DETERMINATION OF FIBER TRACTS USING Q-ROUTING METHOD

by

Mert Efe Yıldız

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Biomedical Engineering

Yeditepe University
2023

DETERMINATION OF FIBER TRACTS USING Q-ROUTING METHOD

APPROVED BY:

Assist. Prof. Dr. Andaç Hamamcı ……
(Thesis Supervisor)
(Yeditepe University)

Assoc. Prof. Dr. Gökhan Ertaş ……
(Yeditepe University)

Assist. Prof. Dr. Ali Bayram ……
(İstanbul University)

DATE OF APPROVAL: ..../..../20..

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name ……

Signature …..…….……..

ACKNOWLEDGEMENTS

I would want to express my sincere gratitude to my professor, Andaç Hamamcı for his assistance and support. I also want to express my gratitude to my family for their unending support.

ABSTRACT

DETERMINATION OF FIBER TRACTS USING Q-ROUTING METHOD

In this study, it is suggested to use the Q-routing method to the diffusion MR fiber tractography problem. Within that way, it is expected to achieve the robustness of global methods while still utilizing the advantages of local streamline methods by taking into account the whole cost of the path rather just local orientations. The suggested Q-routing based approach is assessed using the FiberCup and ISMRM 2015 dataset, a diffusion phantom that simulates the fiber connections of the human cerebral white matter. It showed slightly better results in ismrm 2015 dataset. The outcomes demonstrate the method's potential to the diffusion MR fiber tractography issue.

ÖZET

Q-YÖNLENDİRME METODU KULLLANRAK FİBER YOLAKLARIN BELİRNLENMESİ

Bu çalışmada, difüzyon MR fiber traktografi problemine Q-yönlendirme yönteminin kullanılması önerilmektedir. Bu şekilde, sadece yerel doku yönelimleri yerine yolun tüm maliyetini hesaba katarak, yerel akış çizgisi yöntemlerinin avantajlarını kullanmaya devam ederken, küresel yöntemlerin sağlamlığını elde etmesi beklenmektedir. Önerilen Qyönlendirme tabanlı yaklaşım, insan beyin ana ak maddesinin fiber bağlantılarını simüle eden bir difüzyon fantomu olan FiberCup ve ISMRM 2015 veriseti kullanılarak değerlendirilmiştir. Önerilen yöntem ISMRM 2015 veriseti üzerinde biraz daha iyi sonuç vermiştir. Sonuçlar, yöntemin difüzyon MR fiber traktografi sorununa yönelik potansiyelini göstermektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

G	Amplitude
δ	Time of applied gradients
Δ	Duration beween the paired gradients
$\lambda 1, \lambda 2, \lambda 3$	Diffusion tensor eigenvalues

DWI	Diffusion weighted image
MRI	Magnetic resonance imaging
DTI	Diffusion tensor imaging
SH	Spherical harmonic
ODF	Oriented diffusion function
fODF	Fiber orientation diffusion function
FA	Fractional anisotropy
TE	Time to echo
TR	Repetition Time
DMR-FT	Diffusion MR fiber tractography
CPM	Conventional probabilistic method
CDM	Conventional deterministic method
Q-RPTM	Q-routing probabilistic tractography method
Q-RDTM	Q-routing deterministic tractography method
RL	Reinforcement learning

1. INTRODUCTION

Tractography is a crucial scientific tool for biomedical and neuroscience research because it is the sole non-invasive technique to reconstruct brain fiber bundle pathways in vivo. One of its objectives is to generate coherent patterns of fiber orientations from diffusion magnetic resonance imaging data in order to create mathematical models of fiber bundles. However, this is made difficult by the presence of crossing fibers, partial volume effects, and inherent ambiguities as an inverse problem. However, traditional tractography approaches are still necessary for producing high-fidelity training data and for analyzing unique and unusual cases, such as non-human subjects, post-operative cases, etc. Learning-based approaches are evolving as a sensible approach for segmenting fiber bundles across subjects [1]. It is suggested in this study that the Diffusion MR fiber tractography problem be solved using the Q-routing approach which is sub-branch of most used reinforcement learning algorithm [2]. By taking into account the whole cost of the road rather than individual texture orientations at each point, it is intended to attain the robustness of global approaches while still gaining the benefits of local flowline methods.

1.1. DIFFUSION WEIGHTED IMAGING

A type of MR imaging called diffusion-weighted imaging (DWI) gauges the random Brownian motion of water molecules within a tissue voxel [3]. Generally speaking, tissues that are densely cellular or that have cellular swelling have lower diffusion coefficients. Diffusion is very helpful for identifying tumors and cerebral ischemia [4]. MRI pulse sequence diagram is described in Figure 1 to explain DWI easily.

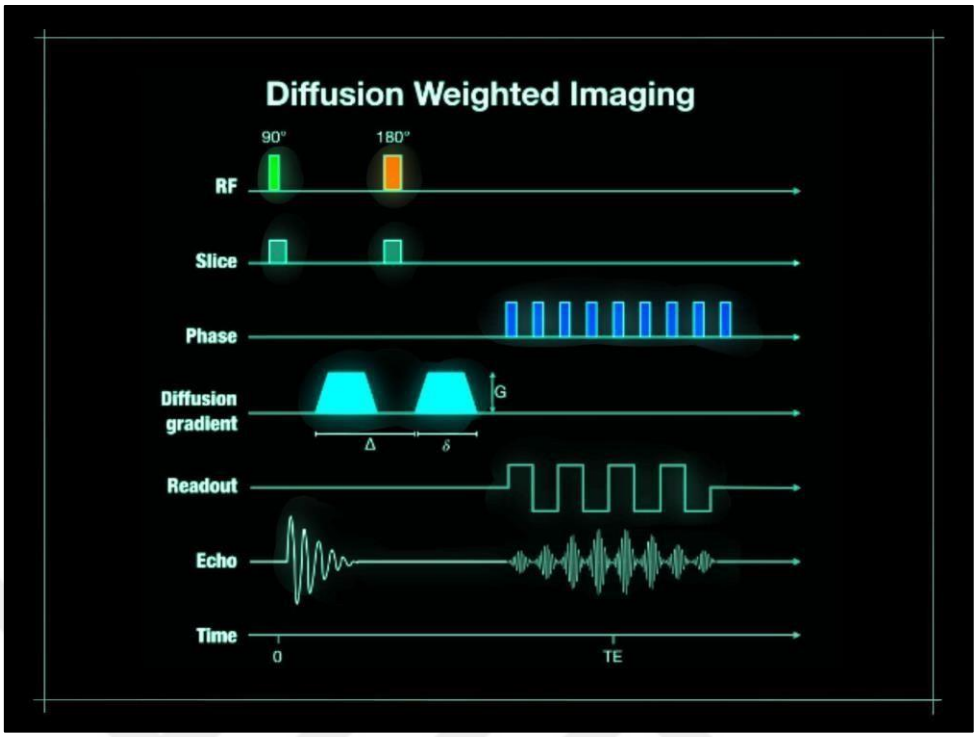

Figure 1.1. MRI pulse sequence diagram

Diffusion-weighted imaging's basic tenet is the attenuation of the T2 signal dependent on how quickly water molecules can diffuse in a given region. Less initial T2 signal will be present the easier water can diffuse, or the farther a water molecule can move during the process. For instance, water in cerebrospinal fluid (CSF) diffuses readily, leaving very little signal and giving the ventricles a dark appearance. In contrast, cell walls in the brain parenchyma prevent water from moving as freely, therefore the brain's original T2 signal is only marginally reduced.In contrast, cell walls in the brain parenchyma prevent water from moving as freely, therefore the brain's original T2 signal is only marginally reduced. The important implication of this is that a region of the brain with no T2 signal cannot display signal on isotropic diffusion-weighted imaging, regardless of the tissue's diffusion properties. The repetition time is the time between subsequent pulse sequences delivered to the same slice (TR). The time (TE) to echo is the elapsed time between the delivery of the RF pulse and the reception of the echo signal. In Figure 2, "short" and "long" are measured in relation to the T1 and T2 values of the problems to demonstrate how the choice of TR and TE might result in different sorts of images [5][6] .

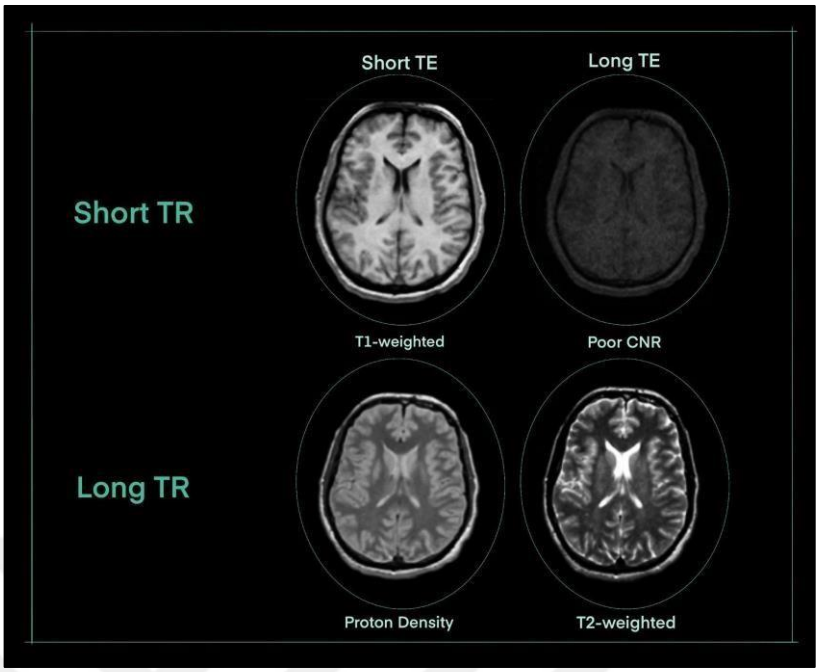

Figure 1.2. Example of T1 and T2 weighted images

To extract diffusion information from the tissue, a T2 weighted image without diffusion attenuation must first be obtained. The b=0 image is what we have here. The ease of water diffusion is then evaluated in a variety of directions; we will utilize the minimum of three orthogonal directions as X, Y, and Z for the remainder of this explanation [7].

This is accomplished by symmetrically applying a strong gradient on either side of the 180degree pulse. The area under the diffusion gradients, which is related to the amplitude and duration of the gradient, as well as the distance between the gradients, are the main determinants of the degree of diffusion weighting. These elements work together to produce the b value. The more significant the signal attenuation caused by diffusion, the higher the number [8].

The first gradient is used to apply phase information to stationary water molecules. However, because they remained in the same place after the 180-degree pulse, the consequences of the initial pulse are completely reversed (since they have flipped 180-degrees). They have therefore kept their signal at the time the echo is generated [9].

On the other hand, moving water molecules pick up phase information from the first gradient, but because they are moving, they are not in the same place when they are exposed

to the second gradient and are not exposed to the exact same gradient after the 180-degree pulse. As a result, they are not rephased and portion of their signal is lost. The more they can move, the more signal will dissipate and the less successfully they will be rephased [10].

By expressing the amplitude (G), time of applied gradients (δ), and duration between the paired gradients (Δ), b value indicates the degree of diffusion weighting applied. It is calculated as described in Formula (1.1).

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \quad (1.1)$$

Therefore, by boosting the gradient's amplitude and duration as well as lengthening the space between its paired gradient pulses, a higher b value is attained [11].

1.2. DIFFUSION TENSOR IMAGING

Instead of using this information alone to assign contrast or color to pixels in a crosssectional image, the magnetic resonance imaging technique known as diffusion tensor imaging (DTI) allows assessment of the limited diffusion of water in tissue to create neural tract images. It also offers helpful structural details on other tissues like muscle, particularly the heart muscle [12].

1.2.1. Spherical harmonic

The input signal provided to the tracking system is made up of spherical harmonic (SH) coefficients that represent fiber ODFs and were derived by a constrained spherical deconvolution, as shown in Fig. 1.3.

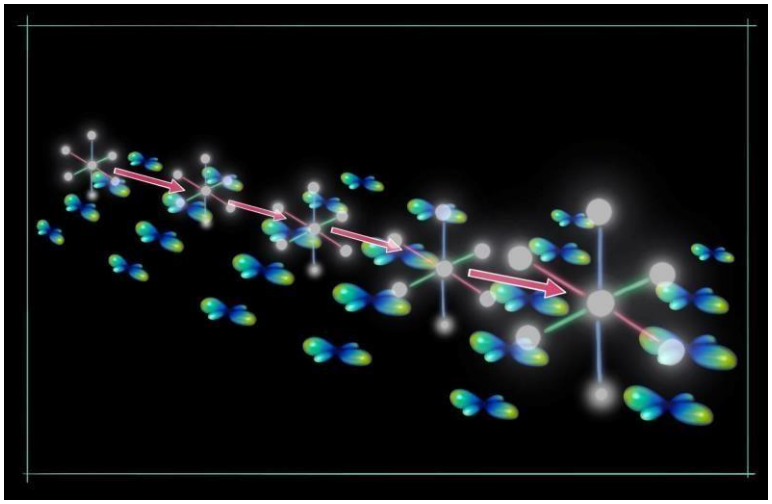

Figure 1.3. Demonstration of the input signal and the directional and surrounding information. fODF glyphs are used to describe SH (spherical harmonic) coefficients. Arrows are used to show the segments of a streamline. The input signal is computed at the locations marked by white circles.

The diffusion rate and the preferred direction of diffusion for each voxel in the DTI have one or more pairs of parameters, both represented in terms of three-dimensional space. The properties of each voxel in a single DTI image are typically calculated using six or more distinct diffusion weighted acquisitions. The diffusion sensitizing gradients were applied in various orientations to obtain each of these acquisitions.

To create a single calculated image data set, some approaches require hundreds of measurements, each of which constitutes an entire image. A DTI voxel's greater information richness makes it highly sensitive to even the most minor forms of brain disease. Tracography is a technique for choosing and following neural routes across the brain that takes advantage of the direction information at a higher level of structure.

Figure 4 provides an explanation of how to create fODF using dMRI measurements. The smoothed diffusion signal X is estimated using discrete dMRI measurements along various b-vectors, and it is then converted into the ODF function and subsequently fODF for structural connectome analysis [13][14].

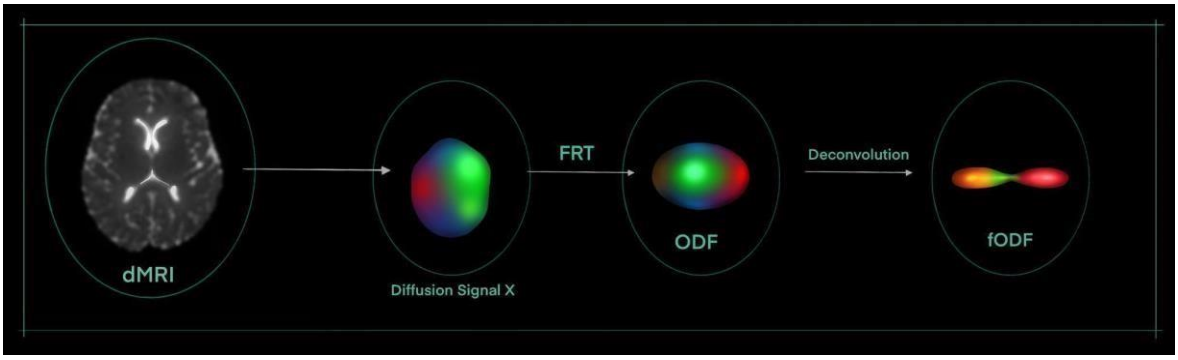

Figure 1.4. Demonstration of how to create fODF

A more detailed explanation of the image capture technique explains that the image intensities at each location are reduced depending on the strength (b-value) and direction of the so-called magnetic diffusion gradient as well as the local microstructure in which the water molecules diffuse. The more the image is attenuated at a particular place, the more diffusion occurs in the direction of the diffusion gradient. To evaluate the tissue's whole diffusion profile, further MR scans must be performed with different diffusion gradient orientations for each scan [15].

1.2.2. Tensors

Diffusion MRI is based on the mathematical and physical interpretations of the tensors shown in Figure 1.5, which are geometrical quantities.

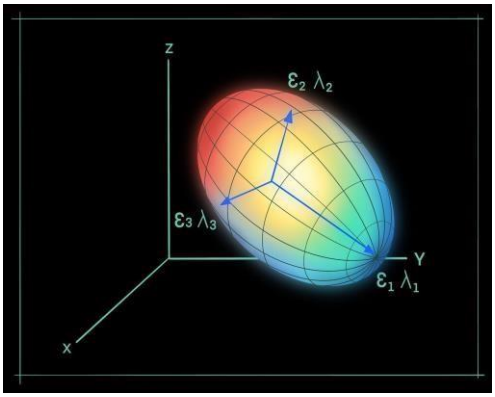

Figure 1.5. Presentation of the tensor on coordinate system

Imaging, based on the idea of a symmetric matrix, is only a specific example of the basic mathematical concept. Tensors do not move when the coordinate system used to represent them is rotated because they have a genuine, physical existence in a substance or tissue [16]. The diffusion tensor's elements are shown in the Formula (1.2).

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{1.2}$$

1.2.3. Fractional Anisotropy

The degree of anisotropy in a diffusion process is indicated by a scalar number called fractional anisotropy (FA), which ranges from zero to one. Diffusion is said to be isotropic when its value is 0, meaning that it is unfettered (or equally limited) in all directions. If the value is 1, diffusion is completely confined in all other directions and only happens along one axis. Low and high FA is illustrated in figure 1.6.

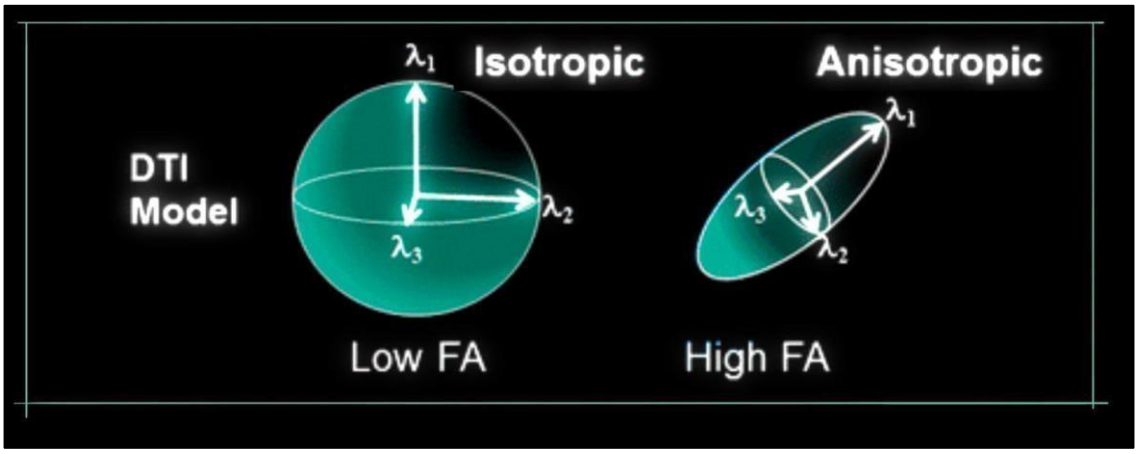

Figure 1.6. Illustration of low and high FA

The Diffusion Tensor, D, fully describes a Diffusion Ellipsoid. FA is derived from the eigenvalues of the diffusion tensor and is given in equation 3**.** FA is determined using the diffusion tensor's eigenvalues, $\lambda 1, \lambda 2, \lambda 3$ as described in formula (1.3).

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1-\lambda)^2+(\lambda_2-\lambda)^2+(\lambda_3-\lambda)^2}{\sqrt{\lambda_1{}^2+\lambda_2{}^2+\lambda_3{}^2}}} \quad (1.3)$$

attenuated the image is at a specified location. Repeating the MR scans with varying diffusion gradient directions for each scan is necessary to measure the tissue's whole diffusion profile [17].

1.3. TRACTOGRAPHY

A common neuroimaging technique known as diffusion MR fiber tractography (DMR-FT) is used to examine neural connections between cortical and subcortical regions and to assess abnormalities in nerve fibers brought on by illnesses or surgical procedures [18].

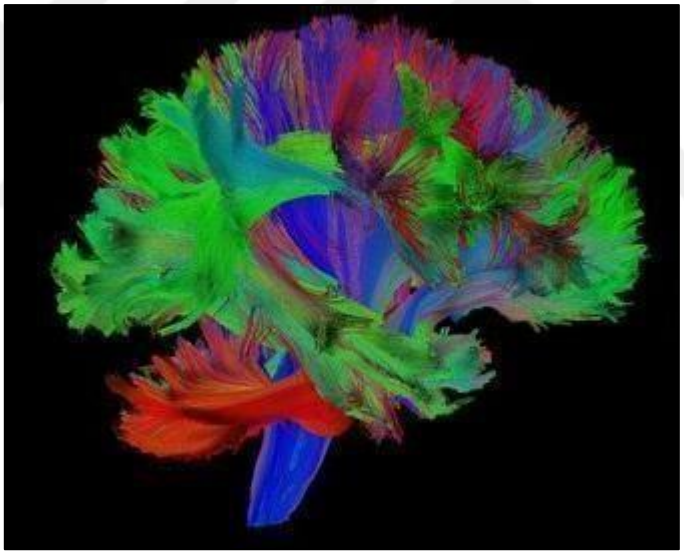

Figure 1.7. Tractometer 2015 challenge groundtruth [19]

The benefit of the Diffusion MR on brain imaging to avoid stroke was proved and the method be admitted in first years of the 90's also. Since that time, Diffusion MRI technology has improved and recent clinical applications opened up. A new perspective to the connectivity of brain white matter came with Diffisuon MRI and hence, tractography [20]. The aim of the diffusion tractography technique is to determine the white matter pathways between distant brain regions by evaluating the changes in the self-diffusion measurements of water depending on the tissue microstructure. The methods used for this purpose can be evaluated

in two main groups as local and global. In local methods such as deterministic or probabilistic flowline; starting from a determined core point and following local tissue orientations, possible pathways are determined [21], [22]. The pattern DTI data to show the probabilistic and deterministic methods is shown in figure 1.8.

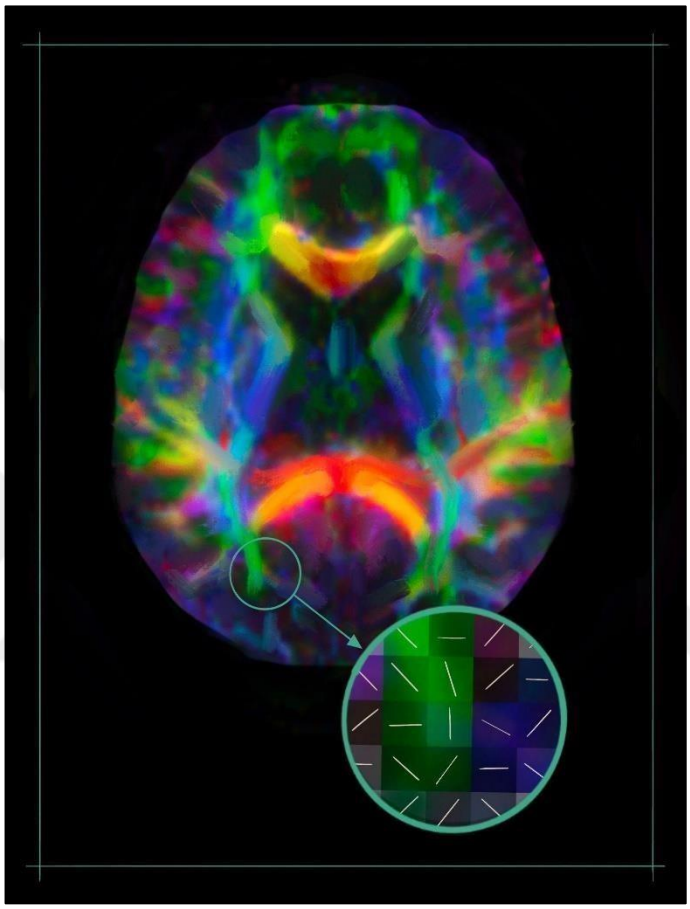

Figure 1.8. The pattern DTI data

The major eigenvector of the diffusion tensor defines the local tract direction in deterministic tractography methods, which are based on streamline algorithms as seen in figure 1.9.

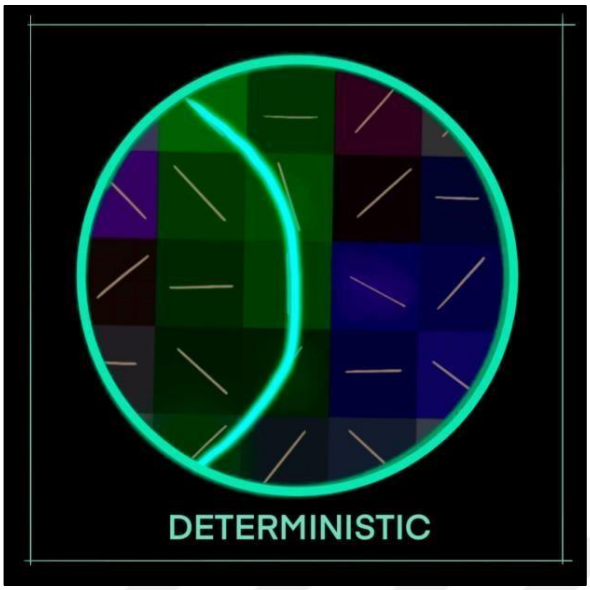

Figure 1.9. Example tracking for deterministic method

However, in probabilistic tracking, the tracking direction is no longer predictable and is instead randomly selected from a distribution of potential directions at each point along the path which is stated in the figure 1.10.

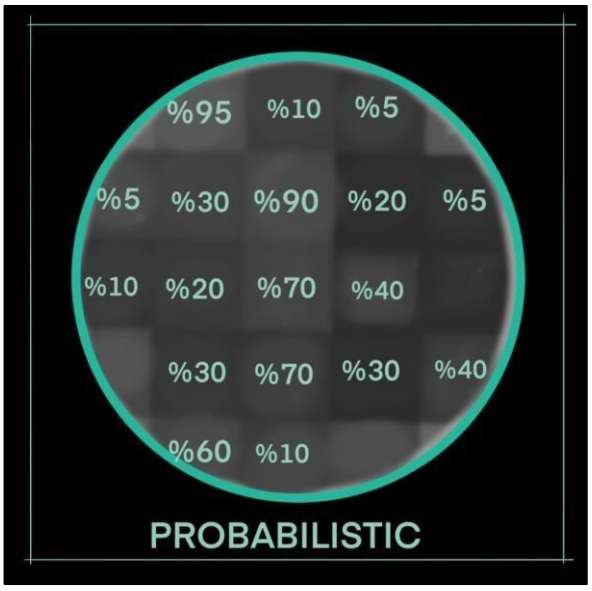

Figure 1.10. Example diagram of ptobabilistic method

These methods are widely used because they are easy to understand and local preliminary information can be easily transferred. However, these methods are sensitive to noise in local

tissue orientations and their success decreases with link length [23]. On the other hand geodetic tractography methods, which are global tractography methods, are based on the shortest path calculation on undirected graphs created using local diffusion measurements [24], [25]. In the shortest path methods, interregional dependencies are obtained, not paths. Local knowledge is difficult to convey [26]. Although they are successful in removing remote connections, they can give false positives by exaggerating the connections [25].

1.3.1. Limitations of tractography

The premise underlying the tractographic methods is that the principal eigenvector's direction in each voxel corresponds to the typical direction of a single fiber bundle. Thus, the method only produces a valid result when the region of interest is homogenous and the direction variations of the fasciculi are on a par with the size of the voxels. As a result, the method performs badly in regions with more bundles of fibers or in areas where they intersect, approach, converge, or diverge. The interpretation of tractographic reconstructions needs previous experience and knowledge since these limitations in clinical practice might result in following pathways that are not present (false positives) or inadequately pursuing paths that are present (false negatives) [27].

Other limitations include the algorithm's inability to discern the function of the neural route, to establish the direction of the neural pathway (afferent from efferent projections), and to recognize the presence of synapses along the same neural pathway [28].

1.4. REINFORCEMENT LEARNING

Machine learning is a branch of artificial intelligence that studies how well a machine can mimic intelligent human behavior. Artificial intelligence (AI) systems are utilized to complete challenging tasks in a manner closer to how people solve problems [29].

Supervised machine learning models are trained on labeled data sets, allowing them to develop and learn over time. For example, an algorithm could be trained using photographs that were all annotated by people, and over time, the computer would develop its own ability to recognize images. Nowadays, supervised machine learning is the most common kind

[30],[31].

A program uses unsupervised machine learning, a branch of machine learning, to find patterns in unlabeled data. Unsupervised machine learning can find trends or patterns that people are not aware of. Unsupervised machine learning algorithms, for instance, may sift through online sales data and identify different client types making purchases [31].

Through experimentation and learning from mistakes, reinforcement machine learning trains machines to perform the best action. By letting the computer know when it has made the right decisions and gradually teaching it what actions to take, reinforcement learning can be used to teach models to play computer games or a robot to walk [32].

Reinforcement learning is a machine learning method that works differently from supervised and unsupervised learning. The agent enters a new situation by choosing an action for himself at each step, and with the reward feedback he receives in each action, he recognizes the environment and chooses his actions accordingly shown at Figure 1.11.

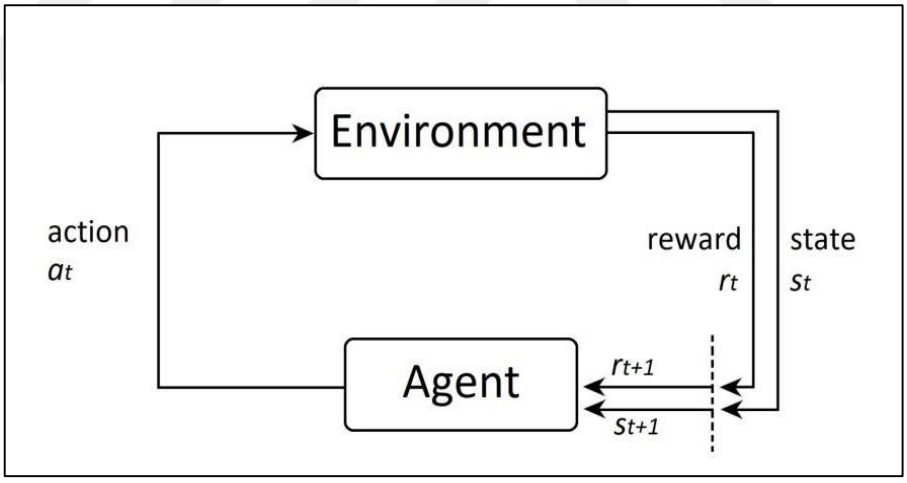

Figure 1.11. Reinforcement learning working schema

In order for the agent to obtain more rewards, the agent must choose actions agent tried in the past and for which agent has received positive rewards. The agent uses actions that has experienced before to obtain rewards, but must also discover if there are actions that can get more reward from in a situation that encountered. Thus, the agent should try various actions and progressively promote the ones they can get the best reward [33].

All reinforcement learning agents have clear goals and are capable of perceiving the characteristics of their environment and choosing actions that affect it. The agent is the one who does the action and learns from it. A behavioral agent can be part of a larger behavioral system. The interaction of reinforcement learning with other engineering disciplines is one of its most important features. Reinforcement Learning has influenced psychology and neuroscience, resulting in significant benefits for both fields [34].

1.4.1. Reinforcement learning elements

In a Reinforcement Learning system, apart from the agent and the environment, there are four elements, one of which is optional:

- Policy
- Reward (reward signal)
- Value/Status Value (value function)
- Environmental model (model)

Policy; determines the action that the agent can take in the situation it is in. It can be thought of as a kind of action-reaction pairing. If the situation is accepted as an effect, the agent gives an action in response. This policy can be defined as a simple action or as a lookup table that meets all situations. Policy can also be characterized as dynamic. The main reason for this is that the agent evaluates the situation he is in and seeks (realizes) the actions he can take.

Prize; It is the rating the agent receives from the environment for a certain action. A reinforcement learning agent's goal is to maximize the long-term rewards it receives. The value of the reward impacts how well or poorly the activity is accepted. The agent changes its policy over time based on these rewards. For example, if a low score is obtained after an action taken, the agent may choose to take a different action when the same situation occurs in the future.

Status value; It is the sum of the rewards the agent can expect from the situation is in and other situations that follow. Rewards refer to what is good and bad momentarily, while status value refers to what is good and bad in the long run. For example; a situation may have a

low reward but a high value. This is because the low reward situation is followed by other high reward situations. The opposite is also possible. There may also be situations with consistently low rewards after a high reward situation. The situation here can be thought of as “forward-sightedness”.

A model is an optionally included element in the system. It is a simulation of the environment and allows the agent to predict the reward and the situation that will result from this action before taking an action. In this way, a change can occur in the behavior of the agent by making a plan [33], [35].

1.4.2. Limitations and scope of RL

Most of the reinforcement learning applications use value function but it is not that necessary. For example, to approach reinforcement learning problems, approaches such as genetic algorithms and programming, as well as other optimization methods, have been employed without ever referring to value functions. These methods assess the long term behavior of a large number of non-learning agents, each with its own policy for interacting with its environment, and choose the ones that are able to gain the most reward. If the policy space is small enough, or can be structured in such a way that good policies are common or easy to identify, or if there is a lot of time available for the search, evolutionary approaches can be useful. Furthermore, evolutionary approaches provide advantages in situations where the learning agent is unable to precisely assess the status of its environment Focus on reinforcement learning approaches, which include learning while interacting with the environment, as opposed to evolutionary methods. Strategies that benefit from the uniqueness of individual behavioral interactions are frequently much more effective than evolutionary techniques. Many of the useful structures of the reinforcement learning problem are disregarded by evolutionary methods, including the fact that the policy they are looking for is a function from states to actions, the states that an individual passes through throughout its lifetime, and the actions that it chooses [36].

Despite the fact that evolution and learning have many similarities and naturally interact, we do not believe evolutionary methods are particularly well adapted to reinforcement learning problems on their own. however, some approaches that do not use value functions, such as evolutionary methods. These methods look for policies in spaces specified by a set of

numerical parameters. They calculate which directions the parameters should be altered in order to increase a policy's performance as quickly as possible [37]. They develop these estimates while the agent is engaging with its environment, unlike evolutionary techniques, and hence may take advantage of the nuances of individual behavioral interactions. This type of strategy, known as policy gradient methods, has proven beneficial in a variety of issues, and it includes some of the most basic reinforcement learning methods. Some of these methods, in fact, use value function estimates to improve their gradient estimates. Overall, there isn't a clear separation between policy gradient approaches and reinforcement learning methods [38]. The relationship between reinforcement learning and optimization approaches is worth mentioning because it is a common source of confusion. When we say that a reinforcement learning agent's goal is to maximize a numerical reward signal, we don't mean that the agent has to actually achieve that goal. Attempting to maximize a quantity does not imply that it will ever be maximized. The notion is that a reinforcement learning agent is continually attempting to get more rewards[39], [35].

1.5. Q-LEARNING

The most used algorithm of reinforcement learning is Q-learning. The agent enters a new situation by choosing an action for himself at each step, and with the reward feedback he receives in each action, he recognizes the environment and chooses his actions accordingly. In Q-learning, the agent keeps the experiences he has gained at each step in the Q table as seen in Figure 1.12 and tries to maximize his reward by making use of this table in his actions [40].

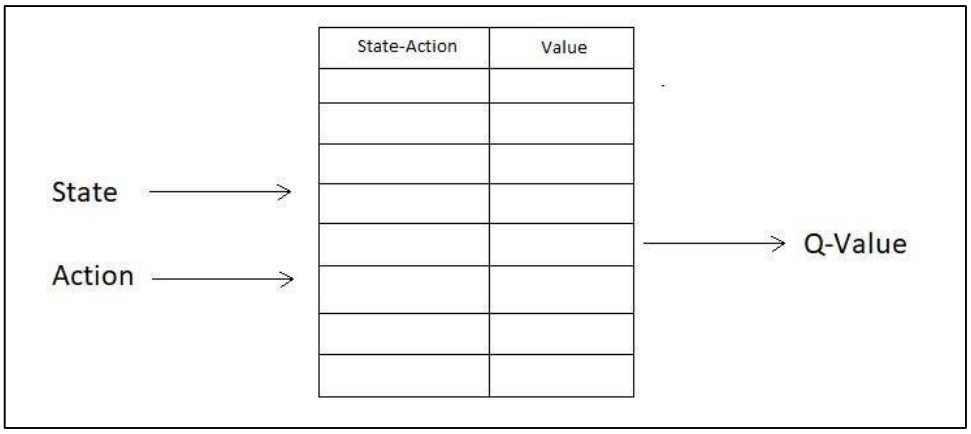

Figure 1.12. Q table schema

1.6. Q-ROUTING

Q-routing is a method that directly applies Q-learning to a packet routing problem in a communication network [41]. Unlike Q-learning, a cost is defined instead of a reward for any action the agent takes, and the agent tries to move forward with the lowest cost. It has been reported that the Q routing method provides faster routing on the network compared to the shortest paths method, especially at high network load [41].

It is suggested in this study that the Diffusion MR fiber tractography problem be solved using the Q-direction method. By taking into account the whole cost of the road rather than local texture orientations at each point, it is intended to attain the robustness of global approaches while still gaining the benefits of local streamline methods.

2. METHOD

2.1. Q-routing

This section describes how to use the Q-routing method to calculate paths on a graph. Let G = (V,E) be an undirected and weighted graph formed by the node set V and the edge set E ∈ V x V. Let each edge in the edge set be assigned a value e ∈ E, w(e) ∈ R, representing the cost. An example diagram is given in Figure 2.1.

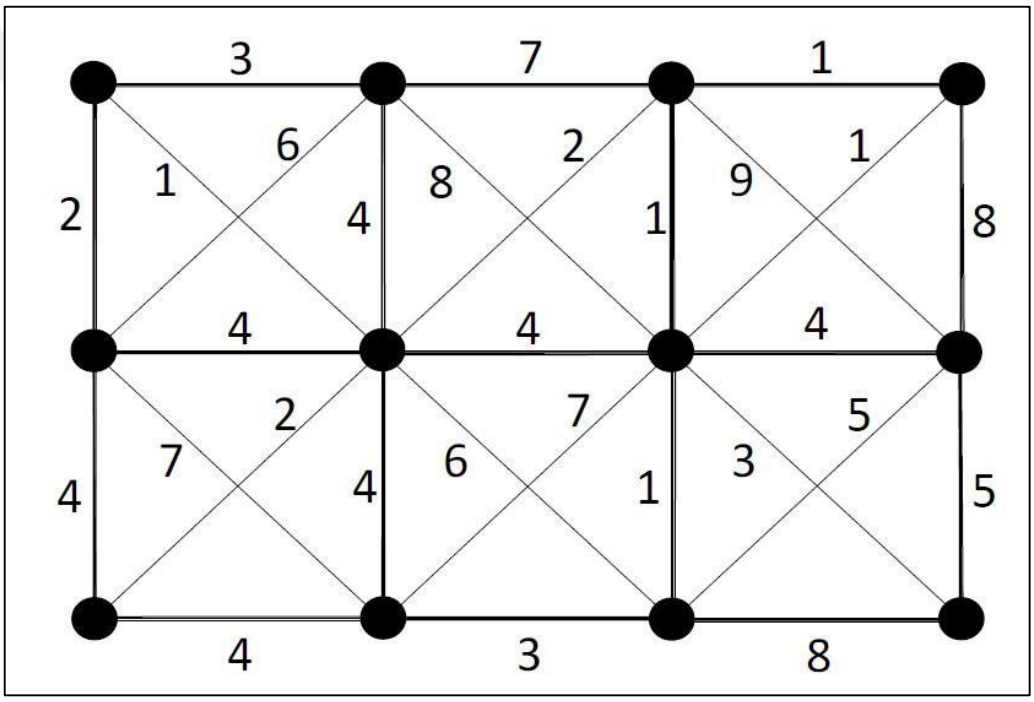

Figure 2.1. Example diagram

Q-direction consists of the terms agent, environment, state, action, and cost, and operates on the basis of the interaction of the agent and the environment. In this study, the case set (D), action set (A) and cost function (M) are defined, see (2.1) ,(2.2), (2.3).

$$D = V \tag{2.1}$$

$$A = \{\leftarrow, \nwarrow, \uparrow, \nearrow, \rightarrow, \searrow, \downarrow, \swarrow\} \tag{2.2}$$

$$M : D \times A \rightarrow R \tag{2.3}$$

State change function $\boldsymbol{\delta} : D \times A \longrightarrow D$, specifies the transition of agent in any state $d \in D$ from the action set to state $d' = \boldsymbol{\delta}(d,a)$ with an action $a \in A$. The cost of this change of state (M(d, a)) is determined by the weight of the edge corresponding to the action.

The purpose of reinforcement learning is to learn the long-term cost of the action that the agent will take in any given situation. For each state, a Q function is defined that specifies the long-term costs of possible actions. By assigning $Q^0(d; a) = 0$ as the initial value to the Q function, the training of the agent, that is, the determination of the Q function, is done as follows: In each training step, a random state ($d \in D$) and action ($a \in A$) are selected for the agent and at step t, the $Q^t(v; a)$ value is updated as shown in formula (2.4).

$$Q^{t+1}(d, a) = (1 - \eta)\, Q^t(d, a) + \eta\, [\, M(d, a) + M'] \tag{2.4}$$

Here η is the learning rate. The term M' is calculated in formula (2.5)

$$M' = min_{a'}\{Q^t\,(d', a') : d' = \delta(d, a), a' \in A \tag{2.5}$$

For the diagram given in Figure 2.1, the Q values created as a result of 50000 cycles are presented in the diagram in Figure 2.2.

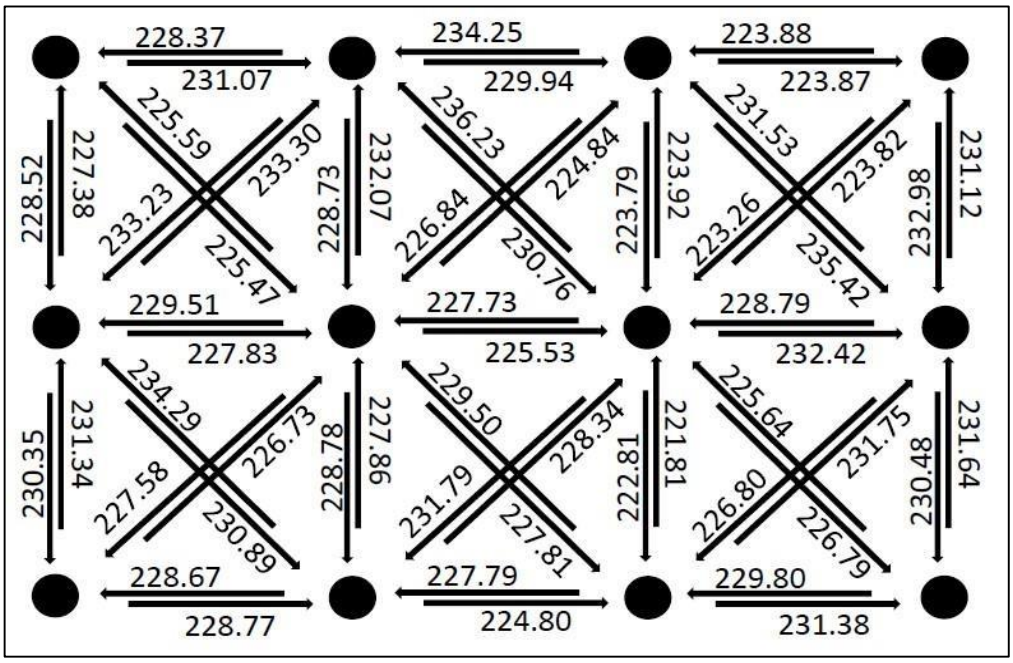

Figure 2.2. Calculated q values for the example diagram

After calculating the function Q, the preferred action, a*, for an agent in state d is determined as seen in formula (2.6).

$$a^* = \arg min_a \{ Q(d,a) : a \in A \} \quad (2.6)$$

Accordingly, the path for the agent starting from state d^0 is determined by iteration, see formula (2.7)

$$d^{t+1} = \delta(d^t, a^*) \quad (2.7)$$

In the diagrams in Figure 2.3 and 2.4, the routes given for the agents started from different initial points.

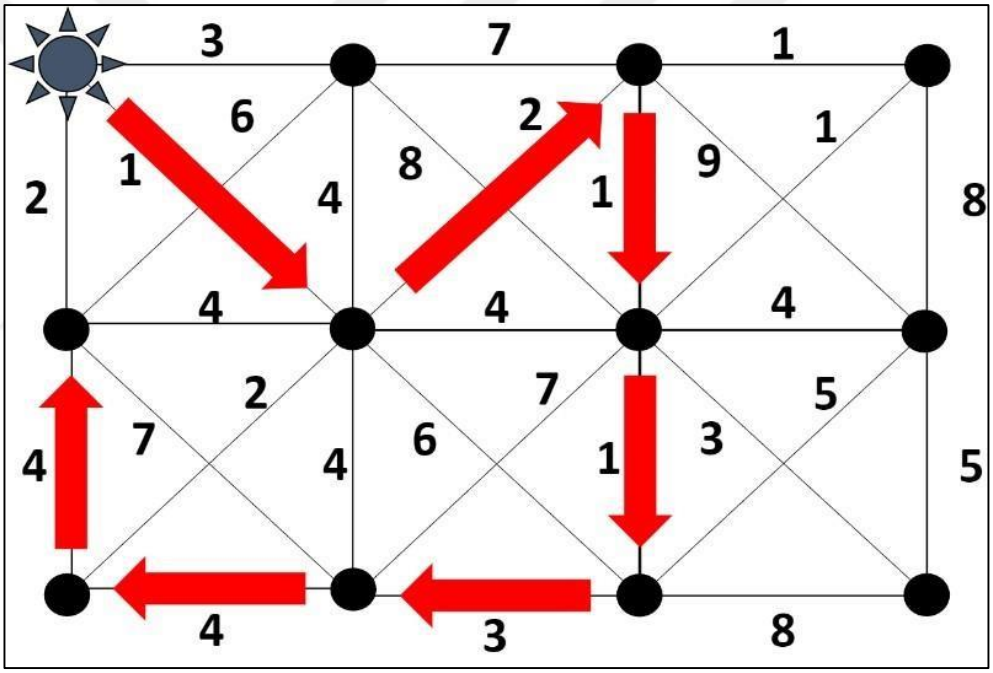

Figure 2.3. Paths obtained with agents launched from first initial state on the example diagram

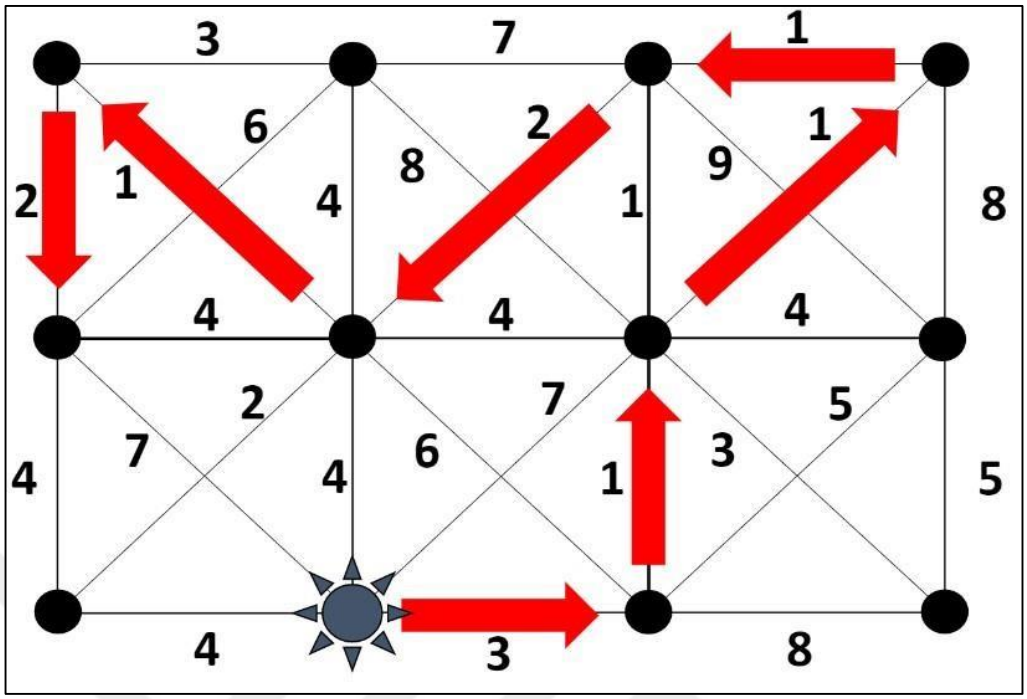

Figure 2.4. Paths obtained with agents launched from second initial state on the example diagram

2.2. APPLICATION OF Q-ROUTING BASED TRACTOGRAPHY ON DIFFUSION PHANTOM

The FiberCup phantom, a realistic diffusion phantom replicating the primary white matter connections of the human brain, is used in this part to test the suggested Q-routing based tractography method. The Tractometer competition dataset [42] consisting of diffusionweighted MR images of the FiberCup phantom with known fiber distribution was used [20]. A 64x64 pixel mid-section of the phantom's average data taken with the value of b = 1500 was used. The actual connection paths on the phantom and the core regions located at both ends are defined and presented in figure 6.

The restricted spherical deconvolution approach was used to calculate the fiber orientation distribution area (fODF) for the phantom [43]. Weights of graph edges (w(e)) are calculated by applying negative logarithms to dependencies localized from fODFs as used in the shortest paths methods and given in the study [44]. At the boundaries, actions out of the data volume are constrained by assigning high values in the cost function. The Q function was

calculated by repeating the iteration given in (2.4) 5x106 times and using a learning rate of 0.7.

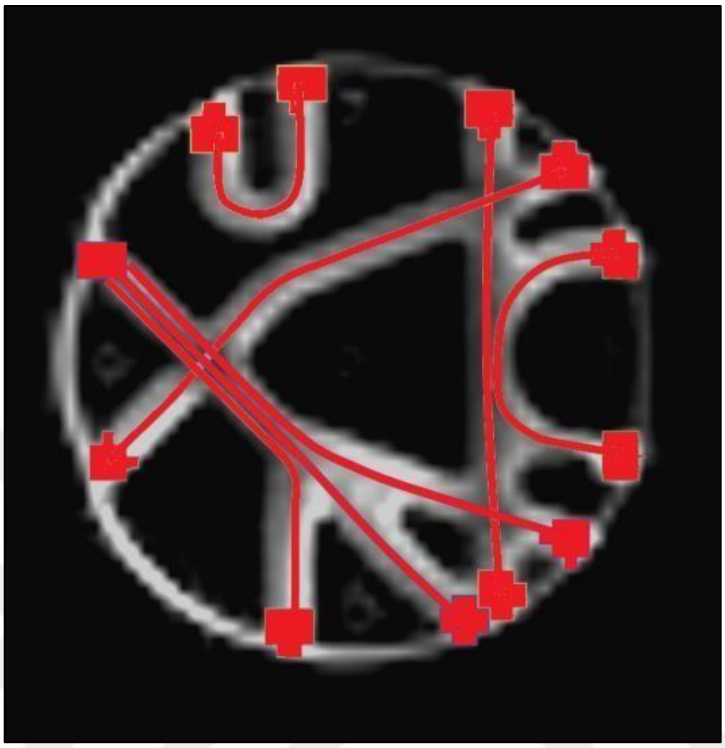

Figure 2.5. Real connection paths on the phantom and seeds at both ends [45]

Initial states of agents are assigned as seeds in determining pathways on the phantom. In determining the preferred route for the agent, the iteration given in formula (2.4) was applied by repeating 30 times.A high value in the Q function is added to the actions that return to the state at each step so that the agent does not enter the positive loop. A high value in the Q function is added to the actions that return to the state at each step so that the agent does not enter the positive loop. Thus, it is ensured that the path obtained does not intersect with itself.

Similarly, a high value in the Q function was added to actions that curved more than a certain angle at each step to limit the bends in the agent's path. Thus, sharp turns in the movement of the agent are prevented.

Here, the results were compared using two different limits, the maximum curl of 45° and 90° degrees. For comparison purposes, the same fiber orientation distribution function (fODF) and deterministic streamline tractography for cores were performed as adapted in the Dipy package [46].

2.3. APPLICATION OF Q-ROUTING BASED TRACTOGRAPHY ON ISMRM 2015 TRACTOMETER DATASET

In this section, conventional probabilistic method on diffusion signal and the proposed Qrouting based tractography method that trained over odf data, not from an empty table applied on ISMRM 2015 Tractometer data [43] to show the effectiveness of the proposed method. Conventional probabilistic method is shorten as CPM and proposed method called Q-RTM. The ISMRM 2015 Tractography Challenge to show the effectiveness of the method. ISMRM 2015 Tractography Challenge intend to evaluate and compare tractography pipelines. Using Fiberfox, a new data generating method that enables the creation of a groundtruth for scoring using the Tractometer online evaluation system, this barrier is overcame. The challenge's objective was to provide the best reconstruction of fiber pathways in a faithfully replicated whole-brain diffusion-weighted MR image with characteristics resembling those of a clinical acquisition. Despite the completion of the challenge, the result was obtained using the development tool, which is still available.

The Q function was calculated by repeating the iteration given in formula (2.4) $7x10^8$ times and using a learning rate of 0.7.

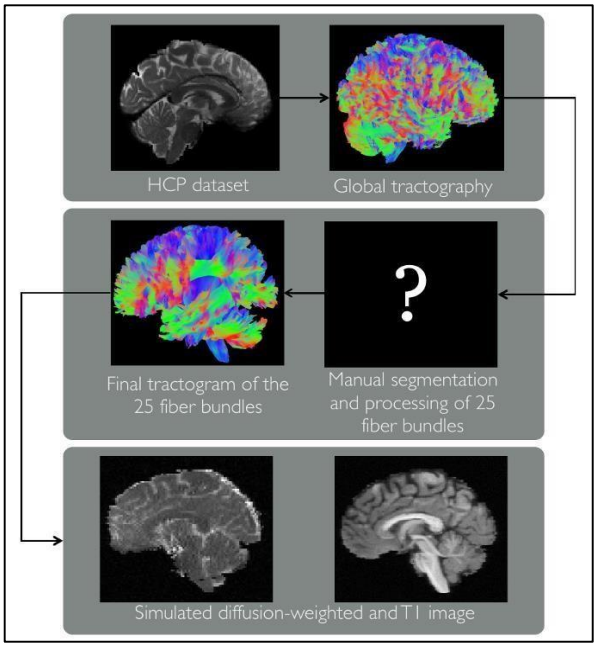

Figure 2.6. Data generation process in a nutshell [42]

A 2mm isotropic diffusion collection with 32 gradient directions and a b-value of 1000 s/mm2 makes up the clinical-style challenge dataset. The phase-encoding direction is inverted also, and it includes one b=0 image and an optional b=0 volume. Another image is offered that a like T1 [42].

The environment generates a reward r_t from the last tracking step and the peaks retrieved from the fODFs at the streamline's head, as well as a state s_t from the diffusion signal. This reward and state is saved on Q-table. Both are delivered to the agent, the state is used as input to create an action, in this case a new tracking step, and the reward is used to teach the agent how to create better actions. The cycle continues with the return of a new state s_{t+1} and a new reward r_{t+1} . This cycle is described in Figure 2.7.

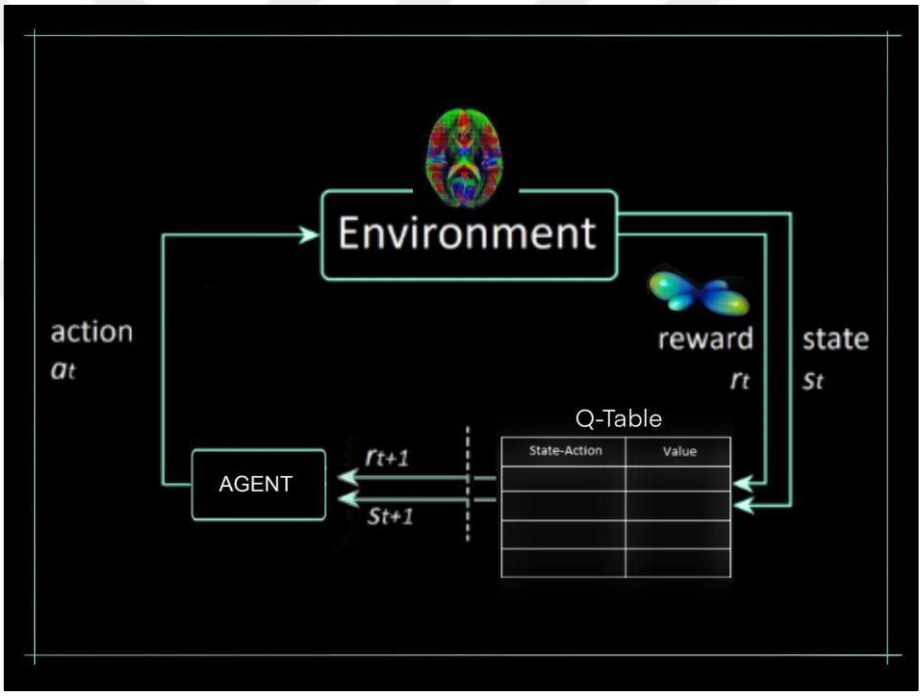

Figure 2.7. Q-learning cycle for the data

And also illustration of the modification in tractograms created by agents while training is shown is Figure 2.7.

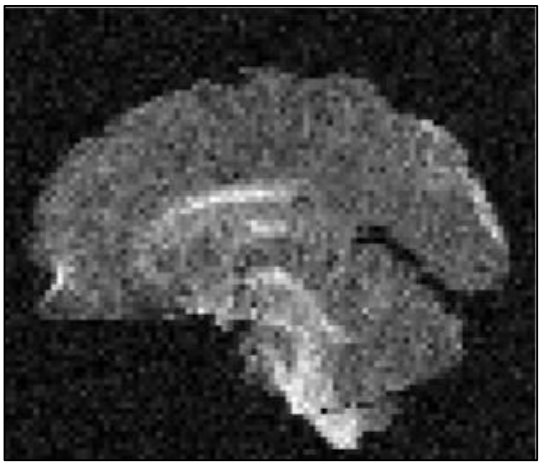

Figure 2.8. Example DWI slice of ismrm 2015 challenge data [42]

2.3.1. Valid Bundles

The total number of valid bundles that both exist in the groundtruth data and were correctly reconstructed in the contestant's entry. 25 groundtruth packages were included in the challenge [19].

2.3.2. Valid Connections

VC is that how many streamlines were included in valid bundles as a percentage.

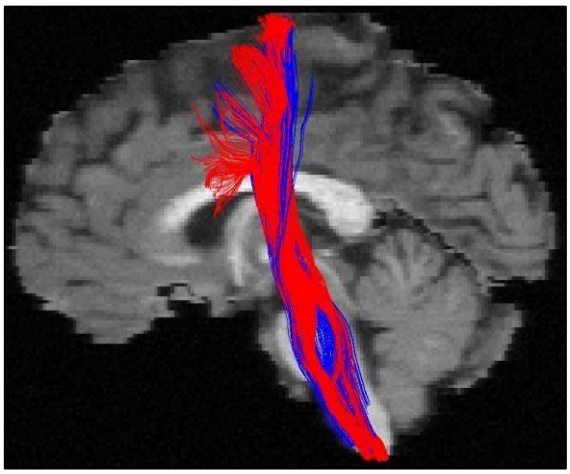

Figure 2.9. Example of valid bundle and conncetions [19]

2.3.3. Invalid Bundles

Invalid bundles is that the number of bundles that were realistic but did not correspond to a groundtruth bundle. The dataset that was submitted has bundles that can be extracted, but they do not match any known bundles [19].

2.3.4. Invalid Connections

Invalid conncetions is that how many streamlines were included in valid bundles as a percentage.

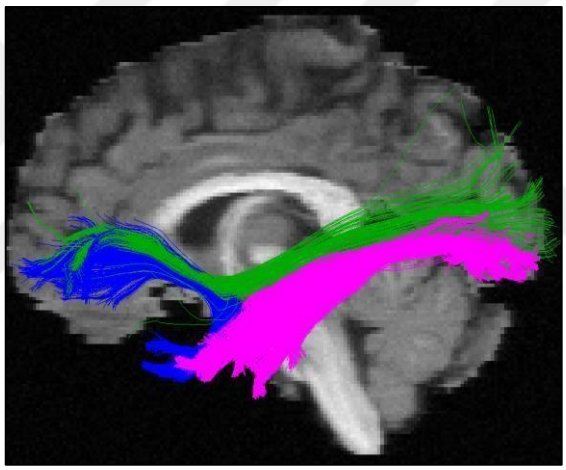

Figure 2.10. Example of invalid bundle and conncetions [19]

2.3.5. No Connection

No conncetions is that a measure of the proportion of streamlines that weren't given to valid conncetion or invalid conncetion. They typically consist of very short streamlines or streamlines that are singular in both location and shape, thus even when grouped together, they remain singletons [19].

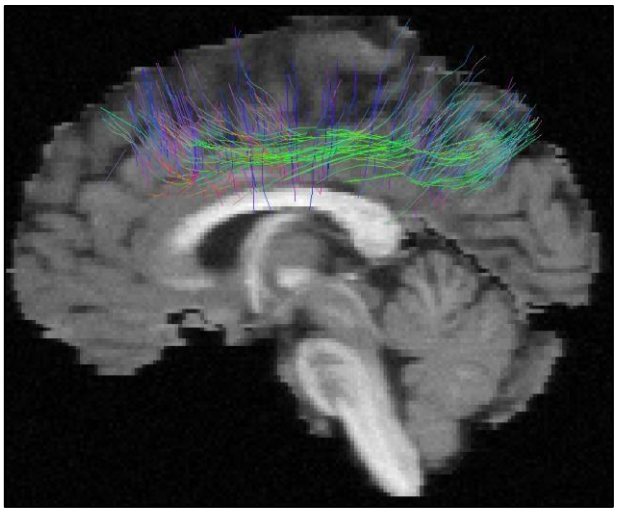

Figure 2.11. Example of no conncetions [19]

In the original Tractometer approach, there was additional statistic known as "Valid connections along a faulty path" (VCWP). Since the current approach does not classify streamlines using masks, this category cannot be determined. These streamlines, however, are frequently allocated to a single Invalid bundle that reaches both endpoints but has a shape that deviates excessively from the groundtruth bundle [19].

2.3.6. Fidelity Metrics Definitions

Additional criteria were used to evaluate the task. And this so fidelity measures are designed to provide a broad picture of how well valid bundles performs in respect to its groundtruth equivalents. Since a submission might identify a particular valid bundle but have extremely limited coverage because it might have only identified a few streamlines to represent that bundle, this is meant to supplement the global connectivity metrics [19].

2.3.7. Bundle Overlap

Bundle overlap is that percent of the voxels in a ground truth bundle's volume that are crossed by at least one valid streamline related to the bundle. This number indicates how well the bundle's original volume is recovered from the tractography result [19].

2.3.8. Bundle Overreach

Bundle oevrreach is that the proportion of all the voxels in a ground truth bundle that are traversed by at least one valid streamline that is connected to the bundle over all the voxels outside the volume of the ground truth bundle. This value represents the amount by which the advisable connections exceed the ground truth bundle's volume. [19].

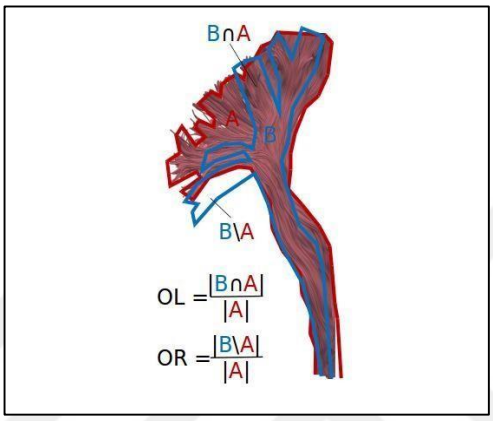

Figure 2.12. Overlap and overreach of the bundle [19]

2.3.9. Angular scores of error

Between the principal local tractogram fiber directions and the corresponding ground truth fiber directions, the mean voxel-wise angular error. Missing directions result in a maximum 90 degree error penalty. Voxels with a single fiber population and voxels with crossed fiber populations make up the two groups into which values are divided [19].

3. RESULTS

3.1. Q-ROUTING BASED TRACTOGRAPHY ON DIFFUSION PHANTOM

In order to measure the convergence of the training of the system, the total Q value is calculated at each step of the iteration given in (1.7) and shown on the graph in figure 3.1.

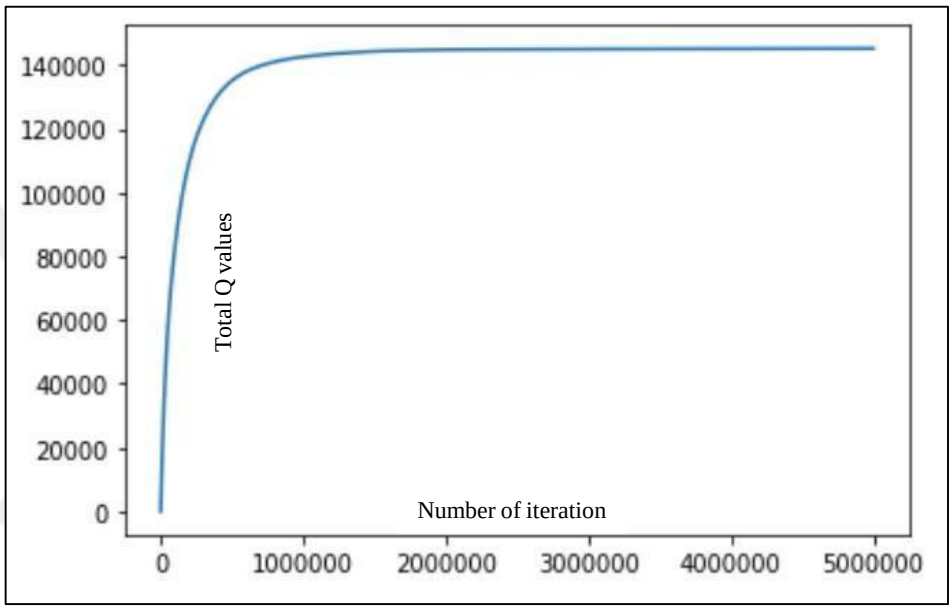

Figure 3.1. Change of total Q values during the training

The results obtained with the Q-routing-based tractography method on the FiberCup diffusion phantom, with different curling constraints for each core, are shown in Figure 3.2-3.4 by plotting on the b0 MR image of the phantom.

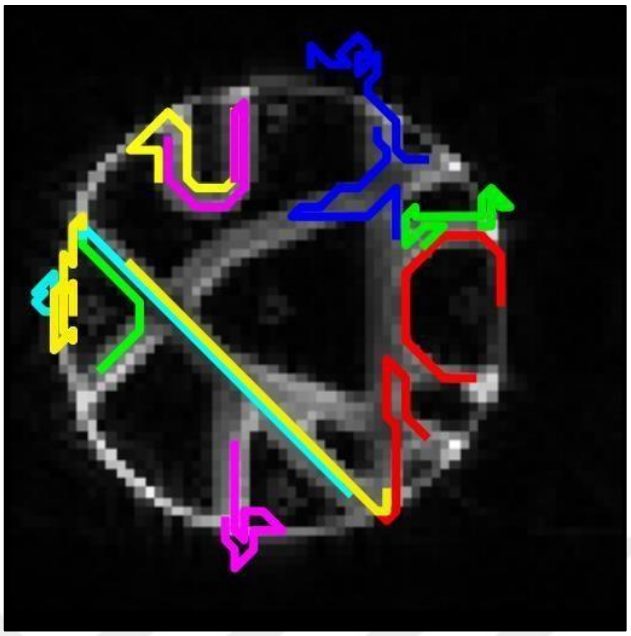

Figure 3.2. Results without curl restriction

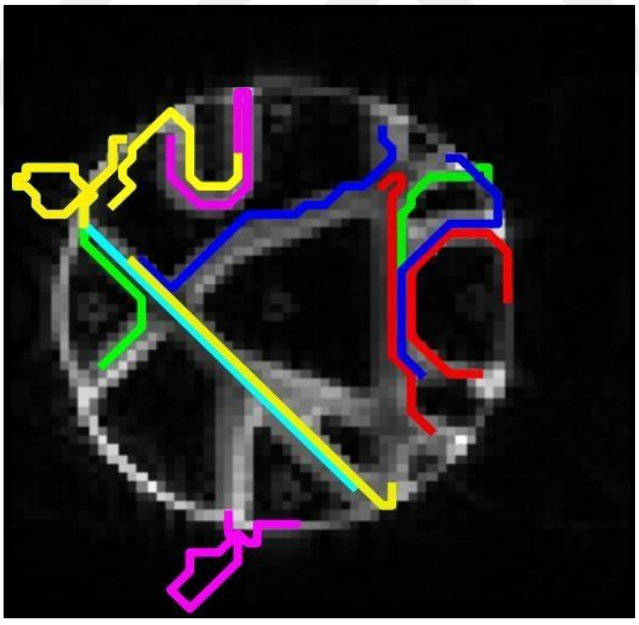

Figure 3.3. Results obtained by restricting curls greater than 90°

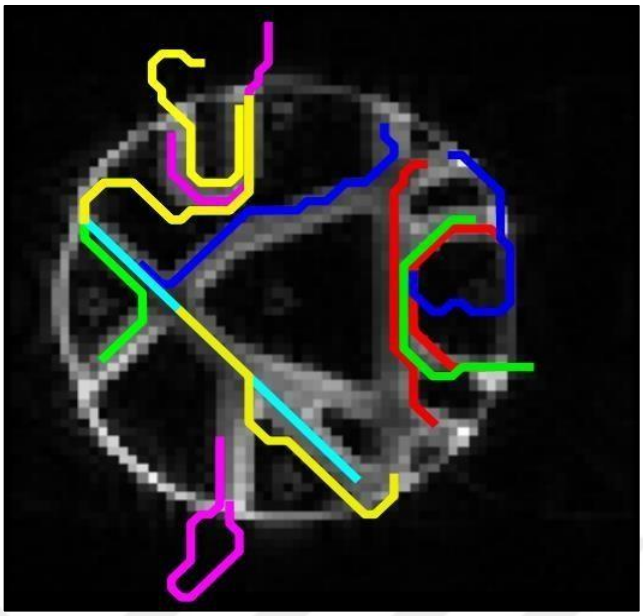

Figure 3.4.Results obtained by restricting curls greater than 45°

3.2. Q-ROUTING BASED TRACTOGRAPHY ON ISMRM 2015 TRACTOGRAPHY CHALLENGE DATASET

The results of Q-RDTM, CDM, Q-RPTM and CPM methods for each bundle, respectively, are given in tables 3.1-3.4 and the obtained tracts are given in figures 3.8-3.11 . In addition, Table 3.5 and 3.6 provides an overall and VB comparison of the metrics. In figure 3.7, the fODF comparsion is shown for r and q table.

Table 3.1. Results of the Q-RDTM

Bundle	Count	% of F1 Scores	Overlap (% of GT)	Overreach (% of GT)
Superior Cerebellar Peduncle (right)	81	28	18.963	45.329
Parieto-Occipito Pontine Tract (left)	65	8	4.674	57.557
Cingulum (right)	177	10	5.423	31.501
Corpus Callosum	4049	32	22.975	46.814

Anterior Commissure	0	0	0.000	0.000
Fronto-Pontine Tract (right)	98	2	6.538	31.615
Posterior Commissure	0	0	0.000	0.000
Inferior Cerebellar Peduncle (left)	322	33	25.515	52.090
Fronto-Pontine Tract (left)	120	11	6.391	43.987
Uncinate Fasciculus (right)	249	37	25.926	33.333
Inferior Longitudinal Fasciculus (right)	341	23	15.187	50.726
Cortico-Spinal Tract(right)	0	0	0.000	0.000
Cortico-Spinal Tract (left)	0	0	0.000	0.000
Inferior Longitudinal Fasciculus (left)	472	33	24.087	45.379
Superior Longitudinal Fasciculus (right)	1155	39	34.565	55.508
Fornix	72	16	9.795	57.773
Parieto-Occipito Pontine Tract (right)	3	2	1.048	37.051
Optic Radiation (right)	12	5	2.579	42.757
Uncinate Fasciculus (left)	247	42	31.875	39.369
Superior Longitudinal Fasciculus (left)	1318	33	30.503	62.951
Optic Radiation (left)	127	19	12.149	50.116
Middle Cerebellar Peduncle	463	24	16.185	51.866
Cingulum (left)	326	17	10.118	33.277
Inferior Cerebellar Peduncle (right)	306	41	35.201	49.371
Superior cerebellar Peduncle (left)	56	22	13.411	42.727

Table 3.2. Results of the CDM

Bundle	Count	% of F1 Scores	Overlap (% of GT)	Overreach (% of GT)
Superior Cerebellar Peduncle (right)	55	23	14.623	39.439
Parieto-Occipito Pontine Tract (left)	136	15	9.071	54.546
Cingulum (right)	181	10	5.945	33.606
Corpus Callosum	3879	31	22.409	47.516
Anterior Commissure	0	0	0.000	0.000
Fronto-Pontine Tract (right)	100	13	7.229	26.033
Posterior Commissure	0	0	0.000	0.000
Inferior Cerebellar Peduncle (left)	305	30	23.516	54.889
Fronto-Pontine Tract (left)	88	13	7.643	40.709
Uncinate Fasciculus (right)	293	41	30.305	32.642
Inferior Longitudinal Fasciculus (right)	369	23	15.576	52.026
Cortico-Spinal Tract(right)	0	0	0.000	0.000
Cortico-Spinal Tract (left)	0	0	0.000	0.000
Inferior Longitudinal Fasciculus (left)	489	32	23.640	46.251
Superior Longitudinal Fasciculus (right)	1124	35	30.488	58.247
Fornix	82	16	10.420	55.916
Parieto-Occipito Pontine Tract (right)	8	3	1.711	36.764
Optic Radiation (right)	87	6	3.545	51.376
Uncinate Fasciculus (left)	232	40	30.604	41.721
Superior Longitudinal Fasciculus (left)	1580	34	32.872	64.242
Optic Radiation (left)	71	16	10.083	51.769

Middle Cerebellar Peduncle	432	28	19.319	42.809
Cingulum (left)	287	17	10.225	33.597
Inferior Cerebellar Peduncle (right)	316	40	34.838	51.030
Superior cerebellar Peduncle (left)	132	26	17.828	50.000

Table 3.3. Results of the Q-RPTM

Bundle	**Count**	**% of F1 Scores**	**Overlap (% of GT)**	**Overreach (% of GT)**
Superior Cerebellar Peduncle (right)	12	24	15.610	43.121
Parieto-Occipito Pontine Tract (left)	41	30	19.791	34.071
Cingulum (right)	138	24	14.803	33.562
Corpus Callosum	1945	52	47.384	40.357
Anterior Commissure	0	0	0.000	0.000
Fronto-Pontine Tract (right)	220	43	31.465	28.481
Posterior Commissure	0	0	0.000	0.000
Inferior Cerebellar Peduncle (left)	165	45	42.327	50.967
Fronto-Pontine Tract (left)	76	29	19.452	35.651
Uncinate Fasciculus (right)	176	54	47.539	35.972
Inferior Longitudinal Fasciculus (right)	250	46	41.954	46.604
Cortico-Spinal Tract(right)	0	0	0.000	0.000
Cortico-Spinal Tract (left)	0	50	0.000	0.000
Inferior Longitudinal Fasciculus (left)	302	41	46.906	45.272
Superior Longitudinal Fasciculus (right)	390	16	41.742	58.294

Fornix	25	33	10.256	57.834
Parieto-Occipito Pontine Tract (right)	47	13	22.822	38.942
Optic Radiation (right)	18	56	7.297	35.887
Uncinate Fasciculus (left)	237	37	53.432	39.936
Superior Longitudinal Fasciculus (left)	517	12	39.171	63.791
Optic Radiation (left)	15	45	7.253	42.783
Middle Cerebellar Peduncle	174	33	35.012	36.992
Cingulum (left)	242	22	22.502	34.538
Inferior Cerebellar Peduncle (right)	118	51	48.104	44.400
Superior cerebellar Peduncle (left)	27	21	20.702	51.130

Table 3.4. Results of the CPM

Bundle	**Count**	**% of F1 Scores**	**Overlap (% of GT)**	**Overreach (% of GT)**
Superior Cerebellar Peduncle (right)	2	5	2.648	41.614
Parieto-Occipito Pontine Tract (left)	23	22	13.525	27.231
Cingulum (right)	54	14	8.085	29.695
Corpus Callosum	1402	51	44.404	38.555
Anterior Commissure	0	0	0.000	0.000
Fronto-Pontine Tract (right)	74	28	17.775	23.426
Posterior Commissure	0	0	0.000	0.000
Inferior Cerebellar Peduncle (left)	120	43	38.760	50.040
Fronto-Pontine Tract (left)	27	19	11.399	28.021
Uncinate Fasciculus (right)	115	51	41.178	30.998

Inferior Longitudinal Fasciculus (right)	122	38	28.964	44.594
Cortico-Spinal Tract(right)	0	0	0.000	0.000
Cortico-Spinal Tract (left)	1	2	1.112	25.688
Inferior Longitudinal Fasciculus (left)	111	42	32.174	36.414
Superior Longitudinal Fasciculus (right)	258	40	37.769	57.121
Fornix	15	13	8.227	53.286
Parieto-Occipito Pontine Tract (right)	23	23	14.283	31.148
Optic Radiation (right)	6	5	2.873	38.024
Uncinate Fasciculus (left)	145	51	45.141	40.498
Superior Longitudinal Fasciculus (left)	303	37	35.732	61.024
Optic Radiation (left)	13	12	6.936	40.722
Middle Cerebellar Peduncle	75	32	21.372	35.785
Cingulum (left)	91	22	13.216	28.211
Inferior Cerebellar Peduncle (right)	85	48	43.144	43.855
Superior cerebellar Peduncle (left)	13	27	17.456	29.764

Table 3.5. Metric comparision of the methods

Method	VB	VC	IB	IC	NC	VCWP
Q-RDTM	21/25	%19.55	104	%54.18	%26.26	0
CDM	21/25	%18.10	119	%48.40	%33.48	0
Q-RPTM	21/25	%29.16	150	%52.65	%18.17	0
CPM	22/25	%25.54	138	%49.44	%25.00	0

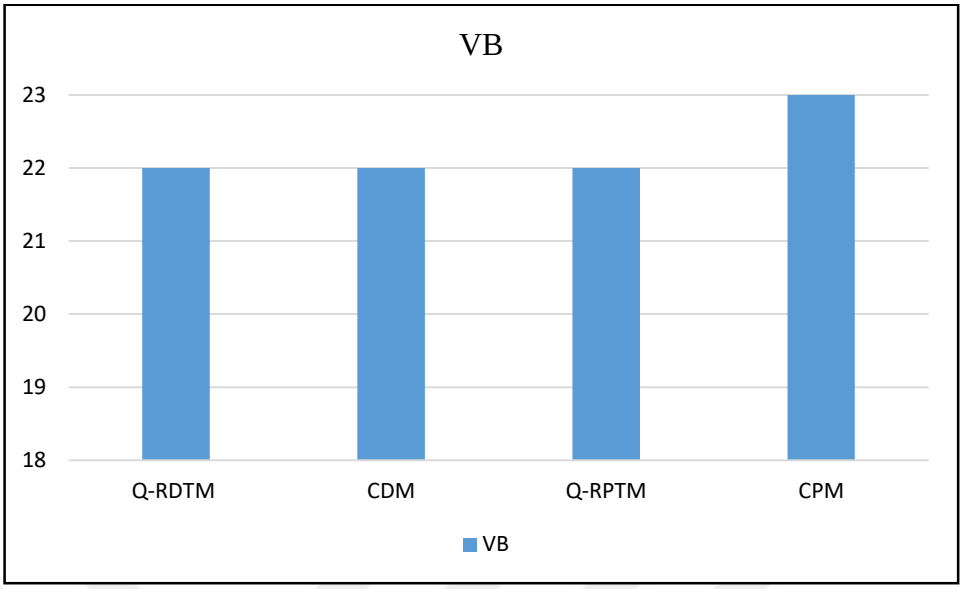

Figure 3.5. Valid bundle column chart

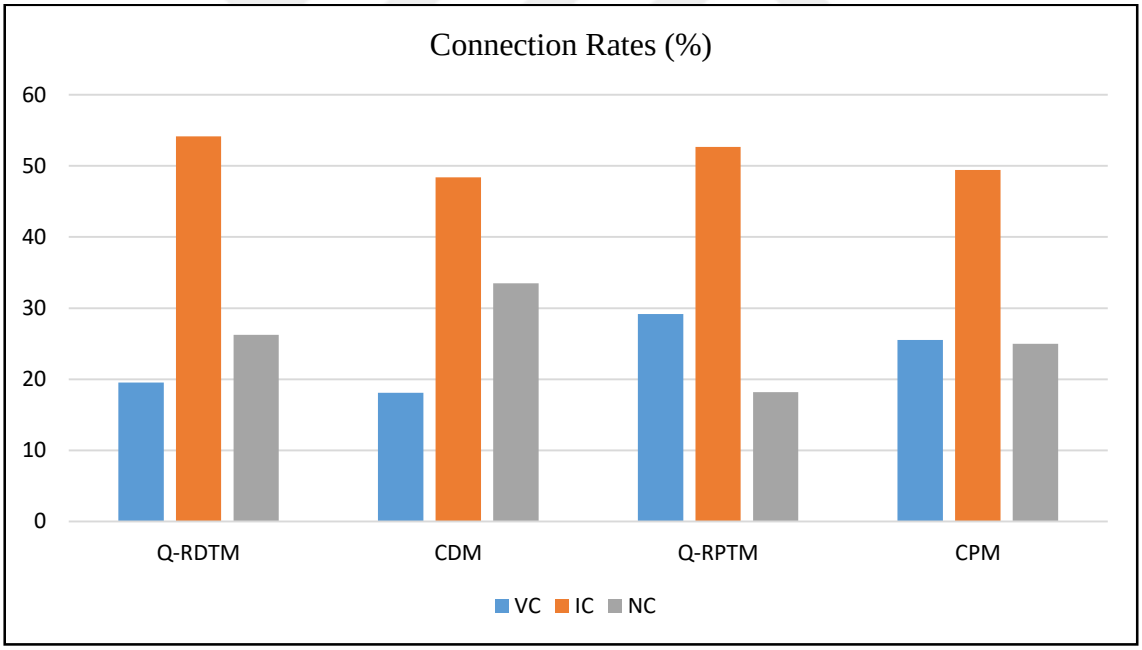

Figure 3.6. Mehods' connection rate column chart

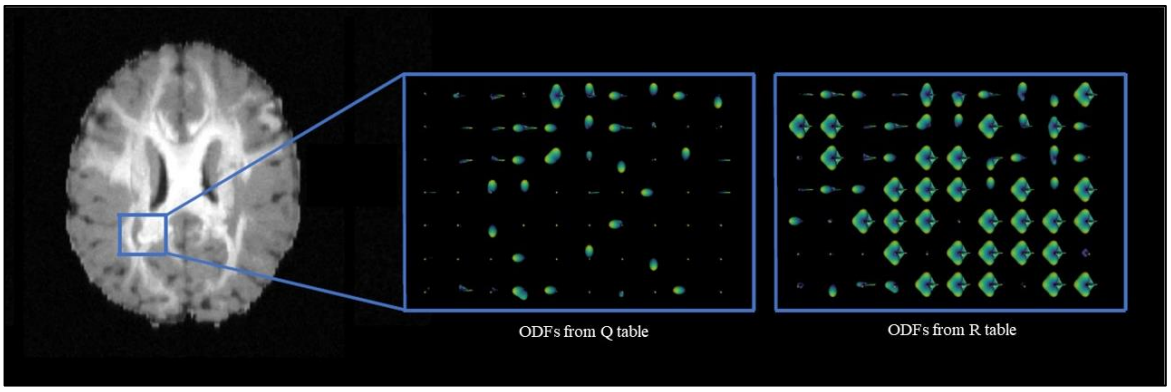

Figure 3.7. Comparison of fODFs obtained from r and q tables by superior-longitudinal-fasciculus and coronal-radiata bundle intersection region

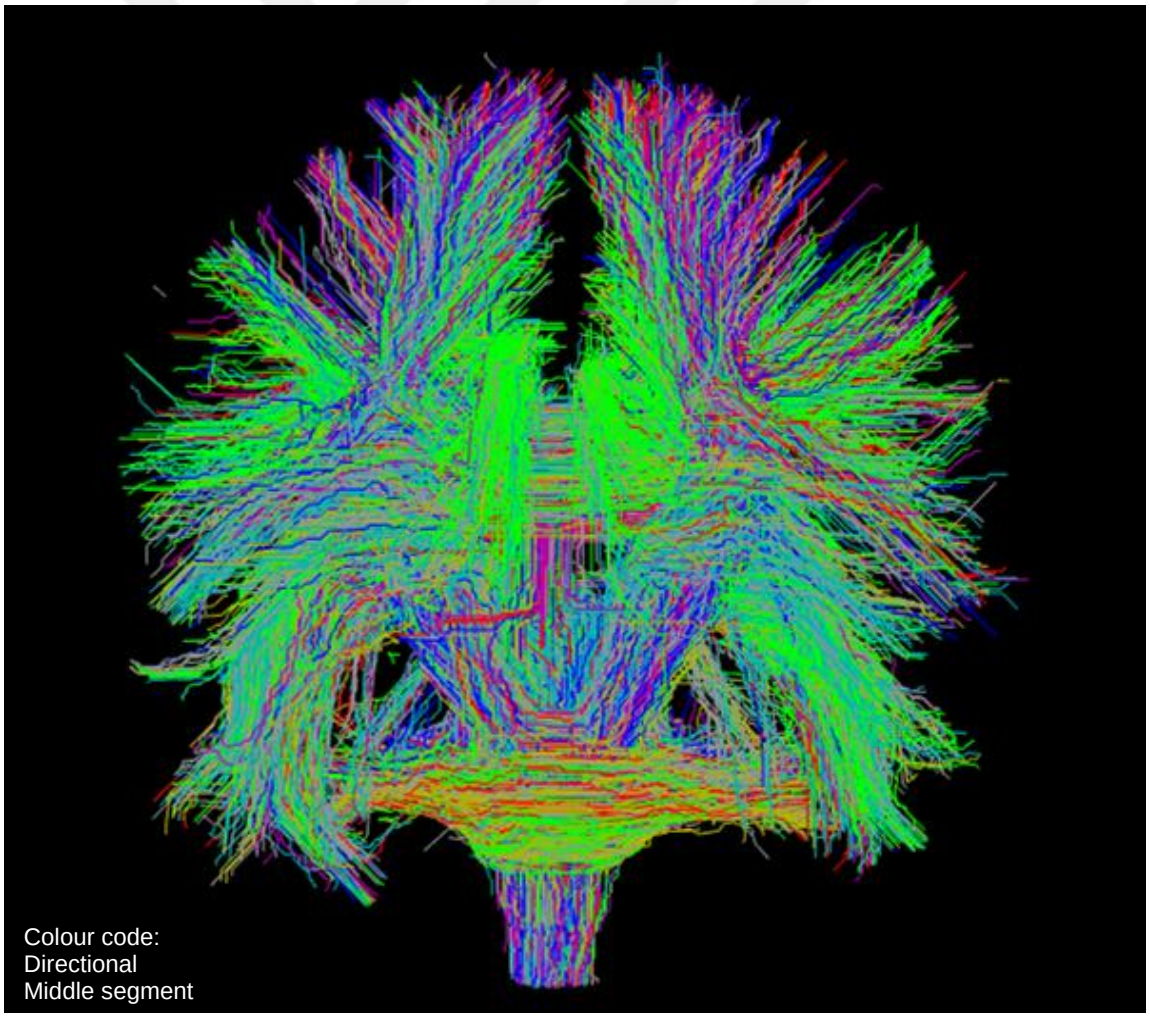

Figure 3.8. Tracts obtained by Q-RPTM

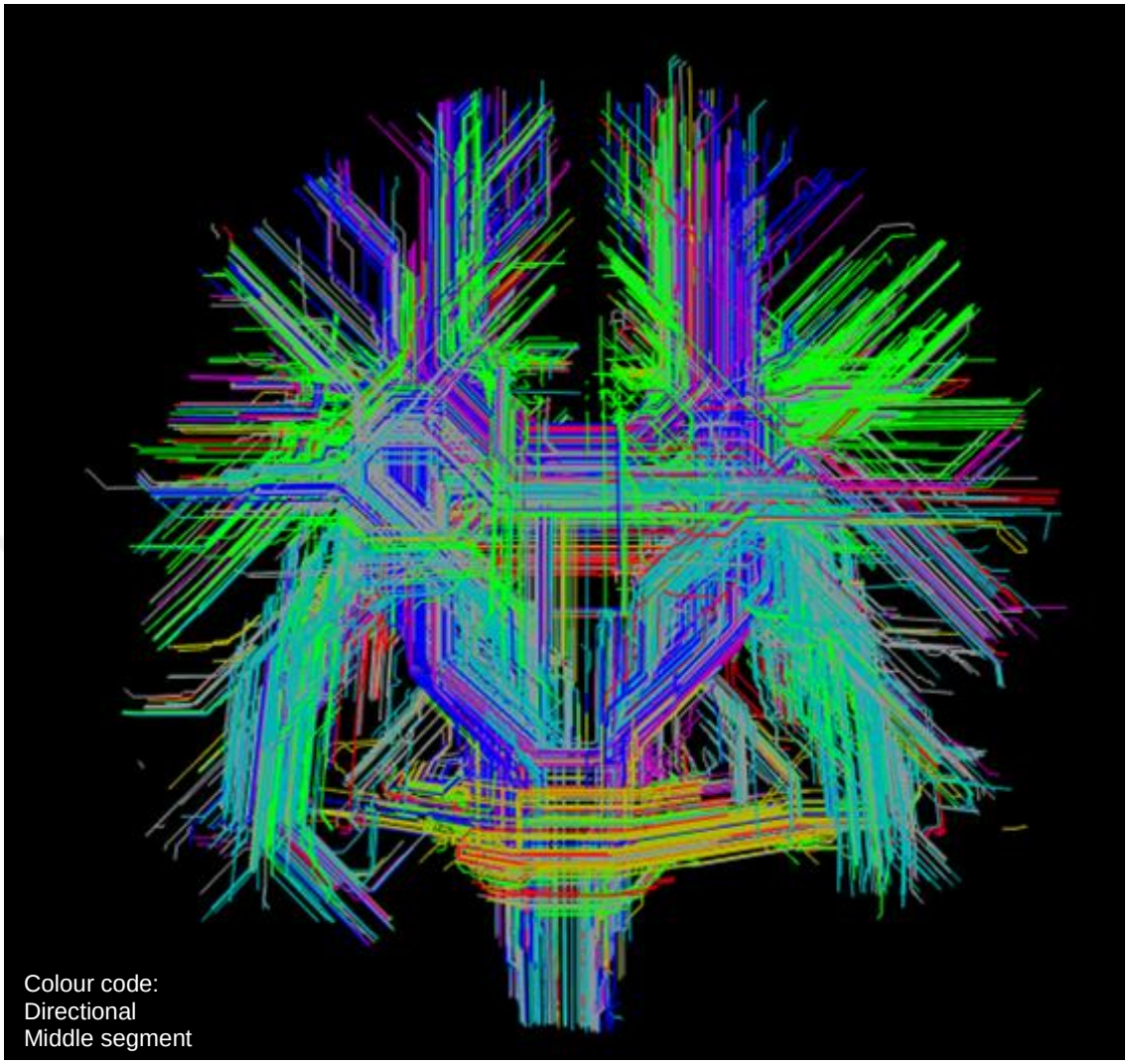

Figure 3.9. Tracts obtained by Q-RDTM

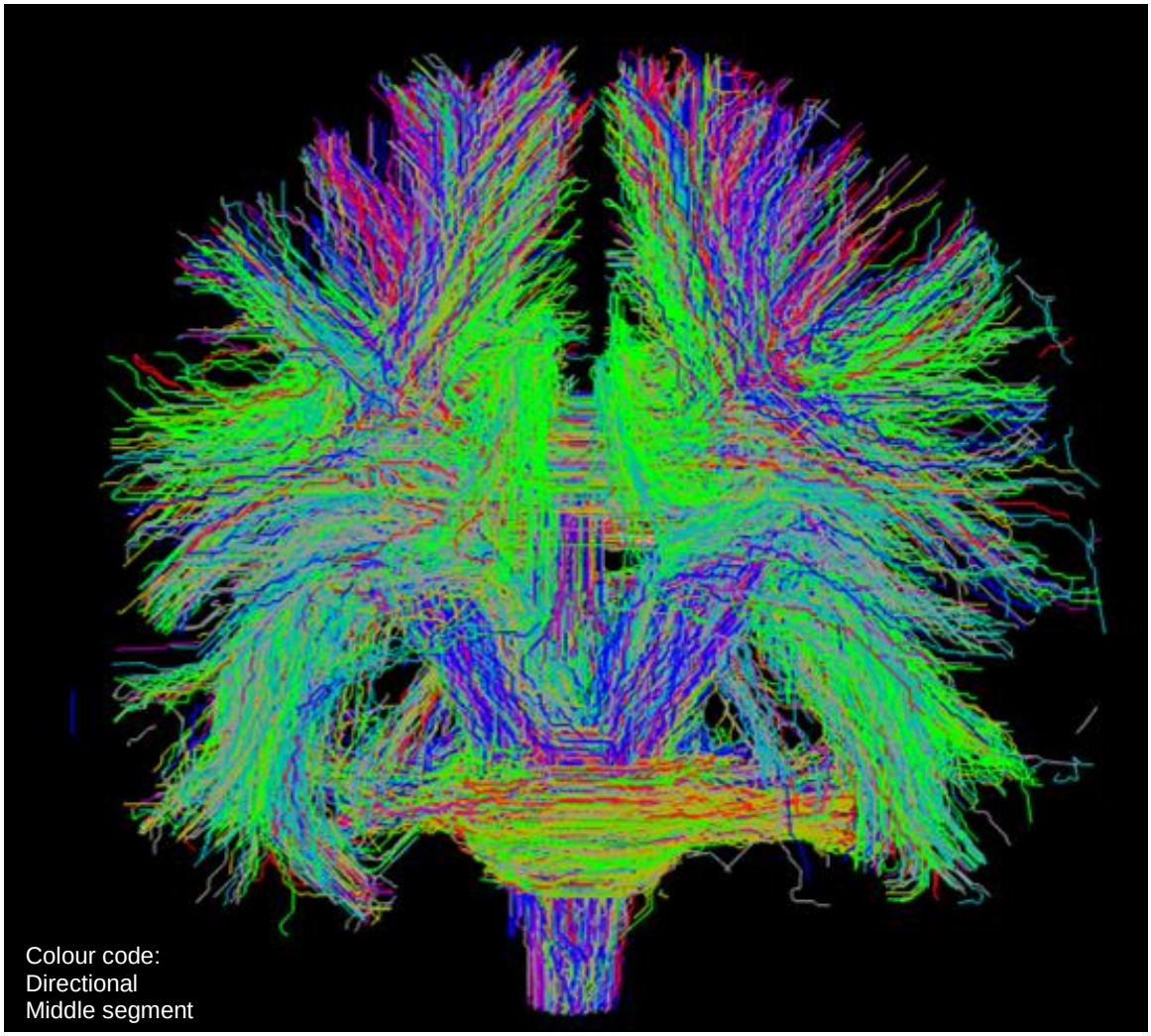

Figure 3.10. Tracts obtained by CPM

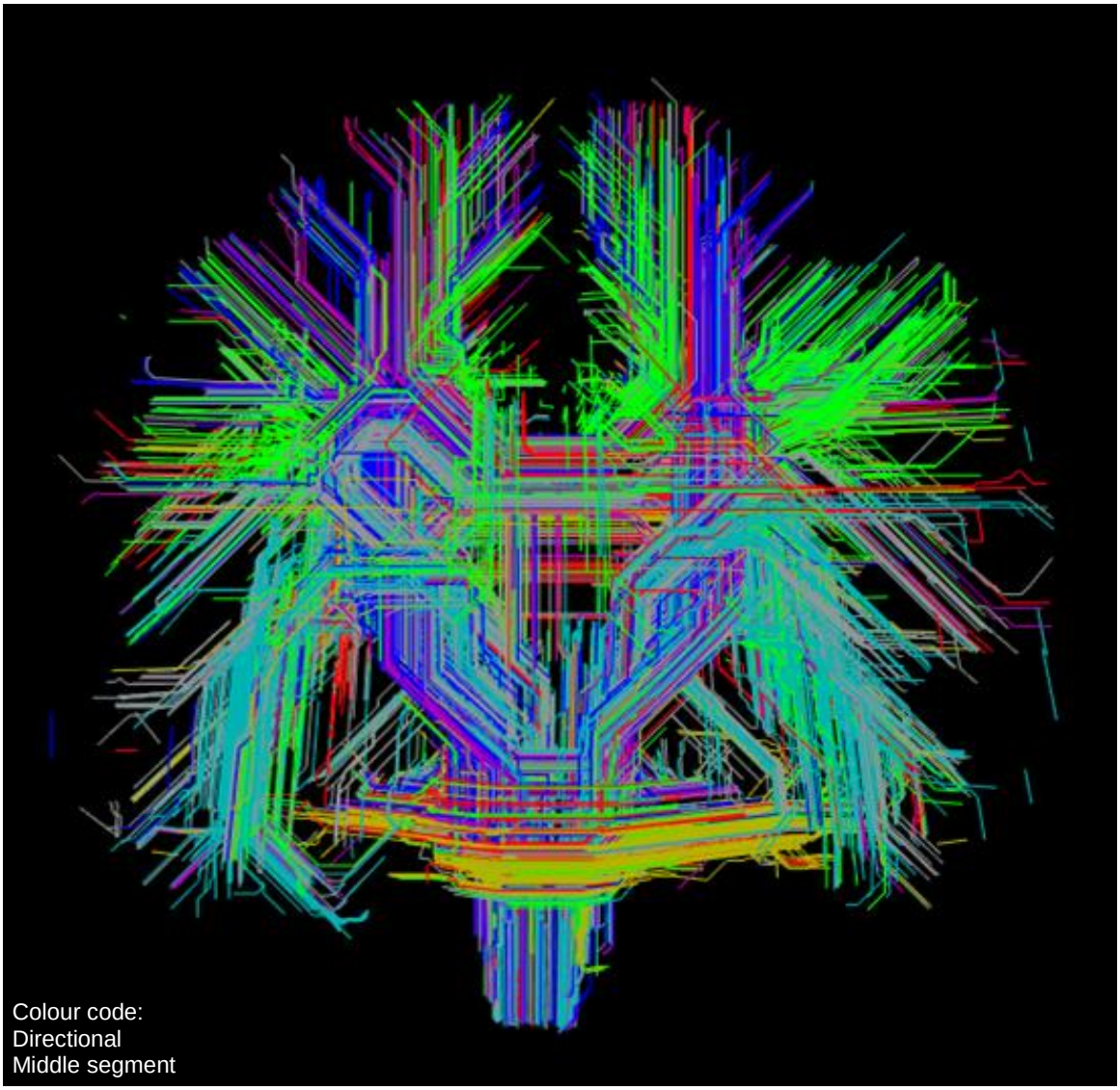

Figure 3.11. Tracts obtained by CDM

4. DISCUSSION

In this study, it is proposed to use the Q-routing method to identify fiber pathways on two different datasets and method compared to conventional methods.

In the first experiment, the diffusion phantom data is used. The results obtained by applying 3 different amounts of curl restriction were compared with the results of the deterministic flowline tractography method applied under the same conditions. Deterministic tractography tends to follow more of the pathways. The closest result was obtained when curls greater than 45° were restricted. However, since the number of steps is used as the stopping criterion in Q-routing, the agents continue their action after passing the target. The restriction of bends greater than 45° and the deterministic flowline methods were able to reach the correct target point when starting from both ends in the 3/7 pathway.

However, in cases where there is no restriction and 90°, correct results could be obtained from only one end in one of these 3 pathways. When fiber intersections were evaluated, all 3 methods failed to show success at intersection points. Supporting the reinforcement learning method with randomness (ε-greedy) may increase success at intersections by finding a more global best solution.

For the second experiment, most of the streamlines were not determined valid even though a visual inspection revealed that they were relatively close to the groundtruth bundles because of the huge number of close bundles in the data and the endpoint masks' voxel thickness is only 1. The fact that the training was carried out on unmasked data also caused the streamline scores of both the traditional and the proposed method to be low. Despite the fact that the traditional probabilistic method finds 1 extra bundle more, it slightly gives better results in all other metrics, revealing that the proposed method has a potential in big and realistic data. When the fODFs at the intersection of the Superior-Longitudinal-Fasciculus and the Coronal-Radiata are compared, the results obtained with the q table tend to follow the Coronal-Radiata, unlike the r table, which supports the effectiveness of the method. Apart from adding the epsilon greedy algorithm, supporting the algorithm with CNN networks can help the method to achieve faster and true positive global results.

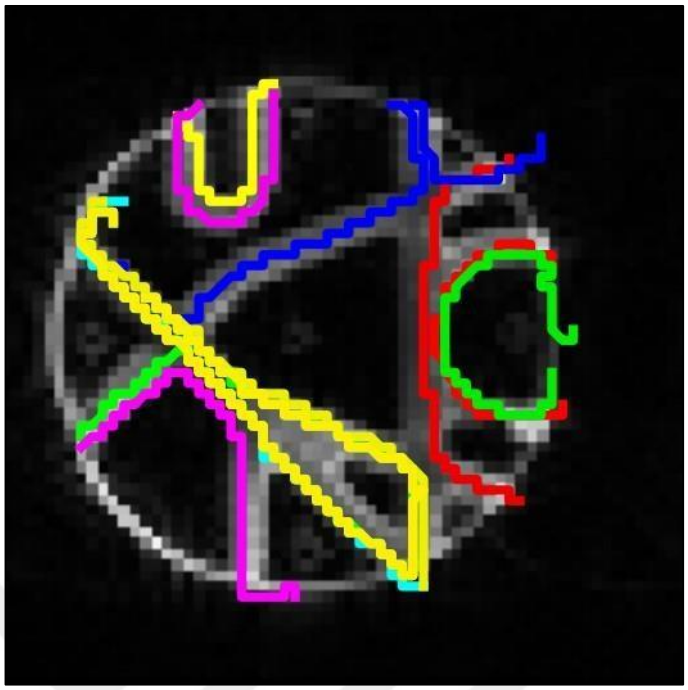

Figure 4.1. Results obtained with the deterministic flowline tractography method

The proposed method based on Q-routing includes both the global features of the shortest paths method and provides an application similar to the streamline methods. The results obtained demonstrate the method's potential for use in the diffusion MR fiber tractography issue.

5. CONCLUSION

There are now three significant families of tractography algorithms, each of which calculates the subsequent tracking step differently. The primary rebuilt direction is always chosen by the deterministic algorithms. In contrast, probabilistic algorithms use a probability density function for the next direction based on the diffusion model. Finally, global tractography algorithms attempt to simultaneously minimize a loss function and iteratively infer all directions. However, selecting a tractography algorithm is not the only decision to be made, as the diffusion signal modeling strategy also affects the tractograms that are generated. Recent studies have concentrated on modifying the diffusion model to better fit the tractography procedure.

The tractography inverse issue is inadequately, which suggests that, in our opinion, reinforcement learning techniques are best suited to handle it. Without the use of challenging and/or biased reference streamlines, RL techniques can take advantage of trial and errors' expressive capabilities. As a result, it is presented the Q-routing algoritm from reinforcement learning. The algorithm enables data-driven enhancements over traditional approaches by requiring a learning agent to iteratively improve its tracking capabilities through trial and error. On different datasets, competitive results presented.

REFERENCES

1. Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vivo fiber tractography using DT-MRI data. Magnetic resonance in medicine. 2000 Oct;44(4):625-32.

2. Kumar S, Miikkulainen R. Dual reinforcement Q-routing: An on-line adaptive routing algorithm. InProceedings of the artificial neural networks in engineering Conference 1997 Nov 9 (pp. 231-238).

3. Tournier JD, Calamante F, Connelly A. MRtrix: diffusion tractography in crossing fiber regions. International journal of imaging systems and technology. 2012 Mar;22(1):53-66.

4. Shinya S, Sasaki T, Nakagawa Y, Guiquing Z, Yamamoto F, Yamashita Y. The usefulness of diffusion-weighted imaging (DWI) for the detection of gastric cancer. Hepato-gastroenterology. 2007 Jul 1;54(77):1378-81.

5. Khoo MM, Tyler PA, Saifuddin A, Padhani AR. Diffusion-weighted imaging (DWI) in musculoskeletal MRI: a critical review. Skeletal radiology. 2011 Jun;40(6):665-81.

6. Baliyan V, Das CJ, Sharma R, Gupta AK. Diffusion weighted imaging: technique and applications. World journal of radiology. 2016 Sep 9;8(9):785.

7. Robitaille PM, Abduljalil AM, Kangarlu A, Zhang X, Yu Y, Burgess R, Bair S, Noa P, Yang L, Zhu H, Palmer B. Human magnetic resonance imaging at 8 T. NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo. 1998 Oct;11(6):263-5.

8. Tyler DJ, Robson MD, Henkelman RM, Young IR, Bydder GM. Magnetic resonance imaging with ultrashort TE (UTE) PULSE sequences: technical considerations. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2007 Feb;25(2):279-89.

9. Reichert IL, Robson MD, Gatehouse PD, He T, Chappell KE, Holmes J, Girgis S, Bydder GM. Magnetic resonance imaging of cortical bone with ultrashort TE pulse sequences. Magnetic resonance imaging. 2005 Jun 1;23(5):611-8.

10. Cawley PJ, Maki JH, Otto CM. Cardiovascular magnetic resonance imaging for valvular heart disease: technique and validation. Circulation. 2009 Jan 27;119(3):468-78.

11. Burdette JH, Durden DD, Elster AD, Yen YF. High b-value diffusion-weighted MRI of normal brain. Journal of computer assisted tomography. 2001 Jul 1;25(4):515-9.

12. Le Bihan D, Mangin JF, Poupon C, Clark CA, Pappata S, Molko N, Chabriat H. Diffusion tensor imaging: concepts and applications. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine. 2001 Apr;13(4):534-46.

13. Bhadelia RA, Price LL, Tedesco KL, Scott T, Qiu WQ, Patz S, Folstein M, Rosenberg I, Caplan LR, Bergethon P. Diffusion tensor imaging, white matter lesions, the corpus callosum, and gait in the elderly. Stroke. 2009 Dec 1;40(12):3816-20.

14. Skudlarski P, Schretlen DJ, Thaker GK, Stevens MC, Keshavan MS, Sweeney JA, Tamminga CA, Clementz BA, O'Neil K, Pearlson GD. Diffusion tensor imaging white matter endophenotypes in patients with schizophrenia or psychotic bipolar disorder and their relatives. American Journal of Psychiatry. 2013 Aug;170(8):886-98.

15. Sullivan EV, Pfefferbaum A. Diffusion tensor imaging and aging. Neuroscience & Biobehavioral Reviews. 2006 Jan 1;30(6):749-61.

16. Brand L. Vector and tensor analysis. Courier Dover Publications; 2020 Apr 15.

17. Kochunov P, Williamson DE, Lancaster J, Fox P, Cornell J, Blangero J, Glahn DC. Fractional anisotropy of water diffusion in cerebral white matter across the lifespan. Neurobiology of aging. 2012 Jan 1;33(1):9-20.

18. Jbabdi S, Woolrich MW, Andersson JL, Behrens TE. A Bayesian framework for global tractography. Neuroimage. 2007 Aug 1;37(1):116-29.

19. Côté MA, Boré A, Girard G, Houde JC, Descoteaux M. Tractometer: online evaluation system for tractography. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention 2012 Oct 1 (pp. 699-706). Springer, Berlin, Heidelberg.

20. Côté MA, Girard G, Boré A, Garyfallidis E, Houde JC, Descoteaux M. Tractometer: towards validation of tractography pipelines. Medical image analysis. 2013 Oct 1;17(7):844-57.

21. Behrens TE, Johansen-Berg H, Woolrich MW, Smith SM, Wheeler-Kingshott CA, Boulby PA, Barker GJ, Sillery EL, Sheehan K, Ciccarelli O, Thompson AJ. Noninvasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature neuroscience. 2003 Jul;6(7):750-7.

22. Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vivo fiber tractography using DT-MRI data. Magnetic resonance in medicine. 2000 Oct;44(4):625-32.

23. Sotiropoulos SN, Bai L, Morgan PS, Constantinescu CS, Tench CR. Brain tractography using Q-ball imaging and graph theory: Improved connectivities through fibre crossings via a model-based approach. Neuroimage. 2010 Feb 1;49(3):2444-56.

24. Zalesky A. DT-MRI fiber tracking: a shortest paths approach. IEEE transactions on medical imaging. 2008 Apr 18;27(10):1458-71.

25. Hamamci A. Cellular automata tractography: Fast geodesic diffusion MR tractography and connectivity based segmentation on the GPU. Neuroinformatics. 2020 Jan;18(1):25-41.

26. Kasenburg N, Liptrot M, Reislev NL, Ørting SN, Nielsen M, Garde E, Feragen A. Training shortest-path tractography: Automatic learning of spatial priors. NeuroImage. 2016 Apr 15;130:63-76.

27. Campbell JS, Pike GB. Potential and limitations of diffusion MRI tractography for the study of language. Brain and language. 2014 Apr 1;131:65-73.

28. Alhourani A, Richardson RM. Inherent limitations of tractography for accurate connectivity maps. Neurosurgery. 2015 Apr 1;76(4):N11-2.

29. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2019 Jul;9(4):e1312.

30. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015 Jul 17;349(6245):255-60.

31. Lloyd S, Mohseni M, Rebentrost P. Quantum algorithms for supervised and unsupervised machine learning. arXiv preprint arXiv:1307.0411. 2013 Jul 1.

32. Li Y. Deep reinforcement learning: An overview. arXiv preprint arXiv:1701.07274. 2017 Jan 25.

33. Dayan P, Niv Y. Reinforcement learning: the good, the bad and the ugly. Current opinion in neurobiology. 2008 Apr 1;18(2):185-96.

34. Sutton RS, Barto AG. Reinforcement learning. Journal of Cognitive Neuroscience. 1999 Jan 28;11(1):126-34.

35. Thrun SB. Efficient exploration in reinforcement learning.

36. Henderson P, Islam R, Bachman P, Pineau J, Precup D, Meger D. Deep reinforcement learning that matters. InProceedings of the AAAI conference on artificial intelligence 2018 Apr 29 (Vol. 32, No. 1).

37. Bravo M, Faure M. Reinforcement learning with restrictions on the action set. SIAM Journal on Control and Optimization. 2015;53(1):287-312.

38. Ostaszewski M, Trenkwalder LM, Masarczyk W, Scerri E, Dunjko V. Reinforcement learning for optimization of variational quantum circuit architectures. Advances in Neural Information Processing Systems. 2021 Dec 6;34:18182-94.

39. Szepesvári C. Algorithms for reinforcement learning. Synthesis lectures on artificial intelligence and machine learning. 2010 Jul 7;4(1):1-03.

40. Mnih V, Badia AP, Mirza M, Graves A, Lillicrap T, Harley T, Silver D, Kavukcuoglu K. Asynchronous methods for deep reinforcement learning. InInternational conference on machine learning 2016 Jun 11 (pp. 1928-1937). PMLR.

41. Boyan J, Littman M. Packet routing in dynamically changing networks: A reinforcement learning approach. Advances in neural information processing systems. 1993;6.

42. The Tractometer competition dataset; [cited 2022 23 August]. Available from:

http://tractometer.org/ismrm_2015_challenge/data

43. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. Neuroimage. 2007 May 1;35(4):1459-72.

44. Hamamci A. Cellular automata tractography: Fast geodesic diffusion MR tractography and connectivity based segmentation on the GPU. Neuroinformatics. 2020 Jan;18(1):25-41.

45. Yo TS, Anwander A, KnAusche T. Fiber cup 2009: Reconstructing fibers from the phantom data. InMICCAI workshop on Diffusion Modelling and the Fiber Cup (DMFC'09), London, United Kingdom 2009 (Vol. 7).

46. Garyfallidis E, Brett M, Amirbekian B, Rokem A, Van Der Walt S, Descoteaux M, Nimmo-Smith I, Dipy Contributors. Dipy, a library for the analysis of diffusion MRI data. Frontiers in neuroinformatics. 2014 Feb 21;8:8.