Brain Tissue Microstructure Characterization Using dMRI Based Autoencoder Neural-Networks

Mauro Zucchelli, Samuel Deslauriers-Gauthier, Rachid Deriche

Brain Tissue Microstructure Characterization Using dMRI Based Autoencoder Neural-Networks

Mauro Zucchelli, Samuel Deslauriers-Gauthier, and Rachid Deriche

Inria, Université Côte D'Azur, France

Abstract. In recent years, multi-compartmental models have been widely used to try to characterize brain tissue microstructure from Diffusion Magnetic Resonance Imaging (dMRI) data. One of the main drawbacks of this approach is that the number of microstructural features needs to be decided a priori and it is embedded in the model definition. However, the number of microstructural features which is possible to obtain from dMRI data given the acquisition scheme is still not clear.
In this work, we aim at characterizing brain tissue using autoencoder neural networks in combination with rotation-invariant features. By changing the number of neurons in the autoencoder latent-space, we can effectively control the number of microstructural features that we obtained from the data. By plotting the autoencoder reconstruction error to the number of features we were able to find the optimal trade-off between data fidelity and the number of microstructural features. Our results show how this number is impacted by the number of shells and the b-values used to sample the dMRI signal. We also show how our technique paves the way to a richer characterization of the brain tissue microstructure in-vivo.

Keywords: Spherical Harmonics · Rotation Invariant Features · Neural Networks · Diffusion MRI · Autoencoder

1 Introduction

Diffusion Magnetic Resonance Imaging (dMRI) has been widely used in research to probe the brain tissue microstructure in the living human brain. To estimate microstructural features from the dMRI signal, it is first necessary to select a biophysical model which can capture the desired features [7]. The choice of the model is limited by the type of dMRI data available, e.g. the type of diffusion sequence, the number of gradients (or shells), the number of echo times, etc [5]. Although this approach has been successful, it has several limitations. The biggest limitation is that the choice of the model completely decides the number and type of microstructural feature which are obtainable from the data [8]. A model-fitting will always provide a result but the accuracy of this result is not simple to verify. Nowadays, several software tools [1,2] easily permit researchers

to fit multiple models on the same data. However, the comparison and validation of such models is still an open question[1].

Neural networks (NN) have become widely popular in medical imaging for their astonishing classification given a sufficiently large and accurate training set [6]. One of the most important features of NN is that they can be considered a model-free approach. Normally, no information on the characteristics of the data that we want to classify is given to the NN, because they are able to learn the classes representation directly from the training data.

NN have been already applied to dMRI data in order to obtain microstructural features [3,15]. In these works, a multi-compartment model has been chosen for generating the microstructural features used as labels for the NN training set. This approach limits the representation power of the NN to the model used for labeling the training data, not fully exploiting the model-free characteristic of the NN.

Autoencoders are a particular family of NN, principally used for denoising and feature extraction [12]. An autoencoder NN is composed of two parts, an encoder, and a decoder. The encoder that takes as input the data and outputs a certain number of features (also known as latent-space), corresponding to the number of neurons in the last layer of the encoder. The decoder takes as input these features and outputs the original data (see Fig. 1). An autoencoder NN is trained by trying to minimize the error between the data and the output of the decoder. In order to prevent the network from learning the identity operation, the number of features produced by the encoder is strictly lower than the size of the input data. This enables the autoencoder NN to codify each input in a small number of features which can then be used for classification or other tasks. Autoencoders have been recently used in dMRI by Vasilev and colleagues to detect abnormal voxels in multiple-sclerosis patients' data [12]. One of the main limitations of using autoencoders with dMRI data is that the dMRI signal is strongly directional-dependent and even a slight change in the orientation in the underlying brain fiber bundles will result in a completely different signal profile. This means that if the dMRI signal samples are used for training the autoencoder at least some of its features will be used to capture the orientation profile. Tissue orientation can be easily retrieved from dMRI data using the popular spherical deconvolution approaches [11], therefore it is generally not included in the scope of microstructural estimation technique. In fact, Rotation Invariant Features (RIF) can be used as an input microstructural models, instead of the row signal samples [9,14].

In this article, we aim to combine high order RIF, derived from the 4^{th} order Spherical Harmonics expansion of the dMRI signal, with autoencoders. Our purpose is to characterize the number of microstructural features which is possible to obtain from dMRI pulse-gradient spin-echo (PGSE) data, regardless of the tissue orientation. We also investigate how this number changes to the number of shells used to sample the data. Our results show that this technique is indeed feasible to model dMRI data and show also that increasing the number

[1] see the recent MEMENTO challenge: `https://my.vanderbilt.edu/memento/`

of shells and the b-value provide a richer characterization of the brain tissue microstructure in-vivo.

2 Materials and methods

2.1 Rotation-invariant features of the dMRI signal

The dMRI signal at each b-value can be represented as a spherical function $f(\mathbf{u})$, where the unit vector $\mathbf{u}$ corresponds to the gradient direction. Considering the diffusion signal is real-valued and antipodal-symmetric, we can expand it into a truncated Spherical Harmonics (SH) series

$$f(\mathbf{u}) = \sum_{l=0,even}^{L} \sum_{m=-l}^{l} c_{lm} Y_l^m(\mathbf{u}) \tag{1}$$

where c_{lm} are the real-valued SH coefficients Y_l^m are the real SH [4] of degree l and order m, and L is the maximum degree. In [14], the authors defined a set of algebraic independent RIF. Each invariant can be calculated as

$$I_{\mathbf{l}}[f] = \sum_{m_1=-l_1}^{l_1} \cdots \sum_{m_d=-l_d}^{l_d} c_{l_1 m_1} \cdots c_{l_d m_d} G(l_1, m_1 | \cdots | l_d, m_d) \tag{2}$$

where G are the generalized Gaunt coefficients [4], namely the integral of d SH and the vector $\mathbf{l}$ is equal to $[l_1, l_2 \cdots, l_d]$. Eq. (2) provides a rotation invariant indices for any set of $\mathbf{l}$. In this work, we focused on the 12 algebraic independent RIF obtained from the SH expansion at $L = 4$ of the diffusion signal [14]. Being invariants, these RIF depends only on the "shape" of the signal and not on the orientation. For example, the value of the RIF change between crossing fibers and single fiber voxels, if two voxels present different intra-axonal signal fractions, or if the water diffusivity changes between the voxels [9]. Since multiple factors contribute to the invariants numerical values they cannot be considered biomarkers per se, but by removing the orientation-dependency they greatly improve the estimation of brain tissue microstructure. In particular, neural network-based microstructural feature estimation considerably benefits from this form of rotation invariant signal representation [15].

2.2 In-vivo dataset

In this work, we consider three subjects of Human Connectome Project (HCP) diffusion MRI young-adults dataset [10]. For each subject, 288 dMRI volumes has been acquired: 18 volumes at b-value 0 s/mm^2, 90 volumes at b-values 1000 s/mm^2, 90 volumes at b-values 2000 s/mm^2, and 90 volumes at b-values 3000 s/mm^2. Each volume is composed of $145 \times 174 \times 145$ voxels with a resolution of $1.25 \times 1.25 \times 1.25$ mm^3. Before calculating the SH expansion, we normalized each volume dividing the diffusion signal by the mean of the b-value 0 s/mm^2 volumes.

2.3 Autoencoder design

Figure 1 shows a graphical representation of the autoencoder used in this work. We design our autoencoder as a fully connected NN for taking as input the 12 RIF calculated from each shell of the HCP dataset. Therefore, the number of input (and output) channels corresponds to 12 for the single-shell b=1000 s/mm^2 data, 24 for the two shells up to b=2000 s/mm^2, and 36 for the complete dataset with three shells up to b=3000 s/mm^2. Following the input layer, we add a first hidden layer composed of 50 neurons combined with a Rectified Linear Unit (ReLU) functions. After this layer, the data is passed to another layer of neurons plus ReLU forming the latent-space layer. The output of the latent-space layer represents the set of microstructural features that characterize the dMRI signal. The cardinality of the latent-space layer is the main variable of interest for our autoencoder design. This number defines the amount of features that the autoencoder uses to represent the input. In this work, we created several neural networks with latent-space ranging from 3 neurons up to the number of the input channels. In this last case, the network should be able to simulate the identity transform perfectly recovering the input, therefore we use it only as a reference limit-case. These first three elements form the *encoder*. The *decoder* takes as input features output of the latent space and passes it to the second hidden layer composed of 50 neurons. As in the case of the first hidden layer, each neuron is followed by a ReLU function. The output of the second hidden layer is forwarded to the neurons of the output layer having the same number as the input of the network. Each autoencoder is optimized trying to minimize the Mean Square Error (MSE) between the input RIF at the different shell and the output of the network. Since each RIF have a different range of values, we scale each invariant according to a factor of 1, 10, or 100 to bring them in the same numerical range. This normalization is necessary in order to avoid that the biggest invariant (I_0, the mean of the signal) dominates the optimization reducing the influence of the other RIF. We trained the network using each of the three HCP subjects, independently, splitting the voxel randomly. Sixty percent of each brain voxels were used for training the network, and forty percent for testing. To test the robustness of the method, we used 10 different random seeds for dividing the data into training and testing, and for the network initialization and optimization. All the autoencoders implemented in this work are written in python using the Pytorch[2] library and trained using the Adam optimizer with a learning rate of 5×10^{-5} and a weight decay of 1×10^{-6}. The training set is split in batches of 1000 elements and we trained the networks for 300 epochs. All of the autoencoder hyper-parameters (including the number of layers, the learning rate, and the weight decay), were optimized by minimizing the MSE of the invariants on one of the HCP subjects.

[2] `https://pytorch.org/`

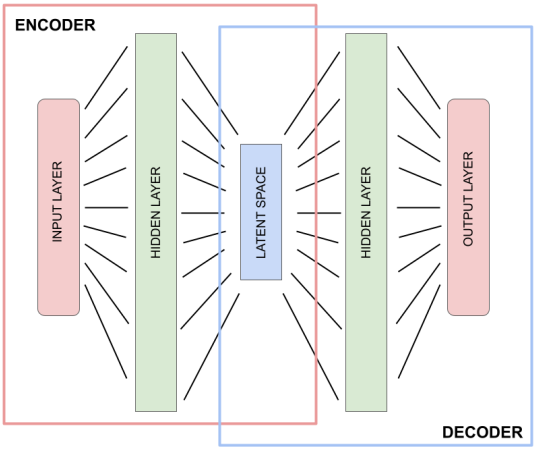

Fig. 1. Graphical representation of the autoencoder design used in this work. The input layer counts 12 $\times$ number of b-values channels, each hidden layer is composed of 50 neurons, and the output layer has the same size of the input. For each data we trained several autoencoder with the number of neurons in the latent space ranging from 3 to the size of the input channel.

3 Results and Discussion

In this work, we used the RIF extracted from the different shells of the dMRI signal as input, and the autoencoder latent-space represents the microstructural features. However, it is unclear how many microstructural features is it possible to extract from PGSE multi-shell dMRI data, in particular, as a function of the number of shells and the b-values used in the acquisitions. Figure 2 shows the log MSE between the original and reconstructed RIF for the testing set voxels of three HCP subjects (in red, green, and blue) as a function of the number of neurons in the latent-space. Each transparent line represents one random-seed instance of the same network, and the solid line the average of the ten networks results. Figure 2 (left) shows the results for the autoencoders trained using the 12 RIF at b-value 1000 s/mm^2. At this b-value, the difference in the reconstruction error between the subjects is in the same range of the changes obtained training the same subject with a different network initialization, in particular, when the number of features of the latent-space is lower than 9. We calculated the elbow-point of the average of the curves for each subject, and we mark it on the graph with black crosses. The elbow-point represents an empirical way to find the "optimal" trade-off between reconstruction error and the number of features. In our experiments, two subjects (160123 and 473952) have reached an optimal number of features with 6 neurons in the latent-space, and one subject (174437) with 9 neurons. These features contain information on both the inner properties of the axons and extracellular space (like the axonal density and water diffusivity) and

information regarding the axons architecture (crossing bundles, fibers fanning, etc). This number of features is higher than the information usually retrieved from single-shell data, which are generally modeled as the three eigenvalues of the diffusion tensor. However, in this case, we used SH of order 4 for modeling the signal, while the diffusion tensor can be considered an order 2 representation. Therefore, the increase in the number of features can be explained by the increase in the order of the signal representation. Figure 2 (center) shows the same graph for the two shells data at b-value 1000 and 2000 s/mm^2. In this case, we have 24 RIF as input (12 for each shell). The autoencoders show a different reconstruction error for each subject when the latent-space grows bigger than 15 neurons. Before this point, the difference between the MSE of the different subjects is in the range of variance induced by the change of the random seed. The optimal number of features, calculated by finding the elbow-points of the curves, increase to 13 for subjects 160123 and 473952 and 14 for subject 174437. This means that there is a minimum increase of 5 features while moving from single-shell low-b-value data, to two shells data. Training the network with all the three shells, up to b-value 3000 s/mm^2 further increase this difference (see Figure 2 right). In this case, we have 36 input channels considering 12 RIF per shell. Using the b-value 3000 s/mm^2 data we observe a clear difference between the reconstruction error of the three subjects. As in the previous cases, this difference increases with the increase of the number of features of the latent-space. It is necessary to underline that each subject has been used for training its autoencoder independently to the other two subjects. This means that there is something in the b-value 3000 s/mm^2 RIF which led the autoencoder to have a systematic different reconstruction error between the subject. This can be due to noise or another hidden acquisition artifact that may be completely uncorrelated to tissue microstructure. Although the errors are different, the trend of the errors is very close to each other. In fact, all three subjects have an optimal number of features at 15 latent-space neurons. The increase to the two-shells data is reduced to only one additional microstructural feature. This means that adding a third shell at b-value 3000 s/mm^2 while adding some subject-specificity and reduce the variance of the results does not increase the number of microstructural features obtainable from the data sensitively. Although we do not have the data, we can imagine that with a further increase in the number of shells we will probably reach a plateau in the number of microstructural features. This result is in line with other studies whose highlight that in order to be sensitive to the full brain tissue microstructure it is necessary to include other dMRI acquisition data, for example, Spherical Encoding [5], shells acquired using multiple diffusion-times [7], or multiple echo-time [13].

Figure 3 shows three examples of the quality of the autoencoder RIF reconstruction for the three shells data with 15 neurons in the latent space (HCP subject 160123). The three considered RIF are the signal mean invariant I_0, the degree 2 power spectrum invariant I_{22}, and one of the new invariant developed by using the original framework presented in [14]: invariant I_{224}. As it is possible to see these three RIF were recovered almost perfectly by the autoencoder.

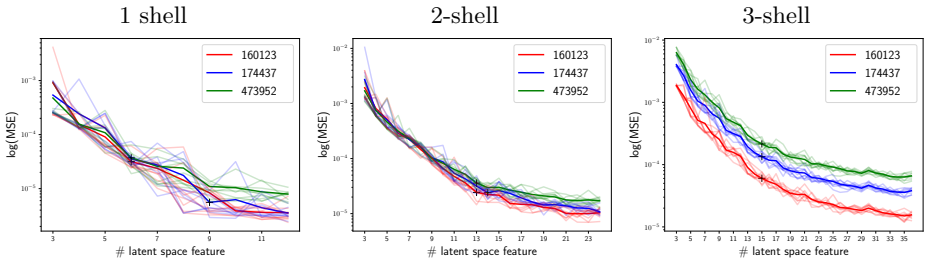

Fig. 2. Log MSE of the difference between the input RIF and the network RIF as a function of the size of the latent-space, for 1 shell (b-value = 1000 s/mm^2, on the left), two shells (up to b-value = 2000 s/mm^2, at the center), and three shells (up to b-value = 3000 s/mm^2, on the right). Each transparent line corresponds to a different random network initialization and the solid line represents the average for each of the three HCP subjects (in red, blue, and green). The black crosses correspond to the elbow point of the log curves.

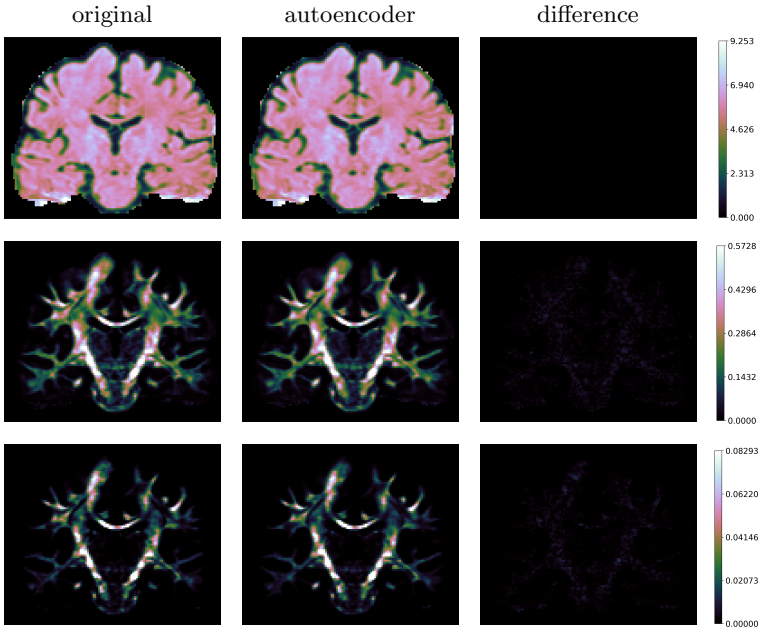

Fig. 3. Example of three RIF (left), their autoencoder reconstruction (center), and the absolute difference between the two (right). The top row represent invariant I_0 at b-value = 1000 s/mm^2, the second row invariant I_{22} at b-value = 2000 s/mm^2, and the third row invariant I_{224} at b-value = 3000 s/mm^2.

Although we only show three examples this level of quality is consistent for all the 36 RIF in the three shells.

Figure 4 shows the 15 microstructural features obtained for the HCP subject 160123 using the 3-shells RIF. As it is possible to see, the different features highlight different parts of the brain tissues. Most of these microstructural features present a high-contrast between white matter and gray matter. In particular, some single-fiber white-matter areas like the corticospinal tract and corpus-callosum present the higher contrast in most of the features, similarly with classical fractional anisotropy maps. We must remind the reader that the RIF are very sensitive to the underlying fiber geometry. Therefore, it is not surprising that most of the autoencoder-based microstructural features can provide high contrast in the different white-matter regions. The work of Vasilev and colleagues [12]] shows that the autoencoder-based features of the dMRI signal can be used to discriminate the voxels affected by multiple sclerosis. Although we based this work only on healthy subjects, we are positive that by combining autoencoders and rotation invariant features the resulting microstructural features might even be more precise in detecting pathological conditions.

4 Conclusions

In this work, we explore the use of the combination of autoencoder neural networks and RIF for the estimation of tissue microstructural features in PGSE dMRI data. There are two main advantages of using this technique. First, the use of RIF permits to eliminate the influence of tissue orientation in the microstructural features derived from the autoencoder latent-space. Second, the autoencoder is a model-free technique. This is particularly important in clinical data where the choice of the wrong model could lead to an incorrect characterization of the microstructural changes induced by pathology. One of the drawbacks of the technique is the reduced interpretability of the microstructural features. While "axonal density" and "extra-axonal diffusivity" are directly linked to the physical properties of the brain tissue, the autoencoder-derived features are more difficult to label. However, the main objective of these features is to help to identify a pathological condition in the brain tissue and be effectively used as biomarkers. Therefore, our future work will be aimed to test our technique on clinical data like multiple sclerosis and Alzheimer's disease.

5 Acknowledgments

This work has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.
This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant agreement No 694665 : CoBCoM - Computational Brain Connectivity Mapping).
Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657)

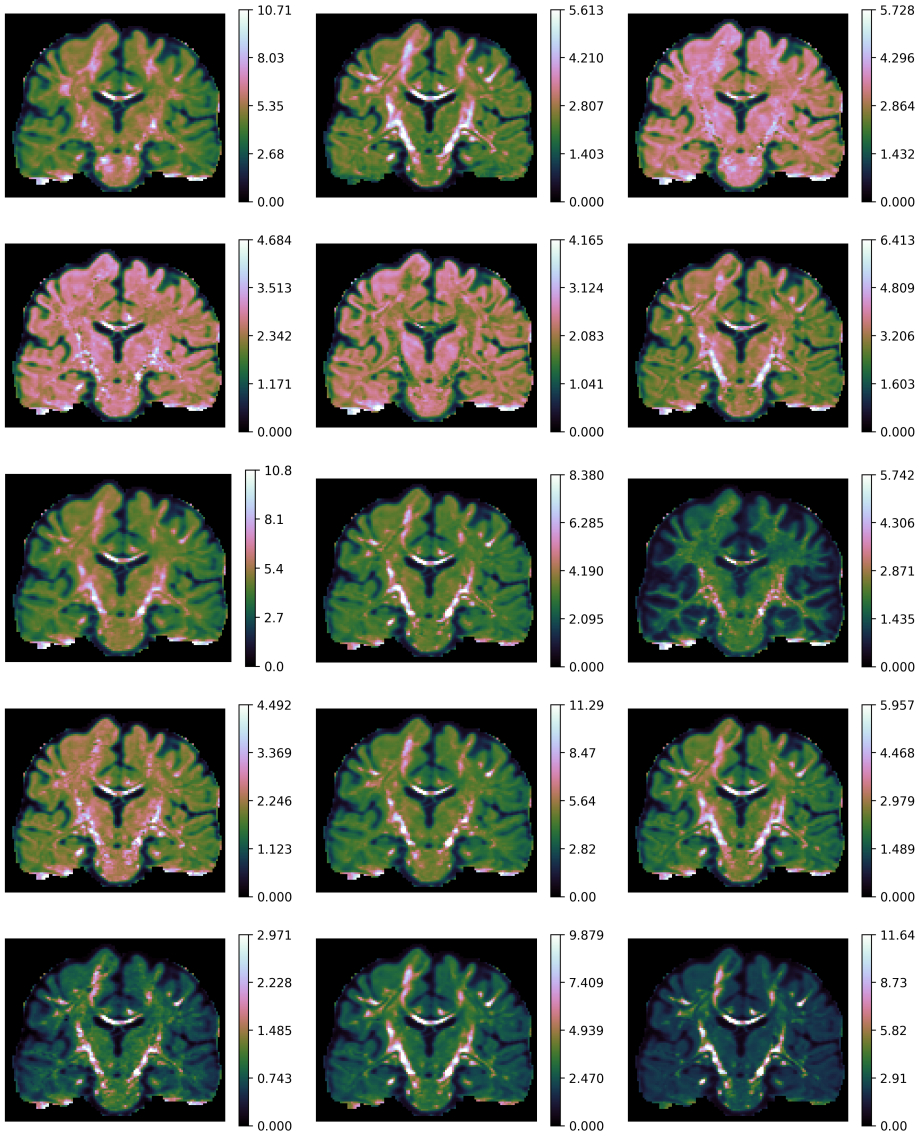

Fig. 4. The 15 autoencoder microstructural features obtained from the 3-shells data for HCP subject 160123.

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.

References

1. Fick, R.H., Wassermann, D., Deriche, R.: The dmipy toolbox: Diffusion mri multi-compartment modeling and microstructure recovery made easy. Frontiers in neuroinformatics **13**, 64 (2019)
2. Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I.: Dipy, a library for the analysis of diffusion mri data. Frontiers in neuroinformatics **8**, 8 (2014)
3. Golkov, V., Dosovitskiy, A., Sperl, J.I., Menzel, M.I., Czisch, M., Sämann, P., Brox, T., Cremers, D.: q-space deep learning: Twelve-fold shorter and model-free diffusion mri scans. IEEE Transactions on Medical Imaging **35**(5), 1344–1351 (2016). https://doi.org/10.1109/TMI.2016.2551324
4. Homeier, H.H., Steinborn, E.: Some properties of the coupling coefficients of real spherical harmonics and their relation to gaunt coefficients. Journal of Molecular Structure: THEOCHEM **368**, 31 – 37 (1996). https://doi.org/https://doi.org/10.1016/S0166-1280(96)90531-X, proceedings of the Second Electronic Computational Chemistry Conference
5. Lampinen, B., Szczepankiewicz, F., Mårtensson, J., van Westen, D., Sundgren, P.C., Nilsson, M.: Neurite density imaging versus imaging of microscopic anisotropy in diffusion mri: A model comparison using spherical tensor encoding. NeuroImage **147**, 517 – 531 (2017). https://doi.org/https://doi.org/10.1016/j.neuroimage.2016.11.053, http://www.sciencedirect.com/science/article/pii/S105381191630670X
6. Liu, X., Faes, L., Kale, A.U., Wagner, S.K., Fu, D.J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J.R., Schmid, M.K., Balaskas, K., Topol, E.J., Bachmann, L.M., Keane, P.A., Denniston, A.K.: A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital Health **1**(6), e271–e297 (2019). https://doi.org/https://doi.org/10.1016/S2589-7500(19)30123-2, https://www.sciencedirect.com/science/article/pii/S2589750019301232
7. Novikov, D.S., Fieremans, E., Jespersen, S.N., Kiselev, V.G.: Quantifying brain microstructure with diffusion mri: Theory and parameter estimation. NMR in Biomedicine **32**(4), e3998 (2019). https://doi.org/https://doi.org/10.1002/nbm.3998, https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3998, e3998 nbm.3998
8. Novikov, D.S., Kiselev, V.G., Jespersen, S.N.: On modeling. Magnetic Resonance in Medicine **79**(6), 3172–3193 (2018). https://doi.org/https://doi.org/10.1002/mrm.27101, https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27101
9. Novikov, D.S., Veraart, J., Jelescu, I.O., Fieremans, E.: Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion mri. NeuroImage **174**, 518 – 538 (2018). https://doi.org/https://doi.org/10.1016/j.neuroimage.2018.03.006
10. Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., et al.: Advances in diffusion MRI acquisition and processing in the human connectome project. Neuroimage **80**, 125–143 (2013)

11. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. NeuroImage **35**(4), 1459–1472 (2007)
12. Vasilev, A., Golkov, V., Meissner, M., Lipp, I., Sgarlata, E., Tomassini, V., Jones, D.K., Cremers, D.: q-space novelty detection with variational autoencoders. In: Bonet-Carne, E., Hutter, J., Palombo, M., Pizzolato, M., Sepehrband, F., Zhang, F. (eds.) Computational Diffusion MRI. pp. 113–124. Springer International Publishing, Cham (2020)
13. Veraart, J., Novikov, D.S., Fieremans, E.: Te dependent diffusion imaging (teddi) distinguishes between compartmental t2 relaxation times. NeuroImage **182**, 360–369 (2018)
14. Zucchelli, M., Deslauriers-Gauthier, S., Deriche, R.: A computational framework for generating rotation invariant features and its application in diffusion mri. Medical Image Analysis **60**, 101597 (2020). https://doi.org/https://doi.org/10.1016/j.media.2019.101597, `http://www.sciencedirect.com/science/article/pii/S1361841519301379`
15. Zucchelli, M., Deslauriers-Gauthier, S., Deriche, R.: Investigating the effect of dmri signal representation on fully-connected neural networks brain tissue microstructure estimation. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 725–728 (2021). https://doi.org/10.1109/ISBI48211.2021.9434046