

RadEx: Integrated Visual Exploration of Multiparametric Studies for Radiomic Tumor Profiling

E. Mörth^{1,2} , K. Wagner-Larsen^{2,3} , E. Hodneland^{2,4} , C. Krakstad^{2,5} , I. S. Haldorsen^{2,3} , S. Bruckner^{1,2} , and N. N. Smit^{1,2} 

¹Department of Informatics, University of Bergen, Norway

²Mohn Medical Imaging and Visualization Centre, Department of Radiology, Haukeland University Hospital, Norway

³Department of Clinical Medicine, University of Bergen, Norway

⁴Norwegian Research Centre (NORCE), Bergen, Norway

⁵Centre for Cancer Biomarkers, Department of Clinical Science, University of Bergen

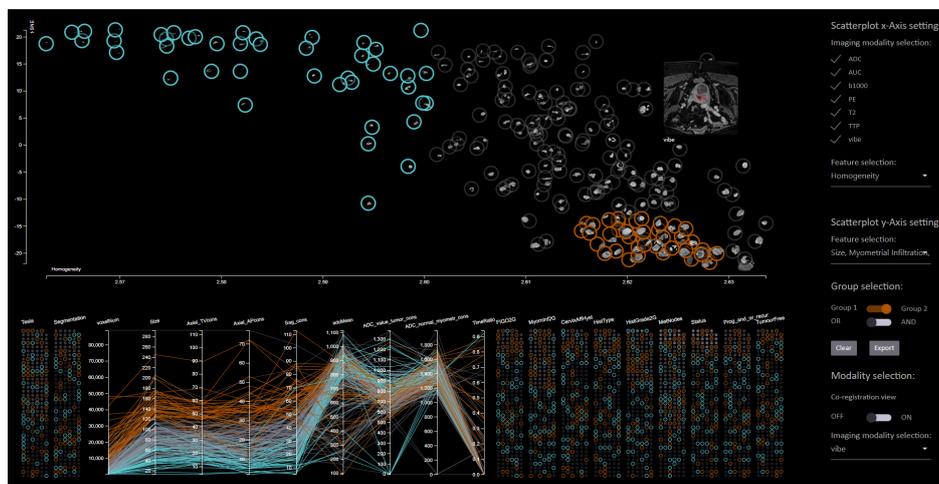


Figure 1: Overview of the RadEx interface, an integrated analysis and exploration platform for gynecological cancer data.

Abstract

Better understanding of the complex processes driving tumor growth and metastases is critical for developing targeted treatment strategies in cancer. Radiomics extracts large amounts of features from medical images which enables radiomic tumor profiling in combination with clinical markers. However, analyzing complex imaging data in combination with clinical data is not trivial and supporting tools aiding in these exploratory analyses are presently missing. In this paper, we present an approach that aims to enable the analysis of multiparametric medical imaging data in combination with numerical, ordinal, and categorical clinical parameters to validate established and unravel novel biomarkers. We propose a hybrid approach where dimensionality reduction to a single axis is combined with multiple linked views allowing clinical experts to formulate hypotheses based on all available imaging data and clinical parameters. This may help to reveal novel tumor characteristics in relation to molecular targets for treatment, thus providing better tools for enabling more personalized targeted treatment strategies. To confirm the utility of our approach, we closely collaborate with experts from the field of gynecological cancer imaging and conducted an evaluation with six experts in this field.

CCS Concepts

• **Applied computing** → **Health informatics**; • **Human-centered computing** → **Visualization design and evaluation methods**;

1. Introduction

The World Health Organization announced in 2018 that cancer is globally the second leading cause of death after cardiovascular dis-

ease [BFS* 18]. There are numerous forms of cancer and they arise in all kinds of cells in the human body. When exploring imaging features, radiomic tumor profiling may be performed [BGE* 18].

Gillies et al. state that the goal of radiomics is to harvest high dimensional data from clinical images which serve as a basis for further analysis, e.g., in terms of predictive value for predicting clinical outcome and response to targeted therapy [GKH16]. This analysis is by no means an easy task and often involves working with multiple modalities or multiple parametric images. For endometrial cancer, which is the most common gynecological tumor in industrialized countries [AMN*05], preoperative staging by multiparametric magnetic resonance images (MRI) and results from endometrial biopsy routinely guide choice of surgical procedure and adjuvant therapy.

Tumor regions of interest (ROIs) can be manually placed on the different MRI sequences to quantify tissue features (e.g., diffusion properties on apparent diffusion coefficient (ADC) maps, perfusion markers on dynamic contrast enhanced (DCE)-MRI/parametric maps), and these tumor characteristics have been shown to aid in predicting metastases or tumor progression [FBYH*18, HSG*14, HS16, BFM*16]. However, manual tumor segmentations and analyses of ROI data are time consuming, and thus presently unfeasible in daily routine. A main challenge when analyzing cohorts of patients is to unravel the most relevant imaging patterns that are linked to clinical parameters and patient outcome. Furthermore, analysis of the high dimensional value domain resulting from the combination of the clinical parameters consisting of histological markers, radiological findings, and outcome related parameters and multiparametric imaging data is challenging.

In current clinical research, imaging modalities are typically analyzed individually, and the tumors are analyzed based on manually placed ROIs. As a part of ongoing research of our collaborators, convolutional neural networks are used to perform a machine learning based segmentation of tumors. In addition, an automatic co-registration of all available modalities is performed. This allows integrative analysis of all voxels of the tumor, with the potential to empower clinical researchers with more targeted and capable analysis platforms. Tumor textural features based on manually placed ROIs and only for single modalities have already been proven to be related to high-risk histological subtypes in endometrial cancer [HSG*14, FBYH*18, YHDL*18]. Analysis tools for different ROI measurements taking multiple modalities and all tumor voxels into account are not available. Nonetheless such an analysis would likely improve comparability between different patients and across different hospitals.

To further support cohort analysis in radiomic tumor profiling, we provide a tool which supports both data preprocessing steps for cohort analysis as well as an integrated dashboard for cohort analysis and hypothesis generation. The application targets cancer imaging research where radiomic tumor profiling is carried out and where multiparametric imaging data is acquired. We provide an interactive analysis platform that enables gynecological cancer researchers to analyze cohort data with the goal of hypothesis formation. The overall aim of the application is to provide a tool for visualization, exploration, and identification of complex relations between radiomic tumor profiles and clinical and histological markers.

Our main contributions are the following: (1) We present an interactive cohort analysis application that targets hypothesis forma-

tion in radiomic tumor profiling workflows that includes imaging data and clinical parameters. (2) We provide a workflow that enables validation of manual or automatic machine learning-based tumor segmentations and validation of automatic co-registration of multi-parametric imaging data. (3) To show the utility of our approach, we evaluated our system with six experts in gynecological cancer imaging research, using the System Usability Scale (SUS) [Bro04] and a qualitative evaluation.

2. Medical Background

Endometrial cancer is the most common pelvic gynecological malignancy in high-income countries. The endometrium comprises the innermost lining of the uterine cavity, and patients typically experience abnormal vaginal bleeding. The diagnosis is confirmed by an endometrial biopsy establishing the histological-/molecular tumor subtype/grade, and subsequently a preoperative pelvic MRI is routinely performed for local staging. However, analyzing complex imaging data in combination with clinical-/tissue data is not trivial. Imaging data is often multiparametric, allowing the visualization of different aspects of tumor physiology related to, e.g., tumor microcirculation and microstructure, which reportedly are closely linked to the observed clinical phenotype in cancer [HSG*14, FBYH*18]. Extracting whole volume multiparametric imaging data simultaneously may also be done through radiomic tumor profiling.

Radiomic tumor profiling plays an emerging important role in the new era of precision medicine [GKH16]. This field of research aims to produce high-dimensional feature vectors from clinical images [RBR*18] to find tumor markers with a higher predictive value. The typical workflow consists of image acquisition, image reconstruction, tumor segmentation, feature extraction and qualification, and finally analysis and model building. One challenge in this regard is feature selection which may be performed a priori or from dimensionality reduction. Radiomics features are extracted from a region of interest representing the tumor. Hence, a high quality tumor segmentation is a crucial step [RBR*18]. Manual segmentation is often used as the ground truth for tumor segmentation although some inter-reader variability is inevitably present [RBR*18].

Tumor texture parameters derived from MRI scans have been shown to be associated with high-risk disease and reduced survival in endometrial cancer [YHDL*18]. Parameters like kurtosis, entropy and mean of positive pixels (MPP) from the ADC map and the T1 contrast-enhanced images have been shown to predict high-risk histological subtypes and advanced stage, e.g., deep myometrial invasion in endometrial cancer [YHDL*18, FBYH*18]. Furthermore, tumor tissue markers have been compared with tumor markers from parametric maps (based on dynamic contrast-enhanced MRI) finding that reduced tumor blood flow in MRI reflects increased microvascular proliferation in the tumor samples and predicts poor survival in endometrial cancer [HSG*14, FBYH*18].

3. Related Work

Traditionally, gynecological cancer biomarker research employs statistical analysis tools like SPSS [GS10] or RStudio [RS10]. Visualization approaches for targeted cancer diagnosis often focuses on techniques for representing biomarkers or imaging data but there is only a small body of work combining clinical cohort data with

multiple imaging sequences per patient. Raidou et al. [RVD*15] introduced a visual analytics approach for the exploration of tumor tissue characterization featuring a 2D t-Distributed Stochastic Neighbor Embedding (t-SNE) [vH08] dimensionality reduction also taking tumor characteristics into account. Although their approach is similar, we differ in multiple aspects in that we do not focus on pharmacokinetic parameter maps but rather visualize high-dimensional radiomic tumor texture features. In contrast to their approach, we use a 1D t-SNE dimensionality reduction and our visual exploration approach aims for the identification of complex relations between radiomic tumor profile and clinical and histological markers. To the best of our knowledge, there is no related literature where a 1D t-SNE dimensionality reduction has been used in combination with a clinically meaningful parameter on the other axis. Compared to the other approaches presented here, we preserve one axis of the 2D visualization to present a meaningful value selection, which is not the case when applying a 2D t-SNE dimensionality reduction.

Image-Centric Cohort Visualization Closely related to our approach is the work by Steenwijk et al. [SMB*10], where the authors used multiple linked views including scatterplots and parallel coordinate plots combined with imaging data for each patient. One major difference is that they are limited to displaying two images per patient without a link to the imaging data in their visualization application. Klemm et al. [KOJL*14] introduced an epidemiological approach which is also image-centric and involves segmentations and hypothesis formulation. However, they do not visualize the original imaging data in their application and mainly focus on model-based visualization in their application. Jönsson et al. [JBF*19] introduced a cohort analysis platform which also allows for group comparison and incorporates imaging data for each patient as well as clinical parameters. In contrast to our approach, they do not use multiparametric imaging data, and they do not work with radiomic tumor features.

Visual Cohort Analysis Preim et al. [PL20] provided an extensive overview of visual analytics approaches for public health data in their survey. Further related work in the field of visual analytics of patient cohorts includes the work of Angelelli et al. [AOH*14]. The authors presented a prototype aiming for cohort-based hypothesis formulation for heterogeneous data. Eckelt et al. [EAZ*19] presented a visual analysis tool enabling statistical analysis of tabular data, cancer drug target discovery, and closing the gap between visualization and statistical analysis. Raidou et al. [RCMA*18] introduced a visual analytics application which allows for analysis on both cohort level and patient level for radiotherapy induced bladder toxicity. Bernard et al. [BSM*15] presented a data centered approach for analyzing large amounts of patients using multiple linked views and selective analysis. When dealing with large amounts of patient data, user guidance could be implemented, as discussed by Ceneda et al. [CGM*17, CGM19]. Further related work in the field of cohort construction includes the work of Krause et al. [KPS16]. In the field of cancer characterization, the work of Turkay et al. [TLS*14] is relevant to our approach. In contrast to these approaches, we combine multiparametric imaging data, radiomic tumor profiling data and clinical parameters in one application.

Co-Registration and Segmentation Validation As we work with co-registered data which partially features machine learning-based segmentation masks, the data needs to be validated by experts before further analysis. Hastreiter et al. [HE98] proposed to use fused visualization methods. Jenkins et al. [JBBS02] suggested that overlaying of prominent edges provides a usable co-registration check. More complex and automatic co-registration validation methods include the work of Schnabel et al. [STCS*03], which validates nonrigid image registration using finite-element methods. Our approach adds a parameter-based pre-selection of cases where the co-registration outcome is suspicious to first present these cases to medical researchers for further analysis.

Automatic segmentation methods are available, but often they do not meet the acceptance criteria needed for usage in clinical cohort studies. Therefore, a validation step has to be employed before using them as a source of analysis [VBK*13]. Von Landsberger et al. [VBK*13] introduced a user guided automatic segmentation method where algorithm parameters are set intuitively by using visual analytics tools. Karimov et al. [KMAB15] introduced an approach for interactive segmentation correction based on histogram dissimilarities. Haehn et al. [HKT*18] proposed a segmentation proofreading technique and demonstrated that expert proofreading has increased performance and speed over manual expert segmentation. In contrast to these approaches, we benefit from various features measured by radiologists which allow us to spot potentially incorrect automatic segmentation. Adding these cases to the training set of automatic algorithms allows for an incremental improvement of the segmentation mask result.

4. Data and Tasks

Endometrial cancer classification is highly dependent on the data used. For our clinical collaborators, multiparametric imaging data and histological data as well as further clinical parameters of the patients are available. Due to our close collaboration with both radiologists and experts in the field of molecular biomarkers we were able to gather insight in this highly specialized field of research. In a clinical setting, potential patients who typically face symptoms such as vaginal bleeding, have an endometrium biopsy which serves as the basis for a histological investigation. If the biopsy confirms an endometrial cancer diagnosis, preoperative pelvic multiparametric contrast-enhanced MRI is routinely performed. Imaging findings guide the choice of treatment, normally consisting of surgery in all cases, followed by adjuvant chemo- and/or radiation therapy in high-risk patients. After treatment, the patients have regular clinical follow-ups to detect recurrent disease/tumor progression, and progression-free survival is recorded.

4.1. Clinical Parameters

The following clinical parameters are available for analysis:

- Tesla: Field strength value of the MRI scanner used for the screening, either 1.5 or 3 Tesla.
- Segmentation: Indicator of manual or machine learning segmentation of the tumor.
- FIGO2G: International Federation of Gynecology and Obstetrics (FIGO) [BBCF*19] classification of the tumor. FIGO I and II are one group and FIGO III and IV represent the other group. The grouping is performed based on tumor aggressiveness.

- MyomInf2G: Myometrial infiltration of the tumor with infiltration of $\leq 50\%$ and $>50\%$.
- CervixAffHyst: Tumor extending to the uterine cervix.
- HistType: Endometrial or non-endometrial subtype.
- HistGrade2G: High-grade and low-grade tumors.
- MetNodes: Histologically confirmed lymph node metastasis, no metastases or not investigated.
- Status: Last known status of the patient, category one combines the following possible states of the patient: alive and well, dead from other causes or dead with but not due to active disease, second category is alive with active disease, and third category indicates dead from disease.
- Prog_and_or_recur: Progression or recurrence of disease after surgery.
- TumorFree: Tumor free at the most recent follow-up meeting.

4.2. MRI Specifics

MRI imaging includes different sequences depicting tumor extent (using T1- and T2- weighted MRI) and microstructural tumor characteristics (e.g., in diffusion weighted imaging (DWI)). First introduced by Rofsky et al. [RLL*99], the Volumetric Interpolated Breath-hold Examination (VIBE) facilitates a 3D gradient-echo sequences that produces T1 weighted images. The advantage of this approach is improved resolution in the Z-axis which enables high-quality multiplanar reconstruction. In DWI, highly cellular tissue features a lower diffusion coefficient [KM04]. A quantitative assessment of the diffusion may be performed with the generation of apparent diffusion coefficient (ADC) values obtained at different b-values [KP06]. When using an intravenous contrast agent, the dynamic contrast-enhanced (DCE) MR perfusion is recorded. Typical measurements during this examination are the peak enhancement (PE) measuring the relative enhancement in contrast after the update of the contrast agent and the time to peak (TTP) [GPME*14]. The area under the peak enhancement curve (AUC) is also a typical measurement [GPME*14]. In total, we have seven MR imaging sequences available.

4.3. Specifics of the Application Domain

Gynecological cancer imaging research consists of multiple steps. Data of several MRI sequences must be analyzed, currently done separately and partly only for specific regions within the tumor. The overall goal of our medical collaborators is to examine and explore tumor biomarkers which potentially have a larger predictive value for clinical outcome than well-established ones. These biomarkers may help to further improve treatment of patients and increase personalization. Recent research of our collaborators already includes the analysis of tumor texture, which relies on prior segmentation of the tumor [HSG*14, FBYH*18, YHDL*18]. The gold standard in this segmentation approach is a manual segmentation performed by a trained radiologist. One aspect to consider in this regard is that this step comprises intra-operator variance. Due to the very time-consuming process of segmenting each volumetric slice this process is only performed using one of the seven MRI sequences. Therefore, our collaborators aim for a complete automation using machine learning algorithms. Another step already performed by our collaborators is the co-registration of all seven sequences present. For both steps our collaborators are looking for a validation possibility to assess data quality before analysis. The analysis

of the imaging data and clinical parameters involves several tasks: group selection, tumor texture feature analysis, data quality validation, and hypothesis formation. To the best of our knowledge there is no application available which combines these tasks in an easy and intuitive way without having to export and import data multiple times.

4.3.1. Machine Learning Segmentation

One of our co-authors applied a 3D convolutional neural network (UNet3D [CAL*16]), using Keras [GP17] and TensorFlow [ABC*16] as backend engine, to facilitate automatic segmentation of the tumor data in endometrial cancer patients. The network was trained using 139 expert segmentations based on preoperative pelvic imaging. The network can retrieve tumor volumes which are comparable to human expert level and a set of segmentation masks with human agreement not differing from inter-rater agreement. Although this algorithm is very promising for further analysis of the segmented tumor volumes and masks, proofreading by a radiologist is still necessary. Common tools used in clinical practice and research allow for such a validation but quickly finding cases where the segmentation might be wrong is desirable but not supported yet. After this processing step, the tumor segmentation is still only available on one sequence. To analyze the tumor on every given sequence, co-registration is necessary. This process was performed automatically but needs to be validated since a segmentation of non-tumor-regions would potentially introduce a critical error in data analysis. Examples of incorrect segmentations can be seen in the accompanying video and figures.

4.3.2. Automatic Co-registration

Our collaborators performed the co-registration automatically using FMRIB's Linear Image Registration Tool (FLIRT) [JS01, JBBS02] without optimization and only performing geometric alignment in scanner coordinates. However, this automatic registration method may not always find a perfect transformation for each modality and therefore must be validated. The employed co-registration algorithm features a relatively low failure rate but for a meaningful analysis spotting cases where it might have failed is crucial. Our medical collaborators request for an intuitive way to find and validate these cases. Having all sequences co-registered and the segmentation prepared, radiomics feature extraction is the next step. Examples of incorrect co-registrations are presented in the additional materials.

4.3.3. Radiomics Feature Extraction

Radiomics feature extraction takes volumetric imaging data and the volumetric tumor mask as input and generates a high-dimensional feature vector describing the tumor in each parametric imaging sequence. We merge the generated data afterwards with the clinical parameters. In recent work by our collaborators [YHDL*18] tumors textural features were analyzed using TexRad [Ltd20] software. The number of features in this approach is limited and the feature generation algorithms are not open source. Therefore, our medical collaborators expressed interest in a transparent and more controllable data handling method. Based on prior research, tumor texture features are interesting measurements, believed to be correlated with aggressiveness of tumors [YHDL*18, FBYH*18]. To further support this hypothesis, we calculate potential features for

homogeneity analysis, including normalized inverse difference moment, contrast, short run emphasis and long run emphasis. The information content contribution of each sequence is not known a priori and therefore an explorative analysis of single sequences and their combinations is of interest for our collaborators.

4.4. Task Abstraction

We performed a task abstraction using the task framework proposed by Brehmer and Munzner [BM13]. We assessed the current status of clinical research in gynecological cancer during multiple interviews with our collaborators. We also encouraged them to envision new workflows including results of convolutional neural network-performed segmentations and the possible parallel analysis of all sequences after co-registration. During the interviews, we identified two phases. Phase one deals with ensuring data quality and phase two with cohort analysis. Two tasks (T1 and T2) handle the need for a segmentation and co-registration validation. Task T3 reflects a common practice in medical research, namely group selection. Finally, tasks T4 and T5 provide analysis functionality. During discussions, our collaborators mentioned that they commonly use R or SPSS for statistical analysis and that they would like to continue doing so. Therefore, we exclude statistical analysis capabilities from our application design.

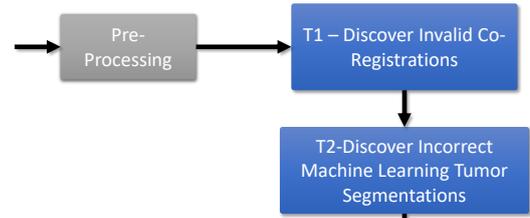
T1–Discover Invalid Co-Registrations Our medical collaborators classify co-registrations as invalid based on the misalignment of the given volume. Bladder filling and other physiological processes in the body cause a shift of the organs and therefore automatic methods may not result in a completely perfect alignment of these images. However, finding cases where automatic co-registration fails to find a sufficient transformation is crucial to support productive and time-efficient analysis. The analysis platform should allow the user to *discover* misaligned volumes and to analyze data in detail to *identify* potential causes of the misalignment.

T2–Discover Incorrect Machine Learning Tumor Segmentations Manual volumetric segmentation of endometrial tumors is a tedious and time-consuming task. This task could be automated using machine learning techniques, for example based on convolutional neural networks. Although the algorithm employed by our collaborators features a low failure rate and a comparable precision as the medical experts, the results still must be validated before further analysis. Spotting cases where the segmentation is potentially wrong is challenging and browsing through all patients is not time efficient. The user wants to *discover* potential faulty cases and *identify* the cause of the incorrect segmentation mask.

T3–Group Selection and Comparison Group identification and selection is a common and important task in clinical research. When analyzing a cohort of patients, it is of great interest to spot patients which share similar features, e.g., in imaging or histological analysis results. During the analysis of such cohort data, different groups can be *selected* and the medical researchers would like to *compare* them with each other.

T4–Homogeneity Analysis Heterogeneity is putatively linked to aggressive cancer phenotype supported by previous studies linking specific textural features to high-risk histological subtypes in endometrial cancer [YHDL*18, FBYH*18]. An exploratory platform allowing an assessment of textural features reflecting tumor heterogeneity derived from the different sequences/parametric

Phase 1: Data Quality check



Phase 2: Search and Query

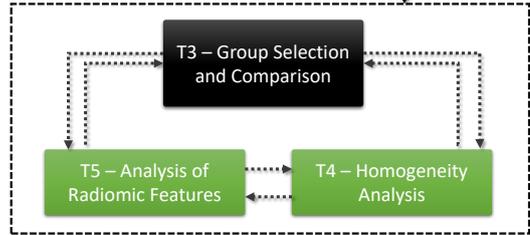


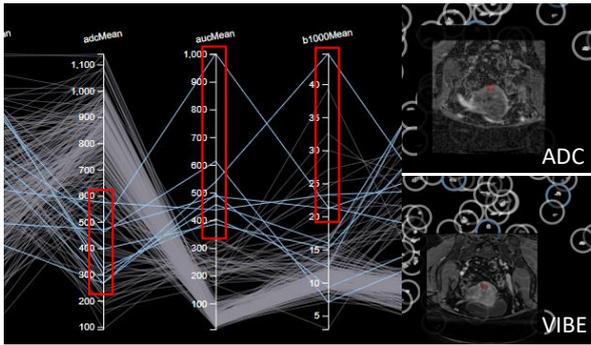
Figure 2: The RadEx workflow based on our task abstraction. The flow is denoted by arrows consists of two phases. Phase One: *Discover* invalid co-registrations and segmentation masks. Phase Two (*Search and Query*): When the data quality is ensured, users are able to search and query the whole cohort. The arrows depict that there is no pre-defined ordering of the tasks: each task can be executed in any order in the search and query section.

maps interactively would potentially be clinically useful. *Browsing* through further homogeneity measurements over all available imaging modalities and *comparison* of predictive value is interesting for our collaborators.

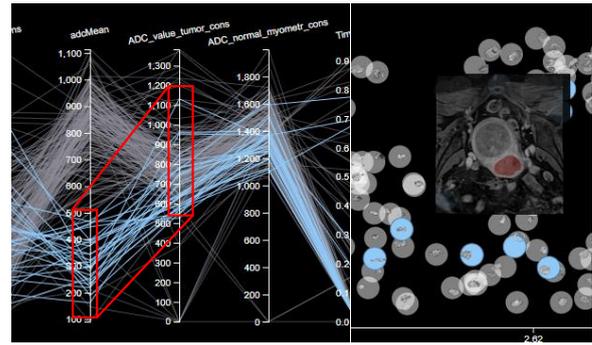
T5–Analysis of Radiomics Features Radiomics feature generation is a promising way to generate characteristics which might facilitate a predictive value of specific clinical parameters. Experts would like to have a look at individual parameters, but also at the analysis of multiple combined parameters. They want to *locate* specific combinations or single parameters to *identify* a predictive value to previously defined groups, based on clinical parameters.

5. RadEx Workflow and Interface

Figure 2 illustrates the workflow of RadEx when analyzing unprocessed cohort data. The first two interactions with the application ensure data quality for further analysis steps. Before analyzing the data within RadEx, a pre-processing step is required pre-calculating all slices per patient and tumor extent. The analysis workflow within the RadEx application starts with co-registration validation, followed by machine learning segmentation validation. These two steps ensure the data quality and deliver valuable feedback to our collaborators to further improve the segmentation and registration output. After this step, clinical researchers can analyze the cohort data. They can focus on different aspects, for example, browsing, exploring, or locating specific feature. Group selection is performed as a first step, users can explore different homogeneity measurements or radiomics feature combinations. If the users are interested in specific characteristics, they can browse for possible groupings. The number of patients currently included in our appli-



(a) Outliers in the co-registration view of the parallel coordinate plot indicate misaligned volumes after co-registrations. If acquisitions are not registered, segmentation output based on a single sequence will not be correct for other sequences. This leads to unexpected derived values.



(b) The ADC map value is comparable between patients and therefore routinely measured in gynecological cancer cases. A correlation analysis to the mean ADC value within the tumor segmentation could raise suspicion with regards to segmentation accuracy.

Figure 3: Typical cases where co-registration (a) or tumor segmentation (b) failed. Finding these cases is not an easy task and browsing through all patients would be very time-consuming. Therefore, the parallel coordinate plot in the co-registration view enables a quick search for potential erroneous cases.

cation is already a high number for this type of studies, therefore we chose our visualization techniques to cope with the given number of patients.

5.1. Central Scatterplot View

The central scatterplot view presented in Figure 1 reveals the whole cohort at a glance. This view plots a homogeneity measurement against a one-dimensional t-SNE dimensionality reduction of selected radiomics features allowing for an overview of the data. Each patient is marked with a gray circle. According to Cleveland et al. [CS87], three factors determine effective scatterplot design: (1) the marks are designed with preattentive features in mind, (2) the detection of individual objects is in focus, and (3) the distance between the objects presents a notion of similarity. We use these features to guide our scatterplot design. As shown in Figure 1, each mark representing a single patient contains a small glyph representing the shape and the size of the tumor. We generate this small image by finding the slice with the largest amount of tumor voxels and extract VIBE pixels within the segmentation mask. If the tumor consists of multiple parts, these are still visible within the circle. An example of these glyphs is shown in Figure 7. Hovering over a glyph reveals a tooltip showing an image slice with the tumor segmentation as a color overlay. The imaging modality as well as the slice can be selected by the user. This gives the clinical experts a direct relation to underlying imaging data and allows for a detailed co-registration and segmentation validation. The tooltip view presents details on demand by holding the Shift key while hovering over a mark. In this detailed version of the tooltip, an overview over all modalities is presented allowing the user to compare all seven modalities (T1). Both tooltip versions are shown in Figure 4. The tooltip is only shown on mouse hover over a glyph of interest representing a single patient. At this point the user is interested in exploring data of this single patient, occlusion of other patient glyphs is therefore less problematic. This method for tooltip display does not introduce a visual focus change for the user and is therefore the

appropriate placing for it. By scrolling through the modalities or by using the detailed tooltip version, a co-registration check and a segmentation validation can be performed (T1-2).

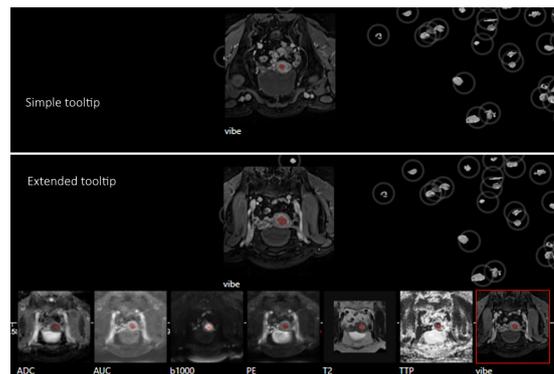


Figure 4: Top: The simple tooltip reveals an image slice with segmentation information as a color overlay. Bottom: The extended tooltip visualizing all available modalities with tumor segmentation overlay and a red border around the selected modality.

A large amount of data and large marks in the scatterplot lead to overplotting, which can be avoided by various methods. Marks or position could be changed locally to reduce overplotting on demand. We use a simple zoom and pan interaction because it is efficient, and our collaborators are already familiar with these interactions. The methods are also easy to understand and execute while keeping the position of the dots in the scatterplot space static. We also added an option to reset the zoom and pan to its original state on demand.

Selection of the axes is crucial in scatterplots. As inhomogeneity is an essential but rather new measurement that is believed to correlate with tumor aggressiveness, we display this on the x-axis (T4). Following the description of Cleveland et al. [CS87] we use

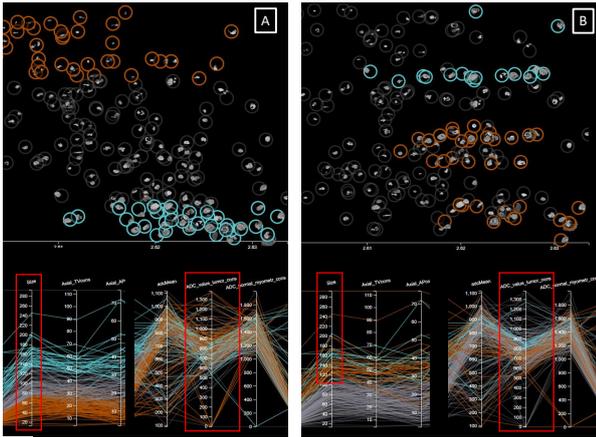


Figure 5: A: *t*-SNE calculation only taking size into account. B: Both size and ADC_value_tumor_cons are taken into account in the 1D dimensionality reduction. Selecting large tumors and different ADC value ranges for both groups indicate that the ADC value in large tumors could correlate with aggressiveness. A pattern is visible revealing in the scatterplot that patients are separated by ADC value.

the x-axis to bring patients with a similar homogeneity closer to each other to enable clustering. For the y-axis we aim to allow for clustering patients according to similarity in the higher dimensional feature space consisting of clinical parameters and radiomic tumor features (T5). To this end, we display a 1D dimensionality reduction result using *t*-SNE [vH08]. During our development process we also used a 1D principal component analysis (PCA) which delivered less convincing results. Therefore we chose *t*-SNE dimensionality reduction for our specific scenario, but this choice might not be the best option for other problem domains. Our scatterplot layout delivers an overview of the interplay between imaging data and clinical parameters for the purpose of radiomic tumor profiling. One example is presented in Figure 3(a). In Figure 5A only the size influences the *t*-SNE and in Figure 5B both size and the ADC value of the tumor are considered. As the y values of the dots change, different groupings are visible. In Figure 5A only size related clusters can be found while in Figure 5B the ADC value has an influence and new clusters are present. This interactive dimensionality reduction enables hypothesis generation relating imaging and clinical parameters.

5.2. Parameter Overview

In addition to the imaging data represented in the central scatterplot view, clinical parameters are also a focal point in cohort analysis. These parameters consist of numerical, ordinal, and categorical data. Visualizing multiple data types together in one visualization can cause problems, because not all data types are compatible with all visualization idioms. Therefore, we opted for splitting these into two different visual representation. For the numerical data, we employ a parallel coordinate plot and for the categorical and ordinal data we use unit charts. As our tool was collaboratively developed with domain experts, we received iterative feedback on their ability to understand and work with selected visualization techniques during development.

Feature Dimension View A parallel coordinate plot (PCP) is an effective tool to analyze correlations between different feature dimensions [ID90]. Every patient represents one line in the parallel coordinate plot and each axis shows one feature dimension. The decision which axes/dimensions to use in the PCP is very important [ID90]. They serve as visual anchor and allow for use of ticks and descriptions. The ordering is also important, because it is difficult to compare dimensions which are further apart in the plot. Therefore, we decided to put specific axes next to each other where the correlation serves a specific purpose, e.g., the size of the tumor measured by the experts in the VIBE modality and the amount of voxels derived from the tumor mask. If the correlation between these two measurements is suspicious for certain patients, there might be something wrong with the data (T2). The PCP dimensions can also be selected for validation purposes specifically. If the user is performing segmentation or registration validation, the mean values of all modalities within the tumor are visualized. This enables detection of outliers, which may be caused by a misaligned segmentation mask (T1, T2).

Clinical Parameter View Unit charts are one of the simplest visualization methods and have already been described by Neurath in the early 1930's [Neu36]. More recent work by Park et al. [PDFE18] states that this type of visualization can provide information that matches the user's mental model and allows for novel interactions. The unit chart representation is used for all categorical and ordinal parameters. Each dot in each column of the visualization represents exactly one patient. The color of each dot represents the value of the parameter for that specific patient. In addition, we use a tooltip to present imaging data when hovering with the mouse over the marks. Missing values are at the bottom of the chart and colored in dark gray. The other dots follow a quantitative grayscale colormap. The values in the unit chart are ordered according to expected outcome severity, meaning that values that have a negative influence on the outcome, e.g., life expectancy and suspected quality of life after treatment, are positioned on top of the chart (T3-5).

5.3. Settings and Interaction Techniques

Our application features a group selection feature where the user can select two different groups (T3). User can select which group is active and if selected patients should be combined using an 'AND' or 'OR' function. This allows for a detailed group selection. Using the 'AND' option, the user can select patients that, e.g., have multiple clinical parameters in common. In contrast, the 'OR' connection allows for selecting patients that, have a large tumor and myometrial invasion, but do not necessarily need to have both properties. The group selection interaction is supported across application views. Selection operations can be performed in the scatterplot by brushing with a rectangular selection box, in the PCP by selecting along an axis, or in the clinical parameter view by clicking one of the dots representing a specific value. Selections can be reset using a clear function. The user is further able to change the x- and y-axis properties. When changing the settings, the scatterplot updates with an animation, to improve context preservation. This allows users to locate parameter ranges to maximize selected target group separation. In the settings, the user is also able to change the modality presented in the tooltip and to swap the PCP to the registration and segmentation validation view.

6. Implementation

Our web-based application is composed of modules. The main part of the preprocessing, namely the data extraction and feature generation, is developed in Python. We use the PyRadiomics, a library developed by Gillies et al. [GKH16]. The library supports first-order statistical features such as voxel-intensity histogram-based features, e.g., the median, the standard deviation or the maximum and the minimum value. In addition, also second-order statistics are supported. These include, for example, features based on the gray level co-occurrence matrix or the gray level run length matrix [GKH16]. We use the Visualization Toolkit from Kitware [SKL06], to create the tumor icons and the tooltip information for further analysis. Numpy [Oli06] is used for working with the high dimensional radiomic tumor feature arrays and data handling within Python. We handle csv data handling using the Python library OpenCV [Bra00].

The web-based part of our application is implemented in Javascript. The scatterplot and the parallel coordinate plot are both implemented using D3 [BOH11]. Our implementation of the unit chart visualization is based on the approach by Park et al. [PDFE18]. For our dimensionality reduction, we use the t-SNE implementation TSNEJS provided by Karpathy [Kar16], which is based on the original work from van der Maaten [vH08].

7. Case Studies

The RadEx application is visible in Figure 1 and consists of multiple components as described in the Section 5. To demonstrate the utility of our application, we showcase its functionality in three case studies, developed in close collaboration with our collaborators. We identified three major areas of application for our tool, namely the exploration of tumor characteristics, the co-registration validation functionality, and the machine learning segmentation check. Data of 330 patients is provided by two of our co-authors. Before including them in the application 12 patients have been excluded because the co-registration did not work, due to imaging quality problems. 97 patients were excluded because the machine learning based segmentation mask was completely misplaced or much too small. In the end 221 patients are included for further analysis. For 92 patients a manual created segmentation mask is available and for 129 patients a machine learning created segmentation mask is used. For every patient seven MRI sequences created by either a 1.5 Tesla or 3 Tesla MRI scanner, ten clinical parameters, and five measures from radiologists are available.

7.1. Explorative Radiomic Tumor Profiling

Radiomic tumor profiling involves the calculation of high-dimensional feature vectors that need to be analyzed to discover tumor characteristics that are marker for possible outcome or to evaluate existing ones. Typical use cases in tumor profiling include homogeneity analysis of the imaging data and analyzing the association between various radiomics features with respect to different clinical parameters. In our application clinical experts can select two groups, one with low aggressiveness and one with higher aggressiveness. Aggressiveness can be measured, e.g., by presence of metastases or if the patient has already died from disease or had recurrence. Also, the time between the surgical removal of the tumor and a possible recurrence is an indicator for aggressiveness.

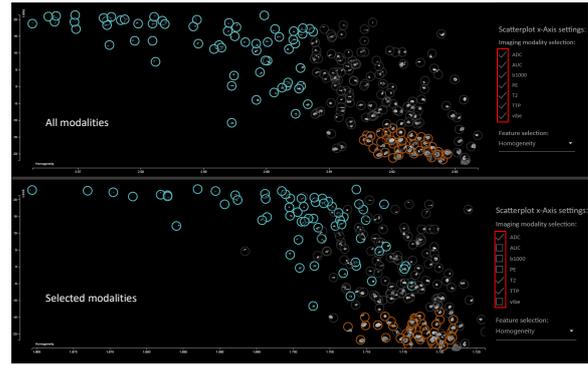


Figure 6: Textural tumor features include several homogeneity measurements believed to correlate with tumor aggressiveness. We support a selection of homogeneity measurements and combine up to seven MRI sequences resulting in one measurement presented on the x-axis. Changes in parameter selection result in animated transitions.

During the exploration of different settings for the x-axis, the group selection stays the same and gives a clear picture if the group separation improves or not, Figure 6 reveals two possible group separations using different x-axis values. Allowing the user to try out different homogeneity measurements enables interactive exploration and assessment of the differences between each of the measures for this patient cohort. While exploring different measures, e.g., the homogeneity measurement based on the normalized inverse difference moment as described by Gillies et al. [GKH16], we found a separation of a group of patients that features a low aggressiveness while having a low homogeneity.

Preliminary exploration revealed that the derived textural homogeneity can separate patients with high-risk disease from these with low-risk disease, demonstrating the usefulness of this tool for identification of imaging markers to be further explored. After selecting different feature combinations for dimensionality reduction, we are able to find different patient clusters that have different distinctive clinical parameters, such as for example the cluster shown in Figure 6 in orange. This cluster features patients presenting with large tumors, which is known to be associated with increased risk of metastases and death from disease. This association is already well known, and our application can show this link. The status for these patients tells that they are either alive with active disease or dead from disease. The presence of metastases which evolve from the primary tumor is also an indicator for an aggressive tumor phenotype.

7.2. Registration Validation

Co-registration of multi-parametric images is an essential part of the data processing step needed for our application. The result of a successful co-registration is a set of perfectly aligned volumes. When dealing with high number of patients where most of the co-registration works well and which fails only in a small portion of cases, it is important to support visualization of likely outliers or error cases. To support such a filtering, the dimension selection visualized in the feature dimension view in the lower center of our application can be used to support the registration check. When doing so, the chart visualizes the median values of the tumor in each of

the seven imaging parameters. This view allows for a quick check of outliers in the graph which might highlight segmentation masks that do not mask the tumor adequately due to misaligned volumes. The segmentation is then also misaligned because the segmentation mask is only available for one modality in our case. One example for such an operation is shown in Figure 3(a), where outliers in the ADC value range revealed potential errors in the co-registration.

7.3. Segmentation Validation

Tumor segmentation masks created by machine learning algorithms need to be validated by experts before using them in the data analysis step. Before validating all segmentation masks, experts could first find outliers where the algorithm did a bad job and those could be used to further improve the algorithm. There are multiple ways to find these outliers. One way is to validate if the number of voxels in the tumor mask align with the size measured by the clinical experts. Both parameters are present in the feature dimension view in neighboring positions. Selecting a small number of voxels and a larger size measured by the experts or vice versa points to potential error cases. To validate if the segmentation is deficient, the user can hover over highlighted glyphs in the scatterplot and slice through the imaging volume. In the view presented in Figure 4, we use a semi-transparent red overlay of the tumor segmentation mask on top of the various imaging sequences. Another possibility to find potential misaligned tumor segmentation masks is to select measured ADC mean values from our radiomics approach and compare them to the representative ADC value measured by the clinical experts. The experts indicate one representative region within the

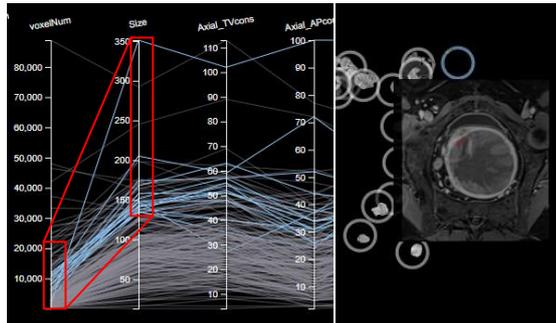


Figure 7: One possible measure to find incorrect segmentation masks is the size, which in our case is measured in our case by radiologists and our application pre-processing phase based on the VIBE image and on the segmentation mask. These measures have a natural correlation and unexpected relations might indicate incorrect segmentation.

tumor and one region in the healthy endometrium to compare these throughout patients. If there is a major discrepancy, these cases should be investigated more closely. One such case is shown in Figures 3(a) and 3(b). A third method to inspect segmentation quality is to inspect the segmented tumor symbols in the scatterplot. When selecting very large tumors and seeing very small symbols on the scatterplot dots, it is an indication that the segmentation may be incorrect. Similarly, any discrepancy between selected features and visible segmentation symbols would lead to detection of questionable segmentation quality. Figure 7 reveals such a case. The segmentation validation feature of our tool is of major interest for

the machine learning experts working on implementing automatic segmentation algorithms since it allows for quick and intuitive validation that spurs on further development of automated methods.

8. Evaluation

To further evaluate the utility of our interactive exploration and analysis platform, we invited six gynecological cancer imaging research experts to validate our tool. E1 is a professor in radiology and expert in gynecologic and abdominal radiology. She has over 14 years experience in this field and is one of the co-authors on this paper. E2 is a radiologist since 2006 and holds a medical doctors degree. She is currently a PhD student in gynecological imaging, has over 10 years of experience in MRI reading, and is a co-author of the paper. E3 is a medical physicist in radiology and is a PhD student in medical physics with over 13 years of experience. E4 holds a masters degree in cell biology since 2010 and a PhD in neuro-oncology since 2015. E5 has 4 years of experience in pelvic imaging and holds a medical doctors degree. E6 has 5 years experience in MRI reading of gynecological cancer, holds a medical doctors degree, and is currently a PhD student.

In the beginning of the evaluation we demonstrated the application to the experts worked through the different use cases. Our application works with data provided by E1 and another co-author which was not part of the evaluation. Afterwards, we invited them to try out the tool themselves. During the evaluation we asked the experts to discuss their experience and to talk about benefits and disadvantages of the system compared to their current workflow. After this phase, which took roughly 40 minutes we asked the experts to fill out a questionnaire with 34 questions discussing different aspects of our application. The questions are structured in the following groups: general (G1-7), tumor visualization (V1-5), group selection (S1-7), homogeneity (H1-5), dimensionality reduction (D1-5), and segmentation and co-registration validation (C1-5). In addition to our evaluation form, the experts filled out the system usability scale (SUS) provided by Brook et al. [Bro04]. All statements are evaluated based on a 5-point Likert scale. We also included negatively formulated questions.

8.1. Evaluation Results

The result of the evaluation is shown in Table 1. Questions marked with a star have originally been negatively formulated and here we present them in their positive form. The results for these questions are also inverted. In general, the application got positive feedback overall. All experts would like to contribute to the future development of the application and 5 out of 6 experts would like to use the application in the future. The tumor icons and tooltips received overall a good feedback. One expert (E2) mentioned that the tooltip pictures could be enlarged. E3 mentioned that the size of the tumor icons made it difficult to perceive shape, however, other participants agree that shape is also visible in our design. E5 also mentioned that the size of the tumor icons makes it difficult to compare based on icons alone.

All experts are in favor of the group selection. All questions, except one have an average value of at least 4.5. Only selecting specific properties in the Unit Chart or clinical parameter view has a value of 4.33. Regarding this point, we received the feedback that the dots used in the chart are challenging to click on. The homogeneity view also received strongly positive feedback overall. E3

Table 1: Response of the experts on a 5-point Likert scale. The meaning of the values on the scale are: 1: strongly disagree, 2: disagree, 3: neither agree nor disagree, 4: agree and 5: strongly agree. Statements marked with a star were rephrased to present the positive form in this table and the scores have been inverted. On the right end of the table the average value over all experts is presented and in the last row the result of the system usability scale questionnaire is presented. ¹ paper co-authors.

	Statements:	E1 ¹	E2 ¹	E3	E4	E5	E6	Avg.
G1	The linked interactions between the scatterplot and the parallel coordinate plot are well established and intuitive	4	5	5	5	5	4	4,67
G2	The linked interactions between the scatterplot and the unit chart are well established and intuitive*	5	5	5	5	5	4	4,83
G3	The selection interactions between the unit chart and the parallel coordinate plot are well established and intuitive	5	5	5	5	5	4	4,83
G4	I see myself using RadEx in the future	3	5	5	4	5	4	4,33
G5	I would like to contribute in the future development of the application*	5	5	5	5	5	5	5,00
G6	I would like to use RadEx for exploring clinical cohort data	4	5	5	5	5	4	4,67
G7	The export functionality helps me to further analyze the group selections in my statistics tool of choice*	4	5	5	5	4	4	4,50
V1	The small tumor icon enables a quick comparison between the tumors of different patients	5	4	4	5	2	4	4,00
V2	The tumor icons give me more information than only the size of it*	5	4	2	5	3	4	3,83
V3	The tooltip allows me to analyze the imaging data and the tumor segmentation*	5	3	5	4	5	4	4,33
V4	The extended tooltip is helpful to validate the tumor segmentation	5	3	5	4	5	4	4,33
V5	The extended tooltip is helpful to validate the co-registration*	5	3	5	5	4	4	4,33
S1	The group selection in the scatterplot view is easy to understand and to carry out	5	5	5	5	4	5	4,83
S2	I can select specific patients and add them to an existing group in the scatterplot*	5	5	5	4	5	5	4,83
S3	Selecting specific properties in the Unit chart view is easy to understand and to carry out	4	5	5	4	4	4	4,33
S4	I can select patients having specific states in different clinical parameters*	5	5	5	5	5	4	4,83
S5	Specifying a patient group including multiple clinical parameter manifestations is easy	5	4	5	5	5	4	4,67
S6	Selecting a group in the parallel coordinate plot is easy to understand and carry out*	4	4	5	5	5	4	4,50
S7	The applications makes it easy to select two different groups	5	4	5	5	4	4	4,50
H1	The homogeneity imaging modality selection in combination with the group selection helps me to identify important modalities*	5	4	2	4	4	4	3,83
H2	Trying different homogeneity measurements is easy and fast*	4	5	5	5	5	4	4,67
H3	The animation of the data when changing settings helps me to track the changes	5	5	4	5	5	4	4,67
H4	Having the important measure homogeneity on the x-Axis of the scatterplot makes interpretation of the visualization easy	4	4	4	5	4	4	4,17
H5	I can imagine using this application to formulate hypothesis for future studies about homogeneity*	5	5	4	5	4	5	4,67
D1	The y-Axis in the scatterplot shows me interesting clusters of patients	5	4	5	5	5	4	4,67
D2	The dimensionality reduction allows me to analyze multiple clinical parameters*	5	5	5	5	5	4	4,83
D3	The selection of dimensions taking into account for the y-Axis allows me to explore my clinical data*	5	5	5	5	5	4	4,83
D4	I can imagine using this application to formulate hypothesis for future studies	4	5	5	5	4	4	4,50
D5	Exploring patients that are clustered by the t-SNE is interesting and potentially valuable for further investigation*	5	5	5	4	5	4	4,67
C1	I can select machine learning performed segmentations and validate their correctness	5	1	5	5	4	4	4,00
C2	Selecting potentially wrong segmentations is possible	5	3	5	4	4	4	4,17
C3	Exporting wrongly segmented patients is possible*	5	3	5	5	5	4	4,50
C4	The co-registration view enables me to spot potential wrong co-registrations*	4	5	5	5	4	4	4,50
C5	The tooltip view helps me to validate segmentations and co-registration results	5	3	5	5	4	4	4,33
SUS	System usability scale result	75,00	90,00	92,50	95,00	87,50	77,50	86,25

mentioned that it is difficult to prove if a modality is important or not and mentioned that the question is formulated too narrowly. However, she is still in favor of the functionality. Our dimensional reduction got the most positive feedback with all average values over 4,5. All experts could imagine using the 1D dimensionality reduction to analyze the cohort data to spot potential groupings of patients. The co-registration and segmentation validation also received positive feedback overall. One expert (E2) mentioned that the tooltip images could be larger to make the validation easier. Another possibility for a more detailed validation could be to use a second screen to show the imaging data of specific patients in a view more like what radiologists are used to (E2, E4-6).

System Usability Scale Scores Our SUS scores are presented at the end of the evaluation result Table 1. The results range from 77,5 to 95. In average our application reached a SUS score of 86,2. Bangor et al. [BKM09], introduced different ways of interpreting SUS scores including the acceptability range, a grade scale (like in education), and an adjective rating scale. Our acceptance rate is: Acceptable (best score), grade scale: A (best score), and an adjective rating of Excellent (best score).

8.2. Evaluation Conclusion

We conclude from our results that our application is valuable for experts in gynecological cancer imaging research. All statement groups received positive feedback and the experts think the features are useful. E1 already thinks about using our application in a research setup to further evaluate machine learning-based segmentation masks and to train radiologists to perform segmentations and compare them to segmentation masks created by experts. E4 can also imagine using the application to validate results with new

imaging series and E6 would like to use the application for his cohort data. Overall, we can say that the application has substantial potential in gynecological cancer imaging research.

9. Discussion

Experimental group selections performed in our application revealed that there is potential in further analysis of different homogeneity measures to separate patients with high-risk disease from those with low risk. Our application is also able to show well known coherence patterns, e.g., between the size of the tumor and clinical phenotype. This shows us that there is a potential to influence future analysis steps in gynecological cancer research and that our application may have an impact in the targeted domain.

For the group comparison we do not offer a feature that determines statistical significance values to prevent p-value significance fishing. Any hypothesis formulated using our application should be validated using an independent study cohort. After performing segmentation validation with our tool, we were able to spot tumor segmentation masks that did not meet the acceptance criteria, e.g., due to the presence of multiple tumors within the same region. The co-registration validation also highlights cases which would not have been suspicious at first sight. The involved machine learning expert therefore also sees potential in working with our application to further refine his machine learning algorithms to deliver even better results. The 1D dimensionality reduction is in the current version only supported by t-SNE but could also be performed using other dimensionality reduction methods such as PCA.

10. Conclusion and Future Work

We present RadEx, an interactive analysis platform and workflow for medical researchers which supports integrated exploration of ra-

diomics and clinical features. By using multiple linked views, interactive group selection methods and custom feature selections, we empower researchers to explore new and validate existing tumor biomarkers. Close collaboration with gynecological cancer imaging and machine learning research experts resulted in several case studies for our application. The benefits brought by our work range from validation of machine learning results to validating and exploring new tumor characteristics. Being able to selectively include and exclude different sequences from the radiomic tumor exploration makes the analysis transparent and intuitive for the medical researchers. Our experimental findings of different patient groups allowed visualization of e.g. well-established association between large tumor size and high-risk disease but also showed potentially interesting new associations. The evaluation of our application revealed a positive response from the target audience reflected both by the qualitative evaluation and SUS score.

Potential future developments include adding an intuitive way to perform a landmark based registration. This could help to easily correct the registration for patients where the automatic co-registration fails. In addition, we plan to widen the scope of the application to different tumor types including cervical cancer. The goal of our efforts is that every single patient should benefit from the findings from cohort analysis to get one step further in the direction of personalized medicine.

Acknowledgments

This research was funded by the Trond Mohn Foundation (Grant '811255', '813558', and 'BFS2018TMT06'), the Western Norway Regional Health Authority (Grant #912263), the Norwegian Cancer Society (Grant #190202), the Norwegian Research Council (Grant #273280), and supported by the Norwegian Research Centre (NORCE).

References

- [ABC*16] ABADI M., BARHAM P., CHEN J., CHEN Z., DAVIS A., DEAN J., DEVIN M., GHAMAWAT S., IRVING G., ISARD M., KUDLUR M., LEVENBERG J., MONGA R., MOORE S., MURRAY D. G., STEINER B., TUCKER P., VASUDEVAN V., WARDEN P., WICKE M., YU Y., ZHENG X.: Tensorflow: A system for large-scale machine learning. In *Proceedings of the USENIX Conference on Operating Systems Design and Implementation* (2016), OSDI'16, p. 265–283. 4
- [AMN*05] AMANT F., MOERMAN P., NEVEN P., TIMMERMAN D., LIMBERGEN E., VERGOTE I.: Endometrial cancer. *Lancet* 366 (2005), 491–505. doi:10.1016/S0140-6736(05)67063-8. 2
- [AOH*14] ANGELELLI P., OELTZE S., HAASZ J., TURKAY C., HODNELAND E., LUNDEVOLD A., LUNDEVOLD A. J., PREIM B., HAUSER H.: Interactive visual analysis of heterogeneous cohort-study data. *IEEE Computer Graphics and Applications* 34, 5 (2014), 70–82. doi:10.1109/MCG.2014.40. 3
- [BBCF*19] BHATLA N., BEREK J. S., CUELLO FREDERES M., DENNY L. A., GRENMAN S., KARUNARATNE K., KEHOE S. T., KONISHI I., OLAWAIYE A. B., PRAT J., SANKARANARAYANAN R.: Revised FIGO staging for carcinoma of the cervix uteri. *International Journal of Gynecology & Obstetrics* 145, 1 (2019), 129–135. doi:10.1002/ijgo.12749. 3
- [BFM*16] BERG A., FASMER K. E., MAULAND K. K., YTRE-HAUGE S., HOIVIK E. A., HUSBY J. A., TANGEN I. L., TROVIK J., HALLE M. K., WOIE K., BJØRGE L., BJØRNERUD A., SALVESEN H. B., WERNER H. M. J., KRAKSTAD C., HALDORSEN I. S.: Tissue and imaging biomarkers for hypoxia predict poor outcome in endometrial cancer. *Oncotarget* 7, 43 (2016), 69844–69856. doi:10.18632/oncotarget.12004. 2
- [BFS*18] BRAY F., FERLAY J., SOERJOMATARAM I., SIEGEL R. L., TORRE L. A., JEMAL A.: Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* 68, 6 (2018), 394–424. doi:10.3322/caac.21492. 1
- [BGE*18] BAKR S., GEVAERT O., ECHEGARAY S., AYERS K., ZHOU M., SHAFIQ M., ZHENG H., BENSON J. A., ZHANG W., LEUNG A. N. C., KADOCH M., HOANG C. D., SHRAGER J., QUON A., RUBIN D. L., PLEVITIS S. K., NAPEL S.: A radiogenomic dataset of non-small cell lung cancer. *Scientific Data* 5, 1 (2018), 180202. doi:10.1038/sdata.2018.202. 1
- [BKM09] BANGOR A., KORTUM P., MILLER J.: Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 4, 3 (2009), 114–123. 10
- [BM13] BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2376–2385. doi:10.1109/TVCG.2013.124. 5
- [BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309. doi:10.1109/TVCG.2011.185. 8
- [Bra00] BRADSKI G.: The OpenCV Library. *Dr. Dobbs's Journal of Software Tools* (2000). 8
- [Bro04] BROOKE J.: SUS-a quick and dirty usability scale. usability evaluation in industry. In *Usability Evaluation In Industry*. CRC Press, 2004, ch. 10, pp. 266–290. 2, 9
- [BSM*15] BERNARD J., SESSLER D., MAY T., SCHLOMM T., PEHRKE D., KOHLHAMMER J.: A visual-interactive system for prostate cancer cohort analysis. *IEEE Computer Graphics and Applications* 35, 3 (2015), 44–55. doi:10.1109/MCG.2015.49. 3
- [ÇAL*16] ÇIÇEK Ö., ABDULKADIR A., LIENKAMP S. S., BROX T., RONNEBERGER O.: 3d u-net: Learning dense volumetric segmentation from sparse annotation. *CoRR abs/1606.06650* (2016). URL: <http://arxiv.org/abs/1606.06650>. 4
- [CGM*17] CENEDA D., GSCHWANDTNER T., MAY T., MIKSCH S., SCHULZ H., STREIT M., TOMINSKI C.: Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 111–120. doi:10.1109/TVCG.2016.2598468. 3
- [CGM19] CENEDA D., GSCHWANDTNER T., MIKSCH S.: A review of guidance approaches in visual data analysis: A multifocal perspective. *Computer Graphics Forum* 38, 3 (2019), 861–879. doi:10.1111/cgf.13730. 3
- [CS87] CLEVELAND W. S., SCHMIEG G. M.: The elements of graphing data. *American Journal of Physics* 55, 8 (1987), 767–767. doi:10.1119/1.15021. 6
- [EAZ*19] ECKELT K., ADELBERGER P., ZICHNER T., WERNITZNIG A., STREIT M.: Tourdino: A support view for confirming patterns in tabular data. In *Proceedings of EuroVA* (2019), pp. 7–11. 3
- [FBYH*18] FASMER K. E., BJØRNERUD A., YTRE-HAUGE S., GRÜNER R., TANGEN I. L., WERNER H. M., BJØRGE L., ØYVIND O SALVESEN, TROVIK J., KRAKSTAD C., HALDORSEN I. S.: Preoperative quantitative dynamic contrast-enhanced mri and diffusion-weighted imaging predict aggressive disease in endometrial cancer. *Acta Radiologica* 59, 8 (2018), 1010–1017. doi:10.1177/0284185117740932. 2, 4, 5
- [GKH16] GILLIES R. J., KINAHAN P. E., HRICAK H.: Radiomics: Images are more than pictures, they are data. *Radiology* 278, 2 (2016), 563–577. doi:10.1148/radiol.2015151169. 2, 8
- [GP17] GULLI A., PAL S.: *Deep learning with Keras*. Packt Publishing Ltd, 2017. 4
- [GPME*14] GORDON Y., PARTOVI S., MÜLLER-ESCHNER M., AMARTEFIO E., BÄUERLE T., WEBER M.-A., KAUCZOR H.-U., RENGIER F.: Dynamic contrast-enhanced magnetic resonance imaging: fundamentals and application to the evaluation of the peripheral perfusion. *Cardiovascular Diagnosis and Therapy* 4, 2 (2014). doi:10.1155/2014.03.01. 4

- [GS10] GREEN S. B., SALKIND N. J.: *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*. Prentice Hall Press, 2010. 2
- [HE98] HASTREITER P., ERTL T.: Integrated registration and visualization of medical image data. *Proceedings of Computer Graphics International* (1998), 78–85. doi:10.1109/CGI.1998.694253. 3
- [HKT*18] HAEHN D., KAYNIG V., TOMPKIN J., LICHTMAN J. W., PFISTER H.: Guided Proofreading of Automatic Segmentations for Connectomics. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2018), 9319–9328. arXiv:1704.00848, doi:10.1109/CVPR.2018.00971. 3
- [HS16] HALDORSEN I. S., SALVESEN H. B.: What is the best preoperative imaging for endometrial cancer? *Current Oncology Reports* 18, 4 (2016), 25. doi:10.1007/s11912-016-0506-0. 2
- [HSG*14] HALDORSEN I. S., STEFANSSON I., GRÜNER R., HUSBY J. A., MAGNUSSEN I. J., WERNER H. M. J., SALVESEN Ø. O., BJØRGE L., TROVIK J., TAXT T., AKSLEN L. A., SALVESEN H. B.: Increased microvascular proliferation is negatively correlated to tumour blood flow and is associated with unfavourable outcome in endometrial carcinomas. *British Journal of Cancer* 110, 1 (2014), 107–114. doi:10.1038/bjoc.2013.694. 2, 4
- [ID90] INSELBERG A., DIMSDALE B.: Parallel coordinates: A tool for visualizing multi-dimensional geometry. In *Proceedings of IEEE Visualization* (1990), p. 361–378. 7
- [JBBS02] JENKINSON M., BANNISTER P., BRADY M., SMITH S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 2 (2002), 825–841. doi:10.1006/nimg.2002.1132. 3, 4
- [JBF*19] JÖNSSON D., BERGSTRÖM A., FORSELL C., SIMON R., ENGSTRÖM M., YNNERMAN A., HOTZ I.: A visual environment for hypothesis formation and reasoning in studies with fMRI and multivariate clinical data. In *Proceedings of Eurographics VCBM* (2019). doi:10.2312/vcbm.20191232. 3
- [JS01] JENKINSON M., SMITH S.: A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 5, 2 (2001), 143–156. doi:10.1016/S1361-8415(01)00036-6. 4
- [Kar16] KARPATY A.: tSNEJS. <https://github.com/karpaty/tsnejs>, 2016. 8
- [KM04] KINGSLEY P. B., MONAHAN W. G.: Selection of the optimum b factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke. *Magnetic Resonance in Medicine* 51, 5 (2004), 996–1001. doi:10.1002/mrm.20059. 4
- [KMAB15] KARIMOV A., MISTELBAUER G., AUZINGER T., BRUCKNER S.: Guided Volume Editing based on Histogram Dissimilarity. *Computer Graphics Forum* 34, 3 (2015), 91–100. doi:10.1111/cgfm.12621. 3
- [KOJL*14] KLEMM P., OELTZE-JAFRA S., LAWONN K., HEGENSCHIED K., VÖLZKE H., PREIM B.: Interactive visual analysis of image-centric cohort study data. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 1673–1682. doi:10.1109/TVCG.2014.2346591. 3
- [KP06] KOH D.-M., PADHANI A. R.: Diffusion-weighted mri: a new functional clinical technique for tumour imaging. *The British Journal of Radiology* 79, 944 (2006), 633–635. doi:10.1259/bjr/29739265. 4
- [KPS16] KRAUSE J., PERER A., STAVROPOULOS H.: Supporting iterative cohort construction with visual temporal queries. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 91–100. doi:10.1109/TVCG.2015.2467622. 3
- [Ltd20] LTD. F. M.: *TexRad (Texture + Radiology)*. Feedback Medical Ltd., 2020. URL: <https://fbkmed.com/textrad-landing-2/>. 4
- [Neu36] NEURATH O.: International Picture Language: The First Rules of Isotype. *Psyche Miniature Series* (1936). doi:(HC)1231.50; cb-pdf. 7
- [Oli06] OLIPHANT T. E.: *A guide to NumPy*, vol. 1. Trelgol Publishing USA, 2006. 8
- [PDFE18] PARK D., DRUCKER S. M., FERNANDEZ R., ELMQVIST N.: Atom: A grammar for unit visualizations. *IEEE Transactions on Visualization and Computer Graphics* 24, 12 (2018), 3032–3043. 7, 8
- [PL20] PREIM B., LAWONN K.: A survey of visual analytics for public health. *Computer Graphics Forum* 39, 1 (2020), 543–580. doi:10.1111/cgfm.13891. 3
- [RBR*18] RIZZO S., BOTTA F., RAIMONDI S., ORIGGI D., FANCIULLO C., MORGANTI A., BELLOMI M.: Radiomics: the facts and the challenges of image analysis. *European Radiology Experimental* 2 (2018), 36. doi:10.1186/s41747-018-0068-z. 2
- [RCMA*18] RAIDOU R. G., CASARES-MAGAZ O., AMIRKhanov A., MOISENKO V., MUREN L. P., EINCK J. P., VILANOVA A., GRÖLLER M. E.: Bladder runner: Visual analytics for the exploration of rt-induced bladder toxicity in a cohort study. *Computer Graphics Forum* 37, 3 (2018), 205–216. doi:10.1111/cgfm.13413. 3
- [RLL*99] ROFSKY N. M., LEE V. S., LAUB G., POLLACK M. A., KRINSKY G. A., THOMASSON D., AMBROSINO M. M., WEINREB J. C.: Abdominal mr imaging with a volumetric interpolated breath-hold examination. *Radiology* 212, 3 (1999), 876–884. doi:10.1148/radiology.212.3.r99se34876. 4
- [RSt20] RSTUDIO TEAM: *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020. URL: <http://www.rstudio.com/>. 2
- [RVD*15] RAIDOU R. G., VAN DER HEIDE U. A., DINH C. V., GHOBADI G., KALLEHAUGE J. F., BREEUWER M., VILANOVA A.: Visual Analytics for the Exploration of Tumor Tissue Characterization. *Computer Graphics Forum* 34, 3 (2015), 11–20. doi:10.1111/cgfm.12613. 3
- [SKL06] SCHROEDER W., KEN M., LORENSEN B.: *Visualization Toolkit: An Object-Oriented Approach to 3D Graphics*, 4 ed. Kitware, Inc., 2006. 8
- [SMB*10] STEENWIJK M. D., MILLES J., BUCHEM M. A., REIBER J. H., BOTHA C. P.: Integrated visual analysis for heterogeneous datasets in cohort studies. *Proceedings of the IEEE VisWeek Workshop on Visual Analytics in Health Care* (2010). URL: <http://research.ihost.com/vahc2010/>. 3
- [STCS*03] SCHNABEL J. A., TANNER C., CASTELLANO-SMITH A. D., DEGENHARD A., LEACH M. O., HOSE D. R., HILL D. L., HAWKES D. J.: Validation of nonrigid image registration using finite-element methods: Application to breast MR images. *IEEE Transactions on Medical Imaging* 22, 2 (2003), 238–247. doi:10.1109/TMI.2002.808367. 3
- [TLS*14] TURKAY C., LEX A., STREIT M., PFISTER H., HAUSER H.: Characterizing cancer subtypes using dual analysis in caleydo stratomex. *IEEE Computer Graphics and Applications* 34, 2 (2014), 38–47. doi:10.1109/MCG.2014.1. 3
- [VBK*13] VON LANDESBERGER T., BREMM S., KIRSCHNER M., WESARG S., KUIJPER A.: Visual Analytics for model-based medical image segmentation: Opportunities and challenges. *Expert Systems with Applications* 40, 12 (2013), 4934–4943. doi:10.1016/j.eswa.2013.03.006. 3
- [vH08] VAN DER MAATEN L., HINTON G.: Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research* 9, nov (2008), 2579–2605. Pagination: 27. 3, 7, 8
- [YHDL*18] YTRE-HAUGE S., DYBVIK J. A., LUNDERVOLD A., ØYVIND O. SALVESEN, KRAKSTAD C., FASMER K. E., WERNER H. M., GANESHAN B., HØIVIK E., BJØRGE L., TROVIK J., HALDORSEN I. S.: Preoperative tumor texture analysis on mri predicts high-risk disease and reduced survival in endometrial cancer. *Journal of Magnetic Resonance Imaging* 48, 6 (2018), 1637–1647. doi:10.1002/jmri.26184. 2, 4, 5