Automatic Intervertebral Disc Localization and Segmentation in 3D MR Images Based on Regression Forests and Active Contours

Martin Urschler[1,2(✉)], Kerstin Hammernik[1], Thomas Ebner[1], and Darko Štern[1,2]

[1] Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria
{urschler,hammernik,ebner,stern}@icg.tugraz.at
[2] Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria

Abstract. We introduce a fully automatic localization and segmentation pipeline for three-dimensional (3D) intervertebral discs (IVDs), consisting of a regression-based prediction of vertebral bodies and IVD positions as well as a 3D geodesic active contour segmentation delineating the IVDs. The approach was evaluated on the data set of the challenge in conjunction with the 3rd MICCAI Workshop & Challenge on Computational Methods and Clinical Applications for Spine Imaging - MICCAI–CSI2015, that consists of 15 magnetic resonance images of the lumbar spine with given ground truth segmentations. Based on a localization accuracy of 3.9 ± 1.6 mm, we achieve segmentation results in terms of the Dice similarity coefficient of 89.1 ± 2.9 % averaged over the whole data set.

1 Introduction

Due to reduced physical activity and working conditions of modern office jobs low back pain (LBP) resembles a very important health problem in the developed countries. It is a leading cause of disability affecting work performance and well-being. Clinical studies indicate correlation between LBP and intervertebral disc (IVD) degeneration [1]. A widely used imaging modality for examining IVD degeneration is magnetic resonance (MR) imaging (MRI), since it provides excellent soft tissue contrast without the need for ionizing radiation. In the diagnosis of MR images of the lumbar spine automatic IVD identification and extraction of quantitative measures of IVD geometry and appearance is of high interest. However, development of such automatic methods for accurate and objective IVD localization and segmentation is challenging and still represents an important research area [2–5]. To objectively compare and analyze IVD segmentation approaches, efforts like the challenge *Automatic Intervertebral Disc (IVD) Localization and Segmentation from 3D T2 MRI Data* in conjunction with the 3rd MICCAI Workshop & Challenge on Computational Methods and Clinical Applications for Spine Imaging - MICCAI–CSI2015 are necessary and crucial for potential future application in clinical practice.

T. Vrtovec et al. (Eds.): CSI 2015, LNCS 9402, pp. 130–140, 2016.
DOI: 10.1007/978-3-319-41827-8_13

In this work we present our novel automatic IVD center localization and segmentation approach. It consists of a machine learning step to predict center locations of both vertebral bodies and IVDs as well as an image processing pipeline to segment IVDs given the located spine landmarks. Segmentation is based on geodesic active contours formulated as convex energy functional. We evaluate our method on the 15 MRI data sets from the MICCAI–CSI2015 challenge and report localization errors and Dice similarity coefficients (DSC) with respect to the provided IVD ground truth segmentations.

2 Methods

Our proposed IVD segmentation algorithm is built upon a powerful machine learning based landmark localization step using regression forests [6,7] together with a high-level Markov random field (MRF) model of the global configuration of the relative landmark positions [8]. After landmark prediction, we attach a three-step image processing pipeline for segmentation. First, we roughly segment vertebral bodies based solely on image gradient information, followed by a merging of pairs of adjacent vertebral bodies to single objects to initialize IVD segmentation. Finally, we formulate the IVD segmentation problem as a convex geodesic active contour optimization task based on edges resembling geometrical similarity to the shape of IVDs. Enabled by the robustness of previous localization, this latter segmentation step requires no a priori information on appearance but only a very rough shape prior. The main algorithm steps are shown in Fig. 1.

2.1 Preprocessing

Since the input MRI data sets contain slightly different absolute intensity values and some of the volumes show intensity inhomogeneities, we first perform an automatic Retinex theory based inhomogeneity correction step similar to the

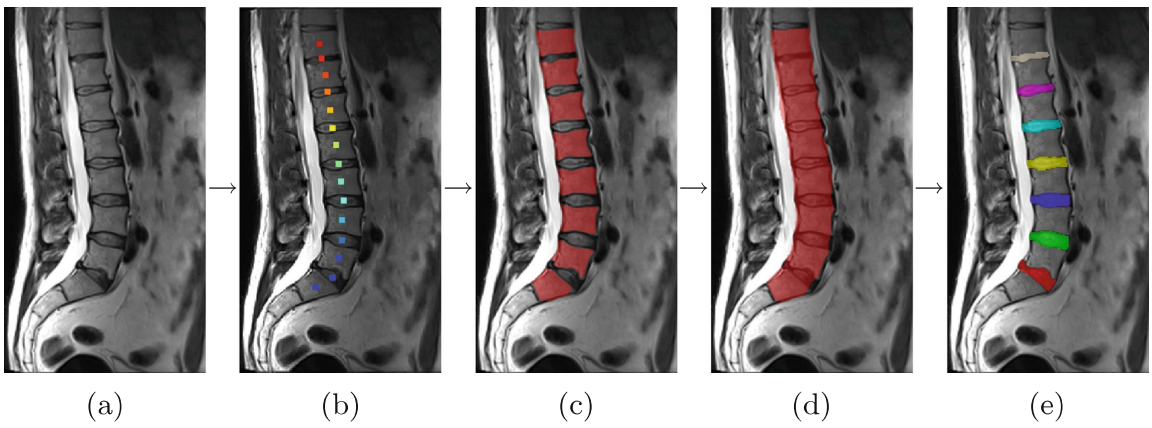


Fig. 1. Algorithm overview. The 3D MR input images (a) are processed by spinal landmark localization (b), vertebral body segmentation (c), merging of vertebral bodies to initialize IVD segmentation (d) and final voxel-wise IVD labeling (e).

work of Ma et al. [9] to remove the smooth bias field due to this imaging artifact. To derive more similar intensity distributions among the MRI data, we additionally apply histogram matching ignoring background voxels as determined by thresholding [10].

2.2 Spinal Landmark Localization Using Regression Forests

Our approach for localization of vertebrae and IVD centers is based on local appearance information of vertebrae and IVDs in spine images. Inspired by the localization method proposed by Donner et al. [8], we use a Hough forest (HF) [11] to generate probability maps $p_l(\boldsymbol{x})$ for each landmark l being at location $\boldsymbol{x}$. Due to the similar, repeating appearance of spinal landmarks, a global geometric model implemented as an MRF with a dynamic programming based solver is used to select the most probable configuration of landmark positions from the set of candidate positions, thus correctly labeling vertebrae and IVDs.

Candidate Position Generation - Hough Forest. For each landmark l we train a HF from manually annotated locations. Each HF consists of K trees with a maximum depth D. Training starts at the root node using all voxels inside a certain radius r around the landmark position. Our node split functions, passing a voxel either to the left or right child node, are based on Haar-like features. The feature response is calculated as the difference between the mean intensity of two cuboids, whose positions are defined relatively to the voxel position. At each node split, T random thresholds, a pool of F random feature boxes with a maximum distance $d_{f,max}$ from the voxel position and a maximum size $s_{f,max}$ are generated. The respective combination of feature and threshold, which maximizes an Information Gain criterion, is used as the node split function and stored in the split node. In a leaf node we calculate a histogram of the voting vectors

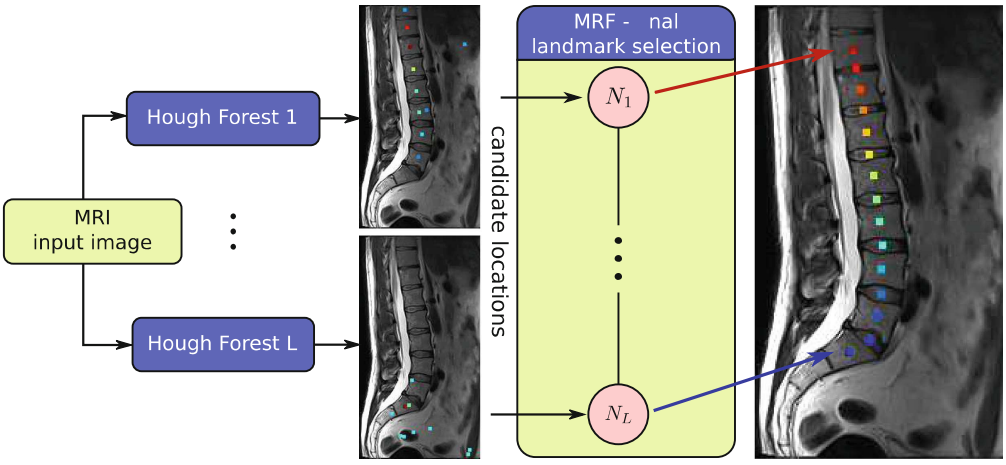


Fig. 2. Overview of the localization pipeline. The colors of the candidate locations indicate the strength of the HF response, where red corresponds to a high value. (Color figure online)

of all voxels arriving at the node and store one single voting vector derived from the maximum of this histogram.

During testing, a random subset of all voxels in an image is pushed through the previously trained trees. All resulting voting vectors from the leaf nodes are summed up in an accumulator data structure, which can be interpreted as probabilities $p_l(\boldsymbol{x})$ of the landmark l being located at position $\boldsymbol{x}$. Using non-maxima suppression we select for each landmark the N strongest local maxima of $p_l(\boldsymbol{x})$ as candidate locations (see Fig. 2).

Candidate Selection - Geometric Model. To select the best candidate for each spinal landmark according to a global geometric model, an MRF is used based on the landmark candidate positions from the previous regression forest step. An MRF is an undirected graph with L nodes and E edges, where each node N_l in the graph corresponds to one landmark. The edges e, connecting the nodes in the graph, are modeling geometric relationships between the landmarks. In our MRF graph only the nodes of neighboring spinal landmarks are connected, thus each vertebral body is connected to its adjacent IVDs and vice versa. Solving the MRF means to select for each node a candidate such that the function

$$\Phi = \prod_{l=1}^{L} \Phi(N_l) \prod_{e=1}^{E} \Phi_n(N_{e_1}, N_{e_2}), \tag{1}$$

based on a product of all node potentials $\Phi(N_l)$ and edge potentials $\Phi_n(N_{e_1}, N_{e_2})$, is maximized. The node potentials $\Phi(N_l)$ are set to the accumulator value $p_l(\boldsymbol{x})$ obtained in the previous step. The edge potentials between nodes N_{e_1} and N_{e_2} and candidate locations c_1 and c_2 are defined as

$$\Phi_n(N_{e_1}, N_{e_2}) = \begin{cases} \Phi_e^x(N_{e_1}, N_{e_2}) \cdot \Phi_e^d(N_{e_1}, N_{e_2}), & \text{if } c_{1,z} \geq c_{2,z} \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

where $\Phi_e^x(N_{e_1}, N_{e_2})$ is a term punishing large deviations in x-coordinates between two landmarks and $\Phi_e^d(N_{e_1}, N_{e_2})$ a term based on the Euclidean distance between candidate locations. The term based on the x-coordinates is defined as

$$\Phi_e^x(N_{e_1}, N_{e_2}) = e^{-\frac{1}{2}\left(\frac{c_{1,x} - c_{2,x}}{\sigma_x}\right)^2}, \tag{3}$$

where σ_x allows to control the allowed deviation in x, which is set empirically to 2 mm. The model for the Euclidean distance is based on statistics from the training data and defined as

$$\Phi_e^d(N_{e_1}, N_{e_2}) = e^{-\frac{1}{2}\left(\frac{\|\boldsymbol{c}_1 - \boldsymbol{c}_2\| - \mu_e}{\sigma_e}\right)^2}, \tag{4}$$

where μ_e and σ_e are mean and variance of the Euclidean distance, respectively.

2.3 Intervertebral Disc Segmentation Using Geodesic Active Contours

From the accurate regression forest based landmark localization approach we achieve a very good initialization for segmenting the IVDs using an image processing pipeline that extracts vertebral bodies and restricts the region where the IVDs are expected. This segmentation pipeline, which is shown in Fig. 3, makes heavy use of a convex 3D geodesic active contour segmentation approach based on total variation (TV), which is described next.

Total Variation Segmentation. Our segmentation framework is based on minimizing following continuous non-smooth energy functional $E_{seg}(u)$ which has previously also been used by Reinbacher et al. [12] and Hammerni et al. [13]. It is a minimal surface segmentation approach formulated as

$$\begin{aligned} \min_u E_{seg}(u) = \min_u \int_\Omega g(x)|\nabla u(x)|\,\mathrm{d}x + \lambda \int_\Omega u(x)\cdot w(x)\,\mathrm{d}x, \\ \text{s.t.} \qquad u \in C_{box} = \{u : u(x) \in [0,1],\, \forall\, x \in \Omega\} \end{aligned} \tag{5}$$

where Ω denotes the image domain and $u \in C^1 : \Omega \mapsto \mathbb{R}$ is smooth. The first term denotes the g-weighted TV semi-norm which is a reformulation of the geodesic active contour energy [14]. The edge function $g(x)$ is defined as

$$g(x) = e^{-\alpha\|\nabla I(x)\|^\beta},\ \alpha, \beta > 0, \tag{6}$$

where $\nabla I(x)$ is the gradient of the input image. The second term in (5) is the data term with w describing a weighting map. The values in w have to be chosen negative if u should be foreground and positive if u should be background. If values in w are set to zero, the pure weighted TV energy is minimized seeking for a minimal surface segmentation. The regularization parameter λ defines the trade-off between our data term and the weighted TV semi-norm. The stated convex problem in (5) can be solved for its global optimum efficiently using the primal-dual algorithm [15]. As the segmentation u is continuous the final segmentation is achieved by thresholding u.

Vertebral Body Segmentation. The first step of our segmentation pipeline is to segment the eight vertebral bodies (T11, T12, L1-L5, S1) individually using (5) (see Fig. 1b). The weighting map w is constructed based on the localization results for vertebral bodies and IVDs. We span a cylinder whose normal vector points from the center of the vertebral body to the center of the IVD located above. This cylinder defines the foreground seed ($w = -\infty$) region. A larger, but again cylindrical region around the foreground seed is set to zero in the weighting map w such that the solution u is influenced by the surrounding image edges. Values farther away in the weighting map are set to background $w = \infty$. The edge function $g(x)$ is defined according to (6), where the image $I(x)$ is the input from the preprocessing step with an additionally applied edge-preserving denoising

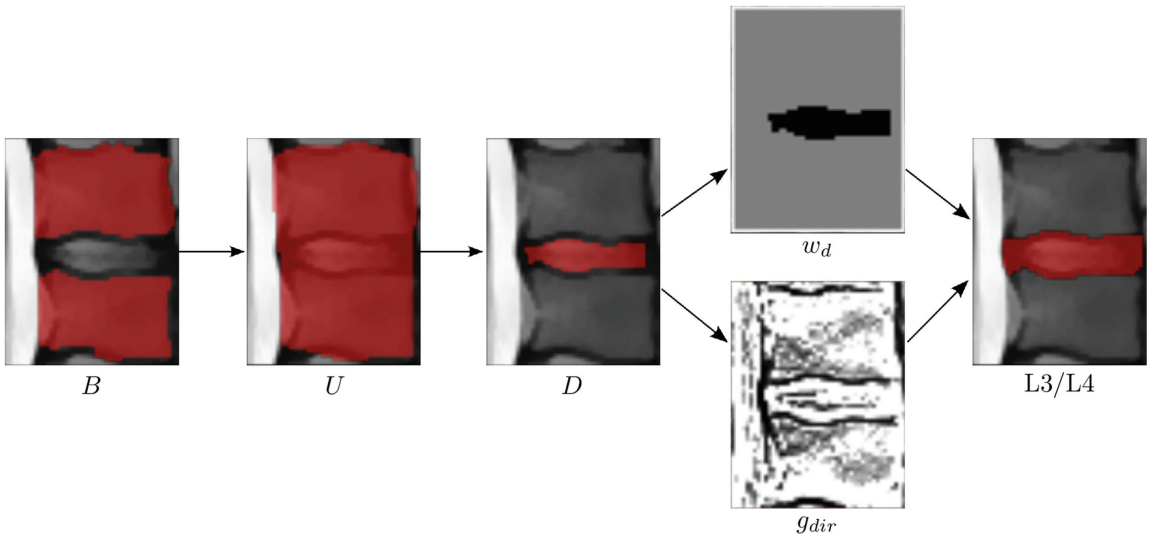


Fig. 3. Exemplary illustration of the segmentation pipeline. The vertebral body segmentation B is used to generate a fused, one-connected structure U containing the IVD. After morphological operations, the difference $D = U - B$ initializes a weighted TV segmentation step using w_d derived from D (black are foreground seeds, white background seeds, grey indicates $w_d = 0$) and g_{dir} as edge indicator function. The final segmentation result shows IVD between L3 and L4.

using the TV-L1 model as explained by Chambolle and Pock [15]. Solving the TV segmentation model in (5) gives us the central part of the vertebral bodies, which is sufficient to constrain the later IVD segmentation.

Fusion of Vertebral Bodies Using the Star Prior Constraint. The second step of our segmentation pipeline is to connect the segmentations of pairs of vertebral bodies to obtain a rough spine segmentation (see Fig. 1c), thus initializing the IVD region. This motivates the use of the star prior constraint introduced by Veksler [16] and extended to multiple star centers by Gulshan et al. [17]. The intention of the star prior is that any ray n sent out from a specified star center is directed in opposite direction of ∇u. This enforces one-connected and star convex objects. The star prior constraint is modeled in terms of a convex set C_{star}

$$C_{star} = \{u : \langle \nabla u(x), n(x) \rangle \leq 0, \forall\, x \in \Omega\}. \tag{7}$$

This constraint on the segmentation u can be handled easily in our segmentation model (5) by minimizing the energy $E_{seg}(u)$ such that $u \in C_{box} \cap C_{star}$. For more details we refer the interested reader to the work of Hammernik [18].

In our segmentation pipeline we use the star prior to connect the gap between adjacent vertebrae as depicted in Fig. 3. This gap region gives a strong hint of the expected IVD location. For each IVD, we use the two neighboring vertebral bodies as foreground seeds setting their weight $w = -\infty$. The border region of the image domain Ω are background seeds ($w = \infty$) and other regions are set to zero. The edge function is derived from the binary segmentation result $B(x)$

of the two vertebral bodies using the edge inversion function from the original geodesic active contour approach [19], i.e.

$$g(x) = \frac{1}{1 + \gamma \|\nabla B(x)\|^2}, \quad \gamma > 0. \tag{8}$$

We use two star centers defined by the localization results of the respective vertebral bodies. The ray direction $n(x)$ is then defined from the star center which is closest to the position x. After solving the weighted TV model in (5) with the star prior constraint (7), we subtract the vertebral body segmentation $B(x)$ from the segmentation result $U(x)$ and perform morphological opening and erosion operations to achieve one final connected component $D(x)$ that is guaranteed to be located inside the IVD region (see D in Fig. 3).

Intervertebral Disc Segmentation. For the final IVD segmentation we again use the weighted TV model for each IVD as described in (5). The foreground seeds are simply defined by the result of the previous step, defining the region where ($w_d = -\infty$). The background is defined by the borders of the image domain Ω, all other regions are set to $w_d = 0$. For the edge information we apply a slightly modified variant of the edge function g in (6) to incorporate a small amount of a priori information on the shape of IVDs. As IVDs are prone to have double edges, we only consider those edges which are aligned with rays that are sent out from a specified disc. This disc is defined by a radius r and a normal vector d which points from the located IVD to the next vertebral body. The rays $n(x)$ are calculated for every point x relative to the closest point on the disc. We define the modified edge function g_{dir} considering directed edges as follows:

$$g_{dir} = e^{-\alpha \|\xi \nabla I(x)\|^\beta}, \ \alpha, \beta > 0 \tag{9}$$

with

$$\xi = \begin{cases} \langle \nabla I(x), n(x) \rangle & \text{if} \quad \langle \nabla I(x), n(x) \rangle > 0 \\ 0 & \text{if} \quad \langle \nabla I(x), n(x) \rangle \leq 0 \end{cases}. \tag{10}$$

3 Experimental Setup

The whole localization and segmentation approach was implemented in C++ and OpenMP, with the exception of the Matlab-based MRF solver. Costly image processing operations were accelerated using Nvidia CUDA to use graphical processing units as parallel numerical co-processors.

Localization. Localization results were obtained using a leave-one-out cross validation on the 15 subjects of the MICCAI–CSI2015 challenge, where centers of the vertebral bodies and IVDs were manually annotated for all subjects by a scientist well experienced in spine image analysis. Hough forests were trained and tested with a voxel size of $1.5 \times 1.5 \times 1.5$ mm. For each HF we trained $K = 64$

trees until a maximum depth $D = 15$ using all voxels in the range $r = 8$ mm around the landmarks. At each node split $F = 10$ candidate feature boxes and $T = 10$ thresholds were generated. The maximum size $s_{f,max}$ and distance $d_{f,max}$ of the feature boxes was set to 20 mm. After non-maxima supression we selected for each landmark the $N = 10$ candidate positions with the highest maxima in the accumulator volume. The MRF made use of a specific statistical model (μ_e, σ_e) computed from the 14 subjects of each cross-validation run and was solved with loopy belief propagation using the publicly available Matlab UGM[1] library.

Segmentation. To provide a realistic, generalizable setup we pooled the individual localization results of each of the 15 leave-one-out cross validation runs to form our input vertebral body and IVD landmarks for segmentation. The further processing pipeline involved a number of parameters. For all edge functions $\alpha = 20$ and $\beta = 0.55$ was selected. We chose a $\lambda = 1.25$ for the edge-preserving TV-L1 smoothing and a radius of 15 mm as well as a height of 7.5 mm to initialize the vertebral body cylinder model. Segmentation of vertebral bodies involved a $\lambda = 0.01$, while the star prior constrained TV segmentation required $\lambda = 1000$ and $\gamma = 0.125$. Finally, IVD segmentation was done using $\lambda = 0.05$ and a radius of 15 mm for the prevention of double edges. All TV segmentation steps were computed until the maximum change of two voxels of subsequent segmentations was below 0.0001 and a threshold of $u = 0.5$ was used to derive a binary result from the convex model in (5).

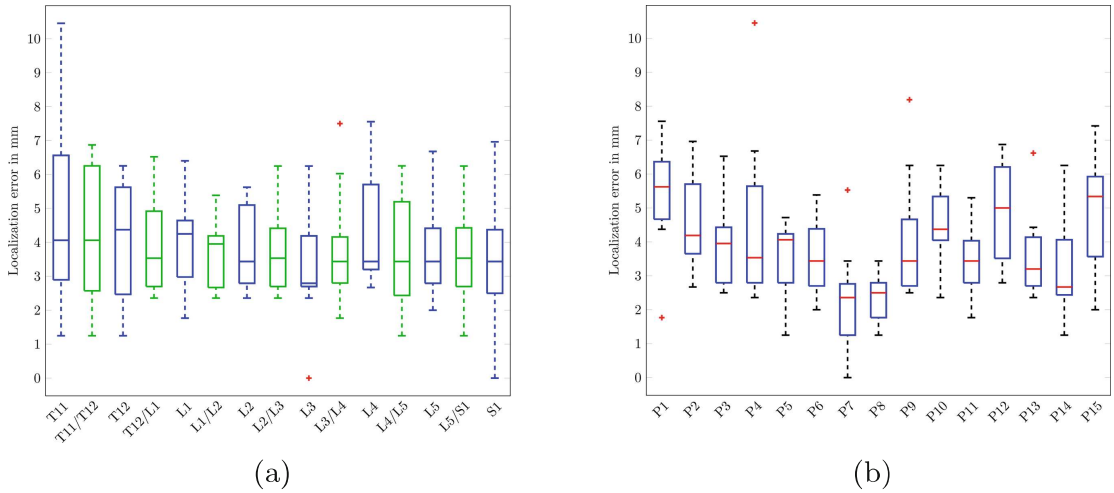
4 Results and Discussion

Localization. Quantitative results of the individual vertebra/intervertebral discs as well as for the 15 different subjects are shown in Fig. 4. We achieve an overall mean localization error ± standard deviation of 3.9 ± 1.6 mm for all 15 landmarks of the 15 subjects compared to our own manual annotation of the landmark centers. For the vertebrae we achieve an average of 4.0 ± 1.7 mm and for the intervertebral discs 3.8 ± 1.5 mm. These are promising results given the image spacing of the input data of $2 \times 1.25 \times 1.25$ mm, i.e. the mean localization error is on the order of a few pixel and standard deviations are reasonable. Our segmentation results also indicate that this localization performance is sufficient to initialize our image processing pipeline.

For the first test set we achieve a mean localization error of 3.97 ± 1.19 mm, with 2.9 % of landmarks below 2 mm, 42.9 % below 4 mm and 94.3 % below 6 mm of distance. For the on-site test set we achieve 4.37 ± 1.17 mm, with 0.0 % below 2 mm, 37.1 % below 4 mm and 91.4 % below 6 mm of distance.

Segmentation. For quantitative evaluation of the overall segmentation algorithm we used the DSC to compute the overlap of our segmentation result with the provided ground truth segmentation from the CSI challenge data set.

[1] Downloaded from http://www.cs.ubc.ca/~schmidtm/Software/UGM.html.

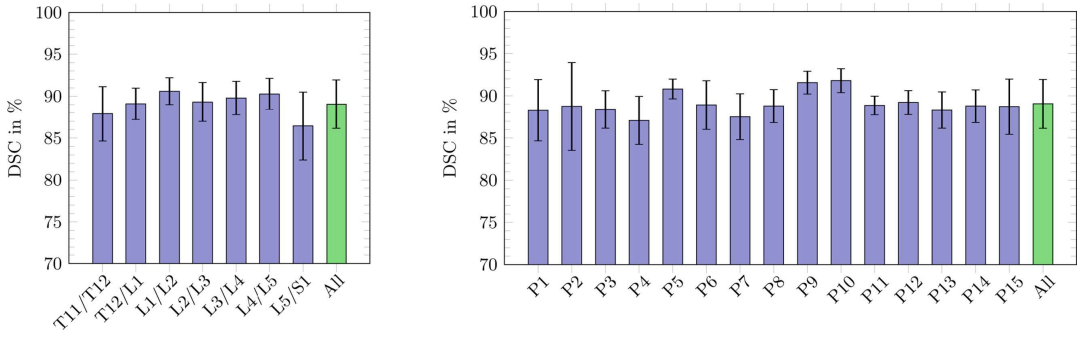


(a) (b)

Fig. 4. Box-whisker plots of the mean localization error of our proposed method. (a) Location error for vertebrae and IVDs showing individual landmarks (blue for vertebral bodies, green for IVDs). (b) Location error for individual subjects. (Color figure online)

We achieved an average DSC of $89.1 \pm 2.9\,\%$ over all IVDs from the 15 subjects. Figure 5 shows more details of the performance of our method according to individual IVDs and subjects.

For segmenting the first five test data sets we achieve a DSC of $87.4 \pm 4.8\,\%$ and a surface distance of $1.47 \pm 0.53\,\text{mm}$. Unfortuantely, during the on-site calculation of the segmentation results, in one of the five data sets a severe segmentation error occurred, leading to merged IVDs of T11/T12 and T12/L1, which prevented our method from being compared to the others.



(a) (b)

Fig. 5. Detailed quantitative Dice similarity coefficients (DSC) results of the IVD segmentation approach (overall: $89.1 \pm 2.9\,\%$). (a) DSC for individual IVDs. (b) DSC for each of the 15 subjects.

5 Conclusion

In this work, a fully automatic localization and segmentation system for intervertebral disc segmentation from MRI data was shown. It builds upon a regression forest together with a simple global geometric model as well as TV based convex active contour segmentation steps extracting vertebral bodies and IVDs in a geometrically constrained manner. Our results on the data of the MICCAI–CSI2015 challenge are located in the lower third of the nine compared approaches.

Acknowledgements. This work was supported by the province of Styria under the funding scheme "HTI:Tech_for_Med" (ABT08-22-T-7/2013-13) and by the Austrian Science Fund (FWF): P28078-N33.

References

1. Neubert, A., Fripp, J., Engstrom, C., Walker, D., Weber, M., Schwarz, R., Crozier, S.: Three-dimensional morphological and signal intensity features for detection of intervertebral disc degeneration from magnetic resonance images. J. Am. Med. Inform. Assoc. **20**(6), 1082–1090 (2013)
2. Štern, D., Likar, B., Pernuš, F., Vrtovec, T.: Automated detection of spinal centrelines, vertebral bodies and intervertebral discs in CT and MR images of lumbar spine. Phys. Med. Biol. **55**(1), 247–264 (2010)
3. Chen, C., Belavy, D., Yu, W., Chu, C., Armbrecht, G., Bansmann, M., Felsenberg, D., Zheng, G.: Localization and segmentation of 3D intervertebral discs in MR images by data driven estimation. IEEE Trans. Med. Imaging **34**(8), 1719–1729 (2015)
4. Law, M., Tay, K., Leung, A., Garvin, G., Li, S.: Intervertebral disc segmentation in MR images using anisotropic oriented flux. Med. Image Anal. **17**(1), 43–61 (2013)
5. Korez, R., Likar, B., Pernuš, F., Vrtovec, T.: Parametric modeling of the intervertebral disc space in 3D: application to CT images of the lumbar spine. Comput. Med. Imaging Graph. **38**(7), 596–605 (2014)
6. Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K.: Regression forests for efficient anatomy detection and localization in computed tomography scans. Med. Image Anal. **17**(8), 1293–1303 (2013)
7. Ebner, T., Stern, D., Donner, R., Bischof, H., Urschler, M.: Towards automatic bone age estimation from MRI: localization of 3D anatomical landmarks. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 421–428. Springer, Heidelberg (2014)
8. Donner, R., Menze, B., Bischof, H., Langs, G.: Global localization of 3D anatomical structures by pre-filtered Hough forests and discrete optimization. Med. Image Anal. **17**(8), 1304–1314 (2013)
9. Ma, W., Morel, J.M., Osher, S., Chien, A.: An L1-based variational model for Retinex theory and its application to medical images. In: Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition - CVPR 2011, pp. 153–160. IEEE (2011)
10. Nyúl, L., Udupa, J., Zhang, X.: New variants of a method of MRI scale standardization. IEEE Trans. Med. Imaging **19**(2), 143–150 (2000)

11. Gall, J., Yao, A., Razavi, N., van Gool, L., Lempitsky, V.: Hough forests for object detection, tracking, and action recognition. IEEE Trans. Pattern Anal. Mach. Intell. **33**(11), 2188–2201 (2011)
12. Reinbacher, C., Pock, T., Bauer, C., Bischof, H.: Variational segmentation of elongated volumetric structures. In: Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition - CVPR 2010, pp. 3177–3184. IEEE (2010)
13. Hammernik, K., Ebner, T., Stern, D., Urschler, M., Pock, T.: Vertebrae segmentation in 3D CT images based on a variational framework. In: Yao, J., et al. (eds.) CSI 2014. LNCVB, vol. 20, pp. 227–233. Springer, Switzerland (2015)
14. Bresson, X., Esedoḡlu, S., Vandergheynst, P., Thiran, J.P., Osher, S.: Fast global minimization of the active contour/snake model. J. Math. Imaging Vis. **28**(2), 151–167 (2007)
15. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011)
16. Veksler, O.: Star shape prior for graph-cut image segmentation. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part III. LNCS, vol. 5304, pp. 454–467. Springer, Heidelberg (2008)
17. Gulshan, V., Rother, C., Criminisi, A., Blake, A., Zisserman, A.: Geodesic star convexity for interactive image segmentation. In: Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition - CVPR 2010, pp. 3129–3136. IEEE (2010)
18. Hammernik, K.: Convex framework for 2D & 3D image segmentation using shape constraints. Master's thesis, Graz University of Technology, Austria (2015)
19. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vis. **22**(1), 61–79 (1997)