

Image Analytics Techniques for Diabetic Retinopathy Detection

Ramon Pires¹, Anderson Rocha¹, Jacques Wainer¹

¹Institute of Computing – University of Campinas (UNICAMP)
13.083-852 – Campinas – SP – Brazil

{pires.ramon, anderson.rocha, wainer}@ic.unicamp.br

Abstract. *Diabetic Retinopathy is the leading cause of blindness in working-age adults. This work aims at enhancing lesion detection, reinforcing referral decisions, and integrating our solutions with low-cost retinal imaging devices. For lesion detection, we proposed a novel coding technique robust to any kind of lesion. For referral decisions, we designed a robust method that does not rely upon lesion detection, proposed an effective data-driven model that significantly improves the performance, designed an accountable model that produces a reliable response and enables pixel-based importance comprehension, and create local descriptors that are encoded into a rich mid-level representation. Our work has invaluable impacts both in biomedical and technical contexts.*

1. Introduction

Every eleventh person in the world suffers from diabetes mellitus, a disorder of sugar metabolism, whose prevalence is expected to reach every tenth person by 2040. With 46.5% of cases currently undiagnosed, a vast amount of people with diabetes are progressing towards complications unawares. Diabetes sufferers are 25 times more likely to suffer from sight loss resulting from diabetic retinopathy (DR), a major long-term microvascular complication, and the leading cause of blindness in high-income countries.

An early DR detection has a huge impact on prognosis and is crucial to prevent vision loss or limit its progression. The World Health Organization recommends annual eye examinations for diabetic patients. However, disfavored communities often lack screening opportunities to adopt frequent consultations and a continuous follow-up program.

In this context, we present herein a complete solution for diabetic retinopathy lesion detection, automated screening, and hardware integration concerned into quality assessment. Our main purpose under the scientific viewpoint was providing experts with higher-level retinal image representations that preserve valuable visual characteristics.

The predominant aim was enhancing the results reported in the literature by designing and reporting methods for handcrafted two-tiered image representation, composing and delineating sophisticated data-driven models employing novel and also consolidated techniques, and ultimately incorporating supervised learning of features with advanced and saliency-oriented mid-level image representations.

1.1. Awards and Partnerships

In the course of this work¹, we established relevant partnerships and gathered significant awards. Our project received the Google Latin American Research Awards twice (2016

¹An extension of the master thesis awarded 3rd place in the XXVII Theses and Dissertations Competition of the XXXIV Brazilian Computer Society Congress

and 2017). With that compelling award, we became a Google's partner.

We have also partnered with the startup Phelcom Technologies². The main purpose was combining the developed image analytics techniques for DR detection with the *Eyer*, a hardware solution that is transforming the smartphone into a portable and low-cost retinal camera. We set forth a partnership, in which we mentored the startup by implementing solutions suitable for embedded systems, and validating with images from the devices.

2. Datasets

In our work, we used eight different retinal image datasets tagged by medical specialists:

- **DR1 dataset** comprises 5,776 images, including poor- and good-quality images, with or without lesions. The dataset is annotated in image level for quality and the commonest DR lesions, and additional auxiliary coordinates that delimit lesions.
- **DR2 dataset** comprises a total of 920 images. The dataset is annotated in image level for quality, presence of DR lesions, and need of referral.
- **Messidor dataset** has 1,200 images. The dataset is annotated in image level for retinopathy grade (0 to 3) and risk of macular edema.
- **Messidor-2 dataset** contains 1,748 images (two per patient). Images were independently graded in image level according to the ICDR scale and macular edema.
- **Inala dataset** comprises 30 images. The dataset is annotated in image level according to presence/absence of red and bright lesions.
- **Kaggle/EyePACS dataset** comprises a total of 88,702 images (two per patient). The dataset is annotated in image level in five classes according to the ICDR scale.
- **IDRiD dataset** comprises 516 images. It is annotated in image level according to ICDR scale, and a portion has additional pixel-level annotations about DR signs.
- **Phelcom dataset** has 600 eye-fundus images. The dataset is annotated in image level for quality assessment.

3. Diabetic Retinopathy Lesion Detection

Assuming the interest in adopting unified approaches suitable for every lesion, our current research regarding DR lesion detection relies on the following questions:

Q1.1: Can we combine the advantages of both *hard* and *soft* codings and provide a good balance for designing efficient and effective DR lesion detectors?

YES: *The proposed semi-soft coding is efficient and effective for DR lesion detection.*

Q1.2: Can we preserve information in the pooling and still obtain satisfactory results?

YES: *Preserving information in the pooling process still produces satisfactory results.*

Q1.3: Can we improve the results employing global codebooks?

NO: *Global codebooks do not improve the results achieved with the class-aware method.*

Q1.4: Can we diagnose patients in an ethnic group (e.g., isolated indigenous communities), even using models trained with data from other different ethnic groups?

YES: *The proposed method is robust for diagnosing people from different ethnic groups.*

For this topic, we have published three papers: two were in the PLOS One journal [Pires et al. 2014b, Pires et al. 2015] for responding questions *Q1.1* and *Q1.4*, respectively; and one was at the IEEE Engineering in Medicine and Biology Society (EMBC'14) conference for responding questions *Q1.2* and *Q1.3* [Pires et al. 2014a].

²<https://www.phelcom.com.br/en>

4. Referable Diabetic Retinopathy Detection

Most existing art focuses on the detection of DR lesions using either visual characteristics specific to each type of lesion or unified DR-lesion. It is much harder (and polemic) to decide automatically to refer or not the patient to the ophthalmologist. Automated referral is a hot topic, because DR risk assessment is complex, based not only on the presence of lesions and their evolution, but also on subtle hints revealed during examination.

Assuming the interest in proposing effective referability assessment methods and preserving relevant information, we crafted a series of questions within the scope:

Q2.1: Can we forgo the DR lesion detection and still have an effective referral decision?

YES: *Forgoing individual DR lesion detection still provides effective referral decisions.*

Q2.2: Can sophisticated handcrafted mid-level features improve direct referral decision?

YES: *Sophisticated mid-level features improve direct referral decisions.*

Q2.3: Can we confirm the suitability of handcrafted direct referral in a second dataset?

YES: *The direct referral is suitable also for independent datasets.*

Q2.4: Can we learn in the data-driven manner highly abstract features that leverage referable DR detection without the need of manual feature engineering?

YES: *Data-driven models leverage referable diabetic retinopathy detection.*

Q2.5: Can we diagnose images taken under different conditions with data-driven models?

YES: *It's possible to pinpoint referral, independently of operators and camera settings.*

Q2.6: Can we transfer knowledge acquired with a different but similar task in the context of diabetic retinopathy screening?

YES: *Fine-tuning has potential to enhance considerably the effectiveness of the solution.*

Q2.7: Can we provide an effective and accountable data-driven solution for referable diabetic retinopathy detection?

YES: *Heatmaps reflecting pixel importance turn data-driven methods self-explainable.*

Q2.8: Can we enhance the performance of the referral model by applying two-tiered data-driven image representation?

YES: *Heatmap is a powerful mechanism also for robust two-tiered image representation.*

Q2.9: Can we combine global data-driven and local saliency-oriented two-tier representations for more accurate decision-making?

YES: *The fusion is promising, reducing both false positive and false negative rates.*

For this topic we have published two papers: one in the IEEE Journal of Biomedical and Health Informatics (J-BHI) [Pires et al. 2017] for responding questions *Q2.1*, *Q2.2* and *Q2.3*; and one in the Elsevier Artificial Intelligence in Medicine (AIIM) [Pires et al. 2019a] for responding questions *Q2.4*, *Q2.5* and *Q2.6*. We were also invited to publish a chapter in the book *Photo Acoustic and Optical Coherence Tomography Imaging: An Application in Ophthalmology* [Pires et al. 2019b]. The manuscript responds questions *Q2.7*, *Q2.8* and *Q2.9*.

5. Retinal Image Quality Assessment

Good quality retinal images are essential for providing reasonable and robust machine learning models (training) and for receiving reliable diagnostic responses (test). For clinical decisions, the quality is merely mandatory. Hence, before diagnosing images from low-cost devices, the images must beforehand be evaluated if it has enough quality.

Assuming the interest in integrating our solutions with simple retinal imaging devices, we crafted a question within the scope of testing with images from low-cost devices:

Q3.1: Can we assess the quality of images from portable devices with data-driven models trained with images from high-cost instruments?

YES: *Mobile models can evaluate the quality of images taken with low-cost devices.*

6. Discussions

In this work, we presented image analytics solution for diabetic retinopathy detection, that encompasses *DR Lesion Detection*, *Referable DR Detection* and *Validation with Images from Portable Devices*. The work incorporates both handcrafted and data-driven approaches, on bottom-up and top-down scenarios, one-stage (global based) and two-tiered (local based) representations, and ends up by combining methodologies proposed at the initial stages and the final stages of the doctoral program, compounding a sophisticated accountable and saliency-oriented mid-level representation.

Throughout the research, we explored prevailing and ongoing computer-vision and machine-learning methodologies, and proposed novel and advanced retinal image analytic approaches, with invaluable impacts both in biomedical and technical contexts.

One of our contributions was the proposal of a new semi-soft coding, exploring advantages of the most traditional hard-sum coding and the soft assignments (*Q1.1*). As we show with ANOVA, the semi-soft coding associated with sparse feature extraction provides a good balance for designing an efficient and effective DR lesion detector.

As another contribution, we mention the use of a cross-training methodology with results obtained from training with images from a particular ethnic group, and testing with a different population: Aboriginal and Torres Strait Islander people (*Q1.4*). This policy has important health screening and treatment impacts as the program can be trained at metropolitan clinics and taken to rural and remote areas for live screening.

In this research, we prioritized investigations of need of consultation. We proposed a novel approach to decide about referral directly from the retinal images and without preliminary lesion detection (*Q2.1*). This decision to forgo lesion detection has both theoretical motivations (use all cogent information instead of just lesion scores) and practical advantages (much simpler to implement, test, and deploy). We highlight that, in the time we proposed the methodology, direct assessment was new for referable DR.

In this work, we also presented a data-driven solution for referable diabetic retinopathy detection (*Q2.4*). We emphasize that our method is based upon investigating, in a rigorous scientific point of view, the advantages and disadvantages of applying novel and also consolidated techniques, and measuring how much it improves the solution. We show that CNNs performance can be boosted by a set of directives.

Another impactful contribution of this work was the proposed accountable and robust framework for DR screening (*Q2.7*). Prior art have been focused on accurate data-driven approaches to effective diagnostic. However, the interpretability of those models is frequently disregarded. In this vein, we proposed the use of saliency maps whose objective is twofold: highlighting regions that potentially influence the decision taken, and capturing regions of interest that could be leveraged for the final model response. In a clinical situation, the superposition could be presented to the ophthalmologists/nurses as well as to the patients to clarify the reasons behind the automated decision (see Fig. 1).

One of the major breakthroughs in the current work is the breakdown of the global

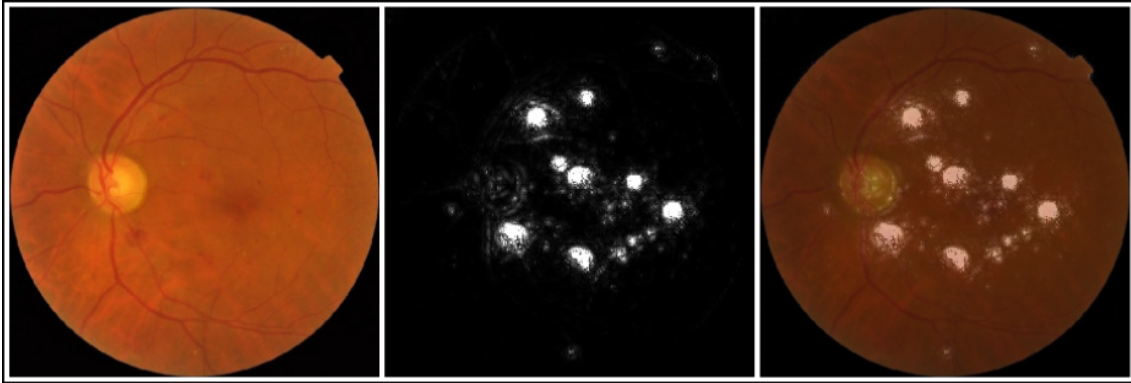


Figure 1. Retinal image; respective saliency map; and the superposition highlighting regions important for the decision

data-driven scenario (*Q2.8* and *Q2.9*). The pipeline involves extracting saliency-oriented regions of interest and combining those information through Fisher Vector (see Fig. 2). Exploring the encoded contextual local-based representations, we reduce the classification error by 27% contrasting with our previous global data-driven method testing with DR2 under cross-dataset validation protocol. By testing with Messidor-2, the robust local saliency-oriented region characteristics had a remarkable improvement, reducing the classification error by 28% over our previous global data-driven proposal (see Fig. 3). By enhancing considerably the performance both over DR2 and Messidor-2, in comparison with the strict global method, we showed that the guided mid-level representation, which emphasizes areas that the CNN model could not assimilate sufficiently, is promising.

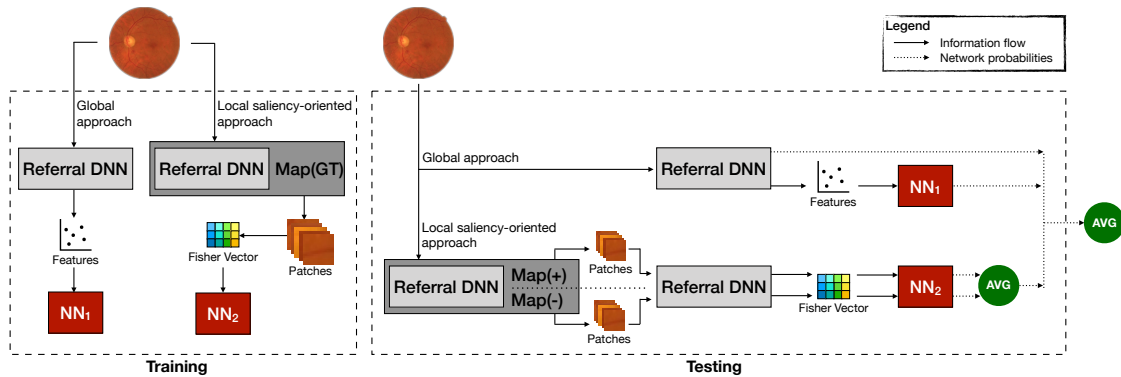


Figure 2. Overview of the proposed method with global and local representation

One paramount purpose of this work was integrating our solutions with simple and portable retinal imaging devices (*Q3.1*). As the images collected during the doctoral program were only graded for quality, we evaluated if they were properly ready for a computer-aided DR screening. With MobileNetV2 model, we achieved promising results. We also figured out that the self-annotation procedure enhanced substantially the results.

Prior research on automated DR screening has followed the natural route of identifying lesions and gathering all the extracted high-level information to evaluate referability or severity. That scenario inherits the common and natural bottom-up fashion: *lesion first, referral later*. Before the explosion of data-driven methodologies, we proposed the direct referral that contradicted previous beliefs regarding the mandatory necessity of de-

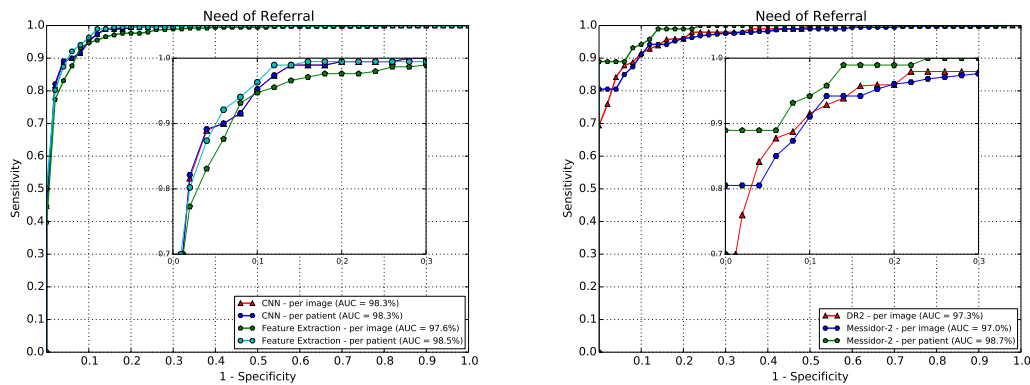


Figure 3. ROC results for referral using (left) global data-driven approach and (right) local saliency-oriented approach, testing with Messidor-2.

tecting lesions. The adoption of deep-learning strategies came to confirm our previous conclusion, dramatically improving the performance and becoming the state of the art for automated DR screening. With the recent demand for accountable solutions, the research has moved toward an opposed *top-down* direction: *referral first, lesion later*.

Moving towards an opposed direction comparing to the natural advance of research in DR diagnostic represents a strong route. Instead of detecting lesions and use the assembled information to decide about severity or referral need, in a bottom-up manner, a possible future work would involve identifying referable DR and use pixel importance for pointing out and recognizing lesions or anatomical retina parts, as a top-down approach.

References

- Pires, R., Avila, S., Jelinek, H. F., Wainer, J., Valle, E., and Rocha, A. (2014a). Automatic diabetic retinopathy detection using bossanova representation. In *Intl. Conference of the IEEE Engineering in Medicine and Biology Society*, pages 146–149.
- Pires, R., Avila, S., Jelinek, H. F., Wainer, J., Valle, E., and Rocha, A. (2017). Beyond lesion-based diabetic retinopathy: a direct approach for referral. *IEEE Journal of Biomedical and Health Informatics*, 21(1):193–200.
- Pires, R., Avila, S., Wainer, J., Valle, E., Abràmoff, M. D., and Rocha, A. (2019a). A data-driven approach to referable diabetic retinopathy detection. *Artificial Intelligence in Medicine*, 96:93–106.
- Pires, R., Carvalho, T., Spurling, G., Goldenstein, S., Wainer, J., Luckie, A., Jelinek, H. F., and Rocha, A. (2015). Automated multi-lesion detection for referable diabetic retinopathy in indigenous health care. *PLoS ONE*, 10(6):e0127664.
- Pires, R., Ferreira, A., Avila, S., Wainer, J., and Rocha, A. (2019b). An accountable saliency-oriented data-driven approach to diabetic retinopathy detection. In Elsevier, editor, *Photo Acoustic and Optical Coherence Tomography Imaging: An Application in Ophthalmology*, chapter 10. Elsevier.
- Pires, R., Jelinek, H. F., Wainer, J., Valle, E., and Rocha, A. (2014b). Advancing bag-of-visual-words representations for lesion classification in retinal images. *PLoS ONE*, 9(6):e96814.