Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth
A. Ugon et al. (Eds.)
© 2018 European Federation for Medical Informatics (EFMI) and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/978-1-61499-852-5-396

SCREEN-DR

Software Architecture for the Diabetic Retinopathy Screening

Micael PEDROSA[a,1], Jorge Miguel SILVA[a], Sérgio MATOS[a], Carlos COSTA[a]
[a] *DETI/IEETA, University of Aveiro, Portugal*

Abstract. Diabetic Retinopathy (DR) is a common complication of diabetes that may lead to blindness if not treated. However, since DR evolves without any symptoms in the initial stages, early detection and treatment can only be achieved through routine checks. This article presents the collaborative platform of the SCREEN-DR project that promotes partnership between physicians and researchers in the scope of a regional DR screening program. The role of researchers is to create classification algorithms to evaluate image quality, discard non-pathological cases, locate possible lesions and grade DR severity. Physicians are responsible for annotating datasets, including the visual delineation of lesions. The collaborative platform collects the studies, indexes the images metadata, and manages the creation of datasets and the respective annotation process. An advanced searching mechanism supports multimodal queries over annotated datasets and exporting of results for feeding artificial intelligence algorithms.

Keywords. Diabetic Retinopathy, Computer-Aided Diagnosis, Image Annotation

1. Introduction

DR is a leading cause of blindness in the industrialized world [1]. A survey conducted by the Portuguese North Health Administration (ARSN[2]), the clinical partner of this work, estimated that 75% of a population of 250.000 diabetic patients who were screened did not show the presence of the disease. The introduction of Computer-Aided Diagnosis (CAD) tools in the current ARSN screening program has the potential of leveraging gains in terms of time-to-decision and resource efficiency. The aim of the SCREEN-DR project is to research and develop an image analysis and machine learning (ML) platform for innovation in DR screening. ML algorithms acquired a significant importance in the radiology field [2]. However, the main difficulty is to find good annotated datasets for developing supervised ML techniques [3]. ARSN manages a centralized repository containing images acquired during the screening program, namely eye fundus images that are a significant source of information for clinical decision. However, those images are not annotated. SCREEN-DR project targets the collaboration between physicians and researchers in the scope of a regional DR screening program and has two main goals. The first is to serve as a platform where ophthalmologists can grade image quality, diagnose cases, and pinpoint lesions present on the images. The second goal is to enable

---

[1] Corresponding Author
[2] http://www.arsnorte.min-saude.pt/portal/page/portal/ARSNorte

the creation of CAD algorithms that prioritize pathological cases in the patient's waiting list to be observed by doctors. This is supported by content discovery and retrieval mechanisms that allow researchers to gather annotated datasets to feed their ML algorithms. This article is focused on the solution for the first goal.

Regarding the DR automatic analysis, although there are positive results in the literature, the majority of the research is highly optimized for small datasets [4,5], compromising the generalization and the required standards for mass screening applications. Projects with larger datasets require the extensive participation of ophthalmologists in the annotation phase [6], and an online accessible platform for them to use. There are already several attempts to construct such platforms such as CrowdFlower[3] or RectLabel[4]. Yet in some cases we need the annotation protocol to be highly optimized for working with thousands of images [7]. Google has created an annotation tool [6] to support the development of automatic DR diagnosis algorithms, but it was designed for annotating a single dataset. Furthermore, it does not allow the annotation of lesions, and therefore it is not possible to validate if their algorithms can identify the regions where lesions are present.

This article describes the software architecture of a collaborative platform that tries to comply with previous demands. The platform collects anonymized studies from the ARSN repository, indexes the images metadata, and allows the creation and management of distinct datasets and respective annotation process (including visual annotation of lesions), according to researchers needs. Moreover, it supports the creation of new datasets based on multimodal selection criteria that may include image metadata as well as information from previous annotations. For instance, a researcher may want a complementary visual annotation on images containing lesions of a specific type. Finally, physicians can have distinct annotation datasets assigned to them.

2. Methods

2.1. Requirements

The software platform must support two main actors, researchers and physicians/annotators. Physicians will be responsible for adding attributes to eye fundus images in accordance with the Annotation Protocol. The platform must accept multiple annotations from different physicians on the same image, to permit the evaluation of agreement for different combinations of annotators and algorithms. Researchers must be able to import/upload DICOM standard files, which should be automatically indexed for future searches. The search mechanism will use indexed DICOM fields and also the annotation fields, retrieving all the annotations for each image. Users should be able to select, from the search results, the ones they want to use to create a dataset or to download locally. For download, a ZIP containing the raw DICOM files and the annotations in JSON format should be generated. To control the access to each feature, a Role-Based Access Control (RBAC) system was implemented. The platform was developed with pure web technologies so it can be accessed anywhere using distinct computing devices, including mobile ones.

[3] https://www.crowdflower.com/use-case/image-annotation/

[4] https://rectlabel.com/

2.2. Annotation Protocol

The annotation process follows the clinical pipeline from the DR Screening manual, and is divided in three steps: Quality Assessment, Grading and Lesion Annotation. In terms of quality assessment, the images will be assigned to one of the following three classes: *Bad* - if the image contains noise or other distortions that prevent the annotator to grade the image. These cases are discarded from the next steps of the pipeline; *Partial* - if the image has some kind of distortion, but it does not jeopardize the physician's grading; *Good* - if the image does not contain distortions or noise.
Regarding the diagnostic grading step, the eye fundus image will be ranked with respect to Retinopathy, Maculopathy and Photocoagulation levels, as well as any other suspected comorbidities, which can be selected from a predefined list. The grading levels in this regard are:

- Retinopathy: *R0* - No DR, *R1* - Mild non-proliferative DR, *R2-M* - Moderate non-proliferative DR, *R2-S* - Severe non-proliferative DR and *R3* - Proliferative DR.
- Maculopathy: *M0* - No Diabetic Macular Edema and *M1* - Diabetic Macular Edema.
- Photocoagulation: *P0* - No Photocoagulation, *P1* - Insufficient Photocoagulation and *P2* - Sufficient Photocoagulation.

Healthy classified images will have a pre-selection with the Retinopathy - *R0* and Maculopathy - *M0* states. In the visual lesion annotation, physicians delineate the regions of the lesions present in the image. In this last stage, only images that were labeled as containing signs of DR (*R1-3*) will be used. Lesions are delineated directly in the pathological images with Vector Graphic tools (ellipses, polygons and line paths), identifying the following types: *MA* - Microaneurysms, *HEM* - Hemorrhages, *HE* - Hard exudates, *SE* - Soft exudates and *NV* - Neovascularizations.

2.3. Architecture

A comprehensive architecture of the SCREEN-DR collaborative platform software is depicted in Figure 1. The DR images will be stored and indexed using Dicoogle Open Source [8], a Picture Archiving and Communication System (PACS) that replaces the traditional relational database with a more agile process of indexing and retrieving mechanism. This software has a plugin based architecture that facilitates the creation of new PACS functionalities. We can take advantage of these features to create the Retinopathy-PACS (R-PACS) plugin, that stores the DICOM model (DIM) in a Relational Database Management System (RDBMS) and indexes it. Furthermore, the R-PACS plugin can also extends the DIM with a generic annotation model, using the PostgreSQL JSONB data type.

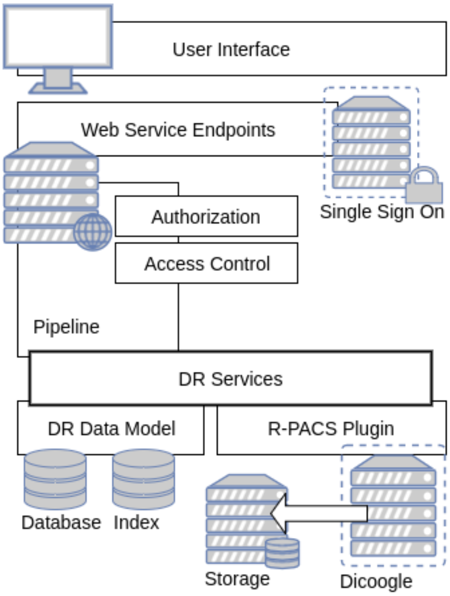

Figure 1. The SCREEN-DR overall architecture.

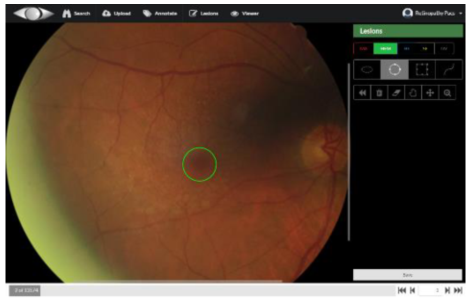

Figure 2. The lesion annotation UI, showing the several Vector Graphic tools available.

The layer of DR Services rests on top of the R-PACS. This layer is comprised of five services: Annotation, Transfer, Index, Search and Dataset. The Annotation service is responsible for fetching, creating and modifying annotations, based on the image reference ID parameter. The Transfer service is used to download or upload DICOM files, where the Index service uses selections of these previously uploaded files to index them in the R-PACS. The Search service uses Lucene like queries[5] using DICOM or annotations fields to retrieve the images from the index. Finally, the Dataset service creates datasets from selected images unique identifiers (UIDs), provides subscription methods to datasets, and also selection of the default dataset to use in the annotation protocol.

The Web Service Endpoints are implemented with an in-house service stack that we call Reactive Through Services (RTS), and exposes asynchronous Java services available at the DR Service layer. RTS can deliver reactive services (e.g. pub/sub messaging) and also common REST services. Finally, an authentication and authorization mechanism was implemented to control the access and permissions for each type of platform user (researchers and annotators). A Single sign-on (SSO) is used to grant JSON Web Tokens (JWT)[6] for the RBAC. These tokens are intercepted and validated in the RTS Pipeline.

The User Interface (UI) was separated in different groups for each required feature: a search tab for querying the system and creating datasets from search results, a grading tab where the annotators can grade the image based on textual annotations, a lesion annotation tab where physicians delineate the regions of lesions in the image and a tab for uploading new DICOM files into the platform. In Figure 2 is an example of the lesion annotation UI.

[5] https://lucene.apache.org/core/2_9_4/queryparsersyntax.html

[6] https://jwt.io

3. Conclusions

This article presents an innovative collaborative platform for supporting a DR research project. It includes an annotation framework with a set of unique functionalities. The platform is extremely important to support research on CAD mechanisms but can also be used for providing teaching cases to undergraduate students. Step-by-step learning protocols [9] can be constructed by reusing the index and search features. It is a Web system for teams that what to experiment different methodologies, in datasets with distinct characteristics. In the future, we plan to extend the platform functionalities and processes for supporting any medical annotation procedure, as also adding other interoperability features like providing annotations in DICOM-SR.

4. Acknowledgments

This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalization - COMPETE 2020 Programme, and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project CMUP-ERI/TIC/0028/2014.

References

[1] Bhavsar AR. Diabetic Retinopathy;. Accessed: 2017-10-10. https://emedicine.medscape.com/
[2] article/1225122-overview.
[3] Suzuki K. Machine learning in computer-aided diagnosis of the thorax and colon in CT: a survey. IEICE transactions on information and systems. 2013;96(4):772–783.
[4] Gentleman R, HuberW, Carey V. Supervised machine learning. In: Bioconductor Case Studies. Springer; 2008. p. 121–136.
[5] Acharya UR, Ng EYK, Tan JH, Sree SV, Ng KH. An integrated index for the identification of diabetic
[6] retinopathy stages using texture parameters. Journal of medical systems. 2012;36(3):2011–2020.
[7] Roychowdhury S, Koozekanani DD, Parhi KK. DREAM: diabetic retinopathy analysis using machine learning. IEEE journal of biomedical and health informatics. 2014;18(5):1717–1728.
[8] Gulshan V, Peng L, Coram M, Stumpe MC,Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama. 2016;316(22):2402–2410.
[9] Papadopoulos DP, Uijlings JR, Keller F, Ferrari V. Extreme clicking for efficient object annotation. arXiv preprint arXiv:170802750. 2017;.
[10] Costa C, Ferreira C, Basti˜ao L, Ribeiro L, Silva A, Oliveira JL. Dicoogle-an open source peer-to-peer PACS. Journal of digital imaging. 2011;24(5):848–856.
[11] Ruble JE, Lom B. Online protocol annotation: a method to enhance undergraduate laboratory research skills. CBE-Life Sciences Education. 2008;7(3):296–301.