DSMS-FCN: A Deeply Supervised Multi-scale Fully Convolutional Network for Automatic Segmentation of Intervertebral Disc in 3D MR Images

Guodong Zeng and Guoyan Zheng(✉)

Institute for Surgical Technology and Biomechanics,
University of Bern, Bern, Switzerland
guoyan.zheng@istb.unibe.ch

Abstract. This paper addresses the challenging problem of segmentation of intervertebral discs (IVDs) in three-dimensional (3D) T2-weighted magnetic resonance (MR) images. We propose a deeply supervised multi-scale fully convolutional network for segmentation of IVDs in 3D MR images. After training, our network can directly map a whole volumetric data to its volume-wise labels. Multi-scale deep supervision is designed to alleviate the potential gradient vanishing problem during training. It is also used together with partial transfer learning to boost the training efficiency when only small set of labeled training data are available. The present method was validated on the MICCAI 2015 IVD segmentation challenge datasets. Our method achieved a mean Dice overlap coefficient of 92.0% and a mean average symmetric surface distance of 0.41 mm. The results achieved by our method are better than those achieved by the state-of-the-art methods.

Keywords: Intervertebral disc · MRI · Segmentation
Deep learning · Fully convolutional network

1 Introduction

Intervertebral disc (IVD) degeneration is a major cause for chronic back pain and function incapacity [1]. In clinical practice, spine magnetic resonance (MR) imaging (MRI) is the preferred modality in diagnosis and treatment planning of various spinal pathologies such as disc herniation, slipped vertebra and so on, not only because MRI is non-invasive and does not use ionizing radiation, but more importantly because it offers good soft tissue contrast that allows for visualization of disc's internal structure [2].

Accurate IVD segmentation from spine MR image is therefore very important for correct diagnosis and treatment planning [1,3]. Traditionally, most quantitative studies on IVD degeneration have been done by manually segmenting the data, which is tedious, time-consuming and error-prone. On the other hand, a

B. Glocker et al. (Eds.): MSKI 2017, LNCS 10734, pp. 148–159, 2018.
https://doi.org/10.1007/978-3-319-74113-0_13

fully-automatic system for IVD identification will significantly reduce the time of the diagnosis. An automatic system might also help reduce errors caused by subjective factors and improve the consistency of diagnosis standards. In this way, it can immediately benefit clinical applications and spinal biomechanics research.

In the literature, different methods have been proposed of IVD segmentation [4–8]. There exist methods based on watershed algorithm [4], atlas registration [5], graph cuts with geometric priors from neighboring discs [6], template matching and statistical shape model [7], or anisotropic oriented flux detection [8]. Most of these methods work only on two-dimensional sagittal images and only a few methods [7] address the challenging three-dimensional (3D) IVD segmentation problem. See [9] for a comprehensive review of existing IVD segmentation methods.

Recently, machine learning-based methods have gained more and more interest. For example, Zhan et al. [10] presented a hierarchical strategy and local articulated model to detect vertebrae and discs 3D MR images and Kelm et al. [11] proposed to use iterated marginal space learning for spine detection in computed tomography (CT) and MR images. A unified data-driven regression and classification framework was suggested by Chen et al. [12] to tackle the problem of localization and segmentation of IVDs from T2-weighted MR data, and Wang et al. [13] proposed to address the segmentation of multiple anatomic structures in multiple anatomical planes from multiple imaging modalities via a sparse kernel machines-based regression.

The more recent development on deep neural networks, and in particular on convolutional neural networks (CNN), suggests another course of methods to solve the challenging IVD segmentation problem [14–22]. Contrary to conventional shallow learning methods where feature design is crucial, deep learning methods automatically learn hierarchies of relevant features directly from the training data [15]. More recently, 3D volume-to-volume segmentation networks were introduced, including 3D U-Net [20], 3D V-Net [21] and a 3D deeply supervised network [22].

In this paper, we propose a deeply supervised multi-scale fully convolutional network (FCN) called "DSMS-FCN" for fully automatic IVD segmentation in 3D T2-weighted MR images. After training, our network can directly map a whole volumetric data to its volume-wise label. Multi-scale deep supervision is designed to alleviate the potential gradient vanishing problem during training. It is also used together with partial transfer learning to boost the training efficiency when only small set of labeled training data are available.

The paper is organized as follows. In Sect. 2, we will describe the proposed architecture and algorithm. The application to the MICCAI 2015 IVD segmentation challenge dataset will be presented in Sect. 3, and we conclude with a discussion in Sect. 4.

2 Method

Figure 1 illustrates the architecture of our proposed neural network for the automatic IVD segmentation in 3D T2-weighted MR images. Our network employs a

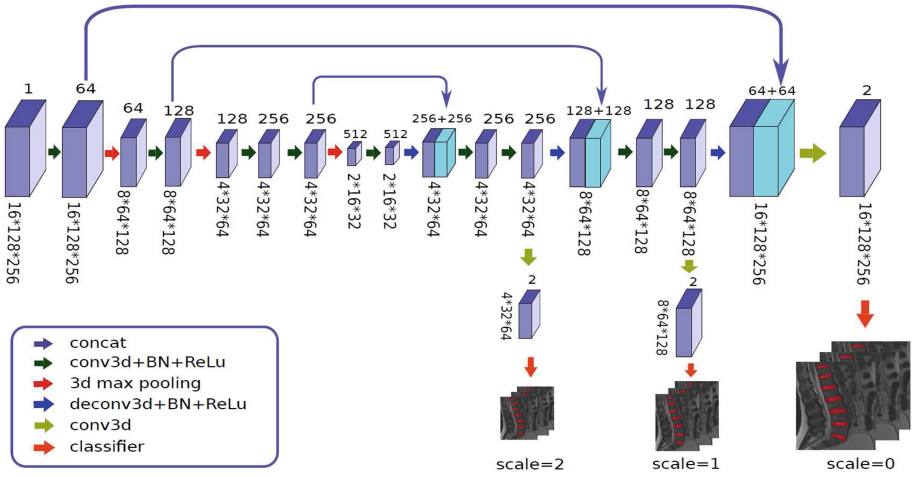

Fig. 1. A schematic illustration of our proposed network architecture. For each block, the digits above indicate the number of feature stack while the numbers below represent the data size.

deeply supervised multi-scale fully convolutional network. In this section, firstly the detailed architecture of our proposed model is elaborated, and then we will introduce the multi-scale deep supervision. Finally, partial transfer learning, which is designed to boost the training efficiency, will be described.

2.1 3D FCN with Skip Connection

Our proposed network is inspired by 3DFCN [23] but with significant differences. Similar to 3DFCN, our network is also a 3D fully convolutional network and consists of two parts, i.e., the encoder part (contracting path) and the decoder part (expansive path). The encoder part focuses on analysis and feature representation learning from the input data while the decoder part generates segmentation results, relying on the learned features from the encoder part. Our network can take arbitrary-sized volumetric data as input and outputs voxel-wise segmentation probability map in the same size as the input.

Different from 3DFCN, long and short skip connections, which help recover spatial context lost in the contracting encoder, are used in our network as shown in Fig. 1. The importance of skip connection in biomedical image segmentation has been demonstrated by previous studies [24]. Skip connections have been widely used in many different convolutional neural networks including Resi-Net [25] and 3D U-Net [20].

In 3DFCN [23], big kernel sizes (e.g. $5 \times 7 \times 7$) are utilized in the convolutional layers. However, previous studies have shown that small kernel size are more helpful for training of deep neural network [26]. For this reason, in our network, kernel size of $3 \times 3 \times 3$ and strides of 1 are utilized for all convolutional layers, and kernel size of $2 \times 2 \times 2$ is used in all max pooling layers. Batch normalization

(BN) [27] and rectified linear unit (ReLU) are adopted to speed up the training and to enhance the gradient back propagation.

2.2 Multi-scale Deep Supervision

Training a deep neural network is challenging. As the matter of gradient vanishing, final loss cannot be efficiently back propagated to shallow layers, which is more difficult for 3D cases when only a small set of annotated data is available. To address this issue, we inject two down-scaled branch classifiers into our network in addition to the classifier of the main network, which is another difference between our network and 3DFCN [23]. By doing this, segmentation is performed at multiple output layers. For the classifier at the coarse scale which is closer to the encoder part, it generates segmentation results with the coarsest resolution, while the classifiers at the middle and the fine scales generate segmentation results with the intermediate and the finest resolutions, respectively. As a result, classifiers in different scales can take advantage of multi-scale context, which has been demonstrated in previous work on segmentation of 3D liver CT and 3D heart MR images [22]. Furthermore, with the loss calculated by the prediction from classifiers from different scales, more effective gradient back propagation can be achieved by direct supervision on the hidden layers.

Specifically, let W be the weights of main network and $w = \{w^0, w^1, ... w^{M-1}\}$ be the weights of classifiers at different scales, where M is the number of classifier branches. For the training samples $S = (X, Y)$, where X represents training sub-volume patches and Y represents the class labels while $Y \in \{0, 1\}$.

$$L_{cls}(X, Y; W, w) = \sum_{m=0}^{M-1} \sum_{(x_i, y_i) \in S^m} \alpha_m l^m(x_i, y_i | W, w^m), \tag{1}$$

where $S = \{S^0, S^1, \ldots, S^{M-1}\}$; S^0 is a sub-volume patch directly sampled from a training image while S^m contains the examples (x_i, y_i) at scale of $m > 0$, which is obtained by downsampling S^0 by a factor of 2^m along each dimension; w^m is the weights of the classifier at scale of m; α_m is the weight of l^m, which is the loss calculated by a training sample x_i, y_i at scale of m.

$$l^m(x_i, y_i | W, w^m) = -\log p(y_i = t(x_i) | x_i; W, w^m), \tag{2}$$

where $p(y_i = t(x_i)|x_i; W, w^m)$ is the probability of predicted class label $t(x_i)$ corresponding to sample $x_i \in S^m$.

The total loss of our multi-scaled deeply supervised model will be:

$$L_{total}(X, Y; W, w) = L_{cls}(X, Y; W, w) + \lambda(\psi(W) + \sum_m \psi(w^m)), \tag{3}$$

where $\psi()$ is the regularization term (L_2 norm in our experiment) with hyper parameter λ.

2.3 Partial Transfer Learning

It is difficult to train a deep neural network from scratch because of limited annotated data. Training deep neural network requires large amount of annotated data, which are not always available, although data augmentation can partially address the problem. Furthermore, randomly initialized parameters make it more difficult to search for an optimal solution in high dimensional space. Transfer learning from an existing network, which has been trained on a large set of data, is a common way to alleviate the difficulty. Previous studies [28] demonstrated that transferring features from another pre-trained model can boost the generalization, and that the effect of transfer learning was related to the similarity between the task of the pre-trained model and the target task. Furthermore, the same study also demonstrated that weights of shallow layers in deep neural network were generic while those of deep layers were more related to specific tasks.

To best utilize the advantage of transfer learning, we need to transfer from a model trained on a related task. In this paper, a pre-trained model in our previous work was adopted [29], which is designed for the task of segmentation of the proximal femur from 3D T1-weighted MR Images. More specifically, the weights of the main network are initialized from our previous model [29], while the weights of all branch classifiers are initialized from a Gaussian distribution ($\mu = 0, \sigma = 0.01$).

2.4 Implementation Details

The proposed network was implemented in python using TensorFlow framework and trained on a desktop with a 3.6 GHz Intel(R) i7 CPU and a GTX 1080 Ti graphics card with 11 GB GPU memory.

3 Experiments and Results

3.1 Data Description

The training data provided by the MICCAI 2015 IVD challenge organizers consist of 15 3D T2-weighted turbo spin echo MR images and the associated ground truth segmentation [9]. These 15 3D T2-weighted MR images were acquired from fifteen patients in two different studies. Each patient was scanned with 1.5 Tesla MRI scanner of Siemens (Siemens Healthcare, Erlangen, Germany). The pixel spacings of all the images are sampled to $2 \times 1.25 \times 1.25\,\text{mm}^3$. There are 7 IVDs T11-L5 to be segmented from each image. Thus, in each image these IVD regions have been manually identified and segmented.

The MICCAI 2015 IVD challenge organizers also released two test datasets. Each test dataset consists of five 3D T2-weighted turbo spin echo MR images. Thus, in this paper, our network was trained on the fifteen 3D training data first, and are then evaluated on the ten test data.

3.2 Training Patches Preparation

In order to enlarge the training samples, data augmentation was utilized. Specifically, each training data was rotated (90, 180, 270) degrees around the y axis of the image and flipped horizontally (taking the z axis as the vertical direction). After that, we got in total 120 images for training our network.

Our network takes a fixed-sized sub-volume as input, and employs end-to-end learning and voxel-wise inference. During training, sub-volume patches with the size of $16 \times 256 \times 128$ was randomly cropped from 120 training examples whose size are about $40 \times 300 \times 300$. In each epoch of training, 120 training images were randomly shuffled and then sub-volume patches was randomly cropped with batch-size of 2 by 5 times from each volumetric training image. Before fed into the network for training, each sub-volume patch was normalized by zero mean and unit variance. In total, for each epoch of training, we trained the network using 1200 ($120 \times 5 \times 2$) sub-volume patches.

3.3 Training

We trained our network for 10,000 iterations after partial transfer learning. All weights were updated by the stochastic gradient descent (SGD) algorithm (momentum $= 0.9$, weight decay $= 0.005$). Learning rate was initialized as 1×10^{-3} and halved by every 3,000 times. In our experiment, we used three branch classifiers at three different scales. The loss weights of three classifiers α_0, α_1 and α_2 are 1.0, 0.67 and 0.33, respectively. The hyper parameter λ was chosen to be 0.005.

3.4 Testing

Our trained models can estimate labels of an arbitrary-sized volumetric image. Given a test volumetric image, we extracted overlapped sub-volume patches with the size of $16 \times 256 \times 128$, and fed them to the trained network to get prediction probability maps. For the overlapped voxels, the final probability maps would be the average of the probability maps of the overlapped patches, which were then used to derive the final segmentation results. After that, we conducted morphological operations to remove isolated small volumes and internal hole.

3.5 Valuation

The segmented results were compared with the associated ground truth segmentation. For each test image, we evaluated both the surface distance as well as the volume overlap measurements of results obtained by different segmentation.

In [9], to compute the average absolute distance (ASD) between the ground truth IVD surface and the automatically segmented surface, surface meshes from binary IVD segmentation were generated first using the Matlab toolbox Iso2mesh [30]. In contrast, in this study, we adopted the average symmetric surface distance (ASSD) as introduced in [31] to measure the surface distance. More specifically,

Table 1. Results on the Test1 dataset from the MICCAI 2015 intervertebral disc segmentation challenge.

Parameters	Mean ± STD
Dice overlap coefficients (%)	91.4 ± 0.5
Jaccard (%)	84.2 ± 1.0
Precision (%)	94.2 ± 2.5
Recall (%)	88.9 ± 2.3
ASSD (mm)	0.44 ± 0.055

STD - standard deviation
ASSD - average symmetric surface distance

ASSD is given in millimeters and based on the surface voxels (instead of surface meshes as in [9]) of two segmentation A and B. Surface voxels are defined by having at least one on-object voxel within their 26-neighborhood. For each surface voxel of A, the Euclidean distance to the closest surface voxel of B is calculated using the approximate nearest neighbor technique [32] and stored. The same process is then applied to surface voxels of B to A in order to symmetry. The ASSD is then defined as the average of all stored distances, which is zero for a perfect segmentation.

Given two binary segmentations of a test image, we compute following volume overlap measurements including Dice overlap coefficient [33], Jaccard coefficient [33], precision and recall.

3.6 Results

Table 1 shows the results of our method when evaluated on the Test1 dataset of the MICCAI 2015 IVD segmentation challenge and Table 2 shows the results of our method when evaluated on the Test2 dataset of the MICCAI 2015 IVD segmentation challenge. A mean Dice overlap coefficient of 92.0% and a mean ASSD of 0.41 mm were achieved by our method. Furthermore, slightly better results

Table 2. Results on the Test2 dataset from the MICCAI 2015 intervertebral disc segmentation challenge.

Parameters	Mean ± STD
Dice overlap coefficients (%)	92.6 ± 1.1
Jaccard (%)	86.4 ± 2.0
Precision (%)	93.8 ± 1.5
Recall (%)	91.4 ± 1.6
ASSD (mm)	0.38 ± 0.045

STD - standard deviation
ASSD - average symmetric surface distance

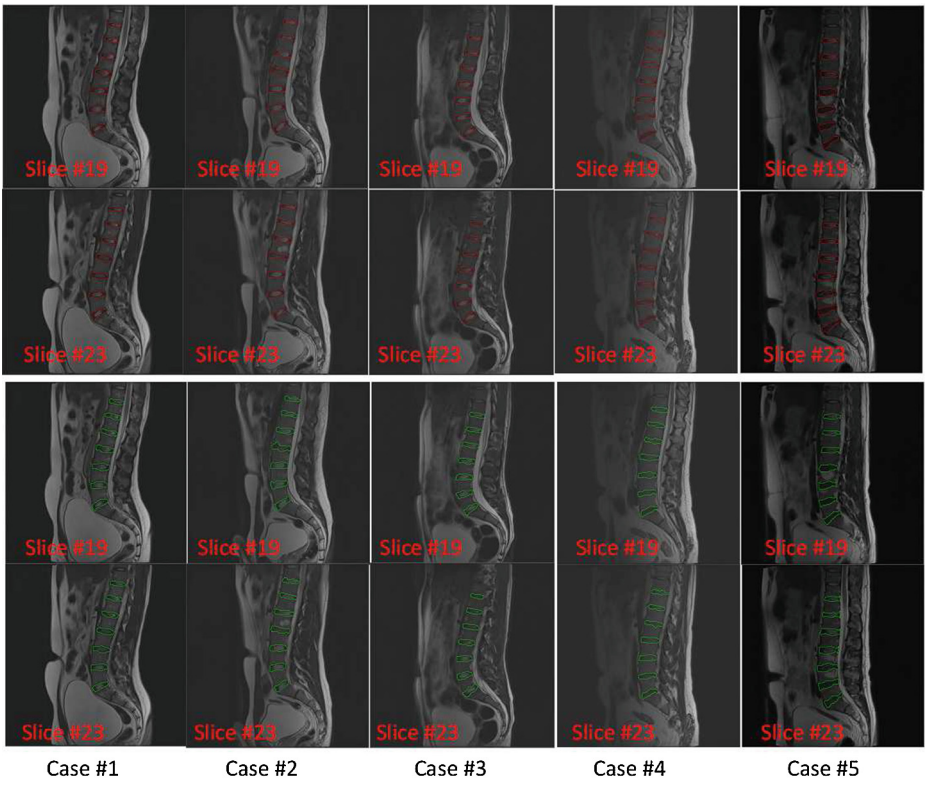

Fig. 2. Qualitative comparison of the results achieved by our method on the Test1 dataset (top two rows) and ground truth segmentation (bottom two rows). For each case, two slices are shown.

were obtained when our method was evaluated on the Test2 dataset than when our method was evaluated on Test1 dataset. Without using any time-consuming registration step or incorporating any advanced shape prior, our method achieved results that were better than those achieved by the state-of-the-art methods [9]. For example, the best segmentation method in the MICCAI 2015 IVD segmentation challenge was the on submitted by Korez et al. [34] where a mean Dice overlap coefficient of 91.8% was reported. Figure 2 shows examples of automatic segmentation achieved by our method on the Test1 dataset and Fig. 3 shows examples of segmentation achieved by our method on the Test2 dataset.

Implemented with Python using TensorFlow framework, our network took about 40 s to test one volumetric MR image with size of $40 \times 300 \times 300$.

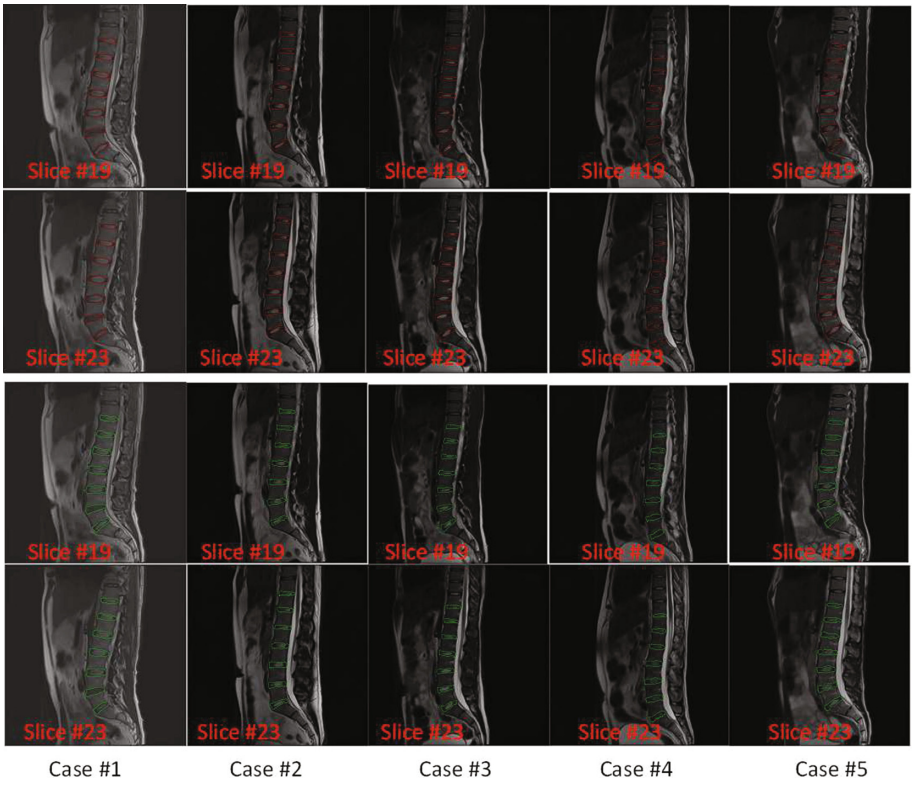

Fig. 3. Qualitative comparison of the results achieved by our method on the Test2 dataset (top two rows) and ground truth segmentation (bottom two rows). For each case, two slices are shown.

4 Conclusion

In this paper, we proposed to use a deeply supervised multi-scale fully convolutional network to solve the challenging IVD segmentation problem. The present method was evaluated on the MICCAI 2015 IVD segmentation challenge datasets and the results achieved by the present method were better than those achieved by the state-of-the-art methods.

In comparison with 3DFCN as introduced by Chen et al. [23], where they incorporate neither skip connection nor multi-scale deep supervision, our method achieved much better segmentation results. More specifically, evaluated on the Test1 and Test2 datasets of the MICCAI 2015 IVD segmentation challenge, their method achieved a mean Dice overlap coefficient of 88.4% and 89.0%, respectively. In contrast, evaluated on the same two datasets, our method achieved a mean Dice overlap coefficient of 91.4% and 92.6%, respectively. The results demonstrated that the incorporation of skip connections and the multi-scale deep supervision, when combined with partial transfer learning, did improve the performance of a 3D FCN.

References

1. Modic, M., Ross, J.: Lumbar degenerative disk disease. Radiology **245**(1), 43–61 (2007)
2. Parizel, P., Van Goethem, J., Van den Hauwe, L., Voormolen, M.: Degenerative disc disease. In: Van Goethem, J., et al. (eds.) Spinal Imaging, pp. 127–156. Medical Radiology. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-68483-1_6
3. An, H., Anderson, P., Haughton, V., Iatridis, J., Kang, J., Lotz, J., Natarajan, R., Oegema, T.J., Roughley, P., Setton, L., Urban, J., Videman, T., Andersson, G., Weinstein, J.: Introduction: disc degeneration: summary. Spine **29**(23), 2677–2678 (2004)
4. Chevrefils, C., Cheriet, F., Aubin, C., Grimard, G.: Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from MR images. IEEE Trans. Inf Technol. Biomed. **13**(4), 608–620 (2009)
5. Michopoulou, S., Costaridou, L., Panagiotopoulos, E., Speller, R., Panayiotakis, G., Todd-Pokropek, A.: Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine. IEEE Trans. Biomed. Eng. **56**(9), 2225–2231 (2009)
6. 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
7. Neubert, A., Fripp, J., Engstrom, C., Schwarz, R., Lauer, L., Salvado, O., Crozier, S.: Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys. Med. Biol. **57**(24), 8457–8376 (2012)
8. 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)
9. Zheng, G., Chu, C., Belavý, D., Ibragimov, B., Korez, R., Vrtovec, T., Hutt, H., Everson, R., Meakin, J., Andrade, I., Glocker, B., Chen, H., Dou, Q., Heng, P., Wang, C., Forsberg, D., Neubert, A., Fripp, J., Urschler, M., Stern, D., Wimmer, M., Novikov, A., Cheng, H., Armbrecht, G., Felsenberg, D., Li, S.: Evaluation and comparison of 3D intervertebral disc localization and segmentation methods for 3D T2 MR data: a grand challenge. Med. Image Anal. **35**, 327–344 (2017)
10. Zhan, Y., Maneesh, D., Harder, M., Zhou, X.S.: Robust MR spine detection using hierarchical learning and local articulated model. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7510, pp. 141–148. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33415-3_18
11. Kelm, M., Wels, M., Zhou, S., Seifert, S., Suehling, M., Zheng, Y., Comaniciu, D.: Spine detection in CT and MR using iterated marginal space learning. Med. Image Anal. **17**(8), 1283–1292 (2013)
12. 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)
13. Wang, Z., Zhen, X., Tay, K., Osman, S., Romano, W., Li, S.: Regression segmentation for M^3 spinal images. IEEE Trans. Med. Imaging **34**(8), 1640–1648 (2015)
14. Bengio, Y.: Learning deep architectures for AI. Found. Trends Mach. Learn. **2**(1), 1–127 (2009)

15. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Pereira, F., et al. (eds.) Proceedings of Neural Information Processing Systems – NIPS 2012, vol. 25, pp. 1097–1105. NIPS (2012)
16. Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M.: Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8150, pp. 246–253. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40763-5_31
17. Roth, H.R., et al.: A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8673, pp. 520–527. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10404-1_65
18. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition - CVPR 2015, pp. 3431–3440 (2015)
19. Roth, H., Yao, J., Lu, L., Stieger, J., Burns, J., Summers, R.: Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications. In: Yao, J., et al. (eds.) Proceedings of 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine CSI 2014, LNCVB, vol. 20, pp. 3–12. Springer (2015). https://doi.org/10.1007/978-3-319-14148-0_1
20. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
21. Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: Proceedings of 4th International Conference on 3D Vision - 3DV 2016, pp. 565–571. IEEE (2016)
22. Dou, Q., Yu, L., Chen, H., Jin, Y., Yang, X., Qin, J., Heng, P.A.: 3D deeply supervised network for automated segmentation of volumetric medical images. Med. Image Anal. **41**, 40–54 (2017)
23. 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
24. Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: Carneiro, G., et al. (eds.) LABELS/DLMIA -2016. LNCS, vol. 10008, pp. 179–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46976-8_19
25. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition - CVPR 2016, pp. 770–778. IEEE (2016)
26. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 (2014)
27. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing. In: Proceedings of 32nd International Conference on Machine Learning - ICML 2015, vol. 37, pp. 448–456. PLMR (2015)
28. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? In: Ghahramani, Z., et al. (eds.) Proceedings of Advances in Neural Information Processing Systems - NIPS 2014, pp. 3320–3328. MIT Press (2014)

29. Zeng, G., Yang, X., Li, J., Yu, L., Heng, P., Zheng, G.: 3D U-Net with multi-level deep supervision:fully automatic segmentation of proximal femur in 3D MR images. In: 8th MICCAI International Workshop on Machine Learning in Medical Imaging - MLMI 2017 (2017)
30. Fang, Q., Boas, D.: Tetrahedral mesh generation from volumetric binary and gray-scale images. In: Proceedings of 6th IEEE International Symposium on Biomedical Imaging - ISBI 2009, pp. 1142–1145. IEEE (2009)
31. Heimann, T., van Ginneken, B., Styner, M., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., Bello, F., Binnig, G., Bischof, H., Bornik, A., Cashman, P., Chi, Y., Cordova, A., Dawant, B., Fidrich, M., Furst, J., Furukawa, D., Grenacher, L., Hornegger, J., Kainmüller, D., Kitney, R., Kobatake, H., Lamecker, H., Lange, T., Lee, J., Lennon, B., Li, R., Li, S., Meinzer, H., Nemeth, G., Raicu, D., Rau, A., van Rikxoort, E., Rousson, M., Rusko, L., Saddi, K., Schmidt, G., Seghers, D., Shimizu, A., Slagmolen, P., Sorantin, E., Soza, G., Susomboon, R., Waite, J., Wimmer, A., Wolf, I.: Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans. Med. Imaging **28**(8), 1251–1265 (2009)
32. Arya, S., Mount, D., Netanyahu, N., Silverman, R., Wu, A.: An optimal algorithm for approximate nearest neighbor searching. J. ACM **45**(6), 891–923 (1998)
33. Karasawa, K., Oda, M., Kitasaka, T., Misawa, K., Fujiwara, M., Chu, C., Zheng, G., Rueckert, D., Mori, K.: Multi-atlas pancreas segmentation: atlas selection based on vessel structure. Med. Image Anal. **39**, 18–28 (2017)
34. Korez, R., Ibragimov, B., Likar, B., Pernuš, F., Vrtovec, T.: Deformable model-based segmentation of intervertebral discs from MR spine images by using the SSC descriptor. In: Vrtovec, T., Yao, J., Glocker, B., Klinder, T., Frangi, A., Zheng, G., Li, S. (eds.) CSI 2015. LNCS, vol. 9402, pp. 117–124. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41827-8_11