# ABNORMALITY DETECTION IN LUMBAR DISCS FROM CLINICAL MR IMAGES WITH A PROBABILISTIC MODEL

*Raja' S. Alomari, Jason J. Corso, Vipin Chaudhary**

Computer Science and Engineering Dept.
University at Buffalo
Buffalo, NY 14260

*Gurmeet Dhillon, MD*

Proscan of Buffalo
Buffalo, NY 14221

## ABSTRACT

*Purpose:* Detection of abnormal discs from clinical T2-weighted MR scans. This aids the radiologist as well as subsequent CAD methods in focusing only on abnormal discs for further diagnosis. Furthermore, it gives a degree of confidence about the abnormality of the intervertebral discs that helps the radiologist in making his decision.

*Material and Methods:* We propose a probabilistic model for detection of abnormality of intervertebral discs. We use three features to label abnormal discs that includes appearance, location, and context. We model the abnormal disc appearance with a Gaussian model, the location with a 2D Gaussian model, and the context with a Gaussian model for the distance between abnormal discs. We use clinical T2-weighted MR volume for each case and inference on the middle slide of each volume. These MR scans are specific for the lumbar area. The ground truth is provided by our collaborating radiologist.

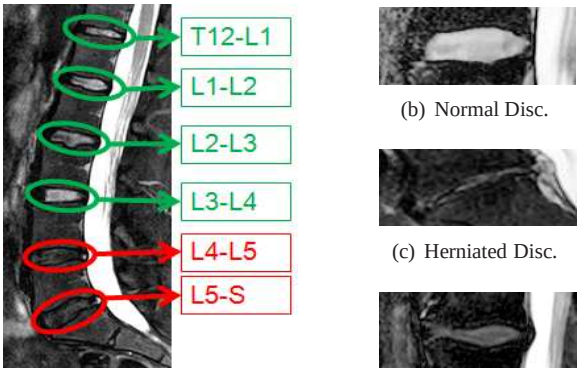
*Results:* We achieve over 91% abnormality detection accuracy in a cross-validation experiment with 80 clinical cases. The experiment runs ten rounds, in every round 30 cases are randomly left out for testing and the rest are used for training.

*Conclusion:* We achieve high accuracy for detection of abnormal discs using our proposed model that incorporates disc appearance, location, and context. We show that our proposed model is extensible for subsequent diagnosis tasks specific to each intervertebral disc abnormality such as desiccation, stenosis, and herniation.

***Index Terms*—** Computer Aided Diagnosis, MRI, lumbar intervertebral disc, Gibbs Distribution.

## 1. INTRODUCTION

Back pain is the second most common neurological ailment in the United States after the headache according to the National Institute of Neurological Disorders and Stroke (NINDS). Americans spend at least $50 billion each year on low back pain and over 12 million Americans have some sort of Intervertebral Disc Disease (IDD) [1]. Increasing demand on diagnosis of back pain diseases justifies seeking full or partial automation of the diagnosis process which usually consists of two main steps: localization of the intervertebral discs and then diagnosis of abnormalities at every disc level. The focus of this paper is on the detection of abnormality in the lumbar area from MRI.




(a) Lumbar disc levels labels. Abnormal lower two levels at $\mathcal{L}4$-$\mathcal{L}5$ and $\mathcal{S}$-$\mathcal{L}5$ (red) and the upper four are normal (green).



(b) Normal Disc.



(c) Herniated Disc.



(d) Degenerative Disc Disease.

**Fig. 1**. Labeling of lumbar area discs and sample abnormalities.

In our previous work [2], we have developed a probabilistic model for localization of the six discs in sagittal T2-weighted MR images for the lumbar area. In our model, we incorporate two levels of information: low- and high-level. In the low-level, we model the local pixel properties of discs, such as appearance. In the high-level, we capture the object-level geometrical and contextual relationships between discs. We estimate the model parameters from manually labeled cases (supervised learning). We tested our model using a dataset of 20 normal cases and showed the extension to an abnormal case. However, in this paper, we use a dataset of 80 clinical cases that contains wide variability in types of abnormalities, patient ages (17 to 81 years old), and patient heights which affect the size and appearance of the discs. Fig. 1(a) shows a sample sagittal view with labeled lumbar disc levels.

In this paper, we propose a method for detection of

*Send correspondence to ralomari@buffalo.edu.

abnormal discs $n_i^*$ in the lumbar area at each disc level $i$ (Fig. 1(a)):

$$n_i^* = \arg\max_{n_i} P(n_i|d_i, \sigma_{\mathtt{I}(d_i)}) \quad (1)$$

where $n_i$ is a binary random variable stating whether it is a normal or abnormal disc and $n_i \in \mathcal{N} = \{n_i : 1 \leq i \leq 6\}$, $d_i \in \mathcal{D} = \{d_i : 1 \leq i \leq 6\}$ is the location of each lumbar disc, and $(\sigma_{\mathtt{I}(d_i)})$ is the intensity of a neighborhood surrounding the disc level $(i)$.

Because abnormal discs vary in characteristics depending on the type of abnormality, our model has the flexibity to model these characteristics. For example, abnormal discs vary in size, shape, height and depend on patient age, patient height and many other issues that help the radiologist decide the abnormality condition of each disc. We can incorporate these variations by incorporating a model for each characteristic of interest.

The remainder of this paper is organized as follows. The background and related work is discussed in section 2. Then we discuss our proposed model in section 3. We then prsent our dataset and the experimental results in sections 4 and 5, respectively. Discussions and future work are presented in section 6.

## 2. BACKGROUND

### 2.1. Abnormalities in the Intervertebral Discs

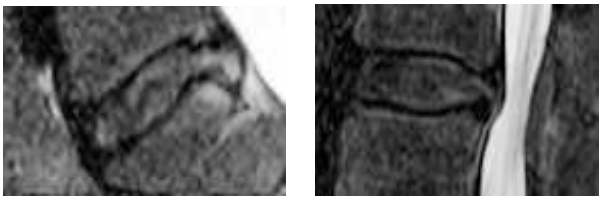
Intervertebral discs are unique structures that absorb shocks between adjacent vertebrae. They act as the ligaments that connect the vertebrae together and the pivot point which allows the spine mobility by bending and rotating. They make about one fourth of the spinal columns length [3].

An inter-vertebral disc is composed of two parts: an outer strong ring called annulus fibrosis and a soft gel-like inner called nucleus pulposus. The nucleus pulposus consists of 80% to 85% water in normal cases. In the lumbar area, there are six discs connected to the five lumbar vertebrae which are labeled top-down as $\mathcal{T}12 - \mathcal{L}1$, $\mathcal{L}1 - \mathcal{L}2$, $\mathcal{L}2 - \mathcal{L}3$, $\mathcal{L}3 - \mathcal{L}4$, $\mathcal{L}4 - \mathcal{L}5$, and $\mathcal{L}5 - \mathcal{S}$ as shown in Fig. 1(a).

Diseases that originate from an intervertebral disc abnormality are the most common diseases in the vertebral column. Most common diseases are: disc herniation, spinal stenosis, degenerative disc disease, disc desiccation, and spinal infection [3].

**Disc herniation** (Fig. 1(c)) is a leak of the nucleus pulposus through a tear in the wall of the annulus fibrosis. This leak presses on the local nerve root causing the pain. The tear in the disc wall usually occur due to aging and/or trauma injury [4, 3].

**Spinal stenosis** is narrowing of the spinal canal and might be caused from different conditions such as disc herniation, osteoporosis, or a tumor. Sometimes, and especially when the reason is a disc herniation, stenosis occurs at same level of the disc. [4, 3].



(a) $\mathcal{L}5 - \mathcal{S}$ disc level: Herniation, DDD, mild Foramina stenosis.

(b) $\mathcal{L}3 - \mathcal{L}4$ disc level: DDD, central stenosis, and central herniation

**Fig. 2**. Sample diagnosis of two Discs with multiple diseases.

**Degenerative disc disease** (Fig. 1(d)) is the gradual deterioration of the disc causing loss of its functions. This disease usually develops over aging or from continuous activities that presses on a disc space. It starts with a small injury in the annulus fibrosis causing damage to the nucleous pulposis and loss of its water contents. Furthur damage causes malfunctioning of the disc and thus collapsing the upper and lower vertebrae. As time passes, the vertebra facet joints twist creating bone spures that grow into the spinal canal and pinching the nerve root (stenosis) [4, 3].

**Disc desiccation** is the drying out of the water contents in the inner pulposis. Usually, it is caused by againg and sudden weight loss [4, 3].

**Spinal infection** occurs when a bacterial infection travels via the bloodstream into an intervertebral disc. This weakens the annulus fibrosis and decays it and might cause collapsing of the disc and thus pressure on the nerve root. Furthur infection might cause fusion of the enclosing vertebrae [4, 3].

It is worth mentioning that existence of one abnormality encourages deveplopment of other abnormalities. For example, spinal stenosis might occur because of existence of a degenerative disc disease or disc herniation for example. Most intervertebral disc diseases diagnosis in our dataset have multiple abnormalities at the same time which complicates the work of subsequent CAD algorithms. Fig. 2.1 shows two sample cases diagnosed with multiple disease. This diagnosis is summarized from the original radiologist report that details both quantitative and qualitative analysis of the diseases.

### 2.2. Related work

Backbone image analysis for various medical imaging modalities has been attracting many researchers in the last two decades. In mid 1980s, Jenkins et al. [5] performed a valuable analysis study on 107 normal and 18 abnormal cases. They analyzed the relation between proton density and age in normal discs. They concluded that quantitative MR analysis may assist in the diagnosis of intervertebral disc degeneration.

An international forum was held in 1995 to discuss methods of management for Lower backpain (LBP). They discussed the possibilities of classification of LBP into spe-

cific categories. Existence of this kind of classification helps developing CAD systems because the basic concept behind detection of abnormalities is an automatic classification problem based on a set of features. Many systems have classified LBP such as [6, 7, 8].

Many researchers have proposed methods for diagnosis of certain abnormailities related to the vertebral column. However, as far as we know, no one has proposed a method for the problem we are targetting in this paper. All related work has been investigating automation of specific abnormalities in various medical imaging modalities.

Bounds et al., [9] utilized a Neural network for diagnosis of backpain and sciatica. They have three groups of doctors to perform diagnosis as their validation mechanism. They claimed that they achieve better accuracy than the doctors in the diagnosis. However, the lack of data forbade them from full validation of their system. Similarly, Vaughn [10] conducted a research study on using Neural network (NN) for assisting orthopaedic surgeons in the diagnosis of lower back pain. Lower backpain is classified into three broad clinical categories: Simple Low Back Pain (SLBP), Root Pain (ROOTP), and Abnormal Illness Behaviour (AIB) and about 200 cases were collected over the period of 2 years with diagnosis from radiologists. Twenty five features are used to train the NN including symptoms clinical assesment results. The NN achieved 99% of training accuracy and 78.5% of testing accuracy.

Tsai et al., [11] used geometrical features (shape, size and location) to diagnose herniation from 3D MR and CT axial (transverse sections) volumes of the discs. They also discussed the diagnosis of 16 clinical cases of various lumbar herniation types and report the follow-up period for 1.8 year. 75% of the patients show excellent outcome after the surgery based on thier diagnosis while the rest 25% ranges between good and no improvement.

Kol et al., [12] proposed a finite element model (FEM) for the $\mathcal{L}4-\mathcal{L}5$ disc and the enclosing vertebrae to invistigate the possible support for medical diagnosis and muscle rehabilitation. They used Nuclear Magnetic Resonance (NMR) and computer tomography (CT) data to build the geometrical FE model. They concluded that there is an indication of supporting diagnosis and muscle rehabilitation decision using their model. Later, Glema et al., [13] invistigated the use of modeling intervertebral discs in the analysis of spinal segments. They used the model of [12] for $\mathcal{L}4-\mathcal{L}5$ and validated it for four loading schemes: axial compression, two bending in vertical plains (sagital and lateral), and torsion. They found that it was possible to verify the validity and quality of the model for disc buldging and some specific other abnormalities.

Chamarthy et al., [14] used k-means to estimate the degree of disc space narrowing with a score ranging between 0 (normal) and 3 (significant narrowing). They perfomed experiments on cervical X-rays and achieved 82% accuracy. Cherukuri et al., [15] used size-invariant, convex hull-based features to discriminate anterior osteophytes (bony growths on vertebrae) in cervical X-ray images and achieved an average accuracy of 86%.

Recent work by Koompairojn et al., [16] used a Bayesian classifier for detection of spinal stenosis using 13 morpohological features. These features include heights of the vertebrae and disc space (anterior, mid and posterior), anteroposterior width of lower and upper spinal canal. They use X-rays from the NHANES II [17] database to train and test their classifier. They achieve accuracy ranging between 75% to 85%.

## 3. PROPOSED MODEL

We caprure the abnormality condition $n_i$ with a Gibbs model:

$$P(n_i|d_i,\sigma_{\mathtt{I}(d_i)}) = \frac{1}{Z[n_i]}\exp^{-\mathtt{E}_{n_i}(d_i,\sigma_{\mathtt{I}(d_i)})} \qquad (2)$$

where $n_i$ is a binary random variable for abnormality of the disc $i$ and $n_i \in \mathcal{N} = \{n_i : 1 \leq i \leq 6\}$, the location of the disc $d_i \in \mathcal{D} = \{d_i : 1 \leq i \leq 6\}$, the $\sigma_{d_i}$ is a neighborhood of pixels around the disc location $d_i$. $\mathtt{E}_{n_i}(d_i,\sigma_{\mathtt{I}(d_i)})$ is the energy function identified by disc location $d_i$ and the intensity of a pixel neighborhood $\sigma_{\mathtt{I}(d_i)}$.

We propose the use of three potentials: the appearance $\mathtt{I}$, the location $d_i$, and the context between discs $(i \sim j)$. This concludes our energy function $\mathtt{E}_{n_i}(d_i,\sigma_{\mathtt{I}(d_i)}))$ to:

$$\begin{aligned}\mathtt{E}_{n_i}(d_i,\sigma_{\mathtt{I}(d_i)}) = \Big[&\beta_1 \sum_{d\in\mathcal{D}} U_{\mathrm{I}}(d_i,\sigma_{\mathtt{I}(d_i)}) && \leftarrow \text{intensity}\\ &+\beta_2 \sum_{d\in\mathcal{D}} U_{\mathrm{D}}(d_i) && \leftarrow \text{location}\\ &+\beta_3 \sum_{(i\sim j)} V_{\mathrm{D}}(d_i,d_j)\Big] && \leftarrow \text{context}\end{aligned} \qquad (3)$$

where $\beta_1, \beta_2$, and $\beta_3$ are the model parameters that control the effect weight of features on the inference. $U_{\mathrm{I}}$ is the appearance potential which is a model of both the location of each disc $d_i \in \mathcal{D}$ and the intensity of the pixel neighborhood $\sigma_{\mathtt{I}}(d_i)$ of that location. $U_{\mathrm{D}}$ is the location potential which is a model of the location of these discs $\mathcal{D}$. $V_{\mathrm{D}}$ is the context potential which is a model of the distance between neighboring discs $(i \sim j)$.

Our model requires two inputs: the locations of the discs $\mathcal{D} = \{d_1, d_2, ..., d_6\}$, and the intensity of a neighborhood surrounding every location $\sigma_{I(d_i)}$. The first input is actually the outcome of the labeling problem which we produce from our previous work [2]. The second input is obtained from the image intensity $\mathtt{I} = \{\text{ Intensity} : 0 \leq \text{Intensity} \leq 2^b - 1\}$ for the disc location and a defined neighborhood $\sigma_{d_i}$ where $b$ is the bit depth of the images, which is 12 bits for our dataset.

Here, we discuss the model for each of the three potential:

**Appearance** potential $U_{\mathtt{I}}(d_i, \sigma_{\mathtt{I}(d_i)})$ models the expected intensity level of the abnormal discs, which we model as Gaussian. After taking the negative log:

$$U_{\mathtt{I}}(d_i, \sigma_{\mathtt{I}(d_i)}) = \frac{\sum_{j \in \sigma_{\mathtt{I}(d_i)}} \left(\mathtt{I}(j) - \mu_{\mathtt{I}}\right)^2}{2\sigma_{\mathtt{I}}^2} \tag{4}$$

where $d_i$ is the location $d_i = (row, col)$ of disc $i$, $\mathtt{I}(d_i)$ is the intensity at location $d_i$, $\sigma d_i$ is some pixel neighborhood of the location $d_i$, $\mu_{\mathtt{I}}$ is the expected intensity levels of the abnormal discs, $\sigma_{\mathtt{I}}^2$ is the variance of the intensity levels of abnormal discs. Both $\mu_{\mathtt{I}}$ and $\sigma_{\mathtt{I}}^2$ are learned from the training data where a set of images are manually labeled (or labeled by our labeling method in [2]).

**Location** potential $U_{\mathrm{D}}(d_i)$ models the expected location of abnormal disc at level $i$. In fact, abnormal discs in general differ in their expected location from normal discs (at the same lumbar level). We model the location as a 2D Gaussian and after taking the negative log, we obtain Mahalanobis distance:

$$U_{\mathrm{D}}(d_i) = \left[ \left(d_i - \mu_{d_i}\right)^{\mathsf{T}} \Sigma_{d_i}^{-1} \left(d_i - \mu_{d_i}\right) \right] \tag{5}$$

where $d_i$ is the location of disc $i$, $\mu_{d_i}$ is the expected location of the abnormal discs at lumbar disc level $i$, $\Sigma_{d_i}$ is the covariance matrix of the abnormal discs at the lumbar disc level $i$. We learn both $\mu_{d_i}$ and $\Sigma_{d_i}$ from the training data.

**Context** potential $V_{\mathrm{D}}(d_i, d_j)$ models the contextual relation between neighboring disc locations $i$ and $j$. We model the distances $e_{ij} = |d_i - d_j|_2$ between neighboring discs at locations $i$ and $j$ as a Gaussian distribution, which concludes after the negative log to:

$$V_{\mathrm{D}}(d_i, d_j) = \frac{\left(e_{ij} - \mu_D\right)^2}{\sigma_D^2} \tag{6}$$

where $d_i$ and $d_j$ are neighboring discs, $\mu_D$ is the expected distance between abnormal discs, $\sigma_D^2$ is the variance of abnormal discs distances. We also learn both $\mu_D$ and $\sigma_D^2$ from the training data.

## 4. AVAILABLE DATA

We use a dataset of 80 clinical MRI volumes containing normal and abnormal cases. Abnormalities include disc herniation, disc desiccation, degenerative disc disease and others. Every single case contains five, six or seven acquisition protocols. Every case contains a full volume of T2-weighted MR beside many other protocols including T1-weighted and T2-weighted Myelo images. We use the T2-weighted volumes for training and testing our proposed model for abnormality detection. We pick the middle slice from every volume to represent that case and use it in our model training and testing.

## 5. EXPERIMENTAL RESULTS

We train our model on T2-weighted modality as disc intensities have better discrimination from other structures in the image as appears in shown Fig. 1(a).

We perform ground truth annotation for our dataset by:

1. Selecting a point inside every disc that roughly represents the center for that disc $d_i$,

2. Determining whether the disc is normal or abnormal $n_{d_i}$ because our model here concerns about discriminating between normal and abnormal discs regardless of the type of abnormalities.

It is worth mentioning that inter-observer error exist in lumbar diagnosis similar to various diagnosis tasks from various imaging modalities including plain radiographs, MRI, CT, SPECT (single-photon emission computed tomography), High Resolution (HR). However, MRI shows high inter-observer reliability compared to plain radiographs in lumbar area diagnosis (e.g., [18]). Mulconrey et al. [19] showed that abnormality detection for degenerative disc and spondylolisthesis with MRI has $\kappa = 0.773$ and $\kappa = 0.728$, respectively, which is considered high in showing interobserver reliability where this reliability is considered perfect when $0.8 \leq \kappa \leq 1$.

We train our model to learn the parameters of the three potentials representing the models for the appearance $\mathtt{I}$, the location $d_i : 1 \leq i \leq 6$, and the context between discs $(i \sim j)$ using the ground truth $(\mathcal{D}, \mathcal{N})$ and the corresponding training images $\mathtt{I}$.

We perfom a cross-validation experiment using the 80 cases to train and test our proposed method. In every round, we separate thirty cases and train on the rest 50 cases. We perfom 10 rounds and every time the cases are selected randomly. Dr. Gurmeet Dhillon provided the ground truth for all the 80 cases to automatically check classification accuracy which we define by:

$$Accuracy_i = 1 - \frac{1}{K} \sum_{j=1}^{K} |g_{ij} - n_{ij}| * 100\% \tag{7}$$

where $Accuracy_i$ represents the classification accuracy at the lumbar disc level $i$ where $1 \leq i \leq 6$, the value $K$ represents the number of cases in every experiment, $g_{ij}$ is the ground truth binary assignment for disc $i$, and $n_{ij}$ is the resulting binary assignment for disc $i$ from the inference on our model. $g_i$ and $n_i$ are assigned the binary values the same way such that:

$$g_i = \begin{cases} 1 & \text{if Disc } i \text{ is Normal} \\ 2 & \text{if Disc } i \text{ is Abnormal} \end{cases} \tag{8}$$

It is worth mentioning that we measure accuracy at every lumbar disc level separately to show the detailed classification accuracy at every level and thus have more understanding of the disc levels and its influence on classification accuracy. This appears in the row before the last in Table 1 where every value is a percentage accuarcy that represents the average of all the rounds in the experiment for every disc level. At the same time, we report the average accuracy for all the discs together for each round; which is the last column in the Table 1, and then the overall average accuracy for discs and for all rounds in the experiment that appears in the bottom-right cell in the same table.

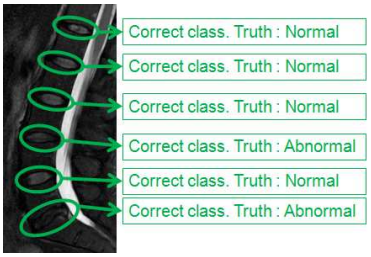
**Table 1**. Classification results for the cross-validation experiment on 80 cases. Row before last shows average accuracy at every lumbar disc level and the last column shows the average accuracy for every round of 30 cases. We achieve over 91% of classification accuracy.

| Set | E6 | E5 | E4 | E3 | E2 | E1 | Accuracy |
|---|---|---|---|---|---|---|---|
| 1 | 27 | 25 | 27 | 29 | 29 | 28 | 91.67% |
| 2 | 26 | 26 | 29 | 29 | 28 | 28 | 92.22% |
| 3 | 26 | 26 | 27 | 27 | 26 | 26 | 87.78% |
| 4 | 28 | 25 | 26 | 27 | 29 | 29 | 91.11% |
| 5 | 27 | 27 | 29 | 28 | 27 | 27 | 91.67% |
| 6 | 25 | 26 | 26 | 27 | 29 | 28 | 89.44% |
| 7 | 25 | 27 | 28 | 26 | 28 | 29 | 90.56% |
| 8 | 28 | 28 | 27 | 28 | 29 | 28 | 93.33% |
| 9 | 27 | 26 | 28 | 27 | 29 | 29 | 92.22% |
| 10 | 27 | 28 | 28 | 28 | 28 | 28 | 92.78% |
| (%) | 88.7 | 88.0 | 91.7 | 92.0 | 94.0 | 93.3 | - |
| **Average Accuracy** | | | | | | | **91.28%** |

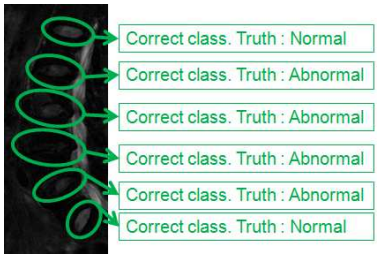
Fig. 3 shows five sample cases of classification output from inferencing on our model. The first three figures show various abnormalities at various levels and full success in abnormality detection. Fig. 3(d) shows a false negative at level $\mathcal{L}2-\mathcal{L}3$ where the disc is labeled as abnormal while its ground truth is normal. Fig. 3(e) shows a false positive at level $\mathcal{L}1-\mathcal{L}2$ where the disc is labeled as normal while its ground truth is abnormal.

## 6. DISCUSSIONS AND FUTURE WORK

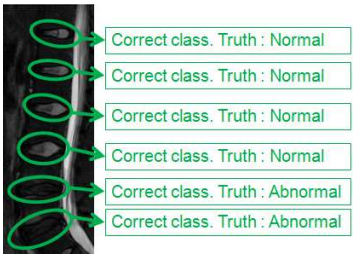
We achieve high abnormality detection accuracy using three main features: appearance, location, and context of discs. However, some abnormal discs are not detected. We find that incorporating a shape model might enhance our detection accuracy. For example, the misclassified disc at level $\mathcal{L}2-\mathcal{L}3$ in Fig. 3(d) appears more compact in shape than other normal discs in the same case. This motivates including a shape model or some geometrical model for height and width of the disc (similar to Koompairojn et al. [16] work for stenosis detection). In general, most abnormal discs are less thickness than normal discs. However, finding a model



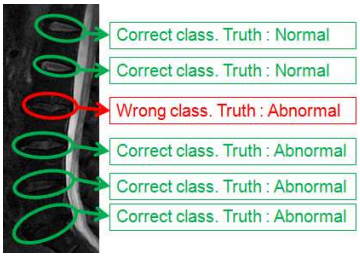

(a) Abnormals levels: $\mathcal{L}3-\mathcal{L}4$ and $\mathcal{L}5-\mathcal{S}$. All levels are correclty classified.



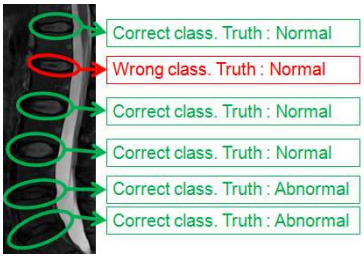

(b) Abnormal levels: $\mathcal{L}1-\mathcal{L}2$, $\mathcal{L}2-\mathcal{L}3$, $\mathcal{L}3-\mathcal{L}4$, $\mathcal{L}4-\mathcal{L}5$. All levels correclty classified.




(c) Abnormal levels: $\mathcal{L}4-\mathcal{L}5$, $\mathcal{L}5-\mathcal{S}$. All levels are correclty classified.




(d) Abnormal levels: $\mathcal{L}2-\mathcal{L}3$, $\mathcal{L}3-\mathcal{L}4$, $\mathcal{L}4-\mathcal{L}5$, $\mathcal{L}5-\mathcal{S}$. Level $\mathcal{L}2-\mathcal{L}3$ is false negative.




(e) Abnormal levels: $\mathcal{L}4-\mathcal{L}5$, $\mathcal{L}5-\mathcal{S}$. Level $\mathcal{L}1-\mathcal{L}2$ is false positive.

**Fig. 3**. Sample abnormality detection from the experiment. Green means it is correctly classified while red means otherwise.

for disc height and width or even shape should not be separate from incorporating a model for patient age and patient height as well. Lumbar area vertebrae and intervertebral discs vary in size depending on patient age and body size. We are working on modeling age of patients and its relation to disc geometrical properties as well as disc shape.

Another focus in solving abnormality detection is the minimization of false negatives. That is, minimization of abnormal discs detected as normal. Having any false negative disc means that this disc will not have the chance for diagnosis by the radiologist or subsequent diagnosis algorithms. However, false positive discs (normal discs detected as abnormal) are not of comparable concern because the only draw here is the needed time for the radiologist (or the subsequent CAD system) to verify that it is a false positive disc.

We are conducting more extensive study on larger dataset to model age and height of the patient and their relation to the geometry and shape of the normal and abnormal lumbar discs. On the other hand, we are working on detection of intervertebral disc diseases such as desiccation, herniation, stenosis, and degenerative disc disease.

## 7. CONCLUSION

We proposed a probabilistic model for incorporating intervertebral disc appearance, location, and context to detect abnormal discs from clinical T2-weighted MRI scans. Our model is extensible for subsequent diagnosis tasks such as diagnosis of desiccation, stenosis, and herniation by incorporating more features emerging from the way that the radiologist make his decision during the diagnosis process. We achieve over 91% accuracy on cross-validation experiment on 80 clinical MRI cases that includes various types of abnormality.

## 8. ACKNOWLEDGEMENT

This work is supported in part by the New York State Foundation for Science, Technology and Innovation (NYSTAR).

## 9. REFERENCES

[1] National Institute of Neurological Disorders and Stroke (NINDS), "Low back pain fact sheet," *NIND brichure*, 2008.

[2] Jason J. Corso, Raja' S. Alomari, and Vipin Chaudhary, "Lumbar disc localization and labeling with a probabilistic model on both pixel and object features.," in *Proc. of MICCAI 2008*. 2008, vol. 5241 of *LNCS Part 1*, pp. 202–210, Springer.

[3] Richard S. Snell, *Clinical Anatomy by Regions*, Lippincott Williams and Wilkins, 8th edition, 2007.

[4] Arthur F Dalley Anne MR Agur, *Atlas of Anatomy*, Lippincott Williams and Wilkins, 11th edition, 2004.

[5] J. P. Jenkins, D. S. Hickey, X. P. Zhu, M. Machin, and I. Isherwood, "Mr imaging of the intervertebral disc: A quantitative study," *British Journal of Radiology*, vol. 58, no. 692, pp. 705–709, 1985.

[6] Bernard TN Jr and Kirkaldy-Willis WH, "Recognizing specific characteristics of nonspecific low back pain," *Clinical orthopaedics and related research*, vol. 217, pp. 266–280, April 1987.

[7] Delitto A, Erhard RE, and Bowling RW, "A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment," *Physical Therapy*, vol. 75, pp. 470–489, 1995.

[8] Bowling RW, Truschel DW, Delitto A, and Erhard RE, "Conservative management of low back pain with physical therapy," pp. 499–594, 1997.

[9] D.G. Bounds, P.J. Lloyd, B. Mathew, and G. Waddell, "A multilayer perceptron network for the diagnosis of low back pain," San Diego, CA, July 1988, vol. 2, pp. 481–489.

[10] Marilyn Vaughn, "Using an artificial neural network to assist orthopaedic surgeons in the diagnosis of low back pain," Department of Informatics, Cranfield University (RMCS), 2000.

[11] Ming-Dar Tsai, Shyan-Bin Jou, and Ming-Shium 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.

[12] Kol W., Lodygowski T., Ogurkowska M.B., and Wierszycki M. and, "Are we able to support medical diagnosis or rehabilitation of human vertebra by numerical simulation," Gliwice, Poland, June 2003.

[13] Glema A., Kakol W., Lodygowski T., Ogurkowska M.B., and Wierszycki M., "Modeling of intervertebra disks in the analysis of spinal segment," Jyvskyl, Finland, July 2004, pp. 24 – 28.

[14] Pavan Chamarthy, R. Joe Stanley, Gregory Cizek, Rodney Long, Sameer Antani, , and George Thoma, "Image analysis techniques for characterizing disc space narrowing in cervical vertebrae interfaces.," *Computerized Medical Imaging and Graphics*, vol. 28, no. 1-2, pp. 39–50, January 2004.

[15] Maruthi Cherukuri, R. Joe Stanley, Rodney Long, Sameer Antani, , and George Thoma, "Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features," *Computerized Medical Imaging and Graphics*, pp. 99–108, 2004.

[16] S. Koompairojn, K.A. Hua, and C. Bhadrakom, "Automatic classification system for lumbar spine x-ray images," *Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on*, pp. 213–218, 2006.

[17] L. Rodney Long, Sameer Antani, Dah-Jye Lee, Daniel M. Krainak, and George R. Thoma, "Biomedical information from a national collection of spine x-rays: film to content-based retrieval," 2003, vol. 5033, pp. 70–84, SPIE.

[18] Sanjeev S Madan and Mersey Deanery, "Interobserver error in interpretation of the radiographs for degeneration of the lumbar spine," *The Iowa Orthopaedic Journal*, pp. 51–56, 2003.

[19] D . Mulconrey, R . Knight, J . Bramble, S . Paknikar, and P . Harty, "Interobserver reliability in the interpretation of diagnostic lumbar mri and nuclear imaging," *The Spine*, vol. 6, pp. 177 – 184, 2006.