Tractography ML

Medical statistics proposal for 2022 Wolfram summer school

Armando Benjamín Cruz Hinojosa
Universidad Nacional Autónoma de México
Mexico City, Mexico
aleph_g@ciencias.unam.mx

Abstract—Current authentication schemes are based on asymmetric cryptography protocols like RSA, with the arrival of cuantum computing and fast factorization algorithms, security is endangered. A-schemes presents an alternative that remains secure regardless of computational power, but there are no tools to measure how secure they are. This proposal intends to create such tool.

I. INTRODUCTION

Tractography is a 3D computational modeling technique that statistically estimates white matter fiber paths on the human brain, using magnetic resonance imaging. The results are arrays of three dimensional trajectories called tractograms fig 1, that can be used to recreate the brain's connection structure in order to understand it's functionality and infer the presence of diseases.
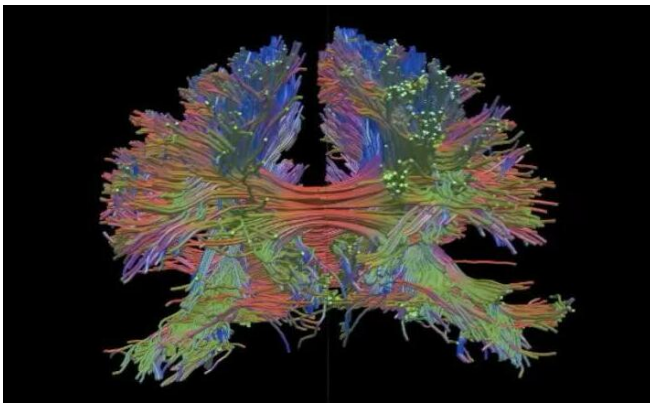

Fig. 1. Tractogram, tractography algorithms main output.

Due to resolution limitations, a tractogram of a brain cannot obtain individual axons, but clusters of axons. Nevertheless each tractogram contains thousands of such clusters, that cannot be processed and filtered for medical research without the help of computers.

The task of determining which clusters are connected in the patient's brain is not a trivial task, since the fibers overlap and the complex structures of the brain's regions. Because of this, all tractography techniques make mistakes and generate noisy tractograms with lots of false positive connections.

Solutions to this problem, based on supervised and unsupervised ML algorithms, has been proposed such as random forest models, recurrent NN, convolutional NN and autoencoders. To date, none of these methods have shown conclusive performance or have been established as a main methodology to follow in the field of tractography.

Last summer I had the oportunity to attend the *Summer in mathematical tools for data sacience*, given at mexican mathematical research center CIMAT. Over one month a team of six students, one instructor and one neurologist (which I was part of) worked on the implementation of four ML proposals in python, to extract of a given tractogram the main brain sturctures.

The training data was curated by Dr Luis Concha from the institute of neurobiology UNAM. From the four models only two were finished, and one correctly trained with satisfactory results. The project came to be named *tractosplit* and is allocated in a public github repository [1].

II. OBJECTIVE

My proposal is to migrate Tractosplit machine learning project from python into wolfram languaje, that implements current tractogram segmentation models and the two unfinished models into an instant API running in Wolfram cloud. Additionally implement a 3D interactive plotting of the segmentated results, inspired on Muhammad Ubadah's last year Wolfram summer school project [2].

III. IMPLEMENTATION

Three things are needed to sucessfully migrate tractosplit and implement 3d interactive plotting.

A. Training

Current data for training is the previously mentioned curated data of 30 patients. Each tractogram streamline is assigned a label of the brain section it belongs to, making it ideal for supervised learning models.

More data can be obtained online from the *International Neuroimaging Data-sharing Initiative* [3] datasets. INDI datasets came from a variety of patients and are curated to study specific diseases like ABIDE: Autism Brain Imaging Data Exchange [4].

B. Model definition

Tractosplit models are a colection of supervised learning architectures whose inputs are tractogram streamlines of size $k \in \{16, 32, 64\}$ (k must be fixed) and return one of 23 labels corresponding to the standard spaces used by XTRACT [1] tractography tool. Current working architectures are

- Long short term memory neural network. This networks was proposed by team's instructure Phd. Jean-Bernard Hayet, due to it's proficency on pedestrian trayectory prediction.
- Autoencoder.

The unfinished models are based on

- Random forest. 3D random forest can be represented as a box segmented by prismatic sections (3-cells), calculus theorems assert that all brain sections can be approximated with such prisms.
- Convolutional neural network. [5] propose a CNN based clasifier that successfully identifies 99% of fibers on avergae.

C. 3D interactive plotting

As mentioned above, segmentation models outputs are a list of several tractogram files, composed of a list of lists size $k \in \{16, 32, 64\}$ of 3D points in brain space.

Each tractogram k-list will be plotted as a trayectory in 3d space, and assigned a color depending on the brain section the it belongs. Interactivity controls will allow the user to move the 3D camera around the brain and focus on specific brain section.

REFERENCES

[1] https://github.com/Aleph-GORY/tractosplit.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.