

# Composite Features for Automatic Diagnosis of Intervertebral Disc Herniation from Lumbar MRI

Subarna Ghosh, Raja' S. Alomari, Vipin Chaudhary and Gurmeet Dhillon

**Abstract**—Lower back pain is widely prevalent in the world today, and the situation is aggravated due to a shortage of radiologists. Intervertebral disc disorders like desiccation, degeneration and herniation are some of the major causes of lower back pain. In this paper, we propose a robust computer-aided herniation diagnosis system for lumbar MRI by first extracting an approximate Region Of Interest (ROI) for each disc and then using a combination of viable features to produce a highly accurate classifier. We describe the extraction of raw, LBP (Local Binary Patterns), Gabor, GLCM (Gray-Level Co-occurrence Matrix), shape, and intensity features from lumbar SPIR T2-weighted MRI and also present a thorough performance comparison of individual and combined features. We perform 5-fold cross validation experiments on 35 cases and report a very high accuracy of 98.29% using a combination of features. Also, combining the desired features and reducing the dimensionality using LDA, we achieve a high sensitivity (true positive rate) of 98.11%.

## I. INTRODUCTION

According to the American Academy of Orthopedic Surgeons (AAOS), four out of five adults experience lower back pain at some point during their lives and many of them have common intervertebral disc disorders.

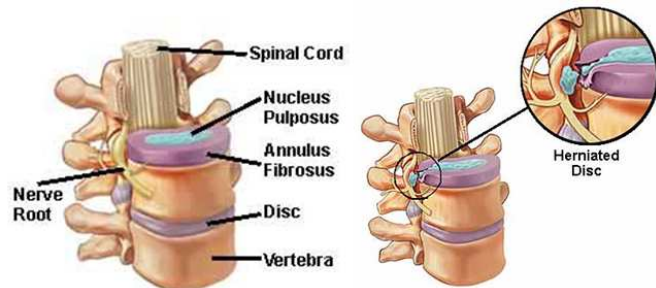


Fig. 1. This figure shows a labeled portion of spine on the left and a magnified view of intervertebral disc herniation on the right.

Intervertebral discs are soft, rubbery pads found between the vertebrae of the spinal column that provide body flexibility. Discs in the lumbar spine (lower-back) are composed of a thick outer ring of cartilage (annulus) and an inner gel-like substance (nucleus) as shown in Fig. 1. A disc herniates when part of the nucleus pushes through the outer edge of the disc and back toward the spinal canal due to

severe trauma, strain or intervertebral joint degeneration. This puts pressure on the nerves leading to pain and change of posture. Statistics show that one-third of adults over the age of 20 show evidence of herniated discs [1] and 90% of herniation occurs in the lumbar and lumbosacral regions of the spine [2]; hence we are motivated to develop a robust and highly accurate system for automatic diagnosis of lumbar disc herniation from lumbar MRI which can not only provide quick screening for patients, but might also detect cases that a radiologist missed due to lack of time [3].

## II. PREVIOUS WORK

Automatic detection of abnormalities from MRI or CT scans has been an active research area over this last decade. The challenges are manifold - ranging from variations in scanner specifications, parameter settings, modalities, differences in body structure and composition, and last but not the least the task of segmentation which is a big challenge in computer vision. Chwialkowski et al. [4] presented a method to detect lumbar pathologies in MR images by first localizing candidate vertebrae with an estimated vertebrae model and then studying the change in gray level intensities in healthy and damaged discs. Tsai et al. [5] detects herniation from 3D MRI and CT volumes of the discs by using geometric features like shape, size and location. Michopoulou et al. [6] achieved 86-88% accuracy for normal vs. degenerated disc classification. The used fuzzy-c means to perform semi-automatic atlas-based disc segmentation and then used a Bayesian classifier. They also reported 94% accuracy using texture features [7] for 50 manually segmented discs. Alomari et al. [8] 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 achieve 92.5% accuracy on 65 clinical MRI cases but a low sensitivity of 86.4%. In our previous work [9], we have used heterogeneous classifiers to achieve an accuracy of 94.85% and sensitivity of 92.45% on 35 cases in a fully automated scheme.

## III. OUR APPROACH

### A. MRI Dataset Used

T2-SPIR Sagittal images from lumbar MRI scans are used for our experiments. They are acquired using a 3 Tesla Philips scanner in clinical settings. Thirty-five anonymized cases are selected such that each case has one or more herniated lumbar disc. Otherwise, they are random with respect to age, sex, symptoms and other lumbar disorders. Radiologist's reports are treated as the ground truth. We use

Subarna Ghosh, Raja' S. Alomari and Vipin Chaudhary are affiliated to the Department of Computer Science and Engineering, State University of New York at Buffalo, NY 14260, USA.

Gurmeet Dhillon(MD) is a Radiologist with ProScan Imaging Inc., Williamsville, NY 14221, USA.

80% of the dataset (i.e. 140 discs) for training and the rest for testing in 5-fold cross-validation experiments.

### B. Automatic disc ROI extraction

We use a probabilistic model for automatic localization and labeling of the discs [10] from each mid-sagittal slice which results in a point inside each disc. We use the label point inside each disc as the initial starting point for our Active Shape Model [11] based segmentation as shown in Fig. 2. Using the ASM boundary we construct an approximate ROI for each disc by adding a few pixels in height and width to the tight bounding box of the disc. This takes care of the cases where ASM does not provide a satisfactory segmentation.

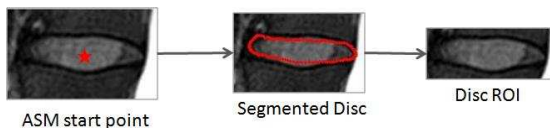


Fig. 2. This figure shows the process of disc ROI extraction starting from labeling each disc in the sagittal MRI, using the label point as the starting point of ASM and finally extracting a bounding box of the disc.

### C. Feature extraction

Literature survey shows that there has not been comprehensive work done on the comparison of feature performance for automatic herniation detection in lumbar MRI scans. Hence we focus on the extraction of various discriminative features like raw, LBP(Local Binary Patterns), Gabor, GLCM (gray-level co-occurrence matrix), intensity and shape features to study their individual and combined performances. The first three features (Raw, LBP and Gabor) gives us values at each pixel of the disc ROI, hence we resize the each disc image into a 20x50 block to maintain a constant feature vector length.

1) *Raw Features*: Raw features are the original pixel intensity information in the disc ROI, vectorized to create the raw feature vector of length 1000.

2) *Local Binary Patterns*: LBP or Local Binary Patterns are a type of feature commonly used for texture classification. For an image  $I$  the pixel-wise LBP is defined as:

$$LBP(x, y) = \sum_{i=0}^7 s(I_i(x, y) - I(x, y)) \cdot 2^i \quad (1)$$

$$\text{such that: } s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Here  $I(x, y)$  is the center pixel and  $I_i(x, y)$  denotes each of the eight neighboring pixels.

3) *Gabor Features*: Gabor features are extracted by convolving Gabor filters with the sample images. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave as:

$$gb(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \quad (3)$$

where

$$x' = x\cos\theta + y\sin\theta \quad (4)$$

$$y' = -x\sin\theta + y\cos\theta \quad (5)$$

Here,  $\lambda$  represents the wavelength of the sinusoidal factor,  $\theta$  represents the orientation of the normal to the parallel stripes of a Gabor function,  $\psi$  is the phase offset,  $\sigma$  is the variance of the Gaussian envelope, and  $\gamma$  is the spatial aspect ratio.  $\gamma$  also specifies the ellipticity of the support of the Gabor function.



Fig. 3. Gabor filter bank used for Normalized Gabor feature.

For our experiments, we create a filter bank of eight Gabor filters with a constant scale of  $\lambda = 3$  and eight orientations ( $\theta = [0 : \pi/8 : 7 * \pi/8]$ ) as shown in Fig. 3. For each disc, we calculate the L2-norm of the superimposed responses which are vectorized to give the Gabor feature vector of length 1000.

4) *GLCM features*: We create the GLCM feature vector exactly as discussed in our previous work [9].

5) *Intensity and Shape features*: We calculate the intensity and shape features as discussed in [9] and concatenate them to form the final feature vector.

### D. Dimensionality Reduction and Classification

We use two popular dimensionality reduction techniques: Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA) in our experiments. On one hand, PCA preserves as much of the variance in the high dimensional space as possible. On the other hand, LDA preserves as much of the class discriminatory information as possible. In our experiments, the within class scatter matrix is often singular and not full rank, so, we first reduce the dimensionality by PCA and then apply LDA. Also, LDA reduces the dimensionality to  $k$  dimensions, such that  $k = n - 1$ , where  $n$  is the number of classes. For our problem, we deal with 2 classes: herniated (positive class) and non-herniated (negative class) disc and hence  $k = 1$ .

We use three popular classifiers for herniation detection: k-Nearest Neighbor (kNN), linear Support Vector Machine (SVM) [12] and Naive Bayes Classifier. For kNN we empirically fix  $k$  as 5.

## IV. EXPERIMENTS AND RESULTS

We divide our 35 cases into 5 non-overlapping folds, each consisting of 5 cases i.e.  $7 * 5 = 35$  lumbar discs to perform 5-fold cross validation experiments. Thus, we ensure that the testing and training datasets are always distinct. Tables I and II show the performance results of the individual features (Section III-C). For each row entry in the tables, the classifier

TABLE I  
PERFORMANCE OF RAW, LBP AND GABOR FEATURES (5-FOLD CROSS VALIDATION)

Classifier	Raw Features			LBP Features			Gabor Features		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
PCA64+5NN	94.29	97.54	86.79	69.71	100	0	86.29	92.62	71.70
PCA64+SVM	79.43	86.07	64.15	84	86.89	77.36	90.28	90.98	88.68
PCA64+Bayes	95.43	95.90	94.34	91.43	91.80	90.57	91.43	90.98	92.45
PCA64+LDA+5NN	96.0	96.72	94.34	93.14	92.62	94.62	83.43	79.51	92.45
PCA64+LDA+SVM	88.57	85.25	96.23	85.71	80.33	98.11	92	94.26	86.79
PCA64+LDA+Bayes	97.71	99.18	94.34	94.86	95.90	92.45	91.43	93.44	86.79

TABLE II  
PERFORMANCE OF GLCM AND SHAPE+INTENSITY FEATURES IN PERCENTAGE (5-FOLD CROSS VALIDATION)

Classifier	GLCM Features			Classifier	Shape+Intensity Features		
	Accuracy	Specificity	Sensitivity		Accuracy	Specificity	Sensitivity
PCA8+5NN	74.86	87.70	45.28	PCA32+5NN	75.43	92.62	35.85
PCA8+SVM	81.14	88.58	64.15	PCA32+SVM	90.28	92.62	84.91
PCA8+Bayes	84.0	81.15	90.57	PCA32+Bayes	94.29	95.08	92.45
PCA8+LDA+5NN	84	86.89	77.36	PCA32+LDA+5NN	93.14	94.26	90.57
PCA8+LDA+SVM	69.71	100	0	PCA32+LDA+SVM	88.0	98.36	64.15
PCA8+LDA+Bayes	85.71	85.25	86.79	PCA32+LDA+Bayes	94.86	96.72	90.57

column can be explained as the dimensionality reduction method followed by the reduced dimension and then the type of classifier used. For example, PCA64+Bayes means that PCA has been used to reduce dimensionality to 64, then a Naive Bayes classifier is used.

We report results for composite features in Table III. We combine the features in two ways: for the first set of experiments (named Concatenated Features), we concatenate all the features in their original form, then perform dimensionality reduction and classification. In the second set of experiments (named Concatenated PCA-reduced Features), we concatenate all the PCA reduced features, then perform dimensionality reduction and classification.

We use both specificity and sensitivity as performance metrics:

$$\text{Specificity} = \frac{\text{TNs}}{\text{TNs} + \text{FPs}} \quad (6)$$

$$\text{Sensitivity} = \frac{\text{TPs}}{\text{TPs} + \text{FNs}} \quad (7)$$

where TNs is the Number of True Negatives, FN is the Number of False Negatives, TPs is the Number of True Positives, and FPs is the Number of False Positives. The x- and y-axis of the ROC curves in Fig. 4, are the False Positive Rate (FPR) and the True Positive Rate (TPR), respectively, defined as:  $\text{FPR} = 1 - \text{Specificity}$  and  $\text{TPR} = \text{Sensitivity}$ .

## V. DISCUSSION

We find that raw features (Table I) perform very well, specially for LDA+Bayes closely followed by the shape+intensity features (Table II). The Gabor and GLCM features do not perform poorly on their own. We also observe that SVM shows lower accuracies than the other classifiers, probably because we did not have a separate validation set

and used default parameters. The 5NN classifier performs well specially when LDA is used for dimensionality reduction. In general, LDA+Bayes seems to show the best performance amongst all the classifiers.

Combining the features by concatenating them boosts the overall performance of the classifiers as shown in Table III. Moreover, we also see that composite features substantially boosts the sensitivity of the classifiers. A robust diagnostic system should not only show a high accuracy, but also a high sensitivity. This is because, while False Positive instances can be quickly rectified by the radiologist, False Negatives might lead to a herniated disc not being diagnosed at all, hence posing a greater penalty. LDA+Bayes and LDA+5NN shows the best performance, with accuracy up to 98.29% and sensitivity upto 98.11%.

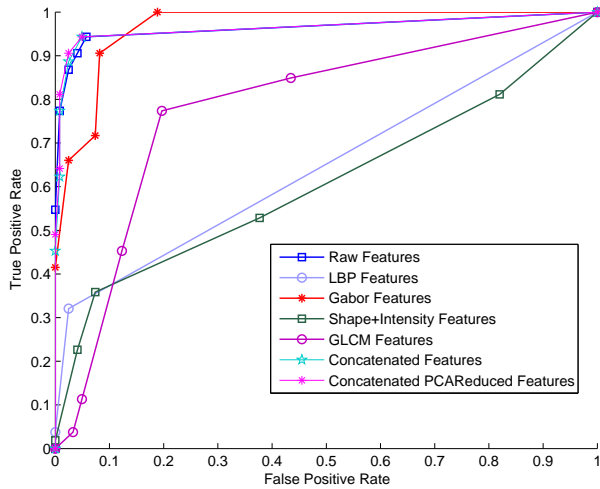
We perform a thorough feature comparison through ROC curves in Fig. 4. To plot ROC for kNN, we vary the threshold from 1 to  $k$  (instead of using the majority rule) and obtain the curves in Fig. 4. Each curve consists of five points in our case. We find that combined features show better ROC compared to individual ones in all the four curves. Comparing Fig. 4(a) with Fig. 4(b), we can clearly see that LDA as a dimensionality reduction works better than PCA.

## VI. CONCLUSIONS AND FUTURE WORK

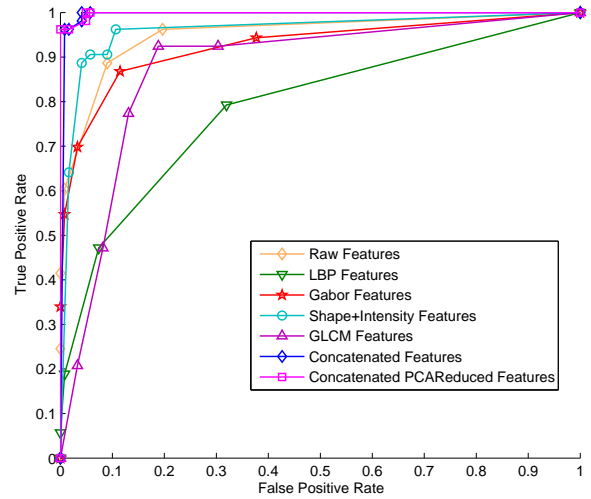
We proposed a fully automated robust system to detect herniated discs from sagittal lumbar MRI by first extracting an approximate ROI (Region Of Interest) for each disc and then using a combination of features to produce a highly accurate classifier. We also performed a thorough analysis of five kinds of features *i.e.* raw, LBP, Normalized Gabor, GLCM, shape, and intensity features; both individually and combined. This leads to the conclusion that concatenating

TABLE III  
PERFORMANCE OF COMPOSITE FEATURES

Classifier	Concatenated Features			Concatenated PCA-reduced Features		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
PCA64+5NN	92.57	99.18	77.36	93.71	99.18	81.13
PCA64+SVM	82.28	86.07	73.58	84.0	85.25	81.13
PCA64+Bayes	94.29	94.26	94.34	94.29	94.26	94.34
PCA32+LDA+5NN	96.57	95.90	98.11	96.0	95.08	98.11
PCA32+LDA+SVM	92.0	88.52	100	91.43	87.7	100
PCA32+LDA+Bayes	98.29	99.18	96.23	98.29	99.18	96.23



(a) Using PCA for dimensionality reduction



(b) Using LDA for dimensionality reduction

Fig. 4. ROC curves for individual and combined features using PCA and LDA for dimensionality reduction respectively and 5NN as classifier.

and combining features gives a better accuracy and simultaneously pulls up the sensitivity. By avoiding an accurate segmentation of the discs, we circumvent a complex problem in computer vision and by extracting desirable and robust features, we present an automatic diagnostic system showing up to 98.29% accuracy and 98.11% sensitivity. We are currently working on associating information from axial MRI slices to detect and localize disc herniation. In addition, we propose to verify our composite features on larger datasets of lumbar MRI to prove its utility in clinical settings.

#### REFERENCES

- [1] "Well-connected report: Low back pain and sciatica," 2000.
- [2] NINDS, "National institute of neurological disorders and stroke : Low back pain fact sheet," 2008.
- [3] M. Bhargavan, J. H. Sunshine, and B. Schepps, "Too few radiologists?," *American Journal of Roentgenology*, vol. 178(5), pp. 1075–1082, 2002.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] S. Ghosh, R. S. Alomari, V. Chaudhary, and G. Dhillon, "Computer-aided diagnosis for lumbar mri using heterogeneous classifiers," in *IEEE International Symposium on Biomedical Imaging*, 2011.
- [10] 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.
- [11] T. Cootes, "Active shape models-their training and application.," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [12] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM: a library for support vector machines*, 2001.