

# **3D Interactive Solution for Neuroanatomy Education**

Thesis submitted in partial fulfillment  
of the requirements for the degree of

*Master of Science*  
*in*  
*Computer Science and Engineering by Research*

by

Mythri V  
2018201028

mythri.v@research.iiit.ac.in



International Institute of Information Technology  
Hyderabad - 500 032, INDIA  
April 2023

Copyright © Mythri V, 2023  
All Rights Reserved

International Institute of Information Technology  
Hyderabad, India

**CERTIFICATE**

It is certified that the work contained in this thesis, titled “3D Interactive Solution for Neuroanatomy Education” by Mythri V, has been carried out under my supervision and is not submitted elsewhere for a degree.

---

Date

---

Adviser: Prof. Jayanthi Sivaswamy

To my family and friends

## **Acknowledgments**

First and foremost, I would like to express earnest gratitude to Prof. Jayanthi Sivaswamy for having faith in me and letting me be a part of the medical image processing lab. I will always be indebted to her for her immense support and guidance in my research and our project. Her insights always focused and centered me on the main problem at hand. I appreciate her motivation and extreme patience with my failures and the consideration she gave me when I had to switch from software development to research. She will always be the most inspirational woman I have ever met, and I aspire to be like her someday.

I would also like to extend my gratitude to Dr. Keshavdas, Dr. Jayadevan, Prof. Doris, Josline, Abhijith, and the rest of the Kerala team who collaborated with me on the "3D Interactive Solution for Neuroanatomy Education" project. I am grateful for all the support and camaraderie I found among members of the MIP group. I want to thank Alphin and Aabhas for their contribution and support while working with me and the other group members, Prathyusha, Abhinav, Naren, Anurag, Aniket, and Soumyasis, for their motivation and insightful discussions.

I am blessed to have an incredibly supportive family who always understood and encouraged me to aim higher. I thank my parents for their constant love and faith in me. I would also like to thank my manager at Softlink, Aditya, who sparked my interest in research. Finally, my journey at IIIT would be incomplete without my friends. I thank Tanu, Prathyusha, Archi, Anoop, Kiruthika, Sowmya, and others for keeping my mind sane, my spirits alive, and for many wonderful memories.

## Abstract

Typically, anatomy is taught through dissection, 2D images, presentations, and cross-sections. While these methods are convenient, they are non-interactive and fail to capture the spatial relationships and functional aspects of anatomy. New visualization technologies such as virtual reality and 3D can compensate for the impediments and provide better understanding while captivating the students. With recent advances in the industry, the methods to provide a 3D experience are economical. In this thesis, we introduce a low-cost 3D-interactive anatomy system designed for an audience of typical medical college students in an anatomy class. The setup used to achieve 3D visualization is Dual projector polarization. While there are other ways to achieve 3D visualization, like alternate frame sequencing and virtual reality, this technique can target a large audience and requires minimum accessories for the setup enabling this to be a low-cost solution for an immersive 3D experience. The 3D interactive Neuroanatomy solution is an end-to-end framework capable of designing anatomy lessons and visualizing the 3D stereoscopic projection of those anatomy lessons. To ensure superior comprehension of students, we incorporate each teacher's unique teaching approach while developing anatomy lessons by providing the ability to create their own lessons. We have created anatomy lessons based on the human brain which is a vital organ and has a complex anatomy. Our aim is to help medical students to understand the complexity of organ systems from not just an anatomical perspective but also a radiological perspective. We use annotations on clinical case data such as MRI, MRA, etc., to create 3D models for anatomy visualization incorporating clinical information and illustrating real cases. Annotations for structures of interest are done using manual, automatic, and semi-automatic segmentation methods.

Manual delineation of the structure boundaries is very tedious and time-consuming. Automatic segmentation is quick and convenient. However, manual annotations were done for the 3D anatomy viewer for small and complex structures due to substandard automatic segmentation. There is a need to improve automatic segmentation performance for those structures. While segmentation is an essential step in 3D modeling, it plays a critical role in many neurological disease diagnoses as well, which are associated with degradation in the sub-cortical region. Therefore accurate algorithms are needed for sub-cortical structure segmentation. Variance in the size of structures is significant, which introduces a performance bias towards larger structures in many deep learning approaches. In this part of the thesis, we aim to remove size bias in sub-cortical structure segmentation. The proposed method addresses this problem with a pre-training step used to learn tissue characteristics, an ROI extraction step that aids in focusing on the local context, and using structure ROIs elevates the influence of smaller structures in the network.

# Contents

Chapter	Page
1 Introduction . . . . .	1
1.1 Purpose . . . . .	2
1.2 Related Works . . . . .	2
1.3 3D Visualization . . . . .	2
1.4 Thesis Focus . . . . .	3
1.5 Contribution . . . . .	4
1.6 Outline of thesis . . . . .	4
2 Use Case: Human Brain . . . . .	6
2.1 Anatomy of human brain . . . . .	7
2.1.1 Cerebrum . . . . .	8
2.1.2 Cerebellum . . . . .	8
2.1.3 Brainstem . . . . .	9
2.1.4 Subcortical Structures . . . . .	9
2.1.5 Blood vessels . . . . .	10
2.1.6 White Matter Fibers . . . . .	10
2.2 Neuroimaging . . . . .	11
2.2.1 Structural NeuroImaging . . . . .	11
2.2.2 Functional NeuroImaging . . . . .	12
2.2.3 Brain Atlas . . . . .	13
3 System Design . . . . .	14
3.1 Overview . . . . .	14
3.2 Hardware Design . . . . .	15
3.2.1 3D stereoscopic projection . . . . .	15
3.2.2 Dual Projector Setup . . . . .	16
3.3 Software Design . . . . .	17
3.3.1 Segmentation . . . . .	17
3.3.2 Model Generation . . . . .	17
3.3.2.1 3D Mesh . . . . .	17
3.3.2.2 Volume Rendering . . . . .	18
3.3.3 Authoring Tool . . . . .	18
3.3.4 Lesson Design Considerations . . . . .	19
3.3.4.1 Anatomical Relationship Grouping . . . . .	19
3.3.4.2 Integration of anatomy and radiology . . . . .	19

3.3.4.3	Visual Properties of Anatomical Structures . . . . .	19
3.3.4.4	Structure Labeling . . . . .	19
3.3.5	3D Anatomy Viewer . . . . .	20
4	Data and Annotation . . . . .	21
4.1	Data . . . . .	21
4.2	Segmentation . . . . .	22
4.2.1	Automatic Segmentation . . . . .	23
4.2.2	Manual Segmentation . . . . .	23
4.2.3	Semi-Automatic Segmentation . . . . .	24
4.2.4	White Matter Tractography . . . . .	24
5	Automatic Segmentation: A Method to remove size bias . . . . .	25
5.1	Introduction . . . . .	25
5.2	Related Work . . . . .	26
5.3	Background . . . . .	26
5.3.1	Normalization . . . . .	27
5.3.2	Activation Layer . . . . .	27
5.3.3	Residual Block . . . . .	27
5.3.4	Dense Block . . . . .	28
5.3.5	Channel Attention (Squeeze and Excitation block) . . . . .	28
5.4	Method . . . . .	29
5.4.1	Proposed Method . . . . .	29
5.4.1.1	Pre-training with tissue segmentation . . . . .	29
5.4.1.2	Atlas-guided ROI extraction . . . . .	30
5.4.2	Loss function . . . . .	30
5.5	Dataset and Implementation details . . . . .	31
5.5.1	Dataset description . . . . .	31
5.5.1.1	IBSR Dataset . . . . .	31
5.5.1.2	MICCAI dataset . . . . .	31
5.5.2	Pre-processing . . . . .	31
5.5.3	Training details . . . . .	32
5.6	Experiments and Results . . . . .	32
5.6.1	Evaluation Metrics . . . . .	32
5.6.2	Ablation Studies . . . . .	33
5.6.3	Qualitative Results . . . . .	36
5.6.4	Comparison with other methods . . . . .	37
5.7	Discussion and Conclusion . . . . .	38
6	Solution Implementation . . . . .	39
6.1	Authoring Tool . . . . .	39
6.1.1	Lesson Creation Workflow . . . . .	40
6.1.1.1	Loading 3D Models . . . . .	40
6.1.1.2	Structure Properties . . . . .	40
6.1.1.3	Labeling . . . . .	41
6.1.1.4	Label Hierarchy . . . . .	41
6.1.1.5	Anatomical Planes . . . . .	42

6.1.1.6	Saving a Lesson . . . . .	43
6.2	Anatomy Viewer . . . . .	43
6.2.1	Lesson Configuration . . . . .	44
6.2.2	Lesson Launch . . . . .	44
6.2.3	3D Models and Views . . . . .	45
6.2.3.1	3D View . . . . .	45
6.2.3.2	Canonical Planes . . . . .	46
6.2.3.3	Others . . . . .	46
6.2.4	Remote Companion Application . . . . .	47
6.2.4.1	Linux Client . . . . .	47
6.2.4.2	Android application . . . . .	47
6.3	Software Specifications . . . . .	47
6.4	System Assessment via User Feedback . . . . .	48
7	Discussion and Future Work . . . . .	49
7.0.1	Future work . . . . .	50
	Bibliography . . . . .	52
	<i>Appendix A: Images of our Solution</i> . . . . .	59
A.1	Photos of the Setup . . . . .	59
A.2	Images of the Authoring Tool and 3D Anatomy Viewer . . . . .	60
	<i>Appendix B: Toolboxes</i> . . . . .	62
B.1	Freesurfer (FSL) . . . . .	62
B.2	ITK-SNAP . . . . .	63
B.3	BrainSuite . . . . .	63
B.4	DSI Studio . . . . .	63
B.5	Blender . . . . .	63

## List of Figures

Figure	Page
2.1 Cross section of brain (Image courtesy: Wikipedia) . . . . .	6
2.2 (a) Anatomy of a neuron (b) White matter and gray matter (c) Cerebrospinal fluid (CSF) (Image courtesy: Wikipedia) . . . . .	7
2.3 Anatomical regions (Image courtesy: Wikipedia) . . . . .	7
2.4 (a) Lobes of the brain (b) Motor and sensory regions of the cerebral cortex (Image courtesy: Wikipedia) . . . . .	8
2.5 (a) Subcortical structures of brain (b) Blood supply to the brain (Image courtesy: Wikipedia)	10
2.6 (a) Different MRI protocols (b) MRA and MR veinogram (Reproduced with permission from Mikulis, D.J. and Roberts, T.P. (2007), Neuro MR: Protocols. J. Magn. Reson. Imaging, 26: 838-847.) (c) CT (Case courtesy of Assoc Prof Frank Gaillard, Radiopaedia.org, rID: 14109) (d) DTI (Image courtesy: Wikipedia) . . . . .	11
2.7 (a) fMRI during language task (Case courtesy of Dr Alexandra Stanislavsky, Radiopaedia.org, rID: 12534) (b) PET and SPECT scans of the healthy and PS patient (Adapted from (Jakobson Mo et al., 2018) under the terms of the Creative Commons Attribution 4.0 International License) . . . . .	12
2.8 Freesurfer (Desikan-Killiany) atlas mapped to pial surface(parcellation) [12] . . . . .	13
3.1 Pipeline of 3D Interactive solution for anatomy education . . . . .	14
3.2 3D Dual Projector Setup . . . . .	16
3.3 Volume Rendering of MR data . . . . .	18
4.1 (a) Manual segmentation of MRI data on ITKSnap (b) Semi-automatic segmentation of blood vessels on CTA data . . . . .	22
4.2 Fiber tracking on DTI data in DSI Studio . . . . .	24
5.1 Proposed 2-phase training framework for sub-cortical structure segmentation. Phase 1: Pre-training with tissue segmentation, Phase 2: Structure segmentation from the atlas guided 3D structure ROI. . . . .	29
5.2 Atlas with subcortical structures annotations only . . . . .	32
5.3 Plots show the % improvement in the Dice score of structures of different sizes for the proposed method on U-Net and Res. U-Net on (a) IBSR dataset (b)MICCAI dataset . .	35
5.4 Qualitative results of the proposed method and its variants.Top to bottom: Coronal, Axial, and sagittal slices. (a) IBSR dataset (b) MICCAI dataset . . . . .	36
6.1 Authoring tool . . . . .	39

6.2	Flow diagram of loading models and setting structure properties . . . . .	40
6.3	Flow diagram for labeling structures and organizing their hierarchy . . . . .	41
6.4	(a) Grouping label feature in authoring tool to create label hierarchy (b) Label Hierarchy of Left ventricle . . . . .	42
6.5	Anatomical planes and positions . . . . .	43
6.6	Viewer configuration . . . . .	44
6.7	Lesson Launcher . . . . .	44
6.8	(a) 3D view of all cortical structures (b) Volume rendering of MR data (c) Canonical planes overlayed on 3D structures . . . . .	45
6.9	(a) 3D view of corpus callosum (b) white matter fibers of corpus callosum s . . . . .	46
6.10	(a) A popup to connect to server's network (b) Android application main view . . . . .	47
A.1	Images of 3D anatomy viewer projected on the silver screen, laptop, and tablet. . . . .	59
A.2	(a) Zoomed-in view of the viewer on the silver screen (b) Dual projectors with polarizers . . . . .	60
A.3	Authoring Tool . . . . .	60
A.4	3D Anatomy Viewer in normal mode . . . . .	61
A.5	3D Anatomy Viewer in stereo mode . . . . .	61

## List of Tables

Table		Page
5.1	Performance analysis of the proposed method. Dice scores averaged over 6 or 5 folds are listed for a baseline Res. U-net and its variants: Trained with Proposed method, only pre-training (PT) and only ROI training (ROI). . . . .	34
5.2	Average performance of sub-cortical structure segmentation for all structures using proposed method(PT+ROI) on various architectures on IBSR dataset and MICCAI dataset.	35
5.3	Comparison of average performance of sub-cortical structure segmentation for all structures between standard (available in toolboxes),state of the art solutions and proposed method on IBSR dataset. . . . .	37
5.4	Comparison of average performance of sub-cortical structure segmentation for all structures between standard (available in toolboxes),state of the art solutions and proposed method on MICCAI dataset. . . . .	37

## *Chapter 1*

### **Introduction**

Generally, undergraduate medical education in anatomy comprises a lecture with PowerPoint slides and occasionally gross anatomical models. In addition, a gross dissection of a cadaver is performed by the instructor. The use of cadavers poses multiple issues; there is limited availability of cadavers, the cost of obtaining and maintaining them is high, and there are ethical issues in using cadavers. There are numerous self-study methods, including websites dedicated to different anatomies and mobile applications. However, in a classroom setting, there is a heavy focus on 2D anatomy methods such as PowerPoint presentations, animations, and 2D atlas, since these teaching aids are easily accessible and cost-effective. However, with these approaches, cognitive leaps between textbooks, verbal lectures, presentations, and dissection are challenging for students as they have varied understanding levels and skills with spatial mappings. Limited time available for such anatomical teaching further hampers their learning. Not much effort has been made to include radiological data and correlate it with the anatomy, which would aid them in diagnoses since radiological data is extensively used for that purpose. Newer technologies are needed to overcome these challenges in anatomy teaching.

With the rapid advancement in technology, better visualization techniques such as virtual reality, augmented reality, and 3D visualization can be utilized to teach anatomy, which not only helps in better understanding but also helps in engaging and motivating the students. Based on various studies and articles, in general, medical students prefer newer three-dimensional visualization methods to traditional methods for learning anatomy [73]. 3D visualization especially helps in learning spatial anatomy compared to traditional 2D visualization. Incorporating 3D interactive visualization tools in addition to traditional methods augments current learning and helps retention of visual information; additional radiological data helps develop diagnostic skill sets in students. Teachers of anatomy are key players who support students' comprehension. Every teacher has their own style and method of instruction, and with our solution, we intend to incorporate their objectives and concepts when designing lessons. Consequently, we propose a 3D interactive visualization tool that is feasible and cost-effective, keeping in mind the needs of anatomy education in a typical classroom of medical colleges in India.

## 1.1 Purpose

The major aim is to boost the intuitive understanding of anatomy in students with newer interactive 3D visualization methods, which are cost-efficient and feasible. Introducing the 3D Interactive solution that incorporates radiology into anatomy teaching as an integral part of the regular curriculum is also a prime objective. Involving the professors in designing the study material is a key aspect in achieving this objective of integrating it into the regular curriculum.

## 1.2 Related Works

Many studies focus on neuroanatomy [1, 4, 11, 38, 62] while others focus on other anatomical structures like ear[53], larynx[72], abdomen [15] etc. De Faria et al. [11] constructed stereoscopic images acquired from the dissected brain and used an anaglyph system to project these stereoscopic images as interactive videos viewed through personal computers. The students were evaluated using a written test, and their subjective opinion was considered. They concluded that their method presented a significant gain of neuroanatomical knowledge compared to traditional methods. Kockro et al. [38] used Dextrobeam, a virtual reality environment projecting multimodality images as 3D interactive objects, for the stereoscopic 3D lecture on the third ventricle for a large group of students, and they rated the 3D lecture as superior in terms of spatial understanding, effectiveness and enjoyability. Agbetoba et al. [1] use a 3D virtual planning software and a standard DICOM software RadiAnt showing 2D cross-sections for 2 groups for evaluation. Various studies Ruisoto et al[62] (neuroanatomy),Nicholson et al[53] (ear), Venail et al [75](temporal bone) including [11, 38, 1], infer that with 3D technologies students score significantly better. However, some studies Tan et al. (Larynx) [72], Donnely et al. (abdomen) [15], Keedy et al. (hepatobiliary) [36] found while there was no significant improvement in test scores with 3D methods compared to traditional methods, the subjective opinion of the students was positive.

## 1.3 3D Visualization

3D visualization of anatomy has evolved from physical models to virtual representations with the progress of technology. Virtual or 3D anatomy representations can be easily developed with images captured using a digital camera at various angles of a specimen. One can also use clinical imaging modalities such as Magnetic resonance imaging (MRI), Computed tomography (CT) to create a 3D model. Computer graphics techniques such as surface meshes and volume rendering can provide appealing 3D effects. These 3D models can be superimposed with the raw modality data (MRI, CT) to provide a greater understanding of anatomy as well as radiology.

Stereoscopy creates the impression of three-dimensional depth from two-dimensional images. This is done by providing offset images separately to the left and right eye, which is perceived as 3D depth

when combined. Visualization of 3D data or objects can be achieved through various techniques. These can be broadly divided into categories which are:

1. **Virtual Reality:** Virtual reality generally refers to technology that provides an immersive experience in an artificial world. The experience can be visual, auditory, and sometimes haptic, usually achieved with a VR headset with a display system and sensors to track the gaze and render the appropriate 3D scene. These headsets provide a 3D experience on an individual level and are more expensive compared to other methods. However, an effort is being made to reduce the cost to achieve the same level of immersive experience as in high-end VR headsets.
2. **Augmented Reality:** Augmented reality, unlike virtual reality, enhances the way we perceive the natural environment by overlaying digital information onto the real world. Perception is not only limited to vision but auditory, olfactory and haptic senses as well. AR can be achieved by devices that can combine the real and virtual worlds, such as smartphones, glasses, and headsets. Similar to VR, it is an individual experience and expensive.
3. **Mixed Reality:** Mixed reality merges real and virtual worlds, where physical and virtual objects interact and co-exist. It is a hybrid of virtual and augmented reality.
4. **3D stereoscopic projection:** This is one of the oldest and most widely used 3D technologies. Most commonly seen in 3D movies. The projection systems can be categorized into active and passive systems. It is intended for multiple viewers and is feasible and cost-effective compared to the above methods.

## 1.4 Thesis Focus

In this thesis, we focus on the main issue of improving the standard of anatomy education. We focus on development of our 3D interactive solution, which provides a framework for anatomy instructors to create content for teaching, the segmentation methods used in the process of content creation, and the issue of size bias in these segmentation methods. Generally, undergraduate medical education in anatomy comprises a lecture with PowerPoint slides and occasionally gross anatomical models. In addition, a gross dissection is performed by the instructor. Current education focuses heavily on 2D anatomy since the teaching aids for such are easily accessible. With the advent of technologies, there is a need to incorporate 3D visualization tools to aid anatomy education. We propose an interactive 3D solution for anatomy education for a typical classroom in medical colleges in India, which is feasible and cost-efficient. The interactive 3D visualization tool makes it easier for students to learn discrete structures and spatial and functional relationships between them and retention of visual information. Integrating raw radiological data with 3D models derived from its annotations is an added benefit.

The annotations needed for the purpose of education have to be sourced from various automatic and manual methods. Manual annotation methods are laborious, and for our purpose, the automatic methods are insufficient. Hence using both methods and combining their results is essential and most effective.

Segmentation and analysis of subcortical structures like the hippocampus, pallidum, and thalamus are of interest in anatomical teaching since they play a vital role in the functioning of humans. They also help in diagnosing some neurological diseases such as Alzheimer’s disease, depression, etc. Segmentation is a challenging task because of brain tissue ambiguity and data scarcity. Deep learning (DL) solutions are widely used for this purpose by considering the problem as a semantic segmentation of the brain. In general, many approaches exhibit a performance bias towards larger structures when training is done on the whole brain, where a significant variance in size exists. We propose a method to address this problem wherein its two phases are a pre-training step and an ROI extraction step. Tissue segmentation is used for pre-training to learn tissue characteristics. A rough ROI extraction step for structures boosts the weightage of smaller structures by focusing on the specifics of structures.

We propose a suite of software tools for anatomy education. An authoring tool to achieve the task of designing an anatomy lesson using the annotations obtained from clinical data, in a meaningful way to create coherent content. We explore various tools that the authoring tool provides to create a lesson. The 3D interactive anatomy viewer loads the lessons created using the authoring tool. It renders the 3D stereo structures of the anatomy using 3D stereoscopic projection to achieve the 3D effect.

## **1.5 Contribution**

The major contributions of this thesis are as follows:

1. 3D interactive system for Neuroanatomy education
  - (a) Authoring tool for lesson creation
  - (b) 3D Anatomy Viewer for interactive 3D visualization
  - (c) Remote companion application for user-friendly touch interface
2. 3D models and lessons created for neuroanatomy
3. A deep learning-based method to remove size bias in subcortical structure segmentation.

## **1.6 Outline of thesis**

This thesis is organized as follows: In chapter 2, we explore specifics of our use case of the human brain to highlight the potential of the 3D interactive system for anatomy instruction. We examine how neuroanatomy is generally taught, and the role neuroimaging plays in capturing anatomical intricacies,

which we further utilize. In chapter 3, we discuss the system design of our solution, providing a high-level overview of the system pipeline. The significance of instructor-guided anatomy education is also emphasized. We further explore the hardware design used to achieve an immersive 3D experience in a classroom setting. We also delve into software design aspects of the system pipeline, the usage of clinical data, annotating structures of interest, various techniques for 3D modeling, and important features required for designing a coherent lesson. In chapter 4, we delve into the details of data acquisition using clinical scanners, manual, automatic, and semi-automatic segmentation methods used for generating annotations. Chapter 5 investigates the problem size bias observed in many automatic segmentation methods, specifically in subcortical structure segmentation. We propose a two-phased approach to address this problem, applicable to different deep architectures and on different datasets to establish the effectiveness of the proposed approach. In chapter 6, we provide in-depth insights into the implementation of our 3D interactive system. We outline tools available to design a lesson and showcase the lesson creation workflow. We also take a look at how anatomy viewer launches lessons created using the authoring tool and implementation of 3D models and views for different anatomical structures. The features and working remote companion application are also examined. The last chapter summarizes the proposed solutions and recommends possible future work.

## Chapter 2

### Use Case: Human Brain

The human brain is one of the most complex organs, which not only controls all body functions but it also helps us interact with the outside world and interpret and analyze the information obtained from the outside world and make decisions based on that. It embodies the essence of human intelligence and creativity. The brain controls everything from the movement of our bodies and the functions of every organ to speech, thoughts, and memories. The brain receives information from the five senses, vision, smell, hearing, touch, and taste, and it determines how to respond to stimuli.

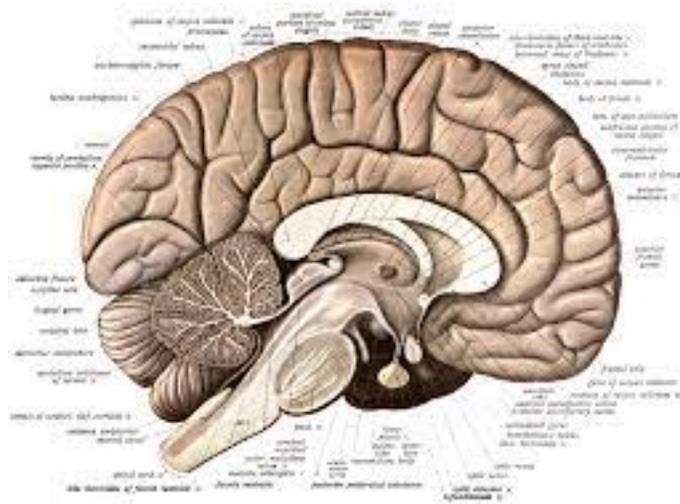


Figure 2.1: Cross section of brain (Image courtesy: Wikipedia)

The composition of the brain includes fats, water, proteins, carbohydrates, and salts. It also consists of blood vessels and billions of nerve fibers. The brain can be divided into different components, white matter, and gray matter.

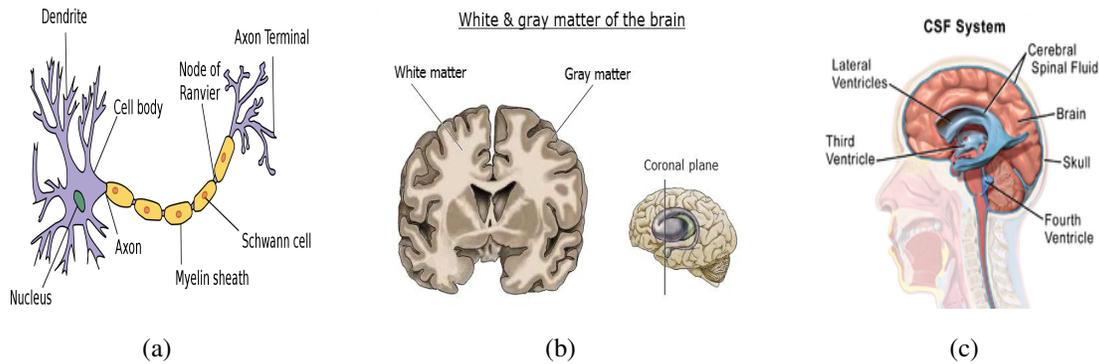


Figure 2.2: (a) Anatomy of a neuron (b) White matter and gray matter (c) Cerebrospinal fluid (CSF) (Image courtesy: Wikipedia)

Gray matter makes up the outer layer of the brain. It consists of neural cell bodies called soma and gets its grey tone due to the high concentration of soma. Gray matter is abundant in the cerebrum, cerebellum, and brain stem. It extends into the spinal cord as well. The white matter area mainly consists of myelinated axons. White matter is predominantly present in deeper areas. Most of the brain's deeper structures are white matter covered by gray matter. Another essential component of the brain is cerebrospinal fluid (CSF). It is a colorless liquid surrounding the brain, acts as a mechanical barrier, and provides lubrication. It also helps maintain the pressure within the skull.

## 2.1 Anatomy of human brain

Anatomically, the three major components of the human brain are the cerebrum, cerebellum, and brain stem.

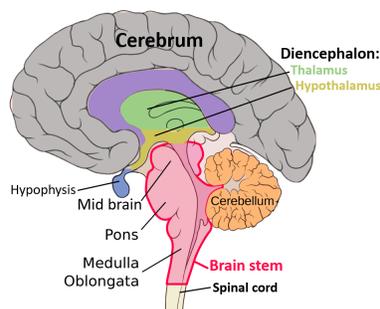


Figure 2.3: Anatomical regions (Image courtesy: Wikipedia)

### 2.1.1 Cerebrum

The cerebrum is the largest part of the human brain. It is divided into left and right hemispheres. The surface of the cerebrum, called the cortex, consists of ridges (gyri) and valleys (sulci). Such folding of the cortex increases the brain's surface area, allowing a large number of neurons to fit, which enables higher functions. Each hemisphere has distinct fissures which divide the brain into four lobes: the frontal, parietal, temporal, and occipital.

The frontal lobes are located directly behind the forehead, while the occipital lobes are at the back of the skull, and parietal lobes are found near the center between the frontal and occipital lobes. The temporal lobe is located close to the base of the skull. The frontal lobe is essential for cognitive functions, such as thoughts, moods and reasoning, and motor functions. The temporal lobe, also called the neo-cortex, processes speech and sensory information, which contributes to the retention of memories and emotions. The parietal lobe integrates and interprets input from various senses and is vital for spatial orientation and navigation. Visual processing takes place in the occipital lobe.

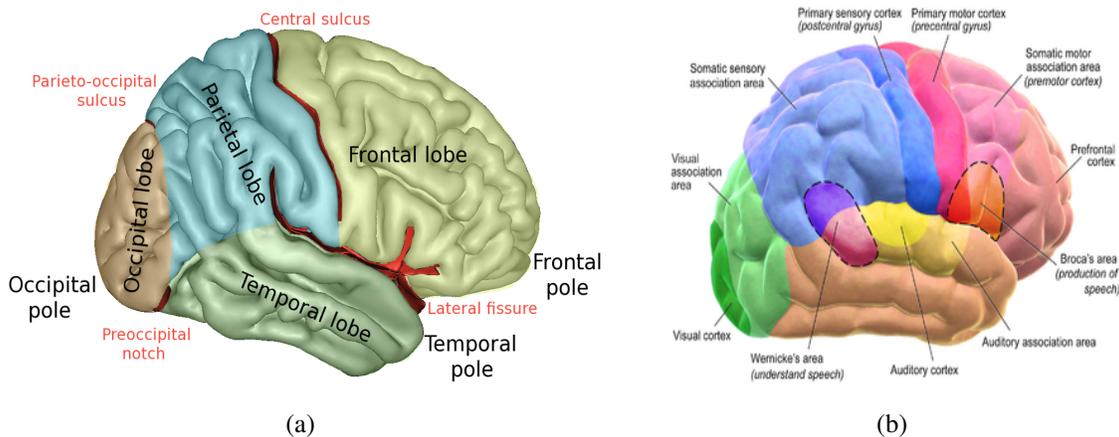


Figure 2.4: (a) Lobes of the brain (b) Motor and sensory regions of the cerebral cortex (Image courtesy: Wikipedia)

### 2.1.2 Cerebellum

The cerebellum is located at the back of the head, below the temporal and occipital lobes, and above the brainstem. Like the cerebrum, it has two hemispheres. The cerebellum communicates with the cerebral cortex and musculoskeletal structures of the body, coordinates these motor signals, and helps control and regulate voluntary muscle movements and maintain posture, gait, and equilibrium. In addition to this, it also plays a role in cognitive functions.

### **2.1.3 Brainstem**

The brainstem connects the cerebrum with the spinal cord. The brainstem consists of the midbrain, the pons, and the medulla.

1. **Midbrain:** The midbrain is a very complex structure with various neuron clusters, neural pathways, and other structures. These features facilitate various functions, from hearing and movement to calculating responses and environmental changes.
2. **Pons:** The pons is the origin point for cranial nerves. It is involved in the regulation of a range of activities such as eye and facial movement, tear production, balance, and hearing. Pons connects the midbrain and medulla.
3. **Medulla:** Medulla is located at the bottom of the brainstem and is connected to the spinal cord. The medulla is essential to survival. The medulla regulates various autonomic functions, including cardiac rhythm, breathing, blood pressure, and reflexive activities such as sneezing, vomiting, and swallowing.

### **2.1.4 Subcortical Structures**

The subcortical structures include the deep gray and white matter structures (such as the corpus callosum, hippocampus, amygdala, thalamus, and putamen). They play a pivotal role in cognitive, affective, and social functions. They also act as a relay station of the nervous system, relaying neural impulses to different areas of the brain.

1. Hypothalamus plays a role in controlling behaviors such as hunger, sleep, and thirst. It also regulates body temperature, blood pressure, and secretion of hormones.
2. Pituitary gland controls other endocrine glands in the body. It secretes hormones that control sexual development, promote bone and muscle growth, and respond to stress.
3. Pineal gland helps regulate the body's internal clock and circadian rhythms.
4. Thalamus serves as an information hub for almost all neural impulses that come and goes to the cortex. It plays a role in pain sensation, attention, alertness, and memory.
5. Basal ganglia includes the caudate, putamen and globus pallidus. These nuclei work with the cerebellum to coordinate fine motions.
6. Limbic system includes the amygdala, cingulate gyri, and hippocampus. It is the center of our emotions, learning, and memory. The amygdala plays a role in emotional reactions and the hippocampus in memory.

Each of the subcortical structures undergoes significant changes through childhood, some persisting well into adolescence. The structural and functional abnormalities of the subcortical regions have been associated with various neurological disorders[39], such as schizophrenia, depression, and autism spectrum disorders [27, 5, 64].

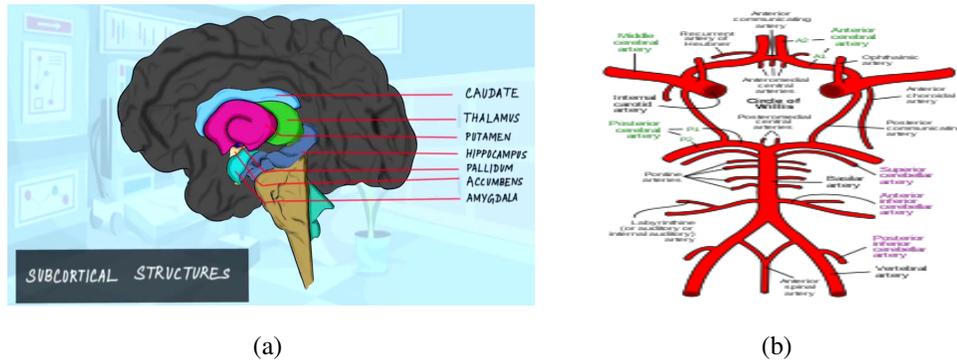


Figure 2.5: (a) Subcortical structures of brain (b) Blood supply to the brain (Image courtesy: Wikipedia)

### 2.1.5 Blood vessels

Blood is supplied to the brain by two arteries, the internal carotid arteries, and the vertebral arteries. The two arteries connect and communicate with each other at the base of the brain, called the Circle of Willis. The internal carotid arteries supply most of the cerebrum. The vertebral arteries supply the cerebellum, brainstem, and underside of the cerebrum.

The venous circulation of the brain is very different from that of the rest of the body. Usually, arteries and veins run together. However, this is not the case in the brain. The major vein collectors are integrated into the dura(outermost layer of the meninges) to form venous sinuses. The venous sinuses collect the blood from the brain and pass it to the internal jugular veins.

### 2.1.6 White Matter Fibers

White matter fibers are essential in connecting various areas of the brain with each other. There are three types of white matter tracts:

1. Projection fibers connect the cortex to other areas such as the brainstem, cerebellum, and spine.
2. Association fibers connect to different areas within the hemisphere
3. Commissural connects cortical areas of each hemisphere

## 2.2 Neuroimaging

Scientific evidence indicates that the brain is highly malleable. The brain adapts its structural and functional organization throughout life based on experience, environment, and age [29]. Neuroimaging offers vital insight into details of brain structure, functioning of the brain, and molecular processes through non-invasive in-vivo techniques. Brain imaging techniques can be classified into two approaches, namely structural imaging and functional imaging.

### 2.2.1 Structural NeuroImaging

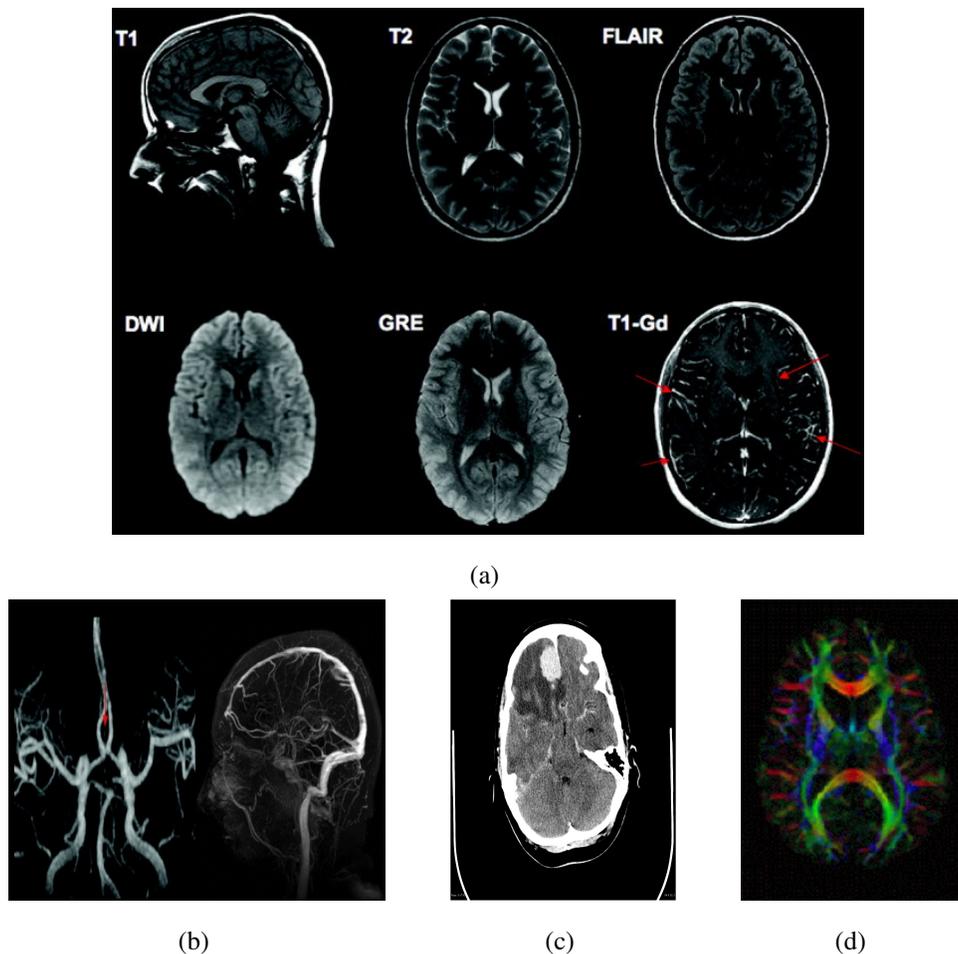


Figure 2.6: (a) Different MRI protocols (b) MRA and MR veinogram (Reproduced with permission from Mikulis, D.J. and Roberts, T.P. (2007), Neuro MR: Protocols. J. Magn. Reson. Imaging, 26: 838-847.) (c) CT (Case courtesy of Assoc Prof Frank Gaillard, Radiopaedia.org, rID: 14109) (d) DTI (Image courtesy: Wikipedia)

Structural Neuroimaging refers to imaging techniques specialized for visualization and analysis of the brain’s anatomical properties. Structural neuroimaging is vital in detecting abnormalities and damage in the brain, as well as in analyzing geometric properties such as the size and volume of structures and cortical thickness. Primary Modalities for structural imaging are Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). Visualization of bone structure and gross brain abnormalities can be obtained through CT, while through MRI, soft tissues of the brain such as white matter and gray matter can be obtained in a higher resolution, which is helpful in visualizing anatomy, abnormalities such as tumors, and mapping white matter tractography. CT uses x-rays to obtain multiple cross-sectional images of the brain. MRI can provide detailed and high-resolution images due to strong magnetic gradients; consequent responses to these gradients differ depending on tissue types and structures. MRI has different modes such as T1-weighted, T2-weighted, fluid attenuation inversion recovery (FLAIR), Proton density (PD), diffusion tensor imaging (DTI), and diffusion-weighted imaging (DWI). DTI provides a detailed reconstruction of white matter connections by tracking the movement of water molecules within the brain, which reflect the underlying white matter tracts. MR angiography (MRA), a type of MRI, is used to visualize blood vessels and identify abnormalities. Angiography can be performed through X-rays and CT as well. However, the contrast which may be used in MRA is less likely to cause an allergic reaction.

### 2.2.2 Functional NeuroImaging

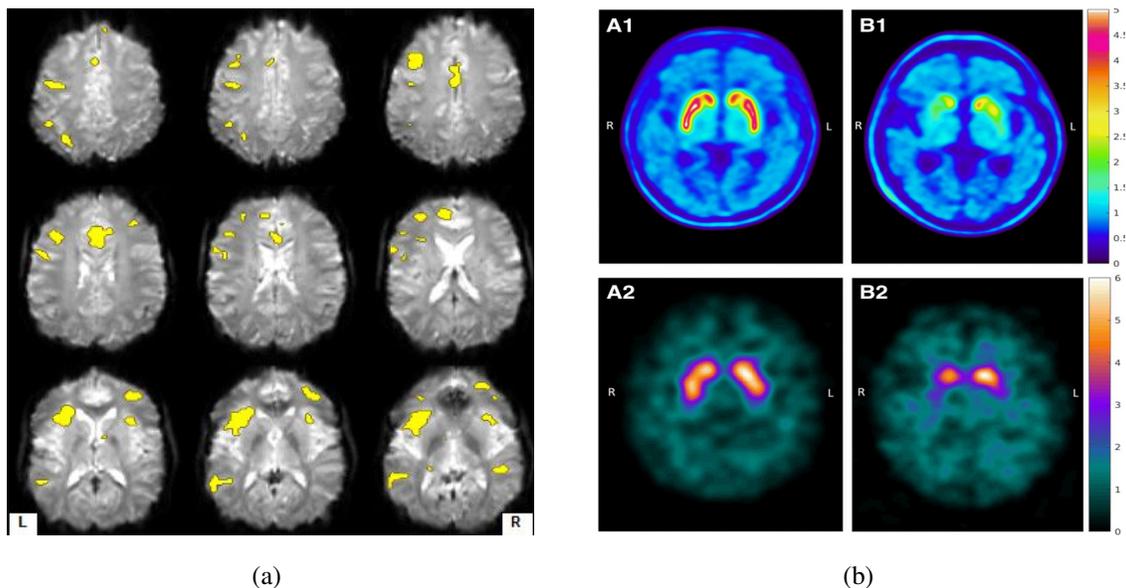


Figure 2.7: (a) fMRI during language task (Case courtesy of Dr Alexandra Stanislavsky, Radiopaedia.org, rID: 12534) (b) PET and SPECT scans of the healthy and PS patient (Adapted from (Jakobson Mo et al., 2018) under the terms of the Creative Commons Attribution 4.0 International License)

Functional neuroimaging is used to localize brain areas and underlying brain functions that are associated with performing a particular cognitive or behavioral activity, such as movement, vision, and sense. Depending on the type of signal being analyzed, inferences between the location of brain activity and brain function can then be determined. Major modalities used for functional neuroimaging are Functional Magnetic Resonance Imaging (fMRI), Positron emission tomography (PET), Single Photon Emission Computed Tomography (SPECT), Electroencephalogram (EEG), and Evoked Related Potentials (ERP).

fMRI tracks the changes in blood flow and oxygen level in the brain associated with performing cognitive or behavioral tasks [50]. An increase in metabolic demand of areas of the brain indicates activity while performing a task, called blood oxygen level dependent (BOLD) response [48], resulting changes in the signal can be detected by the scanner. PET, SPECT uses a radioactive tracer to measure the energy consumed by the brain. EEG directly measures the underlying electrical activity of the brain using electrodes placed on the scalp.

### 2.2.3 Brain Atlas

A brain atlas is an anatomical representation of the brain showing group-wise or study population global or regional brain features [14] and provide a standard coordinate system. Atlases are derived by statistically summarizing, e.g., averaging, voxel-wise, regional, or global brain MRI measures from several individuals. Structural magnetic resonance imaging (MRI) brain atlases are important tools in research and clinical practices, such as registration and segmentation of the brain, statistical mapping, identifying structural changes, and surgical planning. Structural atlases delineate the brain into different regions (coarse (GM, WM, CSF) or fine-grained (180 regions)) through a process called brain parcellation [12].

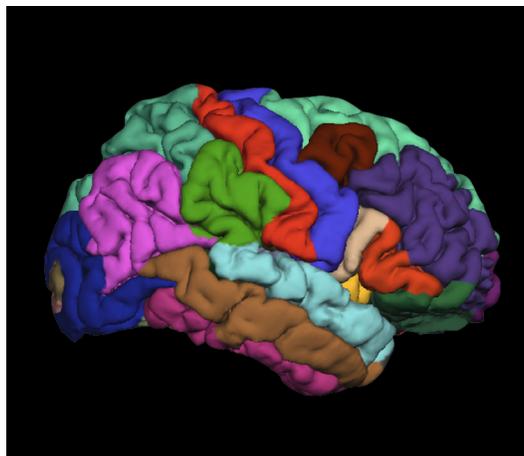


Figure 2.8: Freesurfer (Desikan-Killiany) atlas mapped to pial surface(parcellation) [12]

## Chapter 3

# System Design

### 3.1 Overview

A typical lesson in an anatomy class involves an introduction to a functional or structural system or organ, including a description of various sub-parts of the system. The instructor would provide an explanation of events or sub-structures, their functioning, their role, and spatial context through various anatomical views. We propose a solution with the same format as a typical classroom lecture which augments the syllabus by leveraging the latest technology available. This 3D interactive visualization solution for anatomy teaching is developed considering feasibility and cost-efficiency for an audience of typical medical college students in an anatomy class. A polarized stereoscopic 3D system is used for 3D projection to address and engage the large audience of a classroom. This solution is designed to provide an interactive and immersive environment for students during the teaching session to increase their understanding. We use radiological data from real subjects to create an informative, realistic, and engaging experience for the students. The design calls for the integration of radiology and anatomy, and it involves transforming raw images into realistic 3D models of human anatomy with instructor-guided lesson plans. In order to achieve this, various processes and modules are involved. The pipeline of the system is shown in Fig. 3.1.

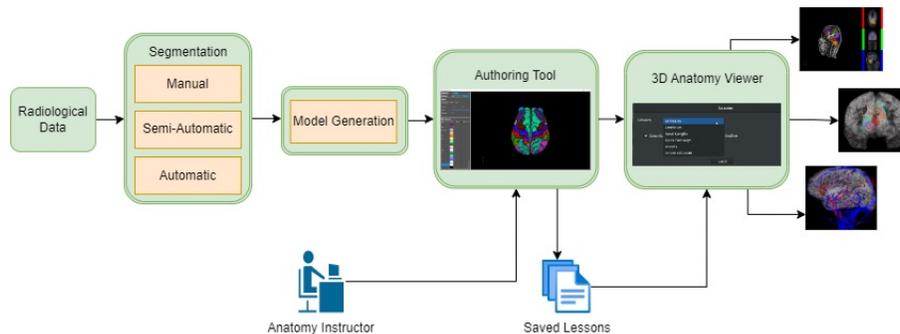


Figure 3.1: Pipeline of 3D Interactive solution for anatomy education

The first step in this system involves procuring radiological data, which is followed by the annotation of raw data for anatomical structures using manual or automatic segmentation. Using modeling techniques, these annotations are then used to create 3D models of anatomical structures. An authoring tool is developed to create and customize lesson plans for different anatomical structures. Instructors can customize the lessons based on their preferences to match their individual teaching styles. The 3D anatomy viewer allows for accessing various lessons created by anatomy instructors for each anatomical structure or system. The 3D viewer visualizes the concepts envisioned by the instructor in 3D, providing an immersive learning experience.

## 3.2 Hardware Design

### 3.2.1 3D stereoscopic projection

3D stereoscopic projection methods are ideal in a classroom environment, where the instructor can control and interact with the 3D content while simultaneously providing an explanation. This ease of use and their economic nature leads us to explore different 3D projection methods. The most popular methods are:

1. **Active shutter 3D:** In this system, the display alternates between left and right eye views, while the glasses (liquid crystal shutter glasses) alternately darken over each eye based on the timing signal in synchronization with the frame rate of the display. The above process achieves the 3D effect. This system offers better quality than the others, but the downside is that there could be noticeable flicker. However, flicker modern glasses eliminate this problem, but liquid crystal shutter glasses are considerably expensive. In addition, the cost of such a system rises with the number of students, making this scenario impractical.
2. **Anaglyph 3D:** In this system, a differently filtered colored scene, which is offset for each eye, is superimposed on the screen to produce the depth effect. The glasses have colored lenses that are chromatically opposite (red and cyan). While this is a low-cost setup, the quality is degraded, and there is significant color distortion due to the colored lenses, making the 3D experience dull.
3. **Polarized 3D:** In this system, the left and right view is overlaid onto the screen through different polarized filters. The glasses consist of lenses with different polarizing filters, which allow only similarly polarized light to pass through, ensuring that each eye sees the corresponding view. This setup is widely used in theatres to view 3D movies. The cost of this system is relatively low, and it scales up well to classrooms with varied audiences to a certain extent, as the glasses are relatively cheaper.

### 3.2.2 Dual Projector Setup

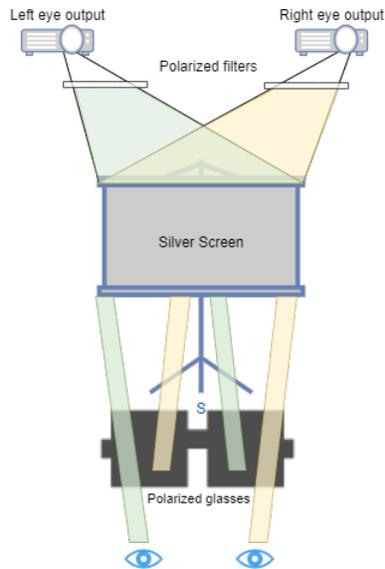


Figure 3.2: 3D Dual Projector Setup

The 3D polarized setup, as discussed earlier, is suitable for its cost efficiency and feasibility. For polarized 3D, a single projector system such as RealD (used in theaters) can be considered, but it increases the setup cost considerably. Consequently, a dual projector system is used to achieve 3D.

The projectors display views of the left and right eyes, each through different polarized filters, which are placed at a distance from the projector's lens. The glasses have corresponding polarized filters to allow only the view intended for that particular eye. There are two types of polarizing filters, circular and linear. Linear polarizers consist of a pair of orthogonal filters with the projectors and glasses having the same orientation, which leads to blocking light of different orientations. Circular polarizers have a pair of filters that have opposite handedness. The right-handed polarizer blocks the light, which is left circularly polarized, and vice versa. Since a circular polarizer allows for tilting of the glasses, it is the better alternative. Both left and right views are superimposed when projected on a screen. However, light reflected from a regular surface or a screen loses some of its polarization. To overcome this problem silver lenticular screen is used. Actual silver is embedded in the screen, making it a highly reflective surface. Refer to Appendix A for the images of our hardware setup.

## **3.3 Software Design**

### **3.3.1 Segmentation**

The annotations of anatomical structures need to be meticulous and are essential to integrate radiology and anatomy. Segmentation can be manual, automatic, or semi-automatic. However, manual segmentation methods are tedious and time-consuming; other conventional automatic segmentation methods, such as Freesurfer's FSL and BrainSuite, take days to provide segmentation, are very specific to the brain, and can work only on certain modalities. Deep learning methods overcome the drawbacks of both these approaches, delivering annotations within seconds once trained and their performance being on par, if not surpassing, other traditional approaches. However, the availability of data for training a deep learning model is challenging, and extensive exploration of methods for each task is necessary. Hence the combination of manual, semi-automatic, conventional automatic, and deep learning based segmentation methods are used to create annotations.

### **3.3.2 Model Generation**

Structure annotations can be transformed into 3D models using various techniques. We primarily use surface modeling, an advanced modeling technique adept in handling complex shapes of anatomical structures. It provides a seamless and smooth surface, allowing for an authentic three-dimensional visual representation of anatomical structures. In addition, volume rendering is also used to represent the volumetric data.

#### **3.3.2.1 3D Mesh**

There are many modeling techniques to obtain surface representation, mesh modeling being one of the most commonly used methods. Rendering of a complex object is simplified by using a polygonal mesh which essentially defines the shape of the complex object. A mesh consists of vertices, faces, and edges.

1. vertex is a single point
2. Edge connects two vertices with a straight line segment
3. Face is a closed flat surface enclosed by edges (polygon). Faces can consist of triangles (3 sides), quadrilaterals (4 sides), or other simple polygons.

The 3D model/mesh is generated using ITKsnap or VTK library, which uses the marching cubes algorithm. It is a simple iterative algorithm that generates 3D surface meshes. The algorithm can be described as follows: 1. Divide the space within the bounds into an arbitrary number of cubes. 2. Test the corners of every cube for whether they are inside the object. 3. For every cube where some corners are inside and some corners are outside the object, the surface must pass through that cube, intersecting

the edges of the cube in between corners of opposite classification. 4. Draw a surface within each cube connecting these intersections. The 3D meshes are saved as STL files (other formats like VTK are also supported). We also use Blender<sup>1</sup> to post-process a few imprecise 3D meshes.

### 3.3.2.2 Volume Rendering

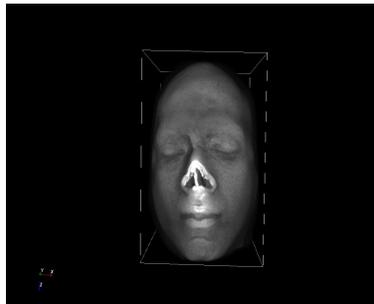


Figure 3.3: Volume Rendering of MR data

Volume rendering displays a 2D projection of the 3D MRI data. This is achieved by directly loading raw data without any processing. Every voxel value is mapped to opacity and color using the transfer function, a piece-wise linear function. Volume ray casting is the technique used to achieve this rendering. This technique uses a simple camera model to generate the ray for each pixel, with the output 2D image lying in between the camera and the volume to be rendered. Volume rendering of the brain using MRI data can be seen in Fig. 6.8(b)

### 3.3.3 Authoring Tool

We want to establish a comparable structured narrative approach to learning by following the curriculum of standard anatomy taught at Indian medical colleges. Designing lessons to match the sequence of the narrative becomes crucial. With the help of the Authoring tool we have developed, the educators have the flexibility to design 3D-based lessons to match their teaching style and the needs of the curriculum. Their expertise makes content meaningful and comprehensible to the learners. The authoring tool helps the user decide the sequence of structures to fit the narrative and manipulate the properties of 3D models of anatomical structures. It has text and other graphical tools to assist in illustrating the context and provide descriptions of particular structures. Through our solution, we place a specific emphasis on instructor-led anatomy teaching. The features of the authoring tool and the lesson creation workflow are discussed further in detail in ch. 6.

---

<sup>1</sup><https://www.blender.org/>

### **3.3.4 Lesson Design Considerations**

#### **3.3.4.1 Anatomical Relationship Grouping**

The human body consists of numerous organs, and each organ is a complex system of smaller structures, such as the brain, which, as we've seen is a complex structure with numerous cortical, and subcortical structures, the blood supply system, and various functional areas which support auditory, and visual functions. In order to understand the working of a system, we divide it into sub-components. This logical division includes spatially and functionally related structures and represents them as groups. While teaching about one structure or an entire system, these relationship groups are also illustrated to provide the necessary context.

#### **3.3.4.2 Integration of anatomy and radiology**

We generate 3D representations based on clinical data. To bring about the seamless integration of anatomy and radiology, we represent radiological data in the 3D anatomical domain. Radiologists generally go through clinical data such as MRIs and CTs by examining 2D slices. We propose to localize these 2D slices or cross sections in 3D space along with the 3D representation of structures to merge radiology and anatomy, helping with cognitive mapping.

#### **3.3.4.3 Visual Properties of Anatomical Structures**

It is difficult to distinguish between structures in a cadaver and radiological scans like MRI and CT for an untrained eye. To train students to discern minute differences between tissues and organs, it is beneficial to provide them with the whole picture so that they can be trained to recognize patterns. So we represent structures with distinguishing visual properties to differentiate between them. Instructors are given the choice of various visual properties of every structure, such as color, opacity, smoothness, texture, etc based on their preferences. Each of these properties can be manipulated independently of each structure.

#### **3.3.4.4 Structure Labeling**

To learn anatomy, it becomes important to identify and relate all small and large structures belonging to a particular anatomical region. Teaching this becomes more convenient and effective by organizing the labels of all structures in a hierarchical manner and presenting them in a logical sequence. It is essential that the display of labels is consistent over the examination of the structure from any angle and that the spacing between the labels is adequate for clarity and understandability. Keeping this in mind, labels are organized into a neat, evenly spaced circular grid in the 3D scene. The labels are designed to be capable of adapting to any view. The techniques used to achieve these key elements are discussed in further chapters.

### **3.3.5 3D Anatomy Viewer**

The 3D anatomy viewer is an interactive visualization tool that presents the 3D lessons created using the authoring tool. It incorporates considerations of lesson design. The viewer represents the visual and structural properties of the 3D structures saved while designing the lesson. It encompasses organ systems' anatomical and functional relationships and integrates radiological data with the 3D structures. The lessons are designed to provide clarity and understanding to students. The interactive 3D interface is designed to provide an immersive experience for a large number of students. The features of the anatomy viewer are discussed further in detail in ch. 6.

## *Chapter 4*

### **Data and Annotation**

The human brain, a complex organ, is chosen as the anatomical structure to be taught. We aim to explore all aspects of neuroanatomy and help students thoroughly examine and study using 3D techniques. In order to do so, we need to obtain 3D representations of brain structures. The first step is to collect 3D volumetric data using various neuroimaging modalities, followed by annotating various structures of the brain on the data available through segmentation approaches and using segmentation data to generate 3D models using modeling techniques.

#### **4.1 Data**

With various neuroimaging modalities available such as T1, FLAIR MRI, DTI, and MRA, various aspects of the brain, such as cortical structures, deep structures, white matter fibers, and blood vessels, are represented in 3D volumes. They also allow for obtaining clinical case data to illustrate the reality rather than the fabricated 3D brain. Naturally, using such data to create 3D representations is a pragmatic approach.

Brain data of a volunteer from various modalities such as MRI, MRA, DTI, and MRV was obtained by SCTIMST, Kerala. The images were acquired in DICOM format using Siemens Magnetom Avanto Tim fit 1.5T scanner and are as follows :

1. Sagittal 3D SPGR , Dimensions: 512 x 512 x 228 , Voxel Spacing : 0.5 x .5x 0.8 mm
2. Sagittal CUBE FLAIR, Dimensions :512 x 512 x 320 , Voxel Spacing : 0.48 x 0.48 x 0.59 mm
3. Axial FLAIR , Dimensions: 512 x 512 x 256, Voxel Spacing : 0.41 x 0.41 x 0.98 mm
4. Axial 3D TOF , Dimensions : 512 x 512 x 196 , Voxel Spacing : 0.42 x 0.42 x 0.6 mm
5. Venogram , Dimensions : 512 x 512 x 480 , Voxel Spacing : 0.46 x 0.46 x 0.7 mm
6. Coronal FLAIR , Dimensions : 512 x 512 x 260 , Voxel Spacing : 0.48 x 0.48 x 0.98 mm

7. Axial DTI(64 Directions) , Dimensions : 256 x 256 x 3757 , Voxel Spacing : 1 x 1 x 2.49 mm

Among the high-resolution data acquired, Axial 3D TOF, Veinogram, Sagittal 3D SPGR, and Axial DTI are used for further processing. The volumes are acquired simultaneously, so they are already registered. They are resampled to 1x1x1 mm voxel spacing and converted to NIFTI format.

## 4.2 Segmentation

Segmentation involves the extraction of regions of interest (ROIs) from the 3D image data we obtained earlier, such as from Magnetic Resonance Imaging (MRI) or Magnetic Resonance Angiography (MRA) scans. The segmentation can be done manually, semi-automatically, and automatically. It is a crucial step in generating 3D models of various anatomical structures. The ROIs, in our case, are the cortical and sub-cortical structures that need to be identified and annotated.



Figure 4.1: (a) Manual segmentation of MRI data on ITKSnap (b) Semi-automatic segmentation of blood vessels on CTA data

Depending on the anatomical area and imaging modality, numerous methodologies, toolboxes, and algorithms are available for each kind of segmentation approach. Manual annotation is more accurate, but the process is too tedious, while automatic segmentation methods are fast and effortless but are not always feasible. It is critical to select a feasible and moderately accurate segmentation method so that subsequent phases in the pipeline work effectively. In the case of the human brain, many automatic segmentation methods are available for cortical and subcortical structure segmentation, which produce reasonable results consistently. However, we employed manual segmentation when automatic segmentation appeared insufficient for certain structures with complex geometries and a few other structures

that were not available via the automatic segmentation technique we used. Semi-automatic methods were also used for certain structures when manual segmentation was extremely tedious and the automatic segmentation methods infeasible.

### 4.2.1 Automatic Segmentation

Automatic segmentation is the go-to method for obtaining a large number of brain structures in one go. There are atlas-based automatic segmentation methods as well as deep learning based methods available for extracting cortical and sub-cortical structures of the brain. Our purpose is educational, where the lessons created using 3D models would be configured and saved well before the classroom presentation. Since there is no time constraint and moderate accuracy is adequate, unlike for clinical diagnoses, we use atlas-based segmentation methods, which are deemed fairly reliable in the case of segmentation of major structures within the brain. The T1 MRI volume is used as its contrast and brightness is appropriate to extract structures of the brain. Two software tools are utilized for automatic segmentation

FreeSurfer, a popular and effective open source software tool <sup>1</sup> was used for analysis and segmentation of neuroimaging data to obtain whole brain segmentation, including the most macroscopic structures as well as get the surface representation of cerebral cortex. The whole brain segmentation method is based on a probabilistic atlas. Segmentation of 40 structures per hemisphere was obtained, including the subcortical structures. In addition, segmentation of hippocampal subfields was also acquired. Freesurfer version 6.0 was used for this process. All the freesurfer algorithms work using commands in the bash shell prompt or bash script. Appendix B contains additional information about Freesurfer.

While freesurfer is known for segmentation, an accurate cerebral cortex model was required. BrainSuite<sup>2</sup> [68], a suite of open-source software tools for analyzing magnetic resonance imaging (MRI) of the human brain, was used for this purpose. One of the primary functions of BrainSuite is to extract and parameterize the cerebral cortex's inner and outer surfaces. Hence we used it to extract the Cerebral Cortex (left and right). Appendix B contains additional information about Brainsuite.

### 4.2.2 Manual Segmentation

Some structures not present in the segmentation obtained using Freesurfer's FSL toolbox were segmented manually from the MRI. The structures whose segmentation was erroneous were either manually segmented from scratch or corrected later using the software. The manual annotation from T1 MRI images was done using a free open source software ITK-SNAP<sup>3</sup> [80] as shown in Fig. 4.1. The structures that were manually segmented are the Optic globe, Optic tract, optic chiasm, Optic nerve, Mammillary

---

<sup>1</sup><http://surfer.nmr.mgh.harvard.edu/>

<sup>2</sup><http://brainsuite.org/>

<sup>3</sup><http://www.itksnap.org/>

Bodies, Interventricular foramen, Third ventricle-supra optic recess, pineal recess, pituitary stalk, pituitary gland, thalamus, Aqueduct of Sylvius, Fourth Ventricle- foramen of Magendi, foramen of luschka, pineal gland, Lateral ventricles, corpus Callosum. The lobes of the brain, i.e., frontal, parietal, occipital, and temporal lobes, were also segmented manually.

### 4.2.3 Semi-Automatic Segmentation

Brain vasculature was extracted from MR Angiography data.MRA data highlights major blood vessels. Veins were separately extracted from an MR Venogram. The annotation for arteries and veins was obtained using a semi-automatic active contour segmentation method on ITKSnap. The threshold was adjusted by visual analysis of the data and seed points were selected for precise vessel detection as shown in Fig. 4.1

### 4.2.4 White Matter Tractography

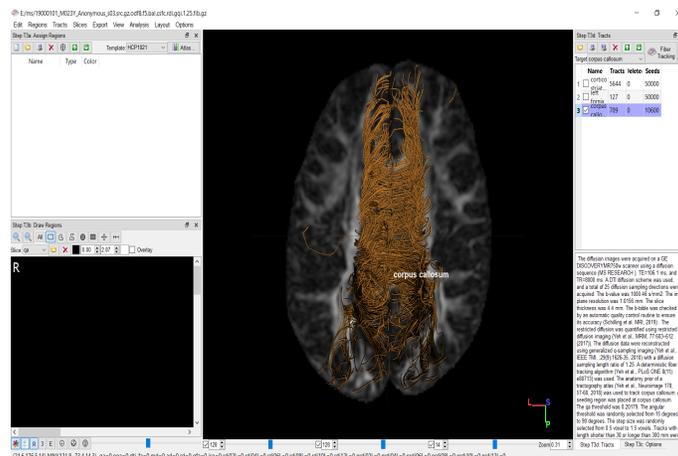


Figure 4.2: Fiber tracking on DTI data in DSI Studio

White matter tracts illustrate the various connections within the brain. While T2W and FLAIR MRIs contain some white matter information, DTI presents intricate details of the white matter fibers, which are essential clinically in analyzing neuropsychological disorders. We use DSI studio <sup>4</sup>, a tractography software to extract white matter fibers as shown in Fig. 4.2. Using the DSI Studio, we employed both manual fiber tracking using a priori anatomical knowledge of fibers and of complex automated fiber tracking, which makes use of normalized anatomical atlases to generate tracts from DTI data. Some of the tracts generated include corpus Callosum and tapetal fibers, corticospinal tracts, optic radiation including Meyer’s loop, arcuate fasciculus, fornix, uncinate fasciculus, and superior longitudinal fasciculus.

<sup>4</sup><https://dsi-studio.labsolver.org/>

## Chapter 5

### **Automatic Segmentation: A Method to remove size bias**

#### **5.1 Introduction**

Subcortical brain structures include deep gray matter structures such as the putamen, thalamus, and hippocampus. They play a pivotal role in activities such as memory, emotion, and hormone production. Structural and functional abnormalities in subcortical regions are associated with degenerative neurological disorders, including schizophrenia, depression, and Alzheimer’s disease[39]. Segmenting subcortical brain structures from Magnetic Resonance Images (MRI) is the first step to diagnosing such disorders. Manual delineation of the structures from MRIs is still used in practice to get accurate segmentation for this purpose. Getting expert marking for the structures is tedious, time-consuming, and inter-rater variability also comes into play. Hence the need for automatic segmentation methods to accurately segment the brain structures.

Segmentation is already a challenging task because of brain tissue ambiguity and data scarcity. In addition, a performance bias can be observed in the segmentation results for bigger structures in many of the proposed solutions, making subcortical segmentation doubly challenging since subcortical structures vary widely in size. The possible factors behind this size bias include i) inconsistency in annotations of large structures vs. small structures across subjects, as the former are easier to annotate compared to the latter, ii) insufficient image resolution leading to more inaccurate boundaries for smaller structures, and iii) learning both large and small structures from many (for the former) versus few (for the latter) voxels is inefficient. While the first two issues cannot be resolved easily, the last one can be resolved by taking a different approach. Towards this end, the proposed method addresses this issue by making the network learn the structures from its local rather than global context, such as the whole brain image. This should aid the network in focusing on each structure separately irrespective of their size, giving an edge to smaller structures. Additionally, a pre-training step is proposed to learn tissue types to aid in discriminating between tissue types, as the local context information alone can be inadequate for efficient segmentation.

## 5.2 Related Work

Sub-cortical structure segmentation is a widely exploited area by both computer vision and medical image processing researchers. Few works are discussed here. Initial attempts were based on label transfer [16, 79, 19]. Expectation Maximization of atlas based prior was used for sub-cortical structure segmentation in [54]. Machine learning methods like SVM [32, 2], graph cut [78] were also used for this purpose. Deformable models also became popular for sub-cortical structure segmentation [46, 25]. F-CNN based segmentation approach [67] was also tried for subcortical structure segmentation.

U-net based segmentation [58] built a new benchmark for the segmentation performance. U-net architecture became popular in medical imaging mainly because of its good performance with limited data in most cases. Many U-net variants were also introduced for this purpose such as [28, 82, 47, 7] with reasonable improvements. A label refinement strategy with two concatenated U-nets was used in [7] and [28] introduced competitive dense blocks and multi-slice information aggregation.[51] proposed M-Net, a memory efficient method to represent 3D context information of a 2D slice.  $\psi$ -net [47] used a densely convolutional LSTM module for selecting and enhancing features.

Although attention mechanisms have been studied widely for various computer vision tasks such as classification, detection, segmentation, image captioning, and visual question answering, they have gained popularity for medical image segmentation tasks only recently. Wang et al.[77] proposed the Residual Attention Network, which uses self-attention to compute 3d attention maps. Hu et al.[30] proposed a compact module squeeze and excitation, which is used to compute channel-wise attention. Roy et al.[61] proposes the generation of separate spatial and channel attention maps, which leverages both information while being computationally less expensive. Zhang et al. [81] introduce split attention blocks, which apply channel wise attention on different network branches. These variants of attention were used for subcortical segmentation in [76, 44]. Wang et al. [76] used squeeze and excitation block, which emphasizes essential features. Lee et al. [44] proposed Split-Attention U-Net (SAU-Net), along with pre-training and fine tuning.

## 5.3 Background

3D U-Net [6] based on the U-Net architecture [59] is utilized. It consists of an encoder and decoder stage with 3D volume as input, each block using 3D operations like 3D convolution, 3D max pooling, and 3D transpose convolution and skip connection between encoder and decoder block. The encoder captures the context of the input volume, and it contracts at each stage. Decoder enables precise localization, and it expands at each stage. Encoder and decoder in U-Net are symmetrical and similar. The introduction of skip connections leads to the recovery of fine-grained details for prediction. They are used to pass features from the encoder path to the decoder path in order to recover spatial information lost during down-sampling. Experiments are performed on variants of 3D U-Net like introducing Residual Block[43] and Dense block[31] at each encoder-decoder level. Residual connections in the

residual block help with the vanishing gradient problem. The dense block uses skip connections by concatenating previous feature maps, which leads to extreme feature reusability. The implementation, in each encoder level, consists of 2 layers, each layer containing a 3D convolution layer, followed by an instance normalization layer and leaky relu.

### 5.3.1 Normalization

During training, the distribution of each layer’s input changes due to the change in parameters of previous layers. This internal covariate shift is addressed by the normalization of the input of each layer. Generally used normalization methods are batch normalization [33] and instance normalization [74]. In the batch normalization method, the normalization step involves computation of the mean and variance of the feature map along the (N, H, W) dimension, while in the Instance normalization method, the mean and variance along (H, W) are computed. Normalization methods give us the advantage of using higher learning rates while also acting as regularizers. Instance normalization makes the network agnostic to the contrast in input volumes and works better with a smaller batch.

### 5.3.2 Activation Layer

After each convolution layer, the activation function is used to determine the activation of that neuron. The common activation functions used are rectified linear unit (ReLU) [52, 71] and leaky rectified linear unit (LeakyReLU) [49]. Nonsaturated activations like ReLU solves the problem of vanishing/exploding gradient and improves convergence. It is also believed the reason for the superior performance of ReLU is its sparsity. Leaky ReLU fixes the dying ReLU problem and speeds up the training. Mathematically ReLU can be represented as,

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (5.1)$$

Leaky ReLU can be represented as,

$$f(x) = \begin{cases} x, & x \geq 0 \\ x/a, & x < 0 \end{cases} \quad (5.2)$$

where  $a$  is a fixed parameter in range  $(1, \infty)$ .

### 5.3.3 Residual Block

With deeper networks, it is observed that accuracy saturates at a certain point and eventually degrades. During backpropagation, gradients are multiplied together. The gradient becomes vanishingly small as we approach shallower layers of a deep network, and learning stops for those layers. In the Residual block, instead of stacking a few layers directly to fit desired underlying mapping, a residual

connection is added to layers 2-3 hops away to fit residual mapping. Backpropagation through this identity function preserves the gradient as well. It is also easier to optimize residual mapping. Formally,

$$\mathcal{F}(x) = \mathcal{H}(x) - x \quad (5.3)$$

where  $\mathcal{H}(x)$  is the desired underlying mapping,  $\mathcal{F}(x)$  is the mapping that stacked layers fit, with the original mapping recast as  $\mathcal{F}(x) + x$ .

### 5.3.4 Dense Block

It is desirable to pass low-level information between input and output directly. In Dense block, maximum information flow between layers is achieved by direct connection of feature maps from one layer to all subsequent layers via concatenation, unlike summation in residual blocks. This leads to high feature reusability, boosting of feature propagation, mitigating the vanishing gradient problem, and reduction in the number of parameters.

### 5.3.5 Channel Attention (Squeeze and Excitation block)

In deep convolutional neural networks, channels represent different convolutional filters that learn various feature maps of input. Each of these feature maps contains information that has a different magnitude of importance. The primary motivation behind channel attention is to prioritize specific channels over others by scaling the more critical channels with higher values. Hu et al.[30] proposed a compact squeeze and excitation module, which is one of the most widely used attention mechanisms. Squeeze and excite block (SE block) consists of a squeeze module, excitation module, and scale module. Global spatial information is encapsulated in a channel descriptor by aggregating information along spatial dimensions by using pooling operations like global average pooling.  $s_c \in \mathbb{R}^C$  represents average pooled features and is calculated as:

$$s_c = A_s q(x) = \frac{1}{HxWxD} \sum_{i=1}^H \sum_{j=1}^W \sum_{k=1}^D x_c(i, j, k) \quad (5.4)$$

where  $x_c \in \mathbb{R}^{HxWxD}$  is the input feature map  $X = [x_1, x_2 \dots x_c]$ .

The output of squeeze operation is then fed into a gating mechanism with 2 fully connected (FC) layers one reduces the dimensionality with  $r$  is the reduction ratio, followed by ReLU and dimensionality increasing layer, and finally a sigmoid activation function. The output is the channel attention map  $e_c$  computed as:

$$e_c = A_e x(s) = \sigma(W_1 \delta(W_2 s_c)) \quad (5.5)$$

where  $\sigma$  denotes sigmoid function,  $\delta$  denotes ReLU function and  $W_1 \in \mathbb{R}^{C/r * C}$ ,  $W_2 \in \mathbb{R}^{C * C/r}$ .

The final output of the block,  $\tilde{X} = [\tilde{x}_1, \tilde{x}_2 \dots \tilde{x}_c]$ , is obtained by rescaling the input feature map as:

$$\tilde{x} = e_c x_c \quad (5.6)$$

## 5.4 Method

### 5.4.1 Proposed Method

The proposed method aims to focus on narrowing the disparity in performance between structures of different sizes. To this end, we propose a 2 phase segmentation framework to reduce the size bias. The two phases are:

1. Pre-training with Grey matter (GM), White matter (WM), Cerebrospinal fluid (CSF) segmentation
2. Sub-cortical segmentation with atlas-guided ROI extraction.

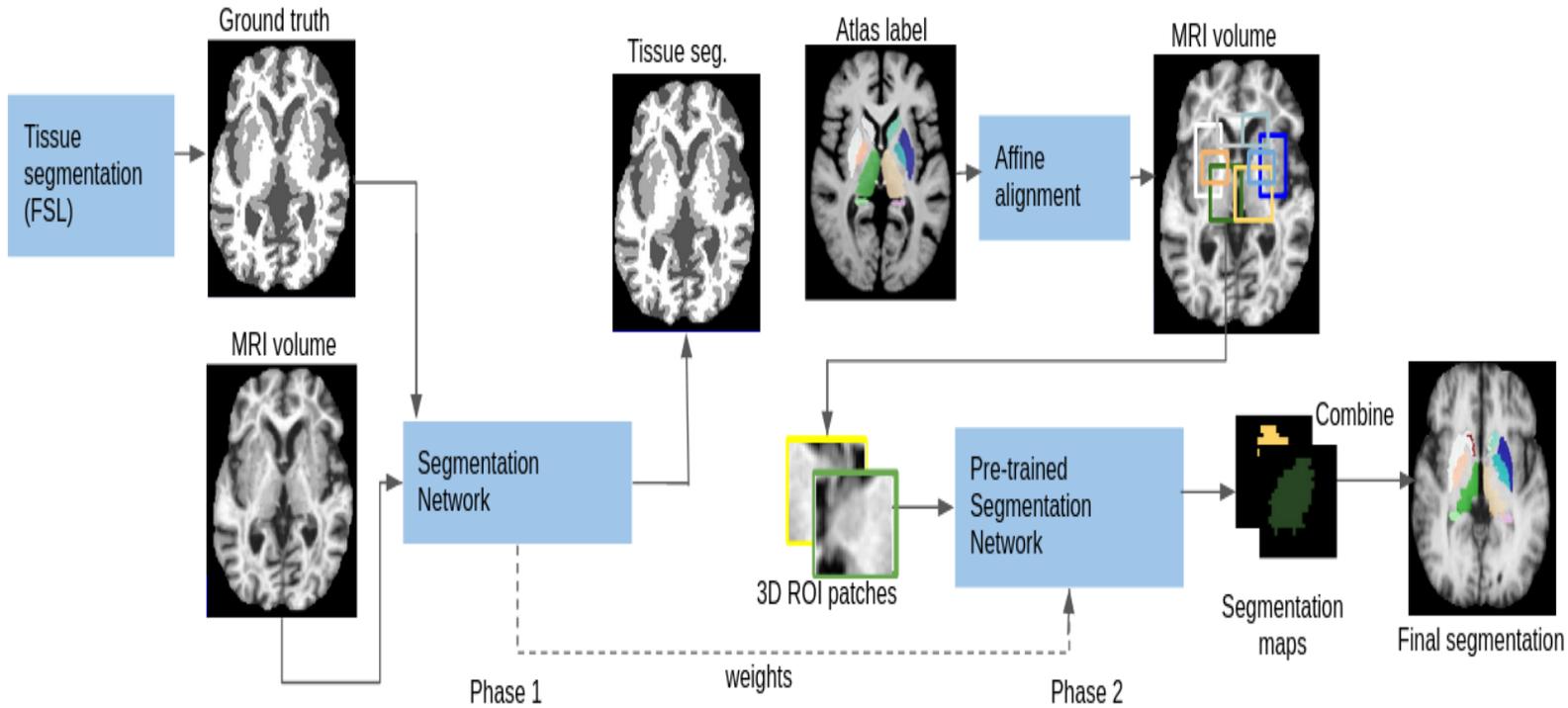


Figure 5.1: Proposed 2-phase training framework for sub-cortical structure segmentation. Phase 1: Pre-training with tissue segmentation, Phase 2: Structure segmentation from the atlas guided 3D structure ROI.

The two phases of training that have been proposed in the segmentation for any network are illustrated in Fig. 5.1.

#### 5.4.1.1 Pre-training with tissue segmentation

There are three main tissue types in the brain Grey matter (GM), White matter (WM), and Cerebrospinal fluid (CSF). Sub-cortical structures are Grey matter embedded in White matter deep inside

the brain. Subsequently, tissue discrimination becomes an important aspect of sub-cortical structure segmentation. Choosing tissue segmentation, i.e., segmentation of Grey matter, white matter, and cerebrospinal fluid as a pre-training task helps the network learn to discriminate between the primary tissue types in the brain, which in turn aids sub-cortical structure segmentation. In this pre-training phase, 3D volume is fed to the network with GM, WM, and CSF labels as ground truth labels. These ground truth labels are computed using FSL[34].

#### 5.4.1.2 Atlas-guided ROI extraction

In this phase, weights of the previous phase are loaded as initial weights, and a rough ROI patch is extracted for each of the sub-cortical structures and fed to the network with the structure labels as ground truth. The output after this phase of training is the probability map for each structure. The probability maps are then combined with maximum probability at each pixel to get the final segmentation map. The motivation behind this strategy is to counteract the variability in sizes of the structures, and it ensures that smaller structures get equal priority while training compared to methods that provide whole brain volume input. In our experiments, the rough ROIs are extracted by an affine alignment of the atlas to the MRI volume. A relaxation step (expanding ROI size by a few pixels) was done for each 3D ROI patch to compensate for imprecise alignment and to prevent under-segmentation.

#### 5.4.2 Loss function

Categorical cross Entropy is used as the loss function for phase 1. The Sum of unweighted soft dice loss and categorical cross-entropy loss is utilized for training the model in phase 2. Categorical cross entropy can be given as

$$L_{ce}(p, g) = - \sum_i^N \sum_j^C g_{i,j} * \log(p_{i,j}) \quad (5.7)$$

Where sum runs over N voxels and C classes of predicted segmentation volume  $p_{i,j} \in P$  and ground truth volume  $g_{i,j} \in G$ . Unweighted soft dice loss can be given as,

$$L_{dice}(p, g) = 1 - \sum_j^C \frac{2 \sum_i^N p_{i,j} g_{i,j} + \epsilon}{\sum_i^N p_{i,j}^2 + \sum_i^N g_{i,j}^2 + \epsilon} \quad (5.8)$$

Where sum runs over N voxels and C classes of predicted segmentation volume  $p_{i,j} \in P$  and ground truth volume  $g_{i,j} \in G$ . The  $\epsilon$  term is used here to ensure the loss function stability by avoiding the numerical issue of dividing by 0.

Final Loss for each structure ROI in phase 2 can be given as,

$$L_{ce+dice} = L_{ce} + L_{dice} \quad (5.9)$$

## 5.5 Dataset and Implementation details

### 5.5.1 Dataset description

Two publicly available datasets were used to perform experiments: IBSR [45], and MICCAI [42] datasets with 18 and 35 volumes, respectively. Both these datasets provide labels for both cortical and subcortical structures.

#### 5.5.1.1 IBSR Dataset

The International Brain Segmentation Repository (IBSR) dataset<sup>1</sup>[45] has 18 volumes of 3D T1 weighted MR images. The voxel size in IBSR is variable, namely,  $0.93 \times 0.93 \times 1.5$  or  $1 \times 1 \times 1.5$ . Manual segmentation of 32 structures is provided by the Center for Morphometric Analysis at Massachusetts General Hospital<sup>2</sup>. 18 volumes are split into six folds with three images per fold. Training is done on four folds; one fold is used for validation and the remaining fold for testing. Model training is done six such that each image is used as a testing image once. For the final results averaging of results of all folds (6x3) is done.

#### 5.5.1.2 MICCAI dataset

This dataset contains 35 volumes of 3D T1 weighted MRI images, released for use in the MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling[42]. It consists of 15 training images and 20 testing images. The original MRIs are from OASIS<sup>3</sup> project. The voxel size is uniformly  $1 \times 1 \times 1.25$ . 134 labels for each volume were provided by Neuromorphometrics, Inc<sup>4</sup>. 15 training images are split into five folds with three images per fold. Training is done on four folds, and the remaining fold is used for validation. Training all the folds gives five models, and we use it to test 20 test images. For final results averaging of results of 100 test images(5x20) is done.

### 5.5.2 Pre-processing

The input volumes are resampled to a uniform voxel size of  $1 \times 1 \times 1mm^3$  and histogram based standardization is performed for each dataset. The subcortical region is smaller in brain MRI scans. For training, following [41], center-cropped the MRI to a region of size 96x96x96 containing subcortical structures for the pre-training step.

---

<sup>1</sup><http://www.nitrc.org/projects/ibsr>

<sup>2</sup><http://www.cma.mgh.harvard.edu/>

<sup>3</sup><http://www.oasis-brains.org/>

<sup>4</sup><http://neuromorphometrics.com/>

### 5.5.3 Training details

3D U-Net based architectures with variants (Residual Blocks[43], Dense block[], SE block), with 3 encoder-decoder blocks and a bottleneck block is employed. Adam[37] optimizer is used with initial learning rate 0.0001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  and weight decay of 0.001, trained for 400 epochs. The code was implemented using PyTorch and training was done on NVIDIA GTX 2080 with 11GB RAM. A fixed atlas<sup>5</sup> [79, 3, 9, 8, 70, 17] was used for extracting ROIs in the training and testing phases which is shown in Fig. 5.2. Implementation is available at [https://github.com/mythri-venkat/subcortical\\_segmentation](https://github.com/mythri-venkat/subcortical_segmentation).

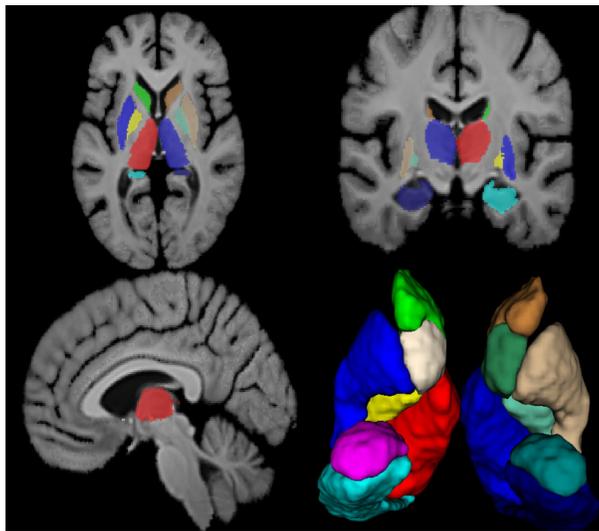


Figure 5.2: Atlas with subcortical structures annotations only

## 5.6 Experiments and Results

A 6-fold validation was done on the IBSR dataset, while a 5-fold validation was done for MICCAI, and the results were averaged across all the folds for reporting. In order to understand the size induced performance variations, analysis was done with a sizewise ordering of structures, i.e., structures are divided into 4 classes: very small, medium, large, and very large. In terms of volume, these are approximately 4%, 20%, and 50% of the average size of the largest structure in the 'very large' category (namely the thalamus)

### 5.6.1 Evaluation Metrics

Quantitative evaluation of the above segmentation methods is done by the most commonly used metric, which is Dice Similarity Coefficient(Dice) [].

<sup>5</sup><https://github.com/voxelmorph/voxelmorph>

Dice Similarity Coefficient(Dice) computes the overlap between the predicted segmentation result and the ground truth:

$$Dice(P, G) = \frac{2 \times |P \cap G|}{|P| + |G|}, \quad (5.10)$$

Where P and G are predicted segmentation and ground truth, respectively,  $|P \cap G|$  denotes the number of pixels in the overlapped region between predicted segmentation and ground truth while  $|P| + |G|$  denotes pixels of both regions. Dice value varies between 0 and 1, where the former constitutes no overlap, and the latter represents complete overlap with ground truth.

## 5.6.2 Ablation Studies

The performance of a Residual U-net with the proposed 2-phase training approach was assessed by comparing against the traditional training with a whole volume. The obtained results are presented in Table 5.1. Dice values are given for the structures at an individual and a group (sizewise) level.

In order to understand the contribution from each phase of training, ablation studies were done. The three variants considered were: Segmenting after pre-training (PT), after ROI extraction (ROI), and with the proposed 2-phase training. These were evaluated on the IBSR and MICCAI datasets, and the results are reported in Table 5.1. Performance degradation of 22% is observed in the Dice value for the very small structure group relative to the very large class group for baseline Res. U-net. The state of the art (SOTA)  $\psi$ -net also reports a 17.6% deterioration [47].

This underscores the need for developing a strategy to overcome size-specific performance variations. The Dice value for smaller structures in both datasets shows improvements with the proposed modifications in training compared to the larger ones.

Another observation from Table 5.1 is that the improvement with the proposed method is significant from both phases 1 and 2 for smaller structures compared to larger ones. Results for the MICCAI dataset appear to be more influenced by pre-training and ROI-based training compared to IBSR. This is possibly due to the difference in data quality.

In MICCAI (relative to the IBSR images), the voxel resolution in the axial direction is better; the contrast between tissues and the quality of the image is also superior. Consequently, the proposed method appears to boost segmentation performance in general with the quantum of boost being affected by the data quality/resolution which is a fundamental issue for segmentation.

	IBSR Dataset					MICCAI Dataset				
	Dice (Mean $\pm$ STD)		Av. Improvement(%)			Dice (Mean $\pm$ STD)		Av. Improvement(%)		
Left & Right	Res. U-net (baseline)	Proposed	PT& ROI	PT	ROI	Res. U-net (baseline)	Proposed	PT& ROI	PT	ROI
<b>Very Small</b>										
R.Accumbens	0.692 $\pm$ 0.045	0.734 $\pm$ 0.038	5.55	2.1	3.3	0.753 $\pm$ 0.094	0.784 $\pm$ 0.048	4.8	1.8	4.05
L.Accumbens	0.701 $\pm$ 0.052	0.736 $\pm$ 0.039				0.750 $\pm$ 0.074	0.781 $\pm$ 0.040			
<b>Medium</b>										
R.Amygdala	0.746 $\pm$ 0.042	0.768 $\pm$ 0.048	3.7	1.325	3.4	0.771 $\pm$ 0.050	0.823 $\pm$ 0.021	4.35	3.5	4.225
L.Amygdala	0.809 $\pm$ 0.066	0.763 $\pm$ 0.047				0.781 $\pm$ 0.042	0.816 $\pm$ 0.028			
R.Pallidum	0.803 $\pm$ 0.027	0.828 $\pm$ 0.021				0.840 $\pm$ 0.085	0.875 $\pm$ 0.025			
L.Pallidum	0.804 $\pm$ 0.022	0.834 $\pm$ 0.020				0.851 $\pm$ 0.067	0.876 $\pm$ 0.024			
<b>Large</b>										
R.Caudate	0.867 $\pm$ 0.016	0.875 $\pm$ 0.019	1.3	0.533	0.617	0.871 $\pm$ 0.09	0.877 $\pm$ 0.055	1.56	0.5	1.45
L.Caudate	0.868 $\pm$ 0.021	0.876 $\pm$ 0.02				0.861 $\pm$ 0.11	0.877 $\pm$ 0.044			
R.Hippocampus	0.809 $\pm$ 0.021	0.825 $\pm$ 0.024				0.844 $\pm$ 0.026	0.865 $\pm$ 0.018			
L.Hippocampus	0.808 $\pm$ 0.019	0.822 $\pm$ 0.022				0.848 $\pm$ 0.03	0.861 $\pm$ 0.025			
R.Putamen	0.884 $\pm$ 0.016	0.894 $\pm$ 0.013				0.891 $\pm$ 0.063	0.908 $\pm$ 0.039			
L.Putamen	0.884 $\pm$ 0.014	0.895 $\pm$ 0.011				0.894 $\pm$ 0.059	0.908 $\pm$ 0.030			
<b>Very Large</b>										
R.Thalamus	0.894 $\pm$ 0.008	0.906 $\pm$ 0.008	1.3	0.75	0.55	0.899 $\pm$ 0.036	0.915 $\pm$ 0.021	1.4	0.95	1.35
L.Thalamus	0.891 $\pm$ 0.008	0.903 $\pm$ 0.009				0.903 $\pm$ 0.033	0.912 $\pm$ 0.020			
<b>Av 3-14</b>	0.832 $\pm$ 0.023	0.849 $\pm$ 0.022				0.855 $\pm$ 0.058	0.875 $\pm$ 0.029			
<b>Av. full</b>	0.814 $\pm$ 0.027	0.834 $\pm$ 0.024				0.839 $\pm$ 0.061	0.862 $\pm$ 0.031			

Table 5.1: Performance analysis of the proposed method. Dice scores averaged over 6 or 5 folds are listed for a baseline Res. U-net and its variants: Trained with Proposed method, only pre-training (PT) and only ROI training (ROI).

Proposed Method	IBSR Dataset	MICCAI Dataset
	Dice(Mean $\pm$ STD)	Dice(Mean $\pm$ STD)
U-Net	0.824 $\pm$ 0.02	0.852 $\pm$ 0.07
Res U-Net	0.833 $\pm$ 0.02	0.862 $\pm$ 0.03
Dense U-Net	0.831 $\pm$ 0.06	0.860 $\pm$ 0.07
SE Net	0.832 $\pm$ 0.06	0.859 $\pm$ 0.06

Table 5.2: Average performance of sub-cortical structure segmentation for all structures using proposed method(PT+ROI) on various architectures on IBSR dataset and MICCAI dataset.

The proposed training strategy is agnostic to underlying deep learning architecture. Since the first phase, which is pre-training for tissue segmentation, as well as the second phase of training the pre-trained network with structure ROIs, can be applied to any network, we know the method is capable of being generalized. In order to further assess the method, we apply both phases of the proposed training regime to various U-Net variants such as U-Net, Residual U-Net, Dense U-Net, and SE Net (Squeeze and Excitation block).

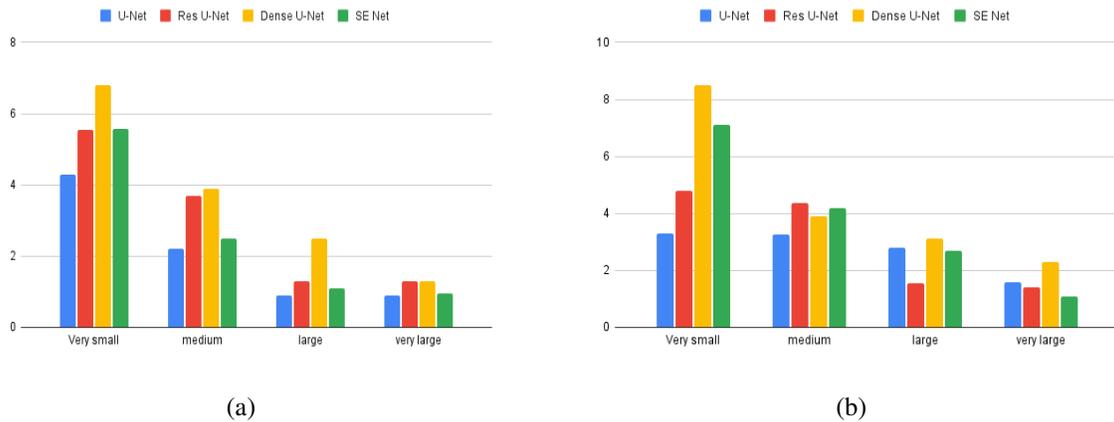
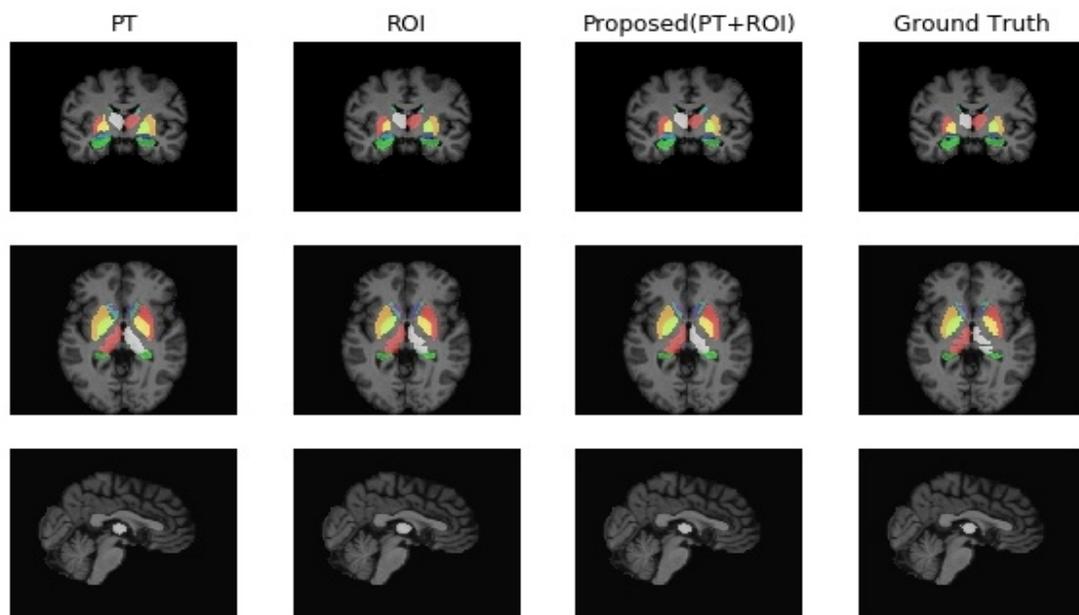


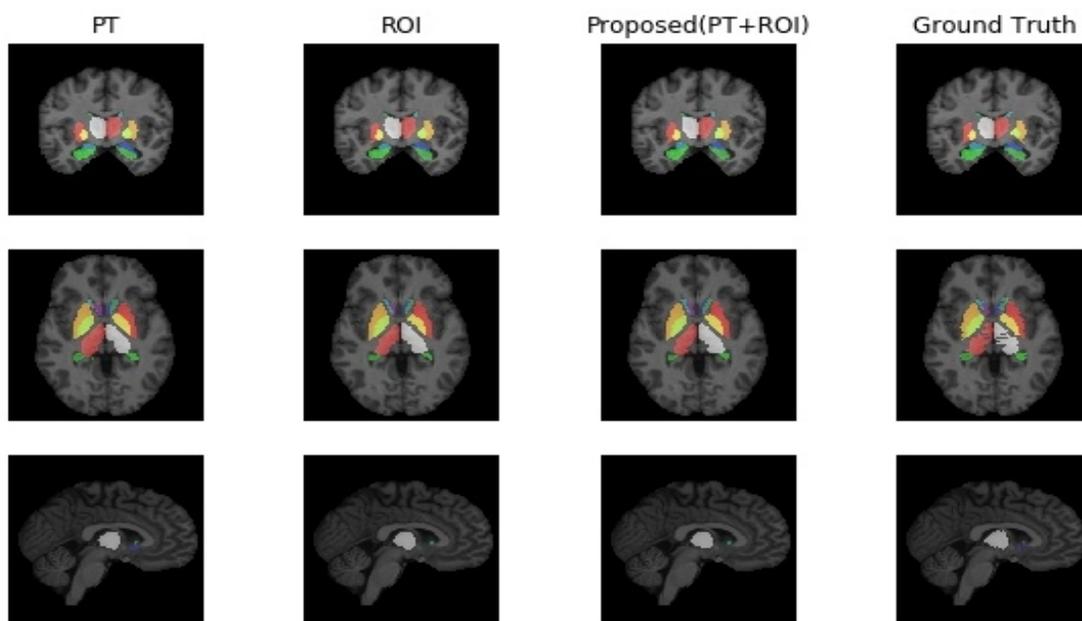
Figure 5.3: Plots show the % improvement in the Dice score of structures of different sizes for the proposed method on U-Net and Res. U-Net on (a) IBSR dataset (b)MICCAI dataset

The performance on IBSR and MICCAI data is shown in Table 5.2. The improvement in Dice scores obtained for the different groups of structures is presented as a bar graph in Fig. 5.3. The improvement in the performance of smaller structures is consistent for all architectures. The trend shows that the boost in performance increases with a decrease in structure sizes, effectively reducing the performance bias of the network due to variance in the size of structures.

### 5.6.3 Qualitative Results



(a)



(b)

Figure 5.4: Qualitative results of the proposed method and its variants. Top to bottom: Coronal, Axial, and sagittal slices. (a) IBSR dataset (b) MICCAI dataset

Sample outputs of the proposed method with Res. U-Net for the IBSR dataset and MICCAI dataset are shown in Fig. 5.4 for visual comparison, and it is observed that the labels obtained with the proposed method are very close to the ground truth with smooth boundaries. Sample of coronal, axial, and sagittal slices of the result of the proposed method (Res. U-Net) and its variants are also shown in Fig. 5.4 for both IBSR and MICCAI datasets. We can observe visually that the result proposed method of combining pre-training using tissue segmentation and ROI extraction is closer to the ground truth compared to only pre-training or only ROI extraction. The same trend follows in qualitative comparison from Table 5.1; it is due to the combined contribution of both the phases, pre-training, which helps the network to learn tissue types, and the ROI extraction, which mitigates the size bias of the network to learn the different structures.

#### 5.6.4 Comparison with other methods

Finally, a comparison is made with baseline segmentation solutions (Freesurfer, FIRST), Res. U-net, SE Net,[76] and the SOTA method ( $\psi$ -net [47]). Since the proposed method uses a batch size of 2 (due to limitations in hardware) whereas a batch size of 8 was used to report results in [47], we retrained the  $\psi$ -net with given hyperparameters on our hardware with a batch size of 2 for a fair comparison and report those results (see Table ??). It is instructive that the proposed method of training for a standard Res. U-net helps it match (and marginally outperform) even the SOTA solution, which employs a more complex LSTM-based strategy.

Method	Freesurfer	FIRST	Res. U-net	SE Net [76]	$\psi$ -net[47]	Proposed (Res U-Net)
<b>Dice (Mean<math>\pm</math>STD)</b>	0.74 $\pm$ 0.11	0.808 $\pm$ 0.08	0.811 $\pm$ 0.03	0.812 $\pm$ 0.07	0.821 $\pm$ 0.03	<b>0.833 <math>\pm</math> 0.02</b>

Table 5.3: Comparison of average performance of sub-cortical structure segmentation for all structures between standard (available in toolboxes),state of the art solutions and proposed method on IBSR dataset.

Method	Freesurfer	FIRST	Res. U-net	SE Net [76]	$\psi$ -net[47]	Proposed (Res U-Net)
<b>Dice (Mean<math>\pm</math>STD)</b>	0.725 $\pm$ 0.13	0.799 $\pm$ 0.09	0.839 $\pm$ 0.06	0.828 $\pm$ 0.07	0.842 $\pm$ 0.05	<b>0.862 <math>\pm</math> 0.03</b>

Table 5.4: Comparison of average performance of sub-cortical structure segmentation for all structures between standard (available in toolboxes),state of the art solutions and proposed method on MICCAI dataset.

## **5.7 Discussion and Conclusion**

A novel method was proposed to remove the size bias in sub-cortical structure segmentation by helping the network to learn to discriminate between the tissue types, then do a focused segmentation by supplying the network with the structure ROIs. With this simple support, the network performed better for the sub-cortical structure segmentation. The proposed modification on different networks shows a consistent improvement across datasets as well as network architectures. The proposed framework has the potential for generalization for other anatomy segmentation problems also where the structures vary widely in size. Integrating the proposed framework into the general segmentation rather than training in two phases and without additional annotation such as GM, WM, CSF labels, and atlas is also an avenue to explore.

## Chapter 6

# Solution Implementation

### 6.1 Authoring Tool

The authoring tool is created to manage numerous anatomical areas of interest since the human brain has many functioning substructures, each being quite complex. With the authoring tool, lessons can be created to show certain anatomical structures' functioning and their spatial relationships. The 3D models obtained from the clinical data through segmentation and 3D modeling are loaded into the authoring tool to create a lesson. A typical lesson is intended to demonstrate the functioning of a system in the human brain. The authoring tool provides a user-friendly GUI for the anatomy instructors to create a lesson and save it effortlessly. Each lesson can then be loaded by the anatomy viewer. The authoring tool was used to design the lesson plans based on input from anatomy teachers at GMC, Kerala. The modules that would benefit the most from the integration of the 3D anatomy viewer were determined based on GMC, Kerala's neuroanatomy curriculum. Basal Ganglia, optic Pathways, ventricles, thalamus, cerebrum, sulci, and gyri are the recommended lessons. Refer to Appendix A for more images of the authoring tool.

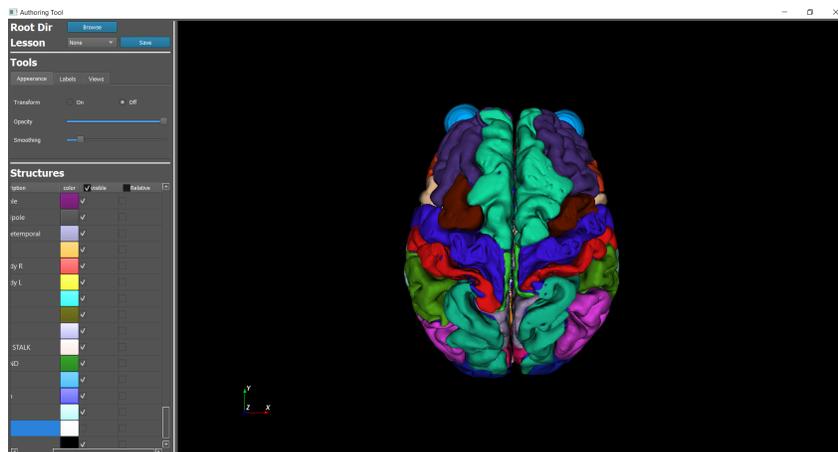


Figure 6.1: Authoring tool

## 6.1.1 Lesson Creation Workflow

### 6.1.1.1 Loading 3D Models

The authoring tool is capable of loading 3D models like surface meshes and tractography files. The browse button opens a file dialog. The users then pick out the directory which contains the 3D models. It detects if a saved lesson is present, then populates the drop-down with all the lessons available. On selection of particular lessons, the models, all the labels, and other settings of that lesson are loaded.

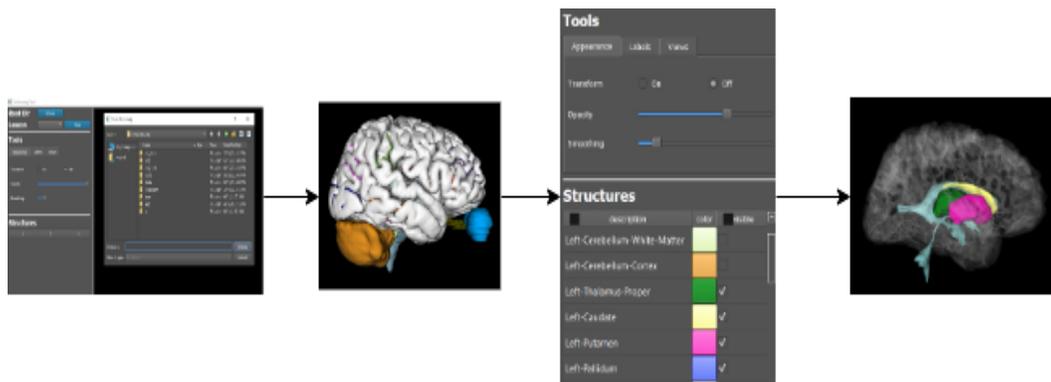


Figure 6.2: Flow diagram of loading models and setting structure properties

### 6.1.1.2 Structure Properties

The visual appearance of the 3D anatomical structures is essential for a seamless 3D effect. The authoring tool provides various tools to manipulate the appearance of the structures, as shown in Fig. 6.2. They are:

1. **Opacity:** The opacity of the structures can be changed to give certain textures or to provide spatial context for smaller structures present within bigger ones. A slider is given to control the opacity of a structure. It ranges from 0 to 1. With zero being transparent and one being opaque. Multiple structures can be selected at once, and their opacity varied at the same time.
2. **Smoothness:** Smoothing operation levels out the irregularities of the 3D structure, making the cells of the surface mesh better shaped and the vertices more evenly distributed. This is done iteratively for each vertex by finding modified coordinates based on its neighboring cells and vertices. A slider is given to control the number of iterations in the smoothing process of a structure. Its range is between 0 to 100. Increasing the slider value increases the smoothness.
3. **Color:** Anatomically different structures can be identified easily by coloring them differently. Coloring structures have the advantage of helping students to commit features of structures and their differences to memory. A color picker is given next to the entry of the structure in the table.

4. **Spatial Transformation:** Spatial transformations like rotation, scaling, and translation can be applied to individual structures manually in case the annotation obtained is incorrect. One must toggle into transformation mode in order to do this. In this mode, the same actions that are used to rotate, translate or zoom the global view can be used to apply the transformation to individual structures.

### 6.1.1.3 Labeling

The 3D labels are dynamically positioned in a circular grid. This is achieved by taking the centroid of all structures as the center of the circle and fixing all the labels at a certain distance(radius) from the center. The 3D labels have 3 degrees of freedom, so they automatically adjust when the view is transformed, and the text of the labels remains static.

To create a label, click on the label button on the side panel of the authoring tool, which opens up a dialog where the text of the label is to be entered, then click on the screen where you intend the label to be placed. Further adjustments can be made to the arrowhead by clicking and dragging it to place it in the exact position. A labeled structure is shown in Fig. 6.10.

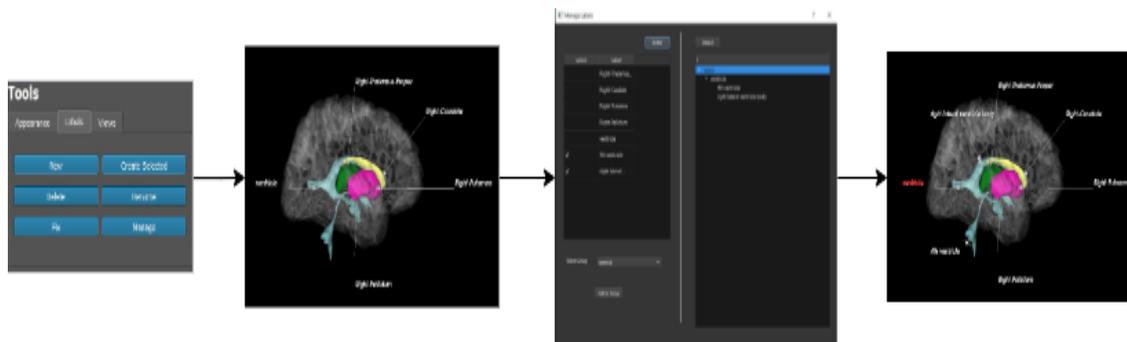


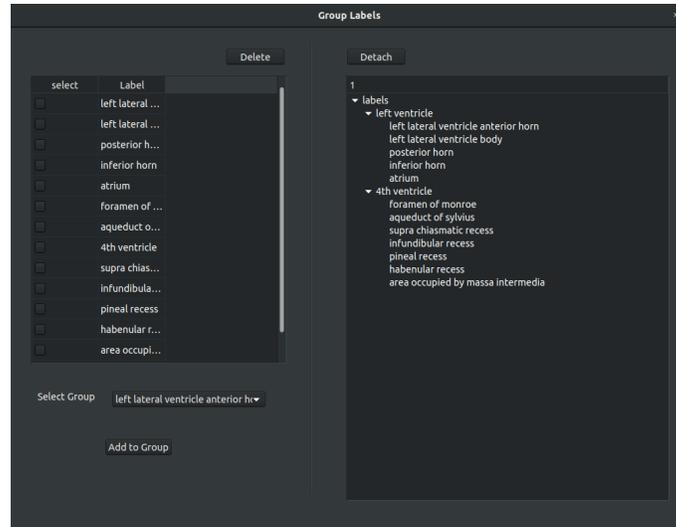
Figure 6.3: Flow diagram for labeling structures and organizing their hierarchy

### 6.1.1.4 Label Hierarchy

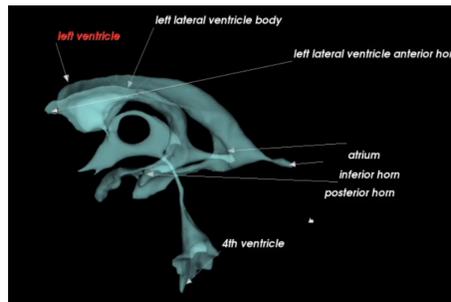
A lesson can consist of many structures, and each structure may contain substructures. There is a need to label them. If all the labels of all the structures are visible at once, it can create confusion and would look chaotic and complex. In order to simplify this, a label hierarchy is created. When a lesson is loaded, only root-level labels are visible. Clicking on a particular label will turn on its descendent labels so that one can focus on teaching that particular structure and its substructures; when done, clicking on the label again will collapse all of its descendants. This is illustrated in Fig. 6.4(b).

To create a label hierarchy, click on the group labels button on the side panel of the authoring tool, which opens up a window as shown in Fig. 6.4(a). This hierarchy works like a forest, which has many n-ary trees. A root node can be created by selecting the root label from the combo box, selecting its

children from the list, and clicking on the group button. By doing this repeatedly, one can create a hierarchy of labels. The hierarchy is also visualized in the right panel as a treeview structure.



(a)



(b)

Figure 6.4: (a) Grouping label feature in authoring tool to create label hierarchy (b) Label Hierarchy of Left ventricle

### 6.1.1.5 Anatomical Planes

To understand and describe anatomy in depth, it is vital to have a standard system. To avoid any ambiguity that can arise when describing the location of structures, we have anatomical planes and directions. Anatomical planes are 2D slices in 3D space that divide the human body orthogonally in the 3D coordinate system. Both anatomically and radiologically, the basic planes are from top to down (axial plane), from front to back (coronal plane), and from side to side (sagittal plane). In anatomy, we have anatomical positions. They describe the position of a structure with respect to the other, or they can be used as directions. We have, Anterior refers to the ‘front’, posterior refers to the ‘back’, Superior refers to ‘higher’, inferior refers to ‘lower’, Medial means towards the midline, lateral means away

from the midline. The authoring tool provides users with these views for instructors to consistently and coherently describe anatomical structures as shown in fig 6.5

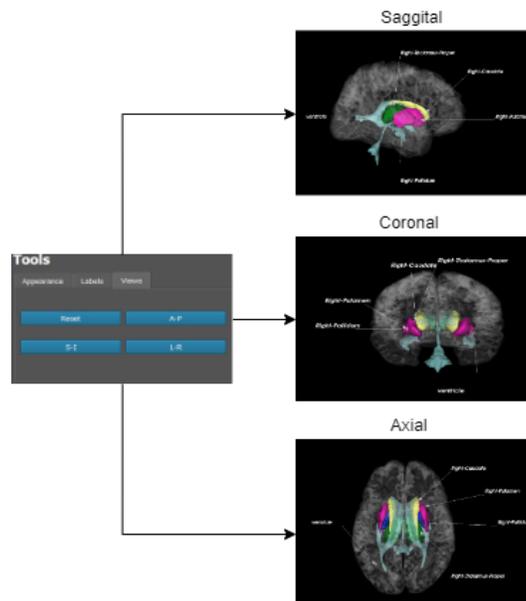


Figure 6.5: Anatomical planes and positions

### 6.1.1.6 Saving a Lesson

After designing the lesson, that is, deciding on the structures, relationships, labels, and other annotations, the lesson can be saved. Clicking on the save button opens up a dialog box for the name of the lesson. Other settings, such as MRI files and segmentation files for canonical planes, can also be set.

## 6.2 Anatomy Viewer

The anatomy viewer showcases the full-fledged lesson that the anatomy teachers designed and developed using the authoring tool. The viewer also integrates radiological data from MRI/CT scans with 3D anatomical structures. By generating a 3D scene of an anatomy lesson in stereo view, it creates an immersive 3D experience for anatomy education. Given that we have a complete 3D model, the left and right views are synthesized within the software using the camera model. It supports several forms of stereo rendering, including anaglyph, crystal eyes, and others. We employ side-by-side horizontal viewports of left and right views for stereo rendering since we use a dual projector polarised 3D configuration, with each projector producing left and right views. Refer to Appendix A for images of the anatomy viewer.

## 6.2.1 Lesson Configuration

On the first launch, a few configurations need to be set. The viewer application asks for the path for lessons that were created using the authoring tool. The other settings can be seen in Fig. 6.6.

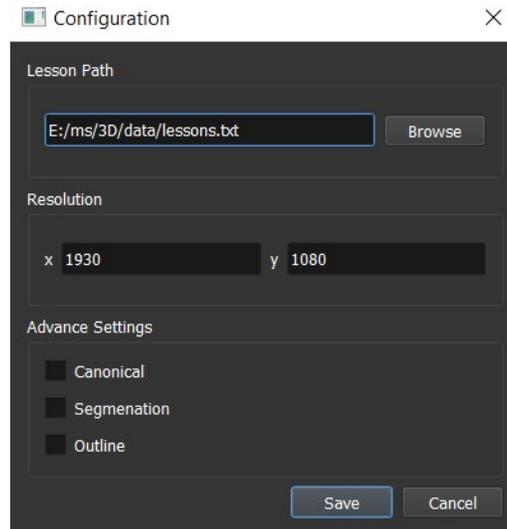


Figure 6.6: Viewer configuration

The default settings for canonical, segmentation, and outline can be set on this window as well as the display resolution of the projectors. Users can also modify the lesson path initially selected.

## 6.2.2 Lesson Launch

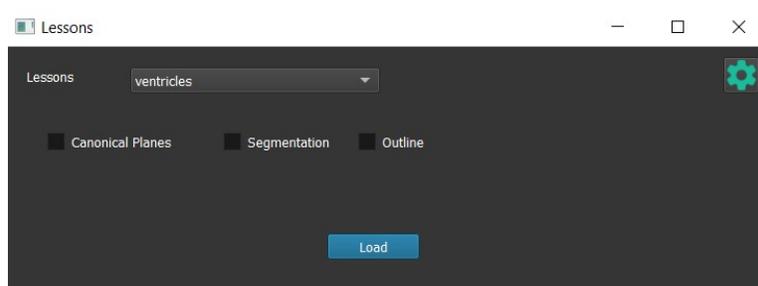


Figure 6.7: Lesson Launcher

After successfully setting the configurations, the main window of the anatomy viewer is revealed first on every subsequent launch of the application. The launcher is shown in Fig. 6.7 The lessons created and saved using the authoring tool are loaded and populated in the drop-down menu. Any lesson can be chosen to launch from the drop-down menu by clicking on the load button. Default options for canonical,

segmentation, and outline can be toggled. The configuration window can be opened by clicking on the settings button.

## 6.2.3 3D Models and Views

### 6.2.3.1 3D View

In this view, which is seen in Fig. 6.8(a), the structures of the brain, which are modeled as 3D meshes, are loaded and can be viewed with an optional outline or bounding box covering them. Here all the structures selected for a particular lesson are visible. A lesson can have relative structures which can be toggled on or off. The objective of relative structures is to help with visualizing the location of the structure to be studied with respect to others. The cortex is generally visualized as a translucent object so that it can encompass deeper structures, visualize the scale, and localize them. Each structure can be selected by clicking on it and will be highlighted, and then one can delete the structure or change its opacity. One can also view all the structures of the lesson and control their visibility. To get a better look at structures one can rotate/pan/zoom the current view. Doing this gives a better handle on how it would be to look at the actual brain in 3D. This view is a great tool for visualizing anatomy and understanding spatial relationships.

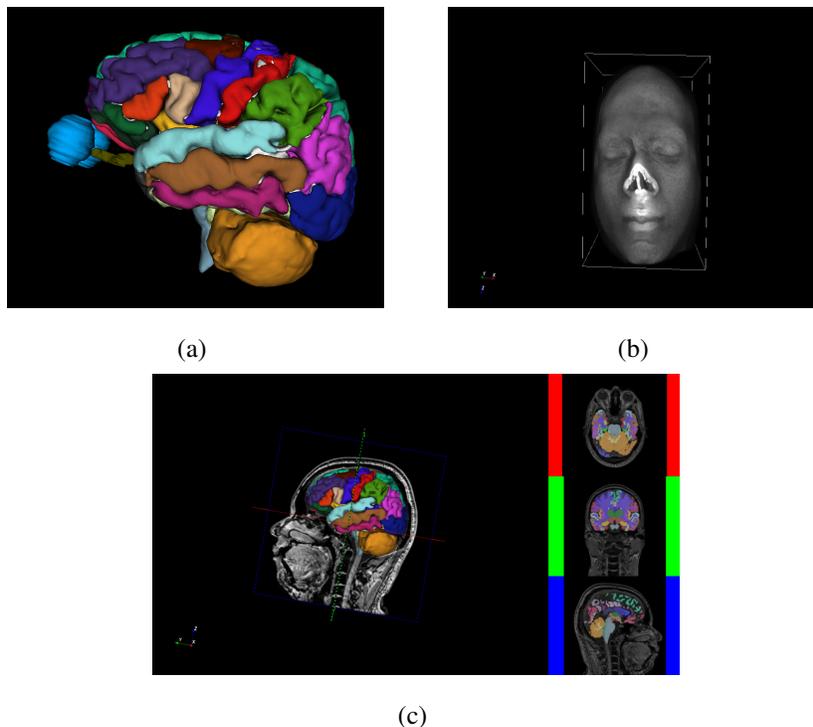


Figure 6.8: (a) 3D view of all cortical structures (b) Volume rendering of MR data (c) Canonical planes overlaid on 3D structures

### 6.2.3.2 Canonical Planes

Radiological anatomy is a crucial facet of learning and applying medical knowledge. Here, radiology is tightly coupled with the 3D anatomical structures from their inception. We showcase that advantage by blending MRI data and 3D surface meshes. This serves the dual purpose of introducing radiology and providing the 2D context for 3D anatomy. We have a 3D viewport that covers part of the screen and contains 3D meshes and the three canonical planes ie. axial, sagittal, and coronal planes. The cross-sections of the MRI are visible on each plane and are superimposed on the 3D model of the brain structures. The other part of the screen contains the 2D planes of the corresponding cross-section of axial, sagittal, and coronal views. The corresponding annotations of the structures present are also overlaid on the 2D cross-section of the MRIs. These can be seen in Fig. 6.8(c). Each plane can be selected, and its corresponding slice can be changed, both the 3D plane and 2D planes are synchronized. So changing the slice on one of them will result in changing the slice on the other.

### 6.2.3.3 Others

Volume rendering is used to visualize 3D volumetric data obtained using MRIs on a 2D display. Volume ray casting is the technique used to achieve this rendering. Volume rendering of the brain using MRI data can be seen in Fig. 6.8(b)

White matter fibers like corpus callosum that are specific to the neuroanatomy are extracted using DSI Studio from DTI images. They can also be part of the lesson. Trackvis library is used to handle the white matter fibers, which are saved as tractography files(track files, .trk files). The saved tractography files which are basically white matter fibers, are modeled into a bundle of polylines that can be colored with specific colors or randomly. Visualization of white matter fibers can be seen in Fig. 6.9(a,b).

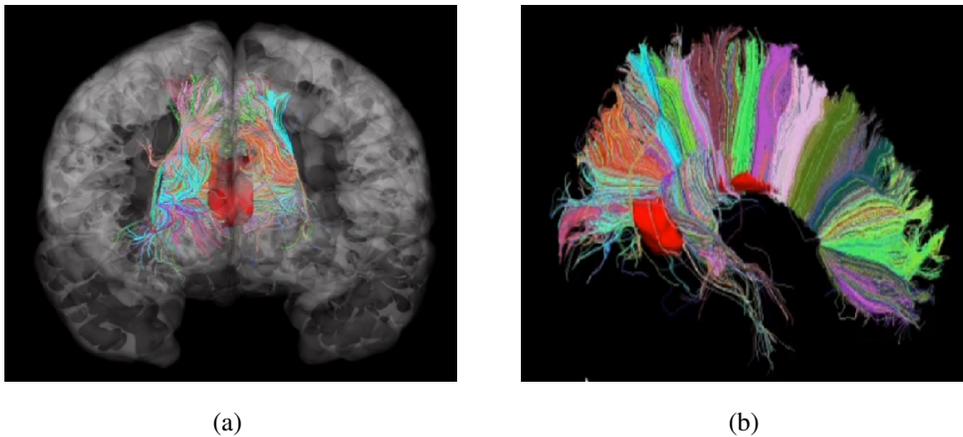


Figure 6.9: (a) 3D view of corpus callosum (b) white matter fibers of corpus callosum s

## 6.2.4 Remote Companion Application

The remote companion application works as an accessory to the anatomy viewer. It provides a smoother user-friendly interface and greater mobility within the classroom for instructors. The crucial elements to be able to remotely control the viewer are the android application and the client on the server where the viewer is present. The android application connects to the server via a network (Wi-Fi /Bluetooth).

### 6.2.4.1 Linux Client

It acts as a web server. It communicates with the android application through Wi-Fi using unencrypted UDP messages. A trusted environment is necessary for the use of Wi-Fi. Bluetooth tethering can also be used to create a mini network for communication with the desktop.

### 6.2.4.2 Android application

The android application provides ease of access to the anatomy viewer through a touch interface. It needs to be connected to the server's network. The various features of the android application are:

1. Lessons
2. Zoom
3. Rotation

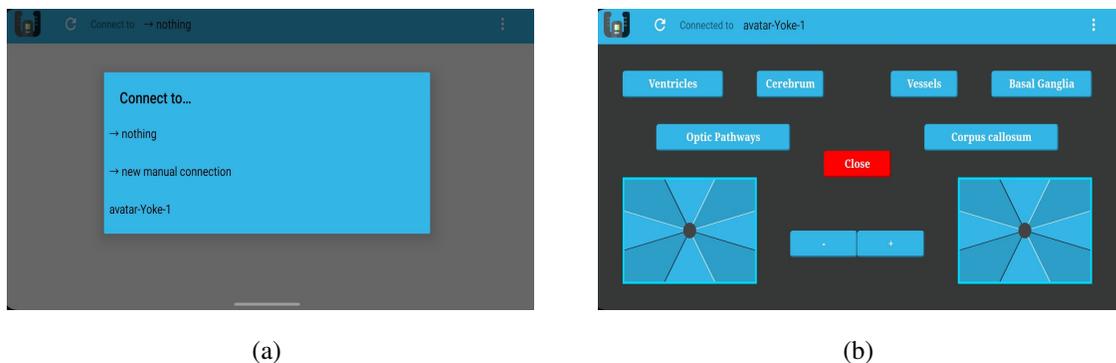


Figure 6.10: (a) A popup to connect to server's network (b) Android application main view

## 6.3 Software Specifications

In order to support multiple platforms, the applications are developed in Python. PyQt 5 library is used for User Interface. The main rendering and stereo visualizations are done by VTK 9.1.0. The

application is currently tested on Ubuntu 16.04 and Windows 10. Remote companion application also uses Python and Java(Android). We have adapted Yoke <sup>1</sup> to suit the anatomy viewer. The code for the anatomy viewer and the authoring tool is available at <https://github.com/mythri-venkat/3DVisualizer>.

## 6.4 System Assessment via User Feedback

The following feedback on the Anatomy Viewer and Authoring tool is provided by the Anatomy Professor at Government Medical College, Kerala, Dr. Doris George Yohannan, MD.

- How do you think this system/tool will help you in teaching anatomy?

My perspective is this will help me show structures better in 3D space in a lecture classroom scenario. The relationship between structures are something that is complex to be shown with 2D diagrams alone. The best place we can teach anatomy is the dissection hall, but the limitation is the limit to show a larger audience and cadaveric specimen availability.

- How do you think it will help the students learn anatomy?

As these are guided demonstrations, students can visualize structures much better with an instructor's guidance. I feel students would also greatly benefit if they can manipulate things on their own too, which can be a further futuristic extrapolation to this tool/system, which would obviously require more facilities in the classroom

- Is an authoring tool helpful in developing more lessons?

Yes. I feel it helps us to customize and curate lessons. As mentioned, neuroanatomical spatial relations are complex and often, structures spanning various conventional "chapters" may need to be visualized together (For instance, corpus callosum - a white matter anatomy structure, with the ventricle; or internal capsule - another white matter structure, with the lentiform nucleus.)

---

<sup>1</sup><https://github.com/rmst/yoke>

## *Chapter 7*

### **Discussion and Future Work**

In this thesis, we discussed the inadequacies of existing methods of anatomy education. The limited time and existing educational aids available such as a 2D atlas, PowerPoint, and cadavers, are insufficient for the students to understand all the critical aspects of anatomy. In order to overcome these challenges, we proposed a 3D interactive visualization solution that captures the students' attention. The progress of technology today makes 3D visualization inexpensive. Based on various other research, 3D interactive visualization is effective and shows significant results in engaging students and helping them grasp complex concepts of anatomy. Using the clinical cases illustrates the reality, and the inherent differences in individual anatomy play a role in greater understanding. Using the data from clinical scanners makes the process of creating 3D models effortless.

In the case of neuroimaging, we explored all kinds of clinical data that can be used to teach various functional and structural aspects of the human brain. Based on the syllabus of GMC, Kerala, and considering the limitations of the neuroimaging modalities, we came up with a cohesive set of lesson plans and their contents with the help of anatomy instructors. These lessons demonstrate the versatility of radiology in capturing tissues, fibers, and fluids within the brain and effectively integrating radiology with anatomy. We also explored various manual, automatic, and semi-automatic segmentation methods to obtain annotations using neuroimaging data collected.

In chapter 5, while further exploring automatic segmentation methods, we looked at the performance bias towards larger subcortical structures such as the thalamus in many of the automatic segmentation methods. We proposed a two-phased method to remove the size bias, which can be applied to any U-Net based architecture. The method has a pre-training phase and an atlas-based ROI extraction phase together to emphasize tissue differences and provide richer local and global information. Experiments on the MICCAI and IBSR datasets showed significant improvement in the performance of smaller structures after applying the proposed framework to all variations of U-Net architecture. The network shows only marginal performance improvement with only the pre-training phase, which helps the network differentiate between tissues. At the same time, a significant performance boost is observed for smaller structures with structure ROIs used for training the network from scratch; ROI extraction plays a vital role in giving equal importance to smaller structures. Combining both phases shows maximum im-

provement in the performance of small structures. A comparison of the proposed training framework on MICCAI and IBSR datasets to many other segmentation methods shows that the performance in terms of dice score is marginally better or on par with the current state-of-the-art methods on both the datasets.

Using the annotations obtained through the segmentation methods, we delved into the process of creating 3D models and sifted through various techniques of 3D modeling applicable to the neuroimaging data. We proposed an authoring tool that uses the generated 3D models and mainly helps anatomy instructors create lessons based on their teaching preferences and the curriculum. The user interface of the authoring tool is intended to offer various tools that simplify the process of developing and managing lessons. We explored the 3D interactive anatomy viewer, which imports the lessons created using the authoring tool and its capabilities in demonstrating the variety of ways it can render in stereo 3D structures, including surface meshes, streamlines, and volume rendering. We evaluated our solution through feedback from one of our key end users, the anatomy instructors, on the usefulness of the authoring tool for designing lessons and the anatomy viewer for instructing students.

### **7.0.1 Future work**

In this thesis, we looked at an economical 3D interactive solution to improve the way anatomy is taught presently to students in medical colleges. Deeper evaluation and analysis are necessary to improve the way 3D technology can be used to teach anatomy intuitively and enhance student understanding. This can be done through a systematic study and demonstration of the application in a live classroom, which is being organized and will take place soon. While we have showcased the use of the 3D solution for medical education, specifically in neuroanatomy, it is easily adaptable to other educational needs, to train professionals, in other large-scale demonstrations, etc.

We have successfully demonstrated the use of clinical data to obtain 3D models for anatomy teaching. However, clinical scanners have limitations; not all anatomical data can be extracted using 3D modalities. Other ways of obtaining microscopic anatomical details need to be explored.

The process of creating a lesson can be further streamlined, and the user experience of the interactive application can be improved based on the reviews from anatomy professors and students. The framework can be extended to provide a suite of software tools to achieve an end-to-end solution for creating 3D models from a variety of segmentation methods, designing lessons, and achieving 3D anatomy visualization.

In chapter 5 we proposed a method to remove size bias in subcortical structure segmentation. With proposed changes to existing architecture, it can marginally outperform the original network and even the state-of-the-art method in terms of dice coefficient. In the future, with the simplicity of the proposed method, it can also be applied to other regions with similar performance biases between their substructures due to their sizes. However, the pre-training step of segmenting tissues of the brain is only specific to the brain, and generalized pre-training can be developed to compensate. The ROI extraction step needs an atlas; while this boosted performance significantly, an automatic method to extract ROIs would be a more generalizable solution, which has its challenges.

## Related Publications

1. **Mythri V**, Alphin J Thottupattu, Naren Akash R J, Jayanthi Sivaswamy, *"A Method to remove size bias in sub-cortical segmentation," 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI 2022).*
2. Josline Joseph, Chandrasekharan Kesavadas, Dr.Jayadevan E.R, **Mythri V**, Jayanthi Sivaswamy, Doris Yohannan, Dr. Oommen, Umesan Govindapillai, Bejoy Thomas, Tirur Kapilamoorthy, *"Virtual Reality as a supportive tool to revolutionize neuro-anatomy education," 22nd Annual conference of Indian Society of Neuroradiology 2020.*

## Bibliography

- [1] A. Agbetoba, A. Luong, J. K. Siow, B. Senior, C. Callejas, K. Szczygielski, and M. J. Citardi. Educational utility of advanced three-dimensional virtual imaging in evaluating the anatomical configuration of the frontal recess. In *International forum of allergy & rhinology*, volume 7, pages 143–148. Wiley Online Library, 2017.
- [2] A. Akselrod-Ballin, M. Galun, M. J. Gomori, R. Basri, and A. Brandt. Atlas guided identification of brain structures by combining 3d segmentation and svm classification. In R. Larsen, M. Nielsen, and J. Sparring, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, pages 209–216, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [3] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019.
- [4] D. N. Brewer, T. D. Wilson, R. Eagleson, S. De Ribaupierre, et al. Evaluation of neuroanatomical training using a 3d visual reality model. In *MMVR*, pages 85–91, 2012.
- [5] L. Cerliani, M. Mennes, R. M. Thomas, A. Di Martino, M. Thioux, and C. Keysers. Increased functional connectivity between subcortical and cortical resting-state networks in autism spectrum disorder. *JAMA psychiatry*, 72(8):767–777, 2015.
- [6] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [7] P. Coupé, B. Mansencal, M. Clément, R. Giraud, B. Denis de Senneville, V.-T. Ta, V. Lepetit, and J. V. Manjon. Assemblynet: A large ensemble of cnns for 3d whole brain mri segmentation. *NeuroImage*, 219:117026, 2020.
- [8] A. V. Dalca, G. Balakrishnan, J. Guttag, and M. R. Sabuncu. Unsupervised learning for fast probabilistic diffeomorphic registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 729–738. Springer, 2018.
- [9] A. V. Dalca, G. Balakrishnan, J. Guttag, and M. R. Sabuncu. Unsupervised learning of probabilistic diffeomorphic registration for images and surfaces. *Medical image analysis*, 57:226–236, 2019.
- [10] A. Dale, B. Fischl, and M. I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179 – 194, 1999.

- [11] J. W. V. de Faria, M. J. Teixeira, L. d. M. S. Júnior, J. P. Otoch, and E. G. Figueiredo. Virtual and stereoscopic anatomy: when virtual reality meets medical education. *Journal of neurosurgery*, 125(5):1105–1111, 2016.
- [12] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- [13] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, and R. J. Killiany. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *NeuroImage*, 31(3):968 – 980, 2006.
- [14] D. A. Dickie, S. D. Shenkin, D. Anblagan, J. Lee, M. Blesa Cabez, D. Rodriguez, J. P. Boardman, A. Waldman, D. E. Job, and J. M. Wardlaw. Whole brain magnetic resonance image atlases: a systematic review of existing atlases and caveats for use in population imaging. *Frontiers in neuroinformatics*, 11:1, 2017.
- [15] L. Donnelly, D. Patten, P. White, and G. Finn. Virtual human dissector as a learning tool for studying cross-sectional anatomy. *Medical teacher*, 31(6):553–555, 2009.
- [16] S. Duchesne, J. Pruessner, and D. Collins. Appearance-based segmentation of medial temporal lobe structures. *NeuroImage*, 17(2):515–531, 2002.
- [17] B. Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.
- [18] B. Fischl and A. M. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20):11050–11055, 2000.
- [19] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.
- [20] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33:341–355, 2002.
- [21] B. Fischl, D. H. Salat, A. J. van der Kouwe, N. Makris, F. Ségonne, B. T. Quinn, and A. M. Dale. Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23(Supplement 1):S69 – S84, 2004. Mathematics in Brain Imaging.
- [22] B. Fischl, M. I. Sereno, and A. Dale. Cortical surface-based analysis: I: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195 – 207, 1999.
- [23] B. Fischl, M. I. Sereno, R. B. Tootell, and A. M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999.
- [24] B. Fischl, A. van der Kouwe, C. Destrieux, E. Halgren, F. Ségonne, D. H. Salat, E. Busa, L. J. Seidman, J. Goldstein, D. Kennedy, V. Caviness, N. Makris, B. Rosen, and A. M. Dale. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, 14(1):11–22, 2004.

- [25] S. Gouttard, M. Styner, S. Joshi, R. Smith, H. Cody, and G. Gerig. Subcortical structure segmentation using probabilistic atlas priors - art. no. 65122j. *Proceedings of SPIE - The International Society for Optical Engineering*, 6512, 03 2007.
- [26] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany, P. Maguire, D. Rosas, N. Makris, A. Dale, B. Dickerson, and B. Fischl. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, 32(1):180–194, 2006.
- [27] C. B. Hartberg, K. Sundet, L. M. Rimol, U. K. Haukvik, E. H. Lange, R. Nesvåg, I. Melle, O. A. Andreassen, and I. Agartz. Subcortical brain volumes relate to neurocognition in schizophrenia and bipolar disorder and healthy controls. *Progress in neuro-psychopharmacology and biological psychiatry*, 35(4):1122–1130, 2011.
- [28] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, and M. Reuter. Fastsurfer - a fast and accurate deep learning based neuroimaging pipeline. *NeuroImage*, 219:117012, 2020.
- [29] G. V. Hirsch, C. M. Bauer, and L. B. Merabet. Using structural and functional brain imaging to uncover how the brain adapts to blindness. *Annals of neuroscience and psychology*, 2, 2015.
- [30] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.
- [31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.
- [32] L. Igual, J. Soliva, S. Escalera, R. Gimeno, O. Vilarroya, and P. Radeva. Automatic brain caudate nuclei segmentation and classification in diagnostic of attention-deficit/hyperactivity disorder. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 36, 09 2012.
- [33] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015.
- [34] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith. Fsl. *Neuroimage*, 62(2):782–790, 2012.
- [35] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. MacFall, B. Fischl, and A. Dale. Reliability in multi-site structural mri studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, 30(2):436 – 443, 2006.
- [36] A. W. Keedy, J. C. Durack, P. Sandhu, E. M. Chen, P. S. O’Sullivan, and R. S. Breiman. Comparison of traditional methods with 3d computer models in the instruction of hepatobiliary anatomy. *Anatomical sciences education*, 4(2):84–91, 2011.
- [37] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

- [38] R. A. Kockro, C. Amaxopoulou, T. Killeen, W. Wagner, R. Reisch, E. Schwandt, A. Gutenberg, A. Giese, E. Stofft, and A. T. Stadie. Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. *Annals of Anatomy-Anatomischer Anzeiger*, 201:91–98, 2015.
- [39] D. Koshiyama, M. Fukunaga, N. Okada, F. Yamashita, H. Yamamori, Y. Yasuda, M. Fujimoto, K. Ohi, H. Fujino, Y. Watanabe, et al. Role of subcortical structures on cognitive and social function in schizophrenia. *Scientific reports*, 8(1):1–9, 2018.
- [40] G. R. Kuperberg, M. Broome, P. K. McGuire, A. S. David, M. Eddy, F. Ozawa, D. Goff, W. C. West, S. Williams, A. van der Kouwe, D. Salat, A. Dale, and B. Fischl. Regionally localized thinning of the cerebral cortex in Schizophrenia. *Archives of General Psychiatry*, 60:878–888, 2003.
- [41] K. Kushibar, S. Valverde, S. Gonzalez-Villa, J. Bernal, M. Cabezas, A. Oliver, and X. Lladó. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. *Medical image analysis*, 48:177–186, 2018.
- [42] B. Landman and S. Warfield. Miccai 2012 workshop on multi-atlas labeling. *Create Space Independent Publishing Platform*, 2, 2012.
- [43] K. Lee, J. Zung, P. Li, V. Jain, and H. Seung. Superhuman accuracy on the snemi3d connectomics challenge. *ArXiv*, abs/1706.00120, 2017.
- [44] M. Lee, J. Kim, R. Kim, H. G. Kim, M. Lee, S.-M. Wang, N. Y. Kim, D. W. Kang, Z. Rieu, J. Yong, D. Kim, and H. K. Lim. Split-attention u-net: A fully convolutional network for robust multi-label segmentation from brain mri. *Brain Sciences*, 10:974, 12 2020.
- [45] M. G. Linguraru, T. Vercauteren, M. Reyes-Aguirre, M. Ángel, M. González Ballester, N. Ayache, and M. Reyes. Segmentation propagation from deformable atlases for brain mapping and analysis. *Brain Research Journal*, 1, 01 2007.
- [46] C. Liu, J. E. Iglesias, and Z. Tu. Deformable templates guided discriminative models for robust 3d brain mri segmentation. *Neuroinformatics*, 11:447–468, 2013.
- [47] L. Liu, X. Hu, L. Zhu, C.-W. Fu, J. Qin, and P.-A. Heng.  $\psi$ -net: Stacking densely convolutional lstms for sub-cortical brain structure segmentation. *IEEE Transactions on Medical Imaging*, 39(9):2806–2817, 2020.
- [48] N. K. Logothetis. The underpinnings of the bold functional magnetic resonance imaging signal. *Journal of Neuroscience*, 23(10):3963–3971, 2003.
- [49] A. L. Maas, A. Y. Hannun, A. Y. Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Citeseer, 2013.
- [50] P. M. Matthews and P. Jezzard. Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(1):6–12, 2004.
- [51] R. Mehta and J. Sivaswamy. M-net: A convolutional neural network for deep brain structure segmentation. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 437–440. IEEE, 2017.

- [52] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.
- [53] D. T. Nicholson, C. Chalk, W. R. J. Funnell, and S. J. Daniel. Can virtual reality improve anatomy education? a randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Medical education*, 40(11):1081–1087, 2006.
- [54] K. Pohl, S. Bouix, R. Kikinis, and W. Grimson. Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework. volume 1, pages 81 – 84 Vol. 1, 05 2004.
- [55] M. Reuter and B. Fischl. Avoiding asymmetry-induced bias in longitudinal image processing. *NeuroImage*, 57(1):19–21, 2011.
- [56] M. Reuter, H. D. Rosas, and B. Fischl. Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4):1181–1196, 2010.
- [57] M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4):1402–1418, 2012.
- [58] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, May 2015.
- [59] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [60] H. D. Rosas, A. K. Liu, S. Hersch, M. Glessner, R. J. Ferrante, D. H. Salat, A. van der Kouwe, B. G. Jenkins, A. M. Dale, and B. Fischl. Regional and progressive thinning of the cortical ribbon in Huntington’s disease. *Neurology*, 58(5):695–701, 2002.
- [61] A. G. Roy, N. Navab, and C. Wachinger. Recalibrating fully convolutional networks with spatial and channel “squeeze and excitation” blocks. *IEEE transactions on medical imaging*, 38(2):540–549, 2018.
- [62] P. Ruisoto, J. A. Juanes, I. Contador, P. Mayoral, and A. Prats-Galino. Experimental evidence for improved neuroimaging interpretation using three-dimensional graphic models. *Anatomical Sciences Education*, 5(3):132–137, 2012.
- [63] D. Salat, R. Buckner, A. Snyder, D. N. Greve, R. Desikan, E. Busa, J. Morris, A. Dale, and B. Fischl. Thinning of the cerebral cortex in aging. *Cerebral Cortex*, 14:721–730, 2004.
- [64] L. Schmaal, D. J. Veltman, T. G. van Erp, P. Sämann, T. Frodl, N. Jahanshad, E. Loehrer, H. Tiemeier, A. Hofman, W. Niessen, et al. Subcortical brain alterations in major depressive disorder: findings from the enigma major depressive disorder working group. *Molecular psychiatry*, 21(6):806–812, 2016.
- [65] F. Segonne, A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl. A hybrid approach to the skull stripping problem in mri. *NeuroImage*, 22(3):1060 – 1075, 2004.
- [66] F. Segonne, J. Pacheco, and B. Fischl. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Trans Med Imaging*, 26:518–529, 2007.
- [67] M. Shakeri, S. Tsogkas, E. Ferrante, S. Lippe, S. Kadoury, N. Paragios, and I. Kokkinos. Sub-cortical brain structure segmentation using f-cnn’s. In *ISBI*, pages 269–272, 2016.

- [68] D. W. Shattuck and R. M. Leahy. Brainsuite: an automated cortical surface identification tool. *Medical image analysis*, 6(2):129–142, 2002.
- [69] J. Sled, A. Zijdenbos, and A. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Trans Med Imaging*, 17:87–97, 1998.
- [70] R. Sridharan, A. V. Dalca, K. M. Fitzpatrick, L. Cloonan, A. Kanakis, O. Wu, K. L. Furie, J. Rosand, N. S. Rost, and P. Golland. Quantification and analysis of large multimodal clinical image studies: Application to stroke. In *International Workshop on Multimodal Brain Image Analysis*, pages 18–30. Springer, 2013.
- [71] Y. Sun, X. Wang, and X. Tang. Deeply learned face representations are sparse, selective, and robust. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2892–2900, 2015.
- [72] S. Tan, A. Hu, T. Wilson, H. Ladak, P. Haase, and K. Fung. Role of a computer-generated three-dimensional laryngeal model in anatomy teaching for advanced learners. *The Journal of Laryngology & Otology*, 126(4):395–401, 2012.
- [73] C. P. Triepels, C. F. Smeets, K. J. Notten, R. F. Kruitwagen, J. J. Futterer, T. F. Vergeldt, and S. M. Van Kuijk. Does three-dimensional anatomy improve student understanding? *Clinical Anatomy*, 33(1):25–33, 2020.
- [74] D. Ulyanov, A. Vedaldi, and V. Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.
- [75] F. Venail, A. Deveze, B. Lallemand, N. Guevara, and M. Mondain. Enhancement of temporal bone anatomy learning with computer 3d rendered imaging softwares. *Medical teacher*, 32(7):e282–e288, 2010.
- [76] C. Wang, S. Ma, Y. Wei, X. Li, and Y. Liu. Subcortical brain segmentation in mr image based on residual fully convolutional networks. In *2020 Chinese Control And Decision Conference (CCDC)*, pages 4834–4837, 2020.
- [77] Y. Wang, Z. Deng, X. Hu, L. Zhu, X. Yang, X. Xu, P.-A. Heng, and D. Ni. Deep attentional features for prostate segmentation in ultrasound. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 523–530. Springer, 2018.
- [78] R. Wolz, P. Aljabar, D. Rueckert, R. A. Heckemann, and A. Hammers. Segmentation of subcortical structures and the hippocampus in brain mri using graph-cuts and subject-specific a-priori information. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 470–473, 2009.
- [79] S. Yousefi, N. Kehtarnavaz, and A. Gholipour. Improved labeling of subcortical brain structures in atlas-based segmentation of magnetic resonance images. *IEEE transactions on bio-medical engineering*, 59:1808–17, 03 2011.
- [80] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.
- [81] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha, et al. Resnest: Split-attention networks. *arXiv preprint arXiv:2004.08955*, 2020.

- [82] Ö. Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. *ArXiv*, abs/1606.06650, 2016.

## *Appendix A*

### **Images of our Solution**

#### **A.1 Photos of the Setup**

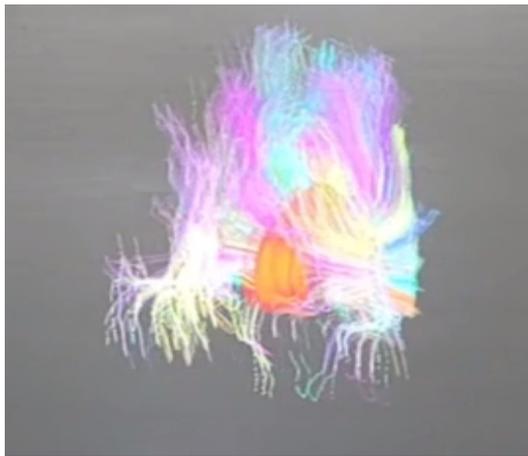


(a)



(b)

Figure A.1: Images of 3D anatomy viewer projected on the silver screen, laptop, and tablet.



(a)



(b)

Figure A.2: (a) Zoomed-in view of the viewer on the silver screen (b) Dual projectors with polarizers

## A.2 Images of the Authoring Tool and 3D Anatomy Viewer

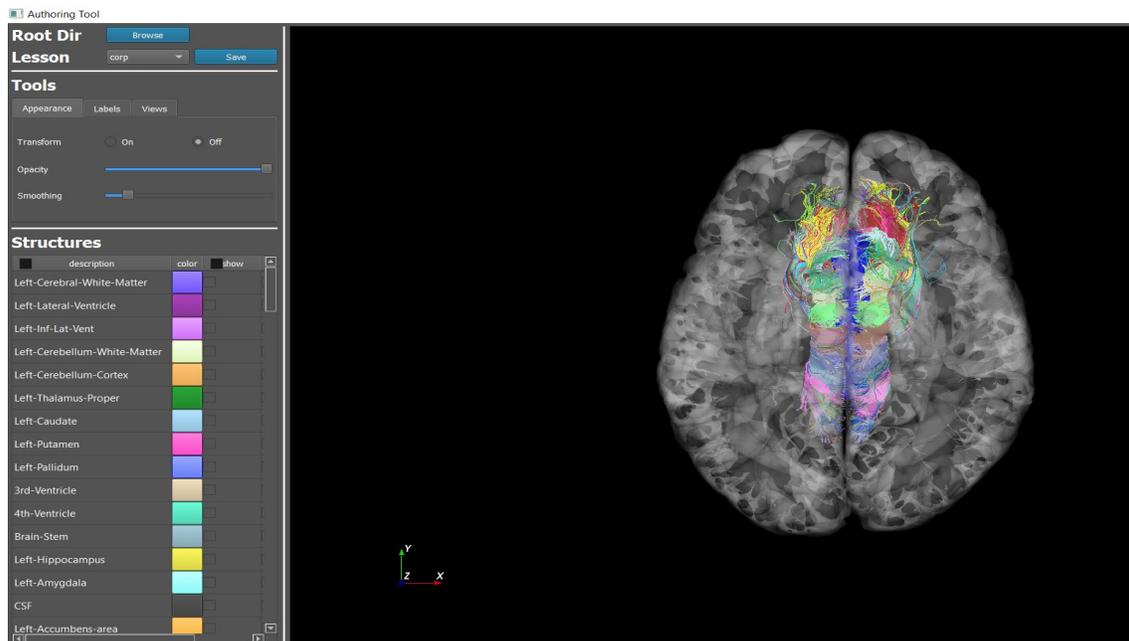


Figure A.3: Authoring Tool

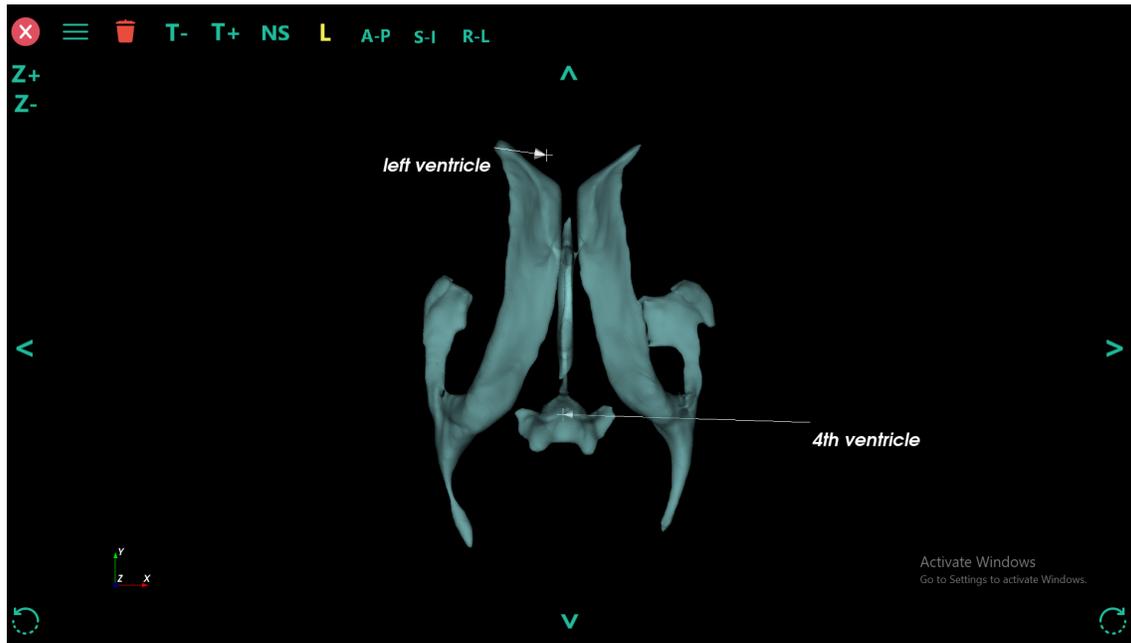


Figure A.4: 3D Anatomy Viewer in normal mode

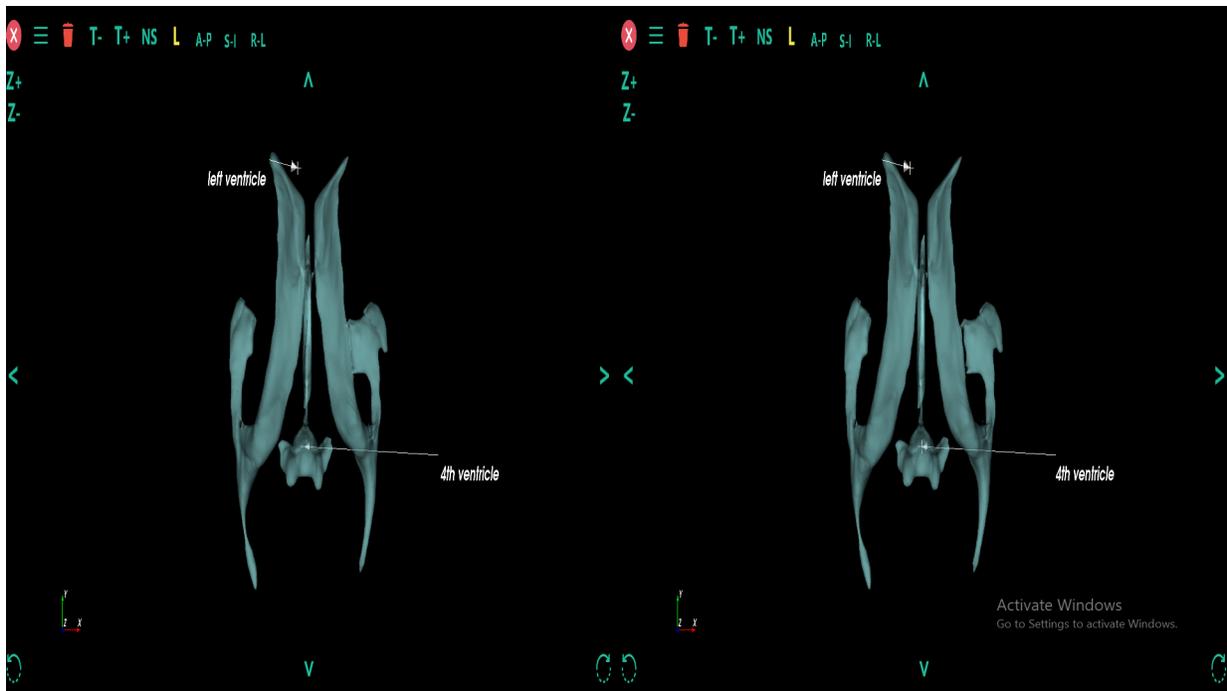


Figure A.5: 3D Anatomy Viewer in stereo mode

## *Appendix B*

### **Toolboxes**

We use the following toolboxes during various stages in our pipeline:

1. FSL: We mainly use it to automatically segment over 40 cortical and sub-cortical brain structures and conversion of data between different formats.
2. ITK-SNAP: We use it for semi-automatic and manual segmentation of structures and also to generate 3D mesh (STL files) from any kind of segmentation.
3. BrainSuite: It is used for cortical surface extraction.
4. DSI Studio: It is used for manual and automatic fiber tracking of white matter fibers in DTI images.
5. Blender: It is used to post-process 3D meshes.

#### **B.1 Freesurfer (FSL)**

FSL<sup>1</sup> is a comprehensive library of analysis tools for FMRI, MRI, and DTI brain imaging data. FreeSurfer is a set of automated tools for cortical and subcortical segmentation, reconstruction of the cortical surface, and other tasks. It uses a probabilistic atlas and other procedures described in publications ([13, 24, 21, 18, 20, 35, 40, 60, 63, 65, 69, 66, 23, 10, 22, 26, 56, 55, 57]). The steps involved in the automated segmentation are as follows: Motion correction and averaging of multiple volumetric T1 weighted images, Removal of non-brain tissue using a hybrid watershed/surface deformation procedure, Automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures, Intensity normalization, tessellation of the gray matter- white matter boundary and automated topology correction, surface deformation following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class and finally followed by parcellation of the cerebral cortex into units with respect to gyral and sulcal structure.

---

<sup>1</sup><http://surfer.nmr.mgh.harvard.edu/>

## B.2 ITK-SNAP

ITK-SNAP<sup>2</sup>[80] is a software that is used to segment structures in 3D medical imaging. ITK-SNAP offers semi-automatic segmentation based on active contour approaches, as well as manual delineation and image navigation. ITK-SNAP provides several supporting utilities in addition to these fundamental functionalities.

## B.3 BrainSuite

BrainSuite<sup>3</sup> [68] is a suite of open-source software tools for analyzing magnetic resonance imaging (MRI) of the human brain. The primary functions of these tools are to extract and parameterize the cerebral cortex's inner and outer surfaces, segment and label gray and white matter regions, and analyze diffusion imaging data. The steps involved in Cortical surface extraction are: Skull stripping, Nonuniformity correction, Tissue classification, Cerebrum labeling, Inner Cortical mask generation, Topology correction, Inner cortical surface extraction, and Pial surface extraction.

## B.4 DSI Studio

DSI Studio<sup>4</sup> is a tractography software that maps white matter fibers and connections in the brain using deterministic fiber tracking methods. It is a blend of several diffusion MRI approaches, including diffusion tensor imaging (DTI), generalized q-sampling imaging (GQI), q-space diffeomorphic reconstruction (QSDR), diffusion MRI connectometry, and generalized deterministic fiber tracking.

## B.5 Blender

Blender<sup>5</sup> is a free and open-source 3D modeling and animation software. It encompasses the entire 3D pipeline, including modeling, rigging, animation, simulation, rendering, compositing, motion tracking, video editing, and game development.

---

<sup>2</sup><http://www.itksnap.org/>

<sup>3</sup><http://brainsuite.org/>

<sup>4</sup><https://dsi-studio.labsolver.org/>

<sup>5</sup><https://www.blender.org/>