COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 130 (2016) 93–105

ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A novel benchmark model for intelligent annotation of spectral-domain optical coherence tomography scans using the example of cyst annotation☆

Ehsan Shahrian Varnousfaderani, Jing Wu, Wolf-Dieter Vogl, Ana-Maria Philip, Alessio Montuoro, Roland Leitner, Christian Simader, Sebastian M. Waldstein, Bianca S. Gerendas, Ursula Schmidt-Erfurth*

Christian Doppler Laboratory for Ophthalmic Image Analysis, Vienna Reading Center, Department of Ophthalmology, Medical University of Vienna, Vienna, Austria

ARTICLE INFO

Article history:
Received 21 August 2015
Received in revised form
5 February 2016
Accepted 10 March 2016

Keywords:
Benchmark dataset
Cyst segmentation
SD-OCT

ABSTRACT

Background and objectives: The lack of benchmark data in computational ophthalmology contributes to the challenging task of applying disease assessment and evaluate performance of machine learning based methods on retinal spectral domain optical coherence tomography (SD-OCT) scans. Presented here is a general framework for constructing a benchmark dataset for retinal image processing tasks such as cyst, vessel, and subretinal fluid segmentation and as a result, a benchmark dataset for cyst segmentation has been developed.

Method: First, a dataset captured by different SD-OCT vendors with different numbers of scans and pathology qualities are selected. Then a robust and intelligent method is used to evaluate performance of readers, partitioning the dataset into subsets. Subsets are then assigned to complementary readers for annotation with respect to a novel confidence based annotation protocol. Finally, reader annotations are combined based on their performance to generate final annotations.

Result: The generated benchmark dataset for cyst segmentation comprises 26 SD-OCT scans with differing cyst qualities, collected from 4 different SD-OCT vendors to cover a wide variety of data. The dataset is partitioned into three subsets which are annotated by complementary readers based on a confidence based annotation protocol. Experimental results show annotations of complementary readers are combined efficiently with respect to their performance, generating accurate annotations.

Conclusion: Our results facilitate the process of generating benchmark datasets. Moreover the generated benchmark data set for cyst segmentation can be used reliably to train and test machine learning based methods.

☆ *Grant*: Austrian Federal Ministry of Science, Research and Economy; National Foundation for Research, Technology and Development.

* *Corresponding author*. Tel.: +43 14040079310.
E-mail address: bianca.gerendas@meduniwien.ac.at (B.S. Gerendas).

http://dx.doi.org/10.1016/j.cmpb.2016.03.012

1. Introduction

In the areas of ophthalmic disease assessment, diagnosis, and treatment planning, analysis of retinal morphology plays an important role [1–5]. The analysis of retinal cysts, sub-retinal fluid, and fluid under the retinal pigment epithelium holds the key to understanding and treating diseases such as age-related macular degeneration (AMD) [6–8], diabetic macular edema (DME) [9] (Gerendas, unpublished data, 2014) and macular edema due to retinal vein occlusion (MEVO) (Waldstein, unpublished data, 2014). Patients are imaged using spectral-domain optical coherence tomography (SD-OCT), a non-invasive modality for acquiring high resolution, 3D cross sectional volumetric images of the retina and the sub-retinal layers [10,11].

Today, SD-OCT is the most important ancillary test for the diagnosis of sight threatening diseases [12]. For these reasons, understanding and measuring the position, size, and composition of cysts is needed, for example. Whereas the delineation of subretinal fluid can be performed easier, the ability for manual delineation of cysts is limited due to the difficulty for human experts to accurately and reproducibly identify them. In addition, when the number of cysts in an image is large, manual identification becomes tedious or even impossible, thus automated systems are preferable.

However, such systems require large volumes of annotated and variable data for training and testing. Until now, there is no publically available dataset featuring expertly annotated and investigated multi-vendor retinal cysts usable as ground truth for the development of automated or semi-automated cyst segmentation methods. As a result, it is difficult to compare methods that have been developed in this area. Such datasets have been created for other purposes in the form of the DRIVE database of digital retinal images for vessel extraction [13], the REVIEW database for the measurement of retinal vessel widths [14], and the STARE database [15] for vessel segmentation, and the database for the coronary artery algorithm evaluation framework [16].

The primary focus of this paper is the development of a novel intelligent method of reader (defined as all people who perform manual annotations for cyst delineation) evaluation, task assignment, and data partitioning of retinal SD-OCT scans based on cyst annotation and the construction of a benchmark dataset for use in training and testing. The data set used here features retinal cysts obtained from 4 of the major SD-OCT scanner vendors, annotated by expert readers at the Vienna Reading Center (VRC) and a subgroup, specially trained for cyst annotation from the Christian Doppler Laboratory for Ophthalmic Image Analysis (OPTIMA) and the evaluation of said annotations. The cysts in question have been manually annotated using a proprietary annotation tool developed for this purpose at OPTIMA. In addition, the reader evaluation and task assignment system has been designed to be applicable for other retinal pathologies other than cysts.

Previous methods of (semi-)automated cyst segmentation [9,17] have primarily used local or private datasets featuring a limited number of scans and/or only a single vendor. In addition, annotations of the cysts have been often carried out by a single non-expert reader without validated reading tools which may result in subjectivity bias or reproducibility errors. Our goal to develop a benchmark dataset of cyst annotations in SD-OCT will allow further development in (semi-)automated cyst segmentation algorithms with the aid of a fully annotated and validated dataset as well as the ability to compare developed methods using a single reference standard. For this purpose, a reader evaluation and task assignment system have been developed to evaluate the annotations of 3 readers and 1 expert supervisor on a dataset consisting of 26 SD-OCT scans from 4 major vendors taken from the VRC and OPTIMA database of patients suffering from AMD and glaucoma. Evaluation of the resulting annotations is performed, along with an intense reader evaluation to judge reader performance.

This paper is organized in 4 sections. In Section 2, the primary method for task assignment, annotation, data evaluation, and reader evaluation is described in detail. In Section 3, results of the annotation process are presented, and summarized and discussed in Section 4.

2. Methods

In this paper, we propose a general framework for constructing benchmark datasets for retinal image processing tasks such as cyst, vessel, and subretinal fluid segmentation as shown in Fig. 1 and as a result, a benchmark dataset for cyst segmentation is developed. The main contributions of this paper are the novel *data selection*, *task assignment*, and *annotation combination*

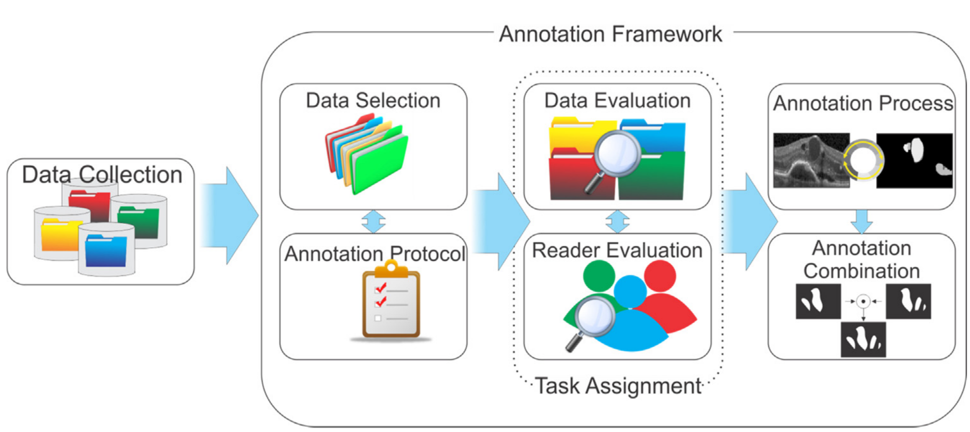

Fig. 1 – Illustration of proposed annotation framework.