Localization of Intervertebral Discs Using Deep-Learning and Region Growing Technique

Sujata Satpute[1(✉)], Ramesh Manza[1], Ganesh Manza[1], and Anjum Shaikh[2]

[1] Dr. Babasaheb Ambedkar, Marathwada University, Aurangabad, India
sujatasatpute7058@gmail.com
[2] Deogiri College, Aurangabad, India

Abstract. Detection and Marking of Intervertebral discs (IVD) of the spinal cord is relevant as it notably enables experts to diagnose spinal cord injury. Many of the experts from medical field do this task manually, therefore there may be risk of wrong labeling of in-vertebral disks and this job is tedious. There are several automated methods are already implemented for CT-SCAN images and MRI images as well. Most of the methods are not freely available and the existing methods fails if the image quality fluctuates. There is another factor that affects the localization go wrong when the algorithms for localization fails to hit discs or it has false positive detection. In this paper we adopted Fully Convolutional Network (FCN), Stacked Hourglass Network with Multi-level Attention Mechanism and region growing technique for vertebral disc localization and segmentation. Deep learning has been used to tackle with false positive detection with the help of pose estimation and semantic segmentation techniques. The accuracy of the results were compared by the ground truth pixel location against predicted pixels location. Spine generic public multi-center dataset was used to evaluate the proposed method.

Keywords: Spinal cord segmentation · IVD localization · Intervertebral Disc · Localization

1 Introduction

The spine is the name given to a bone structure composed of vertebrae that jointly move with each other and extend from the skull to the pelvis. Spine, it starts at the neck and extends to the coccyx. Figure 1 shows structure of spine where it is seen like "S" shaped. It is composed of 33 vertebrae. It is divided into 5 areas: nape (7 vertebrae), back (12 vertebrae), lumbar vertebrae (5 vertebrae), coccyx (sacrum) (5 vertebrae), coccyx (4 vertebrae). The vertebras of the vertebrae join to form the spinal canal. The spinal canal contains the spinal cord. The vertebrae that make up the spine are connected by semi-movable joints. Only the coccyx and coccygeal vertebra have immobile joints between them. The spine protects both the spinal cord and the skeleton from behind. It allows the body to stand upright. It is the cage that protects the lungs through the ribs. It is also a connection point for internal organs. The S-shaped curved structure ensures that the spine is light in jump and balance.

R. Manza et al. (Eds.): ACVAIT 2022, AISR 176, pp. 88–98, 2023.
https://doi.org/10.2991/978-94-6463-196-8_8

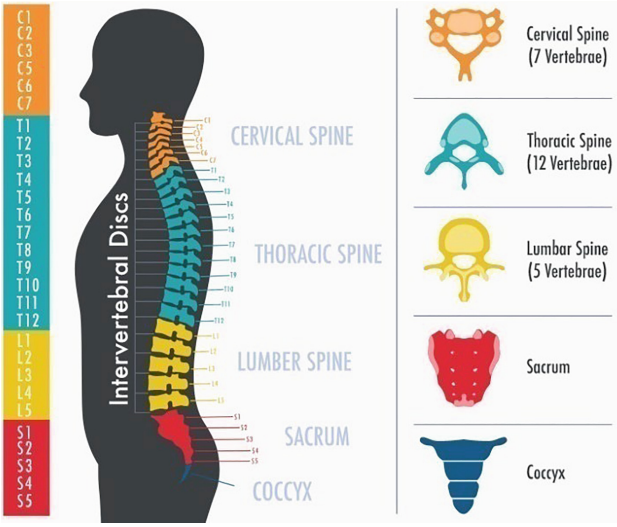

Fig. 1. The Anatomy of the Spine [https://www.motionspineinstitute.com/spine-101/]

Any harm injury to the spinal cord and it's in-vertebral discs lead to major health problems. To diagnose the injury of the spinal cord and in-vertebral discs, images are required. There are two popular methods to acquire images of Intervertebral discs such as Magnetic Resonance Imaging (MRI) and Computerized Tomography (CT). Medical imaging provides images of human spine and its In-vertebral disc to determine the anatomical structure of scoliosis, hernia, disc. It is important for diagnosing diseases such as slippage of disc, disc gaps. Today, the disc in a lumbar MRI image and localization of the vertebrae has been analyzed by radiologists manually. This process takes a long time which may lead to an error. There are several computer aided techniques and tools have been introduced for automatic detection and localization of IVD and Vertebrae. For this reason, many methods have been presented for automatic identification and positioning of the disc. Existing methods are usually based on machine learning. A classifier is used. Recent deep learning methods, men. Most successful in recognizing the structure of the spine. Record the value. In this article, we adopted a fully convolutional network (FCN), Stacked Hourglass Network with a Multi-level Attention Mechanism, and Region Growing techniques for disc localization and segmentation (Fig. 2).

The spine has 23 discs that are situated between the 24 cervical, thoracic, and lumbar vertebrae. Six cervical discs (also known as the cervical spine) are placed in the neck (also known as the cervical spine) between seven cervical vertebrae (C1–C7) just beneath the skull. The lower back (also known as the lumbar spine) has five lumbar discs that are positioned between five lumbar vertebrae.

Each vertebral region performs a specific job in the human body, such as breathing, walking, and protecting the spinal cord. Damage to the in-vertebral disc might result in back discomfort or hypersensitivity in various body areas. The accident or high pressure/tension on the disc usually causes any form of damage to the in-vertebral disc. As a result, diagnosing the in-vertebral disc is critical; assessing the disc form and/or

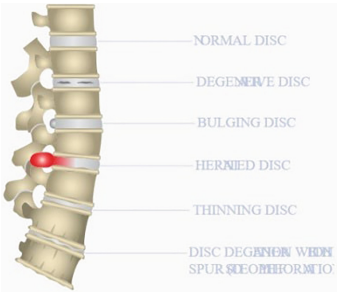

Fig. 2. Damaged IVD

locating the damaged disc may aid in the analysis. The detection of in-vertebral discs is the first step in the diagnosis. Detecting the discs manually is a difficult task. Several strategies have been proposed in existing works to conduct automatic disc recognition and localization mentioned in the Sect. 2.

2 Review of the Literature

Automatic labeling and analysis of the IVD is a critical task. Recently, many methods have been introduced in the field of medical image analysis for Intervertebral disc segmentation and localization [1, 2]. Recent study shows that IVD localization interest have been focused on the deep convolution neural networks (CNNs), and many studies reveals that it outperforming traditional localization techniques. Cohen et. al. [3] proposed 3D Fully Convolutional Network (FCN) to retrieve center coordinates of IVD and segment the disc. In the work proposed in the work proposed by Ji et al. [4] have adopted standard CNN for Intervertebral disc segmentation by using a patch around each pixel. The authors have used 2D patch and impact of vicinity size to evaluate different patch strategies. The authors of [5] adopted deeply supervised multi-scale fully CNN for Intervertebral disk segmentation applied on MR-T2 weighted images. In their work risk of loss of gradients during the training was reduced by use of multi-scale deep supervision in the architecture. In the work proposed by Forsberget. al. [6] clinically annotated spine labels were used for detection and labeling pipelines for cervical and lumbar MR. They have used two distinct pipelines for labeling and detection of vertebrae. And two neural networks (CNNs) were configured for locating lumbar/cervical vertebra. Zhuet. al. [7] introduced a method based on Gabor filter bank for Intervertebral disc localization and segmentation. Alomari et al. [8] employed a two-level probabilistic model for IVD discs localization from MRI images. Michopoulou et al. [9] introduced a semi-automatic method for detection and segmentation of IVDs. In their work they have considered both degenerated and normal lumbar IVDs. Another model based searching method was used to localize entire spine discs by Penget. al. [10]. Castro et al. [11] used active contour model with fuzzy C-means technique to segment the IVDs. Haq et al. [12] used the discrete simplex surface model for segmentation of the IVDs. Anovel anisotropic-oriented flux model employed in the work proposed by Law et. al. [13] to segment the IVDs.

Above mentioned methods needs manual operations or human interaction to refine the results for effective IVDs localization and segmentation. Some studies have proposed

early fusion and late fusion techniques of IVD features for Multi-modal segmentation of IVDs [14–20].

3 Proposed Methodology

3.1 Data

We employed MRI spinal cord dataset [3]. The dataset consist of T1w and T2w MRI data from 235 subjects, the dataset includes inconsistent images since, images were captured from 40 different centers. The network was fed an average of each subject's six center slices as input images. For training, testing, and validation, the dataset was divided into three parts: 75%, 10%, and 15%, respectively. Ground truth data was manually formed through *"labelme"* annotation tool [21].

3.2 Pre-processing

The Spinal Cord Toolbox (SCT) v4.0.1 was used to preprocess 3D volumes of the MRI data [22]. The images were re-sampled at 1 mm isotropic resolution and straightened using the spinal cord segmentation method to produce the spinal cord centerline [23]. The image was cropped to 256 * 256 pixels around the spinal region as part of the straightening procedure [24]. To reduce contrast variability in the image, a Contrast Limited Adaptive Histogram Equalization technique was used [25]. We increased the target size to deal with class imbalance by applying a 10-pixel Gaussian kernel to single-pixel labels.

Further we extract the average of 6 sagittal slices (centered in the middle slice) as a data sample for each subject. We normalize each image to be in range [0, 1] to reduce the effect of data variation. In order to prepare the ground truth data for the training process, first, we extract the intervertebral disc position (single pixel) from the ground truth data then we convolve the image with a Gaussian kernel to generate a smooth ground truth with increased target size (radius 10). We repeat this process for each intervertebral disc separately to produce V channel ground truth, where V is the number of intervertebral discs. Since the Spine Generic dataset consists of samples with variable number of intervertebral discs (between 7–11), we extract 11 intervertebral discs for each subject. For any missing inter vertebral disc we consider unknown position and eliminate its effect on the training process by simply filtering out with the visibility flag on the loss function.

The proposed method starts with the pre-processing for the proposed model. The position of the intervertebral discs were extracted using the pose estimation method with attention mechanism of Haourglass model. Figure 3 depicts the proposed method. Steps involved in the methodology are discussed in the subsequent section.

3.3 Proposed Method

The stacked hourglass network [26] learns the object posture using (N-1) intermediate (shown in 3 as intermediate prediction) and one final prediction, as illustrated in Fig. 3. As a result, the multi-level representation is taken into consideration in terms of

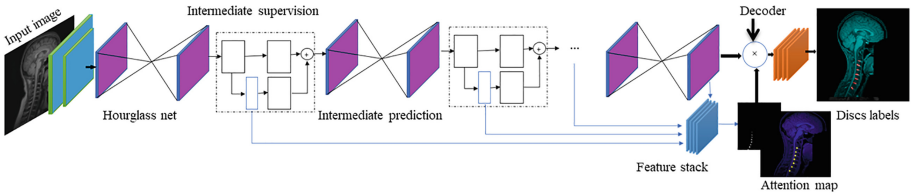

Fig. 3. Proposed stacked hourglass network, An attention mechanism is built into a stacked hourglass network. The model considers the loss function between each hourglass prediction and ground truth mask (intermediate supervision). Further it feds the intermediate representation into an attention layer to produce the attention map (heat map). The attention map guides the decoder layer to focus on the in-vertebral disc.

the N stacked hourglass network. We propose using a multi-level attention technique to enhance the power of representation space. To do this, each hourglass network's intermediate representation (shown in step 2) is concatenated to build a multi-level representation.

This representation can be viewed as a collection of collective knowledge gathered from multiple levels of the network at various scales; consequently, employing this collection of collective knowledge as a supervisory signal to calibrate the final representation can result in a superior representation. We stack all of the intermediate representations to incorporate this supervisor signal. To construct a single channel attention mechanism, this stacked representation is given to the attention block (series of point-wise convolution with sigmoid activation). To re-calibrate the representation space and train the model to pay more attention to the disc position, we multiply this attention channel with the final representation. The attention block fed to region growing technique. Centroid of the block of attention map/ heat map was considered as a seed point for region growing. In the instance of the Region expanding approach, start with a seed pixel and then look at the neighboring pixels. If the neighboring pixels follow the preset rules, that pixel is added to the seed pixel's region, and the procedure continues until no similarity remains. The bottom-up strategy is used in this procedure. The preferred rule might be specified as a threshold in the event of an expanding region. Threshold for region growing is computed using Eq. 1.

$$|Z_{max} - Z_{min}| \leq threshold \tag{1}$$

where, $Z_{max} \rightarrow$ *Maximum pixel intensity value in a region*
And $Z_{min} \rightarrow$ *Minimum pixel intesity value in a region.*

4 Results and Discussion

The Spine Generic Dataset is used to assess the performance of the proposed technique. Each participant has both T1w and T2w contrasts in this sample. Images were taken in 42 various locations across the world. The dataset encompasses a wide range of sample quality, scale, and imaging devices, making it a difficult baseline for intervertebral disc labeling.We use Adam optimization to train the proposed model over 150 epochs with

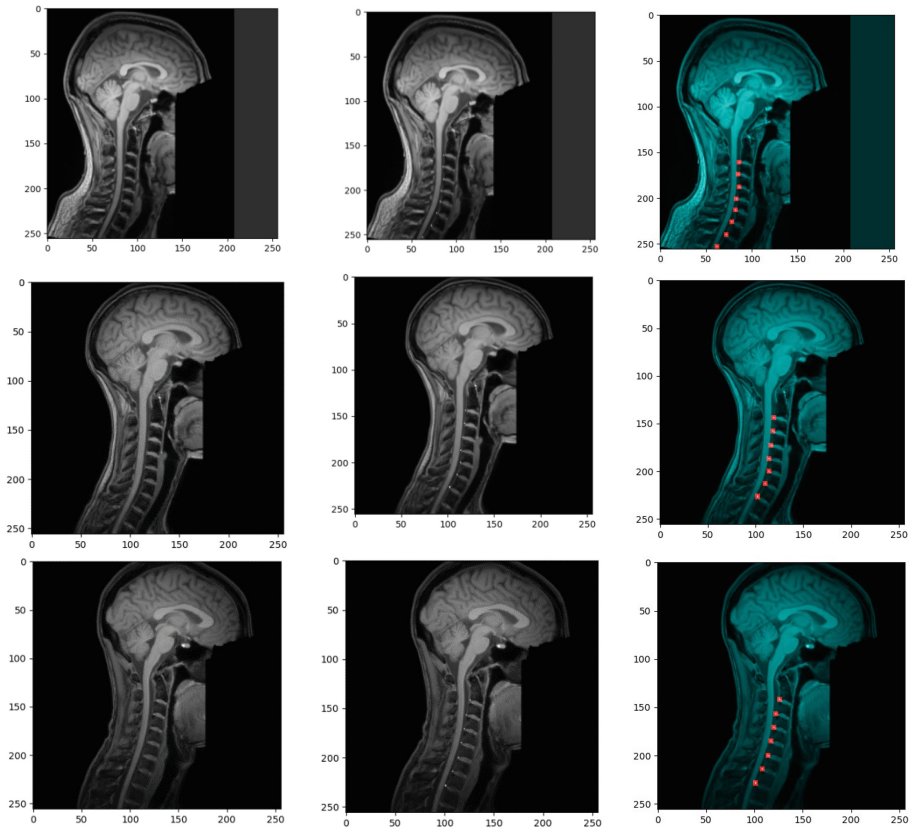

Fig. 4. **(A)** Original image. **(B)** Seed point detected by hourglass network. **(C)** attention map produced by Hourglass Network

a learning rate of 0.00025 and a batch size of 4. In our tests, we found that employing two stacks yielded the greatest results on the validation set. The approach may be easily used via the Spinal Cord Toolbox [22], and the implementation and model training were done in ivadomed [27].

4.1 Evaluation Matrices

Dice Overlap Coefficients

The percentage of successfully segmented voxels is measured using the Dice metric. Dice is calculated using Eq. (1).

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \times 100\% \tag{2}$$

The percentage of successfully segmented voxels is measured using the Dice metric. Where A represents the sets of foreground voxels in the ground-truth data and B

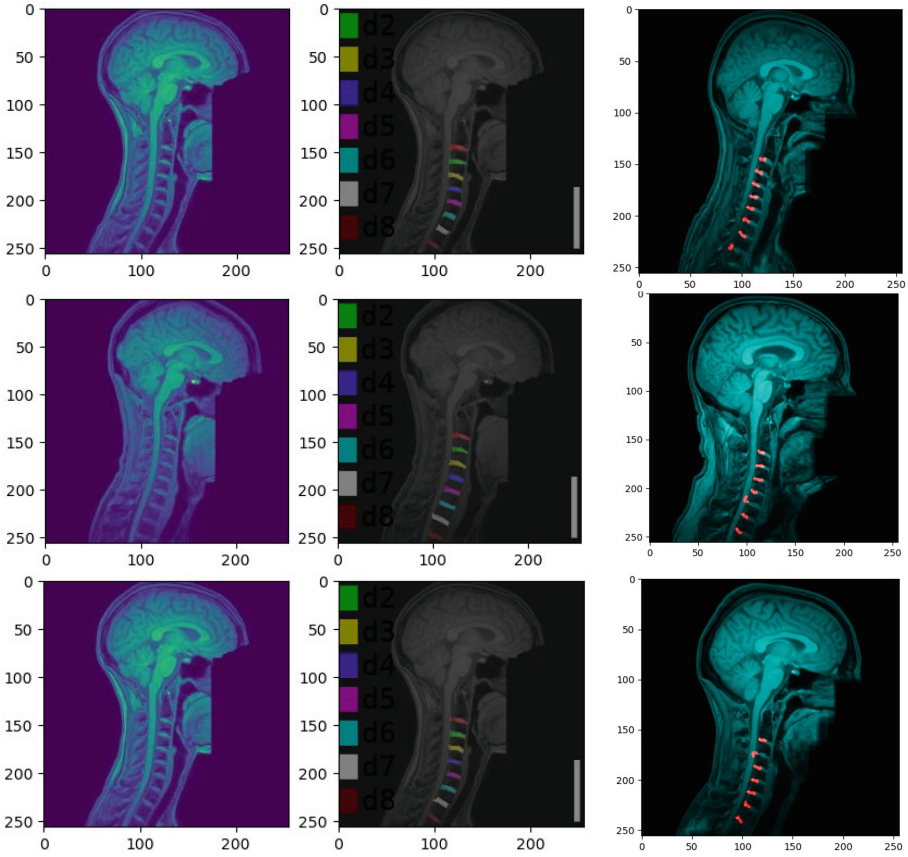

Fig. 5. **A**) Input images (left), **B)** ground truth image (middle), **C**)output predictions of our proposed method (right)

represents the matching sets of foreground voxels in the segmentation result, Dice is calculated. Better segmentation accuracy is associated with a higher Dice metric. Further, dice value was received by confusion matrix as 1 (correctly identified) and 0 (missed) the target and predicted pixel coordinates. Figure 9(a) and (b) depicts the produced confusion matrix from localization distance technique (Figs. 5, 7 and 8).

4.2 Effect of Hourglass Attention Mechanism

The proposed approach makes use of the attention mechanism to re-calibrate the representation space and focus the model's attention to the target region. We trained the model with and without the attention mechanism to see how it affected the results. The results in Fig. 6 show that the attention mechanism model performance in both T1w and T2w modalities. Figure 4(C) shows a sample attention map on the input images to visualize the influence of the network's attention mechanism.

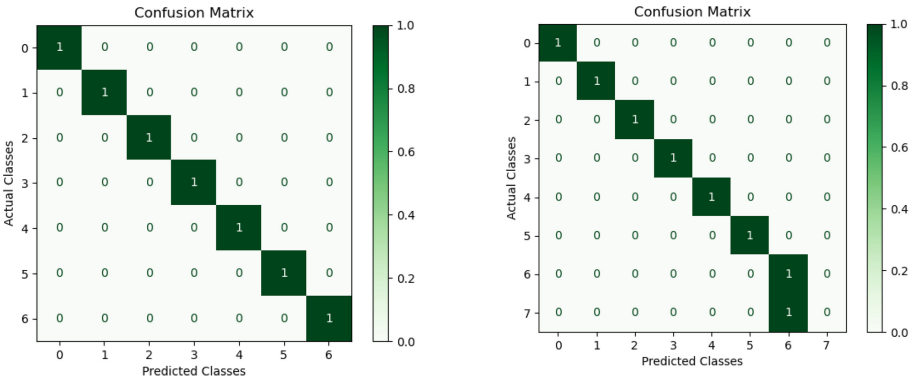

Fig. 6. **(a)** Confusion matrix for testing sample 1, **(b)** Confusion matrix for testing sample 2

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	1
5	1.00	1.00	1.00	1
6	1.00	1.00	1.00	1
accuracy			1.00	7
macro avg	1.00	1.00	1.00	7
weighted avg	1.00	1.00	1.00	7

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	1
4	1.00	1.00	1.00	1
5	1.00	1.00	1.00	1
6	0.50	1.00	0.67	1
7	0.00	0.00	0.00	1
accuracy			0.88	8
macro avg	0.81	0.88	0.83	8
weighted avg	0.81	0.88	0.83	8

Fig. 7. **(a)** Precision recall graph for sample 1, **(b)** Precision recall graph for sample 2

	centroid-0	centroid-1	orientation	area	eccentricity	perimeter
0	162.365385	91.403846	1.346391	52	0.937089	33.106602
1	175.761905	91.000000	1.285165	63	0.939756	35.313708
2	189.793651	90.111111	1.278327	63	0.928895	35.520815
3	202.913793	88.344828	1.426224	58	0.956987	33.656854
4	215.712329	87.917808	1.311686	73	0.936319	38.727922
5	229.090909	83.045455	1.087678	44	0.946110	30.349242
6	245.027397	77.315068	0.924879	73	0.930408	37.284271

Fig. 8. Features extracted from sample 1

	centroid-0	centroid-1	orientation	area	eccentricity	perimeter
0	145.968421	124.284211	1.388515	95	0.958702	47.349242
1	160.250000	122.650000	1.479635	80	0.965572	43.242641
2	175.131579	121.092105	1.259736	76	0.953223	40.935029
3	188.400000	120.013333	1.383193	75	0.954316	39.692388
4	202.390244	119.829268	1.286689	82	0.953312	43.142136
5	216.298701	114.883117	1.285865	77	0.897980	38.627417
6	231.367647	107.514706	1.072349	68	0.917693	36.455844
7	245.471429	97.457143	1.048530	70	0.953265	39.420310

Fig. 9. Features extracted from sample 2

5 Conclusion

In this work we fabricated the IVD localization and labeling through region growing and pose estimation technique. The structural information of the IVDs is employed for training and to localize the true location of the discs.

The proposed approach re-calibrates the representation space to focus more on the intervertebral disc area by utilizing the strength of the attention mechanism. To eliminate the FP and FN detection, we presented a skeleton-based post-processing technique.

A new design for recognizing intervertebral discs is presented in this study. The approach improves localization precision while reducing false positives and negatives. Extending the testing of this model to more "real-life" datasets in patients will be a future goal.

References

1. B. Ayed, K. Punitha kumar, G. Garvin, W. Romano, and S. Li, "Graph cuts with invariant object-interaction priors: application to intervertebral disc segmentation," in Biennial International Conference on Information Processing in Medical Imaging. Springer, 2011, pp.221–232.
2. C. Chen, D. Belavy, W. Yu, C. Chu, G. Armbrecht, M. Bansmann,D. Felsenberg, and G. Zheng, "Localization and segmentation of3D intervertebral discs in MR images by data driven estimation," IEEE transactions on medical imaging, vol. 34, no. 8, pp. 1719–1729,2015.
3. J Cohen-Adad. Spinal cord MRI public database (multi-subjects), July 2019.
4. X. Ji, G. Zheng, D. Belavy, and D. Ni, "Automated intervertebral disc segmentation using deep convolutional neural networks, "in International Workshop on Computational Methods and Clinical Applications for Spine Imaging. Springer, 2016, pp. 38–48.
5. G. Zeng and G. Zheng, "DSMS-FCN: A deeply supervised multiscale fully convolutional network for automatic segmentation of intervertebral disc in 3D MR images," in International Workshop and Challenge on Computational Methods and Clinical Applications in Musculoskeletal Imaging. Springer, 2017, pp. 148–159.
6. Forsberg, Daniel, Erik Sjöblom, and Jeffrey L. Sunshine. "Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data." Journal of digital imaging 30, no. 4 (2017): 406-412.

7. Zhu, Xinjian, Xuan He, Pin Wang, Qinghua He, DandanGao, Jiwei Cheng, and Baoming Wu. "A method of localization and segmentation of intervertebral discs in spine MRI based on Gabor filter bank." *Biomedical engineering online* 15, no. 1 (2016): 1–15.
8. Alomari RS, Corso JJ, Chaudhary V. Labeling of lumbar discs using both pixel-and object level features with a two level probabilistic model. IEEE Trans Med Imaging. 2011; 30:1–10.
9. Michopoulou SK, 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. 2009; 56:2225–31.
10. Peng ZG, Zhong J, Wee W, Lee JH. Automated vertebra detection and segmentation from the whole spine MR images. 2005 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2005; 2527–2530.
11. Castro-Mateos I, Pozo JM, Lazary A, Frangi AF. 2D segmentation of intervertebral discs and its degree of degeneration from T2-weighted magnetic resonance images. Medical imaging 2014.Comput Aided Diagn. 2014; 9035:17.
12. Haq R, Aras R, Besachio DA, Borgie RC, Audette MA. 3D lumbar spine intervertebral disc segmentation and compression simulation from MRI using shape-aware models.Int J Comput Assist Radiol Surg. 2015; 10:45–54.
13. Law MWK, Tay K, Leung A, Garvin GJ, Li S. Intervertebral disc segmentation in MR images using anisotropic oriented flux. Med Image Anal. 2013; 17:43–61.
14. W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, and D. Shen, "Deep convolutional neural Networks for multi-modality isointense infant brain image segmentation," NeuroImage, vol. 108, pp. 214–224, 2015.
15. P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. Benders, and I.Ĭsgum, "Automatic segmentation of MR brain images with a convolutional neural network," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1252–1261, 2016.
16. K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical image analysis, vol. 36, pp. 61–78, 2017.
17. J. Dolz, C. Desrosiers, L. Wang, J. Yuan, D. Shen, and I. Ben Ayed, "Deep CNN ensembles and suggestive annotations for infant brain MRI segmentation," arXiv preprint arXiv:1712.05319, 2017.
18. S. Valverde, M. Cabezas, E. Roura, S. Gonz´alez-Vill'a, D. Pareto, J. C. Vilanova, L. Rami´o-Torrent'a, 'A. Rovira, A. Oliver, and X. Llad´o, "Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach," NeuroImage, vol. 155, pp. 159–168, 2017.
19. N. Srivastava and R. Salakhutdinov, "Multimodal learning with deep boltzmann machines," Journal of Machine Learning Research, vol. 15, pp. 2949–2980, 2014.
20. D. Nie, L. Wang, Y. Gao, and D. Sken, "Fully convolutional networks for multi-modality isointense infant brain image segmentation," in 13th International Symposium on Biomedical Imaging (ISBI), 2016. IEEE, 2016, pp. 1342–1345.
21. LabelMe: a database and web-based tool for image annotation. B. Russell, A. Torralba, K. Murphy, W. T. Freeman. International Journal of Computer Vision, 2007.
22. Benjamin De Leener, Simon L'evy, Sara M Dupont, Vladimir S Fonov, Nikola Stikov, D Louis Collins, Virginie Callot, and Julien Cohen-Adad. SCT: Spinal cord toolbox, an open-source software for processing spinal cord MRI data. Neuroimage, October 2016.
23. Benjamin De Leener, Gabriel Mangeat, Sara Dupont, Allan R Martin, Virginie Callot, Nikola Stikov, Michael G Fehlings, and Julien Cohen-Adad. Topologically preserving straightening of spinal cord MRI, 2017.
24. Charley Gros, Benjamin De Leener, AtefBadji, Josefina Maranzano, Dominique Eden, Sara M Dupont, Jason Talbott, RenZhuoquiong, Yaou Liu, Tobias Granberg, Russell Ouellette,

Yasuhiko Tachibana, Masaaki Hori, Kouhei Kamiya, Lydia Chougar, Leszek Stawiarz, Jan Hillert, Elise Bannier, Anne Kerbrat, Gilles Edan, Pierre Labauge, Virginie Callot, Jean Pelletier, Bertrand Audoin, Henitsoa Rasoanandrianina, Jean-Christophe Brisset, Paola Valsasina, Maria A Rocca, Massimo Filippi, Rohit Bakshi, Shahamat Tauhid, Ferran Prados, Marios Yiannakas, Hugh Kearney, Olga Ciccarelli, Seth Smith, Constantina Andrada Treaba, Caterina Mainero, Jennifer Lefeuvre, Daniel S Reich, Govind Nair, Vincent Auclair, Donald G McLaren, Allan R Martin, Michael G Fehlings, Shahabeddin Vahdat, Ali Khatibi, Julien Doyon, Timothy Shepherd, Erik Charlson, Sridar Narayanan, and Julien Cohen-Adad. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. Neuroimage, 184:901–915, January 2019.

25. Karel Zuiderveld. Contrast limited adaptive histogram equalization. In Graphics gems IV, pages 474–485. Academic Press Professional, Inc., August 1994.
26. Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: European conference on computer vision. pp. 483–499. Springer (2016)
27. Gros, C., Lemay, A., Vincent, O., Rouhier, L., Bourget, M.H., Bucquet, A., Cohen, J.P., Cohen-Adad, J.: ivadomed: A medical imaging deep learning toolbox. Journal of Open Source Software 6(58), 2868 (2021). https://doi.org/10.21105/joss.02868