Multi-class Multi-label Classification and Detection of Lumbar Intervertebral Disc Degeneration MR Images using Decision Tree Classifiers

Isam Abu-Qasmieh*
Department of Biomedical Systems and Informatics Engineering, Yarmouk University, PO box 566, Irbid 21163, Jordan
* E-mail of the corresponding author: iabuqasmieh@yu.edu.jo

Abstract
Evidence-based medicine decision-making based on computer-aided methods is a new direction in modern healthcare. Data Mining Techniques in Computer-Aided Diagnosis (CAD) are powerful and widely used tools for efficient and automated classification, retrieval, and pattern recognition of medical images. They become highly desirable for the healthcare providers because of the massive and increasing volume of intervertebral disc degeneration images. A fast and efficient classification and retrieval system using query images with high degree of accuracy is vital. The method proposed in this paper for automatic detection and classification of lumbar intervertebral disc degeneration MRI-T2 images makes use of texture-based pattern recognition in data mining. A dataset of 181 segmented ROIs, corresponding to 89 normal and 92 degenerated (narrowed) discs at different vertebral level, was analyzed and textural features (contrast, entropy, and energy) were extracted from each disc-ROI. The extracted features were employed in the design of a pattern recognition system using C4.5 decision tree classifier. The system achieved a classification accuracy of 93.33% in designing a Multi-class Multi-label classification system based on the affected disc position. This work combined with its higher accuracy is considered a valuable knowledge for orthopedists in their diagnosis of lumbar intervertebral disc degeneration in T2-weighted Magnetic Resonance sagittal Images and for automated annotation, archiving, and retrieval of these images for later usage.

Keywords: Data Mining, Image Processing, Lumbar Intervertebral Disc Degeneration, MRI-T2, Decision Trees, Multi-class Multi-label Classification.

1. Introduction
Degenerative lumbar discs have been suggested as a potential cause of lower back pain [1, 2]. Therefore, development of a classification system that can detect degeneration at different stages and different positions (discs), helps in archiving and retrieval of Intervertebral Disc Degeneration MR Images. It also provides a robust tool for optimizing the diagnostic decision making. Clinically, T2-weighted magnetic resonance imaging (MRI) at sagittal plane is the modality of choice for diagnosing Degenerative Disc Disease (DDD) where changes in disc MRI signal accurately reflect the presence or absence of degenerative changes seen on discography in patients with low-back pain [3].

Machine learning techniques have been widely and successfully applied in Computer-Aided Diagnosis [4] [5] [6]. Data mining as a machine learning technique has been used as a powerful tool for medical image classification [7-13]. In this study, the decision-tree model is used for classification of the intervertebral disc degeneration. Computerized approaches, based on disc morphology, have been proposed for the characterization of intervertebral disc degeneration [14, 15]. Texture has been widely used in image classification with an application of image retrieval [16]. The most commonly extracted features are Contrast, Energy, Entropy, Homogeneity, and Correlation [17].

In the present study, a texture-based pattern recognition system is proposed for the assessment of lumbar intervertebral disc degeneration. Textural features are generated from MR images of normal and degenerated lumbar intervertebral discs and the extracted features (contrast, energy, and entropy) of the preprocessed MR images are employed in the design of a classification scheme, used for the discrimination between normal and degenerated disc and which disc is the affected one. MR image preprocessing includes image denoising using Gaussian filter, image enhancement using histogram equalization, and image edge detection using Canny edge detection algorithm.
Decision trees are usually used for classification to predict what group a case belongs to [18]. The decision trees generated by C4.5 [19] are used for classification in this study. J48 is an open source Java implementation of the C4.5 algorithm in the Weka data mining tool [20]. J48 in WEKA refers to Quinlan’s C4.5 algorithm with optional pruning. J48 is used to build decision trees from a set of labeled training data using the concept of information entropy. To split the data at each stage of the tree construction, a test is performed to select an attribute with the lowest entropy [21].

2. Methodology

181 T2-MR sagittal images, depicting the lower lumbar spine, were obtained from a local public hospital, corresponding to 89 normal and 92 degenerated discs. The data is randomly divided into 151 training and 30 testing sets.

The images were subjected to preprocessing before starting of training and testing. Initially, all images have been resized to a common resolution (256 x 256) in order to reduce biasing effects. In each of the resized spine images, the region of interest (ROI) were segmented at the lumbar intervertebral discs region by cropping all images automatically at fixed coordinates, assuming that the acquired spine sagittal images are typically depicting the spine discs at the same tissue slice plane and scene (which is the case most of the time). The segmented partition of each image has been histogram equalized to improve the images contrast and to expand the intensity values range. Image denoising and smoothing using Gaussian filtering is performed. Then an embedded operation of Canny edge detection operator which was applied to each segmented and enhanced partition to extract the image boundaries as thin single-pixel lines. Fig1 shows the preprocessing stages of the spine MR image as a primary step in the pattern recognition adopted system which is then followed by features extraction step.

The classification-based automatic image annotation was performed by extracting global texture features using Gray level co-occurrence matrix. The selected global features are entropy, contrast, and energy.

The extracted features are then stored in excel spreadsheet used by Quinlan’s C4.5 classifier to construct the Multi-class Multi-label annotation model. The model is used to classify the lumbar disc images as normal or degenerated and if it is degenerated a sub-classification is made based on which disc is affected. The J48 decision tree generated from WEKA tool was used as an implementation of Quinlan’s C4.5 algorithm.

3. Results and Discussion

151 multi-labeled T2-MR Images where randomly selected for J48 decision tree training algorithm with optional pruning. The J48 pruned tree is converted into if-then-else rules from the set of predefined labeled training images in order to get a decision tree that minimizes the error rate and thus producing the highest number of correctly classified labeled images. The resultant Decision tree model from the input features is illustrated in Fig 2. The Decision tree model can be effectively used to determine the most important attributes in the examined dataset. The attribute which maximizes the gain ratio or information gain the most is selected as the root node or top node. The value of the attribute at which this gain occurs is obviously the split point. As we notice from the tree model, the root node is the energy attribute.

The output report of the decision tree's if-then rules implemented on the testing preprocessed images is shown in Table1.

To generate annotations, each image passes through a series of three-class classifiers for each diagnosis. In our annotation model, each word is assigned independently by the decision tree obtained for each diagnosis from texture feature attributes namely, contrast, entropy, and energy.

Table 2. illustrates the results of the J48 classifier, here 151 spine images are taken for training and 30 images are taken for the testing in both normal and degenerated cases. The results in Table 2. show that the proposed system gives an accuracy of 93.33 percent.

The k-means algorithm is the most commonly used partitional clustering algorithm because it can be easily implemented and is the most time-efficient algorithm in the field of clustering. The central concept in the k-means algorithm is the centroid [22]. In data clustering, the centroid of a set of data observations is the one observation that’s most representative of the group. Fig3 shows that the result of k-means clustering Methods using WEKA tool gives a well separated (non-overlapping) clusters for images with different diagnosis.

Michopoulou et al [23] has achieved an overall classification accuracy of 93.8% in their work of designing a
classification system for Cervical intervertebral discs, where they used the Least Squares Minimum Distance Classifier in combination with four textural features in order to classify the cervical discs as normal or degenerated in 32 sagittal magnetic resonance images. This accuracy was obtained by using the same set of images for training and for testing as well. In our work we have used two different datasets for training and testing (151 images for training and another 30 images for testing) which gives our model a broadest attainable generality of applicability. Moreover, our model is a Multi-class Multi-label annotation model which sub-classifies the degenerated discs based on which disc in the lumbar region is affected.

4. Conclusion
The proposed classifier has been successfully applied in constructing a decision tree model with promising degree of accuracy for Multi-class Multi-label intervertebral disc degeneration T2-MR images classification, detection, and retrieval. This accuracy is assumed to manifest itself in building a decision support system for diagnosis of lumbar degenerative disc disease.

References


---

**Table 1. If-then rules of the adapted decision tree algorithm.**

**J48 pruned tree**

```
class = abnormal
  | contrast <= 49.6812
  |   | energy <= 41: S1 (4.0/2.0)
  |   | energy > 41: L3 (23.0)
  | contrast > 49.6812
  |   | energy <= 41
  |   |   | contrast <= 51.04: S1 (3.0/1.0)
  |   |   | contrast > 51.04: L4 (18.0/1.0)
  |   | energy > 41
  |   |   | contrast <= 53.875
  |   |   |   | energy <= 80
  |   |   |   |   | energy <= 56
  |   |   |   |   |   | entropy <= 0.5297: L5 (8.0)
  |   |   |   |   |   | entropy > 0.5297: L4 (4.0/1.0)
  |   |   |   | energy > 56: L5 (17.0)
  |   |   | energy > 80: L4 (3.0)
  |   | contrast > 53.875: L2 (3.0/1.0)
class = normal: no (80.0/2.0)
```

**Number of Leaves**: 10

**Size of the tree**: 19

**Attribute mappings**

**Model attributes**  **Incoming attributes**

The J48 pruned tree is converted into if-then-else rules from the set of predefined labeled training.
Table 2. The output report of the accuracy result after model training.

--- Evaluation on test set ---

Time taken to test model on supplied test set: 0 seconds

--- Summary ---

| Correctly Classified Instances | 26 | 93.333 % |
| Incorrectly Classified Instances | 2 | 6.667 % |
| Kappa statistic | 0.9084 |
| Mean absolute error | 0.0273 |
| Root mean squared error | 0.125 |
| Relative absolute error | 18.3214 % |
| Root relative squared error | 38.3386 % |
| Coverage of cases (0.95 level) | 93.333 % |
| Mean rel. region size (0.95 level) | 16.5714 % |
| Total Number of Instances | 30 |

--- Detailed Accuracy By Class ---

<table>
<thead>
<tr>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>MCC</th>
<th>ROC Area</th>
<th>PRC Area</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>0</td>
<td>1</td>
<td>0.1</td>
<td>0.099</td>
<td>0.977</td>
<td>0.9</td>
<td>0.633</td>
<td>L9</td>
</tr>
<tr>
<td>0.975</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.975</td>
<td>0.935</td>
<td>0.915</td>
<td>0.906</td>
<td>L4</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>L5</td>
</tr>
<tr>
<td>0</td>
<td>0.033</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>L1</td>
</tr>
<tr>
<td>0</td>
<td>0.033</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>L2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>L3</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.933</td>
<td>0</td>
<td>1</td>
<td>0.929</td>
<td>0.964</td>
<td>0.957</td>
<td>0.942</td>
<td></td>
</tr>
</tbody>
</table>

--- Confusion Matrix ---

The evaluation on a test dataset consisting of 30 images. The result shows an accuracy of 93.33%.

Fig1. The stage T2-MR spine image preprocessing. a) The resized original image., b) Cropped image at the region of Lumbar vertebral region, c) Histogram equalized image of the cropped image in b, and d) Canny edge detection of c.
Fig 2. Output decision tree using J48 algorithm. Figure 2 shows the value of the attribute at which information gain occurs most often which is energy feature. The disc's symbol shown in the figure means the vertebra that is below the disc, for example S1 means the disc known as L5-S1.

Fig 3. k-means clustering method in WEKA for energy feature. The x-axis indicates diagnosis whether it is normal with "no" label or degenerated labeled with the affected disc, and y-axis is for energy. The results show that clusters of normal and upnormal are very well separated.
This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE’s homepage: http://www.iiste.org

CALL FOR JOURNAL PAPERS

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. There’s no deadline for submission. **Prospective authors of IISTE journals can find the submission instruction on the following page:** http://www.iiste.org/journals/ The IISTE editorial team promises to the review and publish all the qualified submissions in a fast manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Recent conferences: http://www.iiste.org/conference/

**IISTE Knowledge Sharing Partners**

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar