
Otto von Guericke University Magdeburg



Faculty of Computer Science
Department of Simulation and Graphics

Master Thesis

Visual Analysis of Brain Lesion Load in Patients with Cerebral Small Vessel Disease

Author:

Sarah Mittenentzwei

March 27, 2021

Advisers:

PD Dr.-Ing. habil.
Steffen Oeltze-Jafra

Department of Neurology
Otto von Guericke University
Medical Faculty
Leipziger Str. 44
39120 Magdeburg, Germany

Dr.-Ing.
Monique Meuschke

Department of Simulation and Graphics
Otto von Guericke University
Faculty of Computer Science
Universitätsplatz 2
39106 Magdeburg, Germany

Mittenentzwei, Sarah:

*Visual Analysis of Brain Lesion Load in Patients
with Cerebral Small Vessel Disease*

Master Thesis, Otto von Guericke University
Magdeburg, 2021.

Contents

Abstract

1 Introduction and Motivation

1.1 Motivation	1
1.2 Aim of this Thesis	2
1.3 Contributions of this Thesis	3
1.4 Structure of this Thesis	3

2 Medical and Technical Fundamentals

2.1 Fundamentals of Cerebral Small Vessel Disease	5
2.2 Visualization in Medicine	7
2.2.1 Direct Volume Rendering	8
2.2.2 Indirect Volume Rendering	10
2.2.3 2D Representations	10
2.3 Registration of Medical Image Data	11
2.3.1 MNI Space and Talairach Space	12
2.3.2 Brain Atlases	12
2.4 Detection and Segmentation of Brain Lesions	13

3 Related Work

3.1 Visual Analytics in Health Care	17
3.1.1 Medical Web Applications	21
3.2 Analysis and Visualization of Neurological Data	24
3.2.1 Brain Visualization Techniques	25
3.3 Research on Cerebral Small Vessel Disease	27
3.4 Summary and Requirements	29

4 Interactive Visual Analysis of Lesion Loads

4.1 Underlying Data Sets	31
4.2 Software	33
4.3 Visualizations and User Interface Components	33
4.3.1 Visual Analytics Concept and Workflow	36
4.3.2 3D Volume and 2D Slice Views	38
4.3.3 Abstract Visualization of the Lesion Load	43
4.4 Data Processing	46

4.4.1	Import of the Data	47
4.4.2	Accumulation of Lesion Data for the Cohort Creation	47
4.4.3	Creation of Surface Models	49
4.4.4	Parcellation of the Brain	50
4.5	Discussion of the Developed Visualizations	53
4.5.1	Comparison of Bullseye Plots with Alternative Visualizations	53
4.5.2	Comparison of Color-coded and Texture-coded Lesions	55
4.5.3	Challenges in Visual Analytics for Health Care	57
5	Evaluation	
5.1	Evaluation Setup	61
5.2	Discussion of the UI Elements and Visualizations	63
5.2.1	Volume Visualization	64
5.2.2	Abstraction as Bullseye Plots	65
5.2.3	Proposed Extensions and Modifications	65
5.3	Analyzed Patient Subsets	67
5.3.1	Comparison of CAA and mixed CAA	67
5.3.2	Comparison of mixed HA and mixed CAA	67
5.4	Conclusion	68
6	Conclusion	
6.1	Summary	73
6.2	Future Work	74
A	List of Figures	
B	Bibliography	

Abstract

Cerebral small vessel disease is known as the most common problem in aging brains, potentially leading to dementia, stroke, and cognitive decline. However, relatively little is known about the disease since the in vivo visualization of the small blood vessels of the brains is challenging. Instead, the occurrences of different lesions serve as markers for the diagnosis and classification of cerebral small vessel disease. Its most common subtypes are cerebral amyloid angiopathy and hypertensive arteriopathy, depending on if the lesions are located in lobar or deep brain regions. Identifying the subtype can give clinicians important clues about fitting treatment options. However, many cases of CSVD show mixed patterns that cannot be easily classified.

To support the further investigation of the lesion load patterns, this work presents an interactive visual application to investigate the lesion loads of a single patient or a cohort of patients. Furthermore, two sub-cohorts can be compared to visualize their differences. The application provides different views including the highlighting of the lesions in the MRI slices, 3D surface models of the lesions, and an abstract representation of the lesion load distribution using bullseye plots. While the 3D surface models and MRI slices allow for a detailed analysis of single lesions, the abstract representation allows for a broad overview and fast comparison of the lesion distribution.

This work was developed based on data from the WMH Segmentation Challenge 2017 and a clinical CSVD cohort from University Hospital Magdeburg. Using the *Think Aloud method*, the developed program was tested and evaluated by three experts. In this way, further suggestions for additions to the created visualizations were collected, which can be the subject of future work. Furthermore, the lesion loads of ten cases of the CSVD cohort were analyzed. The results have been shown to be consistent with previous studies. It was notable that some patients had an asymmetric lesion load in the hemispheres. Asymmetry of lesion load could be expected in rare inflammatory forms of CSVD but has not been thoroughly investigated yet. By using the presented application in combination with a larger data set, more insights could be gained about these cases.

1

Introduction and Motivation

1.1 Motivation

A variety of lesions related to the small blood vessels in the brain are referred to collectively as signs of Cerebral Small Vessel Disease (CSVD). CSVD is the most common aging brain problem and can lead to stroke, dementia, and mortality (SHAABAN and MOLAD, 2020). However, with standard imaging techniques used in clinical practice, it is not possible to visualize the vessels themselves. Instead, the focus is on lesions related to CSVD, like white matter hyperintensities, cerebral microbleeds, and enlarged perivascular spaces (WARDLAW et al., 2013). The localization of these lesions is important for medication decisions, e.g., the use of antithrombotic agents which can prevent stroke but at the same time increase the risk of bleeding (WANG et al., 2014).

Based on brain magnetic resonance imaging (MRI) scans, the lesions are often classified in the two most common CSVD subtypes: cerebral amyloid angiopathy (CAA) which appears only in lobar and hypertensive arteriopathy (HA) which appears mostly in deep gray matter. However, about 40% of participants showed lobar as well as deep lesions, and therefore, cannot be assigned to one of the two groups (PASI et al., 2017; SMITH et al., 2010). Further studies investigated these mixed cases with regard to their tendencies towards either CAA or HA (PASI et al., 2017; SCHEUMANN et al., 2020; TSAI et al., 2019). These studies are based on the analysis of the distribution of the lesions as well as demographics, clinical and genetic data.

1.2 Aim of this Thesis

This thesis aims to develop an application for an interactive visual analysis of lesion load in the brains of patients with CSVD. With regard to the division into CAA, HA, and mixed groups, the anatomical location of the lesions is of particular interest. This includes the relative location to each other as well as to the cortex. Lesions that occur lobar indicate CAA cases, whereas lesions in deep regions are associated with HA. In mixed cases, lesions occur both deep and lobar. The focus of this thesis will be on the three lesion types white matter hyperintensities (WMH), enlarged perivascular spaces (ePVS), and cerebral microbleeds (CMB).

The developed methods should present the lesion loads of single patients and cohorts in an explorable way. Furthermore, visualizations for the comparison of two cohorts will be developed. For this purpose, the lesions must be presented in the spatial context of the brain and individual substructures.

Due to the complexity of the data the cohort representation can quickly become confusing due to the high number of lesions and the general complexity of the brain's morphology. Therefore, more abstract representations of the brain and its substructures should be developed and evaluated.

Finally, the methods and visualizations will be integrated into an already existing visual analytics framework of the working group "MedDigit Medicine and Digitization" at the Department of Neurology (KNEU). The framework provides views from information visualization, such as Scatter Plots and Parallel Coordinates, which are suitable for non-image data as well as attributes derived from image data. The integration will allow a joint visual analysis of image and non-image data from patient cohorts. To use the application efficiently, strategies for the definition of patient groups are developed.

1.3 Contributions of this Thesis

The following list describes the contributions of this thesis:

- **Development of an integrated interactive visual analytics platform for the analysis of brain lesions:** The presented application provides an integrated interactive approach for the visual analysis of brain lesions. Since the visual representations of the image data were integrated into an already existing framework for the statistical analysis of clinical data, the final tool allows a joint analysis of clinical and image data for a single patient, a cohort, or the comparison of two subcohorts. This opens up new possibilities in the research of brain diseases since neurologists are often lacking appropriate tools to analyze cohort data with respect to the image data.
- **Development of a 3D visualization of individual cohorts and comparison of subcohorts:** For the representation of the three-dimensional image data and lesion load, a combination of image slices, 3D visualizations, and interaction techniques were developed. Further details are provided on-demand by selecting the 3D lesions.
- **Linking abstract visualizations with volumetric image data:** An abstract 2D visualization using Bullseye plots allows for a more broader comparison of the lesion loads distribution in single patients and cohorts. The visualizations are linked to support an efficient exploration of the provided data. Selecting brain regions in the Bullseye plot highlights the corresponding brain regions in the image slices and 3D representation, which allows for a quick identification of related lesions. All visualizations support comparative views which highlight differences in the distribution of the lesion load among selected cohorts.

1.4 Structure of this Thesis

For an improved overview, the structure of the content of the thesis is described in more detail. In the following chapter medical and technical basics are described, which are essential for the understanding of the thesis topic. Related publications will be presented and the work on the project

will be described, with a focus on the conception and implementation of the visualizations and the user interface as well as the processing of the data. This is followed by a description of the evaluation and its results. In the last chapter, the results of my work are summarized and possible further developments are discussed.

- **Chapter 2** describes the fundamentals on which this work is based, focusing on medical and technical basics, as well as definitions of terms.
- **Chapter 3** deals with related work and studies. The focus is mainly on the analysis of CSVD and the application of visual analytics in health care.
- **Chapter 4** describes the underlying data sets and the methods used. In particular, the processing and visualization of the data are described. Furthermore, the structure of the user interface and the quantitative data to be extracted are discussed.
- **Chapter 5** discusses the results of the qualitative and quantitative analysis. Furthermore, the evaluation of the results by an expert is described.
- **Chapter 6** summarizes the work and concludes with a discussion of open questions and possibilities for further development.

2

Medical and Technical Fundamentals

This chapter contains definitions and methods that are fundamental to understand this thesis. At the beginning, the different types and characteristics of cerebral small vessel disease are discussed. Hereafter, an overview of the detection and segmentation of the described lesions is presented. The chapter concludes with the analysis of CSVD.

2.1 Fundamentals of Cerebral Small Vessel Disease

Cerebral small vessel disease refers to the sporadic occurrences of different lesions. It is the most common aging brain problem and can cause stroke, dementia, and mortality (SHAABAN and MOLAD, 2020). However, due to the inaccessibility of the small brain vessels in vivo, many aspects of the disease remain unclear.

The most common subtypes of CSVD are cerebral amyloid angiopathy (CAA) and hypertensive arteriopathy (HA) (CHARIDIMOU et al., 2015). Patients are diagnosed with CAA if their lesions are only located in lobar regions of the brain, while HA is associated mainly with deep gray matter lesions. In advanced cases of HA, the lesions can also appear in lobar regions. In 1995 the Boston criteria were introduced and modified in 2010, providing guidelines for the diagnosis of CAA (GREENBERG and CHARIDIMOU, 2018). However, it is unclear how to proceed in case of mixed deep and lobar lesions. In different studies, about 40% of the participants showed mixed characteristics of both CAA and HA (PASI et al., 2017; SMITH et al., 2010). Further studies tried to investigate these mixed cases, based on the analysis of the lesion location, demographics as well as clinical and genetic data (SCHEUMANN et al., 2020; TSAI et al., 2019).

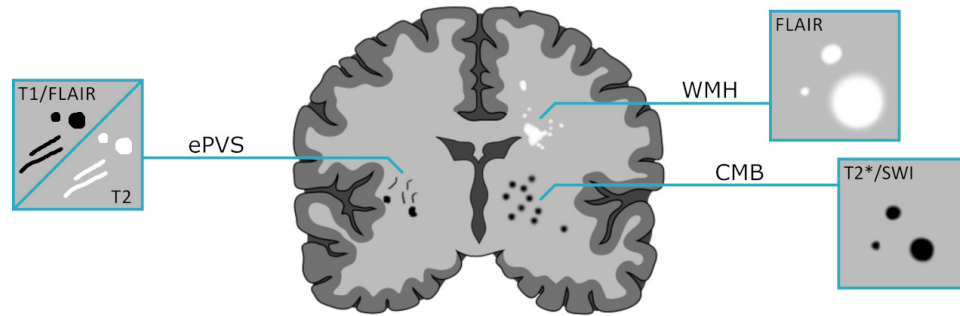


Figure 2.1: A schematic depiction of the different lesion types white matter hyperintensities (WMH), cerebral microbleeds (CMB) and enlarged perivascular spaces (ePVS) and their basic visual characteristics in different MRI techniques. The graphic is adapted from WARDLAW et al. (2013) and WEIDAUER et al. (2020).

The goal of these studies is to assign the mixed cases to one of the well-known subtypes CAA and HA, since the subtype has a strong influence on the treatment decision, e.g. the use of antithrombotic agents which can prevent stroke but at the same time increase the risk of bleeding (WANG et al., 2014).

In 2013, a collaborative work from several international researchers was published with the goal to standardize the research of CSVD under the title STRIVE (STandards for ReportIng Vascular changes on nEuroimaging) (WARDLAW et al., 2013). They defined six lesion types associated with CSVD: recent small subcortical infarct, lacune of presumed vascular origin, white matter hyperintensity of presumed vascular origin, perivascular space, cerebral microbleed, and brain atrophy. Due to the provided data, this work focuses on three of the named lesion types which are therefore described in more detail, see Figure 2.1.

- **White Matter Hyperintensities (WMH)** refers to bright white spots of varying size that can be observed in T2-weighted or fluid-attenuated inversion recovery (FLAIR) images when using MRI. They are caused by damaged white matter and associated with stroke, cognitive decline, and dementia (LIN et al., 2017). Since WMH can also be caused by diseases such as multiple sclerosis, often the term *white matter hyperintensity of presumed vascular origin* is used to specifically refer to WMHs associated with CSVD (DEBETTE and MARKUS, 2010).

- **Enlarged Perivascular Spaces (ePVS)** are associated with hypertension and stroke (POTTER et al., 2013). Perivascular spaces are fluid-filled spaces around certain blood vessels of the central nervous system. They are visible in T2-weighted and FLAIR images, their shape being either punctual or linear depending on the orientation of the image slices. They can be distinguished from other similar types of lesions, such as lacunes, by their small size, which is usually less than 3 mm in diameter but can also be up to 2 cm for giant perivascular spaces (WARDLAW et al., 2013).
- **Cerebral Microbleeds (CMB)** can be visualized using T2*-weighted magnetic resonance imaging. The amount of CMBs that can be detected is depending on different MRI characteristics as well as image post-processing (GREENBERG et al., 2009), their size varying between 2 mm and 10 mm in most cases. They are associated with surrounding tissue damage and therefore cognitive dysfunction and dementia (POELS et al., 2012).

While the three named lesion types are associated with CSVD, their concrete pathogenesis remains unclear (LIN et al., 2017). The location of the lesions, whether they are more lobar or deep, also influences the impact on the motor and cognitive abilities of the affected individuals, as well as treatment decisions based on the classification into CAA and HA (CHARIDIMOU et al., 2015; POELS et al., 2012). Furthermore, the occurrences of the lesion types correlate with each other as can be seen in different studies (CHOWDHURY et al., 2011; DOUVEN et al., 2018; POTTER et al., 2013). Therefore, the analysis of multiple lesion types associated with CSVD with respect to their co-occurrences can not only lead to new insight about their correlation but also allow a more precise investigation of the cerebral small vessel disease.

2.2 Visualization in Medicine

Medical visualizations are widely used nowadays and help with diagnosis, treatment planning, and navigation during surgical interventions (PREIM and BOTHA, 2014b). Even though the small cerebral vessels can not be visualized with standard imaging techniques used in clinical routine, MRI

brain scans are used to analyze the lesions associated with CSVD. There are different ways to visualize volumetric medical image data the two major rendering modes being direct and indirect volume visualization.

However, because medical volume data is often complex and suffers from self-occlusion, neurologists often choose to view the single volume slices instead of generated 3D representations. For investigating a cohort that consists of more than a few cases this approach is not feasible. To overcome this problem, more abstract 2D representations of the data are used. They trade a loss of information for reduced complexity.

2.2.1 Direct Volume Rendering

In direct volume rendering (DVR), three-dimensional objects are represented by grids with corresponding values at the nodes. Many medical imaging methods, e.g. computer tomography and MRI, directly provide such volumes, mostly in the form of slice images. Each value is virtually mapped to physical quantities that describe the emission, absorption, and in some methods also the scattering of light at that point. The volume rendering pipeline specifies the operations and the order they are performed in to generate the 3D image. It often consists of the three steps *sampling*, *classification and illumination* and *compositing*, but DVR techniques differ in the performed operations and their order (PREIM and BOTHA, 2014a). During the volume rendering pipeline, the data set can be traversed either front-to-back or back-to-front. Front-to-back methods are performed pixel by pixel, while back-to-front methods are performed voxel by voxel and each cell is projected onto the image. The mapping is usually done by a transfer function (LJUNG et al., 2016).

The advantage of direct volume rendering is the immediate access to the data and the resulting flexibility. This way, isosurfaces can be displayed without having to extract a geometric representation in advance. Furthermore, several isosurfaces can also be layered on top of each other or the volume can be visualized completely without taking surfaces into account.

Nevertheless, hardware-accelerated and software-optimized algorithms have been introduced that allow direct volume rendering in real-time (RODRÍGUEZ et al., 2014).

There are three main categories of DVR techniques:

- **Image-based** algorithms correspond to ray casting or ray tracing. Here, the contribution of the entire volume to the final color is determined for each pixel of the resulting image. This is done with the help of a ray that is laid through the individual pixels and sampled at discrete points. Ray casting is often performed front-to-back to allow accelerations, such as empty space skipping (WANG et al., 2012).
- **Object-based** methods calculate the contribution of each volume element to the final image, starting at a single voxel and proceeding iteratively. A popular example is the splatting method, which calculates the footprint of each sample after it is "*thrown*" on the view plane (WESTOVER, 1990). The accumulation can be done back-to-front, guaranteeing correct visibility of the elements or front-to-back, which is faster due to acceleration techniques like the early elimination of hidden splats (NEOPHYTOU and MUELLER, 2005).
- **Texture-based** methods exploit the general optimization of graphics hardware for 2D textures (KRUGER and WESTERMANN, 2003). A well-known method is slicing (LACROUTE and LEVOY, 1994). The textures of the volume that should be displayed are placed next to each other as parallel planes. The volume is traversed in a back-to-front manner and cross-sections between the volume and the planes are calculated. There are two types of texture-based volume rendering, 2D texture mapping and 3D texture mapping (ENGEL et al., 2006). In 3D texture mapping, the slice planes are oriented perpendicular to the viewing direction. An interpolation is performed in order to calculate the values for these planes. 2D texture-mapping aligns the slice planes parallel to the coordinate planes. It is suited for hardware that is not able to handle 3D textures. Each slice is rendered as a texture-mapped polygon, containing information about color and opacity, and then the polygons are drawn from back to front. To allow the user to view the volume from all directions, the orientation of the slice planes is adapted based on the viewing angle.

2.2.2 Indirect Volume Rendering

Indirect volume visualization requires pre-processing of the data which can be computationally expensive. This means that a surface extraction must be performed within the given scalar field, assuming that points belonging to a certain structure have the same isovalue. The resulting surface is then rendered, e.g. using the widely used Phong illumination (PHONG, 1975). This surface is referred to as an isosurface.

The most popular algorithm for extracting surface models is marching cubes (LORENSEN and CLINE, 1987). The core principle of marching cubes is to first divide the given voxel model of an object into small cubes and then *march* from one cube to the next to determine how the surface of the object intersects with each cube. Therefore, an isovalue is defined as a threshold. Then, for each cube, the sampled values of its eight vertices are compared to the threshold to check whether a surface corresponding to the isovolume runs through it. The polygons describing the intersection of the isosurface with the cube are generated using a lookup table that contains all 256 possible polygon configurations. The exact intersection point of the polygon with the cube's edge is then calculated by linearly interpolating the values of the two corresponding vertices.

The use of isosurfaces leads to a loss of information since the necessary restriction to surfaces means a serious reduction of the data (TOMANDL et al., 2001). If another isosurface is to be represented, preprocessing must take place again. In conclusion, indirect volume rendering is only applicable to data that can be classified by isovalues and allows only the visualization of a surface and not of the entire volume.

2.2.3 2D Representations

2D representations can reduce the complexity of medical image data and support the analysis, especially for larger patient cohorts. Furthermore, standardized representations like the bullseye plot for cardiovascular imaging allow a comparison of different studies (CERQUEIRA et al., 2002).

PREIM and BOTHA (2014c) classify 2D representations into *projections* and *reformations*. Projections accumulate or flatten the 3D data directly, while

reformations sample the data before flattening it. Based on their core principles, a further subdivision is made into the following three groups:

- **Anatomical unfolding** describes the process of unfolding a surface of an anatomical structure onto the view plane. The geometry depends solely on the anatomy and only measured data is used.
- **Anatomical planar reformations/projections** result in a flattened geometry that is anatomically-guided meaning that it contains only samples from the original data. The missing data points are reconstructed, e.g. by interpolation.
- **Map projections** include all techniques that project 3D shapes onto a 2D plane like it is done to map the globe to a geographic map.

The 2D representations differ in terms of distortion and the use of interpolated values. When deciding for a 2D representation these properties have to be carefully considered since distortions and interpolated values can lead to false conclusions. A survey of flattening-based medical visualization techniques has been performed by KREISER et al. (2018).

2.3 Registration of Medical Image Data

For the integration or voxel-wise comparison of multiple images, it is necessary to transfer them into a common coordinate system. This can be achieved by registering the images. Registration is the process of aligning different images of the same scene or similar objects (HAJNAL and HILL, 2001). One or more object images can be registered to a reference image. Therefore, a transformation is sought that deforms the object images, so that corresponding points are in the same place as they are in the reference image and the images become as similar as possible. In medical applications, registration is often used to match the data of different patients or to align images of the same patient that were taken from different positions or at different times (BEUTHIEN, 2014). Most registration algorithms consist of a transformation model, a correspondence basis, an optimization technique, and an interpolation method (GHOLIPOUR et al., 2007). Based on the algorithm used, the order of these steps can vary.

2.3.1 MNI Space and Talairach Space

The field of image registration is very large and many different techniques exist. In neurology, registration is often used to align brain imaging data in a common space (EVANS et al., 2012). Therefore, a set of standard spaces were developed to allow the comparison of the data between different studies. The two most popular standard spaces are the MNI space and the Talairach space (EVANS et al., 1993; TALAIRACH J, 1988). The Talairach space was created using only one brain, the so-called Talairach Brain, of a post mortem 60-year old woman while the first MNI space, was created by the Montreal Neurological Institute (MNI) using the average of 305 MRI scans and is therefore called MNI305. The MNI space has been established as an international standard by the International Consortium of Brain Mapping (ICBM). The current standard is the MNI-ICBM152, which is based on 152 brain scans that are registered to MNI305 space (MAZZIOTTA et al., 2001a,b). The newest versions are the MNI-ICBM152 Nonlinear atlases (FONOV et al., 2009, 2011). It is a standard MRI template brain based on the ICBM data and has a high spatial resolution as well as signal-to-noise ratio.

Even though the MNI/ICBM templates are originally based on the Talairach brain, they differ in size and shape, which has to be considered when comparing studies that used different spaces or when using maps such as the Brodmann areas. The Brodmann areas are based on the Talairach atlas, therefore a direct transfer of the coordinates of the MNI template to the Brodmann areas can lead to significant errors (BRETT et al., 2002).

2.3.2 Brain Atlases

To divide the registered brain data into different regions, brain atlases are used. They define individual brain regions and serve as a mask that can be applied to registered brain scans. Therefore, they are often used to estimate disease-specific changes of brain regions in clinical cohorts and to compare these cohorts to healthy patients (MANERA et al., 2020).

The Mindboggle-101 is the largest data set of publicly accessible, manually labeled human brain images, consisting of 101 entries (KLEIN et al.,

2017; KLEIN and TOURVILLE, 2012). It was created to serve as a brain atlas for labeling other brains as well as test and training data for the research on automated registration and labeling algorithms. The Cerebrum Atlas (CerebrA) for the state-of-the-art MNI-ICBM152 template was also based on the Mindboggle-101 data (MANERA et al., 2020).

There are two main approaches for cortical labeling, i.e., volume-based and surface-based. During volume-based labeling, the volume data is traversed slice by slice, and the brain regions are marked on each slice (SHATTUCK et al., 2008). Surface-based approaches use inflated or flattened surface meshes instead like it is done for the Mindboggle-101 data (DESTRIEUX et al., 2010).

In addition to cortical brain atlases, atlases for subcortical structures also exist (YOUSEFI et al., 2012). There are many different brain atlases developed for individual purposes, e.g. the analysis of brain activity, gray matter, or white matter. Therefore, an atlas has to be carefully chosen based on its properties that have to match the individual research question.

2.4 Detection and Segmentation of Brain Lesions

Brain lesions are still marked individually by an expert in many cases. This method is not only time-consuming, but it also suffers from inter- as well as intra-observer variability, making it hard to compare different cases (BALLERINI et al., 2018). Therefore, automatic approaches to detect and segment brain lesions are developed.

Different automatic segmentation approaches for white matter hyperintensities have been compared based on the submissions of the WMH segmentation challenge 2019 (KUIJF et al., 2019). The participants were asked to provide a machine learning solution for the segmentation of WMHs based on manually segmented training data. There were a total of 20 submissions, the ten highest ranked approaches all used (deep) neural networks. Furthermore, all the submissions using ensemble methods performed very well including the first placed. Therefore, it can be said that deep learning can be a very promising approach for developing software for the automatic segmentation of WMHs (GHAFOORIAN et al., 2017). Also, commercial software solutions like *mdbrain* are following this approach

to quantify lesions automatically (MEDIARE GMBH, 2021). *mdbrain* is using a deep neural network, trained with thousands of patients, and can even compare the data with previous images of the same patient, creating detailed reports for the user as well as a segmentation of the lesions.

The manual segmentation of cerebral microbleeds is especially time-consuming due to the morphological complexity and the sheer number of lesions. It can take 30 minutes for one observer to select all CMBs in one MRI scan (KUIJF et al., 2019). Therefore, ATEEQ et al. (2018) developed an automatic approach to detect CMBs by using ensemble-classifiers. They first identify potential candidates by a simple threshold and size-based filtering. False positives are then removed by performing a feature extraction and classification of CMBs. For the classification, support vector machines, Quadratic Discriminant Analysis, and the ensemble classifier RUSBOOST are used. All classifiers achieved a sensitivity of at least 90% with Quadratic Discriminant Analysis performing slightly better than the others. As for WMHs, there are also several approaches using convolutional neural networks to detect CMBs, their sensitivity is ranging from 90% , e.g., DOU et al. (2016) to 99%, e.g., HONG et al. (2018).

Similar to CMBs, enlarged perivascular spaces are rather small lesions. Therefore, ePVS are also often quantified manually by classifying them according to severity using a rating scale or by simply counting them (ADAMS et al., 2013; POTTER et al., 2015). To improve the situation, automatic scoring approaches suited for ePVS have been developed recently. DUBOST et al. (2019) compared two neural network-based approaches, one of them estimates the number of ePVS while the other also creates an attention map suited for detection and segmentation of the lesions. Attention maps represent the task-related activation in the input layer of a neural network (JETLEY et al., 2018). They can be used to highlight regions of special interest for the classification task at hand. Both methods showed similar results to the visual scores acquired by two observers. Neural networks are also suggested by SUDRE et al. (2019) for the detection of extremely small objects like ePVS.

Manual detection of lesions is increasingly being replaced by machine learning methods, in particular neural networks. The studies described show that neural networks can provide good results in the detection and segmentation of even extremely small lesions and therefore represent a

reduction in the daily workload of clinicians. While many approaches are still being researched, commercial applications like mdbrain exist and are used in clinical routine and research.

3

Related Work

The work presented in this section is intended to show related research and which findings, as well as problems, have emerged from it. In particular, work on the topic of visual analytics in health care, analysis of neurological data, brain visualization techniques, as well as further research on the cerebral small vessel disease, will be discussed.

3.1 Visual Analytics in Health Care

Medical data is often complex, very large with many missing values and varying quality. Therefore, visual analytics is suited for this application field, since the combined approach of automated data analysis and interactive visualizations allows for a very detailed and individual exploration and analysis of the data (KEIM et al., 2010, 2006). The principle is based on the four components data, models, visualization, and knowledge, see Figure 3.1. Data describes the input data, that can be transformed for further processing by the model or mapping to a visual representation. The model performs an automated data analysis based on parameters adapted by the user by interacting with the visualization. Therefore, user-steered changes in the visualization trigger the (re)calculation of the model, which in turn is followed by an update of the visualization. Based on the interactive exploration of the data, the user can get new insights and gain knowledge.

Visual analytics uses representations from the field of information visualization, where graphics are used to show trends, patterns, and correlations in, often large, data sets (DOS SANTOS and BRODLIE, 2004). The user can interact with the data through visual representations, e.g., by hovering or selecting elements. Multiple representations can be linked to allow

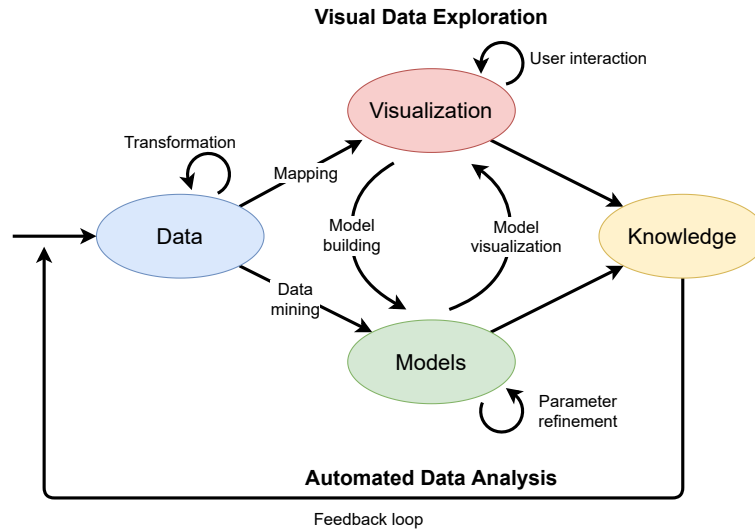


Figure 3.1: Depiction of the the visual analytics process by KEIM et al. (2010). It is characterized by the interaction between the four components data, models, visualization and knowledge, resulting in the combination of interactive visual data exploration and automated data analysis.

the user to investigate them at the same time, following the concept of multiple coordinated views (ROBERTS, 2007). The visual representations can present an overview over complex data sets as well as focus on details. Therefore, coordinated multiple views often follow Shneidermann's mantra "*Overview first, zoom and filter, then details on-demand*" (SHNEIDERMAN, 1996). In many cases, filtering the data to create a subset selection is realized by brushing over the corresponding elements. The other linked views are then automatically updated to show the selected subset. Therefore, this technique is called brushing and linking (ROBERTS, 2007). Following the mantra, the user can direct the analysis according to the individual task (KEIM et al., 2008).

Many graphical user interfaces are based on the model-view-controller (MVC) pattern, depicted in Figure 3.2 (KRASNER and POPE, 1988). As described by TOMINSKI and SCHUMANN (2020) the model consists of the data and the transformations applied to it. This can include statistical or classification analysis. The views include the visual representations and additional information shown to the user. The controller refers to the interaction concepts, e.g., sliders or brushing. When the user performs an interaction, an update request is sent from the controller to the model. After that,



Figure 3.2: A schematic representation of the model-view-controller pattern from TOMINSKI and SCHUMANN (2020).

the views are refreshed showing the updated results. The update process can consist of various different tasks, based on the application and user input, e.g., updating colors, redo a statistical analysis or cluster analysis.

The visual representations are provided to help the user to form mental models of data sets that are complex and therefore hard to grasp TOMINSKI and SCHUMANN (2020). Therefore, it is necessary to follow the human-in-the-loop approach and provide the user with tools to interactively explore and adjust the representations, as explained in this section.

Especially when working with medical data, many challenges arise. SHNEIDERMAN et al. (2013) formulated seven specific challenges in the development of visual analytics applications for health care:

- **Offering busy clinicians timely information in the right format:** In the busy clinical environment, there is often little time to manually analyze large amounts of patient data. Therefore, it is essential that applications summarize the data and present it so that clinicians can efficiently search for anomalies and important details.
- **Moving towards an ecosystem of visual tools:** With regard to the concept of patient-centered medical homes and the increasing shortage of family care doctors, it becomes more and more important to provide visual tools in different languages and for diverse platforms suitable for expert and non-expert users.
- **Facilitating team decision making:** Applications should support team decisions and the needs of many different professions, such as physicians from different disciplines, patients, and nursing staff.

- **Characterizing and understanding similarity:** When faced with difficult cases or rare diseases, many physicians have to refer to related publications to gain more knowledge about possible treatments. This process should be supported by applications that can browse medical literature sources to search for related cases.
- **Visualizing comparative relationships:** Comparative visualizations are essential to study the differences and similarities between different cohorts. This enables the evaluation of treatment methods as well as the evaluation of cause-effect relationships.
- **Presenting risk and uncertainty:** In medicine many treatments and medications have a low risk of complications and side effects. Interactive visualizations have a great potential to present these uncertainties in a comprehensive way.
- **Evaluation:** Errors made in the development of medical applications can have fatal consequences and poor visual designs could result in misunderstandings and inappropriate choices. Therefore, a detailed evaluation under realistic conditions as well as adapting the application according to the feedback from experts is necessary.

Despite these challenges, visual analytics has great potential to improve clinical routine and medical research. JÖNSSON et al. (2019) describe in their study how visual analytics can help in investigating diseases with unknown cause based on the irritable bowel syndrome which is assumingly related to one or more of the following: physical properties in the gut, central mechanisms in the brain, and psychological factors. They combine the analysis of a large range of different measurements from stool samples to MRI data to support scientific reasoning and hypothesis formulation. Visual analytics can also be beneficial for monitoring patient data in clinical routine. BADE et al. (2004) present an application to visualize high-dimensional and time-oriented data of patients in intensive care units. The application is suited for different users and tasks, especially for diagnosis and treatment planning. While it was applied to pulmonary embolism in the publication, the concept can also be used for other medical and non-medical approaches.

In recent years, mobile as well as web technologies have been increasingly used for the development of visual analytics platforms. Based on the principles of *Web 2.0*, the terms *Medicine 2.0* and *Health 2.0* were introduced, describing web-based and mobile applications for patients as well as professionals (EYSENBACH, 2008; HESSE et al., 2010). *Health 2.0* consists of the three main application and research fields personal health, public health, and clinical health (SHNEIDERMAN et al., 2013). Personal health focuses on the patient as a user for devices that allow monitoring the own health status and are intended for self-care (MEYER et al., 2016). On the other hand, public health refers to large databases that are often used by statisticians to investigate new trends and hypotheses on a large and often national scale. A survey about visual analytics in public health has been done by PREIM and LAWONN (2019). Clinical health describes the application field of professionals working in the medical sector. Systems in this area should help in diagnoses and treatment decision processes as well as allowing to investigate patient cohorts for research purposes. Due to the increasing amount of data, it can be difficult for physicians to get an overview of the clinical data of one patient, let alone a cohort. Thus, visual analytics in clinical health can provide a time-efficient solution for generating an overview of patient data as well as offering the possibility to analyze specific aspects in more detail. These opportunities and research challenges are further discussed in the work of CABAN and GOTZ (2015).

3.1.1 Medical Web Applications

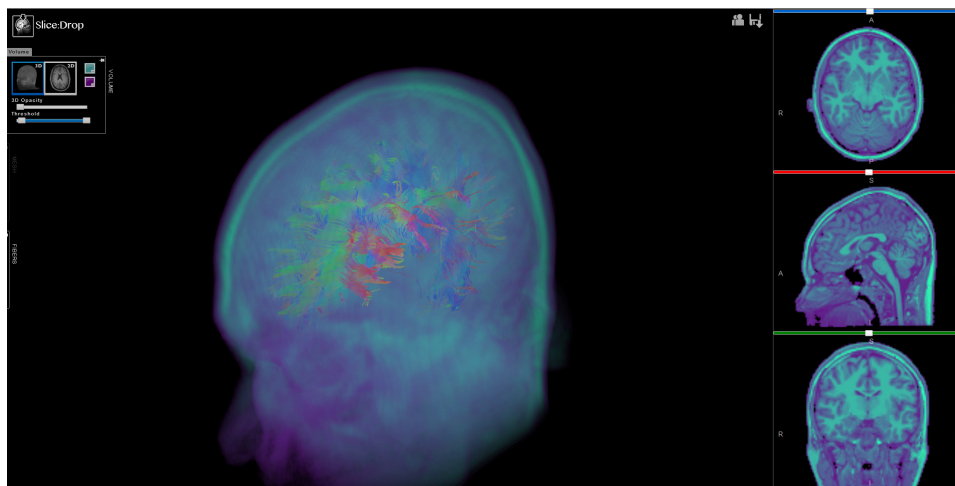
In recent years, web applications for medical analysis, as proposed in the concepts of *Health 2.0*, are gaining more and more popularity as the required technology is becoming more mobile, affordable, and powerful, e.g., tablets, laptops, and cell phones. Technologies like WebGL allow the easy representation and manipulation of 3D data in common browsers. Web-based applications open up new opportunities in science and research since they make applications widely available on the web and for collaborative work. Furthermore, they do not require complex setup processes. Since almost every computer has at least one common web browser installed, many devices meet the requirements regardless of their system specifications or operating system. Implementing the application as part of a client-server structure integrates well with the visual analyt-

ics process. While the client can handle the visualization and interaction interface, the server can take care of the modeling tasks which are more demanding in terms of processing power. Since the visual exploration of one calculated model takes time, a single server can support multiple clients instead of idling between the requests.

The project VICTORIA is a web-based application aiming at generating ground truth data for the ostium reconstruction of cerebral aneurysms (BEHRENDT et al., 2020). Five clinical cases can be segmented on the associated website. The ground truth data can then be generated from the resulting data. Due to the easy accessibility of the application by hosting it on a website, data can be collected from many different participants, regardless of their location and technical equipment.

HUTH et al. (2016) developed an interactive 3D viewer to analyze different brain regions with regards to brain activation based on semantics, see Figure 3.3c. The activity in different brain regions was measured for several subjects listening to different stories. Afterwards, each pixel on the surface model of the brain was colored based on the meaning of the words in the stories that caused the highest brain activity at this point. The tool allows to manipulate the 3D mesh by flattening and unfolding it.

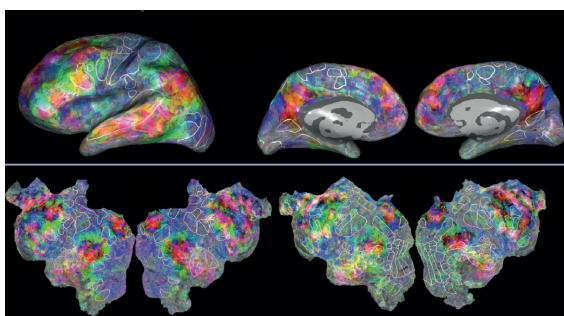
The X Toolkit (XTK) is a WebGL-based toolkit developed for JavaScript with the aim to provide a simple API for the generation of medical visualizations, especially in neuroscience, in the web browser (HAEHN et al., 2014). Several other toolkits exist but they are highly specified on particular anatomical structures (DA CRUZ et al., 2014; RUDOLPH, 2011). XTK is used for different applications, among them AneuryskWeb and Slice:Drop. AneuryskWeb allows the interactive visualization of the data of the Aneurysk database on a web page (ANEURISK-TEAM, 2012). Slice:Drop is a tool that can be hosted on a web page and allows the rendering and collaborative analysis of medical image data of many different formats, see Figure 3.3a (HAEHN, 2013). The data can be shared between different devices and multiple linked browsers can all manipulate the view to allow a cooperative analysis involving multiple users.



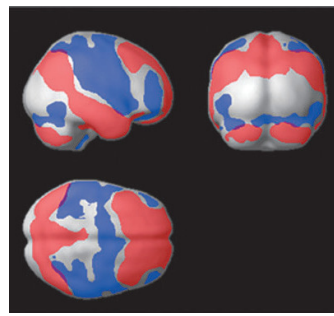
(a) The web application Slice:Drop from HAEHN (2013), visualizing a MRI scan using slices and DVR as well as a 3D surface model of the brain fibers.



(b) The application VisualNeuro by JÖNSSON et al. (2020) offers a workflow that combines statistical analysis with interactive visual exploration.



(c) Visualizations of principal components of voxel-wise semantic models using 3D surface models and flattened representations of the cortex. (HUTH et al., 2016).



(d) Visualization of group-level seed-to-voxel connectivity measures from the toolbox Conn (WHITFIELD-GABRIELI and NIETO-CASTANON, 2012).

Figure 3.3: Exemplary brain visualizations from recent related work.

3.2 Analysis and Visualization of Neurological Data

In the field of neurology, investigations are based on clinical measurement as well as brain scans (SIULY and ZHANG, 2016). The main focus of most applications is on a statistical analysis of this data like it is done by statistical parameter mapping (PENNY et al., 2007) or the connectivity toolbox Conn (WHITFIELD-GABRIELI and NIETO-CASTANON, 2012). While many other applications are focusing on the visualization of the results of statistical analysis only, Conn also includes the image data of the brain scans allowing the user to investigate relations between the clinical measurements and affected brain regions, see Figure 3.3d.

Another advanced tool is FreeSurfer (FISCHL, 2012). FreeSurfer offers a wide range of possible processing steps to view and transform brain data as well as the statistical analysis of the data. However, FreeSurfer is operated by using the command line, which makes it not suited for being used directly by physicians. Furthermore, most of the mentioned tools have long processing times which makes it very time consuming to analyze the data in an iterative way.

In turn, JÖNSSON et al. (2020) developed VisualNeuro, an application that allows an interactive analysis of clinical and scan data in real-time, see Figure 3.3b. VisualNeuro offers multiple linked views to select sub-cohorts and perform a statistical analysis of their data. Additionally, the image data is also included as 2D slices and a 3D model to show the user the correlation of spatial and clinical data as well as group differences between two sub-cohorts.

When developing these tools, analysts are not only confronted with the problems of combining clinical and imaging data but also with the different data types, clinical data consists of as well as missing data. Based on the dual analysis framework *Brushing Dimensions* from TURKAY et al. (2011), an application for the integrated dual analysis of quantitative and qualitative high-dimensional data was developed by MULLER et al. (2021). The application allows a combined statistical analysis of quantitative and qualitative data, while many other tools handle this data separately since many descriptive statistics are not suited for both data types. It was developed using the same clinical CSVD cohort, as used in this work.

3.2.1 Brain Visualization Techniques

Medical imaging is a key aspect in diagnosis and treatment planning. The resulting volume data can depict anatomical structures as well as other modalities like brain function. Over time larger databases of medical image data are acquired and can be used for cohort studies. As discussed in the first part of Section 3.2, it is necessary to include image data in applications that are used for the investigation of neurological diseases with regards to the location of lesions.

Often medical volume data is observed by going through the scan of a specific patient slice by slice. This process can be very time-consuming and is not fitted for the analysis of cohorts. Since the original data is three-dimensional, the most natural representation would be 3D views generated either by indirect or direct volume rendering, see Section 2.2. This also allows the inspection of the data from arbitrary directions. However, medical data is highly complex and suffers from self-occlusion. Nevertheless, it can be seen in Figure 3.3 that 2D slices, 3D models or a combination of both are used in many tools mentioned in Sections 3.1.1 and 3.2 (FISCHL, 2012; HAEHN, 2013; HUTH et al., 2016; JÖNSSON et al., 2020; WHITFIELD-GABRIELI and NIETO-CASTANON, 2012).

The brain is the geometrically most complex structure of the human body, therefore many studies use *brain flattening* to unfold the cortex onto a 2D plane to reduce complexity and occlusions (KREISER et al., 2018). There are different methods for flattening the brain, e.g., transform the brain into a sphere, circle, or plane. This can be applied to the whole brain surface as well as cortical sub-structures. In an approach to project the cortical surface of the brain onto a 2D plane GOEBEL (2000) first inflate the 3D model of the brain and then cut it into several pieces. The pieces can then be projected onto the plane and metric distortions are minimized in a final step. In contrast, HINDS et al. (2008) focus on projecting a specific sub-surface, the primary visual cortex, onto a 2D plane. Therefore, they first extracted the target structure by manually segmenting it in the data slices. On the resulting mesh, pairwise geodesic distances between all vertices were calculated. After that, the planar representation was generated by preserving the shortest distance for each pair while forbidding edge crossings, resulting in a quasi-isometric flattening of the sub-surface. Either way the re-

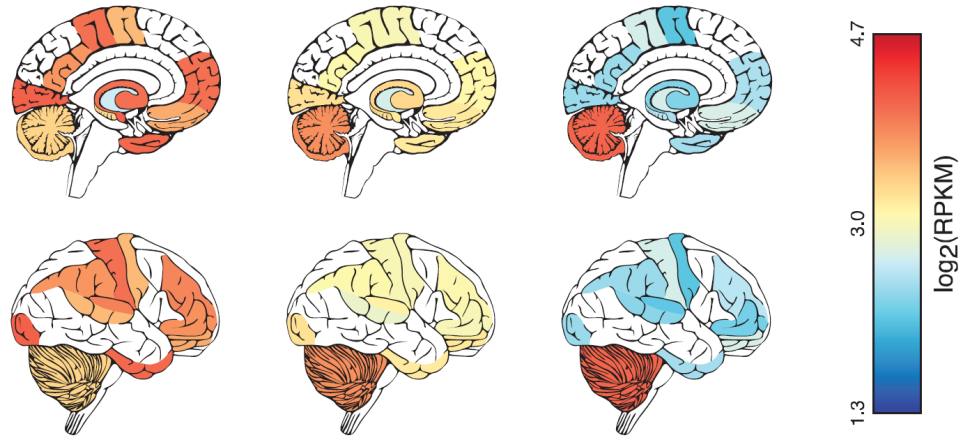


Figure 3.4: A 2D illustrative brain visualization from the cerebroVis toolbox showing the POGZ gene expression in the brain at three developmental time points (BAHL et al., 2016)

sults can be misleading not only due to the emerging distortions but also because points that are originally close together on the original brain surface can be far apart on the flattened result since the surface of the brain is strongly curved (WANDELL et al., 2000). Furthermore, brain flattening is only applicable if the focus of the analysis is on the cerebral cortex, since deeper structures can not be represented with these techniques, see Figure 3.3c.

As an alternative, 2D illustrations of the brain can be used. The different regions in the illustration are color-coded based on scalar values. The toolbox cerebroViz maps spatiotemporal brain data to color-coded vector graphics of the brain, see Figure 3.4 (BAHL et al., 2016). They illustrate the brain from different angles, providing, e.g. sagittal and exterior views. Even though these 2D visualizations do not suffer from self-occlusion, they also do not provide any depth information. Therefore, many individual illustrations of different cross-sections through the brain are needed to visualize information that is located in different depths.

More abstract diagrams from the field of information visualization are used to depict data in different brain regions in one plot. In a study from ROTH et al. (2014) bar charts with error bars were used to plot the brain activity per region. While all the data can be seen without any overlap, it is not possible to intuitively know which bar represents which brain re-

gion. Another approach is to not focus on brain regions but other data, e.g., lesions. Based on derived features from WMH lesions, like volume, shape features, and intensity profiles, JUNG et al. (2021) used the t-SNE algorithm to plot this data in a two-dimensional scatter plot. While the locations of the lesions themselves are not depicted in the plot, clusters of similar lesions can be easily recognized.

SUDRE et al. (2018) suggested a bullseye visualization that is more abstract but still considers the morphology of the brain. In medical visualizations, bullseye plots (BEP) are commonly used for the visualization of the heart, e.g., depicting infarcted myocardial tissue or perfusion data (OELTZE et al., 2006; TERMEER et al., 2007). The brain was divided into different lobes, each lobe being represented by one segment in the BEP. The lobes are further divided based on their distance to the ventricles. Each of the distance-based sections is represented in the BEP as one ring. The resulting cells in the BEP are then color-coded based on the WMH lesion load. The segments of the BEP are roughly arranged based on the position of the respective lobes in the brain, depicting the left and right brain structures in separate segments. The rings are showing the depth, thus the inner ring is representing the regions closest to the ventricles. The same representation was used by BRUGULAT-SERRAT et al. (2019) to depict the correlation between cognition and WMH lesions.

3.3 Research on Cerebral Small Vessel Disease

As described in Section 2.1, the concrete pathogenesis of the lesions associated with CSVD is still unclear. Since the small cerebral vessels cannot be depicted by current image modalities, the studies on CSVD rely on the presence of these lesions and their features (LIN et al., 2017). WARDLAW et al. (2013) introduced the *standards for reporting vascular changes on neuroimaging*, an approach to standardize research on CSVD and define the six lesion types associated with CSVD, see Section 2.1.

Furthermore, patients diagnosed with CSVD are typically assigned to one of the two common subgroups CAA and HA (CHARIDIMOU et al., 2015). Patients with CAA have strictly lobar lesions while the lesions of patients with HA are mostly deep. A popular approach for identifying CAA is the Boston criteria (GREENBERG and CHARIDIMOU, 2018). In the study of

SCHEUMANN et al. (2020) CAA was associated with multiple subcortical WMH spots as well as severe ePVS and ePVS predominance patterns in the centrum semiovale, while HA showed more WMHs in the peri-basal ganglia and severe ePVS in the basal ganglia. Concerning clinical data, patients with CAA tended to have a higher age than those with HA and arterial hypertension and diabetes mellitus were less frequent compared to mixed cases.

Belonging to one of the groups also influences medication and treatment decisions (WANG et al., 2014). Multiple studies show an association between patients using aspirin and the occurrence of lobar CMBs, which are associated with CAA (VERNOOIJ, 2009; WANG et al., 2014). In a study of WILSON et al. (2016) strictly deep CMBs (HA) showed a slightly lower risk of ischemic stroke while strictly lobar CMBs (CAA) have a higher risk of intracerebral hemorrhage. Therefore, for patients with CSVD the use of antithrombotic agents, which can prevent stroke but at the same time increase the risk of hemorrhage, has to be carefully evaluated.

Patients can also have a combination of lobar and deep lesions, making treatment decisions more difficult for clinicians, e.g., mixed cases show a rather high risk of both, ischemic stroke and intracerebral hemorrhage. (SMITH et al., 2010; WILSON et al., 2016). Mixed cases can be either a co-occurrence of both CAA and HA or an advanced stage of HA which can also include lobar lesions (PASI et al., 2017). Studies investigating these mixed cases try to classify the co-occurrences more precisely to be either mixed CAA or mixed HA based on their similarities in image and clinical data to the two main groups (SCHEUMANN et al., 2020; TSAI et al., 2019). In those studies, the mixed group showed more cerebellar CMBs and a higher frequency of anterior subcortical WMHs compared to CAA and more deep as well as cerebellar CMBs and higher occurrence of severe ePVS in the centrum semiovale compared to HA. In the work of SCHEUMANN et al. (2020), one-third of the patients with mixed cases showed ePVS predominance patterns and were therefore classified as mixed CAA. The other patients were classified as mixed HA. Furthermore, mixed CAA cases tended to have more lobar CMBs and less frequently WMHs located in the peri-basal ganglia.

A detailed description of CSVD, its causes, diagnosis, prevention, and treatment can be found in the book of PANTONI and GORELICK (2014).

3.4 Summary and Requirements

Due to the inability to visualize the smaller brain vessels that are affected by the cerebral small vessel disease with current standard imaging techniques, detection, analysis, and treatment decisions of the disease are based on CSVD-related lesions. While the parthenogenesis of these lesions remains unclear their localization is associated with specific symptoms and treatment outcomes. Therefore, CSVD cases are generally classified into the two subgroups cerebral amyloid angiopathy (CAA) and hypertensive arteriopathy (HA), based on whether the lesions are lobar or deep. However, many patients show lobar as well as deep lesions and are categorized as mixed-cases, leading to open questions concerning treatment decisions.

From the preceding discussion of the studies and consultations with Stefanie Schreiber, who is a senior neurologist at the University Hospital in Magdeburg, the following requirements result for this work:

- **Localization of the different lesion types:** The lesion types WMH, CMB, and ePVS should be visualized with respect to their anatomical localization. The localization of the lesions to each other as well as to the cortex is of interest.
- **Visualization of variable cohort sizes and counts:** The application should be able to visualize the lesions of one patient, a cohort of patients as well as a comparison of two cohorts.
- **Abstract visualization:** To allow a comparison of larger cohorts without occlusions that typically appear in 3D visualizations an abstract 2D representation of the brain and its substructures should be created. The 2D representation should visualize the distribution of the lesions instead of depicting each lesion explicitly.

4

Interactive Visual Analysis of Lesion Loads

In this chapter, an overview of the implemented application is given. At the beginning, the software used and the underlying data sets are discussed. This is followed by detailed descriptions of the development of the visualizations and the processing of the data. In addition, the problems encountered during the implementation and possible solutions are discussed.

4.1 Underlying Data Sets

The application was mainly developed using the data set of the WMH Segmentation Challenge (KUIJF et al., 2019) which consists of image data acquired in different hospitals in the Netherlands and Singapore. A total of 60 subjects are included providing T1-weighted and FLAIR images as well as manual segmentation results of the WMHs. For the final development and evaluation of the application the working group *Medicine and Digitalization* (MedDigit) at the Department of Neurology (KNEU) in cooperation with Prof. Stefanie Schreiber from the University Hospital of Magdeburg provided the following data acquired during clinical routine:

- MRI data from a clinical cohort of patients with CSVD,
- ePVS segmentation results of the cohort,
- WMH segmentation results of the cohort,
- CMB segmentation results of the cohort,
- Clinical data of the patients of the cohort.

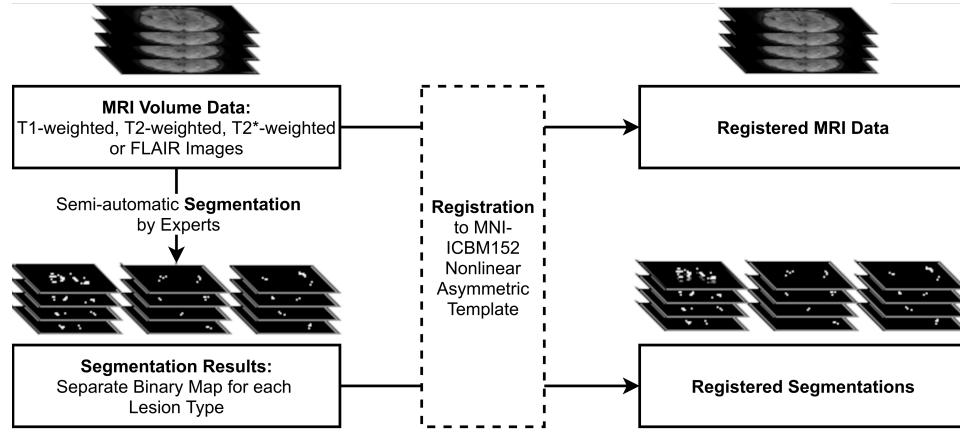


Figure 4.1: A schematic depiction of the preprocessing steps. The MRI scans are segmented by experts, the segmentation result for each lesion type is saved as a separate binary map. The MRI data, as well as the segmentation results, are then registered to the MNI-ICBM152 template.

For the cohort, either T1-weighted, T2-weighted, T2*-weighted or FLAIR images were available. Based on these, the WMH, CMB, and ePVS lesions were segmented semi-automatically by experts. In a study of LESJAK et al. (2017) the intra- and inter-rater variability of manual and semi-automatic segmentations of three experts were compared. The mean dice coefficients for the intra-rater variability was 0.85 for the manual and 0.92 for the semi-automatic approach. The mean inter-rater variability of the manual and the semi-automatic approach were 0.82 and 0.89, respectively. Therefore, semi-automatic approaches seem to reduce the intra- and inter-rater variability. For this work, a fully automated approach was not feasible, as most of the clinical data is of low quality and causes false segmentation results that have to be corrected by an expert. The image data were registered to the MNI ICBM 2009c Nonlinear Asymmetric template (COLLINS et al., 1999; FONOV et al., 2009, 2011) to allow the accumulation of lesions of multiple patients forming a sub-cohort, see Figure 4.1.

Due to the extensive preprocessing required, only a total of ten patients from the CSVD cohort were used for the final development stage and evaluation. In addition to the MRI data and the segmentation results of the sub-cohort, a CSV file with clinical data and the subdivision into the groups CAA, HA, mixed CAA, and mixed HA according to SCHEUMANN et al. (2020) was provided.

4.2 Software

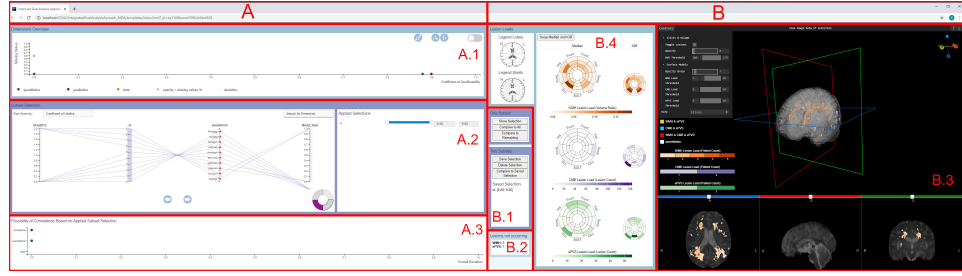
The application was implemented using a combination of JavaScript and Python 3.7 to build an interactive client-server web environment, see Figure 4.2. It was integrated into an already existing visual analytics framework of the working group MedDigit which is based on JavaScript, D3.js, and Python (MULLER et al., 2021). By using JavaScript, no additional software needs to be installed on the PC, as the program can be executed in the web browser, which allows an easy use in clinical practice. The whole application is running locally, meaning there is no network traffic. As a result, the patients' data is not leaving the currently used computer, ensuring data privacy.

Furthermore, the JavaScript libraries D3.js, XTK, and THREE.js were used for generating the visual representations. D3.js is a library for creating interactive data visualizations for the web. Therefore, it generates graphical elements from numerical values using web standards like CSS, HTML, and Scalable Vector Graphics (SVG) which are supported by common web browsers. XTK is a toolkit developed for scientific visualization in the web, as described in Section 3.1.1. It is based on WebGL and is mostly used for 2D and 3D visualizations in neuroscience. Like XTK, THREE.js is mainly based on WebGL. THREE.js allows a quick and easy creation of 3D models, textures, scenes, and lighting.

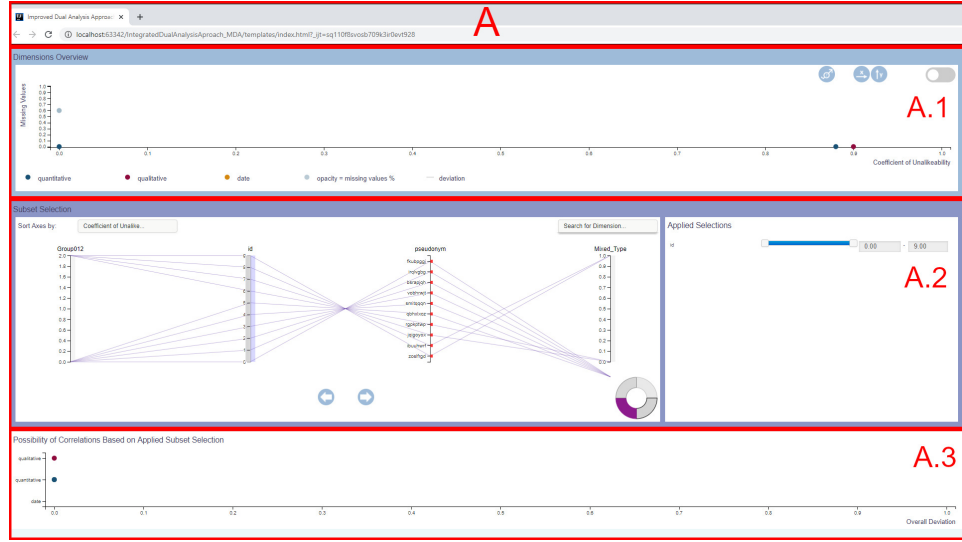
For the parcellation of the brain, the open-source software FreeSurfer was used (FISCHL, 2012). The FreeSurfer output was then processed with the bullseye pipeline from SANROMA (2019) which is further discussed in Section 4.4.4.

4.3 Visualizations and User Interface Components

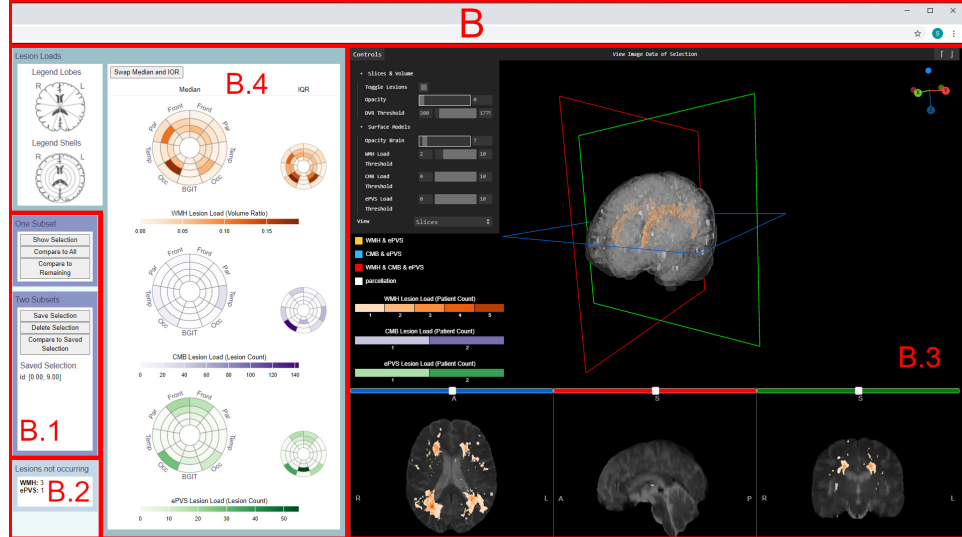
In this section, the different visualizations and components of the user interface (UI) are described. To allow for a parallel analysis of clinical routine data and image data the application was transformed into a two screen solution, showing the initial user interface on the left screen and the extension developed in this work on the right screen. An overview of the UI is shown in Figure 4.2.



(a) The complete two screen application.



(b) The left screen showing the initial framework.



(c) The right screen showing the extension of the framework.

Figure 4.2: A schematic depiction of the final UI. The left screen shows the original framework from MULLER et al. (2021) (A) while the right screen shows the extension developed in this work (B).

A The left screen is showing the initial framework, which performs a statistical analysis of clinical data loaded from a CSV file. It is divided into three plots, each plot is placed in a separate box with varying frame color, see Figure 4.2b.

A.1 A plot showing the statistical measures *coefficient of unlikelihood* (x-axis) and *relative frequency of missing values* (y-axis).

A.2 The parallel coordinates plot depicts each row of the imported CSV file as a dimension. Each row represents one patient and is drawn as a line in the plot. The user can select a subset of patients by brushing over one or more axes. The selection is depicted on the right and can be adjusted there. Afterwards, the statistical plots **A.1** and **A.3** are automatically updated.

A.3 The plot is depicting the possibility of correlations. Sorted by data type, it is the normalized aggregation of the deviations of all descriptive statistical measures.

B The right screen is showing the extension of the initial framework which was developed in this work, see Figure 4.2c. The screen is divided into the visualizations of the image data, the abstract representation of the lesion load, and interaction possibilities for the selection of one or more subsets. While the plots on the left screen update automatically after choosing a subset of patients, the right screen is only updated on demand by pressing the respective button. Generating and importing the image data needs more computation time and is therefore not suited for an automated update after each subset selection.

B.1 The buttons used to start the calculation of the visualizations on the right screen are placed in a box with a frame color similar to the parallel coordinates plot to indicate that both boxes are used to select and display patient subsets. The user can select a single subset and display it or compare it to all data sets including or excluding the selected subset. A subset can also be saved to define a second individual subset afterwards and compare both selections.

B.2 For each lesion type the amount of patients in the selected subset that do not have this lesion is displayed. Therefore, the user

can evaluate how many data sets are forming the accumulated lesion visualizations.

B.3 On the right side of the screen the image data is displayed as 2D slices, 3D surface models, or optional using DVR. An orientation gizmo supports the 3D rotation of the models and a menu allows several interaction techniques, which are further described in Section 4.3.2 .

B.4 Abstract representations of the lesion load of the selected subgroup are depicted as bullseye plots. They are further discussed in sections 4.3.3 and 4.4.4. For WMH the proportional volume of the lesion to the brain region is shown, while for CMB and ePVS the count of individual lesions per region is counted. To explain the abstract visualization to the user, a legend is displayed in the upper left corner.

The views **B.3** and **B.4** from Figure 4.2c are modified based on the selection of the user. The application supports a total of three different views, the inspection of a single patient, a cohort, and the comparison of two selected patients or groups. In the following sections, the visual analytics workflow (see Section 4.3.1) is described. The visual components of the UI are described in more detail (see Section 4.3.2 and 4.3.3), followed by the generation of the visual representations (see Section 4.4). At the end of this chapter, the proposed representations will be discussed and compared to alternative visualization styles, see Section 4.5.

4.3.1 Visual Analytics Concept and Workflow

The user starts the application, by running the python server and opening the HTML file in a browser window. Afterwards the data can be explored by selecting an individual subset in the parallel coordinates plot. The application is based on the *brushing and linking* approach (SHNEIDERMAN, 1996), therefore, a selection of a subset of patients in the parallel coordinates plot automatically leads to an adjustment in the other plots. The selection is performed by brushing over a subset of data entries on one or more axes.

Due to the higher processing time for the visualization of the volume data on the right, the BEPs, as well as the visualization of the volume data, are

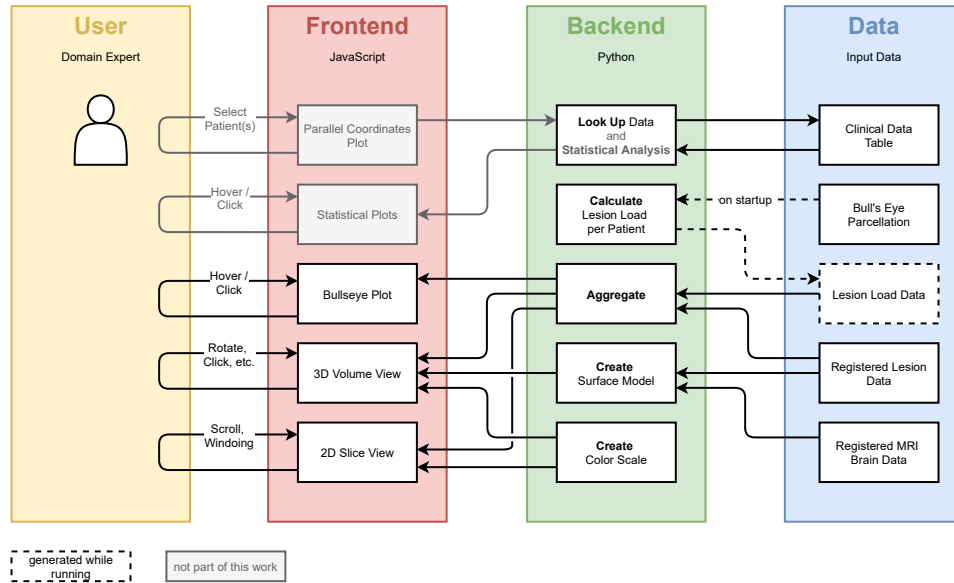


Figure 4.3: A schematic depiction of the application and internal workflow. The application can be divided into the three components: data, backend, and frontend. The fourth component is the user who interacts with the frontend, which leads to further data processing in the backend.

only updated on-demand. For this purpose, different buttons provide various selection options. Either one selected subset, consisting of a single patient or a cohort, can be visualized, or two subsets can be compared. Additional buttons allow a comparative visualization of the subsets either with the whole cohort including the selected subset or the whole cohort excluding the selected subset. To compare a selected subset with another individually selected subset, the user can first save a selection and then select another group. The saved selection will be shown in area B.1 and can be deleted by a press on the respective button.

The application is based on the visualization mantra as described in Section 3.1: *"Overview first, zoom and filter, then details on-demand"* (SHNEIDERMAN, 1996). Therefore, after deciding on a subset of patients, the volume data can be inspected on the right side. Different interaction techniques are available to adjust and filter the visualization. They are further described in section 4.3.2. To analyze the lesion load of the selected subset, the user can interact with the BEPs. By clicking in a cell of one of the

plots, the corresponding brain region is highlighted in the 3D as well as 2D representations of the volume data.

The application follows the visual analytics principle from KEIM et al. (2010) as described in Section 3.1, consisting of the four components *Data* (blue), *Models* (green), *Visualization* (red) and *Knowledge* (yellow). Figure 4.3 depicts the components of the tool developed in this work. It can be seen, that the components fit in the described visual analytics principle. The data is represented by the clinical and image data of the CSVD cohort as well as data derived from it. The models are implemented in the Python backend and process, aggregate, and create data. The results are sent to the JavaScript frontend which includes the visualizations the data is mapped on. Due to interaction with the visualizations, causing the recalculation of the models, the user (in this case a domain expert) can gain knowledge. This principle is also very similar to the model-view-controller pattern, which is described in Section 3.1 as well. The user can manipulate the data through the controller, which consists of the interactive elements in the UI, e.g. sliders and buttons. This causes the model to update and recalculate the data. Afterwards, the view, including all visual representations, is refreshed and the results can be observed by the user.

After this brief overview of the whole prototype, the following sections will focus on the frontend (see Sections 4.3.2 and 4.3.3) as well as the backend (see Section 4.4) components which have been proposed as part of this work.

4.3.2 3D Volume and 2D Slice Views

The image data, being the brain and the segmented lesions, are displayed in 2D as axial, sagittal, and coronal slice images and in 3D using indirect and optionally also direct volume rendering, see Figure 4.4. The views are implemented using the XTK toolkit and their basic functionality has been derived from the application Slice:Drop (HAEHN, 2013; HAEHN et al., 2014). Slice:Drop includes a lot of the functionality needed for the application presented in this work, such as the integrated visualization of slice images, segmentation maps, 3D models, and DVR.

On the bottom, the slice images are displayed. The segmented lesions are highlighted in different colors. A detailed discussion why colors were used

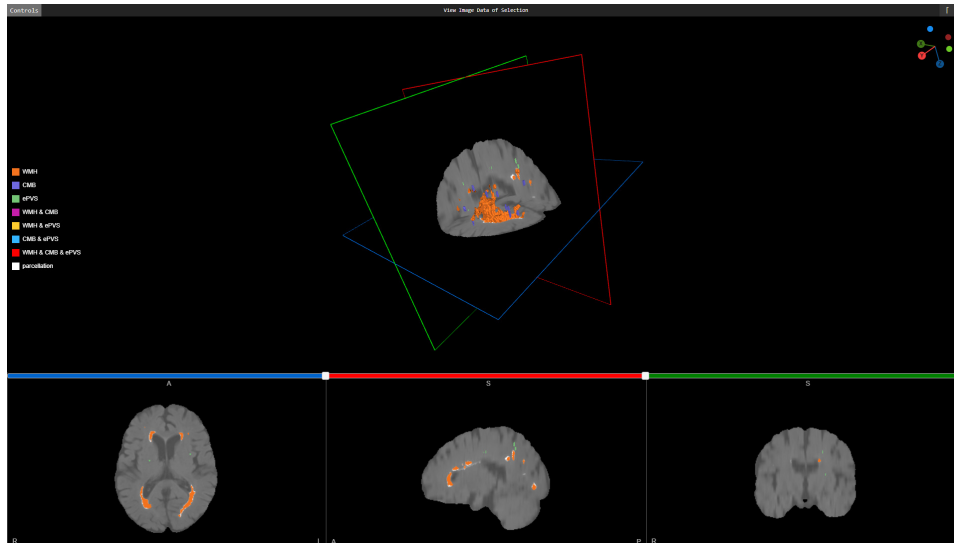


Figure 4.4: The visualization of the volume data of a single patient in 2D and 3D. The different lesions are color-coded based on their type. Overlapping lesions are highlighted with separate colors on the slices. A legend is displayed on the left side. The orientation gizmo in the upper right corner supports the 3D rotation.

instead of textures is done in Section ???. If multiple lesion types overlap, the corresponding regions are colored in a different hue, since in the 2D images, overlaps cannot be analyzed by hiding specific lesions. Furthermore, simply mixing the colors of two or three lesion types to show an overlap would distort the impression of color scales used to encode the lesion load. The user can scroll through the slices using the mouse wheel or the slider on top of each image and adapt the window function by clicking in the image and move the mouse. Up and down movements increase or decrease the brightness by lowering and raising the window level, respectively. Moving the mouse to the left shrinks the window size and increases the contrast in the image while moving the mouse to the right expands the window width and decreases the contrast. This is the interaction typically provided by other applications that allow the exploration of medical image data and therefore familiar to experts in this field.

In the center, the 3D visualizations are displayed. Here, the user can choose to either show the slices in 3D, to see their current position in the volume, or to visualize the whole volume using DVR instead. By using DVR, no surface model of the brain is needed. Furthermore, by extracting

a surface model many details are lost. DVR, on the other hand, offers the possibility to visualize arbitrary details of the original MRI scan by adjusting different parameters, like the transparency and gray value thresholds. The 3D slices have frames of the same colors as the corresponding sliders for the 2D slices to build a visual bridge between them. The rotation widget from JANKOWIAK (2020) is displayed in the upper right corner to support the 3D rotation of the objects. By clicking on it, the user can rotate the objects in the 3D view so that the selected dimension is parallel to the view plane.

Similar to the 2D slices, the lesions in the 3D view are color-coded based on the lesion type. On the left side of the visualization, a legend with the color scales is shown. The color scale is adaptive to the size of the selected subset, see Figure 4.5. When selecting one patient, swatches depicting a single color per lesion type or a combination of lesion types are shown. For the visualization of a cohort, the swatches of singular lesion types are replaced by a discrete color bar. The min and max values of the color bars depend on the current data's range. The color depicts how many patients have the same lesion type in identical voxel locations. Each lesion type is represented by a unique color scale, to ease their distinction, i.e. orange (WMH), purple (CMB), and green (ePVS).

If two subsets are selected to generate a comparative view, diverging color-scales are used. They depict the lesion load dominance of the subset comparison. One subset dominates the other in case more of the included patients have a lesion at the highlighted position. The number difference in patient count describes the amount of dominance of one set over the other. A detailed explanation of the dominance calculation is given in Section 4.4.2. Due to the amount of colors needed, used color scales include hue ranges that are hard to distinguish, e.g., orange and red. The divergent color scales pink-green (WMH), red-blue (CMB), and purple-orange (ePVS) were used. To select colors and color scales with the best possible contrast, *ColorBrewer* was used (HARROWER and BREWER, 2003). *ColorBrewer* is a web tool for selecting colors for maps that are easy to distinguish.

Introducing many colors per color scale increases the visual clutter and makes it hard to interpret the resulting image. For this reason, all color scales implement automatic binning in case the lesion maps include a

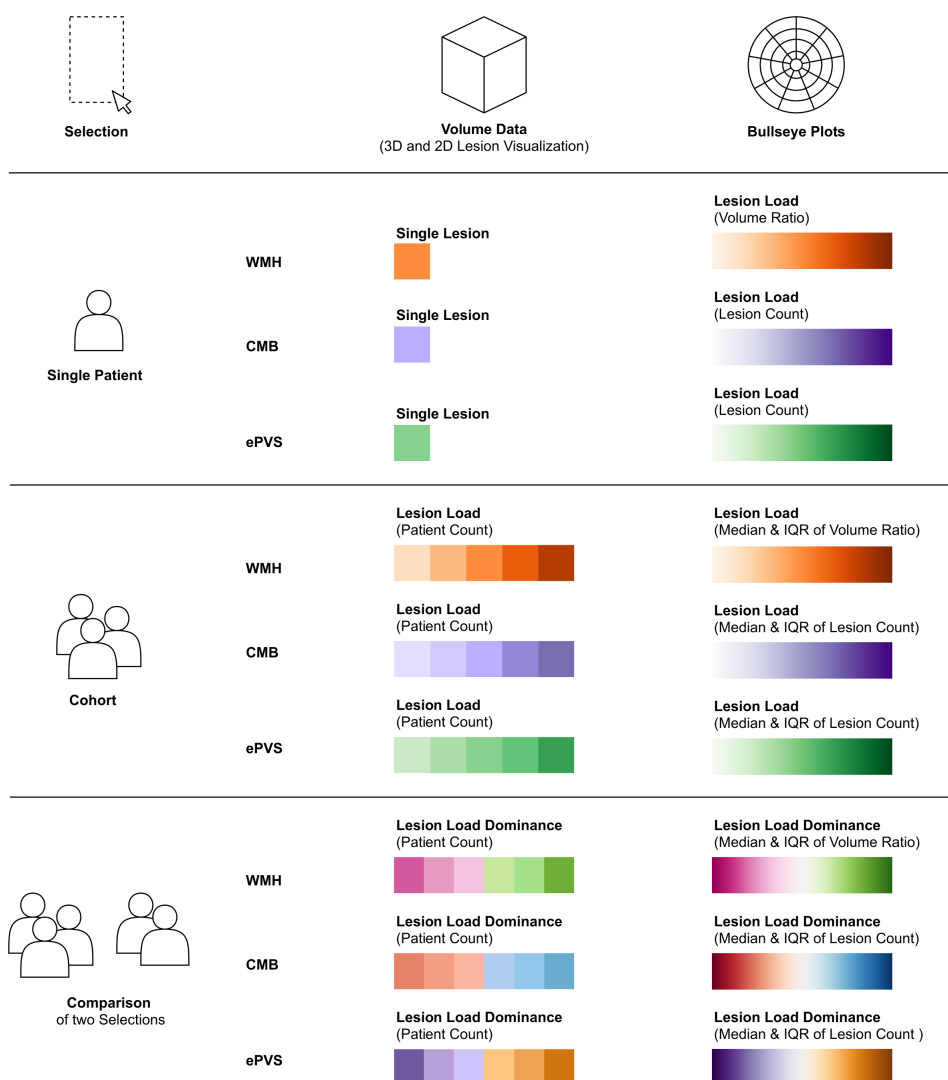


Figure 4.5: The used color scales depend on the user selection. If a single patient is selected, each lesion type in the slices, as well as the 3D lesions, are mapped to a single color. In case the user selects a cohort, the lesion load (in this case the number of patients that have a lesion at the same voxel position) is colored by a discrete color scale. A diverging color scale is used for the comparison of two selections. Each of the two selections can either be a single patient or a cohort. The diverging color scale shows the lesion load dominance, meaning the surplus of patients showing a lesion at the same voxel position in one subset in comparison to the number of patients in the other subset. The bullseye plots use continuous color scales to depict the lesion load for the selection of a single patient, a cohort, or the comparison of two selections.

large number of unique values. For each subset and lesion type, a maximum of six equally sized bins is produced. If more than six shades of the same color are used, it becomes hard for the user to distinguish them.

If the user is unsure about a specific lesion type, clicking on a 3D model of a lesion will provide further information. The pop-up tooltip shows not only the type but also the volume of the selected lesion. Multiple tooltips can be opened to highlight and compare selected lesions. Corresponding meshes will turn yellow when being hovered and remain dark yellow once the user clicks on them. In case the user rotates the scene, all selections will be rescinded and their respective tooltips will be closed. This prevents the tooltips and meshes from being obscured or mismatched. Displaying detailed information on demand for individual lesions simplifies their comparison. Furthermore, the different lesion types can be filtered based on their lesion load or completely hidden to reduce visual clutter in the 3D scene.

A menu at the top left corner displays several opportunities for interaction, see Figure 4.2c. The menu was created using *Guify* (COLE, 2019) and offers the following interaction possibilities:

- **3D Slices or DVR Dropdown:** The user can choose to visualize the volume data either as 3D slices or with DVR using a dropdown menu.
- **Toggle Lesion Map:** A toggle button is used to show and hide the lesion map that highlights the lesion locations in the 2D slides.
- **Volume Opacity Slider:** The opacity of the 3D slices and the DVR can be adjusted by a slider.
- **DVR Threshold Range Slider:** The upper and lower threshold of the gray values that are to be displayed in the DVR can be assigned using a range slider.
- **Brain Surface Model Opacity Slider:** The user can employ a slider to adjust the opacity of the surface model. While the surface model allows a localization of included lesions, it can also conceal relevant information. Modifying its opacity allows to adapt the visualization to the user's preference.

- **Lesion Load Threshold Sliders:** One slider per lesion type allows to show or hide the lesion surface models with a specific lesion load in the 3D view. This allows the user to focus on lesion loads in a certain range, which can be important during the analysis of larger cohorts.

The 2D, as well as 3D views, provide the user with well-known visualizations and a detailed anatomical context. However, the high level of detail can cause visual clutter which makes it hard to recognize individual lesions and patterns, especially in the accumulated cohort views. Therefore, a simplified visualization of the lesion load was included in the software and will be discussed in the following section.

4.3.3 Abstract Visualization of the Lesion Load

The abstract representation of the lesion load as bullseye plots presents an overview of the lesion distribution of a single patient as well as of a cohort, see Figure 4.6 The parcellation of the brain and representation of the lesion load as bullseye plots were introduced by SUDRE et al. (2018). The BEPs are divided into lobes and shells. The lobes refer to widely known brain regions: frontal (Front), parietal (Par), temporal (Temp), occipital (Occ) for each hemisphere as well as one joined region consisting of the basal ganglia, intratentorial regions, and thalami (BGIT). The shells provide information about the depth of the lesions. The inner ring of the BEP is depicting the shell closest to the ventricles while the outer ring is closest to the cortical grey matter. In the application, a legend is used to explain the parcellation for the user, see Figure 4.7. The parcellation is described in more detail in Section 4.4.4.

In the study of SUDRE et al. (2018) the BEPs were only used to depict the location of WMH, while in this work also CMB and ePVS are visualized with this method. The BEPs are showing the volume ratio of the brain region volume and the lesion volume for WMH and the lesion count for CMB and ePVS. The last two lesion types are so small, that their segmentation often consists of only a single voxel and the calculated volume ratio for CMB and ePVS would be extremely small. Therefore, displaying their lesion count is better suited for the quantification of the lesion load.

To create a visual connection of the different graphical representations of the lesions and the lesion loads, the color scale of the BEPs is based on the

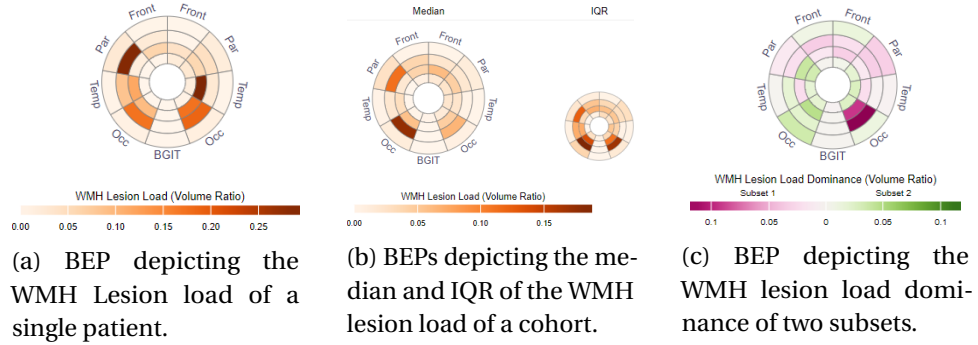
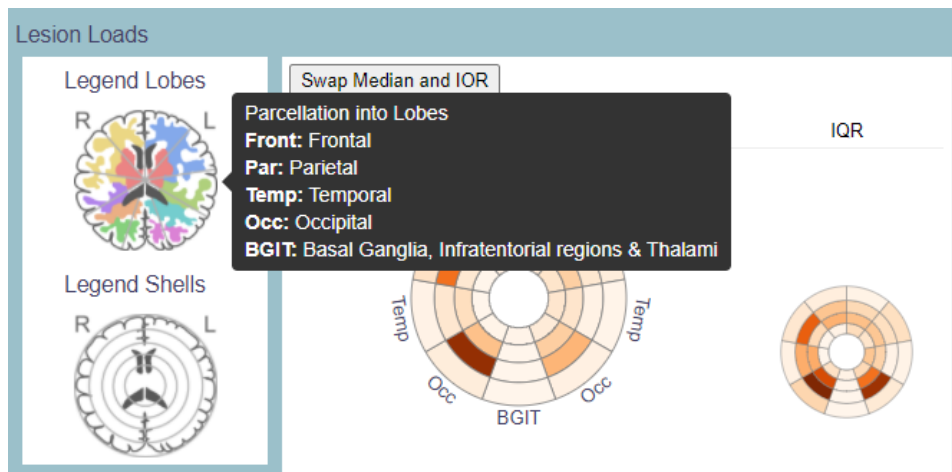


Figure 4.6: Depending on the selection, different BEPs are displayed in the UI. (a) One BEP displays the lesion load of a single patient, while (b) two BEPs depict the median and the interquartile range (IQR) of the lesion load of a cohort. When two cohorts are compared, (c) a BEP showing the lesion load dominance is generated additionally.

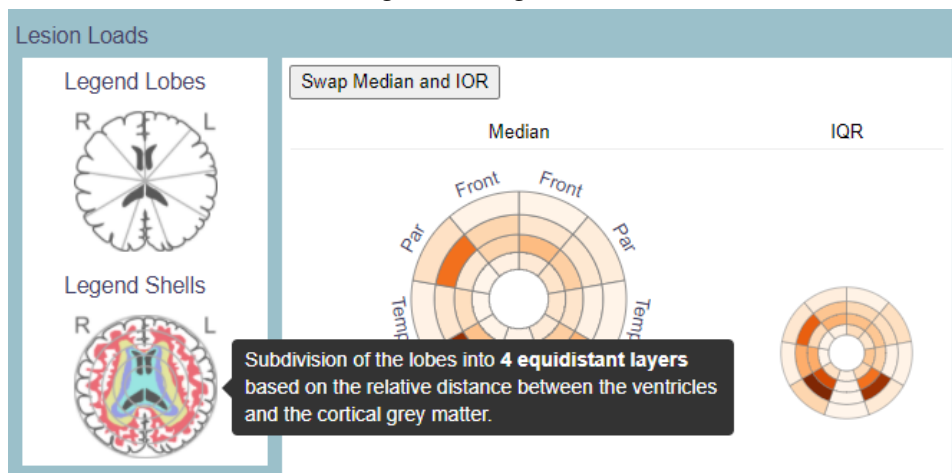
same hues as the color scales used for the 3D and 2D lesions in the volume data, see Figure 4.5. In contrast to the color scales used for the volume data, the BEPs use continuous color scales since they do not represent the number of patients with overlapping lesion areas but their median lesion load per segment.

In the plots, not only the lesion load itself but also other parameters can be displayed, see Figure 4.6. When the user selects a cohort, the mean and the IQR are depicted in two separate BEPs per lesion type. The use of these measures was approved by a neurologist with more than ten years of expertise in the field of cerebral vascular diseases. She stated that the plots allow getting a better impression on the lesion loads distribution over the selected subset of patients. For the comparison of two selections, the mean and IQR for each subset are depicted. Additionally, the lesion load dominance is visualized.

Like in the representation of the 3D and 2D lesions, the BEPs also visualize the lesion dominance when comparing two selections. For each subset, the median is determined. The dominance then indicates how much larger the median of one group is compared to the other. Therefore, the lesion load dominance plot highlights the differences between the two subsets.



(a) Hovering over the legend for the lobes.



(b) Hovering over the legend for the shells.

Figure 4.7: To explain the bullseye parcellation to the user, a legend for the lobes and shells was created. By hovering over the images, a schematic representation of the brain regions is displayed in color and a tooltip with further information opens.

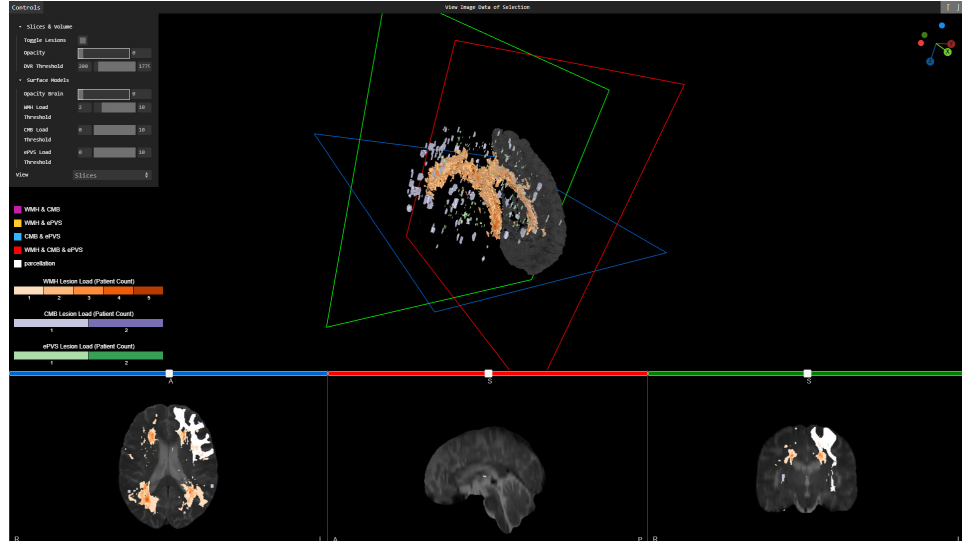


Figure 4.8: After the outer shell of the frontal lobe was selected, the respective brain region is shown as 3D surface model as well as white highlight in the slice images.

To better integrate bullseye plots and volume visualization, the plots were made interactive. Hovering individual cells displays the exact numerical value in a tooltip while clicking on a cell highlights the corresponding brain region in the volume visualization, see Figure 4.8. This allows for a quick identification of lesions contributing to the lesion load of this segment. An arbitrary number of cells can be selected. Only surface models of the selected brain regions are displayed in the 3D view, instead of the entire brain. In the slice images, the corresponding regions are highlighted in white.

4.4 Data Processing

This section describes the processing of the data which is performed in the python backend to generate the visual representations. It includes the import of the data as well as its aggregation which is performed to generate the volume visualizations as well as the bullseye plots.

4.4.1 Import of the Data

The segmentation results of the lesions are stored as separate binary maps for each patient and lesion type. It is saved as a three-dimensional volume with the same size as the registered MRI data, being 256 x 256 x 256 voxels with a spacing of 1 in each dimension. Each voxel where a lesion was detected has a value of 1 while the other voxels have a value of 0. Both, MRI data and segmentation results are saved in the NIfTI file format. To read the NIfTI data files, the Python toolkit SimpleITK was used.

Since XTK is limited to processing a single segmentation map per volume, the segmentation maps are aggregated into a combined segmentation map before being sent to the frontend. Segmentation maps are aggregated by creating a new segmentation map that assigns unique values to the occurrence or overlap of lesion types. The following values were assigned: (0) no lesion, (1) WMH, (2) CMB, (3) ePVS for non-overlapping lesions and (4) WMH and CMB, (5) WMH and CMB, (6) CMB and ePVS, (7) WMH and CMB and ePVS for overlapping occurrences.

Afterwards, the meshes of the lesions are being created using marching cubes, as described in Section 4.4.3. To synchronize the colorization of meshes and the lesions of the segmentation map a colortable that maps the occurrence of each label in the segmentation map to a unique color will be stored as a CSV file.

4.4.2 Accumulation of Lesion Data for the Cohort Creation

To accumulate the lesions for the analysis of a patient cohort, the segmentation results of all patients are summed up, which can be performed as a simple matrix operation, see Figure 4.9. The result is a three-dimensional matrix per lesion type, the entries encode the number of patients who have a lesion at the respective voxel position. Based on this data, the 2D and 3D representations of the lesions are created.

Once again the segmentation maps for all three lesion types are combined into a single map. In contrast to the combination of a single patient's segmentation maps, each map can consist of values larger than 1. To aggregate the maps, the maximum value for each map is determined in the first step. Lets consider \max_{wmh} , \max_{cmb} , and \max_{epvs} to be the maximum

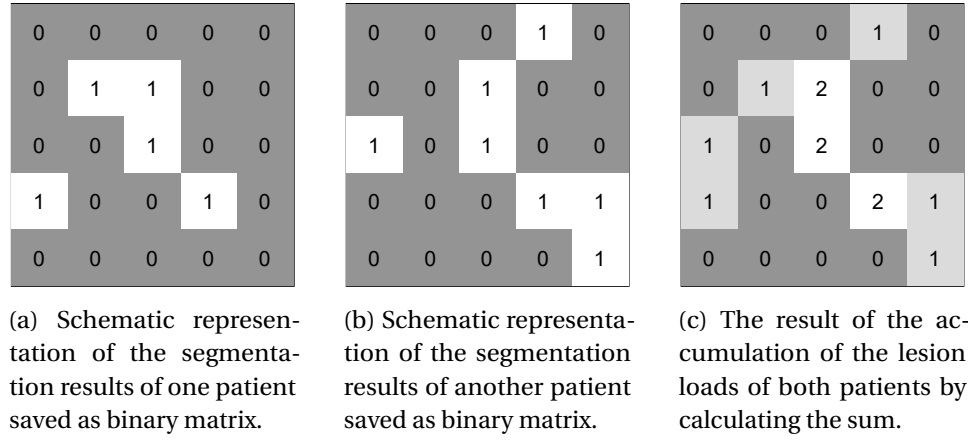


Figure 4.9: Schematic depiction how the segmented lesions of a selected cohort are aggregated to generate the visualizations of the 2D and 3D lesions. The segmentation results are stored in 3D matrices which are summed up to calculate the patient count at each voxel position.

value of the respective lesion type map. The combined aggregated lesion-map maps wmh lesions to the range $[1, \max_{wmh}]$, followed up the cmb lesions to the range $[\max_{wmh} + 1, \max_{wmh} + \max_{cmb}]$ and $epvs$ lesions to $[\max_{wmh} + \max_{cmb} + 1, \max_{wmh} + \max_{cmb} + \max_{epvs}]$.

As described in Section 4.3.2, a discrete color scale with a maximum of six different colors is used for the depiction of the lesion load of a cohort in the volume visualization. If the accumulation of the cohort data leads to values larger than six, it is necessary to group these values. Therefore, binning is performed. The application always tries to generate five bins with each of them holding an equal amount of values. If the number of different values is not divisible by five, a sixth bin is included containing the remaining values.

To compare two cohorts the segmentation data for each selection is first summed up. Afterwards, the results are subtracted to calculate the lesion load dominance, see Figure 4.10. This way a 3D matrix is created, where the entries encode how many more patients of one subset have a lesion at each voxel position. This data is used to generate 3D and 2D representations of the lesions, showing which cohort is dominant at each voxel position as well as how strong the dominance is.

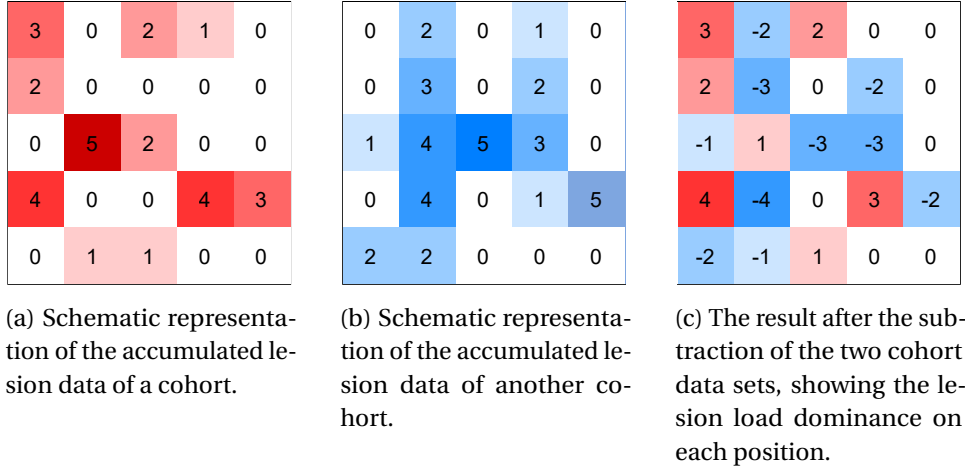


Figure 4.10: Schematic depiction of the creation of the lesion load dominance data for the 2D and 3D lesion visualizations. The two aggregated lesion data sets are subtracted to gain the lesion dominance for each position.

The segmentation maps are only used for the visualization of the lesions as surface models and the highlighting in the volume slices. The lesion load for the bullseye parcels for each patient is calculated when starting the application and saved as a binary file, which will be further explained in Section 4.4.4. When a cohort is selected, the respective binary files are used to calculate the mean as well as the IQR for each parcel. Similar to the segmentation results, the results for each lesion type are saved separately. To generate the dominance plot, the mean of the subsets are subtracted to extract the lesion load dominance for each brain region.

4.4.3 Creation of Surface Models

For the visualizations of the lesions in 3D, as described in Section 4.3.2, three-dimensional models of the lesions have to be created. Therefore, marching cubes was performed on the segmentation maps. Since the segmentations are saved as a binary matrix, an isovalue of 1 can be used to extract the lesions. Through a connected components analysis, the individual models can be extracted separately. They are saved as OBJ files and imported into the 3D scene.

If a cohort is selected, marching cubes is performed on the aggregated data. To reduce the computational effort, all lesions with the same lesion

load are extracted as one model instead of separating each individual lesion. Therefore, the 3D matrix is filtered. For each value represented in the matrix, a temporary binary matrix is created. This way, marching cubes can be performed with the same isovalue as for single patients. The same approach is followed in generating the comparison of two cohorts. Therefore, not the lesions of the two cohorts are visualized in the comparative visualization, but surface models of the lesion load dominance.

Marching cubes was also used to generate a surface model of the MNI-ICBM152 template brain which is used for the 3D visualization. The brain as well as the lesion for each individual patient are saved permanently and do not have to be recalculated when restarting the application. The surface models of the aggregated lesion loads have to be calculated each time a new selection is made. Due to the large amount of possible subsets and combinations, saving all files permanently would require a large amount of disc space. On the other hand, generating the models at run-time leads to longer loading times.

4.4.4 Parcellation of the Brain

To create the bullseye representations of the selection as described in Section 4.3.3, the brain needs to be parcellated into separate regions. The parcellation in this work is based on the code of SANROMA (2019) and the *recon-all* pipeline of FreeSurfer which are recreating the parcellation proposed by SUDRE et al. (2018).

First, a lobar parcellation is done on the registered image data and the segmented lesions to divide them into nine lobes using FreeSurfer. The lobes are the segments of the BEP, the labels are based on their first letters. Frontal (Front), parietal (Par), temporal (Temp), and occipital (Occ) lobes are depicted on the left and right sides. The basal ganglia, thalami, and infratentorial regions (BGIT) were combined in the lower segment.

Second, a concentric parcellation divides the data further into shells, based on the distance to the ventricles. Therefore, a normalized distance map is used, interpolating between the cortex and the ventricles. Four equidistant shells are created for the four corresponding rings of the BEP.

Finally, the bullseye parcellation is created by intersecting the lobes and shells, the result can be seen in Figure 4.11.

The parcellation is performed on the MNI-ICBM152 template brain, resulting in a brain atlas, that can be applied to all data that is registered to the same space. For the generation of the bullseye data, the parcellation is used on the segmentation maps. The volume (WMH) or the count (CMB and ePVS) of the lesions inside of each parcel is then calculated and stored as a binary file for each patient.

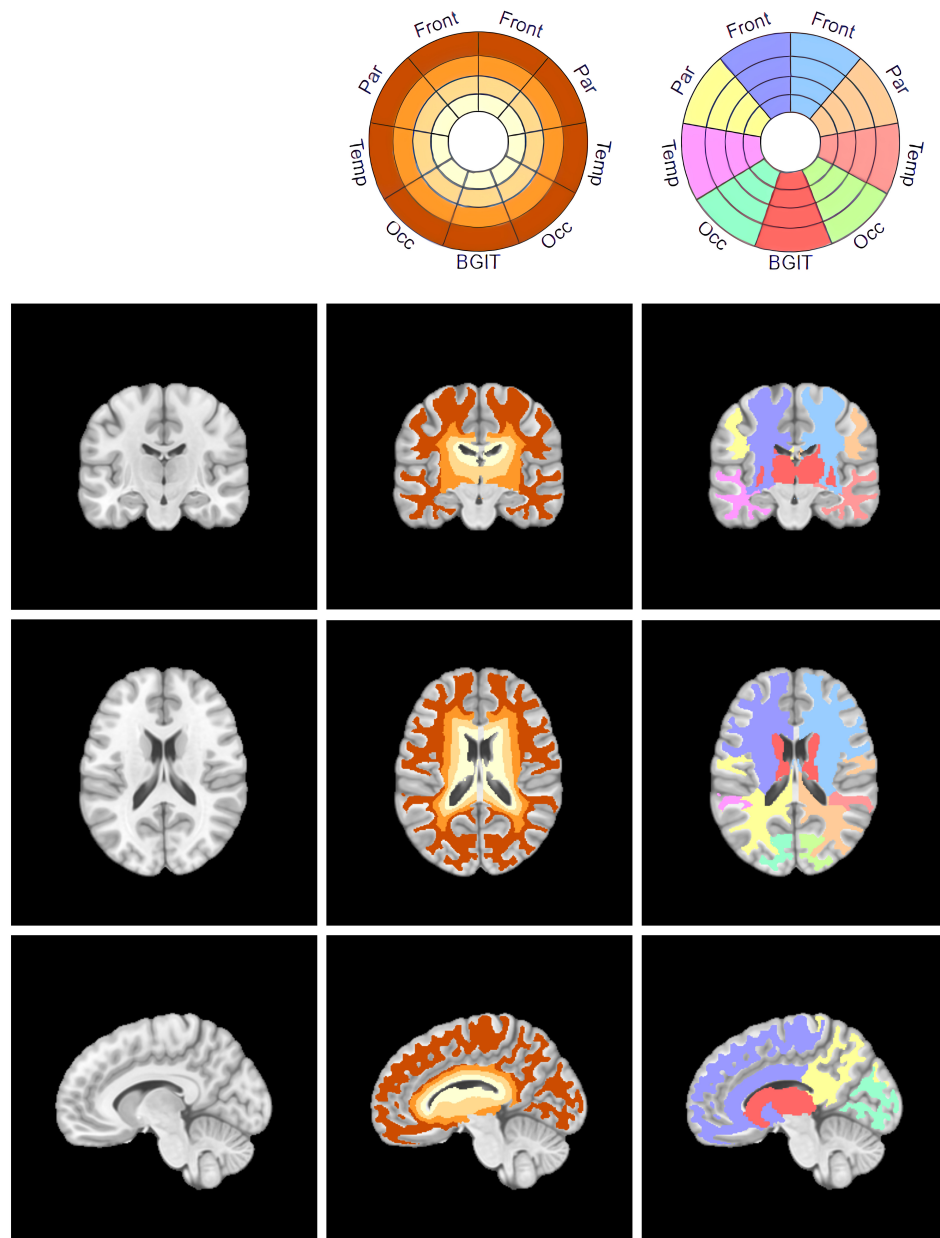


Figure 4.11: The bullseye parcellation mapped to the MNI-ICBM152 template brain. The first column shows the brain without any annotations. In the second column, the shells were drawn on the brain image and the third column shows the position of the different lobes in the brain.

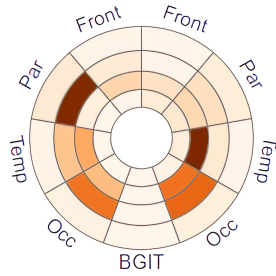
4.5 Discussion of the Developed Visualizations

In this section, the chosen visualizations for mapping the lesion load are discussed and compared with alternative representations. Afterwards, the application as a whole is analyzed with regard to the challenges named in Section 3.1.

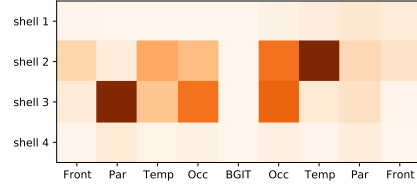
4.5.1 Comparison of Bullseye Plots with Alternative Visualizations

The biggest advantage of the bullseye plot is the large amount of information it can present. A total of 36 brain regions are displayed and arranged inspired by the anatomical location of the lobes in the left and right brain hemispheres as well as their distance to the ventricles. Therefore, the visualization is more intuitive and compact than other plots like e.g., bar charts, heatmaps, or radar charts, see Figure 4.12. To display the same amount of information in a bar chart, 36 bars are necessary. Due to the arrangement of the bars next to each other, the anatomical-based positioning used by the BEPs is being lost. Furthermore, there is no visual distinction between lobes and shells. This could be improved by grouping the bars or using stacked bar charts, see Figures 4.12c and 4.12d. In stacked bar charts, each bar could represent a lobe while the shells are the partitions of the bars. However, stacked bar charts are not easy to read and quickly to misinterpret, since it is hard to compare the different sections of the bar when they are not aligned at a common baseline (STREIT and GEHLENBORG, 2014).

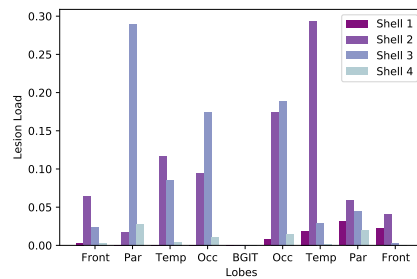
Other radial plots could be used instead to display the data in a more compact way than a bar chart. In radar charts, the different axes could be used to depict the lobes. The shells are represented by rings, similar to the BEP. However, in radar charts, the data is drawn as a line, which can lead to visual clutter when displaying many different entries. Alternatively, one plot has to be drawn for each shell, which would require more space and reduce intuitivity, see Figure 4.12e. Furthermore, the axes of each plot should have similar scaling to make the shells comparable. Shells that have a relatively small lesion load compared to the others are almost not visible in the plots. Other radial plots, e.g., polar area charts suffer from a similar problem, see Figure 4.12f. Additionally, polar area charts are easily misinterpreted due to an inaccurate estimation of the size of one entry



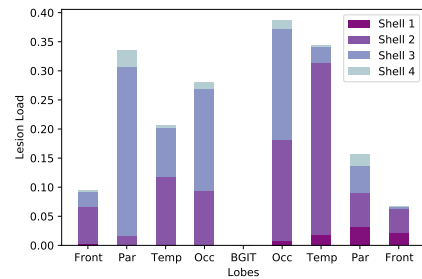
(a) A bullseye plot depicting WMH lesion load of a single patient.



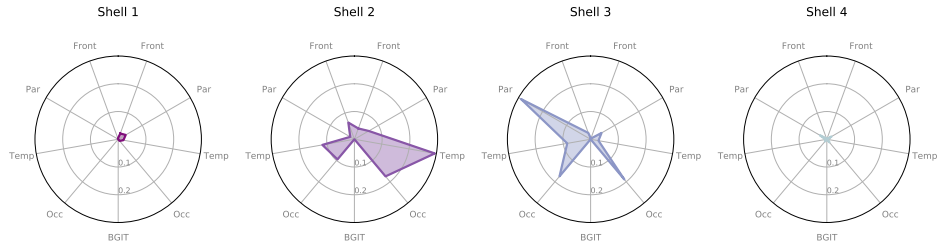
(b) A heatmap depicting WMH lesion load of a single patient.



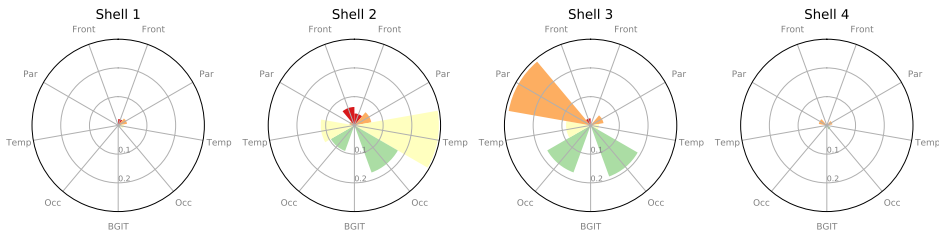
(c) A grouped bar chart depicting WMH lesion load of a single patient.



(d) A stacked bar chart depicting WMH lesion load of a single patient.



(e) Radar charts depicting WMH lesion load of a single patient, where each plot represents a shell.



(f) Polar area charts depicting WMH lesion load of a single patient, where each plot represents a shell.

Figure 4.12: Comparison of (a) a bullseye plot, (b) a heatmap, (c) grouped and (d) stacked bar charts, (e) radar charts and (f) polar area charts depicting the same lesion load data of a single patient.

since the area of each entry is not proportional to the values it represents. Therefore, the value should only be checked on the corresponding axis, which in turn makes the surfaces obsolete and distracting for the viewer.

The BEPs can be easily transformed into a heatmap by placing a cut between two sections and unfolding and transforming the circle into a rectangular shape, see Figure 4.12b. This way, the same amount of data can be displayed in an evenly compact plot. However, the anatomical parallels to the structure of the brain are mostly lost.

Besides using plots from the field of information visualization, 2D representations of the brain are often used to generate an overview of the whole brain. Therefore, different brain flattening techniques can be used. As described in Section 3.2.1, brain flattening is only applicable to the cortex or individual substructures and cannot depict structures in deeper brain regions. Furthermore, flattening highly curved surfaces like the brain can lead to massive distortions. Since the focus of this work is on the distribution of different lesions in the white matter, a flattening of the cortical surface is not sufficient to depict all the necessary information. In contrast, the bullseye representation considers the whole white matter and can differentiate between deep and lobar regions.

Thus, it meets the criteria of PREIM and BOTHA (2014c) described in Section 2.2.3 to be an anatomical planar projection. The data is accumulated for each brain region and depicted in a planar geometry. In contrast to reformations, all of the measured values (in this case the lesion load) are used instead of sampling the data or generating new values due to an interpolation.

The compact and anatomy-related representation allows the user to quickly assess the lesion load in deep as well as lobar regions. The size of the different parcels is not depending on the scalar values, therefore, the structure of the plot stays always the same and multiple BEPs can be easily compared.

4.5.2 Comparison of Color-coded and Texture-coded Lesions

As mentioned in Section 4.3.2, the use of three diverging color scales leads inevitably to the use of rather similar colors. Furthermore, highlighting

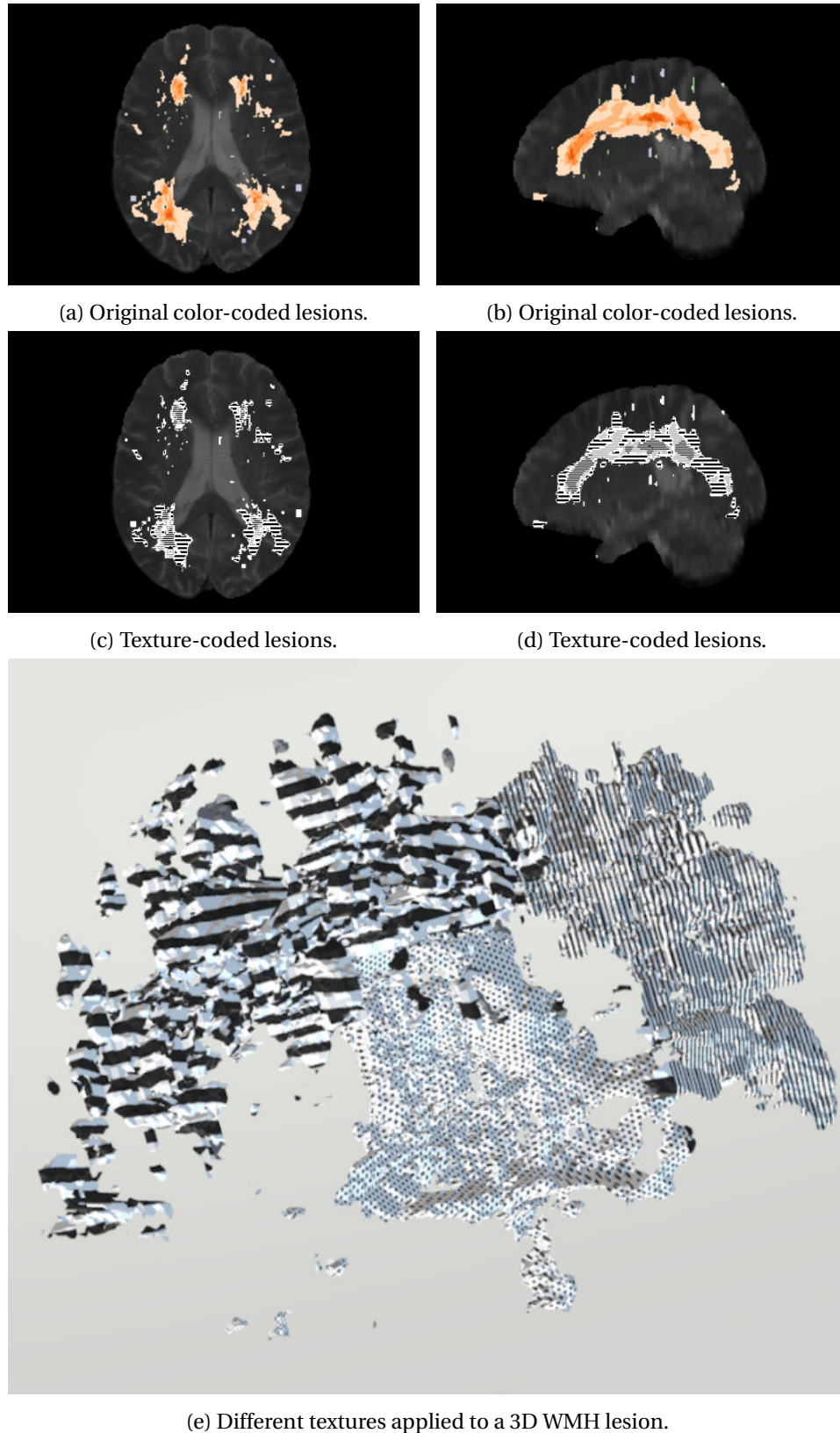


Figure 4.13: Comparison of (a,b) the color-coded lesion load used in this work and (c,d) a schematic depiction of how a texture-coded lesion load could look like. Three different textures were applied on (e) the 3D model of a WMH lesion.

overlaps of different lesion types with additional colors is almost impossible.

Apart from color, scalar values can also be encoded using textures. Not only does this allow even more different values to be encoded, but textures are also readable for people with limited color vision.

However, the slice images are depicted rather small, which makes textures hard to distinguish, see Figure 4.13. Furthermore, CMB and ePVS lesions are very small, so that the user would not be able to recognize the texture.

In addition to this, the 3D models have a very uneven surface. If combined with a texture, it makes the visualization difficult to interpret and tiring to look at. To apply textures to the 3D models, they should be smoothed beforehand. However, this would alter their original shape.

4.5.3 Challenges in Visual Analytics for Health Care

In Section 3.1, SHNEIDERMAN et al. (2013) defined seven challenges in the development of visual analytics applications for health care. This section discusses whether and how this work addresses the challenges.

Offering busy clinicians timely information in the right format: The application presented in this work can summarize MRI data of multiple patients in two ways. First, the lesions are aggregated and depicted in the slice images and as 3D models, providing the user with a high amount of details. However, the calculation of the aggregated volume data takes several seconds, which slows down the workflow. On the other hand, the more abstract visualization of the lesion load as bullseye plots does not only have a very short processing time but also offers a compact and standardized overview. It is suited for the analysis of one patient, a cohort as well as the comparison of multiple patients or cohorts. Overall, the goal of the application is to support clinical research. Therefore, the setting is not as time-critical as in clinical routine and longer loading times are tolerable.

Moving towards an ecosystem of visual tools: Due to the implementation of the application as a client-server structure using Python and JavaScript, it can be used on many different platforms and only a common browser is

needed to run it. However, the application is designed to be used with two 16:9 screens. Therefore, the UI cannot be displayed correctly on small devices. Since the aggregation of the volume data is computationally expensive, the backend has to be running on a machine with enough computing power. The application works with web requests, therefore, the calculations could be done by a local server that returns the data to be visualized to the client. Additional UIs for, e.g., mobile devices could be developed to provide further means of accessing the data and make the visualization available during clinical practice. Furthermore, the application supports multiple languages. While currently only English and German translations are provided, additional languages can be included easily.

Facilitating team decision making: The application works with several visualizations which differ in their suitability for people from different professions. The bullseye plots represent well-known brain regions, that are familiar to many persons with a medical background. However, the BEPs might not be well suited for patients, as they are mostly not familiar with the anatomy of the brain. The visualization of the volume data in 3D and the 2D slices is easier to understand for non-professionals since they are less abstract. However, the application was developed for expert users who are familiar with brain anatomy as well as statistics and most elements are not intuitive for non-experts. Furthermore, the application can only be used on one machine at the moment. The underlying architecture would allow to extend the tool for remote use, as well as cooperative use by multiple users on different machines.

Characterizing and understanding similarity: The presented application does not support the import of unformatted text or the automated search for related information online or offline.

Visualizing comparative relationships: Comparative visualizations are a strong focus of the developed application. Due to the representation of the lesion load dominance, differences in the two data sets are highlighted. The BEPs additionally allow the side-by-side comparison of two sub-cohorts. Therefore, differences in the image data of two subsets can be easily perceived.

Presenting risk and uncertainty: The focus of the application presented is less on treatment decisions and their associated risks, but rather on an understanding of possible patient groups and their commonalities or differences. In this context, descriptive statistical values such as IQR are relevant to assess the significance of observed differences. In addition, the missing data is shown to the user. There are no further depictions of uncertainties included in the visualization of the image data.

Evaluation: The application was evaluated by three experts from the field of neurology. However, the focus of this evaluation is on the technical aspects of this work. A more in-depth and comprehensive evaluation of the application, including a larger data set, is necessary in order to guarantee the application works correctly and maps the data accurately. The evaluation is discussed in detail in the next chapter.

5

Evaluation

This chapter evaluates the developed application for brain lesion analysis. The following sections present a critical evaluation of the results. The application is discussed in the context of an expert study. Therefore, it was presented to three experts who are working frequently with neurological data. The focus of the evaluation were the technical aspects of the application. In addition, the cohort of ten patients which was described in Section 4.1 was assessed by medical experts.

First, the setup of the evaluation is described. Afterwards, the results of the discussion of the application and its visualizations followed by a medical assessment of the data set are presented. The chapter closes with a conclusion that summarizes the evaluation results.

5.1 Evaluation Setup

Data The clinical data described in Section 4.1 were used. It consists of ten patients diagnosed with CSVD. The WMH, CMB, and ePVS lesions were segmented and the data registered to a common space. Six of the patients were classified as CAA, while two have mixed CAA and the other two mixed HA.

Participants Three experts who are working in the field of neurology were included in the evaluation. The first expert is a senior neurologist at the department of neurology at the University Hospital Magdeburg and is doing active research on the CSVD. She was also involved in the development of the application. The second expert is a PhD student from the field of neurology. The third expert comes from the field of control engineering and

modeling and is working on quantitative analysis as well as imaging analysis. All of the experts are working in the same department and are familiar with the CSVD as well as the initial dual analysis framework on which this application is based on.

Procedure A formative evaluation was conducted to discuss improvement opportunities for the prototype (LEWIS, 2012). Therefore, the application was presented to each expert in a separate meeting. The same click paths were used for each run, which consists of the following actions:

- **Selecting a single patient:** A single patient was selected. Based on this selection the general UI was explained as well as the visualization of the volume data and the bullseye plots.
- **Selection of a sub-cohort:** The sub-cohort of patients diagnosed with CAA were selected and displayed. Based on this, the changes in the visualization of the volume data and the bullseye plots were discussed.
- **Comparison of two selections:** The selected subset of patients with CAA were compared to a subset of patients with mixed CAA. Again, the changes in the visualizations were explained.
- **Comparison of two selections:** The two subsets of patients with mixed CAA and mixed HA were compared.

For the evaluation, a semi-structured interview was conducted. Therefore, the so-called *Think Aloud method* (VAN SOMEREN et al., 1994) was used. In this method, the participants are asked to speak out loud their thoughts about the functionalities during the presentation of the application, so that these can be recorded. It ensures that participants can provide detailed feedback on all aspects of the application, as it is not limited to individual questions chosen by the developer.

In order to obtain targeted feedback on the specific components, the experts were asked four questions concerning the application:

- Q1:** Is the 3D visualization a good addition to the classical 2D slides? Is it cluttered or crowded in your opinion?
- Q2:** Are the bullseye plots a good visualization for presenting the lesion load and how understandable and intuitive are they?
- Q3:** Apart from the lesion volume ratio for WMH and the lesion count for CMB and ePVS, are there other parameters that would be interesting for the analysis of the lesion load?
- Q4:** Are there any interaction possibilities or functionalities missing?

The two medical experts, being the senior neurologist and the PhD student, were furthermore asked to analyze the lesion load of the CAA, mixed CAA, and mixed HA subsets:

- Q5:** Does the visualization of the lesion load of CAA and mixed CAA cases look as expected? Do you see anything of special interest?
- Q6:** Does the visualization of the lesion load of mixed CAA and mixed HA cases look as expected? Do you see anything of special interest?

The following sections are structured thematically and address the respective questions.

5.2 Discussion of the UI Elements and Visualizations

All of the experts were familiar with the original framework and stated that including image data is a very good extension to the purely statistical analysis of clinical data. Due to the visual nature of the extension, it is more intuitive. Therefore, one expert said, that the visualization of the imaging data would also be interesting for physicians in clinical practice, while the statistical analysis of the metadata is mainly interesting for research purposes. However, the combination of both is interesting due to the possibility to create and visualize sub-cohorts based on clinical data, e.g., age. Furthermore, the experts said, that they do not know a comparable interactive tool to visually analyze the image data of patient cohorts yet.

The two experts with a medical background said that the application is also suited for a longitudinal study of patients with CSVD. Furthermore,

the principle is also interesting for other diseases and other organs, e.g., the heart or the liver. However, for research purposes, a statistical analysis of the lesion load would be important.

5.2.1 Volume Visualization

The three experts all liked the visualization of the volume data as 3D models and the highlighting of the lesions in the slice images. Since the visualizations are directly based on the MRI data, it is easy to interpret. Concerning **Q1**, the experts described the 3D view as interesting and the semi-transparent brain gives a good anatomical context. It provides a good overview of the overall lesion load without browsing the slices. The interaction techniques are especially helpful to investigate smaller lesions.

The experts also had concrete suggestions for how to improve the volume visualization:

- **Using more saturated colors** to highlight a high lesion load in the cohort view, as these stand out more against the dark background.
- **Including separate sliders** in the comparative view to show or hide each of the subsets individually, as well as buttons to toggle the individual lesion types in the slice images.
- **Looking at the lesion load per region** instead of a voxel-wise analysis might result in more interesting views. Especially for the small lesions, it is very likely that affected voxels are close together but not exactly congruent. Nevertheless, these cases should be counted as overlapping as well.

Concerning the last point, the expert with a more technical background suggested to first smooth the segmentation maps, aggregate them and filter the results with a threshold afterwards, to give each lesion voxel a footprint, close to the principle of the volume rendering technique splatting as described in Section 2.2.1 (WESTOVER, 1990). Furthermore, he argued that the loading times could be reduced by pre-calculating fixed subsets, e.g. the groups of CAA, HA, mixed CAA and mixed HA. But he also mentioned that this would reduce the flexibility of the tool to select arbitrary subsets.

Overall, the threshold filtering of the 3D lesion models reminded the senior neurologist of voxel-based morphometry, an approach where multiple brain images are compared voxel-wise based on quantitative descriptors like, e.g., size, intensity, or shape parameters (ASHBURNER and FRISTON, 2000). A voxel-wise statistical analysis would be beneficial to quantify the results of the analysis for research.

5.2.2 Abstraction as Bullseye Plots

The bullseye plots were approved and the two experts with a medical background found this visualization especially interesting. Concerning **Q2**, they said that the plot is very easy to understand, but the legend is necessary to grasp the meaning of the BGIT region as well as the rings of the plot. All experts thought the legend is explaining the plots well. The plots have been described as very good for the comparison of different patients and cohorts. The two experts with a medical background said, that the plots would also be beneficial as an overview for physicians, providing a good impression of where to look at in the volume data. Looking at the plot is way faster to go through the volume slice by slice. The plots are depicting a reasonable amount of data and anatomical context and none of the experts had any further suggestions for improving the structure of the plots. However, one suggestion was made by the experts:

- **Offer the possibility to show additional parameters** in the bullseye plot, e.g. p-values of a statistical analysis.

The visualization of the lesion distribution as a bullseye plot was considered a good alternative to the presentation of numerical values in a table. Due to the color-coding, it is easier to capture the overall lesion load distribution and compare different cases.

5.2.3 Proposed Extensions and Modifications

During the evaluation with the experts, they mentioned possible extensions for the developed application. In the following, these suggestions will be discussed in more detail. First, the parameters to be added to the analysis process will be described and then other desired functionalities for the analysis process will be mentioned.

Additional Parameters Concerning **Q3** all experts said, that the parameters used were a good choice for displaying the lesion load. For the further development of the application, the following additional parameters were suggested:

- **Showing a relative lesion load** based on the patient count.
- **Fining certain patterns of WMH** lesion distribution, e.g., based on their shape, texture, and size, which might be classified to define subgroups of cognitive diseases (GWO et al., 2019). An interesting aspect of such an analysis for diagnostics is the occurrence of large areal and small point-shaped WMHs. However, this field is currently only addressed in research and not yet relevant in clinical routine.
- **Visualizing the deformation** of specific regions due to the non-linear registration might be an interesting aspect. Since the shape of the lesions could be meaningful in terms of classifying disease patterns, a potential deformation due to registration has to be considered. Another interesting aspect that could be analyzed is the symmetry of the lesion load in both hemispheres.

Additional Functionalities Concerning **Q4**, the experts suggested additional functionalities that would improve the workflow:

- **An option to export images** would be good, as the quality of screenshots is always depending on the resolution of the current screen.
- **Including more atlases** would be beneficial to describe affected areas. Many atlases are more finely resolved than the bullseye parcelation. Moreover, the description by public available atlases makes the results comparable with other studies. Since the data is already registered to a common space, respective atlases could be used.
- **Including a visualization for similarities** for the bullseye plots and the volume visualizations. For the comparison of the mixed cases, it is very interesting so see the differences, like it is done in the lesion load dominance visualizations. On the other hand, when comparing CAA and mixed CAA or HA and mixed HA, the similarities are more important.

5.3 Analyzed Patient Subsets

The two medical experts were asked to give an assessment of the lesion load distribution of the patients with CAA, mixed CAA and mixed HA. Of particular interest was whether the distribution of lesions was as expected and if any abnormalities could be detected.

5.3.1 Comparison of CAA and mixed CAA

The CAA subset includes six patients of which two had no WMH data and one had no ePVS data. The mixed CAA subset consists of two patients, both of them had all three lesion types. The experts stated, that the overall lesion load comparing CAA and mixed CAA was as expected. Particularly noticeable was the fact that in the presented CAA cases, CMBs appear exclusively in lobar regions, see Figures 5.1 and 5.3a. This pattern is strongly associated with CAA (CHARIDIMOU et al., 2017; SCHEUMANN et al., 2020). In the mixed CAA cases, some lesions appear in deeper shells in the frontal as well as the BGIT region, but the main lesion load for CMB follows the typical CAA pattern.

5.3.2 Comparison of mixed HA and mixed CAA

The subsets of mixed HA and mixed CAA each consisted of two patients, while only one of the patients with mixed HA had WMH lesions. The comparison of mixed HA and mixed CAA was also as expected by the experts. Mixed CAA lesions tend to be lobar, whereas mixed HA lesions are more prevalent in deep regions, see Figure 5.2 and 5.3b. Thus, the mixed CAA cases are more similar to CAA and mixed HA cases to HA in their lesion distribution, which matches the findings of SCHEUMANN et al. (2020) who analyzed the same data set. Typical patterns for HA could be found in the mixed HA cases, e.g., a higher lesion load of CMB at the frontal regions and of ePVS at the BGIT region. The patients with mixed CAA had a stronger lesion load in the occipital regions.

It was noticed that the lesion load in the frontal regions was not symmetrically distributed. The experts stated, that in principle a high symmetry of lesion load in the brain can be expected because it is a systemic disease affecting the brain in its entirety at certain predilection sites. Asymme-

try of lesion load could be expected in rare inflammatory forms of CSVD. Studying the asymmetry could help clinicians to diagnose these variants. Furthermore, a hemisphere comparison might be useful to investigate whether a pattern is a resilience factor against CSVD or to investigate the relationship between CSVD load and vascular patterns of the circle of Willis.

5.4 Conclusion

All experts stated, that the tool provided very interesting insights. It serves as a good starting point on which to build further research. The abstraction of the lesion load as bullseye plots provides a good and intuitive overview and the concept of it was quickly grasped by the experts. A more detailed analysis can be performed due to the visualization of the volume data as 3D models as well as slice images. A statistical analysis of the lesion load should be included in the future for further research. Additionally, highlighting not only differences but also similarities would be very interesting for the investigation of the mixed cases of CAA and HA. Moreover, the visualizations might also be interesting for clinicians as well as for longitudinal studies and for different diseases and organs. Thus, the principle of the presented application is not limited to the presented use case. One of the experts also noted that the application should be evaluated using more patient data to also ensure sufficient accuracy of the application.

Even though only a small number of patients was included in the analysis of the lesion load, the comparison of the CAA and mixed CAA as well as mixed HA and mixed CAA cases showed results that are similar to other studies and hints further insights that need to be investigated using a larger set of patients. It was noticeable that the lesion load was not symmetrical in some cases. This could be interesting for research as well as diagnostics and should be further investigated.

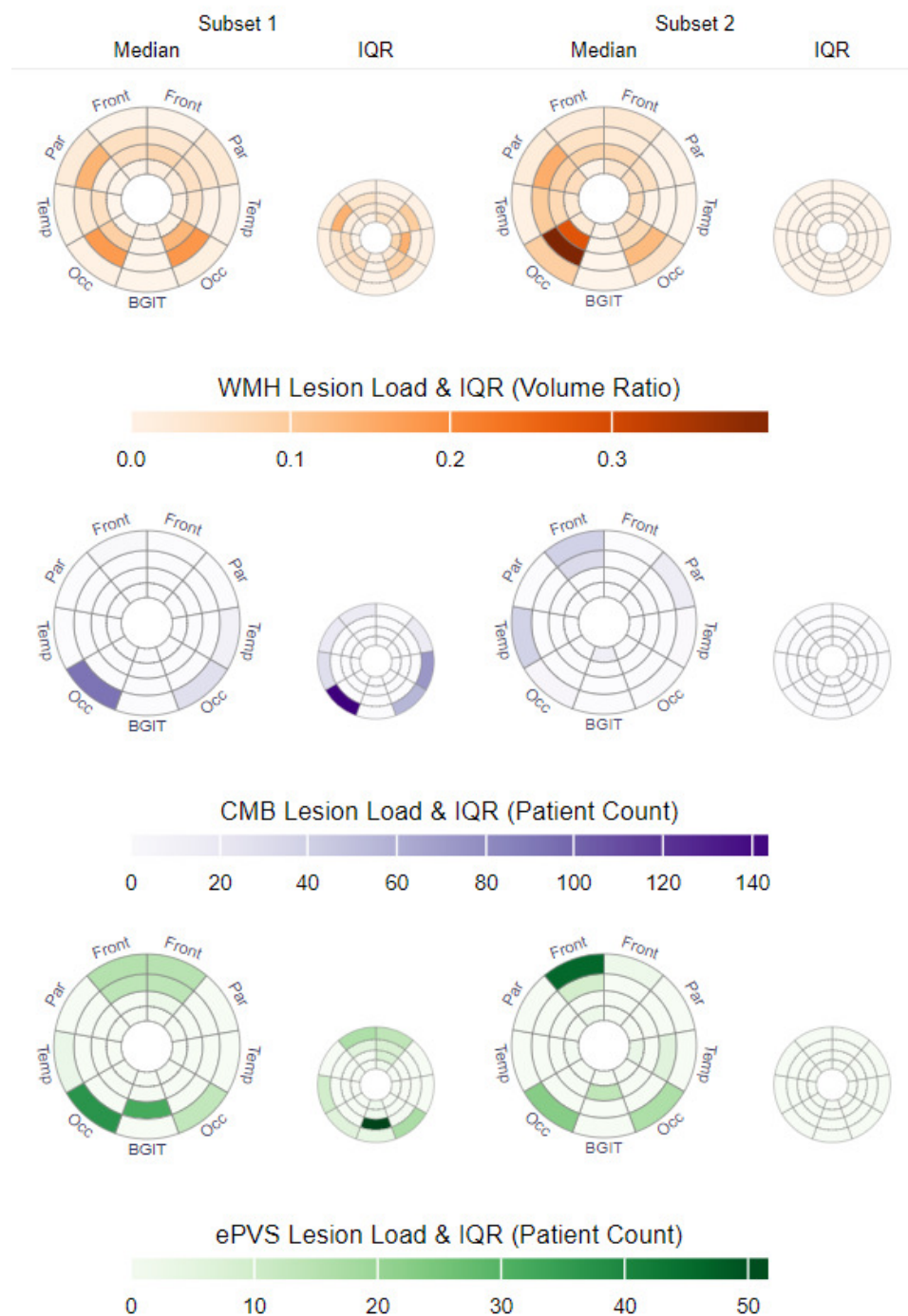


Figure 5.1: A comparison of the lesion loads of the patients with CAA (subset 1, on the left) and mixed CAA (subset 2, on the right).

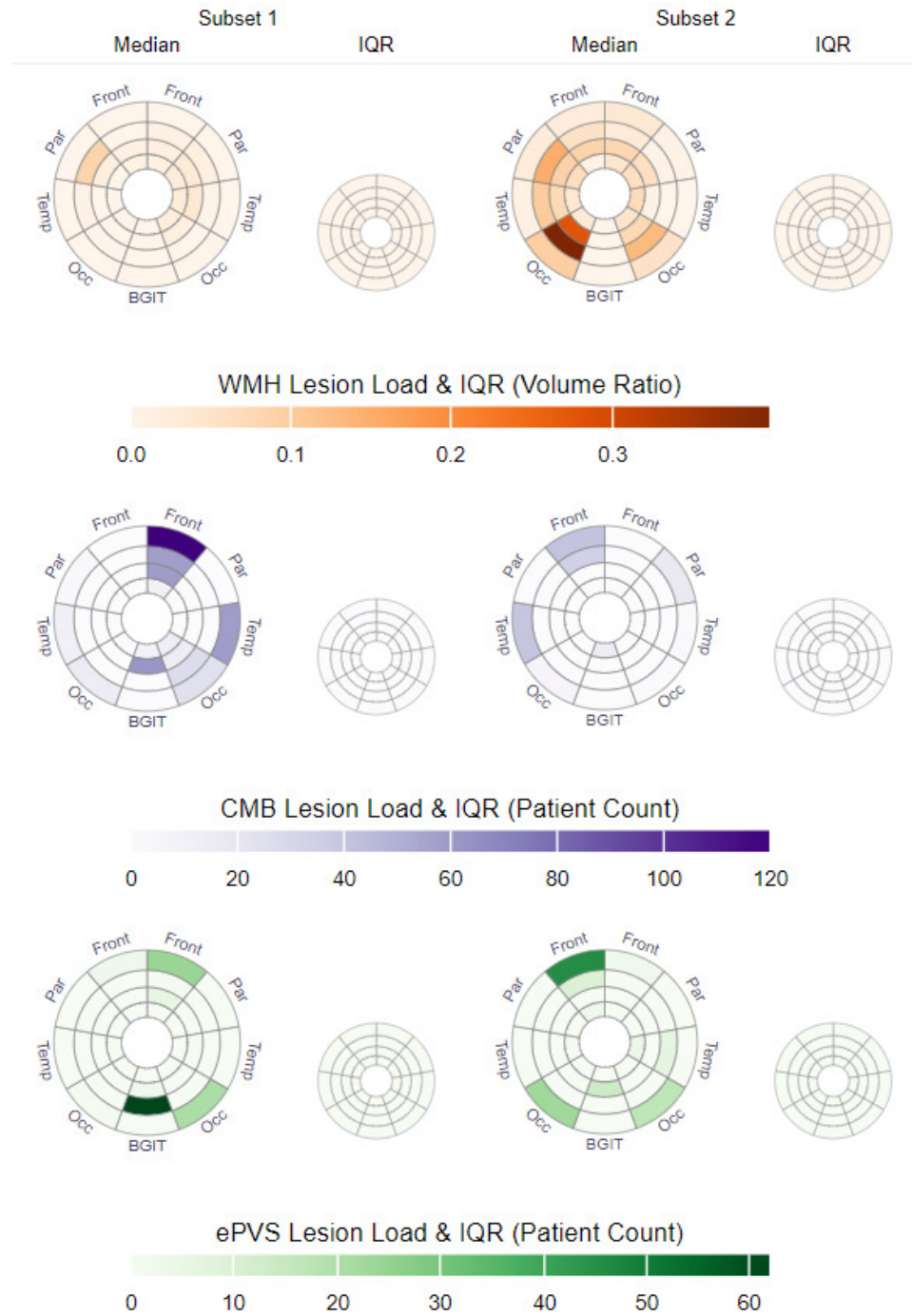
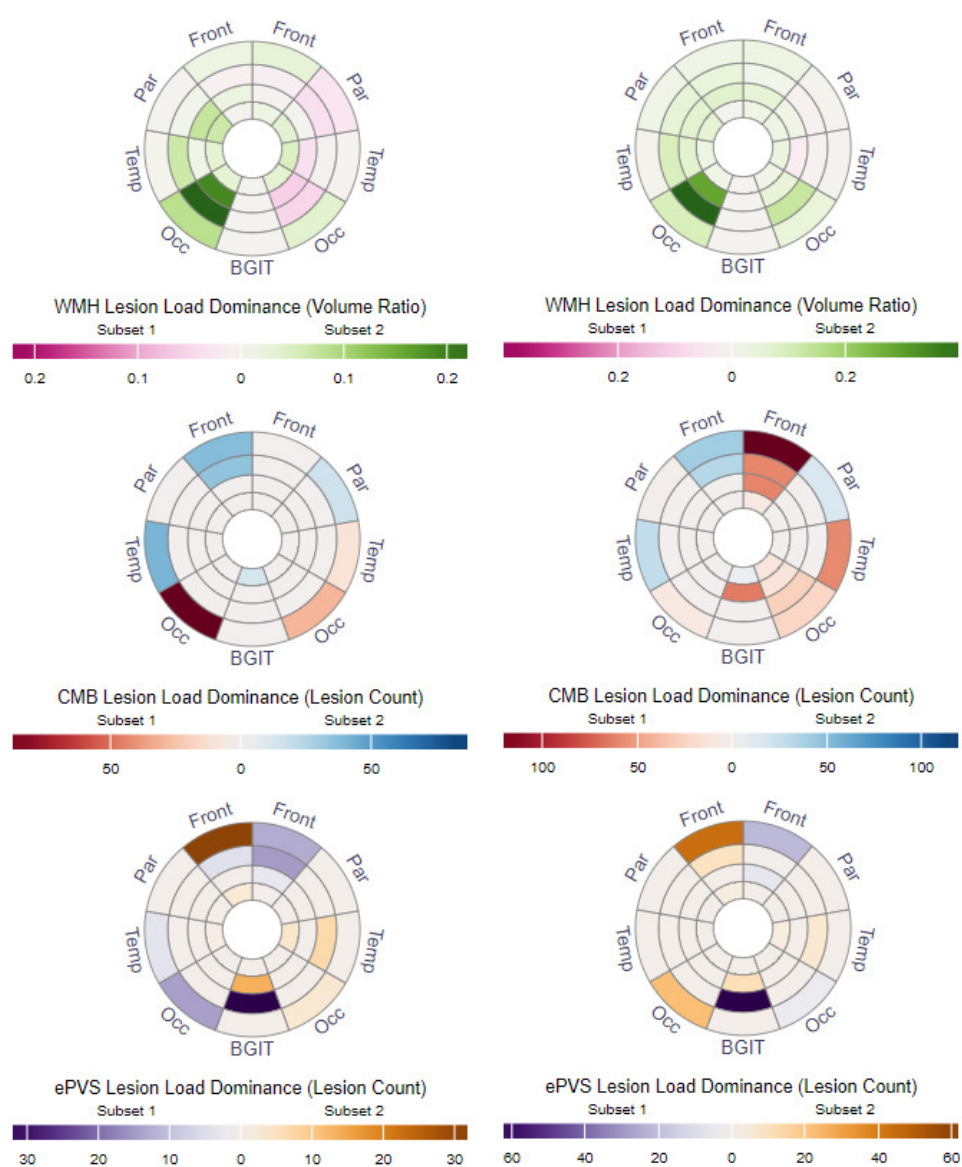


Figure 5.2: A comparison of the lesion loads of the patients with mixed HA (subset 1, on the left) and mixed CAA (subset 2, on the right).



(a) Lesion load dominance plots of the patients with CAA (subset 1) and mixed CAA (subset 2).

(b) Lesion load dominance plots of the patients with mixed HA (subset 1) and mixed CAA (subset 2).

Figure 5.3: Comparison of the lesion load dominance of the patients with (a) CAA and mixed CAA as well as (b) mixed HA and mixed CAA.

6

Conclusion

In the presented work, an application for the analysis of CSVD was developed and integrated into an existing application for the statistical analysis of clinical cohort data. For this purpose, different views for the exploration of individual lesions and the overall lesion distribution as well as an interactive user interface were provided. The results were evaluated by experts from the field of neurology. Furthermore, by analyzing a small data set of patients diagnosed with CSVD, observations about their lesion load distribution could be made.

6.1 Summary

The application allows the analysis of the lesions of one patient, a cohort of patients as well as the comparison of two cohorts or patients. Therefore, the lesions are highlighted in the MRI slices and 3D surface models of the lesions are extracted. This allows the user to investigate the data in detail and focus on individual lesions. Due to the high level of detail, the visualization of the volume data is very complex, and especially accumulated visualizations of multiple patients are easily cluttered. Different interaction techniques allow filtering the data and focus on individual aspects. Since the visualizations are modeled based on the MRI data, the experts find them easy and intuitive to understand. The color scales could be improved to better highlight a high lesion load. Furthermore, the lesion load of voxels that are close together should be accumulated to a joined region to give a more appropriate representation of the lesion load in certain brain areas.

In order to further support the analysis of the lesion distribution, the lesion load in the brain is abstracted and visualized as bullseye plots. The plots provide a standardized representation for the lesion load in different lobar and deep brain regions which is well suited for the comparison of different subsets. In a comparative visualization, the lesion load dominance highlights the differences between two subsets. The plots were well received by the experts. For further analysis of lesion distributions, an equivalent plot for highlighting similarities should also be integrated. Overall, the visualization proved to be intuitive and well-structured and therefore suited for providing an overview of the lesion load as well as lesion load derived parameters. Especially the symmetry of lesion load is very well depicted in the plots.

A main point the experts stated during the evaluation was the inclusion of a statistical analysis of the lesion load. Furthermore, they saw the potential to use the application to analyze lesion load in a longitudinal study or adapt it for other diseases and organs. The general principle of an interactive visual analysis of the medical image data was rated as a very promising approach for further research as well as for diagnostics in clinical practice.

6.2 Future Work

The presented prototype includes a set of functionalities and visualizations useful for the exploration of brain lesions. Potential extensions are discussed in the following.

Integration of additional functionalities and modifications To improve the interaction with the presented application in the future, additional filtering techniques should be provided. This includes filter options for each lesion type in the image slices as well as separate filters for both subsets of the comparative view. Furthermore, the contrast between the colors used for encoding the lesion load is hard to distinguish. An improvement might include a user study to find colors with an optimal contrast. An additional function to export images and views created by the application would be useful to save interesting findings. The option to include additional brain atlases would also allow a more standardized description of those findings. Since all of the data is already registered to the MNI-ICMB152 template, it

would be logical to include the corresponding Cerebrum Atlas (CerebrA) as well (MANERA et al., 2020). The atlas divides the brain into 51 regions per hemisphere and is therefore not only more detailed than the bullseye parcellation but the standardized parcellation allows to compare the findings to the results of other studies.

Limitations of the current implementation Not all of the functionalities suggested by the experts during the evaluation can be easily implemented in the current software. The application was developed using the XTK toolkit (HAEHN et al., 2014). While the toolkit provides a lot of useful functionalities like the rendering of surface models and image slices as well as support for a wide range of data types, using a library also limits the possible functions that can be included in the software. In order to overcome these limitations, a new internal structure for the visualization of volume data could be developed in the future, which is adapted to the individual requirements of the application. On the other hand, the code of the XTK library is publicly available and may be extended to better support the application.

Improving the lesion load analysis During the analysis of the data sets, the neurologists noticed, that the lesion load of the patients is not always symmetrical. This is currently only said to be the case for rare inflammatory cases of CSVD (POHL et al., 2016). A further investigation of the symmetry could be supported by automatically analyzing if the same points in the left and right hemisphere are affected by the same lesion or lesion load. Therefore, affected voxels that are close to each other should be grouped, either by using predefined brain regions like provided by the CerebrA or by giving each voxel a footprint. This could be done by giving each voxel a footprint, similar to the principle of splatting (NEOPHYTOU and MUELLER, 2005). A naive approach would be to smooth the segmentation map using a gaussian filter. Afterwards, the data of multiple patients can be aggregated and low values can be filtered using a threshold. A possible problem of this approach could be, that the smoothing is performed over multiple cerebral convolutions. To prevent this, a mask of the brain, like it is already provided for the MNI-ICBM152 space, can be used to restrain the smoothing directions. As an alternative, a kernel density estimation could

be performed on the lesion maps to calculate the probability of having a lesion at a specific position for a given subset KIM and SCOTT (2012). The grouping of close voxels would also improve the results of the accumulated lesion load that is currently displayed in the application. Especially for small lesions, it is unlikely that the respective voxels will be placed at the exact same position since brains can be shaped very differently. In addition, there are distortions due to the asymmetric non-linear registration. Therefore, lesion voxels are often not congruent, but should still be counted as overlapping.

Smoothing of the surface models Using marching cubes on binary data, such as the segmentation maps, leads to artificial surfaces with a lot of staircase artifacts. However, extracting surface models from the original grayscale images is not possible, since a simple threshold value is not sufficient to extract the lesions. MÖNCH (2014) discussed different approaches to generate smooth anatomical surfaces. Common smoothing techniques are not feasible for medical application, since they alter the geometry leading to inaccuracies. An alternative is context-aware smoothing, which smooths an object depending on the neighboring structures. However, the only existing structures are the lesions themselves and the cortical surface. In most cases, the objects are so close together that hardly any smoothing would be performed. Furthermore, a set of fixed smoothing parameters is not suitable for a wide range of objects of different sizes and shapes such as the lesion models. Visually-guided smoothing, an approach where the user can decide on the smoothing parameters, might be a better solution. Due to the huge differences between the lesions, it would be necessary to apply separate smoothing parameters to, e.g., small and large lesions.

Depiction of distortions The distortions and deformation of the brain regions due to the registration is an interesting parameter, that could be visualized additionally on-demand. This could be represented by a continuous color scale that can be shown on top of the 2D slice images, similar to the visualization of HAN (2011). The deformation of the brain regions is also accompanied by the deformation of the respective lesions. Therefore, the visualization of the deformation can help to correctly assess the lesions, which is particularly important for WMH. The shape of WMH le-

sions is currently investigated in other studies (BLAIR et al., 2020; GWO et al., 2019). Especially the distribution of large areas of WMH and small point-shaped lesions can lead to new insights about different sub-groups of cognitive disease.

Inclusion of statistical analysis The application should be expanded to include a statistical analysis of the lesion load. In neurology, voxel-based morphometry is performed to compare different brains based on extracted parameters like size, shape, or intensity (ASHBURNER and FRISTON, 2000). Therefore, multiple brain images are compared voxel-wise to measure local differences. This principle could be adapted for the lesion maps and visualized in the volume visualization of the presented application. Instead of depicting the patient count, the result would be a statistical parameter map that highlights significant higher or lower lesion loads of one subset. Another approach is to color-code statistical parameters in the volume visualization. JÖNSSON et al. (2020) visualized the spearman correlation of a brain region and clinical parameters in their application by coloring the respective regions in the MRI image slices and a 3D representation of the brain. Alternatively, the statistical significance of a certain sub-cohort having a lesion at the same brain region could be visualized. Again, this could be done voxel-wise, by using the brain regions provided by an atlas or using the broader parcellation used for the bullseye plots. Using the latter method, the results could also be directly visualized in a bullseye plot which might lead to less clutter than color-coding the volume visualization. Instead of using many different plots, a refined bullseye plot could be used to present two parameters (OELTZE et al., 2006). Therefore, statistical parameters could be included in the plot for each parcel. The presentation of two parameters in the same plot facilitates the comparison. Therefore, this technique might also be well suited to compare two subsets. If used for the analysis of a longitudinal study, data of the same patient recorded at different times could be visualized as well.

Clustering of patients with similar lesion load patterns Finding such groups of patients with a similar lesion distribution is of general interest to verify existing classifications, categorize a new case into the known classes or find new intrinsic classes of the data. Therefore, the lesion data could

be projected in a two-dimensional space by a dimensionality reduction technique. This could be realized by deriving features from the data that describe the distribution of lesions in a brain. With these features the dimensionality reduction results in a representation with possible clusters of distribution patterns. JUNG et al. (2021) derived features from WMH data that describe the individual lesions, such as volume, shape features, and intensity profiles. They used t-SNE to reduce the dimensionality of the extracted features and visualize the data in a scatter plot. Following this approach, an interactive scatter plot, like it has been used by RAIDOU et al. (2014) could be included in the application. The scatter plot could then be used as an alternative to the parallel coordinates plot to select and display subsets. The current application highlights the differences in the lesion load of two subsets, which is interesting for finding separate groups, e.g. mixed CAA and mixed HA. On the other hand, to find subsets that should be assigned to a similar group, it would be important to highlight their similarities. Furthermore, only three of the six lesion types associated with CSVD are included in the application. To include more lesion types, no updates of the API have to be made. Only minor changes in the frontend and its interaction workflow are necessary to include additional bullseye plots and legends.

Enhanced evaluation Regardless of the concrete further development of the application, a more detailed study with more patient data and needs to be conducted. Furthermore, experts need to be consulted to validate the results and test for a sufficient accuracy. Therefore, more cases of the CSVD cohort can be processed and the visual results can be compared to the study of SCHEUMANN et al. (2020) who used the same data to perform an analysis of the lesion load distribution. This aspect is also highlighted in the challenges of SHNEIDERMAN et al. (2013) described in Section 3.1. The discussion in Section 4.5.3 shows that the application presented in this work addresses many of the challenges at least in some of their aspects. The only challenge that is not addressed at all in this work is the *Characterizing and understanding similarity*. The challenges *Presenting risk and uncertainty*, as well as *Evaluation*, should be given more attention in the further development of the application. Overall, the application can be easily extended to meet even more points of the challenges in the future.



List of Figures

2.1 Schematic depiction of the different lesion types	6
3.1 Depiction of the the visual analytics process	18
3.2 Principle of the model-view-controller pattern	19
3.3 Exemplary brain visualizations from recent related work . . .	23
3.4 2D illustrative brain visualization from cerebroVis	26
4.1 Schematic depiction of the preprocessing steps	32
4.2 The final user interface	34
4.3 Schematic depiction of the internal workflow	37
4.4 Visualization of the volume data of a single patient	39
4.5 Overview of colors used for different selections	41
4.6 Different bullseye representations of lesion load	44
4.7 Legend of the bullseye parcellation	45
4.8 Selection of a bullseye cell	46
4.9 Aggregation of lesion data for cohort creation	48
4.10 Aggregation of the lesion load dominance	49
4.11 Depiction of the bullseye parcellation	52
4.12 Comparison of different visualization techniques	54
4.13 Comparison of color-coded and texture-coded lesion load . .	56
5.1 Cohort comparison CAA and mixed CAA	69

5.2 Cohort comparison mixed HA and mixed CAA	70
5.3 Comparison of the lesion load dominance	71

B

Bibliography

- [ADAMS et al. 2013] H. H. Adams, M. Cavalieri, B. F. Verhaaren, D. Bos, A. van der Lugt, C. Enzinger, M. W. Vernooij, R. Schmidt and M. A. Ikram. **Rating Method for Dilated Virchow–Robin Spaces on Magnetic Resonance Imaging**. *Stroke*, Vol. 44(6):1732–1735, 2013.
- [ANEURISK-TEAM 2012] **AneuriskWeb project website**, <http://ecm2.mathcs.emory.edu/aneuriskweb>. Web Site.
- [ASHBURNER and FRISTON 2000] J. Ashburner and K. J. Friston. **Voxel-Based Morphometry—The Methods**. *NeuroImage*, Vol. 11(6):805–821, 2000.
- [ATEEQ et al. 2018] T. Ateeq, M. N. Majeed, S. M. Anwar, M. Maqsood, Z. ur Rehman, J. W. Lee, K. Muhammad, S. Wang, S. W. Baik and I. Mehmood. **Ensemble-classifiers-assisted detection of cerebral microbleeds in brain MRI**. *Computers & Electrical Engineering*, Vol. 69:768–781, 2018.
- [BADE et al. 2004] R. Bade, S. Schlechtweg and S. Miksch. **Connecting time-oriented data and information to a coherent interactive visualization**. In: *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*. 2004, ACM Press.
- [BAHL et al. 2016] E. Bahl, T. Koomar and J. J. Michaelson. **cerebroViz: An R package for anatomical visu-alization of spatiotemporal brain data**. *Bioinformatics*, p. btw726, 2016.
- [BALLERINI et al. 2018] L. Ballerini, R. Lovreglio, M. del C. Valdés Hernández, J. Ramirez, B. J. MacIntosh, S. E. Black and J. M. Wardlaw. **Perivas-**

- cular Spaces Segmentation in Brain MRI Using Optimal 3D Filtering.** Scientific Reports, Vol. 8(1), 2018.
- [BEHRENDT et al. 2020] B. Behrendt, S. Voss, O. Beuing, B. Preim, P. Berg and S. Saalfeld. **VICTORIA**. In: Informatik aktuell, pp. 209–214. 2020. Springer Fachmedien Wiesbaden.
- [BEUTHIEN 2014] B. Beuthien. **Restringierte Medizinische Bildregistrierung**. Ph.D. thesis, Institute of Mathematics and Image Computing, University of Lübeck, 2014.
- [BLAIR et al. 2020] G. W. Blair, M. J. Thrippleton, Y. Shi, I. Hamilton, M. Stringer, F. Chappell, D. A. Dickie, P. Andrews, I. Marshall, F. N. Doubal and J. M. Wardlaw. **Intracranial hemodynamic relationships in patients with cerebral small vessel disease**. Neurology, Vol. 94(21):e2258–e2269, 2020.
- [BRETT et al. 2002] M. Brett, I. S. Johnsrude and A. M. Owen. **The problem of functional localization in the human brain**. Nature Reviews Neuroscience, Vol. 3(3):243–249, 2002.
- [BRUGULAT-SERRAT et al. 2019] A. Brugulat-Serrat, , G. Salvadó, C. H. Sudre, O. Grau-Rivera, M. Suárez-Calvet, C. Falcon, G. Sánchez-Benavides, N. Gramunt, K. Fauria, M. J. Cardoso, F. Barkhof, J. L. Molinuevo and J. D. Gispert. **Patterns of white matter hyperintensities associated with cognition in middle-aged cognitively healthy individuals**. Brain Imaging and Behavior, Vol. 14(5):2012–2023, 2019.
- [CABAN and GOTZ 2015] J. J. Caban and D. Gotz. **Visual analytics in healthcare – opportunities and research challenges**. Journal of the American Medical Informatics Association, Vol. 22(2):260–262, 2015.
- [CERQUEIRA et al. 2002] M. D. Cerqueira, N. J. Weissman, V. Dilsizian, A. K. Jacobs, S. Kaul, W. K. Laskey, D. J. Pennell, J. A. Rumberger, T. Ryan and M. S. Verani. **Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart**. Circulation, Vol. 105(4):539–542, 2002.
- [CHARIDIMOU et al. 2017] A. Charidimou, G. Boulouis, M. Pasi, E. Auriel, E. S. van Etten, K. Haley, A. Ayres, K. M. Schwab, S. Martinez-Ramirez, J. N. Goldstein, J. Rosand, A. Viswanathan, S. M. Greenberg and M. E.

- Gurol. **MRI-visible perivascular spaces in cerebral amyloid angiopathy and hypertensive arteriopathy.** *Neurology*, Vol. 88(12):1157–1164, 2017.
- [CHARIDIMOU et al. 2015] A. Charidimou, L. Pantoni and S. Love. **The concept of sporadic cerebral small vessel disease: A road map on key definitions and current concepts.** *International Journal of Stroke*, Vol. 11(1):6–18, 2015.
- [CHOWDHURY et al. 2011] M. H. Chowdhury, A. Nagai, H. Bokura, E. Nakamura, S. Kobayashi and S. Yamaguchi. **Age-Related Changes in White Matter Lesions, Hippocampal Atrophy, and Cerebral Microbleeds in Healthy Subjects Without Major Cerebrovascular Risk Factors.** *Journal of Stroke and Cerebrovascular Diseases*, Vol. 20(4):302–309, 2011.
- [COLE 2019] **Guify**. <https://github.com/colejd/guify>.
- [COLLINS et al. 1999] D. L. Collins, A. P. Zijdenbos, W. F. C. Baaré and A. C. Evans. **ANIMAL+INSECT: Improved Cortical Structure Segmentation.** In: *Lecture Notes in Computer Science*, pp. 210–223. 1999. Springer Berlin Heidelberg.
- [DA CRUZ et al. 2014] L. C. da Cruz, V. D. A. Thomaz and J. C. D. Oliveira. **AICoH 3D: Interactive Atlas of Human Body.** In: *2014 XVI Symposium on Virtual and Augmented Reality*. 2014, IEEE.
- [DEBETTE and MARKUS 2010] S. Debette and H. S. Markus. **The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis.** *BMJ*, Vol. 341(jul26 1):c3666–c3666, 2010.
- [DESTRIEUX et al. 2010] C. Destrieux, B. Fischl, A. Dale and E. Halgren. **Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature.** *NeuroImage*, Vol. 53(1):1–15, 2010.
- [DOS SANTOS and BRODLIE 2004] S. dos Santos and K. Brodlie. **Gaining understanding of multivariate and multidimensional data through visualization.** *Computers & Graphics*, Vol. 28(3):311–325, 2004.
- [DOU et al. 2016] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi and P.-A. Heng. **Automatic Detection of Cerebral Microb-**

- leeds From MR Images via 3D Convolutional Neural Networks.** IEEE Transactions on Medical Imaging, Vol. 35(5):1182–1195, 2016.
- [DOUVEN et al. 2018] E. Douven, P. Aalten, J. Staals, S. H. J. Schievink, R. J. van Oostenbrugge, F. R. J. Verhey and S. Köhler. **Co-occurrence of depressive symptoms and executive dysfunction after stroke: associations with brain pathology and prognosis.** Journal of Neurology, Neurosurgery & Psychiatry, Vol. 89(8):859–865, 2018.
- [DUBOST et al. 2019] F. Dubost, M. Dünnwald, D. Huff, V. Scheumann, F. Schreiber, M. Vernooij, W. Niessen, M. Skalej, S. Schreiber, S. Oeltze-Jafra and M. de Bruijne. **Automated Quantification of Enlarged Perivascular Spaces in Clinical Brain MRI Across Sites.** In: OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging, pp. 103–111. 2019. Springer International Publishing.
- [ENGEL et al. 2006] K. Engel, M. Hadwiger, J. Kniss, C. Rezk Salama and D. Weiskopf. **Sampling a Volume Via Texture Mapping.** In: Real-Time Volume Graphics, pp. 24–36. 2006. ACM Press.
- [EVANS et al. 1993] A. Evans, D. Collins, S. Mills, E. Brown, R. Kelly and T. Peters. **3D statistical neuroanatomical models from 305 MRI volumes.** In: 1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference. 1993, IEEE.
- [EVANS et al. 2012] A. C. Evans, A. L. Janke, D. L. Collins and S. Baillet. **Brain templates and atlases.** NeuroImage, Vol. 62(2):911–922, 2012.
- [EYSENBACH 2008] G. Eysenbach. **Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness.** Journal of Medical Internet Research, Vol. 10(3):e22, 2008.
- [FISCHL 2012] B. Fischl. **FreeSurfer.** NeuroImage, Vol. 62(2):774–781, 2012.
- [FONOV et al. 2009] V. Fonov, A. Evans, R. McKinstry, C. Almli and D. Collins. **Unbiased nonlinear average age-appropriate brain templates from birth to adulthood.** NeuroImage, Vol. 47:S102, 2009.
- [FONOV et al. 2011] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry and D. L. Collins. **Unbiased average age-appropriate atlases for pediatric studies.** NeuroImage, Vol. 54(1):313–327, 2011.

-
- [GHAFOORIAN et al. 2017] M. Ghafoorian, N. Karssemeijer, T. Heskes, I. W. M. van Uden, C. I. Sanchez, G. Litjens, F.-E. de Leeuw, B. van Ginneken, E. Marchiori and B. Platel. **Location Sensitive Deep Convolutional Neural Networks for Segmentation of White Matter Hyperintensities**. Scientific Reports, Vol. 7(1), 2017.
- [GHOLIPOUR et al. 2007] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous and K. Gopinath. **Brain Functional Localization: A Survey of Image Registration Techniques**. IEEE Transactions on Medical Imaging, Vol. 26(4):427–451, 2007.
- [GOEBEL 2000] R. Goebel. **A fast automated method for flattening cortical surfaces**. NeuroImage, Vol. 11(5):S680, 2000.
- [GREENBERG and CHARIDIMOU 2018] S. M. Greenberg and A. Charidimou. **Diagnosis of Cerebral Amyloid Angiopathy**. Stroke, Vol. 49(2):491–497, 2018.
- [GREENBERG et al. 2009] S. M. Greenberg, M. W. Vernooij, C. Cordonnier, A. Viswanathan, R. A.-S. Salman, S. Warach, L. J. Launer, M. A. V. Buchem and M. M. Breteler. **Cerebral microbleeds: a guide to detection and interpretation**. The Lancet Neurology, Vol. 8(2):165–174, 2009.
- [GWO et al. 2019] C.-Y. Gwo, D. C. Zhu and R. Zhang. **Brain White Matter Hyperintensity Lesion Characterization in T2 Fluid-Attenuated Inversion Recovery Magnetic Resonance Images: Shape, Texture, and Potential Growth**. Frontiers in Neuroscience, Vol. 13, 2019.
- [HAEHN 2013] D. Haehn. **Slice:drop**. In: ACM SIGGRAPH 2013 Computer Animation Festival on - SIGGRAPH '13. 2013, ACM Press.
- [HAEHN et al. 2014] D. Haehn, N. Rannou, B. Ahtam, E. Grant and R. Pienaar. **Neuroimaging in the Browser using the X Toolkit**. Frontiers in Neuroinformatics, Vol. 8, 2014.
- [HAJNAL and HILL 2001] J. V. Hajnal and D. L. Hill, Eds. **Medical Image Registration**. CRC Press, 2001.
- [HAN 2011] Z. Han. **Effect of Non-rigid Registration Algorithms on the Analysis of Brain MR Images with Deformation based Morphometry**. Ph.D. thesis, Vanderbilt University, Nashville, Tennessee, 2011.

- [HARROWER and BREWER 2003] M. Harrower and C. A. Brewer. **Color-Brewer.org: An Online Tool for Selecting Colour Schemes for Maps**. *The Cartographic Journal*, Vol. 40(1):27–37, 2003.
- [HESSE et al. 2010] B. W. Hesse, D. Hansen, T. Finholt, S. Munson, W. Kellogg and J. C. Thomas. **Social participation in health 2.0**. *Computer*, Vol. 43(11):45–52, 2010.
- [HINDS et al. 2008] O. Hinds, J. R. Polimeni, N. Rajendran, M. Balasubramanian, L. L. Wald, J. C. Augustinack, G. Wiggins, H. D. Rosas, B. Fischl and E. L. Schwartz. **The Intrinsic Shape of Human and Macaque Primary Visual Cortex**. *Cerebral Cortex*, Vol. 18(11):2586–2595, 2008.
- [HONG et al. 2018] J. Hong, S.-H. Wang, H. Cheng and J. Liu. **Classification of cerebral microbleeds based on fully-optimized convolutional neural network**. *Multimedia Tools and Applications*, Vol. 79(21-22):15151–15169, 2018.
- [HUTH et al. 2016] A. G. Huth, W. A. de Heer, T. L. Griffiths, F. E. Theunissen and J. L. Gallant. **Natural speech reveals the semantic maps that tile human cerebral cortex**. *Nature*, Vol. 532(7600):453–458, 2016.
- [JANKOWIAK 2020] **Three-Orientation-Gizmo**. <https://github.com/jrj2211/three-orientation-gizmo>.
- [JETLEY et al. 2018] S. Jetley, N. A. Lord, N. Lee and P. H. S. Torr. **Learn To Pay Attention**. *CoRR*, Vol. abs/1804.02391, 2018.
- [JÖNSSON et al. 2019] D. Jönsson, A. Bergström, I. Algström, R. Simon, M. Engström, S. Walter and I. Hotz. **Visual Analysis for Understanding Irritable Bowel Syndrome**. In: *Advances in Experimental Medicine and Biology*, pp. 111–122. 2019. Springer International Publishing.
- [JÖNSSON et al. 2020] D. Jönsson, A. Bergström, C. Forsell, R. Simon, M. Engström, S. Walter, A. Ynnerman and I. Hotz. **VisualNeuro: A Hypothesis Formation and Reasoning Application for Multi-Variate Brain Cohort Study Data**. *Computer Graphics Forum*, Vol. 39(6):392–407, 2020.
- [JUNG et al. 2021] K.-H. Jung, K. A. Stephens, K. M. Yochim, J. M. Riphagen, C. M. Kim, R. L. Buckner and D. H. Salat. **Heterogeneity of Cerebral White Matter Lesions and Clinical Correlates in Older Adults**. *Stroke*, Vol. 52(2):620–630, 2021.

-
- [KEIM et al. 2008] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer and G. Melançon. **Visual Analytics: Definition, Process, and Challenges**. In: *Lecture Notes in Computer Science*, pp. 154–175. 2008. Springer Berlin Heidelberg.
- [KEIM et al. 2010] D. Keim, J. Kohlhammer, G. Ellis and F. Mansmann. **Mastering The Information Age – Solving Problems with Visual Analytics**. Eurographics Association, 2010.
- [KEIM et al. 2006] D. Keim, F. Mansmann, J. Schneidewind and H. Ziegler. **Challenges in Visual Data Analysis**. In: *Tenth International Conference on Information Visualisation (IV'06)*. 2006, IEEE.
- [KIM and SCOTT 2012] J. Kim and C. D. Scott. **Robust Kernel Density Estimation**. *J. Mach. Learn. Res.*, Vol. 13(1):2529–2565, 2012.
- [KLEIN et al. 2017] A. Klein, S. S. Ghosh, F. S. Bao, J. Giard, Y. Häme, E. Stavsky, N. Lee, B. Rossa, M. Reuter, E. C. Neto and A. Keshavan. **Mind-boggling morphometry of human brains**. *PLOS Computational Biology*, Vol. 13(2):e1005350, 2017.
- [KLEIN and TOURVILLE 2012] A. Klein and J. Tourville. **101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol**. *Frontiers in Neuroscience*, Vol. 6, 2012.
- [KRASNER and POPE 1988] G. Krasner and S. Pope. **A cookbook for using the model-view controller user interface paradigm in Smalltalk-80**. *Journal of Object-oriented Programming - JOOP*, Vol. 1:26–49, 1988.
- [KREISER et al. 2018] J. Kreiser, M. Meuschke, G. Mistelbauer, B. Preim and T. Ropinski. **A Survey of Flattening-Based Medical Visualization Techniques**. *Computer Graphics Forum*, Vol. 37(3):597–624, 2018.
- [KRUGER and WESTERMANN 2003] J. Kruger and R. Westermann. **Acceleration techniques for GPU-based volume rendering**. In: *IEEE Visualization, 2003. VIS 2003.*, pp. 287–292. 2003, IEEE.
- [KUIJF et al. 2019] H. J. Kuijf, A. Casamitjana, D. L. Collins, M. Dadar, A. Georgiou, M. Ghafoorian, D. Jin, A. Khademi, J. Knight, H. Li, X. Llado, J. M. Biesbroek, M. Luna, Q. Mahmood, R. McKinley, A. Mehrtash, S. Ourselin, B.-Y. Park, H. Park, S. H. Park, S. Pezold, E. Puybareau, J. D.

- Bresser, L. Rittner, C. H. Sudre, S. Valverde, V. Vilaplana, R. Wiest, Y. Xu, Z. Xu, G. Zeng, J. Zhang, G. Zheng, R. Heinen, C. Chen, W. van der Flier, F. Barkhof, M. A. Viergever, G. J. Biessels, S. Andermatt, M. Bento, M. Berseth, M. Belyaev and M. J. Cardoso. **Standardized Assessment of Automatic Segmentation of White Matter Hyperintensities and Results of the WMH Segmentation Challenge**. IEEE Transactions on Medical Imaging, Vol. 38(11):2556–2568, 2019.
- [LACROUTE and LEVOY 1994] P. Lacroute and M. Levoy. **Fast volume rendering using a shear-warp factorization of the viewing transformation**. In: Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH '94. 1994, ACM Press.
- [LESJAK et al. 2017] Ž. Lesjak, A. Galimzianova, A. Koren, M. Lukin, F. Pernuš, B. Likar and Ž. Špiclin. **A Novel Public MR Image Dataset of Multiple Sclerosis Patients With Lesion Segmentations Based on Multi-rater Consensus**. Neuroinformatics, Vol. 16(1):51–63, 2017.
- [LEWIS 2012] J. R. Lewis. **Usability Testing**. In: Handbook of Human Factors and Ergonomics, pp. 1267–1312. 2012. John Wiley & Sons, Inc.
- [LIN et al. 2017] J. Lin, D. Wang, L. Lan and Y. Fan. **Multiple Factors Involved in the Pathogenesis of White Matter Lesions**. BioMed Research International, Vol. 2017:1–9, 2017.
- [LJUNG et al. 2016] P. Ljung, J. Krüger, E. Groller, M. Hadwiger, C. D. Hansen and A. Ynnerman. **State of the Art in Transfer Functions for Direct Volume Rendering**. Computer Graphics Forum, Vol. 35(3):669–691, 2016.
- [LORENSEN and CLINE 1987] W. E. Lorensen and H. E. Cline. **Marching cubes: A high resolution 3D surface construction algorithm**. ACM SIGGRAPH Computer Graphics, Vol. 21(4):163–169, 1987.
- [MANERA et al. 2020] A. L. Manera, M. Dadar, V. Fonov and D. L. Collins. **CerebrA, registration and manual label correction of Mindboggle-101 atlas for MNI-ICBM152 template**. Scientific Data, Vol. 7(1), 2020.
- [MAZZIOTTA et al. 2001] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. L.

- Goualher, D. Boomsma, T. Cannon, R. Kawashima and B. Mazoyer. **A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)**. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, Vol. 356(1412):1293–1322, 2001a.
- [MAZZIOTTA et al. 2001] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. L. Goualher, J. Feidler, K. Smith, D. Boomsma, H. H. Pol, T. Cannon, R. Kawashima and B. Mazoyer. **A Four-Dimensional Probabilistic Atlas of the Human Brain**. Journal of the American Medical Informatics Association, Vol. 8(5):401–430, 2001b.
- [MEDIAIRE GMBH 2021] **mdbrain**. <https://mediaire.de/product/mdbrain/>.
- [MEYER et al. 2016] J. Meyer, A. Kazakova, M. Büsing and S. Boll. **Visualization of Complex Health Data on Mobile Devices**. In: Proceedings of the 2016 ACM Workshop on Multimedia for Personal Health and Health Care. 2016, ACM.
- [MULLER et al. 2021] J. Muller, L. A. Garrison, P. Ulbrich, S. Schreiber, S. Bruckner, H. Hauser and S. Oeltze-Jafra. **Integrated Dual Analysis of Quantitative and Qualitative High-Dimensional Data**. IEEE Transactions on Visualization and Computer Graphics, pp. 1–1, 2021.
- [MÖNCH 2014] T. Mönch. **Context-Aware 3D Model Generation for Biomedical Applications**. Ph.D. thesis, Otto-von-Guericke University Magdeburg, 2014.
- [NEOPHYTOU and MUELLER 2005] N. Neophytou and K. Mueller. **GPU accelerated image aligned splatting**. In: Fourth International Workshop on Volume Graphics, 2005. 2005, IEEE.
- [OELTZE et al. 2006] **Integrated Visualization of Morphologic and Perfusion Data for the Analysis of Coronary Artery Disease**.
- [PANTONI and GORELICK 2014] L. Pantoni and P. B. Gorelick. **Cerebral Small Vessel Disease**. Cambridge University Press, 2014.

- [PASI et al. 2017] M. Pasi, A. Charidimou, G. Boulouis, E. Auriel, A. Ayres, K. M. Schwab, J. N. Goldstein, J. Rosand, A. Viswanathan, L. Pantoni, S. M. Greenberg and M. E. Gurol. **Mixed-location cerebral hemorrhage/microbleeds**. *Neurology*, Vol. 90(2):e119–e126, 2017.
- [PENNY et al. 2007] W. Penny, K. Friston, J. Ashburner, S. Kiebel and T. Nichols. **Statistical Parametric Mapping: The Analysis of Functional Brain Images**. Elsevier, 2007.
- [PHONG 1975] B. T. Phong. **Illumination for computer generated pictures**. *Communications of the ACM*, Vol. 18(6):311–317, 1975.
- [POELS et al. 2012] M. M. F. Poels, M. A. Ikram, A. van der Lugt, A. Hofman, W. J. Niessen, G. P. Krestin, M. M. B. Breteler and M. W. Vernooij. **Cerebral microbleeds are associated with worse cognitive function: The Rotterdam Scan Study**. *Neurology*, Vol. 78(5):326–333, 2012.
- [POHL et al. 2016] D. Pohl, G. Alper, K. V. Haren, A. J. Kornberg, C. F. Lucchinetti, S. Tenenbaum and A. L. Belman. **Acute disseminated encephalomyelitis**. *Neurology*, Vol. 87(9 Supplement 2):S38–S45, 2016.
- [POTTER et al. 2015] G. M. Potter, F. M. Chappell, Z. Morris and J. M. Wardlaw. **Cerebral Perivascular Spaces Visible on Magnetic Resonance Imaging: Development of a Qualitative Rating Scale and its Observer Reliability**. *Cerebrovascular Diseases*, Vol. 39(3-4):224–231, 2015.
- [POTTER et al. 2013] G. M. Potter, F. N. Doubal, C. A. Jackson, F. M. Chappell, C. L. Sudlow, M. S. Dennis and J. M. Wardlaw. **Enlarged Perivascular Spaces and Cerebral Small Vessel Disease**. *International Journal of Stroke*, Vol. 10(3):376–381, 2013.
- [PREIM and BOTHA 2014] B. Preim and C. Botha. **Direct Volume Visualization**. In: *Visual Computing for Medicine*, pp. 269–287. 2014. Elsevier.
- [PREIM and BOTHA 2014] B. Preim and C. Botha. **Introduction**. In: *Visual Computing for Medicine*, pp. 1–11. 2014. Elsevier.
- [PREIM and BOTHA 2014] B. Preim and C. Botha. **Projections and Reformations**. In: *Visual Computing for Medicine*, pp. e1–e17. 2014. Elsevier.

-
- [PREIM and LAWONN 2019] B. Preim and K. Lawonn. **A Survey of Visual Analytics for Public Health**. Computer Graphics Forum, Vol. 39(1):543–580, 2019.
- [RAIDOU et al. 2014] R. G. Raidou, M. P. Moreira, W. van Elmpt, M. Breeuwer and A. Vilanova. **Visual analytics for the exploration of multiparametric cancer imaging**. In: 2014 IEEE Conference on Visual Analytics Science and Technology (VAST). 2014, IEEE.
- [ROBERTS 2007] J. C. Roberts. **State of the Art: Coordinated & Multiple Views in Exploratory Visualization**. In: Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007). 2007, IEEE.
- [RODRÍGUEZ et al. 2014] M. B. Rodríguez, E. Gobbetti, J. I. Guitián, M. Makhinya, F. Marton, R. Pajarola and S. Suter. **State-of-the-Art in Compressed GPU-Based Direct Volume Rendering**. Computer Graphics Forum, Vol. 33(6):77–100, 2014.
- [ROTH et al. 2014] L. Roth, T. Kaffenberger, U. Herwig and A. B. Brühl. **Brain Activation Associated with Pride and Shame**. Neuropsychobiology, Vol. 69(2):95–106, 2014.
- [RUDOLPH 2011] P. Rudolph. **Realtime Visualization of the Connectome in the Browser using WebGL**. Frontiers in Neuroinformatics, Vol. 5, 2011.
- [SANROMA 2019] **Bullseye Pipeline**. https://github.com/gsanroma/bullseye_pipeline.
- [SCHEUMANN et al. 2020] V. Scheumann, F. Schreiber, V. Perosa, A. Assmann, C. Mawrin, C. Garz, H.-J. Heinze, M. Görtler, E. Düzel, S. Vielhaber, A. Charidimou and S. Schreiber. **MRI phenotyping of underlying cerebral small vessel disease in mixed hemorrhage patients**. Journal of the Neurological Sciences, Vol. 419:117173, 2020.
- [SHAABAN and MOLAD 2020] C. E. Shaaban and J. Molad. **Cerebral small vessel disease**. Neurology, Vol. 94(21):909–910, 2020.
- [SHATTUCK et al. 2008] D. W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. L. Narr, R. A. Poldrack, R. M. Bilder and A. W.

- Toga. **Construction of a 3D probabilistic atlas of human cortical structures.** *NeuroImage*, Vol. 39(3):1064–1080, 2008.
- [SHNEIDERMAN 1996] B. Shneiderman. **The eyes have it: a task by data type taxonomy for information visualizations.** In: *Proceedings 1996 IEEE Symposium on Visual Languages*, 1996, pp. 336–343.
- [SHNEIDERMAN 1996] 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, IEEE Comput. Soc. Press.
- [SHNEIDERMAN et al. 2013] B. Shneiderman, C. Plaisant and B. W. Hesse. **Improving Healthcare with Interactive Visualization.** *Computer*, Vol. 46(5):58–66, 2013.
- [SIULY and ZHANG 2016] S. Siuly and Y. Zhang. **Medical Big Data: Neurological Diseases Diagnosis Through Medical Data Analysis.** *Data Science and Engineering*, Vol. 1(2):54–64, 2016.
- [SMITH et al. 2010] E. E. Smith, K. R. Nandigam, Y.-W. Chen, J. Jeng, D. Salat, A. Halpin, M. Frosch, L. Wendell, L. Fazen, J. Rosand, A. Viswanathan and S. M. Greenberg. **MRI Markers of Small Vessel Disease in Lobar and Deep Hemispheric Intracerebral Hemorrhage.** *Stroke*, Vol. 41(9):1933–1938, 2010.
- [STREIT and GEHLENBORG 2014] M. Streit and N. Gehlenborg. **Bar charts and box plots.** *Nature Methods*, Vol. 11(2):117–117, 2014.
- [SUDRE et al. 2018] C. Sudre, B. G. Anson, I. Davagnanam, A. Schmitt, A. Mendelson, F. Prados, L. Smith, D. Atkinson, A. Hughes, N. Chaturvedi, M. Cardoso, F. Barkhof, H. Jaeger and S. Ourselin. **Bullseye's representation of cerebral white matter hyperintensities.** *Journal of Neuroradiology*, Vol. 45(2):114–122, 2018.
- [SUDRE et al. 2019] C. Sudre, B. Gomez Anson, S. Ingala, C. Lane, D. Jimenez, L. Haider, T. Varsavsky, L. Smith, S. Ourselin, R. Jäger and M. Cardoso. **3D multirater RCNN for multimodal multiclass detection and characterisation of extremely small objects.** In: M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal and T. Vercauteren, Eds., *Proceedings of The 2nd International Conference on*

- Medical Imaging with Deep Learning, Vol. 102 of Proceedings of Machine Learning Research, pp. 447–456. 2019, PMLR, London, United Kingdom.
- [TALAIRACH J 1988] T. P. Talairach J. **Co-planar stereotaxic atlas of the human brain**. Thieme, 1988.
- [TERMEER et al. 2007] M. Termeer, J. O. Bescós, M. Breeuwer, A. Vilanova, F. A. Gerritsen and M. E. Gröller. **CoViCAD: Comprehensive Visualization of Coronary Artery Disease**. IEEE Trans. Vis. Comput. Graph., Vol. 13(6):1632–1639, 2007.
- [TOMANDL et al. 2001] B. F. Tomandl, P. Hastreiter, C. Rezk-Salama, K. Engel, T. Ertl, W. J. Huk, R. Naraghi, O. Ganslandt, C. Nimsky and K. E. W. Eberhardt. **Local and Remote Visualization Techniques for Interactive Direct Volume Rendering in Neuroradiology**. RadioGraphics, Vol. 21(6):1561–1572, 2001.
- [TOMINSKI and SCHUMANN 2020] C. Tominski and H. Schumann. **Interactive Visual Data Analysis**. AK Peters Visualization Series. CRC Press, 2020.
- [TSAI et al. 2019] H.-H. Tsai, M. Pasi, L.-K. Tsai, Y.-F. Chen, B.-C. Lee, S.-C. Tang, P. Fotiadis, C.-Y. Huang, R.-F. Yen, J.-S. Jeng and M. E. Gurol. **Microangiopathy underlying mixed-location intracerebral hemorrhages/microbleeds**. Neurology, Vol. 92(8):e774–e781, 2019.
- [TURKAY et al. 2011] C. Turkay, P. Filzmoser and H. Hauser. **Brushing Dimensions - A Dual Visual Analysis Model for High-Dimensional Data**. IEEE Transactions on Visualization and Computer Graphics, Vol. 17(12):2591–2599, 2011.
- [VAN SOMEREN et al. 1994] M. W. Van Someren, Y. F. Barnard, J. A. C. Sandberg and J. A. C. Sandberg. **The Think Aloud Method: A Practical Approach to Modelling Cognitive Processes**. Academic Press, 1994.
- [VERNOOIJ 2009] M. W. Vernooij. **Use of Antithrombotic Drugs and the Presence of Cerebral Microbleeds**. Archives of Neurology, Vol. 66(6):714, 2009.

- [WANDELL et al. 2000] B. Wandell, S. Chial, B. Backus, D. Fleet, D. Heeger, W. Press, G. Sapiro and A. Wade. **Visualization and Measurement of the Cortical Surface**. *Journal of cognitive neuroscience*, Vol. 12(5):739–752, 2000.
- [WANG et al. 2012] Y. Wang, W. Dou and J.-M. Constans. **Accelerating volume ray casting by empty space skipping used for Computer-Aided Therapy**. In: 2012 International Conference on Audio, Language and Image Processing. 2012, IEEE.
- [WANG et al. 2014] Z. Wang, Y. O. Soo and V. C. Mok. **Cerebral Microbleeds: is antithrombotic therapy safe to administer?** *Stroke*, Vol. 45(9):2811–2817, 2014.
- [WARDLAW et al. 2013] J. M. Wardlaw, E. E. Smith, G. J. Biessels, C. Cordonnier, F. Fazekas, R. Frayne, R. I. Lindley, J. T. O'Brien, F. Barkhof, O. R. Benavente, S. E. Black, C. Brayne, M. Breteler, H. Chabriat, C. DeCarli, F.-E. de Leeuw, F. Doubal, M. Duering, N. C. Fox, S. Greenberg, V. Hachinski, I. Kilimann, V. Mok, R. van Oostenbrugge, L. Pantoni, O. Speck, B. C. M. Stephan, S. Teipel, A. Viswanathan, D. Werring, C. Chen, C. Smith, M. van Buchem, B. Norrving, P. B. Gorelick and M. Dichgans. **Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration**. *The Lancet Neurology*, Vol. 12(8):822–838, 2013.
- [WEIDAUER et al. 2020] S. Weidauer, M. Wagner and E. Hattingen. **White Matter Lesions in Adults – a Differential Diagnostic Approach**. *RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, Vol. 192(12):1154–1173, 2020.
- [WESTOVER 1990] L. Westover. **Footprint evaluation for volume rendering**. In: Proceedings of the 17th annual conference on Computer graphics and interactive techniques - SIGGRAPH '90. 1990, ACM Press.
- [WHITFIELD-GABRIELI and NIETO-CASTANON 2012] S. Whitfield-Gabrieli and A. Nieto-Castanon. **Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks**. *Brain Connectivity*, Vol. 2(3):125–141, 2012.
- [WILSON et al. 2016] D. Wilson, A. Charidimou, G. Ambler, Z. V. Fox, S. Gregoire, P. Rayson, T. Imaizumi, F. Fluri, H. Naka, S. Horstmann,

- R. Veltkamp, P. M. Rothwell, V. I. Kwa, V. Thijs, Y.-S. Lee, Y. D. Kim, Y. Huang, K. S. Wong, H. R. Jäger and D. J. Werring. **Recurrent stroke risk and cerebral microbleed burden in ischemic stroke and TIA.** *Neurology*, Vol. 87(14):1501–1510, 2016.
- [YOUSEFI et al. 2012] S. Yousefi, N. Kehtarnavaz and A. Gholipour. **Improved Labeling of Subcortical Brain Structures in Atlas-Based Segmentation of Magnetic Resonance Images.** *IEEE Transactions on Biomedical Engineering*, Vol. 59(7):1808–1817, 2012.

Statement of Authorship

Thesis: Visual Analysis of Brain Lesion Load
in Patients with Cerebral Small Vessel Disease

Surname, Name: Mittenentzwei, Sarah

Matr. No.: 213105

Date of Birth: 02.02.1996

I herewith assure that I wrote the present thesis independently, that the thesis has not been partially or fully submitted as graded academic work and that I have used no other means than the ones indicated. I have indicated all parts of the work in which sources are used according to their wording or to their meaning.

I am aware of the fact that violations of copyright can lead to injunctive relief and claims for damages of the author as well as a penalty by the law enforcement agency.

Magdeburg, 27th March 2021

.....
(Signature)