

VISCERAL —

VISual Concept Extraction challenge in RAdioLogY: ISBI 2014 Challenge Organization

Oscar Alfonso Jiménez del Toro¹, Orcun Goksel², Bjoern Menze², Henning Müller¹
Georg Langs³, Marc-André Weber⁴, Ivan Eggel¹, Katharina Gruenberg⁴,
Markus Holzer³, András Jakab³, Georgios Kontokotsios⁵, Markus Krenn³,
Tomàs Salas Fernandez⁶, Roger Schaer¹, Abdel Aziz Taha⁵,
Marianne Winterstein⁴, Allan Hanbury⁵

University of Applied Sciences Western Switzerland, Switzerland¹
Swiss Federal Institute of Technology (ETH) Zürich, Switzerland²

Medical University of Vienna, Austria³

University of Heidelberg, Germany⁴

Vienna University of Technology, Austria⁵

Catalan Agency for Health Information, Assessment and Quality, Spain⁶

Abstract

The VISual Concept Extraction challenge in RAdioLogY (VISCERAL) project has been developed as a cloud-based infrastructure for the evaluation of medical image data in large data sets. As part of this project, the ISBI 2014 (International Symposium for Biomedical Imaging) challenge was organized using the VISCERAL data set and shared cloud-framework. Two tasks were selected to exploit and compare multiple state-of-the-art solutions designed for big data medical image analysis. Segmentation and landmark localization results from the submitted algorithms were compared to manually annotated ground truth in the VISCERAL data set. This paper presents an overview of the challenge setup and data set used as well as the evaluation metrics from the various results submitted to the challenge. The participants presented their algorithms during an organized session at ISBI 2014. There were lively discussions in which the importance of comparing approaches on tasks sharing a common data set was highlighted.

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

In: O. Goksel (ed.): Proceedings of the VISCERAL Organ Segmentation and Landmark Detection Benchmark at the 2014 IEEE International Symposium on Biomedical Imaging (ISBI), Beijing, China, May 1st, 2014
published at <http://ceur-ws.org>

1 Introduction

Computational approaches that can be scaled to large amounts of medical data are needed to tackle the ever-growing data resources obtained daily from the hospitals [Doi05]. Handling this enormous amount of medical data during clinical routine by health professionals has complexity and scaling limitations. It is also very time-consuming, and hence requires unsupervised and automatic methods to perform the necessary data analysis and processing for data interpretation. There are already many algorithms and techniques for big data analysis, however, most research groups do not have access to large-scale annotated medical data to develop such approaches for medical images. Distributing these big data sets (on the order of terabytes) requires efficient and scalable storing and computing capabilities. Evaluation campaigns and benchmarks can objectively compare multiple state-of-the-art algorithms to determine the optimal solution for a certain clinical task [HMLM14, GSdHKCDF⁺13].

The Visual Concept Extraction Challenge in Radiology (VISCERAL) project was developed as a cloud-based infrastructure for the evaluation of medical image analysis techniques on large data sets [LMMH13]. The shared cloud environment in which the VISCERAL project takes place allows access and processing of these data without having to duplicate the data or move it to participants' side. Since the data are stored centrally, and not distributed outside the cloud environment, the legal and ethical requirements of such data sets can also be satisfied, so also confidential data sets can be benchmarked in this way as only a small training data set can be accessed by participants [EILI⁺10]. The cloud infrastructure is provided and funded by the VISCERAL project. The participants are provided with computationally powerful virtual machines that can be accessed remotely in the shared cloud infrastructure while working on the training data and tuning their algorithms. Participant access is withdrawn during the evaluation phase and only the organizers access the machines. The algorithms are brought to the data to perform automated processing and data mining. The evaluation of the performance of these methods can therefore be done with real clinical imaging data and the outcomes can be reused to improve the methods.

The whole body 3D medical imaging data including manual labels that is provided by VISCERAL includes a small subset with ground truth annotated by experienced radiologists. Through evaluation campaigns, challenges, benchmarks and competitions, tasks of general interest can be selected to compare the algorithms on a large scale. This manually annotated gold corpus can be used to identify high quality methods that can also be combined to create a much larger "reasonably annotated" data set, satisfactory but perhaps not as reliable as manual annotation. Using fusion techniques this silver corpus will be created with the agreement between the segmentations of the algorithms on a large-scale data set. This maximizes the gain of manual annotation and also identifies strong differences between participating systems on the annotated organs.

2 ISBI Challenge Framework

The registration procedure for the ISBI challenge was that of the VISCERAL series benchmark that includes several campaigns. The participants filled information details and uploaded a signed participation agreement form, which corresponds to ethics requests for usage of the data. Since the VISCERAL data set is stored on the Azure Cloud, each participant then received access to an Azure virtual cloud-computing instance. There were 5 operating systems available to choose from including Windows 2012, Windows 2008, Ubuntu Server 14.04 LTS, openSUSE 13.1 and CentOS 6.5. All cloud-computing instances have an 8-core CPU with 16 GB RAM to provide the same computing capabilities to different solutions proposed. The participant gets administrator rights on their virtual machine (VM) and can access remotely to deploy their algorithms and add any supporting library/applications to their VM. The VISCERAL training data set can then be accessed

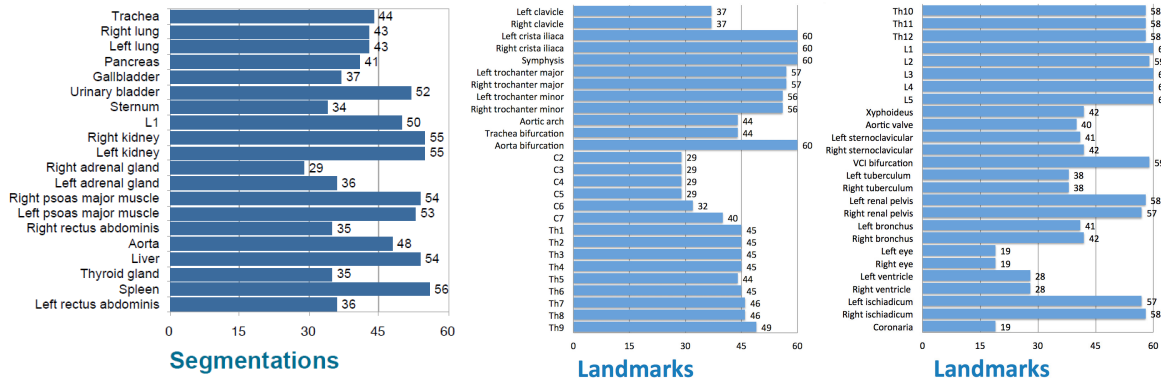


Figure 1: The ISBI training set.

and downloaded securely within the VMs through secured URL links.

2.1 Data Set

The medical images contained in the VISCERAL data set have been acquired during daily clinical routine work. Data sets of children (<18 years) were not included based on the recommendations of the ethical committee. In the provided data sets multiple organs are visible and depicted in a resolution sufficient to reliably detect an organ and delineate its borders. This is to enforce that a large number of organs and structures can be segmented in one data set. The data set consists of computed tomography (CT) scans and magnetic resonance (MR) imaging with and without contrast enhancement to evaluate the participants algorithms on several modalities, contrasts and MR sequence directions, making sure that algorithms are not optimized for one specific machine or protocol.

The available training set from VISCERAL Anatomy² benchmark was used by the participants of the ISBI VISCERAL challenge. The contents of this dataset are elaborated below.

2.1.1 CT Scans

There are 15 unenhanced whole-body CT volumes acquired from patients with bone marrow neoplasms, such as multiple myeloma, to detect osteolysis. The field-of-view spans from and including the head to the knee (see Fig. 2, A). The in-plane resolution ranges between 0.977/0.977 to 1.405/1.405 mm, and the in-between plane resolution is 3 mm or higher.

15 contrast-enhanced CT scans of the trunk that have been acquired in patients with malignant lymphoma are also included. They have a large field-of-view from the corpus mandibulae to the lower part of the pelvis (see Fig. 2, B). They have an in-plane resolution of between 0.604/0.604 and 0.793/0.793 mm, and an in-between plane resolution of at least 3 mm or higher.

2.1.2 MR Scans

15 whole-body MR scans in two sequences (30 in total) are also part of the training set. They were acquired in patients with multiple myeloma to detect focal and/or diffuse bone marrow infiltration. Both a coronal T1-weighted and fat-suppressed T2-weighted or STIR (short tau inversion recovery) sequence of the whole body are available for each of the 15 patients. The field-of-view starts and includes the head and ends at the feet (see Fig. 2, C). The in-plane resolution is 1.250/1.250 mm, and the in-between plane resolution is 5 mm.

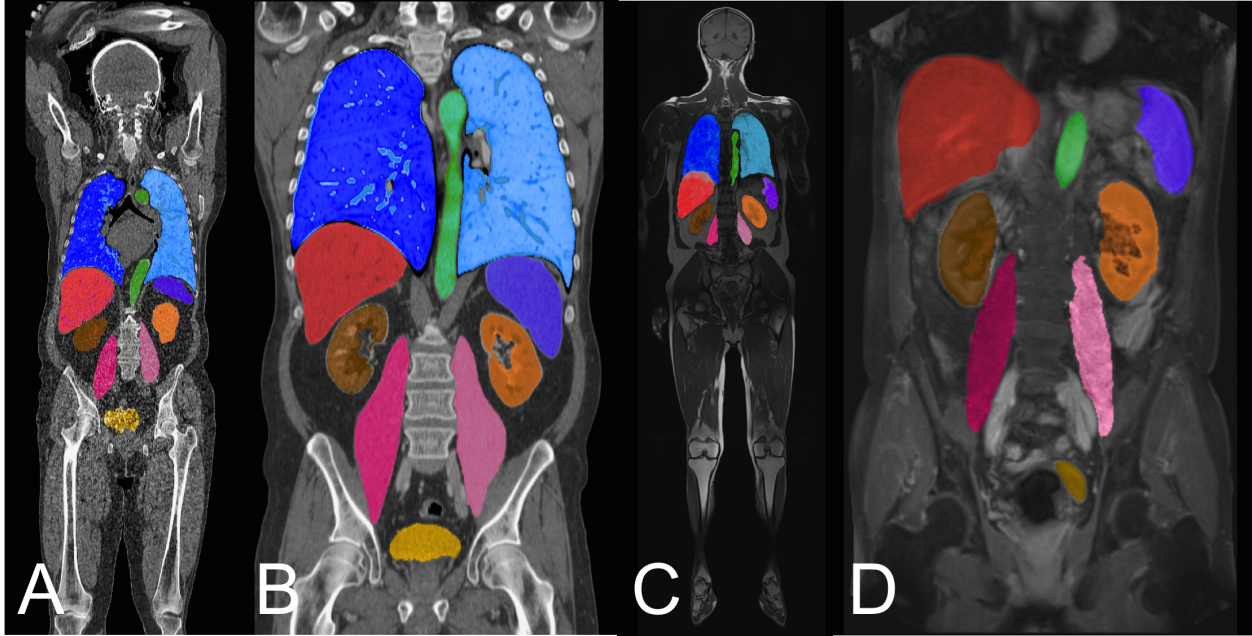


Figure 2: Sample data set volumes. A) Whole-body unenhanced CT; B) contrast-enhanced CT of the trunk; C) whole-body unenhanced MR; D) contrast-enhanced MR of the abdomen

To improve the segmentation of smaller organs (such as the adrenal glands), 15 T1 contrast-enhanced fat saturated MR scans of the abdomen are also included. They were acquired in oncological patients with likely metastases within the abdomen. The field-of-view starts at the top of the diaphragm and extends to the lower part of the pelvis (see Fig. 2, D). They have an in plane resolution of between 0.840/0.804 to 1.302/1.302 mm, and an in-between plane resolution of 3 mm.

2.1.3 Annotated Structures and Landmarks

There are in total 60 manually annotated volumes in this ISBI challenge training set. The available data contains segmentation and landmarks of several different anatomical structures in different imaging modalities, e.g. CT and MRI.

The two categories of annotations and results are:

- **Region segmentations:** These regions correspond to anatomical structures (e.g. right lung), or sub-parts in volume data. The 20 anatomical structures that make up the training set are: trachea, left/right lungs, sternum, vertebra L1, left/right kidneys, left/right adrenal glands, left/right psoas major muscles, left/right rectus abdominis, thyroid gland, liver, spleen, gall-bladder, pancreas, urinary bladder and aorta. Not all structures are visible or within the field-of-view in the images, therefore leading to varying numbers of annotations per structure (see Fig. 1 for a detailed break-down).
- **Landmarks:** Anatomical landmarks are the locations of selected anatomical structures that should be identifiable in the different image sequences available in the data set. There can be up to 53 anatomical landmarks (see Fig. 1) located in the data set volumes: left/right clavicles, left/right crista iliaca, symphysis, left/right trochanter major, left/right trochanter minor, aortic arch, trachea bifurcation, aorta bifurcation, vertebrae C2-C7, Th1-Th12, L1-L5, xyphoideus, aortic valve, left/right sternoclavicular, VCI bifurcation, left/right tuberculum,

left/right renal pelvises, left/right bronchus, left/right eyes, left/right ventricles, left/right ischiadicum and coronaria.

In total the 60 training set volumes containing 890 manually segmented anatomical structures and 2420 manually located anatomical landmarks make up the training set. Some of the anatomical structures in the volumes were not segmented if the annotators considered there was insufficient tissue contrast to perform the segmentation or to locate the landmark. Other structures are missing or not included in the training set because of anatomical variations (e.g. missing kidney) or radiologic pathological signs (e.g. aortic aneurysm). Landmarks are easy and quick to annotate whereas precise organ segmentation is time-consuming even when using automatic tools.

2.1.4 Test Set

The test set contains 20 manually annotated volumes. Each modality (whole-body CT, thorax and abdomen contrast-enhanced CT, whole-body MR and abdomen contrast enhanced MR) is represented by 5 volumes. The anatomical structures and landmarks contained in the selected volumes were used to evaluate the participants' algorithms.

2.2 ISBI VISCERAL Challenge Submission

The participants can select the structures and modalities in which they choose to participate. The outputs are therefore evaluated per structure and per modality. The evaluation of the ISBI challenge has been organized differently than the general VISCERAL evaluation framework to allow for the evaluation results to complete in the given relatively short time-frame. For this challenge, the test set volumes were made available in the cloud some weeks ahead of the challenge. The participants themselves computed the annotations (segmentations and/or landmark locations) in their VMs and stored them on their VM storage. The files could then be submitted within their VM through an uploading script provided to the participants. The script stored their output files in a corresponding cloud container created for the challenge individual for each participant. A list containing the available ground truth segmentations of the test set filtered duplicates or output files with incorrect file names. It also ensured all files were coherent with the participant ID list from the organizers.

2.3 Evaluation Software

To evaluate the output segmentations and landmark locations against the ground truth, the VISCERAL evaluation tool was used. This software was also included in the VM assigned to each participant. This evaluation tool has different evaluation metrics implemented such as (1) distance-based metrics, (2) spatial overlap metrics and (3) probabilistic and information theoretic metrics. The most suitable subset of the metrics was used in the analysis of the results and all metrics were made available to the participants. For the output segmentations of the ISBI challenge the following evaluation metrics were selected:

- DICE coefficient [ZWB⁺04]
- Adjusted Rand Index [VPYM11]
- Interclass Correlation [GJC01]
- Average distance [KCAB09]

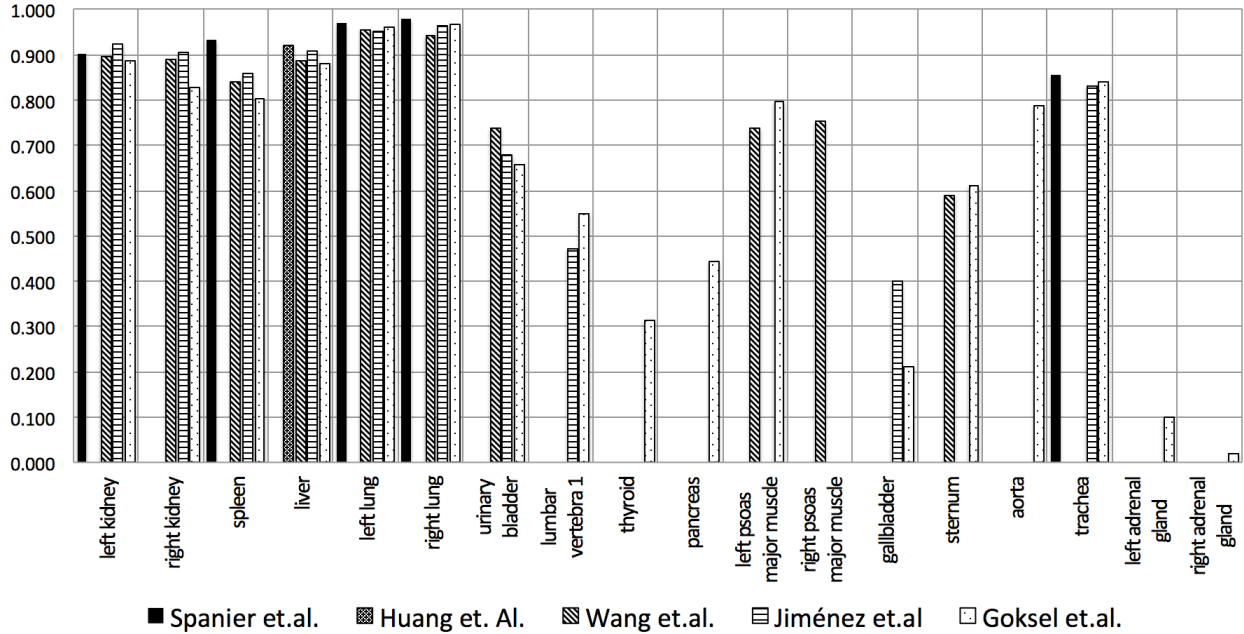


Figure 3: Anatomical structure segmentation task: DICE coefficient results. Contrast-enhanced CT scans of the Thorax and Abdomen.

Only one label is considered per image. The voxel value can be either zero (background) or one for the voxels containing the segmentation. A threshold is set at 0.5 to create binary images in case the output label has a fuzzy membership or a probability map.

For the landmark localization evaluation the same VISCERAL tool measures the landmark-specific average error (Euclidean distance) error between all the results and the manually located landmarks. The percentage of detected landmarks per volume (i.e. landmarks detected / landmarks in the volume) is also computed.

2.4 Participation

The ISBI training and test set volumes were made available through the Azure cloud framework for all the registered participants of the VISCERAL Anatomy² benchmark. In total 18 groups got access to the challenge training set and the 60 training volumes of the data set. The research groups that submitted working virtual machines had a chance to present their methods and results at the "VISCERAL Organ Segmentation and Landmark Detection Challenge" at the 2014 IEEE International Symposium on Biomedical Imaging (ISBI).

A single-blind review process was applied to the initial abstract submissions. The accepted abstracts were then invited to submit a short paper presenting their methods and results in the challenge. There were 5 high-quality submissions accepted and included in these proceedings.

Spanier et al. [SJ14] submitted segmentations for five organs in CT contrast-enhanced volumes. Their multi-step algorithm combines thresholding and region growing techniques to segment each organ individually. It starts with the location of a region of interest and identification of the largest axial cross-section slices of the selected structure. It then improves the initial segmentation with morphological operators and a final step performs 3D region growing.

Huang et al. [HLJ14] proposed a coarse-to-fine liver segmentation using prior models for the shape, profile appearance and contextual information of the liver. An AdaBoost voxel-based classifier creates a liver probability map that is refined in the last step with free-form deformation with

Table 1: Anatomical structure segmentation task: DICE coefficient results table. Contrast-enhanced and unenhanced CT scans submissions.

Participant	Modality	Region	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland	Average	
			29663	29662	86	58	1326	1302	237	29193	7578	170	32249	32248	187	2473	480	1247	30325	30324		
Measure			DICE coefficient																			
Spanier et.al.	Ctce	ThAb	0.902	0.934		0.970	0.979											0.856			0.928	
	Ctce	ThAb	0.922																		0.922	
Huang et.al.	CT	wb	0.911																		0.911	
	CT	wb	0.729	0.777	0.887	0.904	0.971	0.972	0.806				0.722	0.764		0.712						0.824
Wang et.al.	Ctce	ThAb	0.896	0.890	0.842	0.887	0.956	0.942	0.738				0.737	0.752		0.590						0.823
	CT	wb	0.678	0.649	0.677	0.823	0.969	0.967	0.616	0.440						0.271			0.855			0.694
Jiménez et.al.	Ctce	ThAb	0.923	0.905	0.859	0.908	0.952	0.963	0.680	0.472					0.400			0.830				0.789
	Ctce	ThAb	0.885	0.827	0.803	0.882	0.960	0.966	0.657	0.548	0.315	0.442	0.797		0.212	0.612	0.787	0.839	0.099	0.019		0.627
Goksel et.al.	Ctce	ThAb	0.756	0.679	0.684	0.798	0.955	0.965	0.636	0.333	0.439	0.466	0.773	0.780	0.078	0.630	0.724	0.837	0.282	0.133		0.608
	CT	wb																				

a gradient appearance model.

Wang et al. [WS14] segmented 10 anatomical structures in CT contrast-enhanced and unenhanced scans. Their multi-organ segmentation pipeline performs in a top-down approach by a model-based level-set segmentation of the ventral cavity. After dividing the cavity in thoracic and abdominopelvic cavity, the major structures are segmented and their location information is passed to the lower-level structures.

Jiménez del Toro et al. [JdTM14] segmented structures in CT and contrast-enhanced CT scans with a hierarchical multi-atlas approach. Based on the spatial anatomical correlations between the organs, the bigger and high-contrasted organs are first segmented. These then define the initial volume transformations for the smaller structures with less defined boundaries.

Goksel et al. [GGS14] submitted segmentations for both CT and MR anatomical structure segmentation. They also submitted results for the landmark localization task. For the segmentations they use a multi-atlas based technique that implements Markov Random Fields to guide the registrations. A multi-atlas template-based approach fuses the different estimations to detect the landmarks.

3 Results

There were approximately 500 structure segmentations and 211 landmark locations submitted to the VISCERAL ISBI challenge. Four participants submitted results for the segmentation tasks in multiple organs using whole-body CT or contrast-enhanced scans with results presented in Table 1 and Fig. 3. There was one participant that contributed segmentations on both the whole-body MR scans and the contrast-enhanced MR abdomen volumes with results presented in Table 3. Only one participant submitted landmark localization results, with Table 4 showing their evaluation results.

4 Conclusions and Future Work

The VISCERAL project has the evaluation of algorithms on large data sets as its main objective. The proposed VISCERAL infrastructure allows evaluations with private or restricted data, such as electronic health records, without the participants access to the test data by using a fully cloud-based approach. This infrastructure also avoid moving data, which is potentially hard for very large data sets. The algorithms are brought to the data and not the data to the algorithms.

Table 2: Anatomical structure segmentation task: Average Distance results table. Contrast-enhanced and unenhanced CT scans submissions.

Participant	Modality	Region	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	lumbar Vertebra 1	thyroid	pancreas	left psoas major muscle	right psoas major muscle	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland	Average
			29663	29662	86	58	1326	1302	237	29193	7578	170	32249	32248	187	2473	480	1247	30325	30324	
Measure			Average distance																		
Spanier et.al.	Ctce	ThAb	0.23	0.14		0.08	0.04												0.30		0.16
	Ctce	ThAb				0.29															0.29
Huang et.al.	CT	wb				0.31															0.31
	CT	wb	3.63	1.21	0.45	0.46	0.07	0.06	0.78			1.77	1.15		1.24						1.08
Wang et.al.	Ctce	ThAb	0.27	0.28	0.87	0.65	0.15	0.20	1.59			1.75	1.58		4.14						1.15
	CT	wb	3.06	3.31	2.64	1.76	0.05	0.06	2.08	4.18					25.02				0.30		4.24
Jiménez et.al.	Ctce	ThAb	0.15	0.37	0.60	0.37	0.11	0.07	1.69	4.22					4.34				0.58		1.25
	Ctce	ThAb	0.39	1.48	1.08	0.73	0.38	0.92	1.93	2.46	5.54	6.83	0.80		13.50	1.13	0.78	0.71	9.63	15.07	3.73
Goksel et.al.	MRT1	wb	21.23	16.73	0.83	0.90	159.38	0.28	1.44	1.21	1.01		0.94		8.18	5.44	0.45	0.53	12.09	3.64	14.64
	CT	wb	1.42	3.66	2.15	1.60	0.88	0.10	2.32	6.99	3.57	3.36	1.33	1.33	14.55	1.33	1.02	0.87	6.01	6.62	3.28
	MRT1cefs	Ab	0.34	3.76	0.82	0.83			5.91	9.14		5.91	1.63		11.54						4.43
	MRT1cefs	Ab																			

Table 3: Evaluation metrics for the MR scan submissions.

Participant	Modality	Region	left kidney	right kidney	spleen	liver	left lung	right lung	urinary bladder	lumbar Vertebra 1	thyroid	pancreas	major muscle left psoas	major muscle right psoas	gallbladder	sternum	aorta	trachea	left adrenal gland	right adrenal gland	Average
			29663	29662	86	58	1326	1302	237	29193	7578	170	32249	32248	187	2473	480	1247	30325	30324	
Measure			DICE coefficient																		
Goksel et.al.	MRT1	wb	0.548	0.589	0.646	0.817	0.486	0.909	0.577	0.623	0.488		0.765		0.044	0.359	0.783	0.747	0.144	0.268	0.550
	MRT1cefs	Ab	0.888	0.732	0.785	0.861			0.334	0.084		0.356	0.654		0.000						
Measure			Average distance																		
Goksel et.al.	MRT1	wb	21.23	16.73	0.83	0.90	159.4	0.28	1.44	1.21	1.01		0.94		8.18	5.44	0.45	0.53	12.09	3.64	14.64
	MRT1cefs	Ab	0.34	3.76	0.82	0.83			5.91	9.14		5.91	1.63		11.54						

Both gold corpus and silver corpus will be available as a resource to the community. The ISBI test set volumes and annotations are now available and are part of the VISCERAL Anatomy² benchmark training set.

So far, both past VISCERAL anatomy benchmarks have addressed organ segmentation and landmark localization tasks. There are two more benchmarks under development in the VISCERAL project, a retrieval benchmark and a detection benchmark. The retrieval benchmark will be the retrieval of similar cases based on both visual information and radiology reports. The detection benchmark will focus in the detection of lesions in MR and CT images.

In the future, the automation of the evaluation process is intended to reduce the need for intervention from the organizers to a minimum and to provide faster evaluation feedback to the participants. The participants will then be able to submit their algorithms through the cloud virtual machines and obtain the calculated metrics directly from the system. Such a system could then store the results from all the algorithms submitted and perform an objective comparison with state-of-the-art algorithms. Through the involvement of the research community, the VISCERAL framework could produce novel tools for the clinical work flow that has substantial impact on diagnosis quality

Table 4: Landmarks results.

Participant	Modality	Region		aorta_bifurcation	aortic_arch	clavicle_left	clavicle_right	crista_iliaca_left	crista_iliaca_right	symphysis	trachea_bifurcation	trochanter_major_left	trochanter_major_right	trochanter_minor_left	trochanter_minor_right	Average
Measure			Euclidean distance													
Goksel et al.	CT	wb	average	19.05	17.68	9.27	5.69	7.70	6.12	8.01	3.99	34.37	36.18	5.16	4.06	13.11
	Ctce	ThAb	average	36.22	16.18	16.26	32.35	13.93	10.38	15.59	3.35	37.84	38.31	11.22	12.64	20.36
	MRT1	wb	average	252.49	43.67	13.05	23.31	23.29	19.21	122.45	61.20	29.57	44.40	18.51	62.40	59.46
	MRT1cefs	Ab	average	61.28				88.92	57.65	50.86		30.49	59.81	28.54	34.84	51.55

and treatment success. Having all tools and algorithms in the same cloud environment can also help us to combine tools and approaches with very little additional effort, which expectedly yields better results.

5 Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 318068 VISCERAL. We would also like to thank Microsoft research for their financial and information support in using the Azure cloud for the benchmark.

References

- [Doi05] K Doi. Current status and future potential of computer-aided diagnosis in medical imaging. *British Journal of Radiology*, 78:3–19, 2005.
- [EILI⁺10] Bernice Elger, Jimison Iavindrasana, Luigi Lo Iacono, Henning Müller, Nicolas Roudit, Paul Summers, and Jessica Wright. Strategies for health data exchange for secondary, cross-institutional clinical research. *Computer Methods and Programs in Biomedicine*, 99(3):230–251, September 2010.
- [GGS14] Orcun Goksel, Tobias Gass, and Gabor Szekely. Segmentation and landmark localization based on multiple atlases. In *Proceedings of the VISCERAL Challenge at ISBI*, pages 37–43, 2014.
- [GJC01] Guido Gerig, Matthieu Jomier, and Miranda Chakos. A new validation tool for assessing and improving 3D object segmentation. In Wiro J. Niessen and Max A. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001*, volume 2208 of *Lecture Notes in Computer Science*, pages 516–523. Springer Berlin Heidelberg, 2001.
- [GSdHKCDF⁺13] Alba García Seco de Herrera, Jayashree Kalpathy-Cramer, Dina Demner Fushman, Sameer Antani, and Henning Müller. Overview of the ImageCLEF 2013 medical tasks. In *Working Notes of CLEF 2013 (Cross Language Evaluation Forum)*, September 2013.

- [HLJ14] Cheng Huang, Xuhui Li, and Fucang Jia. Automatic liver segmentation using multiple prior knowledge models and free-form deformation. In *Proceedings of the VISCERAL Challenge at ISBI*, pages 22–24, 2014.
- [HMLM14] Allan Hanbury, Henning Müller, Georg Langs, and Bjoern H. Menze. Cloud-based evaluation framework for big data. In Alex Galis and Anastasius Gavras, editors, *Future Internet Assembly (FIA) book 2013*, Springer LNCS, pages 104–114. Springer Berlin Heidelberg, 2014.
- [JdTM14] Oscar Alfonso Jiménez del Toro and Henning Müller. Hierarchical multi-structure segmentation guided by anatomical correlations. In *Proceedings of the VISCERAL Challenge at ISBI*, pages 32–36, 2014.
- [KCAB09] Hassan Khotanlou, Olivier Colliot, Jamal Atif, and Isabelle Bloch. 3d brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. *Fuzzy Sets and Systems*, 160(10):1457–1473, 2009. Special Issue: Fuzzy Sets in Interdisciplinary Perception and Intelligence.
- [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.
- [SJ14] Assaf B. Spanier and Leo Joskowicz. Rule-based ventral cavity multi-organ automatic segmentation in ct scans. In *Proceedings of the VISCERAL Challenge at ISBI*, pages 16–21, 2014.
- [VPYM11] Nagesh Vadaparathi, Suresh Varma Penumatsa, Srinivas Yarramalle, and P. S. R. Murthy. Segmentation of Brain MR Images based on Finite Skew Gaussian Mixture Model with Fuzzy C-Means Clustering and EM Algorithm. *International Journal of Computer Applications*, 28:18–26, 2011.
- [WS14] Chunliang Wang and Örjan Smedby. Automatic multi-organ segmentation using fast model based level set method and hierarchical shape priors. In *Proceedings of the VISCERAL Challenge at ISBI*, pages 25–31, 2014.
- [ZWB⁺04] Kelly H. Zou, Simon K. Warfield, Aditya Bharatha, Clare M.C. Tempany, Michael R. Kaus, Steven J. Haker, William M. Wells III, Ferenc A. Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index¹: scientific reports. *Academic Radiology*, 11(2):178 – 189, 2004.