimage* **132**, 104-114, doi:10.1016/j.neuroimage.2016.02.004 (2016).

8 Mangeat, G., Govindarajan, S. T., Maineroc, C. & Cohen-Adad, J. Multivariate combination of magnetization transfer, T-2* and B0 orientation to study the myelo-architecture of the in vivo human cortex. *Neuroimage* **119**, 89-102, doi:10.1016/j.neuroimage.2015.06.033 (2015).

9 Cercignani, M. & Bouyagoub, S. Brain microstructure by multi-modal MRI: Is the whole greater than the sum of its parts? *Neuroimage* **182**, 117-127, doi:10.1016/j.neuroimage.2017.10.052 (2018).

10 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444, doi:10.1038/nature14539 (2015).

11 Ng, L. *et al.* An anatomic gene expression atlas of the adult mouse brain. *Nat Neurosci* **12**, 356-362, doi:10.1038/nn.2281 (2009).

12 Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207-214, doi:10.1038/nature13186 (2014).

13 Christensen, P. C. *et al.* High-resolution fluorescence microscopy of myelin without exogenous probes. *Neuroimage* **87**, 42-54, doi:10.1016/j.neuroimage.2013.10.050 (2014).

14 van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J Mach Learn Res* **9**, 2579-2605 (2008).

15 Olden, J. D., Joy, M. K. & Death, R. G. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecol Model* **178**, 389-397, doi:10.1016/j.ecolmodel.2004.03.013 (2004).

16 Song, S. K. *et al.* Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage* **17**, 1429-1436, doi:10.1006/nimg.2002.1267 (2002).

17 Does, M. D., Parsons, E. C. & Gore, J. C. Oscillating gradient measurements of water diffusion in normal and globally ischemic rat brain. *Magn Reson Med* **49**, 206-215, doi:10.1002/mrm.10385 (2003).

18 Aggarwal, M., Burnsed, J., Martin, L. J., Northington, F. J. & Zhang, J. Imaging neurodegeneration in the mouse hippocampus after neonatal hypoxia-ischemia using oscillating gradient diffusion MRI. *Magn Reson Med* **72**, 829-840, doi:10.1002/mrm.24956 (2014).

19 Aggarwal, M., Jones, M. V., Calabresi, P. A., Mori, S. & Zhang, J. Probing mouse brain microstructure using oscillating gradient diffusion MRI. *Magn Reson Med* **67**, 98-109, doi:10.1002/mrm.22981 (2012).

20 Insolera, R., Bazzi, H., Shao, W., Anderson, K. V. & Shi, S. H. Cortical neurogenesis in the absence of centrioles. *Nat Neurosci* **17**, 1528-1535, doi:10.1038/nn.3831 (2014).

21 Pautler, R. G., Silva, A. C. & Koretsky, A. P. In vivo neuronal tract tracing using manganese-enhanced magnetic resonance imaging. *Magn Reson Med* **40**, 740-748 (1998).

22 Lee, J. *et al.* The contribution of myelin to magnetic susceptibility-weighted contrasts in high-field MRI of the brain. *Neuroimage* **59**, 3967-3975, doi:10.1016/j.neuroimage.2011.10.076 (2012).

23 Stikov, N. *et al.* In vivo histology of the myelin g-ratio with magnetic resonance imaging. *Neuroimage* **118**, 397-405, doi:10.1016/j.neuroimage.2015.05.023 (2015).

24 Veraart, J. *et al.* Nonivasive quantification of axon radii using diffusion MRI. *Elife* **9**, doi:10.7554/eLife.49855 (2020).

25 Aggarwal, M. *et al.* Spatiotemporal mapping of brain atrophy in mouse models of Huntington's disease using longitudinal in vivo magnetic resonance imaging. *Neuroimage* **60**, 2086-2095, doi:10.1016/j.neuroimage.2012.01.141 (2012).

26 Jespersen, S. N., Kroenke, C. D., Ostergaard, L., Ackerman, J. J. & Yablonskiy, D. A. Modeling dendrite density from magnetic resonance diffusion measurements. *Neuroimage* **34**, 1473-1486, doi:10.1016/j.neuroimage.2006.10.037 (2007).

27 He, X. & Yablonskiy, D. A. Biophysical mechanisms of phase contrast in gradient echo MRI. *Proc Natl Acad Sci U S A* **106**, 13558-13563, doi:10.1073/pnas.0904899106 (2009).

28 Jelescu, I. O. & Budde, M. D. Design and Validation of Diffusion MRI Models of White Matter. *Front Phys-Lausanne* **5**, doi:ARTN 61 10.3389/fphy.2017.00061 (2017).

29 Adler, D. H. *et al.* Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. *Neuroimage* **84**, 505-523, doi:10.1016/j.neuroimage.2013.08.067 (2014).

30 Tward, D. *et al.* Diffeomorphic Registration With Intensity Transformation and Missing Data: Application to 3D Digital Pathology of Alzheimer's Disease. *Front Neurosci* **14**, 52, doi:10.3389/fnins.2020.00052 (2020).

31 Xiong, J., Ren, J., Luo, L. & Horowitz, M. Mapping Histological Slice Sequences to the Allen Mouse Brain Atlas Without 3D Reconstruction. *Front Neuroinform* **12**, 93, doi:10.3389/fninf.2018.00093 (2018).

32 Schilling, K. G. *et al.* Histological validation of diffusion MRI fiber orientation distributions and dispersion. *Neuroimage* **165**, 200-221, doi:10.1016/j.neuroimage.2017.10.046 (2018).

33 Stolp, H. B. *et al.* Voxel-wise comparisons of cellular microstructure and diffusion-MRI in mouse hippocampus using 3D Bridging of Optically-clear histology with Neuroimaging Data (3D-BOND). *Sci Rep* **8**, 4011, doi:10.1038/s41598-018-22295-9 (2018).

34 Chang, E. H. *et al.* The role of myelination in measures of white matter integrity: Combination of diffusion tensor imaging and two-photon microscopy of CLARITY intact brains. *Neuroimage* **147**, 253-261, doi:10.1016/j.neuroimage.2016.11.068 (2017).

35 Novikov, D. S., Kiselev, V. G. & Jespersen, S. N. On modeling. *Magnetic Resonance in Medicine* **79**, 3172-3193, doi:10.1002/mrm.27101 (2018).

36 Christiansen, E. M. *et al.* In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell* **173**, 792-+, doi:10.1016/j.cell.2018.03.040 (2018).

37 Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat Methods* **15**, 917-+, doi:10.1038/s41592-018-0111-2 (2018).

38 Rivenson, Y. *et al.* Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat Biomed Eng* **3**, 466-477, doi:10.1038/s41551-019-0362-y (2019).

39 Ragan, T. *et al.* Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nat Methods* **9**, 255-258, doi:10.1038/nmeth.1854 (2012).

40 Kuan, L. *et al.* Neuroinformatics of the Allen Mouse Brain Connectivity Atlas. *Methods* **73**, 4-17, doi:10.1016/j.ymeth.2014.12.013 (2015).

41 Chung, K. *et al.* Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332-337, doi:10.1038/nature12107 (2013).

42 Renier, N. *et al.* iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* **159**, 896-910, doi:10.1016/j.cell.2014.10.010 (2014).

43 Goodfellow, I. *et al.* in *Conference on Neural Information Processing Systems.* 2672-2680.

44 Wu, D. *et al.* In vivo high-resolution diffusion tensor imaging of the mouse brain. *Neuroimage* **83**, 18-26, doi:10.1016/j.neuroimage.2013.06.012 (2013).

45 Chuang, N. *et al.* An MRI-based atlas and database of the developing mouse brain. *Neuroimage* **54**, 80-89, doi:10.1016/j.neuroimage.2010.07.043 (2011).

46 Ceritoglu, C. *et al.* Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. *NeuroImage* **47**, 618-627, doi:10.1016/j.neuroimage.2009.04.057 (2009).

47 Descloux, A., Grussmayer, K. S. & Radenovic, A. Parameter-free image resolution estimation based on decorrelation analysis. *Nat Methods* **16**, 918-+, doi:10.1038/s41592-019-0515-7 (2019).

48 Han, X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys* **44**, 1408-1419, doi:10.1002/mp.12155 (2017).

49 He, K., Zhang, X., Ren, S. & Sun, J. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770-778 (Las Vegas, NV, 2016).