Exploit ¹⁸F-FDG enhanced urinary bladder in PET data for deep learning ground truth generation in CT scans

Christina Gsaxner, Birgit Pfarrkirchner, Lydia Lindner, Norbert Jakse, Jürgen Wallner, et al.

Christina Gsaxner, Birgit Pfarrkirchner, Lydia Lindner, Norbert Jakse, Jürgen Wallner, Dieter Schmalstieg, Jan Egger, "Exploit ¹⁸F-FDG enhanced urinary bladder in PET data for deep learning ground truth generation in CT scans," Proc. SPIE 10578, Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging, 105781Z (12 March 2018); doi: 10.1117/12.2292706

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Exploit ^{18}F-FDG Enhanced Urinary Bladder in PET Data for Deep Learning Ground Truth Generation in CT Scans

Christina Gsaxner[a,b], Birgit Pfarrkirchner[a,b], Lydia Lindner[a,b], Norbert Jakse[c], Jürgen Wallner[b,c]
Dieter Schmalstieg[a] and Jan Egger[a,b,d]

[a] TU Graz, Institute for Computer Graphics and Vision, Inffeldgasse 16c/II, 8010 Graz, Austria
[b] Computer Algorithms for Medicine (Cafe) Laboratory, 8010 Graz, Styria, Austria
[c] Medical University of Graz, Department of Maxillofacial Surgery, Auenbruggerplatz 12, 8036 Graz, Austria
[d] BioTechMed-Graz, Krenngasse 37/1, 8010 Graz, Austria

ABSTRACT

Accurate segmentation of medical images is a key step in medical image processing. As the amount of medical images obtained in diagnostics, clinical studies and treatment planning increases, automatic segmentation algorithms become increasingly more important. Therefore, we plan to develop an automatic segmentation approach for the urinary bladder in computed tomography (CT) images using deep learning. For training such a neural network, a large amount of labeled training data is needed. However, public data sets of medical images with segmented ground truth are scarce. We overcome this problem by generating binary masks of images of the ^{18}F-FDG enhanced urinary bladder obtained from a multi-modal scanner delivering registered CT and positron emission tomography (PET) image pairs. Since PET images offer good contrast, a simple thresholding algorithm suffices for segmentation. We apply data augmentation to these datasets to increase the amount of available training data. In this contribution, we present algorithms developed with the medical image processing and visualization platform MeVisLab to achieve our goals. With the proposed methods, accurate segmentation masks of the urinary bladder could be generated, and given datasets could be enlarged by a factor of up to 2500.

Keywords: Medical Image Segmentation, Urinary Bladder, Deep Learning, Combined PET/CT, ^{18}F-FDG, Data Augmentation.

1. DESCRIPTION OF PURPOSE

Since imaging modalities like computed tomography (CT) are widely used in diagnostics, clinical studies, treatment planning and evaluation, automatic algorithms for image analysis have become an invaluable tool in medicine. In particular, deep learning approaches have made a large impact on the field of medical image analysis over the past view years. Machine learning algorithms can be applied to problems like image detection or recognition, segmentation, registration or computer-aided diagnosis and disease quantification, just to name a few [1]. Image segmentation algorithms are of special interest, since segmentation plays a vital role in many medical applications, like the quantification of tissue volumes [2]-[5], localization of pathologies [7]-[10], treatment planning (especially in radiotherapy) [11], and computer-integrated surgery [12]. To this day, delineation is often done manually or semi-manually, especially in regions with limited contrast and for organs or tissues with large variations in geometry. This is a tedious task, since it is time-consuming and requires a lot of experience. Furthermore, the process of manual segmentation is prone to errors and, since it is highly dependent on the operator, not reproducible [13].

Therefore, we aim to develop an automatic approach for medical image segmentation, specifically for segmentation of the urinary bladder in CT images, based on deep learning. One limitation of neural networks is that a large amount of labeled training data is needed to provide enough information to specify all the network's connections, especially if the network is large and deep. For general-purpose image processing tasks, this is not a problem anymore, since open databases containing millions of labeled images exist. For medical data, however, this is a major drawback, since public medical databases are generally small [14]. Furthermore, already segmented CT images are even harder to find in open databases.

Therefore, two goals can be defined for this project: Generate binary masks of the urinary bladder as a reference standard for training a deep neural network and enlarge the size of available datasets. We address the first problem by using combined positron emission tomography and computed tomography (PET-CT) scans. PET imaging is based on measuring radiation emitted by a radiotracer injected into the patient. Radiotracers are chosen to accumulate in regions relevant for specific screening, like inflammatory sites or tumor cells. By backtracking the measured radiation to its source,

Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging,
edited by Barjor Gimi, Andrzej Krol, Proc. of SPIE Vol. 10578, 105781Z · © 2018 SPIE
CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2292706

such regions can be imaged [15]. Unlike CT, PET images exhibit high contrast and are therefore comparably easy to segment automatically [16]. Therefore, PET scans can easily be exploited to generate binary masks as a reference standard for training a neural network. We found a simple thresholding sufficient for segmentation

To achieve the second goal, namely, increasing the amount of training data, data augmentation is used. The basic idea of data augmentation is to create plausible changes, which preserve label information, to the existing data [17]. The application of geometric transformations, the addition of noise or color jittering are utilized to generate new, additional training data that is similar to, but not the same as the existing data. In our approach, rotation and scaling are applied to CT images as well as to the masks generated from the corresponding PET images. Furthermore, different noise types can be added to CT images. Since the size and shape of the urinary bladder is highly variable between patients and individual slices, these augmentations seem meaningful. A huge amount of training sets can be generated with this approach.

For the implementation of the proposed algorithms, we used the medical imaging framework MeVisLab [18]-[25]. MeVisLab is a modular framework for the development of medical image processing algorithms and visualization of medical data. It allows development via visual programming using preexisting modules, macro module-creation via Python scripting, and module implementation in C++.

2. METHODS

For this study, the *reference image database to evaluate therapy response* (RIDER) [26] data sets of PET/CT images were used. RIDER contains a total of 65 scans of lung cancer patients. As radiotracer, fluorine-18-labelled fluorodeoxyglucose (^{18}F-FDG), which always accumulates in the urinary bladder, was used in all PET scans. Patient data with very high noise levels in either the CT or PET data or with unusually low contrast in the region of interest were rejected. The general network implemented in MeVisLab can be seen in Figure 1.

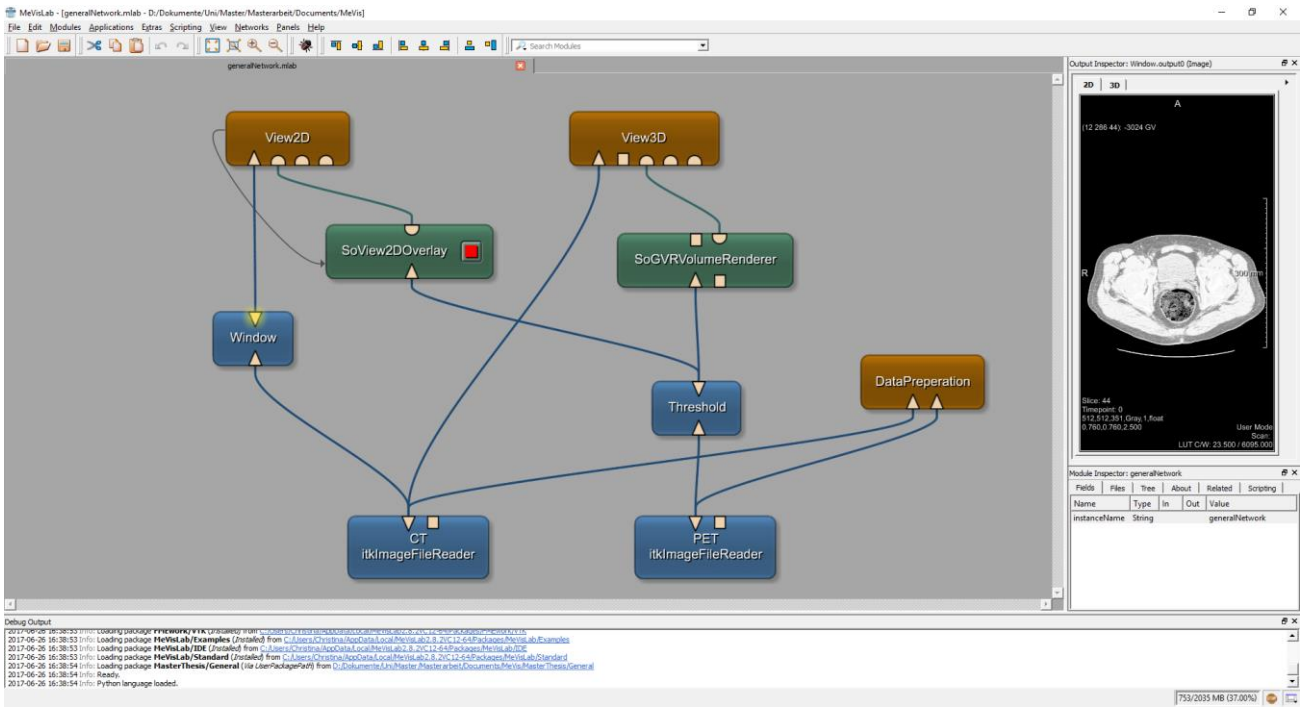


Fig. 1: General Network for loading, processing and visualizing PET and CT data, implemented in MeVisLab.

Corresponding PET and CT data is loaded into the framework and fed into the *DataPreperation* macro-module for processing. The remaining modules are for visualization purposes. The internal network of the *DataPreperation* module can be observed in Figure 2. The network consists of two groups, one for CT and one for PET image processing. First, general information about the input image is extracted via the *Info* module. This module provides information like image size, image type, maximal and minimal pixel value. The *Reformat* module transforms the PET dataset to the local coordinate system of the CT dataset using trilinear interpolation. The *SubImage* module allows the extraction of sub regions from input images. By iterating over the z-coordinate of a dataset, while leaving the other parameters unchanged,

individual transversal slices can be selected. Next, the extracted PET slices are segmented by the application via the *Threshold* module. A fixed threshold is calculated for each dataset in the Python script. Every pixel above the threshold is considered foreground; all other pixels are labeled as background. The output of this module is a binary image of the urinary bladder. Data augmentation of the binary masks and CT slices is performed using the *AffineTransformation2D* module. This module enables the application of several affine transformations in 2D. For this macro module, rotation and scaling is enabled. Further, the *AddNoise* module enables the addition of noise from various distributions to CT slices. The *Scale* module allows scaling and conversion of the image file type. The contrast for soft tissue in CT images is adjusted by applying windowing using the *Window* module. At last, the produced slices are saved in TIFF format using the *ImageSave* module. Interactions between these modules and with inputs and outputs, as well as module parameters, were realized via Python scripting. Furthermore, a user interface was created using the MeVisLab Definition Language (MDL) script file [27]. The code is freely available for download [28]:

https://github.com/cgsaxner/DataPrep_UBsegmentation

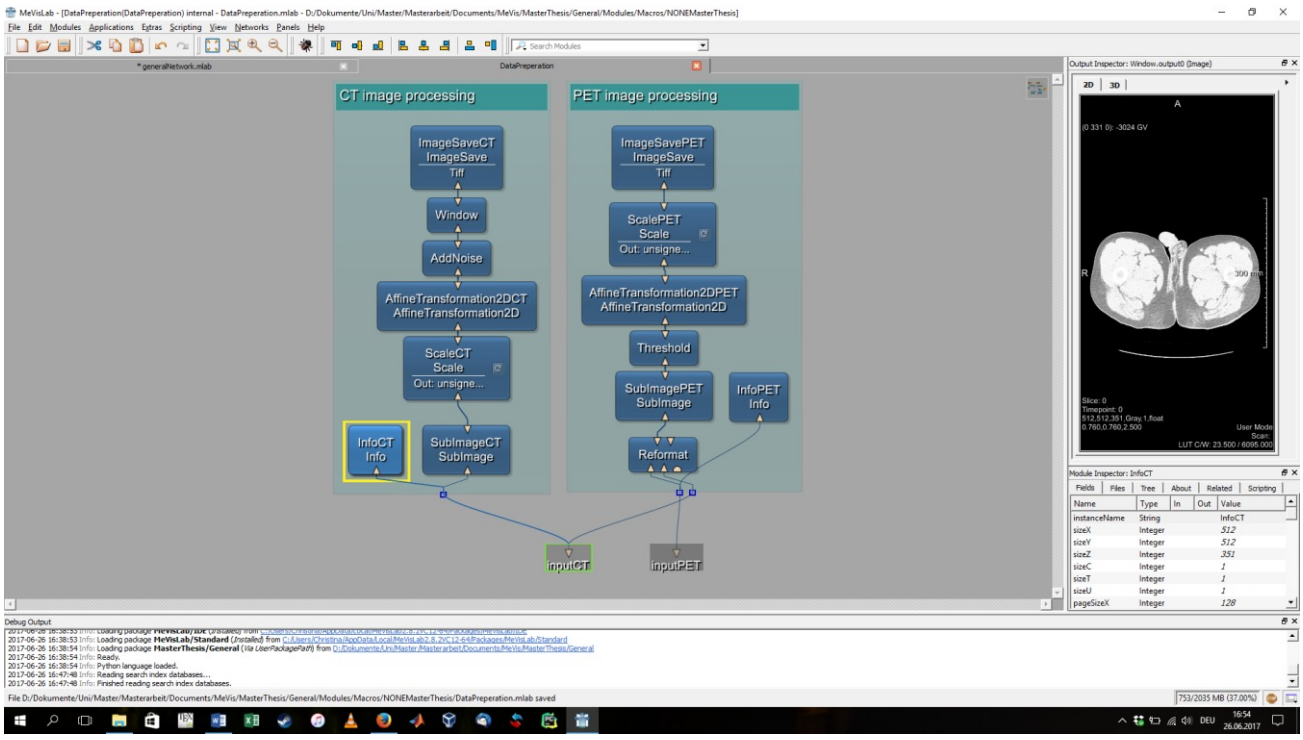


Fig. 2: The internal network of the *DataPreperation* macro module.

3. RESULTS

After removing patient data with low contrast and high noise from the RIDER database, a total of 33 patient datasets were obtained. The CT datasets offer between 122 and 358 transversal slices (297 in average), yielding to a total of 9819 images. Since these scans cover the whole torso, the urinary bladder is only visible in a fractional amount of the images. The average number of slices showing the bladder is 26, resulting in a total of 853 slices. To visualize the results of the PET segmentation using thresholding, Figure 3 shows and overlay of CT slices (in grayscale) and their corresponding binary masks of the urinary bladder generated from the respective PET datasets (in red). Representative datasets were chosen for illustration.

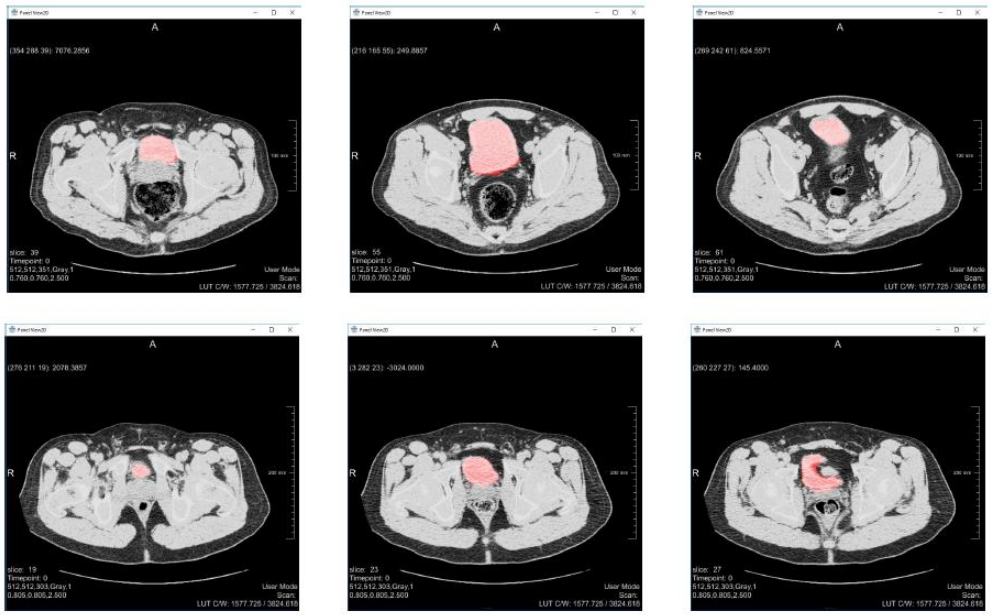
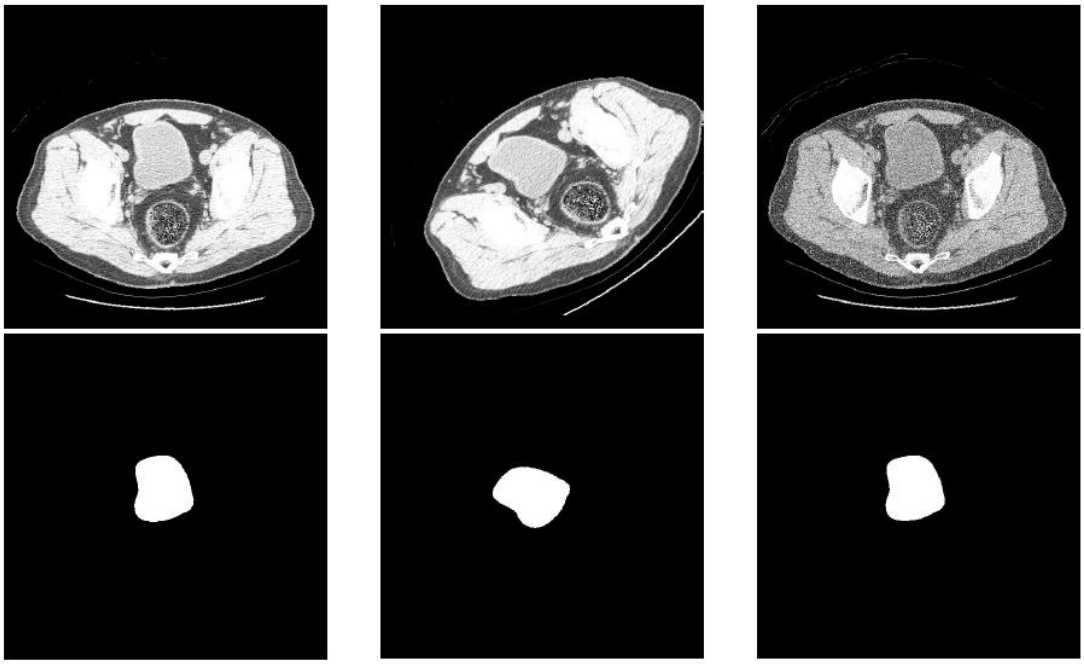


Fig. 3: Overlay of CT data and binary mask of two RIDER PET/CT datasets. The selected slices were taken from the beginning, middle and end of the urinary bladder. The binary mask obtained by thresholding the PET data is shown in red.



(a) Original (b) Rotated and scaled (c) Added noise

Fig. 4: Original and augmented datasets (upper row) and corresponding binary masks of the urinary bladder (lower row, white). Image (a) shows the original dataset without augmentations. Image (b) shows the dataset after applying a rotation of 45° and scaling to a scale factor of 0.9 in x-direction and 1.1 in y-direction to both the CT image and the binary mask. In image (c), zero-mean Gaussian noise with a standard deviation of 500 gray values was added to the CT image. The binary mask remains unchanged by the addition of noise.

With the default parameters specified in Table 1, 32 augmented datasets can be created from each slice in the input data. With the urinary bladder showing an average of 26 slices per dataset, this results in 832 augmented images and binary masks per patient data. With 853 slices of the urinary bladder obtained from the 33 selected original patient datasets, this results in a total of 27,296 positive training datasets. Setting the number of generated slices in each augmentation step to their maximal value results in 2500 augmented datasets obtained from each input slice. With these settings, a total of 2,132,500 positive training datasets could be achieved. Figure 4 shows two examples of augmented training datasets.

Augmentation Type		Parameter	Default	Minimum	Maximum
Rotation		Maximal rotation angle	45	0	180
		Number of rotations per slice	4	1	10
Scaling		Maximal scale factor	0.1	0.05	0.15
		Number of scalings in x-direction	2	0	5
		Number of scalings in y-direction	2	0	5
Noise		Number of 'noisy' slices	4	1	10
	Uniform	Maximal amplitude	1500	500	2500
	Zero-mean Gaussian	Maximal standard deviation	100	500	1000
	Salt & Pepper	Maximal density	0.2	0.05	0.5

Tab. 1: Parameters for data augmentation. This table shows all parameters specifiable by the user, as well as their default, minimal and maximal values.

4. CONCLUSIONS

In this contribution, a macro-module in MeVisLab for the preparation of training data for a deep neural network to segment the urinary bladder in CT images has been presented. For generating the ground truth to train a deep neural network, PET data was automatically segmented using a thresholding algorithm. In addition, the dataset was enlarged by applying data augmentation, specifically by the application of affine transformations (rotation and scaling) and by the addition of noise (uniformly distributed, zero-mean Gaussian and salt-and-pepper noise). Agreement between the binary masks generated from PET image data and CT image data was overall good, with slices showing a large area surface of the urinary bladder yielding accurate results. It can be observed that the shape, size and position within the image of the urinary bladder is highly varying between datasets, which might pose a difficulty for automatic segmentation approaches. With a maximal amount of 2,132,500 training datasets generated, the proposed data augmentation algorithm has the capacity to greatly enlarge the given original database.

Beside these results, there are some areas left for an upcoming future work, in particular the evaluation of the generated data with an actual deep neural network. When too much training data is produced from a small original database using data augmentation, there is the risk that using them as training data will not increase the accuracy of a neural network. Furthermore, the augmented datasets might get distorted too much, consequently losing their anatomical meaningfulness. Moreover, computational effectiveness has to be taken into account. The larger the training dataset, the more computationally expensive and time consuming the training of a neural network will get, while its performance might not increase significantly.

ACKNOWLEDGEMENT

The work received funding from BioTechMed-Graz in Austria ("Hardware accelerated intelligent medical imaging"), the 6th Call of the Initial Funding Program from the Research & Technology House (F&T-Haus) at the Graz University of Technology (PI: Dr. Dr. habil. Jan Egger). The corresponding Macro module and Python source code is freely available under (June 2017): https://github.com/cgsaxner/DataPrep_UBsegmentation

REFERENCES

[1] Zhou, K. et al. "Deep Learning for Medical Image Analysis," Elsevier, pp. 1-458 (2017).
[2] Egger, J. et al. "GBM Volumetry using the 3D Slicer Medical Image Computing Platform," Sci Rep., Nature Publishing Group (NPG), 3:1364 (2013).
[3] Bauer, M. et al. "Boundary estimation of fiber bundles derived from diffusion tensor images," International journal of computer assisted radiology and surgery 6 (1), 1-11 (2011).
[4] Bauer, M. et al. "A fast and robust graph-based approach for boundary estimation of fiber bundles relying on fractional anisotropy maps," 20th International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, pp. 4016-4019 (2010).

[5] Greiner, K. et al. “Segmentation of Aortic Aneurysms in CTA Images with the Statistic Approach of the Active Appearance Models,” Proceedings of Bildverarbeitung für die Medizin (BVM), Berlin, Germany, Springer Press, 51-55 (2008).
[6] Egger, J. et al. “Pituitary Adenoma Segmentation,” In: Proceedings of International Biosignal Processing Conference, Charité, Berlin, Germany (2010).
[7] Lu, J. et al. “Detection and visualization of endoleaks in CT data for monitoring of thoracic and abdominal aortic aneurysm stents,” Proc. of SPIE Vol 6918, 69181F-1 (2016).
[8] Zukic, D. et al. “Robust Detection and Segmentation for Diagnosis of Vertebral Diseases using Routine MR Images,” Computer Graphics Forum, Volume 33, Issue 6, Pages 190–204 (2014).
[9] Schwarzenberg, R. et al. “Cube-Cut: Vertebral Body Segmentation in MRI-Data through Cubic-Shaped Divergences,” In: PLoS One (2014).
[10] Zukic, D. et al. “Segmentation of Vertebral Bodies in MR Images,” Vision, Modeling, and Visualization (VMV), The Eurographics Association, pp. 135-142, (2012).
[11] Egger, J. “Image-guided therapy system for interstitial gynecologic brachytherapy in a multimodality operating suite,” SpringerPlus, 2:395 (2013).
[12] Pham, D. L. et al. “Current methods in medical image segmentation,” Annual review of biomedical engineering 2(1), 315-337 (2000).
[13] Shi, F. et al, “Automatic segmentation of bladder in CT images,” Journal of Zhejiang University-Science A 10: 239-246 (2009).
[14] Ronneberger, O. et al. “U-net: Convolutional networks for biomedical image segmentation,” International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, LNCS, Vol.9351: 234-241 (2015).
[15] Wernick, M. et al. “Emission tomography: the fundamentals of PET and SPECT,” Academic Press, pp. 1-596 (2004).
[16] Foster, B. et al. “A review on segmentation of positron emission tomography images,” Computers in biology and medicine 50:76-96 (2014).
[17] Wong, S. C. et al. “Understanding data augmentation for classification: when to warp?,” Digital Image Computing: Techniques and Applications (DICTA), IEEE, pp. 1-6 (2016).
[18] Egger, J. et al. “Integration of the OpenIGTlink network protocol for image guided therapy with the medical platform MeVisLab,” The international Journal of medical Robotics and Computer assisted Surgery, 8(3):282-390 (2012).
[19] Egger, J. et al. “Interactive reconstructions of cranial 3D implants under MeVisLab as an alternative to commercial planning software,” PLoS ONE 12(3): e0172694 (2017).
[20] Egger, J. et al. “HTC Vive MeVisLab integration via OpenVR for medical applications,” PLoS ONE 12(3): e0173972 (2017).
[21] Egger, J. et al. “Interactive-cut: Real-time feedback segmentation for translational research,” Computerized Medical Imaging and Graphics 38 (4), 285-295 (2014).
[22] Egger, J. et al. “PCG-Cut: Graph Driven Segmentation of the Prostate Central Gland,” PLOS ONE 8 (10), e76645 (2013).
[23] Egger, J. et al. “Preoperative Measurement of Aneurysms and Stenosis and Stent-Simulation for Endovascular Treatment,” IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Washington (D.C.), USA, pp. 392-395, IEEE Press (2007).
[24] Egger, J. “Refinement-Cut: User-Guided Segmentation Algorithm for Translational Science,” Sci Rep 4:5164 (2014).
[25] Egger, J. et al. “Simulation of bifurcated stent grafts to treat abdominal aortic aneurysms (AAA),” Proceedings of SPIE Medical Imaging Conference, Vol. 6509, pp. 65091N(1-6), San Diego, USA (2007).
[26] Armato, S. et al. “The Reference Image Database to Evaluate Response to Therapy in Lung Cancer (RIDER) Project: A Resource for the Development of Change-Analysis Software,” Clinical Pharmacology & Therapeutics, Wiley Online Library, 448-456 (2008).
[27] Gsaxner, C. & Egger, J. “Preparation of Training Data for automatic Urinary Bladder Segmentation using Deep Learning,” Technical Report, Graz University of Technology, May 2017, pp. 1-35 (2017)
[28] Gsaxner, C. & Egger, J. “MeVisLab Macro Module for Preparation of Training Data for Urinary Bladder Segmentation,” GitHub (2017).