International Journal of Science and Technology Research Archive

ISSN: 0799-6632 (Online)

Journal homepage: https://sciresjournals.com/ijstra/

(Review Article)

Check for updates

Deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy: A review

Francisco Javier Corvalan [2], Nathalie Márquez [2], Nathalia Garcia [2], Ankur Seth [1, 3] and Carlos Eduardo Rivera [1, 2, 3]

[1] *Ophthalmology Department, Collective Innovations Colombia, Cali, Colombia.*
[2] *Ophthalmology Department, Javeriana University, Cali-Colombia.*
[3] *Ophthalmology Department, GSR medical center, Cali, Colombia.*

International Journal of Science and Technology Research Archive, 2023, 04(01), 103–112

Publication history: Received on 05 November 2022; revised on 19 December 2022; accepted on 21 December 2022

Article DOI: https://doi.org/10.53771/ijstra.2023.4.1.0163

Abstract

Relevance: Glaucoma is a group of diseases characterized by progressive, bilateral yet asymmetric optic neuropathy, which results in permanent vision loss when is not treated promptly; It is asymptomatic in the early stages; thus, unfortunately, the diagnosis is discovered when the compromise is already severe, and the condition is advanced. Because of this, it is crucial to conduct early screening using technologies that are accessible to the population. Artificial intelligence (AI), particularly deep learning (DL), plays an essential role in this issue. DL may be an efficient approach for glaucoma screenings with the proper training.

Objective: Describe the development of AI and DL over time and their use and significance in glaucoma screening.

Methods: A literature search was conducted in PUBMED/MEDLINE, EMBASE, and manuscript references in English and Spanish between January 2014 to July 2022 on the role and evolution of AI and DL over the years and the usefulness of deep learning for glaucoma diagnosis. Of the 1914 abstracts reviewed, 105 articles were selected that contained information on the history of AI in medicine and the applicability of this tool for the early diagnosis of glaucoma.

Findings and conclusions: We can demonstrate that deep learning can outperform glaucoma specialists in diagnosing the condition through fundus imaging data; DL is an exciting tool in the screening and early diagnosis of glaucoma.

Keywords: Artificial Intelligence; Deep learning; Glaucoma; Optic Nerve; Fundus images

1 Introduction

Glaucoma is a progressive optic neuropathy, usually bilateral, characterized by the loss of retinal ganglion cells (RGCs) and their axons, and associated visual field defects (1). It is a disease that remains asymptomatic in most cases until it leads to loss of vision in advanced stages (2).

It is a disease that affects more than 70 million people worldwide, causing approximately 10% of bilateral blindness, which makes it the leading cause of irreversible blindness in the world (3). The Pan American Association of Ophthalmology (PAAO) estimated that by 2015 one in five people in Latin America and the Caribbean had some degree of vision loss. By 2020, it was predicted that this proportion would rise to 2.2–3.7% in Latin America, which is higher than the global average of 2.0% for blindness and moderate to severe visual impairment (4).

* Corresponding author: Carlos Eduardo Rivera Hoyos; Email: carlosriverahoyos@gmail.com

Its worldwide prevalence has been increasing over the years. In 2013, 64 million cases were reported, rising to 76 million in 2020, and estimated for 2040, 111.8 million cases (2,5). In Colombia, the year 2005, the overall prevalence of glaucoma was 1.1%, being more common in women (1,2), contributing to 2.7% of the total visual impairment during that period. Diagnosis requires a clinical examination combined with quantitative functional and structural measurements, which are not accessible in all populations and can be expensive, tedious, time-consuming, and prone to human error (2). Glaucoma is estimated to cost $2.5 billion of the U.S. healthcare budget annually, with $1.9 billion in direct medical costs (6). Therefore, there is a need to develop an automatic glaucoma detection system that can provide an affordable, accurate, rapid, and interpretable diagnosis (7).

There are multiple methods for the evaluation of glaucoma, including optical coherence Tomography (OCT), visual fields, and optic nerve photography; however, the first two are usually expensive and inaccessible to the majority population since they are generally available only in ophthalmology clinics and some tertiary hospital centers (8). For this reason, artificial intelligence (AI) trained to interpret optic nerve photographs constitutes an exciting and cost-effective approach for early glaucoma diagnosis. DL is the evolution of machine learning within AI, which *uses layered algorithmic architectures to analyze data.*

In the case of this pathology, Deep Learning convolutional neural networks can outperform glaucoma specialists in detecting the disease based on fundus image data, achieving better diagnostic accuracy. With these findings, this technology opens several future possibilities for glaucoma screening, confirmation, and follow-up (1).

2 Methods

A literature search for scientific articles was performed in databases such as PUBMED/MEDLINE and EMBASE and in manuscript references in English and Spanish between January 2014 to February 2022. Of the 1464 abstracts reviewed, 450 articles were selected that contained information on the usefulness of deep learning in diagnosing glaucoma. Of the 450 articles reviewed, 35 articles were selected that contained information on DL as a diagnostic tool for glaucoma, emphasizing diagnosis using fundus photographs.

3 Primary Glaucoma

3.1 Physiopathology

Glaucoma is an optic neuropathy characterized by progressive degeneration of retinal ganglion cells and their axons, generating specific morphological changes. Retinal ganglion cells are neurons of the central nervous system that receive signals from photoreceptors, which they process and transmit in axons through the optic nerve to other brain parts. The degeneration of these nerves leads to the formation of excavation, which may increase and thus generate a progressive loss of vision. Likewise, papillary hypoperfusion generated by high intraocular pressure, low perfusion pressure, and low cerebrospinal fluid pressure lead to structural changes and remodeling of the lamina cribosa, which also generates deterioration of axonal transport in the optic nerve fibers (3,9).

This increasing loss of retinal ganglion cells leads to a progressive deterioration of the visual field, which usually begins mid-periphery and then progresses until only a central or peripheral island of vision remains intact. Other functional alterations include impaired contrast and color perception and difficulty in reading. However, the mechanisms by which retinal ganglion cell losses are not yet fully understood (3).

3.2 Classification and Epidemiology

Primary glaucoma can be divided into two types: primary open-angle glaucoma (POAG) and primary angle-closure glaucoma (PACG). In the United States, most cases (80%) are due to open-angle glaucoma; however, angle-closure glaucoma accounts for a disproportionate number of patients with severe vision loss. Risk factors for developing POAG include ocular hypertension, age, family history of glaucoma, enlarged papillary excavation, decreased central corneal thickness, and high myopia, among others (10). The prevalence of (POAG) increases with age, from 0.4% at 40-44 years to 2.7% at 70-74 years and 10.0% after 90 years of age in people of European descent (3).

The prevalence of (POAG) also varies by ethnic group; African descent people have a glaucoma prevalence 2.8 times higher than Europeans, while angle-closure glaucoma and normal pressure glaucoma are more common in Asians (11). In 2013, the global prevalence of PACG was shown to be 0.5% [95% confidence interval (CI) = 0.11-1.36%]. In addition, the global population with (PACG) was calculated to be 23.36 million by 2020 and 32.04 million by 2040, with Asia being the continent accounting for more than three-quarters of the (PACG) population (8).

3.3 Clinical Manifestations

Glaucomatous changes are manifested by tissue loss at the neuroretinal rim and enlargement of the optic nerve excavation enlargement. Other findings include a non-physiological discrepancy between the optic nerve excavations in the two eyes, hemorrhages at the edge of the optic disc, thinning of the retinal nerve fiber layer, and atrophy of the peripapillary tissue, which can be evaluated by fundus examination (2).

3.4 Diagnosis

Glaucoma is diagnosed based on the distinctive clinical signs of progressive optic neuropathy and the corresponding visual field abnormalities. The American Academy of Ophthalmology currently advises individuals with glaucoma risk factors to undergo routine complete ophthalmologic exams at a frequency that is individually decided, considering age, risk factors, race, and family history (12).

Based on this, diagnosis requires both structural and functional testing. Using serial visual fields assesses the degree of functional impairment resulting from the loss of optic nerve fibers. While structural tests, such as optical coherence tomography, evaluate the existence of nerve fiber layer defects(3). Nerve fiber layer defects concordant with altered visual fields make a diagnosis sure; therefore, it is essential to keep in mind the specificity and sensitivity that both functional and structural studies currently provide us, as both are necessary for the establishment and follow-up of the disease (10).

Among the ophthalmologic tests that aids in diagnosis are nerve fiber and ganglion cell layer OCT evaluation, visual field, and optic nerve photography. In addition, not all these tests are easily accessible to the general population and are operator-dependent, which can lead to errors in reliability. To sum up, the major impediment to glaucoma screening is cost-effectiveness. Because of this, the search for new screening technologies, such as artificial intelligence through its different branches, could positively impact the affordability of glaucoma screening as new screening technologies become available (13).

4 Artificial Intelligence

4.1 Evolvement

Artificial intelligence is described as the ability of a computer to mimic intellectual intelligence, link events to specific causes, generalize, and learn from experience (14).

Alan Turing introduced the idea of using computers to emulate intelligent behavior. The "Turing test," which Turing later popularized, measures a machine's capacity to behave intelligently. The experiment involved evaluating natural language exchange between humans.

Six years later, John McCarthy described artificial intelligence (AI) as "the science and engineering of making intelligent machines." (14).

Early AI focused on developing machines that could make inferences or decisions that only a human could previously realize, as was the first robotic arm in 1961 by General Motors. In 1964 Joseph Weizenbaum created Eliza, which through a natural language processing algorithm, could communicate using pattern matching and substitution methodology to mimic a human conversation (surface communication), serving as a framework for future chatbots (14).

In 1966, Shakey, "the first electronic person" developed at Stanford, was the first mobile robot to interpret more complex instructions. Thanks to these innovations, medicine entered the world of AI (14).

The prototype demonstrating the feasibility of applying AI to medicine was developing a query program for glaucoma using the CASNET model. It was developed at Rutgers University and officially presented at the American Academy of Ophthalmology meeting in Las Vegas, Nevada, in 1976 (14).

In 1986, the University of Massachusetts launched DXplain, a decision support system. Based on the different symptoms entered on the platform, this program could generate a list of possible differential diagnoses. It also served as an electronic medical textbook, providing detailed descriptions of diseases and additional references. When it was first released, DXplain could provide information on approximately 500 diseases. Since then, it has expanded to more than 2,400 diseases (14).

In 2007, IBM created an open-domain question-answering system called Watson, which competed with human participants and won first place on Jeopardy's television game show in 2011. Unlike traditional systems that used prospective reasoning, it was capable of reasoning in hindsight. This technology, called DeepQA, uses natural language processing and various queries to analyze data about unstructured content to generate possible answers. By extracting information from a patient's electronic health record and other electronic resources, DeepQA technology could be applied to provide evidence-based medical answers. As such, it opened new possibilities in evidence-based clinical decision-making (14).

Natural language processing transformed chatbots from external communication (like Eliza) to meaningful conversation-based interfaces. This technology was applied to Apple's virtual assistant Siri in 2011 and Amazon's virtual assistant Alexa in 2014. Pharmabot is a chatbot developed in 2015 to assist in medication education for pediatric patients and their parents. Mandy was created in 2017 to optimize and automate the patient admission process in a primary health care context (14).

As can be deduced from the previous paragraphs, AI has been evolving over several decades to include more complex algorithms that work like the processing of the human brain. AI encompasses many sub-branches like medical specialties, among which Machine Learning (ML), Deep Learning (DL), and Computer Vision (CV) stand out (14).

4.2 Machine learning, Deep learning

The term "Machine Learning" (ML) is used when referring to AI; it refers to the ability of machines to "learn" autonomously, from experience, without being explicitly programmed, and without human intervention or assistance. Arthur Samuel first introduced this term in 1959 (16).

Although we may be unfamiliar with the term, ML technology drives many aspects of our daily lives, such as web searches, content filtering on social networks, recommendations/advertising on e-commerce websites, and is increasingly present in mass consumer products such as smartphones (17,18).

On the other hand, DL marked an important breakthrough for including AI in medicine. Unlike ML, which uses a set number of features and requires human involvement, DL can be trained to classify data using neural networks generated by the algorithm itself. Although DL was first introduced in the 1950s, its application to medicine was limited for several decades due to "overfitting"(14). The term "overfitting" refers to when machine learning focuses too much on a specific data set and cannot accurately process new data sets, usually due to insufficient computing power and lack of training data. These limitations were overcome in the 2000s due to the availability of larger datasets and superior computing power (14).

In 2006, Geoffrey Hinton introduced new architecture associated with Deep Learning. These are convolutional neural networks (CNNs), deep belief networks, and recurrent neural networks, allowing computers to distinguish objects and text in images and videos (14).

These algorithms learn features (or representations) from the data, provided they are previously trained with enough information. The algorithm works by directly entering the relevant features of the references to be analyzed into a DL model, automatically learning the most relevant features to analyze and process the information according to what it has learned (19).

4.3 Image processing

Fundus photography is used to diagnose and monitor the disease or for programs to detect and evaluate symptoms of retinal detachment or ocular disease (20).

Machine learning has allowed in the last decade to increase the ability to process data with high dimensionality through image detection. Kirzhevsky et al. in 2012 exposed the benefits of using CNN networks in a competition to correctly classify more than one million images with the lowest error rate (21).

An image, for example, comes in its native form as a matrix of pixels. The features learned in the first layer generally represent the presence/absence of edges and the particular orientation of the image. The second layer usually detects motifs by detecting specific edge details, and the third layer may assemble motifs into larger combinations, corresponding, for example, to parts of familiar objects. The following layers will be able to detect objects as combinations of these layers already described (22).

The steps for data processing and analysis are learned and implemented implicitly. The only design work for the developer of such systems is the number of layers, their interconnection, and the appropriate choice of images for training (21). For the algorithm to deliver an optimal result for a never-before-seen image, it must be exposed to many images (usually tens of thousands) during the learning phase. This phase has an essential role in the final response of our neural network since it will be able to establish hidden relationships between the input variables, providing greater complexity to its final interpretation (22,23). The final result will be produced by combining all the features of the fully connected layers. Some of the DL algorithms now available are: Le-NET, AlexNet, VGG, GoogLeNet, and ResNet(14,24).

4.4 Deep Neural Networks

CNNs consist of a multilayer neural network and are mainly used for image processing and object detection. Yann LeCun developed the first CNN in 1988, which he called LeNet, which used to recognize characters such as zip codes and digits. CNNs are widely used for identifying satellite images, processing medical images, forecasting time series, and detecting anomalies (24,25).

As mentioned above, CNNs have multiple layers that process and extract features from the input data; each layer will have specific characteristics (26

- Convolution layer: It has several filters to perform the convolution operation. The latter is a mathematical operation that combines two signals to produce a third output signal.
- Rectified linear unit (ReLU): Its function is based on performing operations with the elements of the convolution layer. The output signal is a rectified feature map.
- Clustering layer: The grouping layer converts the two-dimensional matrices resulting from the rectified map into a single, long, continuous linear vector.
- Fully connected layer: This is formed when the one-dimensional matrix from the clustering layer is used as input for image classification and identification.

4.4.1 CNN has multiple architectures, but among the best known, we have

LeNet-5, a 7-level convolutional network developed by LeCun and collaborators in 1998, which classifies digits, was used by several banks to recognize handwritten numbers on digitized checks in 32x32 pixel grayscale images. The ability to process higher resolution images requires a more significant number of convolutional layers (19).

AlexNet was designed in 2012 by the SuperVision group. This network has a very similar architecture to LeNet by Yann LeCun et al. but is deeper, with more filters per layer and stacked convolutional layers (24).

Appearing in 2013, ZFNet managed to adjust the AlexNet hyperparameters while keeping the same structure with additional DL elements (24).

In 2014, GoogLeNet / Inception was inspired by LeNet CNN; however, it implemented a novel element called the inception module, which employing several very small convolutions significantly reduced the number of parameters. Its architecture consisted of a 22-layer deep CNN, whose parameters were reduced from 60 million (AlexNet) to 4 million (24).

Then in 2014, VGGNet appeared, consisting of 16 convolutional layers, with a uniform architecture, like AlexNet (3x3 convolutions), but with more filters. Currently, it is the preferred choice in the community for extracting features from images. VGGNet is publicly available and used in many other applications. However, it consists of 138 million parameters, which makes it difficult to handle (24).

The so-called Residual Neural Network (ResNet) by Kaiming He et al. introduced a new form of architecture with "jump connections." This technique made it possible to train a 152-layer neural network with lower complexity compared to VGGNet. It achieves an error rate that exceeds human-level performance on this dataset (24).

As mentioned above, between 2012 and 2014, different deep neural network architectures were created, such as AlexNet, VGG16, and GoogleNet. Since 2014, very deep convolutional networks have become widespread and successfully applied in computer vision tasks such as object detection and tracking, image classification, and single object localization (26).

In 2015, the Inception model emerged. Its name derives from the scientific article Network in network (27) and the internet meme "we need to go deeper." This architecture was widely acclaimed for its outstanding results in the

ImageNet Large Scale Visual Recognition Challenge 2014 (ILSVRC 2014) classification challenge, a platform used to compare image recognition and detection algorithms with approximately 1 million images and 1,000 object classes. This approach was described in a paper published by Christian Szegedy et al. entitled Going Deeper with Convolutions (28).

Since its publication, the Inception model has undergone different transformations, each one based on the philosophy of reducing the computational cost and modifying its previous architecture. For this, different network optimization techniques are used, such as factored convolutions, regularization, dimension reduction, and parallel computations, all to optimize the network and facilitate the model's adoption (26).

5 Glaucoma and deep learning

5.1 Deep learning in OCT

Studies comparing the sensitivity and specificity of spectral-domain and time-domain OCT were performed. The study by Chang et al. evaluated time-domain OCT with one or more abnormal quadrants, and determined a sensitivity of 96%, and specificity of 76%, while for spectral-domain OCT, it was 98% and 80%, respectively (29).

Asaoka et al. constructed a DL model (CNN) to diagnose early glaucoma (mean deviation>-5 dB) by thickness through OCT obtained with Topcon OCT- 1000 or OCT-2000 devices. They used 8×8 macular CFN and ganglion cell thickness with an inner plexiform layer (19). To avoid overfitting problems, they performed transfer learning. They found that the DL model with transfer learning increased the AUC value from 0.766 to 0.937. The DL model also outperformed two traditional machine learning methods. This study showed the transfer learning method's advantages in improving the DL model's performance (30).

Muhammad et al. developed a hybrid DL method (HDLM) to distinguish between eyes classified as healthy or glaucoma-suspicious subjects or with confirmed mild glaucoma. They considered 102 patients, of which 57 were glaucomatous eyes and 45 were healthy/suspicious eyes (31). This method was based on a pretrained CNN model (AlexNet) whose task was feature extraction and a random forest model whose goal was classification. Six types of images were input: 1. map thickness + retinal ganglion cell (RGC)+; 2. map thickness of NFC; 3. probability map of RGC +; 4. probability map of NFC; 5. facial projection; 6. combination of NFC thickness, RGC + probability map, and probability map of NFC. The results yielded degrees of accuracy of the different types of HDLM between 63.7% and 93.1%, while the values of the area under the curve ranged from 0.742 to 0.973. Overall, the NFC probability map had the best accuracy, and the NFC thickness map had the highest AUC value (31).

On the other hand, Thompson et al. developed a segmentation-free DL algorithm based on 2D OCT circular B-scans and found that it achieved better performance in detecting glaucomatous structural changes compared to conventional nerve fiber layer thickness parameters, with AUC: DL model vs CFN thickness = 0.960 vs. 0.870) (32).

5.2 Deep learning in fundus photographs

2-D fundus photography (FO) is an essential ophthalmologic screening tool widely used worldwide due to its affordability, ease of use, and diagnostic capability (33). Several studies have shown an excellent diagnostic correlation between 2-D fundus images vs. images obtained from high-tech devices (OCT, HRT, GDx) (33,34,35). As in the 2006 report by Reus et al., which concluded that automated HRT (Heidelberg Retinal Tomography) analysis had similar diagnostic accuracy for glaucoma when compared to stereoscopic fundus images interpreted by glaucoma specialists, with a sensitivity, specificity, and overall correct classification of 91.7%, 95%, and 93.2% respectively (33).

One of the main problems with FO images is that the interpretation is operator-dependent and, therefore, lends itself to erroneous or different interpretations from professionals(36). Ophthalmologists with extensive clinical experience overestimate or underestimate glaucomatous damage when evaluating fundus photographs(37). Although standardization of classification methods can improve concordance, the dependence on having expert evaluators to interpret photographs is incompatible with a mass screening at the population level. Because of this problem, research and development began on different artificial intelligence algorithms that seek to evaluate all images under the same parameters and that have the capacity to learn to improve their diagnostic capacity over time, and offer an affordable, accurate, fast, and interpretable diagnosis in the future (37).

Worldwide trends in the medical analysis of diagnostic images have Deep Learning as their primary focus of interest. Different diagnostic algorithms have been validated for detecting diabetic retinopathy, age-related macular degeneration (AMD), and glaucoma, showing even better results than manual fundoscopic interpretation performed by healthcare professionals (32,38).

DL can provide predictions automatically using backpropagation algorithms, recognizing the intricate implicit structure in an extensive database (22). By performing proper training of the DL through prior exposure to FO images, it will be able to process and analyze the data more efficiently and give a more accurate diagnosis.

Early diagnosis of glaucoma is key to preventing blindness. In glaucoma, a series of morphological changes occur at the optic disc (OD) level in typical patterns as the disease progresses. Several studies on training DL algorithms from fundus images used a set of variables that allowed characterization of the OD, thus determining glaucoma risk. Evaluation of the optic nerve (NO) and the nerve fiber layer surrounding the OD is very important for accurate and early diagnosis of glaucoma, as these structural changes may even precede visual field (VC) losses (39). Anshul Thakur et al. conducted a prospective longitudinal study to evaluate the accuracy of using DL to predict the development of glaucoma from the fundus years before disease onset, finding accuracy of the Deep Learning model of 0.77 (95% confidence interval [CI], 0.75-0.79) and an accuracy of 0.95 (95% CI, 0.94 to 0.96) for glaucoma detection after disease onset. This study demonstrated how the DL allows us to predict preclinical signs of the disease, which can help us to complement other extended studies performed for glaucoma evaluation (40).

On the other hand, Alessandro A et al 2019 performed a cross-sectional study where they compared the M2M (machine-to-machine) DL algorithm capability with that of human evaluators to detect eyes with glaucomatous visual field loss, and demonstrated that the M2M DL algorithm predictions had a significantly stronger correlation with visual field metrics compared to human classifications, furthermore, that it could provide an objective and quantitative assessment of neural damage that would allow it to be used for glaucoma diagnosis and detection (37). Medeiros et al. proposed an alternative approach to train DL models for the evaluation of FO photographs in glaucoma, the M2M DL model, which was compared with spectral-domain optical coherence tomography (SDOCT) to discriminate the thickness of retinal nerve fiber layers of a glaucomatous eye from normal ones—finding similar performance between the two with ROC curve areas of 0.940 for the M2M DL model versus 0.944 for the SDOCT (P = 0.724). Thus, another advantage demonstrated by this algorithm was the ability to quantify the amount of neuronal loss, raising the possibility that fundus photographs can be used to detect changes over time in resource-poor locations where OCT is not available (41). Li et al., in their 2018 article, demonstrated the efficacy of Deep Learning in identifying glaucomatous optic neuropathy (GON) based on 48,116 fundus photographs. Such algorithm showed significant results, with a sensitivity of 95.6% and specificity of 92% for GON screening (26). While Yang et al., in their 2020 publication, evaluated the diagnostic accuracy of the Deep Learning ResNet-50 model to differentiate between glaucomatous optic neuropathy (GON) vs. non-glaucomatous optic neuropathy (NGON) images, they obtained a sensitivity of 93.4% and specificity of 81.8% (36). Al-Aswad et al., in their 2019 retrospective study, evaluated the diagnostic ability of DL (Pegasus) to classify FO images according to a binary classification ("glaucoma/non-glaucoma") and compared it with that of a group of 6 ophthalmologists; They observed that Pegasus obtained a sensitivity of 83.7% and a specificity of 88.2%, while the sensitivity of the ophthalmologists ranged from 61.3% to 81.6% and the specificity from 80.0% to 94.1% (5). On the other hand, Rogers et al., in their 2019 retrospective study, evaluated the diagnostic ability of a DL algorithm (Pegasus v1.0) for GON detection from 94 stereoscopic fundus images and compared it with 243 European ophthalmologists and 208 British optometrists. Pegasus was able to detect GON with an accuracy of 83.4% (95% CI 77.5-89.2), being comparable to the average accuracy of the ophthalmologist group 80.5% (95% CI 67.2-93.8) and the average accuracy of the optometrist group 80% (95% CI 67-88) for the same images and demonstrating better accuracy for Pegasus (not statistically significant) when compared to the other two groups (42). In the same year, Pahn et al. investigated the performance of CNNs for looking at discrimination using color fundus images. Their studies included images of glaucoma, suspected glaucoma, and non-glaucomatous eyes that included low-quality (depixelated) images. Their study showed that the area that most discriminates the model is the optic disc. The CNN model showed an area under the curve of 0.9 between eyes with glaucoma compared to non-glaucomatous eyes, as opposed to eyes with suspected glaucoma compared to non-glaucomatous eyes that showed an area under the curve of 0.7 (43).

Diaz-Pinto et al. in 2019 realized that it is widespread and valid to use CNN machine learning models for their high discrimination capability. However, they explored five models trained on ImageNet (VGG16, VGG19, InceptionV3, ResNet50, and Xception) and used 1707 images, yielding an AUC of 0.96, specificity of 0.85, and sensitivity of 0.93 using Xception, thus improving performance. A cross-validation strategy supports high specificity and sensitivity. Thus, models trained with ImageNet are an alternative to this automatic glaucoma detection system (44).

6 Conclusion

Deep Learning (DL), a branch of AI, is the evolution of machine learning within the field of AI, which uses layered algorithmic architectures to analyze data. The architecture is inspired by how biological neurons are interconnected to process information in the brain. Thanks to its evolution over the years, it has enabled earlier diagnosis of certain diseases, such as glaucoma.

In the case of this pathology, Deep Learning convolutional neural networks can outperform glaucoma specialists in detecting the disease based on fundus image data, achieving better diagnostic accuracy. With these findings, this technology opens several future possibilities for glaucoma screening, confirmation, and follow-up.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declared no conflict of interest.

Funding

This research did not receive any type of funding.

Author's contributions

All the authors enumerated, have direct and substantial contribution to the work and approved the version submitted for revision.

References

[1] Girard MJA, Schmetterer L. Artificial intelligence and deep learning in glaucoma: Current state and future prospects. En: Progress in Brain Research [Internet]. Elsevier;2020p.37-64. https://linkinghub.elsevier.com/retrieve/pii/S007961232030100X

[2] Phene S, Dunn RC, Hammel N, Liu Y, Krause J, Kitade N, et al. Deep Learning and Glaucoma Specialists. Ophthalmology. December of 2019;126(12):1627-39.

[3] Schuster AK, Erb C, Hoffmann EM, Dietlein T, Pfeiffer N. The Diagnosis and Treatment of Glaucoma. Dtsch Ärztebl Int [Internet]. 27 march 2020; https://www.aerzteblatt.de/10.3238/arztebl.2020.0225

[4] Barría, F., & Jiménez-Román, J. (2019). Latin American guide to primary open-angle glaucoma for the general ophthalmologist. International Council Ophthalmology, 1-49.

[5] Al-Aswad LA, Kapoor R, Chu CK, Walters S, Gong D, Garg A, et al. Evaluation of a Deep Learning System For Identifying Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. J Glaucoma. december 2019;28(12):1029-34.

[6] Covin, Y. N., Laroche, D. A. N. I. E. L., & Olivier, M. (2014). The societal costs of blindness from uncontrolled glaucoma. Glaucoma Today, 28-29.

[7] Sreng S, Maneerat N, Hamamoto K, Win KY. Deep Learning for Optic Disc Segmentation and Glaucoma Diagnosis on Retinal Images. Appl Sci. 17 july 2020;10(14):4916.

[8] Ahn JM, Kim S, Ahn KS, Cho SH, Lee KB, Kim US. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. Bhattacharya S, editor. PLOS ONE. 27 november 2018;13(11):e0207982.

[9] Weinreb RN, Aung T, Medeiros FA. The Pathophysiology and Treatment of Glaucoma: A Review. JAMA. 14 may 2014;311(18):1901.

[10] Díez, R. C., Román, J. J., & Barbosa, M. J. I. (2014). Concept of suspected open-angle glaucoma: definition, diagnosis, and treatment. Mexican Journal of Ophthalmology, 88(4), 153-160..

[11] Zhang N, Wang J, Chen B, Li Y, Jiang B. Prevalence of Primary Angle Closure Glaucoma in the Last 20 Years: A Meta-Analysis and Systematic Review. Front Med. 18 january 2021;7:624179.

[12] Dietze, J., Blair, K., & Havens, S. J. Continuing Education Activity.

[13] Tan NYQ, Friedman DS, Stalmans I, Ahmed IIK, Sng CCA. Glaucoma screening: where are we and where do we need to go? Curr Opin Ophthalmol. march 2020;31(2):91-100.

[14] Benet D, Pellicer-Valero OJ. Artificial intelligence: the unstoppable revolution in ophthalmology. Surv Ophthalmol. January 2022;67(1):252-70.

[15] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[16] Balyen, L., & Peto, T. Promising Artificial Intelligence-Machine Learning-Deep Learning Algorithms in Ophthalmology. 2019. 3(8):264-72.

[17] Van Assen M, Lee SJ, De Cecco CN. Artificial intelligence from A to Z: From neural network to legal framework. Eur J Radiol. August 2020;129:109083.

[18] Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, et al. A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography. Ophthalmology. september 2018;125(9):1410-20.

[19] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? Adv Neur In. 2014;27:3320–8. 47.

[20] Öhnell, H., Heijl, A., Anderson, H., & Bengtsson, B. Detection of glaucoma progression by perimetry and optic disc photography at different stages of the disease: results from the Early Manifest Glaucoma Trial. Acta ophthalmologica. 2017;95(3):281-7.

[21] Krizhevsky, Sutskever, & Hinton. ImageNet Classification with Deep Convolutional Neural Networks. University of Toronto. 2012;

[22] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 28 may 2015;521(7553):436-44.

[23] Bottou, L. & Bousquet, O. The tradeoffs of large scale learning. In Proc. Advances in Neural Information Processing Systems 20 161–168 (2007).

[24] Biswal, A. (2021). Top 10 deep learning algorithms You should Know in 2021. Simplilearn. Simplilearn. com/tutorials/deep-learning-tutorial/deep-learning-algorithm.

[25] Fradkov AL. Early History of Machine Learning. IFAC-Pap. 2020;53(2):1385-90.

[26] Das, S. CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more. 2019; https://medium.com/analytics-vidhya/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

[27] Lin,M.,Chen,Q.,&Yan,S. (2013). Network In Network. arXiv. https://doi.org/10.48550/arXiv.1312.4400

[28] Christian Szegedy; Wei Liu. Going Deeper with Convolutions. CVF. 2015;

[29] Bočková M, Veselý P, Synek S, Hanák L, Beneš P. Sensitivity and Specificity of Spectral Oct in Patients with Early Glaucoma. Czech Slovak Ophthalmol. 21 october 2019;75(5):260-4.

[30] Asaoka R, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. Am J Ophthalmol. 2019;198:136–45.

[31] Muhammad H, Fuchs TJ, De Cuir N, De Moraes CG, Blumberg DM, Liebmann JM, et al. Hybrid deep learning on single wide- field optical coherence tomography scans accurately classifies glaucoma suspects. J Glaucoma. 2017;1086–94.

[32] Thompson AC, Jammal AA, Medeiros FA. A deep learning algorithm to quantify neuroretinal rim loss from optic disc photographs. Am J Ophthalmol. 2019;201:9–18.

[33] Reus NJ, de Graaf M, Lemij HG. Accuracy of GDx VCC, HRT I, and clinical assessment of stereoscopic optic nerve head photographs for diagnosing glaucoma. Br J Ophthalmol. 1 march 2007;91(3):313-8.

[34] Pablo LE, Ferreras A, Fogagnolo P, Figus M, Pajarin AB. Optic nerve head changes in early glaucoma: a comparison between stereophotography and Heidelberg retina tomography. Eye. January 2010;24(1):123-30.

[35] Saito H, Tsutsumi T, Iwase A, Tomidokoro A, Araie M. Correlation of Disc Morphology Quantified on Stereophotographs to Results by Heidelberg Retina Tomograph II, GDx Variable Corneal Compensation, and Visual Field Tests. Ophthalmology. February 2010;117(2):282-9.

[36] Yang HK, Kim YJ, Sung JY, Kim DH, Kim KG, Hwang JM. Efficacy for Differentiating Nonglaucomatous Versus Glaucomatous Optic Neuropathy Using Deep Learning Systems. Am J Ophthalmol. August 2020;216:140-6.

[37] Jammal AA, Thompson AC, Mariottoni EB, Berchuck SI, Urata CN, Estrela T, et al. Human Versus Machine: Comparing a Deep Learning Algorithm to Human Gradings for Detecting Glaucoma on Fundus Photographs. Am J Ophthalmol. March 2020;211:123-31.

[38] Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology. July 2017;124(7):962-9.

[39] Shibata N, Tanito M, Mitsuhashi K, Fujino Y, Matsuura M, Murata H, et al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. Sci Rep. december 2018;8(1):14665.

[40] Thakur A, Goldbaum M, Yousefi S. Predicting Glaucoma before Onset Using Deep Learning. Ophthalmol Glaucoma. july 2020;3(4):262-8.

[41] Medeiros FA, Jammal AA, Thompson AC. From Machine to Machine. Ophthalmology. April of 2019;126(4):513-21.

[42] Rogers TW, Jaccard N, Carbonaro F, Lemij HG, Vermeer KA, Reus NJ, et al. Evaluation of an AI system for the automated detection of glaucoma from stereoscopic optic disc photographs: the European Optic Disc Assessment Study. Eye. November 2019;33(11):1791-7.

[43] Phan, S., Satoh, S., Yoda, Y., Kashiwagi, K., Oshika, T., & Japan Ocular Imaging Registry Research Group. Evaluation of deep convolutional neural networks for glaucoma detection. Japanese journal of ophthalmology. Japanese journal of ophthalmology. 2019;(63(3)):276-83.

[44] Vessani RM, Moritz R, Batis L, Zagui RB, Bernardoni S, Susanna R. Comparison of Quantitative Imaging Devices and Subjective Optic Nerve Head Assessment by General Ophthalmologists to Differentiate Normal From Glaucomatous Eyes: J Glaucoma. march 2009;18(3):253-61.