

Qualitative visuelle Analyse von Interobserver-Variabilität

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Visual Computing

eingereicht von

Hannah Clara Bayat, BSc

Matrikelnummer 01614855

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Assistant Prof. Dr.in techn. Manuela Waldner, MSc

Mitwirkung: Assistant Prof. Dr.in Renata Raidou

Wien, 7. Dezember 2022

Hannah Clara Bayat

Manuela Waldner



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Multi-faceted Visual Analysis of Inter-Observer Variability

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Visual Computing

by

Hannah Clara Bayat, BSc

Registration Number 01614855

to the Faculty of Informatics

at the TU Wien

Advisor: Assistant Prof. Dr.in techn. Manuela Waldner, MSc

Assistance: Assistant Prof. Dr.in Renata Raidou

Vienna, 7th December, 2022

Hannah Clara Bayat

Manuela Waldner



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Hannah Clara Bayat, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 7. Dezember 2022

Hannah Clara Bayat



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Danksagung

Ich widme diese Arbeit meiner Großmutter, die mich in jeder nur erdenklichen Weise unterstützte.

Der Verlauf dieser Arbeit war nicht einfach. Oftmals verlor ich meine Hoffnung und Motivation und wusste nicht, wie ich weitermachen sollte. Wann immer mir die Kraft fehlte meinen Zielen und Ambitionen nach zu streben, wusste ich die unglaublichen Menschen in meiner Umgebung zu schätzen. Sie zögerten nie, mir zu sagen, was gesagt werden musste und für mich da zu sein, wenn ich mir zu viel Druck machte. Zunächst danke ich meinen Eltern und meiner Schwester. Sie unterstützten mich nicht nur in meiner akademischen Laufbahn sondern zeigten mir unerbittliche Liebe in jeder Lebenssituation. Meine Familie war schon immer mein größtes Vorbild. Sie zeigt mir tagtäglich, dass Erfolg von harter Arbeit stammt und ich hoffe ich kann sie mit meinen Taten stolz machen.

Egal ob Motivation, Ablenkung oder Hilfe bei der Fehlersuche im Code, ich weiß ich kann immer auf Mar zählen. Wann immer ich meine Arbeit aufschob, war sie diejenige, die mir den nötigen Ruck gab um weiterzumachen. Sie machte nie Halt davor, mir die emotionale Unterstützung zu geben, die ich brauchte. Der Abschluss meines Studiums wäre ohne sie sicherlich ein weitaus mühsamerer Weg gewesen. Ähnlich wie ohne den Rückhalt meiner Freunde, die immer offen für meine Beschwerden und Ambitionen waren. Sie waren für mich da, wenn ich eine Ablenkung brauchte, um meinen Kopf frei zu kriegen und mich durch neue Perspektiven neu konzentrieren konnte. Darüber hinaus möchte ich meinen Kolleginnen und Kollegen am Institut danken, die mich so herzlich in ihre Umgebung aufgenommen haben. Über ein Jahr, habe ich unglaublich viel gelernt und Seite an Seite mit wundervollen Menschen gearbeitet.

Letztlich, möchte ich meinen Betreuerinnen Manuela Waldner und Renata Raidou danken, ohne die all dies nicht möglich gewesen wäre. Sie duldeten meine unüblichen Entscheidungen bezüglich Programmiersprachen, halfen mir bei den trivialsten Kleinigkeiten und zeigten Verständnis und Unterstützung über die gesamte Zeit.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

This thesis is dedicated to my grandmother, who always supported me in any way possible.

The process of this thesis was not easy. Oftentimes I would lose hope and motivation and not know how to continue. While I was lacking the strength to accomplish my goals, the incredible people in my surroundings would never refrain to tell me what I needed to hear or be there for me when I was being hard on myself. Firstly, my parents and my sister who have not only supported me in my academic career but have given me unconditional love. My family has always shown me that hard work is essential to a successful life and I hope I can make them proud by following in their footsteps.

A huge force in this process was Mar. Whether it was motivation, distraction or a debugging buddy, I knew I could always count on her. Whenever I was procrastinating she would get me out of it and never hesitate to give me emotional support so I could get work done. Completing my studies would have been a much more tedious path without her. As it would have been without the influence of my friends, who always had an open ear for my complaints and ambitions. They were there for a good distraction when I needed to free my head and get some perspective to find focus again. I want to thank my colleagues at the institute for welcoming me so warmly into their space. Working with them has taught me so much and most importantly, I got to work alongside awesome human beings.

Finally, none of this would have been possible without the undying support of my supervisors Manuela Waldner and Renata Raidou. They put up with my odd programming preferences, never got tired of helping out with simple things and have been understanding and supporting through the entire time.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Trotz der Fortschritte bei den automatischen Algorithmen zur Segmentierung ist die manuelle Segmentierung in der Medizin immer noch notwendig. Die manuelle Segmentierung von Tumoren ist unter Radiolog:innen gängige Praxis und ein kritischer Schritt in der Planung einer Krebsbehandlung. Wir veranschaulichen die Variabilität, die sich aus der Segmentierung medizinischer Scans eines einzelnen Patienten durch mehrere Expert:innen ergibt und als Interobserver-Variabilität bezeichnet wird.

Die Neuheit dieser Arbeit besteht darin, den Prozess der Segmentierung eines Objekts zu erfassen. Der Schwerpunkt liegt darin, Einblicke in den Denkprozess und die Entscheidungsstrategien der Beobachter:innen zu erhalten. Um diese Aspekte der Segmentierung zu untersuchen, führen wir eine Datenerhebung mit Noviz:innen und Expert:innen durch, indem wir ihre Gedanken in einem Think-Aloud-Protokoll und ihre Aufmerksamkeitsbereiche durch Aufzeichnen ihrer Mausbewegungen während des Segmentierungsprozesses erfassen. Diese Daten werden mit unserer **Multi Observer Looking Environment** (Mehrfachbeobachter-Anschauungsumgebung) visualisiert. MOLE erlaubt einen gründlichen Einblick in den Segmentierungsprozess der Beobachter:innen und ermöglicht den Vergleich verschiedener Segmentierungsergebnisse. Mit unseren Visualisierungstechniken heben wir Regionen der Unsicherheit hervor, die bei der Segmentierung mehr Aufmerksamkeit erforderten. Darüber hinaus werden relevante Schlüsselwörter aus dem Think-Aloud-Protokoll extrahiert und mit den Positionen in der Segmentierung abgeglichen, was Aufschluss über den Denkprozess der Beobachter:innen liefert. Wir verknüpfen das Ausgangsbild mit einer dreidimensionalen Darstellung der Segmentierung und stellen bei Bedarf weitere Details des Think-Aloud-Protokolls zur Verfügung.

Unser Programm ist universell auf Segmentierungs-, Aufmerksamkeits- und Gedankenprozessdaten anwendbar, unabhängig von der Natur der Daten. Wir zeigen, wie MOLE sowohl mit einem medizinischen Datensatz als auch mit einem künstlich erzeugten Datensatz verwendet werden kann. Durch die Validierung unseres Programms mit Hilfe eines medizinischen Experten, der aktiv in diesem Bereich arbeitet, definieren wir mögliche Einsätze und Anwendungsfälle in der bestehenden Vorgehensweise der Tumorsegmentierung und Krebsbehandlung.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Despite the advancements in auto-segmentation tools, manual delineation is still necessary in the medical field. For example, tumor segmentation is a crucial step in cancer radiotherapy and is still widely performed by hand by experienced radiologists. However, the opinions of experienced radiologists might differ, for a multitude of reasons. In this work, we visualize the variability originating from multiple experts delineating medical scans of the same patient, known as inter-observer variability.

The novelty of this work consists of capturing the process of segmenting a target object. The focus lies in gaining insight into the observer's thought processes and reasoning strategies. To investigate these aspects of segmenting we conduct a data acquisition with novice users and experts, capturing their thoughts in a think-aloud protocol and their areas of attention by tracking their mouse-movement during the segmentation process. This data is visualized with our **M**ulti **O**bserver **L**ooking **E**nvironment (MOLE). MOLE allows to gain deep insight into the observers' segmentation process and enables to compare different segmentation outcomes and how these occurred. With our proposed visualization techniques we emphasize regions of uncertainty that need more attention when delineating. Additionally, relevant keywords are extracted from the think-aloud protocol and aligned with the positions in the segmentation, providing information about the thought process of an observer. We link the initial image to a three-dimensional representation of the delineations and provide more details of the think-aloud protocol on demand.

Our approach is universal to segmentation, attention and thought process data regardless of the domain of the data. We show how MOLE can be used with a medical dataset as well as an artificially created dataset. By validating our approach with the help of a medical expert actively working in the field, we define potential use cases in the existing pipeline of tumor delineation for cancer treatment.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Kurzfassung	xi
Abstract	xiii
Contents	xv
1 Introduction	1
1.1 Motivation	1
1.2 Aim of the Work	2
1.3 Contributions	2
1.4 Outline	4
2 Background	5
2.1 Visualization in Medicine	5
2.2 Image Segmentation	7
2.2.1 Segmentation Algorithms	8
2.2.2 Inter-Observer Variability	8
3 Related Work	11
3.1 Visualization of Segmentations	11
3.2 Visualization of Processes	13
3.3 Visualization of Inter-Observer Variability	16
4 Methodology	21
4.1 Step 1: Defining Terms and Tasks	22
4.1.1 Uncertainty	22
4.1.2 Inter-Observer Variability	22
4.1.3 Target Users and Task Definition	23
4.2 Step 2: Data Acquisition	24
4.2.1 Reasoning Strategies	24
4.2.2 Attention Areas	25
4.2.3 Segmentation	25
4.3 Step 3: Visualization	26
4.3.1 Scan Data	26
	xv

4.3.2	Segmentation Data	26
4.3.3	Attention Data	29
4.3.4	Reasoning Data	31
5	Data Acquisition	33
5.1	Goal	33
5.2	Participants	33
5.3	Segmentation and Logging Software	34
5.3.1	Data	35
5.3.2	Implementation	36
5.3.3	Data Collection	37
5.3.4	Transcript Collection	38
5.4	Procedure	38
5.4.1	Warm-Up Task	39
5.4.2	Segmentation Task	40
5.4.3	Confidence Survey	41
5.5	Results	41
5.5.1	Improvements	43
5.5.2	Summary	43
6	MOLE	47
6.1	Visualization and Interaction Design	47
6.1.1	Discarded Features	53
6.2	Implementation	54
6.2.1	Deployed Technologies	54
6.2.1.1	Visualization Toolkit (VTK)	54
6.2.1.2	Qt	55
6.2.1.3	Python	55
6.2.2	Data Handling	55
6.2.2.1	Pre-Processing	56
6.2.3	Features	59
6.2.3.1	Overview First	60
6.2.3.2	Zoom & Filter	62
6.2.3.3	Details on Demand	64
6.2.3.4	Colors	65
7	Validation	69
7.1	Usage Scenario	69
7.1.1	Expert Data	69
7.1.2	Novice Data	79
7.2	Domain Expert Feedback	89
8	Discussion & Conclusion	91
8.1	Discussion	91

8.2	Future Work	93
8.3	Summary & Conclusion	94
A	Data Acquisition	97
A.1	Welcome Text	97
A.2	Warm-Up Task	97
B	Cleaned Transcripts	99
B.1	Experts	99
B.1.1	Expert 1	99
B.1.2	Expert 2	100
B.1.3	Expert 3	102
B.1.4	Expert 4	104
B.1.5	Expert 5	106
B.2	Novices	108
B.2.1	Novice 1	108
B.2.2	Novice 2	112
B.2.3	Novice 3	114
B.2.4	Novice 4	118
C	Data Processing Scripts	127
C.1	Process Logfiles	127
C.2	Fill Gaps	129
C.3	Clustering	130
	List of Figures	133
	List of Tables	137
	Bibliography	139



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Introduction

1.1 Motivation

Segmentation of medical images is mostly performed with automated algorithms. However, due to the fact that automated approaches are often still not able to deliver robust and generalizable results, manual segmentation, i.e. slice-by-slice delineation of structures of interest remains common practice for many clinical applications. A significant challenge is that the assessment of pathological structures in medical images can vary significantly across medical experts [1]. This variability can influence the diagnosis and treatment planning of a cancer patient. This phenomenon is known as *inter-observer variability* and is a common source of uncertainty in the medical field [2]. Inter-observer variability can be found in the domain of radiotherapy planning where, despite well-established guidelines for delineations, experts may still provide disagreeing outcomes [3]. The delineations are evaluated in a clinical board of examiners before being used and there are additional mechanisms to counteract them, but the origin of disagreements and their effect on diagnoses and treatments is not understood. Quantifying, raising awareness and visualizing uncertainty has been stressed multiple times within the medical domain [4]. In the medical field the segmentation of tumors, bladders [5], pelvic organs [6], metabolisms [7] or other objects and body parts of variable shapes by multiple experts can lead to inter-observer variability, as shown in Figure 1.1.

Currently, the visualization of segmentations lies in visualizing and presenting the different *segmentation outcomes*. We want to understand why the outcomes vary and sense the potential underlying reasons for inter-observer variability. We need a qualitative approach, revealing the thought processes and reasoning strategies of observers for their delineations. Understanding how clinical experts perceive and assess medical images can shed new light on the origins of inter-observer variability. To gain this desired insight we strive to provide a qualitative analysis into this information. It is therefore highly relevant to

the field of radiology [1] and to the domains of biomedical image analysis [8] as well as uncertainty visualization.

1.2 Aim of the Work

For a comprehensive qualitative analysis, it is not sufficient to analyze single voxel positions, values or ranges, as resulting from the regions of interest that are included in the delineations. It is essential to understand which image areas and slices are essentially investigated and where the observer's attention concentrates at. Furthermore, it is essential to document and understand the experts' reasoning process, while analyzing the data set. This is anticipated to reveal information on whether, how and where the experts' analytical processes differ from each other.

We aim for a joint analysis of points of interest, delineations resulting in quantified variability information and think-aloud protocols. The analysis of such data is highly complex. Qualitative analysis has been suggested to test whether expert performance and its underlying processes transfer to novel tasks within a domain [10]. However, these suggestions require extensive manual pre-processing, such as open coding and identification of pre-defined areas of interest in the analyzed images, which may limit in-depth qualitative analysis.

Our hypothesis is that qualitative analysis can be facilitated through interactive visualization of qualitative observation data, conveying the interplay between imaging information, attention, reasoning and final delineations on visual input, across multiple experts. This thesis advances state-of-the-art in provenance visualization for the field of medical visualization by finding a way to visualize the obtained data, comprising attention data, delineations and variability quantification, think-aloud protocols and interaction logs, without the necessity of defining areas of interests in advance or manual coding of think-aloud protocols. To sum up, the research question of this work is "How can we facilitate the qualitative comparative analysis of visual observations and reasoning strategies of clinical experts, who manually assess volumetric medical data sets"?

1.3 Contributions

Previous related work indicates that the domain of medical visualization has dealt with uncertainty mainly as "error", "imprecision" or "non-specificity" [4], leaving out inter-observer variability. There exists only little related work covering inter-observer variability which is mostly related to visual storytelling [11]. So far the visualization of segmentations has covered comparative analysis of the results of segmentations as discussed in Chapter 3. However, only little importance has been put on the thought and delineation process behind a segmentation.

This thesis explores the potential of visualizing the thought and reasoning process going into a delineation and discusses potential advantages of considering these aspects when

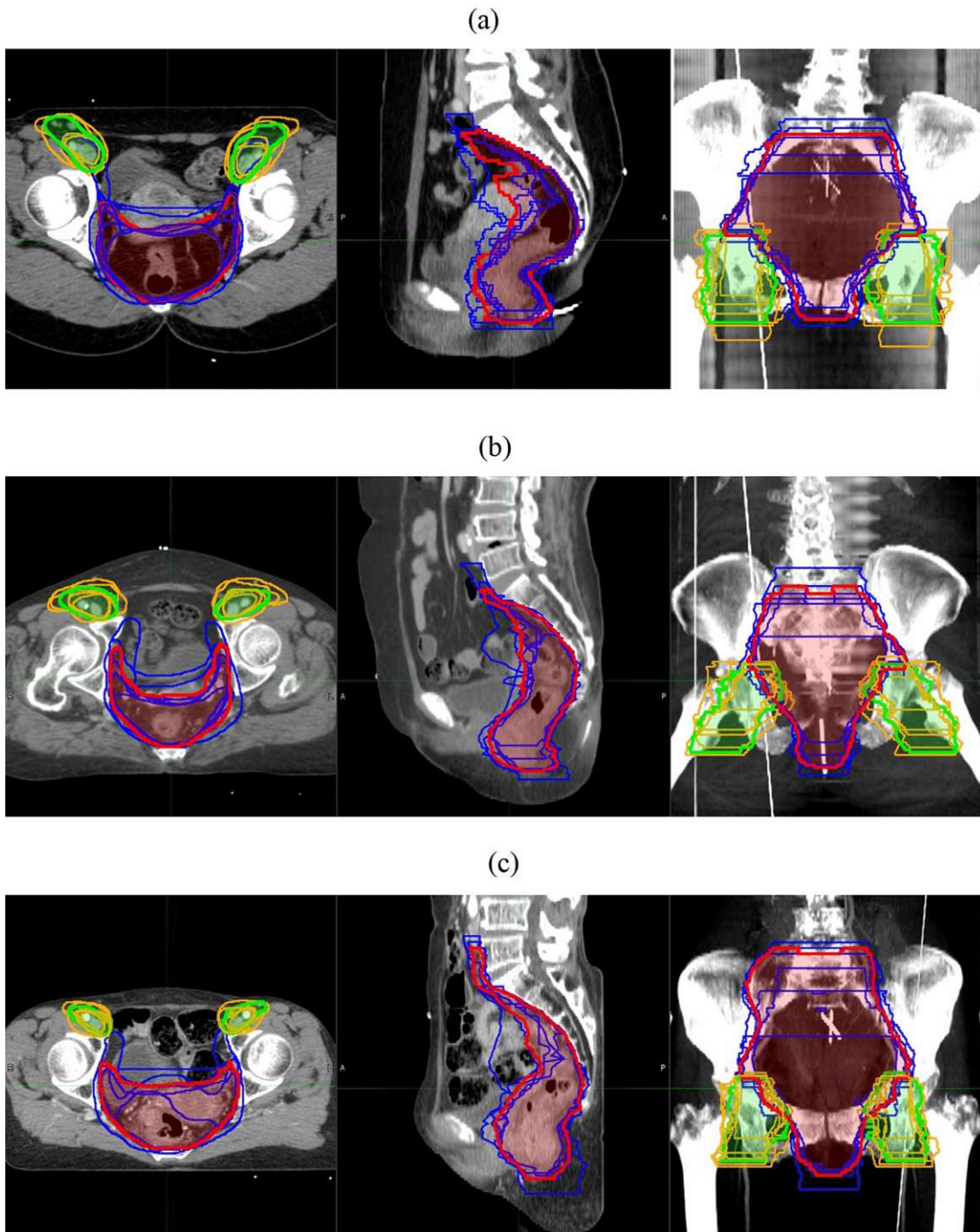


Figure 1.1: Segmentation of the pelvis and inguinal of three different patients (a, b, c) by seven radiation oncologists (blue and orange) and an expectation maximization algorithm (red and green), leading to inter-observer variability [9].

segmenting medical image data. Furthermore, this work provides a tool for comparative functions when analyzing segmentations. The inter-observer variability resulting from manual segmentation tasks and the captured thought and reasoning process is visualized and made accessible using a variety of visualization components, including heatmaps, 3D volumes and billboards. Working with our tool and analyzing the given data shall provide insight and detailed information that help the segmentation process and serve support in the delineation pipeline. Exploring underlying causes of inter-observer variability is crucial in order to understand the outcome and disagreements of the segmentation method. In contrast to existing work the focus and reasoning in the process of delineation is a central part of this work.

The contributions of this thesis are summarized as follows:

- Development of a qualitative data acquisition protocol for remote delineation sessions to be able to reach a large number of clinicians
- Design of a visual exploratory analysis tool that allows the comparison and inspection of the obtained data
- Preliminary assessment of our approach through expert feedback and two applied cases

1.4 Outline

After giving an overview about the goal of this thesis, the scientific contributions and the motivation behind it, Chapter 2 will cover some basic background information that is necessary to understand the contents of this thesis. It discusses visualization in medicine and image segmentation processes, including different medical imaging techniques, the storage of medical data and the procedure of creating segmentations from medical images. Afterwards, Chapter 3 gives an overview of work related to this thesis. We will discuss current and previous work on segmentation visualization, process visualization and visualization of inter-observer variability. Chapter 4, explains the design and visualization choices made for the final outcome of this thesis. The decisions are presented on an abstract level, so they are applicable and reproducible regardless of the desired technologies used.

Next, Chapter 5 dives into the process of acquiring the data needed for our intentions. This chapter illustrates the participants, procedure and outcome of the data acquisition process. Chapter 6 presents the application of this thesis and explains the features and their implementation in detail, showing what our tool is capable of. Before concluding the thesis, Chapter 7 discusses possible usage scenarios of our application as well as their evaluation and possible applications in the real world. Lastly we conclude the thesis in Chapter 8, where we discuss results and possible future work.

Background

This chapter provides the necessary background knowledge for this thesis. We give an overview of different imaging techniques, specifically in the medical field as well as structure delineation procedures.

2.1 Visualization in Medicine

Visualization in medicine is used to explore data, test hypotheses or simulations and present results. With an image acquisition technique, called a *modality*, we acquire medical image, which are either represented as a stack of individual images or concatenated into one image. Each of the individual images represent one slice of the scanned object. They serve to diagnose patients with different medical conditions that are not determinable through the patients' symptoms. Furthermore, medical images are used as the basis for therapy planning, monitoring the patient's status before and after surgical procedures and biomedical research [7]. To be able to talk about parts of the body universally Figure 2.1 shows the accurate jargon for referring to the axes of the body. When referring to the right or left part of the body, one always considers the patient's viewpoint. A person's right lung is positioned on the left side when looking at the person.

Modalities can be classified based on the procedure the image is acquired with. The most common procedures include the utilization of X-rays, radionuclides, ultrasound, nuclear magnetic resonance (NMR), infrared light, electrical impedance and the visible light spectrum [7]. Figure 2.2 shows a classification of medical imaging modalities based on ImageCLEF 2015 [13]. The modality with which a patient, or their body part in question is scanned, is determined by the type of information that they are able to reveal. X-rays and *Computed Tomography (CT)* are used for dense structures such as bones or calcifications. CT is also used for tumors and cancers as well as injuries from trauma. Both of these methods use ionizing radiation. *Magnetic Resonance Imaging (MRI)* is used for soft tissues, vasculature, tumors and also functional imaging. It is applicable

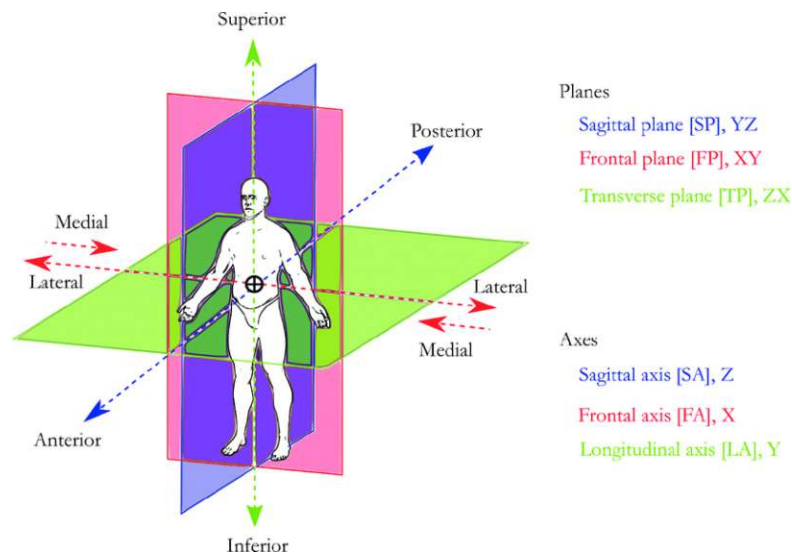


Figure 2.1: Anatomical planes and axes [12].

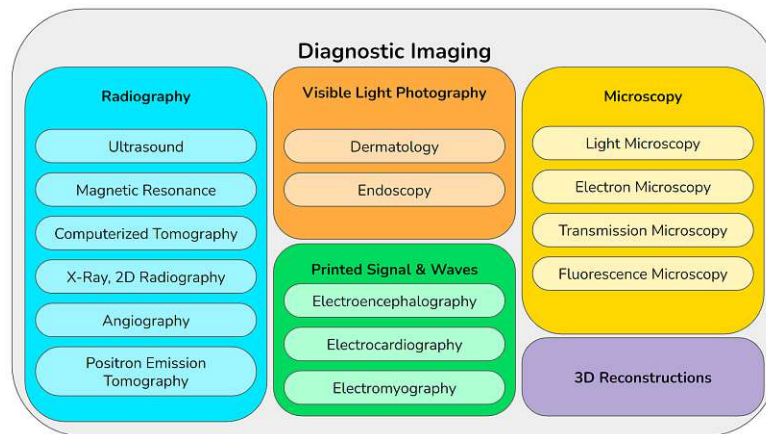


Figure 2.2: Modality classifications according to Elmogy et al. [13]

when looking for *multiple sclerosis (MS)* diagnoses and strokes. Ultrasound is used for live imaging soft tissues and to monitor pregnancies. It uses sound waves to create the image. A scan using *Positron Emission Tomography (PET)* detects metabolic activities such as tumors or brain activity. It is used in Alzheimer’s and Parkinson’s treatment as well as to detect cancer [14]. Since the use case extracted from the medical field in this thesis focuses on CT images, the medical images are acquired using CT and PET. The scans used in the data acquisition process are CT scans and the ground truth defined for the expert task as described in Section 5.4.2 is acquired using PET.

CT data consists of a series of X-ray images, obtained through a detector-emitter system rotating around the object to be scanned. Each image shows the intensities measured



Figure 2.3: CT scanner of model SOMATOM X.ceed by SIEMENS [15].

by the detector. Based on the filtered backprojection and the intensity profiles of a full rotation, the slices are computed. The images are accumulated, representing the data as one volume. Figure 2.3 shows a CT scanner. The torus-shaped part of the CT scanner contains the emitter and detector, rotating around the object.

PET is a modality in nuclear medicine. The scans detect metabolic activities, like metabolic activity and generate three-dimensional images. By injecting a so called “tracer” or “isotope”, a short-lived radio-pharmaceutical substance the target area is made visible. The tracer makes its way to the metabolism where it is processed. The radioactive isotope decays with this process and emits a positron, which when interacting with an electron generates gamma rays in opposite directions. The emitted rays are then detected by the PET scanner. It is possible to localize the source and thus the location in the body, where it is metabolised [7].

Medical image data can be stored in different ways. One commonly used format is the DICOM (Digital Imaging and Communications in Medicine) standard established by the National Electrical Manufacturers Association (NEMA). The DICOM format is constantly adjusted by the DICOM committee including leading manufacturers and several institutions. A scan can either be saved as one single DICOM file or as multiple DICOM images. DICOM files contain information about the patient, acquisition data and study data in a standardized, hierarchical manner that enables the safe storage, sharing, and interoperability of the data across institutions [7].

2.2 Image Segmentation

Preim and Botha define segmentation as “the task of decomposing image data into meaningful structures that are relevant for a specific task” [7]. Segmentation in this thesis happens on the acquired medical image data. When segmenting an image the target object should be determined as precisely as possible. Image segmentation can be done manually or with the help of algorithms. Structures with a unique definition and shape

can be segmented with algorithms and deliver accurate results. However, structures like tumors can change in form and size and do not have a definitive shape. The borders of the delineations are usually unclear due to partial volume effects, resolution limitations or image digitization. Otherwise an auto-segmentation tool would suffice to delineate. Despite the existence of several robust approaches for tumor segmentations, they are not trusted yet and do not offer accountability. Therefore, tumors are commonly segmented manually or partially manual with computer support. Manual segmentations are done by drawing on each individual slice of a scan and delineating the target in every slice [7].

2.2.1 Segmentation Algorithms

There are three types segmentation methods. The method can be either fully manual, semi-automatic or fully automated [16]. Manual segmentations are usually performed by radiologists. Commonly, the delineation is done on two-dimensional slices. Manual segmentation can result in a variability.

Semi-automatic approaches combine supporting algorithms with manual segmentation in order to speed up the delineation process. These algorithms utilize region growing filters for instance to reduce the amount of slices that need to be segmented [16]. With the help of semi-automatic algorithms the aim is to reduce the variability as well. With automatic segmentation methods the results cannot be fully trusted. However, they are reproducible and consistent, which can be of advantage in some cases.

2.2.2 Inter-Observer Variability

Inter-observer variability is defined as the difference in segmentations between multiple observers. For n observers segmenting an object, the result will be n segmentations of different shape and size. For manual tumor segmentations, some observers might mark more generously and cut away more of the tumor. Others might be uncertain about other body parts like vessels and veins and delineate parts that actually do not belong to the tumor [16]. These decisions and complications lead to inter-observer variability. Figure 2.4 shows the inter-observer variability of 10 different segmentations of a lung tumor from the dataset that is used in this thesis. The slices are cut very thick, so the resolution in the vertical axis is quite sparse.

Inter-observer variability shall not be confused with *intra-observer variability*. Intra-observer variability describes the variability of the segmentation outcomes when delineated by the same observer multiple times, e.g. in different days. Thus, looking at multiple segmentation outcomes, one will not be able to distinguish if the differences stem from inter-observer variability or intra-observer variability [17].

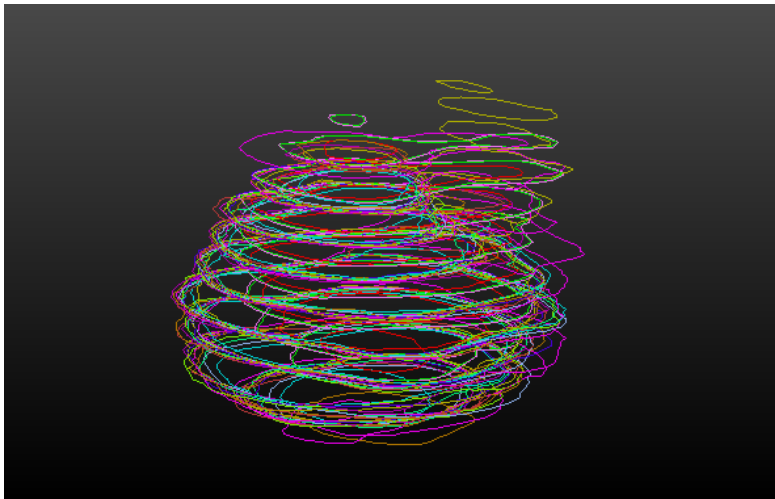


Figure 2.4: Segmentations of a lung tumor as delineated by 10 different observers [18]. Each color represents one observer's delineation.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Related Work

The following chapter provides information about scientific work related to the basis of this thesis. It covers current work about how delineations and segmentations in the medical field are visualized and presented. Related work includes research about the visualization of think-aloud protocols and the visualization of processes as well as eye-tracking and mouse-tracking data. The current approaches to the visualization of inter-observer variability and variable shapes are essential for the basis of this thesis.

3.1 Visualization of Segmentations

This section will cover different approaches to visualizing segmentations and delineations in two and three dimensions. The goal of visualizing segmentations is to show a delineation of a single observer or segmentations in a single target space. Segmentations are often acquired using automated algorithms but often also stem from manually drawn delineations. The data that is dealt with is mostly based on binary masks or binary images indicating the parts that are inside or outside the segmented structure.

For analyzing and manipulating segmentation data stemming from lungs infected with COVID-19, Oulefki et al. developed a VR environment [19]. A representation in two dimensions using different colors to show segmentations on a human foot is used by Strzelecki et al. [20] in their approach, as seen in Figure 3.1. This simple two-dimensional representation is also used in other works [21].

Friese et al. [22] propose a platform for visualizing segmentations in the three-dimensional space called *YaDiV* which stands for “Yet Another DICOM Viewer”. One way to visualize the volume data makes use of iso-surfaces. However, since their goal was a natural looking presentation, geometric smoothing is applied to achieve a more realistic look and to avoid step-like artifacts. Figure 3.2 shows the results of this method. Similarly, Prugnaller et al. [23] use iso-surfaces multiplied with a binary mask in their developed toolbox for

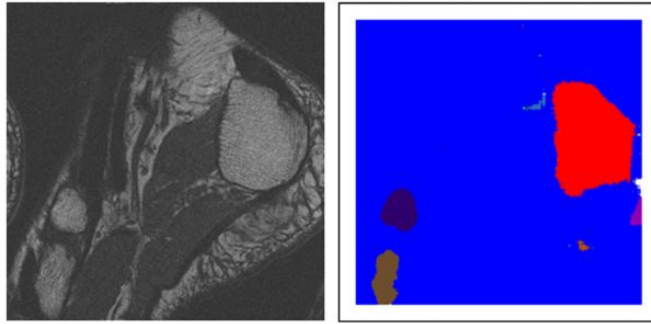


Figure 3.1: Segmentation results (right) stemming from human foot MR visualized in two dimensions (left) [20].

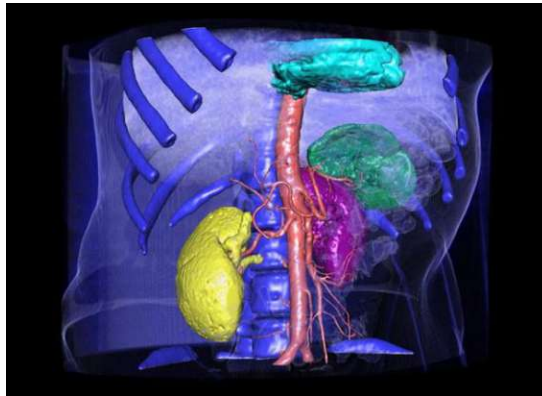


Figure 3.2: Interactive visualization of organ segmentation using YaDiV [22].

visualizing electron microscopy segmentations. Saad et al. use iso-surfaces to represent subsections of the brain to show likelihoods, which correspond to uncertainty extracted from tumor delineations [24]. In the method used by Groenendaal et al., they also stack the two-dimensional images in three-dimensional space, with the difference that the final volume is being registered to the input image [25]. This way, the result becomes smooth as well.

In our approach we reconstruct a three-dimensional volume from manual delineations originating from two-dimensional medical images. A similar method where two-dimensional images are represented with three-dimensional volumes was proposed by Norouzi et al. [26]. The segmentations extracted from two-dimensional images are placed into a three-dimensional space. The data is then visualized using multi-planar reformatting, surface rendering and hardware-accelerated volume rendering [26]. Figure 3.3 shows the initial two-dimensional segmentations and the resulting three-dimensional volume.

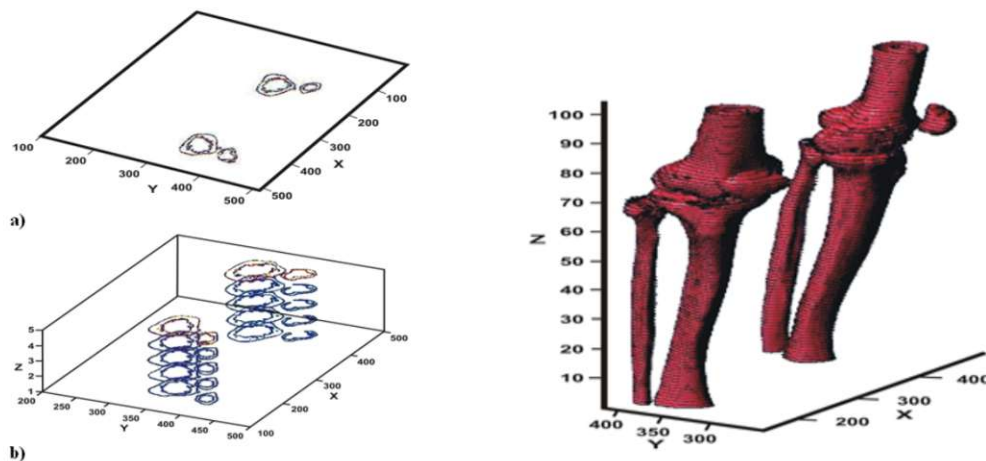


Figure 3.3: Initial slice by slice segmentation (left) and its 3D visualization (right) [26].

3.2 Visualization of Processes

In this thesis we work with components to visualize think-aloud protocols that capture the observers' reasoning strategies and thought process. This section gives an overview of current approaches to visualizing provenance, mouse trajectories and eye-tracking data.

In recent years provenance-related work has increased, which calls for a need of ways to visualize the acquired data. Reasons to visualize provenance data include the need for replication, adaptive systems, report generation as well as understanding the user [27]. Provenance visualization corresponds to visual story telling, as the aim is to present a process. Wohlfart et al. integrate story telling into visualizations in order to present the data and point out different areas of interest to the observer [28]. With this tool, users can explore complex datasets and put a story together. This combines the interactive aspect of a visualization with story telling. Another approach to visualizing operation processes was proposed by Waldner et al. [29]. They present a visualization to review processes and access information sources. This is achieved by creating a visualization combining a node-link diagram and snapshot sequences from significant points of the interaction process.

Similar to the aim of this thesis, Amabili et al. have developed a tool to visually explore processes stemming from diagnostic medical imaging [11]. Their work uses visual story telling based on provenance data for scientific visualization in medical data exploration. This tool ties the data from user interactions and the results together, allowing to connect the interactions with the final outcome. The tool is realized using provenance graphs, as shown in Figure 3.4. Provenance graphs are acyclic and consist of a value and an evaluation label at each node and edge. The provenance graph designed here allows for details on demand. He et al. created a tool called "MediSyn" that allows for five types of interactions [30]. Capturing users' interaction processes while exploring the data with

3. RELATED WORK

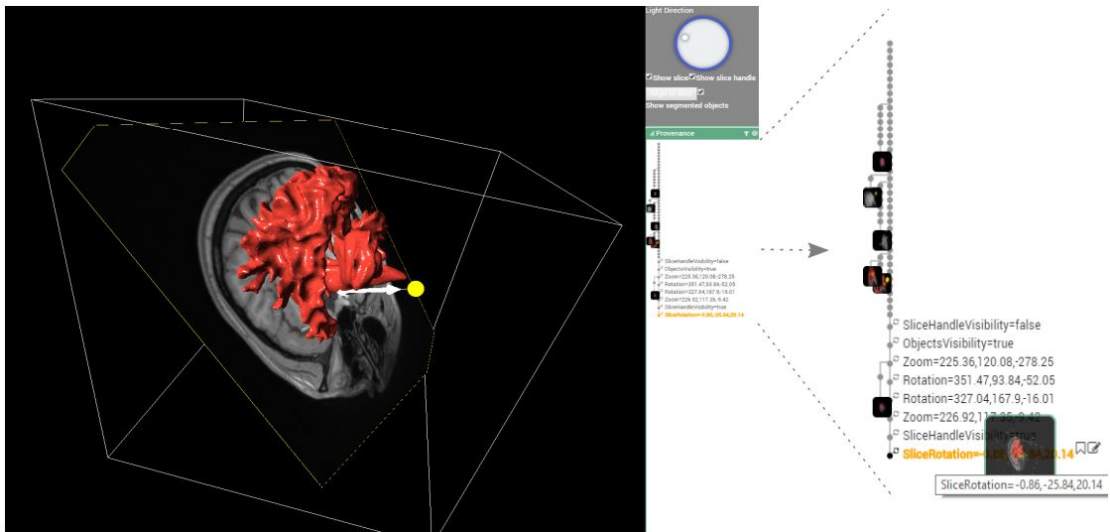


Figure 3.4: Tool developed by Amabili et al. with the main window for 2D and 3D exploration on the left and the provenance graph on the right [11].

MediSyn, they were able to identify seven interaction patterns. These patterns can be correlated to four different types of insight quality, which proves that insight can be qualified through the underlying interactions. In this approach the researchers obtained their own data in two steps, similar to our work. However, our goal with this thesis is to find reasoning strategies of different observers based on their actions while segmenting. Dou et al. investigated how the reasoning processes could be recovered using user interactions [31]. In a different medium, Nguyen et al. [32] have developed “SensePath”, a tool allowing the analysis of sense-making processes to support the understanding of large and complex datasets. Both Blaschek et al. [33] and Stitz et al. [34] have also proposed tools with which provenance stemming from interaction with a tool can be explored. Figure 3.5 shows a user’s transcript, actions and fixation durations, amongst more information is gathered and visualized. The fixation and transcript are aspects, also captured and visualized within the scope of this thesis.

Interaction logs as used by He et al. are visualized by Guo et al. using a color-coded timeline [35]. As shown in Figure 3.6 this timeline consists of different colors, representing different interactions with the tool. In addition, this proposed tool also offers timestamps through brushing over the timeline, which then connect to the captured video, providing details on demand. Fan et al. designed a machine learning model to find speech patterns of think-aloud protocols when solving tasks for usability purposes. The processed data is then visualized using a color-coded timeline as well, adding additional glyphs that update as the video is played as shown in Figure 3.7. In addition Fan et al. present a timeline showing when a problem was solved, a problem description and the video of the recording itself. So, the visualization features change dynamically as the video progresses [36].

Obtaining and visualizing eye-tracking data plays a central role in our application.

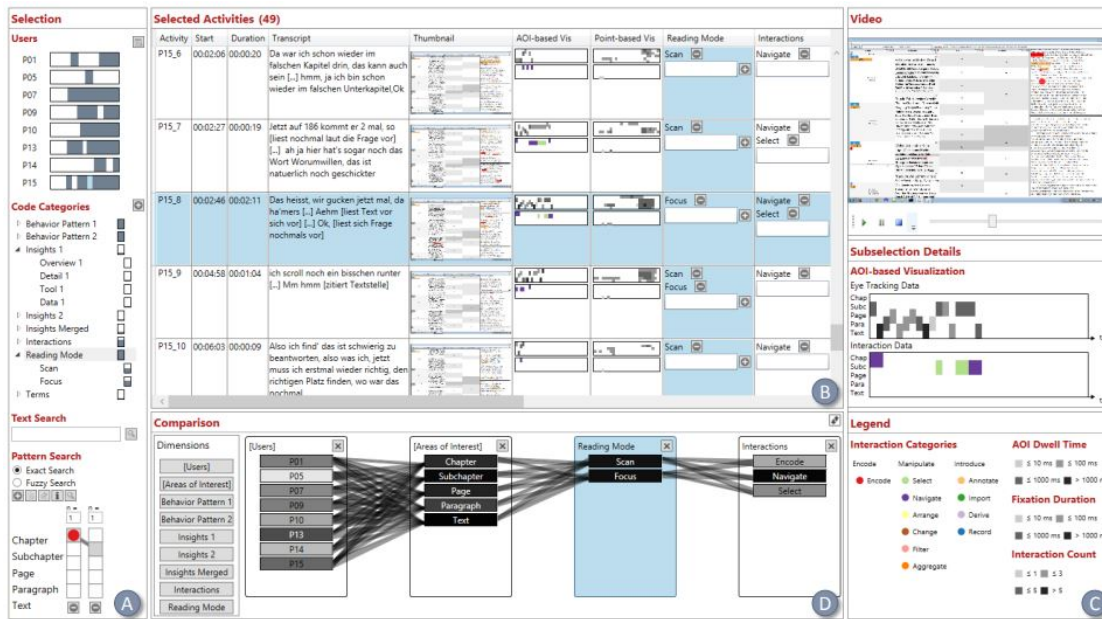


Figure 3.5: Dashboard to analyze provenance data proposed by Blaschek et al. [33]

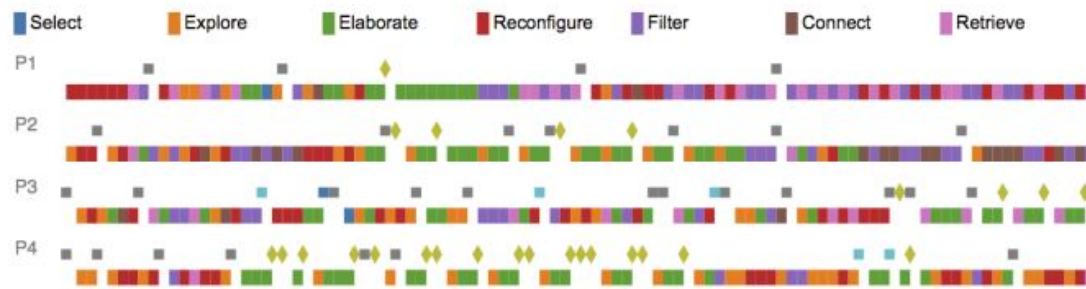


Figure 3.6: Visualization of interaction log with details on demand as proposed by Guo et al. [35]

Blaschek et al. have summarized state-of-the-art approaches to visualizing eye-tracking data [37]. Eye-tracking data can be visualized either in a temporal, spatial or spatio-temporal way, static or animated, two-dimensional or three-dimensional and either be displayed for a single observer or multiple observers. In the state-of-the-art-report, Blaschek et al. present line charts, bar charts and scatter plots as possible visualization techniques. In addition, there are attention maps and scanpaths, as shown in Figure 3.8. Another approach to visualizing eye-tracking data, as also mentioned by Blaschek et al. [37] are space-time cubes. Tominski et al. [38] use the space-time cube to visualize data that can be presented as a trajectory. Figure 3.9 shows a space-time cube displaying the development of traffic jams in San Francisco. This visualization includes spatial and temporal display. With the colors and the bands, spatial information as well as

3. RELATED WORK

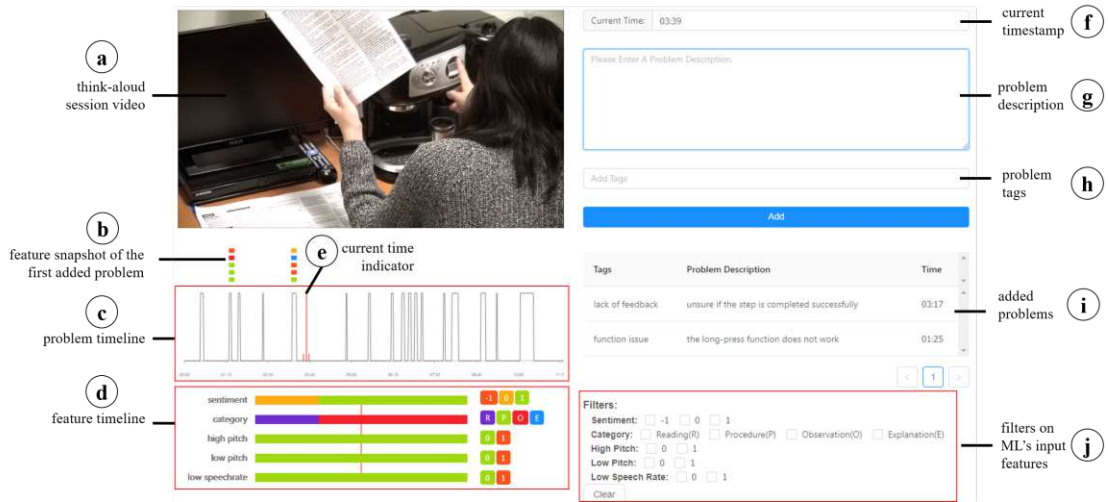


Figure 3.7: Fan et al. visualization of think-aloud protocols using machine learning models and color-coded timelines [36].

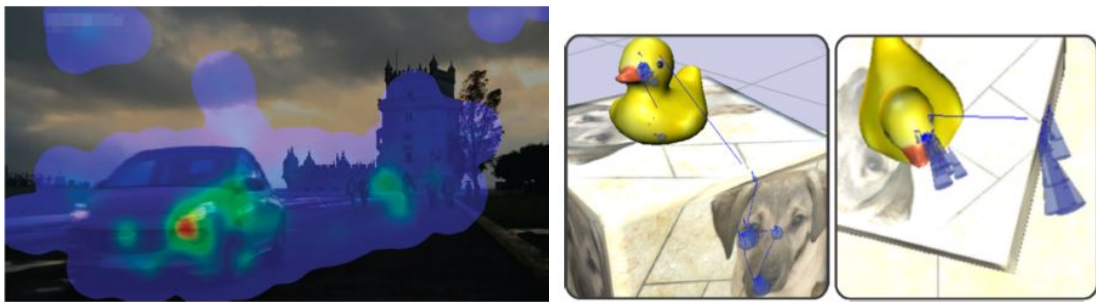


Figure 3.8: Eye-Tracking data visualized with attention maps in two dimensions (left) and scanpaths in three dimensions (right) [37].

temporal information is visualized. In 2016 Blaschek et al. discussed how think-aloud and interaction data can be combined with eye-tracking data to gain deeper insight into user behavior [33]. Because in this thesis we need something more accessible, we equate the mouse trajectory to eye-tracking data on the basis of research done by Navalpakkam et al. [39]. They have shown that the users' mouse-movement is highly correlated with their eye-movement.

3.3 Visualization of Inter-Observer Variability

The visualization of inter-observer variability can be equated to simply visualizing variable shapes. Conceptually, the source of the variability is not of importance to the visualization itself. However, the source gives indications about the nature of the data which can influence decisions such as encodings. The following section discusses the

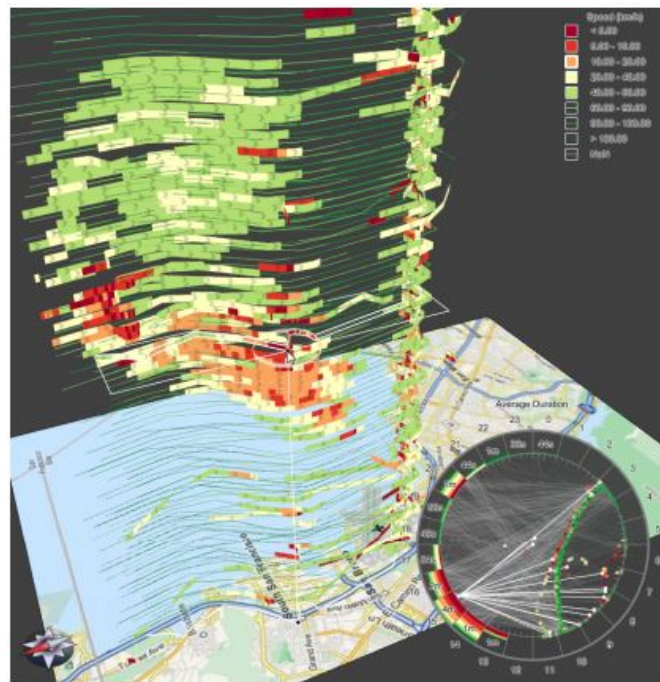


Figure 3.9: Eye-Tracking data visualized with a space-time cube representing the development of traffic jams in San Francisco [38].

visualization inter-observer variability, intra-observer variability and variable shapes in general. Regions of high variability can be correlated with uncertainty.

Recent work by Steenbergen et al. uses different colors, to visualize the different delineations on prostate tumors [3]. Jungo et al. visualize uncertainty originating from medical image segmentation in order to estimate uncertainties [8], using a color scale. These approaches are shown in Figure 3.10. Lelandais et al. use intensity values on a two-dimensional image to convey the plausibility of certain regions in segmentations [40]. The higher the intensity of a pixel, the higher the segmentation rate. Similarly, Linsen et al. use bright colors to highlight regions of uncertainty on segmentations of brain scans [41]. In this approach, darker colors signalize certainty. Prašni et al. [42] apply uncertainty iso-lines which depict uncertainty through the distance between lines.

Saha et al. [43] explored the agreement between observers using the Dice-coefficient in the form of overlaps from segmenting tumors in breasts with a bounding box. Furthermore, Raidou et al. [44] proposed a tool allowing the user to analyze uncertainty in pelvic organ segmentations. Jungo et al. [8] researched the use of fusion methods based on inter-observer data for the creation of ground truth using a synthetic dataset.

Despite the well accepted way to visualize variability in two dimensions, there are also approaches to visualize variability in three dimensions. Reiter et al. partly uses envelopes to represent the variability of the bladder over time [6]. The difference in

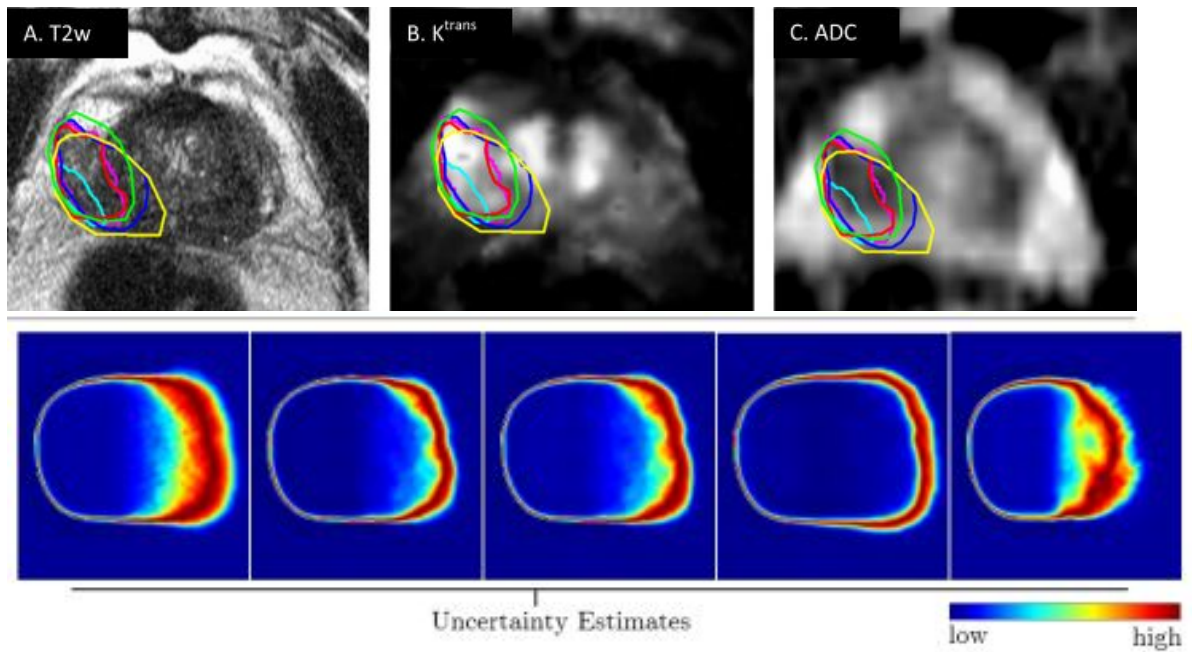


Figure 3.10: 2D visualizations of variable data by Steenbergen et al. [3] (left) and Jungo et al. [8] (right)

this visualization is that regions of high variability are identified and colored, as seen in Figure 3.11. Similarly, work by Grossmann et al. displays the variability of the pelvis using a generalized 3D contour boxplot [5]. The resulting contour boxplots are projected onto the axial planes and so the variability is shown in two dimensions as well. This type of visualization can be seen in Figure 3.11. Another related approach is proposed by Furmanová et al. where the variability is visualized using the color's saturation in three dimensions [45]. In addition, the differences are also visualized in 2D, using cross glyphs and regions of uncertainty highlighted using saliency. This two-dimensional approach is shown in Figure 3.12.

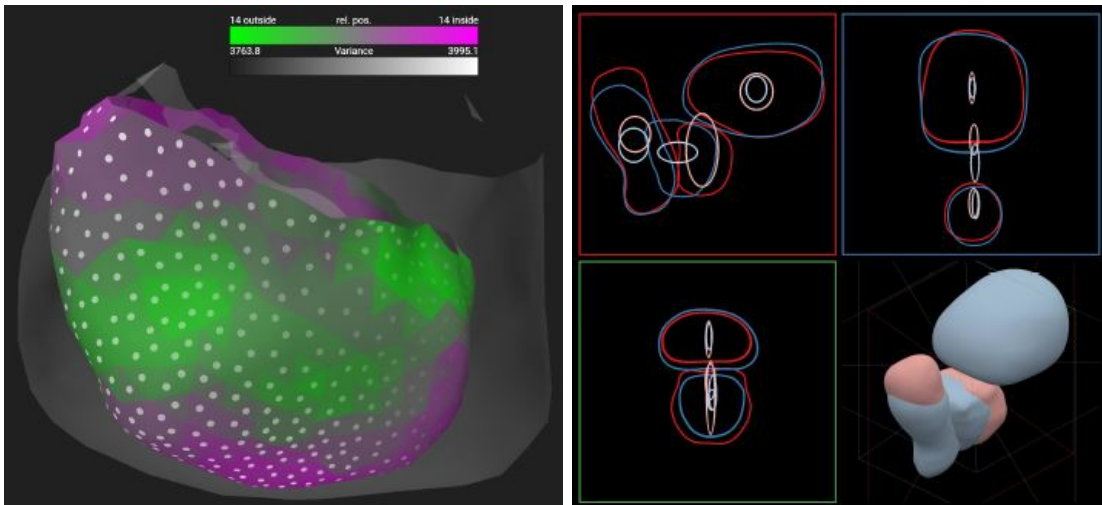


Figure 3.11: 3D overview visualizations by Reiter et al. [6] (left) and Grossmann et al. [5] (right).

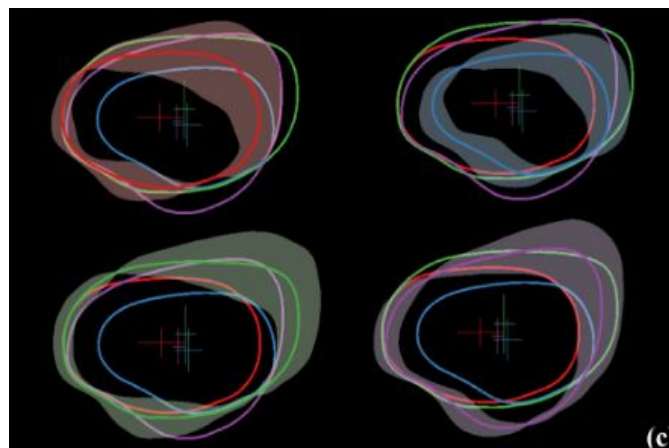


Figure 3.12: Shape variations of bladders from different determined clusters by Furmanová et al. [45]



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Methodology

The following chapter discusses the methodology with which the research question is approached. The pipeline of this paper is shown in Figure 4.1. First, we define the essential tasks and questions that will be answered in the field of inter-observer variability. We want to explore *uncertainty* in the form of inter-observer variability. Therefore, it first needs to be defined, which we do in Section 4.1 After settling for definitions and tasks, we acquire the appropriate data needed to explore the inter-observer variability in respect to the observers' reasoning strategies. This process will be described conceptually in Section 4.2. Lastly, we visualize the data in ways that attempt to answer questions emerging from the resulting inter-observer variability. The visualization step allows us to explore causes for the variability stemming from the thought process, which will be discussed in Section 4.3.

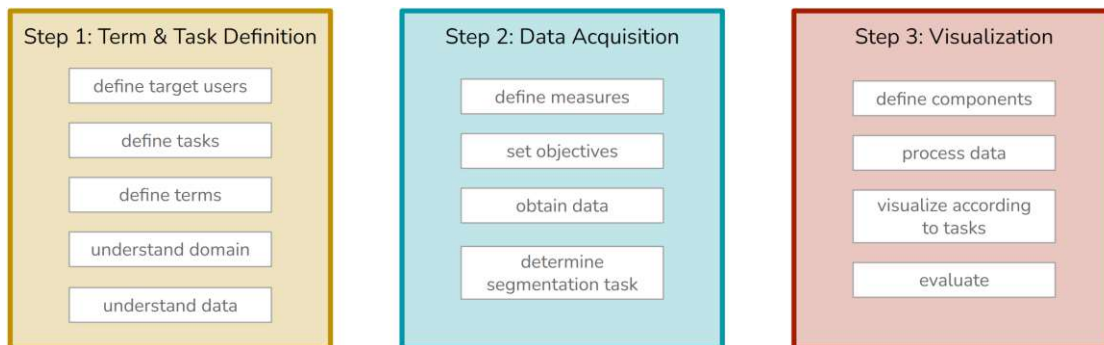


Figure 4.1: Schematic illustration of the methodology applied in this thesis divided into three steps.

4.1 Step 1: Defining Terms and Tasks

In this section we determine the terms and their meanings to our domain, to give a clear and unified understanding of the goal and focus of this thesis.

4.1.1 Uncertainty

In the scope of this thesis we assume two sources of uncertainty:

1. Uncertainty in the form of inter-observer variability and
2. uncertainty taking place during the delineation process.

First, we define uncertainty originating from inter-observer variability as lack of consensus between the different observers in their delineations. This uncertainty is based on the outcome of the segmentation task and is influenced by an observer's thoughts and reasoning process and decisions made during the delineation process. The first type of uncertainty appears in the form of inter-observer variability and can be due to a lack of consensus for instance.

The second type of uncertainty is related to external factors that make the observers follow different decision making processes. In the delineation process, uncertainty usually occurs in areas, where an observer lingered for a longer period of time. This type of uncertainty stems from the observer's choices and thoughts directly. We correlate this uncertainty with the decision-making process. A difficult decision is commonly made with more thought and time. The observer has to weigh out different scenarios and find reasons that support either one or the other option to acquire the desired outcome. The process of making a difficult choice, considering varying outcomes based on one's decision and finding reasons to support either one of the choices will cause an observer to linger on a position. In contrast, an easy decision calls for less consideration and balancing and is therefore done faster and with less uncertainty. In this case one choice is clearly more desirable than the other, due to the outcome. Hence, we directly relate the decision time and consequently the lingering time in an area to uncertainty.

4.1.2 Inter-Observer Variability

As briefly described in Section 2.2.2, inter-observer variability is commonly defined as different segmentation outcomes of multiple observers looking at the same data. In this thesis we do not deviate from this definition. When we address inter-observer variability, we talk about varying *results* from a delineation task, where several observers segmented the same input image. Figure 4.2 shows an example of inter-observer variability on a two-dimensional image. In this thesis we link the variability from the segmentation tasks as well as lingering on a position or revisiting a position to uncertainty. In the thought process uncertainty is shown when words or phrases that convey ambiguity, hesitance or lack of clarity.

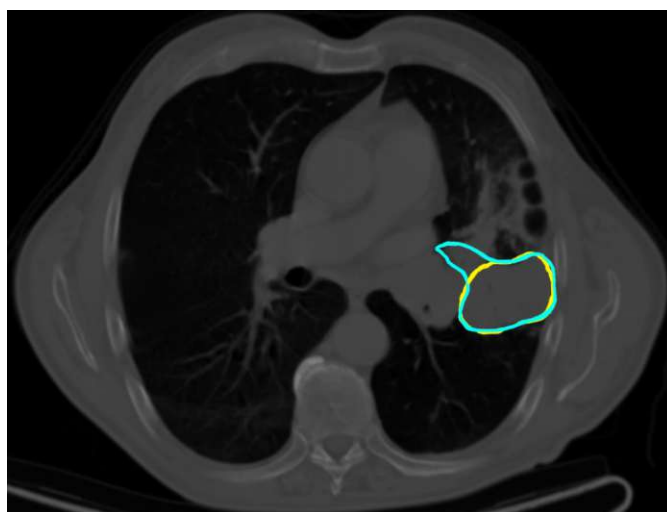


Figure 4.2: Variability of a segmentation originating from two observers (green, blue) on a two-dimensional slice extracted from a CT scan. Data from *TCIA* [18].

4.1.3 Target Users and Task Definition

In order to define tasks, we first need to determine the target users for our application. Generally, the data used for this thesis is medical image data, specifically tumor data. Therefore, the target users mainly include radiologists and aspiring radiologists, since they segment tumors in their daily tasks. The reasoning process and outcome of the segmentation of other radiologists is of high interest between experts. Visualizing inter-observer variability helps to understand different approaches and to be aware of phenomena that can possibly influence the final decision. In our application we use CT data, but the approach can be generalized to any volumetric imaging data set. So, we conduct the data acquisition not only with medical experts but also with novices and imagery from another domain.

With our target users in mind, we can formulate the questions that we want to answer with our visualization. These questions are based on radiologists' needs and questions that arise in the segmentation pipeline:

- (Q1) How does scan and segmentation look like in three dimensions?
- (Q2) Where are possible collisions and complications with other landmarks of the scan?
- (Q3) How do multiple segmentations differ from each other and what do they encompass?
- (Q4) Where is the segmentation in union with the ground truth and where are parts over-segmented and under-segmented?
- (Q5) What parts of the target object can to some degree of certainty be designated as common ground or uncertain?

- (Q6) Where are areas of uncertainty and attention?
- (Q7) What are reasons for lingering and hesitation?
- (Q8) What was the segmentation approach?
- (Q9) What is the thought and reasoning behind the segmentation?

To develop our visualization we define tasks, based on the questions formulated. Beyond answering these questions our tool shall follow the Shneiderman's mantra on visualization: Overview first, zoom and filter, then details-on-demand [46]. The application shall be easy and quick to use independent of the domain. Features include the following: Understand reasoning processes behind delineations, highlight emerging uncertainties, crystallize sources of uncertainty, compare segmentation outcomes among observers, gain insight to segmentation processes and compare segmentation approaches. These requirements stem from the questions formulated previously and user-interaction schemes that are common practice in the field and provide good usability. From the determined requirements we derive the following tasks:

- (T1) Exploratory and comparative visualization of scan and segmentation data in context (Q1-Q5)
- (T2) Collection of additional data from delineation tasks (Q6-Q9)
- (T3) Exploratory visualization of attention data (Q6-Q7)
- (T4) Exploratory visualization of think-aloud protocol with details on demand (Q8-Q9)

4.2 Step 2: Data Acquisition

To be able to visualize the thought process and reasoning strategies of the manual delineation process, the relevant data is needed (T2). In addition to the medical images that provide the data about the scanned object, we also capture the thought process in a think-aloud protocol, the segmentation outcome as well as the observers' attention areas. This is important, because the data gives us valuable information in addition to the segmentation outcomes.

4.2.1 Reasoning Strategies

In the capturing process we encourage the observers to think aloud and talk about their decisions. The focus lies in understanding why an observer chooses to delineate certain parts of the scan and why they exclude other areas. We equate these choices with the observers' reasoning strategies. The verbal commentary to the segmentation task is recorded and transcribed semi-automatically to obtain the necessary data in a form that can later be used in the data processing step. Moving on, we will refer to this think-aloud

protocol as the thought process or reasoning strategies. This data provides the basis for Task 4 (**T4**).

Furthermore, we take the approach to the segmentation task into account for the thought process. In some cases observers choose to scroll through the entire scan to get an overview of the scan and locate the target object. In other cases observers go to the area they assume the object is located in and start assessing the scan from there. When segmenting the target object, the direction in which the slices are delineated vary between observers: some choose to delineate starting from the top slice of the target object, some from the bottom slice and others switch back and forth between different parts of the target object without following a direction. To capture this information we log the position of the pointing device used for the segmentation task.

4.2.2 Attention Areas

The reasoning strategies represent one aspect of the observers' choices. However, when segmenting, the observers' attention might fall to other objects in the scene and deviate from the target object. Other areas that observers focus on equate to areas of interest. The areas can display objects or phenomena that influence the delineation. Therefore it is important to take those areas into account when analyzing the observers' strategies and exploring causes for inter-observer variability. This data provides the basis for Task 3 (**T3**).

The attention can commonly be logged using an eye-tracking device. However, eye-tracking devices are expensive in time and cost, since the acquisition process needs to be conducted in person at a location with the required setup. Another alternative to an eye-tracking device such as depicted in Figure 4.3, would be to use an eye-tracking software such as GazeRecorder [47] and EyeVido [48]. The software itself would have either needed to be installed by the observer on their own device which could have led to problems depending on the participant's system and technical knowledge. In addition, these two examples for eye-tracking software would have been reduced to a time-limited trial. Another option would have been to install eye-tracking software on the administrator's system and create a set-up for the data acquisition process, but this would have again required the participant's to come to the location and increased the time and complexity for one data acquisition session. Therefore we choose to obtain this data using the mouse-movement of the observer. As Navalpakkam et al. have shown in their research from 2013 the users' mouse-movements are highly correlated with their eye-movement [39]. We are aware that this correlation will not be as exact as an eye-tracker has the potential to be. Nonetheless, we determined this method to be the best suiting for our research question and it outweighs an eye-tracker in its advantages.

4.2.3 Segmentation

The resulting segmentations of each individual observer provide the central information to analyze inter-observer variability. To obtain the information about what parts of the

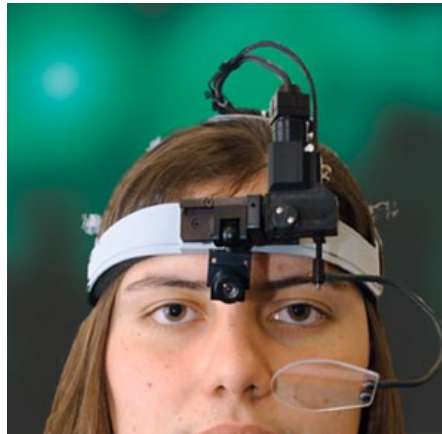


Figure 4.3: Eye-Tracking device [49].

image were delineated by the observer, the segmentation task is done in a logging tool with the ability to log the delineations for each slice combined in one additional file. The delineations temporarily correspond to the mouse-movement which is due to the fact that the delineations are done with the mouse or pointing device. This data provides the basis for Task 1 (**T1**).

4.3 Step 3: Visualization

Each visualization component in our tool is based on the questions previously defined. They depict common problems experts in the field of radiology encounter when dealing with uncertainty originating from inter-observer variability in medical images. We fulfill the tasks (**T1**, **T3**, **T4**) defined and develop visualizations for the acquired data.

4.3.1 Scan Data

To be able to see the segmentations in context (**Q1**), it is necessary to visualize the medical image data. We choose to display the scan data in a three-dimensional space. This shall provide a better understanding of the spatial dimensions of the scan and the segmentations for novice users. In a two-dimensional view, the depth value gets lost and so, it is difficult to tell how large the segmented object is compared to the rest of the scan and its context. The visualization of the scan should allow to threshold the tissue at various values and should therefore be adjustable, depending on the desired tissue value. So, one can change between context visualization of bones, soft tissue or other structures of interest with a specific density value.

4.3.2 Segmentation Data

The delineation of the target object happens manually on two-dimensional slices. Therefore, it is not known how the segmentation looks like in the three-dimensional space. The

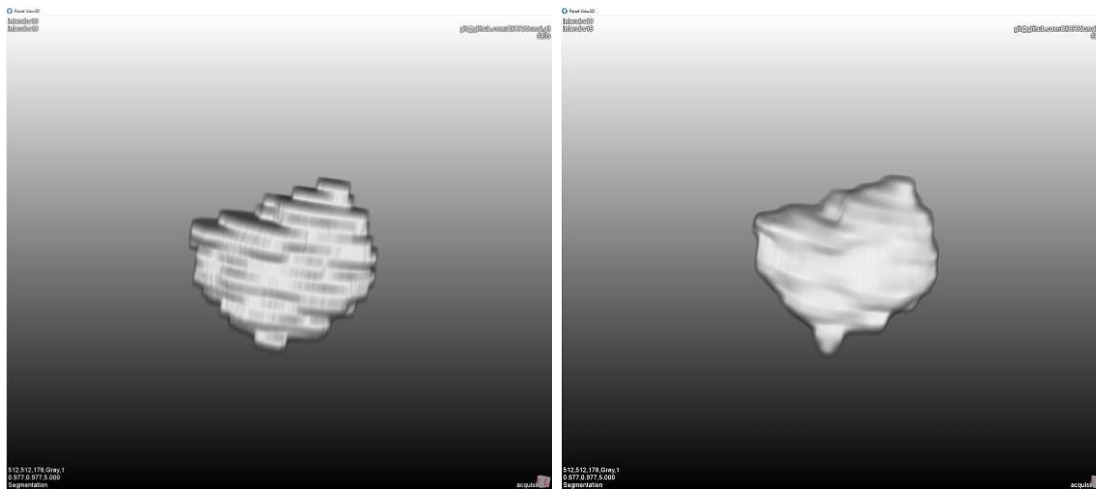


Figure 4.4: Raw segmentation data (left) and segmentation data with interpolation between slices (right). Data from TCIA [18].

location of the segmentation in comparison to the scanned object can be misunderstood if having it seen in two-dimensional slices only. It can be difficult for a novice to tell how the segmentation lies in the scanned object compared to other landmarks in the scan (**Q2**). Doctors have a good spatial understanding in 2D slices. To diminish the uncertainty originating from the exact location of the target object, we display the segmentation outcome in the three-dimensional space. It can so be combined with the scan and be observed in the full three-dimensional space. Previous work, as mentioned in Section 3.1, has also visualized variable shapes three-dimensionally.

To be able to compare two segmentations effectively (**Q3**), we choose to have two different ways to visualize the segmentation data. As described before we show the segmentation in the three-dimensional space. One visualization is a volume placed in the according location. The volume is composed of the segmentation in each slice, displayed with the thickness of one slice in the three-dimensional space. The slices are not interpolated in between, since that would include additional data and show altered data, as depicted in Figure 4.4. The left image shows a surface reconstruction without interpolation and the right one is interpolated. This visualization allows a quick and comprehensive overview of the segmentation outcome.

However, to compare multiple segmentations this representation is not sufficient. Overlaying two segmentations with the same visualization technique, using the same opacity and same color makes comparing them in a static environment very difficult. Displaying the segmentations dynamically — adding and removing a segmentation from the view — allows for a comparison. Additionally, we allow the visualization of a segmentation with its silhouette. The silhouette solely displays the edges of the segmentation and allows for a good comparison with other segmentations. In addition the silhouette enables to see what the segmentation encompasses. In case of a medical image data, one will

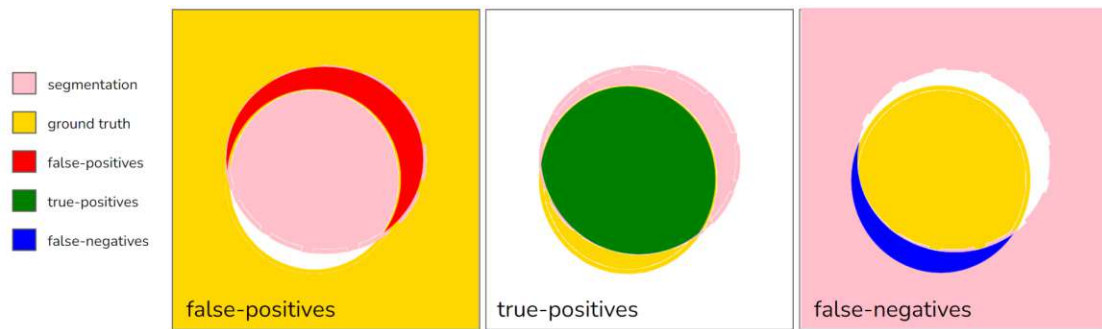


Figure 4.5: Logic of the computation of the ground truth comparison (false-positives, true-positives, false-negatives).

be able to see the tissue that is covered by the segmentation. If the target object is close to other objects in the scan, this visualization enables to see if that object was mistakenly delineated as well. In case a ground truth is determined the task is to see where a segmentation overlaps with the ground truth.

To determine areas where the observer over- and under-segmented, and in which regions the segmentation overlaps with the ground truth (**Q4**), we visualize the voxels labeled as false-positives, false-negatives and true-positives respectively. We use the determined ground truth and an individual segmentation and the inverted values to compute the respective voxels. The false-positives are all those voxels, where the segmentation was delineated but there is no value for the ground truth. The false-negatives are all those voxels, where the segmentation was not delineated, but there is a ground truth value. Figure 4.5 shows the logic for this comparison. The true-positives are all those voxels, where the segmentation and the ground truth overlap. To see where a delineation is over-segmented we look at the false-positives and to see where a delineation is under-segmented we look at the false-negatives. We omit the true-negatives because they are derivable from the other three comparisons and lead to visual.

Comparing a single segmentation with the ground truth delivers information about outliers and common ground of one segmentation only. However, we also want to analyze the entire spectrum of the segmentations and explore the inter-observer variability. With the different delineations, not considering a ground truth, we can draw conclusions and make decisions based on the expertise of the observers only. To answer **Q5** we look at the distribution of outliers of all acquired segmentations. Analyzing which voxels were delineated as part of the target object and which ones were not by every observer respectively, gives the information needed. We count the amount of voxels covered by how many observers. The voxels delineated by every segmentation can be seen as common ground and for target objects of variable shapes possibly also determined as the ground truth. Voxels that were delineated only once or twice represent outliers of the target object and can be seen as uncertain, since they do not represent agreement across observers.

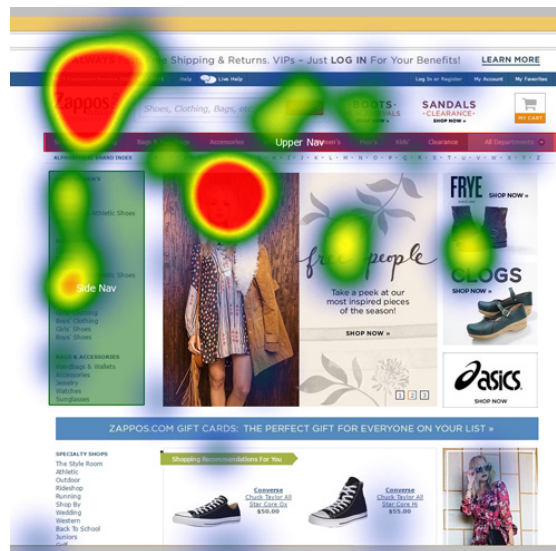


Figure 4.6: Eye-Tracking data visualized using heatmaps [50].

4.3.3 Attention Data

We measure the attention given to voxel positions in milliseconds as the cursor position. The data obtained is sorted by the amount of milliseconds spent on every voxel visited. With more time spent on a voxel, we link that the observer was more invested in that certain voxel or rather area (**Q7**, **Q8**). A higher attention span could be related to uncertainty such that the area in question was not clear and needed more thought and analysis, which equates to uncertainty. With the accumulated times for each position in the scan, we create a heatmap, as is commonly used with eye-tracking data. Figure 4.6 shows an example of eye-tracking data being visualized with heatmaps. The image shows an example of eye-tracking data on a website, which is two dimensional. The data we acquire is two-dimensional for each slice as well. However, accumulated over the whole scan we obtain attention data in three dimensions.

Using the same data, Figure 4.7 shows gazeplots and focusmaps, which present other alternatives to visualize attention data [50]. However, focusmaps are centered around the focus itself and therefore do not allow for the rest of the data to be seen, which in our case is essential for the context. We chose to do heatmaps over gazeplots due to the color-coding, which is easier to read than area-encoding. In addition, gazeplots do not allow for a voxel-exact visualization. The area of attention might be anywhere within the circled area. For our purpose, we need a voxel-exact visualization to be able to identify regions of uncertainty. Gazeplots however, have the possibility to show the sequence of attention, which can be of importance for future work. With the heatmap we can answer the question of what areas were uncertain to the observer and where we can expect variability.

The approach to the segmentation task can be different between observers. We define

4. METHODOLOGY

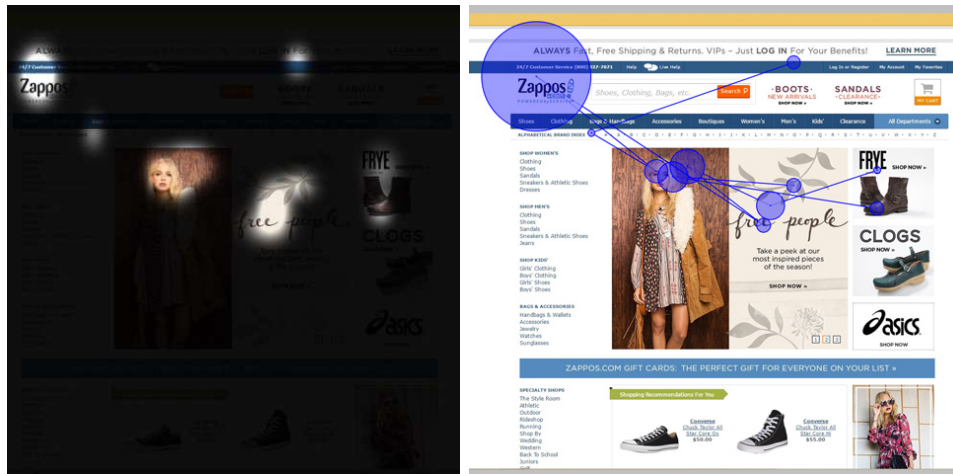


Figure 4.7: Eye-Tracking data visualized using Focus Maps (left) and Gazeplots (right) [50].

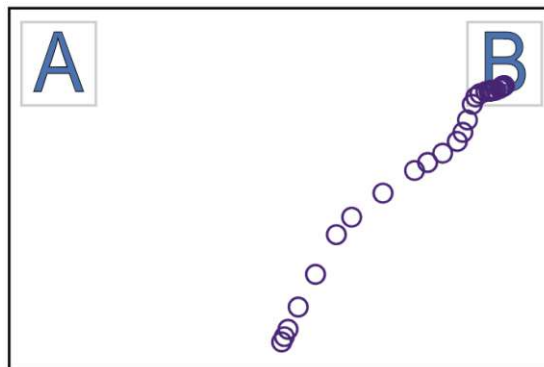


Figure 4.8: Mouse trajectory visualized using circles at every logged position [51] for a binary choice (between A and B).

the approach as the trajectory of the mouse. The trajectory can have different directions depending on where the observer started to segment. The frequency with which an observer switches the slices in the segmentation task may also vary. To visualize these attributes of the attention data and answer the question of what the segmentation approach looks like (Q8), we visualize the mouse-movement as a continuous line according to the logged positions. An alternative is to visualize the positions, where the mouse was logged, which leads to clusters around positions where more time was spent on and a sparser distribution whenever longer distances are covered with one mouse-movement, as seen in Figure 4.8.

4.3.4 Reasoning Data

With the think-aloud protocol we can get an insight into what the observer was reasoning at different positions (**Q9**). Summarizing the transcript and extracting relevant words for the domain can give a good overview about the observers' thoughts. The thoughts, aligned with the positions they appeared at, shall answer the question for reasons of lingering and uncertainty. Extracting relevant phrases and words of the respective domain, will crystallize other objects and phenomena that were mentioned during the segmentation progress. Identifying other relevant topics emerging in the think-aloud protocol allows for related subjects to be highlighted. The related subjects might correlate to hindering objects and phenomena influencing the segmentation outcome. Showing only a small percentage of the whole think-aloud protocol might take utterances out of context. Therefore, we want to give a detailed view and provide the full transcript for slices on demand. This will make information about the detailed and exact thought of the observer at every slice accessible. Thus, enabling to read the full thought process and putting utterances into context as well.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Data Acquisition

To obtain data that captures an observer's thought process and attention points as well as the outcome of the delineation, we conducted a data acquisition. This chapter describes the methodology of this data acquisition step and presents its outcome.

5.1 Goal

The initial data, as will be described in Section 5.3.1 only consists of individual segmentations, but we need the additional info as it will provide insights on the process of the delineations. We needed to reproduce the segmentation process and capture the necessary additional data. The focus of the data acquisition laid on the process of delineating the object and the observer's thoughts and attention points. Thus, the participants were strongly encouraged to think aloud and comment on their every step. To mimic the setting of a manual delineation, a logging tool was created that allowed the participants to segment an object by marking the visible parts in every slice of the CT. With this logging tool we conduct two experiments: one with expert users, knowledgeable in the medical field and one with novice users.

5.2 Participants

The data acquisition was conducted with nine people. We distinguish between observers with a medical background (experts) and observers with no experience in the applied medical field (novices). In our application for the medical example it is necessary to know how to locate a tumor when only given the slices. Splitting the participants among two groups allows for a comparison between two use cases. The results from the novices provide an abstraction to our approach, since the process and task are equal and only the imagery changes. This ensures that the two experiments are comparable. Table 5.1 shows the demographics of the participants. The participants were given one of two tasks

Table 5.1: Data acquisition participants sorted by their status based on the experience in the applied medical field.

	Gender	Age	Highest educational level	Profession
Expert 1	Female	45 - 54	PhD	General physician
Expert 2	Female	18 - 24	High school	Medicine student
Expert 3	Female	18 - 24	High school	Medicine student
Expert 4	Male	18 - 24	High school	Medicine student
Expert 5	Female	18 - 24	High school	Medicine student
Novice 1	Male	25 - 34	Master's degree	University assistant
Novice 2	Male	33 - 44	Master's degree	Researcher of Visual Analytics
Novice 3	Male	25 - 34	Master's degree	Data science
Novice 4	Male	25 - 34	PhD	Computer Science

based on their status. The expert participants were chosen based on whether they would be able to find a tumor in the CT scan, when shown the slices, by their own estimation. Novice participants were given a different dataset, which included a gun in a backpack. The age range of all participants lies between 18 and 54 and the highest educational level ranges from a graduating high school to a PhD. With the exception of Expert 5, where the preferred language was German, the data acquisition was conducted in English. The German session needed some additional data processing steps that will be discussed in Section 6.2.2.1.

5.3 Segmentation and Logging Software

The logging tool used for the data acquisition was created with *MeVisLab*, a modular framework for image processing and development with a special focus on medical imaging [52]. Figure 5.1 shows the interface of the logging tool. On the left side there is a *volume view* that shows the scanned object in 3D. This view can be rotated interactively, allowing the observer to get a view of the complete scan and giving them an overview of other landmarks. This mechanism is usually not provided in a real setting. However, it was important to include, so the observer can allocate the object to delineate faster and orient themselves with other objects in the scan. Even though in case of the human body, the structure would not change significantly between scans, for other arbitrary objects, the structure is not known prior. On the right side of the interface the *slice view* shows the scanned object in slices. By scrolling through this view, the observer can move vertically along the longitudinal axis through the scan. For this approach only one axis was provided. This view is also used for the manual delineation. By using a mouse, the observer can mark the object in every slice where it is visible and thereby create a manual segmentation.

The logging tool logs the observer's mouse movement by saving the two-dimensional

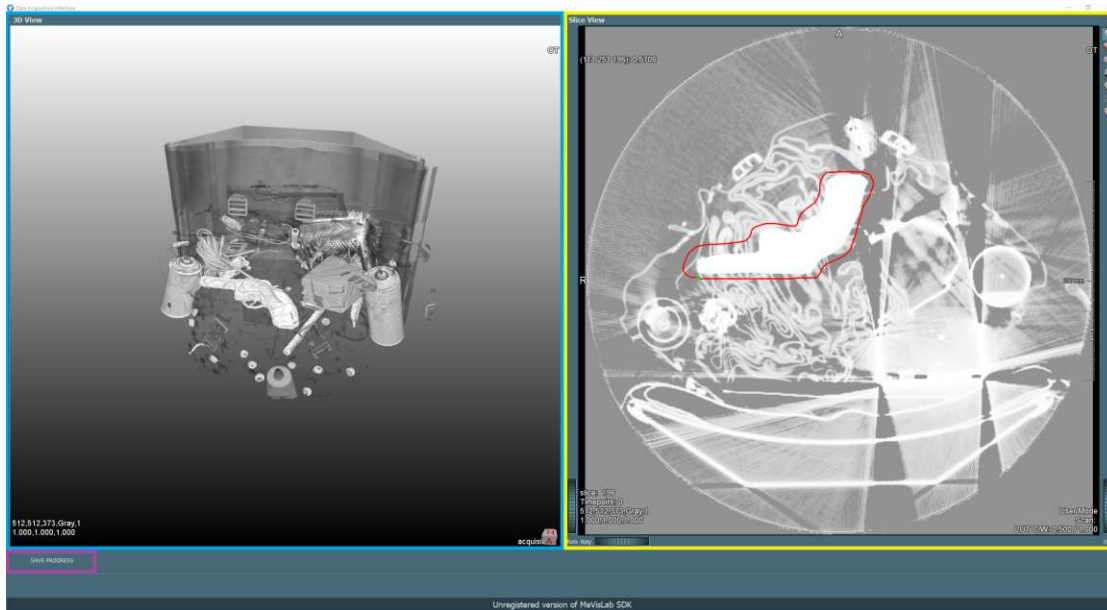


Figure 5.1: Logging tool interface created with *MeVisLab* [52]. Blue: Volume View. Yellow: Slice View. Purple: Save Progress Button.

position on a slice as well as the slice number, such that it results in a three-dimensional coordinate corresponding to one voxel in the scanned object. In addition, the logging tool logs when the observer uses the volume View, the value of the pixel the mouse is currently positioned on as well as if the left mouse key is pressed or not. So, the information is not logged in temporal intervals, but when the information itself changes, which can happen more or less frequently than a preset temporal resolution. Lastly, the interface features a button for the observer to save their delineation progress in the bottom left of the window.

5.3.1 Data

For the two experiments of this data acquisition, two different datasets were created corresponding to the experts and novices. For the **experts** the cancer data was obtained from the *Cancer Imaging Archive (TCIA)*. The retrieved data consists of 22 CT scans from non-small cell lung cancer radiotherapy patients. The segmentations are results of manually adjusted delineations of an auto-segmentation tool that was not further specified in the data description. They were created by five different radiation oncologists, of which two were trainees at the time of delineation [18]. From this dataset, one patient with a tumor in their right lung was chosen. This scan was then used for the acquisition with the experts. The gold standard in this dataset was determined in the given dataset. However, it is not possible to define a definite ground truth for tumor segmentations, because of their nature and inaccessibility. This is why we call it the gold standard for this dataset. The dataset consists of 178 slices of which the tumor is visible in 10.

For the **novices** a more simple dataset was needed with an object to delineate that people without experience in the medical field can identify. To simulate a familiar situation, the scan of a backpack was used and an artificial dataset was created by merging the scan of a backpack with the 3D model of a gun. The scan of the backpack was retrieved from *Open Scientific Visualization Datasets* [53] and the gun was downloaded from *Turbosquid* [54]. The model of the gun was first modified such that it represented a volume with a value of 1 inside the mesh and a value of 0 outside. Then, Perlin Noise was added and a bit of smoothing also was performed such that the scan would look more realistic and simulate a real scan. Lastly the gun was merged into the scan of the backpack using *MeVisLab* [52] and served as the dataset for the novices, making them take the role of TSA officers (Transportation Security Administration). The outline of the gun itself, which is determined through the creation of the artificial dataset, is considered the ground truth. This dataset consists of 372 slices of which the gun is visible in 51.

5.3.2 Implementation

As described above, the logging tool is separated into a *volume view* (left) and a *slice view* (right), showcased in Figure 5.1. The views are implemented using the *View3D* and *View2D* module respectively. The datasets described in Section 5.3.1 serve as the input data for the logging tool pipeline. Before displaying the image, it is run through a *histogram equalization* module, so that the contrast of the input image is automatically adjusted to its values and the image is correctly visible. The adjusted image is then displayed directly using a *View2D* module for the slice view and a *View3D* module for the volume view.

The segmentation feature on the slice view is implemented by adding a *SoView2DCSO-ExtensibleEditor* module with the input from a *SoSCOLiveWireEditor*, which sets the features of the lines drawn by the user and a *SoCSOVisualizationSettings* module, where the colors and shapes of the segmentation line can be set. To store and edit the segmentations, the *SoView2DCSOExtensibleEditor* requires a *CSOManager*. *MeVisLab* saves the segmentations as *Contour Segmentation Objects (.cso)* [55]. Using *MeVisLab*, these files can later be converted to more commonly used file formats, like *DICOM (.dcm)*.

The network including the described modules forms a new, custom module which is then instanced in a new network. To log the activities of an observer, the *MLogFile* module is added to the new network. This module allows the creation of a logfile. The script files of the created networks contain so-called *FieldListeners* that observe variables and in case of a change of value trigger the specified callback functions defined in the python files of the respective module. The variables that are listened to are listed in Table 5.2 and discussed further in Section 5.3.3. The callback function then logs the information using the function `MLAB.log()`. Figure 5.2 shows the scheme of how the callback functions and logging system is implemented in the logging tool. This mechanism is used in combination with the *CSOSave* module to save the segmentations by clicking on the “Save Progress” button in the bottom left of the logging tool interface.

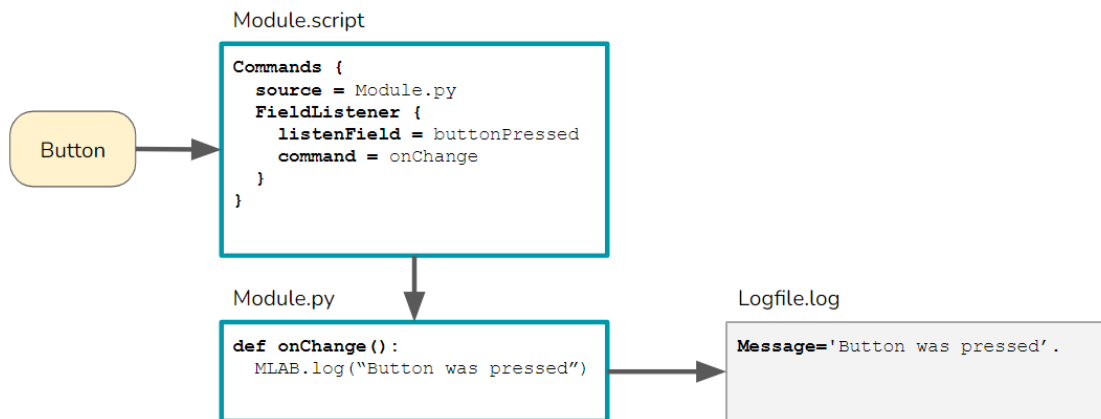


Figure 5.2: Logging scheme in MeVisLab.

Table 5.2: Logged information and their respective triggers.

Trigger	Logged Information
run logging tool	Message=' DD-MM-YYYY, HH:MM:SS.f, Starting prototype' .
Slice View: mouse position changed	Message=' DD-MM-YYYY, HH:MM:SS.f, Position changed, Position_X, Position_Y, Position_Z, Value, Segmentation, Clicked' .
Slice View: value changed	Message=' DD-MM-YYYY, HH:MM:SS.f, Position changed, Position_X, Position_Y, Position_Z, Value, Segmentation, Clicked' .
Slice View: new segmentation started	Message=' DD-MM-YYYY, HH:MM:SS.f, Position changed, Position_X, Position_Y, Position_Z, Value, Segmentation, Clicked' .
Slice View: left mouse key pressed	Message=' DD-MM-YYYY, HH:MM:SS.f, Position changed, Position_X, Position_Y, Position_Z, Value, Segmentation, Clicked' .
Volume View: mouse position changed	Message=' DD-MM-YYYY, HH:MM:SS.f, 3D View' .
Save Progress: button pressed	Message=' DD-MM-YYYY, HH:MM:SS.f, Saving progress' .

5.3.3 Data Collection

The data acquisition was logged using the scheme and modules as described in Section 5.3.2. Table 5.2 shows the information that was collected in the final logfile and their respective triggers. The information is saved in a .txt file. The information is formatted such that it can be converted into a .csv file in little steps, which will be discussed in

Section 6.2.2.1. In addition to the numerical data from the logging tool, the thought process data during the acquisition was logged as well. The participants were encouraged to think aloud. To log the thought process, the data acquisition session was recorded and its transcription delivered the necessary data. It effectively logged the timestamps and the utterances of the participants. The transcript and recording were automatically created with the meeting software used. The transcript contains all the utterances in the format of HH:MM:SS <utterance>. In comparison to the logfile created from the logging tool, the transcript does not log the milliseconds, which is a limitation from the recording software and not a conscious design choice.

5.3.4 Transcript Collection

The data acquisition was administered online. Due to the ongoing COVID-19 pandemic and the participants' availability, online calls provided a great flexibility and it was chosen to conduct the data acquisition via online meetings. In addition, various online meeting software offers automatic transcription of the meeting which was necessary to capture think-aloud protocol. *Zoom* [56] provided the needed tools for the data acquisition process: the possibility of a transcript of the call and the remote control of the administrator's screen. Other online meeting platforms such as *Microsoft Teams* or *Google Hangout* were considered as well. However, the lag in the remote control function of Microsoft Teams would not allow reasonable control over a remote desktop and Google Hangout did not provide such a function at all. It should be mentioned that the transcription function in Microsoft Teams gives more accurate results, but since correcting the transcript manually in a post-processing step was inevitable, Zoom was chosen as the tool to work with.

5.4 Procedure

The participants were invited to a Zoom meeting. Beforehand, each participant was sent a consent form, a short description of the goal and content of the data acquisition, as well as a survey about their background and demographics, of which the results are listed in Table 5.1. Upon entry to the meeting, the participants were again given another quick introduction to how the data acquisition would proceed as well as which data would be collected. After giving consent, the meeting was recorded and the participants were given control over the administrator's desktop. The logging tool, described in Section 5.3, was ready to use on the administrator's setup. The setup of the data acquisition is shown in Figure 5.3. The administrator shared their screen with the logging tool running, using the online-meeting software. Participants were given remote control to the administrator's screen and had the form open separately. The meeting was recorded on the administrator's PC.

First, the participants were given a warm-up task to get familiar with the tool. Next, they had to segment the actual data. Lastly, the observers had to fill in a survey about how confident they felt with their answers. These steps will be described in more detail in the following subsections.

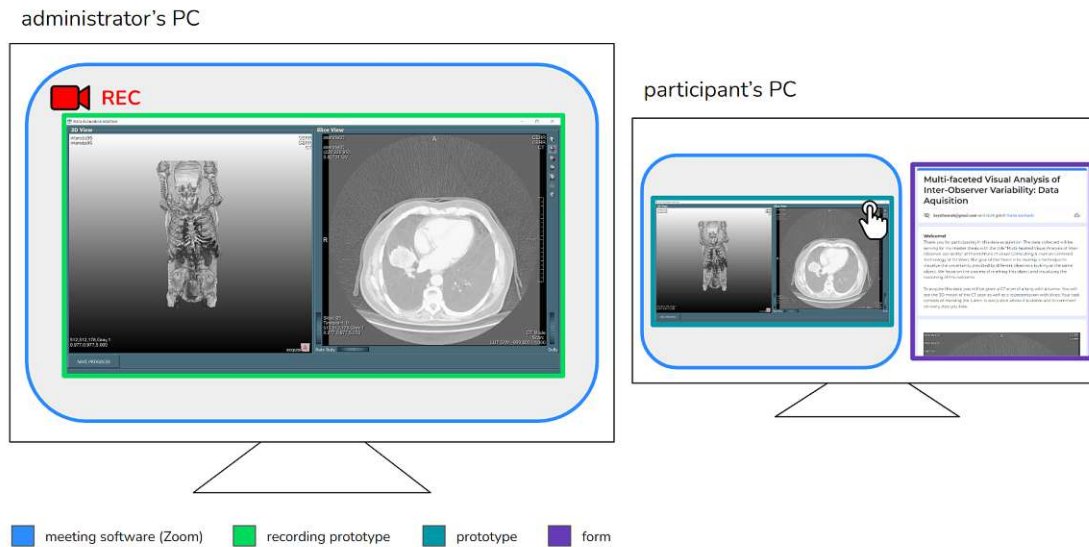


Figure 5.3: Data Acquisition setup with the administrator’s PC and the participant’s PC and the respective open programs.

5.4.1 Warm-Up Task

To get familiar with the controls and the functions of the logging tool, the participants were given a warm-up task. This task used a simple version of the novice dataset as described in Section 5.3.1, consisting of 372 slices and the gun being visible in approximately 30. The dataset was adapted such that the gun was placed orthogonal to all axes in the backpack scan, so that it is easy to locate and segment. The goal for the participants was to segment the gun in every slice where it was visible. With this task the participants were reminded and it was emphasized to think aloud and comment on everything they see and think and explain their reasoning strategies. The warm-up task as well as the welcome message presented to the participants is attached in Appendix A.

By introducing a warm-up task, the objective was to decrease any variability that could come from using the logging tool like missing knowledge about using the interface, lags or hardware complications. In addition, the warm-up task helped in creating a space where the participant feels comfortable with the setting of thinking aloud and essentially talking to themselves. As the administrator tried to intervene as little as possible, so the outcome of the thought process would solely be the participants’ thoughts. Whenever participants felt comfortable thinking aloud, understood the controls of the logging tool and had the goal of the data acquisition in mind, the warm-up task was finished by informing the administrator.

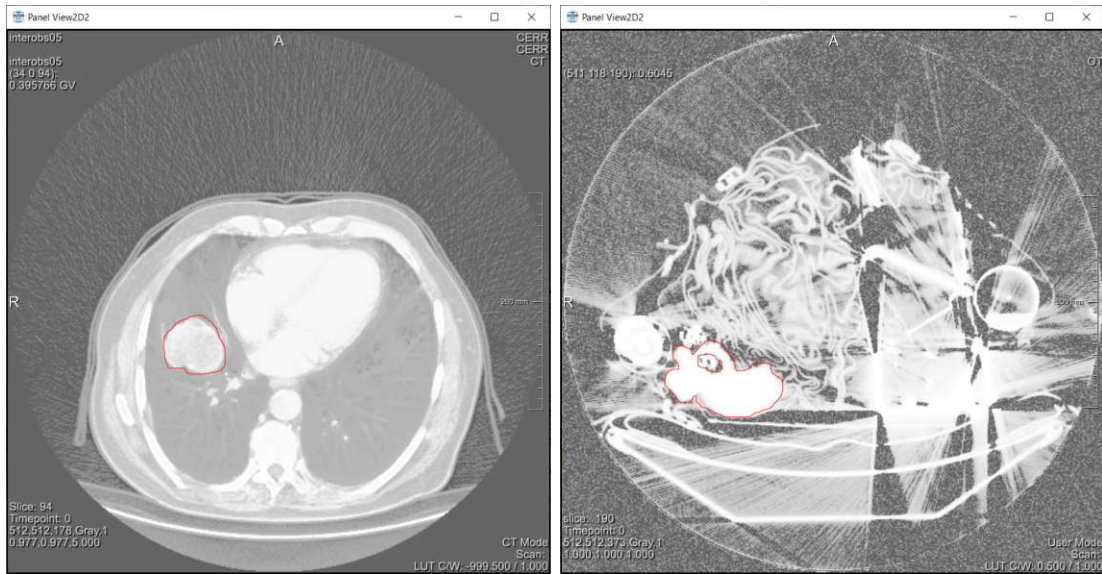


Figure 5.4: Slice View of the Segmentation task with a delineation of the target object (red) for experts (left) and for novices (right).

5.4.2 Segmentation Task

After completing the warm-up task, participants were given a description to the segmentation task in written form. While getting familiar with the next task, the administrator set up the correct dataset according to the participant’s status of novice or expert. The goal of this task was to get a qualitative thought process data of the participants and capture their reasoning, their fixations and the results of their segmentations.

Experts

Experts received the data of a patient with a tumor in their right lung, described in Section 5.3.1. To provide the same environment for all participants the task was set to start on slice 40 and the volume was set to face the coronal plane.

Novices

Novices were given a similar dataset as in the warm-up task with the difference that the gun was now placed arbitrarily and more blurry than before, as described in Section 5.3.1. To provide the same environment for all participants, the task was set to start on slice 170, and the volume was set to face the coronal plane.

Figure 5.4 shows the slice views of the segmentation tasks. This task was considered completed when the participant decided that there were no more slices or parts to segment. They then proceeded by clicking on the “Save Progress” button and were prompted to continue by filling out the confidence survey.

5.4.3 Confidence Survey

The confidence survey was conducted after the segmentation task in regards to the participants' experience with that task. The goal of this questionnaire was to find a correlation between an observer's estimation of their own confidence and their acquired data. The survey consisted of six statements that were to be answered with a five-point Likert scale [57] with the possible answers "strongly agree", "agree", "neutral", "disagree" and "strongly disagree". These statements were composed of the following:

1. I understood the assignment.
2. It was easy to find the gun/tumor.
3. The borders of the gun/tumor were clear.
4. Other objects/parts of the body hindered me from segmenting the gun/tumor.
5. I am confident in my decisions.
6. The segmentation task was easy to solve.

In addition, an open voluntary question was added to further describe a participant's segmenting process:

- Describe your segmentation process when a line or border was unclear to see.

The confidence survey as well as all other introductions and information were communicated through a form created with *Google Forms* [58]. These forms were handled anonymously.

5.5 Results

For each participant, which will be called observers interchangeably, the acquired data is described in Table 5.3. The data is processed in the next step before being used for the visualization. The exact data processing step is described in Section 6.2.2.1.

Figure 5.5 summarizes the answers of the six statements of the confidence survey for all nine participants. All participants either strongly agreed or agreed with understanding, showing that the assignment was clear and there is no variability to be expected from the understanding of an individual observer. Furthermore, observers thought it was easy to find the gun or tumor with the exception of one person that was neutral towards that statement. One person thought the border of the object was rather unclear, while four other people stood neutral towards that statement and four more people agreed. This shows that even though the object was easy to locate, determining its shape proved to be more difficult. Since a total of five people agreed or completely agreed with the fact

Table 5.3: Resulting data from the acquisition process, their respective source and filetype.

Data	Source	Filetype
Segmentation outcome (warm-up task)	MeVisLab	.cso
Segmentation outcome (segmentation task)	MeVisLab	.cso
Logfile (warm-up task)	MeVisLab	.txt
Logfile (segmentation task)	MeVisLab	.txt
Video recording of the meeting	Zoom	.m4a
Transcript of the meeting	Zoom	.txt
Audio recording of the meeting	Zoom	.mp4

that objects were hindering them from segmenting the object, it can be assumed that other objects were one of the reasons why the borders were unclear to some participants. The indications made about how confident the participants were in their decisions and the indications about the difficulty of solving the segmentation task are similar. Based on this trend an assumption can be made that the difficulty of the task is related to the observer's confidence.

The parallel coordinate diagram in Figure 5.6 shows that the experts finished the task faster than the novices, which could also be due to the fact that there are more slices to segment in the segmentation task created for novices. The fastest participant was an expert with a completion time of six minutes. Since, in practice, delineations are done in as little time as possible, shorter times are plotted higher and aligned with a more positive outcome. Looking at the last two axes of the parallel coordinates, there is no obvious relation between the time taken by a participant and their confidence. However, for six participants the confidence in their decisions directly correlates with their assessment about the difficulty of the task. A similar relation can be observed between the understanding of the assignment and the ease with which the object was located.

The answers to the open question show that the experts mostly relied on the slice view of the logging tool and referenced their knowledge of the human anatomy and other physiological parts. One expert reported that they use the color of the scan as well to make decisions about their segmentations. The experts also used consecutive slices to make up the shape of the tumor and construct their segmentation. Looking at consecutive slices was also a strategy used by the novices. In addition, novices explained that they used the volume view to orient and differentiate as well. All participants regardless of their task mentioned that they used the slice view and looked at previous and subsequent slices to understand the object and determine where a delineation shall be made or not. One participant justified this technique by saying "... the noise and artefacts generally change much faster than outlines of solid objects between the slices, so going through slices helped with the slices a lot".

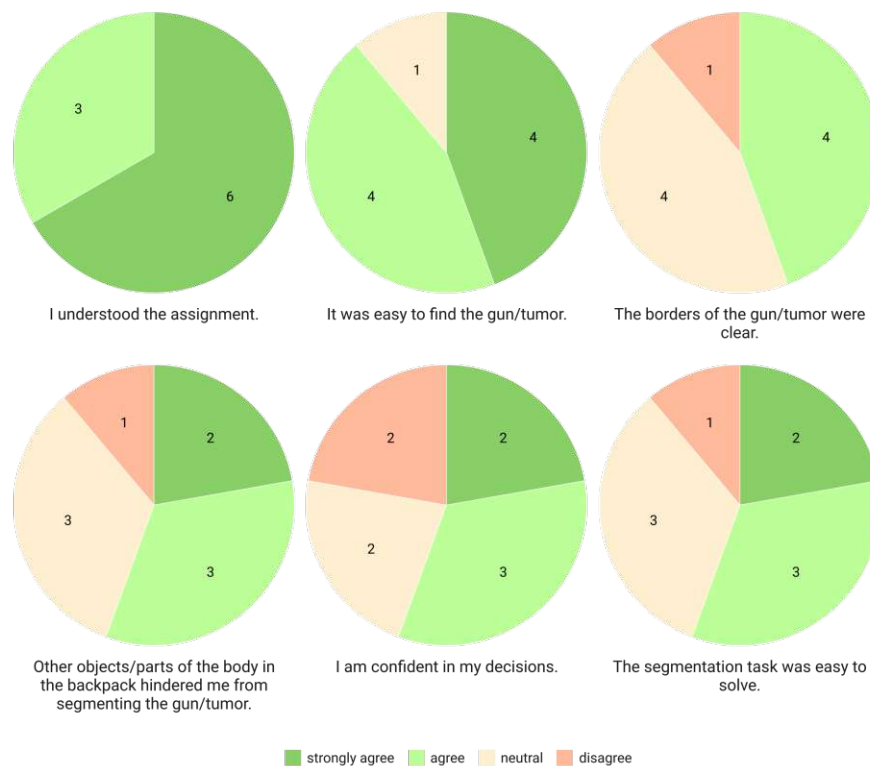


Figure 5.5: Results of the confidence survey of all nine participants.

5.5.1 Improvements

The outcome of the segmentations showed that delineations in some cases were not completed in the process, which resulted in an open curve, rather than a closed surface. A closed surface was desired so it could then be turned into a binary mask, determining the delineation per slice. The possibility for the creation of open curves as seen in Figure 5.7 is a feature of the logging tool. Since the binary masks are necessary in the next steps, an additional pipeline was needed to fix the open curves and create closed surfaces. In retrospect, this additional step could have been avoided by using a *SoCSOSplineEditor* module instead of a *SoCSOLiveWireEditor* module as described in Section 5.3.2, which will close a curve automatically, as soon as the mouse is released and so creates a closed surface. With the used logging tool all delineations needed correction.

5.5.2 Summary

All participants were able to complete their data acquisition sessions. The logging tool crashed one time during the session of Novice 2, so the session took longer, since the segmentation was lost and needed to be redone. This problem did not occur again. In total, for the experts we logged 38.688 interactions, recorded 9197 words and saved 58

5. DATA ACQUISITION

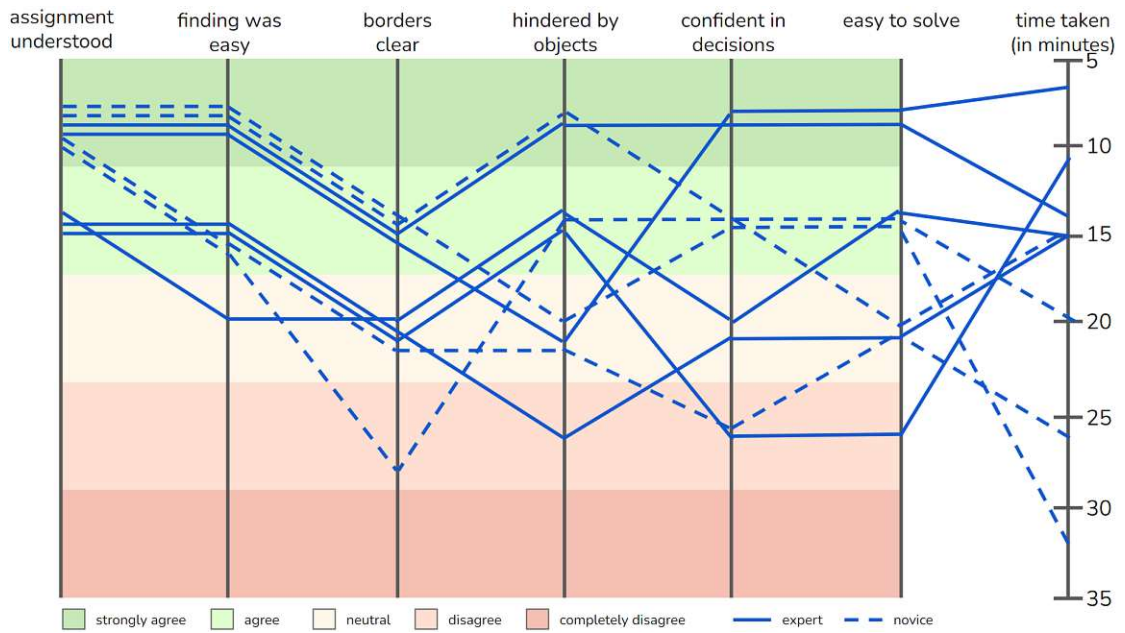


Figure 5.6: Results of the confidence survey of all nine aligned in parallel coordinates. Full lines represent experts. Dotted lines represent novices.



Figure 5.7: Open segmentation curve on the gun, highlighted with the yellow rectangle.

segmentations. For novices we logged 91244 interactions, recorded 13183 words and saved 271 segmentations. The full cleaned transcripts are available in Appendix B.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 6

MOLE

MOLE or *Multi-Observer Looking Environment* is a tool, created to be able to compare and look at segmentation tasks, done by multiple observers. The visualization of each individual observation as well as a combined view of all segmentation data shall give insight to regions with a high factor of variability as well as an overview and detailed information about the process of the delineation. The following chapter will cover both the specifics of the implementation of *MOLE* as well as the design choices made.

6.1 Visualization and Interaction Design

The goal of *MOLE* was to provide a system that allows the user to have a detailed insight of an observer's thought process and the outcome of their delineation. Since the dataset drawn from the TCIA [18] only provided segmentations and no data about the thought process, we conducted a data acquisition as explained in Section 5.3. From the data acquisition we obtain the following three data items:

1. Segmentation data
2. Attention data
3. Reasoning data

We continue with this data and design visualization components that will answer our questions (**Q1** - **Q9**) and fulfill our requirements. The components and their advantages are explained in this Section. Every component is designed to answer at least one of the defined questions.

6. MOLE

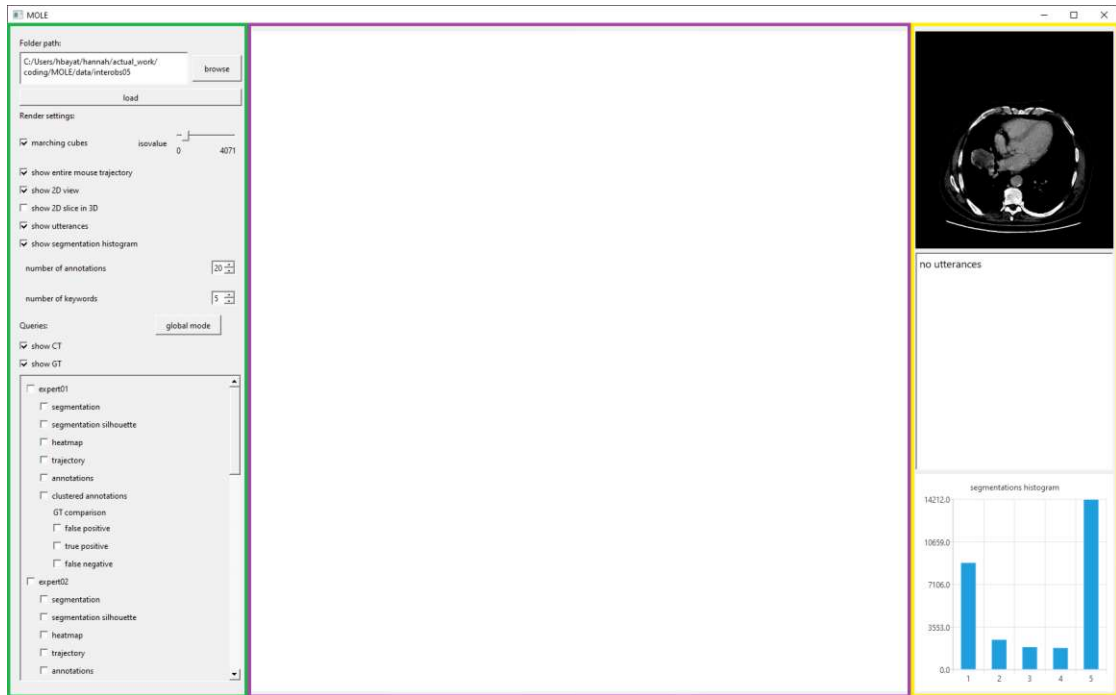


Figure 6.1: MOLE layout overview. Render settings (green). Render window (purple). Detail windows (yellow).

Interface. The interface of our tool shall be clear and easy to use as defined in our requirements. It is separated into three parts as shown in Figure 6.1. On the left side of the window (marked in green), the user can set the settings of how the information is being rendered and set queries on what is displayed. The main window in the center (marked in purple), also referred to as render window, will adjust dynamically to the queries selected on the left, which allows a quick way to compare the observers' outcomes, by adding and removing visualization components. The visualization components are provided for every observer and grouped under one super-ordinate checkbox that allows to show and remove all the components for one observer. On the right side, details will be displayed on demand. The windows on the right (marked in yellow) will only be shown when the according checkbox on the left side is selected. This selective display allows the largest possible space for the render window, when analyzing the information visualized in the render window.

Segmentation. To be able to see segmentations in context, as defined for Task **T1**, we need to find means of visualizing the delineations. The delineation result is visualized as an iso-surface, displaying the slices segmented by the observer as a volume. As described in Section 5.5.1, the logging tool allows for an observer to create open curves which now have to be closed to create binary masks that define an outer area and an inner area of the segmentation. A morphological operation is applied to the segmentations that dilate

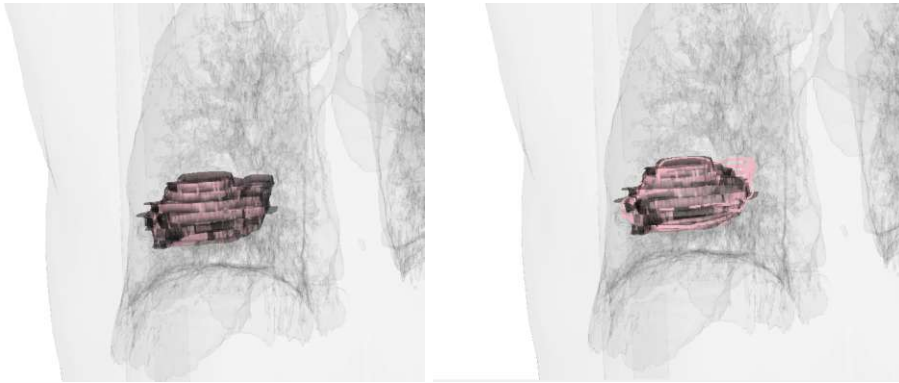


Figure 6.2: Overlay of two segmentations using the marching cube representation for both segmentation (left) and a silhouette representation combined with a marching cube representation (right).

the image which closes potential gaps and erodes again, restoring the original shape of the segmentation. Two morphological operations are applied in sequence: the first being a closing operation and the second an erosion. The closed contour needs to be filled. Again, a closing morphology is applied and afterwards by finding the contours in the image and creating a bounding rectangle around the contour, we compute the filled contour and result in a binary image where the segmentation has a value of 1 and the outer values are 0. We chose to not interpolate the values in-between two slices, because an interpolation would not correspond to a more accurate visualization and would only distort the result. Thus, the segmentations are displayed with a staircase-like transition.

Segmentation silhouette. As comparing multiple delineations is central to Task **T1** the silhouette for the delineation was added. The segmentation by itself was hard to see when layered with other visualization components. The silhouette allows to see the outlines of the segmentation in dependence of the viewing perspective at all times and from all angles, which enables to compare it with other segmentations for instance. Figure 6.2 shows the advantage of overlaying the silhouette of a segmentation to another segmentation, rather than overlaying two segmentations in their iso-surface representation. On the left the two segmentations are harder to distinguish, whereas where one segmentation is represented as its silhouette, they can be compared more easily. It is visible that the silhouetted segmentation is delineated more in the top right of the object than the other segmentation. However, overlaying multiple delineations with the iso-surface representation increases the opacity in areas where the segmentations overlap. Therefore, a lower opacity refers to a higher agreement or simply an overlap.

Ground truth comparison. To be able to have a quantitative measure of how accurate an observer's delineation was, compared to the determined ground truth, we implemented a ground truth comparison. This function allows to see where the observer correctly delineated (true-positives), over-segmented (false-positives) or under-segmented



Figure 6.3: Piecewise colorscale for the heatmap component with linear interpolation.

(false-negatives), which poses as question **Q4**. With this component we can compare how well different observers did, by adding and removing the individual comparisons to the scene. The comparisons are computed for each individual observer. By adding multiple comparisons to the scene, they are accumulated. As with the segmentations, the opacities are added up so a lower opacity, stands for a higher agreement.

Heatmap. To be able to see where the user spent the most time and fulfill Task **T3**, we created a three-dimensional heatmap. This allows to see areas the observer struggled at and other areas of interest. As seen in Figure 6.4, between multiple observers there was a region in the back on the right of the object where more time was spent at, which may imply a difficult area to segment. When displaying multiple heatmaps, the values are not accumulated. So, to compare multiple heatmaps one would dynamically add or remove the heatmaps to see the difference. The initial data includes the amount of milliseconds spent at each logged position. However, visualizing those values gave only very little insight to areas of higher attention. Therefore, the data is blurred using a Gaussian blur. This allows to identify areas of attention, rather than voxels of interest. The heatmaps are available for each observer. To be able to summarize all observers' attention areas and identify common areas of uncertainty and lingering (**Q5**), we provide a global heatmap. This heatmap accumulates all the times spent on all voxels by all observers. The colorscale for the heatmap starts with white for the minimum values, so positions that no time was spent on, are essentially not visualized in the heatmap. It interpolates linearly between the minimum and an eighth of the maximum, where it reaches the color "dodgerblue", shown in Figure 6.3. At one fourth of the maximum the colorscale is set to "gold" and between this value and the maximum, it is interpolated between "gold" and "red". We created this colorscale to fit the distribution of the time values.

Trajectory. Visualizing the trajectory of the mouse enables seeing movement patterns and shows the observer's approach to segmenting the target object (**T3** and **Q8**). Figure 6.5 shows the difference between the mouse trajectory of two observers. One focused on the area the object was located in and the other looked through the scan, before focusing on the area of interest. This component shall give an insight into what areas where visited with the mouse and with high probability been looked at.

Annotations. The full thought process can be very long and not all utterances are of importance to understand reasoning process. Showcasing only a few selected words of the utterances, gives an overview about the thoughts of the observer at the most visited positions. By filtering out nouns, proper nouns and adjectives we attempt to summarize

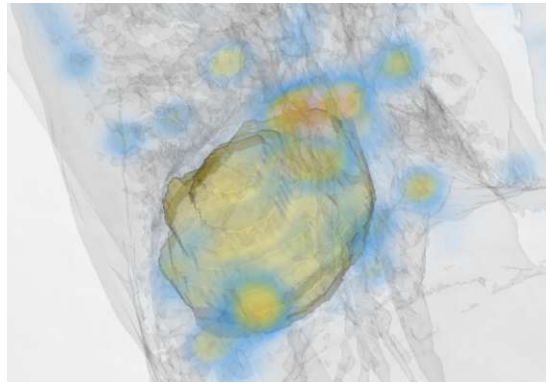


Figure 6.4: Overlay of two heatmaps. The ground truth is displayed in gold.

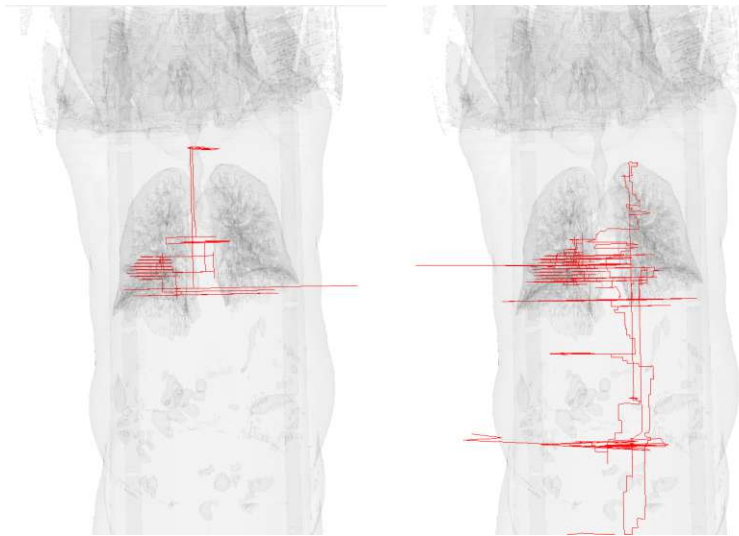


Figure 6.5: Two mouse trajectories with different approaches: observing the area of the object to delineate only (left) and observing the entire scan (right).

the thought process in a few more meaningful words. This shall allow for question **Q9** to be answered. These annotations provide keywords, filtered with a natural language processing tool.

Clustered annotations. The basic annotations, mentioned before are queried from the positions, where an observer spent the most time. These positions often lie very close to each other in space, as seen in Figure 6.6. Therefore, we added clustered annotations that are created with the k-means clustering algorithm. The k-means algorithm separates data into n groups, called clusters of equal size where the sum of the squares of the distance between the data points in one cluster and its center is minimal. Based on this criteria the algorithm results in k centroids which are the positions of the clusters formed. Each data point is associated with one cluster and consequently one centroid. This delivers

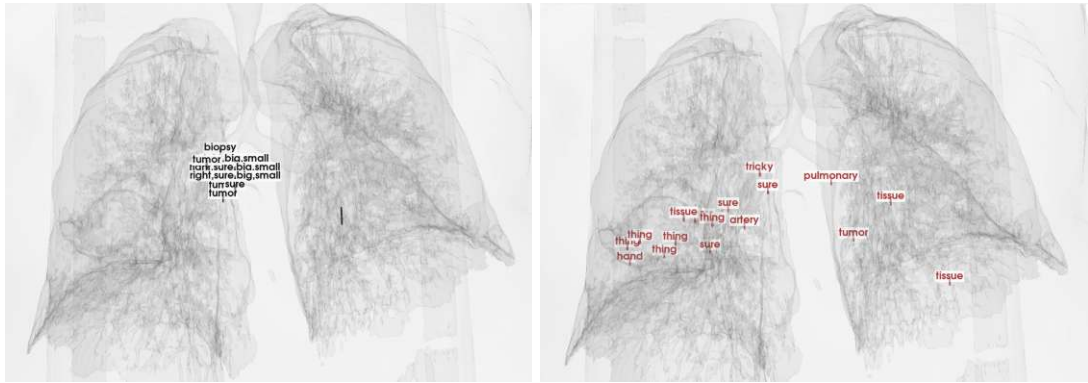


Figure 6.6: Basic annotations positioned very close to each other (left) and clustered annotations spread across the scan (right).

the voxels that form the center of a cluster, which in return corresponds to a point of interest grouping other points visited by the observer in the area of that cluster center voxel. Afterwards, the annotations are obtained by weighing the tokenized utterances with a *term frequency-inverse document frequency* (tf-idf). Tf-idf is a numerical weighing strategy to determine a word's importance in a document, which in our case are the keywords per cluster. To compute the importance of a word, we first compute the term frequency TF (Equation 6.1). Then we compute the inverse document frequency IDF (Equation 6.2). The final tf-idf scores are computed with the amounts counted before and the weights for each word.

$$TF(w) = \frac{\text{number of times word } w \text{ appears in a document}}{\text{total number of words in the document}} \quad (6.1)$$

$$IDF(w) = \log_e \frac{\text{total number of documents}}{\text{number of documents with word } w \text{ in it}} \quad (6.2)$$

The clustered annotations do not match the basic annotations and thus provide another perspective to the keywords provided by an observer's utterances.

2D View. The 2D view of the CT scan, was necessary to incorporate, since the connection between what an observer saw, when segmenting and the 3D visualization of their delineation result was missing. The 2D view in combination with the slice visualization in 3D as seen in Figure 6.7, allows this desired link. This function can provide information on the reasons for a certain delineation or for an extended time spent segmenting (**T1**). This view is linked to the render window. The 2D view provides details that might only be of interest on demand. When not in use, the space is used to maximize the render window and therefore the 2D view must be enabled when needed.

Utterances. Task **T4** requires for details on demand. As mentioned before the thought process can include a lot of text. However, to be able to track the entire thought process and reasoning strategy on demand, we still provide the full transcript. This optional

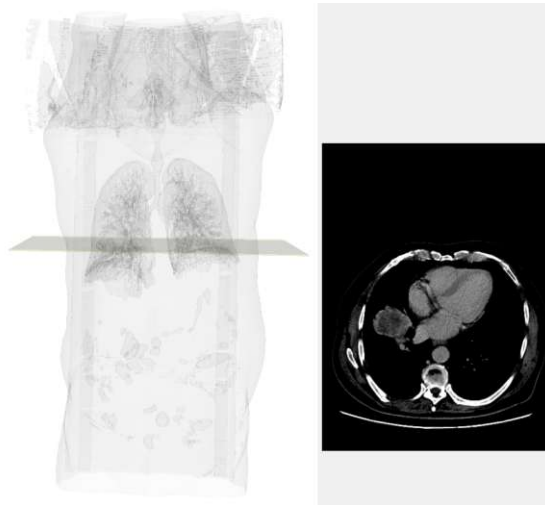


Figure 6.7: Render window with a scan and the selected 2D slice visualized in 3D.

information enables to get an insight to the full utterances. Displaying all the utterances, rather than just the keywords allows to read into the exact thought process on a certain slice. For each super-ordinate checkbox selected, the utterances in the slice, displayed in the 2D View, are accumulated and shown in the text box. Both the slice view in the render window as well as the utterance text box will change dynamically as the slice in the 2D view is changed.

Segmentation Histogram. The histogram presenting the voxels covered by the number of possible segmentations gives an overview about the commonalities of all available segmentations as well as information about the amount of outliers. This component offers another approach to Task **T1**. The histogram shows the amount of voxels that are covered by only one segmentations, two, three, four or all segmentations. By selecting a single bar, the correspondent voxels are shown. In addition, the histogram provides the functionality of analyzing the delineations independent of the observers and is classified by the amount of delineations in agreement. This function is also provided on demand.

6.1.1 Discarded Features

While developing our tool two main features were planned that later were discarded. First, similar to the segmentation histogram there was the idea to display the density values of the CT scan in a histogram as well. The goal was to visualize the density values of the segmentation in the same histogram, essentially comparing the density value of the segmentation with the ones from the CT scan. This required a mask of the segmentation, multiplied with the CT scan that would result in an image with the voxels of the CT scan covered by the segmentation. From the resulting image a histogram is computed. Showing the densities for the whole CT scan and the densities of the segmentation-covered

voxels proved to be impracticable, because the amount of voxels covered and therefore the amount of densities shown was very little compared to the densities of the CT scan, which is the reason it turned out to be unhelpful for our tool.

Second, we wanted to find a way to be able to focus on the segmentation while having more context visualization. The idea was to show an X-ray viewer that would allow an insight view to the tumor. At the time of this idea, the scene was cluttered and lead to confusion when displaying more than two objects in the scene. This was due to the arbitrary color and opacity values that were set at the early stages of developing our tool. By placing a sphere that would act like a magic lens [59], focus in context was desired. This approach however was discarded because there was no need for a better or more detailed context visualization. The CT visualization as it was implemented in the final tool, provides context visualization that is user-defined. Due to the settings of the opacities and colors for the different possible objects placed in the scene increased the visibility and decluttered the scene.

6.2 Implementation

This section focuses on the specific implementation of MOLE. This includes the technologies used to achieve the final outcome, the input data and its preparation and a description of the implemented features.

6.2.1 Deployed Technologies

MOLE was implemented in JetBrains' *CLion* [60] using C++. The user interface was implemented using Qt, and *The Visualization Toolkit* (VTK) was used for the visualizations. *Python* [61] was used for data processing and scripting.

6.2.1.1 Visualization Toolkit (VTK)

The Visualization Toolkit (VTK) is an open source software for image processing, volume rendering and scientific visualization amongst others [62]. It is developed in C++, but there are bindings for other programming languages like Python as well. VTK offers a rendering pipeline that can be integrated into the Qt interface. Furthermore, we made use of *vtkImageData*, which was mainly used to store the scan and segmentation data, *vtkImageMathematics* for the computation of accumulated information and comparisons between the ground truth and single segmentations, *vtkDICOMImageReader* to read DICOM files and visualize them and *vtkImageHistogram* to obtain information about the histogram of an arbitrary image. More functions used from the VTK library will be discussed in Section 6.2.3. The documentation for VTK exists online and there is a documentation for every class, however, the examples made available by *Kitware* [63] cover only a small part of the available classes.

An alternative that was used in the beginning for processing DICOM images was *DCMTK*,

which is a collection of libraries and applications including software for processing DICOM image files [64]. The integration of DCMTK into VTK reduces to a pixel-wise transfer of the output from the DCMTK class into the *vtkImageData* class of VTK. However, DCMTK proved to be redundant, as VTK offers a sufficient way to read and process DICOM images. Therefore, DCMTK was discarded.

6.2.1.2 Qt

The visualization in MOLE is embedded in Qt, which is used for the graphical user interface (GUI). Qt is a cross-platform software that allows the creation of GUIs [65]. Since the well-established version of Qt5 does not include the QtCharts, which is used for the histogram display of overlapping segmentations, MOLE uses Qt6 that includes this function. Qt allows for the integration of the render windows that are provided by VTK and thus was used for MOLE, since the visualizations are done with VTK. The integration of the render windows from VTK is done by promoting a *QWidget* object to a *QVTKOpenGLNativeWidget*. These settings are determined either directly in the XML file that holds the information about the user interface or in *QtCreator* which is the accompanying editor to Qt. QtCreator also allows to quickly prototype an interface by drag and dropping interface elements into the editor. In addition, with Qt it is possible to create interface objects at run-time which was necessary to automatically adjust the interface to the number of observers and create a tool, usable with arbitrary multi-observer data. Qt is a well documented software and widely used in programs such as applications developed by *Autodesk*, a 2D- and 3D- development software, the technology company *LG Electronics* and the vehicle automotive brand *Mercedes Benz* amongst others [66].

6.2.1.3 Python

Python allows for quick data processing, which is why it was used in the data processing steps. The library *pandas* [67] used in the scripts written with Python, allowed a quick processing of the output data from the data acquisition and shaping the data into the correct form. Furthermore, to be able to apply the k-means algorithm the *sklearn* library [68] and *spaCy* [69] was used. Furthermore, the imaging library *PIL* [70] was used to process and correct the images created in the data acquisition process. Python is known to be used for machine learning applications and thus was a simple solution to an otherwise complex problem. Processing the raw data and applying machine learning algorithms in C++ would have been computationally expensive and time-intensive at run-time, which is the reason the processing of data is done beforehand.

6.2.2 Data Handling

The data obtained from the data acquisition needs to be structured as seen in Figure 6.8 for it to be properly read by MOLE. When starting the application, the user can select the folder structured as seen, here called “data”. The “gt”-folder contains the ground

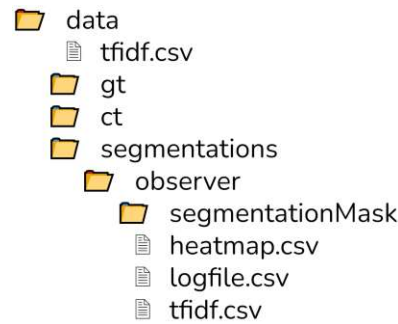


Figure 6.8: Folder and file structure for MOLE data.

truth saved as single DICOM image slices. Similarly, the “ct”-folder contains the ct scan, also in single DICOM image slices. The folder called “segmentations” contains all the observer data — in this case the participants’ results from the data acquisition process. The “segmentationMask”-folder contains the DICOM image slices from an observer’s segmentation. These images are created in the pre-processing step from the contour segment objects resulting from the data acquisition process. The csv-files listed in this structure are the processed files as will be described in Section 6.2.2.1.

6.2.2.1 Pre-Processing

The information gained from the data acquisition, as explained in Section 5.3 now has to be transferred into a shape that then can be processed by MOLE. The logfiles resulting from the data acquisition are transformed into a comma-separated value file (.csv), such that the information is accessed easily with any software or programming language. Since the logfile is already composed in a way that resembles the .csv structure, this step is very simple. In order to align the thought process of an observer to their actions, the transcript of the meeting, needs to be corrected as well. For this task, it is necessary to fix errors made by the transcription tool of Zoom. This is done by listening to the audio of the meeting and manually correcting the errors in the transcript file. For one participant whose data acquisition was conducted in German, the transcript was processed using *Sonix* [71], which is an online tool allowing the English transcription of German Audio. With the corrected transcript file and the raw logfile, it is then possible to create the necessary files that will be processed in MOLE.

As listed in Table 5.3 the acquired segmentations are saved as contour segmentation object files (.cso) which now have to be converted into binary masks, so they can be displayed as a volume. Since the CSO structure is unique to MeVisLab, the segmentation files need to be transformed to DICOM (.dcm) files, which is a universal file format and readable by the technologies used for MOLE.

Logfile. To convert the raw logfile created during the data acquisition process into a .csv file, a script was written that takes the raw output file and the manually corrected

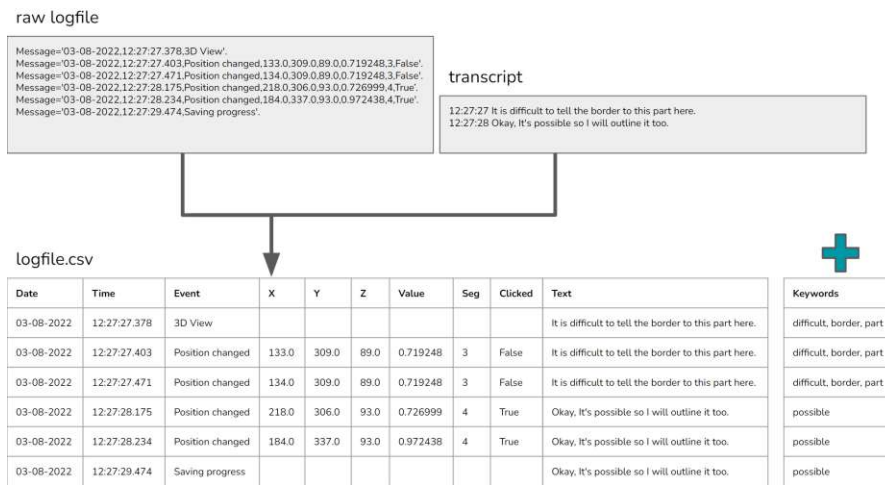


Figure 6.9: Pipeline for the creation of the logfile file for one observer.

transcript file of the meeting as an input and created three files that later will be used to visualize the obtained data in MOLE. The pipeline for this conversion is shown in Figure 6.9. First, the raw logfile is cleaned by removing the redundant text added to the logging messages by MeVisLab. This will result in a comma-separated value file, containing the initially logged data from the prototype. Next, the utterances from the transcript file are aligned to the logfile. The alignment is based on the logged timestamp and the timestamp saved for the utterance. To simplify the aligning process and lack of information about the duration of an utterance, the transcript does not account for how long a participant took for a single utterance and so a single utterance is aligned to a timestamp until a new utterance with a new timestamp is logged. This strategy results in an utterance in every row of the logfile, because a row with a timestamp that does not have an utterance will log the last utterance. After the utterances are aligned to the logfile, the keywords of a single utterance are extracted and added as an additional column. The extraction of the keywords is done using *spaCy*, an open-source natural language processing software library [69]. With the English tokenizer provided by *spaCy* the utterance is being filtered for the keywords. The filtered words from every utterance then serve as the keywords for each row. This table will serve as the logfile used in the implementation of MOLE and is saved as a separate file, named *logfile.csv*.

Second, to display an observer's heatmap, visualizing the positions they spent the most time on, a file accumulating the times spent on each position is needed. The creation of this file is schematized in Figure 6.10. For clearance a separate file is created that will contain the accumulated time per position, the positions' coordinates and the keywords associated with that position. To compute the total amount of milliseconds spent on one position, first the timestamps of two subsequent rows are subtracted, which results in the time spent on a single logged position. The table now consists of the same amount of rows as before but with the additional column that shows how much time was spent on a



Figure 6.10: Pipeline for the creation of the heatmap file sorted by descending time spent on each visited voxel.

logged position. To obtain the desired file for the heatmap, the table is grouped by the position and the columns for the keywords and time spent on a position are accumulated. This file is saved as *heatmap.csv*. The script for this process is attached in Appendix C.1.

Lastly, to obtain the keywords clustered by positions, the logfile is used. Figure 6.11 shows that for clustering a simple k-means clustering was applied, using *sklearn* [68]. The clustering algorithm is applied to the positions in the logfile and 20 clusters are created. The number of the cluster is added in another column and applied to the table. The table is then grouped by the cluster number, with the utterances being accumulated. Since there are 20 clusters the table now has 20 rows, storing all utterances associated with one cluster, the cluster number and the centroid. The number of clusters was set to 20 because it prevents the scene from getting too crowded. However, this number can be changed as desired in the processing step.

Sklearn provides both a function that counts the amount of times a word appears in one cluster as well as a function that computes the weight of a word. The result of this clustering and weighting algorithm are saved as a .csv file called *tfidf.csv*, containing the cluster number, the position of the centroid and the respective keywords of the cluster sorted by their weights in descending order. The script for the clustering pipeline is attached in Appendix C.3.

Segmentations. The participants' delineations after the data acquisition are stored in .cso files. This type of file is only readable by MeVisLab, so to be able to use the data for the visualization in MOLE, they are converted into .dcm files. Before the files can be converted to DICOM format, the potential open curves of the segmentation task need to be closed. This process is done using MeVisLab. Figure 6.12 shows the original segmentation, the outcome after the first morphological operation and the result. The

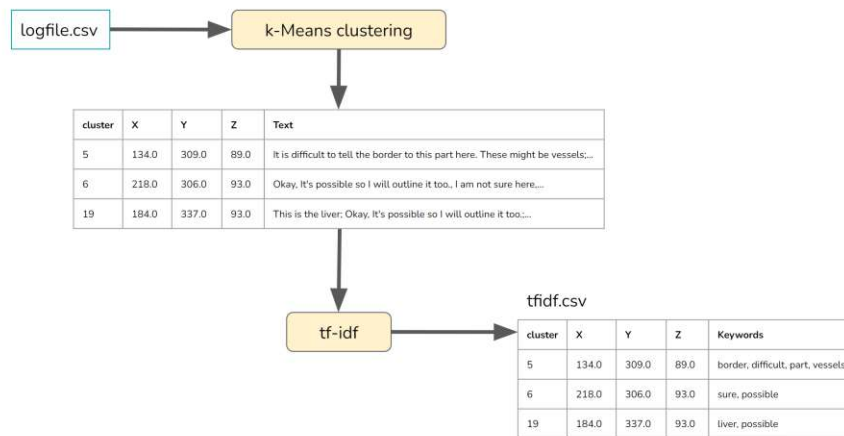


Figure 6.11: Pipeline for the creation of the tfidf file.

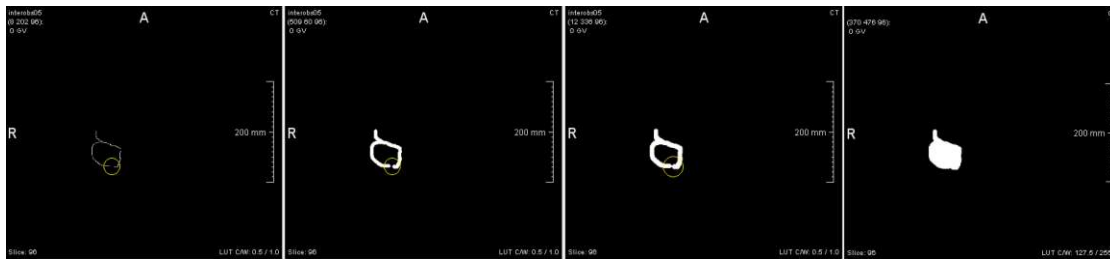


Figure 6.12: A segmentation with a gap (highlighted in yellow) being closed and its final binary mask.

filling of the closed contours is done using a script that uses *opencv*, a computer vision library for python [72]. This code is attached in C.2. With MeVisLab all the converted images can be imported into one DICOM volume, which in another step can then again be saved as single images, with the *SubImage* module of MeVisLab, if necessary.

6.2.3 Features

The features explained in the following sections are displayed individually and can be removed from the render window as such as well. Upon first call the data is stored in an object. When being rendered for the first time the feature is wrapped into a *vtkActor* or *vtkVolume* object which is then added to the renderer by using the function *AddActor()*. VTK does not offer a way to manage the actors and volumes in a scene and only stores them as a collection without the possibility to call a specific actor or volume. Therefore, a work-around is needed to dynamically remove actors and volumes. The work-around used, is a map-object storing the actors and volumes as pointers with a string as the key. When a specific object needs to be removed, the pointer value is retrieved with the key

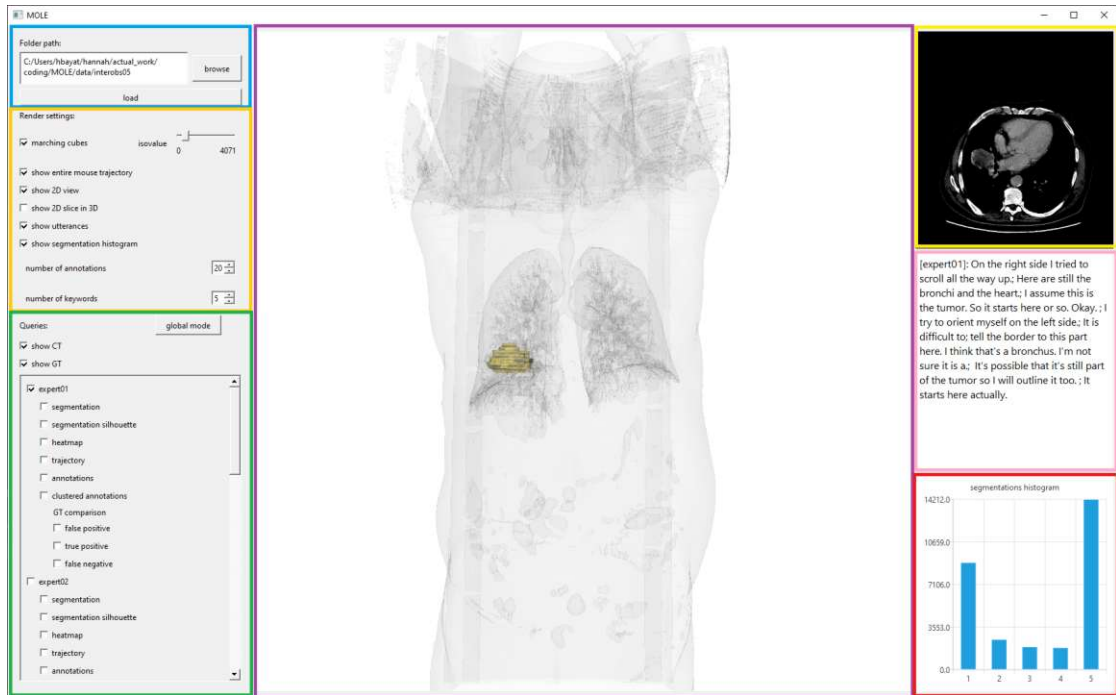


Figure 6.13: MOLE Interface Overview. Data information (blue). Render settings (orange). Query settings (green). Render window (purple). 2D View or Slice view (yellow). Utterances (pink). Segmentation histogram (red).

of the object that needs to be removed and can then be removed from the renderer using the *RemoveActor()*-function provided by VTK.

6.2.3.1 Overview First

After loading the provided data, the user can see an empty render window in the center of the application (Figure 6.13). On the left, MOLE displays the source path, different render settings and the possible queries for each individual observer. The right side of the application shows the slices of the CT scan, the utterances of the selected observers as well as a histogram displaying information about the combined segmentations.

The get an overview first following the requirement of applying Shneiderman’s mantra, we can display the ground truth and the CT scan. This can be done at all times, by selecting the checkbox “show GT” and “show CT” respectively, in the query section of MOLE. The ground truth is always displayed in “gold” and with an opacity of 0.3. The CT scan is displayed with 0.05 opacity and in grey for context visualization. The method with which the CT is being rendered can be changed in the render settings using the checkbox “marching cubes”. If the checkbox is selected, the CT will be rendered with the marching cubes algorithm [73] and the iso-value set in the slider to the right of the checkbox. The slider ranges from the minimum value in the loaded CT scan to



Figure 6.14: MOLE Interface in Global Mode with the switch button (green).

the maximum value. If the checkbox is not selected, the CT will be rendered through ray-casting [74]. The marching cubes algorithm for VTK proves to be faster and delivers a smoother interaction, than the ray tracing approach. Both of these functions are implemented by creating a *vtkImageData* with the DICOM files read from the data folder path given. The *vtkImageData* is then rendered with marching cubes, if not selected otherwise for the CT scan. These settings allow to display the target object within the scan, providing context visualization (**Q1** and **Q2**).

The button displayed in the query section on the left with the label “global mode” changes MOLE into the so-called “global mode”, which is seen in Figure 6.14. In this mode, data for individual observers is not available. Instead, the user can choose to see the global heatmap, which visualizes the accumulated times spent on each voxel of the CT scan by all observers collectively. This data is computed by accumulating all individual heatmaps. The values of each heatmap are added up, ultimately creating one global heatmap. The global annotations are the clustered annotations for all observers together. The data for the annotations in global mode is read from the *tfidf.csv* file located in the data folder path. This information is extracted from a document, accumulating all individual transcripts. When in global mode, the button changes the label to “individual mode”, which lets the user switch back to the initial view. With the global mode we summarize all observers’ attention areas and thoughts and answer question **Q5**.

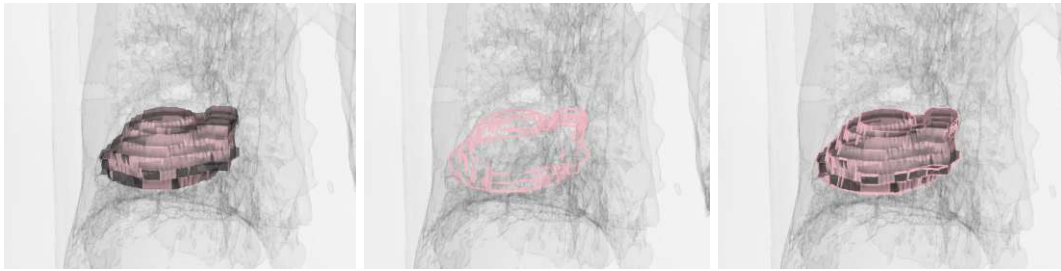


Figure 6.15: Segmentation iso-surface, Segmentation Silhouette and Segmentation iso-surface combined with its Silhouette.

6.2.3.2 Zoom & Filter

Next in Shneiderman’s mantra is to zoom and filter the data. Therefore, we list the possible visualization components in the query settings on the left side of the application. These components are available for each observation. Essentially this regards to the folders in the “segmentations”-path and encompasses the segmentation, its silhouette, the heatmap, visualizing the positions an observer spent time on and the trajectory of the mouse cursor, as well as annotations available on two different bases and a comparison of the segmentation compared with the ground truth, showing the false-positive, true-positives and false-negatives. All of the elements mentioned are displayed if their respective checkbox is selected. Internally, MOLE creates one “Observation”-object for each segmentation, where all the visualization components are stored in. Task **T1** calls for comparative visualization. All these components enable a multi-faceted comparison and allow for a detailed analysis of the segmentation process and results.

The segmentation is rendered and loaded in the same way as the ground truth and the CT: by reading the DICOM files from the given path. However, since the segmentations are only placed in one part of the image, only a small region is of interest for display. Therefore, the region of interest is extracted using *vtkExtractVOI*. This way only the important part of the image is rendered and empty volume does not need to be rendered, which increases the performance immensely, only needing 1.5 seconds to create the *vtkImageData* for the segmentation and 0.004 seconds to render. It is displayed in pink and an opacity of 0.5. These component is implemented to answer questions **Q1** and **Q2**. The silhouette of the segmentation is implemented using *vtkPolyDataSilhouette*, which takes the segmentation as a *vtkPolyData* and creates the silhouette, which is also rendered in pink and with an opacity of 1.0. Figure 6.15 shows the segmentation with and without the silhouette as well as the silhouette by itself. To answer question **Q3** we make use of this visualization.

The comparisons to the ground truth (false-positive, true-positives, false negatives) are also stored in *vtkImageData* objects and rendered with the marching cubes algorithm. However, they are first computed by using *vtkImageMathematics*. *VtkImageMathematics* allows pixel-wise multiplication of two images. We use this functionality to multiply the invert ground truth with the segmentation to acquire the false-positives of one

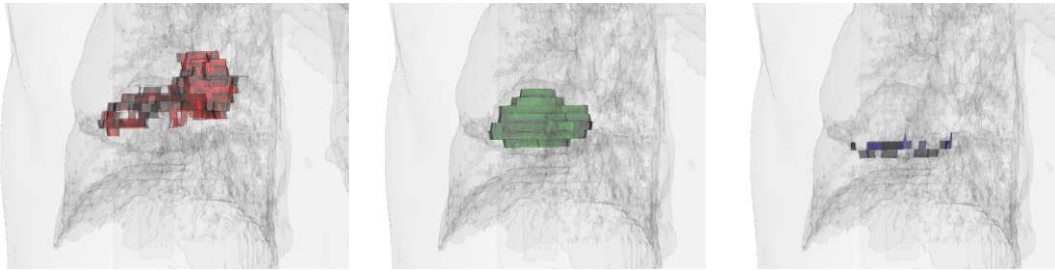


Figure 6.16: ground truth comparison display: false-positives (left), true-positives (middle), false-negatives (right).

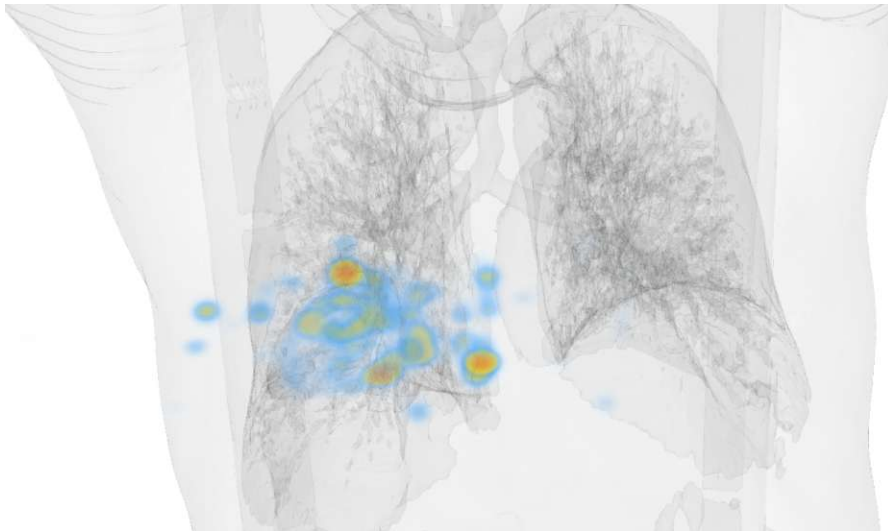


Figure 6.17: Heatmap for one observer.

segmentation. For the true-positive the ground truth and the segmentation are multiplied in their initial form and for the false-negatives the same logic is used as for the false-positives, however this time with the segmentation being inverted and multiplied with the initial ground truth. All the comparison volumes are displayed with an opacity of 0.3 and the colors shown in Figure 6.16, to address question **Q4**.

The heatmap function displays the content of the heatmap.csv file as described in Section 6.2.2.1. A *vtkImageData* object saves the accumulated time value per voxel and is then displayed using the defined colorscale. The opacity ranges from 0.0 for the minimum, which has the accumulated time value of 0 to 0.5 for the maximum accumulated time. The heatmap visualization is shown in Figure 6.17.

Adjustments can be made to how the trajectory of the mouse is displayed, by selecting or deselecting the checkbox “show entire mouse trajectory” in the render settings. This will then show the entire mouse trajectory or only the parts, where the observer pressed the left mouse key. This feature is implemented by creating a *vtkPolyLine* with all the

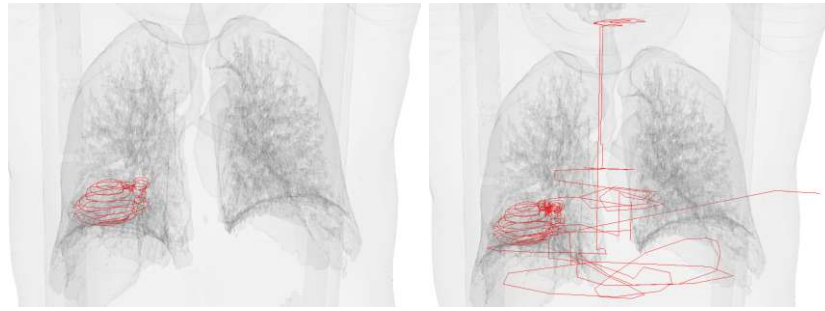


Figure 6.18: Mouse trajectory where the left key was clicked (left) and the whole mouse trajectory (right).

positions logged in the logfile. To display only the parts where the left mouse key was pressed, the column “Clicked” of the logfile is queried. If the value is “True” the point will be added to the polyline, otherwise it is discarded. An example of this feature with its adjustments can be seen in Figure 6.18. This should give a deeper insight into an observer’s approach (Q8). Mostly, the clicked mouse trajectory aligns with the resulting segmentation. However, in parts where the observer edited the segmentation the clicked trajectory will differ.

Further settings for the per-observer functions are the number of annotations. The counter labeled “number of annotations” sets the number of annotations displayed in the image. This correlates to the positions that the observer spent the most time on. For n annotations, n positions with the highest accumulated time spent on are shown. This setting only applies to the basic annotations, as they can be computed at run-time. However, for the clustered annotations always 20 are shown, because the clusters are not given a measure of importance or relevance when being created. This number can be changed in the data processing step. The clusters could be sorted by amount of time spent per cluster. For now, due to simplicity we decided to show 20 unsorted clusters. To declutter the image with the clustered annotations, the number of keywords can be set using the counter with the label “number of keywords”. This number will affect the number of keywords shown for each cluster. If the user is only interested in the top-weighted keyword of each cluster, they can set the value to one, so only one keyword will be displayed. For more keywords, the user shall increase the number. The keywords will be added by their weight — the further on the right the keyword is listed, the smaller is the tf-idf score of that keyword for the corresponding cluster. They can be seen in Figure 6.19.

6.2.3.3 Details on Demand

The 2D View, the utterances and the segmentation histogram on the right side of the overview window shown in Figure 6.13, are only accessible when set in the render settings. The 2D View shows the slices of the CT. By using the mouse and scrolling on the 2D View

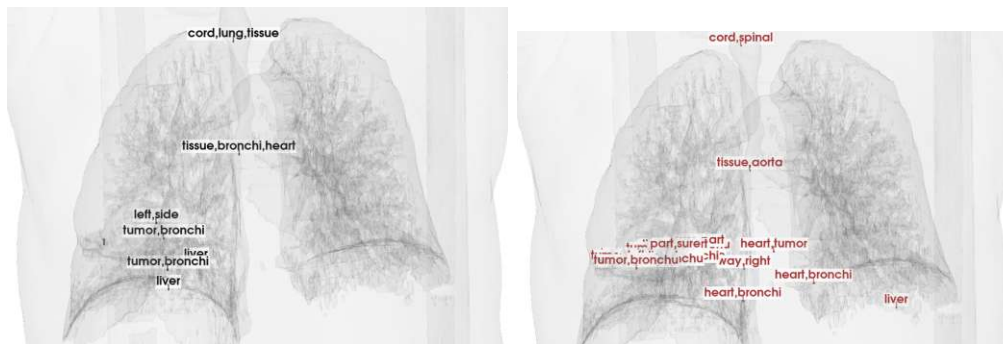


Figure 6.19: Basic annotations (left) and clustered annotations (right).

the user can slice through the CT scan. This feature is implemented as a *vtkImageViewer2*, reading from the folder containing the CT scan. A custom *vtkInteractionStyle* is created to implement the scrolling mechanism. This view is only visible when selecting the checkbox with the label “show 2D view” in the render settings. We keep these components hidden, because they only offer details and providing them at all times is not necessary. Selecting this checkbox enables two more checkboxes with the label “show 2D slice in 3D” and “show utterances”. If the former checkbox is selected, a cuboid with the height of one voxel is displayed in the render window, visualizing the slice that is currently being displayed in the 2D View. By hovering over the 2D View the tooltip of the mouse indicates the number of the slice shown. This feature allows a connection between the slice that observers were seeing when segmenting and the outcome in the render window. By selecting the checkbox labeled “show utterances” a text box underneath the 2D View is shown. These features are shown in Figure 6.20. In order to understand reasoning and causes for the segmentation (**T4**), it is helpful to trace back the outcomes to the origin, which is why we offer the 2D view. The utterances are provided, so the keywords can be put into the full context of the thought process.

The segmentation histogram is a feature added, to see how the segmentations are distributed. It helps to answer question **Q5**, to find outliers and common ground among observers. It is displayed when the checkbox “show segmentation histogram” in the render settings is selected. For example as seen in Figure 6.21 the last bar of the histogram is selected (colored in red). This selection displays all the voxels, where all observers agree.

6.2.3.4 Colors

The colors and opacities used for the visualized objects are listed in Table 6.1. The CT scan is displayed in grey and an opacity of 0.05, since its purpose is to give context and thus should not distract from the other objects in the scene. The ground truth is visualized in gold as the gold standard for the segmentation. The colors for the heatmap are chosen such that positions that have been visited but only very little are colored in shades of blue, which is commonly associated with cold and less interaction. On the other side of the spectrum positions that are visited a lot, are colored in shades of red,

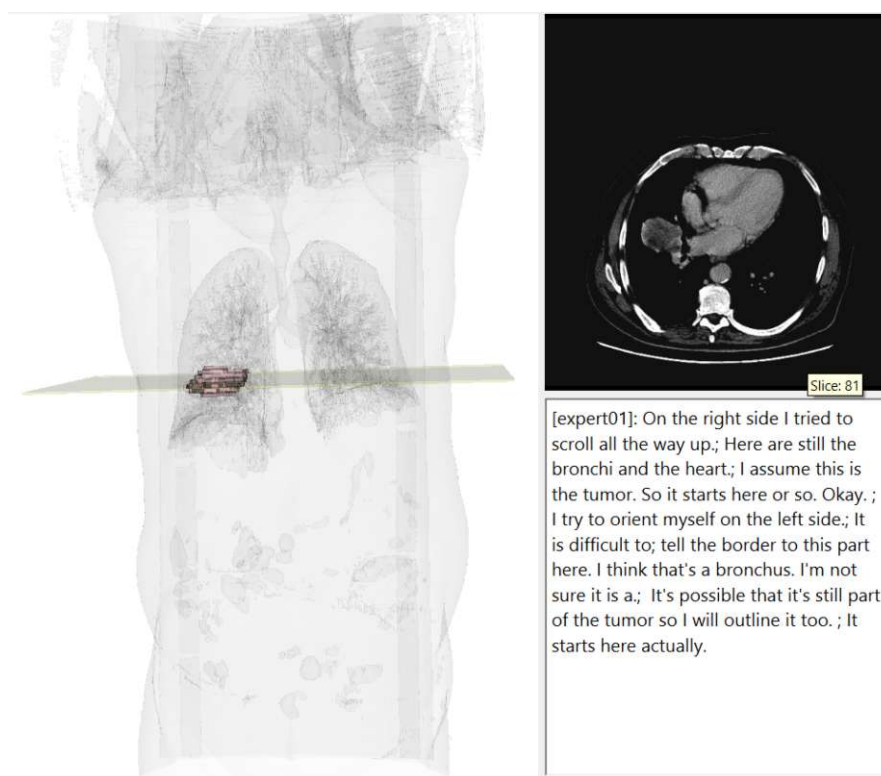


Figure 6.20: Hidden Functions: 2D View, 2D slice in 3D and Utterances.

which is usually associated with heat, visualizing the importance of those positions. The annotations, both clustered and basic are displayed with an opacity of 0.5 and a white background for better readability. To be able to distinguish between the annotations, the basic annotations are rendered in black and the clustered annotations are rendered in brown.

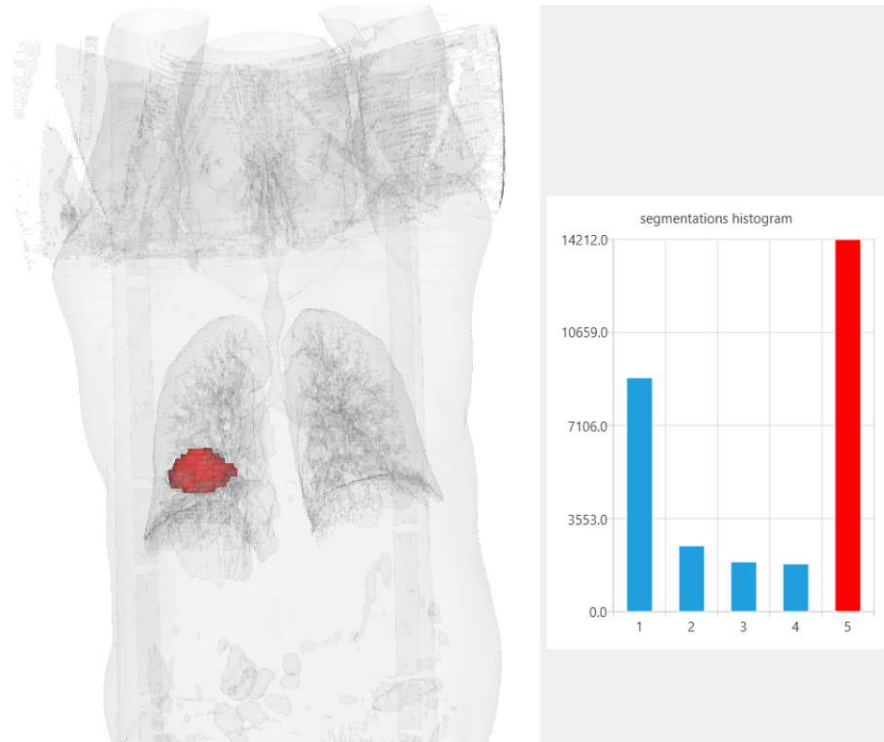


Figure 6.21: Segmentation Histogram with selected bar and its respective display in the render window.

Table 6.1: Visualization components displayed in MOLE, their respective color and opacity.

display object (actor)	color	opacity
ground truth (GT)	Gold	0.3
segmentation	Pink	0.5
segmentation silhouette	Pink	1.0
heatmap	colorscale Figure6.3	0.0 – 0.5
trajectory	Red	1.0
annotations	Black	0.5
clustered annotations	Brown	0.5
false positives	Red	0.3
true positives	Green	0.3
false negatives	Blue	0.3



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Validation

This chapter focuses on the applicable use cases and potential utilization of MOLE. We describe the possible usage scenarios and evaluate the tool with the help of Wouter van Elmpt from the Maastricht University Medical Centre [75]. By guiding our advisor through MOLE and showing him relevant examples with the medical data, we validate the visualization components and the overall usability. This process is described in Section 7.1. In Section 7.2 we summarize the potential application into the current workflow of radiotherapy as evaluated by our domain expert.

7.1 Usage Scenario

By executing a qualitative walkthrough with our application MOLE we, simulate a usage scenario. Using both the medical dataset and the gun dataset as described in Section 5.3.1, we present the applicability of our tool regardless of the segmented target object and input scan. The following sections demonstrate how MOLE is used to gain insight into the segmentation process. We evaluate if our MOLE answers the questions defined in Section 4.1. Upon starting MOLE, the interface displays an empty rendering window on the right, next to a text box with a button saying “browse” and a button underneath labeled “load”. This view is seen in Figure 7.1. The browse-button opens a file browser for the user to select the folder, containing the data they want to analyze. By clicking the load-button, the data is read into the system and the view changes to what is seen in Figure 6.13.

7.1.1 Expert Data

The target object in the expert data is a tumor in the patient’s right lung. This dataset includes the segmentation result of five experts, as discussed in Chapter 5. We start out by looking at the scanned object, which in this dataset is the CT scan of a male patient

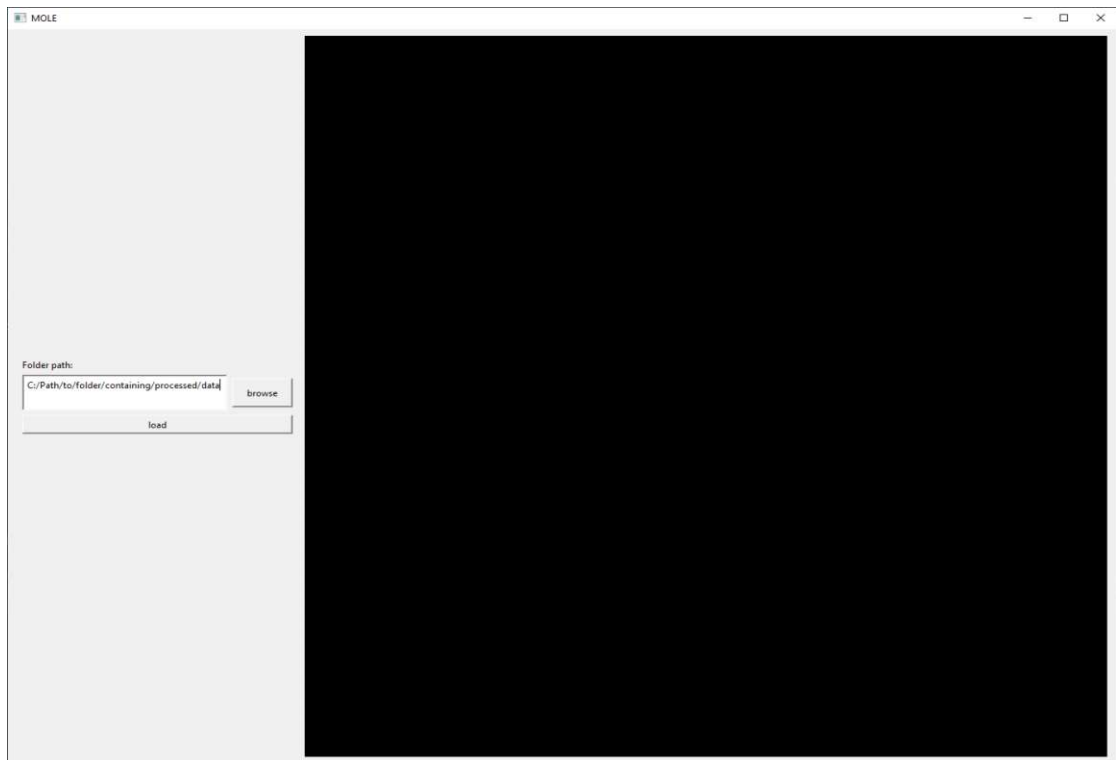


Figure 7.1: Initial MOLE Interface.

ranging from the hips to the head. We add the determined ground truth to the scene to get a feeling of what the scan and target object look like in the three-dimensional space (**Q1**). Figure 7.2 shows this view. The object that was segmented by the observers, the target object, is the tumor. As a user in this example looking at a tumor, one should always be aware that there is no definitive ground truth and the defined gold standard can be loosely considered as another segmentation outcome.

To see how accurate the segmentations are compared to the ground truth, we analyze the false-positives, true-positives and false-negatives of all users (**Q4**). Figure 7.3 shows all segmented voxels that are labeled as true-positives. True-positives are considered to have been segmented correctly. It is clear that a majority of the ground truth volume was segmented. However, the observers identified the lower part of the tumor more often than the upper part which is visible due to the lower opacity in the bottom part.

Comparing the false-positives with the ground truth (**Q2**) emphasizes on the fact that the bottom part of the tumor is consistently missed, as seen in Figure 7.4. In addition the marked section in Figure 7.4 shows that a small part medial to the aorta is missed as well.

To see where observers tend to over-segment, so segment voxels that are not considered part of the ground truth, we use the false-positives (**Q4**). This view, as seen in Figure 7.5,

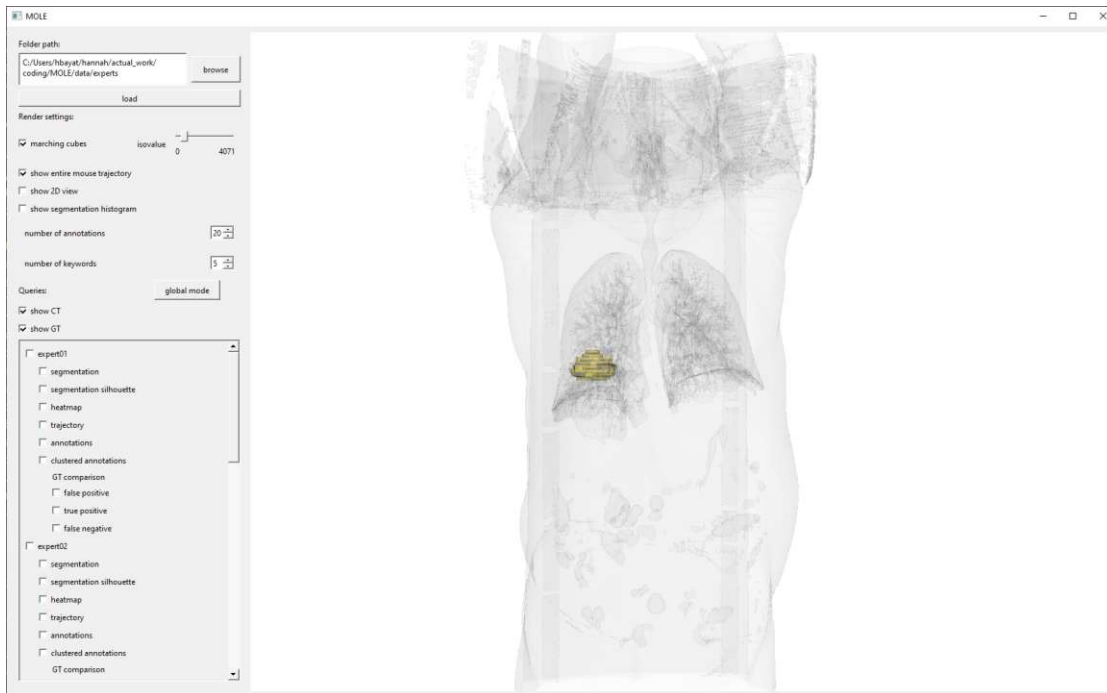


Figure 7.2: Scan of the patient (lightgrey) and the tumor as the target object in the right lung (gold) for the expert task.

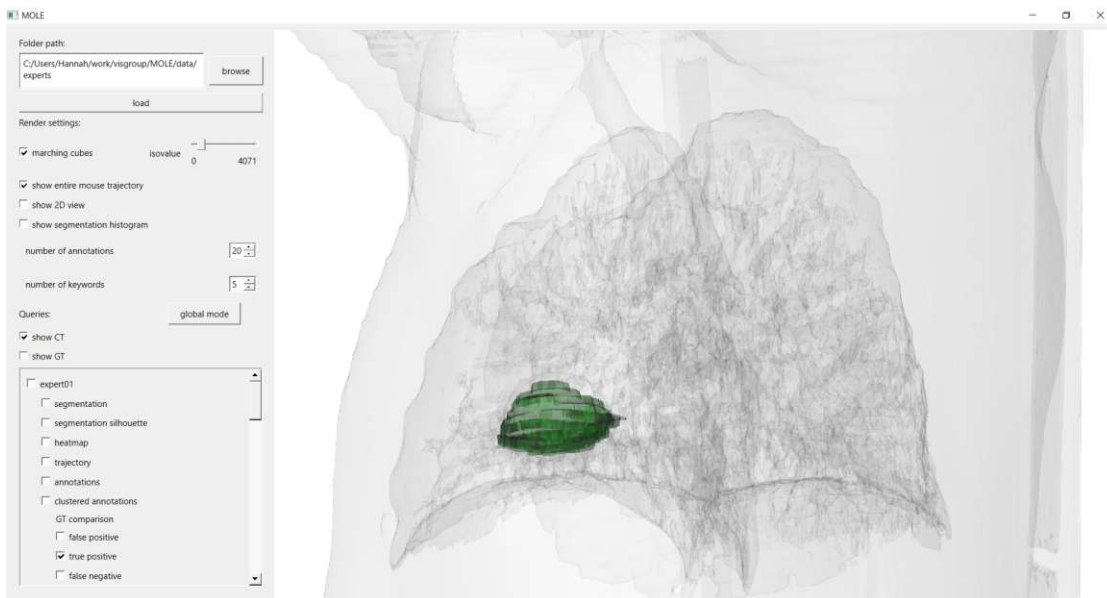


Figure 7.3: The accumulated true-positive voxels of the expert observations.

7. VALIDATION

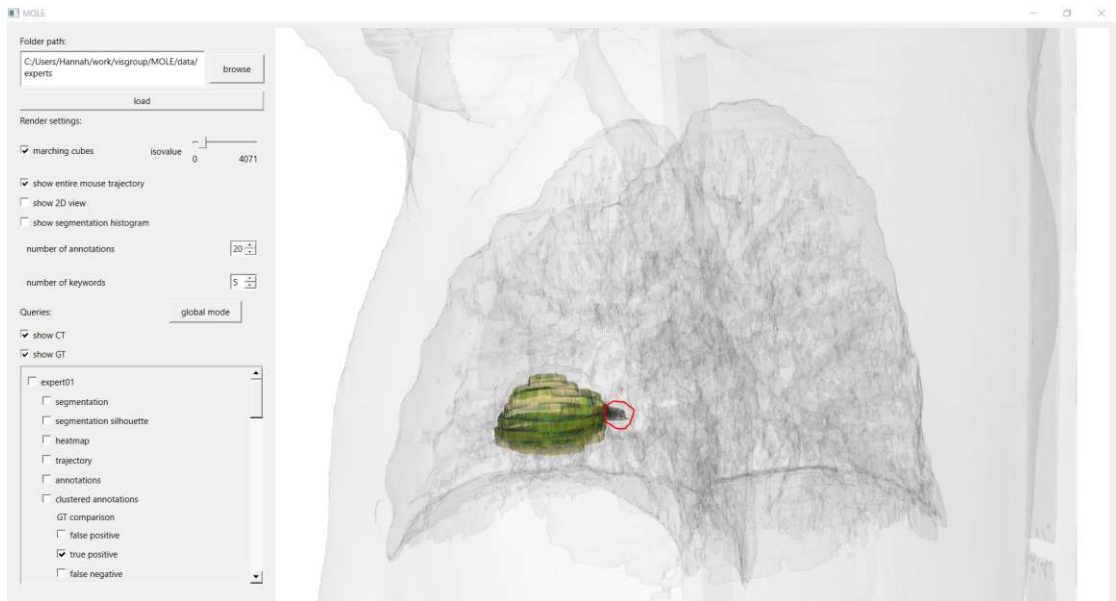


Figure 7.4: The ground truth of the tumor laid over on the accumulated true-positive voxels of the expert observations. Circled in red: the part medial to the aorta.

shows that some observers mistakenly identified a part superior to the tumor. Additionally, we see that observers generally tend to over-segment in the top part of the tumor, which is indicated by the clustered red-colored voxels. The voxels that were missed by all expert observers are shown using the false-negatives. Once again, it is visible that the inferior part of the tumor was constantly missed. This phenomenon is thus worth investigating further. Figure 7.6 shows this supposed error.

To investigate the aforementioned phenomenon, we can use the 2D-View. This view allows us to see the slice, the observers were looking at, when segmenting the tumor. We set the slice in the 2D-View which places the cuboid in the according slice that was missed by all participants. We are able to see the image that was segmented on and the according area in 3D. As shown in Figure 7.7 the image that was segmented on, still shows a bit of possible tumor tissue. Thus the reason for not segmenting this part of the tumor did not lie in the input image. Another possibility to explore reasons for the missed slice, is to look at the observers' thought processes (Q9).

In a next step combining the information gained from a single segmentation outcomes with the ground truth and the according annotations might give more insight (Q3, Q7). This combination of the visualization components is shown in Figure 7.8 for all expert observations. Expert 1 talked about “bronchi” and the “tumor”, so they might have been confused whether this tissue was part of the bronchi or part of the tumor. Expert 2 mentions the heart, the lung, a vein, and an artery. The basic annotations of Expert 3 and Expert 4 do not deliver any contextual information but rather adjectives like “bad” and “big” that suggest struggling on this slice. Expert 5, who over-segmented largely

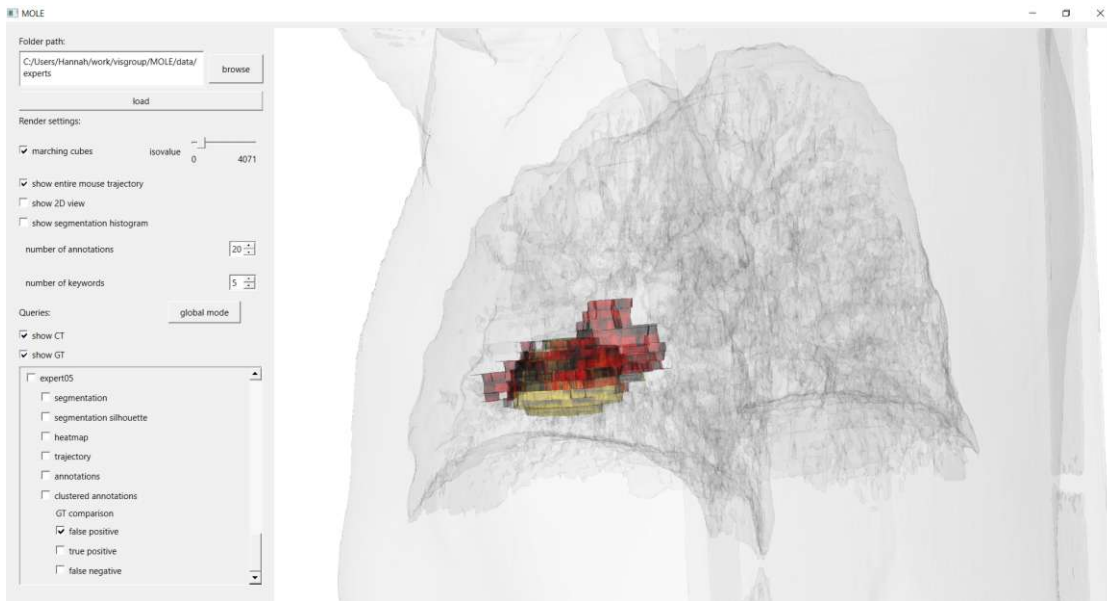


Figure 7.5: The accumulated false-positives of all expert observers placed in comparison with the ground truth.

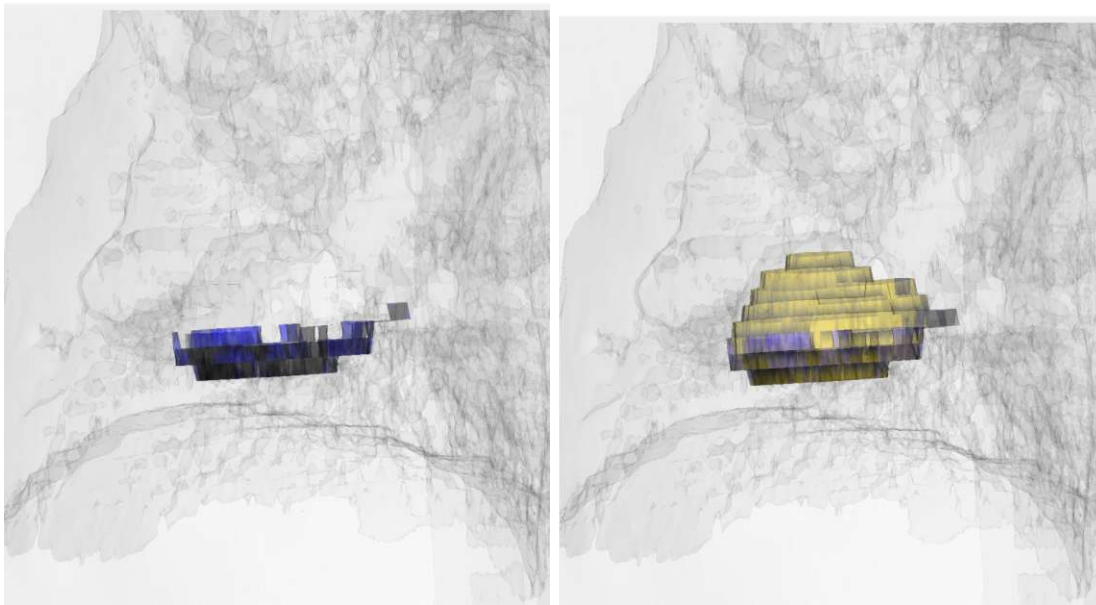


Figure 7.6: The accumulated false-negatives of all expert observers (left). False-negatives placed in comparison with the ground truth (right).

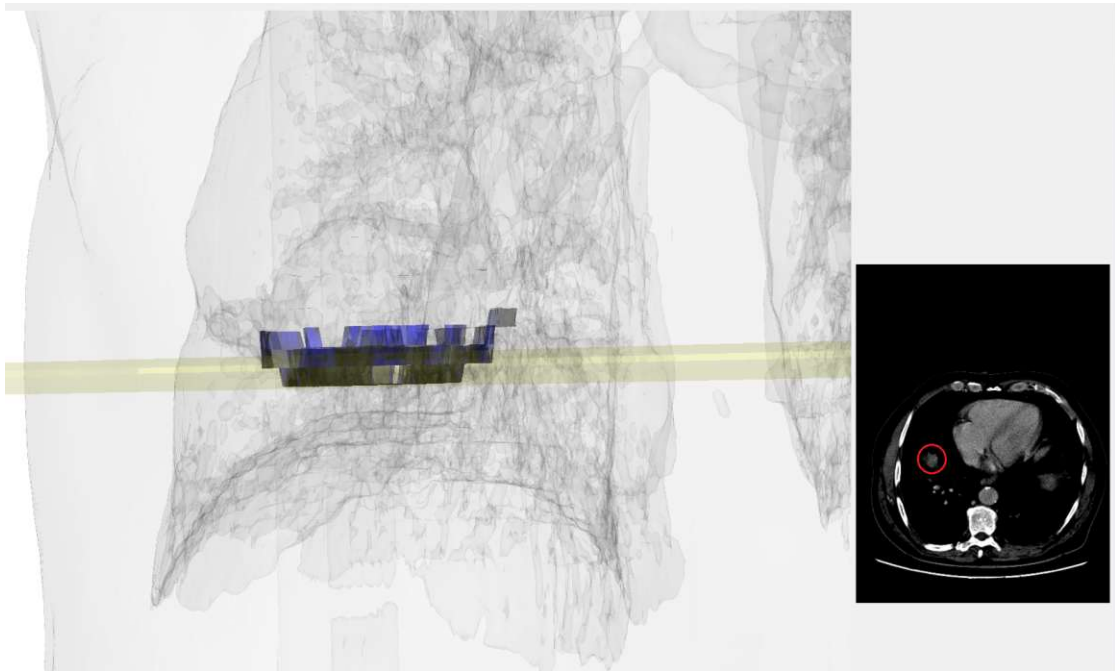


Figure 7.7: Analysis of false-negatives using the Slice View (located on the right). Circled in red: possible tumor tissue.

on the top part of the tumor, has no annotations on the slice in question, which might originate from their focus on the upper part of the tumor.

For Expert 5 we can see some more details using the clustered annotations (Q7). As shown in Figure 7.9, Expert 5 was using the words “tissue”, “normal”, “liver” and “vessel” in the surrounding area of the cluster that has its center in the missed slice. The words are displaying in descending order according to their importance, which is computed as described in Section 6.1. Since, the word “tissue” has the highest importance, there is an indication that the tissue on the slice, as shown in Figure 7.7 was talked about and possibly labeled as “normal”, a part of the “liver” or a “vessel”.

To understand where Expert 3 and 4 were struggling, we use the visualization of their mouse trajectory, which will give us an overview on their approach to the segmentation task (Q8). Figure 7.10 shows how the mouse trajectory of Expert 3 and Expert 4 differ. While Expert 4 shows a clear trajectory that is mainly focused around the height of the tumor, Expert 3 spent time around the tumor, but also looked at other organs and areas of the body. We can assume that locating the tumor proved to be a more difficult task for Expert 3 than for Expert 4. The confidence survey conducted during the data acquisition, as described in Section 5.4.3, supports this hypothesis. The results for the experts are summarized in Table 7.1. Expert 3 said they disagreed with statement 6, whereas Expert 5 was neutral towards it.

Using the heatmap of Expert 4 we see that they spent a lot of time in the bottom of

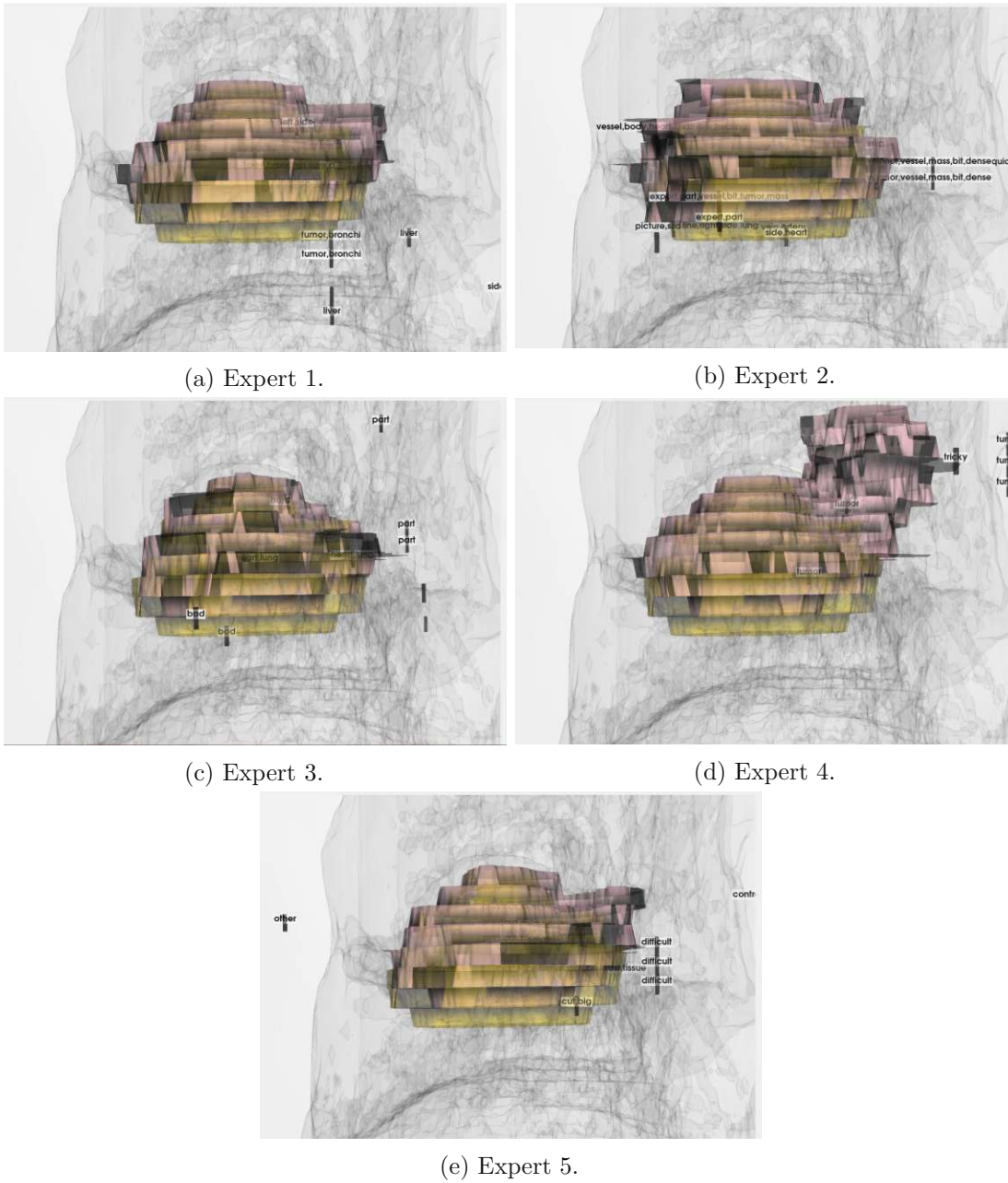


Figure 7.8: Segmentation outcome in combination with the according basic annotations for each expert observer.

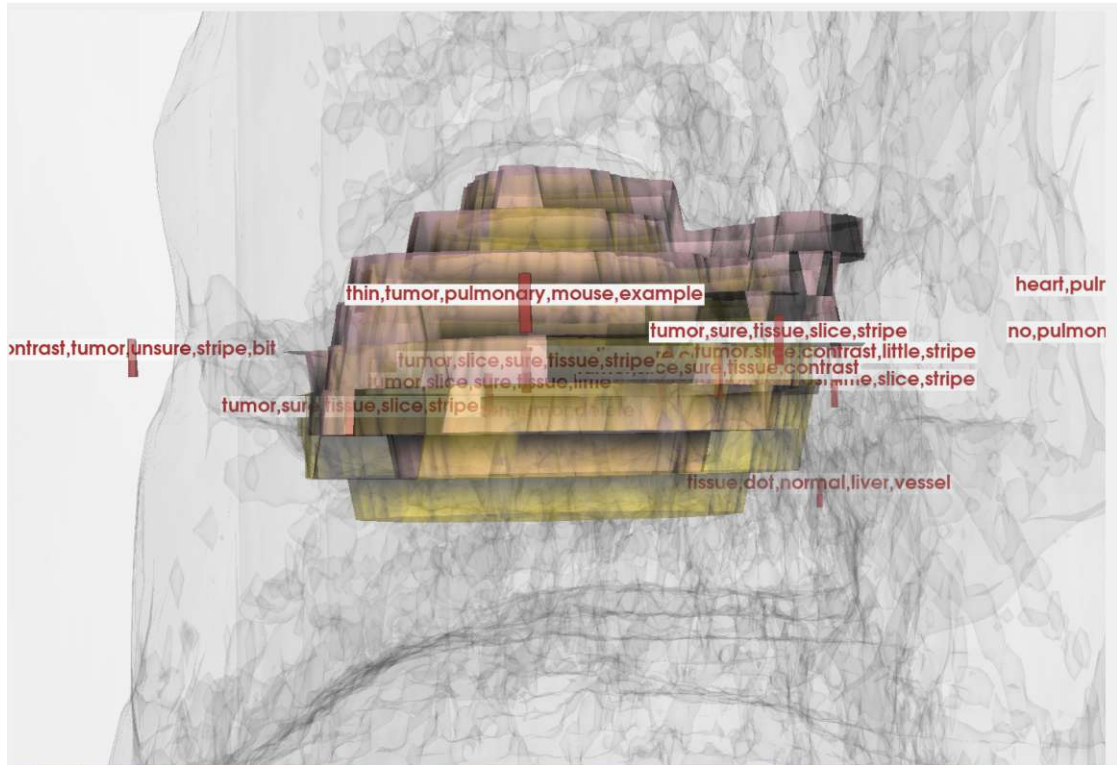


Figure 7.9: Clustered Annotations for Expert 5.

Table 7.1: Confidence survey results for expert participants for the questions as listed in 5.4.3.

	assignment understood	finding was easy	borders clear	hindered by objects	confident in decisions	easy to solve
Expert 1	str. agree	str. agree	agree	neutral	str. agree	str. agree
Expert 2	str. agree	str. agree	agree	str. agree	str. agree	str. agree
Expert 3	agree	agree	neutral	agree	disagree	disagree
Expert 4	agree	agree	neutral	disagree	neutral	neutral
Expert 5	agree	neutral	neutral	agree	neutral	agree

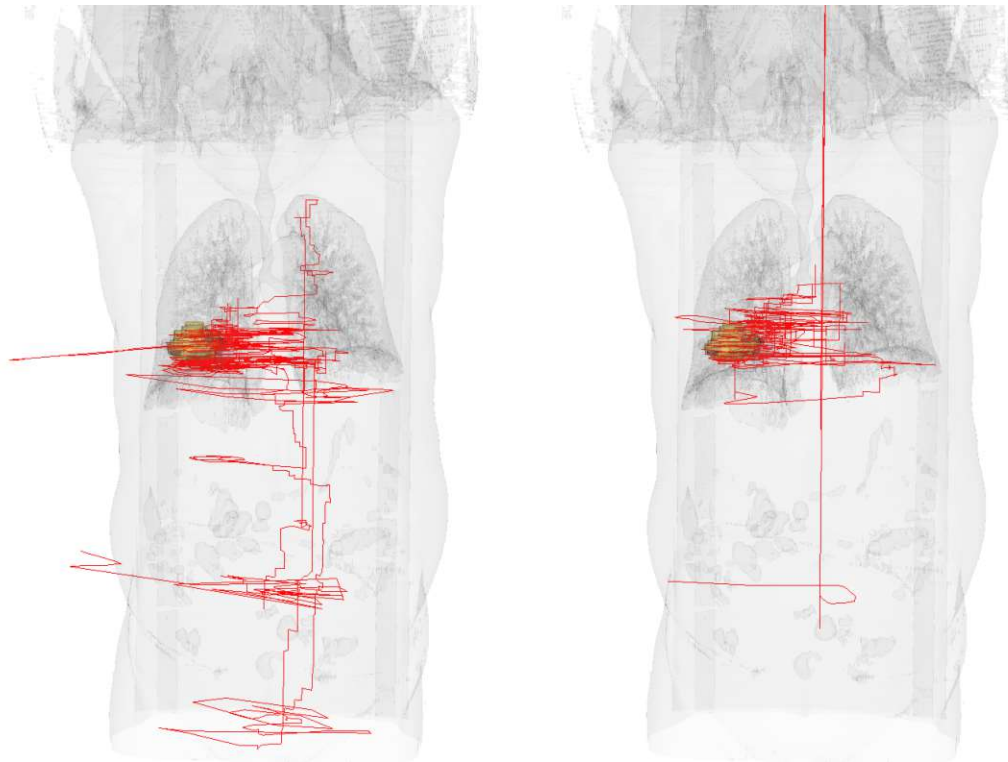


Figure 7.10: Mouse Trajectories for Expert 3 (left) and Expert 4 (right).

the tumor but ended up not segmenting it, as this part is not within the silhouette of the segmentation. This is shown in Figure 7.11. According to the heatmap, the observer spent a lot of time on the top right part of the tumor as well as the bottom part (**Q6**). Comparing these areas with other segmentations, we can observe a high variability and thus assume these areas were unclear when delineating. There is no apparent correlation between the amount of time spent and the area being segmented or not for Expert 4. However, we can determine that the bottom of the tumor, the right and upper right areas are unclear to Expert 4.

We have a further look into the global heatmap to see where all observers combined spent the most time (**Q6**). As seen in Figure 7.12 the most time was spent in the bottom of the tumor, showing that Expert 4 was no exception. As shown before however, the slice where the global heatmap shows the most activity was missed by almost all observers. All observers tended to over-segment, as displayed in Figure 7.13. The segmentation histogram is selected to show all those voxels that were segmented by every observer. In comparison to the ground truth we can see a clear over-segmentation on the top part of the tumor, because the gold standard is visible in the bottom whereas in the top part of the tumor it is fully covered by the segmentations. Since, the gold standard for the tumor is solely another segmentation outcome, the agreement of all five experts can be considered a strong argument in favor of part of the tumor (**Q5**).

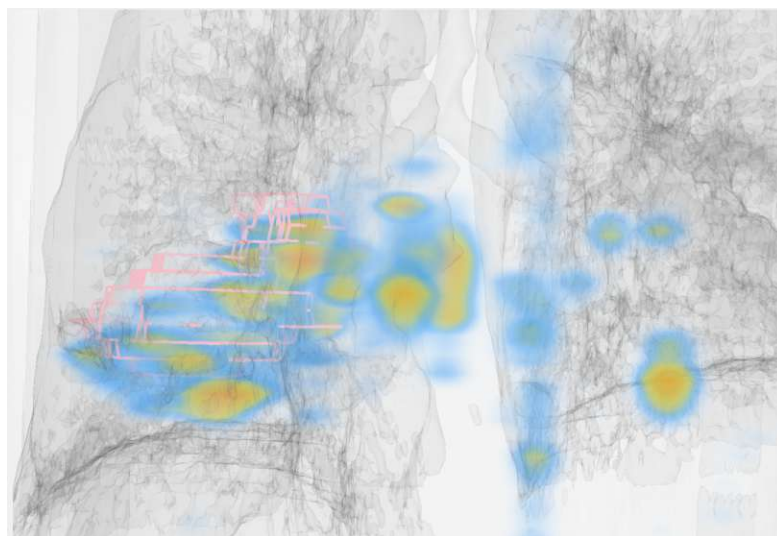


Figure 7.11: Heatmap and segmentation silhouette of Expert 4.

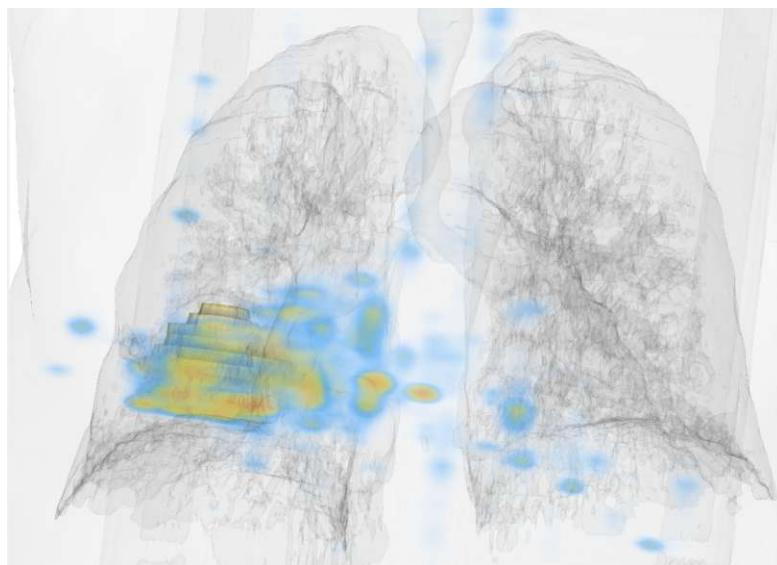


Figure 7.12: Global heatmap of expert observers.

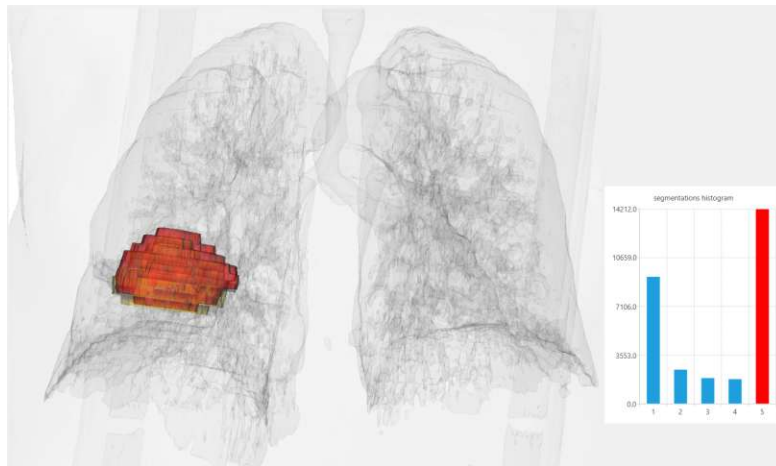


Figure 7.13: Agreement in the segmentation outcomes of all five experts.

Continuing to use the segmentation histogram, we look at the outliers of all the segmentations. These are all those voxels, segmented by only one observer. To gain some more detailed insight into the observers' thoughts we look at the utterances. Figure 7.14 shows the outlier voxels colored in red. In addition the utterances of slice 76 are displayed. According to the utterance text, Expert 1 did not mention a tumor on this slice. Expert 2 talked about the heart, like Expert 1 and shows no uncertainty about the existence of a tumor in this slice. Expert 3 said "Maybe this is part of it", expressing uncertainty. Expert 4 mentions that this part is "tricky" and mentions that they "think that's not tumor" but contradicts themselves at a later point by saying "This has to be tumor, you can see this is tumor too here". While Expert 5 says "This could be something" they decide that the tissue in question resembles pulmonary hiluses based on the connections to the heart, using their knowledge about the human anatomy to make decisions.

7.1.2 Novice Data

The target object in the novice data is a gun, specifically a revolver hidden in a backpack. Due to simplicity the revolver will be referred to as a gun. This dataset includes the segmentation results of four different observers. In order to talk about the target object of this dataset, Figure 7.15 shows the basic jargon used for the different parts of a handgun.

The dataset with its target object, the gun, is shown in Figure 7.16. The novice dataset consists of 372 slices. There were more slices to segment than in the experts' dataset, resulting in smoother segmentations. In addition, the ground truth displayed in 7.16 is definitive, contrary to the example described in the previous Section (Q1).

To see where the novices spent the most time segmenting (Q6), we look at the global heatmap of the dataset. As displayed in Figure 7.17, there was an accumulation around the trigger as well as the upper part of the cylinder. The time spent at the bottom of the

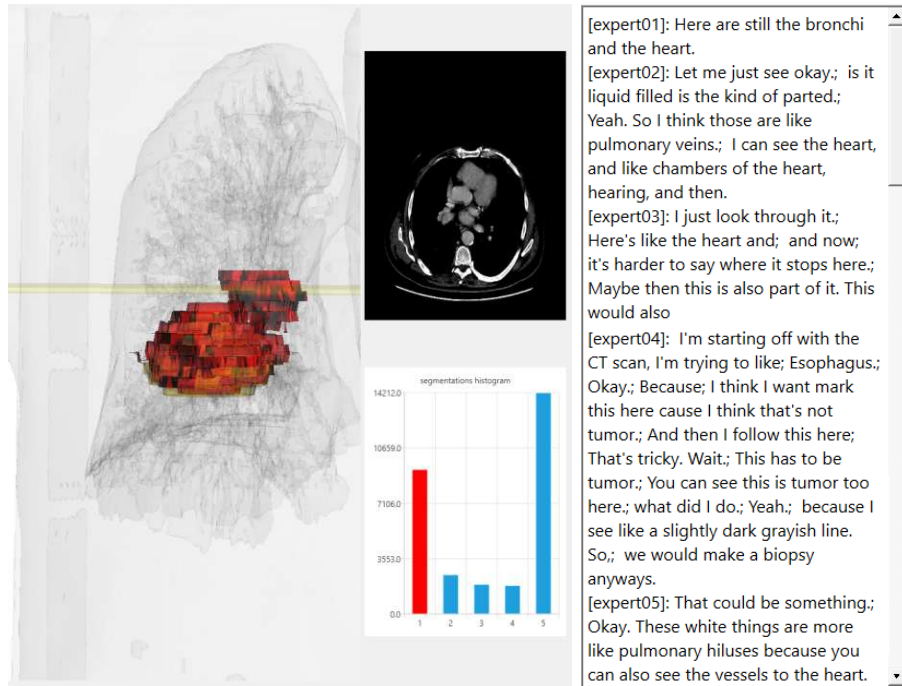


Figure 7.14: Outliers in the segmentation outcomes (left) and the respective expert utterances of slice 76 (right).

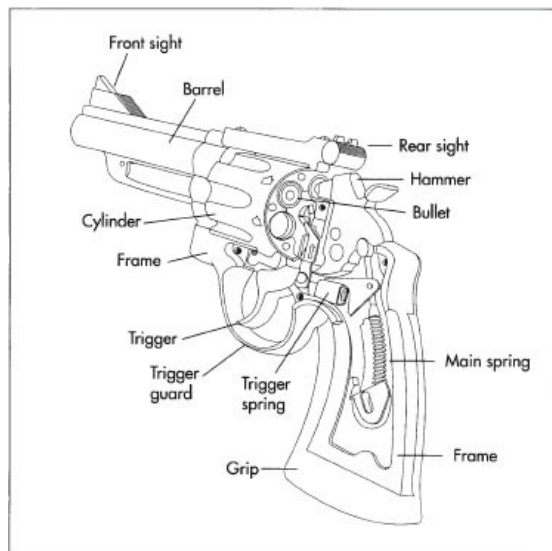


Figure 7.15: Parts of a handgun (revolver) [76].

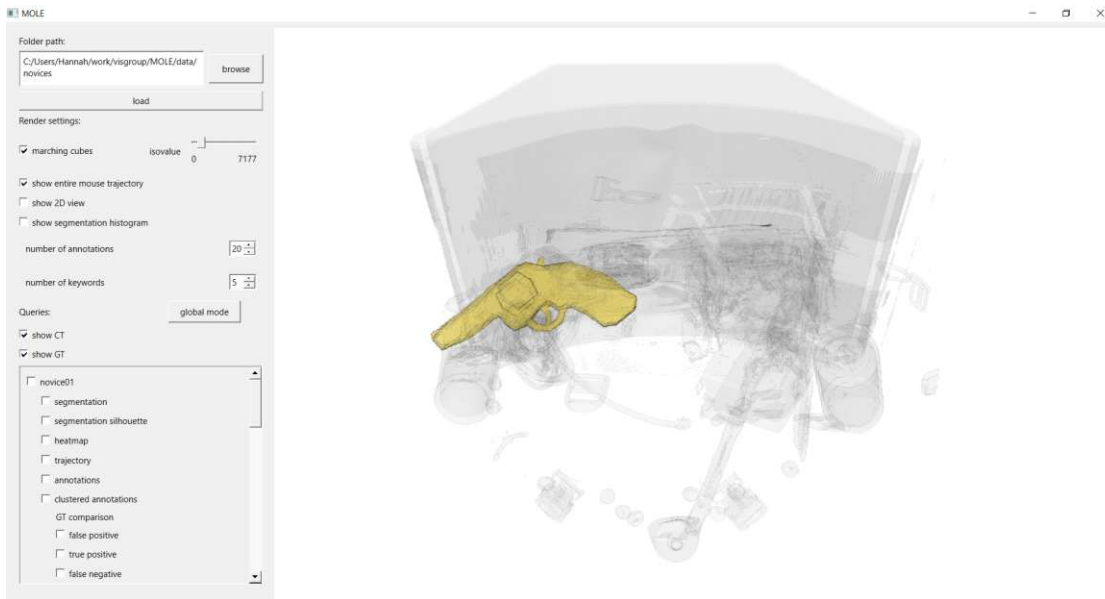


Figure 7.16: Scan of the backpack (lightgrey) and the gun as the target object (gold) for the novice task.

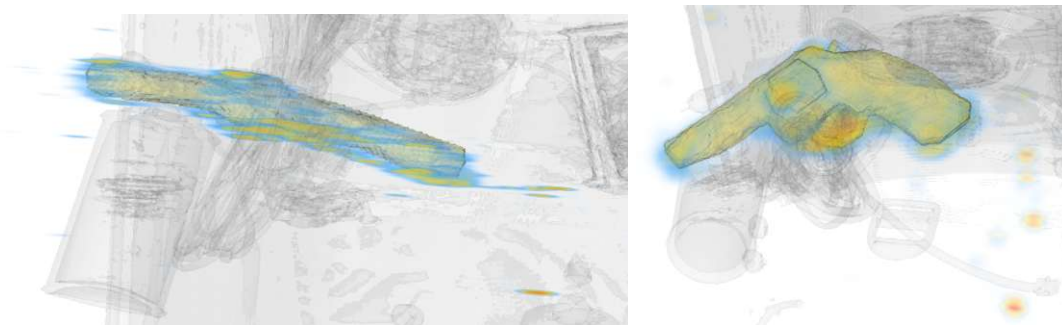


Figure 7.17: Global heatmap for novice observers from different perspectives.

gun on the left side in Figure 7.17 can be explained with the appearance of gun parts. Scrolling through the slices, these are the points, where the part is showing up at first, so the exact slice where the gun is visible needs to be specified. Determining where the gun begins, requires the observer to examine the surroundings, thus making them spend more time in those areas.

The trigger presented as a source of uncertainty as well. We look at the global clustered annotations, as shown in Figure 7.18. The trigger appears in multiple clusters with a high importance. As the word “trigger” is present in the majority of the clusters, we can assume that the observers talked about the trigger often (**Q7**). Therefore, the observers put a lot of importance to the trigger. This might be due to the fact that it was difficult to segment, as it requires to also determine the trigger hole.

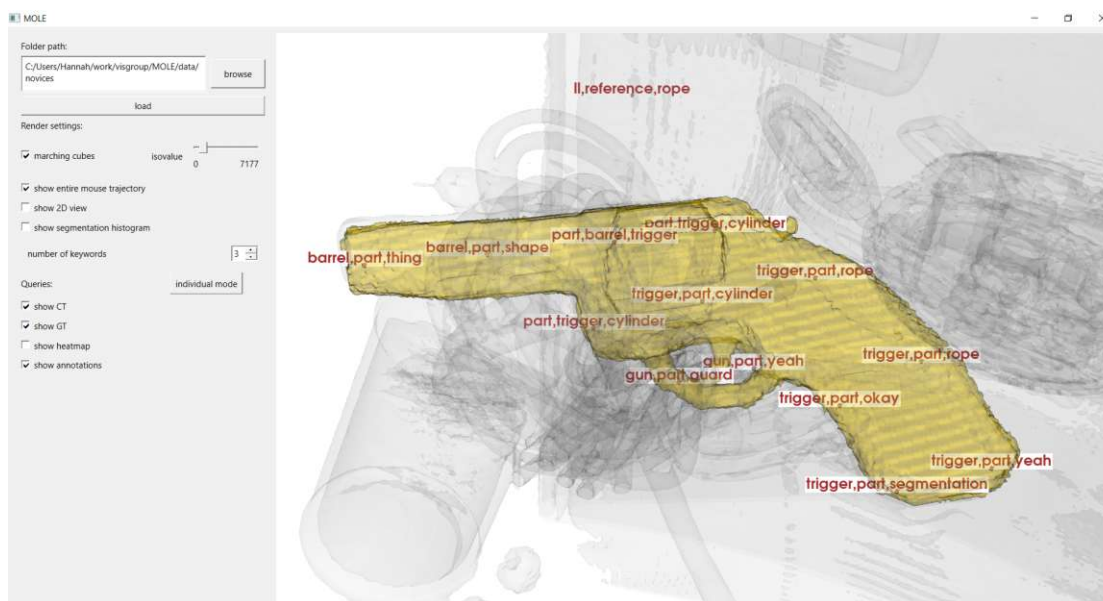


Figure 7.18: Global clustered annotations for novice observers. The clusters display the three most important keywords.

Figure 7.19 shows that the observers over-segmented (**Q4**). Displaying the accumulated false-positive marked voxels we see that the shape of the gun is still visible in all dimensions. However, it is highly over-segmented. In addition, there are some outliers on the bottom part of the grip and the bottom part of the barrel, which stem from single observers and do not match the majority, as seen in Figure 7.29.

Looking at the individual segmentations, displayed in Figure 7.20, it is clearly visible that Novice 4 segmented the closest to the ground truth and over-segmented the least (**Q3**). As mentioned before, the outliers in the false-positives shown in Figure 7.19 stem from Novice 2, who over-segmented in the top part of the front part of the barrel and the grip. Subfigure 7.20a shows that Novice 1 missed the hammer of the gun, which we will investigate on further.

Since Novice 1 missed the hammer, the next step is to see what the reasons for not segmenting this part of the gun were. We display the false-negatives, to see the exact voxels that were missed. As shown in Figure 7.21 Novice 1 also missed a part in the top of the cylinder (circled in purple). By adding the mouse trajectory of Novice 1 it becomes clear that the observer passed the upper part of the cylinder but did not segment it (**Q8**). The observer did not, however, pass the hammer at all. This might be due to the fact that the observer did not expect this part of the gun and did not account for it, since they also did not move near it with their mouse.

The false-negative visualization of Novice 4 shows that even though they segmented the gun most accurately, they missed one slice. As shown in Figure 7.22 the false-negatives of Novice 4 include a small part in the bottom of the trigger guard and the frame, but

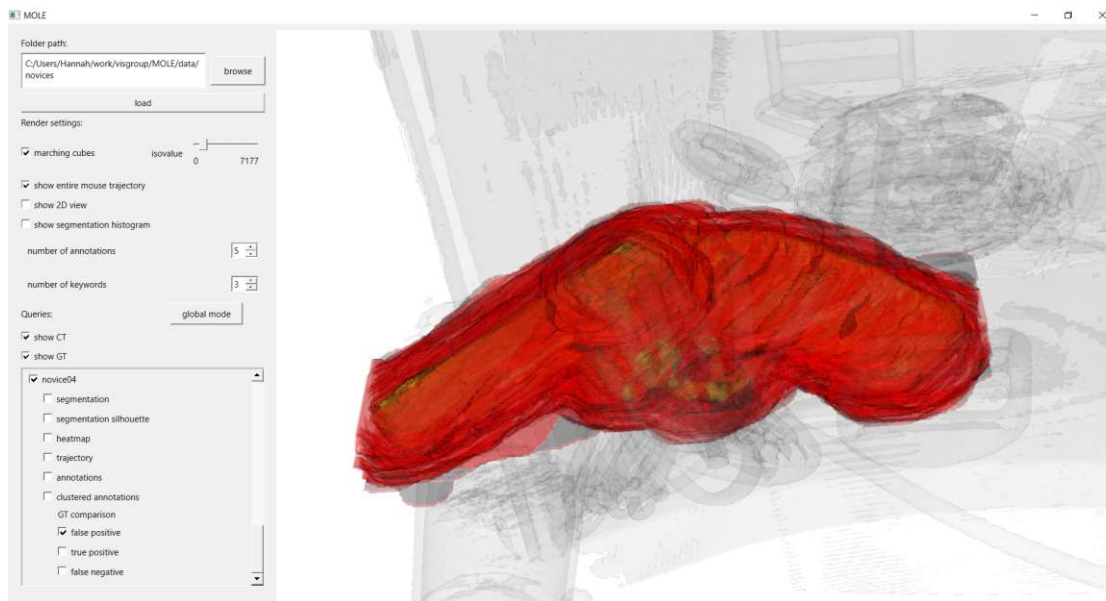


Figure 7.19: Accumulated false-positive voxels for all four novice observers.

also a whole slice, which we will examine in more detail moving on.

As previously shown, we see a clear over-segmentation below the trigger guard. To understand the observers' thoughts better, we use the clustered annotations for Novices 2 and 3 and gain more insight into the process (Q8, Q9). The trigger and trigger guard proved to be quite challenging. Figure 7.23 shows that Novice 2 besides the words "trigger" and "part" was also talking about a "rope" in the area of the trigger and its guard. While Novice 3 labels this part of the gun as "weird". During the segmentation process Novice 2 would refer to the object hindering them in segmenting the trigger as a "rope" (Q2). However, the same object would be called "cables" during the segmentation process of Novice 3. The same object being called differently by multiple observers, can lead to another source of uncertainty. This phenomenon however is unlikely to happen in the dataset presented in Section 7.1.1 since medical scans usually display commonly known objects that are universally labelled. Visualizing the spots Novice 2 spent the most time on, it becomes clear that for them the trigger was the most difficult part as well. By using the basic annotations of Novice 2 in combination with the heatmap we can also see that they were clearly confused with the rope in the area of the trigger. The keywords in this area are "rope" and "confusing", as shown in Figure 7.24. This conclusion is also reflected in the conducted confidence survey. As listed in Table 7.2, Novice 2 agrees strongly with the statement that other objects hindered them in the segmentation task (Q2).

Looking more closely into the thought process of Novice 3 to see where they spent the most time on and what they might have been struggling with, we use the heatmap and the slice view (Q6). By navigating to the slice where the heatmap has higher values, as

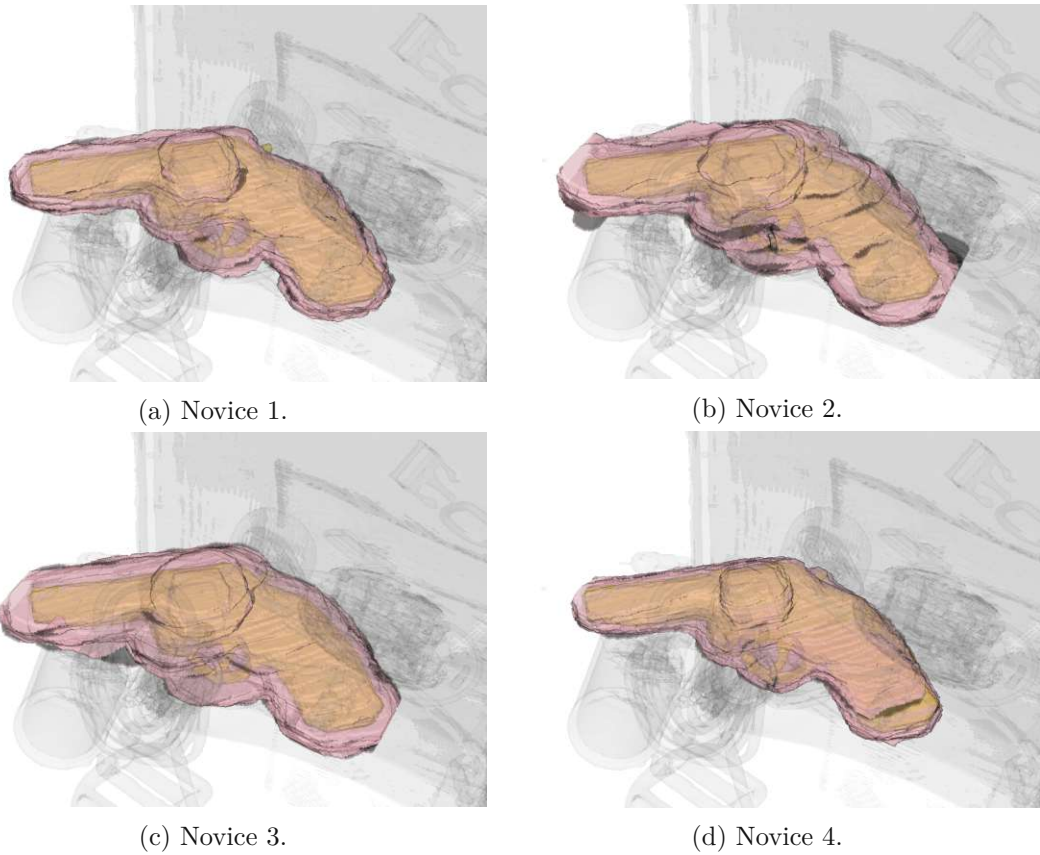


Figure 7.20: Segmentation outcome compared to the ground truth for each novice observer.

Table 7.2: Confidence survey results for novice participants for the questions as listed in 5.4.3.

	assignment understood	finding was easy	borders clear	hindered by objects	confident in decisions	easy to solve
Novice 1	str. agree	agree	disagree	agree	agree	agree
Novice 2	str. agree	str. agree	agree	str. agree	agree	neutral
Novice 3	str. agree	agree	neutral	neutral	disagree	neutral
Novice 4	str. agree	str. agree	agree	neutral	agree	agree

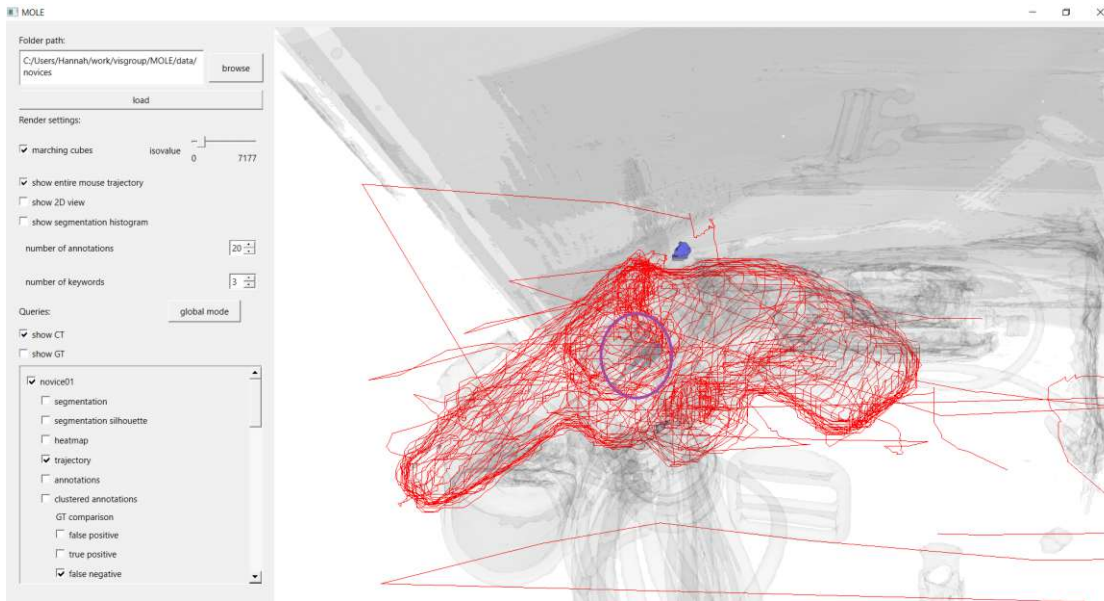


Figure 7.21: Mouse trajectory and false-negatives of Novice 1 with a part missed in the top of the cylinder circled in purple.

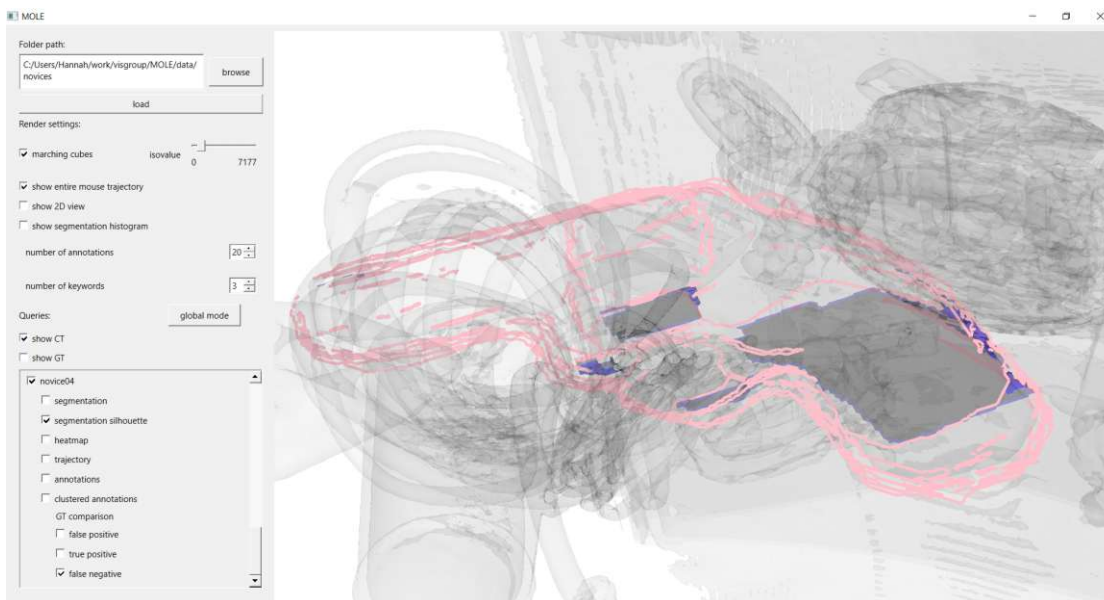


Figure 7.22: Segmentation silhouette and false-negatives of Novice 4.

7. VALIDATION

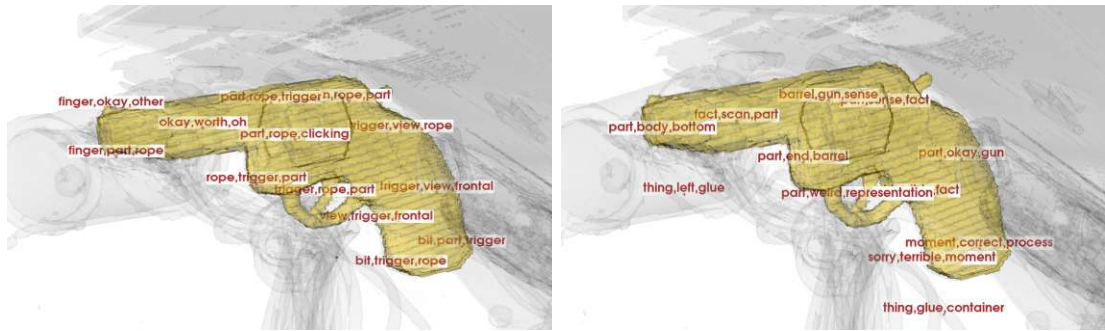


Figure 7.23: Clustered Annotations for Novice 2 (left) and Novice 3 (right) combined with the ground truth.

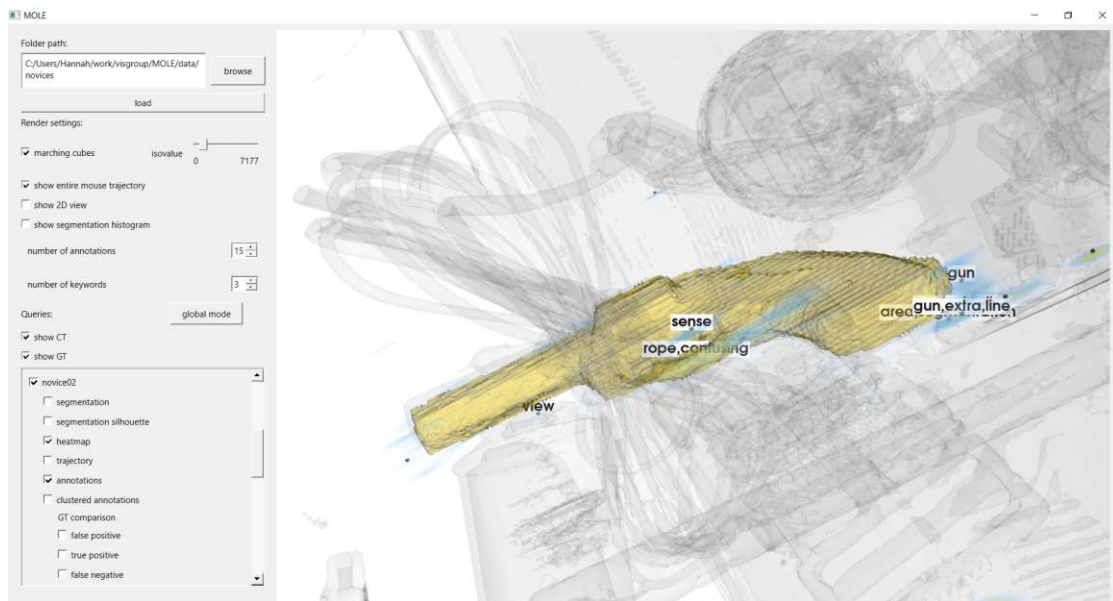


Figure 7.24: Basic annotations of Novice 2 combined with their heatmap.

displayed on the left side of Figure 7.25, we visualize the slice on which the observers were operating. By looking at the previous slice and the slice displayed in Figure 7.25 we see that this slice is the first slice where part of the cylinder is visible. As the part of the cylinder only clearly visible on this slice for the first time, in order to determine the beginning of the cylinder the observer compares previous slices. More information about the thought process is retrieved by looking at the utterances of the slice we marked (Q9). Figure 7.26 shows the utterances for all novice observers. From the utterances of “It comes in here” (Novice 1), “showing already” (Novice 2) and “that’s the gun already” (Novice 4) we can extract that they are talking about the fact that the gun is now showing up. Displayed in Figure 7.27, Novice 3 explicitly mentions that it is “difficult to disentangle from whatever is next to it”, which might refer to the cables. At

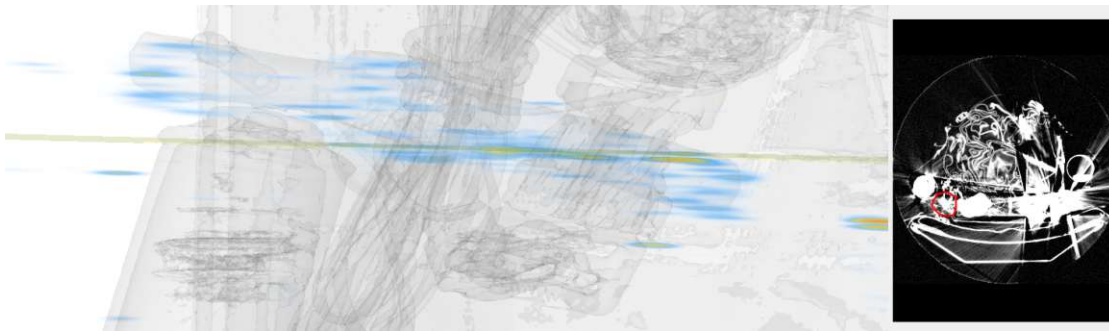


Figure 7.25: Heatmap of Novice 3 with its slice number 178 visualized in 2D and 3D. Marked in red: visible part of the cylinder.

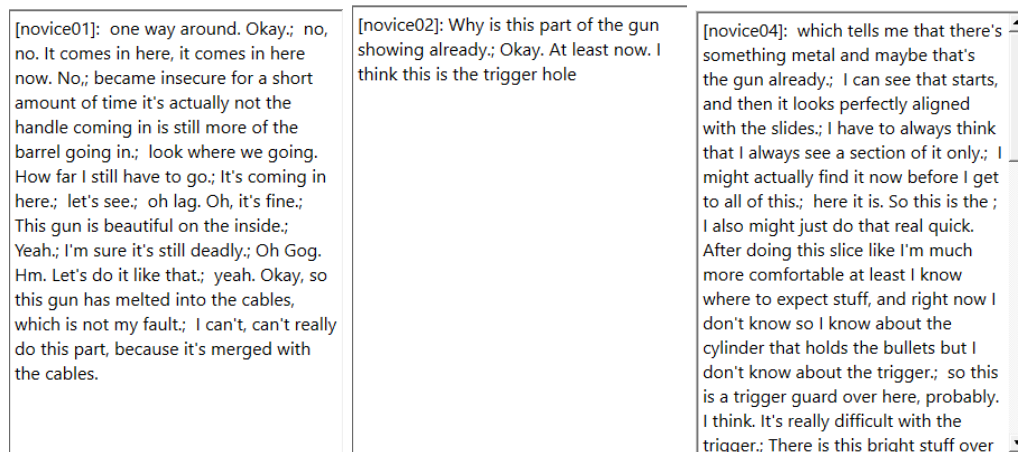


Figure 7.26: Utterances of Novice 1 (left), Novice 2 (middle) and Novice 4 (right) for slice number 178.

a later point, however they say “This is still part of it”, proving that they went back to revise their decisions and segmentations.

Similar to the expert dataset, the outliers are concentrated to the outer parts of the gun. The consensus between all four observers includes more than the ground truth (**Q5**). Contrary to the expert dataset however, the ground truth in this dataset is reliable and so the over-segmentation can be seen as such. The over-segmentation shall not be received as a consideration to change the ground truth, like in the expert dataset. This visualization also emphasizes that the trigger was missed by at least one observer, as it is not included. The voxels that all four novice observers agreed on is visualized in Figure 7.28. The outliers are shown in Figure 7.29.

[novice03]: It's just; shape-wise.; And I don't know what belongs to here. This is quite difficult to disentangle from whatever is next to it; I see there's; I think, I think we were right.; Wait a moment that doesn't seem correct; I am a fool.; This is still part of it; this; Yes I can; Mister software.; I don't know if this makes sense.; So hard to gage.; And it's not clear whether I'm.; And then I think that's it. I think

Figure 7.27: Utterances of Novice 3 for slice number 178.

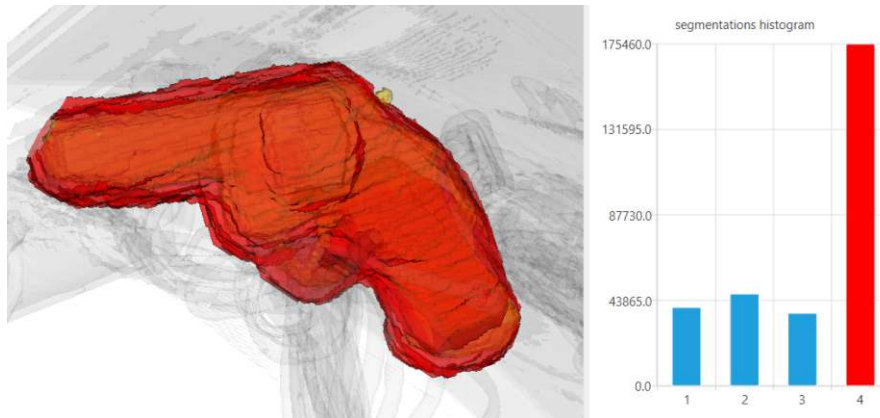


Figure 7.28: Voxels that were segmented by all four novice observers compared to the ground truth.

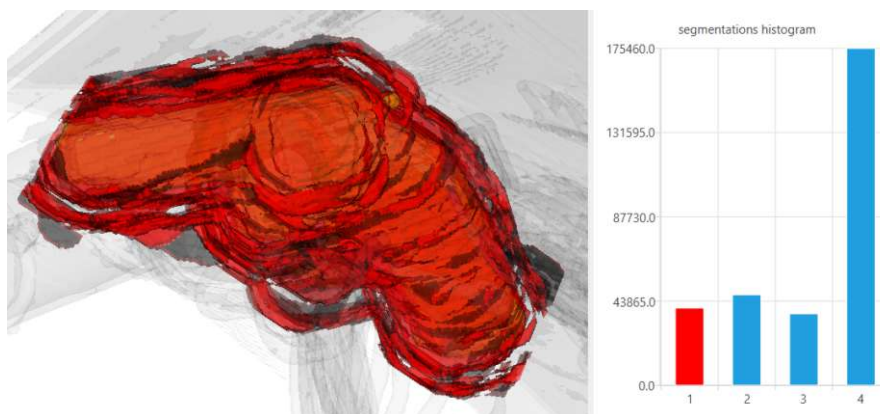


Figure 7.29: Voxels only segmented by one novice observer compared to the ground truth.

7.2 Domain Expert Feedback

Concluding our validation process we summarize the feedback and potential use cases for MOLE, as discussed with domain expert Wouter van Elmpt from the Maastricht University Medical Centre. Currently, our domain expert is an assistant professor as part of the Faculty of Health, Medicine and Life Sciences with a focus on radiotherapy as well as the School of Oncology and Reproduction. Our domain expert's expertise served as input, in order to assess and validate the uses and possible integration into the existing pipeline of radiotherapy. During the walkthrough we showed our domain expert our tool while he commented on it. Afterwards, he was free to explore the tool further by instructing us to navigate through it. The feedback from our domain expert is summarized in the following.

The three-dimensional representation, although working well here, is still unusual to the domain and needs a connection to the underlying image. This is provided in MOLE and delivers information about the original input. A gap that can be filled is to project the mouse trajectory onto the underlying 2D slices. In addition, one could have more information about where the segmentation was edited, where parts were added or removed. Our domain expert found interest in the visualization of the annotations and the mouse trajectory, since they seemed the most novel. The global heatmap showed that most time was spent around the tumor. However, when viewing this data our domain expert considered the time it takes to draw the contour as well. From our visualization our domain expert was able to understand that the observers were not experts in radiotherapy. In practice, the mouse trajectory for experts would only move around the lungs and never go further than the liver. In addition he saw that many people looked at the region in the top right of the tumor and some decided not to delineate this part whereas one person did. "For me it all makes sense what I see", says our domain expert.

As described by our domain expert, target definition is the major uncertainty in radiotherapy. It is the most critical part of the entire workflow and thus tools that can reduce this uncertainty and give insight to it, are of help and need. The segmentation process needs to be efficient, since usually there is only time for a few minutes to look at tumors.

With auto-segmentation tools the delineation of tumors has been simplified and accelerated. However, manual delineations based on the automatic segmentations are still unavoidable, due to the ever-changing nature of tumors. In the pipeline of delineating tumors, MOLE can be used to understand where the auto-segmentation does not work and what regions should be double-checked by the next doctor.

In the clinic, our domain expert says, they have a peer-review process. After one doctor delineates a tumor, another doctor will look over the segmentation and do another delineation. The second delineation can either be done "blindly" without knowing how the first actor segmented or based on the first segmentation. In this stage of the workflow, MOLE can help to highlight those regions where the first doctor was uncertain. For the second doctor's segmentations our tool will explicitly show disagreements. This can be especially helpful for larger tumors because segmenting them takes more time. For

tumors located in the head in the region of the parotid, which is responsible for the production of saliva, the back part is often difficult to see. Thus, if there is a limited amount of time it is best to look at the region in the back of the head, which would be highlighted by MOLE. The uncertainty of experts can be used for training deep learning models and improve auto-segmentation tools. The model can find similarities in regions where there is high variability and be trained with this knowledge.

Cancer cells oftentimes spread to other organs and lymph nodes. Even in case there is only one tumor, but multiple metastases or multiple lymph nodes that are affected, MOLE can be helpful. The affected lymph nodes can spread widely, so it becomes difficult to determine which ones need to be treated. This is difficult to see on the CT, because the soft tissue contrast is not good. When doctors discuss patients, they usually write this information down and discuss it verbally, but with MOLE the lymph nodes in question would be highlighted immediately. The dataset used here, can be segmented quite easily, since it is only one tumor and is not bordering other organs or affected lymph nodes.

MOLE has high potential for educational purposes. People in training for radiotherapy can track their time and results over time and improve based off that information. In addition, seeing their evolution can boost their self-confidence. In a training setting MOLE can be helpful in the discussion with the trainee's supervisors. Tracking what people are looking at, the supervisors can make sure the important parts were looked at and considered. The information gathered from a trainee can be passed on to a trained doctor and be reviewed and fine-tuned by them. In a clinical routine there is usually no time to discuss someone's train of thought. The communication is reduced to lymph nodes or parts of the tumor were a doctor is unsure. Decisions are made very quickly. MOLE can convey not only the regions of uncertainty and lingering but also the thought process and provide more data on demand.

Discussion & Conclusion

8.1 Discussion

We hereby summarize aspects that worked well in the process of this thesis, as well as parts that did not work well. Furthermore, we highlight our major findings and contributions.

In order to answer questions **Q7** and **Q8**, we captured the mouse trajectory of the observers. Capturing the mouse trajectory instead of using an eye-tracking device or software, proved to be of an advantage. It was cheaper in terms of costs, time and complexity. The outcome using this method resulted in useful heatmaps that help to give insight into an observers' areas of lingering. However, since the mouse was tracked at all times, we oftentimes have a hotspot in the positions the observers started out at. These hotspots would need to be accounted for in the heatmap. Deleting them however, would distort the data and take out the position entirely, resulting in a gap.

The initial delineations obtained from the data acquisition differ slightly from the segmentation that is finally being displayed in MOLE. This deviation stems from the procedure to close the gaps in the segmentation results from the data acquisition process. This processing step is needed to fulfill task **T1** and visualize the segmentations. The erosion applied when closing the gaps is responsible for enlarging the segmentation and even though a dilation is applied as well, the shape of the initial segmentation is still slightly changed. We observed that this deviation is bigger in the novice dataset. This might be due to the higher number of slices and the associated decreasing patience. The possibility of gaps emerging originates in the logging tool used in the data acquisition process. The improvement of the logging tool was discovered in the process of producing images for the thesis and is described in Section 5.5.1. Since the visualization of the segmentation data was implemented after the acquisition the problem originating from gaps in the segmentation had to be fixed after the acquisition. The posterior adjustment

of the acquired data is described in Section 6.2.2.1. Anyhow, with fixing the logging tool such that gaps in the segmentation are not possible, the results will be more accurate. Furthermore, the data processing would be reduced by several steps, making the pipeline of visualizing the acquired data more efficient.

Currently, there is no data processing step implemented into MOLE directly. The data is prepared beforehand and then loaded into our tool. This caused for more work in preparing the necessary files to allow for a quick data conversion into the right format. In addition, it prevents us from using the information logged from the data acquisition directly, which is desirable in order to make the pipeline more accessible. Processing the data within MOLE, allows for more statistical information to be computed and visualized, offering more measures to draw conclusions and insight from. The data processing step was an unexpected bottleneck for the visualization, as other ideas for visualization components, would have required additional processing steps. Due to this bottleneck the requirements of MOLE to be easy and quick to use, as described in Section 4.2, is only partially fulfilled.

Usually, segmentations are required from a two-dimensional slice view. We chose to visualize the segmentations originating from two dimensions and a slice number in a three-dimensional space. This allowed a more intuitive visualization. As Novice 3 said “It is very strange to look at, the 2D slice representation of an object whose shape you completely understand in 3D.” However, since the segmentation outcome is highly related to the image the observers were seeing when segmenting, it was important to find a way to link the initial image and the segmentation outcome. This also answers questions **Q6** to **Q9**. The connection to the original input image can crystallize reasons for the specific outcome, since decisions are made on what is seen in the initial image. We introduce this connection, by implementing the slice view on demand that is being linked to the three-dimensional view by the cuboid placed in the render view. Although this component serves the purpose of linking the input image and the segmentation outcome, information that is represented in the three-dimensional space cannot be connected to the two-dimensional origin image precisely. MOLE only displays the original image, so an exact visualization of the segmentation, the trajectory, the heatmap and the annotations is not possible. These connections can be importance, when analyzing the causes for a specific outcome, as reasoning strategies are mostly based on how things are perceived from the original image.

When it comes to evaluating MOLE, we used two different datasets to prove other application areas than the medical field. Looking back, the novice dataset could have been more similar to the expert dataset in terms of the amount of slices that needed to be segmented. This would have made the results of the data acquisition, especially in the confidence survey, more comparable. Receiving feedback from our domain expert, introduced a lot of perspective and ideas in what direction to extend MOLE from a medical perspective. These tips would have been very useful in the beginning of the process. Due to time restraints, we were not able to validate MOLE with other professionals. Ideally, more external experts could give feedback about possible use cases for MOLE in different

work environments and pipelines.

In closing we can say that the step of data processing took longer than expected, due to the gaps in the segmentation files. We needed to find a plausible solution to this problem that would distort the data as little as possible but still deliver binary masks that could be visualized. The data acquisition itself however, proved to be a very easy and well-established process. The sessions went smoothly, without any inconveniences and the data obtained, provided all the information needed to answer the proposed questions **Q1** to **Q9**. In fact, the logging tool we developed to capture the segmentation process, provides a great tool for fast and accessible delineation and generation of attention data. The heatmap and the segmentation ground truth comparisons proved to be especially useful. The heatmap visualizes uncertain areas of the process, which as such is unique for segmentation processes. We were able to successfully identify regions of hesitation, since the attention data was captured was visualized. The ground truth comparisons enable a quick overview to crystallize regions of over- and under-segmentation. Due to the quick computation based on only the ground truth and the segmentations, no additional data was required to be captured. We have shown that with the help of the heatmap question **Q6** can be answered. The segmentation and ground truth comparisons are crucial in order to answer question **Q4**. Above the novelty of these specific components, MOLE uniquely enables a dynamic, multi-faceted exploration tool. All the different components can be combined, added, removed and adjusted as needed, in order to answer the user's questions.

8.2 Future Work

While we focused on exploring if qualitative analysis can be facilitated through interactive visualization of qualitative observation data and found answers in the methods we used, there are still loose ends with potential to extend this work in the future.

In the process of acquiring the data there are two major improvements that would possibly lead to better and more detailed results. As mentioned in Section 5.5.1 the logging tool can be constructed in a way such that there would not be problems with open segmentation lines. In addition, the data acquisition process can be improved by reducing the amount of steps needed to capture the data. For now, the combination of the logging tool created in MeVisLab and the transcription software of Zoom was sufficient. In the future it would be helpful to have all the necessary functions combined in one tool. This advancement would save the step of aligning the transcript accordingly and provide more detailed logs, eliminating sources for error in the pre-processing step.

Although using the mouse to track the observers' attention areas was sufficient for the scope of this thesis, the attention data has room for more accuracy. In this extent using an eye-tracker instead of the mouse to capture the observers' attention areas, may deliver more accurate results.

With better machine learning models the transcript data can be assessed based on its

sentiment, which allows for more insight and a better summary of the transcription data captured. This opens the possibility of an additional visualization component displaying the sentiment of the observer in various regions, leading to a quicker analysis of possible sources of uncertainty.

At the moment, the clustering happens in a processing step prior to the visualization and is only computed for one whole observation. Implementing this processing step in a dynamic way allows for a viewport-based query for keywords. Keywords would be displayed based on the viewport of the user and queried only within the area that is being analyzed. With a dynamic clustering process the amount of clusters would also be adjustable. With access to all data at runtime, normalizing the data accordingly would provide a different insight as well. It is worth investigating if a function to search for words and phrases would be of help when using MOLE. This would allow for the user to search for specific keywords among different observers and find possible connections more easily.

Further visualization components that could either be added or improved include the connection between the two-dimensional and the three-dimensional representation, as well as a time-based visualization of the segmentation process. The connection between the slice view and the information in the three-dimensional space are currently linked by showing the slice of the two-dimensional view in the render window. Therefore, there is room for improvement, like adding the slice view to the render view as a texture. In addition, the heatmap could also be overlaid on the slice-view to directly combine the information of the heatmap and attention areas with the slice view.

Including a way to visualize how the data was produced by combining the mouse trajectory, single segmentation slices and possibly keywords would introduce a more interactive way of analyzing an observer's process. This could be incorporated using a time-slider in combination with the visualization components already provided, but updated according to the time.

Regarding the interface of MOLE, with a more complex dataset, knowing what is being displayed in the render window can become a struggle. This can be improved by implementing a window giving information about what is being displayed. Since the focus of MOLE mostly laid in the visualization components and core functionalities, the interface lacks a design concept that goes beyond the default layout provided by Qt. Therefore, there is room for advancements in simply adjusting the look and feel of MOLE to a more modern and time-appropriate look.

8.3 Summary & Conclusion

In this thesis, we presented MOLE, a tool that allows to analyze reasons and causes of inter-observer variability stemming from segmentation processes of image data, exemplified on medical imaging. The focus lies in exploring and explaining the thought process and reasoning strategies in regards to the segmentation outcome and attention areas using

our visualization. We acquire the attention data, thought process data and segmentation outcomes with a logging tool developed specifically for this thesis. Creating two different datasets, one from the medical domain and one from a general application allowed us to validate MOLE for data regardless of the original domain. The obtained data provides information about regions of uncertainty, of which the causes can be analyzed with MOLE.

Through using different visualization components we allow for a comparison between different segmentation outcomes. In addition, we provide different tools to understand an observer's approach and technique to the segmentation problem. Users can dynamically change the information displayed, enabling a detailed insight into the segmentation process. MOLE allows for a comparison between the segmentations and approaches of multiple observers as well as detailed information about all observers combined. Using heatmaps, keyword annotations, false-positive, false-negative and true-positive visualizations, the mouse trajectory and segmentation outcome itself, we provide multiple components visualizing the acquired data.

With MOLE, it is possible to investigate causes for regions of uncertainty and find reasons for a specific outcome. By effectively combining the provided visualization tools, MOLE can answer questions arising from inter-observer variability. Furthermore, observers are able to re-visit their own work, providing a helpful tool to improve ones own segmentation process and learn from past tasks. We have shown that our tool can help to understand the reasons and causes for uncertainty stemming from inter-observer variability. We prove that qualitative analysis can be facilitated through interactive visualization of multi-variate observation data, conveying the interplay between imaging information, attention, reasoning and final delineations on visual input, across multiple experts.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Data Acquisition

A.1 Welcome Text

Welcome! Thank you for participating in this data acquisition. The data collected will be serving for my master thesis with the title "Multi-faceted Visual Analysis of Inter-Observer Variability" at the Institute of Visual Computing & Human-Centered Technology at TU Wien. The goal of the thesis is to develop a technique to visualize the uncertainty provided by different observers looking at the same object. We focus on the process of marking this object and visualizing the reasoning of the outcome.

To acquire this data, you will be given a CT scan of a backpack, containing a gun. You will see the 3D model of the backpack as well as a representation with slices. Your task consists of marking the gun in every slice where it is visible and to comment on every step you take.

CT scan: A CT scan, also known as computed tomography scan, is a medical imaging technique used in radiology (X-ray) to obtain detailed internal images of the body non-invasively for diagnostic purposes.

A.2 Warm-Up Task

To get familiar with the controls and the tool we have prepared a warm-up task. This is a simpler version of the task you will be given in the second step of this data acquisition. The objective of this task is to get familiar with the tool so play around as much as you need until you feel comfortable using the tool.

There are 372 slices of a backpack containing a gun. Find the gun and mark it in every slice where it is visible. Point at the areas you are looking at and talking about with your mouse cursor. Comment on your thought process as much as possible. Think out loud:

A. DATA ACQUISITION

- why are you marking things the way you are?
- what is confusing about the image?
- what is clear about the image?

Controls On the *left side* of the window you see a 3D View of the scan which you can look at by holding the left mouse button and moving the cursor around.

On the *right side* of the window you see the slices of the scan which you can look through by either scrolling or using the up- and down-arrow keys. By holding the left mouse button and dragging on a slice you can mark the gun. Finish one marking by double clicking.

You can delete a marking by selecting it (left mouse click) and pressing DEL.

Cleaned Transcripts

B.1 Experts

B.1.1 Expert 1

12:26:38 So, I go to the left picture and turn this 3d picture so far that I can see the patient for as if I stand at the foot of the bed.

12:27:01 So I can see here, the heart.

12:27:06 Kidneys.

12:27:10 I'm trying to find out if you can see the tumor here or not.

12:27:28 I cannot locate it. I'm not sure maybe it's here

12:27:31 Let's see.

12:27:40 On the right side I tried to scroll all the way up.

12:27:50 Here we have the spinal cord.

12:28:04 the lung tissue starts here.

12:28:16 and the aorta

12:28:20 and lung tissue

12:28:34 Here are still the bronchi and the heart.

12:29:00 I assume this is the tumor. So it starts here or so. Okay.

12:29:26 I try to orient myself on the left side.

12:29:35 It is difficult to

12:29:40 tell the border to this part here. I think that's a bronchus. I'm not sure it is a.

12:30:12 Okay, It's possible that it's still part of the tumor so I will outline it too.

12:30:20 It starts here actually.

12:30:50 probably here

12:31:08 i'm not sure if this still belongs to the tumor or it is just bronchi

12:32:02 Okay. I go further.

12:32:10 I think this is the liver.

12:32:12 This is the liver.

12:32:23 So, am I looking for just one tumor?

B.1.2 Expert 2

10:35:15 okay right so I can see on the bottom left of the right thing that it says slice 40 so I just want to get to zero.

10:35:30 Right.

10:35:34 Okay.

10:35:37 So I think zero is probably down here, because usually.

10:35:45 It's like shot like shot from.

10:35:48 Well, the way you look at it is from bottom to top. So, and then also I can here see the bones of the femur and the hips.

10:36:00 So that, so this is like the way.

10:36:05 I think I'm looking at it.

10:36:09 Yeah.

10:36:12 And then there's a.

10:36:16 Okay, let's see if I can see a tumor on here, I can't. Okay.

10:36:20 Anyway, let's just move up and I will try and position this guy properly.

10:36:29 So, there we go. Okay, so now I'm just kind of moving up.

10:36:40 up, up, up and the white, so the more, like I said, the most dense, things will be white so those will be the bones.

10:36:52 Yeah, so now I can kind of see that the separation from the legs is kind of disappearing and it's getting more compact so I should be like somewhere in the like hip area.

10:37:08 So those would be the

10:37:11 hips know really know the bones in English but yeah anyway.

10:37:16 So now I'm just kind of moving up, up, up, and then yeah right so here I can see the spine.

10:37:25 Yeah. This will be the spine. And then here. That should be a blood vessel because it's really dark.

10:37:39 So, probably be the aorta

10:37:51 Right.

10:37:55 still going up.

10:38:16 So here. Okay, Perfect.

10:38:21 All right. Okay, so if I go back down.

10:38:29 Let's see. Right, so here on the left and the right side, I can see the kidneys. Those two things, those are the like lower poles of the kidney that are appearing.

10:38:35 And then, so the sides are like flipped, so this is right in this left and then on the right side, this would be the liver.

10:39:15 Okay that's okay. All right, so this, this is liver and the liver has, like a left and right part and then in here there was like a dent where all the ligaments and muscles go so that I can see as well.

10:39:20 Right, so these will be the vessels that connect the kidneys to the aorta

10:39:32 moving up, moving up so the diameter of the kidney is getting smaller and the liver is getting bigger, so I'm like, moving up and up.

10:39:44 And then I think, what is starting to appear here would be the stomach.

10:39:51 So I'm still kind of in the abdomen area, and then

10:39:58 I'm still kind of in the abdomen area, and then that might be the pancreas.

10:40:06 Okay, moving. Right, so now I can see. So right

10:40:11 now I'm just where the diaphragm is so that's what separates the thorax from the abdomen.

10:40:24 And then I can see the heart starting to take shape.

10:40:32 Right, so this will be.

10:40:36 Yeah.

10:40:38 Okay, so this is where the heart kind of starts and then I have the aorta here and the spine here.

10:40:49 And the lungs.

10:40:53 On the left and right side and I can see the like vessels of the lungs.

10:41:04 and this thing should not be here so this I'm guessing is the tumor.

10:41:11 Let me just see okay.

10:41:15 Right, okay.

10:41:18 So, let's see, I'm going to start marking it from the bottom side. Okay, this is where it appears first. Right this is the heart.

10:41:37 This is probably still stomach that's kind of invisible

10:41:43 because it disappears then. Yeah, right. So this is the first slice of the tumor and this will be

10:41:52 the pulmonary veins and arteries, but let me check. Okay.

10:41:59 is the tumor on the left one is it.

10:42:05 Can you see it in on here as well or not.

10:42:11 Okay. All right.

10:42:11 Great. Okay, so let me draw a line so that will be the right side, right lung.

10:42:20 Okay, wait.

10:42:33 So, right. And like I said, the lung, is not as white, as dense because it's air filled, and then the tumor.

10:42:44 Because there's obviously some sort of like tissue growth or something so it's more dense and not air filled so that is how you can tell it apart from the rest.

10:43:09 Um, Yes.

10:43:12 Alright.

10:43:18 So the thing that, like, I'm not like an expert at, is kind of telling a part.

10:43:30 what are vessels and kind of like goes a bit like together in here so I will just kind of try and focus on the tumor mass and just assume that those around there just kind of vessels.

10:43:54 Yeah, so then when we look at a tumor, you can kind of try and see if it's like, filled with like liquids.

10:44:01 So that were vessels in the slide before so I think I'm just kind of kind of leave those out, but I will, like, go back to the

10:44:18 previous slide, kind of check. Okay there I did not
10:44:21 See. Okay, so those are vessels, I think, yeah.
10:44:38 The thing is like the tumor could be invading the vessels as well so that's like the one thing I'm kind of maybe struggling with. Um, let's see I don't know I think I'm just gonna, because the tumor mass is a bit less dense than the vessels.
10:45:15 All right, think they're a bit, bit more grayish. Yeah.
10:45:08 Okay.
10:45:13 Can I delete that wasn't very pretty.
10:45:33 Okay.
10:45:33 All right. Anyways, when you're looking at a
10:45:38 tumor. things that you can consider are like density, is it liquid filled is the kind of parted.
10:45:47 What, what are the, the edges like are they kind of straight or are they more... Is it like more of a like, like synonymous No.
10:46:03 Are they like clean edges or rough so things like that. All right.
10:46:05 Yeah. Okay next slice.
10:46:22 Yeah. So I think those are like pulmonary veins.
10:46:30 Between the heart and the tumor and I yeah so the I've been in the like.
10:46:41 I have already passed, like largest diameter or like extension of the tumor and I'm kind of moving up and it's getting smaller as I go up.
10:47:33 ok so now the tumor has.
10:47:36 Yeah.
10:47:38 So as I go up and up let's just scroll through the entire lung, so that it's like complete.
10:47:46 Yeah.
10:47:54 And then I've got head and arms, brain. Brain. The head is getting smaller. Arms.
10:48:08 Let's scroll back.
10:48:15 and See if like there's anything I missed. So, this will be the apex of the lung.
10:48:22 And then let's just see if the tissue is kind of like the same density and if there is anything that should not be there like smaller than some parts. So those will be vessels.
10:48:41 That's fine, so yeah so now I can see like the connection to the
10:48:51 bigger vessels of the body, I can see the heart, and like chambers of the heart, hearing, and then.
10:49:05 Yes, so this is where the tumor comes in the picture. And let's see, if I marked the tumor in each slide.
10:49:13 Now, as I've moved down. Yeah. Okay, so I think that's good.

B.1.3 Expert 3

11:15:19 So, I start here because it's easier to look at.
11:15:20 Okay, so what do I see,
11:15:22 I see the heart here.

11:15:31 I see something here, like the...
11:15:54 Um
11:16:00 like what is this,
11:16:04 like and organ
11:16:09 Okay.
11:16:15 I don't think I see the tumor here, because
11:16:28 it should be like, I would imagine it to be. Not like this transparent as this thing here.
11:16:38 Maybe it's somewhere here.
11:16:51 Okay. I look at the 2D view again. Is this the top?
11:17:10 Okay.
11:17:21 I just look through it.
11:17:48 Okay, So um
11:17:51 I'm trying to orient myself
11:17:55 orientate
11:17:57 So
11:18:00 those are the lungs.
11:18:03 Here's like the heart and
11:18:10 vessels.
11:18:12 And that's the vertebra.
11:18:14 that's the liver.
11:18:33 The kidneys.
11:18:42 Okay.
11:18:44 What if I never find it?
11:18:49 Okay.
11:18:56 This is like the bottom part. Legs.
11:19:42 Okay.
11:19:50 Okay so it's like this.
11:20:06 Okay I found it. I start at the bottom. I like, I see that the color of the lung is like very different that's actually, I think those are like bronchi and stuff.
11:20:12 Or those are vessels. These are also vessels I think.
11:20:20 And this is too big to be like a vessel or
11:20:25 bronchi or something.
11:20:28 And it's not like if I scroll here I see that there's no connection to something.
11:20:39 At the bottom, like an organ.
11:20:42 And if I scroll through, I see it might connect here.
11:20:46 But I know from anatomy that there's nothing that goes like from the heart into the lung, like this far.
11:20:57 Okay, and now
11:21:03 I don't know, it's harder to say where it stops here.
11:21:12 But it looks like it ends here.
11:21:18 I don't know what this is.
11:21:20 But it's also here maybe the lung vessels.

11:21:24 or something.
11:21:33 Okay, so I start marking it.
11:21:51 here I think that it, like, I don't know.
11:22:01 I think it ends like here.
11:22:01 Because this is like.
11:22:10 This looks a little brighter than this, but I'm not sure.
11:22:13 And this looks strange to me, the connection.
11:22:30 Here it looks like it ends here.
11:22:43 Maybe then this is also part of it. This would also
11:23:55 No I Think not.
11:23:11 I think it's just this.
11:23:47 um.
11:23:54 Oh, I think like from the color, I would say it's just this part but I know.
11:24:05 Because this is so bright, I don't know.
11:24:10 Okay.
11:25:07 Is it bad that I painted like this?
11:25:45 Okay. I don't think I see it here.
11:25:55 I think that's it.

B.1.4 Expert 4

16:20:06 So, to start off with, I'm having a look at this 3D model. I can see both the kidneys very clearly.
16:20:22 And the heart,
16:20:28 the spleen.
16:20:35 Okay.
16:20:54 It's interesting to see how the arteries look here, when you're having the hands above your head. Okay.
16:21:05 Now, I'm starting off with the CT scan, I'm trying to like
16:21:22 get some orientation okay so here's the head I'm starting, beginning from the head.
16:21:22 I know it's a lung tumor but I just wanna have a quick look through the head
16:21:29 Okay, sinuses.
16:21:35 Oral cavity now through to the neck.
16:21:45 and here we can see the beginning of the lungs, the apex, I don't know if it's called to the way in English.
16:21:51 slicing through.
16:22:02 Now the aorta.
16:22:12 Esophagus.
16:22:26 Okay so here
16:22:33 is the beginning of.
16:22:41 So this is liver.
16:22:58 liver, liver and that's stomach.

16:22:53 and behind somewhere there's the spleen.
16:22:58 Okay so, I'd say this thing here must be tumor.
16:23:08 Okay.
16:23:11 Because
16:23:14 when we start here this is heart. Heart. Heart with like to four chambers.
16:23:24 Over here. These must be very pulmonary artery or vein.
16:23:30 And here this must all be lung tissue and then here.
16:23:41 This must be spleen, and stomach and then liver and that's the end of the lung.
Okay, so I would say this thing here.
16:24:15 Um
16:24:28 I'm not quite sure how, how to mark the little spots here I guess this are lymph
nodules,
16:24:30 which often can be bigger.
16:24:37 When there is a tumor.
16:24:41 So, this is tumor here.
16:25:03 Well, it's very solid tumor.
16:25:11 I think I won't mark these white stripes here.
16:25:20 Because I think these are
16:25:26 arteries, or veins, capillaries.
16:25:36 Or maybe I should mark them because they seem like they belong to the tumor.
16:25:48 And I think I won't mark this thing here I'll show you in a second.
16:26:07 I think I want mark this here cause I think that's not tumor.
16:27:11 yeah okay so that's trachea here.
16:27:15 And then I follow this here
16:27:23 Bronchus.
16:27:28 Bronchioli.
16:28:08 I'm not sure but it looks like it could compress the bronchus or even infiltrate,
which would definitely make it harder for the patient to breathe.
16:28:51 Okay so here, this, this thing here is tumor.
16:29:04 And then again, yeah that's tumor.
16:29:15 Yeah.
16:29:16 Okay. Um no.
16:29:33 I can take this.
16:29:39 I'm just thinking if I could also
16:29:47 mark this here. No, no.
16:29:55 That's tricky. Wait.
16:30:16 No tumor here.
16:30:24 Well, this
16:30:28 This has to be tumor.
16:30:46 This has to be tumor.
16:31:15 You can see this is tumor too here.
16:31:59 I think this can't be right because here.
16:32:07 I'm sure about these here, down here, then

16:32:16 This thing here
16:32:17 Yeah, but then it gets bigger again and then like this small. this can't be right.
16:32:30 what did I do.
16:32:33 Okay.
16:32:41 No cause here.
16:32:45 No this is
16:32:56 So this is the left ventricle and
16:33:00 this must be the right one. Yeah, so this is a pulmonary artery here.
16:33:10 Yeah.
16:33:15 Artery.
16:33:16 Artery.
16:33:21 I'm still not. I think this one here.
16:33:21 I'm not sure. I'm really not sure if this is tumor here or not, because I see like a slightly dark grayish line. So,
16:33:39 then maybe this here is tumor too. I don't know,
16:33:45 I'm not sure.
16:33:47 Shall I mark it if I'm not sure or?
16:33:55 I think I'll go with it.
16:34:01 Well, we would make a biopsy anyways.

B.1.5 Expert 5

14:48:15 Good.
14:48:17 I am looking at it on the left side for an overview.
14:48:29 Can I zoom in somehow or does it only work like this.
14:48:33 Oh yeah.
14:48:50 Okay.
14:48:57 Um.
14:48:58 Good. I am going to look at it from the top, so cranially.
14:49:05 I will look at it from all the way cranial.
14:49:08 It's supposed to be a lung tumor so I will look where the lung starts. Which would be here.
14:49:26 Good, here in the right lung there is a little dot.
14:49:33 That could be something.
14:49:37 So a small white dot.
14:49:40 Okay, no. That are probably pulmonary vessels because there are multiple.
14:49:48 Let's see.
14:50:05 Okay.
14:50:26 Okay. These white things are more like pulmonary hiluses because you can also see the vessels to the heart.
14:50:43 Do you see my mouse? Oh okay alright. Then I can show it more easily.
14:50:51 This is the heart.
14:50:55 Here you see the vessels going to and from the lung.

14:51:05 And this is probably the tumor, because there should not be anything here. This is supposed to be lung tissue like here.

14:51:19 I am just going to look through entirely.

14:51:26 Exactly these white dots here are normal. They are pulmonary vessels.

14:51:37 And on the right side you see these dots, they are also normal.

14:51:49 Good. And here we are.

14:51:49 Spleen.

14:51:50 Liver.

14:51:55 Okay. I think I am going to start marking now.

14:52:07 These are very coarse cuts because it starts very big and is gone in the next slice.

14:52:20 Okay.

14:52:32 Yes, they are really coarse cuts because it is suddenly much bigger here.

14:52:42 Oh man.

14:53:11 Now. Oh man, what is happening?

14:53:32 Now phew.

14:53:38 Okay, here I am unsure if these stripes belong to it.

14:53:54 Because it could be that it's simply a vessel that is cut like this.

14:54:00 But. Hm.

14:54:17 So I would not count these two stripes on the side. Because I think they are vessels, but I am not completely sure.

14:54:40 Even though here it looks like it belongs to it again.

14:55:07 Here it is difficult too because of the structures of the hilus and I don't know exactly.

14:55:18 So a different cut would be better, so you can distinguish it. Because from the contrast it looks pretty much the same.

14:55:39 Can I scroll while marking? Probably not. Okay.

14:55:49 Can you delete this?

14:55:54 In the meantime I will look at the other slices.

14:56:10 I will start from here again because here I am sure.

14:56:32 Okay. Here it is still doable. Here we can still distinguish because on the previous slice you can see that it belongs to the vessels.

14:56:56 I am just going back and forth so I can compare

14:57:01 how the image changes.

14:57:07 Hm.

14:57:25 It could be possible that this is part of the tumor because

14:57:37 it squeezes itself over there.

14:57:49 Okay.

14:57:56 I will take the little tip here as well, because of the contrast. There is also a little bit of greyish here too.

14:58:08 And here you can really see that is it completely white.

14:58:17 This should technically go over into the other one.

14:58:32 For me it makes it really difficult that the slices are so thick.

14:58:40 Can I delete this again?

14:58:50 No.

14:58:52 Thanks.

14:59:12 Alright.

14:59:22 Then here. Here it is really very difficult.

14:59:46 So I go through again because I am really unsure.

15:00:27 Okay, here you can see the vessels and in the next slice it is somehow blurred in with the

tumor. That's why I am not sure where to start here. But I will just do it here because of the contrast that looks like tumor tissue or it looks more like tumor tissue.

15:00:57 Then I do it like this because I think these are more like vessels.

15:01:55 I will look through again.

15:01:58 So as I said with the stripes I am really unsure. This I can't tell if they are vessels that are trimmed or if they really belong to the tumor.

15:02:13 Because if they are trimmed very thin the contrast could also be different.

15:02:38 Okay.

15:02:43 Here for example you can also see a pulmonary vessel that is trimmed very thin and this is definitely a pulmonary vessel because it leads here.

15:02:54 So this could be the same in the slice.

15:03:03 I would say it is like that.

15:03:08 Yes. And I think,

15:03:10 I am done.

B.2 Novices

B.2.1 Novice 1

10:23:23 It's not a gun Oh no, is that, no, it's a fucking lighter I'm sorry also for the swearing. There we go.

10:23:31 There we go.

10:23:34 Okay.

10:23:37 It's a chunky boy.

10:23:40 That's definitely. Is that mine? Is that my own thing?

10:23:49 That looks like if I had made it. It looks so bad.

10:23:56 Looks deepfried. I like the deepfried gun

10:24:00 and other things. Alright, alright.

10:24:04 So, I'm gonna see if I can find it with a quick scroll through.

10:24:10 I can see the spraying cans to line up the scan a bit.

10:24:14 Oh, oh no.

10:24:19 I should be looking here. No that's the lighter now I'm

10:24:25 wrong side, one way around. Okay.

10:24:32 There we go. There we go there's our boy.

10:24:36 Yeah.
10:24:39 Maybe look for a job at the airport security.
10:24:42 Okay now of course the gun is split a bit.
10:24:46 which makes it a bit harder.
10:24:48 Can't forget to mark, both parts. I already forget it here. Okay.
10:24:55 See here.
10:25:06 I don't see anything from here, and I will assume this part belongs to the gun because it comes together. Yeah, there we go.
10:25:13 Yeah,
10:25:17 wrong way around.
10:25:22 This part I think also belongs to the gun now it's a bit hard, also because it's bit blending with the cable things around it, which are also mental, so of course they would register in a similar value.
10:25:40 Here we go.
10:25:44 No, I'm not happy with this line,
10:25:57 I keep forgetting that I have to scroll the other way around. Now I did it the other way before
10:26:04 I need to scroll down to go forward,
10:26:13 Would be. I mean, yeah, would be nice to have an indication which direction I'm going but I might have that.
10:26:23 like implicit.
10:26:26 There we go. Yeah.
10:26:32 No, it's coming together.
10:26:37 may even be still split on Oh, I'm gonna say I'm happy with this
10:26:46 Again, cable parts sticking in and around of it.
10:26:55 So I have to, technically assume at points.
10:27:00 Like here.
10:27:04 But I mean, I am familiar with the shape of a gun a bit, because it's like sort of a common imagery.
10:27:10 So I will just
10:27:14 work from my own concept of it.
10:27:21 Okay, so now I have officially lost, what I was going to consider Gun and what not
10:27:33 this part, I think I would add the one from the next image.
10:27:40 Yeah, I just attach another one.
10:27:49 No, No, you tricked me.
10:27:57 no I always get my microphone to close to my mouth.
10:28:02 I was getting my microphone too close to my mouth. Um, I mean this might be part of the hammer or something.
10:28:14 Yeah.
10:28:18 I want to now double check how this gun actually looks like.
10:28:24 I mean yeah it's sort of chunky Uh huh. Yeah, fuck.
10:28:31 Oh, this might be the belt buckle. No, may not be the belt buckle. Its overlap-

ping at some point.
10:28:39 No, no, no. It comes in here, it comes in here now. No,
10:28:46 that might be might be alright still.
10:28:50 This is curious though.
10:28:59 Let be tricked by some. I can see it here though. No, it might be the hammer should be fine.
10:29:08 Gonna do one more scroll through it. Yeah, yeah, it's just because it's so weird late Yeah, now the hammers coming in here.
10:29:25 What's the thing.
10:29:22 Oh well.
10:29:28 Yeah, again, lot of overlap with metal stuff around it now gets a bit harder.
10:29:33 It's very similar. Again, have to work off my own assumptions of how things look.
10:29:41 Don't we all. Delving into Gestalttheory.
10:29:50 This is like there is no way to know what's in and what's out, but this stuff overlaps very close. I've already seen it.
10:29:59 Almost as if someone has added this gun
10:30:06 this image.
10:30:07 So here we go.
10:30:38 I'm starting to get to the handle.
10:30:46 Still always very similar shapes.
10:31:07 Yap, handle is coming in.
10:31:31 Oh,
10:31:20 it's also good advice if you need to had a gun in your luggage.
10:31:26 To put it to the cables.
10:31:29 Because airport person will have a hard time doing the segmentation.
10:31:52 became insecure for a short amount of time it's actually not the handle coming in is still more of the barrel going in.
10:32:02 But I'm fairly sure I'm marking the gun.
10:32:10 I am working from the 3d to the to the 2d image so I wouldn't have been able as easily finding the gun in the slice image, as in the 3d image
10:32:24 was able to locate it quicker in the 3d image.
10:32:32 Hope that line is there now. Yeah.
10:32:46 I feel like that CSI dude.
10:33:00 had a little lag here, but it's fine.
10:33:23 Ah. So I'm counting the shade that it does here, because it looks like it might have been interpolated or something.
10:33:37 Because there's.
10:33:38 Let's see the barrel vanishing here. So I count this too.
10:33:43 Oh, I'm still drawing.
10:33:46 Good old MeVisLab. Yeah, there we go, and I think I'm through with the barrel now.
10:34:12 It's a bit confusing how the starting point of the line was a bit offset here again.

I'm gonna scroll ahead a bit again, look where we going. How far I still have to go.
10:34:27 You never should do that. Yeah, are we coming up there.
10:34:31 Okay.
10:34:35 Yeah, this is really just now I'm off-the-top-of-my-heading this. I don't know, I think I see the trigger coming in later now, I've already screwed up the trigger.
10:34:46 Yeah, like a while before. Oh no.
10:34:49 Let's see well this isn't it.
10:34:52 It's coming in here.
10:34:55 I'm going out here,
10:35:00 and I may want to add this bit like I'm just getting this after the fact. I mean, yeah, now.
10:35:08 There we go.
10:35:11 Got a lot of noise in that area.
10:35:41 No.
10:35:51 Let's see if I'm
10:35:53 Oh God.
10:35:55 Not that consistent.
10:35:59 Can't decide now if I want to count this or not
10:36:04 I'll count it.
10:36:07 It's a thick trigger that's the opposite of a hair trigger. There we go.
10:36:41 Hmm.
10:36:41 Yeah, this is now. I don't know.
10:36:45 probably gonna have another look at it.
10:37:00 Okay, let's see.
10:37:23 Well segmentation is hard.
10:37:40 I like when people knock at to your office door and then enter immediately.
10:37:54 Oh, oh lag. Oh, it's fine.
10:38:07 This gun is beautiful on the inside.
10:38:10 Yeah.
10:38:15 I'm sure it's still deadly.
10:38:29 Oh Gog. Hm. Let's do it like that.
10:38:35 I have to see. I mean, am I making the trigger up now? Oh, it's here. Okay. Yeah, I see the problem, the trigger is. Is it clipping through the?
10:38:48 It might be.
10:38:50 Yeah, yeah. Okay, so this gun has melted into the cables, which is not my fault.
10:38:57 And it's good.
10:39:01 Now, who can say where the gun starts and the cables end.
10:39:17 This is a human fetus.
10:39:25 Yeah, I can't, can't really do this part, because it's merged with the cables.
10:39:39 All right.
10:39:41 Now I guess also we can only see some part. I'll add this.
10:39:51 Yeah, yeah.
10:39:55 I still hope to get to see my segmentation at some point.

10:40:01 Because that always looks fun.
10:40:13 Hey, it managed to close itself this time.
10:40:27 Yeah, it closes if the green dot finds the other green dot
10:40:33 I've only done like 100 slides so far.
10:40:38 No, the dots didn't find each other this time.
10:40:50 well.
10:41:05 No, it doesn't like me.
10:41:38 It even does a little symbol when you find it. Yeah.
10:41:44 Problem is, I'm drawing all the way, when I'm looking for it so I'm also like
10:41:48 Yeah so now I don't want to. Yeah, yeah.
10:41:53 Now.
10:41:56 I see that the green dot isn't always where the line starts which is, you know, MeVisLab
10:42:01 So I don't know if I'm looking for the line. There we go.
10:42:10 I feel happier when the line closes.
10:42:13 Even though the segmentation might not be as good.
10:42:19 No.
10:42:21 You should have let, let me put the Zodiac. Because that's so virgo.
10:42:34 No, I can't find it.
10:42:38 Almost done.
10:42:48 yeah.
10:42:50 Now, I'm also already going off a bit of the previous thing. And the problem is the previous thing also was done by me, so it's not like 100%, I guess.
10:43:03 That's not what segmentation is about.
10:43:09 Yeah.
10:43:11 Tiny bit here.
10:43:18 I'm deciding the gun is done.

B.2.2 Novice 2

11:39:34 I originally found the gun.
11:39:38 Before it crashed. Okay, here is the gun, in this area here.
11:39:46 I'm gonna start my segmentation here okay I start it here.
11:40:00 This is part of the gun.
11:40:07 Okay seems like I'm not very good with the.
11:40:08 Okay.
11:40:18 Why
11:40:18 Okay let's try to delete it now.
11:40:25 Just click to delete.
11:40:30 Yeah.
11:40:35 Okay.
11:40:47 All right.
11:40:51 This is part of the gun, let's try not to have an extra line so we don't have to

delete anything

11:41:09 Deciding where to double-click is a bit confusing, or a bit tricky.

11:41:19 Maybe it's the remote control thing.

11:41:56 Alright.

11:42:12 So, segmenting the gun.

11:42:58 I'm not going to even try to delete it.

11:43:09 I can just leave it.

11:43:22 The gun is inside.

11:43:33 Let's hope that there was a snap.

11:43:50 Okay I think more of the gun is showing.

11:43:55 I think this is part of the gun.

11:44:00 But there's no way to be sure. That is the trigger.

11:44:05 Is that the trigger.

11:44:09 No.

11:44:10 No.

11:44:14 Let's leave it, until I'm hundred percent sure it's the trigger.

11:44:20 Why is this part of the gun showing already.

11:44:43 Okay. At least now. I think this is the trigger hole

11:44:49 Okay. Here is the gun.

11:44:57 Now I need to also mark the empty space.

11:45:36 Okay I still can't tell where is the trigger so I'm not gonna segment it. I think the ropes are sort of confusing.

11:45:56 Confusing. They are confusing me.

11:46:04 Still can't tell 100% if this is the trigger.

11:46:07 Actually I can check from the right view.

11:46:13 Let me check here.

11:46:17 Let's have a frontal view. We fix the frontal view. Okay, here we go.

11:46:28 Then,

11:46:38 We go in. And trigger is there.

11:46:38 The ropes are really making it difficult.

11:46:57 Okay I'm going to assume this is the trigger now.

11:47:11 I don't know if that segmentation makes sense. Whatever.

11:47:24 Okay the gun is already disappearing without me seeing the trigger. The ropes are really blocking the view.

11:47:55 I don't know if the inaccuracy is because I'm bad at this. Or because I'm hungry.

11:48:03 Almost there, almost there.

11:48:21 The gun is appearing here.

11:48:38 I think it's appearing here.

11:48:41 I think it's appearing here.

11:48:46 Nothing from the barrel. There is nothing from the barrel which does not exist in any other ship.

11:48:55 Okay, let's look at it from here.

11:49:07 Okay.

11:49:10 Oh that's why the spot appeared earlier. Okay.
11:49:19 Now it makes sense.
11:50:35 This is part of it maybe.
11:51:00 Well, you get the idea.
11:51:30 Back to the ropes. Oh yeah, these are the ropes.
11:52:00 Alright, alright, alright.
11:52:06 You know what. I need to save the progress.
11:52:17 Just in case I crash it again.
11:52:44 Okay, now we have two parts.
11:53:04 Alright. Almost there. Appearing.
11:53:20 Now my fingers are hurting.
11:53:25 Not even my fingers that I'm using for the clicking.
11:53:32 Oh, no. There is nothing here worth.
11:53:36 Okay, let's check.
11:53:37 Okay, I think I'm done.

B.2.3 Novice 3

12:13:48 So step one is in the 3d view I'm trying to locate the gun.
12:14:10 Oh, I see yes, sorry I'm just blind. So it's located sort of behind.
12:14:25 One of the tube things. Ok good.
12:14:25 I'm currently looking at what seems to be a mess of cables.
12:14:48 So now I'm trying to zoom through the picture and see if I can figure out when I'm looking at.
12:14:54 But what side of it I'm on, is currently it's unclear to me, whether
12:15:02 It would be helpful if I could see like a representation of where my slice is in 3d space.
12:15:20 Oh my Lord. I'm sorry.
12:15:43 Okay, cool. Sorry. Every now. I think you just have a slight input lag and you get slightly disoriented if you don't know what's going on.
12:15:56 I'm still scrolling through everything.
12:16:10 Sorry,
12:16:16 So just to confirm, on your side, am I actually interacting with something because on my side everything it seems to be frozen.
12:16:34 Okay.
12:16:44 Okay so trying to figure out the little tubes.
12:17:19 So we expect the gun to show over here on the right hand side.
12:17:26 Since. That depends on the orientation which is a bit unclear
12:17:38 Okay. We are looking at what seems to be a mess of something. That's good I guess. It means we're getting closer to this like chunky bit at of back.
12:17:51 Maybe some things appear at the top. Okay, good. So that's presumably a glue container maybe.
12:18:16 Sorry, I'm just lost.

12:18:49 I see okay cool. So if you lose the weird glue container to the thing at the top, as a point of reference.

12:18:55 Then, as we zoom inwards for lack of a better word.

12:19:05 Yeah okay, So we're looking at it.

12:19:09 Using the left most representation, we seem to be looking at it from the top down, because the goal.

12:19:24 Why do I not get to know where I am in 3d space. It strikes me as a needless complication.

12:19:46 So as we zoom in the right paint shaker.

12:19:52 Whatever it is comes into view here good. Which means we expect the other one to show up on the left here momentarily.

12:20:00 Yes. Good.

12:20:03 The weird glue thing starts crawling upwards because its place at a slight angle. Yes, good. And so, right after the in this white representation the left paint shaker thing disappears.

12:20:21 We now would expect the gun to appear. I think.

12:20:30 Yeah, so now we see this weird swirl of intestines show up which I assume is this cable mess here, we see this block, which could be.

12:20:44 Not sure what that can be

12:20:48 It's just.

12:21:05 Okay wait so now we see like this weird shadow I think show up which could be the gun.

12:21:12 Maybe

12:21:17 shape-wise.

12:21:32 Yes So shape-wise the fact that the gun is seemingly placed at a slight angle relative to this like scan plane seems to cause some wackiness I think.

12:21:33 Yeah, so the butt of the gun shows up first down here

12:21:42 I say segmenting it very poorly.

12:22:10 I don't know whether this is correct.

12:22:19 If it isn't, I'm sorry,

12:22:28 I'm also encouraged to sort of like voice my thought process. This is hell of tedious.

12:22:40 I feel like. I mean I don't know where they're sort of even just like a simple contrast based detection system will help speed up this process. I don't know.

12:22:53 Oh, sorry.

12:23:05 So, I'm just going to quickly double check.

12:23:10 Double checking all this effort already submitted something yeah I know exactly but this makes sense the fact that we still have the weird cap of the paint shaker, whatever in view.

12:23:21 While we're looking at the butt of the gun does seemingly make, sorry.

12:23:27 Sorry that was a terrible, terrible line.

12:23:32 One more time.

12:24:16 Really hope I don't have to see the ultimate product of my segmentation at the

end. Oh Christ this looks terrible

12:24:32 It is very strange to be forced to look at the, the 2d slice representation of an object whos shape you completely understand in 3d.

12:24:45 But looking at it's like sliced up representation in a way that the slice plane is not completely in line with its, excuse me

12:24:58 the way you expect, makes this like really counter-intuitive. It's kind of cool.

12:25:41 Now, finally a little confirmation that we haven't just been looking at ghosts the entire time. Because now we see our guns, what I'll call it surely will show back up.

12:25:43 And I don't know what belongs to here. This is quite difficult to disentangle from whatever is next to it.

12:25:52 I see there's.

12:25:54 Yes, I think, I think we were right.

12:26:03 Wait a moment that doesn't seem correct

12:26:10 Let's pulled it down.

12:26:12 Ah ha Wait a moment. Yes, I am a fool.

12:26:21 I've been ignoring the fact that of course the weird barrel chamber gun bullet thing would be still

12:26:32 In each of our slices so that means this weird like growth here, I think, is actually part of the gun.

12:26:51 And the fact that it was like.

12:26:53 So I had this come from.

12:26:56 And then here.

12:27:01 This is still part of it.

12:27:09 Let's delete, this.

12:27:14 Yes I can

12:27:18 Mister software.

12:27:31 I don't know if this makes sense.

12:27:32 So hard to gage.

12:28:27 So I think some of the barrel so I'd be

12:28:29 some of the bits, where the bullets actually come out of the end.

12:28:32 I didn't segment 100% completely because it wasn't quite clear what was just like background noise and what was part of the gun

12:28:49 It's also not super clear whether the trigger is located here.

12:28:58 And it's not clear whether I'm.

12:29:22 Handling this, is so awful.

12:29:38 Sorry. That's unintentional.

12:29:43 Also, for some reason I do lose control over my full click.

12:29:54 I don't know if it's just like a lag, which causes Mister program to interpret that I have released left mouse button

12:30:04 that is slightly frustrating

12:30:08 I shall repeat again, for the record, I have no clue where the trigger is.

12:30:19 Essentially, there are like two lines where I'd expect one, because there's only one trigger.

12:30:25 But somehow,
12:30:51 I don't know.
12:30:43 So I'm just like screwed. So I am just zooming on just to make sure I can see the light at the end of the tunnel.
12:31:27 Hello.
12:31:44 Oh God, I've turned into a ghost again.
12:31:55 So here, I'm also not sure anymore.
12:31:59 There sort of our guns starts and ends.
12:32:03 This is a sort of growth, for lack of a better word, here.
12:32:21 I don't know if that's like a piece of a gun or not.
12:32:25 Looking at like this piece of rope.
12:32:28 Yeah, there's like this piece of rope that passes right by our trigger.
12:32:34 So I'm going to make the executive decision and claim that no, this is not part of our gun it's just a piece of string that our backpacker is somehow carrying with him as well.
12:32:49 To presumably hijack a plane
12:32:57 Yeah, I don't know without.
12:33:07 Okay, let's see. I think.
12:34:14 It turns out I'm submitting something completely wrong.
12:34:18 Like I'm actually just looking at.
12:34:35 Okay.
12:34:49 Also I'm going to shield, my
12:35:01 and justify the fact that I'm so terrible at like hand drawing this under the guise of overlap is just to that.
12:35:10 It's not all my fault but I can't draw.
12:35:26 So I think.
12:35:30 So unless I'm segmenting something which really isn't the gun
12:35:33 then I would argue the most challenging part was first finding out where I was in 3d space. And I was only looking at the 2d representation of the scene,
12:35:47 like in this case we have those conceptually, those like useful paint shaker landmarks, to sort of like figure out where we are, but at the same time the fact that the two of them, means that it's unclear.
12:36:03 From what side we're looking at this thing so trying to
12:36:09 figure out whether I'm traversing the backpack top to bottom or bottom to top, as I'm zooming through the scan is quite disorienting.
12:37:03 So I think I Slowly should be done. Slowly slowly.
12:37:07 I think,
12:37:11 you still see that slightly wonky shape.
12:37:21 In the,
12:37:20 the tip of the barrel so that seems to make sense.
12:37:32 But it's also now becoming increasingly unclear what parts are part of like the guns body and to what extent we're just looking at nothing.
12:37:51 So like we expect the rest of the body to eventually sort of phase out to disappear

12:38:01 to disappear and then only for the like yeah exactly here so now I think we still see the chamber the bullet thing, rotating thing.

12:38:12 Because that of course sticks out of the rest of the gun.

12:38:20 Yes.

12:38:18 That seems to make sense.

12:38:21 And then because the barrel, not the barrel, the shooty bits is on a slight angle.

12:38:31 We still see it as well.

12:38:34 Whereas in the previous scan where the gun was perfectly in line with the scan plane it disappeared much quicker now because of that angle it's still here.

12:38:48 That still makes sense.

12:38:52 Good.

12:38:55 Hot Dog, and now I just have the tip of the

12:39:01 barrel and the end of the bullet holder thing

12:39:11 very descriptive

12:39:15 then we have more bullets container and the end of

12:39:28 So there is a little bit left here.

12:39:38 and a little bit left here.

12:39:43 And then I think that's it. I think.

B.2.4 Novice 4

15:15:37 Okay, so I will start with this 3d view I'll try to rotate around will try to put the gun in the 3d for a better idea of what's going on. I mean, this is how I suppose Yeah, the back is oriented in comparison with the 2d view.

15:15:46 So I didn't see it yet.

15:15:50 Is this the? Ah there is the gun okay. So the gun is now positioned differently.

15:15:55 And my slices will be like this.

15:15:58 So that means when I go through them approximately in the middle I should spot the gun somewhere down here,

15:16:07 Ah there is little starburst there, which tells me that there's something metal and maybe that's the gun already.

15:16:14 And now I don't see anything. Ah there is the gun okay so it clearly wasn't the gun. So this very bright thing here, I can see that starts, and then it looks perfectly aligned with the slides.

15:16:27 I have to always think that I always see a section of it only.

15:16:31 And so, we can start. This is still the end of the butt of the gun. You can see it here. You can see it here now.

15:16:47 Now I move couple times up and down through the slides to see and try to identify where the gun ends and where it starts. And I think the last slide that I can see it, if I'm not lying to myself is this.

15:16:57 So I mark that.

15:17:12 And now I try to keep as much over my whatever temporal cone. And we come here to now something wrong and picking out little bit.

15:17:15 up. And yeah, I remember from the last slides. Where did I see the gun, so that I can, you know, continue to see it in the next slide and not lost.

15:17:27 the context because it's very difficult with all these crumbled walls around it. I mean. Good So, I delete.

15:17:38 and try again

15:17:43 and double-click.

15:17:45 Ah now I did, I did the scroll in the right direction and that made me happy a little bit, even though that's probably irrelevant to your study.

15:17:55 I think I'm getting better at scrolling the right direction.

15:18:03 I'm not very good at double-clicking.

15:18:16 So like I expected now it's going to be difficult because kind of like where the gun intersects with the plane.

15:18:23 It kind of fades away and so I'm really not exactly sure where it ends so this left side of the gun.

15:18:32 I'm just guessing like where exactly it should end. But I hope that's okay.

15:18:39 Filling this fade over here, kind of you know turns me out because I'm not really sure

15:18:47 until when I should mark so I just mark somehow in the middle of the gradient.

15:18:52 Oh, no, it closed on it's own. I hope that's okay as well.

15:18:57 Yeah, here it's getting really kind of fady.

15:19:02 So I just go like in the middle.

15:19:13 Oh and I just thought yeah like I'm kind of assuming that this would be like, continuing growing always this part of the gun but I also pay attention that some other parts may show up somewhere else on the slides.

15:19:29 Really so far like the gun is pretty simple shapes so I hope it didn't happen so far I might check.

15:19:35 First I finish drawing this guy.

15:19:42 Okay, so I'm not completely sure if I left something out now. This thing I don't know. This is a really bright spot. It probably doesn't belong to the gun I might check in the 3d view real quick.

15:20:00 That's probably already some weights,

15:20:02 some other stuff.

15:20:04 I mean there is this cylinder which shows the ammunition is inside out.

15:20:14 So I mark it like this.

15:20:16 So, now, I think this is the best part.

15:20:22 It's not totally there from the amount of damage. But it's separated once.

15:20:27 We are still here. And this part of the gun.

15:20:30 All.

15:20:33 Pay attention to that.

15:20:37 So, I might actually find it now before I get to all of this.

15:20:42 Ah yes, here it is. So this is the

15:20:49 It just popped into my mind so I wanted to check

15:21:08 I don't know what it's called so I'm just saying part.

15:21:24 It starts to curb a little bit.

15:21:45 Also, at some point, I will have to figure out where the trigger and the trigger guard is because it's also going to be pretty complicated in this angle.

15:21:56 So, I also might just do that real quick. After doing this slice like I'm much more comfortable at least I know where to expect stuff, and right now I don't know so I know about the cylinder that holds the bullets but I don't know about the trigger.

15:22:12 Okay, so this is a trigger guard over here, probably. I think. It's really difficult with the trigger.

15:22:17 There is this bright stuff over there.

15:22:24 In this slice it looks like two separate triggers.

15:22:29 And I assume from the size of it that the right one is actually the trigger and the other is just something else.

15:22:31 I am really not sure.

15:22:35 Yeah, especially because this thing that I thought was the trigger, it now separated from the gun.

15:22:40 So I'm going to guess where is the trigger.

15:22:49 Okay, so.

15:22:53 Ah, so this is the other stuff that wasn't the trigger

15:22:59 and that is the hold.

15:23:11 Super weird.

15:23:26 It still doesn't explain why here there is a gap that confuses me. But it might be that that's not the trigger anymore.

15:23:38 That there is just something touching the trigger from the other side, maybe it's the rope. Okay so at some point.

15:23:47 Ah at some point the trigger continues into the rope and. I am really confused as to

15:23:50 Why this has happened, but okay. Also I don't know from which side the slices are.

15:23:59 Okay so they are going from the bottom because the buck started over here.

15:24:05 I am going into the true direction. And so I see at some point I see rope. And then later I will see a trigger at the same slope.

15:24:19 Alright. I am not completely sure.

15:24:25 But I think this is too far. Okay it's pretty good.

15:24:41 Okay, so.

15:25:24 This thing over here, I don't think it belongs to the gun but then again I didn't see anything else on the scan so maybe then again I don't see these glows in the 3d as well.

15:25:39 So I just kind of ignore it and just like assume what the shape of the, of the grip is, because I think it's a bit too far out.

15:25:48 Maybe not.

15:25:51 Right. Maybe I'm wrong, I should have maybe marked this.

15:25:56 Yeah, I mean it's really unclear where exactly the gun ends in this part because it's bright, everywhere else.

15:26:04 Clear, but here these strips. So I kind of just ignore it.
15:26:09 Okay.
15:26:11 And we're again this is, I think, here it already starts my, my cylinder thingy.
15:26:17 So this is definitely the gun. This is the gun.
15:26:21 I think this is still the gun.
15:26:23 And here I'm not sure anymore.
15:26:25 And Here there is no gun.
15:26:27 So this is where it starts actually kind of like i mean it's it's a cylinder so you'll start as a triangle and then it will become a square later point, because the gun has this angle.
15:26:39 And here.
15:26:42 Yeah, I think this is already a gun so I'll just mark it. Maybe I'm wrong.
15:26:49 It looks like it's already the cylinder.
15:26:54 And I think this is already part of the trigger guards. And this might be also part of the trigger guard and this is the cylinder.
15:27:09 So this guy.
15:27:15 Now,
15:27:19 is like marking in one continuous stroke will be fine.
15:27:30 Okay,
15:27:34 so once again.
15:27:47 Good.
15:28:00 It's going to be much easier once these guys connect. I think this is also part of the gun can I also mark in the previous slides I'm gonna assume yes okay yes.
15:28:08 So this is also part of the gun pretty sure. This was difficult to mark because it's so small.
15:28:18 Okay.
15:28:26 Okay. So, now I see that this is also maybe a gun.
15:28:35 This is too far. This is definitely something the rope going into so maybe I was actually wrong with either.
15:28:35 But, yeah, I mean the trigger starts there so might be as well trigger.
15:28:46 Here we can see the rope crossing the trigger so it will kind of part of it so I'll mark around it so this is the cylinder thingy, if I'm not mistaken. This looks like they're only connected to the grip.
15:28:56 This looks like trigger guard. And here it ends. And that part over there is the rope.
15:29:02 trigger guards. And here it is hard, over there, if they're all.
15:29:10 Okay, and yeah I don't know really about the trigger I think it's not there yet, only here the trigger starts couple of slides down the stack so I'll just keep marking.
15:29:22 I'm going up and down the stack, a lot to kind of maintain the context, the 3d context of the gun because if I will be looking for one one slides for for very long time.
15:29:34 It would be very easy for me to kind of lost the context and start marking something else and then five slides down I would realize I'm actually marking some rope that sticks out of the gun that's why I'm doing this check regularly.

15:29:47 So here we have the gun that cylinder is are already connected to the rest of it. This part actually I don't know that goes up.

15:29:57 It looks pretty solid, but it could be also the rope. So for now, I will not mark it I will go around it and I will look at the 3d.

15:30:08 But I played myself here.

15:30:12 That's embarrassing.

15:30:13 Okay.

15:30:17 Let's see what I can do here.

15:30:20 I can just start over I guess.

15:30:27 Just in case I do like this this.

15:30:31 like this this. Okay, let me, let me just start marking the trigger guard separately.

15:30:39 Yeah, I think that could trigger guards.

15:30:49 This is the grip.

15:30:55 Here is the cylinder.

15:30:57 And this I believe it's not the gun

15:31:02 Oh no.

15:31:04 Okay, once again.

15:31:23 Now checking the 3d again, but I believe it's a bunch of ropes that I are sticking out here.

15:31:30 Yeah, really hard to tell. So, I mean I don't, I don't see I don't from this image stack.

15:31:38 So here's going to be again complicated here at some point it will merge into the gun so, really no idea where, but like this round shape, like I don't see where that on the gun is more like, because square-ish rather, but I might be wrong. I don't know.

15:31:56 I'll just keep assuming that this part is the rope, and then I just have to decide where is the rope here. I think this is actually the gun already. This looks pretty angular over here.

15:32:06 But on this slice 191 I think, yeah that might be already the gun. But this slice 190 I think this is still the rope or something else. Okay and the trigger I think it's somewhere here.

15:32:23 So again 192 I remember. Okay, now i'm just marking.

15:32:25 So, here's the cylinder thingy.

15:32:28 The grid.

15:32:29 The guard.

15:32:35 Here I think in order to make like a full circle and go like that.

15:33:10 Yeah, here I have no idea where the trigger starts, but I think

15:33:15 Wait. Its angle is like that. We will cross.

15:33:24 Yeah we will see a sample in the cross section, where the trigger will separate from the gun. So this part isn't really trigger but it will not be connected with the gun.

15:33:35 I see the gun on these slides yet. So, I will mark

15:33:42 the inside of the trigger guard first.

15:33:47 And then I believe this part is the trigger.

15:33:53 And then again.

15:33:58 The rest of the trigger guard.
15:33:59 The grip.
15:34:00 And the cylinder.
15:34:01 And here yeah I don't know.
15:34:11 And now I think the trigger already connects to the body of the gun so I will mark the whole gun.
15:34:26 And then again, outside of the trigger guard.
15:34:47 Yeah, Here I think it's the same situation so the trigger is connected to the trigger guard to the body of the gun.
15:35:02 trigger guard.
15:35:14 Somewhere started seeing the barrel over here somewhere.
15:35:49 So.
15:36:11 The trigger guard maybe
15:36:21 that I would look in the previous slice. And this would be inside of trigger guard and this is the trigger.
15:36:24 But let me check with the previous slides now. Yeah that looks pretty coherent to me.
15:36:33 Yeah. And then on the next slide, the trigger that would be even less so maybe I'm still here, and this is not the gun anymore but this is still the trigger guard but we don't see the trigger anymore.
15:36:41 So I think this is just last part of the trigger. I will start there, because there I can just use my
15:36:47 mouse where I believe the trigger guard will end in the next slide.
15:36:53 And then I can mark from there because I can't see anything in this area. The rest of it is kind of clear. This is pretty clear.
15:37:02 Also this part. And here again it starts to be a little bit unclear to me
15:37:08 triggered guard, just like a stop. Last time that we've seen. If I go through slides Looks kind of coherent now. looks like a child I think that's as good as it gets. Ah, and we are seeing the barrel now. The barrel over here.
15:37:33 So it's pretty clear in most of the slices device and it has kind of like this kind of shape.
15:37:37 And that is it.
15:37:43 This whole thing is the gun.
15:37:53 This was actually part of the barrel already I believe, so I should mark it probably.
15:37:55 Also here but this is so, so, thin that I will not even try to mark it over there, but I think this is already part of the barrel over there.
15:38:11 This is such a small part but I will mark it anyway.
15:38:30 And here.
15:38:39 All right, and here this is already the barrel.
15:38:46 The grip.
15:38:49 And here the rest of the body of the gun.
15:38:59 Here again, to barrel starts to be pretty visible but still it's under the angle so

it's a weird shape.

15:39:12 Right now I kind of realized.

15:39:21 Again, a thought that just came to my mind that I've never seen a hammer of the gun, so I have to check. And it is actually there, and I totally ignored it. Maybe it didn't come yet so I will check it now. Because on none of the previous slides I drew it, 15:39:36 but now I see okay so this little thing is the hammer. So I luckily didn't miss it.

But I have to keep it in mind.

15:39:43 Now this thing goes through a nice curve.

15:39:55 But it will probably be broken by I think it's this slide.

15:39:59 Yeah so on the previous slide I didn't mark it, so I'll mark it.

15:40:07 Alright so there is the hammer. Another thing to mark would be the aiming thing, but I don't see it so this gun doesn't have it or the CT didn't pick it up, so I can assume the gun is pretty flat

15:40:19 when I am outlining the other slides so these slides should always be just straight

15:40:25 Alright, so we have a trigger oh sorry the hammer. So now it's clear to mark it together with the rest of the gun.

15:41:28 At some point I will probably will strike, where the grip ends, because then in that area over here, it's kind of unclear. I'm not sure if, and 3d cannot really help me because I don't see all this glow in it, whatever this folded stuff is.

15:41:35 So I'll again try to do the thing where I go through couple slices and try to maintain this temporal vision of how the gun ends so here I see this, this part kind of advancing, which is the, where we go through the gun.

15:41:51 And on this slice it's still like curved little bit so we see more of the, of the handle, but here we don't anymore so it starts at some point in the slice it starts to be kind of straight.

15:42:06 Also, I'll check, since I'm checking Now, where does the trigger end. I that's pretty clear Okay that will be very visible for me, not really the hammer.

15:42:15 Okay, so let's continue with marking.

15:42:21 I Now, keep starting in this area. Which would be the least clear area for me so it will help me to see the previous slides and start somehow, where I see the gun is.

15:42:36 Yeah, it's much easier to see when I do it immediately like the rest of the gun is pretty clear so there is no issue there.

15:42:58 I think this is one of the last slides that we will see the hammer, then it's going to be easier to draw.

15:43:07 Yeah, and here it's not at all, so.

15:43:10 Now the shape of the gun is very simple.

15:43:32 Which is a bad thing because when it's simple

15:43:35 I get very lazy and very comfortable and then my outlines aren't going to be as nice, but I'll try to prevent that, I think this is what happened over there, even though I don't know why, but it's really like different than to be in.

15:43:53 In this area over here so it looks like they're is a step.

15:44:06 Or maybe it's the other side of the cylinder that we see Oh yeah, exactly. So this, the other side of the cylinder which was covered here by the body of the gun here it

starts to like kind of show through.

15:44:11 Okay.

15:44:14 I want to change this because it was an anomaly which I didn't know why it happened and if I'm just incorrect or, it's part of the gun but yeah, it seems to be, this is how it's supposed to be.

15:44:28 Now we maintain this little thing or here is if it's part of the gun, it really looks like it's all mark it like that.

15:44:38 Now we really just see the top of the barrel and the cylinder.

15:44:54 Now I'm thinking whether I will ever get to see.

15:45:05 Whatever I drew, in 3d, and I hope I will, and I think that'll be fun.

15:45:17 Now the cylinder starts to be kind of separated from the rest of the gun which tells me we are almost done

15:45:23 but it's still possible, to draw in one stroke.

15:45:49 Okay now I will draw in separate strokes because the connection between the barrel and the cylinder is really thin.

15:46:06 Okay, once again.

15:46:49 Now, I'm just trying to keep both the cylinder and the rest of the barrel in my vision as I'm drawing so I don't lose this context for the next slice, because it gets really difficult to see between all the other stuff.

15:47:11 Here I don't see it anymore.

15:47:13 This part might be the cylinder. I think it is because of the shape, yeah should be triangle.

15:47:22 The Barrel yeah I will get back to afterwards. So this I believe was the last slice with the cylinder.

15:47:27 And the barrel would be somewhere here. I'll mark it.

15:47:32 Well, I'm not really sure which part is actually the barrel.

15:47:37 And that's the end of it.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Data Processing Scripts

C.1 Process Logfiles

```
1 import pandas as pd
2 from datetime import datetime, timedelta, time
3 import spacy
4 import gensim
5 from gensim.corpora import Dictionary
6 from gensim.models import LdaModel, CoherenceModel, LsiModel, HdpModel
7 import gensim.models.nmf as NmfModel
8 from gensim.models import nmf
9 from sklearn.feature_extraction.text import CountVectorizer
10
11 def tokenize(utterance):
12     utterance = str(utterance)
13     doc = nlp(utterance)
14     doc = [token.lemma_ for token in doc if token.pos_ == "NOUN" or token.
15           pos_ == "PROPN" or token.pos_ == "ADJ"]
16     return doc
17
18 print("processing data")
19
20 print("creating logfile")
21 logfile_raw = open("logfile_segmentation.log", "r")
22 logfile = open("logfile.csv", "w")
23
24 logfile_header = "Date,Time,Event,Position_X,Position_Y,Position_Z,Value,
25                 Segmentation,Clicked\n"
26 logfile.write(logfile_header)
27
28 for line in logfile_raw:
29     line = line[9:-3]
30     while (line.count(",") < 8):
31         line += ", "
```

```

30     logfile.write(line + "\n")
31 logfile.close()
32 print("logfile created")
33
34 print("aligning transcript")
35 transcript_raw = open("closed_caption.txt", "r")
36 logfile = pd.read_csv("logfile.csv", sep=",")
37 logfile.insert(9, column = "Text", value = "")
38 audio = ""
39 for index, timestamp in logfile.iterrows():
40     for text in transcript_raw.readlines():
41         if timestamp['Time'][:-4] == text[:8]:
42             audio = text[9:-1]
43             logfile.at[index, 'Text'] = audio
44             transcript_raw.seek(0)
45 logfile.to_csv("logfile.csv", sep=",", index=False)
46 print("transcript aligned")
47
48 print("align and extract keywords")
49 nlp = spacy.load("en_core_web_sm")
50 logfile.insert(10, column = "Keywords", value = "")
51 for index, row in logfile.iterrows():
52     text = logfile.at[index, "Text"]
53     keywords = tokenize(text)
54     keys = ""
55     if len(keywords) > 0:
56         for keyword in keywords:
57             if keyword != 'nan':
58                 keys += keyword + ","
59             logfile.at[index, "Keywords"] = keys
60
61 logfile.to_csv("logfile.csv", sep=",", index=False)
62 print("keywords aligned and extracted")
63
64 print("creating heatmap data")
65 logfile = pd.read_csv("logfile.csv", sep=",")
66 heatmap = logfile[['Position_X', 'Position_Y', 'Position_Z', 'Keywords']]
67 heatmap.insert(3, column = "accTime", value = 0)
68
69 format = "%H:%M:%S.%f"
70 for index, row in logfile[:-1].iterrows():
71     currentTime = datetime.strptime((logfile.at[index, 'Time']), format)
72     nextTime = datetime.strptime((logfile.at[index+1, 'Time']), format)
73
74     dt = nextTime - currentTime
75     addup = int(dt.total_seconds()/timedelta(microseconds=1000).total_seconds
76     ())
77     heatmap.at[index, 'accTime'] = addup
78
79 heatmap = heatmap.groupby(['Position_X', 'Position_Y', 'Position_Z']).agg({'
80     accTime': ['sum'], 'Keywords': ['sum']}).reset_index()

```



```

81 for index, row in heatmap.iterrows():
82     if row['Keywords'] != 0:
83         keywords = row['Keywords'].split(',')
84         keys = []
85         if len(keywords) > 0:
86             for keyword in keywords:
87                 if keyword != '':
88                     if not keyword in keys:
89                         keys.append(keyword)
90
91         words = ""
92         for key in keys:
93             words += key + ","
94         heatmap.at[index, "Keywords"] = words[:-1]
95
96 heatmap.to_csv("heatmap.csv", sep=",", index=False)
97 print("heatmap data created")
98 print("data processed")

```

C.2 Fill Gaps

```

1 import cv2
2 import matplotlib.pyplot as plt
3 import numpy as np
4 import os
5 from PIL import Image
6
7 print("filling segmentations")
8
9 for root, dirs, files in os.walk(r'segmentation_slices'):
10     for file in files:
11         if file.endswith('.tif'):
12             # load image
13             img = Image.open(os.path.join(root, file))
14
15             # threshold
16             thresh = np.array(img, dtype=np.uint8)
17
18             # apply close morphology
19             kernel = cv2.getStructuringElement(cv2.MORPH_ELLIPSE, (5,5))
20             thresh = cv2.morphologyEx(thresh, cv2.MORPH_CLOSE, kernel)
21
22             # get bounding box coordinates from the
23             # one filled external contour
24             filled = np.zeros_like(thresh)
25             contours = cv2.findContours(thresh, cv2.RETR_EXTERNAL, cv2.
CHAIN_APPROX_SIMPLE)
26             contours = contours[0] if len(contours) == 2 else contours[1]
27             if len(contours) > 0:
28                 the_contour = contours[0]
29                 x,y,w,h = cv2.boundingRect(the_contour)
30                 cv2.drawContours(filled, [the_contour], 0, 255, -1)

```

```

31
32         # crop filled contour image
33         result = filled.copy()
34         result = result[y:y+h, x:x+w]
35
36         # write result to disk
37         slicenumber = file[5:-4]
38         while len(slicenumber) < 4:
39             slicenumber = "0" + slicenumber;
40         cv2.imwrite(os.path.join(root, "segmentation_slices_filled\slic"
+ slicenumber + ".tif"), filled)
41
42 print("segmentations filled")

```

C.3 Clustering

```

1 import pandas as pd
2 import numpy as np
3
4 import matplotlib.pyplot as plt
5 from sklearn.cluster import KMeans
6 from sklearn import datasets
7 from sklearn.feature_extraction.text import TfidfVectorizer, TfidfTransformer
, CountVectorizer
8 import spacy
9
10 def tokenize(utterance):
11     utterance = str(utterance)
12     doc = nlp(utterance)
13     doc = [token.lemma_ for token in doc if token.pos_ == "NOUN" or token.
pos_ == "PROPN" or token.pos_ == "ADJ"]
14     return doc
15 nlp = spacy.load("en_core_web_sm")
16
17
18 df = pd.read_csv('logfile.csv')
19 df = df.drop(df[df['Event'] != 'Position changed'].index)
20 df.drop(labels=['Date', 'Time', 'Event', 'Value', 'Segmentation', 'Clicked',
'Keywords'], axis=1, inplace=True)
21 # calculate clusters and their centroids
22 np.random.seed(0)
23 clusters = 20
24 est = KMeans(n_clusters = clusters)
25 est.fit(df[['Position_X', 'Position_Y', 'Position_Z']])
26 labels = est.labels_
27 centroids = est.cluster_centers_
28
29 df.drop(labels=['Position_X', 'Position_Y', 'Position_Z'], axis=1, inplace=
True)
30 df['cluster'] = labels
31 df.loc[df['Text'].isnull(), 'Text'] = ""
32 df = df.drop_duplicates(keep='first')

```

```

33 df.to_csv('tfidf.csv', index=False)
34 df_grouped = df.groupby('cluster').agg({'Text':','.join}).reset_index()
35
36 df_grouped['Position_X'] = centroids[:,0]
37 df_grouped['Position_Y'] = centroids[:,1]
38 df_grouped['Position_Z'] = centroids[:,2]
39
40 # tokenize text into cleaned keywords
41 s = ", "
42 for index, row in df_grouped.iterrows():
43     text = df_grouped.at[index, "Text"]
44     tokens = tokenize(text)
45     df_grouped.at[index, 'Keywords'] = s.join(tokens)
46 df_grouped.drop(labels=['Text'], axis=1, inplace=True)
47
48 # count how many times a word appears in a cluster
49 vectorizer = CountVectorizer()
50 # matrix with amount times words appear per cluster
51 word_count_vector = vectorizer.fit_transform(df_grouped['Keywords'])
52
53 # train idf on word count matrix
54 tfidf_transformer = TfidfTransformer(smooth_idf=True, use_idf=True)
55 tfidf_transformer.fit(word_count_vector)
56
57 # train idf on clusters
58 count_vector = vectorizer.transform(df_grouped['Keywords'])
59 tf_idf_vector = tfidf_transformer.transform(count_vector)
60 feature_names = vectorizer.get_feature_names_out()
61
62 # compute tfidf scores per cluster
63 for index, row in df_grouped.iterrows():
64
65     df_tfidf = pd.DataFrame(tf_idf_vector[index].T.todense(), index=
66     feature_names, columns=['tfidf'])
67     df_tfidf = df_tfidf.sort_values(by=['tfidf'], ascending=False)
68
69     keywords = df_tfidf[df_tfidf['tfidf'] > 0].index
70     df_grouped.at[index, 'Keywords'] = ",".join(keywords)
71
72 # this document saves the tfidf weighted keywords in descending order
73 df_grouped.to_csv('tfidf.csv', index=False)

```



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Figures

1.1	Segmentation of the pelvis and inguinal of three different patients (a, b, c) by seven radiation oncologists (blue and orange) and an expectation maximization algorithm (red and green), leading to inter-observer variability [9].	3
2.1	Anatomical planes and axes [12].	6
2.2	Modality classifications according to Elmogy et al. [13]	6
2.3	CT scanner of model SOMATOM X.ceed by SIEMENS [15].	7
2.4	Segmentations of a lung tumor as delineated by 10 different observers [18]. Each color represents one observer's delineation.	9
3.1	Segmentation results (right) stemming from human foot MR visualized in two dimensions (left) [20].	12
3.2	Interactive visualization of organ segmentation using YaDiV [22].	12
3.3	Initial slice by slice segmentation (left) and its 3D visualization (right) [26].	13
3.4	Tool developed by Amabili et al. with the main window for 2D and 3D exploration on the left and the provenance graph on the right [11].	14
3.5	Dashboard to analyze provenance data proposed by Blaschek et al. [33] .	15
3.6	Visualization of interaction log with details on demand as proposed by Guo et al. [35]	15
3.7	Fan et al. visualization of think-aloud protocols using machine learning models and color-coded timelines [36].	16
3.8	Eye-Tracking data visualized with attention maps in two dimensions (left) and scanpaths in three dimensions (right) [37].	16
3.9	Eye-Tracking data visualized with a space-time cube representing the development of traffic jams in San Francisco [38].	17
3.10	2D visualizations of variable data by Steenbergen et al. [3] (left) and Jungo et al. [8] (right)	18
3.11	3D overview visualizations by Reiter et al. [6] (left) and Grossmann et al. [5] (right).	19
3.12	Shape variations of bladders from different determined clusters by Furmanová et al. [45]	19
4.1	Schematic illustration of the methodology applied in this thesis divided into three steps.	21
		133

4.2	Variability of a segmentation originating from two observers (green, blue) on a two-dimensional slice extracted from a CT scan. Data from <i>TCIA</i> [18].	23
4.3	Eye-Tracking device [49].	26
4.4	Raw segmentation data (left) and segmentation data with interpolation between slices (right). Data from <i>TCIA</i> [18].	27
4.5	Logic of the computation of the ground truth comparison (false-positives, true-positives, false-negatives).	28
4.6	Eye-Tracking data visualized using heatmaps [50].	29
4.7	Eye-Tracking data visualized using Focus Maps (left) and Gazeplots (right) [50].	30
4.8	Mouse trajectory visualized using circles at every logged position [51] for a binary choice (between A and B).	30
5.1	Logging tool interface created with <i>MeVisLab</i> [52]. Blue: Volume View. Yellow: Slice View. Purple: Save Progress Button.	35
5.2	Logging scheme in <i>MeVisLab</i>	37
5.3	Data Acquisition setup with the administrator's PC and the participant's PC and the respective open programs.	39
5.4	Slice View of the Segmentation task with a delineation of the target object (red) for experts (left) and for novices (right).	40
5.5	Results of the confidence survey of all nine participants.	43
5.6	Results of the confidence survey of all nine aligned in parallel coordinates. Full lines represent experts. Dotted lines represent novices.	44
5.7	Open segmentation curve on the gun, highlighted with the yellow rectangle.	44
6.1	MOLE layout overview. Render settings (green). Render window (purple). Detail windows (yellow).	48
6.2	Overlay of two segmentations using the marching cube representation for both segmentation (left) and a silhouette representation combined with a marching cube representation (right).	49
6.3	Piecewise colorscale for the heatmap component with linear interpolation.	50
6.4	Overlay of two heatmaps. The ground truth is displayed in gold.	51
6.5	Two mouse trajectories with different approaches: observing the area of the object to delineate only (left) and observing the entire scan (right).	51
6.6	Basic annotations positioned very close to each other (left) and clustered annotations spread across the scan (right).	52
6.7	Render window with a scan and the selected 2D slice visualized in 3D. . .	53
6.8	Folder and file structure for MOLE data.	56
6.9	Pipeline for the creation of the logfile file for one observer.	57
6.10	Pipeline for the creation of the heatmap file sorted by descending time spent on each visited voxel.	58
6.11	Pipeline for the creation of the tfidf file.	59
6.12	A segmentation with a gap (highlighted in yellow) being closed and its final binary mask.	59

6.13 MOLE Interface Overview. Data information (blue). Render settings (orange). Query settings (green). Render window (purple). 2D View or Slice view (yellow). Utterances (pink). Segmentation histogram (red).	60
6.14 MOLE Interface in Global Mode with the switch button (green).	61
6.15 Segmentation iso-surface, Segmentation Silhouette and Segmentation iso-surface combined with its Silhouette.	62
6.16 ground truth comparison display: false-positives (left), true-positives (middle), false-negatives (right).	63
6.17 Heatmap for one observer.	63
6.18 Mouse trajectory where the left key was clicked (left) and the whole mouse trajectory (right).	64
6.19 Basic annotations (left) and clustered annotations (right).	65
6.20 Hidden Functions: 2D View, 2D slice in 3D and Utterances.	66
6.21 Segmentation Histogram with selected bar and its respective display in the render window.	67
7.1 Initial MOLE Interface.	70
7.2 Scan of the patient (lightgrey) and the tumor as the target object in the right lung (gold) for the expert task.	71
7.3 The accumulated true-positive voxels of the expert observations.	71
7.4 The ground truth of the tumor laid over on the accumulated true-positive voxels of the expert observations. Circled in red: the part medial to the aorta.	72
7.5 The accumulated false-positives of all expert observers placed in comparison with the ground truth.	73
7.6 The accumulated false-negatives of all expert observers (left). False-negatives placed in comparison with the ground truth (right).	73
7.7 Analysis of false-negatives using the Slice View (located on the right). Circled in red: possible tumor tissue.	74
7.8 Segmentation outcome in combination with the according basic annotations for each expert observer.	75
7.9 Clustered Annotations for Expert 5.	76
7.10 Mouse Trajectories for Expert 3 (left) and Expert 4 (right).	77
7.11 Heatmap and segmentation silhouette of Expert 4.	78
7.12 Global heatmap of expert observers.	78
7.13 Agreement in the segmentation outcomes of all five experts.	79
7.14 Outliers in the segmentation outcomes (left) and the respective expert utterances of slice 76 (right).	80
7.15 Parts of a handgun (revolver) [76].	80
7.16 Scan of the backpack (lightgrey) and the gun as the target object (gold) for the novice task.	81
7.17 Global heatmap for novice observers from different perspectives.	81
7.18 Global clustered annotations for novice observers. The clusters display the three most important keywords.	82

7.19	Accumulated false-positive voxels for all four novice observers.	83
7.20	Segmentation outcome compared to the ground truth for each novice observer.	84
7.21	Mouse trajectory and false-negatives of Novice 1 with a part missed in the top of the cylinder circled in purple.	85
7.22	Segmentation silhouette and false-negatives of Novice 4.	85
7.23	Clustered Annotations for Novice 2 (left) and Novice 3 (right) combined with the ground truth.	86
7.24	Basic annotations of Novice 2 combined with their heatmap.	86
7.25	Heatmap of Novice 3 with its slice number 178 visualized in 2D and 3D. Marked in red: visible part of the cylinder.	87
7.26	Utterances of Novice 1 (left), Novice 2 (middle) and Novice 4 (right) for slice number 178.	87
7.27	Utterances of Novice 3 for slice number 178.	88
7.28	Voxels that were segmented by all four novice observers compared to the ground truth.	88
7.29	Voxels only segmented by one novice observer compared to the ground truth.	88

List of Tables

5.1	Data acquisition participants sorted by their status based on the experience in the applied medical field.	34
5.2	Logged information and their respective triggers.	37
5.3	Resulting data from the acquisition process, their respective source and filetype.	42
6.1	Visualization components displayed in MOLE, their respective color and opacity.	67
7.1	Confidence survey results for expert participants for the questions as listed in 5.4.3.	76
7.2	Confidence survey results for novice participants for the questions as listed in 5.4.3.	84



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Bibliography

- [1] T. Watadani, F. Sakai, T. Johkoh, S. Noma, M. Akira, K. Fujimoto, A. A. Bankier, K. S. Lee, N. L. Müller, J.-W. Song, *et al.*, “Interobserver variability in the ct assessment of honeycombing in the lungs,” *Radiology*, vol. 266, no. 3, pp. 936–944, 2013.
- [2] G. Ristovski, T. Preusser, H. K. Hahn, and L. Linsen, “Uncertainty in medical visualization: Towards a taxonomy,” *Computers & Graphics*, vol. 39, pp. 60–73, 2014.
- [3] P. Steenbergen, K. Haustermans, E. Lerut, R. Oyen, L. De Wever, L. Van den Bergh, L. G. Kerkmeijer, F. A. Pameijer, W. B. Veldhuis, F. J. Pos, *et al.*, “Prostate tumor delineation using multiparametric magnetic resonance imaging: Inter-observer variability and pathology validation,” *Radiotherapy and Oncology*, vol. 115, no. 2, pp. 186–190, 2015.
- [4] G. Ristovski, T. Preusser, H. K. Hahn, and L. Linsen, “Uncertainty in medical visualization: Towards a taxonomy,” *Computers & Graphics*, vol. 39, pp. 60–73, 2014.
- [5] N. Grossmann, “Pelvis runner - comparative visualization of anatomical changes,” Master’s thesis, Research Unit of Computer Graphics, Institute of Visual Computing and Human-Centered Technology, Faculty of Informatics, TU Wien, Favoritenstrasse 9-11/E193-02, A-1040 Vienna, Austria, aug 2019.
- [6] O. Reiter, “Comparative visualization of pelvic organ segmentations,” Master’s thesis, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11/E193-02, A-1040 Vienna, Austria, mar 2018.
- [7] B. Preim and C. Botha, *Visual Computing for Medicine (Second Edition)*. Boston: Morgan Kaufmann, 2014.
- [8] A. Jungo, R. Meier, E. Ermis, M. Blatti-Moreno, E. Herrmann, R. Wiest, and M. Reyes, “On the effect of inter-observer variability for a reliable estimation of uncertainty of medical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 682–690, Springer, 2018.

- [9] K. K. Kyung Su Kim, Kwang-Ho Cheong, “Interobserver variability in clinical target volume delineation in anal squamous cell carcinoma,” *Scientific Reports*, vol. 11, 2021.
- [10] A. Gegenfurtner and M. Seppänen, “Transfer of expertise: An eye tracking and think aloud study using dynamic medical visualizations,” *Computers & Education*, vol. 63, pp. 393–403, 2013.
- [11] L. Amabili, J. Kosinka, M. A. van Meersbergen, P. M. van Ooijen, J. B. Roerdink, P. Svetachov, and L. Yu, “Improving provenance data interaction for visual storytelling in medical imaging data exploration.,” in *EuroVis (Short Papers)*, pp. 43–47, 2018.
- [12] V. Richard, *Multi-body optimization method for the estimation of joint kinematics: prospects of improvement*. PhD thesis, 06 2016.
- [13] F. Ahmed Abd EL-Ghani, M. Elmogy, and A. Atwan, “Current trends in medical image registration and fusion,” *Egyptian Informatics Journal*, 10 2015.
- [14] R. UVA Health and M. Imaging, “Different imaging tests, explained,” 2019. <https://blog.radiology.virginia.edu/different-imaging-tests-explained/> Last accessed 2 October 2022.
- [15] SIEMENS Healthcare GmbH, “Somatom x.ceed mit myexam companion,” 2022. <https://www.siemens-healthineers.com/de/computed-tomography/single-source-ct-scanner/somatom-xceed> Last accessed 2 October 2022.
- [16] M. P. Starmans, S. R. van der Voort, J. M. Castillo Tovar, J. F. Veenland, S. Klein, and W. J. Niessen, “Chapter 18 - radiomics: Data mining using quantitative medical image features,” in *Handbook of Medical Image Computing and Computer Assisted Intervention* (S. K. Zhou, D. Rueckert, and G. Fichtinger, eds.), The Elsevier and MICCAI Society Book Series, pp. 429–456, Academic Press, 2020.
- [17] B. H. Meesters A.M.L., Ten Duis K., “What are the interobserver and intraobserver variability of gap and stepoff measurements in acetabular fractures?,” *Clinical Orthopaedics and Related Research*, vol. 478, pp. 2801–2808, dec 2020.
- [18] L. Wee and A. Dekker, “Data from head-neck-radiomics-hn1,” *Cancer Imaging Arch*, vol. 10, 2019.
- [19] A. Oulefki, S. Agaian, T. Trongtirakul, S. Benbelkacem, D. Aouam, N. Zenati-Henda, and M.-L. Abdelli, “Virtual reality visualization for computerized covid-19 lesion segmentation and interpretation,” *Biomedical Signal Processing and Control*, vol. 73, p. 103371, 2022.
- [20] M. Strzelecki, P. Szczypinski, A. Materka, and A. Klepaczko, “A software tool for automatic classification and segmentation of 2d/3d medical images,” *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 702, pp. 137–140, 2013. PET/MR and SPECT/MR: New Paradigms for Combined Modalities in Molecular Imaging.

- [21] G. Floros and B. Leibe, “Joint 2d-3d temporally consistent semantic segmentation of street scenes,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2823–2830, 2012.
- [22] F.-E. W. Karl-Ingo Friese, Philipp Blanke, “Yadiv—an open platform for 3d visualization and 3d segmentation of medical data,” *The Visual Computer*, vol. 27, 2010.
- [23] S. Pruggnaller, M. Mayr, and A. S. Frangakis, “A visualization and segmentation toolbox for electron microscopy,” *Journal of Structural Biology*, vol. 164, no. 1, pp. 161–165, 2008.
- [24] A. Saad, G. Hamarneh, and T. Möller, “Exploration and visualization of segmentation uncertainty using shape and appearance prior information,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1366–1375, 2010.
- [25] G. Groenendaal, M. R. Moman, J. G. Korporaal, P. J. van Diest, M. van Vulpen, M. E. Philippens, and U. A. van der Heide, “Validation of functional imaging with pathology for tumor delineation in the prostate,” *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 145–150, 2010. Selected papers from the 10th Biennial ESTRO Conference on Physics and Radiation Technology for Clinical Radiotherapy.
- [26] T. S. M.S.M. Rahim, Alireza Norouzi, “3d bones segmentation based on ct images visualization,” *Biomedical Research (India)*, vol. 28, pp. 3641–3644, 01 2017.
- [27] C. W. Kai Xu, Alvitta Ottley, “Survey on the analysis of user interactions and visualization provenance,” *Computer Graphics Forum*, vol. 39, no. 3, pp. 757–783, 2020.
- [28] M. Wohlfart and H. Hauser, “Story Telling for Presentation in Volume Visualization,” in *Eurographics/ IEEE-VGTC Symposium on Visualization* (K. Museth, T. Moeller, and A. Ynnerman, eds.), The Eurographics Association, 2007.
- [29] M. Waldner, S. Bruckner, and I. Viola, “Graphical histories of information foraging,” in *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, pp. 295–304, 2014.
- [30] C. He, L. Micalef, L. He, G. Peddinti, T. Aittokallio, and G. Jacucci, “Characterizing the quality of insight by interactions: A case study,” *IEEE Transactions on Visualization and Computer Graphics*, 2020.
- [31] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang, “Recovering reasoning processes from user interactions,” *IEEE Computer Graphics and Applications*, vol. 29, no. 3, pp. 52–61, 2009.

- [32] P. Nguyen, K. Xu, A. Wheat, B. Wong, S. Attfield, and B. Fields, “Sensepath: Understanding the sensemaking process through analytic provenance,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, 09 2015.
- [33] S. K. Tanja Blascheck, Markus John, “Triangulating user behavior using eye movement, interaction, and think aloud data,” ETRA '16, (New York, NY, USA), p. 175–182, Association for Computing Machinery, 2016.
- [34] H. Stitz, S. Gratzl, H. Piringer, T. Zichner, and M. Streit, “Knowledgepearls: Provenance-based visualization retrieval,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 120–130, 2018.
- [35] H. Guo, S. R. Gomez, C. Ziemkiewicz, and D. H. Laidlaw, “A case study using visualization interaction logs and insight metrics to understand how analysts arrive at insights,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 51–60, 2015.
- [36] J. Z. Mingming Fan, Ke Wu, “Vista: Integrating machine intelligence with visualization to support the investigation of think-aloud sessions,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 343–352, 2020.
- [37] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl, “State-of-the-art of visualization for eye tracking data.,” in *EuroVis (STARs)*, 2014.
- [38] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko, “Stacking-based visualization of trajectory attribute data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2565–2574, 2012.
- [39] V. Navalpakkam, L. Jentzsch, R. Sayres, S. Ravi, A. Ahmed, and A. Smola, “Measurement and modeling of eye-mouse behavior in the presence of nonlinear page layouts,” in *Proceedings of the 22nd International Conference on World Wide Web, WWW '13*, (New York, NY, USA), p. 953–964, Association for Computing Machinery, 2013.
- [40] B. Lelandais, I. Gardin, L. Mouchard, P. Vera, and S. Ruan, “Dealing with uncertainty and imprecision in image segmentation using belief function theory,” *International Journal of Approximate Reasoning*, vol. 55, no. 1, pp. 376–387, 2014.
- [41] L. Linsen, A. Al-Taie, G. Ristovski, T. Preusser, and H. K. Hahn, “Uncertainty and reproducibility in medical visualization,” in *4th EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization*, pp. 1–3, Eurographics Association, 2016.
- [42] J.-S. Prassni, T. Ropinski, and K. Hinrichs, “Uncertainty-aware guided volume segmentation,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1358–1365, 2010.

- [43] A. Saha, M. R. Harowicz, and M. A. Mazurowski, “Breast cancer mri radiomics: An overview of algorithmic features and impact of inter-reader variability in annotating tumors,” *Medical physics*, vol. 45, no. 7, pp. 3076–3085, 2018.
- [44] R. G. Raidou, F. J. J. Marcelis, M. Breeuwer, E. Gröller, A. Vilanova, and H. M. M. v. d. Wetering, “Visual Analytics for the Exploration and Assessment of Segmentation Errors,” in *Eurographics Workshop on Visual Computing for Biology and Medicine* (S. Bruckner, B. Preim, A. Vilanova, H. Hauser, A. Hennemuth, and A. Lundervold, eds.), The Eurographics Association, 2016.
- [45] K. Furmanová, N. Grossmann, L. P. Muren, O. Casares-Magaz, V. Moiseenko, J. P. Einck, M. E. Gröller, and R. G. Raidou, “VAPOR: Visual Analytics for the Exploration of Pelvic Organ Variability in Radiotherapy,” *Computers & Graphics*, vol. 91, pp. 25–38, 2020.
- [46] B. Shneiderman, “The eyes have it: a task by data type taxonomy for information visualizations,” in *Proceedings 1996 IEEE Symposium on Visual Languages*, pp. 336–343, 1996.
- [47] GazeRecorder, “Gazerecorder - webcam eye tracking,” 2022. <https://gazerecorder.com/> Last accessed 17 October 2022.
- [48] EYEVIDO GmbH, “Eyevido - usability testing tool,” 2022. <https://eyevido.de/en/> Last accessed 17 October 2022.
- [49] Wikipedia, “Eye tracking,” 2007. https://en.wikipedia.org/wiki/Eye_tracking#/media/File:EyeTracker1.jpg Last accessed 17 October 2022.
- [50] J. Sauro PhD, “Essential eye-tracking visualizations and metrics,” 2016. <https://measuringu.com/eye-tracking/> Last accessed 18 October 2022.
- [51] Z. A., A. A., and P. P. T., “Decision landscapes: visualizing mouse-tracking data,” p. 12, 11 2017.
- [52] MeVis Medical Solutions AG, “Mevislab,” 2022. <https://www.mevislab.de/> Last accessed 1 September 2022.
- [53] Scientific Computing and Imaging Institute, “Open scientific visualization datasets,” 2022. <https://klacansky.com/open-scivis-datasets/category-ct.html> Last accessed 3 September 2022.
- [54] TurboSquid, “Turbosquid,” 2022. <https://www.turbosquid.com/> Last accessed 5 September 2022.
- [55] MeVis Medical Solutions AG, “Cso overview,” 2022. <https://mevislabdownloads.mevis.de/docs/current/MeVisLab/Standard/Documentation/Publish/Overviews/CSOoverview.html> Last accessed 3 September 2022.

- [56] Zoom Video Communications, Inc., “Zoom,” 2022. <https://zoom.us/> Last accessed 3 September 2022.
- [57] R. Likert, “A technique for the measurement of attitudes.,” *Archives of psychology*, 1932.
- [58] Google, “Google forms,” 2022. <https://forms.google.com/> Last accessed 5 September 2022.
- [59] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose, “Toolglass and magic lenses: The see-through interface,” in *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, (New York, NY, USA), p. 73–80, Association for Computing Machinery, 1993.
- [60] JetBrains, “Clion,” 2022. <https://www.jetbrains.com/de-de/clion/> Last accessed 8 September 2022.
- [61] Python Software Foundation, “Python,” 2022. <https://www.python.org/> Last accessed 8 September 2022.
- [62] B. L. Ken Martin, Will Schroeder, “Vtk,” 2008. <https://vtk.org/> Last accessed 8 September 2022.
- [63] Kitware, “vtk-examples,” 2022. <https://kitware.github.io/vtk-examples/site/> Last accessed 8 September 2022.
- [64] OFFIS, “DcmTk,” 2022. <https://dicom.offis.de/> Last accessed 8 September 2022.
- [65] The QT Company, “Qt,” 2022. <https://www.qt.io/> Last accessed 8 September 2022.
- [66] The QT Company, “Qt world,” 2022. <https://www.qt.io/qt-world> Last accessed 8 September 2022.
- [67] NumFOCUS, “pandas,” 2022. <https://pandas.pydata.org/> Last accessed 8 September 2022.
- [68] scikit-learn developers, “scikit-learn,” 2022. <https://scikit-learn.org/stable/> Last accessed 7 September 2022.
- [69] Explosion, “Spacy,” 2022. <https://spacy.io/> Last accessed 7 September 2022.
- [70] Alex Clark and Contributors, “Pillow,” 2022. <https://pillow.readthedocs.io/en/stable/> Last accessed 8 September 2022.
- [71] I. Sonix, “Sonix,” 2022. <https://sonix.ai/> Last accessed 7 September 2022.

- [72] OpenCV Team, “Opencv,” 2022. <https://opencv.org/> Last accessed 7 September 2022.
- [73] W. E. Lorensen and H. E. Cline, “Marching cubes: A high resolution 3d surface construction algorithm,” *SIGGRAPH Comput. Graph.*, vol. 21, p. 163–169, aug 1987.
- [74] S. D. Roth, “Ray casting for modeling solids,” *Computer Graphics and Image Processing*, vol. 18, no. 2, pp. 109–144, 1982.
- [75] G. Scholar, “Wouter van elmpt,” 2022. https://scholar.google.com/citations?user=rJ8_V2EAAAAJ&hl=en Last accessed 10 October 2022.
- [76] How Products are Made - Volume 1, “Revolver,” 2022. <http://www.madehow.com/Volume-1/Revolver.html> Last accessed 3 November 2022.