# A supervised approach towards segmentation of clinical MRI for automatic lumbar diagnosis

Subarna Ghosh[1], Manavender R. Malgireddy[1],
Vipin Chaudhary[1], and Gurmeet Dhillon[2]

[1]Department of Computer Science and Engineering
State University of New York at Buffalo,Buffalo, NY 14260
[2]Proscan Imaging Inc.,Williamsville, NY 14221

**Abstract.** Lower back pain(LBP) is widely prevalent in people all over the world. It is associated with chronic pain and change in posture which negatively affects our quality of life. Automatic segmentation of intervertebral discs and the dural sac along with labeling of the discs from clinical lumbar MRI is the first step towards computer-aided diagnosis of lower back ailments like desiccation, herniation and stenosis. In this paper we propose a supervised approach to simultaneously segment the vertebra, intervertebral discs and the dural sac of clinical sagittal MRI using the neighborhood information of each pixel. Experiments on 53 cases out of which 40 were used for training and the rest for testing, show encouraging Dice Similarity Indices of 0.8483 and 0.8160 for the dural sac and intervertebral discs respectively.

**Keywords:** clinical MRI, computer-aided diagnosis, Lumbar MRI segmentation, Automatic Disc Localization, Dural sac segmentation

## 1 Introduction

Lower back pain is the second most common neurological ailment in the United States after headache [1] with more than $50 billion spent annually on rehabilitation and healthcare. In the past decade there has been a severe shortage of radiologists [2] and projections show that by the year 2020 there will be a significant boom in the ratio of their demand and supply. This concern motivates us to automatically detect and diagnose various lumbar abnormalities from clinical scans to reduce the average time for diagnosis and help to curtail excessive burden on radiologists.

CT and MRI are two popular modalities used to diagnose causes of lower back pain. While on one hand MRI is more expensive, it is non-invasive and also much better in terms of soft tissue detailing. Fig. 1 illustrates intervertebral disc herniation diagnosed via the sagittal and axial slices of lumbar MRI. Requirements for CAD systems of the lumbar region are unique since we need to segment the dural sac and/or localize, label and segment the lumbar intervertebral discs before we can diagnose any abnormalities.

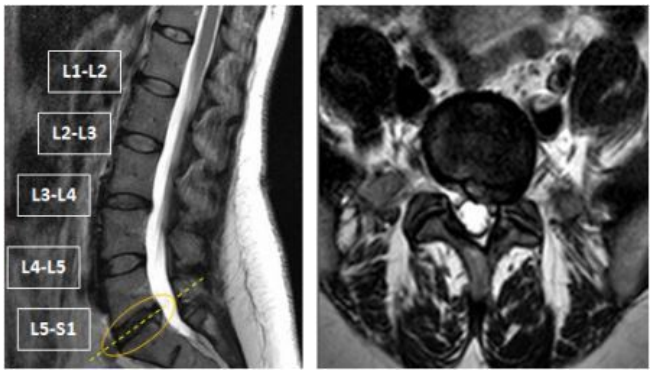



**Fig. 1.** (Left) Sagittal view of a lumbar MRI showing an L5-S1 disc herniation and (Right) the corresponding axial view of the lumbar MRI confirming a left sided herniation.

The lumbar vertebrae are the five vertebrae between the rib cage and the pelvis which are designated L1 to L5, starting at the top. The lumbar vertebrae help support the weight of the body and permit movement. The intervertebral discs are fibrocartilaginous cushions serving as the spine's shock absorbing system, which protect the vertebrae, brain, and other structures. They are named depending on the vertebral bodies above and below, e.g., the disc in between L1 and L2 is named L1-L2 and so on. Dural sac is the membranous sac that encases the spinal cord within the bony structure of the vertebral column as shown in Figure 2. The human spinal cord extends from the foramen magnum and continues through to the conus medullaris near the second lumbar vertebra, terminating in a fibrous extension known as the filum terminale. The dural sac usually ends at the vertebral level of the second sacral vertebra.

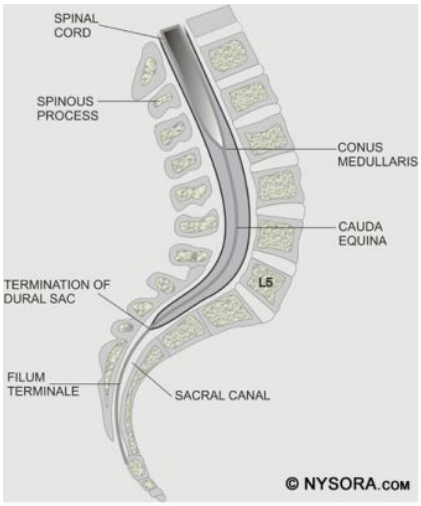



**Fig. 2.** This figure illustrates a cross section of the lumbar vertebrae and spinal cord. The position of the conus medullaris, cauda equina, termination of the dural sac and filum terminale are shown.

In general, MRI scans are very difficult to segment, since they suffer from partial volume effects and bias fields which might blur the delineation between different kinds of tissues. Moreover, localization of lumbar discs is a challenging

problem due to a wide range of variabilities in the size, shape, count and appearance of discs and vertebrae. Similarly accurately segmenting the dural sac is also difficult due to variations in grayscale values and distortion in shape due to various abnormalities like stenosis. To this end we propose an automatic method to simultaneously segment the vertebra, intervertebral discs and the dural sac of clinical sgittal MRI using the neighborhood information of each pixel. In the subsequent sections, we discuss in detail previous research (Section 2), our approach (Section 3) and experimental results (Section 4). Finally we draw our conclusion and discuss the scope for future work in Section 5.

## 2 Related Work

There has been quite some research in the direction of automatic dural sac segmentation [3–5], labeling and localization of intervertebral discs [6–10] and diagnosis of abnormalities [11] from lumbar MRI.

Schmidt et al. [6] introduced a probabilistic inference method using a part-based model that measures the possible locations of the intervertebral discs in full back MRI. They achieve upto 97% part detection rate on 30 cases. Bhole et al. [7] presented a method for automatic detection of lumbar vertebrae and discs from clinical MRI by combining tissue property and geometric information from T1W sagittal, T2W sagittal and T2W axial modalities. They achieve 98.8% accuracy for disc labeling on 67 sagittal images. Alomari et al. [8] proposed a two-level probabilistic model that captures both pixel- and object-level features to localize discs. The authors use generalized EM (Expectation Maximization) attaining an accuracy of 89.1% on 50 test cases. Oktay et al. [9] proposed another approach using PHOG(pyramidal histogram of oriented gradients) based SVM and a probabilistic graphical model and achieved 95% accuracy on 40 cases. In all these works, the authors have concentrated on localizing the vertebrae and/or intervertebral discs, i.e. they provide a point within the structure. Ghosh et al. [10] presented another approach using heuristics and machine learning methods to provide tight bounding boxes for each disc achieving 99% localization accuracy on 53 cases.

Koh et al. [3] presented an automatic method for the segmentation of the dural sac using Gradient Vector Flow Field which achieved a similarity index of 0.7 on 52 cases. Horsfield et al. [4] proposed a semi-automatic method for the segmentation of the spinal cord from MRI utilizing an active surface model to assess multiple sclerosis. Koh et al. [5] also proposed an unsupervised and fully automatic method based on an attention model and an active contour model, achieving 0.71 Dice Similarity Index on 60 cases.

## 3 Proposed Approach

In most of the previous work, segmentation of the dural sac and the intervertebral discs have been handled separately which might lead to overlapping tissue regions. Moreover, some techniques depend on shape models which might

lead to errors in case of high variability in appearance. Hence, in our proposed method, we adopt a unified approach where we simultaneously label each pixel as belonging to one of four class labels (vertebra, intervertebral disc, dural sac or background) using a probabilistic atlas and two decision trees based on the neighborhood information of each pixel.

### 3.1 Our Clinical Dataset

Clinical lumbar MRI used by our group is procured using a 3T Philips MRI scanner at Proscan Imaging Inc. It consists of manually co-registered T2 and T1 weighted sagittal views and T2 weighted axial views. We randomly pick 53 anonymized cases, all of which have one or more lumbar disc abnormalities. According to the radiologist's report there are a total of 65 herniated discs, 27 bulging discs, 26 desiccated discs, 60 degenerated discs and 73 disc levels having mild to severe stenosis.

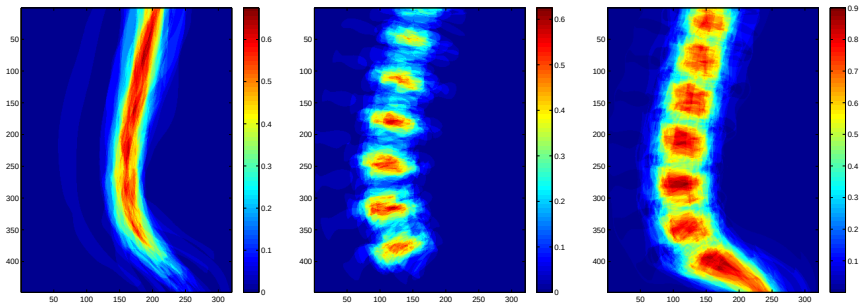
For our experiments we use the T2 weighted mid-sagittal slice, each image having a resolution of 512 x 512. We use our own labeling tool for manual segmentation, which performs B-spline interplolation to interactively give a smooth outline of segmented regions, as shown in Fig. 4 (b) and Fig. 5 (b). We randomly select 40 cases for training and the rest is kept aside for testing.

Let us denote $\boldsymbol{X} = \{x_i : i \in \{1, 2..., n\}\}$ as the set of pixel grayscale values in the mid-sagittal image. Our approach treats the segmentation of lumbar MRI as a 4-class problem where each pixel can belong to any one of the following categories : vertebra, intervertebral disc, dural sac and background. The class labels are denoted by the set $\boldsymbol{L} = \{l : l \in \{1, 2, 3, 4\}\}$ and the set of pixel labels $\boldsymbol{Y} = \{y_i : i \in \{1, 2..., n\}, y_i \in \{\boldsymbol{L}\}\}$ where $y_i$ is the output class label for the $i$th pixel.

### 3.2 Training Phase

The training phase consists of the following three steps :

**1) Creation of a probabilistic atlas** : We create a simple probabilistic atlas (probability map) by combining the label information from manual segmentation of the 40 training images as illustrated in Fig. 3. Since the vetebral column is centrally located in the 512x512 image, we avoid a registration step which can be complicated due to high variability in intensities and shapes of structures in the lumbar region. The atlas is thus a $r$ x $c$ x 4 matrix where $r$ and $c$ are the total num of rows and columns respectively and $n = r$ x $c$ is the total number of pixels. Thus the probabilistic atlas gives us : $P_{atlas_i} \propto P(y_i = l|row_i, col_i)$ where $l$ is the class label assigned to the $i$th pixel and ($row_i$,$col_i$) gives the location of the $i$th pixel in the image.



(a) (b) (c)

**Fig. 3.** Probabilistic Atlas of the lumbar region : (a), (b) and (c) shows the atlas for the dural sac, the intervertebral discs and the vertebra respectively

**2) Training a HOG tree** : We train a classification tree [12] based on a pixel's neighborhood HOG(Histogram of Oriented Gradients) [13]. HOG are feature descriptors popularly used in computer vision and image processing for the purpose of object detection. This technique counts occurrences of gradient orientation in localized portions of an image. For our experiments, given an $h$ x $w$ neighborhood around a pixel, we divide it into 3 x 3 = 9 sub-windows and fix the bin size to 9 resulting in a vector of length 81. We empirically fix $h = w = 27$ and train the HOG tree using HOG feature vectors and pixel class labels obtained from our 40 training images. The hog tree gives us :

$$P_{hogTree_i} \propto P(y_i = l | hog_{nhood_i}),$$

where $hog_{nhood_i}$ is the HOG calculated from the 27 x 27 image neighborhood around the $i$th pixel.

**3) Training a label tree** : We train another classification tree [12] based on a pixel's 27 x 27 neighborhood class labels. Hence the feature length is 27x27-1 = 728. The label tree gives us :

$$P_{labelTree_i} \propto P(y_i = l | label_{nhood_i}),$$

where $label_{nhood_i}$ is the class label information of the 27 x 27 neighborhood around the $i$th pixel.

### 3.3 Testing Phase

We implement two methods to segment our 13 test images.

**1) Method 1** : In this maximum-likelihood method we assign a class label to each pixel according to its location in the image (using the probabilistic atlas) and its neighborhood HOG information (using the HOG Tree). Mathematically, given a new image we assign a class label to each pixel as :

$$y_i = \underset{l}{\operatorname{argmax}}\, P(y_i = l|nhood_i),$$

where $P(y_i = l|nhood_i) \propto P_{hogTree_i} * P_{atlas_i}$.

**2) Method 2** : In this method we assign a class label to each pixel according to its location in the image (using the probabilistic atlas), its neighborhood HOG information (using the HOG tree) and its neighborhood label information (using the label tree). Given a new image we randomly assign a class label to each pixel and then utilise Gibbs sampling to sample a label for each pixel given its neighborhood HOG feature vector and all the other pixel labels. The update equation used is as follows :

$$P(y_i|nhood_i, \neg y_i) \propto (P_{hogTree_i} * P_{atlas_i} * P_{labelTree_i}).$$

We run Gibbs sampling for 200 iterations and use the last 100 iterations to decide the final class label, i.e. allow 100 iterations as burn in period.

**Morphological Post-Processing** : We finally apply binary morphological post-processing operations like closing, opening and hole filling on the resulting label maps to generate smoother segmentations.

## 4 Experimental Results

We use the Dice Coefficient as a Similarity Index to evaluate the validity of the automatic segmentation results. The Dice Coefficient $D(G, M)$ is defined as the ratio of twice the intersection over the sum of the two segmented results, the gold standard $G$ and our automated result $M$ :

$$Dice(G, M) = \frac{2 * n\{G \cap M\}}{n\{G\} + n\{M\}},$$

where $n\{G\}$ is the number of elements in set $G$. This measure is derived from a reliability measure known as the kappa ($\kappa$) statistic to evaluate the inter-observer agreement in regard to categorical data. According to this $D > 0.8$ indicates near-perfect agreement and $0.6 < D \leq 0.8$ represents substantial agreement and $0.4 < D \leq 0.6$ moderate agreement [14].

Tables 1 and 2 tabulate the Dice Similarity Indices of our automatic segmentation with respect to the expert manual segmentation. Table 1 lists the indices achieved by the two methods before morphological post-processing and Table 2 shows the indices after post-processing. We observe that the average results before post-processing fall in the category of 'substantial agreement', while after

**Table 1.** Results : Dice Similarity Indices before morphological post-processing

| **Case Num** | **Method 1** | | | **Method 2 (Gibbs)** | | |
|---|---|---|---|---|---|---|
| | Dural sac | Disc | Vertebra | Dural sac | Disc | Vertebra |
| 1 | 0.8203 | 0.6490 | 0.5730 | 0.8375 | 0.7271 | 0.8046 |
| 2 | 0.7949 | 0.6754 | 0.6784 | 0.8291 | 0.7622 | 0.8302 |
| 3 | 0.8245 | 0.6960 | 0.6934 | 0.8597 | 0.7938 | 0.8494 |
| 4 | 0.6467 | 0.5465 | 0.6401 | 0.6494 | 0.6136 | 0.8041 |
| 5 | 0.6918 | 0.4535 | 0.5961 | 0.6735 | 0.4716 | 0.6962 |
| 6 | 0.8136 | 0.6971 | 0.7004 | 0.8523 | 0.7799 | 0.8052 |
| 7 | 0.8312 | 0.5691 | 0.5897 | 0.8419 | 0.6189 | 0.7880 |
| 8 | 0.7436 | 0.6051 | 0.5676 | 0.7856 | 0.6804 | 0.7673 |
| 9 | 0.7506 | 0.5950 | 0.6069 | 0.7033 | 0.6431 | 0.7968 |
| 10 | 0.7433 | 0.6695 | 0.6280 | 0.8025 | 0.7803 | 0.7944 |
| 11 | 0.7416 | 0.6945 | 0.6948 | 0.7463 | 0.7510 | 0.7431 |
| 12 | 0.7710 | 0.5896 | 0.5740 | 0.8170 | 0.6769 | 0.7378 |
| 13 | 0.7024 | 0.6990 | 0.6863 | 0.6810 | 0.7689 | 0.8181 |
| Avg | 0.7597 | 0.6261 | 0.6330 | 0.7753 | 0.6975 | 0.7873 |

**Table 2.** Results : Dice Similarity Indices after morphological post-processing.

| **Case Num** | **Method 1** | | | **Method 2 (Gibbs)** | | |
|---|---|---|---|---|---|---|
| | Dural sac | Disc | Vertebra | Dural sac | Disc | Vertebra |
| 1 | 0.8765 | 0.8324 | 0.6853 | 0.8765 | 0.8023 | 0.8350 |
| 2 | 0.8881 | 0.8608 | 0.8130 | 0.8907 | 0.8595 | 0.8713 |
| 3 | 0.9107 | 0.8672 | 0.8329 | 0.9088 | 0.8703 | 0.8892 |
| 4 | 0.7419 | 0.7112 | 0.8197 | 0.7070 | 0.6849 | 0.8621 |
| 5 | 0.7844 | 0.5972 | 0.7166 | 0.7371 | 0.5704 | 0.7446 |
| 6 | 0.8985 | 0.8663 | 0.8644 | 0.9028 | 0.8533 | 0.8894 |
| 7 | 0.8960 | 0.7811 | 0.7245 | 0.8824 | 0.7362 | 0.8324 |
| 8 | 0.8427 | 0.8091 | 0.7284 | 0.8411 | 0.7955 | 0.8161 |
| 9 | 0.8522 | 0.7912 | 0.7934 | 0.7671 | 0.7532 | 0.8571 |
| 10 | 0.8272 | 0.8844 | 0.8071 | 0.8530 | 0.8728 | 0.8490 |
| 11 | 0.8090 | 0.9006 | 0.8731 | 0.7829 | 0.8479 | 0.8295 |
| 12 | 0.8843 | 0.8310 | 0.7700 | 0.8848 | 0.8165 | 0.8108 |
| 13 | 0.8167 | 0.8762 | 0.8444 | 0.7545 | 0.8530 | 0.8797 |
| Avg | 0.8483 | 0.8160 | 0.7902 | 0.8299 | 0.7935 | 0.8435 |

morphological operations the result indicates 'near-perfect agreement'. While Method 1 depends on the post- processing stage for its enhanced performance, before post-processing Method 2 (Gibbs Sampling) performs better than Method 1 since neighborhood label information is included within it. Hence, we could potentially improve its performance by designing better neighborhood masks and by adding some shape information.

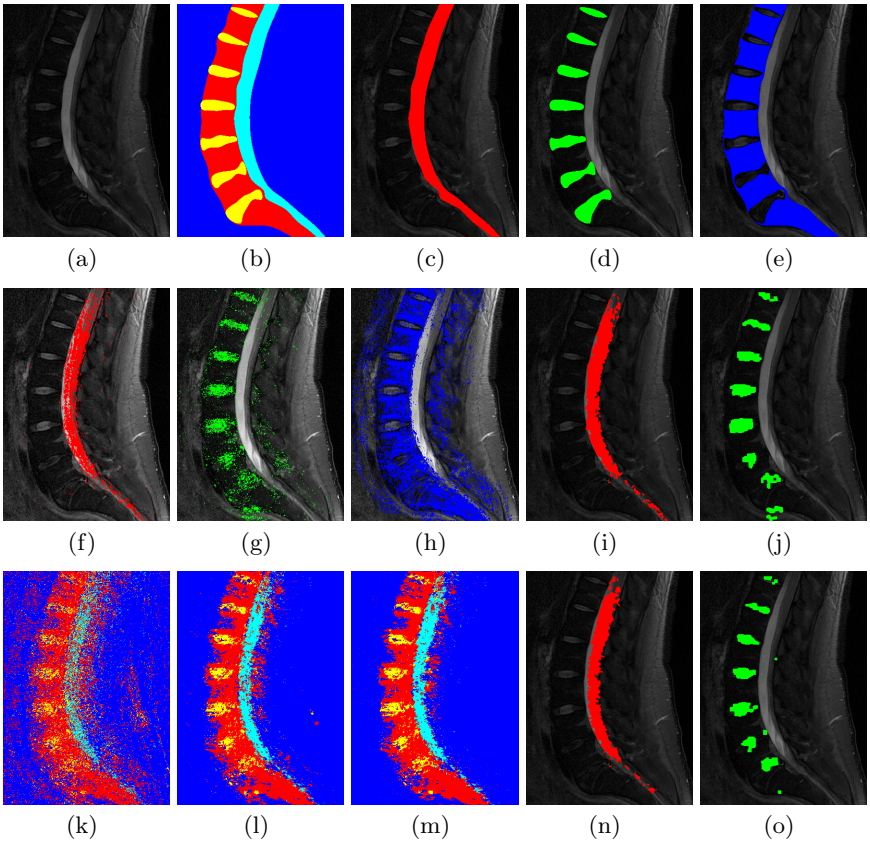



**Fig. 4.** Illustration of a challenging case showing low dice similarity indices (Test case 5): (a) shows the original mid-sagittal MRI, (b), (c), (d) and (e) show the manual segmentation (ground truth), (f), (g) and (h) show the label maps for the dural sac, disc and vertebra respectively using Method 1, while (i) and (j) show the dural sac and disc segmentation after morphological post processing. (k), (l) and (m) show the label maps generated at the end of iteration number 1, 6 and 200 respectively using Method 2 (Gibbs Sampling), while (n) and (o) show the dural sac and disc segmentation after morphological post processing.

The segmentation results of two test cases are illustrated in Figures 4 and 5. Fig. 4 illustrates a relatively challenging case (Test case 5) which shows low similarity indices in the automatic segmentation. Not only does the patient have an abnormal intervertebral disc (L5-S1), the intensity variations make automatic segmentation very difficult. Fig. 5 illustrates another case (Test case 6) which shows good automatic segmentation results (high dice similarity indices) inspite of having an abnormal intervertebral disc (L5-S1).

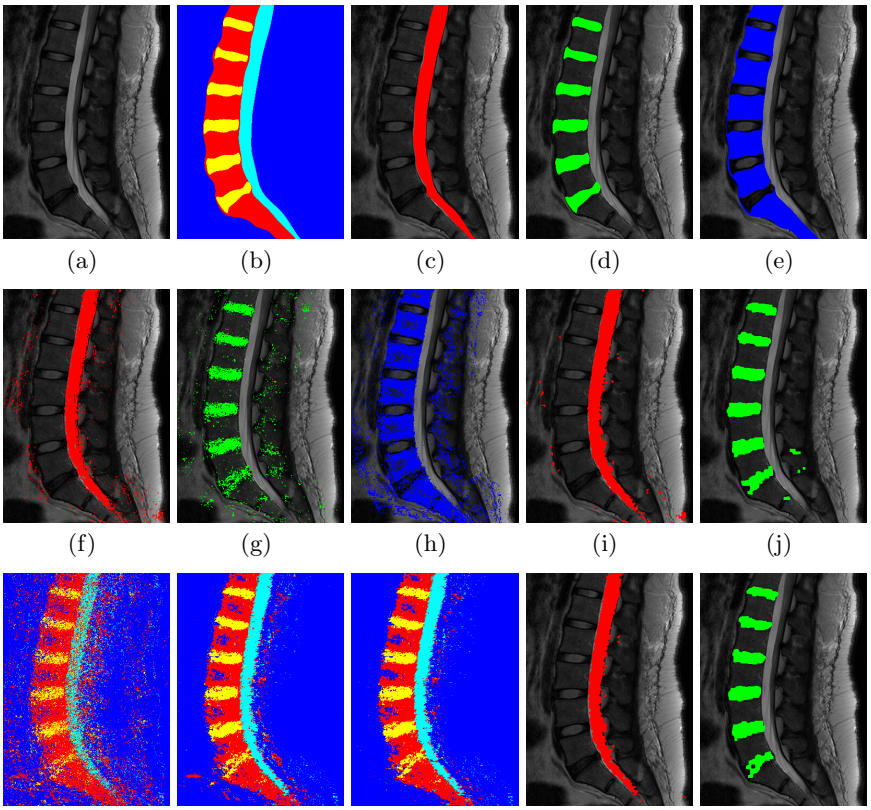



**Fig. 5.** Illustration of a case with high dice similarity indices (Test case 6): (a) shows the original mid-sagittal MRI, (b), (c), (d) and (e) show the manual segmentation (ground truth), (f), (g) and (h) show the label maps for the dural sac, disc and vertebra respectively using Method 1, while (i) and (j) show the dural sac and disc segmentation after morphological post processing. (k), (l) and (m) show the label maps generated at the end of iteration number 1, 6 and 200 respectively using Method 2 (Gibbs Sampling), while (n) and (o) show the dural sac and disc segmentation after morphological post processing.

## 5 Conclusion and Future Work

We have proposed a supervised and unified approach towards complete segmentation of lumbar MRI. Using this approach we can simultaneously segment a sagittal slice into 4 class labels : dural sac, intervertebral disc, vertebra and background. We have also provided validation of our method using 53 clinical cases out of which 40 were used for training and the rest for testing. On an average, we achieved greater than 0.8 Dice Similarity Indices for both the dural sac and the intervertebral dics. Keeping in mind our encouraging results, we propose to experiment on larger datasets and also enhance our approach by incorporating shape and better neighborhood information into our model.

**Acknowledgements** : This research was funded in part by NSF Grants DBI 0959870 and CNS 0855220 and NYSTAR grants 60701 and 41702.

## References

1. Cherry, D.K., Hing, E., Woodwell, D.A., Rechtsteiner, E.A.: National ambulatory medical care survey: 2006 summary. National Health Statistics Reports **3** (2008) 1–39
2. Bhargavan, M., Sunshine, J.H., Schepps, B.: Too few radiologists? American Journal of Roentgenology **178**(5) (2002) 1075–1082
3. Koh, J., Kim, T., Chaudhary, V., Dhillon, G.: Automatic segmentation of the spinal cord and the dural sac in lumbar mr images using gradient vector flow field. In: Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC. (2010) 2117–2120
4. Horsfield, M., Sala, S., Neema, M., Absinta, M., Bakshi, A., Sormani, M., Rocca, M., Bakshi, R., Filippi, M.: Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: Application in multiple sclerosis. Neuroimage (2010)
5. Koh, J., Scott, P.D., Chaudhary, V., Dhillon, G.: An automatic segmentation method of the spinal canal from clinical mr images based on an attention model and an active contour model. In: Proceedings of the 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI. (2011) 1467–1471
6. Schmidt, S., Kappes, J., Bergtholdt, M., Pekar, V., Dries, S., Bystrov, D., Schnoerr, C.: Spine detection and labeling using a parts-based graphical model. In: Proceedings of the 20th international conference on Information processing in medical imaging, IPMI. Volume 4584. (2007) 122–133
7. Bhole, C., Kompalli, S., Chaudhary, V.: Context-sensitive labeling of spinal structures in MRI images. In: The proceedings of SPIE medical imaging. (2009)
8. 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 Transactions on Medical Imaging **30**(1) (2011) 1–10
9. Oktay, A.B., Akgul, Y.S.: Localization of the Lumbar Discs Using Machine Learning and Exact Probabilistic Inference. In: MICCAI (3). (2011)
10. Ghosh, S., Malgireddy, M.R., Chaudhary, V., Dhillon, G.: A new approach to automatic disc localization in clinical lumbar MRI: Combining machine learning with heuristics. In: Proceedings of the 9th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI. (2012) 114–117

11. Ghosh, S., Alomari, R.S., Chaudhary, V., Dhillon, G.: Computer-aided diagnosis for lumbar mri using heterogeneous classifiers. In: Proceedings of the 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI. (2011) 1179–1182
12. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Wadsworth and Brooks, Monterey, CA (1984)
13. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: International Conference on Computer Vision & Pattern Recognition. Volume 2. (2005) 886–893
14. Kundel, H.L.: Measurement of observer agreement. In: RSNA. (2003) 303–308