Artificial Intelligence in Medicine 62 (2014) 79–90

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

From spoken narratives to domain knowledge: Mining linguistic data for medical image understanding

Xuan Guo[a,*], Qi Yu[a], Cecilia Ovesdotter Alm[b], Cara Calvelli[c], Jeff B. Pelz[d], Pengcheng Shi[a], Anne R. Haake[a]

[a] College of Computing & Information Sciences, Rochester Institute of Technology, 20 Lomb Memorial Drive, Rochester, NY 14623, USA
[b] College of Liberal Arts, Rochester Institute of Technology, 92 Lomb Memorial Drive, Rochester, NY 14623, USA
[c] College of Health Sciences & Technology, Rochester Institute of Technology, 90 Lomb Memorial Drive, Rochester, NY 14623, USA
[d] Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY 14623, USA

ARTICLE INFO

Article history:
Received 2 June 2013
Received in revised form 29 July 2014
Accepted 10 August 2014

Keywords:
Unified Medical Language System
Lexical consensus
Semantic relatedness
Clustering algorithm
Image-based diagnostic reasoning
Medical data analysis

ABSTRACT

Objectives: Extracting useful visual clues from medical images allowing accurate diagnoses requires physicians' domain knowledge acquired through years of systematic study and clinical training. This is especially true in the dermatology domain, a medical specialty that requires physicians to have image inspection experience. Automating or at least aiding such efforts requires understanding physicians' reasoning processes and their use of domain knowledge. Mining physicians' references to medical concepts in narratives during image-based diagnosis of a disease is an interesting research topic that can help reveal experts' reasoning processes. It can also be a useful resource to assist with design of information technologies for image use and for image case-based medical education systems.
Methods and materials: We collected data for analyzing physicians' diagnostic reasoning processes by conducting an experiment that recorded their spoken descriptions during inspection of dermatology images. In this paper we focus on the benefit of physicians' spoken descriptions and provide a general workflow for mining medical domain knowledge based on linguistic data from these narratives. The challenge of a medical image case can influence the accuracy of the diagnosis as well as how physicians pursue the diagnostic process. Accordingly, we define two lexical metrics for physicians' narratives—lexical consensus score and top *N* relatedness score—and evaluate their usefulness by assessing the diagnostic challenge levels of corresponding medical images. We also report on clustering medical images based on anchor concepts obtained from physicians' medical term usage. These analyses are based on physicians' spoken narratives that have been preprocessed by incorporating the Unified Medical Language System for detecting medical concepts.
Results: The image rankings based on lexical consensus score and on top 1 relatedness score are well correlated with those based on challenge levels (Spearman correlation >0.5 and Kendall correlation >0.4). Clustering results are largely improved based on our anchor concept method (accuracy >70% and mutual information >80%).
Conclusions: Physicians' spoken narratives are valuable for the purpose of mining the domain knowledge that physicians use in medical image inspections. We also show that the semantic metrics introduced in the paper can be successfully applied to medical image understanding and allow discussion of additional uses of these metrics.

1. Introduction

Image understanding is an important topic studied in imaging, computing, and the cognitive sciences and incorporates the domain knowledge of target images, human vision and psychophysics, and data mining. In order to better perform computational image understanding tasks, such as object detection [1–3], shape estimation [4], or depth estimation [5], common knowledge from

* Corresponding author at: B. Thomas Golisano College of Computing and Information Sciences, 20 Lomb Memorial Drive, Rochester, NY 14623, USA. Tel.: +1 585 201 2213.
E-mail addresses: andyroddick1989@gmail.com, xxg3358@g.rit.edu (X. Guo).

http://dx.doi.org/10.1016/j.artmed.2014.08.001
0933-3657/

humans is borrowed and injected into a variety of algorithms. However, hard-coded human knowledge cannot be easily and directly applied for complex tasks such as image classification and retrieval, which involve a wide range of images and require human expertise. For example, content-based image retrieval (CBIR), as an application of image understanding, is subject to the "semantic gap" between visual image features and the richness of human understanding [6–8]. Because image-feature representations often require high dimensionality and are described by complex semantics, incorporating human knowledge and high-level image representation into computational image classification is necessary [9]. Since language is the primary conduit of expressing meaning, researchers are currently integrating verbal metadata with image features [10] and moving toward new directions in CBIR, such as association-based image retrieval (ABIR) and perception-based image retrieval (PBIR) [11–13]. We take these insights in new directions by using novel linguistic data for medical image understanding.

1.1. The importance of expert knowledge in medical image understanding

Classification of medical images can benefit from experts' input of domain knowledge, because the semantics are not necessarily described by salient low-level visual features captured by computer vision algorithms [14]. Physicians in the specialties such as radiology and dermatology have developed perceptual expertise. They know where to look and what to look for and perform better than unsupervised algorithms that lack guidance from medical knowledge [15,16]. However, it is time consuming and impractical for physicians to manually annotate medical images since these images are usually stored in large-scale image databases with a large and rapidly growing number of digital images. We address this issue by collecting data from physicians as they engage in medical image inspection. Our approach to gathering data for analysis during physician image inspection involves a more natural task for experts to perform in contrast to asking them to conform to predefined image annotation labels and rules. Therefore, properly tracing the use of human knowledge—in our case physicians' knowledge during the diagnostic reasoning process—becomes an interesting and meaningful topic. It also furthers the theoretical understanding of experts' reasoning and decision making in medical domains and thus facilitates the adequate use of medical images in medical applications. In order to make these contributions, we conducted an experiment that recorded physicians' spoken descriptions during medical image inspection. The descriptions were then transcribed to diagnostic narratives and preprocessed by programs that use the Unified Medical Language System (UMLS), an ontology of medical concepts [17]. We subsequently applied computational techniques to the narratives obtained in order to benefit medical image understanding.

1.2. Challenges in language data analysis for domain knowledge extraction

Since many medical images are inherently complex and noisy due to both photographic inconsistency and different presentations of even the same medical condition, grouping relevant medical images into semantically related and meaningful groups has been a long-standing challenge. Researchers have made efforts to incorporate domain knowledge in image clustering [18]. However, understanding physicians' use of knowledge remains a challenging task that must be studied in the cognitive and biomedical domains. This is especially difficult in our case as we try to trace human knowledge used during image-based diagnostic reasoning, because visual diagnostic reasoning is a complex interaction of domain knowledge, perceptual expertise, reasoning processes [19], and idiosyncratic visual information in the image case being inspected by physicians. The scope of this paper is not to understand the reasoning process itself but to analyze experts' expressions of domain knowledge during the diagnostic reasoning process, because it is difficult to know and clearly define which reasoning strategies are used in a given reasoning process [20].

Currently, studies such as work by Evered et al. [21] use non-specific measures, such as participating physicians' response time, to predict diagnostic accuracy and thus overlook much detailed information from physicians during diagnostic reasoning. We exploit human experts' knowledge to facilitate medical image grouping by applying a methodology that is more objective and automated than current research [21–23]. The intuition is that the meaning of a medical image is expected to be mirrored by the spoken narrative of a physician when s/he describes the image during a diagnostic process. In this way, we incorporate physicians' domain knowledge obtained from years of systematic study and clinical training to achieve more effective medical image grouping. However, the language data—our collected materials that contain expert knowledge—still provide multiple challenges:

- Language is, by nature, sparse [24]. In most linguistic data sets, the vast majority of lexical items tend to occur rarely, and speakers can express similar meaning in a variety of ways, both syntactically and lexically.
- Semantic ambiguity occurs in language data. The understanding and successful professional interpretation of ambiguous language data depends on knowledge in the field. For example, the medical term *plaque* refers to a primary lesion type in the dermatology domain, while it could be interpreted as dental plaque by a dentist.
- Segmentation of medical text, such as in lexical tokenization, requires a combination of language processing techniques and medical knowledge. For example, *contact dermatitis* is a multiword expression jointly referring to a specific medical condition as opposed to two individual word tokens.
- Medical concepts[1] have different levels of semantic scope and specificity. For example, *contact dermatitis* is a representative example of *dermatitis* that is manifested by reactivity to materials or substances coming in contact with the skin.[2] In order to distinguish the diagnoses from broad (e.g., *dermatitis*) to specific (e.g., *contact dermatitis*), we need ontological knowledge and an understanding of relations between concepts.
- The difference in narration styles among multiple human experts results in variability that obscures common strategies of diagnostic reasoning.
- Language data are influenced by the mode in which they were produced. Naturally occurring speech data differ substantially from standard written text data.

1.3. Contributions

In this paper, we propose a novel framework that exploits physicians' domain knowledge in a systematic fashion to facilitate the understanding of highly complex and oftentimes noisy medical images. In particular, we leverage physicians' spoken narratives during the diagnostic process as sensors for domain knowledge collection and introduce two new lexical metrics that derive new insights into medical images. The use of spoken narratives also lead to novel ways of clustering medical images, which open a gate for organizing large-scale images generated in the medical domain on a

[1] By *concepts* in this paper, we mean the level of meaning.
[2] https://uts.nlm.nih.gov.