

Visual Assessment of Growth Prediction in Brain Structures after Pediatric Radiotherapy

C. Magg¹, L. Toussaint², L. P. Muren², D. J. Indelicato³, R. G. Raidou¹

¹TU Wien, Austria, ²Danish Centre for Particle Therapy, Aarhus, Denmark, ³University of Florida, Jacksonville, USA

Abstract

Pediatric brain tumor radiotherapy research is investigating how radiation influences the development and function of a patient's brain. To better understand how brain growth is affected by the treatment, the brain structures of the patient need to be explored and analyzed pre- and post-treatment. In this way, anatomical changes are observed over a long period and are assessed as potential early markers of cognitive or functional damage. In this early work, we propose an automated approach for the visual assessment of the growth prediction of brain structures in pediatric brain tumor radiotherapy patients. Our approach reduces the need for re-segmentation and the time required for it. We employ as a basis pre-treatment Computed Tomography (CT) scans with manual delineations (i.e., segmentation masks) of specific brain structures of interest. These pre-treatment masks are used as initialization, to predict the corresponding masks on multiple post-treatment follow-up Magnetic Resonance (MR) images, using an active contour model approach. For the accuracy quantification of the automatically predicted post-treatment masks, a support vector regressor (SVR) with features related to geometry, intensity, and gradients is trained on the pre-treatment data. Finally, a distance transform is employed to calculate the distances between pre- and post-treatment data and to visualize the predicted growth of a brain structure, along with its respective accuracy. Although segmentations of larger structures are more accurately predicted, the growth behavior of all structures is learned correctly, as indicated by the SVR results. This suggests that our pipeline is a positive initial step for the visual assessment of brain structure growth prediction.

CCS Concepts

• *Applied computing* → *Life and medical sciences*; • *Human-centered computing* → *Visualization*;

1. Introduction

Radiotherapy (RT) is a well-established and essential treatment modality for children with brain tumors. It uses high-energy ionizing radiation to kill tumor cells, but surrounding normal tissues might also be affected. This can lead to anatomical, cognitive, or functional impairment in young patients, who are particularly sensitive to radiation [BKL*16]. To investigate how RT influences the growth and cognitive development of a child's brain, numerous brain structures of interest need to be segmented and visualized over time in a longitudinal manner [TIS*19]. This facilitates tracking and analyzing anatomical changes, during and after the treatment period. However, during this follow-up period, the anatomy of the patient, but also scanning protocols and/or modalities change. This implies that the follow-up data need to be re-segmented and re-assessed—incurring significant overhead.

In this early work, we propose an approach for the visual assessment of the growth prediction of brain structures in pediatric brain tumor RT patients, which mitigates the overhead of re-segmentation. The contributions of this work are: (1) The design of a *pipeline to automatically predict segmentation masks* for selected brain structures in post-treatment MR images, based on pre-

treatment CT data; (2) an approach to *assess the accuracy* of the predicted segmentations of the brain structures; and (3) an interface to *visualize structure growth*, focusing on how one structure changes over treatment time, along with the *prediction accuracy*.

2. Related Work

A recent survey discusses how visual computing approaches can support different steps within the RT planning workflow [SRM*19]. Although there is no specific previous work on the visual assessment of brain growth after RT, we position our work close to approaches revolving around the use (i.e., support, enhancement, assessment) of (semi-) automated segmentation algorithms within RT. De Geus et al. [DGW96] propose an approach for the detection, modeling, and visual stylization of structures of interest from CT images. Zindy et al. [ZMBL00] propose assisted contouring based on scattered data interpolation methods. Raidou et al. use visual analytics to facilitate the assessment of outcomes and errors of segmentation methods for cohort and individual patient investigation [RMB*16]. Schlachter et al. [SFA*17] propose a visualization framework for rapid quality assessment of segmentation, targeting temporal fusion data. Other recent related works address the exploration and analysis of anatomical variabil-

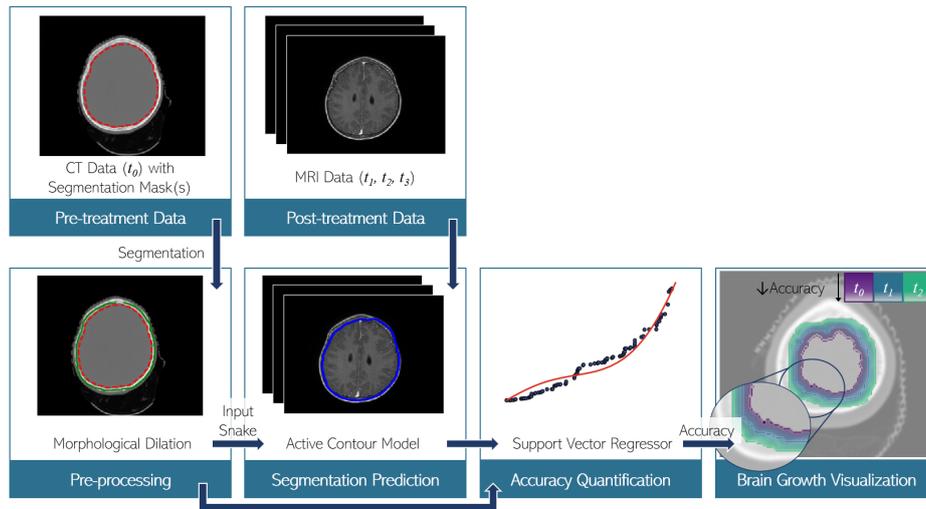


Figure 1: The workflow employed for the visual assessment of predicted brain structure growth: After pre-processing, the segmentation of brain structures is predicted in the post-treatment imaging ($t_1 - t_3$) using the pre-treatment segmentation (t_0) through an active contour model. Then, the prediction accuracy is quantified with a support vector regressor (SVR). Finally, the predicted segmentation is visualized together with its accuracy on the underlying CT/MRI data.

ity [KLR*13,RCMA*18,FGM*20] and the prediction of anatomical motion through treatment [FMC*21]. Finally, uncertainty and its quantification has been tackled multiple times in relation to the medical domain [RPHL14,GSWS21].

3. Background

Current Workflow: Pediatric brain tumor RT investigates the effects of radiation doses delivered to different brain structures [BKL*16,TIS*19]. The response of distinct brain structures to different RT doses and treatment modalities is studied through radiobiological models that evaluate different treatment strategies and their respective dose/volume parameters. The *quantitative* pre-treatment assessment of the dose delivered to pre-defined structures ensures that radiation limits are respected, thereby limiting the risk of side effects for the patient. The *qualitative* (i.e., visual) pre- and post-treatment assessment of the growth patterns of these structures enables to follow the patient's cognitive and functional development over time. Both approaches require that the brain structures of interest are accurately delineated. This is most commonly done manually by medical experts (e.g., radiation oncologists) and is a very time-consuming procedure—especially, with a high number of structures in a long follow-up period. Automating this step could reduce significantly processing time.

Dataset: The available dataset consists of 20 folders, corresponding to 20 pediatric patients with previously treated brain tumors. The data includes CT pre-treatment scans (referred to as t_0 , in this paper) and 1–3 additional pre- and post-treatment MRI data (referred to as $t_1 - t_3$). All the data are registered to the coordinate system of each patient. For the pre-treatment CT scans, segmentation masks of over 100 structures are available. For this project, 21 brain structures were selected as the most relevant, based on previous work that associates these structures with cognitive side effects [TIS*19].

Tasks: The current workflow could improve significantly through the integration of:

- [T1] **Segmentation Prediction:** Predict the segmentations of brain structures post-treatment, based on pre-treatment data.
- [T2] **Accuracy Quantification:** Quantify the accuracy of the predicted segmentations.
- [T3] **Comprehensive Visualization:** Visualize structure evolution over time (pre-, post-treatment), along with its accuracy.

4. Visual Assessment of Predicted Brain Structure Growth

In this section, we describe the workflow for the visual assessment of the predicted growth of brain structures in pediatric RT patients. The workflow is schematically depicted in Figure 1. First, a *pre-processing* step is conducted. This is followed by the *segmentation prediction*, where an active contour model is initialized with the dilated version of the pre-treatment segmented structures to predict the post-treatment segmentations. Here, a hyperparameter search is conducted to obtain adequate parametrizations of the algorithm. Subsequently, the *prediction accuracy* is quantified using a support vector regressor (SVR), trained on the pre-treatment data. Finally, the *visualization* of the predicted segmentation and its accuracy supports the visual assessment of brain structure growth.

Pre-processing: The CT and MRI images are stored as DICOM files. The first step is to extract the voxel data and convert their world coordinates to image coordinates, considering translation and scaling with voxel spacing (default: 0.97 mm x 0.97 mm x 1 mm). The segmentation masks for the pre-treatment CT scans are also stored using the DICOM format, and for each structure a binary mask is available. The folders 1–15 are used for developing the pipeline, while folders 16–20 are locked and kept for testing.

Segmentation Prediction: Segmentation masks are only available for pre-treatment CT scans. From these, the segmentation masks

for post-treatment MRI data need to be predicted [T1]. Conceptually, the growth of brain structures is anticipated to be similar to a “dilation”. An active contour model, initialized with a dilated pre-treatment segmentation, fits well to the anticipated growth pattern. Figure 2 illustrates the pipeline for the predicted segmentation of a brain structure. First, the segmentation contour from the pre-treatment CT scan (t_0) undergoes a morphological dilation, to generate a larger mask. The reason for this is that all structures are expected to grow over time. Then, the dilated mask is used as the initialization (initial snake coordinates) for the active contour optimization in the post-treatment MRI scans. This is possible, as the data are co-registered.

Brain structures range from very small to almost area-filling shapes and might comprise a single or multiple components. Additionally, most brain structures have a clear boundary, but adjacent tissues have similar intensity values. To identify a method that can handle all these variations, we empirically tested several approaches, among which the basic active contour model [KWT88], the Chan Vese algorithm [CV99], morphological active contours without edges (MorphACWE) [MBA14], and morphological geodesic active contours (MorphGAC) [MBA14]. MorphACWE and MorphGAC do not yield satisfying results, given the similarity of intensity values and the effort required for their initialization, respectively. The most feasible and satisfying initial results are achieved with the active contour model.

The active contour parameters and the kernel size of the dilation are hyperparameters, which need to be adequately chosen. This choice is not trivial [DR12]. Three parameters—kernel size of dilation k , smoothness of the snake shape β , and maximal number of iterations used to optimize the snake n —influence mainly the result. We perform a sparse grid search on half of the data (folders 1–10) to determine favorable settings for each structure.

Since the ground truth information is available, we can quantify its deviation from the predicted segmentation within the hy-

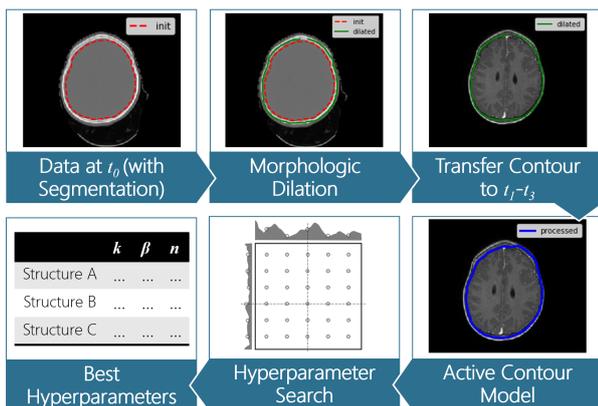


Figure 2: The predicted segmentation pipeline: The pre-treatment segmentation mask (t_0 , red dotted line) is dilated (green) and employed as initialization for segmentation of the post-treatment data ($t_1 - t_3$) using active contours (result shown in blue). Then, a hyperparameter search is employed to fine-tune the parametrization.

perparameter search for each structure. For this, we use the *Dice Coefficient*, defined as $\frac{2|A \cap B|}{|A| + |B|}$, where A and B are the volumes of the predicted structure segmentation and the ground truth data, respectively. This procedure is performed for each structure and each patient, in two rounds. In the first round, we investigate all possible combinations of the following parametrizations: $k \in \{5, 10\}$, $\beta \in \{0.05, 0.1, 0.2\}$, and $n \in \{2, 5, 8, 15, 20, 30\}$. This results in 36 different settings for the sparse grid search. In the second round, the local neighborhood of the best values is revisited. To obtain the optimal parameter setting for the entire segmentation workflow, the final parameter values per structure are derived by a majority vote over the data of all training patients (i.e., folders 1–10). This ensures that our pipeline will also work for unseen data. The final values for all structures are included in the supplementary material.

Accuracy Quantification: After the prediction, we need a quantitative measurement for the segmentation accuracy [T2]. We employ a support vector regressor (SVR), trained on the pre-treatment data, as illustrated in Figure 3. Our approach is based on the learning approach by Kohlberger et al. [KSA*12]. First, the segmentation of the pre-treatment data (t_0) are re-generated using the active contour model. Together with the ground truth information, they build the training data for the SVR. At test time, the accuracy of post-treatment segmentations ($t_1 - t_3$) is predicted.

The SVR is set up similarly to the work of Kohlberger et al. The variables are calculated per structure, and for each data slice. The 35 employed independent variables are a combination of geometry, intensity and gradient features, and selected ratios of those. Since the feature values have a large range and SVR is not scale-invariant, they are standardized by subtracting the mean and dividing by standard deviation. Five error metrics are implemented as dependent variables: Jaccard distance, Dice Coefficient, Hausdorff distance, modified Hausdorff distance, and average surface error [KSA*12]. The *scikit-learn* implementation of SVR with a radial basis function kernel and $\epsilon = 0.01$ is used. The best value for the regularization parameter $C \in [1, 99]$ is determined with 5-fold cross validation [HCL03]. Therefore, the dataset is split into five parts. One part is used for testing, and the remaining four are used for fitting the SVR with a specific C value. This is done five times, so that

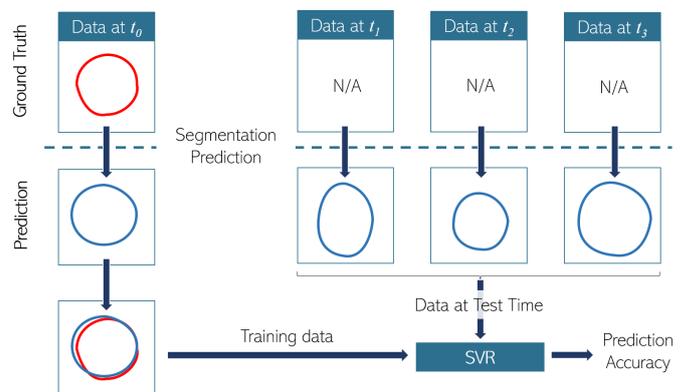


Figure 3: Our approach for training the support vector regressor (SVR) and for quantifying the accuracy of the prediction.

each partition is used once as test set. Then, the results for each C are averaged and compared to the results for other C values for each structure (see also supplementary material). To evaluate the SVR performance, the coefficient of determination R^2 of prediction is used. In addition, we test which dependent variable combination works best per brain structure labels (21 in total). The best single-output SVR was with Jaccard distance (10/21 labels), followed by the Dice Coefficient (6/21 labels). The best multi-output combination is Dice Coefficient with Jaccard distance (14/21 labels).

Comprehensive Visualization: For the visual assessment of the predicted brain structure growth, we display structural changes over time for each structure for one patient, along with the prediction accuracy [T3]. We opt for a simple 2D slice-based view on the three anatomical planes, due to its simplicity and widespread use in the clinical setting. We refrain from employing a 3D rendering, due to visibility challenges—especially, for inner structures.

To get the distances between a timestep and its successor, distance transform is applied to the segmentation mask of one timestep, and is clipped to the area enclosing the other timestep. If a structure grows in one direction and shrinks in another, the distance transform needs to be applied to both segmentation masks to obtain the distances in all directions. The distances are encoded by per-voxel intensity values (0 to 255, where 255 corresponds to the largest distance between two timesteps). Our approach only takes timelines into account—not the point-correspondence between two timesteps. Thus, note that voxel values do not vary from 0 to 255 following a streakline; instead, the maximal value on a streakline is scaled with respect to the maximal distance of the overall mask.

After generating the distance masks, a lookup table is applied to the data, to map distance values to respective colors. We select the HSV color space to encode the timestep transition, using the hue, and the prediction accuracy, using the formula for the saturation s :

$$s(a) = \begin{cases} 0.2, & \text{for } a \leq 0.5 \\ 0.5, & \text{for } 0.5 < a \leq 0.75 \\ 0.9, & \text{for } 0.75 < a \leq 0.9 \\ 1.0, & \text{for } 0.9 < a \leq 1 \end{cases}$$

where a is the prediction accuracy. The cut-offs and values for the saturation formula are empirically tested, and they are chosen in a non-linear way to highlight the segmentation masks with higher accuracy (i.e., more certain predictions are more prominent). For example, the initial timesteps (t_0) are always fully saturated, as they are not predictions, but actual segmentation data. The user can choose between three different color maps: a qualitative RGBO color map following the ColorBrewer guidelines, a binned plasma, and a binned viridis colormap. All three colormaps are defined for a maximum of four bins (i.e., timesteps), which is the maximum in our application. For a small number of timesteps higher than four, the approach can be easily extended, but for a high number of timesteps, it might need to be reworked in the future. If a dataset contains less timesteps, then only the first few colors are used.

The segmentation contours are additionally displayed to provide further context of the structure boundaries at the scanning time-points. The colors match the colors assigned to their respective segmentation mask and they are shown with full saturation, to make the boundary more prominent and to provide a link between the

segmentation mask and the boundary. The segmentations and their contours are overlaid on top of the respective CT/MRI data following common clinical practice, to provide anatomical contexts and the additional possibility to verify the predicted segmentation masks. This part is implemented in Python, using PyQT and the Visualization Toolkit (VTK), and is presented in Figure 4.

5. Results and Discussion

The predicted error metrics of our segmentation pipeline show a wide range for different brain structures. The average Jaccard distance is higher for large structures, such as the brain (0.81), temporal lobe (0.74), and scalp (0.66), as compared to smaller structures, such as the thalamus anterior (0.09), hypothalamus (0.11), and cingulum (0.10). The same behavior is observed for other error metrics. However, in the k -fold cross validation, the SVR is able to predict the Jaccard distance with a mean squared error between 0.03 – 0.08, which means that the behavior observed in the training dataset is learned correctly. The SVR scores and the best error metrics for the k -fold cross validation are presented in detail for all investigated brain structures in the supplementary material.

Although our approach delivers satisfactory results, regardless the underlying image modality, a drawback is that the segmentation prediction and feature calculation for feeding the SVR cannot be performed in real-time, due to computation duration and used memory storage. Our next steps will focus on obtaining a more efficient implementation that could speed up the process and facilitate in-line integration. As an alternative, a deep learning approach for semantic segmentation could replace the prediction and the accuracy quantification [RFB15], given the built-in probability of certainty that a voxel belongs to a specific prediction class in neural networks. This could also resolve the speed and memory issues. Yet, a sufficient amount of data and a dedicated training phase would be a strong pre-requisite.

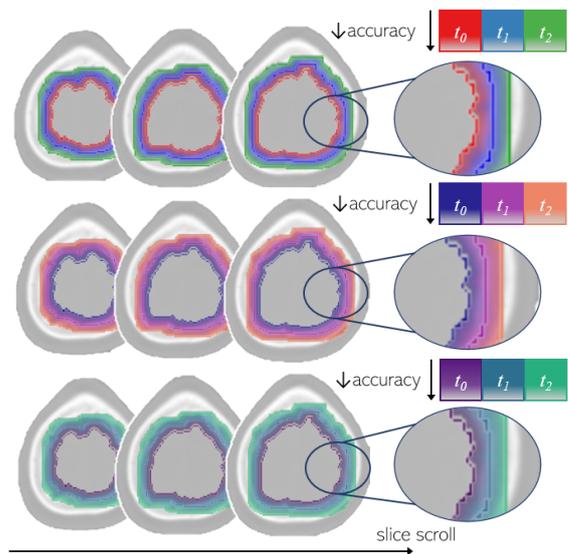


Figure 4: 2D views for the visual assessment of the predicted segmentations and their accuracy with different color schemes. Hue indicates the timestep and saturation the accuracy.

The visualization could also be improved and/or extended. Streaklines to show the point-correspondence between segmentations could be employed in addition to the timelines, by utilizing a 2D color map [SBM*15], or by integrating other state-of-the-art visual designs, which could handle better 3D views. For this, in the future, we intend to look further into the domain of illustrative flow visualization, which could provide significant inspiration for the representation of growth [BCP*12]. Additionally, the extension to the simultaneous visualization of multiple structures for multiple patients, also within a comparative context would be an interesting direction to support cohort analysis and patient stratification. This case would be particularly useful and challenging (in terms of visualization design) if combined with co-occurring assessment of RT plans, to analyze and predict which strategy would be more suitable for each patient case [FMCM*21]. Finally, a thorough quantitative evaluation with a higher number of patients and a thorough user study would be required to assess the suitability of our approach for use within the clinical workflow.

6. Conclusion

The main contribution of this work is the design of a reusable pipeline to predict segmentations for new follow-up post-treatment data on the basis of pre-treatment scans, and the quantification of the respective prediction accuracy. These are conducted using, respectively, an adapted active contours methodology and a SVR approach. To visualize the predicted segmentations along with their prediction accuracy on top of imaging data, we provide a simple, proof-of-concept visualization for the time evolution of the segmentation masks. This is an initial step towards the visual assessment of the predicted growth of brain structures of pediatric RT patients, with good initial results and interesting future directions.

References

- [BCP*12] BRAMBILLA A., CARNECKY R., PEIKERT R., VIOLA I., HAUSER H.: Illustrative Flow Visualization: State of the Art, Trends and Challenges. In *Eurographics 2012 - State of the Art Reports* (2012). 5
- [BKL*16] BRINKMAN T. M., KRASIN M. J., LIU W., ARMSTRONG G. T., OJHA R. P., SADIGHI Z. S., GUPTA P., KIMBERG C., SRIVASTAVA D., MERCHANT T. E., GAJJAR A., ROBISON L. L., HUDSON M. M., KRULL K. R.: Long-term neurocognitive functioning and social attainment in adult survivors of pediatric CNS Tumors: Results from the St Jude Lifetime Cohort Study. *Journal of Clinical Oncology* 34, 12 (2016), 1358–1367. 1, 2
- [CV99] CHAN T., VESE L.: An active contour model without edges. In *Scale-Space Theories in Computer Vision* (1999), pp. 141–151. 3
- [DGW96] DE GEUS K., WATT A.: Three-dimensional stylization of structures of interest from computed tomography images applied to radiotherapy planning. *International Journal of Radiation Oncology* Biology* Physics* 35, 1 (1996), 151–159. 1
- [DR12] DARVISH A., RAHNAMEYAN S.: Optimal parameter setting of active-contours using differential evolution and expert-segmented sample image. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 16 (09 2012), 667–686. 3
- [FGM*20] FURMANOVÁ K., GROSSMANN N., MUREN L. P., CASARES-MAGAZ O., MOISEENKO V., EINCK J. P., GRÖLLER M. E., RAIDOU R. G.: VAPOR: Visual Analytics for the Exploration of Pelvic Organ Variability in Radiotherapy. *Computers & Graphics* 91 (2020), 25–38. 2
- [FMCM*21] FURMANOVÁ K., MUREN L. P., CASARES-MAGAZ O., MOISEENKO V., EINCK J. P., PILSKOG S., RAIDOU R. G.: Previs: Predictive visual analytics of anatomical variability for radiotherapy decision support. *Computers & Graphics* 97 (2021), 126–138. 2, 5
- [GSWS21] GILLMANN C., SAUR D., WISCHGOLL T., SCHEUERMANN G.: Uncertainty-aware Visualization in Medical Imaging - A Survey. *Computer Graphics Forum* (2021). 2
- [HCL03] HSU C.-w., CHANG C.-c., LIN C.-j.: A practical guide to support vector classification. *Tech. Rep., Department of Computer Science, National Taiwan University* (11 2003). 3
- [KLR*13] KLEMM P., LAWONN K., RAK M., PREIM B., TÖNNIES K. D., HEGENSCHIED K., VÖLZKE H., OELTZE S.: Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data. In *Vision, Modeling and Visualization* (2013), pp. 121–128. 2
- [KSA*12] KOHLBERGER T., SINGH V., ALVINO C., BAHLMANN C., GRADY L.: Evaluating segmentation error without ground truth. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012* (2012), pp. 528–536. 3
- [KWT88] KASS M., WITKIN A., TERZOPOULOS D.: Snakes: Active contour models. *International Journal of Computer Vision* 1, 4 (1988), 321–331. 3
- [MBA14] MÁRQUEZ-NEILA P., BAUMELA L., ALVAREZ L.: A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36, 1 (2014), 2–17. 3
- [RCMA*18] RAIDOU R. G., CASARES-MAGAZ O., AMIRKHANOV A., MOISEENKO V., MUREN L. P., EINCK J. P., VILANOVA A., GRÖLLER M.: Bladder Runner: Visual Analytics for the Exploration of RT-Induced Bladder Toxicity in a Cohort Study. *Computer Graphics Forum* 37, 3 (2018), 205–216. 2
- [RFB15] RONNEBERGER O., FISCHER P., BROX T.: U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (2015), vol. 9351. 4
- [RMB*16] RAIDOU R. G., MARCELIS F. J. J., BREEUWER M., GRÖLLER E., VILANOVA A., WETERING H. M. M. v. D.: Visual Analytics for the Exploration and Assessment of Segmentation Errors. In *Eurographics Workshop on Visual Computing for Biology and Medicine* (2016), pp. 193–202. 1
- [RPHL14] RISTOVSKI G., PREUSSER T., HAHN H. K., LINSEN L.: Uncertainty in medical visualization: Towards a taxonomy. *Computers & Graphics* 39 (2014), 60–73. 2
- [SBM*15] STEIGER M., BERNARD J., MITTELSTÄDT S., HUTTER M., KEIM D., THUM S., KOHLHAMMER J.: Explorative analysis of 2D color maps. In *Proceedings of WSCG* (2015), vol. 23, Eurographics Association, pp. 151–160. 5
- [SFA*17] SCHLACHTER M., FECHTER T., ADEBAHR S., SCHIMEK-JASCH T., NESTLE U., BÜHLER K.: Visualization of 4D multimodal imaging data and its applications in radiotherapy planning. *Journal of applied clinical medical physics* 18, 6 (2017), 183–193. 1
- [SRM*19] SCHLACHTER M., RAIDOU R., MUREN L., PREIM B., PUTORA P., BÜHLER K.: State-of-the-Art Report: Visual Computing in Radiation Therapy Planning. *Computer Graphics Forum* 38, 3 (2019), 753–779. 1
- [TIS*19] TOUSSAINT L., INDELICATO D. J., STOKKEVÅG C. H., LASSEN-RAMSHAD Y., PEDRO C., MIKKELSEN R., PINTO M. D., LI Z., FLAMPOURI S., VESTERGAARD A., PETERSENN J. B. B., SCHRÖDER H., HØYER M., MUREN L. P.: Radiation doses to brain substructures associated with cognition in radiotherapy of pediatric brain tumors. *Acta Oncol.* 58, 10 (2019), 1457–1462. 1, 2
- [ZMBL00] ZINDY E., MOORE C., BURTON D., LALOR M.: Morphological definition of anatomic shapes using minimal datasets. In *2000 IEEE Conference on Information Visualization. An International Conference on Computer Visualization and Graphics* (2000), IEEE, pp. 366–370. 1