

# A Deep Learning Approach to Accurately Discriminate Between Optic Disc Drusen and Papilledema on Fundus Photographs

Kanchalika Sathianvichitr, MD, FICO<sup>1</sup>; Raymond P. Najjar, PhD<sup>1,2,3</sup>; Tang Zhiqun, PhD<sup>1</sup>; J. Alexander Fraser, MD<sup>4</sup>; Christine Wen Leng Yau, MBBS(S'pore), MMed(Ophth)<sup>5</sup>; Michael Julien Alexandre Girard, PhD<sup>2,6,7</sup>; Fiona Costello, MD<sup>8</sup>; Mung Yan Lin, MD<sup>9,10</sup>; Wolf Alexander Lagrèze, MD<sup>11</sup>; Catherine Vignal-Clermont, MD<sup>12</sup>; Clare L. Fraser, MBBS, FRANZCO<sup>13</sup>; Steffen Hamann, MD, PhD<sup>14</sup>; Nancy J. Newman, MD<sup>10,15,16</sup>; Valérie Biousse, MD<sup>10,15</sup>; Dan Milea, MD, PhD<sup>1,2,5,14</sup>; for the BONSAI Group

## Affiliation Information for All Authors:

1. Visual Neuroscience Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
2. Duke-NUS Medical School, National University of Singapore, Singapore
3. Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
4. Departments of Clinical Neurological Sciences and Ophthalmology, Western University, London, Ontario, Canada
5. Neuro-Ophthalmology Department, Singapore National Eye Centre, Singapore
6. Ophthalmic Engineering & Innovation Laboratory, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
7. Institute for Molecular and Clinical Ophthalmology, Basel, Switzerland
8. Departments of Clinical Neurosciences and Surgery, University of Calgary, Alberta, Canada
9. Department of Medicine, Emory University School of Medicine, Atlanta, Georgia, USA
10. Department of Ophthalmology, Emory Eye Center, Emory University School of Medicine, Atlanta, Georgia, USA
11. Eye Center, Medical Center-University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
12. Hopital Fondation Adolphe de Rothschild, Paris, France
13. Save Sight Institute, Faculty of Health and Medicine, The University of Sydney, New South Wales, Australia
14. Department of Ophthalmology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark
15. Department of Neurology, Emory University School of Medicine, Atlanta, Georgia, USA
16. Department of Neurological Surgery, Emory University School of Medicine, Atlanta, Georgia, USA

## Word count:

Title (characters with space): 115 characters

Abstract: 311 words

Manuscript: 3198 words

Number of references: 49

Number of tables: 5

Number of figures: 4

**Corresponding author:** Dan Milea, MD, PhD  
Singapore Eye Research Institute.  
20 College Road Discovery Tower, Level 6. The Academia, 169856  
Tel: +65 63228313. Fax: +65 62263395. Email: [dan.milea@singhealth.com.sg](mailto:dan.milea@singhealth.com.sg)

**Funding:** Supported by the Singapore National Medical Research Council (Clinician Scientist Individual Research grant CIRG18Nov-0013), the Duke-NUS Medical School, Ophthalmology and Visual Sciences Academic Clinical Program grant (05/FY2019/P2/06-A60).

**Conflict of interest:** no conflicts of interest related to this work.

1 **Abstract**

2 **Objective:** To assess the performance of a deep learning system (DLS) to discriminate between  
3 optic disc drusen (ODD) and papilledema caused by intracranial hypertension, using standard  
4 color ocular fundus photographs collected in a large international multi-ethnic population.

5 **Design:** Retrospective study.

6 **Participants:** The study included 4,508 color fundus images in 2,180 patients from 30 neuro-  
7 ophthalmology centers (19 countries) participating in the Brain and Optic Nerve Study with  
8 Artificial Intelligence (BONSAI) Group.

9 **Methods:** We trained, validated, and tested a dedicated DLS for binary classification of ODD  
10 vs. papilledema (including various subgroups within each category), on conventional mydriatic  
11 digital ocular fundus photographs. For training and internal validation, we used 857 ODD  
12 images and 3,230 papilledema images, in 1,959 patients. External-testing was subsequently  
13 performed on an independent dataset (221 patients) including 207 images with ODD (96 visible  
14 and 111 buried), provided by 3 centers of the Optic Disc Drusen Studies Consortium, and 214  
15 images of papilledema (92 mild-to-moderate and 122 severe) from a previously validated  
16 study.

17 **Main outcome measures:** Area under the receiver operating characteristic curve (AUC),  
18 accuracy, sensitivity, and specificity were used to discriminate between ODD and papilledema.

19 **Results:** Overall, the DLS could accurately distinguish between all ODD and papilledema (all  
20 severities included): AUC 0.97 (95% confidence interval [CI], 0.96 to 0.98), accuracy 90.5%  
21 (95% CI, 88.0% to 92.9%), sensitivity 86.0% (95% CI, 82.1% to 90.1%), and specificity 94.9%  
22 (95% CI, 92.3% to 97.6%). The performance of the DLS remained high for discrimination of  
23 buried ODD from mild-to-moderate papilledema: AUC 0.93 (95% CI, 0.90 to 0.96), accuracy

24 84.2% (95% CI, 80.2%-88.6%), sensitivity 78.4% (95% CI, 72.2% to 84.7%), and specificity  
25 91.3% (95% CI, 87.0% to 96.4%).

26 **Conclusions:** A dedicated DLS can accurately distinguish between ODD and papilledema  
27 caused by elevated intracranial pressure, even when considering buried ODD vs mild-to-  
28 moderate papilledema. Future studies are required to validate the utility of this DLS in clinical  
29 practice.

## 30 **Introduction**

31           Papilledema (optic disc swelling from intracranial hypertension) is an important ocular  
32 fundus finding, visible on funduscopy examination or standard retinal photographic images.  
33 Its clinical identification is not always easy, especially when attempted by non-ophthalmic  
34 healthcare providers (neurologists, general practitioners, etc.).<sup>1</sup> Detection of papilledema  
35 prompts urgent critical investigations (brain imaging, cerebrospinal fluid examination, etc.) in  
36 order to identify potentially life- or vision-threatening conditions; failure to detect papilledema  
37 in a timely manner can lead to visual loss and irreversible neurological sequelae.

38           Other optic nerve head abnormalities can mimic papilledema (i.e., pseudopapilledema),  
39 leading sometimes to unnecessary, invasive and costly investigations.<sup>2,3</sup> Optic disc drusen  
40 (ODD) represent a common cause of misdiagnosed papilledema, especially if the ODD are not  
41 visible on funduscopy examination, so-called “buried” ODD. Accurate diagnosis of ODD,  
42 which are present in 1.0-2.0% of the general population,<sup>4-6</sup> is important in clinical practice, not  
43 only to discriminate ODD from papilledema, but also because of the possible visual  
44 complications from ODD.<sup>7</sup> Diagnosing ODD is clinically straightforward in the presence of  
45 visible, calcified lesions. However, discriminating ODD from true papilledema can be  
46 challenging, especially when the optic disc elevation is moderate. Several ocular imaging  
47 modalities are helpful in this context, including B-scan orbital ultrasonography,<sup>8,9</sup> fundus  
48 autofluorescence, and more recently, spectral domain enhanced depth imaging optical  
49 coherence tomography (OCT).<sup>10-12</sup> Indeed, OCT has become a valued investigation for ODD  
50 diagnosis, although it requires expertise to facilitate image interpretation. Recently, the  
51 application of deep learning (DL) methods to OCT imaging has improved the performance of  
52 OCT in distinguishing ODD from papilledema, an approach which requires further  
53 validation.<sup>13</sup>

54 Standard color photography of the ocular fundus is an easily accessible modality that  
55 can be implemented in non-ophthalmic environments,<sup>14,15</sup> and is potentially useful for  
56 discriminating ODD and papilledema.<sup>16,17</sup> Fundus photography has recently regained new  
57 interest for detection of optic nerve disease, with some successful applications of DL for  
58 accurate diagnosis.<sup>18</sup> A large international study group (Brain and Optic Nerve Study with  
59 Artificial Intelligence, BONSAI) showed that a DLS could detect papilledema and determine  
60 its severity on color fundus images, with a performance which was comparable to that of expert  
61 neuro-ophthalmologists.<sup>19-21</sup> The main aim of the initial BONSAI study was to distinguish  
62 papilledema from normal discs, and from a group of “other” optic disc abnormalities, including  
63 a mix of ODD, optic atrophy, and anterior ischemic optic neuropathies. Although this  
64 classification can be very helpful as the initial approach in a patient with optic disc  
65 abnormalities, there is an additional clinical need to discriminate accurately between  
66 pseudopapilledema (buried ODD being its main cause) and true papilledema, because the  
67 vision and life-threatening implications for the patients may be radically different.

68 The aim of our current study was to develop, train and test a new dedicated DLS, able  
69 to specifically discriminate between ODD and true papilledema on standard color ocular  
70 fundus images. In addition, we evaluated the classification performance of the DLS with  
71 various clinically relevant subclasses of papilledema severity and ODD visibility.

72

## 73 **Methods**

### 74 Standard Protocol Approvals, Registrations, and Patient Consents

75 The study was approved by the Centralized Institutional Review Board of SingHealth,  
76 Singapore as well as by ethical committees of each contributing institution, and was conducted

77 in accordance with the Declaration of Helsinki. Informed consent was exempted, given the  
78 retrospective nature of the study and the use of deidentified medical information and ocular  
79 fundus images.

80

## 81 Study Population

82 The primary training and internal validation dataset consisted of 4,087 ocular fundus  
83 images, among which there were 857 images of optic discs from patients with confirmed ODD  
84 and 3,230 images of optic discs with confirmed papilledema, retrospectively collected from  
85 1,959 patients from 30 participating centers of BONSAI (Table 1).

86 The external testing dataset consisted of 221 patients and 421 images (207 with ODD  
87 and 214 with papilledema). The ODD images were provided by three independent expert  
88 centers participating in the Optic Disc Drusen Studies (ODDS) Consortium<sup>10,22</sup>: 1/ University  
89 of Calgary (Alberta, Canada), 2/ Rigshospitalet, University of Copenhagen (Denmark) and 3/  
90 Western University (Ontario, Canada). Images with papilledema were selected at random from  
91 four participating centers: 1/ Ramathibodi Hospital, Mahidol University (Thailand), 2/  
92 University of Freiburg (Germany), 3/ Farabi Eye Hospital, Tehran University of Medical  
93 Sciences (Iran), and 4/ Angers University Hospital (France), as previously described  
94 elsewhere.<sup>21</sup> All the images included in the external-testing dataset were independent from the  
95 images used in the training dataset.

96 We excluded patients with concurrent ophthalmic pathologies (e.g., ODD co-existing  
97 with optic disc swelling or ODD co-existing with other optic nerve or retinal disease). We also  
98 excluded images with poor quality or with poorly centered optic discs. Thus, among the 4,574  
99 initially available images, we excluded 66 (1.4%) images (Fig 1).

## 100 Image Acquisition

101 Fundus images and relevant clinical information were retrospectively collected in a mix  
102 of consecutive and convenience samples in multiple international neuro-ophthalmology  
103 centers. Fundus images were obtained in eyes with pharmacologically dilated pupils, using  
104 various commercial desktop digital fundus cameras (Table S1, Supplemental Material) by  
105 neuro-ophthalmologists who routinely obtain fundus images and who have access to the  
106 patients' medical records. The fundus images were centered on either the macula or the optic  
107 disc at various fields of view (subtending 20° to 45°). Fundus images obtained by wide-field  
108 fundus camera or with multicolor imaging mode were excluded from this study. Deidentified  
109 unaltered fundus images were transferred to the Singapore Eye Research Institute for further  
110 analysis.

111

## 112 Classification of ODD

113 Patients with ODD were enrolled by the expert neuro-ophthalmology providers only if  
114 the diagnosis was confirmed by at least one of the following imaging modalities<sup>10</sup>: 1/ spectral  
115 domain enhanced depth imaging optical coherence tomography (EDI-OCT) of the optic nerve  
116 head, in agreement with the ODDS Consortium acquisition protocol<sup>22</sup>; 2/ swept-source OCT  
117 (SS-OCT) of the optic nerve head; 3/ other ODD diagnostic methods, including B-scan orbital  
118 ultrasonography, and fundus autofluorescence.

119 After diagnosis confirmation, ODD were further classified into “visible ODD” and  
120 “buried ODD.” Visible ODD were defined by the presence of visually identifiable refractile  
121 bodies on the optic disc surface; the remaining confirmed ODD by ancillary investigations  
122 were classified as “buried.”<sup>23-25</sup> In the primary training dataset, the classification of “visible”



123 vs “buried” ODD was made post-hoc, individually, by two neuro-ophthalmologists at the  
124 Singapore Eye Research Institute, on the 857 available color fundus images, using a dedicated  
125 semi-automated image presentation software.<sup>26</sup> A senior neuro-ophthalmologist adjudicated  
126 the 41(4.8%) images with discordant diagnosis to obtain consensus.

127 In the external-testing dataset, the classification “visible” vs “buried” was provided  
128 directly by the local neuro-ophthalmology experts from the three centers participating in the  
129 ODDS Consortium, based on a combination of the appearance of the ODD on the color fundus  
130 images and the results of ancillary imaging investigations (most often, EDI-OCT).

131

### 132 Classification of papilledema severity

133 Papilledema images were obtained in patients with confirmed intracranial hypertension  
134 due to 1/ a known secondary cause (i.e. intracranial mass, hydrocephalus, cerebral venous  
135 thrombosis, medication, etc.) or 2/ idiopathic intracranial hypertension, according to the  
136 modified Dandy criteria and abnormally elevated cerebrospinal fluid opening pressure on  
137 lumbar puncture.<sup>27</sup> These images were further classified according to the papilledema severity  
138 by two expert neuro-ophthalmologists into two classes, as previously described.<sup>21</sup> In brief, the  
139 standard Frisén 5-grade severity classification was simplified into a 2-grade classification: (1)  
140 mild-to-moderate papilledema, corresponding to Frisén grades 1-3, and (2) severe papilledema,  
141 corresponding to Frisén grades 4 and 5. The images with discordant severity were adjudicated  
142 by 2 additional neuro-ophthalmologists and a consensus was obtained for all images.

143

144

## 145 Development of the Deep-Learning Classification Model

### 146 Data splitting

147           The primary dataset was randomly split, according to common practices for deep  
148 learning model development, into training/validation (80%) and internal-testing datasets  
149 (20%). A stratified split was also applied to reduce technical bias among centers. Among the  
150 4,087 images used for this purpose, 3,357 images (82.1%) were used for training/validation,  
151 and 730 images (17.9%) were used for the internal evaluation of the model. A stratified split  
152 was also used to divide the training/validation dataset. Eventually, 2,856 of 3,357 images  
153 (85.1%) were used for training the model and 501 (14.9%) for validation or model parameter  
154 tuning.

### 155 Image Segmentation

156           The segmentation network was based on U-Net architecture with a ResNet-34  
157 encoder,<sup>28</sup> using an image resolution of 256x256 pixels, to identify the optic disc and  
158 peripapillary region (region of interest, ROI). The automatically segmented ROIs were then  
159 resized to 456x456 pixels as the input images for the classification network.

### 160 Image Preprocessing and Sampling

161           After image segmentation, we used data augmentation techniques including random  
162 rotation, horizontal, warp, zoom, translation shifts, and random drop out of certain input  
163 regions.<sup>29</sup> Oversampling was applied to the minority class (ODD images).

### 164 Model Training

165           We used a classification network (EfficientNet-B5), initialized using weights pretrained  
166 on ImageNet<sup>30</sup> and fine-tuned in an end-to-end manner to achieve the best performance. The

167 EfficientNet-B5 architecture was chosen due to its high efficiency for image classification,  
168 while using fewer resources and computing power.<sup>31</sup> The systematically compound scale-up  
169 algorithm composed of mobile inverted bottleneck convolution (MBCConv) blocks (Fig 2) has  
170 improved performance compared to other DLS.<sup>19</sup> Label smoothing was used to minimize the  
171 cross-entropy loss function. The classification network was trained and tuned on 3,357 fundus  
172 images (training and validation dataset) to automatically classify the ROIs into specific classes  
173 of discs with ODD or discs with papilledema.

#### 174 Model Selection, performance evaluation and heatmaps generation

175 After selecting the best predictive model (based on preliminary evaluation of the  
176 internal-testing dataset), we evaluated its performance on 421 fundus images from the  
177 independent external-testing dataset. We first assessed the model's performance for  
178 discriminating between all ODD images (with both visible and buried ODD) and all  
179 papilledema images (irrespective of their severity). Then, we evaluated more specifically the  
180 performance of the same DLS for tasks with clinically increasing difficulty: 1/ discrimination  
181 between visible ODD and severe papilledema, 2/ discrimination between visible ODD and  
182 mild-to-moderate papilledema, 3/discrimination between buried ODD and severe papilledema,  
183 4/ discrimination between buried ODD and mild-to-moderate papilledema.

184 Finally, we generated heatmaps using the class-specific gradient information (Gradient-  
185 weighted Class Activation Mapping, Grad-CAM) extracted from the final convolutional layer  
186 of the CNN model to visualize the relative relevance of pixels in the input images for the  
187 classification task.<sup>32</sup>

188

189

## 190 Statistical Analysis

191 To evaluate the performance of the DLS to distinguish between discs with ODD and  
192 discs with papilledema (and their respective subclasses), we calculated various metrics,  
193 including the area under the receiver operating characteristic curve (AUC), sensitivity,  
194 specificity, and accuracy. Bootstrapping (2,000 times) with patient as the sampling unit was  
195 used to estimate 95% confidence intervals (CI) of the performance metrics.

196

## 197 **Results**

### 198 **Patient and Image Characteristics**

199 For training and internal validation purposes, the study included 4,087 images,  
200 collected from 1959 patients (91.2% with images of both eyes, 9.9% with images on follow-  
201 up visits) recruited by 30 sites participating in the BONSAI consortium for developing the  
202 model. We included 857 ODD images and 3,230 papilledema images. The external-testing  
203 dataset included images from seven international sites. In total, 421 fundus images from 221  
204 patients (90.5% with images of both eyes, none of the images duplicative of the same eye) were  
205 enrolled. Among them, 207 ODD images (111 buried ODD and 96 visible ODD, Table 2) were  
206 collected from three sites participating in the ODDS Consortium; and 214 papilledema images  
207 (92 images of discs with mild-to-moderate papilledema and 122 images of discs with severe  
208 papilledema), which were previously validated as described above, were included.

209 Patient demographics and image characteristics are described in Table 2. Altogether,  
210 the ratio between visible/buried drusen varied according to the age of patients, with higher  
211 proportions of buried drusen at lower ages. In the training/validation group, the frequency of

212 buried ODD was 33.3% in patients over 40, reaching 96.9% in patients under 11; in the  
213 external-testing dataset, the frequency of buried ODD in those age groups was 44.3% and  
214 83.3%, respectively (Table 3).

215

## 216 **Overall Classification Performance**

217 In the validation, internal-testing, and external-testing datasets, the DLS was able to  
218 discriminate discs with confirmed ODD from discs with papilledema with AUCs of 0.99 (95%  
219 CI, 0.98 to 1.00), 0.98 (95% CI, 0.97 to 0.99), and 0.97 (95% CI, 0.96 to 0.98) (Table 4). In the  
220 internal-testing dataset, the DLS achieved a sensitivity of 91.9% (95% CI, 90.1% to 94.3%)  
221 and specificity of 93.9% (95% CI, 90.1% to 97.3%) for an overall accuracy of 93.2% (95% CI,  
222 91.1% to 94.7%). In the external-testing dataset, the DLS had an overall sensitivity of 86.0%  
223 (95% CI, 82.1% to 90.1%), a specificity of 94.9% (95% CI, 92.3% to 97.6%), and an accuracy  
224 of 90.5% (95% CI, 88.0% to 92.9%). Examples of representative heatmaps allowing for the  
225 visualization of pixels with the most discriminative value for classification tasks by the DLS  
226 are shown in Fig 3.

227

## 228 **Subclass performance (buried/visible ODD versus mild-moderate/severe papilledema)**

229 Next, we assessed the performance of the DLS in the external-testing datasets,  
230 specifically for the two subgroups of ODD (visible and buried) and the two classes of  
231 papilledema severity (Table 5). The DLS performance followed a gradient, with the highest  
232 parameters for distinguishing visible ODD and severe papilledema (AUC 0.99 [95% CI, 0.98  
233 to 1.00], accuracy 96.3% [95% CI, 94.4% to 98.6%]), followed by marginally decreased

234 performance for distinguishing visible ODD from mild-to-moderate papilledema (AUC 0.96  
235 [95% CI, 0.94 to 0.98], accuracy 93.1% [95% CI, 90.2% to 96.1%]). The performance of  
236 discriminating buried ODD from papilledema at all stages showed only slightly reduced  
237 performance (AUC 0.96 [95% CI, 0.95 to 0.98], accuracy 89.2% [95% CI, 86.4% to 92.2%]).  
238 Unsurprisingly, the performance of the DLS was excellent in discriminating buried ODD from  
239 severe papilledema (AUC 0.99 [95% CI, 0.98 to 1.00], accuracy 88.4% [95% CI, 85.0% to  
240 92.1%]). The lowest performance of the DLS was noted for distinguishing buried ODD from  
241 mild-to-moderate papilledema (AUC 0.93 [95% CI, 0.90 to 0.96], accuracy 84.2% [95% CI,  
242 80.2% to 88.6%]) (Fig 4).

243

#### 244 **Post-hoc analysis of classification errors**

245 Forty retinal fundus images (9.5% of the 421 images) in the external-testing dataset  
246 were wrongly diagnosed by the DLS compared to the ground truth. More specifically, the DLS  
247 misclassified 11 papilledema images as ODD, in 5.1% of the total of 214 papilledema images.  
248 The majority of these misclassified images represented mild papilledema (n=8, 72.7%). Among  
249 the 207 images of confirmed ODD, 29 images (14.0%) were wrongly diagnosed as  
250 papilledema. The large majority of the misdiagnosed ODD were buried ODD (24 images,  
251 82.8%).

252

253

254

255

## 256 **Discussion**

257           The main finding of this study is that a trained DLS achieved excellent performance at  
258 discriminating ODD from true papilledema on standard color ocular fundus images, without  
259 any additional clinical information. The performance of the DLS across various subgroups of  
260 ODD and papilledema images followed a decreasing trend which was compatible with the  
261 reality of everyday clinical practice. Indeed, just as clinicians do in real life, the DLS displayed  
262 a higher performance in discriminating visible ODD from severe papilledema. Conversely, the  
263 performance of the DLS dropped (but only marginally) in the most difficult clinical situation  
264 of mild elevation of the optic disc (typically seen in buried ODD and mild-to-moderate  
265 papilledema).

266           Accurate distinction between ODD and papilledema is crucial to avoid unnecessary,  
267 invasive, and costly procedures or, conversely, to detect vision- or life-threatening conditions.  
268 Among several ancillary investigations aiming to discriminate between ODD and papilledema,  
269 fluorescein angiography has been reported to have high accuracy and good interobserver  
270 agreement,<sup>12,33</sup> but it remains an invasive procedure, reducing its applicability outside  
271 ophthalmic departments. Fundus autofluorescence is a cost-effective, non-invasive method  
272 requiring little subjective interpretation for detecting superficial ODD. However, it doesn't  
273 reliably detect ODD located in the deeper portions of the optic nerve head.<sup>34</sup> Ultrasonography  
274 was until recently considered the most accurate investigation for diagnosing ODD, and, more  
275 generally, for distinguishing true papilledema from pseudopapilledema.<sup>35</sup> However, it is less  
276 able to detect poorly calcified ODD or buried ODD,<sup>36</sup> and is also a highly operator-dependent  
277 procedure, difficult to implement outside dedicated clinics.

278 OCT is commonly used for ODD diagnosis and is considered as gold standard by the  
279 international ODDS Consortium.<sup>22</sup> The advent of EDI-OCT has facilitated visualization of  
280 deep structures in the optic nerve head, including ODD that are deeply located underneath  
281 Bruch's membrane opening. In discs with papilledema, transverse axial and en face OCT reveal  
282 structural changes, including anterior displacement of Bruch's membrane opening,  
283 peripapillary folds or wrinkles, and peripapillary fluid accumulation.<sup>37-39</sup> Quantitative OCT  
284 measures of the peripapillary RNFL thickness also may be helpful in the evaluation and  
285 monitoring of the optic disc swelling of papilledema.<sup>40-43</sup> However, there are several limitations  
286 in clinical use, including the considerable overlap of RNFL thickness values among normal  
287 optic discs, discs with deep ODD or mild papilledema,<sup>44,45</sup> absence of normative RNFL  
288 thickness values in children and in ocular conditions such as high myopia, falsely normative  
289 values of RNFL thickness when there is concurrent optic disc edema and optic atrophy, and  
290 false positively thickened RNFL from other disc abnormalities (e.g., gliosis, hyperopia).<sup>46</sup>

291 In 2020, the BONSAI Study Group reported that an artificial intelligence-based DLS  
292 can be successfully applied to standard ocular fundus color photograph for the detection of  
293 papilledema with an excellent AUC of 0.96, a sensitivity of 96.4%, and a specificity of 84.7%.<sup>19</sup>  
294 However, that DLS did not address specifically the clinically relevant discrimination between  
295 entities that can mimic each other such as true papilledema and ODD. For this purpose, we  
296 subsequently developed a next-generation DLS, after inclusion of further robust imaging and  
297 clinical data provided by international expert neuro-ophthalmologists. The model was  
298 developed using a recent advanced deep convolutional neural network, EfficientNet-B5, which  
299 already has been successfully applied in the fields of radiology<sup>47</sup> and, more recently, in  
300 ophthalmology to detect vision-threatening diabetic retinopathy.<sup>48</sup>



301           The advantages of our study are its considerably large and multiethnic population,  
302    which is, to our knowledge, the largest studied population comparing papilledema with ODD.  
303    We also believe that the ground truth for ODD and papilledema in this study was robustly  
304    accurate. This is particularly true for patients with papilledema, as the presence of elevated  
305    intracranial pressure was confirmed in every case, either by results of brain imaging or by the  
306    cerebrospinal fluid opening pressure. Furthermore, we have confidence in our results regarding  
307    the diagnosis of ODD and its subgroups given that: 1) the distribution of the ODD (visible vs  
308    buried) according to age groups was comparable in the training and the testing datasets, and  
309    also aligned with previous findings in the literature<sup>49</sup>; indeed, as in our study (Table 3), the  
310    frequency of visible ODD increases with age; and 2) the three referring centers of ODD images  
311    are members of the international ODDS Consortium, a group of clinicians with a specific  
312    interest and expertise in the detection and study of ODD.

313           Regarding limitations, this is a retrospective study based on convenience samples,  
314    acquired by expert neuro-ophthalmologists who provided high-quality images obtained with  
315    mydriatic desktop cameras; therefore, we cannot yet know if these results can be extrapolated  
316    to images obtained with non-mydriatic or handheld cameras. Additionally, we have not  
317    assessed patients with coexistent optic disc pathologies (i.e., ODD and papilledema), although  
318    clinically, it is likely a rare occurrence.

319           **In conclusion**, our DLS applied to color ocular fundus images accurately discriminates  
320    between eyes with true papilledema and eyes with ODD at all severities; its performance  
321    decreases only marginally when challenged to make the clinically difficult distinction between  
322    eyes with buried ODD and eyes with mild-to-moderate papilledema. Further prospective  
323    studies are needed to confirm the applicability of such a DLS in real clinical settings.

324 **Acknowledgements:** we thank Megan Tay Mei Chen and Jodi Ling Wei Yan for their precious  
325 assistance during this study.

## 326 References

- 327 1. Biousse V, Bruce BB, Newman NJ. Ophthalmoscopy in the 21st century: The 2017 H. Houston  
328 Merritt Lecture. *Neurology*. 2018;90(4):167-175. doi:10.1212/WNL.0000000000004868
- 329 2. Fisayo A, Bruce BB, Newman NJ, Biousse V. Overdiagnosis of idiopathic intracranial  
330 hypertension. *Neurology*. 2016;86(4):341-350. doi:10.1212/WNL.0000000000002318
- 331 3. Kovarik JJ, Doshi PN, Collinge JE, Plager DA. Outcome of pediatric patients referred for  
332 papilledema. *Journal of American Association for Pediatric Ophthalmology and Strabismus*  
333 *{JAAPOS}*. 2015;19(4):344-348. doi:10.1016/j.jaapos.2015.05.007
- 334 4. Skougaard M, Heegaard S, Malmqvist L, Hamann S. Prevalence and histopathological signatures  
335 of optic disc drusen based on microscopy of 1713 enucleated eyes. *Acta Ophthalmol*.  
336 2020;98(2):195-200. doi:10.1111/aos.14180
- 337 5. Malmqvist L, Li XQ, Eckmann CL, et al. Optic disc drusen in children: The Copenhagen Child  
338 Cohort 2000 Eye Study. *Journal of Neuro-Ophthalmology*. 2018;38(2):140-146.  
339 doi:10.1097/WNO.0000000000000567
- 340 6. Friedman AH, Henkind P, Gartner S. Drusen of the optic disc. A histopathological study. *Trans*  
341 *Ophthalmol Soc U K*. 1975;95(1):4-9.
- 342 7. Fraser JA, Rueløkke LL, Malmqvist L, Hamann S. Prevalence of optic disc drusen in young patients  
343 with nonarteritic anterior ischemic optic neuropathy: A 10-year retrospective study. *J*  
344 *Neuroophthalmol*. 2021;41(2):200-205. doi:10.1097/WNO.0000000000000974
- 345 8. Neudorfer M, Ben-Haim MS, Leibovitch I, Kesler A. The efficacy of optic nerve ultrasonography  
346 for differentiating papilloedema from pseudopapilloedema in eyes with swollen optic discs. *Acta*  
347 *Ophthalmologica*. 2013;91(4):376-380. doi:10.1111/j.1755-3768.2011.02253.x
- 348 9. Kohli AA, Pistilli M, Alfaro C, et al. Role of ocular ultrasonography to distinguish papilledema  
349 from pseudopapilledema. *Journal of Neuro-Ophthalmology*. 2021;41(2):206-211.  
350 doi:10.1097/WNO.0000000000000984
- 351 10. Costello F, Rothenbuehler S, Sibony P, Hamann S, the Optic Disc Drusen Studies Consortium.  
352 Diagnosing optic disc drusen in the modern imaging era: A practical approach. *Neuro-*  
353 *Ophthalmology*. 2021;45(1):1-16. doi:10.1080/01658107.2020.1810286
- 354 11. Costello F, Malmqvist L, Hamann S. The role of optical coherence tomography in differentiating  
355 optic disc drusen from optic disc edema. *Asia Pac J Ophthalmol (Phila)*. 2018;7(4):271-279.  
356 doi:10.22608/APO.2018124
- 357 12. Chang MY, Velez FG, Demer JL, et al. Accuracy of diagnostic imaging modalities for classifying  
358 pediatric eyes as papilledema versus pseudopapilledema. *Ophthalmology*. 2017;124(12):1839-  
359 1848. doi:10.1016/j.ophtha.2017.06.016
- 360 13. Girard MJ, Panda S, Tun TA, et al. Discriminating Between Papilledema and Optic Disc Drusen  
361 Using 3D Structural Analysis of the Optic Nerve Head. *Neurology*. Published online September  
362 29, 2022:10.1212/WNL.0000000000201350. doi:10.1212/WNL.0000000000201350

- 363 14. Bruce BB, Thulasi P, Fraser CL, et al. Diagnostic accuracy and use of nonmydriatic ocular fundus  
364 photography by emergency physicians: phase II of the FOTO-ED study. *Ann Emerg Med.*  
365 2013;62(1):28-33.e1. doi:10.1016/j.annemergmed.2013.01.010
- 366 15. Thulasi P, Fraser CL, Biousse V, Wright DW, Newman NJ, Bruce BB. Nonmydriatic ocular fundus  
367 photography among headache patients in an emergency department. *Neurology.* 2013;80(5):432-  
368 437. doi:10.1212/WNL.0b013e31827f0f20
- 369 16. Gili P, Flores-Rodríguez P, Yangüela J, Orduña-Azcona J, Martín-Ríos MD. Sensitivity and  
370 specificity of monochromatic photography of the ocular fundus in differentiating optic nerve head  
371 drusen and optic disc oedema: optic disc drusen and oedema. *Graefes Arch Clin Exp Ophthalmol.*  
372 2013;251(3):923-928. doi:10.1007/s00417-012-2223-1
- 373 17. Lee KM, Woo SJ, Hwang JM. Differentiation between optic disc drusen and optic disc oedema  
374 using fundus photography. *Acta Ophthalmologica.* 2017;95(4):e329-e335. doi:10.1111/aos.13338
- 375 18. Leong YY, Vasseneix C, Finkelstein MT, Milea D, Najjar RP. Artificial intelligence meets neuro-  
376 ophthalmology. *Asia Pac J Ophthalmol (Phila).* 2022;11(2):111-125.  
377 doi:10.1097/APO.0000000000000512
- 378 19. Milea D, Najjar RP, Jiang Z, et al. Artificial intelligence to detect papilledema from ocular fundus  
379 Photographs. *New England Journal of Medicine.* 2020;382(18):1687-1695.  
380 doi:10.1056/NEJMoa1917130
- 381 20. Biousse V, Newman NJ, Najjar RP, et al. Optic disc classification by deep learning versus expert  
382 neuro-ophthalmologists. *Ann Neurol.* 2020;88(4):785-795. doi:10.1002/ana.25839
- 383 21. Vasseneix C, Najjar RP, Xu X, et al. Accuracy of a deep learning system for classification of  
384 papilledema severity on ocular fundus photographs. *Neurology.* 2021;97(4):e369-e377.  
385 doi:10.1212/WNL.0000000000012226
- 386 22. Malmqvist L, Bursztyn L, Costello F, et al. The Optic Disc Drusen Studies Consortium  
387 Recommendations for diagnosis of optic disc drusen using optical coherence tomography. *J*  
388 *Neuroophthalmol.* 2018;38(3):299-307. doi:10.1097/WNO.0000000000000585
- 389 23. Jørgensen M, Malmqvist L, Hansen AE, Fraser JA, Hamann S. Volumetric measurement of  
390 peripapillary hyperreflective ovoid masslike structures in patients with optic disc drusen.  
391 *Ophthalmology Science.* 2022;2(1):100096. doi:10.1016/j.xops.2021.100096
- 392 24. Katz BJ, Pomeranz HD. Visual field defects and retinal nerve fiber layer defects in eyes with buried  
393 optic nerve drusen. *Am J Ophthalmol.* 2006;141(2):248-253. doi:10.1016/j.ajo.2005.09.029
- 394 25. Malmqvist L, Wegener M, Sander BA, Hamann S. Peripapillary Retinal Nerve Fiber Layer  
395 Thickness Corresponds to Drusen Location and Extent of Visual Field Defects in Superficial and  
396 Buried Optic Disc Drusen. *J Neuroophthalmol.* 2016;36(1):41-45.  
397 doi:10.1097/WNO.0000000000000325
- 398 26. Milea L, Najjar RP. Classif-Eye: A Semi-automated image classification application; 2020. GitHub  
399 repository. [github.com/milealeonard/Classif-Eye/](https://github.com/milealeonard/Classif-Eye/).
- 400 27. Friedman DI, Liu GT, Digre KB. Revised diagnostic criteria for the pseudotumor cerebri syndrome  
401 in adults and children. *Neurology.* 2013;81(13):1159-1165. doi:10.1212/WNL.0b013e3182a55f17

- 402 28. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image  
403 segmentation. *arXiv:150504597 [cs]*. Published online May 18, 2015. Accessed May 4, 2022.  
404 <http://arxiv.org/abs/1505.04597>
- 405 29. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *Journal of*  
406 *Big Data*. 2019;6(1):60. doi:10.1186/s40537-019-0197-0
- 407 30. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image  
408 database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition.*; 2009:248-255.  
409 doi:10.1109/CVPR.2009.5206848
- 410 31. Tan M, Le QV. EfficientNet: Rethinking model scaling for convolutional neural networks.  
411 *arXiv:190511946 [cs, stat]*. Published online September 11, 2020. Accessed May 4, 2022.  
412 <http://arxiv.org/abs/1905.11946>
- 413 32. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual  
414 explanations from deep networks via gradient-based localization. In: *2017 IEEE International*  
415 *Conference on Computer Vision (ICCV)*. 2017:618-626. doi:10.1109/ICCV.2017.74
- 416 33. Pineles SL, Arnold AC. Fluorescein angiographic identification of optic disc drusen with and  
417 without optic disc edema. *J Neuroophthalmol*. 2012;32(1):17-22.  
418 doi:10.1097/WNO.0b013e31823010b8
- 419 34. Loft FC, Malmqvist L, Wessel Lindberg AS, Hamann S. The Influence of Volume and Anatomic  
420 Location of Optic Disc Drusen on the Sensitivity of Autofluorescence. *J Neuroophthalmol*.  
421 2019;39(1):23-27. doi:10.1097/WNO.0000000000000654
- 422 35. Carter SB, Pistilli M, Livingston KG, et al. The role of orbital ultrasonography in distinguishing  
423 papilledema from pseudopapilledema. *Eye*. 2014;28(12):1425-1430. doi:10.1038/eye.2014.210
- 424 36. Wilkins JM, Pomeranz HD. Visual manifestations of visible and buried optic disc drusen. *J*  
425 *Neuroophthalmol*. 2004;24(2):125-129. doi:10.1097/00041327-200406000-00006
- 426 37. Sibony PA, Kupersmith MJ, Feldon SE, Wang JK, Garvin M, OCT Substudy Group for the  
427 NORDIC Idiopathic Intracranial Hypertension Treatment Trial. Retinal and choroidal folds in  
428 papilledema. *Invest Ophthalmol Vis Sci*. 2015;56(10):5670-5680. doi:10.1167/iovs.15-17459
- 429 38. Sibony PA, Kupersmith MJ, Kardon RH. Optical coherence tomography neuro-toolbox for the  
430 diagnosis and management of papilledema, optic disc edema, and pseudopapilledema. *J*  
431 *Neuroophthalmol*. 2021;41(1):77-92. doi:10.1097/WNO.0000000000001078
- 432 39. Reggie SN, Avery RA, Bavinger JC, et al. The sensitivity and specificity of retinal and choroidal  
433 folds to distinguish between mild papilloedema and pseudopapilledema. *Eye*. 2021;35(11):3131-  
434 3136. doi:10.1038/s41433-020-01368-y
- 435 40. Jivraj I, Cruz CA, Pistilli M, et al. Utility of spectral-domain optical coherence tomography in  
436 differentiating papilledema from pseudopapilledema: A prospective longitudinal study. *J*  
437 *Neuroophthalmol*. 2021;41(4):e509-e515. doi:10.1097/WNO.0000000000001087
- 438 41. Thompson AC, Bhatti MT, El-Dairi MA. Bruch's membrane opening on optical coherence  
439 tomography in pediatric papilledema and pseudopapilledema. *J AAPOS*. 2018;22(1):38-43.e3.  
440 doi:10.1016/j.jaapos.2017.09.003

- 441 42. Fard MA, Fakhree S, Abdi P, Hassanpoor N, Subramanian PS. Quantification of peripapillary total  
442 retinal volume in pseudopapilledema and mild papilledema using spectral-domain optical  
443 coherence tomography. *Am J Ophthalmol*. 2014;158(1):136-143. doi:10.1016/j.ajo.2014.03.008
- 444 43. Sarac O, Tasci YY, Gurdal C, Can I. Differentiation of optic disc edema from optic nerve head  
445 drusen with spectral-domain optical coherence tomography. *J Neuroophthalmol*. 2012;32(3):207-  
446 211. doi:10.1097/WNO.0b013e318252561b
- 447 44. Flowers AM, Longmuir RA, Liu Y, Chen Q, Donahue SP. Variability within optic nerve optical  
448 coherence tomography measurements distinguishes papilledema from pseudopapilledema. *Journal*  
449 *of Neuro-Ophthalmology*. 2021;41(4):496-503. doi:10.1097/WNO.0000000000001137
- 450 45. OCT Sub-Study Committee for NORDIC Idiopathic Intracranial Hypertension Study Group,  
451 Auinger P, Durbin M, et al. Baseline OCT measurements in the idiopathic intracranial hypertension  
452 treatment trial, part II: correlations and relationship to clinical features. *Invest Ophthalmol Vis Sci*.  
453 2014;55(12):8173-8179. doi:10.1167/iovs.14-14961
- 454 46. Chen JJ, Kardon RH. Avoiding clinical misinterpretation and artifacts of optical coherence  
455 tomography analysis of the optic nerve, retinal nerve fiber layer, and ganglion cell layer. *J*  
456 *Neuroophthalmol*. 2016;36(4):417-438. doi:10.1097/WNO.0000000000000422
- 457 47. Marques G, Agarwal D, de la Torre Díez I. Automated medical diagnosis of COVID-19 through  
458 EfficientNet convolutional neural network. *Applied Soft Computing*. 2020;96:106691.  
459 doi:10.1016/j.asoc.2020.106691
- 460 48. Chetoui M, Akhloufi MA. Explainable diabetic retinopathy using EfficientNET. *Annu Int Conf*  
461 *IEEE Eng Med Biol Soc*. 2020;2020:1966-1969. doi:10.1109/EMBC44109.2020.9175664
- 462 49. Malmqvist L, Lindberg ASW, Dahl VA, Jørgensen TM, Hamann S. Quantitatively Measured  
463 Anatomic Location and Volume of Optic Disc Drusen: An Enhanced Depth Imaging Optical  
464 Coherence Tomography Study. *Investigative Ophthalmology & Visual Science*. 2017;58(5):2491-  
465 2497. doi:10.1167/iovs.17-21608