Localization of the Lumbar Discs Using Machine Learning and Exact Probabilistic Inference

Ayse Betul Oktay and Yusuf Sinan Akgul

GIT Vision Lab
Department of Computer Engineering, Gebze Institute of Technology
Kocaeli, 41400, Turkey
{oktay,akgul}@bilmuh.gyte.edu.tr
http://vision.gyte.edu.tr

Abstract. We propose a novel fully automatic approach to localize the lumbar intervertebral discs in MR images with PHOG based SVM and a probabilistic graphical model. At the local level, our method assigns a score to each pixel in target image that indicates whether it is a disc center or not. At the global level, we define a chain-like graphical model that represents the lumbar intervertebral discs and we use an exact inference algorithm to localize the discs. Our main contributions are the employment of the SVM with the PHOG based descriptor which is robust against variations of the discs and a graphical model that reflects the linear nature of the vertebral column. Our inference algorithm runs in polynomial time and produces globally optimal results. The developed system is validated on a real spine MRI dataset and the final localization results are favorable compared to the results reported in the literature.

Keywords: lumbar disc detection, graphical models, exact probabilistic inference, object detection.

1 Introduction

There are many intervertebral disc and vertebra localization methods for the vertebral column in the literature [9,14,5]. Schmidt et al. [13] introduce a probabilistic inference method that measures the possible locations of the intervertebral discs. The approach uses a part-based model that describes the disc appearances by employing a tree classifier [8]. Similar to [13], Alomari et al. [1] use a graphical model that assumes local and global levels with latent variables. The inference on their graphical model is based on the expectation maximization method which is an approximate and iterative inference technique.

This paper introduces a novel method for the automatic localization and labeling of the lumbar discs from T1-weighted sagittal MR images. At the local level, the method uses recent machine learning methods to locally search the disc positions by employing Pyramidal Histogram of Oriented Gradients (PHOG) [4] with Support Vector Machines (SVM). At the global level, the method takes advantage of the chain-like structure of the spine by assuming latent variables for

the disc positions. A polynomial time exact inference method on the chain-like graphical structure [3] is employed to find the final lumbar disc positions and labeling.

Our method has two important advantages. First, our graphical model does not directly use any image intensity information which varies greatly. Instead, we use the PHOG based SVM detector results as the observed data in our model. Also, many local and semi-global features can be conveniently incorporated into the SVM without complicating the global level inference mechanism.

For the second major advantage of our system, we exploit the one dimensional nature of the spine and form a chain like graphical model that contains only the latent variables for the disc positions. The literature includes very efficient polynomial time inference algorithms for chain-like graphical models that we employ to find robust disc locations even for pathological cases. In addition, the chain-like structure of our model makes it possible to include extra information for the end discs of the lumbar region, which are the most problematic sections in terms of localization. As a result, our model does not assume any image dependent spatial disc positions and hence it does not require manual initialization.

The rest of this paper is organized as follows. The disc scoring with PHOG based SVM is described in Section 2. Section 3 includes the chain-like probabilistic graphical model and exact inference. The validation of the method is presented in Section 4. Finally, we conclude in Section 5.

2   Disc Scoring with PHOG Based SVM

The lumbar discs vary in the size, location, shape, and appearance because of pathologies and individual variations. We observe that the most invariant property of a disc image is its edges and orientation rather than the intensity, location, and shape which are used in the disc detection methods in the literature [1,5]. Therefore, we use the PHOG descriptors for feature extraction which are more robust to disc and vertebrae abnormalities.

In the PHOG descriptor extraction, a shape is represented by a histogram of gradient orientations which are quantized into a number of bins. Each bin shows the number of edges that have orientations within a specific angular range. PHOG combines the local image shape [6] with the spatial pyramid kernel [7].

The gradient values are first computed and their orientations are calculated. Then, a PHOG descriptor is calculated for each candidate region. Extracting the PHOG descriptors in the original images (512x512) has high computational cost. Therefore, we use the integral histogram technique [12] to speed up the feature extraction process.

After the extraction of the PHOG descriptors, they are trained with Sequential Minimal Optimization [10] for SVM. We use a sliding window approach for the scoring process. The pixel in the center of the window is assigned a score that indicates if the window contains a disc. These scores are generated by fitting a logistic regression model to the outputs of SVM [11].

Let $d = \{d_1, d_2 \ldots, d_6\}$ denote the labels for the lumbar intervertebral discs (T12-L1, L1-L2, L2-L3, L3-L4, L4-L5, and L5-S1 which are shown in Figure 1).

Each disc d_i is trained and scored separately. Therefore for a given image I, each pixel is assigned 6 scores, one for each disc. For further processing, the pixels that have scores less than 0.5 are eliminated.

Consider $p_{d_i}(l_k)$ as the score assigned to the pixel at location l_k for the disc d_i by SVM. For the disc d_i, the assigned scores are denoted as

$$p^i = \{p_{d_i}(l_1), p_{d_i}(l_2) \ldots, p_{d_i}(l_m)\}, \tag{1}$$

where $1 < i < 6$ and m is the number of pixels in the image. The next section explains how we use these scores in the inference algorithm.

3   Disc Center Localization with the Graphical Model

We propose a probabilistic graphical model for the final localization of the disc centers. The local features of the discs are captured by the disc scoring process with PHOG and SVM, however it is not sufficient to discriminate the discs in the lumbar region. More contextual global information like positional and orientational differences between the discs and distance to the spinal cord are necessary for the final localization of discs on the spine. We use a chain-like graphical model that combines such global information with the local information gathered from the disc scoring process.

Our graphical model is a chain consisting of 6 nodes and 5 edges where each node represents a lumbar disc (Figure 2). Let $x_k \in \Re^2$ be a random variable that assigns node k to its image location. The optimal configuration $x' = \{x_1, x_2 \ldots, x_6\}$ assigns all discs $d = \{d_1, d_2 \ldots, d_6\}$ to their exact locations.

Our objective is finding the optimal localizations of the disc centers with the maximum a posteriori (MAP) estimate

$$x' = \arg \max_{x} P(x|p, \alpha), \tag{2}$$

where α represents the parameters learned from the training set and $p = \{p^1, \ldots, p^6\}$ are the assigned scores with Eq. 1. $P(x|p, \alpha)$ captures the probability of being a disc and the relation with the neighboring discs. The Gibbs distribution of $P(x|p, \alpha)$ is defined as

$$P(x|p, \alpha) = \frac{1}{Z} exp \left\{ - \left[\sum \psi_L(I, x_k) + \lambda \sum \psi_G(x_k, x_{k+1}, \alpha) \right] \right\}, \tag{3}$$

where the potential function $\psi_L(I, x_k)$ carries local information about the discs and the potential function $\psi_G(x_k, x_{k+1}, \alpha)$ includes more global information like the distance and orientation. λ is a weighting parameter which is selected as 0.5.

In the local potential function $\psi_L(I, x_k)$, we directly use the scores $p^i = \{p_{d_i}(l_1), p_{d_i}(l_2) \ldots, p_{d_i}(l_m)\}$ generated with Eq. 1 .

The global potential function $\psi_G(x_k, x_{k+1}, \alpha)$ is defined as

$$\psi_G(x_k, x_{k+1}, \alpha) = U(x_k, x_{k+1}, \alpha) R(x_k, x_{k+1}, \alpha) D(x_k, \alpha), \tag{4}$$

where the functions U and R capture the positional and orientational differences between neighboring disc variables x_k and x_{k+1}, respectively, and $D(x_k, \alpha)$ captures the Euclidean distance to the spinal cord whose position is calculated using the method of [2].

Let $y^i = \{y_1, y_2, \ldots, y_t\}$ be the Euclidean distances between the disc variables x_i and x_{i+1} of the training set used in the scoring process where t is the number of samples in the training set. The distance function $U(x_k, x_{k+1}, \alpha)$ is

$$U(x_k, x_{k+1}, \alpha) = \begin{cases} \frac{|E - \mu(y^k)|}{w}, & \text{if } E \in [min(y^k) - \sigma, max(y^k) + \sigma] \\ \infty, & \text{else,} \end{cases} \quad (5)$$

where $\mu(y^k)$ is the mean of y^k measured from the training set, $w = max(y^k) - min(y^k)$, σ is a threshold, and E is the Euclidean distance between the disc variables x_k and x_{k+1}.

The angular differences between the discs define the curve-like shape of the lumbar region. In order to handle this information, we use the angle information between the discs. Let $r^i = \{r_1, r_2, \ldots, r_t\}$ be the angles between the neighboring disc variables x_i and x_{i+1} measured from the training set. The orientation term $R(x_k, x_{k+1}, \alpha)$ is defined as

$$R(x_k, x_{k+1}, \alpha) = \begin{cases} \frac{|O - \mu(r^k)|}{f}, & \text{if } O \in [min(r^k) - \tau, max(r^k) + \tau] \\ \infty, & \text{else,} \end{cases} \quad (6)$$

where O gives the angle differences between the disc variables x_k and x_{k+1}, $\mu(r_k)$ is the mean orientation learned from the training set, τ is a threshold, and $f = max(r^k) - min(r^k)$.

Fig. 1. An uncropped T1-weighted MRI mid-sagittal view of the lumbar vertebrae and intervertebral discs
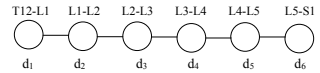
Fig. 2. Our graphical model

3.1 Exact Inference on the Chain

Given a target image I, our objective is to infer the optimal configuration x' by maximizing the Equation 2. In our graphical model, a node is conditionally

dependent to its neighboring nodes. Instead of using all of the pixels in the image I, we only use the candidate disc centers for the disc localization process. The search space for the optimal x' is decreased by conditional dependence and detected candidates. This allows us to find a globally optimal solution with recursive message passing based dynamic programming [3].

The computation time of our inference method is much better than the inference method of [13] which is based on A* search algorithm. Although A* search can produce the globally optimal solution with an appropriate heuristic, it is potentially intractable. The alternative inference method of expectation maximization [1] is computationally cheaper, but it may not find the globally optimal solution and has known initialization problem.

4   Experimental Results

The developed system is validated on a real MRI dataset for the lumbar spinal column. A 3D MRI volume is 512x512x12 voxels in size. The sagittal view includes the 6 discs of the lumbar vertebrae and some discs from the thoracic vertebrae and sacrum. In order to improve our localization results for the end-lumbar discs, we extended our graphical model with the T11-T12 and S1-S2 discs. However, we do not report their localization results because our purpose is localizing the lumbar discs only.

The dataset consists of MRI volumes of 40 different subjects where 4 of the subjects are pathology-free and the remaining 36 subjects have pathologies like disc degeneration, herniation, and scoliosis, etc. There are totally 240 lumbar intervertebral discs in the dataset and 97 of them have pathologies.

The disc detection is performed on T1-weighted sagittal slices. We use the mid-sagittal slice for both disc detection and disc center localization as it is commonly used in the medical practice. We asked an expert to mark the image region and center point of each disc for the verification and training.

4.1   Disc Scoring Results

For the disc scoring (Section 2), we perform a subset of leave-10-out cross validation. We randomly divide the dataset into 4 subsets each containing 10 subjects. In each sub-experiment, 30 MRI slices from 3 subsets are trained and 10 slices in the other subset are tested. We perform totally 4 sub-experiments, so each slice in the dataset is tested once.

For the SVM training, 90 positive samples and 900 negative samples are used for each disc. For scoring, the features are extracted from the target images. The window size is selected between the minimum and maximum window size in the training set for that disc.

The average SVM classification rates for the testing of 4 subsets are shown in the Table 1. Normally, we produce scores for the windows with logistic model fitting to SVM. If a window is assigned a probability value greater than 0.5, it is evaluated as a detected disc and otherwise it is evaluated as not a disc. The

disc classification is considered as correct if the expert marked disc is completely contained within the window. The classification rate is the number of correctly classified windows over the number of all windows. The misclassified samples are mainly caused by the confusion between the neighboring discs. The borders of the neighboring discs have nearly same orientations, so this causes the detection of the neighboring disc as the target disc most of the times. Note that, the confusion between the neighboring discs are expected because disc scoring is a completely local process.

The detection results reported by [13] has an average classification rate of 70.16%. Our average classification rate is 97% and it shows the robustness of using modified PHOG detector in the disc detection. Note that, the method of [13] and our method are tested on different datasets and [13] uses image sizes of 512x1024 pixels which contain the whole spine.

4.2 Disc Localization Results

Our disc center localization method with graphical model runs on the scoring results of the SVM. The same training and test subsets of the disc scoring process are used for the disc localization.

In order to evaluate the disc localization performance of our system, we use two different methods. In the first method, a disc is evaluated as correctly located if the localized disc center is inside the disc contour. The average disc localization accuracy of our system is 95.42%. The end discs (L4-L5 and L5-S1) have higher error rates than other discs because their positions vary greatly and pathologies generally exist in these discs. The system of Alomari et al. [1] uses 512x512 lumbar region images and it reports an accuracy of 90.7% which is lower than our accuracy. Note that, the method of [1] and our method are not tested on the same dataset.

Table 1. The average classification rates (%) of our system with SVM

	Disc detection rates(%)					
	T12-L1	L1-L2	L2-L3	L3-L4	L4-L5	L5-S1
S1	98.04	97.20	97.85	98.45	97.06	97.99
S2	97.44	97.08	97.73	98.17	98.08	98.88
S3	96.41	96.50	96.71	97.89	96.25	98.58
S4	97.12	97.57	97.78	98.89	97.93	98.96
Avg	97.25	97.09	97.52	98.35	97.33	98.59

Table 2. The mean Euclidean distances to ground truth for each subset

	Mean Euclidean distances (mm) for each subset					
	T12-L1	L1-L2	L2-L3	L3-L4	L4-L5	L5-S1
S1	2.06	2.08	3.03	3.89	3.14	4.10
S2	2.77	2.12	2.40	2.75	4.58	3.85
S3	3.28	2.78	2.60	2.67	4.11	4.07
S4	2.29	2.18	1.75	2.72	3.24	2.52
Avg	2.60	2.29	2.45	3.01	3.77	3.63

The second evaluation method for the disc center localization is the Euclidean distance to the center labeled by an expert. The mean of the Euclidean distances to the ground truth disc centers are shown in Table 2. The plot box of the Euclidean distances for the lumbar vertebrae connected discs are also shown in Figure 3. In the box plot, the centerline of the box is the median, the top and bottom of lines of the box are 25th and 75th percentiles and the pluses are the statistical outliers. Figure 4 shows the median of the Euclidean distances of our
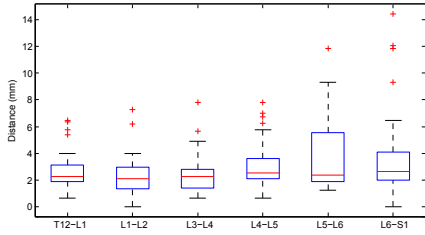
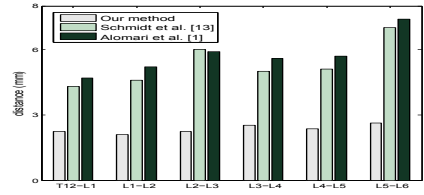**Fig. 3.** The box plot shows the Euclidean distances (in mm) to the disc centers

Fig. 4. Median errors of our system, the method of [13] and [1]
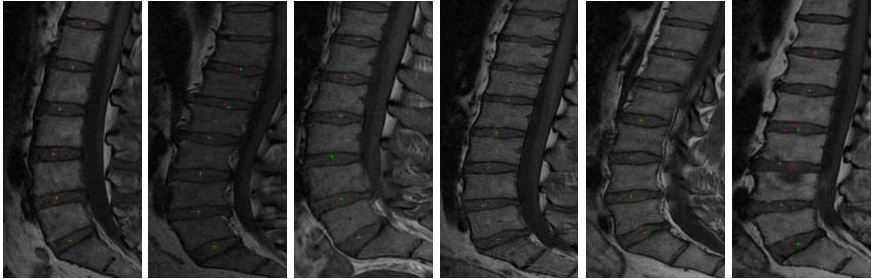

Fig. 5. The results of the disc localization. Red pluses are the ground truth and green pluses are the disc centers localized by our system. The images are cropped for better visualization.

method, the method of [1] and [13]. Our median of the Euclidean distances are always lower than the medians of the other methods [1] and [13]. Note that, [1] and [13] are tested on different datasets. We also show a few visual results of the localizations from our system (Figure 5). It is obvious that our method can localize the lumbar discs in pathological cases.

5   Conclusions

We presented a lumbar intervertebral disc localization method by employing PHOG based SVM and exact inference on a probabilistic graphical model. Our disc scoring method is more robust than other methods in the literature as verified by the experiments. In addition, other local information about the discs such as T2-weighted image features can be easily incorporated into the system. We rely on the robustness of our scoring process to eliminate some disc positions before they enter the dynamic programming, which makes our system more efficient. Also, our dynamic programming based inference mechanism can locate the disc centers in polynomial time without requiring manual initialization. Finally, the method can be extended to localize the whole intervertebral discs of

the whole spine. Our future work includes handling of the missing detected discs and providing scale independency.

References

1. Alomari, R.S., Corso, J.J., Chaudhary, V.: Labeling of Lumbar Discs Using Both Pixel- and Object-Level Features With a Two-Level Probabilistic Model. IEEE Trans. Med. Imaging 30, 1–10 (2011)
2. Bhole, C., Kompalli, S., Chaudhary, V.: Context-sensitive Labeling of Spinal Structures in MRI Images. In: SPIE Medical Imaging (2009)
3. Bishop, C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, Heidelberg (2007)
4. Bosch, A., Zisserman, A., Munoz, X.: Representing Shape with a Spatial Pyramid Kernel. In: Proceedings of the International Conference on Image and Video Retrieval (2007)
5. Carballido-Gamio, J., Belongie, S.J., Majumdar, S.: Normalized Cuts in 3-D for Spinal MRI Segmentation. IEEE Trans. Med. Imaging 23, 36–44 (2004)
6. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: International Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 886–893 (June 2005)
7. Lazebnik, S., Schmid, C., Ponce, J.: Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories. In: IEEE Conference on Computer Vision and Pattern Recognition (2006)
8. Lepetit, V., Fua, P.: Keypoint Recognition Using Randomized Trees. IEEE Trans. Pattern Anal. Mach. Intell. 28, 1465–1479 (2006)
9. Peng, Z., Zhong, J., Wee, W., Huei Lee, J.: Automated Vertebra Detection and Segmentation from the Whole Spine MR Images. In: Conf. Proc. IEEE Engineering in Medicine and Biology Society (2005)
10. Platt, J.C.: Advances in Kernel Methods, pp. 185–208. MIT Press, Cambridge (1999)
11. Platt, J.C.: Advances in Large Margin Classifiers, pp. 61–74 (1999)
12. Porikli, F.: Integral Histogram: A Fast Way To Extract Histograms in Cartesian Spaces. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 829–836 (2005)
13. Schmidt, S., Bergtholdt, M., Dries, S., Schnorr, C.: 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)
14. Seifert, S., Wachter, I., Schmelzle, G., Dillmann, R.: A Knowledge-based Approach to Soft Tissue Reconstruction of the Cervical Spine. IEEE Trans. Med. Imaging 28(4), 494–507 (2009)