



**HAL**  
open science

# Vertebrae localization, segmentation and identification using a graph optimization and an anatomic consistency cycle

Di Meng, Eslam Mohammed, Edmond Boyer, Sergi Pujades

## ► To cite this version:

Di Meng, Eslam Mohammed, Edmond Boyer, Sergi Pujades. Vertebrae localization, segmentation and identification using a graph optimization and an anatomic consistency cycle. Chunfeng Lian; Xiaohuan Cao; Islem Rekik; Xuanang Xu; Zhiming Cui. *MLMI 2022: Machine Learning in Medical Imaging*, 13583, Springer, pp.307-317, 2022, Lecture Notes in Computer Science, 978-3-031-21013-6. 10.1007/978-3-031-21014-3\_32 . hal-03779423

**HAL Id: hal-03779423**

**<https://inria.hal.science/hal-03779423>**

Submitted on 16 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Vertebrae localization, segmentation and identification using a graph optimization and an anatomic consistency cycle

Di Meng, Eslam Mohammed, Edmond Boyer, and Sergi Pujades

Inria, Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, France  
di.meng@inria.fr

**Abstract.** Vertebrae localization, segmentation and identification in CT images is key to numerous clinical applications. While deep learning strategies have brought to this field significant improvements over recent years, transitional and pathological vertebrae are still plaguing most existing approaches as a consequence of their poor representation in training datasets. Alternatively, proposed non-learning based methods take benefit of prior knowledge to handle such particular cases. In this work we propose to combine both strategies. To this purpose we introduce an iterative cycle in which individual vertebrae are recursively localized, segmented and identified using deep networks, while anatomic consistency is enforced using statistical priors. In this strategy, the transitional vertebrae identification is handled by encoding their configurations in a graphical model that aggregates local deep-network predictions into an anatomically consistent final result. Our approach achieves the state-of-the-art results on the VerSe20 challenge benchmark, and outperforms all methods on transitional vertebrae as well as the generalization to the VerSe19 challenge benchmark. Furthermore, our method can detect and report inconsistent spine regions that do not satisfy the anatomic consistency priors. The code and model are available for research purposes. ([https://gitlab.inria.fr/spine/vertebrae\\_segmentation](https://gitlab.inria.fr/spine/vertebrae_segmentation))

**Keywords:** Spine · Vertebrae segmentation · Vertebrae labelling · Transitional vertebrae.

## 1 Introduction

Segmenting, identifying and localizing vertebrae in CT images play an essential role in many clinical applications [16]. These tasks are intrinsically interdependent and a common issue comes from inconsistencies that can propagate between them. For instance neighboring vertebrae share similar shapes which makes their identification uncertain. Moreover pathological spines can present abnormal shapes or the number of transitional vertebrae can be different among patients. These transitional vertebrae, *i.e.*, the absence of T12 or occurrence of T13 or L6, are common and are reported to affect between 15% and 35% of the general population [32]. However, they only impact one vertebrae among the

24 in a spine and their occurrence percentages in standard vertebrae datasets appear to be therefore much lower, down to a few percentages in practice. As a consequence, the current state-of-the-art methods, which are mostly based on deep networks trained over standard datasets, tend to experience important performance drops in the presence of transitional vertebrae.

Besides augmenting datasets, an alternative solution is to exploit prior knowledge on the full structure of the spine, as initially introduced in non-learning based methods. In this work, we investigate the combination of such a strategy with a deep learning approach. Precisely we propose to iteratively cycle between the localization, segmentation and identification tasks, which uses deep networks, while enforcing anatomical consistency with statistical priors. On one hand, the priors are used to localize vertebrae: they leverage learned statistics of vertebrae volume and inter-vertebral distances which exhibit more robustness to pathological cases than shape or appearance models. On the other hand, in order to handle transitional vertebrae in the identification, we propose to encode the admissible configurations in a graphical model. Such a model leverages local deep-network predictions and aggregates them into an anatomically consistent result. Our experiments demonstrate that this strategy successfully handles local inaccurate predictions. Therefore it performs better than other methods with transitional vertebrae while providing state-of-the-art results on standard benchmarks such as VerSe20.

To solve for vertebrae localization, segmentation and identification, traditional methods use shape [5] or graphical models [10, 4]. Surface [15, 2, 5, 23] or image based appearance models [10, 4], which make assumptions on exact shapes, are not robust to pathological cases (such as fractures or metal implants). However relative positions of vertebrae [10, 11, 33] are less sensitive to these cases. In our work we propose to learn statistics on vertebrae relative positions and volumes. The volume quantifies a *part of the vertebra shape* that is robust to morphological deviations. Furthermore we learn statistics conditioned on the spine level (cervical, thoracic, lumbar), increasing hence their performance. Moreover we relax the need for a fixed threshold to detect vertebrae [25] and compute adaptative thresholds that automatically vary across spine levels and patients.

For the segmentation task, existing methods use primarily statistical models [15, 26], active shape models [12, 3, 1], graph cuts [2], or level sets [31]. Several methods achieve very good performances by assuming that a specific part of the spine is observed [26, 14], yet making them sensitive to arbitrary field of views. For this task, recent learning based approaches have demonstrated the best performances [17, 25, 21] with a local approach. Our segmentation strategy builds on the segmentation network of Payer *et al.* [25] and the refinement strategy of Lessman *et al.* [17].

As for the vertebrae identification task, fully automatic methods rely on machine learning and convolutional neural networks (CNN) in which hand-crafted features [11, 7] tend to be replaced by learned ones. Several methods [22, 13, 25, 7] use a two-step approach that first detects the vertebrae before identifying them. Other networks, such as Btrfly [28], bidirectional recurrent network

(Bi-RNN) [18], FCN [8] or the Spine-transformer [30] output directly vertebrae locations and labels simultaneously. All these methods have difficulties with transitional vertebrae as a consequence of their sparsity in training data. Another group of methods proposes to enforce the identification consistency as a post process by using chain-structure graphical models [34], conditional random fields [20] or Hidden Markov model [8]. They anyway do not explicitly model transitional vertebrae. In our work, we take a two-step approach to localize and identify vertebrae. To identify individual vertebra we build on Mohammed *et al.* [24] and use the segmented vertebrae shape instead of the raw CT image. In addition we also contribute a new graphical model to select a global optimal configuration among all the individual predictions, within which the transitional vertebrae are explicitly modeled.

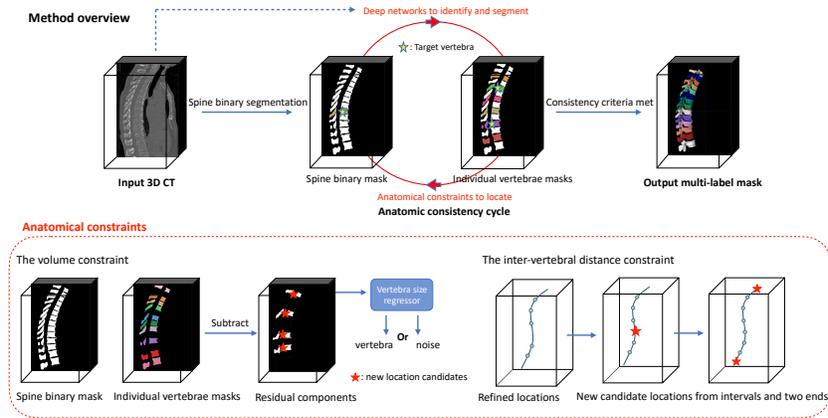
## 2 Method

### 2.1 Anatomic Consistency Cycle

We propose to cycle through the three tasks of localization, segmentation and identification in order to enforce anatomic consistencies among them. Our method starts by segmenting a *spine mask* (details in Sup Mat.) which enables to locate individual vertebrae (Sec. 2.2). From the obtained locations, individual vertebrae segmentation masks are estimated with an individual vertebrae segmentor [25] and refined with an iterative location-segmentation refinement scheme [17]. Vertebrae are further identified using the locations and segmentation masks (Sec. 2.3). The obtained identifications are in turn used to enforce finer anatomic consistency constraints (Sec. 2.2) and detect possible new candidate locations. In the latter case, the new vertebrae go through the segmentation and identification steps described before. The process ends when the proposed consistency criteria are satisfied. It also stops when the set of detected locations, segmentation masks and identifications does not change over the cycle. The remaining inconsistencies, such as a failure in the location-segmentation refinement or an inconsistency in the anatomic constraints, are reported in the result. Figure 1 illustrates the method overview.

### 2.2 Anatomic consistency constraints for vertebrae localization

Vertebrae are naturally ordered in the spine and their relative locations and consecutive sizes are heavily correlated. With the objective to exploit such prior information we compute statistics over inter-vertebral distances and individual vertebrae volumes. Given these priors we enforce two consistency criteria. First, distances between all detected locations should follow the inter-vertebral distances statistics. Second, the *spine mask* should be similar to the union of individual vertebrae masks. The segmentation of a *spine mask* is, in general, a straightforward task which in turn allows the areas where individual vertebrae should lie to be identified. In practice, at each cycle the difference between the

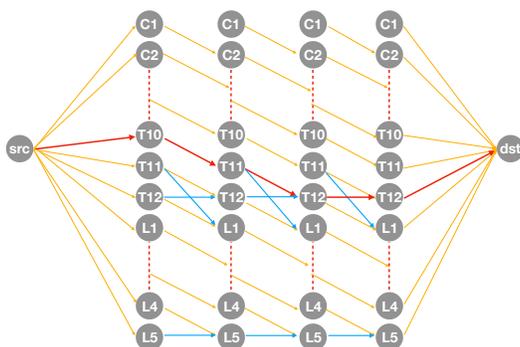


**Fig. 1.** The method overview. Given a 3D CT as input, the spine is segmented as a reference to locate the individual vertebrae. The anatomical constraints are leveraged with deep networks for localization and segmentation. Once the location is stable, its identification is obtained given its segmentation. The set of location, segmentation and identification is cycled through until the consistency criteria are met.

*spine mask* and all individual vertebrae masks is computed to obtain residual connected components. (The residual connected component is the whole *spine mask* at the first cycle.) Vertebrae volume statistics are then considered to decide whether a residual is a vertebra or just noise.

**The vertebrae volume constraint.** Given a residual connected component, the idea is to check its volume by considering the volumes of the neighboring vertebrae. To this purpose a regressor is trained over consecutive vertebrae to predict the volume of the previous or next vertebrae along the spine. If the residual volume size is larger than a fraction of the predicted size (50% in our experiments), it is considered as a vertebra and its location at its mass center is added to the vertebrae list. The residual is otherwise discarded as noise.

**The inter-vertebral distance constraint.** This constraint builds on the fact that distances between locations are well structured. To this aim, we use two statistical models of the distance between vertebrae. The first is a Gaussian distribution for each anatomic group capturing the mean and variance of the distances between consecutive vertebrae. It is used as a prior to detect abnormally large distances. In addition, linear regressors conditioned on each anatomic group are used to predict the inter-vertebral distance from the neighbouring ones. When a larger inter-vertebral distance is detected, new candidate locations are added in between. The number of new candidates is adapted depending on the current distance and the predicted one. We also leverage identification to check if the spine extremes are complete. If C1 (or L5) is not yet found, a location is predicted up (or down) using the two most top (down) locations. If it is inside the image field of view, it is added.



**Fig. 2.** Spine identification global graph. An example with 4 vertebrae. Edges in orange are regular connections between nodes. Edges in blue connecting T12 to T12, and L5 to L5, allow the presence of T13 and L6. The edge between T11 and L1 allows T12 to be absent. The result with red edges corresponds to a spine where T10, T11, T12 and T13 are observed. (Best viewed in color)

### 2.3 Vertebrae identification

To identify vertebrae we propose to combine both a local and global reasoning. The local reasoning is based on an individual classification network that predicts one of the 24 possible labels for each vertebra. On top of it, the global reasoning assumes that consecutive vertebrae are *sorted* in the spine and aggregates the local predictions into a global, consistent identification.

**Individual vertebra classification.** Our individual vertebra classification module builds on [24] where a 3D VGG network is used to classify a single vertebra into one of the 24 categories. The input of the network is a binary volume of size  $128 \times 128 \times 128$  that contains a vertebra segmentation mask. In our work, we extend this method [24] by including the neighboring vertebrae to the input, while keeping the target vertebra in the center of the cube, which effectively adds more context for the prediction. At inference time, two predictions are performed: one with the *spine mask* and a second with the union of all individual vertebrae segmentations. Both predictions are averaged and used as input to the graph optimization. When an individual segmentation mask is empty, *e.g.* metal implants can disturb the individual segmentor, a uniform distribution over all vertebrae is used. In order to deal with transitional vertebrae, such as T13 and L6, for which only few samples appear in the training set - 1 T13 and 13 L6 in practice - we consider T13 as T12 and L6 as L5 and leave their fine identification to the global reasoning.

**Global graph optimization.** Some vertebrae can still be misidentified by the local approach. In order to improve the classification, we propose a global strategy that enforces the natural ordering of vertebrae, *i.e.* consecutive vertebrae should have consecutive labels. We cast the problem as a shortest path search in a graph which identifies an optimal set of labels, as illustrated in Fig. 2.

Given a list with  $n$  consecutive locations, a graph with  $n \times 24$  nodes is created. Each of the  $n$  columns represents a located vertebra, and each of the 24 rows represents a vertebra class. Both a unary and a binary costs are then considered for the optimal path search. We populate the unary cost of each node with the cost obtained from the individual local probability. Additionally, as the 3-group predictions are robust, we also add a cost to each unary to prevent group swapping. The binary costs are defined on the edges between the nodes in the graph. These edges encode the fact that two consecutive vertebrae need to have consistent classes. In practice, node  $N_i^j$  with  $i \in [1, n]$  and  $j \in [1, 24]$  is only connected with  $N_{i+1}^{j+1}$ . These edges are shown in orange in Fig. 2. They effectively enforce that consecutive vertebrae get consecutive labels.

Three configurations require special attention in this process, which correspond to transitional vertebrae: i) The presence of T13, ii) The lack of T12 and iii) The presence of L6. To handle the presence of T13, an edge is added between two consecutive T12 nodes. For the lack of T12, an edge is added between two consecutive nodes belonging to T11 and L1. The presence of L6 is dealt with an edge between two consecutive L5 nodes. These three special edges are shown in blue in Fig. 2 and are given a higher cost. To complete the graph, two extra nodes with no cost are added that are used as the source (*src*) and the destination (*dst*) in the optimization. The *src* node is connected to all nodes of the first vertebrae and the *dst* node is connected to all nodes of the last vertebrae.

Once the graph is created and the costs are populated, we compute the shortest path in the graph using the classical Dijkstra algorithm [9]. As a post process, we check if repeated instances of T12 (or L5) are obtained and adjust the class of the second node to T13 (or L6) accordingly.

### 3 Experiments

We evaluate our method using the VerSe20 MICCAI challenge dataset [19, 29, 27], which is currently the largest spine dataset. It is split into a public training set, a public testset and a hidden testset. Each set has 100 CT scans. The challenge has closed so the annotation of the hidden testset is also released for the public. We follow the challenge setting and use the training data to set up our method: 80 random CT scans from the training set are used to train the networks and estimate the anatomical prior statistics, the 20 left for validation. We first evaluate the novel graphical model performance and then we quantitatively evaluate our full method on the public and hidden testsets.

**Vertebrae identification.** To evaluate the benefit of using the context and the graph (Sec. 2.3), we compare four identification settings: Using a single vertebra binary mask [24] (a), adding the neighbouring vertebrae (b) and adding the neighbouring vertebrae and using the graph (c). To compare to image based identification approaches we trained a classification method using the CT image as input (d). For a) the identification accuracy is 70.50%. When including the neighboring vertebrae masks, the accuracy is increased to 84.55%. After the graph optimization, 97.36% accuracy is obtained. When using the CT image

	public testset				hidden testset			
	labelling results		segmentation results		labelling results		segmentation results	
	ID rate(%)	MLD(mm)	DSC(%)	HD(mm)	ID rate(%)	MLD(mm)	DSC(%)	HD(mm)
Zhang A.	94.93	2.99	88.82	7.62	96.22	2.59	89.36	7.92
Payer C. [25]	95.06	2.90	91.65	<b>5.80</b>	92.82	2.91	89.71	<b>6.06</b>
Chen D. [6]	95.61	<b>1.98</b>	91.72	6.14	<b>96.58</b>	<b>1.38</b>	<b>91.23</b>	7.15
ours	<b>96.59</b>	2.21	<b>92.53</b>	7.03	95.38	2.43	91.11	6.69
ours(a)	96.02	2.22	92.06	6.82	94.93	2.34	90.85	6.53
ours(b)	93.61	3.10	89.76	7.80	89.41	3.81	85.45	8.39

**Table 1.** Quantitative comparison of our method with the top ranked methods on the VerSe20 challenge[27] as well as with alternative (a) in which the anatomic consistency constraints are not conditioned on the identification and alternative (b) where the anatomic consistency priors are not used for detection.

patch as input instead of its mask, a lower identification accuracy (82.36%) is obtained compared to the binary mask of the vertebrae (84.55%). We believe that this is due to the fact that the masks provide shape discriminative cues on the vertebrae.

**VerSe20 challenge results.** We quantitatively evaluate our method on the VerSe20 Challenge public and hidden testsets with 200 CT subjects in total and adopted the metrics of the challenge evaluation protocol [27]: The *Dice coefficient (DSC)* and the *Hausdorff Distance (HD)* evaluate the correctly identified segmentation performance in terms of voxel and surface similarity respectively; The *Identification rate (ID rate)* and the *Mean Localization Distance (MLD)* evaluate the accuracy of the labelling task, where MLD measures the Euclidean distance of the predicted location to the GT location and 20mm is the criterion for a valid identification. In the challenge metrics, when a vertebra is missed, MLD and HD are not computed. Thus there is a trade off between ID rate and MLD (same for DSC and HD). While more positive detections lead to a better ID rate, a worse MLD can be obtained if the detections are less accurate.

Table 1 shows the results of the proposed method and the top scoring methods. Our method is on par with the best performing method in the leaderboard [6] while being significantly more robust to transitional cases, as shown in Table 2, which is our main objective. In that respect, we observe that Chen D. et al. [6] did not win the challenge, as one important criterion in the challenge was the performance in handling the transitional vertebrae - 6 cases with T13 (2/2/2 in Train/Public/Hidden), 47 cases with L6 (15/15/17) and 8 cases with absent T12(3/4/1). Following the challenge guidelines we computed the Dice score only on scans with transitional vertebrae and the obtained results are presented in Table 2. Our method consistently outperforms all existing methods.

To evaluate the generalization of our method, we perform the experiment of the VerSe20 challenge [27] where the approach developed and trained on the VerSe20 training set is applied to the VerSe19 hidden testset. Table 3 shows that our method outperforms all approaches and generalizes well to VerSe19.

	public testset	hidden testset
ours	<b>91.04</b>	<b>89.70</b>
Payer C. [25]	85.96	89.59
Zhang A.	87.15	87.35
Chen D. [6]	84.21	87.01

**Table 2.** Methods evaluation (Dice score %) on transitional vertebrae. [27]

	DSC(%)
ours	<b>90.84</b>
Chen D. [6]	86.44
Payer C. [25]	84.11
Zhang A.	85.42

**Table 3.** Generalization performance on VerSe19 hidden testset. [27]

Inference times on public testset vary between 1.5 mins to 248 mins depending on the scan size, with a median runtime of 26 mins. Most of the computation time is devoted to the spine segmentation, which could be drastically improved by considering parallelization over patches.

Qualitative results and typical failure cases of our method are provided in the supplementary material.

**Ablation study.** To highlight the benefits of the different parts of the proposed method we perform two experiments. In the first one (*ours(a)* in Tab. 1) we do not use the identification to condition the vertebrae statistics. Instead of learning specialized regressors for each spine level, we use a fixed threshold to determine if a residual mask is a potential true vertebra (we use half of smallest vertebra volume size in the training set:  $7820/2$ ). To decide if there is a missing vertebra given a distance between 2 we use a fixed threshold of 50mm [25]. The overall scores show that the improvement when using conditional statistics is minor but not negligible: without conditioning 8 vertebrae are missed in the public testset and 6 in the hidden test set while with the identification information no vertebra is missed in public testset and only 4 in hidden testset. In the second experiment (*ours(b)* in Tab. 1) the anatomic consistency priors (Sec. 2.2) are not used. Without the constraints on vertebrae volume and the relative locations the performance significantly drops.

## 4 Conclusion

In this work we presented a new approach to localize, segment and identify vertebrae in CT images. An anatomic consistency cycle is proposed that aggregates task-specific deep networks while enforcing statistical priors. As seen with the generalization on the VerSe19 dataset, it improves robustness in the results. In addition, we also present a graphical model that enforces global consistency among individual predictions, hence being able to handle transitional vertebrae as well as pathological cases. The evaluation on the standard VerSe20 challenge benchmark demonstrates the interest of the proposed strategy, in particular with transitional vertebrae which are present in a meaningful part of the population. The extension of the proposed anatomic consistency cycle to other anatomic structures with similar specific cases is currently under investigation.

## References

1. Al Arif, S.M.R., Gundry, M., Knapp, K., Slabaugh, G.: Improving an active shape model with random classification forest for segmentation of cervical vertebrae. In: International Workshop on Computational Methods and Clinical Applications for Spine Imaging. Springer (2016)
2. Aslan, M.S., Ali, A., Chen, D., Arnold, B., Farag, A.A., Xiang, P.: 3d vertebrae segmentation using graph cuts with shape prior constraints. In: ICIP. IEEE (2010)
3. Benjelloun, M., Mahmoudi, S., Lecron, F.: A framework of vertebra segmentation using the active shape model-based approach. IJBI (2011)
4. Bromiley, P.A., Kariki, E.P., Adams, J.E., Cootes, T.F.: Fully automatic localisation of vertebrae in ct images using random forest regression voting. In: International Workshop on Computational Methods and Clinical Applications for Spine Imaging. Springer (2016)
5. Cai, Y., Osman, S., Sharma, M., Landis, M., Li, S.: Multi-modality vertebra recognition in arbitrary views using 3d deformable hierarchical model. TMI (2015)
6. Chen, D., Bai, Y., Zhao, W., Ament, S., Gregoire, J., Gomes, C.: Deep reasoning networks for unsupervised pattern de-mixing with constraint reasoning. In: ICML. PMLR (2020)
7. Chen, H., Shen, C., Qin, J., Ni, D., Shi, L., Cheng, J.C., Heng, P.A.: Automatic localization and identification of vertebrae in spine ct via a joint learning model with deep neural networks. In: MICCAI. Springer (2015)
8. Chen, Y., Gao, Y., Li, K., Zhao, L., Zhao, J.: vertebrae identification and localization utilizing fully convolutional networks and a hidden markov model. TMI (2019)
9. Dijkstra, E.W., et al.: A note on two problems in connexion with graphs. Numerische mathematik (1959)
10. Glocker, B., Feulner, J., Criminisi, A., Haynor, D.R., Konukoglu, E.: Automatic localization and identification of vertebrae in arbitrary field-of-view ct scans. In: MICCAI. Springer (2012)
11. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae localization in pathological spine ct via dense classification from sparse annotations. In: MICCAI. Springer (2013)
12. Graham, J., Cooper, D., Taylor, C., Cootes, T.: Active shape models their training and applications. CVIU (1995)
13. Jakubicek, R., Chmelik, J., Jan, J., Ourednicek, P., Lambert, L., Gavelli, G.: Learning-based vertebra localization and labeling in 3d ct data of possibly incomplete and pathological spines. Computer methods and programs in biomedicine
14. Janssens, R., Zeng, G., Zheng, G.: Fully automatic segmentation of lumbar vertebrae from ct images using cascaded 3d fully convolutional networks. In: ISBI (2018)
15. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in ct images. MedIA (2009)
16. Knez, D., Likar, B., Pernuš, F., Vrtovec, T.: Computer-assisted screw size and insertion trajectory planning for pedicle screw placement surgery. TMI (2016)
17. Lessmann, N., Van Ginneken, B., De Jong, P.A., Išgum, I.: Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. MedIA (2019)

18. Liao, H., Mesfin, A., Luo, J.: Joint vertebrae identification and localization in spinal ct images by combining short-and long-range contextual information. *TMI* (2018)
19. Löffler, M.T., Sekuboyina, A., Jacob, A., Grau, A.L., Scharr, A., El Hussein, M., Kallweit, M., Zimmer, C., Baum, T., Kirschke, J.S.: A vertebral segmentation dataset with fracture grading. *Radiology: Artificial Intelligence* (2020)
20. Mader, A.O., Lorenz, C., von Berg, J., Meyer, C.: Automatically localizing a large set of spatially correlated key points: a case study in spine imaging. In: *MICCAI*. Springer (2019)
21. Masuzawa, N., Kitamura, Y., Nakamura, K., Iizuka, S., Simo-Serra, E.: Automatic segmentation, localization, and identification of vertebrae in 3d ct images using cascaded convolutional neural networks. In: *MICCAI*. Springer (2020)
22. McCouat, J., Glocker, B.: Vertebrae detection and localization in ct with two-stage cnns and dense annotations. *arXiv preprint arXiv:1910.05911* (2019)
23. Meng, D., Keller, M., Boyer, E., Black, M., Pujades, S.: Learning a statistical full spine model from partial observations. In: *International Workshop on Shape in Medical Imaging*. Springer (2020)
24. Mohammed, E., Meng, D., Pujades, S.: Morphology-based individual vertebrae classification. In: *International Workshop on Shape in Medical Imaging*. Springer (2020)
25. Payer, C., Stern, D., Bischof, H., Urschler, M.: Coarse to fine vertebrae localization and segmentation with spatialconfiguration-net and u-net. In: *VISIGRAPP (5: VISAPP)* (2020)
26. Rasouljan, A., Rohling, R., Abolmaesumi, P.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+ pose model. *TMI* (2013)
27. Sekuboyina, A., Hussein, M.E., Bayat, A., Löffler, M., Liebl, H., Li, H., Tetteh, G., Kukačka, J., Payer, C., Štern, D., et al.: Verse: A vertebrae labelling and segmentation benchmark for multi-detector ct images. *MedIA* (2021)
28. Sekuboyina, A., Rempfler, M., Kukačka, J., Tetteh, G., Valentinitich, A., Kirschke, J.S., Menze, B.H.: Btrfly net: Vertebrae labelling with energy-based adversarial learning of local spine prior. In: *MICCAI*. Springer (2018)
29. Sekuboyina, A., Rempfler, M., Valentinitich, A., Menze, B.H., Kirschke, J.S.: Labeling vertebrae with two-dimensional reformations of multidetector ct images: An adversarial approach for incorporating prior knowledge of spine anatomy. *Radiology: Artificial Intelligence* (2020)
30. Tao, R., Liu, W., Zheng, G.: Spine-transformers: Vertebra labeling and segmentation in arbitrary field-of-view spine cts via 3d transformers. *MedIA* (2022)
31. Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W.E., Willsky, A.: A shape-based approach to the segmentation of medical imagery using level sets. *TMI* (2003)
32. Uçar, D., Uçar, B.Y., Coşar, Y., Emrem, K., Gümüşsuyu, G., Mutlu, S., Mutlu, B., Çaçan, M.A., Mertsoy, Y., Gümüş, H.: Retrospective cohort study of the prevalence of lumbosacral transitional vertebra in a wide and well-represented population. *Arthritis* (2013)
33. 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. In: *Proceedings of the IEEE/CVF CVPR* (2021)
34. Yang, D., Xiong, T., Xu, D., Huang, Q., Liu, D., Zhou, S.K., Xu, Z., Park, J., Chen, M., Tran, T.D., et al.: Automatic vertebra labeling in large-scale 3d ct using deep image-to-image network with message passing and sparsity regularization. In: *IPMI*. Springer (2017)