

Extended Abstract

Learning Retinal Patterns from Multimodal Images [†]

Álvaro S. Hervella ^{1,2,*}, José Rouco ^{1,2}, Jorge Novo ^{1,2} and Marcos Ortega ^{1,2}

¹ CITIC—Research Center of Information and Communication Technologies, University of A Coruña, 17051 A Coruña, Spain; jrouco@udc.es (J.R.); jnovo@udc.es (J.N.); mortega@udc.es (M.O.)

² Department of Computer Science, University of A Coruña, 17051 A Coruña, Spain

* Correspondence: a.suarezh@udc.es; Tel.: +34-981-167-000

[†] Presented at the XoveTIC Congress, A Coruña, Spain, 27–28 September 2018.

Published: 17 September 2018

Abstract: The training of deep neural networks usually requires a vast amount of annotated data, which is expensive to obtain in clinical environments. In this work, we propose the use of complementary medical image modalities as an alternative to reduce the required annotated data. The self-supervised training of a reconstruction task between paired multimodal images can be used to learn about the image contents without using any label. Experiments performed with the multimodal setting formed by retinography and fluorescein angiography demonstrate that the proposed task produces the recognition of relevant retinal structures.

Keywords: self-supervised; multimodal; retinography; angiography

1. Introduction

In clinical practice routine, patients are typically subjected to multiple imaging tests, producing complementary visualizations of the same body parts or organs. This leaves available large sets of paired multimodal images. This paired data can be used to train a neural network to predict one modality from other. If the transformation between modalities is complex enough, the network will have to learn about the objects represented in the images to solve the task. This domain-specific knowledge can be used to complement the training of additional tasks in the same application domain, reducing the amount of labeled data required.

We applied the described paradigm to the multimodal image pair formed by retinography and fluorescein angiography. These image modalities are complementary representations of the eye fundus. The angiography has additional information about the vascular structures due to the use of an injected contrast. This also makes this modality invasive and less employed. We train a neural network to predict the angiography from a retinography of the same patient and demonstrate that the network learns about relevant structures of the eye with this self-supervised training [1].

2. Methodology

The multimodal reconstruction is trained with a set of retinography-angiography pairs obtained from the public Isfahan MISP database. This dataset includes 59 image pairs from healthy individuals and from patients diagnosed with diabetic retinopathy.

The multimodal image pairs are aligned following the methodology proposed in [2] to produce a pixel-wise correspondence between modalities. After the image alignment, a reconstruction loss can be directly computed between the network output and the target image. This allows the self-supervised training of the multimodal reconstruction, which will generate a pseudo-angiography representation for any retinography used as input to the network. Three difference functions are considered to obtain the reconstruction loss: L1-norm, L2-norm and SSIM.

The multimodal reconstruction is based on a U-Net network architecture. The network training is performed with the Adam algorithm with an initial learning rate of $\alpha = 0.0001$, which is reduced by a factor of 0.1 when the validation loss plateaus. Spatial data augmentation is used to reduce the overfitting.

3. Results and Conclusions

Examples of generated pseudo-angiographies are depicted in Figure 1. It is observed that the best results are obtained training with SSIM, in which case the network has learned to adequately transform relevant retinal structures. An additional experiment is performed to specifically measure the ability to recognize the retinal vasculature. A global thresholding is applied to produce a rough vessel segmentation from both the pseudo-angiography and the original retinography. This experiment is performed in the public DRIVE dataset, which comprises 40 retinographies and their ground truth vessel segmentation. The results are evaluated with the Receiver Operator Characteristic (ROC) curves. The measured Area Under Curve (AUC) values are 0.5811 for the retinographies and 0.8183 for the pseudo-angiographies generated after training with SSIM. This improvement demonstrates that the multimodal reconstruction provides additional information about the retinal vasculature.

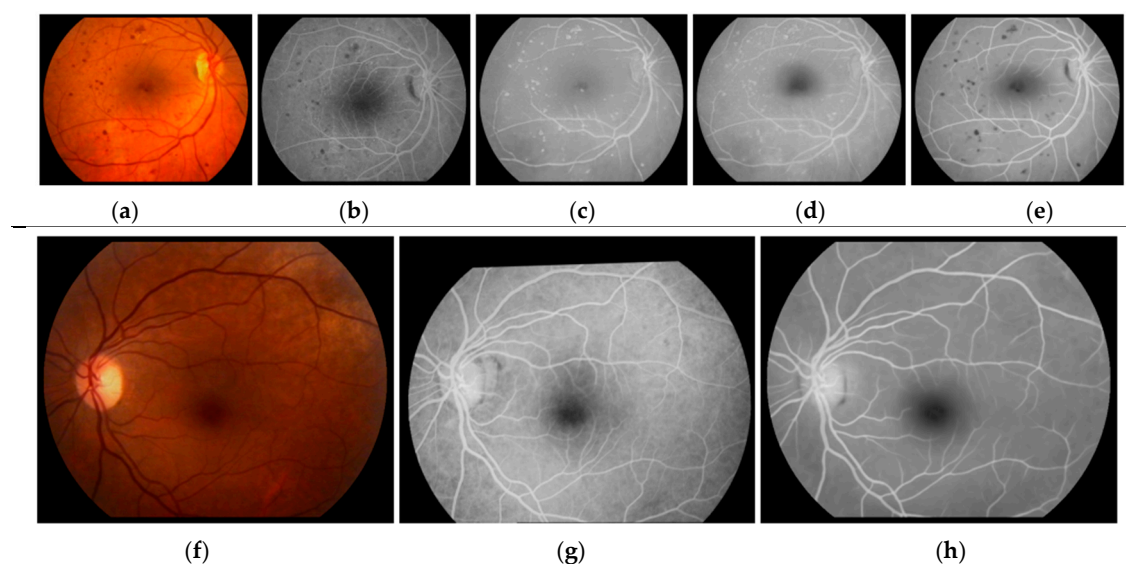


Figure 1. Examples of generated pseudo-angiographies: (a,f) original retinographies; (b) original angiography for (a); (c–e) pseudo-angiographies generated from (a) after training with L1 (c), L2 (d) and SSIM (e); (g) original angiography for (f); (h) pseudo-angiography from (f) after training with SSIM.

The results indicate that the proposed task can be used to produce a pseudo-angiography representation and learn about the retinal structures without requiring any annotated data.

Author Contributions: A.S.H., J.R. and J.N. contributed to the analysis and design of the computer methods and the experimental evaluation methods, whereas A.S.H. also developed the software and performed the experiments. M.O. contributed with domain-specific knowledge, the collection of images, and part of the registration software. All the authors performed the result analysis. A.S.H. was in charge of writing the manuscript, and all the authors participated in its critical revision and final approval.

Acknowledgments: This work is supported by I.S. Carlos III, Government of Spain, and the ERDF of the EU through the DTS15/00153 research project, and by the MINECO, Government of Spain, through the DPI2015-69948-R research project. The authors of this work also receive financial support from the ERDF and ESF of the EU, and the Xunta de Galicia through Centro Singular de Investigación de Galicia, accreditation 2016-2019, ref. ED431G/01 and Grupo de Referencia Competitiva, ED431C 2016-047 research projects, and the predoctoral grant contract ref. ED481A-2017/328.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Hervella, A.S.; Rouco, J.; Novo, J.; Ortega, M. Retinal Image Understanding Emerges from Self-Supervised Multimodal Reconstruction. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI), Granada, Spain, 16–20 September 2018.
2. Hervella, A.S.; Rouco, J.; Novo, J.; Ortega, M. Multimodal Registration of Retinal Images Using Domain-Specific Landmarks and Vessel Enhancement. In Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES), Belgrade, Serbia, 3–5 September 2018.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).