ICEVis: *Interactive Clustering Exploration* for tumor sub-region analysis in multiparametric cancer imaging

Eric Mörth
Dept. of Informatics / MMIV
UiB / HUS, Norway

Tanja Eichner
Faculty of Informatics / MMIV
TU Vienna / UiB, Austria / Norway

Ingfrid Haldorsen
MMIV / Dept. of Clinical Medicine
HUS / UiB, Norway

Stefan Bruckner
Dept. of Informatics / MMIV
UiB / HUS, Norway

Noeska N. Smit
Dept. of Informatics / MMIV
UiB / HUS, Norway

ABSTRACT

Tumor tissue characteristics derived from imaging data are gaining importance in clinical research. Tumor sub-regions may play a critical role in defining tumor types and may hold essential information about tumor aggressiveness. Depending on the tumor's location within the body, such sub-regions can be easily identified and determined by physiology, but these sub-regions are not readily visible to others. Regions within a tumor are currently explored by comparing the image sequences and analyzing the tissue heterogeneity present. To improve the exploration of such tumor sub-regions, we propose a visual analytics tool called ICEVis. ICEVis supports the identification of tumor sub-regions and corresponding features combined with cluster visualizations highlighting cluster validity. It is often difficult to estimate the optimal number of clusters; we provide rich facilities to support this task, incorporating various statistical measures and interactive exploration of the results. We evaluated our tool with three clinical researchers to show the potential of our approach.

CCS CONCEPTS

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

1 INTRODUCTION

Cancer is globally the second leading cause of death according to a report published by the World Health Organization in 2018 [2]. Different tumor types exist, ranging from low-grade tumors with a favorable prognosis to high-grade tumors associated with a high risk of relapse and cancer spread. Surgery is the most common

treatment type. Early diagnosis and treatment planning are crucial for optimizing the therapy of endometrial cancer [1]. A challenge that can arise when dealing with endometrial tumors, as well as other types of cancer, is tumor heterogeneity [1]. The number and composition of tumor sub-regions may vary between different patients. Still, their identification may be an essential factor for diagnosis, tissue sample (biopsy) selection, and/or therapy planning. At present, tissue heterogeneity is examined by the acquisition of histopathological data from tissue biopsies and, in recent research, by analyzing the tumor imaging data as a whole using radiomic tumor profiling [6, 12, 22]. Feature detection from medical imaging data alone could potentially replace invasive biopsies and improve the expected quality of life of the patients after treatment [27, 34]. The clinical imaging routine for endometrial cancer detection and therapy includes standard MRI sequences. In some cases, additional sequences that might better distinguish tumor patterns are acquired for research purposes. This enables the consideration of different tumor characteristics that are visible in various imaging sequences [18]. The combination of multiple MRI sequences is often referred to as multiparametric medical imaging data. To enable clinical researchers to explore and evaluate tumor sub-regions in cancer research, we propose ICEVis. ICEVis is part of an ongoing collaboration with two clinical experts and enables endometrial cancer researchers to interactively explore hierarchical clustering results based on their multiparametric medical imaging data. The main contribution is a visual analytics approach that supports the assessment of individual distinct tumor sub-regions and relevant features in single patients across imaging sequences. Our approach employs a hierarchical clustering embedded in a visual analytics platform which allows for detailed tumor sub-part analysis. To confirm the utility of our approach, we conducted a qualitative evaluation including three experts in machine learning and gynecological cancer imaging research.

2 MEDICAL BACKGROUND

One of the main research goals of our collaborators is the exploration of endometrial and cervical cancer imaging biomarkers to improve predictions and personalized treatment strategies. Ongoing research aims to analyze relevant imaging features such as texture as well as to develop automatic methods for segmentation and classification [5, 9, 11, 34]. For this reason, the tumor is segmented before the analysis phase, which is currently done mainly manually, taking only one of the several sequences into account. Our collaborators acquired multi-parametric scans of 92 patients
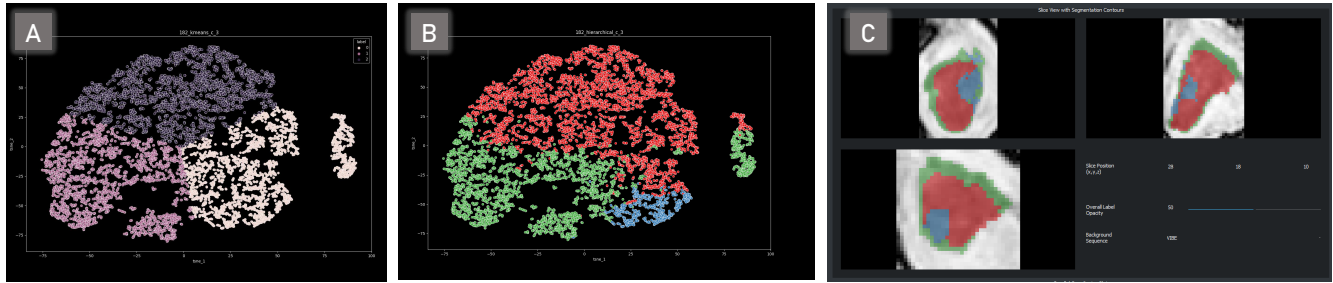
Figure 1: A: a t-SNE result based on all available sequences with subsequent k-Means clustering (k=3). B: hierarchical clustering applied before dimensionality reduction and visualized in the same t-SNE result. C: The hierarchical clustering results in a slice-based view as an overlay on the medical imaging data. The last image shows a necrotic core within the tumor in blue, which would not have been found using k-Means on the t-SNE result.

with endometrial cancer and performed manual segmentation. The seven available sequences for this work are the apparent diffusion coefficient (ADC), Diffusion-weighted magnetic resonance imaging (b1000), peak enhancement (PE), time to peak (TTP), area under the peak enhancement curve (AUC), T2, and contrast-enhanced VIBE (T1-weighted sequence). These multiple sequences provide complementary information, e.g., about the blood flow within the tumor, vascularity, and other physiological properties. They are all aligned via co-registration based on data derived from the scanner. Together with two involved medical collaborators, we analyzed the requirements for our application:

- R1: Find relevant tumor sub-regions based on imaging characteristics per patient
- R2: Support the decision of how many tumor sub-regions are most likely present in the given patient's data
- R3: Enable spatial analysis of the clustering results
- R4: Enable exploration of the results concerning the given imaging characteristics

3 RELATED WORK

In the medical domain, several clustering approaches deliver remarkable results in tissue classification. Juan-Albarracín et al. [15, 16] analyzed malignant brain tumors and managed to improve results obtained by supervised methods. They evaluated five unsupervised voxel classification methods, including K-means clustering, fuzzy K-means, Gaussian Mixture Models (GMM) as a generalization of K-means, and the Gaussian Hidden Markov Random Field (GHMRF) model. Ng et al. [23] used K-Means for medical image segmentation and additionally combined it with an improved watershed algorithm. Raidou et al. [27] applied dimensionality reduction and K-means clustering for cohort analysis of different tumors. Informed by these approaches, we analyzed our data by applying the same techniques but found the results unsatisfactory with our data. After careful consideration, we decided on using a hierarchical clustering approach instead. A comparison with the t-SNE and K-means approaches is visible in Figure 1.

Hierarchical Clustering: Selvan et al. [29] discussed how hierarchical clustering could aid diagnostic imaging data analysis. They outline examples such as the interpretation of X-ray mammography and multi-parametric prostate MRI. In the implementation, they chose the appropriate number of clusters by minimizing the dissimilarity within each cluster. In contrast to their work, we use several

statistical parameters to identify a meaningful number of clusters providing the user with additional decision support.

Visualization for Tumor Tissue Exploration: In 2019, Qu et al. [26] published a review of visual analytics tools for genomic and cancer data, comparing traditional and new methods, among them scatter plots, cluster visualization, and networks. Existing systems for visual analytics in oncologic tissue exploration focus either on the exploration of perfusion data or the distinction between tissues belonging to the tumor and unaffected, healthy regions [27]. In most of the cases [7, 8, 24, 25], perfusion data is used to distinguish between healthy and malignant tissue. We do not have this data available in our approach and work with multi-parametric data of up to seven modalities instead. Yu et al. [35] presented an approach using a heatmap in combination with a dendrogram to analyze the validity of the calculated clusters. In contrast to their approach, we analyze tumor sub-regions instead of the whole tumor volume. Our work is inspired by the approach of Raidou et al. [27]. They propose a visualization tool for the visual examination of the feature space resulting from pre-defined imaging-derived tissue characteristics. Like a related approach [28], they support the assessment of the effect on the parameter space due to different model-based variations in DCE-MRI. A central dimensionality-reduced scatter plot in combination with a density plot guides the interactive exploration. In comparison to their work, we enable clustering of the tumor imaging data before a dimensionality reduction, potentially leading to more accurate clustering results.

4 ICEVIS

Based on the requirements defined together with our medical collaborators, we propose ICEVis. ICEVis consists of a pre-processing pipeline and a visual analytics platform to calculate potential tumor sub-regions. Furthermore, we assist clinical researchers in identifying the optimal number of sub-regions based on imaging characteristics. The pre-processing step of ICEVis employs hierarchical clustering taking all available sequences into account. It is carried out using all available data within the tumor segmentation without prior dimensionality reduction. This step fulfills requirement **R1**. Informed by previous work, we analyzed the results of different clustering techniques (k-Means on the entire dataset, k-Means on the t-SNE result, and hierarchical clustering on the full dataset). We concluded that hierarchical clustering on the full dataset delivered the most promising results. One example of our analysis is presented in Figure 1, where k-Means clustering on the t-SNE
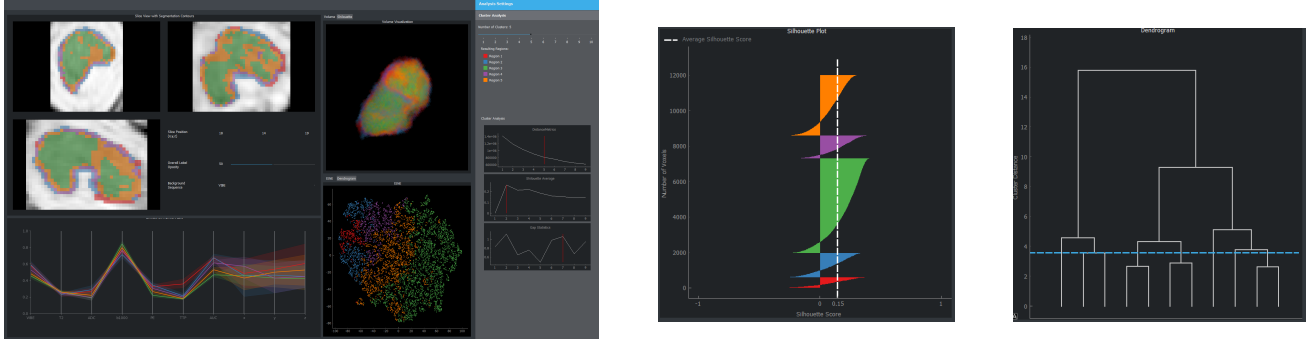
Figure 2: ICEVis consists of a 3D tumor view and calculated cluster information. A parallel coordinate plot, a dendrogram, a silhouette plot, a t-SNE representation of the tumor segmentation, and three optimal cluster number plots are embedded to support the clustering exploration process.

development is juxtaposed with the hierarchical clustering result before applying t-SNE. As an example of the benefits of our approach, a potential histo-pathologically relevant part of the tumor (a necrotic core) is depicted in blue. This important feature was not identified correctly when using t-SNE followed by clustering.

Hierarchical Clustering. To perform hierarchical clustering, different sub-steps are necessary. We first pre-process the data and create feature vectors based on tumor segmentation voxels. Then, we calculate a distance measure by defining the criterion according to which the clusters are split. We normalize the data using min-max normalization. The voxel intensities of each of the seven sequences represent the feature vectors. To compute cluster proximity, different metrics are available. After analyzing the performance of other methods applied to our data (single link, average link, and Ward's method [21]), we decided to use Ward's method [21] for our application. A centroid-based approach defines the proximity between two clusters by calculating the increase of the sum of squared errors (SSE) when potentially merging those two clusters. Like K-means, it minimizes the sum of the squared distances of data objects from their corresponding group centroid.

Determining the Number of Clusters. To support the user in determining an optimal number of clusters and therefore to fulfill requirement **R2**, we present the user with the result of three different measures: the elbow method, the average silhouette method, and the gap statistic. The elbow method is a function of the within-cluster sum of squares. The optimal number of clusters here is the amount where adding another cluster would not improve the result [31]. Kaufman and Rousseeuw [17] introduced the average silhouette method, defining the optimal number of clusters like the one which maximizes the average silhouette score. The last method we support is the gap statistic method by Tibshirani et al. [33]. They propose comparing the within-cluster variation with the expected variation under a reference null distribution. The optimal value is derived by the maximized gap statistic, meaning the cluster structure has the most significant difference in a uniform distribution of points. As there is no definitive answer to how many clusters constitute the optimal solution, we provide the expert with an ensemble of different methods to support the decision process. In Figure 2 on the right side of the first image, all three methods are visualized with the optimal number of clusters highlighted. Note that there is not necessarily agreement among these three methods,

so there is a manual process involved in finding interesting tumor sub-regions by exploring the different cluster number results.

Visualizations. Our tool consists of spatial and non-spatial visualizations, as presented in Figure 2, enabling medical experts to find the optimal number of tumor sub-regions as well as inspect how they change in the imaging data. In close collaboration with our clinical experts, we found these visualization idioms to be the most effective and meaningful use within the proposed application. To enable the analysis of the resulting clusters concerning the underlying imaging data, we employ three orthogonal slice views and a 3D visualization of the data. The clusters are presented as an adaptable semi-transparent overlay, and the underlying medical data is presented in a grayscale colormap. Users can choose which sequence they would like to see. The 3D visualization reveals cluster information in the same colors used in the rest of the tool. This visualization can be used to analyze the spatial relationship between the clusters. These two visualizations fulfill requirement **R3**. To meet requirement **R4** and therefore to characterize clusters by their imaging data content, we add a parallel coordinates plot (PCP). Each feature vector is represented as a polyline, and the axes correspond to the individual features. The PCP is further enriched inspired by box plots [19]. We chose this design over drawing a line for each voxel within the cluster to avoid visual clutter. For each region, the median value of voxel intensities is marked on the axes of the modalities. Their connection forms a line and is visualized in an opaque manner. The more transparent bands around the median lines represent the inter-quartile area to give a better impression of the value range within the clusters.

Cluster Representation. To ensure a meaningful representation of each clustered region, we assign a specific color consistent throughout the different views within the application. We chose the colors using the common *ColorBrewer* [10] tool. To visualize the hierarchical clustering, we created a dendrogram, as shown in Figure 2 on the right side. It shows the hierarchical structure as a tree with distance-based connections between sub-clusters. The length of the vertical lines represents the distance between those clusters, measured by using the Euclidean distance [32]. Furthermore, a cutting line is included based on the selected number of clusters. The vertical lines crossed by the cutting line represent the associated clusters. In addition, the clustering result is presented on top of a t-SNE embedding of the imaging data within the tumor. The feature vectors for the embedding are the seven image sequences
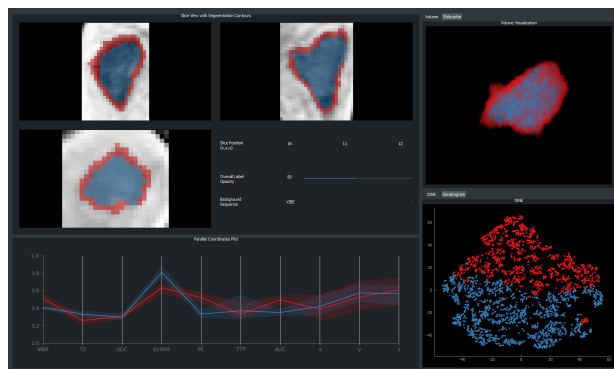
(a) In this particular example, the optimal amount of clusters is two. According to our collaborators, the red area represents the tumor periphery border.
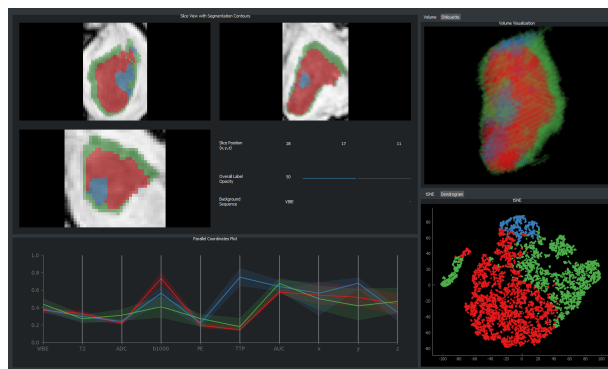
(b) In this example, three clusters differ in ADC, b1000 and in TTP according to the parallel coordinate plot. The blue cluster highlights the necrotic core of the tumor, a region consisting of dead cells within the tumor.

Figure 3: ICEVis use cases presenting interesting cases defined by our clinical collaborators. (a) is showing a necrotic core and (b) is presenting the peripheral border of the tumor.

available for exploration. This allows the medical experts to see if there is a clear difference in clustered voxels within the tumor based on the distance between the clusters in the t-SNE plot.

5 IMPLEMENTATION

Our tool was implemented in Python. For data handling and working with high-dimensional feature arrays, we use the *NumPy* library and for hierarchical clustering we employ the computationally efficient library *fastcluster* [20]. The dendrogram and label extraction are calculated using the *SciPy* [14] cluster hierarchy package. The application was developed with *PyQt5* [4] using the *matplotlib* [13] library and the *PyQtGraph* [3] package for visualizations. Furthermore, the 3D visualization uses *pyvista* [30].

6 RESULTS

Figure 2 depicts the complete user interface for our tool. The number of clusters can be selected via a slider in the analysis settings on the right sidebar. In this case, two clusters have been selected. Below, detected regions with their assigned colors are listed. A slice view allows you to scroll through the slices with a segmentation mask overlay in x-, y-, and z-direction. The slice position is shown numerically in addition and updated when scrolling through the volume. To examine the regions or the imaging data in more detail, the opacity of the segmentation can be adjusted with a slider or by entering a percentage in the corresponding text field. The background sequence can be changed to one of the seven available sequences. The dendrogram, silhouette plot, 3D view, PCP, and t-SNE view, as introduced in Section 4, assist in the exploration and update when changing the number of clusters.

Case Studies and Expert Feedback. We describe four different case studies for our tool based on expert feedback. The first interesting case we found is shown in Figure 3(a). In this case, we found a shell around the tumor. Our experts hypothesized that this region might comprise an inflamed region outside the tumor. This result could impact the treatment decision and may guide the extent to which surgical treatment is carried out. In the second case study presented in Figure 3(b), our medical experts found several three clusters to potentially correlate with histo-pathological properties of the tumor, one of them comprising a necrotic core, which is a distinct pattern within the tumor composed of dead tumor cells. The last case study is included as part of the supplementary material. We

present a case where multiple sequences are not aligned correctly, i.e., misregistration. This can be checked within the ICEVis tool by analyzing the imaging data together with the PCP when using just two clusters. If there is a region that has very distinct values in the PCP in a coordinate axis and any modality, the modality might be wrongly registered. Our collaborating experts see an enormous potential in using ICEVis to spot such cases and then re-registering these using landmark-based registration methods. The result after the correct registration can be seen in Figure 2.

We invited two radiologists involved in gynecological cancer imaging research and one expert in machine learning for medical imaging data to give us feedback on our approach to evaluate our application. All invited experts are not co-authors of this paper but are part of ongoing collaborative research. We got positive feedback on our application, and the experts found exciting and compelling use cases for our application. The two radiologists stated that the regions could be linked to histo-pathological properties, which would have to be confirmed by correlating our results with corresponding surgical findings. When this pattern is established, our application could have an essential impact as such histo-pathological properties could then be considered at the time of imaging and treatment decision. Furthermore, the radiologists would like to use our tool to find areas of the tumor which seem to be the most aggressive parts based on imaging data to consider these for further analysis, e.g., targeted biopsies, especially in case of cervical or prostate cancer where conservative treatment in low-risk patients is more common. The machine learning expert sees a potential use case in further analyzing the results of our approach in a radiomic tumor profiling setup to correlate to clinical variables, e.g., the aggressiveness of the tumor. Overall, all three experts find the helpful tool for different purposes and would like to use it for analyzing their data.

7 CONCLUSION

We proposed ICEVis, an interactive clustering exploration tool with a hierarchical clustering approach embedded in a visual analytics application. The tool includes visualizations to explore clusters and their validity and the characterization of cluster content. We developed our application in close collaboration with cancer imaging experts and conducted a qualitative evaluation including three experts from different fields. Our results show two interesting cases

where the clinical experts were able to find histo-pathological interesting regions within the tumor, which they will compare with the histo-pathological results after the surgery of the tumor. In a future continuation of this work, we would like to enable the clinical experts to analyze the results in an even more interactive way by improving the linking and brushing capabilities of ICEVis. In addition, the application could be used to assist in manual segmentation. Finally, this work represents a step toward patient-specific tumor sub-region exploration for endometrial cancer. It could, in the future, potentially serve as a supportive tool guiding more personalized treatment strategies.

8 ACKNOWLEDGMENTS

REFERENCES

[1] Frederic Amant, Philippe Moerman, Patrick Neven, Dirk Timmerman, Erik Van Limbergen, and Ignace Vergote. 2005. Endometrial cancer. *The Lancet* 366, 9484 (2005), 491–505. https://doi.org/10.1016/S0140-6736(05)67063-8

[2] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. 2018. 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. https://doi.org/10.3322/caac.21492

[3] Luke Campagnola. 2020. PyQtGraph. http://www.pyqtgraph.org/. [Online; accessed 05-October-2020; Version: 0.11.0].

[4] Riverbank Computing. 2020. PyQt5. https://www.riverbankcomputing.com/software/pyqt/. [Online; accessed 05-October-2020; Version: 5.15.0].

[5] Kristine E Fasmer, Atle Bjørnerud, Sigmund Ytre-Hauge, Renate Grüner, et al. 2018. Preoperative quantitative dynamic contrast-enhanced MRI and diffusion-weighted imaging predict aggressive disease in endometrial cancer. *Acta Radiologica* 59, 8 (2018), 1010–1017. https://doi.org/10.1177/0284185117740932

[6] Kristine E. Fasmer, Erlend Hodneland, Julie A. Dybvik, Kari Wagner-Larsen, Jone Trovik, Øyvind Salvesen, Camilla Krakstad, and Ingfrid H.S. Haldorsen. 2021. Whole-Volume Tumor MRI Radiomics for Prognostic Modeling in Endometrial Cancer. *Journal of Magnetic Resonance Imaging* 53, 3 (2021), 928–937. https://doi.org/10.1002/jmri.27444

[7] Sylvia Glaßer, Uta Preim, Klaus Tönnies, and Bernhard Preim. 2010. A visual analytics approach to diagnosis of breast DCE-MRI data. *Computers & Graphics* 34, 5 (2010), 602–611. https://doi.org/10.1016/j.cag.2010.05.016

[8] Sylvia Glaßer, Sophie Roscher, and Bernhard Preim. 2014. Adapted Spectral Clustering for Evaluation and Classification of DCE-MRI Breast Tumors. In *Bildverarbeitung für die Medizin 2014, Algorithmen - Systeme - Anwendungen, Proceedings des Workshops vom 16. bis 18. März 2014 in Aachen (Informatik Aktuell)*, Thomas Martin Deserno, Heinz Handels, Hans-Peter Meinzer, and Thomas Tolxdorff (Eds.). Springer, 198–203. https://doi.org/10.1007/978-3-642-54111-7_39

[9] Ingfrid S Haldorsen, Ingunn Stefansson, Renate Grüner, JA Husby, Inger Johanne Magnussen, et al. 2014. 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. https://doi.org/10.1038/bjc.2013.694

[10] Mark Harrower and Cynthia A. Brewer. 2003. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal* 40, 1 (2003), 27–37. https://doi.org/10.1179/000870403235002042

[11] Erlend Hodneland, Julie A. Dybvik, Kari S. Wagner-Larsen, Veronika Šoltészová, Antonella Z. Munthe-Kaas, Kristine E. Fasmer, Camilla Krakstad, et al. 2021. Automated segmentation of endometrial cancer on MR images using deep learning. *Scientific Reports* 11, 1 (Jan. 2021). https://doi.org/10.1038/s41598-020-80068-9

[12] Erling A. Hoivik, Erlend Hodneland, Julie A. Dybvik, Kari S. Wagner-Larsen, Kristine E. Fasmer, Hege F. Berg, Mari K. Halle, Ingfrid S. Haldorsen, and Camilla Krakstad. 2021. A radiogenomics application for prognostic profiling of endometrial cancer. *Communications Biology* 4, 1 (Dec. 2021). https://doi.org/10.1038/s42003-021-02894-5

[13] John D Hunter. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9, 3 (2007), 90–95.

[14] Eric Jones, Travis Oliphant, Pearu Peterson, et al. 2001. SciPy: Open source scientific tools for Python. (2001).

[15] Javier Juan-Albarracín, Elies Fuster-Garcia, Germán A García-Ferrando, and Juan M García-Gómez. 2019. ONCOhabitats: A system for glioblastoma heterogeneity assessment through MRI. *International Journal of Medical Informatics*

[16] 128 (2019), 53–61. https://doi.org/10.1016/j.ijmedinf.2019.05.002

[16] Javier Juan-Albarracín, Elies Fuster-Garcia, Alexandre Pérez-Girbés, Fernando Aparici-Robles, et al. 2018. Glioblastoma: vascular habitats detected at preoperative dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging predict survival. *Radiology* 287, 3 (2018), 944–954. https://doi.org/10.1148/radiol.2017170845

[17] Leonard Kaufman and Peter J. Rousseeuw (Eds.). 1990. *Finding Groups in Data*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470316801

[18] K. Lawonn, N.N. Smit, K. Bühler, and B. Preim. 2018. A Survey on Multimodal Medical Data Visualization. *Computer Graphics Forum* 37, 1 (2018), 413–438. https://doi.org/10.1111/cgf.13306

[19] Mahsa Mirzargar, Ross T. Whitaker, and Robert M. Kirby. 2014. Curve Boxplot: Generalization of Boxplot for Ensembles of Curves. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2654–2663. https://doi.org/10.1109/TVCG.2014.2346455

[20] Daniel Müllner. 2013. fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. *Journal of Statistical Software* 53, 9 (2013), 1–18. https://doi.org/10.18637/jss.v053.i09

[21] Fionn Murtagh and Pierre Legendre. 2014. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification* 31, 3 (Oct. 2014), 274–295. https://doi.org/10.1007/s00357-014-9161-z

[22] E. Mörth, K. Wagner-Larsen, E. Hodneland, C. Krakstad, I. S. Haldorsen, S. Bruckner, and N. N. Smit. 2020. RadEx: Integrated Visual Exploration of Multiparametric Studies for Radiomic Tumor Profiling. *Computer Graphics Forum* 39, 7 (2020), 611–622. https://doi.org/10.1111/cgf.14172

[23] HP Ng, SH Ong, KWC Foong, PS Goh, and WL Nowinski. 2006. Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE Southwest Symposium on Image Analysis and Interpretation*. IEEE, 61–65. https://doi.org/10.1109/SSIAI.2006.1633722

[24] Steffen Oeltze-Jafra, Helmut Doleisch, Helwig Hauser, Philipp Muigg, and Bernhard Preim. 2007. Interactive Visual Analysis of Perfusion Data. *IEEE transactions on visualization and computer graphics* 13 (11 2007), 1392–9. https://doi.org/10.1109/TVCG.2007.70569

[25] Uta Preim, Sylvia Glaßer, Bernhard Preim, Frank Fischbach, and Jens Ricke. 2012. Computer-aided diagnosis in breast DCE-MRI—Quantification of the heterogeneity of breast lesions. *European Journal of Radiology* 81, 7 (2012), 1532–1538. https://doi.org/10.1016/j.ejrad.2011.04.045

[26] Zhonglin Qu, Chng Wei Lau, Quang Vinh Nguyen, Yi Zhou, and Daniel R Catchpoole. 2019. Visual Analytics of Genomic and Cancer Data: A Systematic Review. *Cancer Informatics* 18 (Jan. 2019). https://doi.org/10.1177/1176935119835546

[27] R.G. Raidou, U.A. van der Heide, C.V. Dinh, G. Ghobadi, J.F. Kallehauge, M. Breeuwer, and A. Vilanova. 2015. Visual Analytics for the Exploration of Tumor Tissue Characterization. *Computer Graphics Forum* 34, 3 (2015), 11–20. https://doi.org/10.1111/cgf.12613

[28] Renata G. Raidou, Marcel Breeuwer, Anna Vilanova, U. A. van der Heide, and P. J. van Houdt. 2014. The iCoCooN: Integration of Cobweb Charts with Parallel Coordinates for Visual Analysis of DCE-MRI Modeling Variations. In *Eurographics Workshop on Visual Computing for Biology and Medicine*, Ivan Viola, Katja Buehler, and Timo Ropinski (Eds.). The Eurographics Association. https://doi.org/10.2312/vcbm.20141178

[29] Arul N Selvan, Laura M Cole, Lynne Spackman, Sarah Naylor, and Chris Wright. 2017. Hierarchical cluster analysis to aid diagnostic image data visualization of MS and other medical imaging modalities. In *Imaging Mass Spectrometry*. Springer, 95–123. https://doi.org/10.1007/978-1-4939-7051-3_10

[30] C. Bane Sullivan and Alexander Kaszynski. 2019. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software* 4, 37 (May 2019), 1450. https://doi.org/10.21105/joss.01450

[31] Robert L. Thorndike. 1953. Who belongs in the family? *Psychometrika* 18, 4 (Dec. 1953), 267–276. https://doi.org/10.1007/bf02289263

[32] Michael Christoph Thrun. 2018. Approaches to Cluster Analysis. In *Projection-Based Clustering through Self-Organization and Swarm Intelligence*. Springer, 21–31. https://doi.org/10.1007/978-3-658-20540-9

[33] Robert Tibshirani, Guenther Walther, and Trevor Hastie. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63, 2 (2001), 411–423. https://doi.org/10.1111/1467-9868.00293

[34] Sigmund Ytre-Hauge, Julie A Dybvik, Arvid Lundervold, Øyvind O Salvesen, Camilla Krakstad, Kristine E Fasmer, Henrica M Werner, Balaji Ganeshan, Erling Høivik, Line Bjørge, et al. 2018. 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. https://doi.org/10.1002/jmri.26184

[35] Lina Yu, Hengle Jiang, Hongfeng Yu, Chi Zhang, Josiah Mcallister, and Dandan Zheng. 2017. iVAR: Interactive visual analytics of radiomics features from large-scale medical images. In *2017 IEEE International Conference on Big Data*. IEEE, 3916–3923. https://doi.org/10.1109/BigData.2017.8258398