Reinforcement Learning Based Tractography With SO(3) Equivariant Agents

Fabian Sinzinger* FABIANSI@KTH.SE and **Rodrigo Moreno** RODMORE@KTH.SE
Division of Biomedical Imaging, Department of Biomedical Engineering and Health Systems, KTH Royal Institute of Technology, Stockholm, Sweden.

Keywords: Tractography, Streamlines, Reinforcement Learning, Spherical CNN, SO(3) Equivariance

1. Introduction

Diffusion MRI-based tractography is a promising noninvasive brain connectivity and neural pathway approximation method (Sotiropoulos et al. (2010)). The ill-posed nature of the problem, Mangin et al. (2013), and a lack of biologically reliable ground truth damages the credibility of automatically generated tractograms. Instead of relying on only a single voxel during the streamline tracking, incorporating neighboring regions has improved tractography Rowe et al. (2013). Recent research suggests a data-driven approach to learning the complex relationship between tractography update directions and local neighborhoods. This work extends the reinforcement learning framework Tract-to-learn (TTL) proposed by Théberge et al. (2021) by integrating spherical harmonics CNN (Ha and Lyu (2022); Cohen et al. (2018); Esteves et al. (2017)) into the actor models. While the original TTL does not preserve the inherent rotational equivariance between the input state and the predicted output (figure 1), our method enforces this equivariance by design. This increases the reliability of the method when exposed to unseen data by decoupling the directional dependence of the predictions from the training data and also improves the explainability of the prediction.

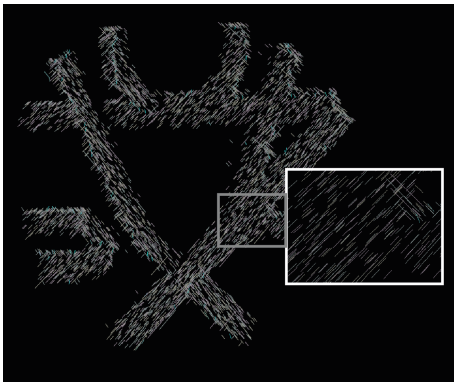

Figure 1: Streamlines from Tract-to-learn (Théberge et al. (2021)) at an early stage of the training (epoch 5). It's apparent that the predictions collapsed towards a few global directions.

2. Methods

2.1. Tractography by Reinforcement Learning

During training, the environment emits rewards to the agent as a reaction to previous actions. In RL-based tractography, the rewards usually guide the agents towards streamlines

* corresponding author

that follow the underlying diffusion MRI data. The rationale here is that the agent maximizes the cumulative rewards received during an episode (i.e., from the start to the end of a streamline) by selecting a feasible action based on a current state. The state is modeled here by a local subset of the input data.

2.2. State: Spherical Local Neighbourhood

We assume here that this data is provided as fiber orientational distribution functions (fodf) at every voxel of the MRI image. In this sense, a dMRI signal can be represented by a function $f_{\text{dMRI}} : \mathbb{Z}^3 \rightarrow \text{S2}$. More precisely, the fodfs are represented in the spherical harmonic basis (SpHarm). The state is formed by a local neighborhood around the current particle location. Therefore, the receptive field of the agent is limited to nearby fodfs. Note that f_{dMRI} as a discrete signal only defines values for the center of each voxel. Since the agent can traverse space continuously, some aggregation scheme is required to assemble a local signal from nearby voxels. Instead of the bilinear interpolation/nearest neighbor method, we apply a custom aggregation method as follows. A bounded cubical region is taken around the current location. The fodfs in this region are then weighted based on their distance to the current particle location. Signals close to the center contribute the most to the estimation of the next direction. Local noise, irregularities, and well-known problems of tractography methods in crossing regions require information about the more distant signal components. We, therefore, propose to weight the signals by a gaussian pdf centered at different distances and assemble the average of the resulting spHarm signals as individual input channels for the actor models (figure 2). Effectively, we decompose the neighborhood signals into different spherical shells. In addition, we append one more spHarm channel to the model inputs to indicate the direction of the previous location of the agent. While this disrespects the strict Markov property of RL, it is required to ensure that the agent can navigate with antipodal fodfs (in other words, this provides information on which way is forward and backward).

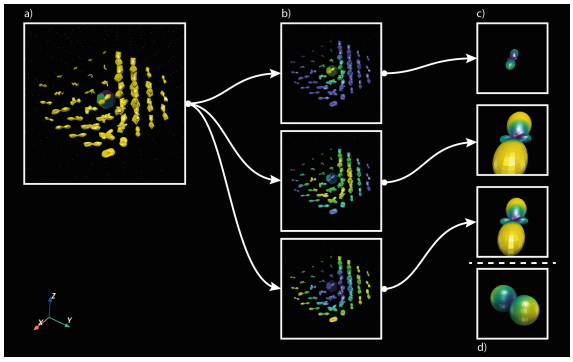

Figure 2: Illustration of the state assembly. A local neighborhood (a) within the diffusion image is weighted (b) dependent on the distance to the center. The fods are then accumulated and form the state (c) together with a spHarm function indicating the previous direction.

2.3. Action: Spherical Harmonics to Directional Update

The central motivation of this work is the exploitation of the rotational equivariance relationship between the spherical state and the direction of the action. Therefore, we train a spherical harmonics model based on the implementation by Ha and Lyu (2022). The model receives the state and outputs a spHarm signal of degree 1. The location of the maximum of

Figure 3: Tractograms of the proposed method after 100 epochs. Left: original tractogram from the model training. Left to right: inferences after rotating the input image by 0, 90, 180, and 270 degrees.

this output signal is then translated into a cartesian vector for the direction of the next step. Moreover, by following the work of Théberge et al. (2021), we employ an adapted version of Twin-Delayed Deep-Deterministic Policy Gradient (TD3) (Fujimoto et al. (2018)).

3. Results

In order to access the rotational equivariance property of the method, we train a model on the "Fiber Cup" dataset (Côté et al. (2013)) and test it with rotated versions. Preliminary qualitative (figure 3) and quantitative (table 1) evaluations support the thesis that our pipeline indeed inherits the rotational equivariance from the spherical models in use.

Table 1: Averaged streamline length in mm of tractograms from the proposed method.

Rotation	Avg. Length
0 degree (training)	95.18 +/- 21.48
90 degree (testing)	95.24 +/- 22.96
180 degree (testing)	95.15 +/- 22.11
270 degree (testing)	96.24 +/- 22.95

4. Discussion

This work established a proof of concept of how rotational equivariance can be integrated into reinforcement learning-based tractography via spherical harmonical CNNs. We disclaim here that this project is still ongoing at the moment of writing, and the results should therefore be considered preliminary. One limitation regards the data that we used for the current experiments. While the evaluated fiber cup phantom serves as a valuable test case during method design, more sophisticated experiments or clinical dMRI brain images are expedient for an actual performance assessment. Another limitation of the current framework concerns the reward function of the RL environment. Concretely, models were reinforced based on three criteria of the resulting tractogram, namely 1. the alignment of the predicted direction with the underlying dMRI data, 2. the smoothness of the streamline, and 3. the total length of the streamline. However, a tractogram that is optimal concerning those three criteria is not necessarily desirable for actual medical problems. In the following stages of this project, we, therefore, intend to introduce a reconstruction reward that measures the capability of the method to produce tractograms that can be mapped back to the original dMRI images, inspired by (Christiaens et al. (2015)) among others.

5. References

References

Daan Christiaens, Marco Reisert, Thijs Dhollander, Stefan Sunaert, Paul Suetens, and Frederik Maes. Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model. *NeuroImage*, 123:89–101, 12 2015. ISSN 1053-8119. doi: 10.1016/J.NEUROIMAGE.2015.08.008.

Taco S. Cohen, Mario Geiger, Jonas Koehler, and Max Welling. Spherical CNNs. In *6th International Conference on Learning Representations (ICLR)*, 2018.

Marc Alexandre Côté, Gabriel Girard, Arnaud Boré, Eleftherios Garyfallidis, Jean Christophe Houde, and Maxime Descoteaux. Tractometer: towards validation of tractography pipelines. *Medical image analysis*, 17(7):844–857, 10 2013. ISSN 1361-8423. doi: 10.1016/J.MEDIA.2013.03.009. URL https://pubmed.ncbi.nlm.nih.gov/23706753/.

Carlos Esteves, Christine Allen-Blanchette, Ameesh Makadia, and Kostas Daniilidis. Learning SO(3) Equivariant Representations with Spherical CNNs. 11 2017. URL http://arxiv.org/abs/1711.06721.

Scott Fujimoto, Herke Van Hoof, and David Meger. Addressing Function Approximation Error in Actor-Critic Methods. *35th International Conference on Machine Learning, ICML 2018*, 4:2587–2601, 2 2018. doi: 10.48550/arxiv.1802.09477. URL https://arxiv.org/abs/1802.09477v3.

Seungbo Ha and Ilwoo Lyu. SPHARM-Net: Spherical Harmonics-based Convolution for Cortical Parcellation. *IEEE Transactions on Medical Imaging*, pages 1–1, 2022. ISSN 0278-0062. doi: 10.1109/TMI.2022.3168670.

J. F. Mangin, P. Fillard, Y. Cointepas, D. Le Bihan, V. Frouin, and C. Poupon. Toward global tractography. *NeuroImage*, 80:290–296, 10 2013. ISSN 1095-9572. doi: 10.1016/J.NEUROIMAGE.2013.04.009. URL https://pubmed.ncbi.nlm.nih.gov/23587688/.

Matthew Rowe, Hui Gary Zhang, Neil Oxtoby, and Daniel C. Alexander. Beyond crossing fibers: tractography exploiting sub-voxel fibre dispersion and neighbourhood structure. *Information processing in medical imaging : proceedings of the ... conference*, 23:402–413, 2013. ISSN 1011-2499. doi: 10.1007/978-3-642-38868-2{_}34. URL https://pubmed.ncbi.nlm.nih.gov/24683986/.

Stamatios N. Sotiropoulos, Li Bai, Paul S. Morgan, Cris S. Constantinescu, and Christopher R. Tench. Brain tractography using Q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach. *NeuroImage*, 49(3):2444–2456, 2 2010. ISSN 1053-8119. doi: 10.1016/J.NEUROIMAGE.2009.10.001.

Antoine Théberge, Christian Desrosiers, Maxime Descoteaux, and Pierre Marc Jodoin. Track-to-Learn: A general framework for tractography with deep reinforcement learning. *Medical Image Analysis*, 72:102093, 8 2021. ISSN 1361-8415. doi: 10.1016/J.MEDIA.2021.102093.