



# Intervertebral Disc Segmentation and Localization from Multi-modality MR Images with 2.5D Multi-scale Fully Convolutional Network and Geometric Constraint Post-processing

Chang Liu<sup>1,2</sup> and Liang Zhao<sup>1</sup>(✉)

<sup>1</sup> SenseTime, Beijing, China

liu.luckpp@gmail.com, zhaoliang@sensetime.com

<sup>2</sup> Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China

**Abstract.** The intervertebral discs (IVDs) segmentation and localization on medical images are important for the clinical diagnosis and research of spine diseases. In this work, we proposed a robust automatic method based on 2.5D multi-scale fully convolutional network (FCN) and geometric constraint post-processing for IVD segmentation and localization on 3D multi-modality Magnetic Resonance (MR) scans. Firstly, we designed a 2.5D multi-scale FCN. And the ensemble outputs of such three networks are used as the IVD prediction maps. The final segmentation and localization of IVDs are generated from these prediction maps with a geometric constraint post-processing method. This work ranked the first in the on-site test of MICCAI 2018 Challenge on Automatic Intervertebral Disc Localization and Segmentation from 3D Multi-modality MR Images (IVDM3Seg).

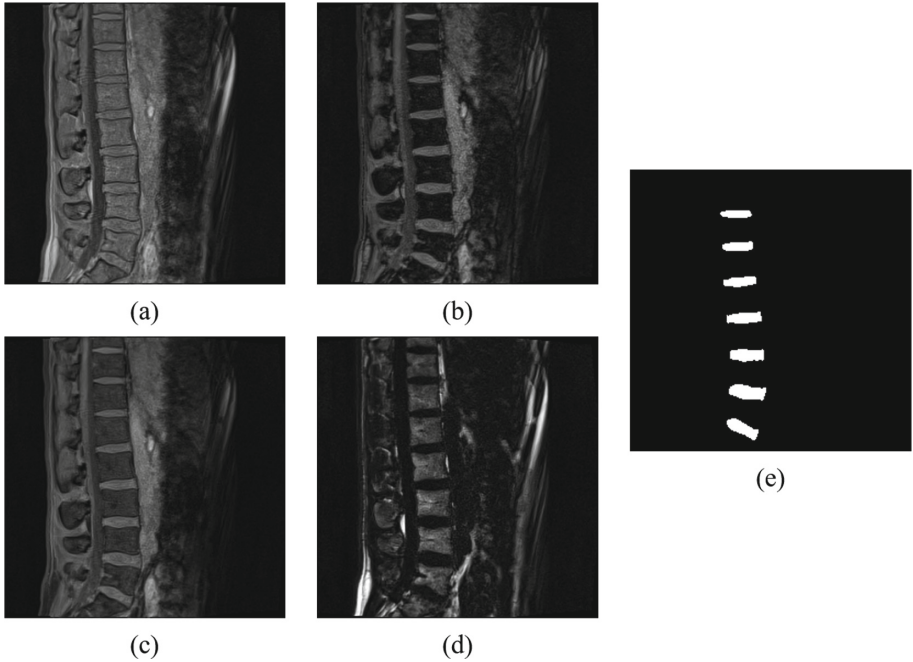
**Keywords:** Intervertebral disc · Segmentation · Localization · Fully convolutional network

## 1 Introduction

The intervertebral disc (IVD) is a cartilaginous joint that lies between adjacent vertebrae. It plays a crucial role in the shock absorption of vertebral movement [1, 2]. In modern society, back pain is becoming a common healthy problem, which causes the pain, stiffness and loss of independency of patients. According to the international studies, the point prevalence of back pain is between 12% and 35%, while the lifetime prevalence is up to 49% to 80% [3]. For this disease, degeneration of the intervertebral disc is considered as a major cause [4].

Magnetic Resonance Imaging (MRI) is a commonly used imaging technique in the diagnosis of IVD degeneration and many other diseases, which provides non-invasive assessment to human body. Compared to other medical imaging methods, such as Computed Tomography (CT) imaging, MRI could provide excellent contrast in soft

tissue without ionizing radiation. Besides, the MR scans could be obtained with different modalities, and provide more information about tissue structure. In this work, four MRI modalities (i.e. in-phase, opposed-phase, water, fat) were used for the segmentation and localization of IVDs. Figure 1 shows an example of these four modalities. It should be noticed that only the 7 IVDs between the twelfth thoracic vertebra and sacrum are delineated manually as the targets.



**Fig. 1.** An example of 3D multi-modality images provided by MICCAI 2018 Challenge on IVDM3Seg. (a) to (d) are in-phase, opposed-phase, water, and fat modality in order while (e) is the manually delineated labels for 7 IVDs between the twelfth thoracic vertebra and sacrum.

The research on IVD degeneration usually needs the segmentation of IVDs. Traditionally, the IVD labels are delineated manually. However, this job is always time-consuming and may be biased for inter- and intra-observer variabilities [5, 6]. For this matter, automatic IVD segmentation and localization methods have great significance to the study of IVD degeneration.

There are three main challenges for automatic IVD segmentation and localization on multi-modality images. Firstly, distinguishing different IVDs is difficult due to the intra-subject similarity of IVDs. Secondly, the intensity of IVD boundary resembles that of the neighborhood tissues, which makes the IVD contour fuzzy. Thirdly, how to harness the multi-modality information effectively in medical image processing remains to be explored.

## 1.1 Previous Work

There are many segmentation and localization methods proposed in previous research, which are based on traditional hand-crafted features [7–13]. Besides, some popular graph-based methods, such as graph cut [10] and statistical shape model [7], were also applied to IVD segmentation. For localization, some graphical models were proposed to take IVD geometric relationship into account [13]. With the reference to the local parts shape and neighborhood anatomical structures, the accuracy of IVD localization improved in some degree.

In recent years, machine learning has drawn extensive attention in many fields. Some classical machine learning algorithms, such as marginal space learning (MSL) [14], Adaboost [15], and sparse kernel machine [16], were also adopted to IVD segmentation and localization. And these methods have shown excellent performance.

More recently, deep learning techniques achieved great success in computer vision. Many researchers began to attempt deep learning algorithms in medical image processing. And these methods have proven effective. In the past few years, all the state-of-art methods on MICCAI IVD segmentation and localization challenge were deep learning-based [17, 18].

Multi-modality images are not only available for IVD segmentation and localization. How to utilize multi-modality information is a common issue in medical image processing, such as MRI-based brain tissue [19] and brain tumor segmentation [20]. Generally, the harness of multi-modality data could improve the performance more or less.

## 1.2 Our Contribution

We propose a 2.5D multi-scale deep learning network for segmentation and localization of IVDs on multi-modality MR scans. Our method achieved the state-of-art performance in the MICCAI 2018 Challenge on IVDMSeg.

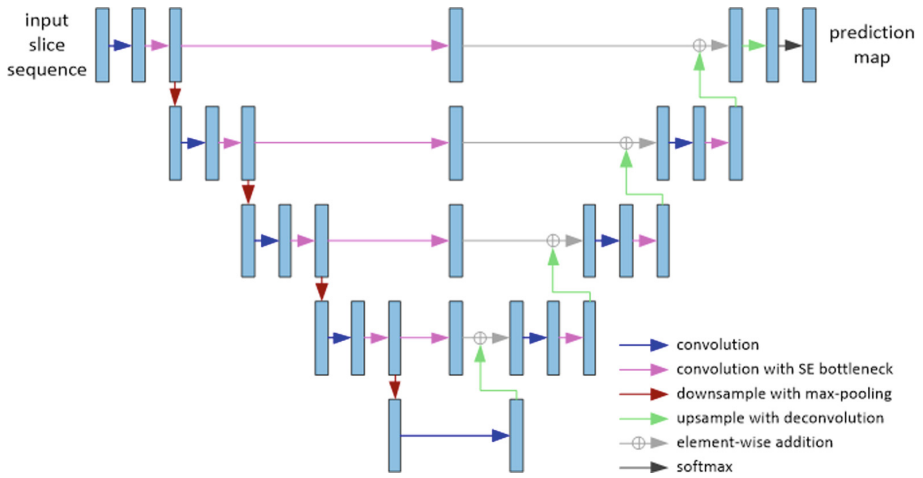
Our main contributions are summarized below:

1. We proposed a multi-scale 2.5D fully convolutional network (FCN) for IVD segmentation and localization on multi-modality MR scans. The back bone of the proposed network is a U-Net [21] like architecture. The input of the 2.5D network is a few adjacent slices from multi-modality MR scans, while the output of this network is a 2D slice corresponding to a certain layer of the input. For the purpose of make full advantage of multi-modality information, Squeeze-and-Excitation (SE) modules [22] are added in the skip connections.
2. We proposed a model fusion strategy to improve accuracy and robustness of IVD prediction. In this work, we trained three different 2.5D networks. The predictions of these models are corresponding to the middle, the rightmost, and the leftmost slices of the input sequence. For the slices located at the middle of 3D images along Z-axis, the average outputs of these models are taken as the final predictions. For the slices near the both edges, IVD predictions are generated by the model, which is corresponding to either the rightmost or the leftmost slice of the input sequence.
3. We proposed a geometric constraint post-processing method to generate accurate IVD localization results. This method takes the intra-subject geometric relationship

of IVDs into account. In our experiments, the false positive regions on the prediction maps are well eliminated by this method.

## 2 Methodology

The detail of IVD segmentation and localization method is elaborated in this section. We start by illustrating the architecture of proposed 2.5D multi-scale FCN for IVD segmentation. Furthermore, we explain the way to harness multi-modality images with this network. To improve the robustness and accuracy of prediction, an ensemble strategy is employed in this work. In order to correct the false positive regions in prediction maps, we proposed a post-processing pipeline, which takes geometric constraint of 7 specified IVDs into account. The final results of segmentation and localization are generated by this post-processing method.



**Fig. 2.** The details of our proposed 2.5D multi-scale FCN for IVD segmentation. The input slice sequence includes 44 slices, which consists of 11 consecutive slices from four modalities with the same corresponding position. The prediction map is corresponding to the middle, the leftmost, or the right most slice of the input sequence.

### 2.1 2.5D Multi-scale FCN for IVD Segmentation

The detail structure of proposed network is shown in Fig. 2. The back bone of this network is a U-Net like architecture, which has achieved great success in medical image processing since it was proposed in 2015. To utilize multi-modality images, the architecture of U-Net is slightly adapted from the origin version. The input of this network is expanded up to 44 (11 slices \* 4 modalities) channels to harness the multi-modality data, while the output is corresponding to a certain position of the input sequence. Besides, residual connections are added between feature maps with the same

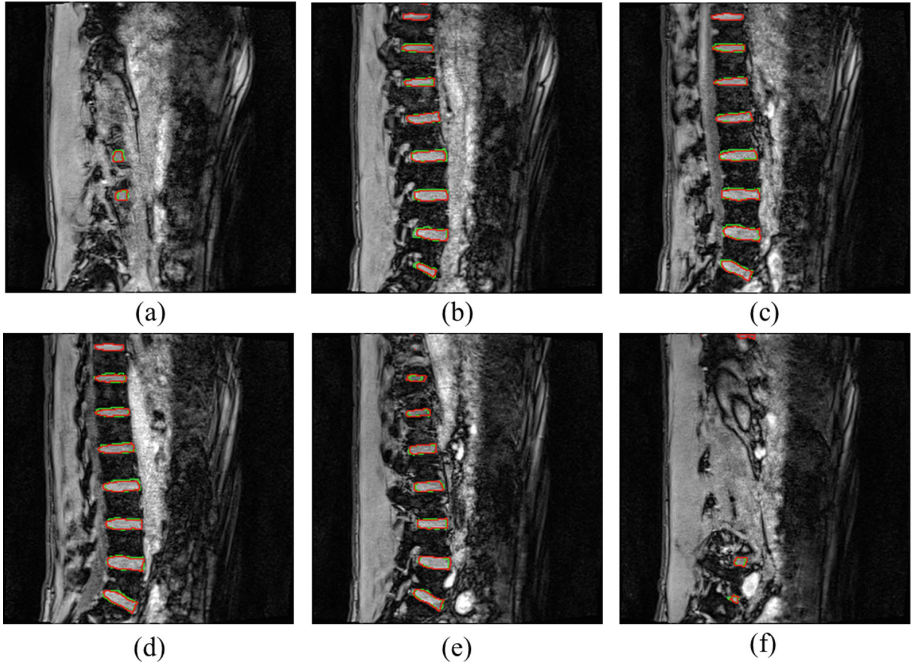
scale. And SE modules are also inserted in skip connections between the contracting path and the expansive path. The reduction ratio used in SE modules is set to be 16.

## 2.2 2.5D Multi-scale FCN Ensemble Strategy

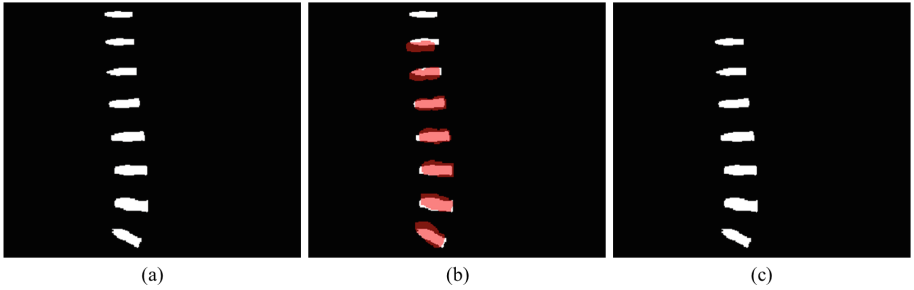
All the multi-modality images used in this work are in the same size of  $256 * 256 * 36$ . For each study, 11 consecutive slices from four modalities with the same corresponding position are extracted and concatenated as the input sequences. And there are 26 such consecutive sequences for each image. These input sequences are utilized to train three 2.5D multi-scale FCNs. The prediction of these models is corresponding to different layers respectively, which are the middle, the leftmost and the rightmost slices in the input sequence. We use  $m_{middle}$ ,  $m_{left}$  and  $m_{right}$  to denote these three models in the following content. The ensemble outputs of these models are produced as prediction results, which are more accurate and robust. For the simplicity of description, a mono-modality 3D image  $V$  is picked as an example. Slices in  $V$  from left to right are denoted as  $S_i (i \in \{1, 2, \dots, 26\})$ . For  $S_6$  to  $S_{31}$ , the average outputs of  $m_{middle}$ ,  $m_{left}$  and  $m_{right}$  are taken as the prediction of IVD segmentation. For  $S_1$  to  $S_5$  and  $S_{32}$  to  $S_{36}$ , the prediction of IVD segmentation is generated by  $m_{left}$  and  $m_{right}$  respectively.

## 2.3 Geometric Constraint Post-processing

Although model ensemble can improve the accuracy and robustness of segmentation results to a certain extent, there are still some obvious false positive regions in the prediction maps. These false positive areas could be categorized as two types, the isolated noise points, and the IVD segmentation above the twelfth thoracic vertebra. Figure 3 visualizes some ensemble prediction maps on opposed-phase. The isolated noise can be well eliminated by excluding the small connected regions in prediction maps. For IVDs above the twelfth thoracic vertebra, we proposed a post-processing method with geometric constraint for removal. Firstly, we picked the ground truths from training set, and aligned them to the segmentation result with reference to the centroid of the last IVD. These ground truths are then registered to the segmentation result with affine transformation. The best fitted one is then selected as the mask. Remove all the connected regions that have no intersection with this mask. The remaining content is right the final prediction of 7 expected IVDs. For the robustness of post-processing, the registered ground truth was dilated before being applied as the mask (Fig. 4).



**Fig. 3.** Examples of prediction map on opposed-phase without post-processing. (a) to (f) are 6 slices extracted from a study. Green contours indicate the boundary of the ground truths. And the ensemble prediction of IVDs is delineated by red lines. (Color figure online)



**Fig. 4.** Illustration of the geometric constraint post-processing. (a) is the ensemble prediction of proposed networks. The red mask in (b) is the chosen registered ground truth with binary dilation. And (c) is the final IVD segmentation result of our method. (Color figure online)

### 3 Experiments and Results

#### 3.1 Data

The performance of our method was evaluated on multi-modality MR scans provided by MICCAI 2018 Challenge on IVDM3Seg. These data were collected from 8 subjects

at two time points of prolonged bed rest study. For each study, four MR scans acquired with different modalities (i.e. in-phase, opposed-phase, water, fat) were enrolled. And the IVDs between the twelfth thoracic vertebra and sacrum are delineated manually as the ground truth. Figure 1 shows an example of these multi-modality images and the corresponding ground truth.

### 3.2 Pre-processing and Data Augmentation

The multi-modality images were pre-processed with some commonly used methods. Firstly, N4 correction algorithm was applied to correct the bias field of MR scans. In the next stage, intensity distribution of the corrected images was normalized as zero mean and unit variance. For the inadequacy of training data, some data augmentation methods (i.e. random scale, rotate, translation, and deformable transformation) are applied during the training stage.

### 3.3 Evaluation Metrics

The segmentation and localization results are evaluated with the following three quantitative metrics:

1. Dice overlap coefficient. The Dice metric is one of the most popular assessments for semantic segmentation, which measures the percentage of true positive voxels in prediction. The definition of Dice can be expressed by the following formula:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \times 100\% \quad (1)$$

Where A is the set of foreground voxels in the ground truth and B denotes the corresponding set in the prediction of foreground.

2. Average absolute distance (ASD). For IVD segmentation task, ASD is the average absolute distance between disc surface of ground truth and segmentation result. Smaller ASD means a better segmentation result.
3. Localization distance. This metric is used for measuring the localization results. It is calculated by the equation below:

$$R = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \quad (2)$$

Where  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  are the absolute distance between the identified IVD centroids and the corresponding ground truth along X-, Y- and Z-axis. It is obvious that a smaller localization distance means a more accurate localization.

### 3.4 Results of MICCAI 2018 On-site Challenge

Tables 1, 2, and 3 list the on-site test results of MICCAI 2018 Challenge on IVD3Seg with proposed method. Our method achieved the state-of-art performance with the respect of all the three quantitative metrics (i.e. Dice, ASD, and Localization distance) among nine participating teams.

**Table 1.** Dice of on-site test results

Dice	Disc_01	Disc_02	Disc_03	Disc_04	Disc_05	Disc_06	Disc_07
Test_01	0.887	0.904	0.927	0.911	0.896	0.890	0.869
Test_02	0.909	0.934	0.941	0.944	0.930	0.925	0.924
Test_03	0.896	0.898	0.901	0.866	0.896	0.817	0.887
Test_04	0.920	0.939	0.938	0.915	0.912	0.885	0.924
Test_05	0.900	0.911	0.917	0.918	0.868	0.898	0.926
Test_06	0.865	0.914	0.930	0.909	0.897	0.898	0.892
Test_07	0.905	0.931	0.931	0.915	0.904	0.890	0.865
Test_08	0.905	0.928	0.924	0.925	0.910	0.907	0.889

**Table 2.** ASD of on-site test results

ASD (mm)	Disc_01	Disc_02	Disc_03	Disc_04	Disc_05	Disc_06	Disc_07
Test_01	0.73	0.67	0.49	0.58	0.63	0.61	0.67
Test_02	0.54	0.48	0.43	0.37	0.43	0.41	0.41
Test_03	0.71	0.80	0.73	0.87	0.58	0.89	0.53
Test_04	0.54	0.47	0.41	0.49	0.47	0.51	0.29
Test_05	0.57	0.70	0.69	0.61	0.98	0.67	0.37
Test_06	0.85	0.66	0.56	0.76	0.78	0.69	0.54
Test_07	0.63	0.52	0.52	0.63	0.62	0.64	0.64
Test_08	0.72	0.63	0.61	0.57	0.62	0.57	0.58

**Table 3.** Localization distance of on-site test results

Localization (mm)	Disc_01	Disc_02	Disc_03	Disc_04	Disc_05	Disc_06	Disc_07
Test_01	0.41	1.42	0.52	0.78	1.37	1.02	1.09
Test_02	0.30	0.73	0.27	0.50	0.72	0.40	0.17
Test_03	0.64	0.31	1.34	2.12	0.97	1.30	0.43
Test_04	0.81	0.52	0.62	1.30	0.49	1.22	0.08
Test_05	0.79	0.30	0.13	1.02	0.92	0.97	0.41
Test_06	1.23	0.47	0.45	0.09	0.65	1.09	0.76
Test_07	0.56	1.09	0.94	1.22	1.19	0.58	1.08
Test_08	0.36	1.31	0.72	0.91	1.04	0.34	0.56



## 4 Discussion

Some common spine diseases, such as low back pain (LBP), have proven to be associated with IVD degeneration [23]. IVD segmentation and localization have important significance in clinical diagnosis and research. In this work, we proposed an automatic IVD segmentation and localization method on multi-modality MRI with 2.5D multi-scale FCN and geometric constraint post-processing.

In the MICCAI 2018 Challenge on IVD3Seg, the deep neural network is the most popular algorithm. For 3D multi-modality MR images, processing with a 3D network is a straightforward approach. Compared to 2D networks, 3D architectures could generate more discriminative spatial features. And these architectures were employed by some teams in this challenge. Due to the plenty of parameters in deep neural networks, a huge amount of data is demanded in training stage. However, there were only 16 studies provided by MICCAI 2018 Challenge on IVD3Seg, which were collected from 8 subjects at two time points. Considering the inadequacy of 3D multi-modality images, we proposed a 2.5D multi-scale FCN architecture as a tradeoff between the capacity of network and the amount of training data. The on-site test results of MICCAI 2018 Challenge on IVD3Seg shows that the performance of 2D networks was better than that of 3D networks in general with limited training data. And our 2.5D FCN surpassed both 2D and 3D architectures.

The intra-subject morphology and topology relationship between IVDs are similar inter-subjects. And it is potential to be utilized for IVD localization. However, this relationship is hard to be captured by FCN. To take this information into account, we proposed a geometric constraint post-processing method based on registration. And it shows great performance in on-site test of MICCAI 2018 Challenge on IVD3Seg. It should be noticed that our registration-based post-processing relies on the inter-subject consistency of IVD intra-subject geometric relationship. If this consistency was destroyed by some severe spine diseases, this method may produce wrong cases. The IVD localization method with better robustness remains to be explored in the future work.

## References

1. An, H.S., et al.: Introduction: disc degeneration: summary. *Spine* **29**, 2677–2678 (2004)
2. Urban, J.P., Roberts, S.J.A.R.T.: Degeneration of the intervertebral disc. *Arthritis Res. Ther.* **5**, 120 (2003)
3. Maniadakis, N., Gray, A.J.P.: The economic burden of back pain in the UK. *Pain* **84**, 95–103 (2000)
4. Luoma, K., Riihimäki, H., Luukkonen, R., Raininko, R., Viikari-Juntura, E., Lamminen, A.J.S.: Low back pain in relation to lumbar disc degeneration. *Spine* **25**, 487–492 (2000)
5. Niemeläinen, R., Videman, T., Dhillon, S., Battié, M.: Quantitative measurement of intervertebral disc signal using MRI. *Clin. Radiol.* **63**, 252–255 (2008)
6. Violas, P., Estivalezes, E., Briot, J., de Gauzy, J.S., Swider, P.: Objective quantification of intervertebral disc volume properties using MRI in idiopathic scoliosis surgery. *Magn. Reson. Imaging* **25**, 386–391 (2007)
7. Neubert, A., et al.: Automated 3D segmentation of vertebral bodies and intervertebral discs from MRI. In: 2011 International Conference on Digital Image Computing Techniques and Applications (DICTA), pp. 19–24 (2011)

8. Corso, J.J., Alomari, R.S., Chaudhary, V.: Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008. LNCS, vol. 5241, pp. 202–210. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-85988-8\\_25](https://doi.org/10.1007/978-3-540-85988-8_25)
9. Chevreteffs, C., Chéret, F., Grimard, G., Aubin, C.-E.: Watershed segmentation of intervertebral disk and spinal canal from MRI images. In: Kamel, M., Campilho, A. (eds.) ICIAR 2007. LNCS, vol. 4633, pp. 1017–1027. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-74260-9\\_90](https://doi.org/10.1007/978-3-540-74260-9_90)
10. Ben Ayed, I., Punithakumar, K., Garvin, G., Romano, W., Li, S.: Graph cuts with invariant object-interaction priors: application to intervertebral disc segmentation. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 221–232. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-22092-0\\_19](https://doi.org/10.1007/978-3-642-22092-0_19)
11. Raja's, A., Corso, J.J., Chaudhary, V.: Labeling of lumbar discs using both pixel-and object-level features with a two-level probabilistic model. *IEEE Trans. Med. Imaging* **30**, 1–10 (2011)
12. Chevreteffs, C., Chéret, F., Aubin, C.É., Grimard, G.: Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from MR images. *IEEE Trans. Inf Technol. Biomed.* **13**, 608–620 (2009)
13. Schmidt, S., et al.: Spine detection and labeling using a parts-based graphical model. In: Karssemeijer, N., Lelieveldt, B. (eds.) IPMI 2007. LNCS, vol. 4584, pp. 122–133. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-73273-0\\_11](https://doi.org/10.1007/978-3-540-73273-0_11)
14. Kelm, B.M., et al.: Spine detection in CT and MR using iterated marginal space learning. *Med. Image Anal.* **17**, 1283–1292 (2013)
15. Huang, S.-H., Chu, Y.-H., Lai, S.-H., Novak, C.L.: Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. *IEEE Trans. Med. Imaging* **28**, 1595–1605 (2009)
16. Wang, Z., Zhen, X., Tay, K., Osman, S., Romano, W., Li, S.: Regression segmentation for  $M^3$  spinal images. *IEEE Trans. Med. Imaging* **34**, 1640–1648 (2015)
17. Li, X., et al.: 3D multi-scale FCN with random modality voxel dropout learning for Intervertebral Disc Localization and Segmentation from Multi-modality MR Images. *Med. Image Anal.* **45**, 41–54 (2018)
18. Chen, H., Dou, Q., Wang, X., Qin, J., Cheng, J.C.Y., Heng, P.-A.: 3D fully convolutional networks for intervertebral disc localization and segmentation. In: Zheng, G., Liao, H., Jannin, P., Cattin, P., Lee, S.-L. (eds.) MIAR 2016. LNCS, vol. 9805, pp. 375–382. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-43775-0\\_34](https://doi.org/10.1007/978-3-319-43775-0_34)
19. Zhang, W., et al.: Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage* **108**, 214–224 (2015)
20. Havaei, M., Guizard, N., Chapados, N., Bengio, Y.: HeMIS: hetero-modal image segmentation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 469–477. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46723-8\\_54](https://doi.org/10.1007/978-3-319-46723-8_54)
21. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
22. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. *arXiv preprint arXiv:1709.01507*, vol. 7 (2017)
23. Kjaer, P., Leboeuf-Yde, C., Korsholm, L., Sorensen, J.S., Bendix, T.: Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine* **30**, 1173–1180 (2005)