Neurocomputing 287 (2018) 185–195

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Automated neural foraminal stenosis grading via task-aware structural representation learning

Xiaoxu He, Stephanie Leung, James Warrington, Olga Shmuilovich, Shuo Li*

The digital imaging group of London , Department of Medical Imaging, Western University, London, ON, N6A 3K7, Canada

ARTICLE INFO

Article history:
Received 23 August 2017
Revised 17 December 2017
Accepted 30 January 2018
Available online 8 February 2018

Communicated by Shaoting Zhang

Keywords:
Automated grading
Structure feature learning
Saliency map
Neural foramina stenosis
MR spinal image

ABSTRACT

Neural foraminal stenosis (NFS) is the most common spinal disease in elderly patients, greatly affecting their quality of life. Efficient and accurate grading of NFS is extremely vital for physicians as it offers patients a timely and targeted treatment according to different grading levels. However, current clinical practice relies on physicians' visual inspection and manual grading of neural foramina (NF), which brings the annoying inefficiency and inconsistency. A fully automated system is highly desirable but faces many technical challenges (e.g., the inefficiency in localizing NF candidates, and the severe ambiguities in grading). In this paper, an automated and accurate localization and grading clinical framework is proposed. By our framework, both localization and grading tasks are handled as multi-class classification problem: two-class classification (NF/non-NF) and four-class classification (normal/slight/marked/severe). To achieved it, a newly proposed saliency-biased Ncuts (SBNcuts) is utilized for efficient localization, and a novel task-aware structural representation learning (TASRL) model is developed for accurate localization and grading. Specifically, SBNcuts creatively incorporates saliency map as a preliminary guess of NF's locations to refine the generated possible NF candidates with the preserved intact structure of NF. TASRL incorporates task labels (e.g., NF object label and four NFS grade labels) into manifold learning to obtain a discriminative, low-dimensional, and structural image representation, which enables similar appearance sharing among images with the same task label and different appearance among images with different task labels. The superior performance in localization and grading, with very high ($>$ 0.89) accuracy, specificity, sensitivity, and F-measure, have been demonstrated by experiments on 110 subjects. With our method, physicians could offer an efficient and consistent clinical grading for NFS.

1. Introduction

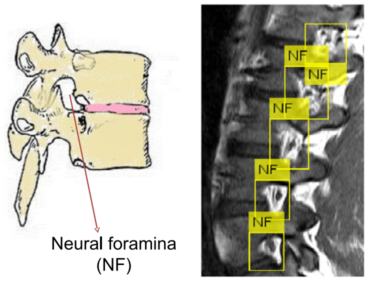
Neural foraminal stenosis (NFS) is defined as the narrowing of the bony exit (see Fig. 1(a)) of the spinal nerve root, caused by osteoarthritis, disc degeneration, and ligament thickening. It commonly appears among people and brings low back pain, neuropathy, even disability [1,2]. For example, about 80% of the adult population suffers from one symptom of NFS-low back pain [3,4]. Efficient and accurate grading is a vital step because treatment plan varies with different gradings of NFS (as shown in Fig. 1(b)). For examples, for Grade 1, normally physical therapy and exercise will be the first option for treatment, while for more severe grades, patients may need to receive surgical treatments such as decompression [3]. However, existing clinical practice relies on physicians' visual inspection and manual grading of NF images [4–6], which is not only laborious for physicians but also inefficient in providing the timely treatment. Hence, an automated clinical tool is highly needed for clinics.

However, the technical implementation of this automated framework faces great challenges, due to the diversity of NF (as shown in Fig. 2(a)) in appearance, shape, size, and pose. For example, popular sliding window detectors [7,8] relying on fixed window size may jeopardize NF's structural integrity and lower the localization accuracy;conventional image-to-image matching methods [9–12] judge the similarity of NF images based on pixel-level representations (e.g., intensity, texture and gist) and are easily mislead by the intra-object(NF)/grade difference and inter-grade ambiguities.

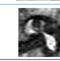
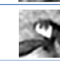
In this paper, we propose a novel fully automated clinical framework. It formulates clinical localization and grading tasks as a two-class (NF object/non-NF object) classification and four-class (normal/slight/marked/severe) classification, respectively. To solve the challenges mentioned above, it introduces a novel Saliency-Biased Ncuts (SBNcuts) and a novel Task-Aware Structural Representation Learning (TASRL) model (as shown in Fig. 2(b)).

* Corresponding author.
E-mail addresses: xiaoxuhe@mail.ustc.edu.cn (X. He), sli287@uwo.ca (S. Li).

https://doi.org/10.1016/j.neucom.2018.01.088
0925-2312/



(a)

Grade	Stage	Example
grade 0	normal	
grade 1	slight stenosis	
grade 2	marked stenosis	
grade 3	severe setnosis	

(b)

Fig. 1. The anatomy of neural foramina (a) and clinical MRI grading system for NFS (b).

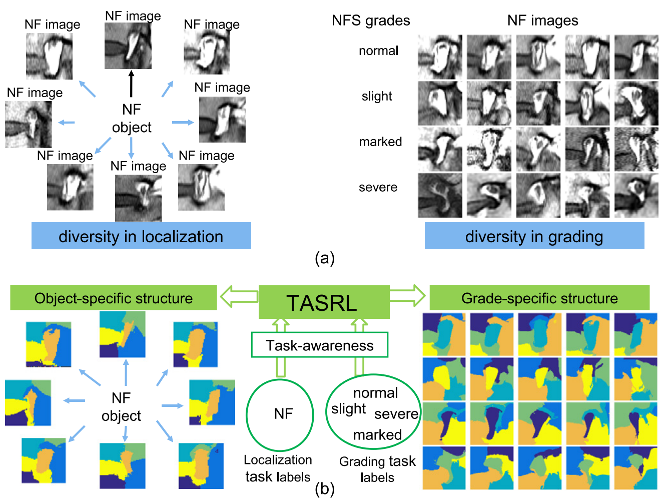


Fig. 2. Challenges from diversity of NF images for automated localization and grading(a), and the advantages (e.g., robustness and discrimination) of the learned task-specific structural representation by our TASRL(b) (Section 3.2).

SBNcuts employs salience map to capture a preliminary guess of NF's locations, and combine it with the optimal Ncuts region partition of spinal image to generate the non-overlapping NF candidates, the well-preserved NF structure in these candidates enable the efficient and accurate localization. TASRL introduces task labels as task-aware term to guide NF's structural representation learning, the generated task-specific structure not only handles the intra-object/grade incompatibility, but also overcomes inter-grade ambiguities as different grades correspond to different grade tasks and have different grade-specific structure, hence it ensures the accuracy in localization and grading task.

The novelty of our work includes the following two aspects:

1. Clinical novelty: the first automated NFS grading framework is proposed for greatly reducing the heavy burden of physicians and enabling the efficient clinical practice.
2. Technical novelty:
   - Saliency-based localization: an efficient and accurate NF detector SBNcuts is proposed, it firstly introduces saliency map of spinal MR image as the rough estimation of the NF's distribution to constraint the segmentation of spinal MR preserving the intact structure of NF region.
   - Task-aware: a novel TASRL model is proposed to learn the optimal representation customized by different tasks. Specifically, robust object-specific structural representations would be learned for localization, and discriminative grade-specific structural representations would be learned for grading.
   - Label-specific: a novel TASRL model is proposed to learn the intra-label robust representation for tolerating the intra-object difference in localization and intra-grade difference in grading task to achieve robust performance.
   - Discriminative: a novel TASRL model is proposed to enlarge the inter-label discrimination by reducing the intra-label difference, which overcomes the inter-grade ambiguities in the grading task to achieve accurate grading.

2. Background

Although popular sliding window object detection and image-to-image matching methods have made great successes, they are not suitable for dealing with the complex and diverse NF images.

2.1. Sliding-window object detection and its limitations

Sliding-window object detection is a popular technique for identifying and localizing objects in an image [7,8]. It firstly scans the image with a fixed-size rectangular window then applies a classifier to the sub-image defined by the window. The classifier extracts image features from within the window (sub-image) and returns the probability that the window (tightly) bounds to a particular object.

However, it is not suitable for our problem due to the following limitations: (1) the various size and shape of NF bring sliding window method a great difficulty in setting the suitable value to fix the window size; (2) the highly overlapping sub-images greatly lower the efficiency as the number of irrelevant sub-images is so huge to be judged.

2.2. Image-to-image matching and its limitations

Accurate image matching method is critical to image classification problems (e.g., object localization and object refined classification) as it directly affects the accuracy of these tasks [9,13–16]. Existing methods for performing image matching are usually based on image-to-image matching which compares two images based on their global or local image representations (e.g., raw pixel intensity or color values [11], shape-based descriptors using extracted edges [10], and Semantic context [17,18]).

Although image-to-image matching has made great successes [9–12], it is not suitable for our problem. Specifically, strict image matching in localization is sensitive to the disturbance from diverse size, poses, intensity profiles, and shapes; weak discriminant ability in grading is sensitive to the difference for NF images with the same grades and easily confused by inter-grade ambiguities; the limited samples of NFS, caused by the difficulty in collection and label, poses big challenges for some supervised learning methods relying on massive training [19,20].