

Diagnosing glaucoma from frequency doubling technology perimetry using supervised machine learning classifiers

J.P. Pascual¹, Z. Zhang^{2,1}, A.J. Hughes¹, J. Hao^{2,1}, T.-W. Lee², T. Sejnowski², M.H. Goldbaum³, R.N. Weinreb¹ and P.A. Sample¹

¹ Hamilton Glaucoma Center, University of California, San Diego, La Jolla, CA

² Institute for Neural Computation, University of California, San Diego, La Jolla, CA

³ Ophthalmic Informatics Laboratory, University of California, San Diego, La Jolla, CA

Commercial Relationships: J.P. Pascual, None; Z. Zhang, None; A.J. Hughes, None; J. Hao, None; T. Lee, None; T. Sejnowski, None; M.H. Goldbaum, None; R.N. Weinreb, Carl-Zeiss Meditec, Inc. F; P.A. Sample, Welch Allyn F; Carl-Zeiss Meditec, Inc. F.

Grant Identification: NEI Grant EY08208 (PAS), NIH EY013928(MHG), NIH Grant EY13235 (RNW)

Abstract

Purpose: To determine if machine learning classifiers can learn to interpret frequency doubling technology (FDT) perimetry and to compare their performance with the global indices for both FDT II Matrix pattern 24–2 and FDT pattern N30.

Methods: Support vector machine with Gaussian kernels (SVMg) and mixture of Gaussian (MoG) classifiers were trained and tested using cross validation on the numerical values of absolute sensitivity plus age from one FDT 24–2 test each from 104 normal and 81 eyes with glaucomatous optic neuropathy (GON) or one FDT N30 test each from 201 normal and 293 eyes with GON. GON was determined by masked consensus reviews of simultaneous stereophotos. For $\hat{\alpha} = 0.05$, a Bonferroni adjustment for 6 comparisons ($P \leq 0.008$) was used to compare the areas under the ROC curves for the machine classifiers and global indices.

Results: For FDT 24–2, the areas under the ROC curves for SVMg, MoG, MD, and PSD were $0.840 \hat{\alpha} \pm 0.029$, $0.785 \hat{\alpha} \pm 0.033$, $0.703 \hat{\alpha} \pm 0.039$, and $0.712 \hat{\alpha} \pm 0.038$ (mean $\hat{\alpha} \pm$ SE), respectively. At a specificity of 0.90, SVMg, MoG, MD, and PSD yielded sensitivities of 0.580, 0.457, 0.321, and 0.296. Based on presence of GON vs. healthy, SVMg had a specificity and sensitivity of 82% and 72%, respectively; MoG of 89% and 46%.

For FDT N30, the areas under the ROC curves for SVMg, MoG, MD, and PSD were $0.802 \hat{\alpha} \pm 0.019$, $0.795 \hat{\alpha} \pm 0.020$, $0.727 \hat{\alpha} \pm 0.023$, and $0.686 \hat{\alpha} \pm 0.024$. At a specificity of 0.90, SVMg, MoG, MD, and PSD yielded sensitivities of 0.529, 0.532, 0.386, and 0.334. SVMg had a specificity and sensitivity of 78% and 68%, respectively; MoG of 67% and 76%.

For both FDT 24–2 and N30, the areas under the ROC curves were greater in both SVMg and MoG than in MD or PSD; SVMg and MoG were not significantly different.

Conclusions: Supervised machine learning classifiers are capable of discriminating between normal and GON eyes using FDT results as well as, if not better than, the tests' global indices. This is consistent with the superior performance of SVMg and MoG for a variety of visual function (SAP, SWAP) and imaging (HRT) data.