Int J CARS (2010) 5:287–293
DOI 10.1007/s11548-009-0396-9

ORIGINAL ARTICLE

Computer-aided diagnosis of lumbar disc pathology from clinical lower spine MRI

Raja' S. Alomari · Jason J. Corso · Vipin Chaudhary · Gurmeet Dhillon

Received: 10 January 2009 / Accepted: 13 August 2009 / Published online: 22 September 2009

Abstract
Purpose Detection of abnormal discs from clinical T2-weighted MR Images. 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.
Materials and methods We propose a probabilistic classifier for the detection of abnormality of intervertebral discs. We use three features to label abnormal discs that include 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 infer on the middle slice of the T2-weighted MRI volume for each case. These MRI scans are specific for the lumbar area. We obtain our gold standard for the ground truth from our collaborating radiologist group by having the clinical diagnosis report for each case.
Results We achieve over 91% abnormality detection accuracy in a cross-validation experiment with 80 clinical cases. The experiment runs ten rounds; in each round, we randomly leave 30 cases out for testing and we use the other 50 cases 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 the extendability of our proposed model to subsequent diagnosis tasks specific to each intervertebral disc abnormality such as desiccation and herniation.

Keywords Computer-aided diagnosis · MRI · Lumbar intervertebral disc · Gibbs distribution

R. S. Alomari (✉) · J. J. Corso · V. Chaudhary
Computer Science and Engineering Department,
University at Buffalo, Buffalo, NY 14260, USA
e-mail: ralomari@buffalo.edu

G. Dhillon
ProScan Imaging of Buffalo, Williamsville, NY 14221, USA

Introduction

Lower 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 lower back pain and over 12 million Americans have some sort of intervertebral disc disease (IDD) [1]. The higher growth for the number of patients each year (about 8%) compared to the growth of radiologists (about 1%) 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 each disc level. The focus of this paper is on the detection of abnormality in the lumbar area from MRI. In our previous work [2], we developed a two-level probabilistic model for localization of the six discs in the sagittal T2-weighted MRI scans for the lumbar area. We incorporated two levels of information: low- and high-level. In the low-level, we modeled the local pixel properties of discs, such as appearance. In the high-level, we captured the object-level geometrical and contextual relationships between discs. We estimated 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 (17–81 years old), and patient heights which affect the size

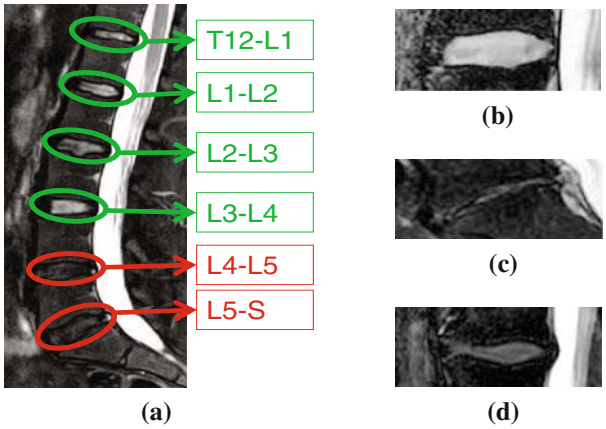

Fig. 1 Labeling of lumbar area discs and sample abnormalities. **a** Lumbar disc levels labels. Abnormal lower two levels at L4-L5 and L5-S (*red*) (also called L5-S1) and the upper four are normal (*green*). **b** normal disc, **c** herniated disc, **d** degenerative disc disease

and appearance of the discs. Figure 1a shows a sample sagittal view with labeled lumbar disc levels.

In this paper, we propose a method for detection of abnormal discs in the lumbar area from clinical T2-weighted MRI. We model disc appearance, location and context and incorporate them in a probabilistic classifier by introducing the random variable n_i and solving:

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

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.

The remainder of this paper is organized as follows: The background and related work is discussed in next section. Then we discuss "Proposed model". We then describe "Clinical data description", "Clinical ground truth" and "Experimental results" respectively. "Discussions and future work" and "Conclusion" are discussed in the last two sections.

Background

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 up about one-fourth of the spinal column's length [3].

An intervertebral disc is composed of two parts: an outer strong ring called annulus fibrosus and a soft gel-like inner called nucleus pulposus. The nucleus pulposus consists of 80–85% water in normal cases. In the lumbar area, there are six discs connected to the five lumbar vertebrae which are labeled top-down as T12-L1, L1-L2, L2-L3, L3-L4, L4-L5, and L5-S1 as shown in Fig. 1a.

Abnormalities in the intervertebral discs

Various diseases that affect the vertebral column are usually painful and influence the patient's everyday life. In our work, we are concerned with the clinical lumbar abnormalities. Fardon et al. [4] presented a nomenclature and a classification of the lumbar disc pathology for standardization of the language and defining the various abnormalities for the lumbar area intervertebral discs. They extended the work of Milette [5] in coordination with the North American Spine Society (NASS), the American Society of Spine Radiology (ASSR), and the American Society of Neuroradiology (ASNR). It was also endorsed by many other worldwide spine societies [6]. We discuss the most popular clinical abnormalities in light of Fardon et al. [4] nomenclature.

Disc herniation (Fig. 1c) is a leak of the nucleus pulposus through a tear in the wall of the annulus fibrosus. This leak presses on the local nerve root causing the pain. Tears in the disc wall usually occur due to aging and/or trauma [3,7].

Spinal stenosis is the narrowing of the spinal canal and can be caused by different conditions such as disc herniation, osteoporosis, or a tumor. Sometimes, and especially when the reason is a disc herniation, stenosis occurs at the same level of the disc [3,7].

Degenerative disc disease (Fig. 1d) is the gradual deterioration of the disc causing loss of its functions. This disease usually develops with aging or from continuous activities that press on the disc space. It starts with a small injury in the annulus fibrosus causing damage to the nucleus pulposus and loss of its water contents. Further damage causes malfunctioning of the disc and thus collapsing the upper and lower vertebrae. As time passes, the vertebra facet joints twist creating bone spurs that grow into the spinal canal and pinching the nerve root (stenosis) [3,7].

Disc desiccation is the drying out of the water contents in the inner pulposus. Usually, it is caused by aging and sudden weight loss [3,7].

Spinal infection occurs when a bacterial infection travels via the bloodstream into an intervertebral disc. This weakens the annulus fibrosus and decays it and might cause collapsing of the disc and thus pressure on the nerve root. Further infection might cause fusion of the enclosing vertebrae [3,7].

Existence of one abnormality encourages development of other abnormalities. For example, spinal stenosis might occur because of existence of a degenerative disc disease or disc herniation. Most intervertebral discs in our dataset have multiple abnormalities at the same time which complicates the work of subsequent CAD algorithms. Figure 2 shows two sample cases diagnosed with multiple diseases.

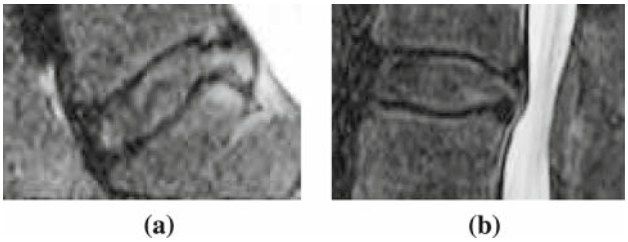

Fig. 2 Sample diagnosis of two Discs with multiple diseases. **a** L5-S1 disc level: Herniation, DDD, mild foramina stenosis, **b** L3-L4 disc level: DDD, central stenosis, and central herniation

This diagnosis is summarized from the original radiologist report that details both quantitative and qualitative analysis of the diseases.

Related work

Backbone image analysis for various medical imaging modalities has been attracting many researchers in the last two decades. In the mid-1980s, Jenkins et al. [8] 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 MRI analysis may assist in the diagnosis of intervertebral disc degeneration.

An international forum was held in 1995 to discuss methods of management for lower back pain (LBP). They discussed the possibilities of classification of LBP into specific 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 [9–11].

Many researchers have proposed methods for diagnosis of certain abnormalities related to the vertebral column. However, as far as we know, no one has proposed a method for the problem we are targeting in this paper. All related work has been investigating automation of specific abnormalities in various medical imaging modalities. Besides abnormality diagnosis, many efforts have been investigating the localization, detection, and segmentation problems for the intervertebral discs [2,12–14].

Bounds et al. [15] utilized a neural network for the diagnosis of lower back pain and sciatica. They had three groups of doctors perform diagnosis as their validation mechanism. They achieved better accuracy than the doctors in the diagnosis. However, the lack of data forbade them from full validation of their system. Similarly, Vaughn [16] conducted a research study on using neural networks (NN) for assisting orthopedic surgeons in the diagnosis of lower back pain. They classified lower back pain into three broad clinical categories: simple low back pain (SLBP), root pain (ROOTP), and abnormal illness behavior (AIB). They collected nearly 200 cases over the period of 2 years with diagnoses from radiologists. They used twenty NN including symptoms and clinical assessment results. The NN achieved 99% training accuracy and 78.5% testing accuracy which shows NN overfitting on training data.

Tsai et al. [17] used geometrical features (shape, size and location) to diagnose herniation from 3D MRI 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 years. About 75% of the patients showed excellent outcomes after the surgery based on their diagnosis while the remaining 25% ranges between good and no improvement.

Kol et al. [18] proposed a finite element model (FEM) for the L4-L5 disc and the enclosing vertebrae to investigate the possible support for medical diagnosis and muscle rehabilitation. They used nuclear magnetic resonance (NMR) and computer tomography (CT) data to build the geometrical FEM. They concluded that there is an indication of supporting diagnosis and muscle rehabilitation decisions using their model. Later, Glema et al. [19] investigated the use of modeling the intervertebral discs in the analysis of the spinal segments. They used the model of [18] for L4-L5 and validated it for four loading schemes: axial compression, bending in two vertical planes (sagittal and lateral), and torsion. They found that it was possible to verify the validity and quality of the model for disc bulging and some other specific abnormalities.

Chamarthy et al. [20] used k-means to estimate the degree of disc space narrowing with a score ranging between 0 (normal) and 3 (significant narrowing). They performed experiments on cervical X-ray radiographs and achieved 82% accuracy. Cherukuri et al. [21] 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. [22] used a Bayesian classifier for detection of spinal stenosis using 13 morphological features. These features include heights of the vertebrae, disc space (anterior, mid and posterior), and anteroposterior width of lower and upper spinal canal. They used X-rays from the NHANES II [23] database to train and test their classifier. They achieved accuracy ranging between 75 and 85%.

Proposed model

We capture the abnormality condition n_i with a Gibbs model:

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

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 $d_i \in \mathcal{D} = \{d_i : 1 \leq i \leq 6\}$ and the σ_{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 the 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 ($i \sim j$). Thus our energy function $\mathtt{E}_{n_i}(d_i, \sigma_{\mathtt{I}(d_i)}))$ is:

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

where β_1, β_2, and β_3 are the model parameters that control the effect weight of features on the inference. U_{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_{D} is the location potential which is a model of the location of each disc in $\mathcal{D}$. V_{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 $\mathcal{D} = \{d_1, d_2, \ldots, d_6\}$ and the intensity of a neighborhood surrounding each location $\sigma_{I(d_i)}$. The first input is the outcome of the labeling problem which we produce from our previous work [2]. We obtain the second input from the image intensity $\mathtt{I} = \{\text{Intensity: } 0 \leq \text{Intensity} \leq 2^b - 1\}$ for the disc location and a defined neighborhood σ_{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 potentials: *Appearance potential* $U_I(d_i, \sigma_{\mathtt{I}(d_i)})$ models the expected intensity level of the abnormal discs, which we model as Gaussian. After taking the negative log:

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

where d_i is the location $d_i = (row, col)$ of disc i, $\mathtt{I}(d_i)$ is the intensity at location d_i, σ_{d_i} is some pixel neighborhood of the location d_i, $\mu_{\mathtt{I}}$ is the expected intensity levels of the abnormal discs and $\sigma_{\mathtt{I}}^2$ 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 labeled by our labeling method in [2].

Location potential $U_D(d_i)$ models the expected location of abnormal discs 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 the 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, μ_{d_i} is the expected location of the abnormal discs at lumbar disc level i and Σ_{d_i} is the covariance matrix of the abnormal discs at the lumbar disc level i. We learn both μ_{d_i} and Σ_{d_i} from the training data.

Context potential $V_D(d_i, d_j)$ 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 results, after the negative log, in:

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

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

Our proposed model shows flexibility in the addition of various features that allows solving various diagnostic tasks depending on the abnormality type. We have other recent work where we utilize our model to diagnose specific disc abnormality. For example, in our work [24], we jointly model the appearance and the intensity level contextual information for diagnosis of disc desiccation from clinical lumbar MRI.

Clinical data description

The clinical standard uses magnetic resonance imaging (MRI) for lower lumbar diagnosis. Radiologists use MRI to diagnose discs as well as vertebrae abnormalities. However, in very rare cases, they might require CT if they suspect unclear vertebral problems due to clearer imaging of bones (vertebrae) in CT. A typical MRI system consists of: (1) a large magnet for magnetic field generation, (2) shim coils for achieving homogeneity of the magnetic field, (3) a radio-frequency (RF) coil for transmitting radio signals into the organ under imaging, (4) a receiver coil for detecting the bouncing radio signals, (5) gradient coils to provide spatial localization of the signals, and (6) a reconstruction protocol for building the final image.

Four main parameters control the appearance (intensity) of the resulting MR image: (1) proton density, (2) longitudinal relaxation time (T1), (3) transverse relaxation time (T2), and (4) the flow. The proton density is the concentration of protons in the tissue in the form of water and macromolecules (proteins, fat, etc). Both T1 and T2 relaxation times define the way the protons revert back to their resting states after the initial RF pulse. The most common effect of the flow is the loss of signal from rapidly flowing arterial blood.

Table 1 T2-weighted MRI parameters for our dataset

Parameter	Value
Repetition time (TR)	3,157
Echo time (TE)	100
Flip angle	90

The pulse sequence parameters control the contrast of the resulting MR image. The MRI technician sets the specific number, strength, and timing of the radio-frequency (RF) and gradient pulses.

There are two common pulse sequences for MRI imaging: T1-weighted and T2-weighted spin-echo sequences. The T1-weighted sequence uses a short TR (repetition time) and a short TE (echo time) (TR $\leq$ 1,000 ms, TE $\leq$ 30 ms) while the T2-weighted sequence uses a long TR and a long TE (TR $\geq$ 2,000 ms, TE $\geq$ 80 ms).

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. Each case contains a full volume of the T2-weighted MRI on which we base our classifier. Table 1 shows the acquisition parameters for our dataset.

Clinical ground truth

We obtain the clinical diagnosis report for each case that contains a diagnosis at each lumbar disc level. These reports are generated by agreement inside our collaborative radiology center between at least three radiologists. We consider these reports as our gold standard.

Mulconrey et al. [25] showed that MRI has high inter-observer reliability for degenerative disc and spondylolisthesis diagnosis with $\kappa = 0.773$ and $\kappa = 0.728$, respectively. This and other efforts such as [26] show that inter-observer error is small in diagnosis of lumbar abnormality from MRI modality. The perfect inter-observer reliability happens when $0.8 \leq \kappa \leq 1$ [25]. This motivates us to consider the clinical reports provided by our collaborating radiology center as our gold standard.

We perform gold standard annotation for our dataset by: (1) selecting a point inside each disc that roughly represents the center for that disc d_i, and then (2) determining the abnormality condition n_{d_i}.

Experimental results

Intervertebral discs have better discrimination from other structures in T2-weighted MR images compared to the T1-weighted [27,28]. Thus we base our model on T2-weighted MRI. We 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 gold standard $(\mathcal{D}, \mathcal{N})$ and the corresponding training images $\mathtt{I}$.

We perform a cross-validation experiment using the 80 cases to train and test our proposed method. In each round, we separate thirty cases and train on the other 50 cases. We perform ten rounds and each time the cases are selected randomly. We check the classification accuracy by comparing our classification results with our gold standard (illustrated in "Clinical ground truth") by defining the accuracy at each disc level i as:

$$\text{Accuracy}_i = \left(1 - \frac{1}{K}\sum_{j=1}^{K} |g_{ij} - n_{ij}|\right) \times 100\% \tag{7}$$

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 each experiment, g_{ij} is the gold standard 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:

$$g_i = \begin{cases} 0 & \text{if disc } i \text{ is normal} \\ 1 & \text{if disc } i \text{ is abnormal} \end{cases} \tag{8}$$

We measure accuracy at each lumbar disc level separately to show the detailed classification accuracy at each level and thus have more understanding of the disc levels and their influence on classification accuracy. This appears in the second to last row in Table 2, where each value is a percentage accuracy that represents the average of all the rounds in the experiment at each 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 Table 2. We also include the overall average accuracy for all discs and for all rounds in the experiment in the bottom-right cell in the same table.

Figure 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. Figure 3d shows a false negative at level L2-L3 where the disc is labeled as abnormal while its gold standard is normal. Figure 3e shows a false positive at level L1-L2 where the disc is labeled as normal while its gold standard is abnormal.

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

Table 2 Classification results for the cross-validation experiment on 80 cases

Set	E6	E5	E4	E3	E2	E1	Accuracy
1	27	25	27	29	29	28	92
2	26	26	29	29	28	28	92
3	26	26	27	27	26	26	88
4	28	25	26	27	29	29	91
5	27	27	29	28	27	27	92
6	25	26	26	27	29	28	89
7	25	27	28	26	28	29	91
8	28	28	27	28	29	28	93
9	27	26	28	27	29	29	92
10	27	28	28	28	28	28	93
(%)	88.7	88.0	91.7	92.0	94.0	93.3	–
Average Accuracy							91.3%

The second to last row shows the average accuracy at each lumbar disc level and the last column shows the average accuracy for each round of 30 cases. Disc level E6 corresponds to the L5-S1 disc level while E1 corresponds to T12-L1. We achieve over 91% classification accuracy

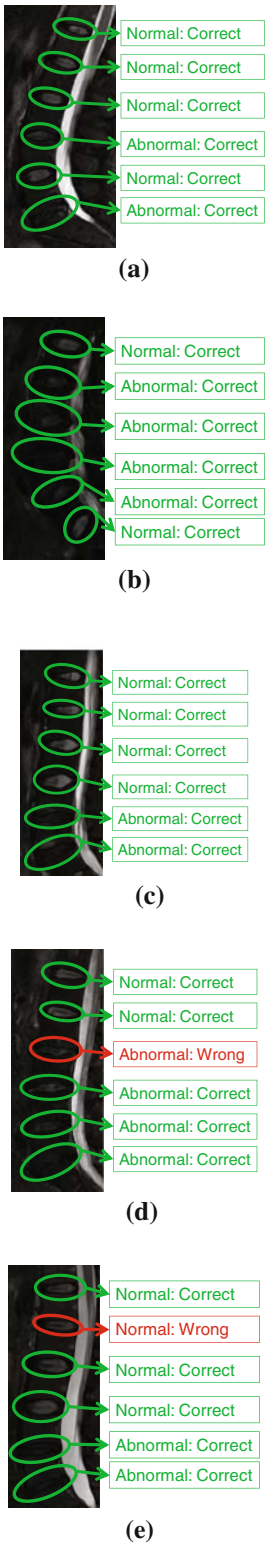

Fig. 3 Sample abnormality detection from the experiment. Green means it is correctly classified while red indicates otherwise. **a** Abnormals levels: L3-L4 and L5-S. All levels are correctly classified. **b** Abnormal levels: L1-L2, L2-L3, L3-L4, L4-L5. All levels correctly classified. **c** Abnormal levels: L4-L5, L5-S1. All levels are correctly classified. **d** Abnormal levels: L2-L3, L3-L4, L4-L5, L5-S1. Level L2-L3 is false negative. **e** Abnormal levels: L4-L5, L5-S1. Level L1-L2 is false positive

accuracy. For example, the misclassified disc at level L2-L3 in Fig. 3d 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. [22] work for stenosis detection). In general, most abnormal discs have less thickness than normal discs. 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 this disc will not have the chance for diagnosis by the radiologist or subsequent diagnosis algorithms. On the other hand, false positive discs (normal discs detected as abnormal) are not of comparable concern because their only penalty is the time needed for the radiologist (or the subsequent CAD system) to verify that it is a false positive disc.

We are conducting a more extensive study on a 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. We are also working on the detection of intervertebral disc diseases such as desiccation, herniation, stenosis, and degenerative disc disease.

Conclusion

We proposed a probabilistic model for incorporating intervertebral disc appearance, location, and context to detect abnormal discs from clinical T2-weighted MR images. Our probabilistic classifier models each disc level in the lumbar area and decides its abnormality condition upon the joint features. Our model has the flexibility to incorporate other domain knowledge features such as age and patient related information. We have demonstrated the clinical applicability

of our proposed model using clinical data from our collaborating radiologist along with the clinical diagnosis report for each case. Our model is extendable for subsequent diagnostic tasks such as diagnosis of desiccation, stenosis, and herniation by incorporating suitable features depending on the abnormality type. We have shown an example of disc desiccation diagnosis from our recent work. We achieved over 91% accuracy on a cross-validation experiment on a set of 80 clinical MRI cases that includes various types of abnormality and a wide range of patient ages and conditions.

Acknowledgments This work is supported in part by the New York State Foundation for Science, Technology and Innovation (NYSTAR). The authors would like to thank both Jeffrey Delmerico and James Evanko for their English-based reviews.

References

1. National Institute of Neurological Disorders and Stroke (NINDS) (2008) Low back pain fact sheet. NIND brochure, available at http://www.ninds.nih.gov/disorders/backpain/
2. Corso JJ, Alomari RS, Chaudhary V (2008) Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: The proceedings of medical image computing and computer assisted intervention (MICCAI'08). LNCS Part 1, vol 5241. Springer, Berlin, pp 202–210
3. Snell RS (2007) Clinical anatomy by regions, 8th edn. Lippincott Williams and Wilkins, Philadelphia
4. Fardon DF, Milette PC (2001) Nomenclature and classification of lumbar disc pathology. Spine 26(5):E93–E113
5. Milette PC (1997) The proper terminology for reporting lumbar intervertebral disk disorders. AJNR 18:1859–1866
6. Fardon DF, Milette PC (2003) Nomenclature and classification of lumbar disc pathology. Website: www.asnr.org/spine_nomenclature/, February
7. Dalley Arthur F, Agur Anne MR (2004) Atlas of anatomy, 11th edn. Lippincott Williams and Wilkins, Philadelphia
8. Jenkins JP, Hickey DS, Zhu XP, Machin M, Isherwood I (1985) MR imaging of the intervertebral disc: A quantitative study. Br J Radiol 58(692):705–709
9. Bernard TN Jr, Kirkaldy-Willis WH (1987) Recognizing specific characteristics of nonspecific low back pain. Clin Orthop Relat Res 217:266–280
10. Delitto A, Erhard RE, Bowling RW (1995) A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Phys Ther 75:470–489
11. Bowling RW, Truschel DW, Delitto A, Erhard RE (1997) In: Erdil M, Dickerson OB (eds) Conservative management of low back pain with physical therapy. Von Norstrand Reinhold, New York, pp 499–594. ISBN-13: 9780471284727
12. Michopoulou S, Costaridou L, Speller R, Panagiotopoulos E, Todd-Pokropek A (2008) Segmenting the degenerated lumbar intervertebral disc from mr images. In: Proceedings of IEEE nuclear science symposium and medical imaging conference
13. Michopoulou S, Speller R, Todd-Pokropek A, Costaridou L, Kazantzi A, Panagiotopoulos E, Panayiotakis G (2009) Computer aided diagnosis of lumbar intervertebral disc degeneration in spine MRI. In: Proceedings of computer aided radiology and surgery (CARS), Berlin, Germany, June
14. Michopoulou S, Costaridou L, Panagiotopoulos E, Speller R, Panayiotakis G, Todd-Pokropek A (2009) Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine. IEEE Trans Biomed Eng (to appear)
15. Bounds DG, Lloyd PJ, Mathew B, Waddell G (1988) A multilayer perceptron network for the diagnosis of low back pain. In: Proceeding of the IEEE international conference on neural networks, San Diego, vol II. IEEE, New York, pp 481–489
16. Marilyn V (2000) Using an artificial neural network to assist orthopaedic surgeons in the diagnosis of low back pain. Department of Informatics, Cranfield University (RMCS)
17. Tsai M-D, Jou S-B, Hsieh M-S (2002) A new method for lumbar herniated inter-vertebral disc diagnosis based on image analysis of transverse sections. Comput Med Imaging Graph 26(6):369–380
18. Kol W, Lodygowski T, Ogurkowska MB, Wierszycki M (2003) Are we able to support medical diagnosis or rehabilitation of human vertebra by numerical simulation. Gliwice, Poland
19. Glema A, Kakol W, Lodygowski T, Ogurkowska MB, Wierszycki M (2004) Modeling of intervertebra disks in the analysis of spinal segment. Jyvskyl, Finland, pp 24–28, July
20. Chamarthy P, Joe Stanley R, Cizek G, Long R, Antani S, Thoma G (2004) Image analysis techniques for characterizing disc space narrowing in cervical vertebrae interfaces. Comput Med Imaging Graph 28(1–2):39–50
21. Cherukuri M, Joe Stanley R, Long R, Antani S, Thoma G (2004) Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features. Comput Med Imaging Graph 28(1–2):99–108
22. Koompairojn S, Hua KA, Bhadrakom C (2006) Automatic classification system for lumbar spine X-ray images. Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on, pp 213–218
23. Rodney L, Sameer A, Lee D-J, Krainak DM, Thoma GR (2003) Biomedical information from a national collection of spine X-rays: film to content-based retrieval, vol 5033, pp 70–84, SPIE
24. Alomari RS, Corso JJ, Chaudhary V, Gurmeet D (2009) Desiccation diagnosis in lumbar discs from clinical mri with a probabilistic model. In: The 2009 (Sixth) IEEE International Symposium on Biomedical Imaging (ISBI'09) Boston, MA, USA, June (to appear)
25. Mulconrey D, Knight R, Bramble J, Paknikar S, Harty P (2006) Interobserver reliability in the interpretation of diagnostic lumbar MRI and nuclear imaging. The Spine 6:177–184
26. Madan SS, Rai A, Harley JM (2003) Interobserver error in interpretation of the radiographs for degeneration of the lumbar spine. Iowa Orthop J 23:51–56
27. An HS, Anderson PA, Haughton VM, Iatridis JC, Kang JD, Lotz JC, Natarajan RN, Oegema TR, Roughley P, Setton LA, Urban JP, Videman T, Andersson GBJ, Weinstein JN (2004) Disc degeneration: summary. Spine 29:2677–2678
28. Videman T, Nummi P, Battie MC, Gill K (1994) Digital assessment of MRI for lumbar disc desiccation: A comparison of digital versus subjective assessments and digital intensity profiles versus discogram and macroanatomic findings. Spine 19:192–198