

Automatic Intervertebral Discs Localization and Segmentation: A Vertebral Approach

Amir Jamaludin¹(✉), Meelis Lootus¹, Timor Kadir², and Andrew Zisserman¹

¹ University of Oxford, Oxford, UK
{amirj,meelis,az}@robots.ox.ac.uk
² Mirada Medical, Oxford, UK
timor.kadir@mirada-medical.com

Abstract. This paper describes an automatic system for intervertebral discs (IVDs) localization and segmentation in three-dimensional magnetic resonance imaging scans. The system builds upon the localization and segmentation system first introduced by Lootus et al. with several improvements to the localization step. The system was trained on T1 and T2 scans of 341 patients obtained from various sources. The proposed system achieved a mean localization error of 1.1 ± 0.6 mm and a mean Dice overlap coefficient of 84.0 ± 1.5 % on the 15 training data of the challenge on IVD localization and segmentation at the 3rd MICCAI Workshop & Challenge on Computational Methods and Clinical Applications for Spine Imaging - MICCAI–CSI2015.

1 Introduction

The objective of this work is the automated localization and segmentation of intervertebral discs (IVDs), and to this end we propose a system that was first presented in the work of Lootus et al. [1] and improve it with ideas from the work of Jamaludin et al. [2]. The whole system comprises five main steps: 1. vertebrae detection and labelling, 2. corner localization of detected vertebrae, 3. detection of the extent of the vertebrae in sagittal slices, 4. IVDs segmentation via graph cuts, and 5. localization of IVDs centres.

2 Methodology

2.1 Vertebrae Detection and Labelling

To detect and label the vertebrae, we use the detection and labelling scheme proposed by Lootus et al. [3] which uses a combination of a deformable part model (DPM) detector [4] and labelling via graphical model. The input to this stage is a three-dimensional (3D) magnetic resonance (MR) volume and the output is a series of approximate bounding boxes with the vertebrae labels from T11 to the combined sacrum (S1 and S2). The detector uses two different groups of histogram of oriented gradients (HOG) templates one for the combined sacrum

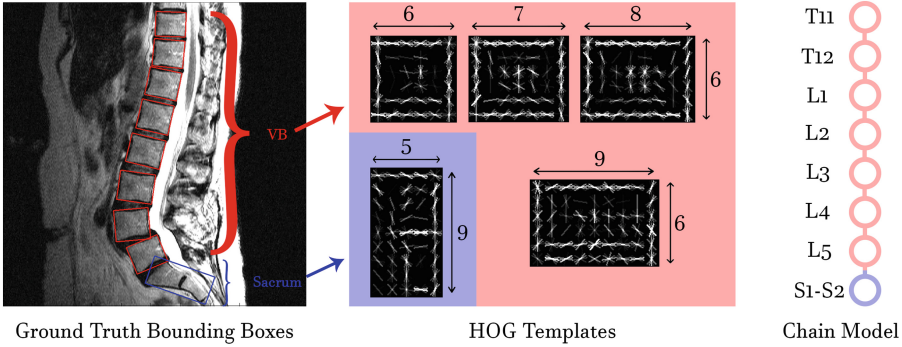


Fig. 1. Examples of the ground truth bounding boxes that were used to train the HOG templates models are shown here in cell units where one cell is made up of 8×8 pixels. Only one template was trained for the sacrum while four different templates of varying ratio were trained for the other vertebrae.

(S1-S2) and the other for T11-L5 vertebrae detections. The graphical model is a chain graph with eight vertices, one for each vertebra (T11 to S1-S2), with the edges describing the geometrical relationships of one vertebra and the next. Both the HOG templates and geometrical relationships of the vertices are trained with annotated ground truth bounding boxes with labels as described in the work of Lootus et al. [3]. Examples of annotated ground truths, the trained HOG templates, and the graph of the chain model can be seen in Fig. 1 while an example of the input and output can be seen in Fig. 2.

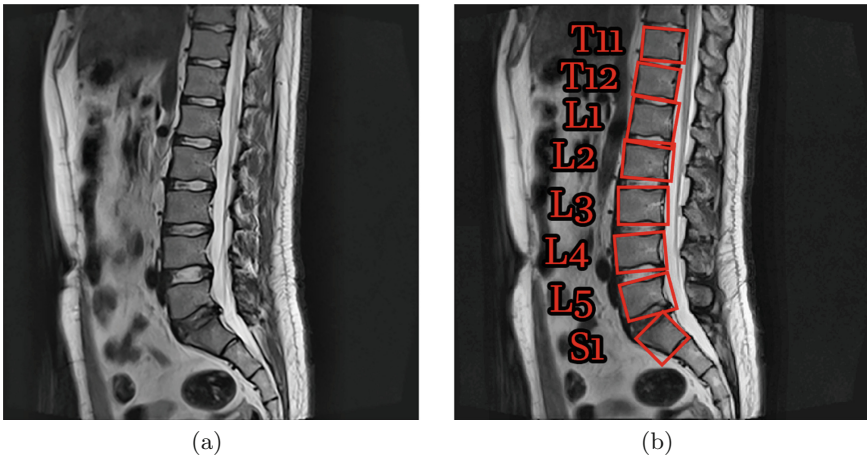


Fig. 2. (a) A mid-sagittal slice of the 3D MR scan (input). (b) The same scan superimposed with the bounding boxes and their corresponding labels (output). Note: the S1-S2 bounding box is truncated to just S1.

2.2 Corner Localization

We then refine the localization of these bounding boxes such that the resulting quadrilaterals are more consistent and tightly fit the vertebrae. This is achieved by regressing to the corner points of the vertebrae contained in the bounding boxes. We adapt the supervised descent method (SDM) by Xiong et al. [5] originally developed for the detection of facial landmarks. Implementation details and experimentation results of the regression of the corner points can be found in the work of Jamaludin et al. [2]. Examples of corner localized vertebrae with corresponding bounding box inputs can be seen in Fig. 3.

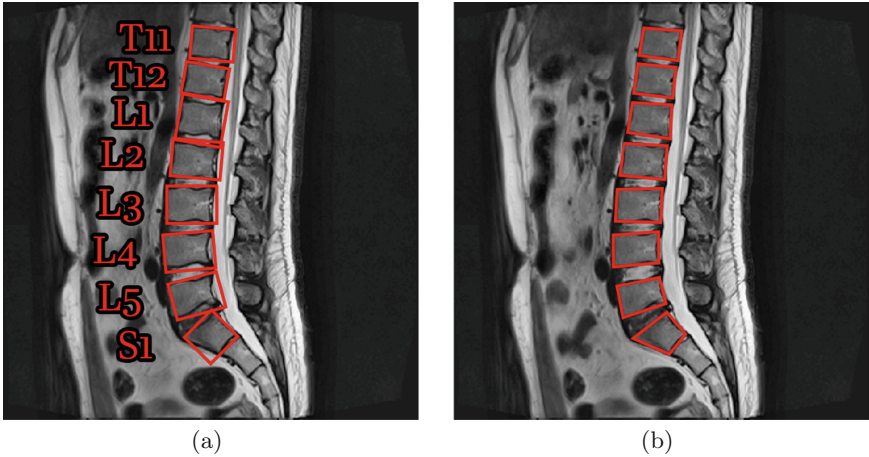


Fig. 3. (a) Input. (b) Output. Note: the quadrilaterals are tighter in terms of fit compared to the original bounding boxes.

2.3 Detection of the Extent of the Vertebrae

All the previous steps are performed on each sagittal slice of the scan, however, it is also necessary to determine the vertebra start and end. This is important since the positions of the vertebrae in a scan are initially unknown and there exist slices which contain only partial volumes of the vertebrae, largely containing other non-vertebral tissue. Such partial vertebrae are problematic because they should be considered to be part of the background class during segmentation. To this end we utilise a binary classifier to distinguish non-vertebrae and vertebrae quadrilaterals.

We follow the method proposed by Chatfield et al. [6], where the steps are: 1. dense scale-invariant feature transform feature extraction over the quadrilaterals, 2. Fisher vector encoding of the features, 3. spatial tiling of the features in the image and 4. classification via linear support vector machines. This is done on a per slice basis on every slice where the quadrilaterals are classified as either vertebra or non-vertebra. Examples can be seen in Fig. 4.

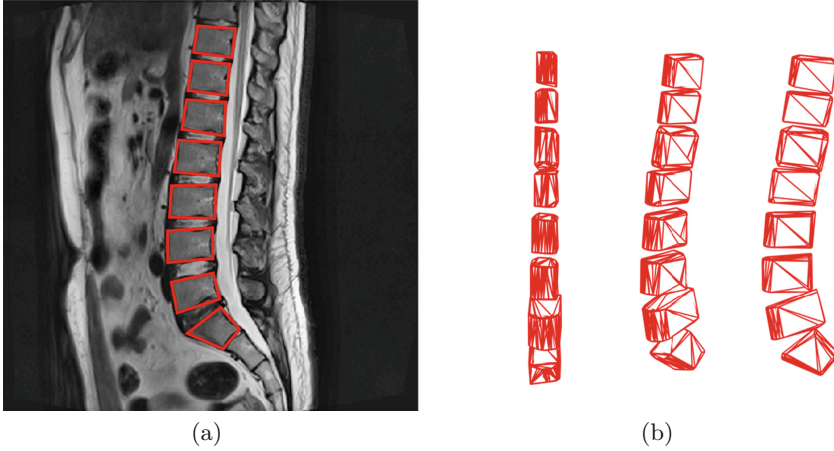


Fig. 4. (a) Tight quadrilaterals of the midsagittal slice (input). (b) Triangular mesh plots made from quadrilaterals in the 3D volume that were classified as vertebrae (output). The plots shown is of a single scan but at three different orientations.

2.4 IVD Segmentation

We follow the segmentation scheme proposed by Lootus et al. [3] which uses a standard graph cuts algorithm. We therefore segment twice, once for the vertebrae and then once more for the IVDs. The placement of the foreground and background seeds are automatically generated according to tight quadrilaterals similar to the work of Lootus et al. [3]. This two-step segmentation proves to be better than segmenting the IVDs directly from the quadrilaterals due to the fact that accurate foreground seed placement is less demanding than vertebrae segmentation.

There are two main differences between our implementation and that discussed in the work of Lootus et al. [3]. First, the seeds are set to be the biggest at the midsagittal point, determined from the extent detector, and smallest at the sagittal edge of the vertebrae extent. Also, for the IVD segmentation, we combine the sagittal segmentation with its coronal segmentation by flipping the third axis of the 3D volume with the first axis and segmenting it again. The joint segmentation result is the final IVDs segmentation. Example segmentations of the vertebrae and IVDs can be seen in Fig. 5.

2.5 Localization of IVDs Centres

To localize the IVDs centres we combine three different localization predictions: 1. the centroid of adjacent vertebrae corner points, 2. a linear regression of adjacent vertebrae corner points, and 3. the centroid of segmented binary mask of each IVD.

The first localization prediction is the centroid of the corner points of adjacent vertebrae to a specific IVD. We assume that this centroid is a close approximation

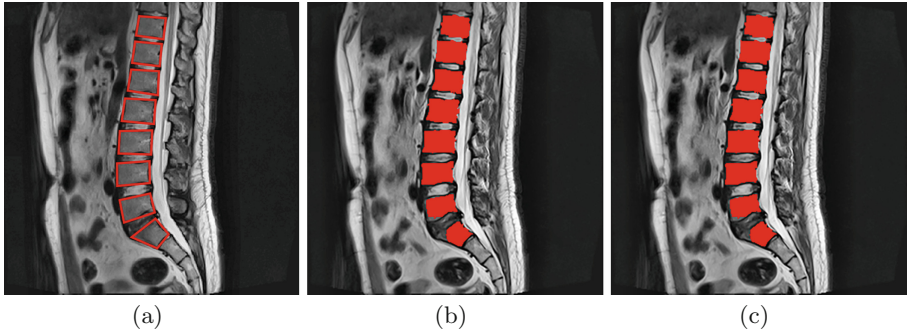


Fig. 5. (a) Tight quadrilaterals of the midsagittal slice, now classified as either vertebra or non-vertebra according to extent detector (input). For (b) vertebrae segmentation and (c) IVD segmentation are shown the resulting segmentation masks.

to the centroid of the IVD since an IVD will be bounded by the adjacent vertebrae. The second localization prediction is essentially the output of a linear regressor using the corner points as features. The linear regressor is trained by leave-one-out cross validation of the whole training set. The final localization prediction is the centroid of the segmentation binary mask. All three predictions are then averaged to give the final localization prediction. Through experimentations, we found averaging the three predictions give us a more accurate prediction overall. Example of IVDs centres localization can be seen in Fig. 6.

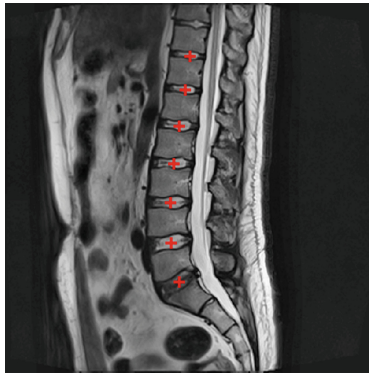


Fig. 6. Shown is a single sagittal slice and the predicted IVDs centres. In practice, the localization predictions predict the centres in 3D space.

3 Results

Results for the corner localization and the extent detector can be seen in the work of Jamaludin et al. [2]. To test our segmentation and IVD localization we use the 15 training data provided as part of the challenge on IVD localization and

segmentation [7] at the 3rd MICCAI Workshop & Challenge on Computational Methods and Clinical Applications for Spine Imaging - MICCAI–CSI2015. As per the challenge we use mean and standard deviation of the results to the ground truth for the localization task and the mean Dice overlap for the segmentation task. Results can be seen in Table 1. Besides the 15 training data provided we also tested our approach with the five test data provided in the challenge. For segmentation, we obtain a dice overlap of $82.3 \pm 3.2\%$ and an absolute distance of 1.57 ± 0.20 mm. Similarly for localization, we achieve a mean localization of 1.02 ± 0.47 mm. Our localization results on the challenge dataset is good and the system manages to achieve sub-voxel accuracy on average. However, our segmentation results can be improved upon, possibly by means of a true 3D graph cut segmentation algorithm.

Table 1. Localization and segmentation results.

Task	Mean \pm STD	Median	Min	Max
Localization (mm)	1.1 ± 0.6	1.0	0.2	2.9
Segmentation (Dice, %)	84.0 ± 1.5	84.2	79.8	86.4

4 Conclusion

This paper has presented an automatic IVDs localization and segmentation system. The proposed system managed to achieve good localization and segmentation accuracy on the challenge data which is impressive considering the system was mostly trained on a totally different dataset. This indicates the robustness of our system.

Acknowledgements. We are grateful for discussions with Prof. Jeremy Fairbank, and Dr. Jill Urban. This work was supported by the RCUK CDT in Healthcare Innovation (EP/G036861/1). The data used in this research was obtained during the EC FP7 project HEALTH-F2-2008-201626.

References

1. Lootus, M., Kadir, T., Zisserman, A.: Automated radiological grading of spinal MRI. In: Yao, J., et al. (eds.) CSI 2014. LNCVB, vol. 20, pp. 119–130. Springer, Switzerland (2015)
2. Jamaludin, A., Kadir, T., Zisserman, A.: Automatic Modic changes classification in spinal MRI. In: Vrtovec, T., et al. (ed.) Proceedings of 3rd MICCAI 2015 Workshop & Challenge on Computational Methods and Clinical Applications for Spine Imaging, MICCAI–CSI 2015, pp. 14–26 (2015)

3. Lootus, M., Kadir, T., Zisserman, A.: Vertebrae detection and labelling in lumbar MR images. In: Yao, J., et al. (eds.) CSI 2013. LNCS, vol. 17, pp. 219–230. Springer, Switzerland (2014)
4. Felzenszwalb, P., Girshick, R., McAllester, D.: Cascade object detection with deformable part models. In: Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, pp. 2241–2248. IEEE (2010)
5. Xiong, X., de la Torre, F.: Supervised descent method and its applications to face alignment. In: Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2013, pp. 532–539. IEEE (2013)
6. Chatfield, K., Lempitsky, V., Vedaldi, A., Zisserman, A.: The devil is in the details: an evaluation of recent feature encoding methods. In: Proceedings of 2011 British Machine Vision Conference, BMVC 2011, pp. 76.1–76.12. BMVA Press (2011)
7. 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)