Author’s response to reviews

Title: Full Accuracy of Machine Learning for Differentiation between Optic Neuropathies and Pseudopapilledema

Authors:

Jin Mo Ahn (jmoa9203@gmail.com)

Sangsoo Kim (sskimb@gmail.com)

Kwang-Sung Ahn (kwangsung.ahn@gmail.com)

Sung-Hoon Cho (sunghoon.cho88@gmail.com)

Ungsoo Kim (ungsookim@kimeye.com)

Version: 2 Date: 07 May 2019

Author’s response to reviews:

Tobias Elze (Reviewer 1): This work investigates the accuracy of different convolutional neural networks to classify fundus images of normal eyes, optic neuropathies, and pseudopapilledema (PPE).

The manuscript is well presented and easy to understand. I have just a few minor points:

(1) In the abstract and similarly later on in the manuscript it is stated: "This study is to evaluate the accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema (PPE)." This suggests a binary classification task (neuropathies vs. PPE). Instead, there are three different classes: optic neuropathies, PPE, and normals, which is a strength of the work that should be appropriately communicated already in the abstract. I suggest rephrasing these parts.

□ We revised as your recommendation

(2) p. 5 l. 24: "ae" -> a
□ Corrected

(3) p. 5 l. 29: "was" -> were
□ Corrected

(4) p. 5 l. 54 "Carl Zeiss Meditec Inc., Dublin, CA and Heidelberg)" CZM does not have a dependency in Heidelberg. The company is based in Jena, Germany, and has a dependency in Dublin, CA. Do the authors mean Spectralis instead of Cirrus, as the latter is indeed based in Heidelberg?

We used two types of OCT in this study. Thus, the sentence is revised as “Cirrus HD-OCT, Carl Zeiss
2. Learned weights trained on Imagenet don't seem to be a good choice. While Transfer Learning is a sensible choice, could not the CNNs have been trained on eye/fundus Images for better learning instead of Imagenet? Did the authors try training on retinal images from
datasets like say STARE/DRIVE/DRISHTI?

We think there was a misunderstanding in the manuscript. Although we have used learned weights from ImageNet, our final model has weights learned from our data (fundus image). ImageNet weights is just a good starting weights instead of random weights since it has already learned many features such as edges, curves and etc. Therefore, it will give more generalized model when there is not enough data to be trained. We have clarified in the manuscript that ImageNet weights were used as initial weights to be trained. We are sorry for the confusion.

Although it would have been great if we trained the model with fundus image learned weights. However, there is currently no state-of-the-art CNN model trained on large fundus photography. Therefore, we have used the famous ImageNet trained weights as our initial weights and trained our dataset to get a tuned final weights for our model.

3. Was there no previous work done in literature on PPE Vs other optic neuropathies classification? There is no comparison with current techniques in the paper.

The comparison between this study and previous study was inserted into discussion. “The accuracy in this study was 95.89%. Chang MY. et al.22 reported that differentiation using fluorescein angiography was 97% and it showed similar performance with this study. However, the differentiation method using machine learning is much safer and easier than FAG. In addition, other modalities such as B-scan ultrasonography, fundus photography and OCT revealed high misinterpretation rate.”