Prediction of Geographic Atrophy progression by Deep Learning applied to retinal imaging

#A0136

Author(s): Guillaume Normand[1], Gwenolé Quellec[2], Ronan Danno[3], Bruno Laÿ[3], Georges Weissgerber[4], Nadia Zakaria[4], Sudeep Chandra[1]

Affiliations: [1]Clinical and Translational Imaging, BMD/TM, Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States; [2]INSERM UMR1101, Brest, France; [3]ADCIS SA, Saint-Contest, France; [4]OPH TM, Novartis Institutes for Biomedical Research, Basel, Switzerland

Introduction

Problematic

Geographic atrophy (GA) is one of the advanced forms of Age-related Macular Degeneration (AMD) and is characterized by the progressive atrophy of the retinal pigmented epithelium and photoreceptors leading to loss of vision. Progression of GA is currently manually assessed by lesion size growth rate by Fundus Autofluorescence (FAF) imaging and is highly variable between patients (**Fig 1**).
The aim of this study was to predict the lesion growth rate at month 12 from baseline images using a Deep Learning approach through Convolutional Neural Network (CNNs).

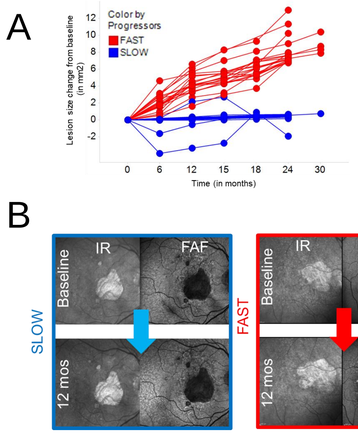

Figure 1 A. Variability of lesion size growth rate between patients B. Example of slow and fast growth in 12 months

Material and Methods

Datasets

Data were pooled from several GA studies conducted by Novartis/Alcon (GATE, GAP, PJMR0092103, CLFG316A2003) representing about 236,822 total images over a period of 1-4 years. Infrared Reflectance (IR) and Fundus Autofluorescence (FAF) images with corresponding lesion size measurements and with a follow-up at 6, 12 or 18 months were selected. 2,708 eyes with IR and 2,204 eyes with FAF follow-up were then split as follows: 80% for CNN training, 10% for the fusion training and 10% for testing. Both eyes from same patient were kept in the same set.

Approach

After image pre-processing (trimming, resizing), several pre-trained CNNs, namely VGG-16/19[2], Inception-v4[3], NASNet[4] and ResNet-101/152[5] were fine-tuned on the IR or FAF datasets. Performance of the CNNs were then evaluated using Pearson's correlation coefficient. We then conducted a late fusion training, which consisted of a Multilayer Perceptron based on features from the prediction of 4 CNNs using FAF, the prediction of 1 CNN using IR as well as the age variable (**Fig 2**).

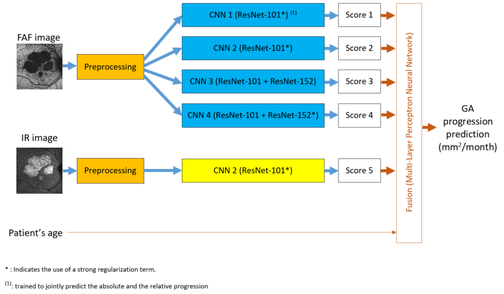

Figure 2 Workflow of the late fusion approach

Activation heatmaps

One of the challenges for CNNs' adoption is to understand and validate the learning process, which leads to the prediction ('black boxes'). We thus implemented a modified sensitivity analysis approach, which was previously used for diagnosis of referable diabetic retinopathy[6], in order to visualize the regions, at the pixel-level, which play a role in the prediction of GA growth.

Results

Performance of prediction

The late fusion approach, which took the best CNN results in input, yielded an overall Pearson correlation coefficient of 0.59 (**Fig 3A**).
The test dataset was then separated into slow and fast progressors based on the average growth rate for all the pooled trials (0.13mm²/month) to work on a binary classification problem.

The late fusion thus yielded an Area Under the Curve (AUC) of 0.8174 (**Fig 3A**) and with an accuracy of 76.7% and a positive/negative predictive rate of 65.9% and 84.9%, respectively (**Fig 3B**).

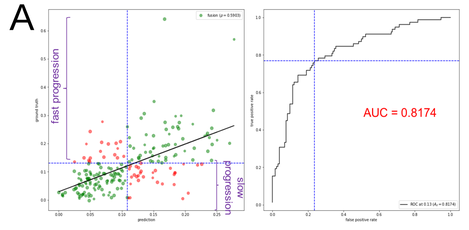

Figure 3 A. Correlation plots of ground truth and predicted growth rate with 0.13mm²/month threshold and AUC B. Confusion matrix

Visualization of the data

The activation maps showed that the prediction is mainly based on the lesion itself (**Fig 4A**). Evaluation of the t-distributed stochastic neighbor embedding (t-SNE) map on the training set suggests that the shape complexity positively may be the most important risk factor for growth rate (**Fig 4B**).

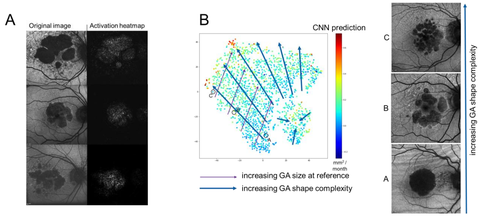

Figure 4 A. Examples of activation maps. B. t-SNE map with representative images from slow to fast progression

Conclusions

We showed here for the first time that baseline retinal images indeed contain predictive information about the GA lesion growth rate at follow-up visits. The high negative predictive value indicate the possibility of screening out slow progressors while modest positive predictive value suggest that additional parameters may be needed to improve the prediction of fast progressors. Furthermore, the visualization features enabled the validation of the algorithm and provided new insights into the natural GA progression.

References

1. Fleckenstein M et al. *Fundus autofluorescence and spectral-domain optical coherence tomography chraracteristics in a rapidly progression form of geographic atrophy* (2011) 52: 3761-66
2. Simonyan K et al. *Very deep convolutional networks for large-scale image recognition*. Proc ICLR (2015)
3. Szegedy C et al. *Rethinking the inception architecture for computer vision*. In: Proc IEEE CVPR (2016) 2818-26.
4. Zoph B et al. *Learning transferable architectures for scalable image recognition*. arXiv (2018) 1707.07012
5. He K et al. *Deep residual learning for image recognition*. Proc CVPR. (2016) 770-8.
6. Quellec et al. Deep image mining for diabetic retinopathy screening. Med Imag Analysis (2017) 39: 178-193

Disclosure: GN, GW, NZ and SC are Employees of Novartis; RD and BL are Employees of ADCIS; GQ is Consultant for ADCIS