



Application number/Title: 41044: Classification and automatic detection of OCT layer segmentation failure modes

Applicant PI: Dr Ryan Amelon

Application Institution: IDX, LLC

Keywords provided by the Applicant PI to describe the research project:

Convolutional-Neural-Network, Deep-Learning, Failure-Modes, Image-Quality, Layer-Segmentation, OCT

Application Lay Summary:

Automated AI diagnostic solutions are expected to properly function and yield accurate results all of the time, but no solution will ever have perfect accuracy. It is important to consider failure modes of automated diagnostics components, both when they fail and how. The detection and proper characterization of these failure modes can make automated diagnostics safer and more efficient.

Glaucoma in the retina is characterized, in part, by a loss of retinal nerve fiber layer thickness. Optical coherence tomography (OCT) imaging results in a three-dimensional view of the retina which allows for measurement of retinal nerve fiber layer thickness. The usual first step in detecting Glaucoma using OCT is segmentation of the layers of the retina. Typical assessment of these algorithms consider the accuracy of the segmentation, which is important when considering algorithm utility. However, most assessments ignore, at a patient level, how often the layer segmentations fail and what is causing the failure. For automated algorithms to become common-place, it must be able to both perform the task it was trained for as well as determine when the task cannot be adequately performed. Further, if an algorithm can determine why it failed, a patient visit may be correctly salvaged by specifying the necessary follow-up steps required for an accurate result.

This work will shed light on the mechanisms of layer segmentation failures that developers should be aware of, how common such failures may occur, and the ability to which they may be automatically detected. This study will provide a proof-of-concept failure mode assessment that will increase device safety and efficacy by correctly identifying algorithm failures and suggest follow-up measures.