Deep learning – the ophthalmological care of the future?

Prof. Dr. Daniel H. Scorsetti
Dr. Luciano D. Scorsetti
Prof. Dr. Leopold Schmetterer

Introduction

Deep learning is a type of machine learning that aims to train a computer to perform tasks that are typically performed by humans based on artificial neural networks. Recent advances in technology have shown that artificial neural networks can be applied to fields such as speech and audio recognition, machine translation, board games, drug design, and medical image analysis. The development of these techniques has been extremely fast over the recent years and artificial neural networks nowadays outperform humans in many of these tasks. Artificial neural networks were inspired by the function of biological systems such as the brain and connected nodes within these networks model neurons. The principle of such networks is that they are trained with datasets where the ground truth is known. As an example the network shall be trained to identify images where a bicycle is depicted. This requires a large number of images where bicycles are manually labeled (the so-called ground truth) that are then analyzed by the computer. If enough images with either bicycle or no bicycle are used the artificial neural network can be trained to identify bicycles in other image sets.

In medical imaging classical approaches include either extraction of semantic features defined by human experts or agonistic features defined by equations. Semantic features may provide good specificity for disease diagnosis, but may differ between different doctors dependent on their level of experience, are time consuming and costly. Agonistic features may have limited specificity, but offer the advantage of high reproducibility. Deep learning takes a different approach. A training dataset is required where the ground truth, in this case the diagnosis, is known. The number of data required is high and usually 100000 images or more are used. Once the artificial neural network is trained it can be applied to a validation dataset in which the diagnosis is also know, but not told to the computer. The output of the artificial neural network is in the simplest case either disease or no disease that can be compared to the ground truth. The agreement with the ground truth is quantified using measures such as area under the curve (AUC, can take values between 0 and 1 with 1 meaning perfect discrimination between health and disease), specificity (can take values between 0% and 100% and quantifies the proportion of actual negatives that are correctly identified) and sensitivity (can take values between 0% and 100% and quantifies the proportion of actual positives that are correctly identified). If high sensitivity or high specificity is required depends on the disease, the prevalence of the disease as well as the actual clinical setting where this network should be employed.

Deep learning in ocular disease

Using retinal fundus photographs from diabetic patients several authors have shown that algorithms based on deep learning provide excellent performance for detecting referable diabetic retinopathy. The first study that showed high performance was published by Google and included more than 120 000 images in the training dataset (Gulshan et al. 2016). Later several authors reported convolutional neural networks that showed high discriminative power for diabetic retinopathy based on fundus photographs (Ting et al. 2017, Gargeya and Leng 2017). Direct comparison between the performance reported in these publication is, however, difficult since this may strongly depend on the characteristics of the study population including ethnicity, severity of disease, the reference standard defined as graded by retina specialists (ground truth) and the quality of the fundus photographs. Recently it was also shown that convolutional neural networks can also replace human grading in clinical studies including epidemiological datasets (Ting et al. 2019a). Whereas most of the features of diabetic retinopathy can be discovered from fundus photographs diabetic macular edema may require optical coherence tomography (OCT) based approaches (Roy et al. 2017, Schlegl et al. 2018).

In age-related macular degeneration the definition of referable disease is not as straightforward as in diabetic retinopathy, because several grading systems for the disease have been proposed (Ferris et al. 2013, Klein et al. 2014). Deep-learning based classification systems were proposed for referability (Burlina et al., 2017) as well as estimation of 5-year risk of conversion to late-stage AMD (Burlina et al. 2018). Further efforts are required to clarify how such systems can be implemented into clinical practice. One study used deep learning to predict the outcome after anti VEGF treatment and found that the amount of intraretinal fluid is associated with baseline visual acuity as well as the visual outcome after 12 months (Schmidt-Erfurth et al. 2018a). Whether artificial intelligence-based retreatment criteria for ant-VEGF injections can be used to guide the frequency of treatment remains to be proven.

In glaucoma artificial intelligence approaches are still in their infancy. This is partially related to the problems in defining the disease based on structural and functional measurements (Casson et al. 2012) and reflected in the uncertainty of the clinical diagnosis defining patients as glaucoma suspects and following them up longitudinally to ascertain diagnosis. Approaches were published based on optic disc photographs showing sufficient specificity and sensitivity (Shibata et al. 2018), but it is unlikely that the complex morphological changes in the optic nerve head that are characteristic for glaucoma can be adequately detected with a 2-dimensional imaging technique. For OCT-based artificial intelligence approaches to glaucoma both macular scans (Asaoka et al. 2019) and optic disc scans have been used (Maetschke et al. 2019), but the number of included images is relatively small and the algorithms have not yet been validated in large-scale multi-ethnic populations. Alternatively visual field data can be used to train convolutional neural networks and different approaches have been proposed (Cai et al. 2017, Wang et al. 2018), but the high variability of data still limits clinical applicability. Applications of artificial intelligence-approaches for glaucoma are not straightforward. Glaucoma screening has so far been considered cost-ineffective (Momont and Mills 2013) and it is yet to be shown whether deep-learning approaches can be employed for cost effective programs. As mentioned above diagnosis in glaucoma is a complex clinical task including many different examination modalities and unlikely to be replaced by deep learning in the near future. Progression analysis may an attractive area for using deep learning approaches, but basing treatment decisions on artificial intelligence will most require randomized controlled trials.

Artificial intelligence algorithms cannot only be used for disease classification, but also for segmentation of images and enhancement of image quality. For the development of deep learning segmentation algorithms a large-scale dataset of manually annotated images is required that can serve as ground truth. Algorithms have been described for segmentation of intraretinal fluid in OCT images (Lee et al., 2017) and segmentation of pigment epithelium detachment (Xu et al., 2017, Schmidt-Erfurth et al., 2018a). Segmentation algorithms have also been proposed for the optic nerve head region with special emphasis on glaucoma (Devalla et al. 2018a,b) and the cornea in an effort for early detection of keratoconus (Bata et al. 2016, D’Aranha et al. 2019). Deep learning has also been used for denoising OCT images (Halupka et al. 2018) thereby artificially producing high quality images from low quality OCT scans. Since these networks have been trained in scans from healthy subjects it is not yet clear whether they can be successfully applied to images from patients with retinal and/or optic nerve head pathology.

Conclusions

Artificial intelligence will revolutionize ophthalmological care in the 21st century. There is still a need to better define how these convolutional neuronal networks shall be utilized in clinical routine. The most advanced application is the use in diabetic retinopathy screening programs to identify patients that are at high risk of getting blind. For more information the reader is referred to some recent in-depth reviews that provide more details on the technical details and clinical applications of this innovation (Schmidt Erfurth et al. 2018b, Ting et al. 2019b,c).

Asaoka R, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, Kanamoto T, Ikeda Y, Mori K, Iwase A, Shoji N, Inoue K, Yamagami J, Araie M. (2019) Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. Am J Ophthalmol. 198:136-145.

Bata AM, Witkowska KJ, Wozniak PA, Fondi K, Schmidinger G, Pircher N, Szegedi S, Aranha Dos Santos V, Pantalon A, Werkmeister RM, Garhofer G, Schmetterer L, Schmidl D. (2016). Effect of a matrix therapy agent on corneal epithelial healing after standard collagen cross-linking in patients with keratoconus: A randomized clinical trial. JAMA Ophthalmol. 134:1169-1176.

Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. (2017) Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol. 135:1170-1176.

Burlina PM, Joshi N, Pacheco KD, Freund DE, Kong J, Bressler NM. (2018) Use of deep learning for detailed severity characterization and estimation of 5-year risk among patients with age-related macular degeneration. JAMA Ophthalmol. 136:1359-1366.

Cai S, Elze T, Bex PJ, Wiggs JL, Pasquale LR, Shen LQ. (2017) Clinical correlates of computationally derived visual field defect archetypes in patients from a glaucoma clinic. Curr Eye Res. 42:568-574.

Casson RJ, Chidlow G, Wood JP, Crowston JG, Goldberg I. (2012) Definition of glaucoma: clinical and experimental concepts. Clin Exp Ophthalmol. 40:341-349.

Devalla SK, Renukanand PK, Sreedhar BK, Subramanian G, Zhang L, Perera S, Mari JM, Chin KS, Tun TA, Strouthidis NG, Aung T, Thiéry AH, Girard MJA. (2018a) DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomed Opt Express. 9:3244-3265.

Devalla SK, Chin KS, Mari JM, Tun TA, Strouthidis NG, Aung T, Thiéry AH, Girard MJA. (2018) A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head. Invest Ophthalmol Vis Sci. 59:63-74

Dos Santos VA, Schmetterer L, Stegmann H, Pfister M, Messner A, Schmidinger G, Garhofer G, Werkmeister RM. (2019) CorneaNet: fast segmentation of cornea OCT scans of healthy and keratoconic eyes using deep learning. Biomed Opt Express. 10:622-641.

Ferris FL 3rd, Wilkinson CP, Bird A, Chakravarthy U, Chew E, Csaky K, Sadda SR; Beckman Initiative for Macular Research Classification Committee. (2013) Clinical classification of age-related macular degeneration. Ophthalmology. 120:844-851.

Gargeya R, Leng T. (2017) Automated identification of diabetic retinopathy using deep learning. Ophthalmology. 124:962-969.

Halupka KJ, Antony BJ, Lee MH, Lucy KA, Rai RS, Ishikawa H, Wollstein G, Schuman JS, Garnavi R. (2018) Retinal optical coherence tomography image enhancement via deep learning. Biomed Opt Express. 9:6205-6221.

Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 316:2402-2410.

Lee CS, Tyring AJ, Deruyter NP, Wu Y, Rokem A, Lee AY. (2017) Deep-learning based, automated segmentation of macular edema in optical coherence tomography. Biomed Opt Express. 8:3440-3448.
Maetschke S, Antony B, Ishikawa H, Wollstein G, Schuman J, Garnavi R. (2019) A feature agnostic approach for glaucoma detection in OCT volumes. PLoS One. 14:e0219126.
Momont AC, Mills RP. (2013) Glaucoma screening: current perspectives and future directions. Semin Ophthalmol. 28:185-90.
Roy AG, Conjeti S, Karri SPK, Sheet D, Katouzian A, Wachinger C, Navab N. (2017) ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed Opt Express. 13:3627-3642.
Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, Podkowinski D, Gerendas BS, Langs G, Schmidt-Erfurth U. (2018a) Fully automated detection and quantification of macular fluid in OCT using deep learning. Ophthalmology. 125:549-558.
Schmidt-Erfurth U, Waldstein SM, Klimscha S, Sadeghipour A, Hu X, Gerendas BS, Osborne A, Bogunovic H. (2018a) Prediction of individual disease conversion in early AMD using artificial intelligence. Invest Ophthalmol Vis Sci. 59:3199-3208.
Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. (2018b) Artificial intelligence in retina. Prog Retin Eye Res. 67:1-29
Shibata N, Tanito M, Mitsuhashi K, Fujino Y, Matsuura M, Murata H, Asaoka R. (2018) Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. Sci Rep. 8:14665.
Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, San Yeo IY, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY. (2017) Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA. 318:2211-2223.
Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, Tan GSW, Schmetterer L, Keane PA, Wong TY. (2019b) Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 103:167-175.
Ting DSW, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, Schmetterer L, Pasquale LR, Bressler NM, Webster DR, Abramoff M, Wong TY. (2019c) Deep learning in ophthalmology: The technical and clinical considerations. Prog Retin Eye Res. 72:100759.
Wang M, Pasquale LR, Shen LQ, Boland MV, Wellik SR, De Moraes CG, Myers JS, Wang H, Baniasadi N, Li D, Silva RNE, Bex PJ, Elze T. (2018) Reversal of glaucoma hemifield test results and visual field features in glaucoma. Ophthalmology. 125:352-360.
Xu Y, Yan K, Kim J, Wang X, Li C, Su L, Yu S, Xu X, Feng DD. (2017) Dual-stage deep learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy. Biomed Opt Express. 8:4061-4076.