



# Novel Deep Learning Approaches for Analyzing Diffusion Imaging Data

Yousef Sadegheih<sup>1</sup> · Leon Weninger<sup>3</sup> · Dorit Merhof<sup>1,2,3</sup> 

Received: 31 August 2023 / Accepted: 3 September 2023  
© The Author(s) 2023

## Abstract

Diffusion magnetic resonance imaging (dMRI) is developing into one of the most important non-invasive tools for clinical brain research. This development is supported by a project funded by the German Research Foundation, in which four major obstacles related to dMRI data were addressed: (1) the lack of transferability of dMRI data between clinical sites, (2) the lack of training and label data, (3) the potential of complex diffusion data, and (4) the integration of spherical signals in neural networks to improve accuracy. To overcome the problem of different MRI systems producing slightly varying data, the project developed a method for harmonizing MRI signals. To address the issue of limited ground truth data, a framework was developed to synthesize individual diffusion data and complete datasets based on important diffusion characteristics and statistics. The integration of complex signals, often discarded during acquisition, to improve reconstruction was also explored. Finally, new methods were developed to preserve the spherical character of the diffusion data in the DL model. The resulting methods are intended to improve the usability of diffusion imaging data and to enable the creation of processing pipelines for dMRI data in clinical studies and clinical practice.

**Keywords** Diffusion imaging · Signal harmonization · Data synthesis · Spherical signals · Deep learning

## 1 Introduction

Diffusion imaging is a non-invasive technique that provides information on the diffusion of water molecules in biological tissues, which can be used to infer the underlying tissue microstructure [1]. However, analyzing high-dimensional diffusion data requires complex image processing techniques, which can be time-consuming and computationally demanding. Therefore, the use of deep learning (DL) methods has gained attention in recent years to accelerate

the analysis of diffusion data and improve the accuracy of the reconstruction [2–4]. This project aimed to develop an innovative and disruptive DL pipeline for diffusion imaging that can handle complex spherical signals and achieve optimal reconstruction quality. To achieve this goal, four fundamental research questions were identified and addressed in this project:

1. *Lack of Transferability*: The transferability of DL models between clinical sites has been a significant challenge due to the variability in image acquisition protocols, scanners, and patient populations. A model trained on a specific dataset may not perform well on new datasets with different acquisition protocols. This problem is particularly challenging in diffusion imaging, where the diffusion properties vary depending on the tissue type and pathology. To address this problem, the project explored supervised, semi-supervised, and unsupervised harmonization approaches to transfer datasets between different scanners or sites. The results showed that using harmonization approaches could improve the transferability of the DL model and reduce the reliance on large datasets for single clinical sites.

---

✉ Dorit Merhof  
dorit.merhof@ur.de

Yousef Sadegheih  
yousef.sadegheih@ur.de

Leon Weninger  
weninger@lfb.rwth-aachen.de

<sup>1</sup> Faculty of Informatics and Data Science, University of Regensburg, Regensburg, Germany

<sup>2</sup> Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany

<sup>3</sup> Faculty of Electrical Engineering and Information Technology, RWTH Aachen University, Aachen, Germany

2. *Lack of Training and Label Data:* The availability of training and label data is a significant challenge in diffusion imaging, particularly in small-scale studies or rare diseases. The lack of ground-truth data makes it challenging to train DL models and evaluate their accuracy. In this project, the use of synthetic data based on novel diffusion models was explored to generate training and label data for the DL pipeline. The results showed that synthetic data could provide accurate ground-truth data for training and evaluating DL models, particularly in small-scale studies or rare diseases.
3. *Potential of Complex Diffusion Data:* The diffusion signal in magnetic resonance imaging (MRI) is a complex signal consisting of amplitude and phase information. The phase information contains valuable information about tissue microstructure, but its use in diffusion imaging is challenging due to its sensitivity to motion and susceptibility artifacts. This project explored integrating phase information in the DL model, which may improve the reconstruction quality of complex diffusion data.
4. *Spherical Signals in Neural Networks:* Diffusion data is a spherical signal, and the gradient directions are highly correlated with the gradients in their immediate neighborhood. Therefore, the spherical character of the diffusion data should be preserved in the DL model to improve the accuracy of the reconstruction. In this project, the use of spherical harmonics was explored to preserve the spherical character of the diffusion data in the DL model.

The above research questions were addressed in a research project funded by the German Research Foundation (DFG) under grant number 417063796. In this project report, all five work packages of the project are presented, and for each work package, a summary of the methodological approaches as well as the results is provided, which have previously been presented in several publications [5–10]. Also, an overall discussion and conclusion are presented.

## 2 Work Packages—Methods and Results

### WP1—Signal Harmonization

Diffusion MRI (dMRI) data from different sites exhibit considerable variability due to various factors, including hardware, acquisition settings, and reconstruction algorithms. The variability can be in the same range as biological variability, which complicates clinical studies, and reliable harmonization methods are needed for multicenter studies. Two approaches to diffusion harmonization are presented here: an unsupervised algorithm that can perform harmonization without paired data and an approach that requires

paired data and incorporates a diffusion tensor imaging loss for harmonization. The cyclic neural network architecture forms the basis of the unsupervised algorithm. Unsupervised learning is possible with cyclic networks because of their unique capacity to generate images without the need for paired data. Training can occur in unsupervised, semi-supervised, or supervised settings depending on the amount of paired data available. The methods and results were previously published in [5, 6]. In the first step, the unsupervised harmonization algorithm was created and evaluated on the HCP dataset. A test group of 86 subjects who were not part of the training procedure was used to evaluate the trained generators. The mean squared error (MSE) of the dMRI raw data, mean diffusivity and fractional anisotropy were chosen for evaluation. The proposed cyclic network was trained using a completely unsupervised method (without paired images), a fully supervised method (using paired images acquired in WP3), and a nested combination of the unsupervised and supervised training methods. Furthermore, a baseline was incorporated for comparison. The baseline data underwent the same preprocessing procedures across all subjects, including registering paired data and resampling diffusion directions. These procedures introduced some degree of smoothing, thereby reducing the mean squared error (MSE) between the two datasets. Table 1 shows the baseline and all outcomes.

The development of this algorithm was then evaluated on the data collected in WP3. The proposed approach was compared to SHResNet [11], the standalone MICA method [12], and the baseline between the two images to evaluate its performance. The 24 individuals were randomly split into four-subject groups for a six-fold cross-validation as part of the evaluation process. Before applying the complete pipeline,

**Table 1** Mean squared error between target and harmonized data for supervised, unsupervised, and mixed training in the 3T range (see also Table 1 in [6])

Training	Raw (in $10^{-3}$ )	FA (in $10^{-3}$ )	MD (in $10^{-8}$ )
Baseline	$1.25 \pm 0.51$	$0.82 \pm 0.32$	$0.58 \pm 0.47$
Unsupervised	$0.95 \pm 0.39$	$0.71 \pm 0.29$	$0.53 \pm 0.44$
2% sup	$0.85 \pm 0.40$	$0.68 \pm 0.29$	$0.49 \pm 0.44$
4% sup	$0.83 \pm 0.38$	$0.68 \pm 0.29$	$0.49 \pm 0.44$
10% sup	$0.81 \pm 0.39$	$0.67 \pm 0.29$	$0.47 \pm 0.44$
20% sup	$0.79 \pm 0.38$	$0.67 \pm 0.30$	$0.47 \pm 0.44$
30% sup	$0.76 \pm 0.37$	$0.66 \pm 0.29$	$0.44 \pm 0.44$
50% sup	$0.74 \pm 0.38$	$0.65 \pm 0.29$	$0.42 \pm 0.44$
80% sup	$0.72 \pm 0.38$	$0.64 \pm 0.30$	$0.41 \pm 0.42$
Supervised	$0.68 \pm 0.35$	$0.63 \pm 0.30$	$0.38 \pm 0.42$

In the unsupervised setting no matching images are available. x% supervision denotes the percentage of subjects with matching images during training (i.e. 2% supervision refers to a single subject with matching images, given a training set of 50 subjects)

cross-validation was carried out, and it was maintained for every single step. Separate but identical harmonization pipelines were employed for each of the three diffusion shells ( $b = 1000 \text{ s/mm}^2$ ,  $b = 2000 \text{ s/mm}^2$ ,  $b = 3000 \text{ s/mm}^2$ ) for MICA and the proposed method. The final product is a merger of the three harmonized shells. The study conducted three distinct evaluations: (1) a comparison of the single shell metrics utilized for training the proposed neural network, including the raw diffusion attenuation and DTI metrics FA and MD, (2) an analysis of the variations in multi-shell microstructure modeling using NODDI [13], and (3) an assessment of the perception-based similarity using the LPIPS metric [14]. The Wilcoxon signed-rank test was used to determine if a statistically significant number of subjects showed significantly better harmonization results than with the previously available methods. This was done for all evaluations in comparison to the best baseline result or the second-best approach. First, the raw diffusion attenuation harmonization performance and the metrics derived from DTI were compared for each shell. The raw diffusion attenuation signal's mean squared error (MSE) was assessed independently for each of the three distinct b-value shells. Individual FA and MD maps were generated for each diffusion shell. The findings are presented in Table 2. On the three derived metrics, our proposed method beats all other approaches, and it is the only examined method that did better than the baseline on all measures. The SHResNet approach, on the other hand, exhibited the best results in terms of raw diffusion attenuation error. Second, we utilized NODDI [13] to measure intra-head coil impacts and harmonization performance on microstructural estimations by fitting it to the whole multi-shell data. The MSE of the neurite orientation dispersion index (ODI) and neurite density index (NDI) were calculated for each method. In addition, we compared the effect on fiber direction by calculating the average orientation of the Watson distribution modified for the neurite compartment. Our algorithm yields statistically significant gains in NDI

and ODI image similarity ( $p < 0.005$ ), with NDI seeing the greater of the two effects [5]. The fiber orientation similarities between the two acquisitions have also been enhanced, but not to a statistically significant degree.

Third, given that basic signal smoothing techniques have demonstrated enhancements in the majority of signal comparison assessments [9], we proceeded to assess the disparities in raw diffusion-weighted images using a perception-based metric. The selection of the LPIPS metric was based on its ability to replicate human perception closely [14]. The calculation was performed on a per-individual basis across all diffusion-weighted images and subsequently averaged across both images and subjects. The approach we propose exhibits superior performance compared to both the baseline and the two comparison algorithms across all three b-value shells (see Table 3 in [5]). It is worth mentioning that the Mean Squared Error (MSE) reduces as the b-value increases for raw data, Fractional Anisotropy (FA), and Mean Diffusivity (MD) owing to the lower signal intensity (see Table 2). Simultaneously, the LPIPS scores exhibit an upward trend with an increase in the b-value, suggesting a heightened level of dissimilarity among images with more robust diffusion weighting. The results of the three distinct harmonization approaches are shown in Figure 3 in [5]. The proposed method and MICA significantly reduce signal distortion compared to SHResNet. The MICA technique is incapable of signal smoothing since it does not intrinsically consider location-dependent effects. However, our suggested approach takes local context into account. Figure 3 in [5] shows an example of this impact, shown by the green circles.

*WP2—Development of Methods for Data Synthesis*

The method of data synthesis relies significantly on the generation of training data using voxels obtained from each subject individually. In this study, using an eroded Fast segmentation mask [15], cerebrospinal fluid (CSF) and gray matter (GM) voxels were extracted from a registered T1 image. The diffusion tensors are calibrated for individual

**Table 2** Difference between the two head coil acquisitions before and after harmonization for the three b-value shells, according to [5], Table 2

	Before	SHResNet	MICA	Proposed
Raw data b1000 (in $10^{-3}$ )	9.39 ± 0.79	<b>5.82 ± 0.49*</b>	8.85 ± 0.75	8.58 ± 7.07
Raw data b2000 (in $10^{-3}$ )	7.27 ± 0.55	<b>4.33 ± 0.32*</b>	6.82 ± 0.52	6.62 ± 0.50
Raw data b3000 (in $10^{-3}$ )	6.18 ± 0.49	<b>3.71 ± 0.32*</b>	5.71 ± 0.44	5.48 ± 0.41
MD b1000 (in $10^{-9}$ )	29.6 ± 6.54	33.9 ± 7.42	29.5 ± 6.60	<b>28.2 ± 6.36*</b>
MD b2000 (in $10^{-9}$ )	9.01 ± 1.75	9.83 ± 1.78	8.71 ± 1.62	<b>8.40 ± 1.53*</b>
MD b3000 (in $10^{-9}$ )	4.50 ± 0.75	4.75 ± 7.00	3.92 ± 0.58	<b>3.80 ± 0.60*</b>
FA b1000 (in $10^{-3}$ )	8.09 ± 0.88	9.34 ± 1.06	7.87 ± 0.84	<b>7.62 ± 0.82*</b>
FA b2000 (in $10^{-3}$ )	6.92 ± 0.62	8.45 ± 0.84	6.99 ± 0.60	<b>6.74 ± 0.59*</b>
FA b3000 (in $10^{-3}$ )	6.79 ± 0.51	8.12 ± 0.76	6.84 ± 0.51	<b>6.53 ± 0.50*</b>

Mean squared error (MSE) of diffusion attenuation (raw data), FA, and MD. The best outcomes are highlighted in bold, and those statistically significant ( $p < 0.005$ ) better than the second-best strategy are denoted with an asterisk\*

b-value shells in order to accommodate these specific voxels. The mean diffusivity of cerebrospinal fluid (CSF) and gray matter (GM) are determined based on these voxels. Subsequently, these values are employed to produce artificial training data. A single diffusion tensor may prove inadequate in representing the white matter (WM) diffusion signal throughout the brain, given that multiple microstructural compartments may impact the diffusion within a voxel. The development of a single-fiber response function in constrained spherical deconvolution (CSD) [16] is used to extract single-fiber WM voxels from voxels within the corpus callosum with an FA larger than a pre-defined threshold ( $FA > 0.7$ ). In comparison to the approach of constrained spherical deconvolution (CSD), our method involves the preservation of all single-fiber voxels extracted from the white matter. Then, for each voxel and shell, we fit prolate diffusion tensors, predicting three diffusion tensors for a three-shell acquisition, identical to the diffusion sequence used in the Human Connectome Project (HCP). Importantly, these tensors are estimated independently of one another. This leads to a range of diffusion tensors that represent various white matter microstructures. This method relies on the assumptions that GM and CSF diffusion are isotropic and that diffusion can be defined by single fibers with prolate tensors for a single b-value. Because diffusion characteristics are evaluated separately for each shell, b-value-dependent effects (such as kurtosis) are incorporated in the synthetic data. In this study, an autoencoder [17], a network architecture originally created for denoising purposes, is used to evaluate the quality of synthetic diffusion data. When synthetic data is utilized for training and real data is utilized for evaluation, it can serve as a metric for assessing the quality of synthetic data. In cases where the generative model of synthetic data aligns with only a subset of the features present in the in vivo data, the trained autoencoder may face limitations in its ability to generalize across the entire spectrum of in vivo data. Conversely, if the generative model induces excessive diversity, the autoencoder's performance may be compromised due to the neural network's loss of specificity. Consequently, the utilization of synthetic data that is more appropriate will result in enhanced reconstruction performance in in-vivo data [8]. The autoencoder proposed in this study is designed to be self-adaptive to the shape of the input in order to optimize the balance between the generality and specificity of the data, regardless of the type of diffusion acquisition. Consequently, the dimensions of the architecture are dynamically adapted to the number of diffusion directions. Four distinct synthetic diffusion data models were subjected to training and evaluation on single-voxel data in order to assess their respective quality. We evaluate the autoencoder's reconstruction performance on various synthetic datasets using the diffusion-weighted raw signal and metrics obtained from adaptive diffusion tensors.

In order to evaluate the reconstruction performance, first of all the diffusion-weighted signal is considered; this is the diffusion signal divided by the  $b_0$  value. A detailed evaluation is provided in [8], where Figure 1 depicts the standard deviation of the reconstructed signal from dMRI images of the brain. Reconstruction performance for Syn, Sample & Mean, and Sample are almost comparable and display no statistically significant variations in HCP and study data, whereas RandomWM mode performs far worse. The performance of reconstruction varies between the HCP acquisitions and study data. The voxel sizes employed in the HCP study (1.25 mm) and the local study (2.4 mm) provide an explanation for this difference in results. Accordingly, one voxel in the local study data corresponds to  $(2.4/1.25)^3 \approx 7$  voxels in the HCP data. Without taking into account variations caused by other acquisition parameters and the pre-processing pipeline, an SNR difference between the two datasets of  $(2.4/1.25)^3 \approx 2.66$  is expected. The observed difference in the raw reconstruction performance is precisely reflected by the mean absolute error values obtained for the three multi-tensor simulations, which are 0.014 and 0.037 for the local study data and HCP data, respectively. Subsequently, the reconstruction performance was evaluated utilizing metrics that were derived from adapted diffusion tensors. The study found insignificant variation in fractional anisotropy (FA) among the three multi-tensor modes, namely Syn, Sample & Mean, and Sample. The mean absolute error for the three modes was 0.0075 in the local study data and 0.0104 in the HCP data. In contrast, the RandomWM model exhibited considerably poorer performance, with a mean absolute deviation of 0.17 in the local study data and 0.22 in the HCP data. Figure 2 in [8] illustrates notable dissimilarities among the multi-tensor modes in their ability to reconstruct the main fiber direction. Significant dissimilarities are apparent between the Syn and WM sampling models. In contrast, variations in the average or sampling of GM and CSF diffusion attenuation do not exert a comparable level of impact on the outcomes. The RandomWM method's performance, as indicated by the local study ( $42.4 \pm 3.2$ ) and HCP ( $42.7 \pm 2.8$ ), is notably inferior to the other methods. Consequently, it has been omitted in the figure to enhance clarity.

This approach for generating synthetic single-voxel data was successfully applied for training free-water correction approaches in diffusion imaging data [18, 19].

*WP3—Collection of diffusion data for the evaluation of the project*

Thirty healthy subjects were scanned on different MRI machines using high-resolution diffusion sequences. For this purpose, two different 3T MRI systems (a Siemens Prisma Scanner and a Siemens Prisma Fit Scanner with coils with 20 and 64 channels, respectively) from the University Hospital Aachen were used. The Siemens Prisma Fit Scanner is identical in design to the Siemens Prisma

Scanner. For this study, 30 healthy male subjects were recruited through lectures and social media. The subjects' ages were divided into two groups (21-25 and 26-30). Identical MRI measurements of the head were performed on both MRI systems. An anatomical measurement of the brain was performed at the beginning of each scan, which was necessary to examine different brain areas. Then, a diffusion imaging sequence was used based on the Human Connectome Project sequence. This very high-resolution sequence allows for optimal evaluation of newly developed methods for reconstructing nerve pathways. Furthermore, this recording can simulate a clinical recording, which does not need to be additionally taken. Measurements on both MRI systems were conducted on the same day, where possible. If this was not possible, the measurements were performed on two consecutive days and as close together as possible at the same time of day. This excluded side effects caused by growth or cyclic hormone effects. In a separate Zoom meeting before the first scan, the subjects were comprehensively informed about the study and its risks. The first measurement was then taken at the first MRI machine, and the second measurement was taken at the second appointment. Each of the two measurements took 1 h and 30 min. In total, it is assumed that the participation in the experiment (4 h) was compensated with 50€. A written consent form for participation in the study will be obtained from all subjects. This study was conducted in accordance with the Declaration of Helsinki. Further information on the recording sequence and study participants can be found in [20], Chapter 6. The data preprocessing was implemented using a Nipype pipeline [21]. In the first step, the DICOM raw images were converted to the NIFTI format using `dcm2niix` [22]. Skull-stripping and segmentation are performed with ANTs [23] on two T1 recordings, one from the 64 channel head coil session and one from the 20 channel head coil session. These processes were carried out jointly with the "CorticalThickness" tool, which uses established criteria for brain extraction and tissue segmentation to carry out an anatomical T1 brain processing. The brain is segmented into cerebrospinal fluid, cortical gray matter, deep gray matter, and white matter. FSL [24] tools are used to process the dMRI scans. The anterior-posterior scans are corrected for susceptibility-induced distortions using the posterior-anterior data with FSL Topup [25]. The brain is extracted using BET [26]. Eddy currents are removed using FSL Eddy [27] to correct motion artifacts. The segmentation map of the brain tissue is transformed into the diffusion space. Dipy [28] is used to fit diffusion tensors to the dMRI data and derive fractional anisotropy (FA) maps. The FA maps are used for affine registration to the T1 scan using the T1 scan from the same scanner. All registrations are performed using the special ANTs registration toolbox [29]. Finally,

the four separate dMRI recordings for each subject must be merged in order to analyze differences in scanner setup and subsequent harmonization.

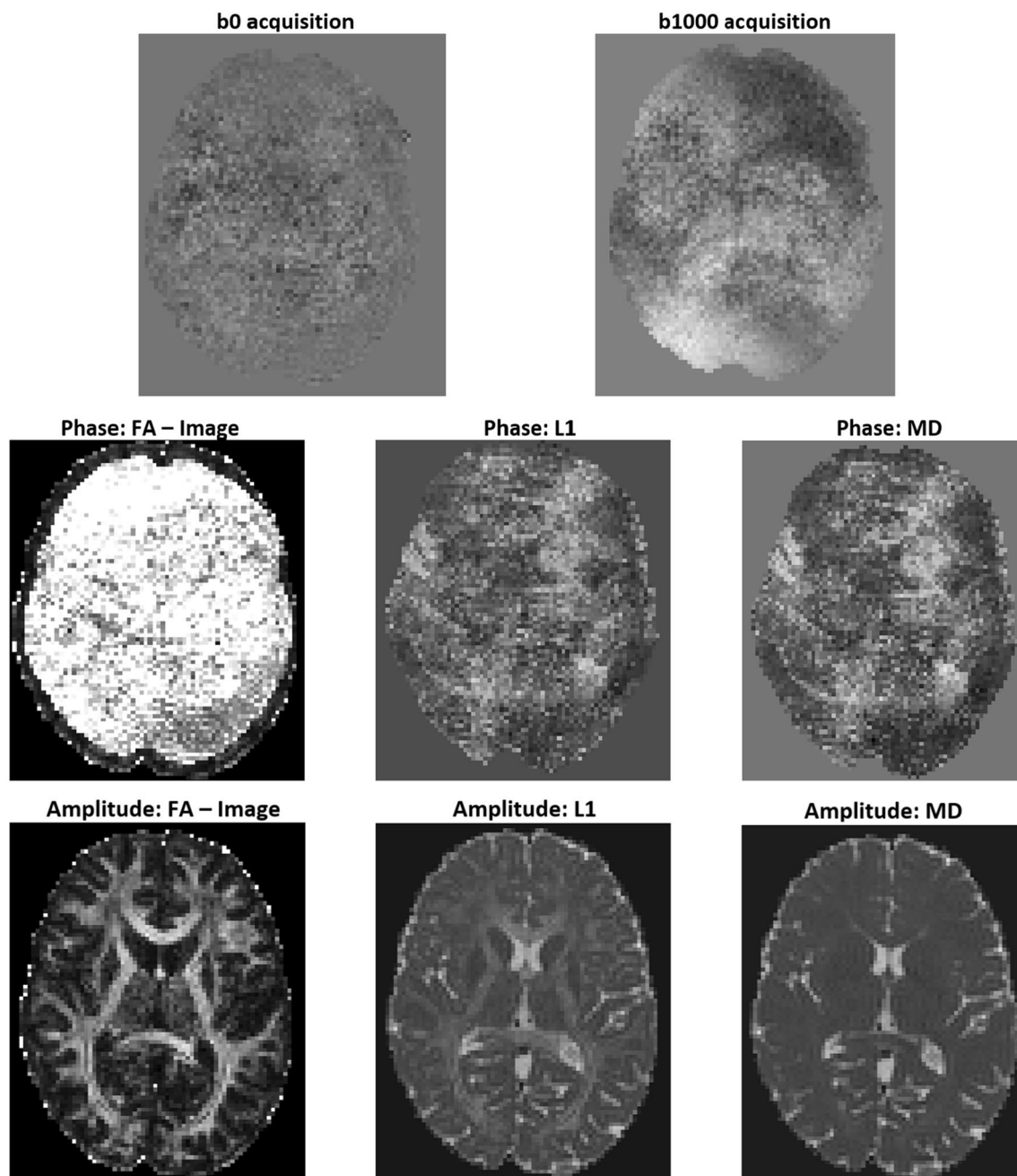
#### *WP4—Complex Signals (Amplitude and Phase)*

The  $B_0$  field strength shows small local changes depending on the magnetizability of the tissue. For example, blood (which is rich in iron) has a different susceptibility than surrounding tissue. In  $T2^*$  weighted images, the phase image provides a clear contrast because  $T2^*$  weighted images are sensitive to susceptibility artifacts and, therefore, to phase. This is mainly due to the gradient-echo sequence used for  $T2^*$  weighted images. To suppress these  $T2^*$  effects, pulse-echo sequences are used for T1 and T2 weighted images, which suppress dephasing. Therefore, the remaining phase is very noisy because it is close to zero. In the case of diffusion-weighted EPI sequences, a multi-echo GRE sequence is used during acquisition and a spin-echo sequence during excitation. The spin-echo sequence's echo occurs during acquisition (Fig. 1). For this reason, only a noisy phase can be measured precisely at TE. Although the ratio of phase to noise is better in the higher-frequency parts of the k-space, they contain only a few image details. Thus, a classical DWI sequence in its current form is unsuitable for obtaining phase information. Changing the weighting and acquisition sequence could improve the phase data's noise. However, a research scanner would be necessary for this purpose. Therefore, whether phase information provides added value cannot be definitively determined. It is known that susceptibility depends on the applied  $B_0$  field. However, the  $B_0$  field is almost unchanged during a DWI sequence because the diffusion gradients are two orders of magnitude smaller than the applied  $B_0$  field.

#### *WP5—Spherical Signals*

Recent results in DL suggest that incorporating additional information improves the results of the respective DL approach [30, 31]. Although novel network architectures such as the UNet [32], ResNet [33], or DenseNet [34] can improve performance, it is commonly agreed that it is still crucial to develop sensible training strategies, individualized loss functions and individual layers [35]. In order to integrate the spherical character of a measured signal into a DL network, novel spherical DL layers were developed in this work package and analyzed with respect to a diffusion signal to ensure optimal processing of the diffusion signal in a deep neural network.

With respect to diffusion imaging data, the effect of different activation functions on the diffusion signal (Fig. 2), as well as on the Fourier signal of a diffusion signal (Fig. 3), was analyzed. This is particularly important since the spherical character of the signal has to be maintained within a DL network. The first activation function (Fig. 2 left) shows the Rectified Linear Unit (ReLU). This activation function shows the input signal without modification, as a ReLU only



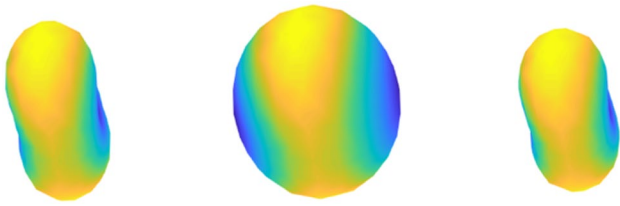
**Fig. 1** *Upper row:* the phase in diffusion-weighted images. *Middle row:* the phase in diffusion tensor maps. *Bottom row:* in comparison, the corresponding representations of the amplitude images

sets negative values to zero. Since the diffusion signal has positive values only, this type of activation has no effect. This changes when a sigmoid or a hyperbolic tangent (TanH) function are chosen as activation functions. Here, a slight contrast reduction (sigmoid) as well as a contrast enhancement (TanH) can be observed after applying these functions.

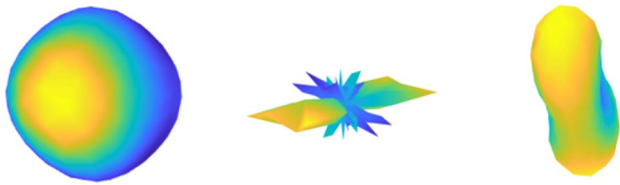
A slightly different situation arises when the activation function is applied in the Fourier space. A regular diffusion signal in the Fourier space usually has many values close

to zero. Only the DC component of the signal has larger absolute values. A comparison of the different activation functions applied in the Fourier space can be seen in Fig. 3.

As illustrated in Fig. 3, a ReLU activation leads to reduced contrast as well as a drastic change in the signal, as any negative coefficients are removed. In case of the sigmoid function, negative values are converted to positive values, which destroys parts of the diffusion signal and therefore has a substantially stronger effect on the



**Fig. 2** Visualization of the diffusion signal for three different activation functions (from left to right: ReLU activation, sigmoid activation, and TanH activation), applied in the signal space. (Images courtesy Simon Koppers, Aachen)



**Fig. 3** Visualization of the diffusion signal for three different activation functions (from left to right: ReLU activation, sigmoid activation, and TanH activation), applied in the Fourier space. (Images courtesy Simon Koppers, Aachen)

signal. In contrast, the TanH function accomplishes a contrast-enhancing effect by altering small values minimally, but large values significantly. This mainly affects the DC component of the signal, which is reduced by the function.

To evaluate the reconstruction accuracy of different spherical layers, a denoising autoencoder [36] was employed. Diffusion signals were artificially enhanced with noise, and the different layer-specific autoencoders were then used to remove this noise. Different combinations of spherical layers and activation functions were used to evaluate the individual autoencoders. Subsequently, the performance was assessed through three distinct evaluations: runtime, signal denoising performance, and the impact on reconstruction accuracy in state-of-the-art reconstruction methods. Twenty subjects from the Human Connectome Project were used for evaluation purposes. This database provides highly resolved diffusion images of healthy subjects. In detail, each diffusion image has a voxel dimension of  $1.5 \text{ mm}^3$ , three different b-values with 90 gradient directions each, and an additional set of 18 non-weighted diffusion images. The different methods were evaluated only on the white matter region of the individual brains. All results, as well as a detailed description of the experiments, were published in [11]. In general, it can be said that spherical layers lead to a significant gain of information within a DL network.

### 3 Discussion

The development of a DL analysis pipeline for diffusion imaging data is an exciting and innovative approach that has the potential to significantly improve the efficiency and accuracy of clinical brain research. The project consists of four methodological work packages (WP1, WP2, WP4, and WP5), each addressing a fundamental research question related to the application of DL in diffusion imaging data, plus one work package (WP3) for diffusion data acquisition. The first work package focuses on the lack of transferability of diffusion MRI data between clinical sites. This is a major barrier to the widespread adoption of diffusion MRI in clinical practice, as different MRI systems can produce significantly different results. The development of an optimal method for harmonizing MRI signals is essential to ensure that the DL pipeline can be used effectively across different clinical sites. The second work package addresses the issue of the lack of training and label data for DL. This is a common problem in medical imaging, as ground truth data is often difficult to obtain. The proposed framework for synthesizing diffusion data based on important diffusion characteristics and statistics is a promising solution to this problem. This approach will allow for the creation of large single voxel datasets with corresponding ground truth that can be used to train DL pipelines for diffusion imaging data analysis. The fourth work package explores the potential of complex diffusion data. While the phase information is often discarded during acquisition, complex MRI signals comprising amplitude and phase may carry important tissue information that could potentially be used to improve the accuracy of reconstruction. However, our results were not conclusive yet and would need sequence development on a research MRI scanner for further exploration. The fifth work package focuses on the integration of spherical signals in neural networks to improve accuracy by explicitly considering the spherical character of the signals. Previous DL methods have not been able to incorporate angle-related diffusion signals per voxel, which is why new methods are needed to adopt previous DL methods to spherical signals. This approach will allow for the inclusion of neighboring information within the signal as well as between signals to ensure optimal reconstruction. Overall, the proposed DL approaches for diffusion imaging data have the potential to further support clinical brain research by improving the accuracy and efficiency of diffusion MRI analyses. The presented work packages address fundamental research questions related to the application of DL in diffusion imaging, and the results of this project could have a significant impact on the future application of diffusion imaging data for clinical purposes.

## 4 Conclusion

In conclusion, the development of an innovative and disruptive DL pipeline for diffusion imaging data is a significant step towards improving clinical practice in brain research. The identified research questions, including the lack of transferability of diffusion MRI data between clinical sites, the lack of training and label data, the potential of complex diffusion data, and the integration of spherical signals in neural networks, have been effectively addressed in this project. The results show that the developed pipeline can handle harmonization issues, synthesize single-voxel diffusion imaging data, and integrate spherical signals into DL models to improve the reconstruction quality. Future research may further refine and expand upon these developments to enhance the capabilities of DL in the field of dMRI for clinical application.

**Acknowledgements** This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) under grant number 417063796 (project ME3737/19-1).

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data Availability** Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Descoteaux M (2015) High angular resolution diffusion imaging (HARDI). Wiley, New York, pp 1–25
- Luo S, Zhou J, Yang Z, Wei H, Fu Y (2022) Diffusion MRI super-resolution reconstruction via sub-pixel convolution generative adversarial network. *Magn Reson Imaging* 88:101–107
- Tanno R, Worrall DE, Kaden E, Ghosh A, Grussu F, Bizzi A, Sotiropoulos SN, Criminisi A, Alexander DC (2021) Uncertainty modelling in deep learning for safer neuroimage enhancement: demonstration in diffusion MRI. *Neuroimage* 225:117366
- Zhang F, Breger A, Cho KIK, Ning L, Westin C-F, O'Donnell LJ, Pasternak O (2021) Deep learning based segmentation of brain tissue from diffusion MRI. *Neuroimage* 233:117934
- Weninger L, Romanzetti S, Ebert J, Reetz K, Merhof D (2023) Harmonization of diffusion MRI data obtained with multiple head coils using hybrid CNNs. In: *Computer Vision—ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VII*, Springer, Berlin, pp 385–396
- Weninger L, Ahmad M, Merhof D (2022) From supervised to unsupervised harmonization of diffusion MRI acquisitions. In: *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp 1–5
- Weninger L, Ecke J, Na C-H, Jütten K, Merhof D (2022) Diffusion MRI specific pretraining by self-supervision on an auxiliary dataset. In: *Bildverarbeitung für die Medizin 2022: Proceedings, German Workshop on Medical Image Computing, Heidelberg, June 26–28, 2022*, Springer, Berlin, pp 147–152
- Weninger L, Drobnjak M, Na C-H, Jütten K, Merhof D (2021) Autoencoder-based quality assessment for synthetic diffusion-MRI data. In: *Bildverarbeitung für die Medizin 2021: Proceedings, German Workshop on Medical Image Computing, Regensburg, March 7–9, 2021*, Springer, Berlin, pp 338–343
- Ning L, Bonet-Carne E, Grussu F, Seppehrband F, Kaden E, Veraart J, Blumberg SB, Khoo CS, Palombo M, Kokkinos I et al (2020) Cross-scanner and cross-protocol multi-shell diffusion MRI data harmonization: algorithms and results. *Neuroimage* 221:117128
- Koppers S, Merhof D (2018) Delimit pytorch—an extension for deep learning in diffusion imaging. [arXiv:1808.01517](https://arxiv.org/abs/1808.01517)
- Koppers S (2019) Signal enhancement and signal reconstruction for diffusion imaging using deep learning. PhD thesis, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen
- Wrobel J, Martin M, Bakshi R, Calabresi P, Elliot M, Roalf D, Gur R, Gur R, Henry R, Nair G, Oh J, Papinutto N, Pelletier D, Reich D, Rooney W, Satterthwaite T, Stern W, Prabhakaran K, Sicotte N, Shinohara R, Goldsmith J (2020) Intensity warping for multisite MRI harmonization. *Neuroimage* 223:117242
- Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC (2012) NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* 61(4):1000–1016
- Zhang R, Isola P, Efros AA, Shechtman E, Wang O (2018) The unreasonable effectiveness of deep features as a perceptual metric. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 586–595
- Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20(1):45–57
- Tournier J-D, Calamante F, Connelly A (2007) Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35(4):1459–1472
- Goodfellow I, Bengio Y, Courville A (2016) *Deep learning*. MIT press, Cambridge
- Weninger L, Koppers S, Na C-H, Juetten K, Merhof D (2020) Free-water correction in diffusion MRI: a reliable and robust learning approach. In: *Computational Diffusion MRI*, Springer International Publishing, Berlin, pp 91–99
- Weninger L, Na C-H, Jütten K, Merhof D (2020) Analyzing the effects of free water modeling by deep learning on diffusion MRI structural connectivity estimates in glioma patients. *PLoS ONE* 15
- Weninger L (2023) Deep Learning-based analysis of Diffusion MRI image data from multicenter studies and from glioma



- patients. PhD thesis, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen
21. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011) Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front Neuroinform* 13
  22. Li X, Morgan PS, Ashburner J, Smith J, Rorden C (2016) The first step for neuroimaging data analysis: DICOM to NIFTI conversion. *J Neurosci Methods* 264:47–56
  23. Avants BB, Tustison N, Song G et al (2009) Advanced normalization tools (ANTs). *Insight J* 2(365):1–35
  24. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE et al (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23:S208–S219
  25. Andersson JL, Skare S, Ashburner J (2003) How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 20(2):870–888
  26. Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17(3):143–155
  27. Andersson JL, Sotiropoulos SN (2016) An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125:1063–1078
  28. Garyfallidis E, Brett M, Amirbekian B, Rokem A, Van Der Walt S, Descoteaux M, Nimmo-Smith I, Contributors D (2014) Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform* 8:8
  29. Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 12(1):26–41
  30. Bronstein MM, Bruna J, LeCun Y, Szlam A, Vandergheynst P (2017) Geometric deep learning: going beyond Euclidean data. *IEEE Signal Process Mag* 34(4):18–42
  31. Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*, pp 2980–2988
  32. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18, Springer, Berlin, pp 234–241
  33. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
  34. Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K (2014) DenseNet: implementing efficient ConvNet descriptor pyramids. [arXiv:1404.1869](https://arxiv.org/abs/1404.1869)
  35. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH (2019) “No New-Net,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II* 4, Springer, Berlin, pp 234–244
  36. Vincent P, Larochelle H, Bengio Y, Manzagol P-A (2008) Extracting and composing robust features with denoising autoencoders. In: *Proceedings of the 25th international conference on Machine learning*, pp 1096–1103