University of Minnesota

Deep Learning and Virtual Reality in the Surgical Sciences

by

Erik Gaasedelen

A dissertation submitted to the Graduate School
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

Bioinformatics and Computational Biology

Advisor: Dr. Paul A. Iaizzo, PhD, FHRS

March 2024

Acknowledgments

This dissertation reflects the kindness and support given to me by my community. I extend my gratitude to Paul for his mentorship, teaching, and for demonstrating that the seemingly impossible is indeed possible. I am also thankful to my colleagues for their role in my development as a researcher and communicator. Above all, I owe everything to my parents, who enabled me to concentrate on my education and take every opportunity available.

Abstract

The development of visualization tools has seen remarkable expansion, with Virtual Reality (VR) and Augmented Reality (AR) emerging as groundbreaking technologies that hold significant potential in the medical field. However, their application is currently limited by the need for further advancements in creating content that merits visualization. This dissertation delves into various applications of these technologies. Initially, we will leverage deep learning to derive insights from data, which will subsequently facilitate the creation of 3D models. Following this, we will investigate the visualization strategies themselves. We aim to demonstrate how such models can serve a wide range of medical applications, from educational purposes to pre-surgical planning and assistance during surgical procedures.

Contents

Acknowledgments **i**

Abstract **ii**

List of Tables **vi**

List of Figures **ix**

Introduction **1**

I Deep Learning For Cardiac Data Analysis and Segmentation **4**

Introduction to Part I **5**

1 The Application of Deep Learning for the Classification of Internal Human Cardiac Anatomy **6**

1.1 Background 7
1.2 Methods 8
1.3 Results 9
1.4 Interpretation 9
1.5 Conclusion 12

2 Deep Learning for the Segmentation and Model Generation of Cardiac Atria from Non-Contrast CT Scans **13**

2.1 Abstract 15
2.2 Introduction 15
2.3 Methods 17
2.3.1 Data 17
2.3.2 Preprocessing 17
2.3.3 Architecture 18
2.3.4 Training 21
2.3.5 Evaluation 21
2.4 Results 22

2.5 Conclusions 24

3 Automated Multiclass Cardiac Volume Segmentation and Model Generation 25
3.1 Abstract 26
3.2 Introduction 27
3.2.1 Anatomical Segmentation 27
3.2.2 Deep Learning 29
3.2.3 Autoseg using deep learning 31
3.3 Methods 33
3.3.1 Dicom Preprocessing 33
3.3.2 Architecture 34
3.3.3 Training 34
3.3.4 Evaluation 34
3.3.5 3D Model Generation 36
3.4 Results 36
3.5 Discussion/Conclusion 39

II Virtual Reality in the Visible Heart Laboratories 42

Introduction to Part II 43

4 3D Graphics to Virtual Reality in Medicine - Opportunities and Prospective 44
4.1 Abstract 45
4.2 Introduction 46
4.3 Simulations and Developed Movie Realism 50
4.4 Real Time Rendering 54
4.5 Expansion of visualization hardware 57
4.6 Virtualized Computation 61
4.7 Prospects 64
4.8 Conclusion 67

5 Using Smartphone-Based Virtual Reality to Explore Internal Anatomy of 3D Heart Models 69
5.1 Background 70
5.2 Methods 71
5.3 Results 72
5.4 Interpretation 74
5.5 Acknowledgements 76

6 Virtual Reality and Visualization of 3D Reconstructed Medical Imaging: Learning Variations within Detailed Human Anatomies 77
6.1 Abstract . . . 79
6.2 Introduction . . . 79
6.3 Description . . . 81
6.4 Discussion . . . 84
6.4.1 Education . . . 84
6.4.2 Presurgical Planning . . . 87
6.4.3 Devices and Industry . . . 90
6.5 Future Directions . . . 91
6.6 Acknowledgments . . . 93

7 Immersive Anatomical Scenes that Enable Multiple Users to Occupy the Same Virtual Space: A Tool for Surgical Planning and Education 94
7.1 Abstract . . . 95
7.2 Introduction . . . 96
7.3 Methods . . . 98
7.3.1 3D Model Development . . . 98
7.3.2 Required Hardware and Software . . . 99
7.3.3 Development of Multiplayer Functionality . . . 100
7.3.4 User Interface . . . 102
7.4 Results . . . 105
7.4.1 Anatomical Education . . . 105
7.4.2 Pre-Surgical Planning . . . 107
7.4.3 Medical Devices . . . 110
7.5 Discussion . . . 110
7.6 Conclusion . . . 113

Afterward 115

Bibliography 117

List of Tables

1.1 Neural network architecture 9

3.1 Comparison of mIOU for each chamber and ensemble 36

List of Figures

1.1 Images depicting 12 of the anatomical classes used for network training. . . . 8
1.2 Confusion matrix demonstrating interclass accuracy. Diagonal represents correct labels. 10

2.1 Examples from a batch showing images where the atria are highlighted. Note most of the images do not contain atria since most of the viewable volume is background. 19
2.2 U-Net Architecture 20
2.3 Jaccard Metric 22
2.4 Example 3D model generated from a validation sample. First column - ground truth. Second column - prediction. Third column - overlay. First row - Anterior view of atria. Second row - Posterior view of atria. 23

3.1 Flowchart diagram of proposed mixed reality methodology. (a) A DICOM scan in MIMICS depicting a cross sectional view of a heart. (b) 2D mask of the heart tissue generated through thresholding functions. (c) The resulting 3D model of the heart created by the 2D masks which can then be 3D printed. (d) The same 3D model of the heart is imported into Unity and a virtual reality environment of this given heart is generated. 28
3.2 U-Net Architecture 30
3.3 Resnet-34 architecture 30
3.4 Screenshot of manual segmentation software, Mimics, developed by Materialise 32
3.5 Dataset Example 33
3.6 Example Augmented Data 35
3.7 Collage of input scan, predicted mask, and ground truth mask 37
3.8 3D Model of network segmented heart. Right atrium (upper left), left atrium (upper right), right ventricle (lower left), and left ventricle (lower right). . . . 38

4.1 First images stored in a computer, first use of vector graphics, and ENIAC. . 47
4.2 Market trends for video games and image from Universe Sandbox, a popular modern videogame which simulates planetary bodies. 49
4.3 Figure from Turing's paper on morphogenesis and an image from ANSYS showing morphogenesis method for creating a structurally stable roof[41]. . . 51

4.4 Living Heart project simulation results and figure form Pixar paper on solid-fluid interaction. 53
4.5 Dr. Daniel Keefe with virtual reality system, 3D model of conjoined twin model (developed and printed by Alex Mattson from the Visible Heart Laboratory), and investigators looking at twin model in virtual reality. 57
4.6 Cadaver model in virtual reality, Micra delivery looking up the tricuspid valve, and mitral/aorta shot using Google Daydream. 59
4.7 Detection of persons in a picture, point cloud information, convolutional neural network to detect diabetic retinopathy. 62
4.8 The Davinci surgical robot and a reinforcement-learning based robot locomotion from Google [20]. 65
4.9 Projected growth of robotics in the medical field[1] and an anatomy demo from the Microsoft Hololens augmented reality headset. 66

5.1 VR Headset Setup 73
5.2 VR view of mitral valve 74
5.3 Apex view 75

6.1 Perfusion-fixed human heart (Heart 229) from the specimen library at the Visible Heart Laboratories. This heart was placed in agar gel, MRI scanned, 3D computationally modeled, and then simulations were incorporated into virtual reality scenes. 82
6.2 (a) DICOM scans were loaded into Mimics and detained anatomies were masked utilizing various thresholding functions. (b) 3D models of anatomy were generated from these masks. (c) Resultant 3D models were imported into Unity and a virtual reality environment of this specific heart was generated. 83
6.3 (a) External image of human heart reanimated on the Visible Heart® apparatus. (b) Videos captured with endoscopes placed in various locations inside beating hearts; this image displays both mitral (right) and aortic (left) valves. (c) Example of placement of functional internal anatomy video beside the virtual anatomy, to demonstrate how valves function within a beating human heart. 85
6.4 (a) Virtual reality scene of the anterior mediastinum of a congenital thoracic model depicting various anatomies, generated from a clinical CT scan. (b) Posterior view of the same congenital heart model with the lung and skeletal anatomies removed. (c) Typical scene of a mitral valve model we can examine in virtual reality; this view would allow a physician to accurately size the mitral valve annulus for either a virtual or subsequent procedural valve repair. 89

6.5 (a) Culotte stenting technique within a human coronary artery bifurcation. Actual stenting procedure was performed by a physician in a swine heart using Visible Heart® methodologies; stents were imaged with micro-CT, modeled, and then digitally placed within a human heart with similar coronary bifurcation anatomies. (b) Placement of leadless Micra pacemaker delivery tool system (Medtronic, Minneapolis, MN, USA) within the right ventricular apex of a human heart 3D model; image was taken from within the tricuspid annulus. (c) Depiction of Micra delivery taken from the right ventricular apex. 92

7.1 Flowchart diagram of proposed mixed reality methodology. (a) A DICOM scan in MIMICS depicting a cross sectional view of a heart. (b) 2D mask of the heart tissue generated through thresholding functions. (c) The resulting 3D model of the heart created by the 2D masks which can then be 3D printed. (d) The same 3D model of the heart is imported into Unity and a virtual reality environment of this given heart is generated. 98
7.2 Image of a user wearing VR headset next to their virtual avatar. 101
7.3 (a) The network state created for our avatar. This state contains the sphere transform property we defined. (b) A bolt entity was added to our avatar, defining its network state. 102
7.4 (a) The menu displaying the scenes the user can enter. Depicted in (b) is the change in transparency of the button, highlighting a collision of the laser and the current button. 103
7.5 Method for UI Button Scene Selection 104
7.6 (a) An external view of a computationally modeled human heart in the virtual reality environment. (b) Shown here, two users in the left atrium are studying the anatomy of the mitral valve. The instructor in the scene is using their laser to highlight the large papillary muscle while explaining its function to the other users. 106
7.7 (a) An example of an instructor teaching anatomy to others on a cadaver model, at native scale. (b) The scale of the cadaver has been enlarged and users can be seen navigating within the spinal canal. 108
7.8 (a) External view of a computationally modelled pediatric congenital heart disease case. (b) The lungs and skeleton have been removed to better depict the features of this congenital heart. (c) The clinical care team can analyze the inversion of the great vessels. (d) Collaborators standing in the right ventricle study the ventricular septal defect. 109
7.9 (a) A collaborator (e.g., medical device developer) showing the encapsulation of the Micra tines in the right ventricular myocardium to other collaborators within this unique human heart environment. (b) Collaborators in the right atrium looking at the Micra and extraction tool through the tricuspid valve. (c) A user in the right ventricle highlighting the lasso on the distal end of the extraction tool that is employed to snare the Micra. 111

Introduction

This text aims to discuss the theories, applications, and opportunities presented by deep learning and virtual reality within the realm of cardiac surgical sciences. It will delve into these subjects through the specific perspective of the Visible Heart® Laboratories ecosystem and my personal experiences therein. While our lab primarily focuses on the systems physiology of the heart, its projects and interactions cover a broad spectrum of topics that intersect academia, industry, and clinical practice.

The Visible Heart® Lab, initiated in 1997 under Dr. Paul Iaizzo's guidance, embarked on pioneering studies involving the novel reanimation of large mammalian heart specimens, including humans. This breakthrough fostered a long-standing collaboration between the University of Minnesota and Medtronic, a leading medical device company. Research at the lab has spanned cardiac device design, physiology, pulmonology, transplantation, gross anatomy, biomechanics, electrophysiology, studies on black bear hearts, 3D printing, deep learning, virtual reality, and more.

My tenure in the lab has afforded me numerous opportunities to ponder future medical technologies, informed by my academic background. I pursued neuroscience in college and am currently completing a PhD in Bioinformatics and Computational Biology. The lab generally hosts students from Biomedical Engineering, with occasional contributions from Mechanical Engineering. A notable observation is the extensive collaboration between surgical sciences and engineering, contrasted with a lesser engagement with computer science, an unexpected finding given the latter's burgeoning mainstream popularity and success in tech industries. Why does medicine diverge in this respect?

Software development in medicine presents unique challenges, differing significantly from the typical coder's ethos. The target audience is narrower, feedback harder to come by, domain knowledge a requisite, FDA approvals often necessary, and development and maintenance costly. This environment has led to the emergence of two distinct groups: software engineers, proficient in various programming languages and frameworks yet unfamiliar with clinical practice, and medical professionals, deeply knowledgeable about physiology, pharmacology, and medical devices but lacking in software engineering skills. Bridging communication between these groups is challenging but improving gradually.

I advocate for a future heavily influenced by a software-first approach to medical care and device design. While this text will not delve into specifics, such as wearable devices, genome sequencing, and comprehensive medical histories, it highlights the critical need for systems that not only collect data but make it actionable. The emphasis is not on the volume of data collected but on the extraction of valuable insights. Medicine has historically excelled at data collection; the challenge now is to develop software that automates tasks, reveals previously obscured information, or transforms data to enhance medical practitioners' work.

My initial project inspiration in the lab was to simplify the conversion of DICOM scans into 3D models for printing, a process traditionally requiring extensive manual effort. This project, a significant portion of my dissertation, addresses the broader issues of data automation and transformation. Despite the lab's success in collaborating with physicians on 3D printing clinical anatomy, widespread adoption remains limited. This raises questions about opportunity costs and the barriers to broader implementation of clinical 3D models.

The Visible Heart® Laboratories is a dynamic environment, attracting visitors from cardiac specialists to middle school students. Demonstrating the value of our projects to a diverse audience, however, can be challenging. Successful ideas must be easily communicable and implementable, a testament to their immediate utility or "first-order utility" – improvements to a system that don't introduce additional dependencies.

This brings us to deep learning (DL) and virtual reality (VR), technologies that have seen a surge in popularity and are at the heart of addressing data competence through automation, extraction, and transformation of information. My contributions aim to enhance first-order utility, with a focus on aiding others rather than pursuing commercial ventures or intellectual property.

The following sections will explore the science, development, and potential of these technologies, encouraging the acquisition of skills, generation of innovative ideas, and conviction in the value of further exploration. While some of my views may be speculative and possibly outdated by the time of reading, it's important to base opinions on evidence rather than conjecture. This document seeks to provide such evidence, underscoring the importance of these technologies as catalysts for a shift towards rigorous and adaptable data practices in medicine. This represents both an opportunity and a challenge; those unable to adapt will inevitably fall behind, mirroring trends already observed in the tech industry. Automation spares no one unprepared, making the mastery of these technologies not just beneficial but essential for future medical practitioners.

Part I

Deep Learning For Cardiac Data Analysis and Segmentation

A simple way to frame my dissertation is by dividing it into two main components: analyses and visualizations. If we adopt this perspective, the current section falls under analyses, which I believe offers the broadest scope and the greatest potential for future impact across medical disciplines. At the heart of any visualization effort lies a sophisticated computational backend, responsible for conducting the intricate numerical processing that yields polished, interactive outcomes. Here, my focus is on the role of deep learning as a foundational element for the visualization sphere, though its application is not strictly necessary. For example, medical diagnostics might not always necessitate an advanced graphics engine; sometimes, a straightforward output indicating the presence or level of a specific measurable entity suffices. Moreover, the significance of data automation cannot be overstated, a topic we will delve into more deeply in the second chapter. The methodologies we are set to examine have implications for a multitude of areas. Therefore, I encourage readers to consider the broad applicability of deep learning across various domains: reflect on how these tools might be integrated into your own field or how they are already being utilized in others.

Chapter 1

The Application of Deep Learning for the Classification of Internal Human Cardiac Anatomy

Authors:

Erik Gaasedelen[1], Alex Deakyne[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

I developed and showcased this project at the Design of Medical Devices conference. In our laboratory, we often reanimate human hearts to capture internal footage of various anatomical structures. Reviewing this footage subsequently can be an arduous and time-consuming task. Herein, I describe a deep learning application that demonstrates the potential for automating such processes, enabling us to enhance the productivity of the Visible Heart® Laboratories. However, even seemingly straightforward tasks proved to be more complex than merely devising a model. This is a common oversight I've observed, even among large corporations. The journey from data collection to the final output necessitates considerable planning around infrastructure. A critical question arises: how does the deep learning model integrate within a broader technological framework to achieve the desired outcome? Constraints such as time, budget, and personnel often render this crucial phase challenging. Yet, if the aim is to achieve direct and practical benefits from this project, developing such an infrastructure is the most straightforward route to our objectives.

In my paper, I advocate for the use of this technology in robot-assisted surgery or for enhancing localization techniques, though these applications may seem distant. However, continued research could make them feasible. For example, regional 2-D segmentation could aid in identifying specific anatomical features like a valve plane, which could be targeted during procedures such as a transseptal puncture. The main challenge lies in devising a simplified model that could be effectively paired with actuators within a realistic timeframe, likely requiring interdisciplinary collaboration.

1.1 Background

The applications of sensing and localization are becoming more sophisticated in many invasive and non-invasive surgical procedures and there is great interest to apply them to the human heart. Ideally, such tools could be indispensable for allowing physicians to spatially understand relative tissue morphologies and their associated electrical conduction. Yet today there remains a steep divide between the creation of spatial environment models and the contextual understandings of adjacent features. To begin to address this, we explore the problem of anatomical perception by applying deep learning to the identification of internal cardiac anatomy images.

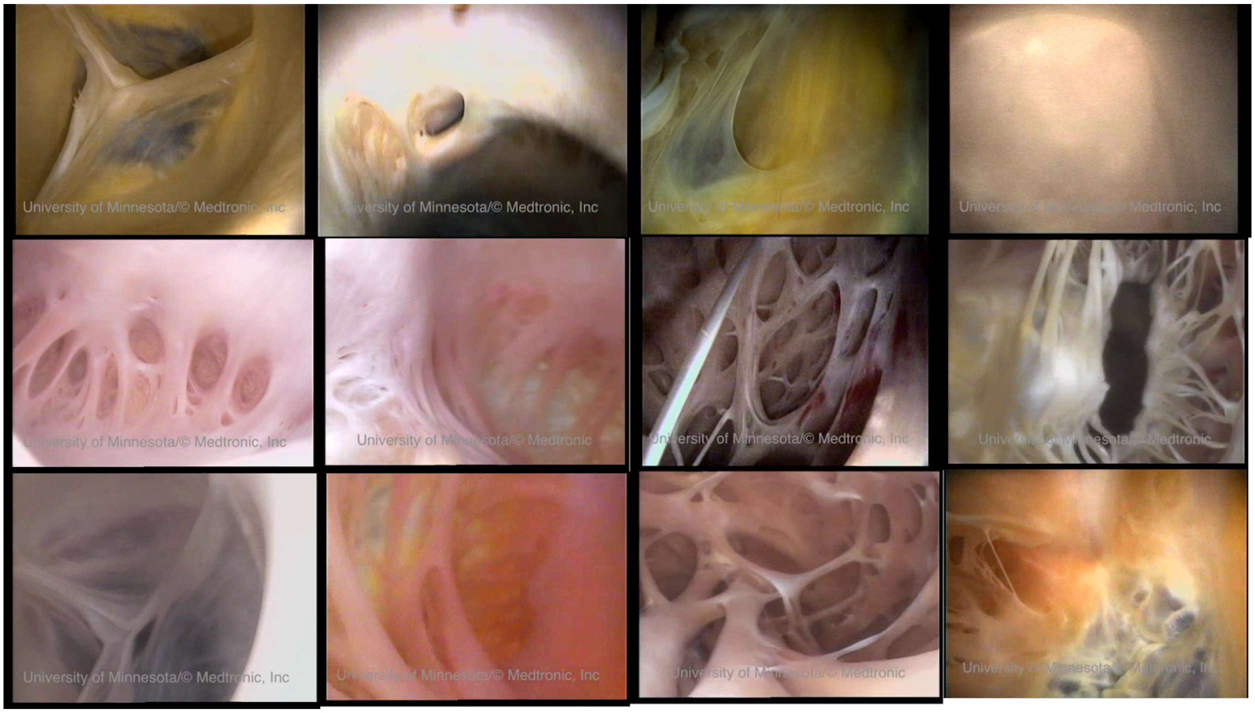


Figure 1.1: Images depicting 12 of the anatomical classes used for network training.

1.2 Methods

A dataset of 339,048 images of internal functional cardiac anatomy was taken from videos of reanimated human hearts using Visible Heart® methodologies.[25] Subsequently, 91,688 images taken from a subset of human hearts were used as the validation set. The remaining images were used for model training. The separation of data based on unique hearts prevented mixing of highly correlated images between training and validation sets. These images were separated into 13 different classes, attempting to define unambiguous features of cardiac anatomy. (This heuristic is not entirely justifiable in practice. Anatomical regions are not well delineated from each other in reality.) The anatomy classes included: aortic valve, coronary sinus ostium, crista terminalis, fossa ovalis, left atrial appendage, lateral wall, mitral valve, pulmonic valve, right atrial appendage, right ventricular apex, triangle of Koch, and tricuspid valve. Examples from the first twelve of these anatomies are shown in column major order in Figure 1.1. Videos for these can be found on the Atlas of Human Cardiac Anatomy[25]. (http://www.vhlab.umn.edu/atlas/)

Images were interpolated to 224x224 pixel density with zero mean intensities. Bottleneck features from the last fully connected layer of a pretrained VGG network[44][40] trained on the ImageNet dataset were calculated for each of these resulting images in the training set. These intermediate features were then trained on a three layer fully connected network; using categorical cross entropy loss with batch normalization[27] on each layer with significant dropout. (Table 1.1)

Table 1.1: Neural network architecture

Layer	**Units**	**Dropout**
1 (Input)	512 (ReLU-BN)	0.5
2	256 (ReLU-BN)	0.7
3	128 (ReLU-BN)	0.7
4 (Output)	13 (Softmax)	None

1.3 Results

Top-1 and top-3 validation accuracies were 0.68 and 0.91 respectively. Error modes can be observed through a confusion matrix. (Figure 1.2)

Training was performed using an Adam optimizer for 3 sets of 5 epochs with successive learning rates of 0.01, 0.001, and 0.0001. Any additional training led to worse performance on the validation set.

1.4 Interpretation

Our early utilization of a simple deep learning approach worked remarkably well for this cardiac image dataset. Specifically, complicated features could be correctly labeled even given the limitations of low image resolutions, ambiguous anatomical classes, and relatively noisy

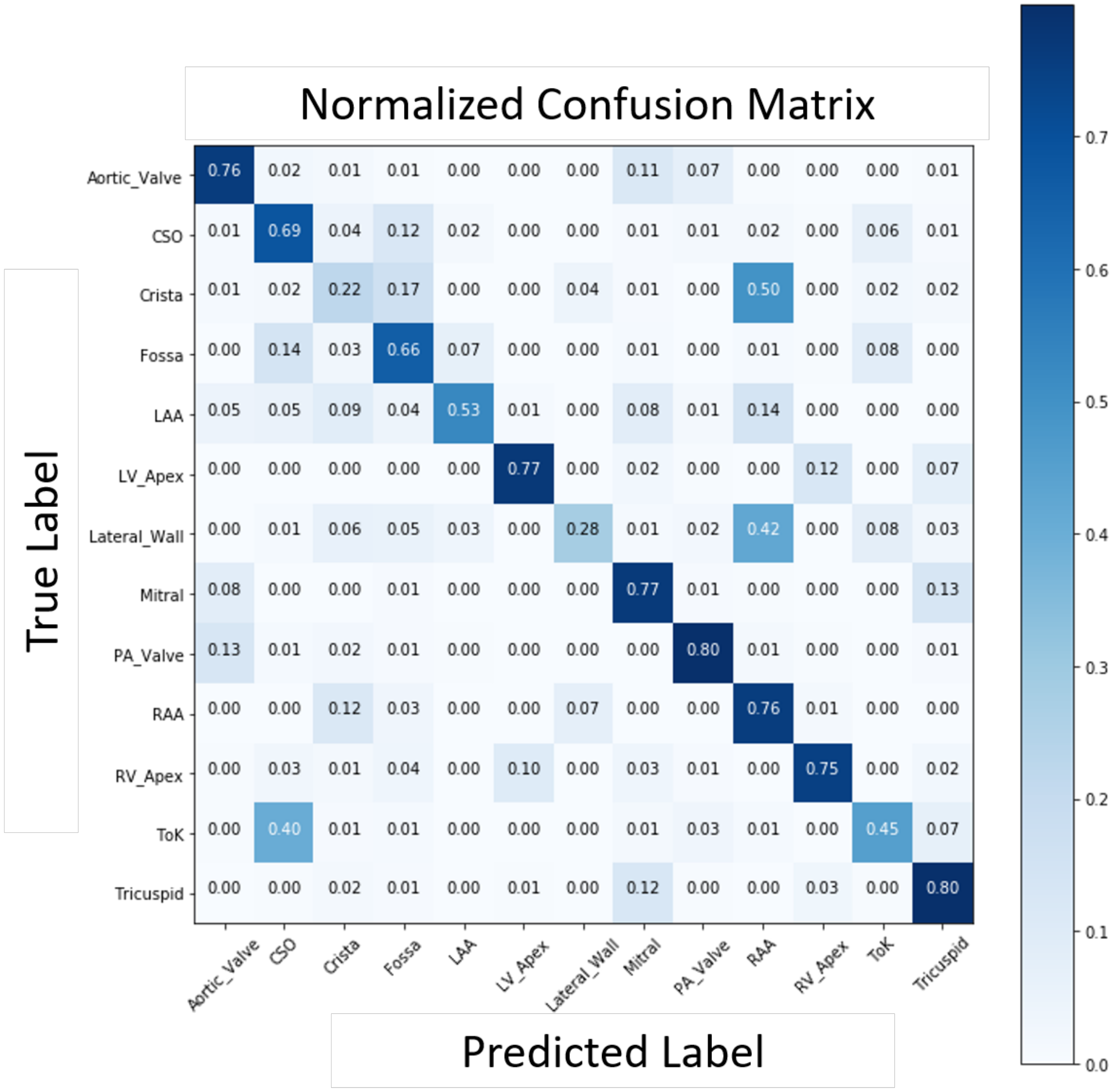


Figure 1.2: Confusion matrix demonstrating interclass accuracy. Diagonal represents correct labels.

training and validation datasets. For instance, anatomically the aortic valve and pulmonic valve have leaflet features which can be difficult even for the human eye to distinguish if not shown in context of associated structures. Yet these valves were identified quite well by the network. A similar argument can be made for identification of tricuspid and mitral valves. One of the most ambiguous classes, the triangle of Koch (an endocardial region within the right atrium), even achieved accuracy of 0.45, though being often confused with other anatomical features in the atria: i.e., the coronary sinus ostium and the tricuspid valve annulus.

Two main factors work against the enhanced performance of this network: manually labelled anatomical ambiguities and data noise. Insights as to the magnitude of the former, can be observed by determining which pairs of anatomical classes were commonly mistaken for each other by the network. Crista terminalis, for example, was frequently misidentified as right atrial appendage. This is not surprising as they lie in close proximity to one another, making it up to interpretation which anatomical feature was the primary focus of the picture. The latter source of lower network performance was likely a consequence of these images being obtained from a video source file. Heart beats can cause the videoscope to briefly look at structures that were not the main focus of the selected video. Consequently, this combined with correlated video images causes rapid overfitting on the training set, hence the need for a large amount of dropout.

Opportunities for further improvements of applied deep learning for our described application would likely come through architectural changes, although they come at the cost of training efficiency. Resnet[21], Inception[46], and Xception[15] are other candidates for transfer learning networks that could be used independently or in an ensemble for cardiac image deep learning. Further, adding data augmentation could also help against data correlation errors, but this in turn would take away the flexibility to precompute intermediate features over the whole dataset. Downsampling these images, which is necessary to input to

VGG, is also a potential error source for discrimination of some of these structures.

Going forward, further optimization could use a sequence of obtained video frames as inputs so to better inform/characterize classifications of an anatomical structures. However, this in turn would require developing a larger architecture, which may be less able to take advantage of a transfer learning model.

A limitation of our application of deep learning on this dataset may be the style of classification used which can lead to a lack of interpretability in label assignment. This can be remediated partially through the use of fully convolutional networks which can highlight salient features of the image at the cost of accuracy. Exploring the applications of these techniques relative to deep learning on this cardiac image dataset will be pursued in our future work.

1.5 Conclusion

Here we have explored applying deep learning strategies for identifying anatomical regions of internal human cardiac anatomies. This required us to treat labeled anatomies as atomic units of given semantic descriptions. In practice this is not true. Nevertheless, the impressive performance of our network demonstrates that our limitation is not in the computational or recognition abilities, but in how we initially frame the problem. In deep learning approaches like the one described here, which require greater understanding of relative human cardiac morphologies, it will be necessary to have a more fluid understanding of where the endoscopic camera was operated in 3D space. We hope, for example, this could be solved by treating localization through classification as an embedding space, an approach frequently seen in fields such as zero-shot learning[38].

Chapter 2

Deep Learning for the Segmentation and Model Generation of Cardiac Atria from Non-Contrast CT Scans

Authors:

Erik Gaasedelen[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

Following the exploration of classification in the preceding chapter, we delve into the more challenging task of segmentation, specifically focusing on binary segmentation of non-contrast CT scans. This venture stems from a collaboration with Medtronic aimed at automating the creation of 3D models for a variety of applications, from statistical shape modeling to aiding in localization efforts. Segmentation, especially in the realm of medical imaging—a field inherently three-dimensional—presents a complex challenge.

Despite its complexity, this project lays a solid groundwork for tackling many 3D segmentation issues, assuming the availability of a suitable dataset. Often, the most daunting aspect of such projects is acquiring the right dataset, a hurdle we fortunately overcame due to prior work. Subsequent challenges included data preprocessing and evaluation, underscoring the importance of robust infrastructure, a theme introduced in the context of anatomical classification. Adding functionalities for virtual reality and 3D printing was deemed crucial, topics that are further explored in later chapters.

The medical implications extend beyond the project itself, offering avenues for modeling various anatomical features such as bones, brains, and tissues across different imaging modalities. An immediate extension of this work is multi-class cardiac segmentation, which is discussed in the following chapter.

From a theoretical standpoint, the multi-axis ensembling strategy adopted here relies on a linear correlation of confidence across axes, a simple yet effective approach. However, I explored an alternative method that leverages cardiac masks from adjacent, verified slices as seeds for expanding segmentation into neighboring slices. Despite its challenges, this technique proved effective in extending the segmentation to vessels initially overlooked. This approach draws parallels with the deep learning subfield focused on attention models, which enhance the interaction of long-range dependencies, mirroring the human process of manually segmenting DICOM images.

A word of caution: overly complex methodologies can become resource-intensive and hinder timely outcomes. Our experience reaffirmed the value of maintaining simplicity to ensure prompt results.

2.1 Abstract

Computer graphics has grown to be an integral part of medical care. Artificial images which accurately reflect state of a procedure or structure of an individual's anatomy can be useful not only for preclinical screening, but for intra-surgical operation as well. Unfortunately, there is an expensive gap between clinically measurable information, such as through CT or MRI, and categorical spatial meshes which ultimately show desirable detailed data in 3D. Currently, deep learning is a field which has made many strides in semantic segmentation, but the lack of 3D modality-based strategies have limited their practical use in the medical field. Here I describe the task of atria segmentation on non-contrast CT scans using a multi-axis ensemble approach to properly address such a 3D modality problem.

2.2 Introduction

Clinical imaging tools must frequently rely on expert anatomical and physiological domain knowledge to make proper decisions. Surgical disciplines in particular require deep understanding of anatomy, even to the extent of an individual's patient anatomy. Although this kind of geometric information can be collected through medical imaging, it is difficult to relate voxel-based data to underlying tissue types and surface models.

A good application for high resolution anatomy labeling is 3D printing.[35] Over the years, its utility has increased, providing a modality for clinical and research work. However, scaling this technology is difficult since voxel labeling is still largely performed manually by humans. Consequently, many useful applications for 3D anatomical model generation are forgoed because of time and expense constraints.

Fortunately, the growing field of deep learning has many effective techniques for pixel-wise image segmentation problems. The most well-known benchmark dataset for semantic segmentation is the Pascal VOC[18] dataset which has been used to develop many effective algorithms. High performing neural network architectures are usually variants of the U-Net architecture.[43] These techniques have been universally successful in domains such as biomedical cell segmentation, autonomous vehicles, natural image manipulation, and more.

Importantly, most medical imaging needs involve processing of 3D-based data; this increases complexity in multiple respects. 3D data is usually much larger to store and the number of samples per unit storage decreases dramatically.[29][28][52] This is further compounded by the diversity of problems that emerge as a result of 3D modalities. In medical imaging, the third dimension takes units of space, while in video classification the third dimension takes units of time. In either case, correct labeling of a slice may require knowledge of slices in the prior or the latter portions of the stack. In the case of medical imaging segmentation, correct identification of vessel may require knowing the vessel connects a chamber, visible only in a different slice.

Here we approach a confined, yet challenging problem of binary, atria segmentation of non-contrast CT cardiac scans. Although its complexity is less than just identifying the names of specific arteries, which frequently requires long distance cross referencing, it still has many of the hallmarks of a difficult medical 3D modality problem as outlined above. By taking advantage of some common sense techniques for remediating the data and complexity problem, we are able to harness the benefits of standard semantic segmentation networks, runnable on most run of the mill machine learning workstations.

2.3 Methods

2.3.1 Data

Data was collected from 119 de-identified patients receiving non-contrast CT scans. 80% was used for training and 20% for validation. All data was generously provided by Medtronic. All voxel information was converted to Hounsfield units before further processing.

2.3.2 Preprocessing

DICOM scans are notoriously diverse as their size, shape, resolution, and even recorded pixel intensities are functions of the scanner, gating, operator, patient, and the pathology of interest. A naive approach for training a neural network on DICOM data is to take whatever sizes and intensities are provided by the system and use those directly to train a network. Besides the inconsistency of intensity values that are inherent in this approach, it also makes learning very difficult as the clinically useful range for most applications is very narrow for DICOM intensities.

Therefore our first step was to preprocess these values such that they match the useful range of our problem. We initially convert all values provided by the system to Hounsfield units, a standard unit in radiological care. We then clip these values such that the tissues of interest fall in between a narrower range. For example, a lead will normally appear in the range of over 1000 in Hounsfield units. This will throw off the network if we do not address it, so our clipping had the effect of placing it in the region of bone intensity. For this application we clipped our pixels to be within the range of -200 to 100. This provided the best contrast between blood and soft tissues. With a new, narrower, and consistent range of pixel intensity values, Hounsfield units could be mapped to standard 0-255 RGB pixel intensities and normalized accordingly allowing use of transfer learned networks.

The next difficulty was dealing with inconsistency of pixel resolution and spacing. Training on extensively distorted DICOM scans could be okay, but the small number of scans caused by the small sample size problem makes further generalizability a game of chance. This could be remediated by resizing with respect to inter-voxel distance, not number of pixels, which is standard practice for deep learning image tasks. An important feature of this was that proper resizing the the z-direction (head to toe) must not be performed based on z voxel size but rather on inter-slice distance. Once these parameters were taken into account, consistently shaped tensors were the output and fed to a network. For our experiments we used a pixel spacing of 0.75 millimeters in all directions.

It was also important for us to maximize the amount of data available for training and evaluation. One strategy was to segment slices that were present in the axial plane. This was not ideal because, as mentioned above, but it was considered important to have information from other slices when correctly identifying the atrial. We therefore produced data on the axial, transverse, and sagittal planes and trained one model that could work on all of them. During evaluation we use predictions from all slices to make a final prediction.

2.3.3 Architecture

The architecture used was a standard U-Net implemented by the FastAI library. A U-Net architecture, in general, combines an encoder and a decoder to provide a mapping between input image and an identically resolved segmentation mask. Originally this task was performed by a convolution-deconvolution architecture.[5] The convolution portion acted as an encoder while the deconvolution portion acts as a decoder. The encoder functioned similarly to any VGG-like[44] architecture, producing an output just prior to flattening and fully connected layers. The decoder, meanwhile took the form of transposed convolutions, converting the intermediate representation back into a native sizing. This functioned poorly until the introduction of U-Net whose principal innovation was to provide skip connections

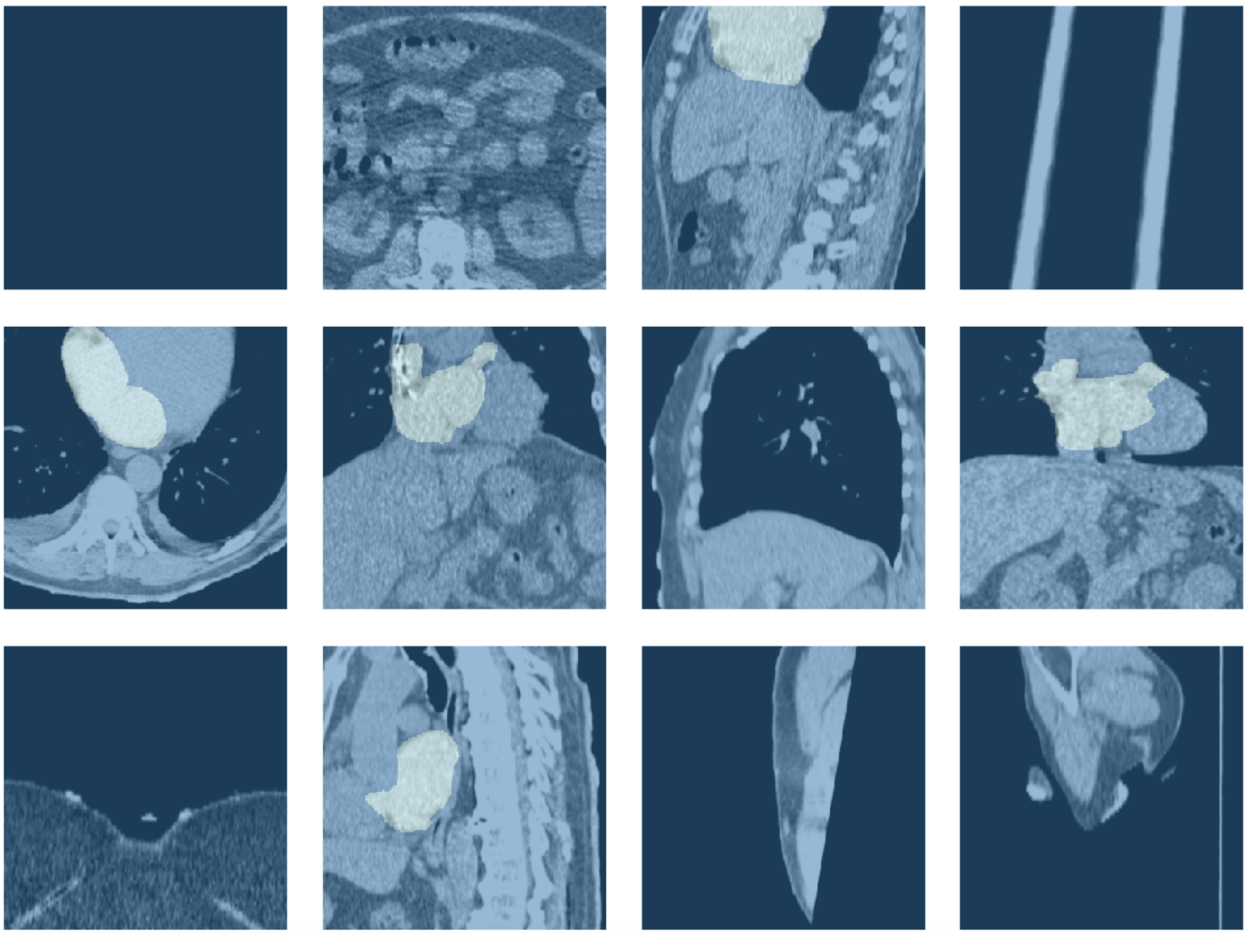


Figure 2.1: Examples from a batch showing images where the atria are highlighted. Note most of the images do not contain atria since most of the viewable volume is background.

between the encoder and decoder, allowing the network to retrieve spatial information during the decoding phase.

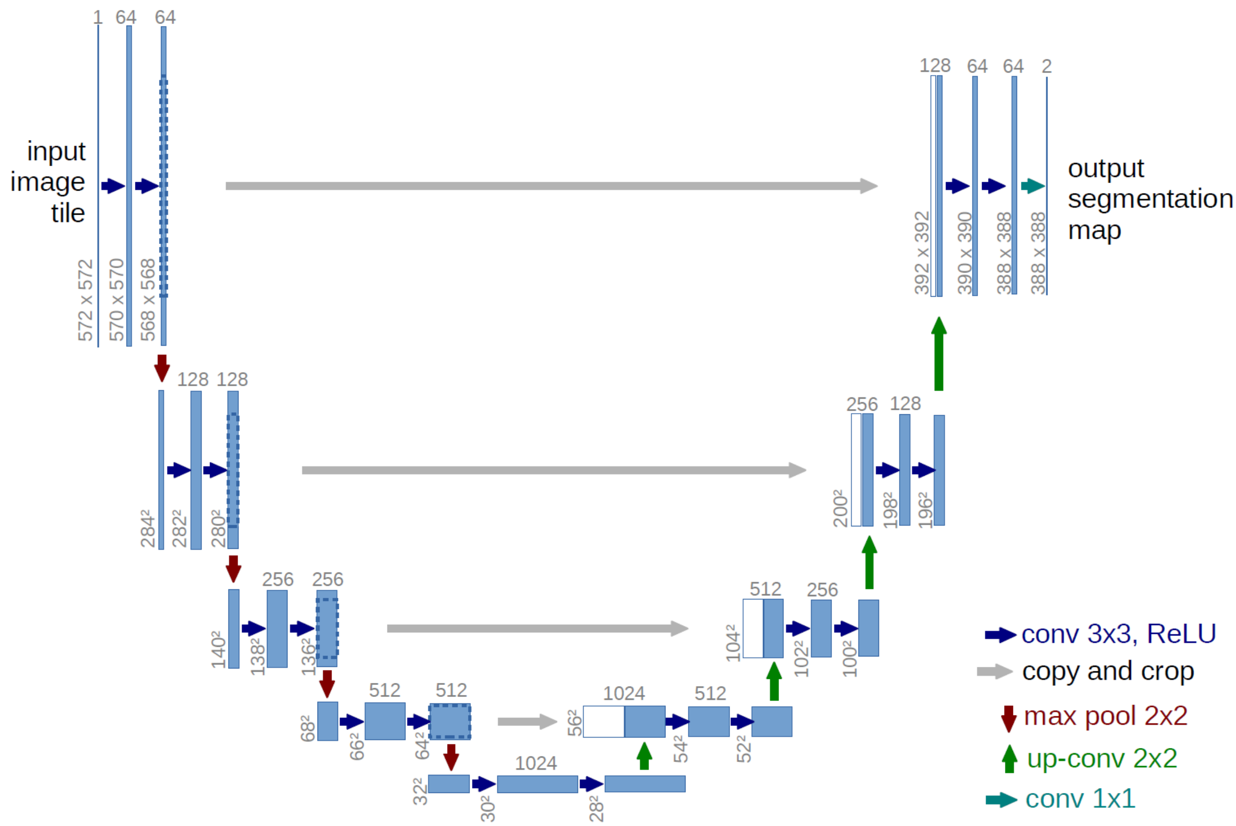


Figure 2.2: U-Net Architecture

A nice feature of a U-Net architecture was that its encoder can be a pre-trained neural network which helps to speed up training considerably. In our case we take the simplest and most manageable network, Resnet-34.[21]Other networks like Resnet-50 and Resnet-152 have higher capacity and would have higher performance potential at the cost of increased memory and training time requirements. We choose Resnet-34 for a couple reasons. First we were constrained by GPU memory and training time. Second, Resnet-34 is a fast network and allows for fast evaluation. Its speed and interoperability also mean that training was well within the capacity for most low-end deep learning machines, allowing for more prototyping accessibility.

The decoder network performs a smaller number of convolutions during the upsample, but has a principal role of combining information from lower areas of the encoder into higher layers.

2.3.4 Training

Image augmentation is an important facet of any successfully trained network. Typical transformations include resizing functionality, usually taking an input image and stretching it so that it fits within a window of interest. This was not desirable for our problem since each slice can have a drastically different size and the slices frame size in millimeters can disrupt the careful steps taken during preprocessing to ensure a consistent sizing. Therefore it was important to instead take consistent crops during training instead.

During training we pick a specific crop size with padding from our preprocessed images. From here we could allow other image transformations, these included: contrast, noise, slight rotations, slight affine transformations, but not horizontal or vertical flips. Since a given heart is not symmetric across the transverse axis, this was useful domain knowledge that we want to provide for the model during training. Here we choose 320 pixel crop sizes which winds up being 24 cm x 24 cm in any slice.

Our training regiment follows a one cycle learning policy,[45] training for four 4 epochs using frozen pretrained weights, followed by 3 epochs using unfrozen weights, allowing exponentially higher weights toward the output of the network. Our loss was simply categorical cross entropy.

2.3.5 Evaluation

For evaluation we have the benefit of evaluation on the entire image since we do not need to calculate the backward pass and we could tune the batchsize to use less resources. Our preprocessing was fortuitous here because our pixel spacing was still 0.75 mm for each direction for every voxel. This happens to be a nice feature of fully convolutional networks like U-Net, that they can be used on different sized images without hindrance to performance.

One important thing to note is that image size in native resolution is not necessarily divisible by 32, a requirement for U-Net. To get around this we simply resize the images to the nearest factor of 32 in log scale. We then resize the output segmentation back to native size.

The next important part was consolidating results from each plane into one final 3D segmentation mask. Evaluation proceeds on axial, saggital, and coronal slices, which, after resizing, can be average together into a final confidence threshold. This is reminiscent of using one model as an ensemble when making predictions. As shown below, we have found this to work better than any axis in isolation.

Our last step was to take the resulting 3D mask and convert it to a 3D mesh. For this it was as simple as using the marching cubes algorithm provided by the Scikit Image library. Here we could also specify the pixel spacing as 0.75, allowing us to output an STL file with native sizing.

2.4 Results

For our metric, we choose to use the intersection over union (IoU) or Jaccard index. This was a metric whose score was maximized at one if the intersection of two sets and was identical to the union of the sets. Any points which were false positives or false negatives penalize this metric.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

Figure 2.3: Jaccard Metric

Accuracy was not a good metric to use since most pixels in an image did not belong to the DICOM stack: they were background pixels. At the end of training we were able to achieve an ensembled mean IoU of 0.852 over 22 studies held out as a validation set. This was a reasonably good score in its own right, but it is important to note that the choices on where the mask goes changes depending on the judgement of the person doing the segmentation. For example, the pulmonary veins, or the superior vena cava could be cut short based on user preferences.

Sometimes the best assessment of performance was just to visually inspect a random output example. Here is one:

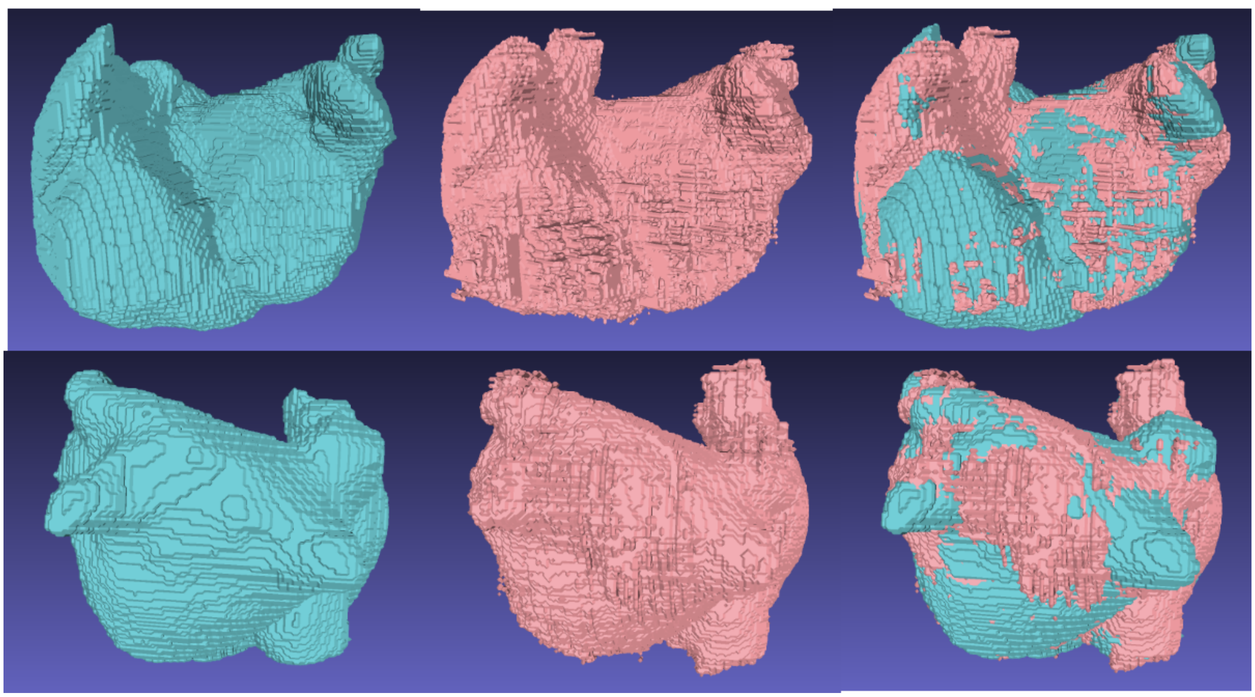


Figure 2.4: Example 3D model generated from a validation sample. First column - ground truth. Second column - prediction. Third column - overlay. First row - Anterior view of atria. Second row - Posterior view of atria.

In general, the computationally generated models were very similar in their construction. It was difficult to tell the difference between the human annotated ones and the computer segmented ones. An important facet of this problem was that it is a non-contrast scan and the borders were not necessarily well defined. This shows a limitation in the employed dataset, not the segmentation algorithm. With that said, in the future, it is still important to show that a segmentation algorithm can also work on a problem with better resolved features. We will leave this to later work.

Do notice, however, that there was disagreement as to the amount of protrusion of the pulmonary veins. This was especially difficult for the network as choosing the cutoff was a bit arbitrary between scans. Nevertheless, additional training on larger crops could prove beneficial for this important clinical parameter.

2.5 Conclusions

Here I have demonstrated a proof of concept for the automatic segmentation and model generation of non-contrast enhanced CT scans of cardiac atria. In the interests of time, there are a few things we have yet to try; one is simply using larger models. This should likely be performed on computers with more graphics cards, more compute time, and with larger crops. I suspect limited returns using this approach, but some improvement would be expected.

Additional next steps would be to investigate harder problems, like those which require multi-class labeling instead of simply binary class. New architectures would also be worth exploring.[30] Attention mechanisms have recently been a possible addition to segmentation architectures to boost performance.[37] This would, however, need to be made compatible with 3D DICOM modalities.

The most important next step, however, is probably simply in the area of infrastructure development, allowing these models to be used more fluidly with APIs designed for more specific use cases.

Chapter 3

Automated Multiclass Cardiac Volume Segmentation and Model Generation

Authors:

Erik Gaasedelen[1], Alex Deakyne[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

This chapter marks a significant advancement in our autosegmentation efforts, moving beyond the binary classification of atrial tissues to a more comprehensive segmentation that encompasses all four heart chambers. This expansion translates into a five-class segmentation challenge, encompassing the left atrium, right atrium, left ventricle, right ventricle, and the background. A notable shift in this phase of research is the use of contrast-enhanced CT scans, which afford a superior resolution compared to the modalities previously employed, enhancing the accuracy and detail of our segmentation results.

The development of an improved methodology for addressing the intricacies of 3D data representation significantly bolstered our segmentation capabilities. This innovation not only advanced our current project but also retroactively enhanced the methodologies outlined in the preceding chapter, showcasing the iterative nature of scientific inquiry and technological advancement.

While the segmentation of the heart's four chambers serves as a robust foundation, the field of anatomical segmentation presents a much broader spectrum of complexity. Drawing parallels with the CityScapes dataset, which challenges segmentation algorithms with thirty diverse classes ranging from urban infrastructure to pedestrians, it becomes evident that scalability and performance in segmentation tasks diminish as the number of classes increases. This limitation underscores the necessity for advanced architectures beyond the conventional U-Net framework to maintain high performance across a wider array of classes. As we move forward, the exploration of these advanced techniques becomes imperative. I advocate for extending our segmentation efforts to encompass an even greater variety of anatomical features, employing innovative architectures to overcome the constraints observed in traditional models. This direction not only promises to refine our understanding and capabilities within autosegmentation but also sets the stage for future breakthroughs in the field.

3.1 Abstract

Many strides have been made in semantic segmentation of multiple classes within an image. This has been largely due to the advancements in deep learning and convolutional neural networks (CNNs). The features within a CNN are automatically learned during training, which allows for the abstraction of semantic information within the images. These networks are powerful enough to handle the segmentation of multiple classes without the need for mul-

tiple networks. Despite these advancements, few attempts have been made to automatically segment multiple anatomical features within medical imaging datasets obtained from CT or MRI scans. This offers a unique challenge since these datasets contain three dimensional data. We propose a pipeline that not only performs multiple class segmentation of the 4 cardiac chambers, but also ensembles the information from all three axis to ensure no spatial information is withheld from the network. This technique has increased the mean IOU of our predictions over the entire DICOM dataset. (don't like how this is said and whole section needs work on fluidity).

3.2 Introduction

3.2.1 Anatomical Segmentation

Medical Imaging datasets, such as DICOM image datasets produced from CT or MRI scans, contain a wealth of information about a patient and their anatomy. This information is invaluable to medical device companies and physicians alike. However, this information, especially when dealing with complex anatomical features, is not easily digested. DICOM datasets are two dimensional images that are stacked to create a three dimensional representation of the patient anatomy. Analyzing a DICOM dataset is traditionally done by sweeping through these two dimensional slices in order to get an idea of the three dimensional anatomy. This has lead to the development of tools that can be used to segment anatomical features of interest within dicom scans. All three axis of the scan can be viewed simultaneously to provide the user with all the spatial information in the scan while segmenting. Segmentations not only highlight the anatomical feature, but certain software packages allow for the creation of 3D models from the 2D segmentations. An example of such a pipeline is depicted in Figure 1 below. These models have wide use in both pre-surgical planning and medical device design. They offer a clear representation of the anatomical feature while retaining all the spatial information in the scan. These models can be easily used to generate anatomical measurements and generate statistics about a certain patient population.

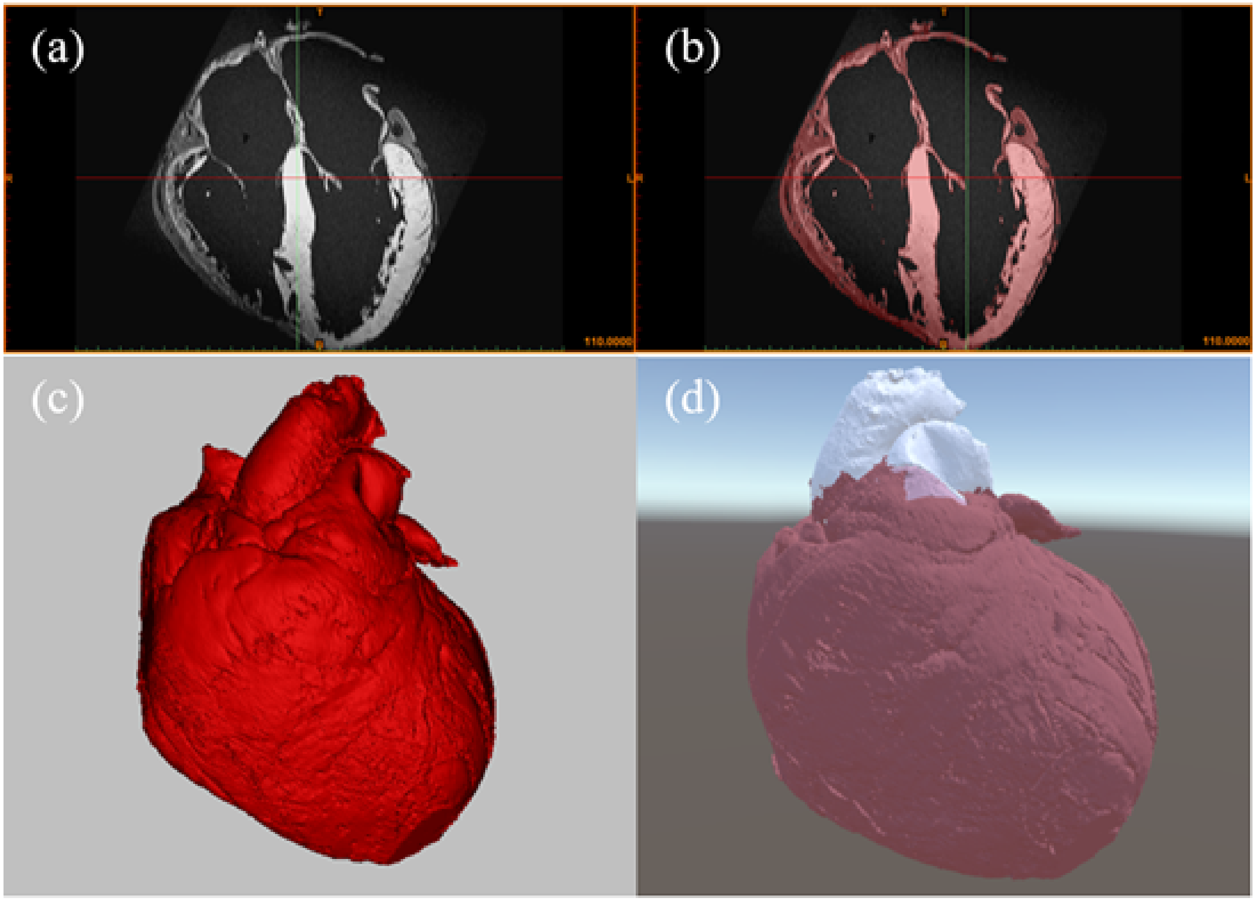


Figure 3.1: Flowchart diagram of proposed mixed reality methodology. (a) A DICOM scan in MIMICS depicting a cross sectional view of a heart. (b) 2D mask of the heart tissue generated through thresholding functions. (c) The resulting 3D model of the heart created by the 2D masks which can then be 3D printed. (d) The same 3D model of the heart is imported into Unity and a virtual reality environment of this given heart is generated.

Although anatomical segmentations are becoming an invaluable information source in the medical field, they are incredibly difficult to produce. Interpretation of simple anatomical features within a DICOM dataset takes time and practice. Often, there are artifacts present in the scans and the distinction between anatomical features can be nebulous. It takes an expert in the relevant anatomies to produce quality segmentations. Also, there are few anatomical segmentation tools that work well, and their availability is gated behind incredibly expensive licensing fees. These restrictions have granted few people with the necessary skill set to accurately interpret DICOM scans and produce quality anatomical models.

Unfortunately, anatomical segmentation is a time consuming process, even for highly skilled individuals. This creates a prohibitive bottleneck in many workflows that wish to utilize anatomical segmentations and 3D models. For surgical planning, a physician must receive a segmented model of the patient anatomy between the pre-procedure scan and the start of the given procedure. To date, this is often logistically impossible which in turn could be considered to restrict the quality of care available to the patient. Further, in the medical device industry, the anatomy of a patient population must be well understood in order to make informed decisions about how to design a new device. This often requires the production of hundreds of anatomical segmentations, which are then used to make measurements that are used to model the statistical information of the patient population. This can be such an arduous process that many companies are forced to make their decisions on fewer segmentations than they would like.

3.2.2 Deep Learning

Deep learning has established itself as the current state of the art for supervised image segmentation. The application emerged by combining traditional classification networks with a deconvolutional layer set. While initially ineffective, the addition of skip connection between downsampling and upsampling passes in an architecture called U-Net significantly improved upon the conv-deconv model, producing state of the art results on diverse datasets.[43]

The skip-connection discovery closely parallels the development of residual connections in traditional classification networks. As demonstrated in research on resnet and densenet architectures, these strategies not only allow for deeper networks without vanishing gradients, but significantly speed up model convergence and improve model accuracy.[21][24]

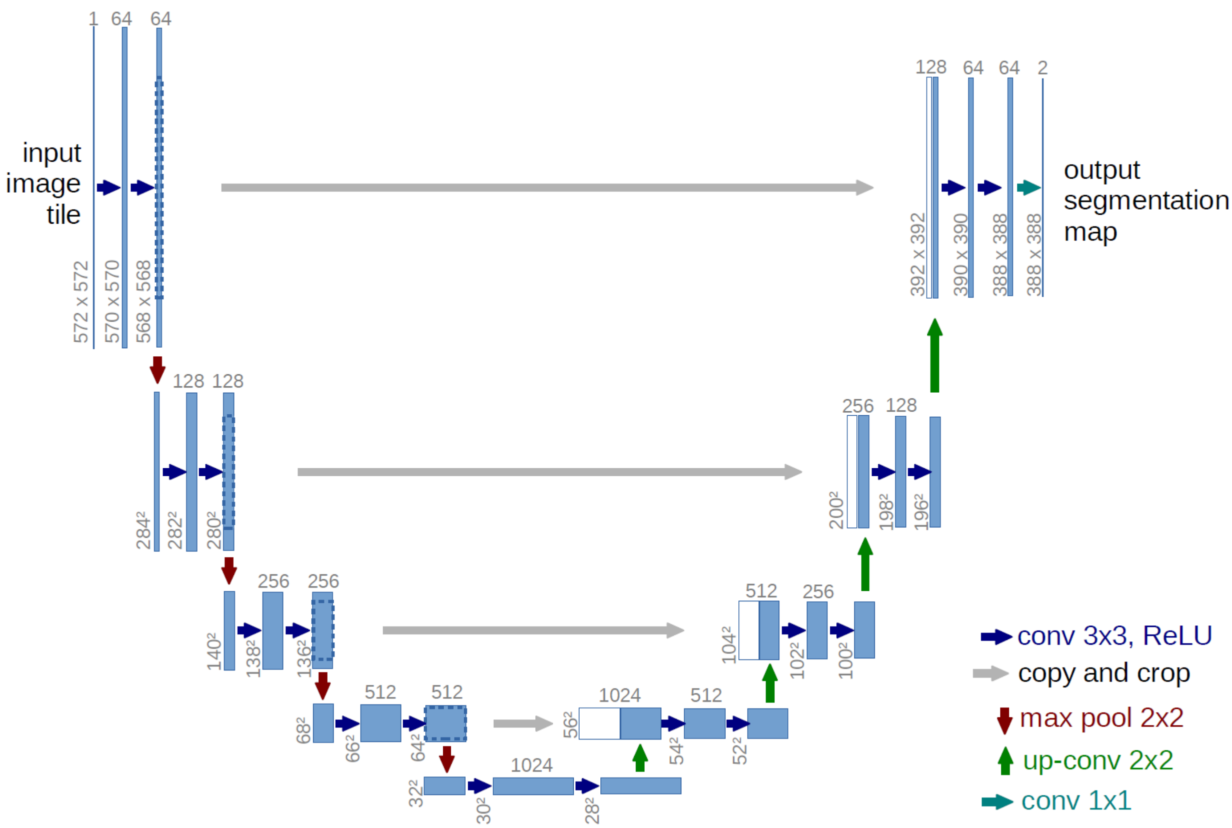


Figure 3.2: U-Net Architecture

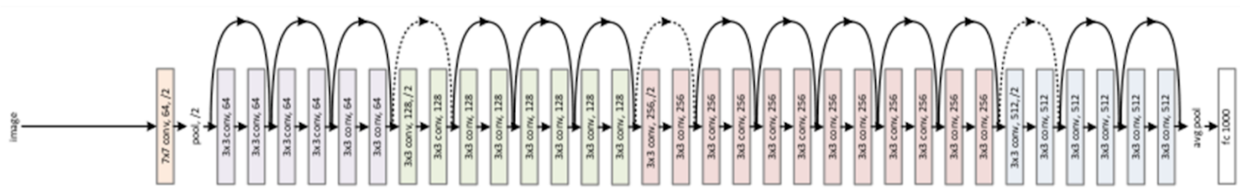


Figure 3.3: Resnet-34 architecture

For binary segmentation, the standard U-Net architecture works well for most domains of interest. For multi-class segmentation, state of the art models resort to some additional modifications. These models are usually variants of the feature pyramid network architecture. [32] The benefit of feature pyramid networks usually becomes apparent as the number of classes increases. Unfortunately, feature pyramids require more GPU memory, are slower to train, and typically produce outputs that can be significantly downsampled compared to the original image. It is unclear why U-Nets and feature pyramid networks have such a performance discrepancy in the multi-class domain.

3.2.3 Autoseg using deep learning

Medical imaging problems commonly require that one extends the 2D segmentation problem to a 3D space. Considering just the desired output of a 3D segmentation problem, it is natural to consider the utilization of a 3D U-Net by simply replacing two-dimensional layers with corresponding three-dimensional layers.[52] The downsides of this strategy is that it is prohibitively expensive for GPU memory and limits the usability of transfer learning for lack of pretrained models on 3D datasets.

Performing a multi-class task on modest hardware resources is not without hope, however. For manual segmentation, software packages provide views of the axial, coronal, and sagittal axes. This provides users with enough information to create good segmentation masks, even though it may take significant time to cross check all planes and neighboring depths.

The dataset we analyzed here included 46 contrast-enhanced cardiac CT scans with each of the four chambers of the heart segmented. Normally 3D medical imaging datasets comprise of MRI scans which usually have much better resolution and are safer to acquire. While MRI uses a magnetic field to construct an image through hydrogen spin resonance, CT technology uses x-ray radiation to determine at what depths radiation permeates and scatters. Since bones and radio-opaque contrast scatter x-rays more, their pixel values are

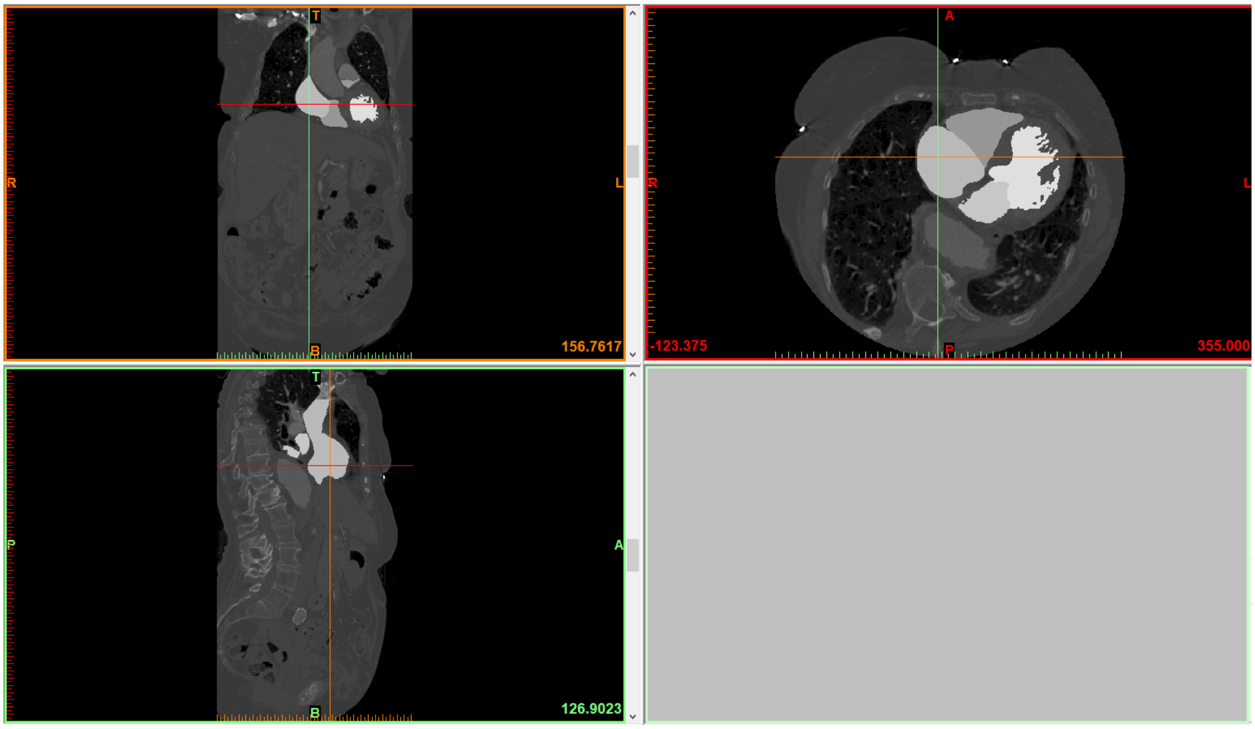


Figure 3.4: Screenshot of manual segmentation software, Mimics, developed by Materialise

brighter in the resulting CT image.

CT scans were then recorded on the Hounsfield scale, represented by an integer between -1000 and 30000. For pretrained neural networks, images were usually normalized after converting 8 bit (0 through 255 valued) images to be between zero and one. Pixel values between -1000 and 3000 have a much larger range and therefore must be addressed to make easily compatible with pretrained models.[10]

Here we attempted to address all of these difficult problems while retaining the speed, efficiencies, and pretrained weights of a traditional U-Net constructions.We found that through simple domain guidance, pragmatic dataset construction, and axis averaging, we achieved highly accurate results, with the abilities to correctly segment and label minute details of cardiac anatomies.

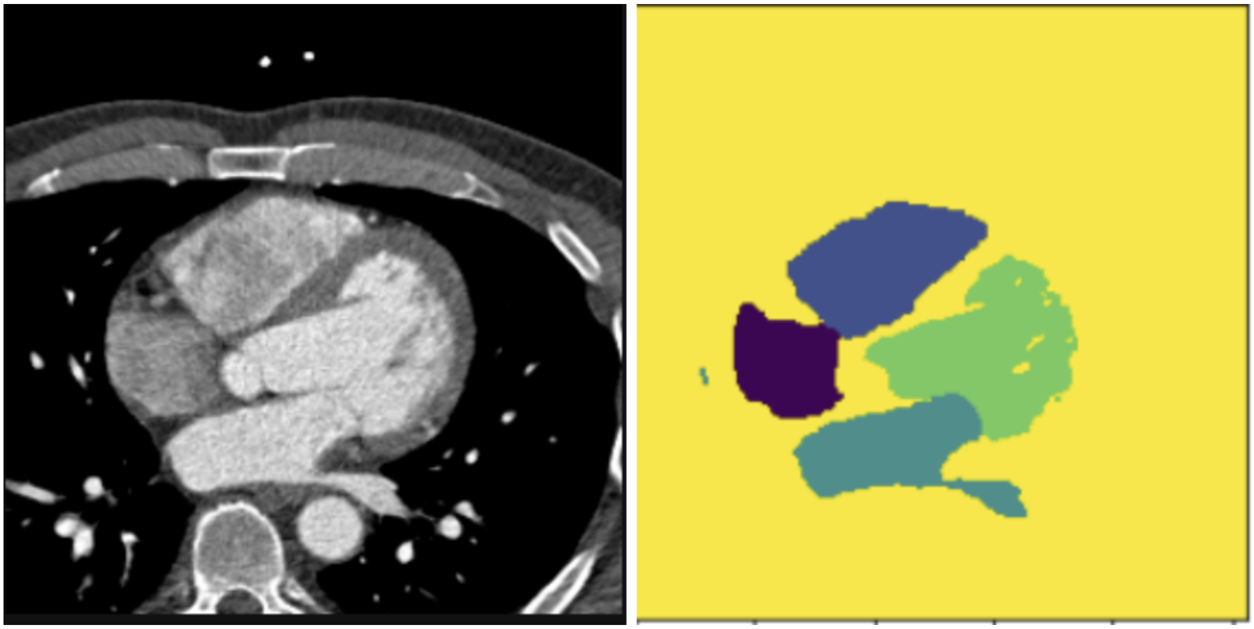


Figure 3.5: Dataset Example

3.3 Methods

3.3.1 Dicom Preprocessing

First we must address the problem of the large spread of the Hounsfield scale. The purpose of consistent data normalization in a transfer learning scenario is not necessarily to match color distributions of the previously trained imageset, but rather to allow color perception to be consistent between datasets. This analogy becomes difficult in the case of a CT scan where its brightness units are most analogous to grayscale color values. Fortunately, we do not need the complete spread of Hounsfield units; only a small range of brightness values that would be most important for distinguishing the volumes of interest. We therefore clipped the Hounsfield units between -200 and 500, before allowing that range to take the values between 0 and 1 prior to normalization. These values roughly correspond to airy tissue like lung (-200) up to cortical bone (500), which places contrast dye nicely between these values.

Such a step is very important as it counteracts under or over saturation in the images which would be difficult for a pretrained network to overcome.

For ease of processing we also converted all DICOMs and masks to PNG files. Masks were produced simply by assigning the pixel values in each slice to 0, 1, 2, 3, and 4 corresponding to the right atrium, right ventricle, left atrium, left ventricle, and background respectively.

3.3.2 Architecture

For our architecture we choose a U-Net. The U-Net used a pretrained resnet-34 network as the encoder, as implemented in the Fast.ai library. We choose resnet-34 out of resource considerations. We only had one Nvidia 1080ti for training. Although, with more computing resources, a resnet-50 or feature pyramid networks could have an added benefit. These considerations are left for the discussion

3.3.3 Training

For training we used the one-cycle policy which is included in the fastai library.[7] We first trained the network with frozen weights for 10 epochs. We then lowered the learning rate and continued training for 10 epochs with unfrozen weights. For image augmentations we included ten degree rotations, zooms, lighting changes, and random crops at 256x256. To do this we sampled from our dataset on all axes including axial, coronal, and sagittal slices. This led to some distortions as a result of inconsistent voxel sizes, however it did not appear to harm the results of our model. We used cross entropy loss for our optimization objective.

3.3.4 Evaluation

For each study we evaluated each slice on all axes. Before running them through the network, however, we enforced a maximum axial dimension of 256x256 by resizing and interpolating

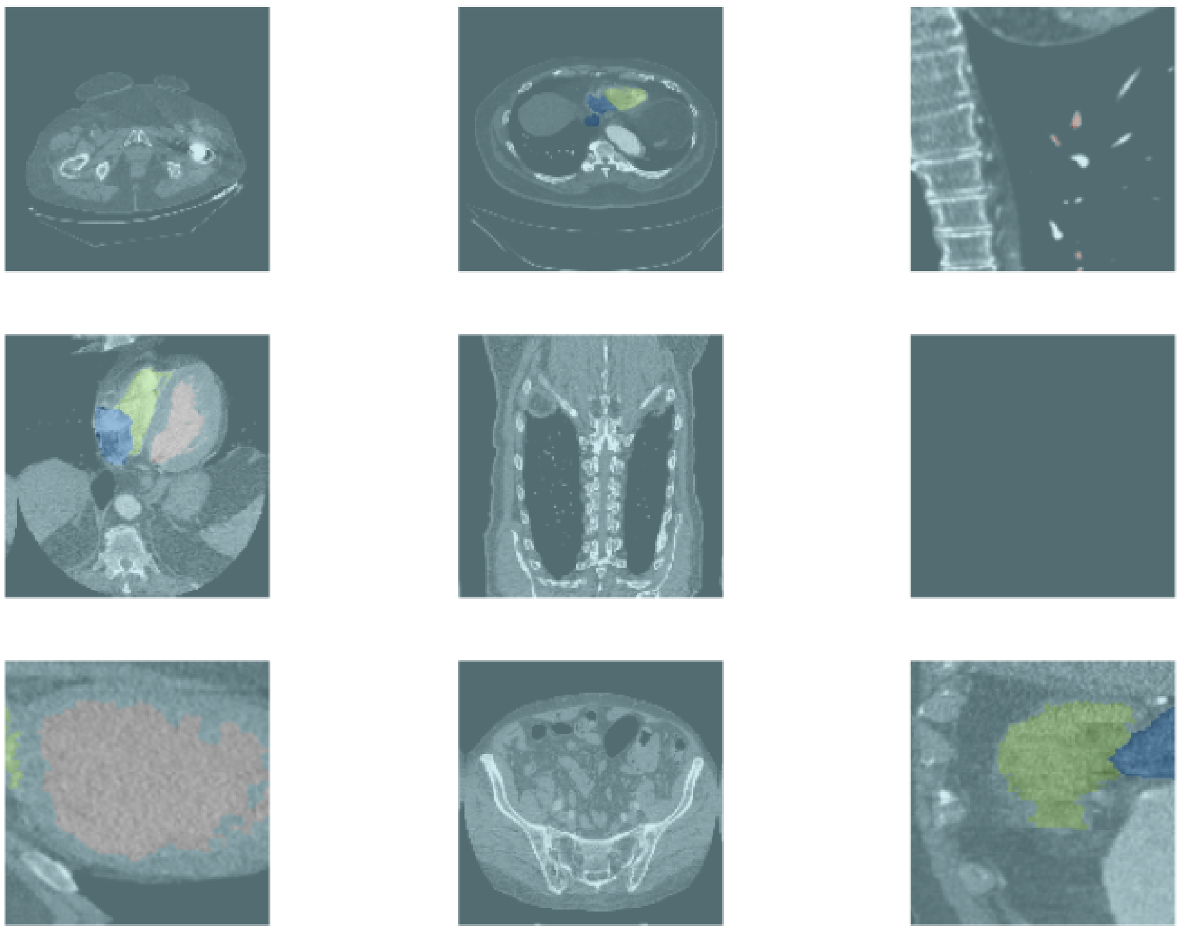


Figure 3.6: Example Augmented Data

the pixels. For the depth axis we resized the number of slices in proportion to the axial resize. This was followed by a slight resize to ensure each dimension was divisible by 32 to ensure proper divisibility of shape throughout the network. Once all slices from each dimension were evaluated, the resulting activations were averaged between all planes. This output was then used to determine the label for every pixel in the 3D tensor, and then subsequently resized back to its original shape before computing metrics.

3.3.5 3D Model Generation

We used the marching cubes implementation in the scikit-image library to create a surface mesh for each chamber of the heart. These meshes were then used to 3D print select examples. The DICOM metainformation for pixel spacing and distance between slices was also used to output meshes that were of consistent sizing to the patient's native anatomy.

3.4 Results

Our primary method for achieving highly accurate results was through our multi-axis ensembling approach. When we allowed each axis to average into the final output we achieved a mean intersection over union of 0.857. To compare performance without such averaging, we also calculated the same metric between axial, coronal, and sagittal exclusively. None of the axes performed as well in isolation with the axial axis achieving the next best score with a 0.851 mIOU.

Table 3.1: Comparison of mIOU for each chamber and ensemble

Method	**Right Atrium**	**Right Ventricle**	**Left Atrium**	**Left Ventricle**	**Full Heart**
Ensemble	0.835741	**0.829308**	0.877934	**0.886810**	**0.857448**
Axial	**0.837558**	0.812199	**0.879793**	0.873627	0.850794
Coronal	0.803602	0.778569	0.836553	0.859823	0.819637
Sagittal	0.722848	0.784649	0.835081	0.861965	0.801136

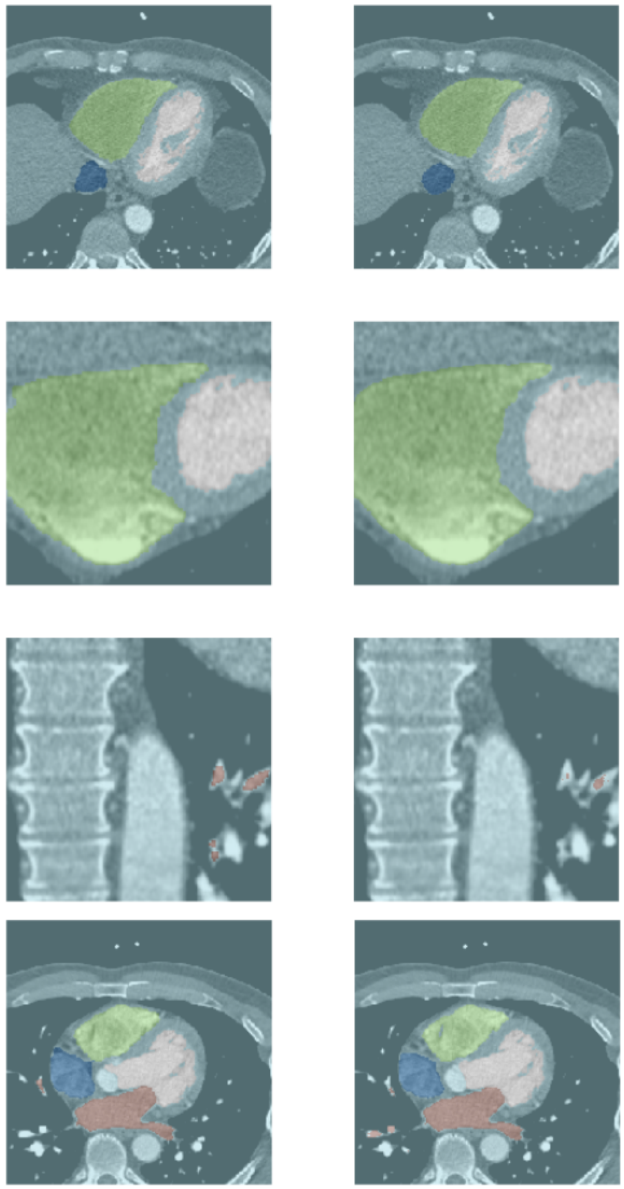


Figure 3.7: Collage of input scan, predicted mask, and ground truth mask

Figure 3.8: 3D Model of network segmented heart. Right atrium (upper left), left atrium (upper right), right ventricle (lower left), and left ventricle (lower right).

Each chamber performed with different average mIOUs throughout our validation dataset. Although ensemble performed better on the entire heart, it's performance degraded only slightly when compared to the axial view for ventricles. The ventricular chambers typically performed better than the atrial chambers. This is probably because the ventricles are better defined when it comes to where its boundaries are delineated. For atria, decisions must be had as to where they end with respect to venous inlets. The left atrium, for instance, is connected to many branching pulmonary veins. For a human segmenter, it can be ambiguous as to where these veins should no longer be segmented.

3D model inspection, however, is the most clear when it comes to showing the minute details our model is able to pick out of the DICOMs. The 3D models we generated were able to clearly show fine details of such structures of the left atrial appendage, semilunar valves, and the extensive trabeculae and papillary muscles that occupy the free wall of the left ventricle.

3.5 Discussion/Conclusion

Creating an unambiguous metric for a task like anatomical segmentation can be difficult. Currently, a human manual segmenter needs to make many decisions as to where the boundaries of a chamber are. For example, superior and inferior vena cava can be inconsistently segmented as well as the pulmonary veins. It can also be difficult to decide when to segment over what looks like muscular bands within the heart, especially when image quality varies between scan. Therefore using human segments as a substrate for objective minimization and metric evaluation can be tricky. Nevertheless, neural networks are robust to errors and variation that are contained within datasets. Although we achieved results that matched very well with human generated masks, ultimately our performance must be taken into context of visual inspections and whatever applications are appropriate: e.g., for anatomical education these are highly useful today, but for pre-clinical planning further evaluation is needed..

It should be noted, that a major consideration that went into our model design was the resource limitations of our computer. A U-Net with a resnet-34 backbone was more than sufficient for creating highly accurate results, but there are some opportunities for improvement. Larger encoder networks such as resnet-50, resnet-101, or squeeze excitation networks could give a boost to the accuracy of our model.[23]

Feature pyramid networks are also promising substitutes for U-Nets. In our problem description we only needed to care about five unique classes. For anatomical datasets with ten or more classes, we suspect feature pyramid networks would begin to see benefits that justify the acquisition of more compute power.

Cardiac segmentation is a great example where macroscopic and fine details must be integrated together to get a proper understanding of the underlying anatomies. While the left ventricle is large, its internal structure is composed of delicate muscle bands which can be an important factor for medical device design and surgical procedures. A further challenge will be datasets where classes are imbalanced between very large structures and very small structures. This is where tricks like feature pyramid networks and class specific sampling frequency will be especially important.

The applications for scalable and fast anatomical segmentation are vast. Presurgical planning and human centered medical device design just scratch these surfaces. Opportunities for large simulation datasets, live surgical visualizations, statistical shape modeling, and mesh-based automated diagnosis all become approachable when such systems become readily available. Important next steps will be to improve our model accuracy with more computational resources, expand our dataset to additional organs of interest, investigate opportunities for rapid visualization and analytical investigations of generated models, and develop additional tools to detect or classify different anatomical regions of interest in a more diagnostic mode.

Part II

Virtual Reality in the Visible Heart Laboratories

Having covered the deep learning segment, we now shift our focus to the realm of virtual reality (VR) and its applicability within our laboratory and the broader domain of cardiac science. This section builds upon the foundation established by its predecessor, proposing that VR could significantly benefit clinical practice. However, the efficacy of VR is contingent upon the achievements of the prior sections. Our discussion will pivot away from deep learning to consider the intrinsic value of VR, necessitating a bit of creative thinking. Unlike game development, where a single product is distributed widely, medical applications of VR require personalized products for each patient.

Despite these challenges, VR has proven to be exceptionally beneficial in our lab, particularly for educational purposes. We have been able to effectively teach students about cardiac anatomy using a limited set of models. While scaling clinical cases may be impractical without the aid of automation tools, we still dedicate efforts to develop models for specific cases. This approach aligns with the concept of first-order utility introduced earlier: education is feasible, preclinical planning is challenging, and live surgical guidance currently remains beyond reach without automation. Although the upcoming sections will primarily focus on educational applications, they will not overlook the significance of advanced applications. Moreover, I encourage readers to venture into this evolving field. If I were to do my PhD over starting today, this is where I would begin.

Chapter 4

3D Graphics to Virtual Reality in Medicine - Opportunities and Prospective

Authors:

Erik Gaasedelen[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

This chapter was contributed to the "Engineering in Medicine" book published by the Institute of Engineering in Medicine. It primarily focuses on Virtual Reality (VR) but also provides a historical overview of 3D graphics in medical applications. Understanding the evolution of 3D graphics is crucial for evaluating VR's utility in medicine, given the fluctuating levels of enthusiasm that typically accompany emerging technologies. Distinguishing between genuine innovation and mere hype becomes easier with this historical context.

The narrative of technological advancement, including VR, often follows a so-called "hype cycle," characterized by phases of exaggerated expectations, followed by disillusionment, and eventually, realistic appreciation. While this model may serve investors well, it can be somewhat limiting for innovators. Technologies, in essence, harbor the potential for creative evolution. For example, a book's intrinsic value lies not in the object itself but in its ability to facilitate the transfer of information. This principle holds true for other mediums such as films, television shows, and video games, which should not be underestimated simply because they are established forms of media.

The critical takeaway from this chapter is the recurring theme of innovation beyond technological constraints. Historical advancements in 3D graphics for medicine, among other fields, illustrate that practical limitations of a particular era do not necessarily restrict creativity. Previous methods and media can inspire or be adapted into novel, creative projects. Consider the simulations developed for the Manhattan Project and imagine applying those techniques with contemporary VR technology. Such a juxtaposition offers a unique opportunity to create visually captivating and educational content that leverages past innovations, demonstrating the timeless value of integrating historical perspectives with modern technological capabilities.

4.1 Abstract

Today, in medicine, performing a procedure on a human requires complex anatomical and physiological knowledge, including 3D relational and physical properties. All commonly employed imaging modalities attempt to recapitulate this information in some useable form. Each technique, however, can only retrieve a small slice of spatial, temporal, or conceptual knowledge of the body area of interest. Furthermore, in some cases, like computed tomography, a large amount of information is gathered without a clear strategy for man-

ual or automated analysis. Current software developments are slowly seeing a convergence of medical imaging methodologies toward one of greater realism. For our purposes of this chapter, this coming together of disciplines will be delineated as the field of 3D graphics. Although this is a simplification, 3D graphics will be the primary grounding for a much larger hybridized field whose aim will be the creation of informative and accurate medical spatial representations. To understand how this will affect the clinical fields, we will need to first explore 3D graphics more holistically, delving into areas of physical simulation, 3D movie rendering, video game development, and the expansion of 3D visualization hardware. As we explore these fields, we will relate each of these subdomains to analogous examples in the medical field, also documenting opportunities for commercialization and innovation.

4.2 Introduction

The first developments in computer visualization were primitive applications developed by some of the earliest computer pioneers. In 1957, an engineer named Russell Kirsch, first demonstrated the ability to scan a photo and save it to one of the first programmable computers[31]. Photo storage is currently so ubiquitous that it is taken for granted, but at the time it set the foundation for computerized image representation. Three years later William Fetter, a Boeing engineer working on airplane ergonomics, showed wireframe pilots in the cockpit of an airplane[12]. This was the first instance of 3D graphics being used for a commercial purpose.

It was not until ten years later, in the 1970s, when wireframe graphics began to be implemented and utilized within the more mainstream movie industry. Although first implemented in the 1976 movie Futureworld, it was the release of the 1977 movie Star Wars which brought such technological innovation to a worldwide audience. Consequently, the George Lucas Company, Industrial Light and Magic, paved the way for animation innovations in

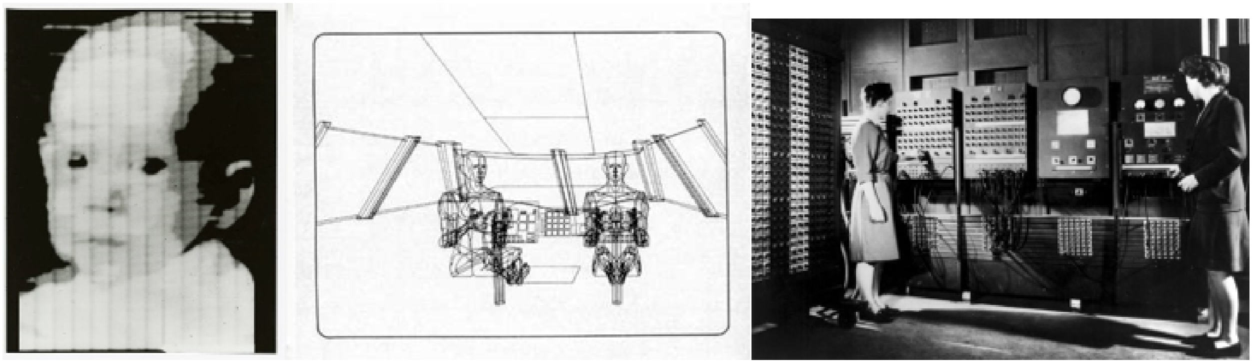


Figure 4.1: First images stored in a computer, first use of vector graphics, and ENIAC.

the movie industry; these technological advances would later be applied to medicine.

While the computational resources of the entertainment industry slowly grew, physicists and natural scientists were trying to make sense about the world around them. The complexity and lack of closed form solution to many important physical phenomena led naturally to the utilization of simulation. Therefore, the value for more advanced computers became apparent in both research and politics, with some of the first major computer advances stemming from the Manhattan project. For example, to solve their simulation problems, scientists like Enrico Fermi, John Von Neumann, and Richard Feynman initially used analog, punch-card computers to perform calculations for problems like artillery-firing tables. These innovative computational advances from these scientific giants further led to the development of the earliest digital computers, called ENIAC and MANIAC[34]. These early, massive systems, relied on a general computation framework called the Von Neumann architecture which is still seen within the functional separation between CPU, RAM, and hard drive roles of our modern-day computers

The simulation and visualization fields evolved side by side, both attempting to reach points of realism, yet, with one side approaching this goal from a physical perspective and another from an artistic perspective. The mixing of these two modes of thinking ultimately led to the creation of the first video games. Despite their extremely simplistic interfaces and primitive underlying computers, games like Pong set the stage for massive future developments[51]. This synthesis of simulated physics and an interactive graphics display began its exponential rise, finding additional uses in the entertainment and education industries which continue today.

Meanwhile, the field of computational simulation was becoming ubiquitous in domains like epidemiology, molecular physics, aerospace, and finance. It soon became clear that the world of artistic realism and rigorous physical realism would ultimately collide, creating fascinating new applications in the movie and marketing industry. It was the focus on general realism that was the key for further medical applications. For example, as movie rendering and physical rendering reside in disparate areas, there continues to be fantastic opportunity for the cross fertilization of ideas toward a medical audience. Nevertheless, these approaches, come at a large computational cost, which can be infeasible in a medical setting requiring real-time interaction.

In recent years, the massive growth of computer games has made the field synonymous with computer graphics. This label does not go undeserved. As of 2016, the global game market has been \$101.1 billion with a projected growth of 6.2% per year[2]. This investment has led to a massive increase in graphical capability of consumer grade computers (including smart phones). The innovations in parallel computation for graphical processing and simulation have even contributed to important strides in artificial intelligence where highly distributed models have become commonplace.

The impact of these patterns in the consumer market cannot be ignored. They continue to represent the life blood for future developments. Currently, the massively scaled capability of the world wide web is pushing the envelope for what can be done with 3D graphics. In other words, visualizations must be small enough to download, streamlined enough to run on a smartphone, and agile enough to communicate immediately with other users. These needs have afforded us advances in compression algorithms, communication protocols, API

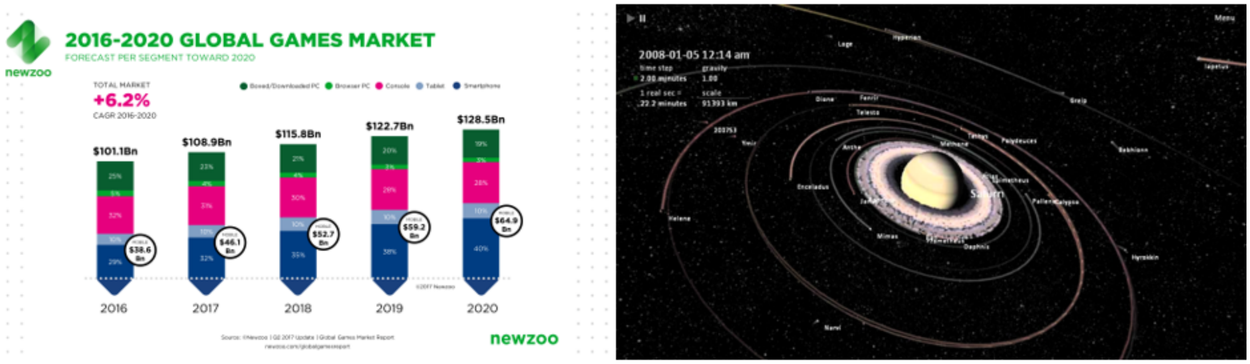


Figure 4.2: Market trends for video games and image from Universe Sandbox, a popular modern videogame which simulates planetary bodies.

designs, and much more. Luckily for our medically related purposes, we will not need to speculate on where the field is progressing in the realm of 3D graphics itself. It is sufficient to speak purely of what is possible in the field now and to use it for our own needs.

This is where the field of medicine importantly enters the picture. Its goals are not technological by design. Instead they aim to take practices into the clinical setting and optimize them for the best patient benefit (i.e., in diagnoses, treatments, education and prevention, etc.,). While a commendable framework to employ, its patient-centered focus results in a communication bottleneck between it and the other technological fields. Nevertheless, within any inefficiency is an unmet opportunity. The challenge is to now create a conceptual framework to translate consumer technological progress to its medical counterpart, setting the stage for a common direction and ridding as much computational redundancy as possible.

To this end, we categorize subdomains of medical 3D graphics as a tradeoff between computational complexity and intractability for the respective application. As a mean of illustration, we will first explore advances in computationally intensive simulation and rendering. More specifically, here will lie opportunities for highly accurate in silico experiments. This affords individuals working in these areas capabilities in physical and visual realism, but for most medical applications this will require faster user input cycles, making these techniques prohibitive. Compromises in physical and visual realism lead us to the field of interactive 3D graphics, a blossoming subdomain whose applications have scope in: 1) medical education, 2) pre-surgical planning, and 3) intraoperative surgical navigation. Such proliferation of visualization technology leads us to examine the overall computational ecosystem, taking note of the hardware advances in greenfield products like virtual reality and augmented reality. An important concept hiding in the background of these techniques is the transformation of new visual or physical understanding to clinical action. This sensing-acting feedback loop, commonly attributed to robotics, will be an important theme, beginning with human medical assistance with gradual transformation into more extensive automation.

4.3 Simulations and Developed Movie Realism

When one initially considers the construction of 3D models of human physiology, they are tempted to make as few compromises about physical realism as possible. This could involve the investigation of neural connections, tumor development, or the complex contraction of a beating heart. On the other hand, sometimes one doesn't need to represent the spatial nature of these problems. For example, neural communication can be represented as a graph structure rather than an accurate 3D rendering. Nevertheless, complexities greatly increase, when one tries to move to models of tumor growth and metastasis. Within this problem, there is an inherent need to consider volumetric modulations over time. Meanwhile, heart contraction combines the complexities of almost every other problem that can be reduced, requiring consideration for contraction, conduction, tissue mechanics, and fluid flow.

Perhaps the most fundamental tool necessary for any such analysis is the differential equation. Although rarely solvable in closed form for a sufficiently complex system, there is a prevalence of literature for the iterative simulation of dynamical systems. This began with first order techniques like Euler's method and has progressed into more complex Runge-Kutta methods which, although having been first introduced in the early 1900s, were not properly

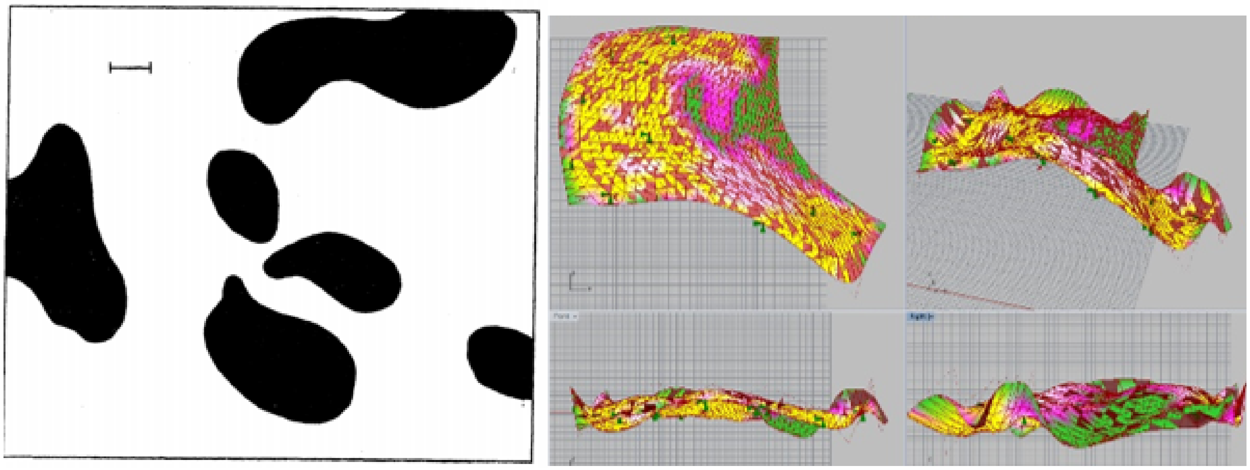


Figure 4.3: Figure from Turing's paper on morphogenesis and an image from ANSYS showing morphogenesis method for creating a structurally stable roof[41].

investigated until the 1970s. Numerical methods for biologic applications first appeared in Alan Turing's seminal work "A Chemical Basis of Morphogenesis", a remarkable paper which predicted function of genes, established the field of theoretical biology, and introduced the world to theoretical underpinnings of chaos theory [48]. The generalizability of these methods has led to a wealth of ordinary differential equation solvers, available for commonly used programming languages.

At some point, the 3D structure of an object becomes important to the solution of a clinical problem. For instance, you may be a medical device designer and need to know how stress and strain effect the deflection of a catheter. This likely requires you to take advantage of the field of finite element analysis. For example, by defining a material's vertices and local interacting forces, complex models can be generated to show how 3D objects interact with their environments. The complexities of these problems become greater when considering the modeling of fluids within a given system. In any case, the predominant software package of choice for these calculation is the ANSYS software package. Meanwhile, the construction of these 3D models is typically delegated to software such as SOLIDWORKS for mechanical objects and Materialise's Mimics for anatomical structures.

These methods become even less tractable, however, when one considers the interactions between dynamic solids and fluids. Even this would be a simplistic representation for the heart as its physiology involves a contracting myocardium and an intrinsic electrical conduction system. Luckily, many strides in related computational areas have come through novel research pertaining to animated movies. Particularly, both Pixar and Disney have strong research laboratories which have advanced technologies needed to produce realistic interaction between solids and fluids. For example, in 2016, Pixar, Disney, and the University of California Santa Barbara (UCSB), published a paper on "Eulerian Solid-Fluid Coupling", a technique that, although targeted toward the movie industry, can have clear implications for the medical field [47]. Furthermore, from vorticle simulation to physics of human hair, the animated movie industry has and will provide rich opportunities for medical computational utilization [4][26].

Despite the large amount of complexity, simulation groups have already begun to investigate how to simulate a full beating heart model, even with the added parameterization of fluid flow. Perhaps the largest collaborative effort on this front is the Living Heart Project, which aims to bring large clinical studies from in vivo to in silico[6]. So far, their efforts have led to interesting finding regarding the prediction of disease pathologies like dilated cardiomyopathy, modeling of the electro-mechanical system, and modeling of the fluid-structure interaction [39].

Although these methods work well for physically realistic representations, there lingers the problem of visual realism. Focusing on visual perception of simulated models can result in better interpretation, yet such can also be important for situations less closely tied to clinical practice, like marketing and education animations. The relative quality of perception is the primary purpose for companies like Disney and Pixar so again these groups can used as examples of effectively developed rendering practices. For instance, sampling from a movie scene can require Monte Carlo rendering techniques. In a 2017 paper, UCSB, Pixar,

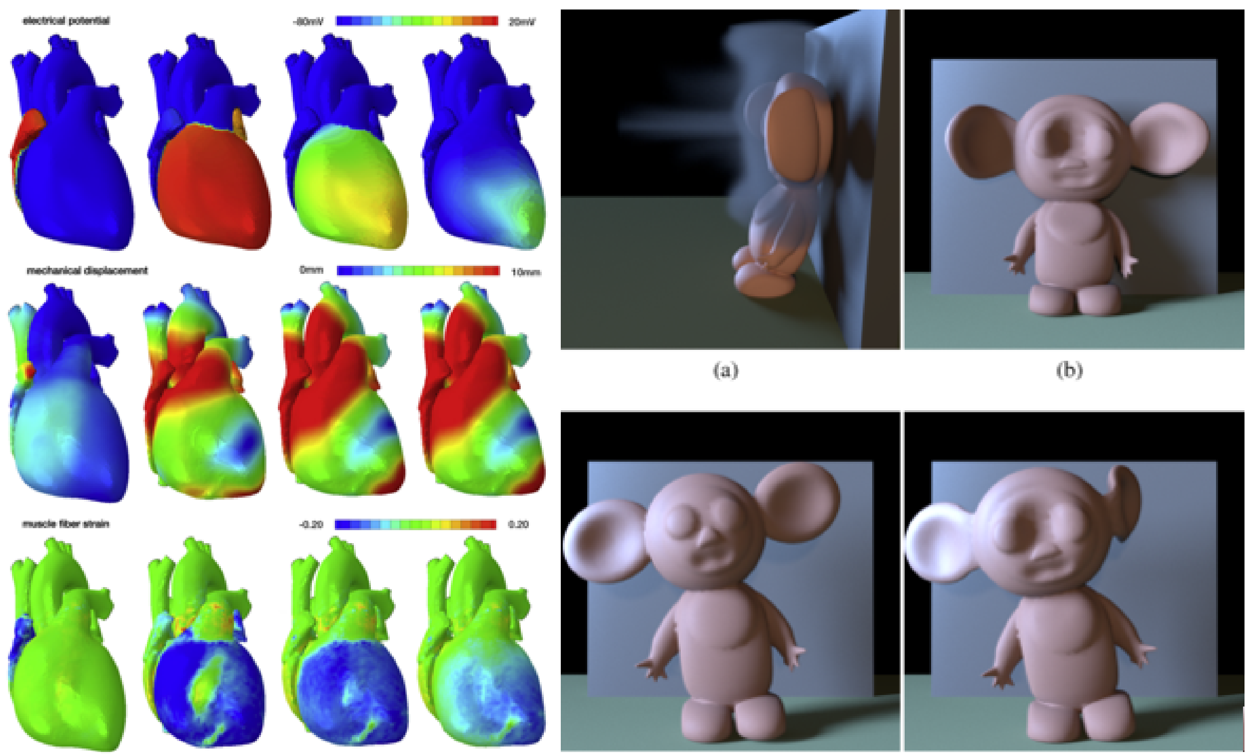


Figure 4.4: Living Heart project simulation results and figure form Pixar paper on solid-fluid interaction.

and Disney reported on the use of a convolutional neural network to de-noise output from such methodologies[7]. In other words, more attention in realistic medical rendering could provide more accessibility for the consumer market, spur more interest, and result in more investment.

4.4 Real Time Rendering

Sometimes high realism in either a physical or visual sense are neither practical nor applicable to clinical applications. When medical decisions must be made quickly, faster interfaces must be created to accommodate these needs. This is where the field of interactive 3D graphics has and will become an integral part of the investigative pipeline. This field has had unprecedented growth in the video game market which in turn has led to rapid increases in technological innovation and availability of development tools.

Today, the diversity of computational requirements for interactive 3D graphics is extensive. Video games must be run from handheld devices to high end gaming computers. Consequently, platforms have been built in accordance with many use cases. Currently, the most extensively used game engine is Unity3D. Its flexibility has allowed it to accommodate many of the bleeding edge technologies like virtual reality and augmented reality. The Unreal engine is also a popular choice, frequently used in the larger budget video games where photorealism is a goal. Most remarkably, however, is that these platforms have been opened to the public for free use. This is a good signal for the power behind the 3D graphics market.

Meanwhile, the intense growth of the internet has required advances in 3D rendering as well. This is difficult, however, because it requires 3D graphics developers to work with algorithmic and file size efficiency suitable for browsers on many devices. WebGL has been the main pioneer in this regard, bringing model and shader programming to large audiences. WebGL was initially created as a low-level graphics language. As a result, graphics frameworks like Three.js have been built on top of WebGL to make graphics programming more intuitive. Further, Unity3D is also built with functionality to compile to WebGL code.

Fundamentally, these tools all solve the problem of complex spatial reasoning. Whereas traditional textbooks present anatomical information in 2D form, like a map, 3D graphics allows the user to better understand the overall spatial environment, but not necessarily the physics involved, which requires more computational resources. This is akin to the divide between the disciplines of physiology and anatomy relative to medicine: simulation is suited for the physiological sciences, while real time 3D graphics is suited for anatomical understanding.

Consequently, perhaps the clearest application for 3D medical spatial reasoning is anatomical education. This is promising for medical students who frequently rely on donated cadavers for their instruction. While 3D graphics will never fully replace cadaver study and dissections, their uses may be supplemental for comparative anatomies. Despite these opportunities, to date, there has not been a major 3D medical education company that has entered the market with major prominence. Mobile phones and tablets are perhaps the most popular medium for anatomy students in a medical school context[11]. This has the advantage of being easily shared to many people, but the promise of its ubiquity is limited. The winner in this field will need to tap into the culture of the mediums being widely utilized. For instance, providing guided understanding of anatomical knowledge to the typical student/consumer would make progress toward this goal. Until then, encyclopedic resources will continue to rank highest in an arena of unfulfilled potential.

The connection between 3D anatomical education and medical practice at first glance may seem slim, but this is not the case for either pre-surgical planning and industrial development. There are frequently situations where the three planar views typical in MRI and CT are insufficient for proper procedural guidance. For instance, complex congenital heart surgeries can exhibit extreme variation between patients and it is difficult to image such throughout the cardiac cycle in a small non-cooperative child. A similar argument can be made for obtaining an adequate anatomical database for the development of medical devices. More specifically, it is important to know how a catheter may fit and may deflect within varied cardiac anatomy. Further, one has to consider pathological anatomies and how they may reverse remodel after successful therapies. The downside for MRI or CT approaches is that they take substantial time and investment to have a skilled anatomist convert medical image scans to their 3D analogs, such as 3D prints. Thus, the aspiring entrepreneur may consider creating an automated medical segmentation system for such purposes.

At the University of Minnesota, progress has been made in making such a virtual modeling pipeline a reality. For example, Dr. Daniel Keefe, in the Interactive Visualization Laboratory, has developed advanced visualization interfaces for the inspection of medical anatomies, involving in-house systems for 3D virtual reality viewing[16]. This became particularly useful when Dr. Keefe in collaboration with the Medical Device Center under the direction of Prof. Arthur Erdman and University of Minnesota pediatric congenital surgeons (Drs. Saltzman and Azakie) visualized a 3D model of the hearts of conjoined twins that were to undergo a separation surgery. This proved to be extremely useful for safe planning of the procedure, which resulted in very positive outcomes[22]. Although the modeling involved required a significant amount of investment, innovations in these fields suggest that these same steps will be compatible with a standard of practice.

Currently, the next big venue for real time graphics in the medical field, is its incorporation into the OR suite. Surprisingly, this is already quite prevalent, but resides more in the background. For example, cardiac navigation systems require real time rendering to show catheter or lead being positioned in the patient. Mapping systems like Medtronic's CardioInsight must both display a 3D representation of the morphology of a heart and simultaneously overlay a heat map of electrical activations. Implicit in these technologies is the challenge of localization and coordinate system synchronization. Although difficult in

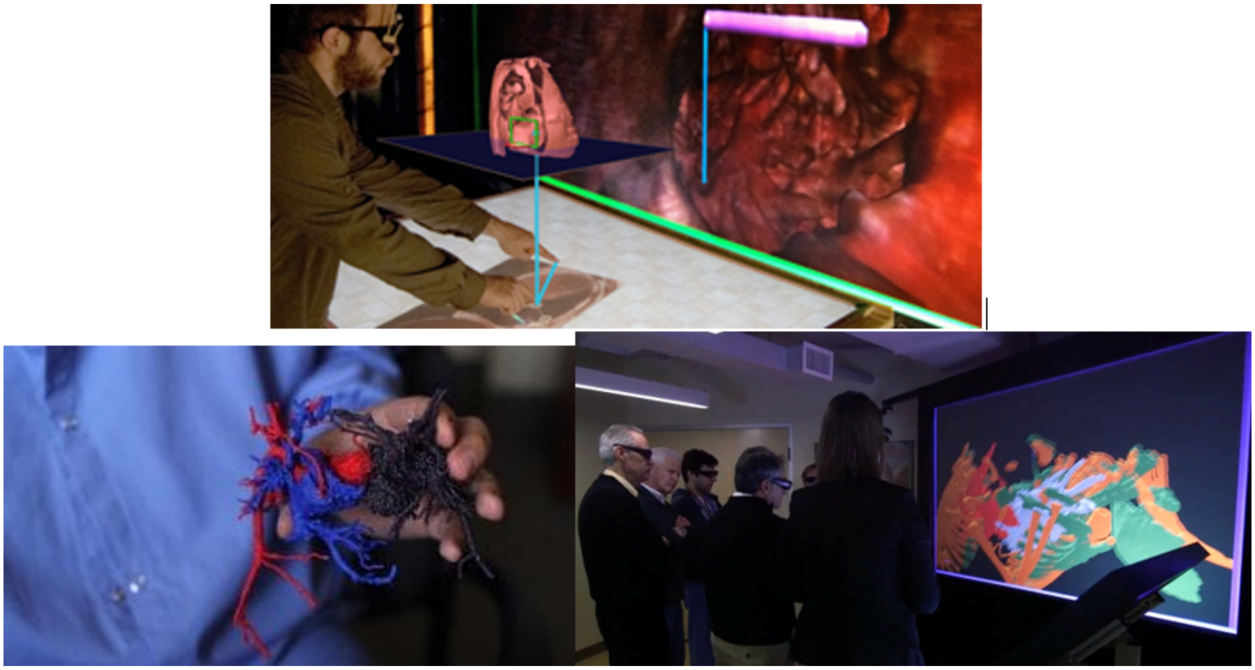


Figure 4.5: Dr. Daniel Keefe with virtual reality system, 3D model of conjoined twin model (developed and printed by Alex Mattson from the Visible Heart Laboratory), and investigators looking at twin model in virtual reality.

practice, its benefits would be numerous. Today, using standard, fluoroscopic navigation for cardiac interventions, for instance, requires the patient to be subjected to relatively high doses of radiation throughout the procedure. This disadvantage is compounded by the fact that the healthcare providers must also be subject to doses of radiation while wearing heavy lead protective wear which can cause spine problems over time. Combinations of interactive visualization and procedural innovations will provide fertile ground for those looking to develop new products for numerous medical fields.

4.5 Expansion of visualization hardware

It used to be the case that virtual reality systems were limited to a small supply of expensive systems, powered by expensive computers, and restricted to niche industry and academic institutions. It has not been until recently that virtual and augmented reality has entered the consumer market. These innovations have, again, been pushed primarily by the large video game market, leading to product development for computer, phone, or video game console. Currently, the major virtual reality systems on the market are the HTC Vive and the Oculus Rift. Yet it should be noted that the hardware sector for virtual reality is developing rapidly, so many more competitors will be entering this market in the future.

While virtual reality is characterized by placing the user in a fully virtual environment, augmented reality functions by superimposing holograms within the user's immediate, visual space. Although the underlying physics of these problems are similar, augmented reality poses the complexity of needing to understand the spatial contours of the immediate area. To date, the only major player in the augmented reality hardware field has been Microsoft with their Hololens technology. It has been used in some industry circumstances, but its standalone nature requires the headset to have a very small field of view. Like with virtual reality, augmented reality is an active field of development and more products should be expected to enter the market in the very near future.

The innovations afforded by these modern visualization technologies parallel the applications outlined for standard computer graphics, including education, presurgical planning, and intrasurgical navigation. At the University of Minnesota's Visible Heart® Laboratory, we have been investigating applications for education and pre-surgical planning utilizing high end personal computers as well as mobile phones [19]. One of the lab's developed applications allow students to explore either a heart or a cadaver model at different zoom scales.; this has been useful in learning general anatomy. Additionally, these simulations bring to light anatomical features of that may affect how a clinical problem could be solved with a given medical device. In other words, learning anatomy becomes more like learning the layout of a home, rather than a 2D map or the structure of a handheld model.

Although virtual reality has its place in an intraoperative setting, augmented reality has more prospects in this regard. A surgeon can safely wear an augmented reality headset during

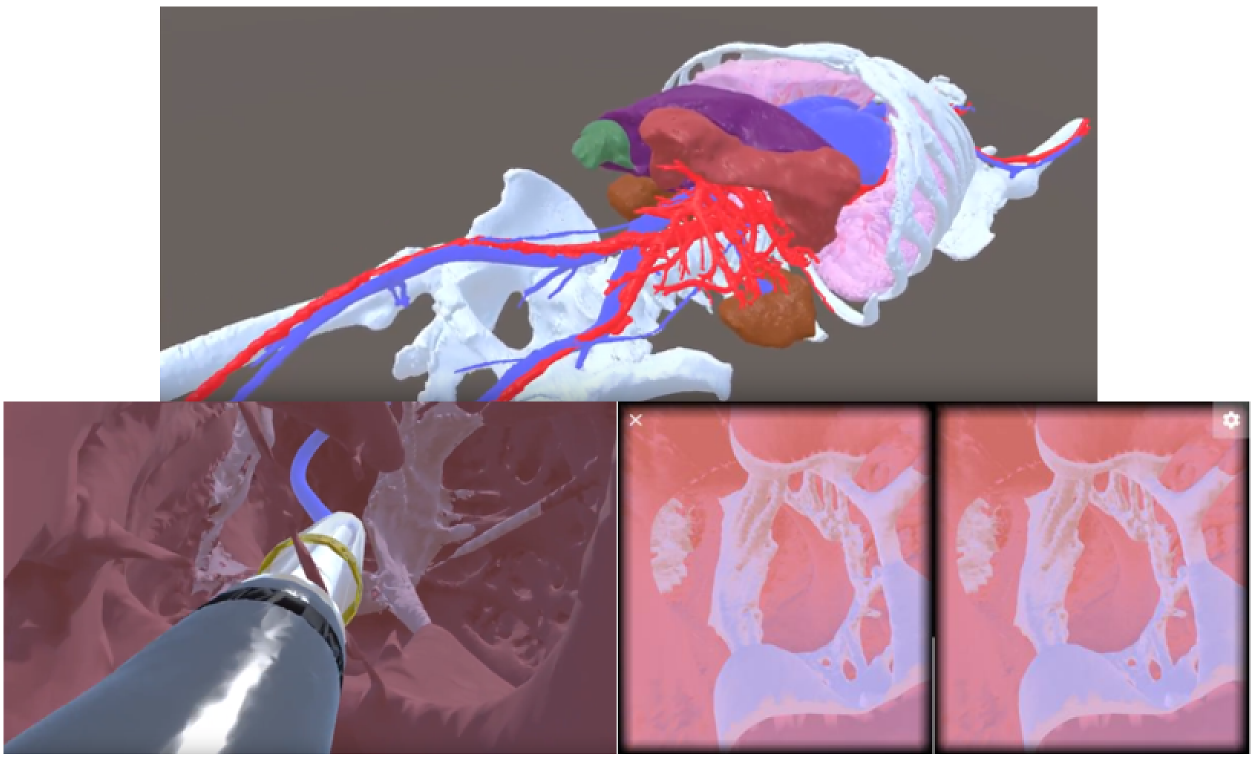


Figure 4.6: Cadaver model in virtual reality, Micra delivery looking up the tricuspid valve, and mitral/aorta shot using Google Daydream.

a procedure while maintaining adequate view of the procedure. In the education sector, VR and AR are both useful tools. The same can be said for industry-based visualization and pre-surgical planning. Even when not considering 3D graphics, the Hololens has functionality for projecting screens arbitrarily within the user's local space. This could act as a replacement for bulky monitors that take up excess space like with a fluoroscopy unit. A reach goal of these field, however, would be to project a patient's internal anatomy onto the patient during the procedure. This would be as if the physician can see through the skin and visualizing organs and internal devices directly [9]. This requires addressing of the same types of localization and interaction challenge as highlighted earlier.

Even in such an immersive scenario, there is still a desire to better be able to incorporate virtual and augmented reality into experiences that would convincingly simulate a real surgery for training purposes. Yet, to do this well, this will require more investigations/applications relative to the current very young field of haptics. Robotics is perhaps the most applicable field for the development of convincing, realistic, haptic systems. It should be noted, that there are products on the market for simulating laparoscopic procedures, but there have been few other use cases that have extended beyond applications that recapitulate procedures that use a controller and monitor [8]. The major challenge in this field, is to convincingly create a model system that is flexible relative to human manual dexterity. To date, products in the virtual reality space have been primarily utilizing gloves to track hands, but there has been little work in converting this information to haptic feedback.

This new hardware arena is particularly interesting, because of its needs for a diverse skillset. Physicians, roboticists, programmers, video game developers, systems engineers, teachers, and more are all required to bring these technologies to their full potential. These separate uses would benefit from an abstracted framework for faster development cycles. Consequently, freelance and corporate partnerships will be important for the pace of development for this sector.

4.6 Virtualized Computation

Fundamentally, enhanced visualization technology allows for better decision making and execution by the medical operator. But facilitating human operation also leads to the facilitation of computerized automation. Although there has been much research in the past about sense-action feedback loop systems, it continues to be an active field of research. Such a feedback system could be: 1) a fully autonomous system for clinical therapies; 2) an error checking system; or 3) monitors of general human awareness on a procedural level. Unfortunately, taking a 3D spatial map of an anatomical structure and converting that to actionable knowledge by a computer, in real-time, is a difficult challenge. Therefore, the intersection between these visualization fields and what could be labeled as artificial intelligence will be important as these areas develop.

From a historic perspective, the analysis of spatial and semantic content can be traced back to the field of computer vision. Face detection, gradient, and optical flow calculations were all major successes in this field. More specifically, for a given problem of interest, this would involve manually engineering features that would detect, highlight, or distort some object of interest. Many strides in these areas were made around the early 2000s, when international security concerns entered a major portion of the public eye. This required innovative systems that could detect, track, and account for various persons of interest in pictures, video, or particle maps. These technological advances were then able to transcend into the entertainment industry, with products like the Microsoft Kinect. Further, the Kinect was of interest relative to software developers, as it provided an inexpensive open sourced particle map device for both consumers and researchers.

Feature engineering is not robust enough to work on high dimensional spaces where conceptual information becomes important. For instance, it is difficult to create manually engineered features to distinguish between dogs and cats in natural images of these animals.

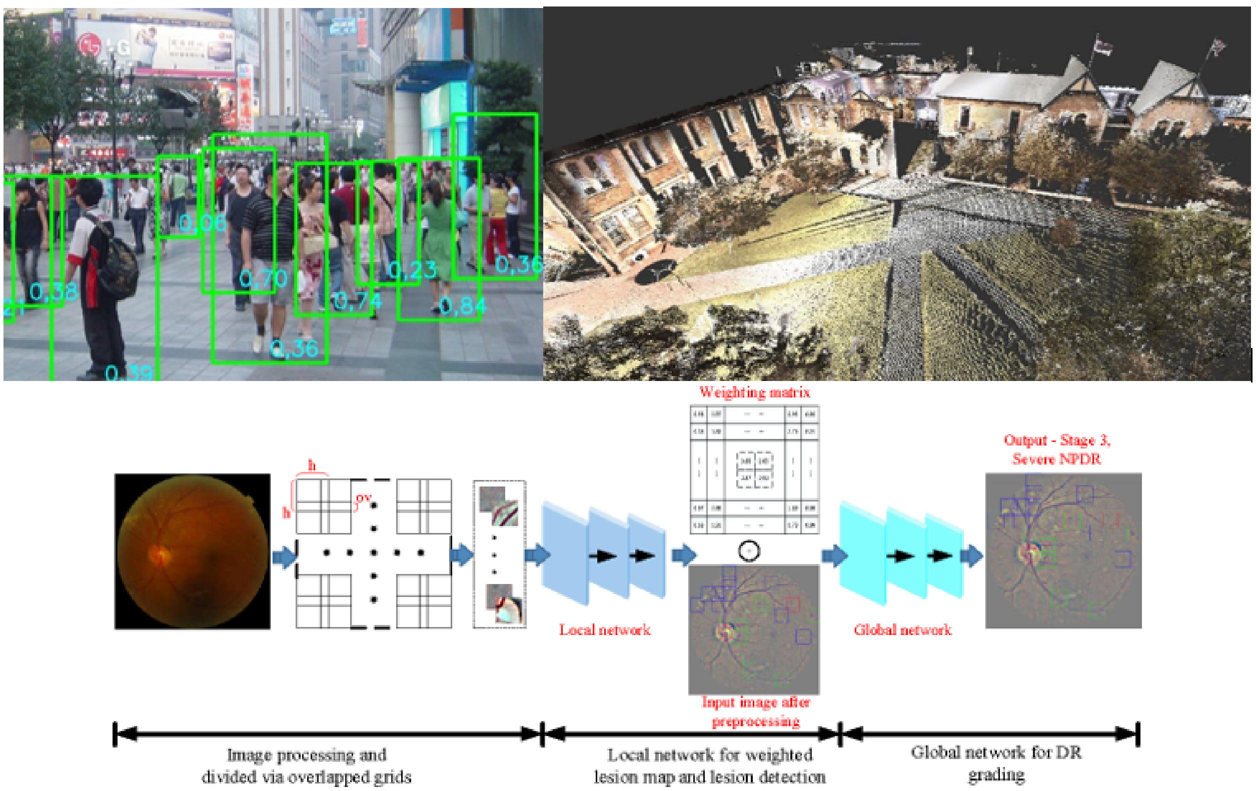


Figure 4.7: Detection of persons in a picture, point cloud information, convolutional neural network to detect diabetic retinopathy.

An application of recent interest, where this also becomes extremely important, is the self-driving car. To detect lane lines, one can create thresholds based off color and hue and computationally fit a polynomial to the resulting lines. However, this may quickly break down in different lighting conditions, weather, or in situations where there are simply no lines provided.

More recent machine learning methods like the support vector machine, random forest, and boosting methods have started to bridge the gap between feature engineering and semantic understanding. Further, the emergence of deep learning and convolutional neural networks have taken the problem of computer vision and made it close to a solved problem. Interestingly, these same deep learning algorithms can be just as useful for tabular data, natural language, and audio data. This flexibility in data representation is particularly useful as it allows multi-modal models to be created for applications like image captioning[17]. Yet, it remains a challenging task to apply data automation to the field of medicine. For a clinical problem, for example, it requires copious quantities of domain knowledge in addition to diverse measuring modalities including imaging and biomarker readings. Consolidation of all these factors has resided within the purview of physician judgement, but these new deep learning algorithms could have enormous potential for solving such problems quickly (in a needed medical timeframe).

A special type of neural network, called a recurrent neural net, has already been used within the context of pediatric mortality risk assessment at Children's Hospital in Los Angeles[3]. Another application of neural networks for the medical field has been in segmentation and feature detection[43]. This is extremely useful for multi-modal image co-merging. A simple strategy for this is to register coordinates between imaging ahead of time, allowing for images to be overlaid. For example, Phillips has developed a product called the EchoNavigator which combines fluoroscopic images with trans-esophageal echocardiography, and then overplayed their results which allows for easier navigation for cardiac interventionalists. Yet, for some medical applications, this will need to be performed with content information alone and not with pre-registered coordinates. Thus, a deep learning approach combined with the techniques required for simultaneous localization and mapping will become crucial in these endeavors. Interestingly, this is very similar to how the self-driving car industry has been evolving.

With rapid advances in these computational factors in play, there are many opportunities for entrepreneurs of diverse backgrounds to enter the engineering in medicine market. Programmers, roboticists, and graphic artists will be the most important careers to merge with these future products. This expertise will not be useful, however, if it cannot be properly integrated with the complex domain knowledge within the medical fields. Consequently, there will need to be effective cross-communication, a task best suited for the medical device innovation leaders. Interestingly, these skills are best aligned with those in entertainment related fields, particularly video game development. This is perhaps where the most opportunity for competitive advantage still exists. Inventors who can embrace this reality and make connections to disparate fields will be able to create more effective medically related products than those who do not. We have not even begun to cover the how effect of scaled networking systems, sometimes called the internet of things, will push all such technological trends. No matter what modern technologies appear on the market, there is and will continue to be plenty of unique opportunities for creative thinking when considering how to combine the strengths of all of them.

4.7 Prospects

Part of predicting how the medical technology associated with augmented or mixed reality fields will go, will require the considerations as to what a hypothetical investor would be willing to fund. Recently, we have observed that the problems relative to3D graphics in the medical field, have been much broader than one would have initially expected. For example, one must not only consider the tradeoffs between visualization and simulation capabilities, but also consider the system, and how customs change as automation contributes to the feedback loop. This global outlook, relative to the somewhat short product life cycle in the medical field, is likely a useful perspective to adopt.

It is interesting to consider today, that the integration of intelligent robotics into a surgical suite is perhaps the precipice of medical challenges, requiring a computer to have adequate 3D visualization, manipulation, anatomical understanding, as well as spatial motion planning. It can be seen already that robotic incorporation into a surgical suite is not dependent on these factors combined. For example, if one utilized existing surgical robotic and further combined them with computer vision or haptic technology would be highly influential. It should be noted that one should keep an eye out for improvements to these technologies in the future. Importantly, medical device developers with familiarity in control theory, electrical engineering, machine learning, and surgery will be required throughout this progression.

Figure 4.8: The Davinci surgical robot and a reinforcement-learning based robot locomotion from Google [20].

These challenges play well into our focus of 3D graphics. Medical 3D modeling with limited automation has already been successful in the global market. Materialise's Mimics platform has been the standard, allowing researchers and physicians to incorporate technology like 3D printing into their pipeline[36]. Despite its current limitations, the construction of 3D geometries generally improves the mental saliency of the clinical problem at hand. Whether your interest is in cardiology, orthopedic surgery, general surgery, or urology, there are opportunities for quickly incorporating visualization into clinical practice or device industry related research. As a result, it is prudent for all players in these associated fields to keep track of graphics technology, hardware, and how it is beginning to be incorporated within standard practice. As simulation becomes more prominent, it will also be likely that the technology plays more toward remote servers performing computation, whereas if real time rendering becomes more prominent, then it will be likely that medical imaging systems will include more graphics hardware in the OR. Unless some unforeseen alternative technology that enters the market supersedes the technologies discussed here, it can be expected that this is where the field will go. Hence, the more likely question is just about time to be applied broadly in engineering in medicine.

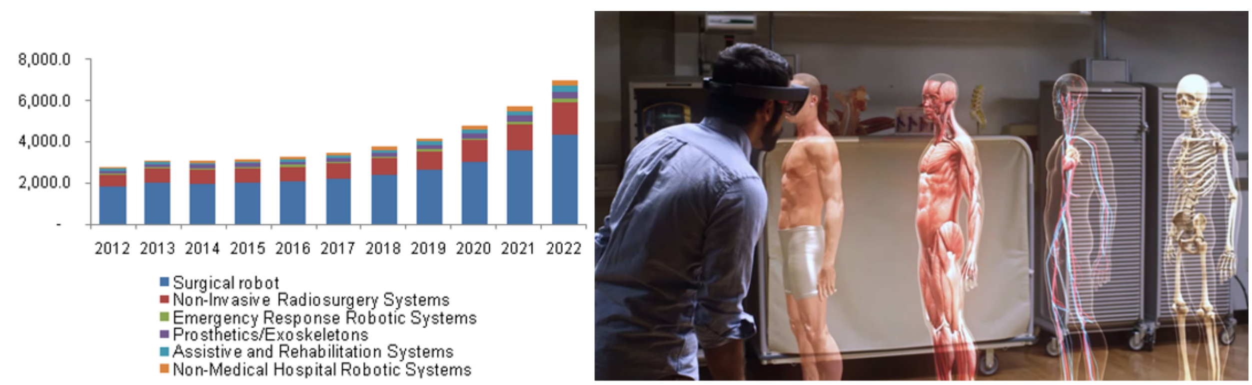


Figure 4.9: Projected growth of robotics in the medical field[1] and an anatomy demo from the Microsoft Hololens augmented reality headset.

To many, these predictions can seem like a major paradigm shift in the current medical landscape. This is not necessarily the case, as 3D graphics technology can incorporate nicely within currently utilized technologies in most healthcare areas. Thus, it will not be necessary to pivot to a completely new field, but rather find ways in which current products can be enhanced by more informative computational representations. For example, within the Visible Heart® Lab, we have recently observed a diverse demand for such tools. To date, within just the virtual reality realm, we had computational visualization requests relative to; 1) transcatheter valve delivery; 2) placements of delivery systems and devices within a beating heart; 3) marketing visualizations for numerous medical device; 4) 3D models and prints for pre-surgical planning; 5; complementary educational tools; and others. What becomes necessary going forward in the medical field, is for the proposed entrepreneur to choose a specific niche and make 3D rendering seamless for specific clinical/educational medical needs.

Being an effective implementer of such technologies going forward, will require one to be familiar with the emerging markets as well. Therefore, it will be necessary for a simulation practitioner to become competent with data analytics, machine learning, or the Pixar model of graphical rendering. Further, web development familiarity will become more and more useful for incorporating skills with interactive design and version control. Importantly, an overwhelming amount of additional information appears daily, so gaining the skill to identify experts and form collaborations is critical for the implementation of lofty goals of employing computational modeling and visualizations within the medical field.

4.8 Conclusion

The virtualization of the medical field is less about 3D graphics and more about a trend toward better sensing and processing of complex anatomical and pathophysiological information. Real-time sophisticated graphics are merely a vehicle for achieving these ends. Further, these applications and ongoing computational needs will lead to a great expanse of associated careers that must intertwine to allow for these modern technologies. Depending on how these innovations blossom, the landscape for clinical and industry practice could completely change. Perhaps then, risk falls on the employee who has decided not to be a lifelong learner, because automation might not take very long to make some careers, even medical specialties, become irrelevant. This is not to say that employment instabilities will be a zero-sum game, but rather that there will need to be a major shift to computational thinking, leaving nonparticipants in a state of need and adopters as the better collaborators within a larger economy.

We are in exciting times, where innovations from the technological industries are not only easily accessible, but largely detached from the medical field. Rarely can one affect such an important market with such low friction. We suggest, that for anyone wanting to continue to participate in the fields of medical visualization an associated technology, the best place to start is to find a technological hobby, perhaps with video game development, and find a way to make it useful for the medical industry. Chances are you will see something very similar in a related or unrelated niche field, now or in the future: hence, find a useful sector, target that group of users, and quickly make that technology practical.

Chapter 5

Using Smartphone-Based Virtual Reality to Explore Internal Anatomy of 3D Heart Models

Authors:

Erik Gaasedelen[1], Alex Deakyne[1], Tinen Iles[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

This section recounts my involvement in a project for the Design of Medical Devices conference, initiated during my inaugural week of exploring virtual reality (VR). The project stemmed from a spontaneous idea, fueled by a conversation with a colleague and significant sleep deprivation. We transitioned from concept to execution in just three days, facilitated by my possession of a Nexus 6p, which was compatible with Google's then-novel VR framework.

Despite its brevity, the technical overview emphasizes hardware with the greatest potential impact. Smartphones, integral to the pervasive device ecosystem, facilitate rapid and straightforward consumer targeting. This project, therefore, represents a prime commercial opportunity.

More broadly, this initiative marked the first application of VHL's cardiac STL models for comprehensive 3D inspection. Initially, my rudimentary Unity skills limited me to visualizing these models in a basic matte white. As my proficiency advanced, I began to enhance the models with specific colors for different parts, such as white chordae and red myocardium.

While this dissertation does not cover all subsequent developments stemming from this project, it is important to note the ongoing work. Recently, I transitioned to using Google Daydream, improving user interaction with the heart models. Additionally, I adapted these models for use with Google's AR toolkit, enabling users to project the heart onto surfaces like floors or desks. These advancements led to further exploration with high-end headsets, detailed in subsequent chapters.

The key takeaway from this chapter is the recognition of smartphones' potential. Although they may not match the capabilities of high-end systems, they offer a remarkable opportunity for users to engage with and explore complex human anatomy. Whether for educational purposes or preclinical planning, this scalable infrastructure presents a viable avenue for widespread application.

5.1 Background

The recent and rapid developments of immersive, interactive 3D environments have been critical in advancing interfaces for entertainment, design, and education. For cardiovascular research, our laboratory and others have been able to use such software tools for the construction of heart models from DICOM files. These models can then be printed in hard or soft plastic from a 3D printer. In general, such models are considered useful for surgical planning and education; these modalities are being applied as critical tools in the field of cardiovascular research.

Recently, the development of virtual reality (VR) has introduced a new modality for exploring 3D virtual structures with high resolution, high flexibility, and fast turn-around times. Until recently, the adoption of these technologies has been hindered by the high costs of VR goggles and the complexities in their setup. New developments in phone software and hardware, however, have alleviated some of these difficulties by allowing smartphone screens, graphics units, and gyroscopes to provide the necessary technologies for VR. In this way, phones can be placed inside a headset holder and used freely, without being connected to the computer. Here we explore the utility of using this VR setup in the context of internal heart anatomy visualization.

5.2 Methods

Hearts were first received from an organ procurement agency called LifeSource. They were subsequently perfusionfixed to elicit and end-diastolic shape and then scanned using MRI or CT. From generated DICOM datasets, 3D models were constructed using a software package called MIMICS (Materialise, Belgium). These whole-heart tissue models were then exported to a stereo-lithography (STL) file format and later converted to an object file.

The creation of a virtual environment for cardiac imaging was setup by taking advantage of recent developments in Google's Cardboard and Daydream projects. First the Unity video game engine was installed along with the Google Daydream VR plugin. Within Unity, a new scene was created, allowing for a heart model object file asset to be imported. At this point the first person perspective can be placed in any chamber of the heart. For our demo shown here, we decided on the left ventricle near the mitral valve and left ventricular outflow tract.

Outputting Unity's environment to an external device required installation and configuration of Android software development kits and application programming interfaces. First Android was installed with the proper packages for Android API version 7.0 (Nougat). The Java developer kit version 8.0 needed to be installed as well as it is a dependency of Android. Once these were available to Unity, it could be configured to build and run under this operating system.

A Nexus 6p running Android 7.1 was used for testing. After activating its developer mode and allowing USB debugging, the scene from Unity could be built and run on the device. This produces a stereoscopic view where the scene is duplicated on two halves of the device. In order to view the scene, the phone was placed in a headset with built in lenses to show the correct half of the phone to the respective eye. In our case we used a Utopia 360 headset, but other headsets like Google Cardboard follow the same design specification and could just as easily be used.

Once the setup is completed and the user is wearing the headset, the phone's gyroscope will sense head motions and react as if the person was located within the heart chamber of choice.

Using 360-degree panoramic picture/video functionalities of Unity, static internal shots can be easily uploaded to an online server and viewed within the browser. Without the overhead of rendering the 3D model, time to VR is faster and more convenient at the cost of user interaction. A demo of these views can be found at our Atlas of Human Cardiac Anatomy. (http://www.vhlab.umn.edu/atlas/vr/)

5.3 Results

For our demo, we placed the user in the left ventricle of a fully modeled human heart.

In Figure 5.2, the mitral valve can be seen on the right of the screen while the left ventricular outflow tract, along with the aortic valve, can be seen at the left of the image. Large

Figure 5.1: VR Headset Setup

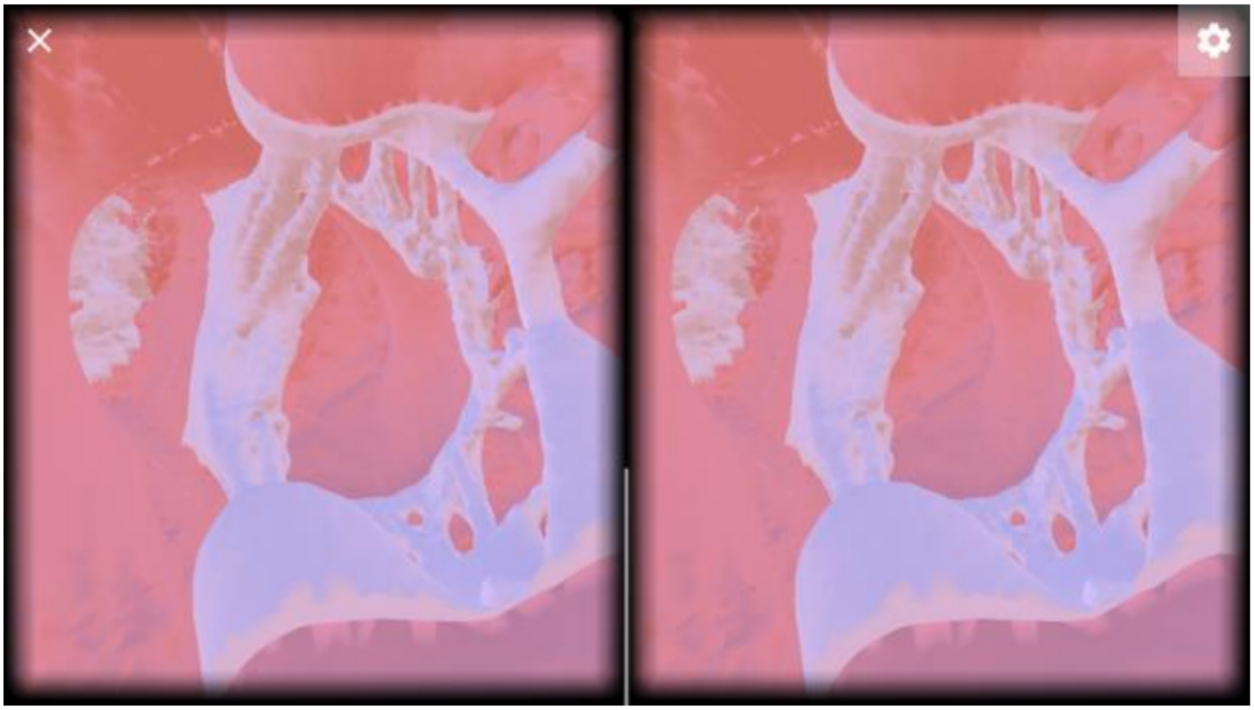


Figure 5.2: VR view of mitral valve

papillary muscles can also be seen extending from the top and bottom of the screen. The valve structures are presented in an off-white color, while the myocardial tissue is depicted as a light red.

The apex of the heart (Figure 5.3) can be seen if the user turns their head around. Trabeculations can be clearly visualized in this view. Any pixelation of the muscle is a result of DICOM to model conversion and not of the VR setup itself.

5.4 Interpretation

An advantage of the mobile approach for VR is that the cost of material and time for development cycle is relatively cheap. In this way, exploration of 3D anatomical structures can be easily distributed to anyone with a capable mobile phone and compatible VR headset. Unfortunately, there are some technical difficulties. Running a VR scene is GPU intensive,

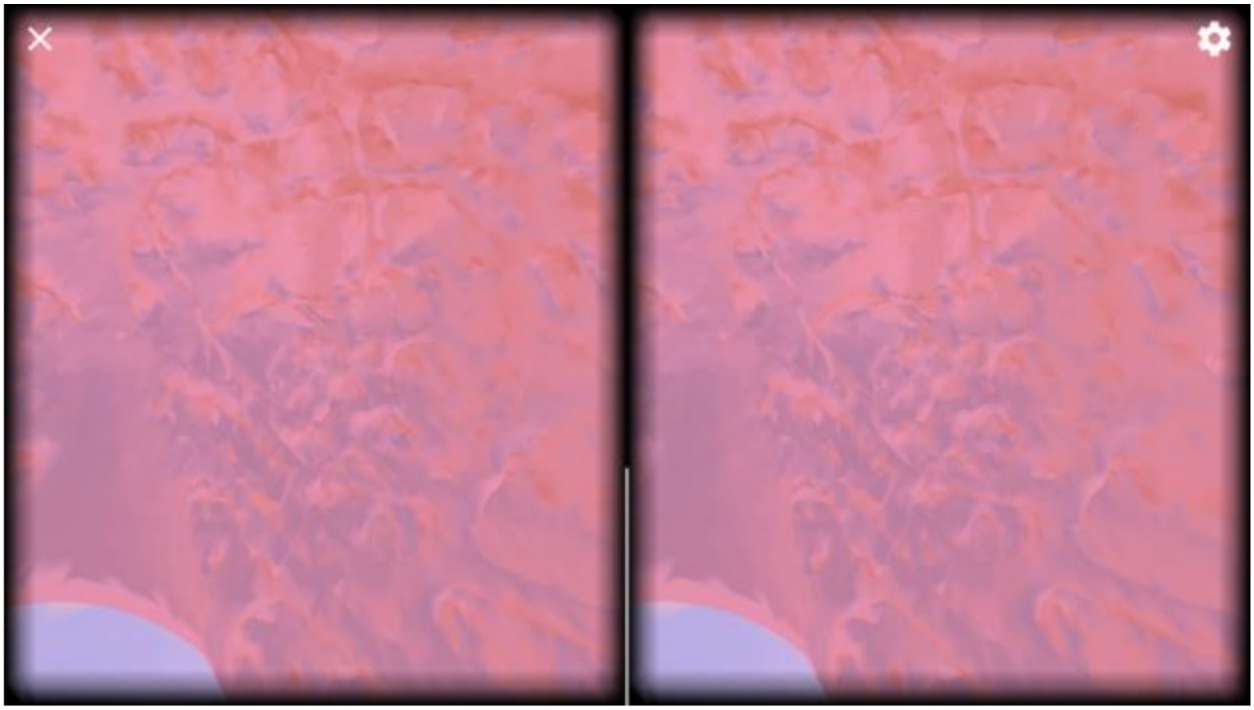


Figure 5.3: Apex view

drains mobile phone battery, and requires a framerate which can cause flickering. These problems are less severe, or remediated entirely, with high end VR headsets like the Oculus Rift or HTC Vive. However, a powerful computer is required to operate these VR headsets. Our workaround for this was to implement 3D static panorama shots which are not affected by the rendering overhead, providing clearer pictures on less capable phones. This is at the cost of user interaction with the model. In any case, advances in technology will continue to make VR experiences more convincing, even for smartphones.

Given the ease of distribution for a smartphone based VR system, we believe this setup is very well suited for education. High schools, colleges, and medical schools would all be possible markets for a product like this that can offer an immersive educational experience. Since most people in these consumer groups already own a smart phone capable for VR, teachers would only need to purchase a headset to give their students an interactive VR anatomy tutorial of the heart.

Surgical guidance is another possible application. For example, patient DICOM files can be modeled and viewed on a computer monitor, therefore these same models could be viewed with VR. This offers a much more compelling modality, to have 360 degree views from inside internal structures. Surgeons would be able to explore and focus on specific anatomy in a more immersive modality prior to performing the procedure.

Virtual prototyping of medical devices is a field of rapid growth, with complex VR systems being developed by multiple centers (including the Medical Device Center at the University of Minnesota). The VR approach described here allows for additional applications complementing the more sophisticated approach for medical device prototype simulation. Even with the system we describe, medical devices could be virtually implanted into heart models utilizing advanced VR technologies. This in turn would offer a new way to envision the implantation of a device and how the device would interact with the heart post implant.

In conclusion, mobile VR of internal heart or other anatomies may offer a unique and cheaper modality to enhance anatomic learning. Because of the ongoing advances in VR technologies and their reasonable costs, this approach could become an excellent educational tool that could be readily utilized for academic education, medical device design, and surgical planning.

5.5 Acknowledgements

Thank you to Megan Schmidt, Lars Mattison, Alex Mattson, and all the rest at the Visible Heart Lab® for your assistance in heart modeling and stewardship of donor hearts.

Chapter 6

Virtual Reality and Visualization of 3D Reconstructed Medical Imaging: Learning Variations within Detailed Human Anatomies

Authors:

Erik N. Gaasedelen[1,2], Alex J. Deakyne[1,2], Alexander R. Mattson[2,3], Lars M. Mattison[2,3], Mikayle A. Holm[2,3], Jorge Zhingre Sanchez[2,3], Megan M. Schmidt[2,3], Michael G. Bateman[2], Tinen L. Iles[2], Paul A. Iaizzo[2,3]

[1]Department of Bioinformatics and Computational Biology,

[2]Department of Surgery,

[3]Department of Biomedical Engineering,

University of Minnesota

Preface

This chapter details enhancements to our VR platform, particularly through the incorporation of advanced headsets like the HTC Vive, which provide a significant upgrade in resolution. The improved visual quality, achieved through modern graphics cards and high-throughput HDMI streaming, surpasses that of mobile devices. However, a crucial advancement is the enriched user interaction enabled by Vive controllers, allowing for the integration of menus and additional elements that facilitate a more intuitive understanding of anatomy.

We explored various applications, such as displaying medical devices and diverse anatomical models alongside movie clips and interactive menus. While pinpointing a specific large-scale application remains challenging, this ambiguity isn't inherently negative. The technology proves highly beneficial for educational purposes and preclinical research. In the lab, we opted against over-specialization to accommodate a broad audience spectrum, from middle-school students to seasoned cardiac surgeons, as focusing too narrowly could shift our efforts towards production rather than innovation.

This paper primarily explores the potential applications of this technology, leaving out certain aspects like dynamic imaging in VR, which involves examining cardiac motion through 3D models generated from multi-phase CT scans. While promising, this approach and the static method it enhances face limitations, such as the absence of a corresponding physical model. Advancing towards the creation of physical models or intra-procedural visualization tools would be beneficial. However, these developments hinge on integrating more sophisticated robotics technologies, a feat that requires considerable attention and resources.

6.1 Abstract

The emerging field of virtual reality has many promising new applications for the medical sciences. For example, by converting magnetic resonance and tomography-based images into 3D models, users can visually inspect individualized anatomic reconstructions at clinically useful high resolutions. Yet, adequate development of these tools will require a wide breadth of associated expertise to take advantage of current video game technologies while maintaining relevance for clinical use. Our laboratory has begun to implement such system approaches for the exploration of hearts, cadaveric specimens, and medical device/tissue interactions. Here we demonstrate several aspects of the potential applicability of virtual reality to serve both clinical science and education, and we additionally discuss future prospects.

6.2 Introduction

With continued rapid advances in magnetic resonance (MR), computerized axial tomography (CT), and ultrasonic imaging modalities, there has been great effort put forth into transforming collected raw data (e.g., DICOM datasets) into formats that are more easily utilized and processed by physicians, patients, and/or computers. Initially these large images were comprehended by reducing them to a suite of parametric measurements, like anatomical volume and thickness, which could be used in conjunction with statistical analyses, similar to what was done for biomarkers[33]. Such methods are inexpensive, quick to utilize, and benefit from developments in mathematical fields, however a large portion of acquired spatial information is lost when casting these full scans into a small parameter set. Further, the inherent disadvantages of these induced reductions in dimensionality become readily apparent in imaging for surgical fields, where physicians must be keenly aware of the complex spatial intricacies of individual clinical cases. As a result, the field of medical 3D reconstruction has gained momentum, and advancements have trended more so with those in computer graphics.

Initially 3D medical reconstructions were intended to offer unique clinical models to be utilized by masters of anatomical illustration, to create idealized color images of associated structures for purposes of education and training. More recently, dramatic improvements in the speed of obtaining scans and the broad availability of computational clinical scans have enabled the broad adoption of new software tools like Mimics (Materialise, Leuven Belgium) and Slic3r (free access software, GNU Affero General Public License), which take image masks from DICOM datasets and convert them to 3D models. In other words, computer-generated models are a vast improvement relative to the artistic renderings created by medical illustrators, and further allow for reconstructions of personalized clinical cases (Salmi et al., 2012). These custom 3D computational reconstructions have important applications in both simulation and preprocedural (surgical) planning.

In this same era, the field of virtual reality (VR) and video game development has grown exponentially, aiming to present visual data to users while maximizing both immersion and realism. Investigators have already begun considering and applying VR for medical education (Chang et al., 2018) and surgical planning[42] however, to date, its adoption can be considered as narrow. Partially this is due to the relative expense of these technologies and also the necessity of expert anatomists to utilize these tools to create high-quality realistic models. Thus, there remain vast computational opportunities left untapped, particularly in the fields of medical device design, surgical planning, telemedicine, and medical education.

Within the Visible Heart Laboratories at the University of Minnesota, we have the unique privilege of receiving hundreds of fresh human hearts, heart/lung blocs, and cadaveric specimens which we investigate in collaboration with our extensive network of physicians, engineers, anatomists, and computationalists. The goal of this paper was to explore the novel use of human anatomical specimens for computational utility, including VR, within various medical contexts. Interactive and video demonstrations of our system are available at https://goo.gl/myFnSW, and we plan to release more interactive demonstrations in the near future.

6.3 Description

Donated human hearts considered nonviable for transplantation were received fresh from LifeSource (Minneapolis, MN, USA) within 4-20 hours of organ recovery. In some cases, these hearts were reanimated using Visible Heart® methodologies on an external perfusion apparatus, and internal functional anatomical images were recorded[14]. Heart specimens were perfusion fixed in formalin as a means to simulate an end-diastolic state. To enable high-resolution MRI or CT scanning, heart specimens were encased in agar gel (Figure 6.1).

For each heart specimen, DICOM scans were obtained and then imported into Mimics segmentation software to create high-resolution (0.1 mm) 3D models (Figure 6.2). We then incorporated these unique models into a video game engine, primarily using the Unity3D computational gaming engine (Unity Technologies, San Francisco, CA, USA); note that other engines such as Unreal Software would also work. For most of our VR applications, we focused on using the HTC Vive headset (HTC Corp., New Taipei City, Taiwan), but other headsets such as the Oculus Rift (Oculus VR, LLC, Irvine, CA, USA) could be utilized as well.

Further improvements to our generated VR scenes will in part depend on objectives for the planned visualizations. For instance, traversing large anatomy models requires mechanics for travel. One possible solution would be to implement this VR functionality by using the remote controller with the SteamVR package (Valve Corp., Bellevue, WA, USA) and Unity3D scripting application programming interface (API). We have utilized the scripting API for other interactive effects such as positional video playback. These generated Unity scenes can

Figure 6.1: Perfusion-fixed human heart (Heart 229) from the specimen library at the Visible Heart Laboratories. This heart was placed in agar gel, MRI scanned, 3D computationally modeled, and then simulations were incorporated into virtual reality scenes.

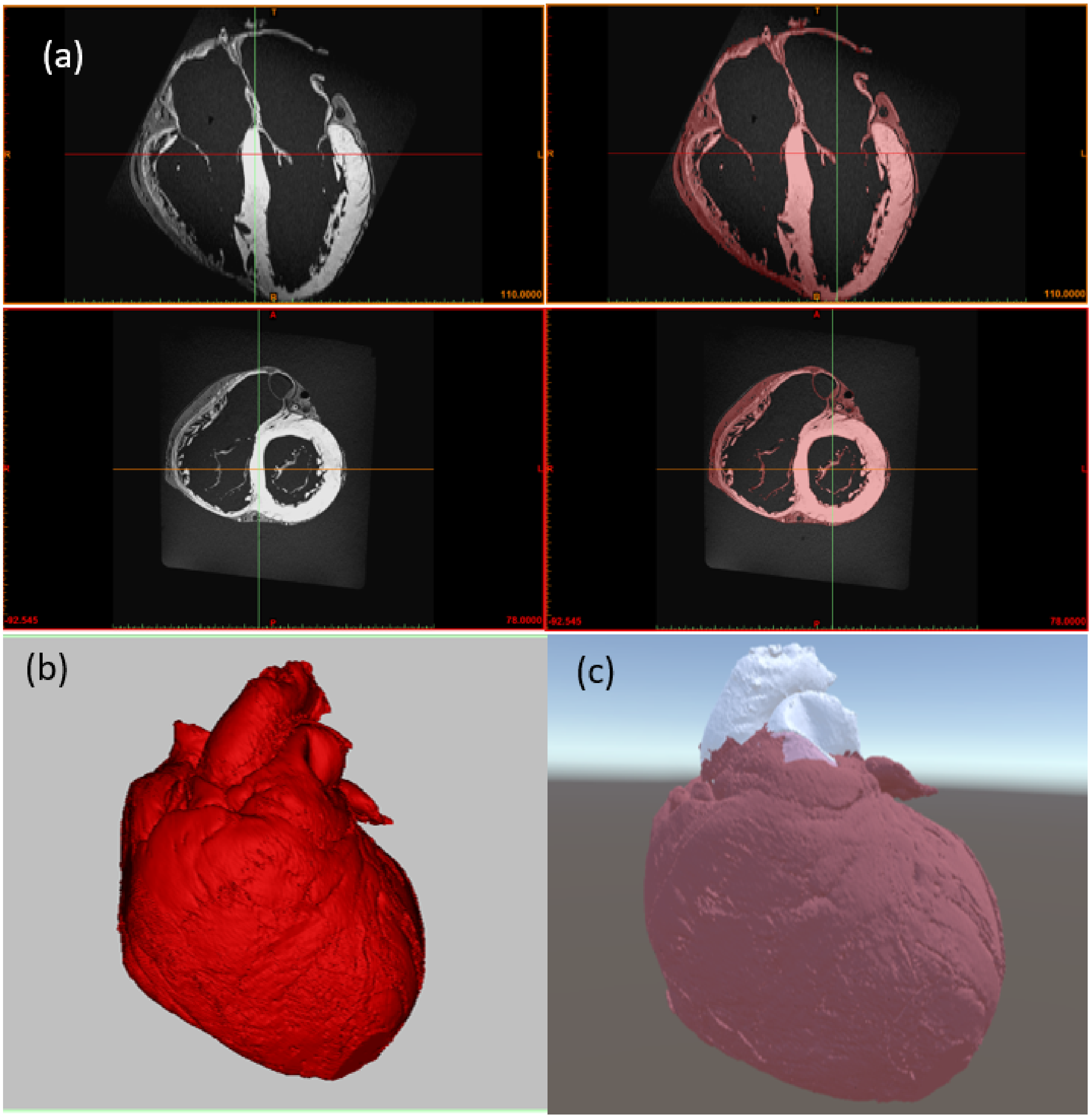


Figure 6.2: (a) DICOM scans were loaded into Mimics and detained anatomies were masked utilizing various thresholding functions. (b) 3D models of anatomy were generated from these masks. (c) Resultant 3D models were imported into Unity and a virtual reality environment of this specific heart was generated.

be readily exported as executable files that can be run on any computer with appropriate specifications.

6.4 Discussion

6.4.1 Education

Medical education in general can be considered a broad field with innumerable applications, and evaluation metrics differ markedly depending on the specific audience and purpose of training. To determine presentation needs, therefore, one must first define the user scope. For example, our laboratory has considered developing educational platforms for introductory anatomy students in either high school or college settings. For these students, images and cadaveric dissections provide the basic forms of instruction. In cardiac studies, this can be especially difficult as complex 3D fluid flow patterns typically must be considered within the context of static two-dimensional representations. By combining modalities, we can present educational models that edge closer to realistic clinical representations by embedding 2D videos of internal functional cardiac anatomies, appropriately placed inside high-resolution static heart models, which users can further explore. Individuals can obtain educational videos from the Atlas of Human Cardiac Anatomy (Visible Heart Laboratories, 2019), a free-access collection of videos and tutorials related to cardiac anatomy and physiology[25] (Figure 6.3).

It is important to note that one of the most vital components of these 3D models visualized in VR is the unique perspective gained when navigating freely within a cardiovascular model at the scale one chooses. For example, by observing the scale differences between the atrioventricular valves relative to the semilunar valves, one can obtain a unique anatomical appreciation. Further, because our laboratory utilizes high-resolution MRI scans of a

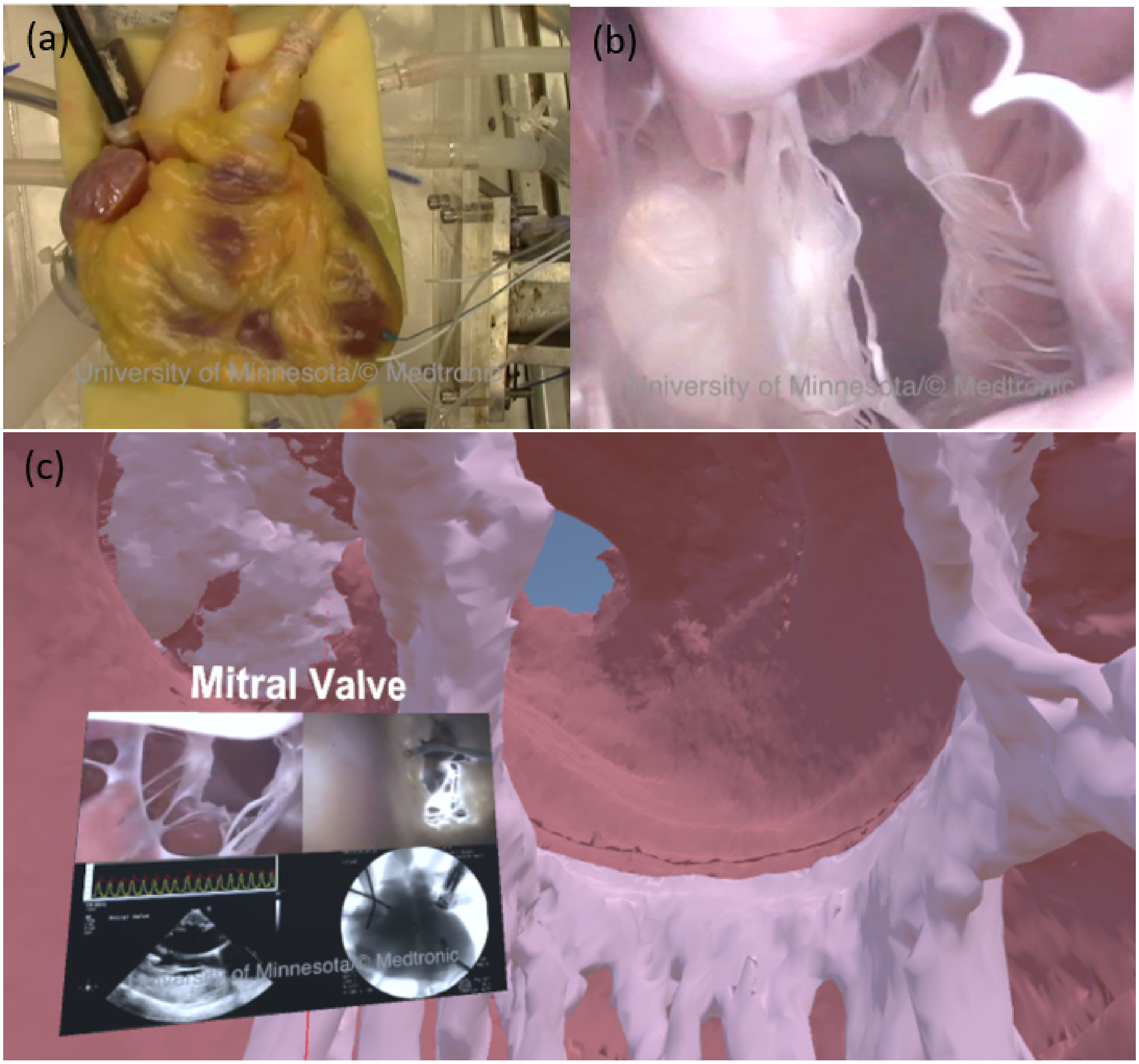


Figure 6.3: (a) External image of human heart reanimated on the Visible Heart® apparatus. (b) Videos captured with endoscopes placed in various locations inside beating hearts; this image displays both mitral (right) and aortic (left) valves. (c) Example of placement of functional internal anatomy video beside the virtual anatomy, to demonstrate how valves function within a beating human heart.

large population of human hearts, one can even navigate through a heart's vasculature; both coronary arteries and veins can be readily viewed and studied. However, this does not preclude the exploration of anatomical structures at realistic scales. As an example, we developed a whole body cadaveric model from detailed CT scans, with and without large volumes of contrast injected at multiple locations into the vasculature. Other companies and institutions have created VR cadaver models as well, however we feel that we possess a distinctive privilege to incorporate high-resolution images from hearts, heart/lung blocs, and whole body specimens gifted to us fresh and then carefully prepared for high-resolution scanning. In other words, we model real anatomies from both healthy and diseased human hearts, offering numerous advantages for a variety of teaching settings. Note that these 3D models can be smoothed or idealized (more like artistic renderings), which sacrifices some of the more intricate anatomical details but allows for other computational uses (see below).

Medical student, resident, and physician training can all benefit considerably from having accurate models of varied human anatomies. This is important because detailed models of individual anatomies uncover structures that can be frequently overlooked in idealized illustrations; no two human hearts are the same. Further, because natural idiosyncratic deviations from normal anatomies can be critical in diagnosing and predicting pathology, we have made efforts to develop and include 3D computational models from multiple hearts with different congenital heart defects as well as various levels of pathological hypertrophy. We have observed that human anatomical information presented in this manner can also be enhanced by audio or textual physiological explanations. To date, we have implemented numerous forms of textual representation, but there are opportunities to develop additional approaches for teaching and learning.

Our educational efforts have not been limited to the realm of academia, as we have developed educational programs specifically geared for the medical device industry, including specific domain knowledge pertaining to the problems and challenges engineers are tasked to solve. More specifically, we have generated 3D cardiac device models incorporating various devices in different positions within a variety of human heart anatomies. These VR visualizations have been extremely useful to understand and assess relationships between important anatomical landmarks and specific device interactions; device designers also gain insights relative to how cardiac anatomy varies across different patient populations. We have already added interactivity by allowing the user to manipulate the position of specific medical devices in a collection of generated models.

6.4.2 Presurgical Planning

Our approach to VR has already been applied to presurgical planning approaches. Cardiothoracic surgeons and interventional cardiologists need to be aware of anatomical intricacies within their patients to optimize individualized care plans. Importantly, exploring these patient-specific anatomies in VR applications can dramatically clarify important questions that may be relevant to making decisions on procedural strategies.

The visualization of pediatric congenital heart defects is one example where we have repeatedly demonstrated value in the preprocedural 3D modeling visualization realm; note that creating complementary 3D printed models also provides added value. Using these visualization experiences, we have enabled physicians to see intricate heart defects—to view them from inside out or outside in, to greatly enlarge them, or even to virtually modify them. Our lab has worked with collaborating physicians to create 3D models for identification and study of ventricular septal defects and co-arctations of the great vessels, and we even assisted with one case involving the separation of conjoined twins[22]. By utilizing VR visualizations combined with 3D printed models, we can go one step further: predicting how a specific patient pathology may interact with a virtually deployed device.

An additional benefit of preprocedural planning is the opportunity to consult with other physicians using virtual models who in turn may provide important/critical perspectives on challenging cases, thus improving the likelihood of better patient outcomes. We have already implemented a networking feature that allows one or more physicians to occupy the same virtual reality space, as well as to virtually point to anatomies/defects that might be clinically important. This approach might be particularly unique, as it allows for the paring of inspective freedom and user collaboration, even when consulting physicians may be located across the country or the world. Consequently, there is great opportunity for building better user interfaces to facilitate this rapidly developing field of telemedicine (Figure 6.4).

The 3D modeling and VR approaches we describe here are not limited in scope to the field of cardiology. For example, through whole body CT scans of either fixed or fresh cadaveric specimens, we have reconstructed an individual's full vertebral column. This allows a user to navigate the spinal canal without the cord being present, and observe the virtual spaces occupied by intervertebral disks as well as the arterial vasculature that feeds the intercostal muscles. This demonstrates the value of our methodologies in the field of orthopedics, and we have also started to make scenes for the fields of dentistry and neuroscience. Part of futuristic preprocedural planning will likely involve determining the optimal positions for device placement within the virtual patient, either in their present state or following therapeutic improvements (e.g., cardiac reverse remodeling). Further, we have begun to implement systems/scenes where a user can interact with objects representative of their device delivery systems. For instance, we have tested the manipulation of valves and leadless pacemakers in multiple detailed human heart models, however this only skims the surface for what is possible. By exploiting the spatial information provided by precise device placement, these visualizations can be used to simulate other relevant surgical image modalities. One could simulate fluoroscopic imaging by changing the relative transparency of a specific heart, thus creating a virtual guide map to be implemented surgically.

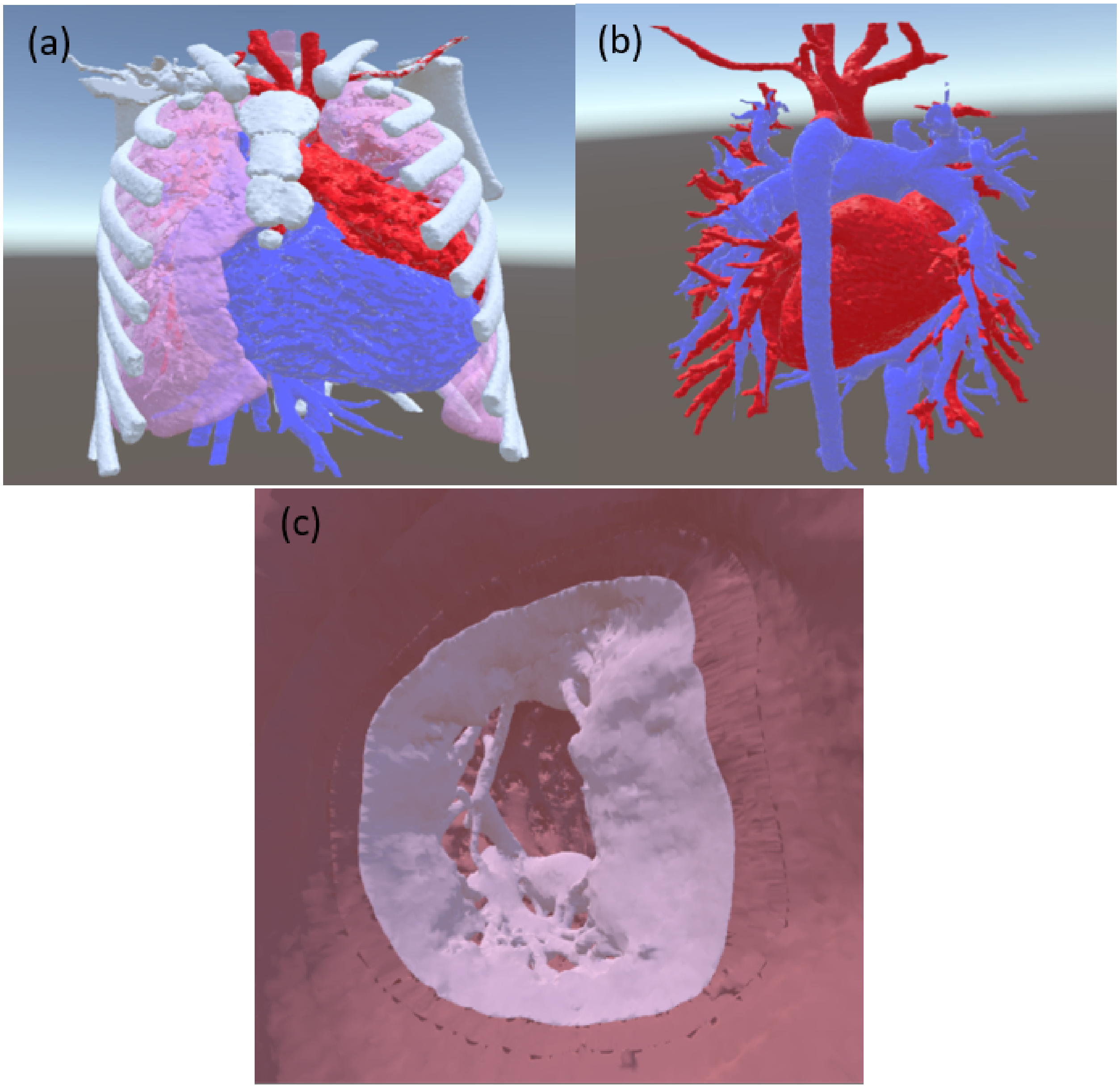


Figure 6.4: (a) Virtual reality scene of the anterior mediastinum of a congenital thoracic model depicting various anatomies, generated from a clinical CT scan. (b) Posterior view of the same congenital heart model with the lung and skeletal anatomies removed. (c) Typical scene of a mitral valve model we can examine in virtual reality; this view would allow a physician to accurately size the mitral valve annulus for either a virtual or subsequent procedural valve repair.

6.4.3 Devices and Industry

Medical device companies are also searching for novel methods to visualize anatomy, for both development and physician training purposes [50][25]. These applications can be exploratory or goal oriented in nature, involving measurements of delivery spaces and optimizing final positioning of delivered devices. Product development can be costly, so it is within the innovative team's best interest to use as much existing knowledge as possible to create functional or virtual prototypes. Many groups around the world are beginning to utilize virtual or augmented reality (AR) approaches to enhance users' spatial relationships between real anatomies and purported device designs. The Food and Drug Administration, National Science Foundation, and National Institutes of Health have hosted numerous joint computer modeling and validation medical devices workshops. Further, the Medical Device Innovation Consortium (MDIC) and the Earl E. Bakken Medical Devices Center (MDC) at University of Minnesota, under the Direction of Professor Art Erdman, have endorsed the long-term vision of developing technology platforms that can support a paradigm shift in simulation-based design of medical devices. These have included not only high-performance computing, but also the generation of collaborative design environments that utilize multidimensional data visualization technologies and advanced human-computer interfaces.

Our laboratory works closely with the MDIC and MDC to create demonstrations showing unique structures within numerous human cardiac anatomies and devices such as catheters, leadless pacemakers, heart valves, and left ventricular assist devices. In one approach for generating a medical device scene, we obtained image sets from patients with prior implanted devices, yet these devices sometimes induce significant amounts of artifacts. To alleviate this problem, we can either use domain knowledge to manually segment only what we believe to be the device, or we can use more appropriate modalities for scanning. One example is the use of micro-CT scanning to secure high-resolution imaging of stents. On the other hand, creating a complete heart scan using micro-CT is not data efficient, but we can utilize 3D models/scenes created from a different scanning session and apply them to other 3D models. To date, using this approach we have generated a variety of different stenting bifurcation technique scenes within the coronary vasculature of the non-treated heart (Figure 6.5).

Continued exponential advancements in video game design and tools imply that our educational anatomical approaches will also improve with the benefits of intractability and immersion. Interfaces, graphical user interfaces (GUIs), and controllable objects are all programmable within the frameworks we have developed. Further, our initial approaches to these educational strategies allow for medical devices to be visualized from within numerous real human heart anatomies, to be translated and rotated with an effort to assess positioning. We have also made significant progress in recapitulating the deflections seen in catheters that can be controlled within VR by the user. Yet, numerous challenges remain in order to take full advantage of physics engines to mimic collisions of delivery systems with tissue.

6.5 Future Directions

Here we describe many of the advantages of VR pertaining to a specific subfield of medical innovation, benefits that could carry over to other innovation fields as well. For instance, we discussed using VR networking for telemedicine applications between physicians for presurgical planning, yet VR or AR networking would be useful in an educational context as well; we have started to develop networking with other innovation centers in the United States. This emphasizes the broad opportunity available for VR software within an array of applications within the medical field, including education, training, procedural planning, and/or virtual prototyping to name a few. This paper aimed to broadly cover some of the main applications we have started to develop within our laboratory, but it should be noted that there are many others that are outside the scope of this discussion. We expect to see more utilization of VR

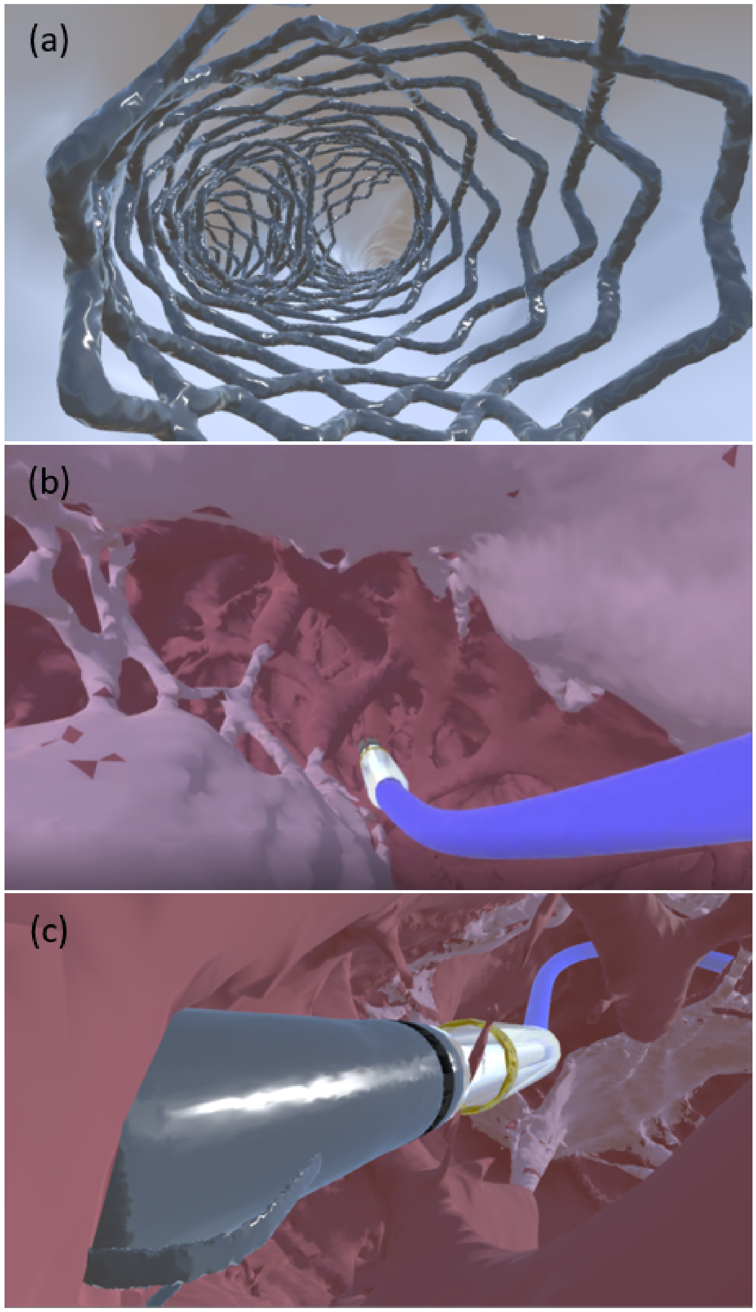


Figure 6.5: (a) Culotte stenting technique within a human coronary artery bifurcation. Actual stenting procedure was performed by a physician in a swine heart using Visible Heart® methodologies; stents were imaged with micro-CT, modeled, and then digitally placed within a human heart with similar coronary bifurcation anatomies. (b) Placement of leadless Micra pacemaker delivery tool system (Medtronic, Minneapolis, MN, USA) within the right ventricular apex of a human heart 3D model; image was taken from within the tricuspid annulus. (c) Depiction of Micra delivery taken from the right ventricular apex.

software and equipment within the medical field as these technologies become cheaper and more mainstream. Consequently, other fields such as artificial intelligence will benefit greatly from modalities that allow for user-based anatomical navigation. One problem, however, that deserves great attention is the simulation-visualization divide. Virtual reality provides an easy way for the user to manipulate devices, but linking up these interactions with physical models has so far been largely unexplored. Advances in this field would expedite many steps within the prototyping and perhaps even the clinical trial pipelines. Nevertheless, the entire field of medical VR is open for many types of innovations, with unlimited opportunities for academicians, clinicians, and industry practitioners.

6.6 Acknowledgments

The authors wish to acknowledge the organ donors and their families, LifeSource, and the University of Minnesota Bequest program. We thank the MRI and CT staff at the University of Minnesota Medical Center. The authors report no conflicts of interest.

Chapter 7

Immersive Anatomical Scenes that Enable Multiple Users to Occupy the Same Virtual Space: A Tool for Surgical Planning and Education

Authors:

Alex Deakyne[1],Erik Gaasedelen[1], Paul Iaizzo[1]

[1]Department of Surgery, University of Minnesota

Preface

This work, in my view, holds considerable potential for commercialization, epitomizing the integration of current technology while establishing a basis for subsequent innovations. I anticipate these tools becoming integral to comprehensive robotics and visualization frameworks. Although this project initially focuses on using networking for human remote communication, the concept extends to various subsystems within a larger technological entity. For example, it involves tracking the motion of surgical tools and integrating this information with pre-existing cardiac anatomy maps.

My involvement began with the original concept, followed by significant contributions to the implementation. Our lab experience, particularly demonstrating various anatomies through VR to physicians and students, highlighted a limitation: the necessity for physical presence at the Visible Heart Labs to engage with our work. This observation sparked the development of a system that significantly enhances current communication capabilities.

Currently, the most immediate and compelling applications of this technology lie in education and pre-surgical planning. However, I am particularly interested in extending these applications to the field of robotics. In this domain, the Robot Operating System (ROS) offers an asynchronous framework for collecting and processing sensor data, which represents an area ripe for further research. I advocate for increased interdisciplinary collaboration, leveraging the Visible Heart Labs' proficiency in physiology, mechanical engineering, and computer science. If such integration is feasible anywhere, it would be here.

7.1 Abstract

3D modeling is becoming a well-developed field of medicine, but its applicability can be limited due to the lack of software allowing for easy utilizations of generated 3D visualizations. By leveraging recent advances in virtual reality, we can rapidly create immersive anatomical scenes as well as allow multiple users to occupy the same virtual space: i.e., over a local or distributed network. This setup is ideal for pre-surgical planning and education, allowing users to identify and study structures of interest. We demonstrate here such a pipeline on a broad spectrum of anatomical models and discuss its applicability to the medical field and its future prospects.

7.2 Introduction

As the numbers of distributed internet technologies grow, innovation increasingly aims to solve the problem of allocation of resources. Depending on the application, such technologies can be used to serve more users by reducing the effective distances between the distributor and customer. The global physician population, who's specialization has become more narrow and complex, is one such example where such technological applications have been under-optimized. Yet, proper medical care is frequently reliant on primary physician, consultants, specialized imaging equipment, and a given patient being present in the same place. This naturally leads to the question of how to decouple these spatial dependencies while maintaining or even enhancing the 'quality of care'.

To date, companies participating in the telemedicine market have primarily focused on developing smartphone applications which have allowed the physician-patient interview to take place over large distances. These applications can increase their analytical capabilities by acquiring data from sensing devices such as stethoscope, EKG, or EEG. For fields like interventional cardiology or surgery, a fully remote interfaces have yet to become applicable for tasks which require high human dexterity. However, physician collaborations has been improved through technologies like Google Glass, which can stream a live clinical procedures to a remotely consulting physician. But typically to date, these kinds of interactions force the remote physician to have a narrower perceptual awareness of the procedure, utilizing video with a small field of view and, presumably, a written preclinical plan. Consequently, greater technological utilizations are not yet commonplace.

Today advancement in medical imaging have been coupled with the field of radiology to successfully allow for the practical separation of patient diagnoses into remote observations and distributed interpretations. Yet to date, for example, when such imaging approaches were so to be used for surgical planning, the utility of remotely consulting physicians can become diminished. We propose that this is because descriptions concerning spatial awareness requires a greater communication bandwidth between physician and consultant, placing a cognitive strain that prevents widespread adoption. This could be much easier, however, if there were modalities that could supplement spoken communication with rich spatial awareness and gestural abilities.

The fields of virtual and augmented realities are rapidly maturing even within the timeframe of months, driven primarily by the video gaming market. A wealth of tools dedicated to creating effective virtual reality environments has followed, even allowing developers to incorporate internet-based multiplayer features into these scenes.

Virtual and augmented reality systems have differences in how they display information to the user. Virtual reality utilizes a headset which completely occludes the user's native field of view, replacing it with high resolution screens that create an artificial space, based on the headsets spatial coordinates. Augmented reality, in contrast, allows the user to maintain their environmental awareness of the real world, by instead projecting holograms in their native space. Depending on the applications and constraints set by AR/VR systems, each can have their own advantages and disadvantages.

Meanwhile, the field of medical 3D modeling continues to boom, aiming to allow physicians to have a more intimate understanding of their patients' relational anatomies. This creates a natural progression to combine the advancements of virtual reality and anatomical 3D modeling, into a single pipeline. Researchers have already considered VR for surgical planning[42]. However, its adoption has been limited in part because there are not many tools yet available to take advantage of these modalities. However, by combining networking functionalities, 3D clinical models, physical prints, and virtual reality, we can fulfill many of the requirements set above for an effective distributed surgical preplanning and consultation tool: a mixed reality approach. We propose that developments of such technologies will not only allow for more effective communication between physicians, but will provide a substrate for novel collaborative pipelines in both patient care and education. (Figure 7.1)

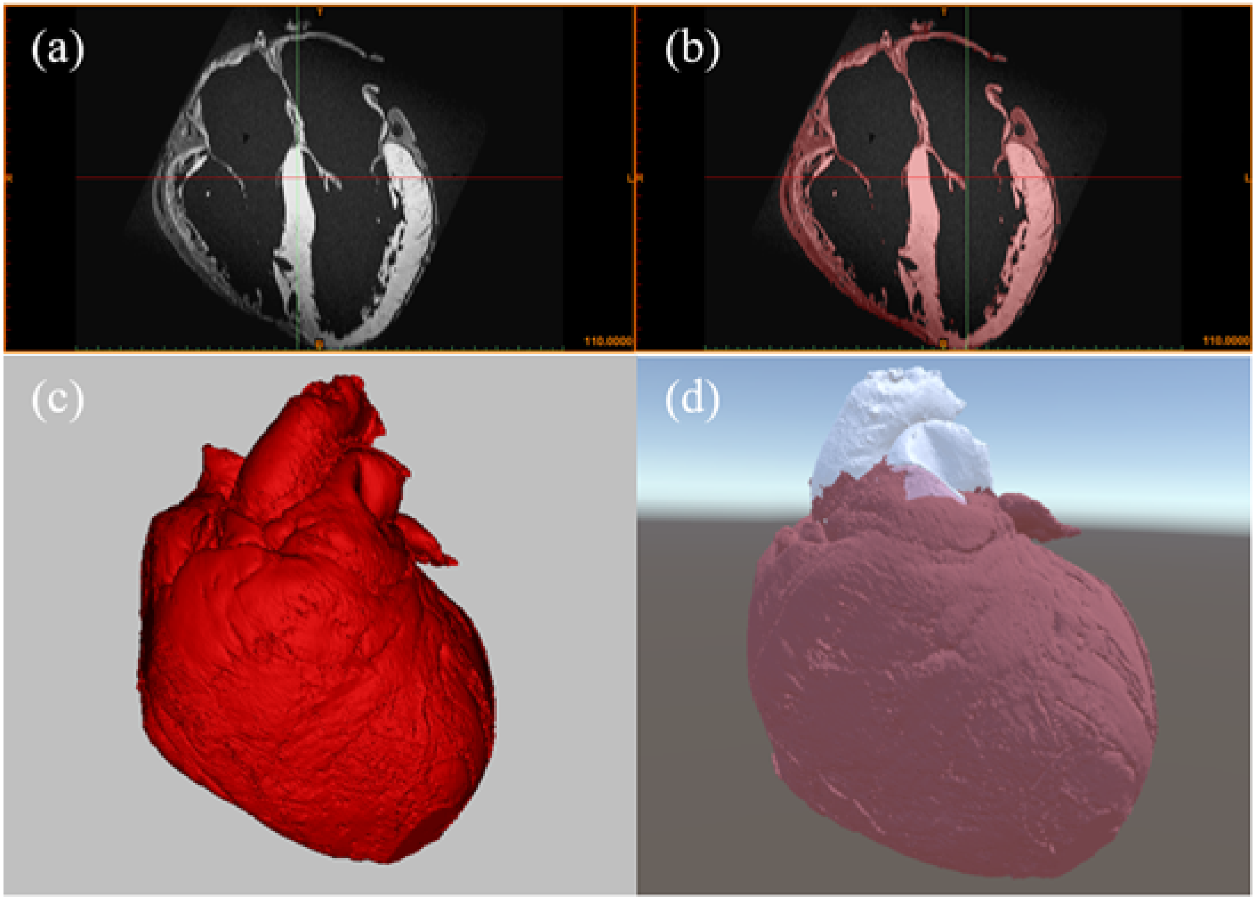


Figure 7.1: Flowchart diagram of proposed mixed reality methodology. (a) A DICOM scan in MIMICS depicting a cross sectional view of a heart. (b) 2D mask of the heart tissue generated through thresholding functions. (c) The resulting 3D model of the heart created by the 2D masks which can then be 3D printed. (d) The same 3D model of the heart is imported into Unity and a virtual reality environment of this given heart is generated.

7.3 Methods

7.3.1 3D Model Development

Voxel-based medical imaging can be retrieved from any compatible CT or MRI scanner: these are required to create corresponding 3D models of the patient's anatomy. Yet, these differing modalities can affect the resultant qualities of the final models because of differing contrasts and resolutions. In our research group, we can scan static hearts that have been explanted using MRI. We can also visualize scans from cadaveric sources in which we can inject contrast to better model the vasculature. This does not affect the technical contributions of this report, but shows the diversity for which this process may be applicable.

Segmentation of various obtained medical imaging was performed using the Mimics (Materialise, Leuven Belgium). This process was characterized by the creation of masks which either represented the complete model or specified anatomical features. The creation of these masks typically required exceptional care, as those who are not experienced with medical imaging and anatomical identification may produce different results. Additionally, it is also required that anatomical structures which must be either manipulated or colored uniquely within the virtual reality environment, thus must be segmented with its own mask. 3D models can be generated from these masks, but to achieve realistic models, some post processing was typically performed. This commonly involves filling of holes, or the smoothing of rough edges. The validation of such operations and their effects on anatomical realism is an active field of investigation and advances will provide immediate improvements to our pipeline.

7.3.2 Required Hardware and Software

To allow these models to be intractable in virtual reality, the Unity3D (Unity Technologies, San Francisco, CA, USA) video game engine was used along with the SteamVR API (Valve Corp., Bellevue, WA, USA) for the HTC Vive . Yet, neither Unity3D nor the HTC Vive were solely required for the creation of such virtual reality scenes. For example, the Unreal Engine (Epic Games, Cary, NC, USA) and the Oculus Rift (Oculus VR, LLC, Irvine, CA, USA) are both competing technologies which offer as a comparable alternative. Nevertheless, interactive components required the creation of scripts which allowed the user to change their coordinate position or navigate to a different virtual environment. For our setup, a computer with a capable graphics card and CPU were necessary to render objects which were intractable in real time. In our case we used an Nvidia GTX-1070 graphics card with Intel core-i7 7700k CPU.

7.3.3 Development of Multiplayer Functionality

Multiplayer functionality required the specific programming of network features. For this we used the Photon Bolt (https://www.photonengine.com/bolt) networking engine which provides simple abstractions for monitoring remote states. To properly set this up within a virtual reality scene, multiple steps had to be taken. First, an avatar was created that would represent each player's headset and controller to every client connected to the server. In this case we simply used a circle and rectangular prism to represent a face with a visor, and a rectangular prism with laser to represent the controller (Figure 7.2).

A custom script was created that spawns an avatar on the server whenever a user enters a new scene. This script utilizes hooks for the headset and controller that allows them to be called when they first appear in a scene. Once they can be referenced, an avatar is instantiated on the server at the user's current position and was replicated to all other connected VR users/clients. This script also sets the avatar transform to that of the user, ensuring the transform information of the avatar was always the same as the user's headset and controller.

Next, a network state must be created for our avatar to define its network properties. This was easily setup through Bolt. In this project, we needed to further define the transform as a property of our avatar's state. We set the transform to be replicated to everyone so that a user's avatar was replicated to every client connected to the server (Figure 3a). Lastly, we add a Bolt Entity component to the avatar and set it equal to the transform state that we had previously defined. Now, once the avatar was spawned on the network at the user's location, any local movement done by the user was replicated by their avatar to all clients

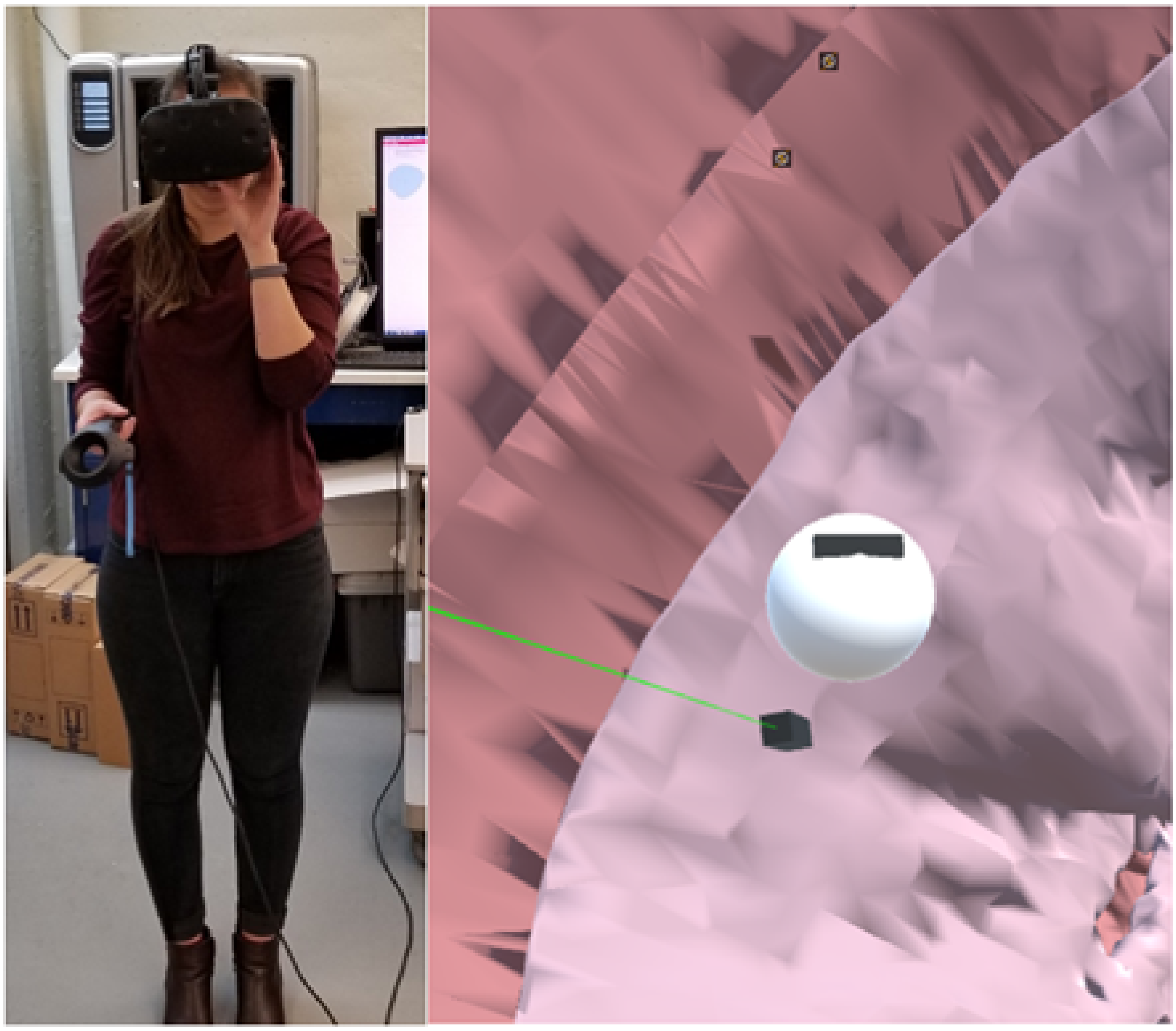
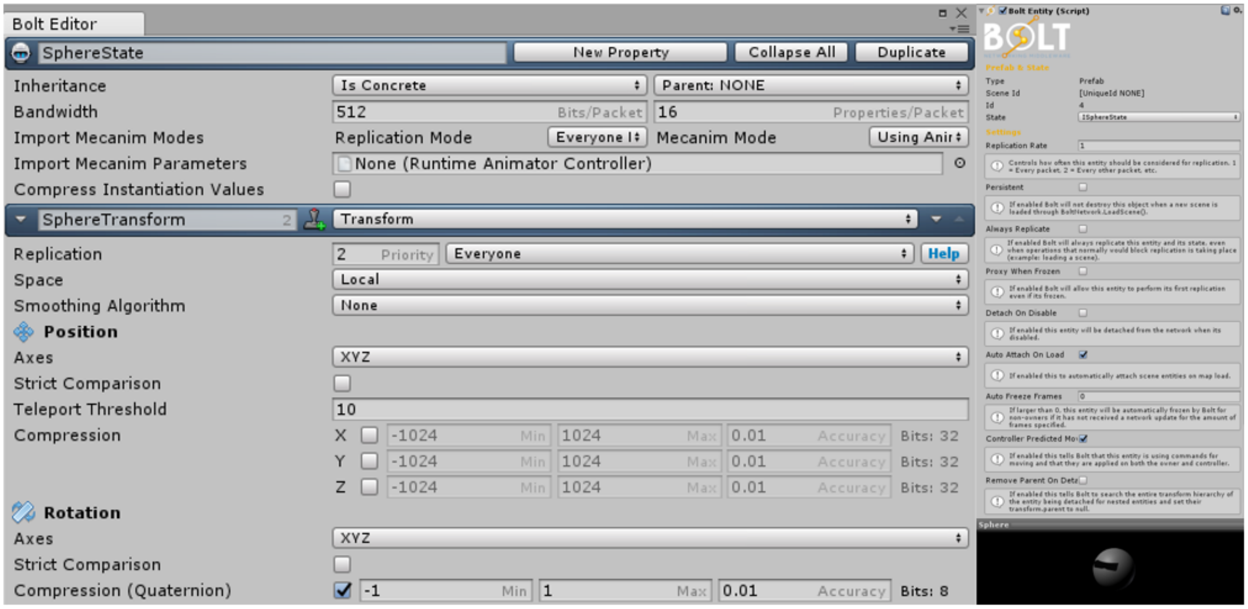


Figure 7.2: Image of a user wearing VR headset next to their virtual avatar.

connected to the server (Figure 7.3b).



(a) (b)

Figure 7.3: (a) The network state created for our avatar. This state contains the sphere transform property we defined. (b) A bolt entity was added to our avatar, defining its network state.

7.3.4 User Interface

A VR scene was created that displays a Menu. This user interface (UI) allows users to select different virtual scenes to explore (Figure 7.4). A script was created that utilizes the laser pointer on the controller to select a scene to enter. The laser is a ray-cast object extending from the tip of the HTC Vive controller and the buttons on the menu detect collision with the laser. When a collision occurs, the button becomes opaque.

A script and event handler manages the functionalities of users entering different virtual environments. An event handler is an empty game object that can be referenced from a user interface button in a unity scene. A script with custom methods was attached to the game object that specifies what happens when a button is selected from the menu (Figure 7.5).

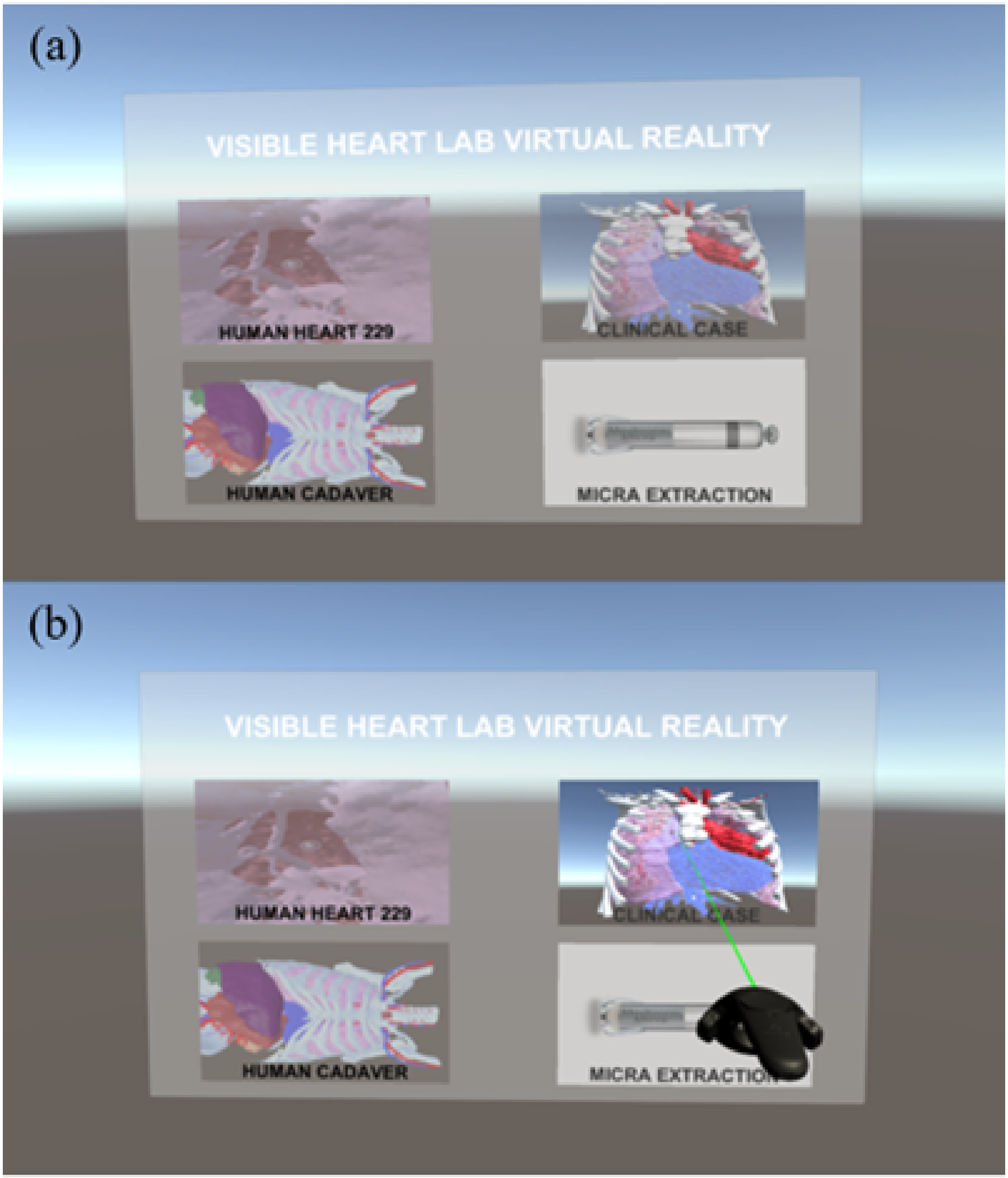


Figure 7.4: (a) The menu displaying the scenes the user can enter. Depicted in (b) is the change in transparency of the button, highlighting a collision of the laser and the current button.

Clicking the back trigger on the Vive controller while highlighting a button calls a method in the event handler that starts the server (if one has not been created) or connects to the server (if a server has already been created). Once a server has been created, or connected to, the even handler loads the selected scene. Once within the scene, the user's transform information is replicated by their respective avatars to every client/connect user. A given user can seamlessly return to the main menu at any time by pressing the menu button on the Vive controller. This removes their avatar from the scene, but has no other effect on the remaining clients.

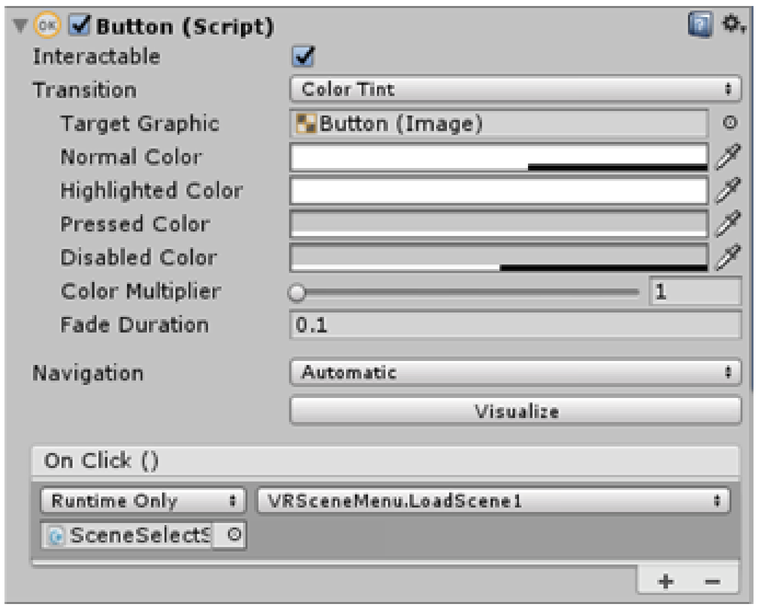


Figure 7.5: Method for UI Button Scene Selection

Finally, a custom script was written that allows the user to move throughout a given virtual reality environment. This script allows the user to move toward the laser on their controller by pressing down on the top half of the steam controller trackpad. Consequently, the user can also move backward by pressing the bottom half of the pad.

7.4 Results

7.4.1 Anatomical Education

The developed system offers many benefits for medical and anatomical education or physician interactions. Work by others has already been done to study the effect of virtual reality on medical learning outcomes[13] with promising results. Currently, most students learn anatomy by reading textbooks and studying two-dimensional pictures. This is disadvantageous because you lose any three-dimensional spatial relationships of these complex anatomical features. The virtual reality system I am developing here within the Visible Heart Laboratory, offers an accessible, immersive environments to learn complex human anatomies. Since the system user is in a three-dimensional environment, they can learn both varied anatomies and the spatial relationships of anatomical features simultaneously. The scale of the given anatomical models in the virtual scenes can be increased, even to a scale where the user can fly through them. This allows users to easily study the critical details of anatomical features.

As one system example, we created a virtual reality environment that contains a 3D model of a human heart (Figure 7.6). The scale of the human heart was increased so the user can easily fly through and explore the whole endocardial and epicardial features of the heart, granting the freedom to analyze features however they please. We believe and have obtained numerous user feedback from students, residents and experienced physicians that new insights can be gained by viewing these cardiac anatomies in this unique perspective. The laser pointer can highlight anatomical features to others in the scene, allowing for real-time instruction and collaborative discovery in an immersive virtual environment.

We can also create scenes which are closer to surgical realism as well. For example, we have placed a model, created from a full human cadaver CT scan, into a virtual environment.

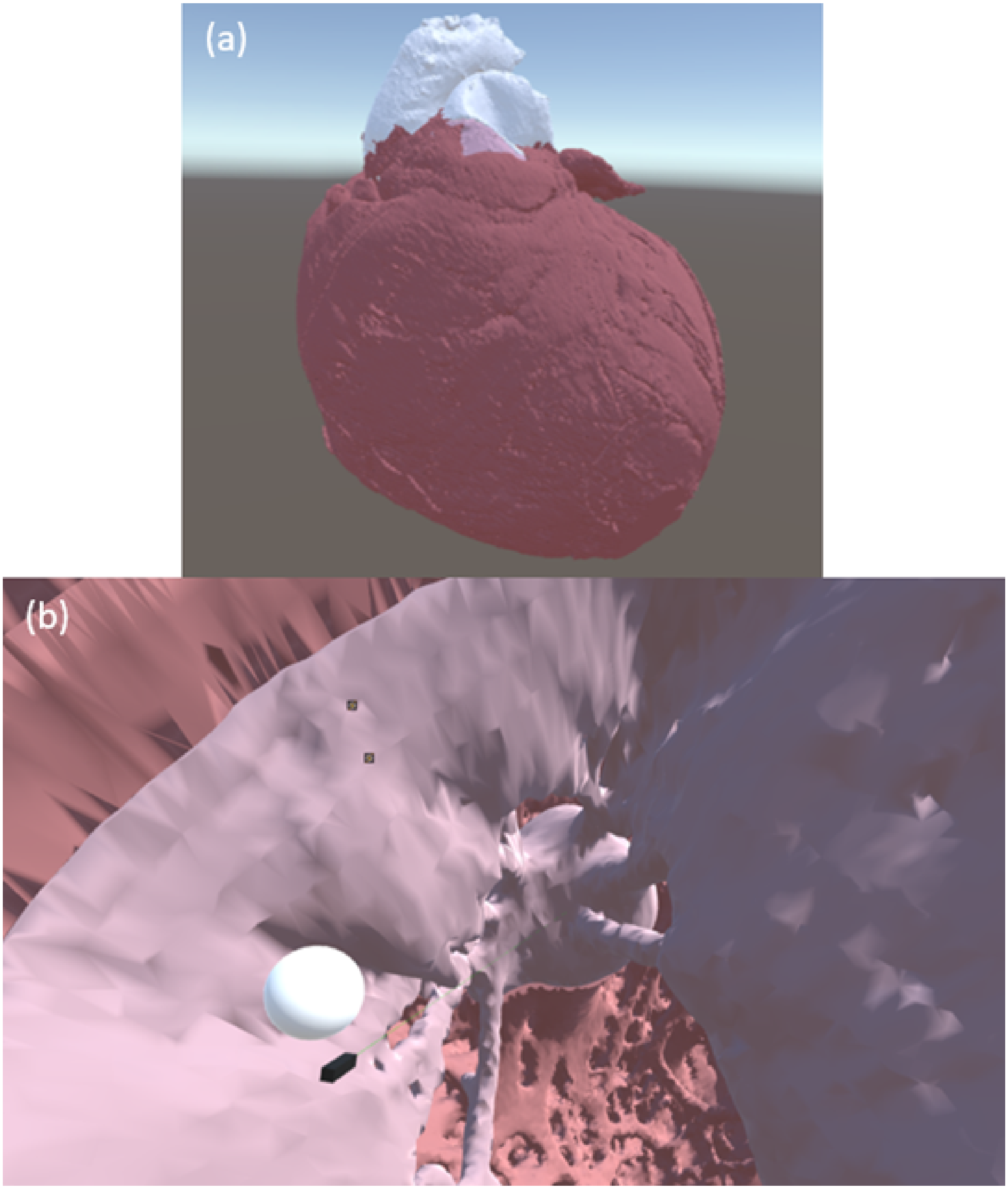


Figure 7.6: (a) An external view of a computationally modeled human heart in the virtual reality environment. (b) Shown here, two users in the left atrium are studying the anatomy of the mitral valve. The instructor in the scene is using their laser to highlight the large papillary muscle while explaining its function to the other users.

This allows users to move around this cadaver at native scale as if it were lying on a surgical table (Figure 7.7a). This offers information on the actual spatial relationships between the many anatomical features modeled. Given the complex nature of human anatomies, allowing for a paired student-teacher interface can be highly instructive; implicating this setup as an education system as well. Further, scaling functionality would also advantageous in the teaching paradigm. For example, users could enlarge the size of the cadaver to the point where users can fly through the spinal canal (Figure 7.7b).

7.4.2 Pre-Surgical Planning

For pre-surgical planning, currently the most relevant applications involve patients with relatively complex anatomy. For example, pediatric cardiac surgery related to congenital heart defects is one such example; due to both their reduced sizes and required complex repairs. To demonstrate this, we modeled a clinical scan from a pediatric patient with a large ventricular septal defect and inversion of the great vessels (Figure 7.8). This was created from pre-procedural clinical CT angiography, allowing for a clear image of the patient's blood volumes in the given heart chambers. From this perspective, it is easy to see the persisting patent ductus arteriosus as well as the ventricular septal defect. The virtual environment can be manipulated, allowing for certain objects to be shown or hidden. In this scene, the clinical care team can view the heart with the skeleton and lungs modeled. This offers an immersive look at the relationships between this patient's heart and other anatomical features. The rest of the abdomen can also be hidden, offering an unobstructed look at the heart anatomy and clinical features of the congenital defects. This model can also be arbitrarily scaled allowing subtle anatomical features to be more easily visualized.

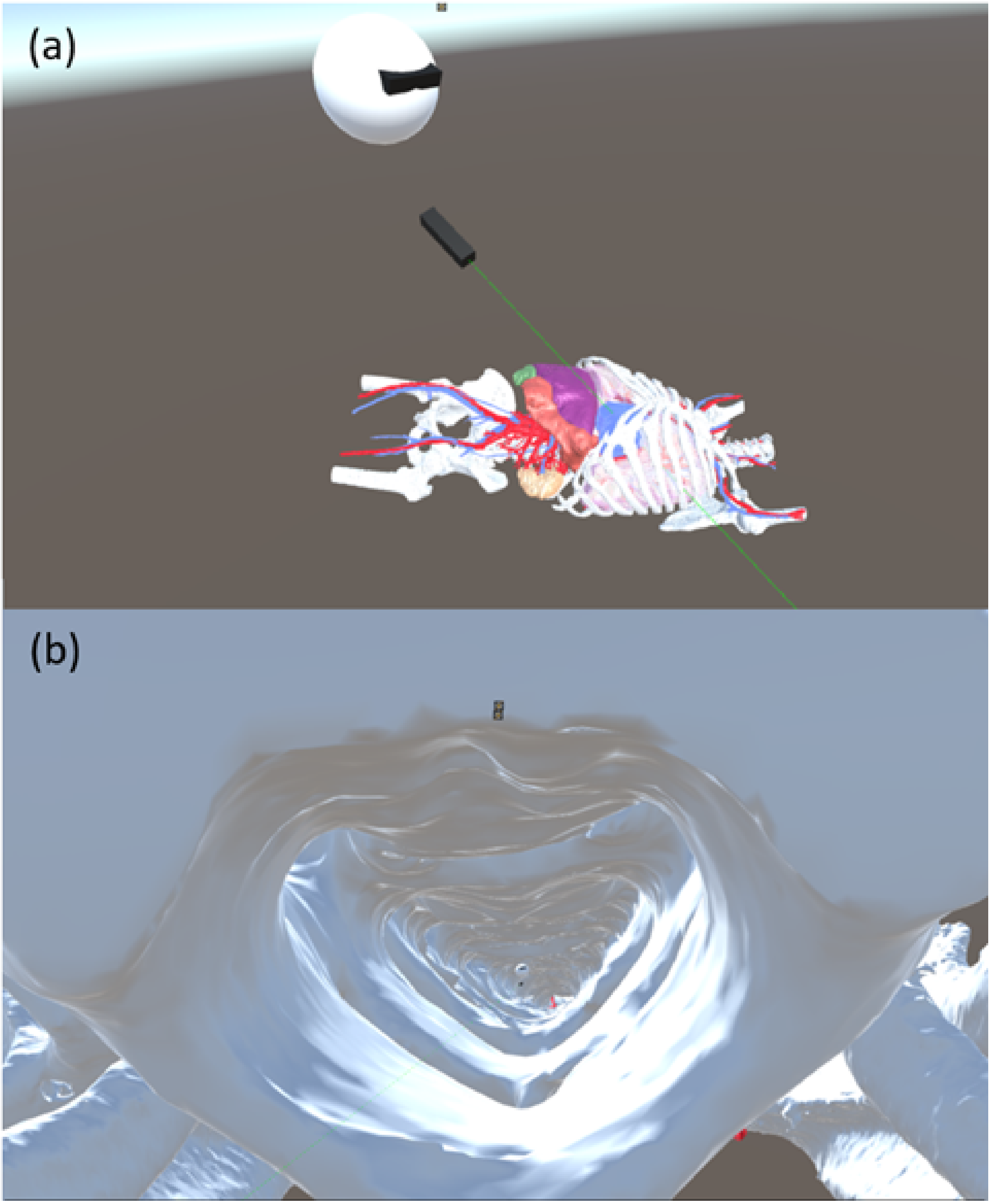


Figure 7.7: (a) An example of an instructor teaching anatomy to others on a cadaver model, at native scale. (b) The scale of the cadaver has been enlarged and users can be seen navigating within the spinal canal.

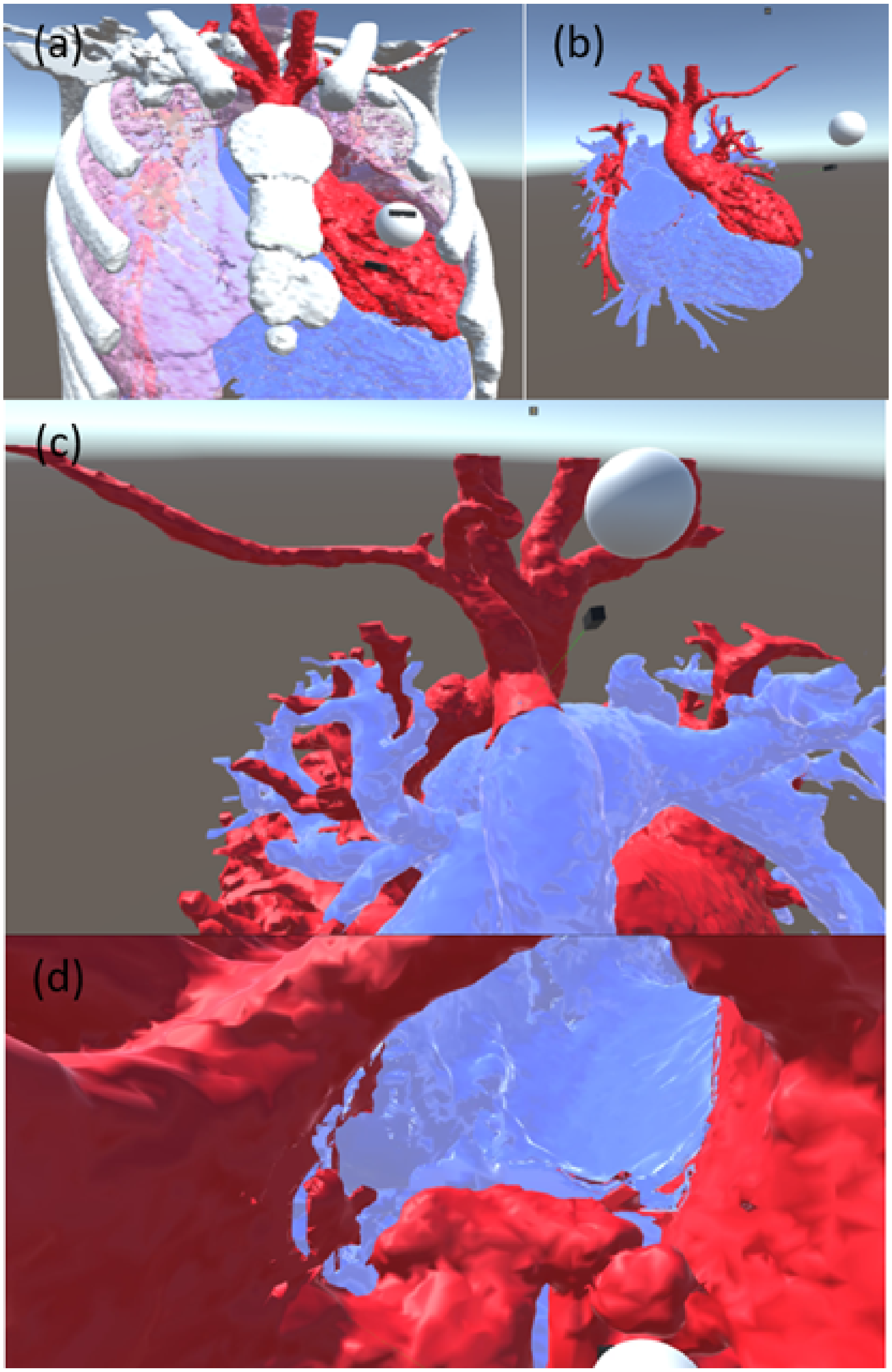


Figure 7.8: (a) External view of a computationally modelled pediatric congenital heart disease case. (b) The lungs and skeleton have been removed to better depict the features of this congenital heart. (c) The clinical care team can analyze the inversion of the great vessels. (d) Collaborators standing in the right ventricle study the ventricular septal defect.

7.4.3 Medical Devices

Frequently, understanding the detailed spatial relationships between medical devices and patient anatomies can be informative for both advancing medical device refinements and utilizations. This pipeline can use models obtained by imaging of anatomies with devices implanted or can also allow for devices to be computationally implanted in detailed anatomical models, and then either can be analyzed by collaborators in virtual environments. More specifically, medical device collaborators can enter a virtual environment to view their device interacting with the anatomy in these novel developed scenes. Biomedical Engineers can also utilize this collaborative methodology for proposing their device to physicians, either explaining the planned procedures in an immersive environment or to receive crucial feedback for improving their devices. To demonstrate this, we have compiled a 3D human heart model with a computationally added Medtronic Micra and extraction tool (Figure 7.9). This environment depicts a Micra extraction procedure[49]. This modality yields a new perspective for studying this emerging procedure; i.e., the devices involved, and their interactions with a given patient's cardiac anatomy.

Maintaining acceptable network performance is critically important for allowing and maximizing these collaborative interfaces. Luckily, our developed pipeline only requires management of simple transform information, making data transfer overhead relatively minimal.

7.5 Discussion

The ability to visualize, manipulate, and synchronize environment states with multiple participating users has applications for a large class of both educational and clinical problems. Broadly, it represents the novel abilities to distribute more salient information for faster human processing and communication. Pre-clinical planning and education represent a couple

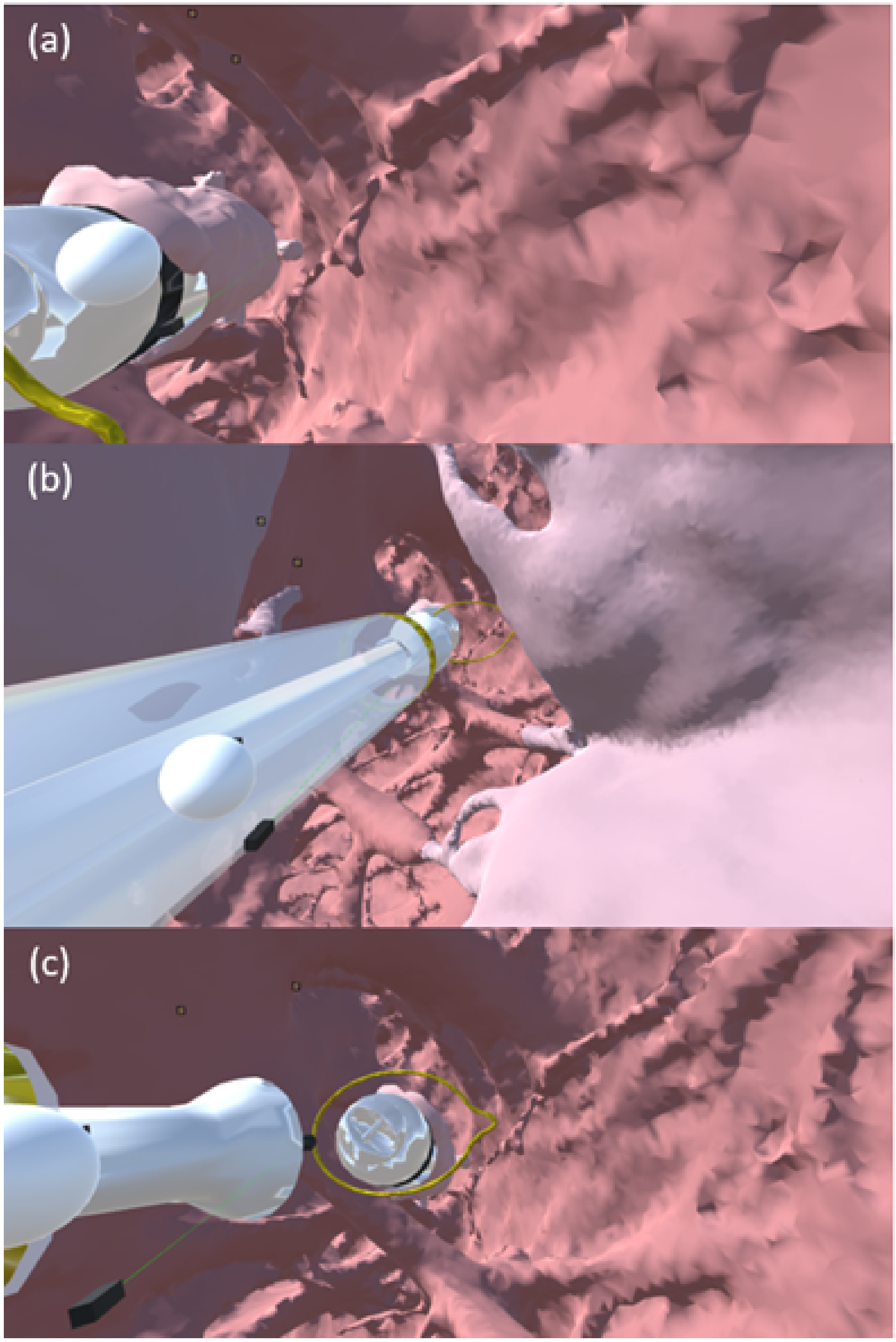


Figure 7.9: (a) A collaborator (e.g., medical device developer) showing the encapsulation of the Micra tines in the right ventricular myocardium to other collaborators within this unique human heart environment. (b) Collaborators in the right atrium looking at the Micra and extraction tool through the tricuspid valve. (c) A user in the right ventricle highlighting the lasso on the distal end of the extraction tool that is employed to snare the Micra.

of important examples that are the most approachable given the current state of technology. However, there are technological limitations that prevent optimized utilizations of such a system. Segmentation is a time-consuming task, requiring the skills of highly trained anatomists, and characterized by uncertainty regarding human variance. Scaling of such a system would require greater optimization of this task through advances in computer vision or artificial intelligence. It is also unknown to what extent post-processing creates models which diverge from the morphologies of their native anatomies. More research into validation schemes will also be necessary to overcome this problem. Nevertheless, all generated scenes can be shared as educational tools for those who could benefit from their unique applications. Yet, if one can fill these dependencies, frameworks for clinical collaborations will be more feasibly scaled.

These applications are not limited to preclinical or anatomical educational tasks. The emergence of the field of surgical robotics, although in its infancy, has the difficult objective of creating systems which can localize and navigate human anatomy. Ultimately, this will require algorithms which can understand and act upon real three-dimensional spaces. However, gaining information about human anatomy is difficult without the availability of training data. From this perspective, our setup has utility for making progress for such in two main ways: it creates an interactive system for a computer agent to explore, and it creates a substrate for live 3D updates during a purported intra-surgical procedure. Although this has ultimate utility for automated technologies, more informative navigation is immediately useful for practicing physicians as well.

In the field of 3D anatomical modeling, it is frequently useful to take measurements of specific structures to inform the proper selection of a medical device to be employed. For example, selection of an artificial heart valve could be informed by the detailed analyses and 3D understanding of the dimensions and relative shapes of the patient's native annulus. Other specialized 3D measurement tools would be important in a variety of fields such as neurosurgery, orthopedic surgery, oral surgery, and more. Given the flexibility of tools present within the Unity3D or other video game engines, these technologies could easily be incorporated to increase utilities. Yet, additional discussions relative to such future work within each respective field is beyond the scope of this thesis.

To date, our laboratory has not attempted to share any information across the network other than transformation information. However, collaboration is highly dependent of vocal communication, which we do not provide. Yet, this can be solved simply by using a phone or voice chat app in tandem with our system. There are also solutions which allow voice data to be shared easily over the same network which was used for transforming data. We again leave refinements of this system to our group's future work.

While the field of virtual reality and augmented reality has progressed exponentially, their mutual utilizations by multiple uses has been minimal. Yet, recent projects have developed scenes which allow a Vive and HoloLens (Microsoft Corporation, Redmond, WA, USA) to occupy the same set of spatial coordinates. Such a setup has promise for taking advantage of the benefits of either or both virtual and augmented realities. This would allow for collaboration between modalities or even for a HoloLens user to manipulate objects using Vive controllers. This has already been implemented as a demo, but their utilizations in the medical field have so far absent. As new products emerge from this market, more possibilities, like this, will become available, expanding the repertoire for how visualization problems will be solved.

7.6 Conclusion

Uses of virtual reality as a tool in telemedicine is a new and promising technique for the facilitation of physician collaboration and education over long distances. These technologies will permeate many aspects of medical inquiry, including live surgical cases, but there are many technological dependencies that must addressed before their uses becomes ubiquitous. This paper outlines just one necessary contribution so to expand these 3D environments. Fortunately, the breadth of applications for this technology is extensive enough to allow for substantial creativity. The hardware and software necessary for building these tools will also improve quickly and decrease in their costs, given the current attention given to the field, especially in the field of video game development. Consequently, it will be important to apply advances from external domains into the field of medicine.

Further work will now require customizations of scenes such as those presented here: i.e., through the development of specialized networking and manipulation tools. Currently, cardiac surgery, orthopedic surgery, neurosurgery, and oral surgery appear to be good candidates, however, these are still limited by lack of automation regarding segmentation and modeling of clinical anatomies. Nevertheless, our current pipeline provides a foundation that is immediately applicable to investigators with complex cases or researchers seeking novel techniques to visualize and interact with anatomical structures.

Afterward

The Visible Heart Labs offer an extraordinary range of experiences that defy simple categorization. The sheer diversity of projects undertaken here makes it challenging to pin down the lab's field of work. This diversity is both exhilarating and daunting, making the lab a place of endless discovery. Far from being confined to scientific research alone, the skills and insights gained here span a broad spectrum.

During my tenure, I engaged in numerous initiatives, including swine surgery experiments, reanimating human hearts for study, mentoring students at various levels, and presenting our findings to audiences ranging from peers in academia to the general public. My journey also took me to remote locations for research on hibernating black bears, to both local and international conferences, and allowed me to meet patients at critical junctures in their heart transplant journeys. Observing a variety of cardiac surgeries, collaborating with industry leaders like Medtronic, expanding my professional network, forging friendships, and deepening my scientific knowledge only scratch the surface of my experiences. Notably, these experiences are separate from the work discussed in this document.

These experiences have underscored the rapid evolution of skill sets required for success in any field. Mastery of material is insufficient without the ability to communicate effectively across diverse settings, from scientific conferences to elementary school science fairs.

Perhaps the most profound lesson has been the interplay between humility and scientific pursuit. Knowledge acquisition invariably leads to new questions, highlighting the collaborative essence of scientific inquiry. It's the supportive ecosystem of resources and collaboration that makes asking these questions possible.

The best way to describe the Visible Heart Labs is as a vibrant ecosystem. It's a dynamic, sometimes chaotic environment where ideas are born, grow, and sometimes perish, only to give way to new transformations. While this document reflects my contributions, the true enablers of this work are the collective resources and teamwork that make such research possible.

The Visible Heart Labs is a testament to the abundance of opportunities it offers to its students, standing out not just for the volume of work, but for the quality of opportunities for growth and discovery. My hope is for the lab to continue thriving and expanding within the global scientific community, adapting to technological advancements and uncovering new ways to enhance healthcare.

The journey through the lab can be challenging, yet it's crucial to maintain perspective. The collaborative spirit among staff and students fuels progress, transforming individual research efforts into collective achievements. As I advance in my career, I look forward to continuing collaborations with the lab, contributing to the field of medicine.

At its heart, the Visible Heart Labs possesses a certain magic, reanimating hearts and undertaking projects that challenge our perceptions of the possible. This sense of wonder can quickly become routine, yet it's vital for future students to retain their awe and optimism. My experiences have taught me that it's this collective passion and expertise that ignites the most significant advancements.

Bibliography

[1] Medical robotic systems market analysis by product and segment forecasts to 2022. `https://www.grandviewresearch.com/industry-analysis/medical-robotic-systems-market`, 2015. Market Research Report - Grand View Research, 129.

[2] The global games market 2017. `https://newzoo.com/insights/articles/the-global-games-market-will-reach-108-9-billion-in-2017...`, 2017. Retrieved December 9, 2017.

[3] M. Aczon, D. Ledbetter, L. Van Ho, A. Gunny, A. Flynn, J. Williams, and L. K. Whittier. Dynamic mortality risk prediction dynamic mortality risk predictions in pediatric critical care using recurrent neural networks. `https://arxiv.org/pdf/1701.06675.pdf`, n.d.

[4] A. Angelidis. Vorticle fluid simulation technical memo 15-01. `http://graphics.pixar.com/library/VorticleFluidSim/paper.pdf`, n.d.

[5] V. Badrinarayanan, A. Kendall, and R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[6] B. Baillargeon, N. Rebelo, D. D. Fox, R. L. Taylor, and E. Kuhl. The living heart project: A robust and integrative simulator for human heart function. *European Journal of Mechanics / A Solids*, 48:38–47, 2014.

[7] S. Bako, T. Vogels, B. Mcwilliams, M. Meyer, J. Novák, A. Harvill, et al. Kernel-predicting convolutional networks for denoising monte carlo renderings. *ACM Trans. Graph.*, 36(97), 2017.

[8] E. Z. Barsom, M. Graafland, and M. P. Schijven. Systematic review on the effectiveness of augmented reality applications in medical training. *Surgical Endoscopy*, 30, 2016.

[9] A. Bauer, D. R. Neog, A.-H. Dicko, D. K. Pai, F. Faure, O. Palombi, and J. Troccaz. Anatomical augmented reality with 3d commodity tracking and image-space alignment. `https://hal.archives-ouvertes.fr/hal-01633632/file/Paper.pdf`, 2017.

[10] Y. Bengio. Deep learning of representations for unsupervised and transfer learning. pages 17–37, 2011.

[11] Biodigital. Biodigital human - 3d anatomy - android apps on google play. `https://play.google.com/store/apps/details?id=com.biodigitalhuman.humanAndroid&hl=en`, 2017.

[12] W. Carlson. A critical history of computer graphics, 2003.

[13] P. W. Chang, B. C. Chen, C. E. Jones, K. Bunting, C. Chakraborti, and M. J. Kahn. Virtual reality supplemental teaching at low-cost (vrstl) as a medical education adjunct for increasing early patient exposure. *Med Sci Educ*, 28:3, 2018.

[14] E. Chinchoy, C. L. Soule, A. J. Houlton, W. J. Gallagher, M. A. Hjelle, T. G. Laske, J. Morissette, and P. A. Iaizzo. Isolated four-chamber working swine heart model. *Ann Thorac Surg*, 70:1607–1614, 2000.

[15] F. Chollet. Xception: Deep learning with depthwise separable convolutions. arXiv preprint arXiv:1610.02357, 2016.

[16] D. Coffey, N. Malbraaten, T. B. Le, I. Borazjani, F. Sotiropoulos, A. G. Erdman, and D. F. Keefe. Interactive slice wim: Navigating and interrogating volume data sets using a multisurface, multitouch vr interface. *IEEE Transactions on Visualization and Computer Graphics*, 18(10):1614–1626, 2012.

[17] B. F. Duffy and D. R. Flynn. A year in computer vision. `http://www.themtank.org/pdfs/AYearofComputerVisionPDF.pdf`, 2017.

[18] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010.

[19] E. Gaasedelen, A. Deakyne, T. Iles, and P. Iaizzo. Using smartphone-based virtual reality to explore internal anatomy of 3d heart models. In *2017 Design of Medical Devices Conference*, page V001T09A009. ASME, 2017. Available: `https://asmedigitalcollection.asme.org/DMD/proceedings/DMD2017/57731/V001T09A009/229171`.

[20] S. Gu, E. Holly, T. Lillicrap, and S. Levine. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. `https://arxiv.org/pdf/1610.00633.pdf`, n.d.

[21] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[22] P. Holley. How doctors used virtual reality to save the lives of conjoined twin sisters. The Washington Post, 2017. URL: http://tinyurl.com/y268r2me [accessed 5 February 2019].

[23] 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.

[24] 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.

[25] P. A. Iaizzo. The visible heart® project and free-access website 'atlas of human cardiac anatomy'. *EP Europace*, 18(suppl_4):iv163–iv172, 2016.

[26] H. Iben, M. Meyer, L. Petrovic, O. Soares, J. Anderson, and A. Witkin. Artistic simulation of curly hair pixar animation studios pixar technical memo #12-03b. `http://graphics.pixar.com/library/CurlyHairB/paper.pdf`, n.d.

[27] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.

[28] S. Ji, W. Xu, M. Yang, and K. Yu. 3d convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):221–231, 2012.

[29] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 1725–1732, 2014.

[30] Alexander Kirillov, Ross Girshick, Kaiming He, and Piotr Dollár. Panoptic feature pyramid networks. *arXiv preprint arXiv:1901.02446*, 2019.

[31] R. A. Kirsch, L. Cahn, C. Ray, and G. H. Urban. Experiments in processing pictorial information with a digital computer. In *Papers and discussions presented at the December 9-13, 1957, eastern joint computer conference: Computers with deadlines to meet*, pages 221–229, New York, New York, USA, 1958. ACM Press.

[32] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie. Feature pyramid networks for object detection. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 2117–2125, 2016.

[33] C. H. Liu. Anatomical, functional and molecular biomarker applications of magnetic resonance neuroimaging. *Future Neurol*, 10:49–65, 2015.

[34] S. McCartney and Scott. *ENIAC, the triumphs and tragedies of the world's first computer*. Walker, 1999.

[35] M. H. Michalski and J. S. Ross. The shape of things to come: 3d printing in medicine. *Jama*, 312(21):2213–2214, 2014.

[36] X. Mottart and P. Slagmolen. 3d lab in a hospital environment 3 key reasons why a professional desktop printer is a great fit. `http://www.materialise.com/en/system/files/resources/Medical_WhitePaper_Formlabs_v3.pdf`, n.d.

[37] Hyeonseob Nam, Jung-Woo Ha, and Jeonghee Kim. Dual attention networks for multimodal reasoning and matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 299–307, 2017.

[38] M. Norouzi, T. Mikolov, S. Bengio, Y. Singer, J. Shlens, A. Frome, et al. Zero-shot learning by convex combination of semantic embeddings. arXiv preprint arXiv:1312.5650, 2013.

[39] J.-I. Okada, T. Washio, M. Nakagawa, M. Watanabe, Y. Kadooka, T. Kariya, et al. Multi-scale, tailor-made heart simulation can predict the effect of cardiac resynchronization therapy. 2017.

[40] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[41] A. Pugnale, M. Sassone, A. Pugnale, P. Student, P. Di, and T. Torino. Aalborg universitet morphogenesis and structural optimization of shell structures with the aid of a genetic algorithm morphogenesis and structural optimization of shell structures with the aid of a genetic algorithm, n.d.

[42] M. Robiony, I. Salvo, F. Costa, N. Zerman, M. Bazzocchi, F. Toso, C. Bandera, S. Filippi, M. Felice, and M. Politi. Virtual reality surgical planning for maxillofacial distraction osteogenesis: the role of reverse engineering rapid prototyping and cooperative work. *J Oral Maxillofac Surg*, 65:1198–1208, 2007.

[43] 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, 2015.

[44] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[45] L. N. Smith. A disciplined approach to neural network hyper-parameters: Part 1–learning rate, batch size, momentum, and weight decay. *arXiv preprint arXiv:1803.09820*, 2018.

[46] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016.

[47] Y. Teng, D. I. W. Levin, and T. Kim. Eulerian solid-fluid coupling. n.d.

[48] A. M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72, 1952.

[49] P. J. Vatterott, M. D. Eggen, A. R. Mattson, P. K. Omdahl, K. E. Hilpisch, and P. A. Iaizzo. Retrieval of a chronically implanted leadless pacemaker within an isolated heart using direct visualization. *HeartRhythm Case Reports*, 4(5):167–169, 2018.

[50] G. G. Wang. Definition and review of virtual prototyping. *J Comput Inf Sci Eng*, 2:232–236, 2003.

[51] M. J. P. Wolf. *The video game explosion: a history from PONG to Playstation and beyond*. Greenwood Press, 2008.

[52] 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, Cham, October 2016.