Title: Explainable Deep Neural Diagnostics: Vision Transformer-Based Retinal Disease Classification from OCT Images Using Gradient-Driven Visual Attribution

Author names: Md. Sakibul Hassan Omi, K. M. Arafat Islam, Shovan Samanta Turzo, Pranab Kumar Dhar & Md. Hasan Imam Bijoy

Abstract: The research introduces the Convolutional Neural Network (CNN) model that automatizes the process of classifying retinal diseases based on the Optical Coherence Tomography (OCT) images. This model will categorize four retinal disorders namely: Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), Drusen Macular Degeneration (DMD) and normal retinal conditions. In order to make the predictions of the CNN more interpretable, we will use Grad-CAM (Gradient-weighted Class Activation Mapping) Explainable AI (XAI) in the state of the art. Grad-CAM also provides a visual representation of what parts of the image influence the most the classification decision of the CNN this way clinicians can get some kind of insight as to why the model is producing its outputs. That transparency can be especially useful in medical settings, where the rationale behind a decision can be of the great importance in a clinical verification. Hyperparameters are also optimized by using random search to tune on important variables including batch size and dropout rate and the optimizer. In the suggested CNN model, the classification accuracy reaches 96.01. As compared to others, the proposed method is more accurate and more interpretable than the other methods that are available and validates its utility in large-scale screening and decision support in diagnosing retinal diseases. This model can be a valuable addition to the toolkit of the ophthalmologist because it conveys high diagnostic performance and interpretability of the main datasets used during its production.

http://link.springer.com/chapter/10.1007/978-3-032-15346-3_17