COMPUTER-AIDED DIAGNOSIS FOR LUMBAR MRI USING HETEROGENEOUS CLASSIFIERS

Subarna Ghosh, Raja' S. Alomari, Vipin Chaudhary

Department of Computer Science and Engineering
University at Buffalo, SUNY
Buffalo, NY 14260

Gurmeet Dhillon, MD

Proscan Imaging Inc.
Williamsville, NY 14221

ABSTRACT

In this paper we propose a robust and fully automated lumbar herniation diagnosis system based on clinical MRI data which will not only aid a radiologist to make a decision with increased confidence, but will also reduce the time needed to analyze each case. Our method is based on three steps : 1) We automatically label the five lumbar intervertebral discs in a sagittal MRI slice using a probabilistic model and then extract an ROI for each disc using an Active Shape Model. 2) We generate relevant intensity and texture features from each disc ROI. 3) We construct five different classifiers (SVM, PCA+LDA, PCA+Naive Bayes, PCA+QDA, PCA+SVM) and combine them in a majority voting scheme. We perform 5-fold cross-validation experiments and achieve an accuracy of 94.85%, specificity of 95.9% and sensitivity of 92.45% for 35 clinical cases *i.e.* a total of 175 lumbar intervertebral discs.

Index Terms— CAD, Lumbar MRI, Lumbar herniation

1. INTRODUCTION

Healthcare facilities in the U.S. have been battling a severe shortage of radiologists for quite some time [1]. While PACS (Picture Archiving and Communication System) has solved the visualization part of the problem, a CAD (Computer Aided Diagnosis) system to generate diagnostic results from MRI and CT scans would not only reduce the burden on a radiologist, but also boost the confidence on a diagnosis. Occasionally, a CAD system, might also detect a disorder that a radiologist could have missed due to insufficient time to analyze a case. This realization motivates us to develop a highly accurate and fully automated system to detect herniation in lumbar intervertebral discs.

The term *herniated disc* initially meant a focal extension of the nucleus pulposus beyond the margin of the disc. Over time, this impractical definition gave way to the more realistic definition proposed by Herzog [2]: 'focal displacement of nuclear, annular, or end plate material beyond the normal peripheral margins of the disc delimited by the margins of the vertebral body end plates' as shown in Fig 1.

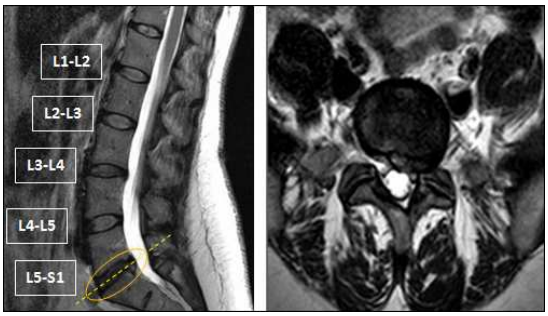

Fig. 1. (Left) Sagittal view of a lumbar MRI showing an L5-S1 disc herniation and (Right) the corresponding axial view of the lumbar MRI confirming a left sided herniation.

Currently MR imaging with its numerous modalities is the most accurate non-invasive imaging technique to diagnose a disc herniation and to determine its exact location [3]. Hence we concentrate on CAD systems from MRI as an preliminary diagnostic aid.

In this paper, we propose a robust lumbar intervertebral disc herniation detection CAD system that does not require an exact disc segmentation. We use T2-SPIR sagittal images from 35 clinical cases to extract suitable intensity and texture features from each of the five intervertebral discs in the lumbar region. Each of these cases has at least one out of five lumbar discs that is herniated as shown in Fig. 2. Otherwise, they are random with respect to age, sex, symptoms and other lumbar disorders.

We use a probabilistic model for automatic localization and labeling of the discs [4] from each sagittal slice which results in a point inside each disc. Subsequently, using the angle of the corresponding axial MRI images, we orient each disc horizontally and then we use an Active Shape Model [5] to get a bounding box of each disc. We proceed with this ROI to extract intensity and texture features from each disc. Then we construct five classifiers by running heterogeneous

learning algorithms(SVM, PCA+LDA, PCA+Naive Bayes, PCA+QDA and PCA+SVM) to detect if a disc is herniated or not. Finally, we combine them in a majority voting scheme which results in a robust diagnostic system.

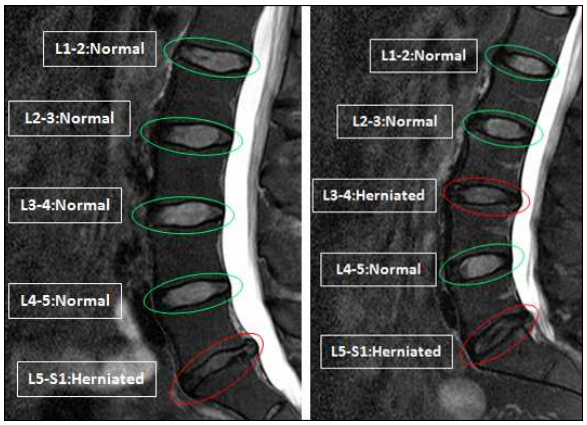

Fig. 2. Two lumbar MRI cases (T2-SPIR sagittal view) labeled with the ground truth. (Left) Case has one (L5-S1) herniated lumbar disc. (Right) Case has two (L3-L4 and L5-S1) herniated lumbar discs.

2. RELATED WORK

There has been a growing interest in the research community for automatic diagnosis of lumbar abnormalities from MRI and CT scans. Chwialkowski et al. [6] presented a method to detect lumbar pathologies in MR images. This algorithm localizes candidate vertebrae with an estimated vertebrae model, and also studies the change in gray level intensities in healthy and damaged discs. Tsai et al. [7] detects herniation from 3D MRI and CT volumes of the discs by using geometric features like shape, size and location. However, it is a computationally expensive method and serves better for visualization. Michopoulou et al. [8] showcased the classification of intervertebral discs into normal or degenerated, by using fuzzy-c means to perform semi-automatic atlas-based disc segmentation and then used a Bayesian clssifier. They achieved 86-88 % accuracy on 34 cases. The same group reported 94 % accuracy for a normal vs. degenerated discs classifier using texture features [9]. However they used 50 manually segmented discs for their experiments. In our previous work, Alomari et al. [10] presented a fully automated herniation detection system using GVF snake for an initial disc contour and then trained a Bayesian classifier on the resulting shape features. They achieved 92.5% accuracy on 65 clinical MRI cases but a low sensitivity of 86.4%.

3. PROPOSED APPROACH

Researchers have mostly concentrated on the intervertebral disc intensity levels and shape features for automatic herniation detection from sagittal MRI. This makes detection dependent on accurate segmentation of the disc. In-depth observations show that, a non-herniated disc might just as well have shape and intensity similar to a herniated one due to disc degeneration, desiccation and other abnormalities. Hence we focus on both intensity and texture features in our approach. Moreover, we try to finalize on features that are not dependent on an accurate segmentation of each disc, but rather works on a rectangular bounding box of the disc. From each disc ROI we extract a series of features as described below.

3.1. Feature Extraction

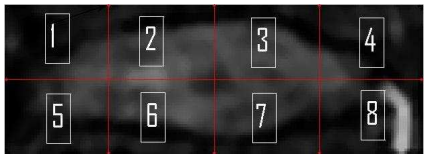

Fig. 3. The disc ROI is divided into 8 equal parts for feature extraction.

3.1.1. Intensity features

First, we calculate the general intensity features like mean, min and max intensity from the disc ROI. Then we divide the ROI into 8 parts as shown in Fig. 3 and calculate the feature set **X** as :

$$\mathbf{X} = \{ \quad \frac{I(i)}{I(j)} \quad | \quad 1 \leq i, j \leq 8 \text{ and } i \neq j \ \} \qquad (1)$$

where $I(i)$ is the average intensity of the ith part of the ROI and $\mathbf{X} = < x_f >$; $1 \leq f \leq 56$. These contextual intensity features are very essential to discriminate between a herniated and a non-herniated disc.

3.1.2. Shape features

We use an important, but simple shape feature R given by $R = a/b \approx w/h$, where a and b are the lengths of the major axis and the minor axis of the disc, respectively; whereas w and h are the width and height of the disc ROI bounding box, respectively. We do not use any other contour features that can be generated from the ASM step, since we deliberately avoid a precise segmentation of each disc.

3.1.3. Texture features

We calculate a series of texture features from the GLCM (Gray level co-occurrence) matrix of the disc ROI in 8 directions. Mathematically, a GLCM matrix G is defined over an n x m image I, parameterized by an offset $(\Delta x, \Delta y)$ as :

$$G_{\Delta x \Delta y}(i,j) = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 1, & \text{if } I(p,q) = i \text{ and} \\ & I(p+\Delta x, q+\Delta y) = j \\ 0, & otherwise \end{cases} \quad (2)$$

We calculate five well-known texture features from the normalized GLCM G_n : Contrast, Correlation, Energy, Homogeneity and Entropy [11]. Similar features are also generated from the right quarter of the disc ROI marked as 4 and 8 as shown in Fig. 3. This step is necessary because a posterior herniation has distinct seepage near the spinal sac and hence distinct texture features.

3.2. Classification

After features are extracted from the disc ROIs, we build six individual classfiers using existing dimensionality reduction techniques and modeling methods. The first classifier is an SVM (Support Vector Machine) [12, 13], implemented using a linear kernel. The second classifier uses PCA(Principal Component Analysis) for dimensionality reduction followed by LDA (Linear Discriminant Analysis) as classifier. Similarly the third, fourth and fifth ones are a combination of PCA and a Naive Bayes Classifier; a combination of PCA and QDA (Quadratic Discriminant Analysis) as classifier and a combination of PCA and an SVM classifier, respectively. The sixth classifier is a KNN (k Nearest Neighbor) classifier, where k has been empirically fixed to 5. Finally, we construct a majority voting classifier as described in the following section.

4. EXPERIMENTS AND RESULTS

4.1. Experimental Setup

We experiment on 35 lumbar MRI cases which have corresponding ground truth in the form of a radiologist's report. We perform 5-fold cross-validation experiments on each of our six classifiers and observe that while the first five show well above 90% accuracy, the kNN classifier shows around 80-85% accuracy. Hence, we construct a majority voting classifier, by combining the results from our first five classifiers (i.e. SVM, PCA+LDA, PCA+Naive Bayes, PCA+QDA and PCA+SVM) and selecting the label given by three or more classifiers as the final result. Majority voting not only guarantees better performance than the worst individual classifier, it also avoids classifier evaluation steps used in complex fusion algorithms. It has also been shown experimentally [14] and mathematically [15] that a majority voting classifier can show improvement over individual classifier accuracy. The performance results of all our classifiers are shown in Fig. 4 and Table 1.

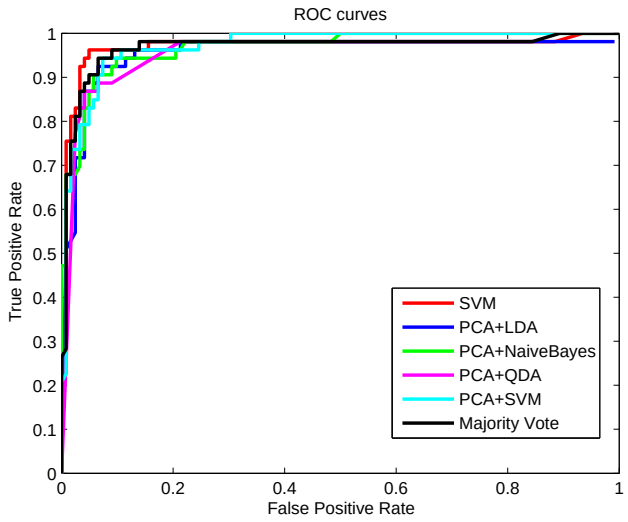

Fig. 4. ROC for the individual classifiers and the Majority Vote classifier.

The performance metrics, specificity and sensitivity are defined as follows :

$$\text{Specificity} = \frac{\text{TNs}}{\text{TNs + FPs}} \quad \text{Sensitivity} = \frac{\text{TPs}}{\text{TPs + FNs}} \quad (3)$$

where TNs is the Number of True Negatives, FNs is the Number of False Negatives, TPs is the Number of True Positives. FPs is the Number of False Positives. The x-axis and y-axis of the ROC curve in Fig. 4, are the False Positive Rate(FPR) and the True Positive Rate(TPR) respectively, defined as:

$$\text{FPR} = 1 - \text{Specificity} \quad \text{and} \quad \text{TPR} = \text{Sensitivity} \quad (4)$$

Table 1. Classifier performance results in percantage for 5-fold cross validation

Classifier	Accuracy	Specificity	Sensitivity
SVM	94.29	96.72	88.68
PCA+LDA	93.14	93.44	92.45
PCA+Bayes	92.0	92.62	90.57
PCA+QDA	92.57	95.08	86.79
PCA+SVM	93.14	95.08	88.68
5-NN	80.57	89.34	60.37
Majority Vote	94.86	95.90	92.45

4.2. Discussion

Amongst the individual classifiers, we observe that, SVM, PCA+LDA and PCA+SVM have better ROC curves (Fig. 4)

and higher accuracies (Table 1) than PCA+Naive Bayes and PCA+QDA. Also, PCA+LDA and PCA+Naive Bayes shows high sensitivity (or low FNs) which is very essential for a lumbar diagnosis system. This is because, while high FPs can be quickly rectified by the radiologist, high FNs might lead to a herniated disc not being diagnosed at all, and hence a greater penalty. Thus a majority voting scheme combining these five individual classifers is a good choice to build a high accuracy system, that maintains a high sensitivity as well. From Table 1 we observe that Majority Vote has a higher sensitivity than SVM, ie. SVM it has too many undesirable FNs and the majority voting scheme helps to bring that number down. Hence we see that Majority Voting Classifier not only shows a slightly increased accuracy (94.86%) compared to the individual classifiers, it also maintains a high value of sensitivity (92.45%) and specificity (95.9%).

5. CONCLUSION

We have proposed a fully automated system to detect herniated discs from sagittal lumbar MRI using robust intensity and texture features. The major advantage of this system is that, it does not require precise segmentation of the lumbar intervertebral discs. Also it does not require the MRI sagittal slice to be in the middle with a perfect view of the spinal sac. Moreover, this approach extracts good features which is evident from the high accuracies of the individual classifiers. Another added advantage is the fact that the final majority voting classifier not only shows a high accuracy and sensitivity; it also boosts the confidence of the diagnosis. As an extension to this approach, we will be working on associating information from the MRI axial modality to further decrease the diagnostic error rates. Moreover, based on the disc herniation location, we will also work on automatic classification of herniation into lateral right, central and lateral left herniation from MRI axial views.

6. REFERENCES

[1] M. Bhargavan, J. H. Sunshine, and B. Schepps, "Too few radiologists?," *American Journal of Roentgenology*, vol. 178(5), pp. 1075–1082, 2002.

[2] R.J. Herzog, "The radiologic assessment for a lumbar disc herniation," *Spine*, vol. 21, 1996.

[3] G. Buirski and M. Silberstein, "The symptomatic lumbar disc in patients with low-back pain: Magnetic resonance imaging appearances in both a symptomatic and control population," *Spine*, vol. 18, 1993.

[4] R. S. Alomari, J. J. Corso, and V. Chaudhary, "Labeling of lumbar discs using both pixel- and object-level features with a two-level probabilistic model," *IEEE Transactions on Medical Imaging*, 2010, in press.

[5] T. Cootes, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[6] M. P. Chwialkowski, P. E. Shile, R. M. Peshock, D. Pfeifer, and R. W. Parkey, "Automated detection and evaluation of lumbar discs in mr images.," *IEEE Engineering in Medicine and Biology*, vol. 2527-2530, 1989.

[7] M. Tsai, S. Jou, and M. Hsieh, "A new method for lumbar herniated inter-vertebral disc diagnosis based on image analysis of transverse sections," *Computerized Medical Imaging and Graphics*, vol. 26, no. 6, pp. 369–380, 2002.

[8] S. Michopoulou, L. Costaridou, E. Panagiotopoulos, R. Speller, G. Panayiotakis, and A. Todd-Pokropek, "Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine," in *IEEE Transactions on Biomedical Engineering*, 2009, vol. 56, pp. 2225–31.

[9] S. Michopoulou, I. Boniatis, L. Costaridou, D. Cavouras, E. Panagiotopoulos, and G. Panayiotakis, "Computer assisted characterization of cervical intervertebral disc degeneration in mri," *Journal of Instrumentation*, vol. 4, pp. 287–293, 2009.

[10] R. S. Alomari, J. J. Corso, V. Chaudhary, and G. Dhillon, "Toward a clinical lumbar cad: herniation diagnosis.," *International Journal of Computer Aided Radiology and Surgery*, 2010.

[11] R. M. Haralick, "Statistical and structural approaches to texture," in *IEEE Proc.*, 1979, vol. 67, pp. 786–804.

[12] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[13] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM: a library for support vector machines*, 2001.

[14] Muhammad A. Khany, Zahoor Jan, and Anwar M. Mirzaz, "Performance analysis of classifier fusion model with minimum feature subset and rotation of the dataset," in *Proc. of 6th international conference on Fuzzy systems and knowledge discovery*, 2009, pp. 251–255.

[15] L. I. Kuncheva, C. J. Whitaker, and R. P. W. Duin, "Limits on the majority vote accuracy in classifier fusion," *Pattern Analysis and Applications*, vol. 6, pp. 22–31, 2003.