

# VertNet: Accurate Vertebra Localization and Identification Network from CT Images

Zhiming Cui<sup>1,2,5</sup>, Changjian Li<sup>3</sup>, Lei Yang<sup>2</sup>, Chunfeng Lian<sup>4</sup>, Feng Shi<sup>5</sup>, Wenping Wang<sup>2</sup>, Dijia Wu<sup>5</sup>, and Dinggang Shen<sup>1,5(⊠)</sup>

<sup>1</sup> School of Biomedical Engineering, ShanghaiTech University, Shanghai, China dgshen@shanghaitech.edu.cn

 $^{2}\,$  Department of Computer Science, The University of Hong Kong, Hong Kong, China

<sup>3</sup> Department of Computer Science, University College London, London, UK

<sup>4</sup> School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

<sup>5</sup> Shanghai United Imaging Intelligence Co. Ltd., Shanghai, China

Abstract. Accurate localization and identification of vertebrae from CT images is a fundamental step in clinical spine diagnosis and treatment. Previous methods have made various attempts in this task; however, they fail to robustly localize the vertebrae with challenging appearance or identify vertebra labels from CT images with a limited field of view. In this paper, we propose a novel two-stage framework, VertNet, for accurate and robust vertebra localization and identification from CT images. Our method first detects all vertebra centers by a weighted voting-based localization network. Then, an identification network is designed to identify the label of each detected vertebra in leveraging the synergy of global and local information. Specifically, a bidirectional relation module is designed to learn the global correlation among vertebrae along the upward and downward directions, and a continuous label map with dense annotation is employed to enhance the feature learning in local vertebra patches. Extensive experiments on a large dataset collected from real-world clinics show that our framework can accurately localize and identify vertebrae in various challenging cases and outperforms the stateof-the-art methods.

# 1 Introduction

Localizing and identifying each vertebra from CT images are two essential steps for clinical practice such as surgical planning [7,8], pathological diagnosis [5] and post-operative assessment [9], as the shape of the spine can serve as an important anatomical reference for other organs in these practices. To this end, doctors usually need to manually localize and identify the vertebrae in CT images, which is laborious and time-consuming. In this regard, a fully automatic method with high precision is practically demanded.

Automating these steps have long been the goal for medical imaging researches (e.g., [6,15,19]) but it remains a challenging task. This is because © Springer Nature Switzerland AG 2021

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12905, pp. 281–290, 2021. https://doi.org/10.1007/978-3-030-87240-3\_27



**Fig. 1.** Typical challenging cases: (a) Spine with pathological fracture; (b) Image with metal artifacts; (c) Adjacent vertebrae having similar shape appearance; (d) Image with a limited field of view.

many spines could be pathological (Fig. 1(a)) or with metal artifacts in the CT images (Fig. 1(b)). More recently, deep learning algorithms that can exploit the large-scale data have shown promising results for these two tasks. One line of the previous works [1,17] employ a one-stage framework to directly segment each of the vertebrae with the corresponding label. Yet, these methods are prone to produce segmentation artifacts in challenging cases, especially where adjacent vertebrae are similar in appearance (e.g., the 7th to the 9th thoracic vertebrae as shown in Fig. 1(c)). To improve the performance, a two-stage approach [4, 10, 13, 16, 18] has been proposed to first localize the vertebrae and then identify the label of each vertebra. These methods utilize the Recurrent Neural Network (RNN) or Long Short-Term Memory (LSTM) network for modeling the relationship of neighboring vertebrae. However, such a model may not fully capture the global dependency among all vertebrae and usually limit to local regions, which is important for handling vertebra identification from CT images with a limited field of view (Fig. 1(d)).

To address the aforementioned issues in challenging cases, we present a novel two-stage learning framework for automatic vertebra localization and labeling. Firstly, in the vertebra localization stage, instead of only utilizing the Gaussianlike heatmap to represent the vertebra centers, we combine it with 3D vertebra center offsets to generate more reliable vertebra positions with the supervision of the Chamfer distance. Then, in the second stage of vertebra identification, with the guidance of the detected vertebra centers, vertebra proposals are generated and fed to the identification network for vertebra labeling. As both global and local information are essential for accurate vertebra identification, we propose a bidirectional relation module to capture the global relationships among vertebrae using a self-attention mechanism. Moreover, we also introduce a continuous label map to parameterize the sequence of discrete vertebrae labels, and formulate the prediction of the continuous label map as an additional task for learning fine-grained features in local proposal patches. Our framework was extensively evaluated on a large dataset collected from clinics, which includes many challenging cases. Compared with the state-of-the-art performance, our proposed approach achieved superior results qualitatively and quantitatively, giving the high usability in real-world clinical practices.



**Fig. 2.** An overview of the proposed *VertNet* for vertebra localization (Sect. 2.1) and identification (Sect. 2.2) from input CT image.

### 2 Method

An overview of our *VertNet* for vertebra localization and identification in CT images is shown in Fig. 2, which consists of the localization and identification sub-networks. We elaborate the framework in this section.

#### 2.1 Vertebrae Localization

To localize each vertebra in CT images accurately, we formulate it as a vertebra center point prediction problem. An intuitive solution is to directly regress the 3D vertebra heatmap representing the centers, but it is prone to fail, especially around the cervical vertebrae that pack tightly and are hard to be separated.

Formally, as shown in Fig. 2, the localization network with two output branches takes as input a 3D CT image to predict a one channel Gaussianlike 3D heatmap H and a three channel offset map O, respectively. The former is centered at the vertebra center points with a small standard deviation  $\delta=3$ voxel-size, while the later indicates the 3D offset vectors of each voxel pointing to its nearest vertebra center.

To localize each vertebra, we first obtain all foreground voxels F from the 3D heatmap H (H > 0.2). Then, for each foreground voxel  $f^i \in F$ , we consider its 3D offset as a vote to the vertebra center and treat the associated 3D heatmap

value as the weight of this vote. Thus, the weighted vote map M is formed by accumulating all the weighted votes. For example, the foreground voxel (0, 0, 1)with an offset vector (0, 1, 0) and a heatmap value 0.7 would contribute to the voxel (0, 1, 1) = (0, 0, 1) + (0, 1, 0) in the weighted vote map with a weight 0.7. Finally, we directly adopt a fast peak search clustering method [14] to find and localize the density peaks in M as the predicted vertebra center points, denoted as V. The rationale is that the clustering vertebrae centers usually have relatively high density values (i.e., weighted votes) and large distance to the nearest voxel with a higher density value, which is defined as:

$$V = (M^i > \delta) \cap (D^i > \lambda), \tag{1}$$

where  $D^i$  refers to the distance between voxel *i* and its nearest voxel with a higher density value than  $M^i$ . The thresholds  $\delta$  and  $\lambda$  are set as 2.0 and 5.0.

To train the localization network, we propose several loss terms to supervise the learning process. Specifically, in the learning of 3D heatmap H and offset O, the smooth L1 loss is employed to calculate the regression error, denoted as  $\mathcal{L}_{H}^{smoothL1}$  and  $\mathcal{L}_{O}^{smoothL1}$ , respectively. In addition, to robustly regress the centers, we introduce the Chamfer distance [3,12] to minimize the bidirectional distance between the candidate center set  $\widehat{C}$  before clustering (i.e., any voxel with a value higher than 2.0 in the weighted vote map) and the ground-truth center set C, defined as:

$$\mathcal{L}_{CD} = \sum_{\hat{c}_i \in \hat{C}} \min_{c_k \in C} ||\hat{c}_i - c_k||_2^2 + \sum_{c_k \in C} \min_{\hat{c}_i \in \hat{C}} ||c_k - \hat{c}_i||_2^2.$$
(2)

Finally, the total loss  $\mathcal{L}_{loc}$  of the localization network is formulated as  $\mathcal{L}_{loc} = \mathcal{L}_{H}^{smoothL1} + \mathcal{L}_{O}^{smoothL1} + \beta \mathcal{L}_{CD}$ , where  $\beta$  is the balancing weight and empirically set to 0.2 for all experiments.

#### 2.2 Vertebrae Identification

As shown in the bottom part of Fig. 2, with all the detected vertebra centers, we further assign the label of each vertebra using an identification network. Considering both global and local information are important for accurate vertebra identification, we model the inter-vertebra relationships via a bidirectional relation module at the global scale, and further introduce a novel continuous vertebra label map to enhance the feature learning in each local vertebra patch.

**Vertebra Proposal Generation.** We generate vertebra proposals guided by the detected vertebra centers. First, we select one vertebra point and crop an image patch from the input 3D CT image I. We then generate an equal-sized center patch using a Gaussian filter centered at the selected vertebra point with a small standard deviation  $\delta = 3$  voxel-size, which serves as a guidance signal. Finally, each image patch is concatenated with the corresponding center patch, yielding a two-channel proposal for the identification network.



**Fig. 3.** The process of continuous label map generation, including (a) vertebra centers with labels, (b) B-spine curve with labels, and (c) continuous label map.

**Bidirectional Relation Module.** For each two-channel vertebra proposal  $p_i \in P$ , we utilize an shared encoder to extract the feature  $f_i$ . Since vertebrae are sequentially connected together from top to bottom and neighboring vertebrae share similar appearance, the contextual clues of the neighboring vertebrae are quite important for accurate identification. Thus, given a proposal  $p_i$ , we utilize the self-attention mechanism to obtain the correlation features  $r_i^{up}$  in upward direction and  $r_i^{down}$  in downward directions to encode the relationship with the other vertebrae in corresponding directions. We take the upward direction as an example to show the calculation of feature  $r_i^{down}$ . Here,  $r_i^{up}$  is a weighted sum of features extracted from upward vertebra proposals defined as  $r_i^{up} = \sum_{k=1}^{i-1} w_{i(k)}^{shape} \cdot f_k$ , where  $w_{i(k)}^{spatial}$  refers to the spatial location weight and  $w_{i(k)}^{shape}$  refers to the shape similarity weight, defined in the following:

$$w_{i(k)}^{spatial} = 1.0 - \frac{\exp(d_k)}{\sum_{j=1}^{i-1} \exp(d_j)}, \qquad w_{i(k)}^{shape} = \frac{\exp(f_k^T f_i)}{\sum_{j=1}^{i-1} \exp(f_j^T f_i)}.$$
 (3)

 $d_k$  measures the distance between centers of the k-th and the i-th proposals, while the shape similarity weight measures the feature similarity. Lastly, we combine the correlation features  $r_i^{up}$  and  $r_i^{down}$  with the proposal feature  $f_i$  to derive the corresponding label of vertebra i by several Fully-Connected (FC) layers.

**Continuous Label Prediction.** The straightforward classification approach with sparse, discrete labels for different proposals receives insufficient supervision and usually leads to inaccurate results. To tackle this problem, we introduce a continuous label map, a novel representation of the vertebra labels, to enhance the network to learn more reliable, fine-grained features. Specifically, as shown in Fig. 3, we first fit a B-spline curve to the 3D coordinates of vertebra centers. Then, we sample a dense set of points on the curve and assign a floating label by linear interpolation of the nearest upward and downward vertebra center labels. Finally, the label values on the curve is mapped to the whole 3D image space based on the nearest neighbor searching. Finally, we add an extra decoder branch to regress the voxel-wise continuous labels in each vertebra patch.

Overall, the training loss  $\mathcal{L}_{id}$  for identification is defined as  $\mathcal{L}_{id} = \mathcal{L}_{cls} + \eta \mathcal{L}_{reg}$ , where  $\mathcal{L}_{cls}$  is the cross-entropy loss for classifying the vertebra labels, and  $\mathcal{L}_{reg}$  is the smooth L1 loss for regressing the continuous label map. The hyper-parameter  $\eta$  is empirically chosen as 0.5 to balance the loss terms.

**Post-label Voting.** Although the identification network already achieves excellent performance, we still observed a few incorrect predictions of the vertebra labels due to insufficient image quality such as motion effects. To correct the prediction error, we employ a voting-based post-processing based on two facts that: 1) vertebra labels always monotonically increase or decrease, and 2) most of our labels are correct. Because every vertebra also implicitly carries the labels of others based on their own labels and adjacency, we finally use the voted label as the resulting label for each vertebra.

# 2.3 Implementation Details

We employed V-Net [11] as the network backbone of our two-stage framework. In the localization network, all CT scans were randomly cropped into the same input size of  $256 \times 256 \times 256$ , and the cropped patch size of the identification network was set as  $96 \times 96 \times 96$  to ensure that the target vertebra can be entirely enclosed. We used the Adam optimizer with an initial learning rate of 0.01 divided by 10 every 5000 iterations. Both networks were trained 20K iterations. It took about 10 h to train the localization network and 12 h for the identification network on a single Tesla M40 GPU.

# 3 Experiments

# 3.1 Dataset and Evaluation Metric

We have extensively evaluated our framework on 1000 chest CT images collected from real-world clinics which contain many cases with severe pathological spine problems. These scans have been pre-processed with isotropic resolution of  $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ . All vertebra centers and labels are annotated by experienced radiologists. We randomly split the 1000 scans into 600 for training, 100 scans for validation, and the remaining 300 scans for testing.

To quantitatively evaluate the performance of our framework, we first measure the localization error (mm) as the distance of each predicted vertebra center to its nearest ground-truth vertebra center. As for the vertebra labeling task, we compute the identification accuracy (%) at the vertebra and patient levels (denoted as  $Id_{acc}$ -V and  $Id_{acc}$ -P, respectively). The former measures the percentage of the correctly identified vertebrae per-patient, while the later measures the percentage of the patients whose vertebrae are all correctly identified. For both networks, we also reported the specific metrics of cervical vertebrae (Cer.), thoracic vertebrae (Tho.), lumbar vertebrae (Lum.), as shown in Tables 1 and 2.

| Method         | Localizati    | on [mm] $\downarrow$ | Identification [%] $\uparrow$ |               |                                       |               |
|----------------|---------------|----------------------|-------------------------------|---------------|---------------------------------------|---------------|
|                | Cer.          | Tho.                 | Lum.                          | All           | $\mathrm{Id}_{acc}\text{-}\mathrm{V}$ | $Id_{acc}$ -P |
| $bNet_{vl}$    | $3.1\pm3.1$   | $3.8 \pm 1.9$        | $3.3\pm2.5$                   | $3.3\pm2.3$   | 82.1                                  | 51.3          |
| $bNet_{vl}$ -W | $1.7 \pm 1.1$ | $2.4 \pm 1.5$        | $2.1 \pm 1.6$                 | $2.0 \pm 1.3$ | 86.4                                  | 57.3          |

Table 1. Quantitative vertebra localization results of alternative networks. The identification task is performed using the baseline network  $bNet_{id}$ .

**Table 2.** Quantitative vertebra identification results of alternative networks. The higher the percentage value (%), the better the identification accuracy ( $\uparrow$ ).

| Method                                   | Cer. | Tho. | Lum. | $Id_{acc}-V$ | $Id_{acc}$ -P |
|--|------|------|------|--------------|---------------|
| $bNet_{id}$                              | 87.2 | 86.0 | 86.9 | 86.4         | 57.3          |
| $bNet_{id}$ -R                           | 91.3 | 89.0 | 90.2 | 90.5         | 82.3          |
| $\mathrm{bNet}_{id}\text{-}\mathrm{R-C}$ | 96.8 | 94.7 | 96.1 | 96.1         | 92.3          |
| FullNet                                  | 99.1 | 98.6 | 98.7 | 98.9         | 98.7          |

#### 3.2 Ablation Study of Key Components

We conduct ablative experiments to demonstrate the effectiveness of the proposed components. We first build the baseline networks for the vertebra localization and identification tasks, denoted as  $bNet_{vl}$  and  $bNet_{id}$ , respectively.  $bNet_{vl}$ directly detects vertebrae using a 3D heatmap, while  $bNet_{id}$  simply builds upon a multi-label classification network without the bidirectional relation module and continuous label prediction in our identification task. All alternative networks are derived from the baseline networks by augmenting different components.

Benefits of Weighted-Voting Scheme. Unlike the direct heatmap regression, we utilize the weighted-voting scheme by combining a 3D offset map with the 3D heatmap with the Chamfer loss as an additional supervision. For validation, we augment the baseline bNet<sub>vl</sub> with an extra output branch to generate 3D offsets followed by the post-clustering, denoted as  $bNet_{vl}$ -W. The quantitative results are shown in Table 1. Compared to  $bNet_{vl}$ ,  $bNet_{vl}$ -W consistently improves the localization and identification accuracy. The mean and variance of the localization error are decreased, leading to more robust results. For cervical vertebrae smaller in size and tightly packed, benefit of the weighted-voting scheme is much clearer (reduced by 1.4 mm). This results in an increase of 4.3% (and 6.0%) in Id<sub>acc</sub>-V (and Id<sub>acc</sub>-P) with bNet<sub>id</sub>.

**Benefits of Bidirectional Relation Module.** We utilize the  $bNet_{vl}$ -W as the localization network and augment the identification baseline network ( $bNet_{id}$ ) with the bidirectional relation module, denoted as  $bNet_{id}$ -R. Table 2 shows that the  $bNet_{id}$ -R consistently improves the identification accuracy, 25.0% in  $Id_{acc}$ -P, illustrating high efficacy of this module for the sequential prediction task.



**Fig. 4.** Comparison between our results (blue) and those by Deep-HMM (yellow) against the ground-truth (GT) (red). Seven typical examples are presented: metal artifacts (1, 7), pathological spines (2, 3), and limited field of view (4, 5, 6, 7). The GT label is annotated if incorrect prediction occurs. (Color figure online)

 Table 3. Quantitative comparison with state-of-the-art methods.

| Method         | Localization                    | $\mathrm{Id}_{acc}\text{-}\mathrm{V}$ | $\mathrm{Id}_{acc}\text{-}\mathrm{P}$ |
|----------------|---------------------------------|---------------------------------------|---------------------------------------|
| J-CNN [1]      | $7.6\pm12.4$                    | 86.7                                  | 60.7                                  |
| ML-LSTM $[13]$ | $2.7\pm2.8$                     | 90.0                                  | 79.3                                  |
| Deep-HMM $[2]$ | $2.5\pm2.3$                     | 95.2                                  | 89.3                                  |
| Ours           | $\textbf{2.0} \pm \textbf{1.3}$ | 98.9                                  | 98.7                                  |

**Benefits of Continuous Label Prediction.** We augment  $bNet_{id}$ -R with this extra task, denoted as  $bNet_{id}$ -R-C. As shown in Table 2, compared to the  $bNet_{id}$ -R,  $Id_{acc}$ -V improves significantly from 90.5% to 96.1%. This shows the task of continuous label prediction can effectively facilitate the network to learn more discriminative features for fine-grained classification.

**Benefits of Post-label Voting.** In our FullNet (*VertNet*), a voting-based postprocessing step is added on top of bNet<sub>id</sub>-R-C, to generate consistent and correct labels. As presented in Table 2, FullNet obtains the best performance. Notably, the Id<sub>acc</sub>-P is significantly boosted to 98.7%, which suggests the potential applicability of our framework in real-world clinical scenarios. A set of typical visual results are shown in Fig. 4 with centers projected to the specific CT slice.

# 3.3 Comparison with State-of-the-Art Methods

We implemented three state-of-the-art vertebra localization and identification methods for comparison, including a one-stage method (J-CNN [1]) and two-stage methods (ML-LSTM [13] and Deep-HMM [2]). Instead of using the CSI

2014 dataset, we train and test all the networks on our newly collected dataset which has a larger sample size and contains more challenging cases. Table 3 shows ML-LSTM and Deep-HMM outperform J-CNN by a large margin. Compared with ML-LSTM, our network with the bidirectional relation module and continuous label prediction generates more reliable vertebra labels. Compared to our bidirectional module for modeling the global relationship among vertebrae, Deep-HMM employs the Markov modeling of vertebra labels which is limited to short-range relationships. As a result, our method significantly outruns DeepHMM in the identification accuracy (3.7% and 9.4% increases on Id<sub>acc</sub>-V and Id<sub>acc</sub>-P, respectively). Figure 4 shows qualitative results between ours (blue) and those by DeepHMM (yellow) against GT labels (red) on challenging cases with metal artifacts, pathological spines, or limited field of views, further supporting our design choices.

# 4 Conclusion

We investigated the problem of vertebra localization and labeling from CT images. The proposed two-stage framework accurately detects all the vertebra centers and successfully identifies all the vertebra labels with satisfactory high accuracy. We qualitatively and quantitatively evaluated our method on our representative clinical dataset and compared against the state-of-the-art approaches. The superior performance suggests the potential applicability of our framework in real-world clinical scenarios.

# References

- Chen, H., et al.: Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 515–522. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24553-9\_63
- Chen, Y., Gao, Y., Li, K., Zhao, L., Zhao, J.: vertebrae identification and localization utilizing fully convolutional networks and a hidden Markov model. IEEE Trans. Med. Imaging 39(2), 387–399 (2019)
- Cui, Z., et al.: TSegNet: an efficient and accurate tooth segmentation network on 3D dental model. Med. Image Anal. 69, 101949 (2021)
- 4. Cui, Z., Li, C., Wang, W.: Toothnet: automatic tooth instance segmentation and identification from cone beam ct images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6368–6377 (2019)
- 5. Haldeman, S., et al.: The global spine care initiative: classification system for spinerelated concerns. Eur. Spine J. **27**(6), 889–900 (2018)
- Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. 13(3), 471–482 (2009)
- Knez, D., Likar, B., Pernuš, F., Vrtovec, T.: Computer-assisted screw size and insertion trajectory planning for pedicle screw placement surgery. IEEE Trans. Med. Imaging 35(6), 1420–1430 (2016)

- Kumar, R.: Robotic assistance and intervention in spine surgery. In: Li, S., Yao, J. (eds.) Spinal Imaging and Image Analysis. LNCVB, vol. 18, pp. 495–506. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-12508-4\_16
- Létourneau, D., et al.: Semiautomatic vertebrae visualization, detection, and identification for online palliative radiotherapy of bone metastases of the spine. Med. Phys. 35(1), 367–376 (2008)
- Liao, H., Mesfin, A., Luo, J.: Joint vertebrae identification and localization in spinal CT images by combining short-and long-range contextual information. IEEE Trans. Med. Imaging 37(5), 1266–1275 (2018)
- Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571. IEEE (2016)
- Qi, C.R., Su, H., Mo, K., Guibas, L.J.: PointNet: deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017
- Qin, C., Yao, D., Zhuang, H., Wang, H., Shi, Y., Song, Z.: Residual block-based multi-label classification and localization network with integral regression for vertebrae labeling. arXiv preprint arXiv:2001.00170 (2020)
- Rodriguez, A., Laio, A.: Clustering by fast search and find of density peaks. Science 344(6191), 1492–1496 (2014)
- 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
- Sekuboyina, A., Rempfler, M., Valentinitsch, A., Kirschke, J.S., Menze, B.H.: Adversarially learning a local anatomical prior: vertebrae labelling with 2D reformations. arXiv preprint arXiv:1902.02205 (2019)
- Wang, F., Zheng, K., Lu, L., Xiao, J., Wu, M., Miao, S.: Automatic vertebra localization and identification in CT by spine rectification and anatomically-constrained optimization. arXiv preprint arXiv:2012.07947 (2020)
- Yang, D., et al.: Deep image-to-image recurrent network with shape basis learning for automatic vertebra labeling in large-scale 3D CT volumes. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 498–506. Springer, Cham (2017). https://doi.org/10. 1007/978-3-319-66179-7\_57
- 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