Diffusion-Simulated Connectivity Challenge: Structured description of the challenge design

CHALLENGE ORGANIZATION

Title

Use the title to convey the essential information on the challenge mission.

Diffusion-Simulated Connectivity Challenge

Challenge acronym

Preferable, provide a short acronym of the challenge (if any).

DiSCo

Challenge abstract

Provide a summary of the challenge purpose. This should include a general introduction in the topic from both a biomedical as well as from a technical point of view and clearly state the envisioned technical and/or biomedical impact of the challenge.

The methodological development in the mapping of the brain structural connectome from diffusion magnetic resonance imaging (dMRI) has raised many hopes in the neuroscientific community. Indeed, the knowledge of the connections between different brain regions is fundamental for studying brain anatomy and its function (for instance via electroencephalography or functional MRI). The reliability of the structural connectome is therefore of paramount importance. In the search for accuracy, researchers have given particular attention to linking the white matter tractography methods used for generating the connectome with information about the microstructure of the nervous tissue. This "quantitative connectome" has shown promising results (e.g. Daducci et al., 2014, Smith et al., 2015). However, the lack of realistic numerical phantoms hindered the development and validation of methods in this framework. While previous phantoms were dedicated to either validate tractography or microstructure, a better assessment of the reliability of the connectome estimation on the one hand and its adherence to the actual microstructure of the nervous tissue, on the other hand, is necessary.

The main goal of this challenge is to evaluate the performance of quantitative connectivity methods. For this scope, we have designed three datasets of simulated diffusion-weighted images from three numerical phantoms and an evaluation framework for quantitative tractography. The phantoms are composed of a large collection of synthetic tubular fibers with diameters ranging from 1.4um to 4.2um (approximately 13,000 fibers), connecting distant Regions of Interest (ROIs). The simulation substrates have a micrometric resolution and an unprecedented size of 1 cubic millimeter to mimic an image acquisition matrix of 40x40x40 voxels. Within each voxel, the signal is simulated using Monte-Carlo simulations of spins dynamics using Monte Carlo sampling with a density of one sample per micrometer cube (1,000,000,000 samples total). This is the first time this technique was used to create phantoms of such size and complexity. After the Monte-Carlo simulation of the signal, the image is upscaled by a factor of 100, resulting in a final image size of 10cm and a voxel size of 2.5mm isotropic, compatible with conventional diffusion tractography methods. The simulated images capture the microscopic properties of the tissue (e.g. fiber diameter, water diffusing within and around fibers, free water compartment), while also having

desirable macroscopic properties resembling the anatomy, such as the smoothness of the fiber trajectories. Each phantom has 16 ROIs, forming 120 possible connections (distinct pairs of ROIs). The participants will be provided with a label map of the ROIs defining the connectivity endpoints and will have to submit a weighted connectivity matrix. The estimated connection strength between any two ROIs will be compared to the ground truth total cross-sectional area of the axon-like structures connecting both ROIs. Participants will have access to one phantom for training and one phantom for validation. Participants will have to submit their estimated connectivity on a third testing phantom.

For all three phantoms, participants will have access to the noisy diffusion signal for various diffusion acquisition parameters, a mask of the white matter volume, and a label map of the endpoint connectivity. The diffusion protocol includes 360 diffusion-weighted images and 4 non-diffusion-weighted images (b=0s/mm2). The diffusion-weighted measurements are distributed over 4 b-shells (b=1000, 1925, 3094, 13191 s/mm2). Those correspond to the 3 b-shells of the ActiveAx protocol (Alexander et al., 2010) with an additional shell at b=1000s/mm2 (echo time of 0.0535s). Each shell is sampled using 90 uniformly distributed gradient directions on the sphere. For the training phantom, participants will have access to the map of intra-/extra-fiber signal fractions, the noise-less diffusion signal, the ground truth connectivity matrix, the distribution of simulated fiber diameters and trajectories, and the evaluation script to train their connectivity estimation method. Participants will also have access to a validation dataset with its corresponding connectivity matrix to evaluate the performance of their method.

Ideally, quantitative connectivity estimation for pairs of ROIs with few interconnecting synthetic fibers should be weaker than for pairs of areas with several interconnecting synthetic fibers. An evaluation and ranking of the submission in this sense could prompt the improvement of tractography methods for quantitative structural connectivity estimation in clinical or neuroscientific settings, where we expect to find connectivity estimates directly related to the number of interconnecting axons. Accordingly, participants have to prepare their submission using relevant connectivity pipelines and refrain from developing new methods tailored to the datasets of the challenge. Overall, the challenge will provide unique datasets and analysis to foster the development of tractography and connectivity estimation methods using the rich information available from the complex microstructural organization of axons.

Challenge keywords

List the primary keywords that characterize the challenge.

Diffusion MRI, Tractography, Structural Connectivity, Monte-Carlo simulations, Microstructure

Year

The challenge will take place in …

2021

FURTHER INFORMATION FOR MICCAI ORGANIZERS

Workshop

If the challenge is part of a workshop, please indicate the workshop.

CDMRI

Duration

How long does the challenge take?

Half day.

Expected number of participants

Please explain the basis of your estimate (e.g. numbers from previous challenges) and/or provide a list of potential participants and indicate if they have already confirmed their willingness to contribute.

40-100 participants, 10-20 teams/labs

The ISMRM 2015 tractography challenge had twenty participating teams/labs (~85 participants). The 3D-VoTEM tractography challenge 2018 (ISBI) had nine participating teams/labs (~40 participants). The Irontract challenge 2019 (MICCAI) had twelve participating teams/labs (~45 participants). The proposed challenge targets the same potential participants.

Publication and future plans

Please indicate if you plan to coordinate a publication of the challenge results.

We plan to publish the detailed result analysis in a journal publication following the challenge. All participants will be invited to contribute to the writing of the paper and added to the co-authors. The training, validation and testing datasets will remain available on the challenge website for researchers to experiment with.

Space and hardware requirements

Organizers of on-site challenges must provide a fair computing environment for all participants. For instance, algorithms should run on the same computing platform provided to all.

This is an online challenge. The evaluation script will be provided to the participants to evaluate their results. The testing dataset results will be submitted through the challenge website.

TASK: Quantitative connectivity estimation

SUMMARY

Abstract

Provide a summary of the challenge purpose. This should include a general introduction in the topic from both a biomedical as well as from a technical point of view and clearly state the envisioned technical and/or biomedical impact of the challenge.

See the challenge abstract.

Keywords

List the primary keywords that characterize the task.

Diffusion MRI, Tractography, Structural Connectivity, Monte-Carlo simulations, Microstructure

ORGANIZATION

Organizers

a) Provide information on the organizing team (names and affiliations).

Gabriel Girard, University Hospital Center (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland
Emmanuel Caruyer, Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U-1228, Rennes, France
Jonathan Rafael-Patino, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland
Marco Pizzolato, Technical University of Denmark, Kongens Lyngby, Denmark
Raphaël Truffet, Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U-1228, Rennes, France
Jean-Philippe Thiran, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

b) Provide information on the primary contact person.

Gabriel Girard, gabriel.girard@epfl.ch

Life cycle type

Define the intended submission cycle of the challenge. Include information on whether/how the challenge will be continued after the challenge has taken place.Not every challenge closes after the submission deadline (one-time event). Sometimes it is possible to submit results after the deadline (open call) or the challenge is repeated with some modifications (repeated event).

Examples:

- One-time event with fixed conference submission deadline
- Open call (challenge opens for new submissions after conference deadline)
- Repeated event with annual fixed conference submission deadline

One time event with fixed submission deadline.

Challenge venue and platform

a) Report the event (e.g. conference) that is associated with the challenge (if any).

MICCAI.

b) Report the platform (e.g. grand-challenge.org) used to run the challenge.

n/a

c) Provide the URL for the challenge website (if any).

http://hardi.epfl.ch/challenge

Participation policies

a) Define the allowed user interaction of the algorithms assessed (e.g. only (semi-) automatic methods allowed).

Semi automatic, Fully interactive, Fully automatic.

b) Define the policy on the usage of training data. The data used to train algorithms may, for example, be restricted to the data provided by the challenge or to publicly available data including (open) pre-trained nets.

No policy defined.

c) Define the participation policy for members of the organizers' institutes. For example, members of the organizers' institutes may participate in the challenge but are not eligible for awards.

May participate but not eligible for awards and not listed in leaderboard.

d) Define the award policy. In particular, provide details with respect to challenge prizes.

No challenge prizes are currently planned. The leaderboad and scores will be publicly available after the challenge event.

e) Define the policy for result announcement.

Examples:

- Top 3 performing methods will be announced publicly.
- Participating teams can choose whether the performance results will be made public.

Participating teams will be announced publicly. Participating teams may withdraw from the challenge at any time.

f) Define the publication policy. In particular, provide details on …

- … who of the participating teams/the participating teams' members qualifies as author
- … whether the participating teams may publish their own results separately, and (if so)
- … whether an embargo time is defined (so that challenge organizers can publish a challenge paper first).

All participating teams' members will qualify as co-author.
Teams must include with their submission a contribution statement for all team members. Participating teams may not publish their results before the publication of the challenge results. All data and evaluation code will be made publicly available after the publication of the challenge results.

Submission method

a) Describe the method used for result submission. Preferably, provide a link to the submission instructions.

Examples:

- Docker container on the Synapse platform. Link to submission instructions: <URL>
- Algorithm output was sent to organizers via e-mail. Submission instructions were sent by e-mail.

Algorithm output and description will be sent to organizers via an online form. Submission instructions will be

available on the challenge website.

b) Provide information on the possibility for participating teams to evaluate their algorithms before submitting final results. For example, many challenges allow submission of multiple results, and only the last run is officially counted to compute challenge results.

Teams will need to register before submitting their results.
Teams will be able to evaluate the performance of their algorithm on a validation dataset.
Teams may submit up to 10 results (connectivity matrices) for the challenge. Only the submission with the highest score will be used for ranking.

Challenge schedule

Provide a timetable for the challenge. Preferably, this should include

- the release date(s) of the training cases (if any)
- the registration date/period
- the release date(s) of the test cases and validation cases (if any)
- the submission date(s)
- associated workshop days (if any)
- the release date(s) of the results

Release date of the training dataset: April 1st, 2021
Release date of the validation and test datasets: June 1st, 2021
Registration date: by September 1st, 2021
Submission date: by September 13th, 2021

Ethics approval

Indicate whether ethics approval is necessary for the data. If yes, provide details on the ethics approval, preferably institutional review board, location, date and number of the ethics approval (if applicable). Add the URL or a reference to the document of the ethics approval (if available).

n/a

Data usage agreement

Clarify how the data can be used and distributed by the teams that participate in the challenge and by others during and after the challenge. This should include the explicit listing of the license applied.

Examples:

- CC BY (Attribution)
- CC BY-SA (Attribution-ShareAlike)
- CC BY-ND (Attribution-NoDerivs)
- CC BY-NC (Attribution-NonCommercial)
- CC BY-NC-SA (Attribution-NonCommercial-ShareAlike)
- CC BY-NC-ND (Attribution-NonCommercial-NoDerivs)

CC BY NC.

Code availability

a) Provide information on the accessibility of the organizers' evaluation software (e.g. code to produce rankings). Preferably, provide a link to the code and add information on the supported platforms.

Evaluation code will be available on the challenge website.

b) In an analogous manner, provide information on the accessibility of the participating teams' code.

Teams may or may not distribute the code used to generate their results. To validate their submission, each team must submit a text document describing the preprocessing of the diffusion signal, the local reconstruction, the tractography and the connectivity methods they have used. The document must also include all relevant references to methods and softwares used for the submission. A template will be provided to harmonise the methodological description across teams.

Conflicts of interest

Provide information related to conflicts of interest. In particular provide information related to sponsoring/funding of the challenge. Also, state explicitly who had/will have access to the test case labels and when.

no conflicts of interest to disclose.

MISSION OF THE CHALLENGE

Field(s) of application

State the main field(s) of application that the participating algorithms target.

Examples:

- Diagnosis
- Education
- Intervention assistance
- Intervention follow-up
- Intervention planning
- Prognosis
- Research
- Screening
- Training
- Cross-phase

Research.

Task category(ies)

State the task category(ies).

Examples:

- Classification
- Detection
- Localization
- Modeling
- Prediction
- Reconstruction
- Registration
- Retrieval
- Segmentation
- Tracking

Prediction.

Cohorts

We distinguish between the target cohort and the challenge cohort. For example, a challenge could be designed around the task of medical instrument tracking in robotic kidney surgery. While the challenge could be based on ex vivo data obtained from a laparoscopic training environment with porcine organs (challenge cohort), the final biomedical application (i.e. robotic kidney surgery) would be targeted on real patients with certain characteristics defined by inclusion criteria such as restrictions regarding sex or age (target cohort).

a) Describe the target cohort, i.e. the subjects/objects from whom/which the data would be acquired in the final biomedical application.

The challenge will evaluate methods designed for in-vivo brain structural connectivity estimation.

b) Describe the challenge cohort, i.e. the subject(s)/object(s) from whom/which the challenge data was acquired.

The geometrical design of the simulation phantoms was done through an iterative optimization process starting from an overlapping set of large tubular fibers (30um of diameter), with both endpoints located in regions of interest on the surface of a sphere of 1mm in diameter. Each fiber was set to follow a trajectory of an arc of a circle going through the center of the sphere. Those fibers were slowly separated using energy terms controlling for length, curvature, and overlap with other fibers. This optimization was done using the Numerical Fiber Generator (Close et al., 2009). Then, the tubular fibers were resampled with a smaller diameter (15um), and their trajectories re-optimized. Finally, a set of gamma-distributed diameters were sampled and placed within each 15um-fiber to form the final set of fibers (~13,000). Those were optimized for an additional iteration to further remove overlapping fibers.

The final set of fibers (trajectories and diameters) were used to generate a mesh of the substrate. For each fiber, a tubular mesh was generated to represent the outer surface of the axon-like structure. An additional inner tubular mesh following the same trajectory but with a diameter of 0.7 times the outer diameter, representing the inner surface of the fiber. The meshes were used as input to the Monte-Carlo simulator. All samples initiated inside the inner-mesh were considered as "intra-axonal" water, those within the outer-mesh but outside the inner mesh were considered "myelin" water, and those outside the outer-mesh were considered "extra-axonal". The simulation parameters for the "intra-axonal" and "extra-axonal" water were identical. The samples initiated inside the "myelin" compartment did not contribute to the generated signal (this is expected for diffusion MRI protocol). The Monte Carlo sampling was done using a density of one sample per micrometer cube (1,000,000,000 samples total per substrate).

All datasets are generated using the same substrate optimization procedure and simulated using the same microstructure parameters and properties, but simply using a different seed for the random generation of connectivity matrices. Although the connectivity matrices have similar controlled sparsity levels, the distribution of the connectivity weights between pairs of ROIs in the matrices are different. This will inevitably affect the fiber trajectory optimization process done using the Numerical Fiber Generator since, for each dataset, the initial fiber configurations that reflect the connectivity matrix are different. We kept the optimization weights and convergence threshold identical for all datasets. Synthetic fiber diameters are also drawn from the same distribution for all datasets.

The three simulation substrates designed for the challenge have a micrometric resolution and an unprecedented size of 1 cubic millimeter to mimic an image acquisition matrix of 40x40x40 voxels. This is the first time this technique was used to create phantoms of such size and complexity. After the Monte-Carlo simulation of the signal, the image is upscaled by a factor of 100, resulting in a final image size of 10cm and a voxel size of 2.5mm isotropic, compatible with conventional diffusion tractography methods. The simulated images capture the microscopic properties of the tissue (e.g. fiber diameter, water diffusing within and around fibers, free water compartment), while also having desirable macroscopic properties resembling the anatomy, such as the smoothness of the fiber trajectories.

Imaging modality(ies)

Specify the imaging technique(s) applied in the challenge.

Diffusion MRI

Context information

Provide additional information given along with the images. The information may correspond …

a) … directly to the image data (e.g. tumor volume).

n/a

b) … to the patient in general (e.g. sex, medical history).

n/a

Target entity(ies)

a) Describe the data origin, i.e. the region(s)/part(s) of subject(s)/object(s) from whom/which the image data would be acquired in the final biomedical application (e.g. brain shown in computed tomography (CT) data, abdomen shown in laparoscopic video data, operating room shown in video data, thorax shown in fluoroscopy video). If necessary, differentiate between target and challenge cohort.

All data are simulated.

b) Describe the algorithm target, i.e. the structure(s)/subject(s)/object(s)/component(s) that the participating algorithms have been designed to focus on (e.g. tumor in the brain, tip of a medical instrument, nurse in an operating theater, catheter in a fluoroscopy scan). If necessary, differentiate between target and challenge cohort.

The challenge will test methods designed for in-vivo brain structural connectivity estimation. These include algorithms for fiber orientation distribution function estimation algorithms (e.g. spherical deconvolution methods),

and tractography algorithms (e.g. probabilistic tractography), clustering and pruning of tracts (e.g. outlier removal), and algorithms for the calculation of connectivity (e.g. length-scaling). The challenge aims at improving clinical and neuroscientific applications of structural connectivity estimated with diffusion MRI. Thus, participants should prepare their submission using clinically relevant connectivity pipelines and refrain from developing new methods tailored to the datasets of the challenge.

Assessment aim(s)

Identify the property(ies) of the algorithms to be optimized to perform well in the challenge. If multiple properties are assessed, prioritize them (if appropriate). The properties should then be reflected in the metrics applied (see below, parameter metric(s)), and the priorities should be reflected in the ranking when combining multiple metrics that assess different properties.

- Example 1: Find highly accurate liver segmentation algorithm for CT images.
- Example 2: Find lung tumor detection algorithm with high sensitivity and specificity for mammography images.

Corresponding metrics are listed below (parameter metric(s)).

Accuracy, Specificity, Precision, Sensitivity, Robustness.

DATA SETS

Data source(s)

a) Specify the device(s) used to acquire the challenge data. This includes details on the device(s) used to acquire the imaging data (e.g. manufacturer) as well as information on additional devices used for performance assessment (e.g. tracking system used in a surgical setting).

Data simulation

b) Describe relevant details on the imaging process/data acquisition for each acquisition device (e.g. image acquisition protocol(s)).

The diffusion protocol includes 360 diffusion-weighted images and 4 non-diffusion-weighted images (b=0). The diffusion-weighted measurements are distributed on 4 b-shells (b=1000, 1925, 3094, 13191 s/mm2). Those correspond to the 3 b-shells of the ActiveAx protocol (Alexander et al., 2010) with an additional shell at b=1000s/mm2 (echo time of 0.0535s). Each shell is sampled using 90 uniformly distributed gradient directions on the sphere.

c) Specify the center(s)/institute(s) in which the data was acquired and/or the data providing platform/source (e.g. previous challenge). If this information is not provided (e.g. for anonymization reasons), specify why.

The datasets were simulated at École Polytechnique Fédérale de Lausanne (Lausanne, Switzerland) using the high performance computing resources.

d) Describe relevant characteristics (e.g. level of expertise) of the subjects (e.g. surgeon)/objects (e.g. robot) involved in the data acquisition process (if any).

n/a

Training and test case characteristics

a) State what is meant by one case in this challenge. A case encompasses all data that is processed to produce one result that is compared to the corresponding reference result (i.e. the desired algorithm output).

Examples:

- Training and test cases both represent a CT image of a human brain. Training cases have a weak annotation (tumor present or not and tumor volume (if any)) while the test cases are annotated with the tumor contour (if any).
- A case refers to all information that is available for one particular patient in a specific study. This information always includes the image information as specified in data source(s) (see above) and may include context information (see above). Both training and test cases are annotated with survival (binary) 5 years after (first) image was taken.

The training, validation and testing phantoms represent a simulated axonal white matter substrate. For each phantom, participants will have to predict the connectivity scores of 120 pairs of regions of interest. Each phantom is independant, with distinc label map, signal, and network organisation.

b) State the total number of training, validation and test cases.

There is one training, one validation and and one testing dataset/phantom. Each dataset concist of 120 cases (connectivity weight predictions).

Please note that the 120 cases are not statistically independent within one phantom, as they share the same endpoints. Although more phantoms would be an interesting addition, we had to limit the challenge to three phantoms as it required an unprecedented computational time to generate each of them.

c) Explain why a total number of cases and the specific proportion of training, validation and test cases was chosen.

Participants should not design algorithms specifically for the challenge but rather use state-of-the-art methods for in-vivo brain imaging. The training and validation cases are provided for tuning parameters to the diffusion properties of the simulated signal. The test cases follow the same substrate design and simulation procedure but have a different connectivity among the 16 ROIs.

d) Mention further important characteristics of the training, validation and test cases (e.g. class distribution in classification tasks chosen according to real-world distribution vs. equal class distribution) and justify the choice.

The ground truth connectivity matrices of the 16 ROIs were generated with a parameter controlling the sparsity of the 16x16 matrix. This resulted in approximately 20% of the pairs of ROIs connected by at least one tubular fiber (20% positive, 80% negative classes). This was selected to be sparser than the densely connected neighboring region networks, but more densely connected than long-range connection networks. For instance, in Girard et al. (2020), authors reported 39% of connected ROIs of the short-range frontoparietal network of the macaque brain. Maier-Hein et al., (2016) studied the connectivity matrix of long-range connections, with less than 10% of interconnected ROIs.

Annotation characteristics

a) Describe the method for determining the reference annotation, i.e. the desired algorithm output. Provide the information separately for the training, validation and test cases if necessary. Possible methods include manual image annotation, in silico ground truth generation and annotation by automatic methods.

If human annotation was involved, state the number of annotators.

The ground truth is from the in silico phantom.

b) Provide the instructions given to the annotators (if any) prior to the annotation. This may include description of a training phase with the software. Provide the information separately for the training, validation and test cases if necessary. Preferably, provide a link to the annotation protocol.

n/a

c) Provide details on the subject(s)/algorithm(s) that annotated the cases (e.g. information on level of expertise such as number of years of professional experience, medically-trained or not). Provide the information separately for the training, validation and test cases if necessary.

n/a

d) Describe the method(s) used to merge multiple annotations for one case (if any). Provide the information separately for the training, validation and test cases if necessary.

n/a

Data pre-processing method(s)

Describe the method(s) used for pre-processing the raw training data before it is provided to the participating teams. Provide the information separately for the training, validation and test cases if necessary.

All simulated dataset provided to the participants will be corrupted with Rician noise (SNR=20). Participants will have access to the noiseless data of the training dataset.

Sources of error

a) Describe the most relevant possible error sources related to the image annotation. If possible, estimate the magnitude (range) of these errors, using inter-and intra-annotator variability, for example. Provide the information separately for the training, validation and test cases, if necessary.

n/a

b) In an analogous manner, describe and quantify other relevant sources of error.

n/a

ASSESSMENT METHODS

Metric(s)

a) Define the metric(s) to assess a property of an algorithm. These metrics should reflect the desired algorithm properties described in assessment aim(s) (see above). State which metric(s) were used to compute the ranking(s) (if any).

- Example 1: Dice Similarity Coefficient (DSC)
- Example 2: Area under curve (AUC)

1. We will use the Pearson correlation coefficient between the ground truth weighted connectivity matrix and the participant submitted matrix. The value of the ground truth matrix corresponds to the cross-sectional area of the simulated axon between pairs of regions of interest.

2. We will also perform a Receiver Operating Characteristic (ROC) analysis, in particular highlighting the Area Under Curve (AUC) of the method submitted by participants, to evaluate the performances of the submitted connectivity matrices at predicting which connections have non-zero connectivity weights. The AUC won't be

used for the challenge ranking. This will be done after the challenge, for the planned follow-up paper.

b) Justify why the metric(s) was/were chosen, preferably with reference to the biomedical application.

The goal of the challenge is to test which algorithm can better predict the relative connectivity strength of pairs of regions of interest. The Pearson correlation coefficient has been used in the connectivity literature to compare diffusion connectivity estimation with tracers cell labeling (e.g. Delettre, C., Messé, A., Dell, L. A., Foubet, O., Heuer, K., Larrat, B., ... & Borrell, V. (2019). Comparison between diffusion MRI tractography and histological tract-tracing of cortico cortical structural connectivity in the ferret brain. Network Neuroscience, 3(4), 1038-1050.).
The ROC analysis was also used to assess the sensitivity and specificity of connectom estimation (e.g. Thomas, C., Ye, F. Q., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., & Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proceedings of the National Academy of Sciences of the United States of America, 111(46), 16574–16579.)

Ranking method(s)

a) Describe the method used to compute a performance rank for all submitted algorithms based on the generated metric results on the test cases. Typically the text will describe how results obtained per case and metric are aggregated to arrive at a final score/ranking.

The correlation between the ground truth connectivity matrix and the matrix provided by the participants will be evaluated using the Pearson's correlation coefficient, and the R value will be used to rank submissions.

The highest R value of each team submissions (up to 10 per team) will be used for ranking.

b) Describe the method(s) used to manage submissions with missing results on test cases.

All submissions must consist of a full 16x16 connectivity matrix describing the connectivity strength of all regions to all other regions (datasets have 16 ROIs). The submission format will be a text file (CSV) with 16 rows and columns indicating the connectivity values of the corresponding regions of interest.

c) Justify why the described ranking scheme(s) was/were used.

The selected ranking method is easy to interpret and visualize. Further analysis will be performed in the planned paper following the challenge (e.g. ROC, area under curve, mean squared error).

Statistical analyses

a) Provide details for the statistical methods used in the scope of the challenge analysis. This may include

- description of the missing data handling,
- details about the assessment of variability of rankings,
- description of any method used to assess whether the data met the assumptions, required for the particular statistical approach, or
- indication of any software product that was used for all data analysis methods.

n/a

b) Justify why the described statistical method(s) was/were used.

n/a

Further analyses

Present further analyses to be performed (if applicable), e.g. related to

- combining algorithms via ensembling,
- inter-algorithm variability,
- common problems/biases of the submitted methods, or
- ranking variability.

n/a

ADDITIONAL POINTS

References

Please include any reference important for the challenge design, for example publications on the data, the annotation process or the chosen metrics as well as DOIs referring to data or code.

Alexander, D. C., Hubbard, P. L., Hall, M. G., Moore, E. a., Ptito, M., Parker, G. J. M., & Dyrby, T. B. (2010). Orientationally invariant indices of axon diameter and density from diffusion MRI. NeuroImage, 52(4), 1374–1389.

Close, T. G., Tournier, J.-D., Calamante, F., Johnston, L. a, Mareels, I., & Connelly, A. (2009). A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. NeuroImage, 47(4), 1288–1300. https://doi.org/10.1016/j.neuroimage.2009.03.077 parameter choice on the reproducibility of results. Frontiers in Neuroinformatics, 14, 8.

Daducci, A., Dal Palu, A., Lemkaddem, A., & Thiran, J.-P. (2014). COMMIT: Convex Optimization Modeling for Micro-structure Informed Tractography. IEEE Transactions on Medical Imaging, 34(1).

Delettre, C., Messé, A., Dell, L. A., Foubet, O., Heuer, K., Larrat, B., ... & Borrell, V. (2019). Comparison between diffusion MRI tractography and histological tract-tracing of cortico-cortical structural connectivity in the ferret brain. Network Neuroscience, 3(4), 1038-1050.

Girard, G., Caminiti, R., Battaglia-Mayer, A., St-Onge, E., Ambrosen, K. S., Eskildsen, S. F., ... Innocenti, G. M. (2020). On the cortical connectivity in the macaque brain: A comparison of diffusion tractography and histological tracing data. NeuroImage, 221, 117201. https://doi.org/10.1016/j.neuroimage.2020.117201

Maier-Hein, K. H., Neher, P. F., Houde, J. C., Côté, M. A., Garyfallidis, E., Zhong, J., ... & Reddick, W. E. (2017). The challenge of mapping the human connectome based on diffusion tractography. Nature communications, 8(1), 1-13.

Schilling, K. G., Nath, V., Hansen, C., Parvathaneni, P., Blaber, J., Gao, Y., ... & Schiavi, S. (2019). Limits to anatomical accuracy of diffusion tractography using modern approaches. NeuroImage, 185, 1-11.

Rafael-Patino, J., Romascano, D., Ramirez-Manzanares, A., Canales-Rodríguez, E. J., Girard, G., & Thiran, J. P. (2020). Robust Monte-Carlo Simulations in Diffusion-MRI: Effect of the substrate complexity and parameter choice on the reproducibility of results. Frontiers in Neuroinformatics, 14, 8.

Smith, R. E., Tournier, J. D., Calamante, F., & Connelly, A. (2015). SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. NeuroImage, 119, 338–351.

https://doi.org/10.1016/j.neuroimage.2015.06.092

Thomas, C., Ye, F. Q., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., & Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proceedings of the National Academy of Sciences of the United States of America, 111(46), 16574–16579

Further comments

Further comments from the organizers.

The results of the 2015 ISMRM tractography challenge sparked several research projects targeting the reduction of false-positive connections identified by tractography methods (google scholar citations: 505). This challenge proposal builds on this effort by, first, increasing the complexity of the simulated data, and, second, adding a quantitative aspect to the connectivity predictions. The datasets will allow further analysis of the connectivity prediction power of diffusion tractography algorithms. Moreover, researchers will be able to test novel methods addressing the false-positive problem of tractography, on datasets with complex microstructure and macrostructure properties.