Hierarchic Anatomical Structure Segmentation Guided by Spatial Correlations (AnatSeg-Gspac): VISCERAL Anatomy3

Oscar Alfonso Jiménez del Toro
oscar.jimenez@hevs.ch

Yashin Dicente Cid
yashin.dicente@hevs.ch

Adrien Depeursinge
adrien.depeursinge@hevs.ch

Henning Müller
henningmueller@hevs.ch

University of Applied Sciences Western Switzerland
University and University Hospitals of Geneva, Switzerland

Abstract

Medical image analysis techniques require an initial localization and segmentation of the corresponding anatomical structures. As part of the VISCERAL Anatomy segmentation benchmarks, a hierarchical multi–atlas multi–structure segmentation approach guided by anatomical correlations is proposed (AnatSeg-Gspac). The method defines a global alignment of the images and refines locally the anatomical regions of interest for the smaller structures. In this paper, the method is evaluated in the VISCERAL Anatomy3 benchmark in twenty anatomical structures in both contrast–enhanced and non–enhanced computed tomography (CT) scans. AnatSeg-Gspac obtained the lowest average Hausdorff distance in 19 out of the 40 possible structure scores in the test set CT scans.

1 Introduction

Medical image analysis and computer–aided diagnosis initially require an accurate location and segmentation of the anatomical structures present. The time expensive task of manually annotating the current large amounts of medical image data daily produced restricts the implementation of further analysis by computer algorithms [Doi05]. Different approaches have been proposed to automatically detect multiple or single anatomical structures within the patient images [LSL+10, CRK+13]. The VISual Concept Extraction challenge in RAdioLogy project (VISCERAL[1]) organizes public

Copyright © by the paper's authors. Copying permitted only for private and academic purposes.

In: O. Goksel (ed.): Proceedings of the VISCERAL Anatomy Grand Challenge
at the 2015 IEEE International Symposium on Biomedical Imaging (ISBI), New York, NY, Apr 16^{th}, 2015
published at http://ceur-ws.org

[1] http://www.visceral.eu/, as of 1 April 2015

benchmarks to test multiple segmentation approaches on the same available medical dataset for an objective evaluation of the algorithms [JdTGM$^+$14]. The VISCERAL data set has been manually annotated by radiologists and includes real medical images obtained from clinical routine in hospitals. The benchmarks are set up in a cloud environment platform designed to host large amounts of medical data with equal computing instances for the participating research groups [LMMH13].

A hierarchic Anatomical structure Segmentation Guided by spatial correlations (AnatSeg-Gspac)[JdTM13, JdTGM$^+$14, JdTM14b] has been previously proposed and tested in the first two VISCERAL Anatomy benchmarks. This approach requires no interaction from the user and generates a robust segmentation for multiple anatomical structures with short re–training phase for new scan parameters or additional structures [JdTM14a]. The evaluation and results of AnatSeg-Gspac in the VISCERAL Anatomy3 benchmark are presented in the following sections.

2 Materials and Methods

2.1 Dataset

For the VISCERAL Anatomy3 benchmark 20 CT contrast–enhanced of the trunk (CTce) and 20 CT whole body unenhanced (CTwb) with their manual annotations (up to 20 anatomical structures), were provided to the participants for training. For the implementation of AnatSeg-Gspac in this benchmark a subset of volumes (7) with all or the majority of manual annotations were selected per modality as atlases. Further information on the VISCERAL data set can be found in [JdTGM$^+$14].

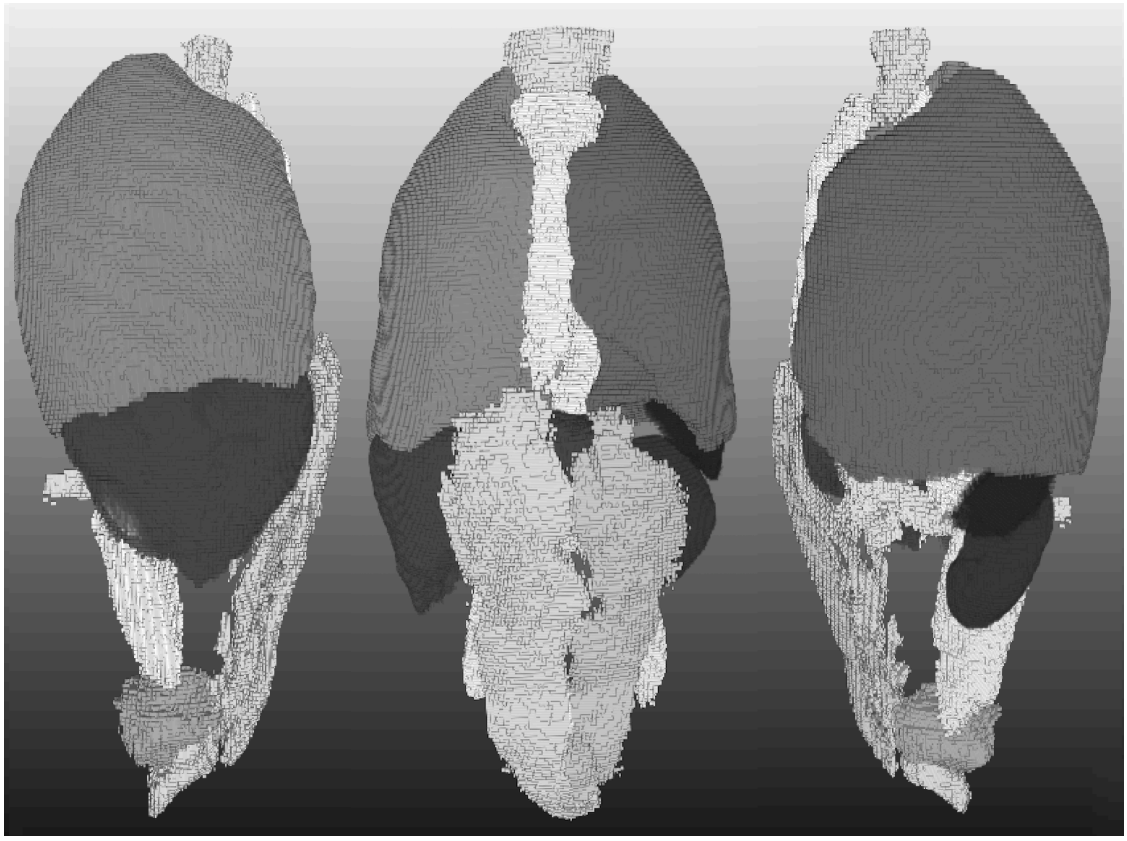

Figure 1: 3D rendering sample output using AnatSeg-Gspac.

2.2 AnatSeg-Gspac

The proposed method performs a hierarchic multi-atlas multi-structure segmentation defining anatomical regions of interest in their spatial domain. The bigger and high contrast anatomical structures are used as reference for smaller structures with low contrast, which are consequently harder to segment. The registration pipeline has been optimized to reduce the amount of registrations needed for the smaller structures obtaining also a robust localization. In Figure 1, a sample segmentation output for one unused training volume including all the anatomical structures evaluated in the VISCERAL benchmarks is shown. Further information on the AnatSeg-Gspac method can be found in the previously referenced papers.

3 Evaluation

For the Anatomy3 benchmark the test set included 10 CTce volumes and 10 CTwb scans. Twenty different evaluation metrics are provided to the participants about their algorithm performance for each anatomical structure. The evaluation phase is performed in the Azure cloud by the organizers with no intervention from the participants.

Table 1: Average Hausdorff distance results in trunk CTce test set of the VISCERAL Anatomy3 benchmark (Anatomy3 Leaderboard, http://www.visceral.eu/Leaderboard/, as of 1 April 2015). Competitive scores were obtained for kidneys and lungs compared to other organ–specific methods.

Average Hausdorff Distance																				
CTce	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	muscle body of left rectus abdominis	muscle body of right rectus abdominis	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland
	29663	**29662**	**86**	**58**	**1326**	**1302**	**237**	**40358**	**40357**	**29193**	**7578**	**170**	**32249**	**32248**	**187**	**2473**	**480**	**1247**	**30325**	**30324**
Dicente et al.					0.050	0.052														
Jia et al.	0.171	0.131	0.385	0.203	0.069	0.078														
JiménezdelToro et al.	0.172	0.243	2.005	0.514	0.071	0.065	1.879	**6.068**	**6.600**	**4.504**	**3.337**	**3.804**	**0.742**	0.757	**3.603**	**0.899**	**1.094**	**0.223**	**3.115**	**2.660**
Wang et al.	0.137	0.072	0.573	0.174	0.076	0.070	0.485						0.934	0.734						

Table 2: Average Hausdorff distance results in CT whole body (CTwb) test set of the VISCERAL Anatomy3 benchmark (Anatomy3 Leaderboard, http://www.visceral.eu/Leaderboard/, as of 1 April 2015). The method AnatSeg-Gspac generates robust segmentations for the big structures like the lungs (best benchmark scores highlighted). Moreover, it also shows overall better results particularly for small structures like thyroid, gallbladder and both adrenal glands.

Average Hausdorff Distance																				
CTwb	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	muscle body of left rectus abdominis	muscle body of right rectus abdominis	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland
	29663	**29662**	**86**	**58**	**1326**	**1302**	**237**	**40358**	**40357**	**29193**	**7578**	**170**	**32249**	**32248**	**187**	**2473**	**480**	**1247**	**30325**	**30324**
Dicente et al.					0.050	0.046														
Jia et al.			0.360	0.239	0.101	0.094														
JiménezdelToro et al.	1.209	1.307	1.974	0.780	**0.043**	**0.038**	1.457	3.550	4.032	**1.953**	**1.466**	5.521	0.595	0.775	**5.938**	0.938	1.193	0.103	**2.672**	**3.445**
Kahl et al.	0.147	0.229	0.534	0.299	0.045	0.043	1.057	1.614	1.922	2.424	2.163	4.478	0.443	0.550	9.617	0.542	0.798	0.083	3.298	7.046
Wang et al.	0.758	2.751	0.248	0.191	0.356	0.096	2.155						0.705	0.812						

The proposed method obtained the lowest average Hausdorff distance of the Anatomy3 benchmark in 12/20 structures in CTce (Table 1) and 7/20 structures in CTwb (Table 2). The DICE coefficient scores are also presented for all the methods submitted in the benchmark (Table 3 and Table 4).

Table 3: DICE coefficient results in the test set trunk CT contrast–enhanced (CTce) of the VISCERAL Anatomy3 benchmark (Anatomy3 Leaderboard, http://www.visceral.eu/Leaderboard/, as of 1 April 2015). The proposed AnatSeg-Gspac(Jiménez del Toro et al. in light grey) was the only submitted method that segmented all available anatomical structures in both CT modalities (enhanced and unenhanced).

DICE Coefficient																				
CTce	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	muscle body of left rectus abdominis	muscle body of right rectus abdominis	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland
	29663	**29662**	**86**	**58**	**1326**	**1302**	**237**	**40358**	**40357**	**29193**	**7578**	**170**	**32249**	**32248**	**187**	**2473**	**480**	**1247**	**30325**	**30324**
Dicente et al.					0.974	0.973														
Jia et al.	0.910	0.922	0.869	0.933	0.966	0.966														
JiménezdelToro et al.	0.910	0.889	0.730	0.887	0.959	0.963	0.679	**0.474**	**0.453**	**0.523**	**0.410**	**0.423**	0.794	0.799	**0.484**	**0.721**	**0.762**	**0.855**	**0.331**	**0.342**
Wang et al.	0.945	0.959	0.909	0.949	0.972	0.971	0.856						0.830	0.845						

Table 4: DICE coefficient results in the test set unenhanced CT of the whole body (CTwb) of the VISCERAL Anatomy3 benchmark (Anatomy3 Leaderboard, http://www.visceral.eu/Leaderboard/, as of 1 April 2015). Highlighted are the best DICE overlap scores obtained in the benchmark by AnatSeg-Gspac in 7 clinically relevant anatomical structures: left and right lungs, thyroid, pancreas, gallbladder, left and right adrenal gland.

DICE Coefficient																				
CTwb	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	muscle body of left rectus abdominis	muscle body of right rectus abdominis	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland
	29663	**29662**	**86**	**58**	**1326**	**1302**	**237**	**40358**	**40357**	**29193**	**7578**	**170**	**32249**	**32248**	**187**	**2473**	**480**	**1247**	**30325**	**30324**
Dicente et al.					**0.972**	0.974														
Jia et al.			0.874	0.923	0.952	0.957														
JiménezdelToro et al.	0.784	0.790	0.703	0.866	**0.972**	**0.975**	0.698	0.551	0.519	0.718	**0.549**	**0.408**	0.806	0.787	**0.276**	0.761	0.753	0.920	**0.373**	**0.355**
Kahl et al.	0.934	0.915	0.870	0.921	**0.972**	**0.975**	0.763	0.746	0.679	0.775	0.424	0.383	0.861	0.847	0.190	0.847	0.830	0.931	0.282	0.220
Wang et al.	0.896	0.796	0.910	0.936	0.961	0.970	0.713						0.828	0.817						

4 Discussion and Conclusions

The proposed method showed robustness in the segmentation of multiple structures from two different imaging modalities using a small training set. Both the distance and overlap scores in this and the previous Anatomy benchmarks show AnatSeg-Gspac outperforms other algorithms in some of the smaller anatomical structures (e.g. both adrenal glands, gallbladder). It can also obtain the best overlap for bigger and high contrasted structures like the lungs.

A limitation of the method is the computation cost, mainly for the B–spline non-rigid registrations. Although the number of registrations and the size of registered regions are reduced using anatomical correlations, the execution time is around 13 hours for a complete CT volume. A faster

code implementation and better selection of the relevant atlases may reduce the number of needed registrations and thus the execution time of the method.

The method can be extended to the other imaging modalities and include more anatomical structures with short re–training phases. This is particularly important for its application with new or different scanners contained in large not annotated data sets. Further clinical image analyses, that may require the location of additional structures, might also benefit from this feature or include the output locations of the method as an initialization step.

5 Acknowledgments

This work was supported by the EU/FP7 through VISCERAL (318068).

References

[CRK+13] Antonio Criminisi, Duncan Robertson, Ender Konukoglu, Jamie Shotton, Sayan Pathak, Steve White, and Khan Siddiqui. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical Image Analysis*, 17(8):1293–1303, 2013.

[Doi05] K Doi. Current status and future potential of computer–aided diagnosis in medical imaging. *British Journal of Radiology*, 78:3–19, 2005.

[JdTGM+14] Oscar Alfonso Jiménez del Toro, Orcun Goksel, Bjoern Menze, Henning Müller, Georg Langs, Marc-André Weber, Ivan Eggel, Katharina Gruenberg, Markus Holzer, Georgios Kotsios-Kontokotsios, Markus Krenn, Roger Schaer, Abdel Aziz Taha, Marianne Winterstein, and Allan Hanbury. VISCERAL – VISual Concept Extraction challenge in RAdioLogy: ISBI 2014 challenge organization. In Orcun Goksel, editor, *Proceedings of the VISCERAL Challenge at ISBI*, number 1194 in CEUR Workshop Proceedings, pages 6–15, Beijing, China, May 2014.

[JdTM13] Oscar Alfonso Jiménez del Toro and Henning Müller. Multi–structure atlas–based segmentation using anatomical regions of interest. In *MICCAI workshop on Medical Computer Vision*, Lecture Notes in Computer Science. Springer, 2013.

[JdTM14a] Oscar Jiménez del Toro and Henning Müller. Hierarchic multi–atlas based segmentation for anatomical structures: Evaluation in the visceral anatomy benchmarks. In *MICCAI workshop on Medical Computer Vision*, Lecture Notes in Computer Science. Springer, 2014.

[JdTM14b] Oscar Alfonso Jiménez del Toro and Henning Müller. Hierarchical multi–structure segmentation guided by anatomical correlations. In Orcun Goksel, editor, *Proceedings of the VISCERAL Challenge at ISBI*, CEUR Workshop Proceedings, pages 32–36, Beijing, China, May 2014.

[LMMH13] Georg Langs, Henning Müller, Bjoern H. Menze, and Allan Hanbury. Visceral: Towards large data in medical imaging – challenges and directions. *Lecture Notes in Computer Science*, 7723:92–98, 2013.

[LSL+10] Marius George Linguraru, Jesse K. Sandberg, Zhixi Li, Furhawn Shah, and Ronald M. Summers. Automated segmentation and quantification of liver and spleen from CT images using normalized probabilistic atlases and enhancement estimation. *Medical Physics*, 37(2):771–783, 2010.