View Reviews

Paper ID

95

Paper Title

Identification and Visualization of the Underlying Independent Causes of the Diagnostic of Diabetic Retinopathy made by a Deep Learning Classifier

REVIEWER #1

REVIEW QUESTIONS

1. Please provide a summary of the paper (few lines)

An algorithm combining independent component analysis (ICA) and layer-wise relevance propagation (LRP) is proposed to visualize and explain what a convolutional neural network (CNN) has learnt. The idea is to include an ICA layer right before the fully-connected layers of the CNN and to apply LRP from that layer in order to obtain one heat map per independent component. The proposed algorithm is applied to diabetic retinopathy (DR) severity assessment.

2. Please list the major strengths of the paper (bulleted list)

* Visualization techniques for CNNs are important to facilitate acceptance by clinicians. Current systems simply indicate which pixels influence the CNN's decision (i.e. a single heat map). This paper tries to go one step further by including a factor analysis step into the process.
* The paper is applied to one classification problem of great interest for the medical image analysis community and, in case of success, it can be easily applied to any (medical) image analysis problem.

3. Please list any major weaknesses of the paper (bulleted list)

* By design, I assume the three independent components are independent. However, as seen in Fig. 5 in particular, the pixels activating these independent components are not. Therefore, the proposed visualization scheme is not really helpful. I expected that each component would be associated with one lesion or group of lesions: it does not seem to be the case. Maybe these components describe different modes of variations of the same lesions. As this is the key contribution of the paper, the authors need to explain what they think each independent component represents and to motivate the relevance of the proposed visualization.
* The introduction section assumes very little knowledge about neural networks and CNNs. I think all readers will know all these concepts very well. On the other hand, many readers may have forgotten about ICA, which they probably heard of ten years ago. In conclusion, the introduction to CNNs should be shortened, the introduction to ICA should be expanded a little bit and the remaining space should be used to expand the discussion.
* The authors propose to re-train the end of the CNN based on the extracted ICs (Fig. 1). I don't think this is useful. In fact, this is even harmful: performance decreases from 0.8 to 0.79. Why not simply using the ICA for visualization purposes?

4. Please provide detailed and constructive comments for the authors. Please also refer to our Reviewer's guide on what makes a good review: http://www.miccai2018.org/files/downloads/MICCAI2018-Reviewers.pdf

* References on visualization techniques for CNNs applied to fundus images are missing:
* Worrall DE, Wilson CM, Brostow GJ. Automated retinopathy of prematurity case detection with convolutional neural networks. Deep Learning and Data Labeling for Medical Applications. 2016:68-76.
* Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. Med Image Anal. 2017 Jul;39:178-193.
* It is not clear whether or not the proposed CNN, including the ICA layer, is trained from end to end. It should be clarified.
* "QWK" has not been defined. I assume this is the quadratic weighted Kappa, used in the Kaggle DR challenge. If so, this is not very good: in 2015, the winners of that challenge reported a QWK of 0.85, even though the training set was smaller than the one used in this paper. I assume this is because a single CNN was used, rather than an ensemble of CNNs as commonly done. The authors should comment on that limitation and maybe explain if/how their framework could be adapted to ensembles of CNNs.

* Fig. 2: what is happening for classes 1 and 3? No explanation is given anywhere.
* Fig. 4: relevance maps for ICAs 0 and 2 are highly (negatively) correlated: they don't seem to highlight independent components. The relevance map for ICA 1 is different though. It would me useful to see the input fundus image to see which lesions these relevance maps correspond to.
* Fig. 5: the input pixels activating the three ICA are more or less the same. Therefore, it seems that the ICAs do not depend on the lesion type, but maybe rather on modes of variations of the same lesions? Interpretations should be given in the paper.

5. Please rate the clarity and organisation of this paper
Satisfactory

6. Please rate whether the paper introduces significant scientific innovation or can be expected to make significant scientific impact.
Satisfactory

7. Please rate whether this work is likely to make significant clinical, or more broadly, biomedical impact.
Satisfactory

8. Please state your overall opinion of the paper.
Weak Accept

9. If you recommend acceptance, why should the MICCAI community hear about this paper?
Although the results are not very convincing, the paper addresses an important problem and provides interesting ideas for solving it.

REVIEWER #2

REVIEW QUESTIONS

1. Please provide a summary of the paper (few lines)
The authors propose the use of dimensionality reduction techniques, such as PCA or ICA, to provide visual explanations when using CNN models for diabetic retinopathy.

2. Please list the major strengths of the paper (bulleted list)
- Relevance of topic for the conference, being diabetic retinopathy one of the main causes of blindness.
- Importance of visual explanations in medical problems, specially when using deep learning models for classification tasks.

3. Please list any major weaknesses of the paper (bulleted list)
- The related work does not include any work focused on diabetic retinopathy, and some relevant issues related to visualization have not been mentioned (e.g. saliency maps).
- The technical contribution is minimal, since the idea consists in applying dimensionality reduction techniques to the deep features extracted by a CNN.
- There is no comparison between their method and the state of the art.

4. Please provide detailed and constructive comments for the authors. Please also refer to our Reviewer's guide on what makes a good review: http://www.miccai2018.org/files/downloads/MICCAI2018-Reviewers.pdf
The authors propose the use of dimensionality reduction techniques, such as PCA or ICA, to provide visual explanations when using CNN models for diabetic retinopathy. Some specific comments:
- Introduction. It explains basic concepts of deep learning (DL) models and neurons, but nothing is said about CNNs which are the basic deep learning approach when dealing with images. Additionally, I would suggest to include some references about popular CNNs instead of four references [1-4] about general deep learning.
- Related work. It should include the state-of-art methods for diabetic retinopathy, as well as other relevant works for visual explanations, such as “Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 618-626)”.
- Methods. The technical contribution of the paper is minimal, since the authors only propose to use dimensionality reduction techniques to the last layer of a CNN model, an idea already used in many other papers and problems.
- Results. The experimentation does not include enough quantitative results, and there is no comparison with other methods.

- Conclusions. The authors state that their method allows the classification of retinographies, but there is no classification performance in the experimental results.
- There are some typos to be corrected; only two examples: “a samples” instead of “a sample” (page 1), “this negative components” instead of “these negative components” (page 6).

5. Please rate the clarity and organisation of this paper

Satisfactory

6. Please rate whether the paper introduces significant scientific innovation or can be expected to make significant scientific impact.

None

7. Please rate whether this work is likely to make significant clinical, or more broadly, biomedical impact.

Poor

8. Please state your overall opinion of the paper.

Strong Reject

REVIEWER #3

REVIEW QUESTIONS

1. Please provide a summary of the paper (few lines)

In this paper, the authors tried to differentiate and identify the independent causes, which made a deep learning model that can classify different diseased eyes. They utilized a combination of Independent Component Analysis (ICA) with a score visualization technique to visualize the result.

2. Please list the major strengths of the paper (bulleted list)

1. The authors try to identify, separate, and visualize the input and hidden space of the independent components responsible of a particular DR classification decision.
2. The paper uses a suitable dataset with clear classes.

3. Please list any major weaknesses of the paper (bulleted list)

1. The paper structure is hard to follow.
2. The Related Work section did not reflect the current research in this topic.
3. The paper needs proofreading.
4. There are many missing details in the paper.

4. Please provide detailed and constructive comments for the authors. Please also refer to our Reviewer's guide on what makes a good review: http://www.miccai2018.org/files/downloads/MICCAI2018-Reviewers.pdf

1. The title of the paper is inconvenient and has wrong structure. The authors should rephrase it in a suitable way, such as “Underlying Independent Causes Identification/Visualization of Diabetic Retinopathy Diagnosis Based on Deep Learning Classifier”
2. The contributions of the paper are not addressed explicitly in the Abstract section. Then, the authors should rewrite the abstract section in a concise way to reflect the problem, their contribution(s), and the archived results.
3. In the Introduction section, the authors reported some statistical data without refereeing to the source, as in “Studies reported that 90% of the cases”. Therefore, appropriate reference citations are needed in the Introduction section.
4. The second paragraph of the Introduction section is generic, and it is well known.
5. The structure of the Related Work section is not consistent. In addition, it did not discuss the current work in this research topic.
6. The authors mentioned at the end of the Methods section that “The final classification is achieved linearly combining the low dimensional independent components layer.” This sentence is not correct if they used the multilayered artificial neural network as a classifier.
7. What are the degrees of DR that are corresponding to the classes of the fundus images? The authors should report them clearly.
8. All abbreviations should be defined first even if they are a well-known, such as CNN on page 5.
9. In Model Modifications section, what is the used linear classifier ???
10. In Fig. 2, how did ICA components contribute to classes 1 and 3?

11. The paper needs English proofreading. It has many grammar and typo errors, such as:
a. In the Abstract section, write “technique for visualizing them” instead of “technique for visualize them”
b. In the Abstract section, write “standardised classes” instead of “standarised classes”
c. In the Introduction section, write “cases could be prevented” instead of “cases can be prevented”
d. In the Introduction section, write “lesions related to this disease” instead of “lesions related with this disease”
e. In the Introduction section, write “requires acquiring expertise in daily practice” instead of “requires to acquire expertise by daily practice”
f. In the Introduction section, “Section 3 we present the methods” ????
g. ……..
12. The authors wrote many long, inconsistent sentences that are hard to follow, such as
a. In Introduction section, “These models are able to identify and extract the statistical regularities present in data that are important for optimizing a defined loss function, with the final objective of mapping a high-multidimensional input into a smaller multidimensional output.”
b. In Methods section, “We study the last layer feature space, previous to the output layer linear combination in order to identify its properties and try to isolate the independent elements that are causing a particular classification.”
13. The paper misses many commas.

5. Please rate the clarity and organisation of this paper

Poor quality

6. Please rate whether the paper introduces significant scientific innovation or can be expected to make significant scientific impact.

Poor

7. Please rate whether this work is likely to make significant clinical, or more broadly, biomedical impact.

Satisfactory

8. Please state your overall opinion of the paper.

Weak Reject