A Deep Learning Approach for Intra-Voxel Structure Analysis of Diffusion Weighted Images

Jaime F. Aguayo-González and Mariano Rivera
Centro de Investigación en Matemáticas A.C. (CIMAT), Guanajuato, Mexico
jaime.aguayo@cimat.mx

Objective

We introduce a self-supervised learning method for recovering fiber tract structure within a voxel from diffusion-weighted magnetic resonance images (DW-MRI) data. The proposed method is based on a novel deep learning architecture that uses only synthetic data for training. This work extends a previous graduate thesis written by Ehrlich [1], improving her results by the use of a dedicated architecture that accomplishes better performance with fewer trainable parameters.

Our work shows the capability of our proposal on inferring the intra-voxel structure of fibers. We show that our method competes with some popular models in the quality of its results, and depicts advantages over other deep learning approaches: is light-weighted and the evaluation schema can be parallelized. We also show that our method is robust in the presence of high levels of noise when compared with other methods such as Constrained Spherical Deconvolution (CSD). In addition, we reassure the convenience of using the Earth Moving Distance (EMD) as a metric for comparing the quality of different methods on this task.

Method

Our method builds upon three core components:

- The procedure for generating the training dataset.
- A novel neural network architecture.
- The encoding of the intra-voxel structure into the ground truth labels.

Let us explain each of these elements.

We extend the procedure for generating synthetic data used by Ehrlich [1]. Our data consist of signal neighborhoods of $3 \times 3 \times 3$ voxels, containing information of up to three axonal bundles crossing at random orientations. These data are realistic enough to be the only source for the correct training of our neural network.

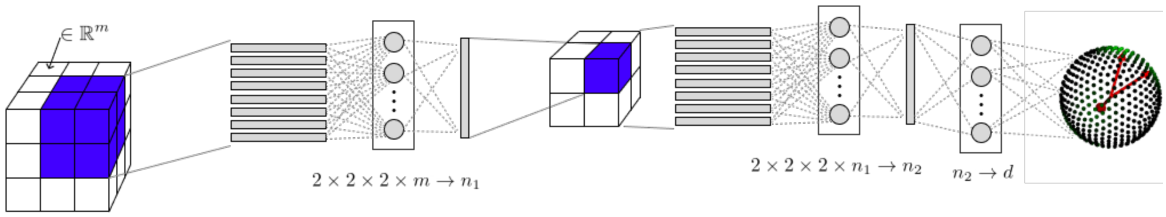

Figure 1: *Proposed model architecture.*

The proposed neural network takes the information of $3 \times 3 \times 3$ patches of voxels to infer the internal structure of the central voxels. The architecture resembles a convolutional neural network in the way layers work, but we substitute the convolutional filters with a small dense layer: a well-studied function approximator. Indeed, the first layer of our network is a dense layer that encodes the information of the 3x3x3 neighborhood into a 2x2x2 voxels neighborhood. This step is done by taking the four local neighborhoods of size 2x2x2 sharing the central voxel and building a descriptor using a dense layer of size n_1 (see Fig. 1). The way this layer works can be seen as a 3D-convolutional layer that uses n_1 filters of size $(2, 2, 2)$ to process data of size $(2, 2, 2)$ with m filters (the number of signals), at a stride of 1. The second layer is analog but takes the previous output as input.

The third core element is the way we generate the ground-truth labels, following the *Gaussian labels* defined by Ehrlich [1]. This representation introduces an additional more sophisticated encoding that codifies a confusion matrix on the orientations that reduce penalizing minor orientation errors.

Results

We performed two experiments. The first experiments focuses on selecting a model by evaluating the performance of networks with a different number of neurons. Table 2 shows the mean squared error (MSE) obtained by the models on an evaluation dataset. The best-ranked model uses a first layer of size 256 and a second layer of size 512. This model packs 1.3 million parameters; a simple Multi-Layer Perceptron capable of getting this level of performance needs 19 million parameters.

	Size of layer 2 (n_2)					
Size of layer 1 (n_1)	**128**	**256**	**512**	**768**	**1024**	**1536**
128	1.99e-05	1.13e-05	1.02e-05	9.42e-06	9.68e-06	1.01e-05
256	3.24e-05	1.18e-05	**9.25e-06**	9.62e-06	9.87e-06	9.98e-06
512	3.83e-05	1.21e-05	1.02e-05	1.01e-05	1.05e-05	1.05e-05
768	3.83e-05	1.23e-05	1.04e-05	1.04e-05	1.06e-05	1.05e-05
1024	5.05e-05	1.26e-05	1.08e-05	1.07e-05	1.05e-05	1.04e-05

Table 2: *Performance our architecture for various number of neurons:* MSE over synthetic-generated signals.

The second experiment consisted of the evaluation of our model using a realistic simulated DW image. The simulated data is based on the fibers of a real brain, using a signal generation model [3] that differs from the training data. We evaluated the quality of our network's predictions against the real orientations and volume fractions using the EMD as an evaluation metric. We contrasted these results against the quality of the ODF predicted by the CSD implementation of the DIPY library [2]. Table 3 summarizes our results.

Noise Level	Proposed method	CSD
Noiseless	0.0582	0.0468
SNR of 30	0.0623	0.1287
SNR of 20	0.0624	0.1337

Table 3: *Performance of our model and CSD method.* EMD is reported; less is better.

Our model predictions almost double the performance of the predictions obtained by the CSD method. Moreover, we observed that the same level of performance is obtained for data with a signal-to-noise ratio of 20 and 30, while the performance of the CSD method deteriorated.

Conclusion

The results we show are encouraging and invite us to use this model for intra-voxel analysis. This procedure may be useful for tractography, especially in those zones with a crossing of axonal bundles. We also discover the potential of neural networks that are trained with synthetic data, eliminating the need for a Golden Standard. We also highlight the convenience of using extra information from the neighboring voxels to predict the structure of a voxel and discuss why we consider the Earth Moving Distance as a suitable metric for this task.

In our research, we explain how our method can be extended for high-resolution images (our current research goes in this direction) and further work on improving the time needed to train the model.

References

[1] Hanna Paola Ehrlich López (2021). Self-Supervised Deep Learning Methods For Intra-Voxel Structure Analysis From Diffusion Weighted Images, Thesis at CIMAT, Guanajuato, Mexico.

[2] Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, Nimmo-Smith I and Dipy Contributors (2014). DIPY, a library for the analysis of diffusion MRI data. Frontiers in Neuroinformatics, vol.8, no.8.

[3] Neher PF, Laun FB, Stieltjes B, Maier-Hein KH. Fiberfox: facilitating the creation of realistic white matter software phantoms. Magn Reson Med. 2014 Nov;72(5):1460-70. doi: 10.1002/mrm.25045. Epub 2013 Dec 9. PMID: 24323973.