Review Article - Narrative Review

Taiwan J Ophthalmol 2024;14:473-485

Access this article online

Quick Response Code:

Website:
http://journals.lww.com/TJOP

DOI:
10.4103/tjo.TJO-D-24-00064

Discriminative, generative artificial intelligence, and foundation models in retina imaging

Paisan Ruamviboonsuk[1]*, Niracha Arjkongharn[1], Nattaporn Vongsa[1], Pawin Pakaymaskul[1], Natsuda Kaothanthong[2]

Abstract:

Recent advances of artificial intelligence (AI) in retinal imaging found its application in two major categories: discriminative and generative AI. For discriminative tasks, conventional convolutional neural networks (CNNs) are still major AI techniques. Vision transformers (ViT), inspired by the transformer architecture in natural language processing, has emerged as useful techniques for discriminating retinal images. ViT can attain excellent results when pretrained at sufficient scale and transferred to specific tasks with fewer images, compared to conventional CNN. Many studies found better performance of ViT, compared to CNN, for common tasks such as diabetic retinopathy screening on color fundus photographs (CFP) and segmentation of retinal fluid on optical coherence tomography (OCT) images. Generative Adversarial Network (GAN) is the main AI technique in generative AI in retinal imaging. Novel images generated by GAN can be applied for training AI models in imbalanced or inadequate datasets. Foundation models are also recent advances in retinal imaging. They are pretrained with huge datasets, such as millions of CFP and OCT images and fine-tuned for downstream tasks with much smaller datasets. A foundation model, RETFound, which was self-supervised and found to discriminate many eye and systemic diseases better than supervised models. Large language models are foundation models that may be applied for text-related tasks, like reports of retinal angiography. Whereas AI technology moves forward fast, real-world use of AI models moves slowly, making the gap between development and deployment even wider. Strong evidence showing AI models can prevent visual loss may be required to close this gap.

Keywords:

Discriminative artificial intelligence, foundation models, generative artificial intelligence, retinal imaging, vision transformer

Introduction

Recent advances of artificial intelligence (AI) in retinal imaging found its application in two major categories: discriminative and generative AI.[1] This classification is practically based on the task of an AI model: discrimination of contents or generation of new contents.

In general, discriminative AI relies on labeled trained data tailored for specific tasks (supervised learning approach),[2,3] such as the classification of severity of diabetic retinopathy (DR)[4] or the prediction of referrable DR from retinal images,[5] without novel content generation. It learns the decision boundary to classify the existing data. Support Vector Machine, Decision Tree, or Logistic Regression are the common AI techniques for training on labeled data to perform discriminative tasks.[2,3]

Generative AI, on the other hand, aims to create new data, such as reports of fluorescein angiography (FA)[6] or generating indocyanine green angiography images from color retinal photographs.[7] The

[1]*Department of Ophthalmology, College of Medicine, Rangsit University,* [2]*Sirindhorn International Institute of Technology, Thammasat University, Bangkok, Thailand*

***Address for correspondence:**
Prof. Paisan Ruamviboonsuk, Department of Ophthalmology, College of Medicine, Rangsit University, Bangkok, Thailand.
E-mail: paisan.trs@gmail.com

Submission: 06-06-2024
Accepted: 07-08-2024
Published: 28-11-2024

How to cite this article: Ruamviboonsuk P, Arjkongharn N, Vongsa N, Pakaymaskul P, Kaothanthong N. Discriminative, generative artificial intelligence, and foundation models in retina imaging. Taiwan J Ophthalmol 2024;14:473-85.

task can be done with minimal, or without, labeled trained data (semi-supervised or unsupervised learning approach).[2,3] Generative Adversarial Network (GAN) is a common generative technique applied in retinal imaging.[8] Convolutional neural network (CNN),[9] although well known for discrimination, can be set up for content generation. The difference between discriminative and generative AI is summarized in Table 1.

Another term, foundation models, is recently well recognized due to the popularity of Chat Generative Pre-Trained Transformer (ChatGPT). The term was established by Stanford Institute for Human-Centered AI.[10] A foundation model is trained on board data to be used in a wide range of applications.[11] Some foundation models are built in specific fields, such as astronomy,[12] radiology,[13] and retinal imaging.[14] ChatGPT, built on Generative Pre-trained Transformer (GPT) technique,[15] is considered a large language model (LLM), a category of foundation models which has been trained on billions of web pages [Figure 1]. A foundation model can be adapted (e.g., fine-tuned) for a wide range of downstream tasks,[16] either discrimination or generation.

In this narrative review, we focused on the recent advances of discriminative AI, generative AI, and foundation models in retinal imaging.

Methods

We conducted a systematic search following the Preferred Reporting Items for Systematic reviews and Meta-Analyses guideline from the databases of PubMed and Scopus and included mainly recent articles published in the past 6 months, from November 1, 2023, to April 30, 2024. The search flow diagram is in Figure 2.

The search terms included retinal imaging, DR, age-related macular degeneration, retinal vein occlusion (RVO), retinopathy of prematurity, AI, LLM, foundation model, GAN, and vision transformer (ViT).

Our inclusion criteria were original peer-reviewed publications and papers in computer engineering in the English language. The exclusion criteria were conference abstract, short commentary, and no full text availability. Three authors independently screened papers and collected data. Articles on oculomics were not included.

Discriminative Artificial Intelligence

Convolutional neural network

Color fundus photographs

Diabetic retinopathy

CNN models for the discrimination of referable or sight-threatening DR have been available for some time. Many researchers still explored and investigated various CNN techniques to improve performances and address the challenges of the models. A study proposed CNNs to address two primary challenges, (1) insensitivity to minority classes due to imbalanced data distribution and (2) neglecting the relationship between the left and right eyes by utilizing the fundus image of only one eye for training in DR datasets.[17]

CNN models for predicting the progression of DR generally have the accuracy of the prediction around 0.7–0.8 using color fundus photographs (CFPs) alone. The accuracy is somewhat improved by adding risk factors into the models; however, the accuracy of this prediction is still lower than that for detecting referable DR.[5]

A recent DR prediction model was developed using a large dataset of more than 700,000 CFPs and validated from more than 100,000 CFPs of multiethnic datasets. The 5-year prediction of DR progression from this model had the area under the receiving operator characteristics curve (AUROC) in the range of 0.72–0.86. The authors suggested that this model could potentially extend the mean screening interval from 12 to 31.97 months when integrated into clinical workflows.[18]

Another recent study prospectively implemented a DR prediction model in a clinical workflow for prioritizing patients to be screened for DR in their next screening visits. The patients with no DR and mild nonproliferative DR (NPDR) were then called for the next visits according

Table 1: The difference between discriminative and generative artificial intelligence

Factors	Discriminative AI	Generative AI
Objective	Aim to model the decision boundary of the classes in the dataset	Capture the actual distribution of the classes in the dataset
Data dependency	Depends on labeled and structured data	Capable to handle unstructured data
Training	Learn the conditional probability - $P(y\|x)$	Predict the joint probability distribution - $P(x, y)$
Application	Classification or regression based on existing data	Generating novel or innovative contents
Computational complexity	Computationally cheaper	Computationally more expensive
Learning approach	Useful for supervised learning	Useful for unsupervised learning
Strengths	Robust to outliers	Effectively deal with uncertainty and changing or scarce data. More creativity, prompt and flexible

AI=Artificial intelligence

to the rank provided by the prediction model, whereas the ground truth was based on the examination of DR grades in the next screening visits in the clinic. Many ranking strategies were compared. The outcome measure was the sensitivity for detecting referrals, defined as moderate NPDR or worse, based on the ground truth, in the first 50% of patients in the next visits in each ranking strategy. The rank proposed by the model had the highest sensitivity of 90.38%. The rank by the DR grades and HbA1C levels from the previous visits had the sensitivity of 86.54%, whereas the rank by the DR grades from the previous visits only had the sensitivity of 68.95%. The sensitivity based on random visits was 50%.[19]

Other diseases

For age-related macular degeneration (AMD), an ML-based computer-aided diagnosis framework was developed to classify CFPs as normal, intermediate AMD, geographic atrophy (GA), and wet AMD, based on the extraction of both global and local markers on the CFPs. This study employed many ML classifiers and applied weighted majority voting on the best classifiers to improve the performance, resulting in an accuracy of 96.85%, sensitivity of 93.72%, and specificity of 97.89% for the classification.[20] In another recent study, EfficientNet_b2 model and GradCAM were applied to train and explain an AI tool in differentiating normal and GA on CFPs.[21]

A recent study on AI for retinopathy of prematurity (ROP) screening using CFPs (a model from the Imaging and Informatics in ROP Study) was in the phase of validation on the external multinational datasets. The investigators showed that the model, with further training and calibration, could achieve AUROC of 83.5% and 82.2% for more-than-mild ROP (mtmROP) and type 1 ROP, respectively, for the dataset from the Stanford University Network for Diagnosis of ROP which had >76,000 images. The validation was also conducted on another dataset from Aravind Eye Hospital with >69,000 images, the AUROC of 80.8% and 87.8% for mtmROP and type 1 ROP, was achieved respectively. These performances suggested the readiness for the deployment of the model.[22]

Although the pandemic of myopia is an issue of interest worldwide, there have not yet been too many AI models for detecting myopic maculopathy (MM) on CFPs. A recent study applied EfficientNet to develop an AI model, which achieved sensitivities of 96.86%, 75.98%, 64.67%, and 88.75% for classifying tessellated fundus, diffuse chorioretinal atrophy, patchy chorioretinal atrophy, and macular atrophy, according to the International Photographic Classification of MM,[23] whereas the specificity for each severity level was higher than 93%.[24]

Optical coherence tomography

For optical coherence tomography (OCT) images, there have been more discriminative AI studies for AMD than DR. Recent publications of discriminative models on OCT images for either DR or diabetic macular edema (DME) were still on improving the classification of DR,[25] segmentation of macular area in DME,[26] and prediction of response to anti-VEGF medications in DME.[27]

For AMD, real-world OCT scans were used for training and testing an AI classifier to classify the entire OCT volume as normal, intermediate AMD, GA, or wet AMD. Initially, a 2D ResNet50 was trained to identify the disease category on an individual OCT B-scan, then, four smaller ResNet models were trained to use the concatenated B-scan-wise output from the initial stage to classify the entire OCT volume.[28] Another recent study demonstrated high accuracy for the segmentation of retinal fluid and sub-retinal pigment epithelium fluid by DL.[29]

In another segmentation task on OCT scans, a previously published DL for detecting retinal fluid in AMD, DME, and RVO,[30] was further trained to quantify macular fluid volumes at baseline and under therapy of patients with neovascular AMD (nAMD); high concordance of more than 0.83–0.95 of AUROC between the DL model and human experts was found for both intraretinal and subretinal fluid.[31] For predictive task, a CNN model was found to be able to predict the progression from intermediate AMD to GA within a month providing the AUROC of 0.94, sensitivity of 91%, and specificity of 80%.[32] Whereas the prediction of recurrence of nAMD within 3 months after 1-month dry-up and the prediction of the first recurrence after three loading doses achieved the accuracy around 60%–70%, respectively.[33,34]

There were attempts to train CNN to classify epiretinal membrane (ERM). In the first paper, ERM was classified by a model as no ERM, small ERM, large ERM, with AUROC more than 0.9 in each category.[35] In the second paper, ERM was classified as normal and five severity levels; the accuracy for classifying normal and stage 3 severity was 98% and 84%, respectively, whereas the overall accuracy was 81.3%.[36]

Another self-supervised model fine-tuned with labeled data for macular telangiectasia type II (MacTel Type II), achieved the area under precision-recall curve (AUPRC) of 0.971, AUROC of 0.970, accuracy of 0.898%, sensitivity of 0.898, and specificity of 0.949 for detecting the disease from OCT scans.[37] Another group of investigators used CNN to classify the severity of MacTel Type II, originally proposed as 7-class scheme[38] into a uniform continuous scale. Kappa statistics for agreement between the model

and ophthalmologists for grading according to this scale were 0.56–0.63.[39]

Other imaging

Other recent publications on discriminative AI for other retinal image domains included optical coherence tomography (OCTA) and ultra-widefield (UWF) images. EfficientNet was used for classifying normal, dry, active wet AMD, and remitted wet AMD in OCTA.[40] For UWF, ResNet50 was recently applied for grading the severity of DR,[41] and for detecting early peripheral retinal degeneration, hyperpigmentation, and white without pressure areas.[42] EfficientNet-b was used for detecting retinal breaks and retinal detachment in UWF.[43]

A ResNet101 model was used for differentiation between extensive macular atrophy and pseudodrusen-like appearance,[44] a rare clinical entity with marked absence of choriocapillaris flow on OCTA,[45] and macular atrophy in AMD on fundus autofluorescence images.[46] Another study found VGG-16 provided the best performance for detecting rhegmatogenous retinal detachment from ophthalmic ultrasound images.[47]

For fundus FA, there was an interesting paper applied three different CNN models to perform automated standardized labeling on FA images. The predicted labels included image quality, lesion location, laterality of eye, phases of the angiography, and five lesions identification (microaneurysms, nonperfusion [NP] areas, leakage, laser scar, and hemorrhage). ResNet18 was identified as the best performance among the three models.[48] In another paper for UWF fluorescein angiography (UWFA), DeepLab v3+ networks (based on ResNet-18, ResNet-50, Xception, InceptionResNet-v2, and Inception-v3) and SegNet networks (based on VGG-16 and VGG-19) were trained to detect retinal capillary NP and neovascularization (NV) from UWFA of patients with severe NPDR and proliferative DR (PDR). DeepLab v3+ gave the best performance, whereas the highest accuracy for detecting NP and NV was 0.8208 and 0.8338, respectively.[49]

A study for multimodal retinal imaging (CFP, OCT, and FA) trained five ML models (support vector machine, random forest, extreme gradient boosting, multilayer perceptron neural network, and lasso) for prediction of 2-year visual response to anti-VEGF treatments (ranibizumab and bevacizumab) in nAMD. The dataset was from the comparison of Age-related Macular Degeneration Treatments Trials.[50] In average, the models gave an R^2 of 0.33–0.38 for predicting visual acuity (VA) change, an R^2 of 0.37–0.45 for predicting actual VA at 2 years, and AUROCs of 0.85–0.87 and 0.67–0.79 for predicting 15-letter visual gain and 15-letter visual loss, respectively.[51]

Vision transformer

Transformers are ML architectures developed for Natural Language Processing (NLP).[52] They learn context and track relationships between the sequence components of words in a sentence.[53] When applied to recognize images, an image is split into patches, the sequence of linear embeddings is provided to these patches as inputs to a transformer. These image "patches" are then treated the same way as "words" in NLP[54] [Figure 3]. This so-called ViT can be pretrained on large amounts of data and transferred to mid-sized or small image recognition benchmarks while requiring substantially fewer computational resources for training.[54] ViT has recently played a dominant role in medical imaging.[55]

Vision transformer for color fundus photographs

Diabetic retinopathy

In one of the first studies on ViT for DR grading on CFPs, ViT was compared with classic CNN models, such as ResNet50 and ResNet101, and shown a better performance with an accuracy of 91.4%, whereas the sensitivity was 0.926 and specificity was 0.977.[56] The authors suggested that ViT might replace the conventional CNN in the future. Another study showed that ViT, pretrained with Masked Autoencoders with 100,000 publicly available retinal images, could outperform conventional CNN, pretrained with much larger datasets with weights from ImageNet for grading DR severity, despite the smaller training datasets.[57]

Another study used ViT and residual attention to classify DR in the five severity scales, the performance for the four DR levels, except mild NPDR, was better than five conventional CNNs: VGG-16, ResNet-18, GoogLeNet, DenseNet-121, and SE-BN-Inception, despite the fact that the two open-sourced training and testing datasets[58,59] had imbalanced and limited labeled data.[60] Another study also aimed to overcome the imbalanced classes in a DR dataset by training ViT with a novel category attention block that enhanced feature information within a DR class, and a lesion relation attention block that captured relationships between lesions.[61] This method achieved better performance than CNNs: ResNet-50, MobileNet 1.0, Xception and Inception V3.

Saliency-guided Self-supervised Image Transformer (SSiT), which is another ViT with self-supervised learning (SSL) pre-trained without annotation outperformed other SSLs at least 9.48% with a high kappa score of 81.88% for DR classification. This SSiT model can be generalized for AMD detection and segmentation for pathologic myopia.[62] Generalizability of this type of model was demonstrated in another study when a Symmetric Mask Pre-training Vision Transformer model, initially trained to classify histological findings in colon

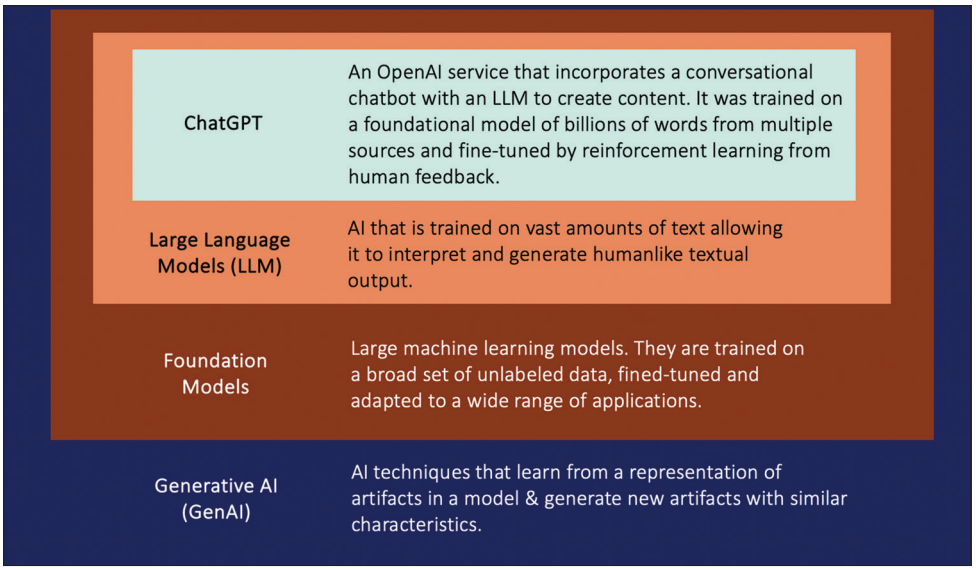

Figure 1: Generative artificial intelligence and its derivatives. LLM: Large language models, ChatGPT: Chat Generative Pretrained Transformer, AI: Artificial intelligence

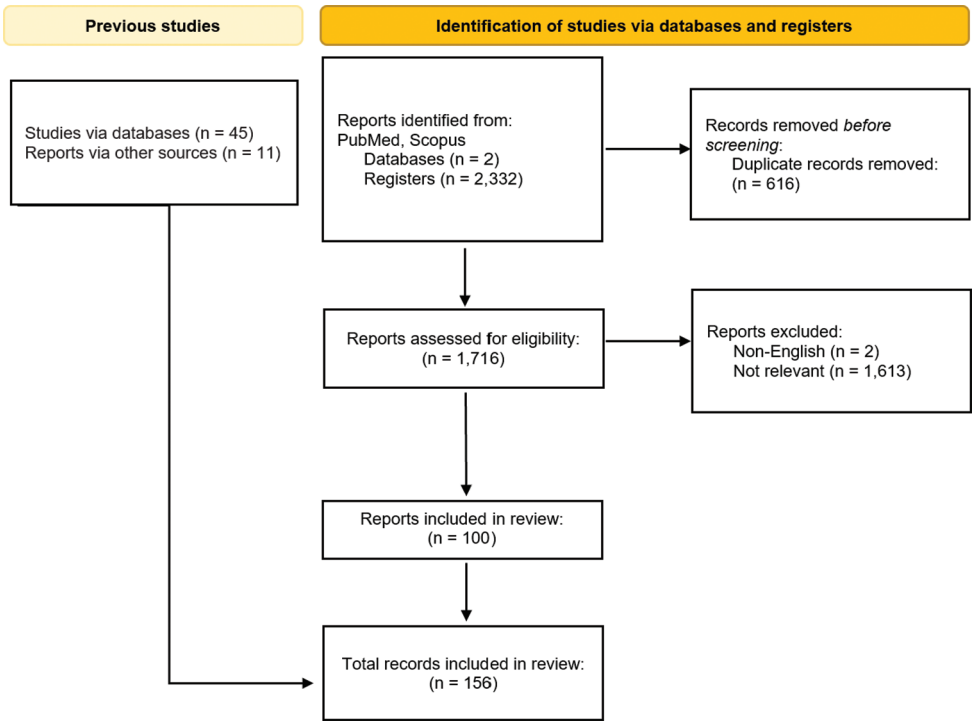

Figure 2: Preferred Reporting Items for Systematic reviews and Meta-Analyses flow diagram

cancer, could be generalized to classify DR severity in CFPs with F-1 scores of 86.91% and 72.85%, for colon cancer and DR, respectively.[63]

Segmentation in CFPs to identify various lesions of DR is another task for ViT and segmentation of hard exudates was found to achieve the highest accuracy among the lesions in a study.[64] Another study incorporated hyperbolic embeddings and a spatial prior module in ViT for the segmentation. The hyperbolic embeddings were used to classify feature matrices in CFPs at the pixel level whereas the spatial prior module was for image convolution and feature continuity.[65]

Age-related macular degeneration and other diseases

For AMD, a hierarchical ViT-based model that integrated data augmentation techniques and SwinTransformer was used to classify dry and wet AMD, and wet AMD into type I and type II macular NV in CFPs. The AUROC for the latter classification could reach 99.36%.[66] On the other hand, a head-to-head comparison between

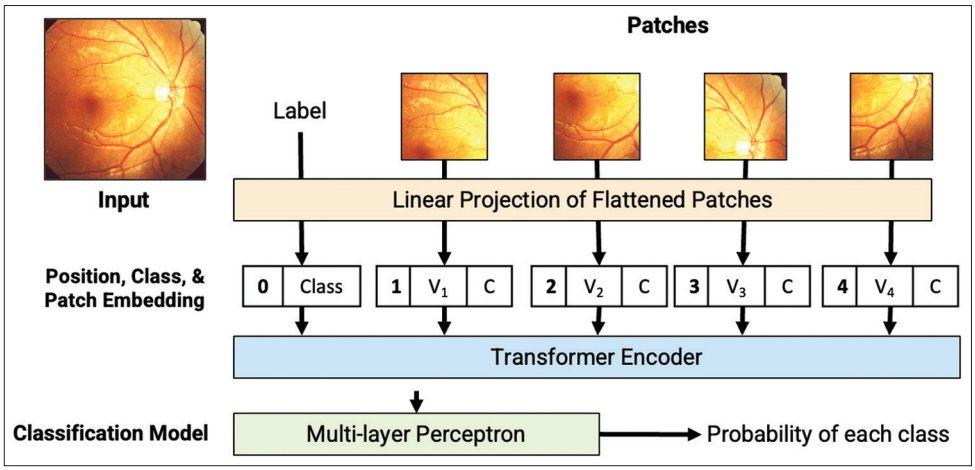

Figure 3: Architecture of vision transformer

8 CNN models and 9 ViT models to classify referrable and non- referrable AMD in CFPs found all the CNN models performed better than all the ViT models, with sensitivity and specificity of the CNN models reaching 90% or more, whereas the ViT models could achieve 63%–94% sensitivity and 24%–48% specificity. The multi-classed classification of the severity of AMD (no AMD, early AMD, intermediate AMD, and advanced AMD) by the CNN and ViT models in the same study found similar trends. The authors concluded that the performances of the CNNs were boosted by test-time augmentation for both the binary and multi-class models.[67]

Vision transformer for optical coherence tomography

ViT was trained to classify diabetic maculopathy as early DME, advanced DME, severe DME, and atrophic maculopathy, in OCT scans achieving an AUROC of 0.96 and the accuracy of 90% or more.[68] Another ViT model was trained for segmenting and classifying drusen and double layer sign in macular OCT scans. This model achieved 82% and 90% sensitivity and specificity, respectively, when tested on a separate dataset, whereas the agreement with senior human graders provided a kappa score of 0.83.[69]

Other studies demonstrated that a hybrid fusion model of ViT and CNN could achieve better performance than an end-to-end model of ViT or CNN alone for feature extractions or classifying drusen, choroidal neovascularization (CNV), and DME in OCT scans.[70-72] One of these studies could extend the classification by the hybrid model to other diseases in the new dataset, such as drusen, macular hole (MH), DR, central serous chorioretinopathy (CSC), and normal with the average accuracy of 97.11%.[72] These hybrid models took advantages of better extraction of local features, such as textures and shapes, by CNN, which may be more useful in OCT scans.

Another group of investigators proposed a novel Structure-Oriented Transformer (SoT) framework to construct the relationship between lesions and retina in OCT scans of a clinical dataset of nAMD in which they were able to effectively classify MNV into type I and type II. The same model was able to classify DME and early AMD in another dataset.[73] Another ViT technique, multiscale Model-based Transformer (MBT), was proposed for the classification of multiple diseases: normal, MH, AMD, CSC, DR, DME, and ERM, in OCT videos. The MBT outperformed several conventional state-of-the-art CNNs.[74]

Segmentation of fluid in OCT scans was also conducted using ViT in fusion with CNN in many recent studies. These included segmentation of retinal layers in healthy individuals and DME,[75] segmentation of retinal fluid[76] to create fluid score,[77] and segmentation of pigment epithelial detachment.[78] Another study applied ViT for assisting in OCT image acquisition and registration to reduce misalignment and improve generalizability in different devices in different modes similar to brain MRI.[79]

Vision transformer for ultra-widefield and multimodal imaging

Other retinal image domains for ViT included UWF in which a study found ViT outperformed most basic ResNet50 models for DR detection.[80] Another study employed ViT for classifying UWF images as normal or abnormal, which included AMD, DR, ERM, and RVO, whereas some images contained more than one disease. ResNet152, was found to perform better than ViT and other CNNs with the AUROC of 96.47% for classifying normal and the retinal diseases in this study.[81] Another study found ViT to detect retinal tears from B-scan ultrasound images with accuracy of 83.8%, sensitivity of 82.7%, and specificity of 82.4%.[82]

A study performed retinal vessel segmentation in CFPs applying a hybrid CNN and ViT trained with data from

both CFPs and OCTA. The training with CFPs and tested with OCTA yielded the highest accuracy of 92.51%, whereas the training with CFPs and tested with both CFPs and OCTA, and the training with OCTA and tested with CFPs, had the accuracy >90%.[83]

Another study compared three CNN models and eight ViT models for classifying DR in both CFPs and OCT scans. The investigators not only found that the performance of ViT exceeded CNN in certain cases but also found that the attributable heatmaps generated by ViT obtained higher clinical acceptance by ophthalmologists.[84]

Generative Artificial Intelligence

Large language model

The public launch of ChatGPT (OpenAI, San Francisco, CA, USA) on November 30, 2022, was considered another milestone in information technology revolution.[85] Within a year, a simple search using the term "ChatGPT" in PubMed database yielded more than 2000 results. In the year 2023, there were already at least four major review articles on ChatGPT in ophthalmology. ChatGPT can be used for advanced data analysis since it is equipped with Data Analyst features.[86]

A recent study assessed the use of ChatGPT to assist in developing AI models for DR detection and DR severity classification without coding requirement.[87] The authors in this paper demonstrated how they queried the Chatbot step-by-step for CFP preprocessing; the processed CFPs were then classified into two datasets, one for DR detection and another for DR severity classification, for training two AI models using online automated ML platform via VertexAI platform in Google Cloud. ChatGPT was also used for generating computer language scripts, and analyzing data to generate diagnostic parameters, such as sensitivity and specificity. Trained on only 1700 CPFs from Messidor-2 database, the first model for the severity of DR achieved an AUPRC of 0.81, a precision of 81.81% and recall of 72.83%. The second model for the detecting DR achieved a precision and recall of 84.48%, and an AUPRC of 0.90. Limitations of ChatGPT for coding included inconsistent and errored scripts. If programming errors occurred when running the scripts, a query to ChatGPT to request a revised script was still possible.[87]

The typical use of ChatGPT was for consultations. These included recommendation for DR screening,[88] management of DME,[89] intravitreal injections for nAMD,[90] general queries for common retinal diseases,[91] such as DR, AMD, RVO, CSC, and vitreoretinal surgery.[92] The agreements with clinicians in these papers varied from fair to substantial depending on research methods and outcome measures. Another study assessed the diagnosis of six common retinal diseases generated from the input of Chinese FFA reports through ChatGPT using Chinese Prompt and English Prompt. The authors found that the English prompt performed better than Chinese prompt in terms of reasoning and diagnoses.[6] Hallucination was found frequently and should be aware for ChatGPT.

Generative adversarial network

GAN consists of two competing types of CNN: a generator and a discriminator for reinforcement. The generator and the discriminator compete during training to generate more authentic new data from a given training dataset. GAN demonstrated remarkable performance in image synthesis and image-to-image translation. A comprehensive review published in early 2022 found the most common tasks of GAN in ophthalmology were segmentation, data augmentation, denoising, domain transfer, super-resolution, postintervention prediction, and feature extraction. The four most common ocular imaging domains for GAN were CFPs, macular OCT, retinal angiography, and UWF.[93]

Generative adversarial network for color fundus photographs

Many recent studies of GAN in retinal imaging still focused on CFPs and DR, for which one of the tasks was high-resolution CFPs synthesis of severe NPDR and PDR since these severity levels are scarce in general training datasets.[94] In another study, to overcome the similar problem of imbalanced data, the investigators proposed a novel framework, Class-Imbalanced Semi-Supervised Learning-GANs, by leveraging a dynamic class-rebalancing sampler. This framework exploited the property that the classifier trained on class-imbalanced data to produce high-precision pseudo-labels on minority classes to leverage the bias inherent in pseudo-labels.[95]

In another study, the investigators developed a model integrated ResNet50 with UGAN (GAN for feature extraction), and convolutional block attention module to enhance the performance of the ResNet50 to screen DR at a community level.[96] Another study trained GAN for segmentation of subretinal fluid on the 2-D CFPs in CSC.[97] Another study trained GAN to mark drusen on CFPs for AMD detection,[98] whereas enhancing the detection of ERM on CPFs was also conducted using GAN.[99]

Another study developed 16 AI models to discriminate referable DR in CFPs, 12 were GANs and 4 were CNNs. The training dataset in this study contained only retinal images of nonreferrable DR, whereas the testing was conducted on the images of both nonreferrable and

referrable DR. The best performance across all the models was from InfoStyleGAN, a self-supervised network, with an AUROC of 0.808.[100] Methods of synthesizing CFPs using GAN for training models for discriminating ROP were also investigated.[101]

Generative adversarial network for optical coherence tomography

GAN was also used to address a problem of imbalanced dataset for OCT. The models trained on the synthesis-balanced public dataset were found to outperform the models trained on the unbalanced public dataset in the training set, validation set, fivefold cross validation, and external test set for classifying normal, drusen, DME, and CNV.[102] Another study also showed that performances of the model trained with the mixed synthetic and authentic OCT scans for classifying normal, AMD, DME, was better than that trained with only authentic images; the accuracy, sensitivity, and specificity were improved by 5.56%, 8.89%, and 2.22%, respectively.[103]

Other studies used GAN to predict treatment outcomes, such as generating posttreatment macular OCT scans from pretreatment scans, after three loading doses of anti-VEGF treatment in nAMD.[104] Another study used pretreatment OCT scans to generate longitudinal evolution of the macula in nAMD after anti-VEGF treatments.[105]

GAN was also used for generating color-coded macular OCT scans from corresponding FA images of DME and vice versa.[106] A newly proposed segmentation strategy called a Dual Stream Segmentation network was studied for embedding into a conditional GAN to improve the accuracy of retinal lesion segmentation.[107]

Techniques other than GAN that may be used in generative AI include Variational Autoencoder (VAE) and Latent Diffusion Model. Compared to GAN, they are used much less frequently in retinal imaging. A recent paper applied VAE to generate macular OCT scans after full-thickness MH surgery based on preoperative scans.[108] In another recent paper, VAE-based model was used to determine clinically relevant latent spaces for retinal disease diagnosis, particularly wet AMD, CSC, and polypoidal choroidal vasculopathy. The input to this model was a patient profile vector containing clinical examination findings and demographic information.[109]

Foundation Models

The first paper demonstrating the application of a foundation model in ophthalmology was published in 2023,[14] 7 years after the breakthrough performances of CNN in detecting referable DR were published.[110,111] The foundation model in this paper, RETFound, was pretrained with 1.6 million retinal images of CFPs and OCT, then fine-tuned and validated in many small datasets for specific tasks, such as detecting referrable DR, AMD, glaucoma, conversion to nAMD, and predicting systemic diseases. This self-supervised model was found to outperform traditional supervised and self-supervised models pretrained on ImageNet datasets for the same tasks in both internal and external validations.[14] Another study on RETFound, validated in real-world retinal images, also found a robust performance.[112] The codes and datasets of this model are available online.[113]

In another study, a large pretrained text-to-image foundation model, SD V.1.4,[114] was fine-tuned on UWF using DreamBooth[115] to produce novel, synthetic UWFs to train trainee orthoptists to diagnose common retinal diseases in a web-based course. After finishing the course, the students significantly improved their diagnostic accuracy from 43.6% to 74.1%.[116]

Another emerging foundation model is the open-sourced Segment Anything Model (SAM) developed by Meta. SAM has been trained on a huge number of images (11 million) and masks (1 billion), and built on an architecture that contained image encoder, prompt encoder, and mask decoder. The aim of SAM is to perform segmentation on any kinds of images by anybody without training.[117] SAM has the potential to be very useful for medical image segmentation. An early work on SAM in retinal imaging found its use as a basis of GlanceSeg, an AI framework which enabled ophthalmologists to detect inconspicuous or minute microaneurysms (MAs) in real time. This framework integrated the gaze maps for rough localization of the MAs by ophthalmologists and saliency maps generated based on the located region of interest. These maps provided prompt points to assist the model to efficiently segment the MAs. The human-in-the-loop experiment of GlanceSeg on two public datasets demonstrated its feasibility and superiority through visualized illustrations and quantitative measures. This AI framework improved annotation efficiency for ophthalmologists and further enhanced segmentation performance on small lesions on CFPs for detecting early DR in real time.[118]

Challenges of Artificial Intelligence in Retinal Imaging

As of the end of April 2024, there have been only three ophthalmic AI devices, IDx-DR, EyeArt, AEYE-DS, approved by the U.S. Food and Drug Administration (FDA). Considering the first AI model approved was in 2021, this means there is an approval rate of a model a year. All these models were approved

for DR screening without human supervision and approved for specific cameras. Steps beyond approval would be deliveries of the AI models for patient care in the real world. A comprehensive review on prospective application of AI models for DR screening published in 2023[119] revealed only eight publications,[120-127] and some of these might not be prospective implementation in the real world. Whereas the number of studies on new techniques of AI in retinal imaging keeps increasing exponentially in recent years, the adoption of AI in ophthalmic care, not specifically for retinal patients, increases at a much lower rate in comparison. This means the gap between development and deployment is widening.

If DR screening is a case study, there are at least four challenging areas in deployment of AI: (1) the lack of more head-to-head comparisons of the available models,[128] (2) no clear evidence of cost-effectiveness of AI compared to human screeners although many studies suggested the trend towards more cost-effectiveness of AI,[129] (3) equity and bias issues, and (4) medicolegal considerations. To leverage on the equity of AI in health care, the AI tools should be inclusive and accessed by people disregard for population, socioeconomic status, and geographical area. The data for training and testing of AI tools should be representative of the population they serve; the reports of their performances should follow the standard guidelines. A robust regulatory framework is pivotal to ensure AI tools adhere to rigorous equity standards.[130] This framework should not include only FDA but also national or international bodies of policy makers who are driving health-care forward. One of the main concerns in AI deployment rests on a possibility that AI may enhance, instead of addressing, health disparities.[131] In medicolegal standpoint, it is still unclear how responsibility and liability will be shared among clinicians, developers, model sellers, and regulating bodies if there is a serious misdiagnosis made by AI.[119]

Democratizing AI may be an answer to some of these challenges.[132] However, the term is complicated and may require interpretation. The simpler term with a similar meaning may be "general increase the accessibility of AI"[132] and this involves increasing AI development, use, profits, and most important, increasing accessibility to AI governance.[133]

In a bigger picture, a real challenge of AI in ophthalmology may lie not on the aspect of engineering research but on the aspect of clinical research in proving that AI models can prevent visual loss. At presence, research on AI is trying to prove that AI is on par with human experts in various tasks; very few uses vision as an outcome measure. In DR screening, for example, well-designed clinical research proving that populations screened with AI are less likely to have visual loss compared with manual screening is essential. Such research may finally be able to close the gap between development and deployment.

Conclusions

Many recent studies on AI in retinal imaging still focused on enhancing performances of discriminative AI using different techniques, with DR being the primary disease and CFPs and OCT being the image domains of choice. ViTs have emerged as powerful AI techniques for image recognition such as transformers for texts in NLP. Many studies found ViTs outperformed CNNs for common discriminative tasks in retinal imaging. Many studies fused CNNs with ViTs taking advantages of both techniques.

For generative AI, LLMs are useful for text-related tasks, whereas GAN is useful for image synthesis to fill the gap of imbalanced and inadequate datasets for model development. With huge amount of pretrained data, foundation models are becoming game changers for their generalizability to be fine-tuned for specific tasks with unlabeled, smaller datasets of images.

The trend toward democratizing AI would further advance research in this field and tentatively reduce AI disparities. The rapid advance in AI research may unfortunately widen the gap between "development" and "deployment" even more, considering only a few AI models are deployed in the real world. To close this gap, the evidence showing AI models could prevent visual loss, particularly in marginalized populations, might be required.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Financial support and sponsorship

Nil.

Conflicts of interest

The authors declare that there are no conflicts of interests of this paper.

References

1. Tortora L. Beyond discrimination: Generative AI applications and ethical challenges in forensic psychiatry. Front Psychiatry 2024;15:1346059. [doi: 10.3389/fpsyt.2024.1346059].
2. Taylor E. Artificial Intelligence Techniques: Explained in Detail; 2023. Available from: https://www.theknowledgeacademy.com/blog/artificial-intelligence-techniques/#:~:text=Understand%20the%20fundamentals%20of%20AI, they%20shape%20the%20AI%20landscape. [Last accessed on 2024 Jul 26].
3. Nelson D. Generative versus Discriminative Machine Learning Models; 2021. Available from: https://www.unite.ai/generative-

vs-discriminative-machine-learning-models/. [Last accessed on 2024 Jul 26].
4. Raumviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, *et al.* Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. NPJ Digit Med 2019;2:25.
5. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, *et al*. Predicting the risk of developing diabetic retinopathy using deep learning. Lancet Digit Health 2021;3:e10-9.
6. Liu X, Wu J, Shao A, Shen W, Ye P, Wang Y, *et al*. Uncovering language disparity of ChatGPT on retinal vascular disease classification: Cross-sectional study. J Med Internet Res 2024;26:e51926.
7. Tanachotnarangkun P, Marukatat S, Kumazawa I, Chanvarasuth P, Ruamviboonsuk P, Amornpetchsathaporn A, *et al*. A framework for generating an ICGA from a fundus image using GAN. 19th International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON); 2022 May 24-27; 2022. p. 1-4.
8. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, *et al*. Generative adversarial nets. Adv Neural Inf Process Syst 2014;27:139-44.
9. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, editors. Rethinking the inception architecture for computer vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. [doi: arXiv: 1512.00567].
10. Stanford University Human-Centered Artificial Intelligence. Available from: https://hai.stanford.edu/. [Last accessed on 2024 Jul 26].
11. Foundation model. Available from: https://en.wikipedia.org/wiki/Foundation_model. [Last accessed on 2024 Jul 26].
12. Nguyen TD, Ting YS, Ciucă I, O'Neill C, Sun ZC, Jabłońska M, *et al*. AstroLLaMA: Towards specialized foundation models in astronomy. arXiv:2309.06126. 2023. Available at https://doi.org/10.48550/arXiv.2309.06126. [Last accessed on 2024 Sep 28].
13. Tu T, Azizi S, Driess D, Schaekermann M, Amin M, Chang P-C, *et al*. Towards generalist biomedical AI. NEJM AI 2024;1:AIoa2300138. [doi: 10.1056/AIoa2300138].
14. Zhou Y, Chia MA, Wagner SK, Ayhan MS, Williamson DJ, Struyven RR, *et al*. A foundation model for generalizable disease detection from retinal images. Nature 2023;622:156-63.
15. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, *et al*. Language models are few-shot learners. Adv Neural Inf Process Syst 2020;33:1877-901.
16. Bommasani R, Hudson DA, Adeli E, Altman R, Arora S, von Arx S, *et al*. On the opportunities and risks of foundation models. arXiv:2108.07258. 2021. Available at https://doi.org/10.48550/arXiv.2108.07258. [Last accessed on 2024 Sep 28].
17. Hai Z, Zou B, Xiao X, Peng Q, Yan J, Zhang W, *et al*. A novel approach for intelligent diagnosis and grading of diabetic retinopathy. Comput Biol Med 2024;172:108246.
18. Dai L, Sheng B, Chen T, Wu Q, Liu R, Cai C, *et al*. A deep learning system for predicting time to progression of diabetic retinopathy. Nat Med 2024;30:584-94.
19. Bora A, Tiwari R, Bavishi P, Virmani S, Huang R, Traynis I, *et al*. Risk stratification for diabetic retinopathy screening order using deep learning: A multicenter prospective study. Transl Vis Sci Technol 2023;12:11.
20. Abd El-Khalek AA, Balaha HM, Alghamdi NS, Ghazal M, Khalil AT, Abo-Elsoud MEA, *et al*. A concentrated machine learning-based classification system for age-related macular degeneration (AMD) diagnosis using fundus images. Sci Rep 2024;14:2434.
21. Sarao V, Veritti D, De Nardin A, Misciagna M, Foresti G, Lanzetta P. Explainable artificial intelligence model for the detection of geographic atrophy using colour retinal photographs. BMJ Open Ophthalmol 2023;8:e001411.
22. Coyner AS, Murickan T, Oh MA, Young BK, Ostmo SR, Singh P, *et al*. Multinational external validation of autonomous retinopathy of prematurity screening. JAMA Ophthalmol 2024;142:327-35.
23. Oh0no-Matsui K, Kawasaki R, Jonas JB, Cheung CM, Saw SM, Verhoeven VJ, *et al*. International photographic classification and grading system for myopic maculopathy. Am J Ophthalmol 2015;159:877-83.e7.
24. Zheng B, Zhang M, Zhu S, Wu M, Chen L, Zhang S, *et al*. Research on an artificial intelligence-based myopic maculopathy grading method using EfficientNet. Indian J Ophthalmol 2024;72:S53-9.
25. Wu P, Qu Y, Zhao Z, Cui Y, Xu Y, An P, *et al*. An adaptive weighted ensemble learning network for diabetic retinopathy classification. J Xray Sci Technol 2024;32:285-301.
26. Liu H, Gao W, Yang L, Wu D, Zhao D, Chen K, *et al*. Semantic uncertainty guided cross-transformer for enhanced macular edema segmentation in OCT images. Comput Biol Med 2024;174:108458.
27. Meng Z, Chen Y, Li H, Zhang Y, Yao X, Meng Y, *et al*. Machine learning and optical coherence tomography-derived radiomics analysis to predict persistent diabetic macular edema in patients undergoing anti-VEGF intravitreal therapy. J Transl Med 2024;22:358.
28. Leingang O, Riedl S, Mai J, Reiter GS, Faustmann G, Fuchs P, *et al*. Automated deep learning-based AMD detection and staging in real-world OCT datasets (PINNACLE study report 5). Sci Rep 2023;13:19545.
29. Borrelli E, Oakley JD, Iaccarino G, Russakoff DB, Battista M, Grosso D, *et al*. Deep-learning based automated quantification of critical optical coherence tomography features in neovascular age-related macular degeneration. Eye (Lond) 2024;38:537-44.
30. Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, *et al*. Fully automated detection and quantification of macular fluid in OCT using deep learning. Ophthalmology 2018;125:549-58.
31. Pawloff M, Gerendas BS, Deak G, Bogunovic H, Gruber A, Schmidt-Erfurth U. Performance of retinal fluid monitoring in OCT imaging by automated deep learning versus human expert grading in neovascular AMD. Eye (Lond) 2023;37:3793-800.
32. Dow ER, Jeong HK, Katz EA, Toth CA, Wang D, Lee T, *et al*. A deep-learning algorithm to predict short-term progression to geographic atrophy on spectral-domain optical coherence tomography. JAMA Ophthalmol 2023;141:1052-61.
33. Jung J, Han J, Han JM, Ko J, Yoon J, Hwang JS, *et al*. Prediction of neovascular age-related macular degeneration recurrence using optical coherence tomography images with a deep neural network. Sci Rep 2024;14:5854.
34. Jang B, Lee SY, Kim C, Park UC, Kim YG, Lee EK. Preliminary analysis of predicting the first recurrence in patients with neovascular age-related macular degeneration using deep learning. BMC Ophthalmol 2023;23:499.
35. Ayhan MS, Neubauer J, Uzel MM, Gelisken F, Berens P. Interpretable detection of epiretinal membrane from optical coherence tomography with deep neural networks. Sci Rep 2024;14:8484.
36. Yan Y, Huang X, Jiang X, Gao Z, Liu X, Jin K, *et al*. Clinical evaluation of deep learning systems for assisting in the diagnosis of the epiretinal membrane grade in general ophthalmologists. Eye (Lond) 2024;38:730-6.
37. Gholami S, Scheppke L, Kshirsagar M, Wu Y, Dodhia R, Bonelli R, *et al*. Self-supervised learning for improved optical coherence tomography detection of macular telangiectasia type 2. JAMA Ophthalmol 2024;142:226-33.
38. Chew EY, Peto T, Clemons TE, Sallo FB, Pauleikhoff D, Leung I, *et al*. Macular telangiectasia type 2: A classification system using multimodal imaging mactel project report number 10. Ophthalmol Sci 2023;3:100261.
39. Wu Y, Egan C, Olvera-Barrios A, Scheppke L, Peto T, Charbel Issa P, *et al*. Developing a continuous severity scale

for macular telangiectasia type 2 using deep learning and implications for disease grading. Ophthalmology 2024;131:219-26.
40. Heinke A, Zhang H, Deussen D, Galang CM, Warter A, Kalaw FG, *et al*. Artificial intelligence for optical coherence tomography angiography-based disease activity prediction in age-related macular degeneration. Retina 2024;44:465-74.
41. Larsen TJ, Pettersen MB, Nygaard Jensen H, Lynge Pedersen M, Lund-Andersen H, Jørgensen ME, *et al*. The use of artificial intelligence to assess diabetic eye disease among the Greenlandic population. Int J Circumpolar Health 2024;83:2314802. [doi: 10.1080/22423982.2024.2314802].
42. Wu T, Ju L, Fu X, Wang B, Ge Z, Liu Y. Deep learning detection of early retinal peripheral degeneration from ultra-widefield fundus photographs of asymptomatic young adult (17-19 Years) candidates to airforce cadets. Transl Vis Sci Technol 2024;13:1.
43. Christ M, Habra O, Monnin K, Vallotton K, Sznitman R, Wolf S, *et al*. Deep learning-based automated detection of retinal breaks and detachments on fundus photography. Transl Vis Sci Technol 2024;13:1.
44. Hamel CP, Meunier I, Arndt C, Ben Salah S, Lopez S, Bazalgette C, *et al*. Extensive macular atrophy with pseudodrusen-like appearance: A new clinical entity. Am J Ophthalmol 2009;147:609-20.
45. Kovach JL. Extensive macular atrophy with pseudodrusen imaged with OCT angiography. Case Rep Ophthalmol Med 2018;2018:8213097. [doi: 10.1159/000526970].
46. Chouraqui M, Crincoli E, Miere A, Meunier IA, Souied EH. Deep learning model for automatic differentiation of EMAP from AMD in macular atrophy. Sci Rep 2023;13:20354.
47. Wang H, Chen X, Miao X, Tang S, Lin Y, Zhang X, *et al*. A deep learning model for detecting rhegmatogenous retinal detachment using ophthalmologic ultrasound images. Ophthalmologica 2024;247:8-18.
48. Gao Z, Pan X, Shao J, Jiang X, Su Z, Jin K, *et al*. Automatic interpretation and clinical evaluation for fundus fluorescein angiography images of diabetic retinopathy patients by deep learning. Br J Ophthalmol 2023;107:1852-8.
49. Lee PK, Ra H, Baek J. Automated segmentation of ultra-widefield fluorescein angiography of diabetic retinopathy using deep learning. Br J Ophthalmol 2023;107:1859-63.
50. CATT Research Group, Martin DF, Maguire MG, Ying GS, Grunwald JE, Fine SL, *et al*. Ranibizumab and bevacizumab for neovascular age-related macular degeneration. N Engl J Med 2011;364:1897-908.
51. Chandra RS, Ying GS. Predicting visual acuity responses to anti-VEGF treatment in the comparison of age-related macular degeneration treatments trials using machine learning. Ophthalmol Retina 2024;8:419-30.
52. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, *et al*. Huggingface's transformers: State-of-the-art natural language processing. arXiv:1910.03771. 2019. Available at https://doi.org/10.48550/arXiv.1910.03771. [Last accessed on 2024 Sep 28].
53. What Are Transformers in Artificial Intelligence?. Available from: https://aws.amazon.com/what-is/transformers-in-artificial-intelligence/#:~:text=Transformers%20are%20a%20type%20of, tracking%20relationships%20between%20sequence%20components. [Last accessed on 2024 May 22].
54. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, *et al*. An image is worth 16×16 words: Transformers for image recognition at scale. arXiv:2010.11929. 2020. Available at https://doi.org/10.48550/arXiv.2010.11929. [Last accessed on 2024 Sep 28].
55. Li J, Chen J, Tang Y, Wang C, Landman BA, Zhou SK. Transforming medical imaging with transformers? A comparative review of key properties, current progresses, and future perspectives. Med Image Anal 2023;85:102762.
56. Wu J, Hu R, Xiao Z, Chen J, Liu J. Vision transformer-based recognition of diabetic retinopathy grade. Med Phys 2021;48:7850-63.
57. Yang Y, Cai Z, Qiu S, Xu P. Vision transformer with masked autoencoders for referable diabetic retinopathy classification based on large-size retina image. PLoS One 2024;19:e0299265.
58. Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabuddhe V, *et al*. Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. Data 2018;3:25.
59. Li T, Gao Y, Wang K, Guo S, Liu H, Kang H. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Inf Sci 2019;501:511-22.
60. Gu Z, Li Y, Wang Z, Kan J, Shu J, Wang Q. Classification of diabetic retinopathy severity in fundus images using the vision transformer and residual attention. Comput Intell Neurosci 2023;2023:1305583. [doi: 10.1155/2023/1305583].
61. Zang F, Ma H. CRA-Net: Transformer guided category-relation attention network for diabetic retinopathy grading. Comput Biol Med 2024;170:107993.
62. Huang Y, Lyu J, Cheng P, Tam R, Tang X. SSiT: Saliency-guided self-supervised image transformer for diabetic retinopathy grading. IEEE J Biomed Health Inform 2024;28:2806-17.
63. Zhang C, Chen C, Chen C, Lv X. SMiT: Symmetric mask transformer for disease severity detection. J Cancer Res Clin Oncol 2023;149:16075-86.
64. Huang S, Li J, Xiao Y, Shen N, Xu T. RTNet: Relation transformer network for diabetic retinopathy multi-lesion segmentation. IEEE Trans Med Imaging 2022;41:1596-607.
65. Wang Z, Lu H, Yan H, Kan H, Jin L. Vison transformer adapter-based hyperbolic embeddings for multi-lesion segmentation in diabetic retinopathy. Sci Rep 2023;13:11178.
66. Xu K, Huang S, Yang Z, Zhang Y, Fang Y, Zheng G, *et al*. Automatic detection and differential diagnosis of age-related macular degeneration from color fundus photographs using deep learning with hierarchical vision transformer. Comput Biol Med 2023;167:107616.
67. Domínguez C, Heras J, Mata E, Pascual V, Royo D, Zapata MÁ. Binary and multi-class automated detection of age-related macular degeneration using convolutional- and transformer-based architectures. Comput Methods Programs Biomed 2023;229:107302.
68. Cai L, Wen C, Jiang J, Liang C, Zheng H, Su Y, *et al*. Classification of diabetic maculopathy based on optical coherence tomography images using a vision transformer model. BMJ Open Ophthalmol 2023;8:e001423.
69. Kihara Y, Shen M, Shi Y, Jiang X, Wang L, Laiginhas R, *et al*. Detection of nonexudative macular neovascularization on structural OCT images using vision transformers. Ophthalmol Sci 2022;2:100197.
70. Dutta P, Sathi KA, Hossain MA, Dewan MA. Conv-ViT: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection. J Imaging 2023;9:140.
71. Ma Z, Xie Q, Xie P, Fan F, Gao X, Zhu J. HCTNet: A hybrid ConvNet-transformer network for retinal optical coherence tomography image classification. Biosensors (Basel) 2022;12:542.
72. He J, Wang J, Han Z, Ma J, Wang C, Qi M. An interpretable transformer network for the retinal disease classification using optical coherence tomography. Sci Rep 2023;13:3637.
73. Shen J, Hu Y, Zhang X, Gong Y, Kawasaki R, Liu J. Structure-oriented transformer for retinal diseases grading from OCT images. Comput Biol Med 2023;152:106445.
74. Ait Hammou B, Antaki F, Boucher MC, Duval R. MBT: Model-based transformer for retinal optical coherence tomography image and video multi-classification. Int J Med Inform 2023;178:105178.
75. Zhang Y, Li Z, Nan N, Wang X. TranSegNet: hybrid CNN-vision transformers encoder for retina segmentation of optical coherence

tomography. Life (Basel) 2023;13:976.
76. Niu Z, Deng Z, Gao W, Bai S, Gong Z, Chen C, *et al*. FNeXter: A multi-scale feature fusion network based on ConvNeXt and transformer for retinal OCT fluid segmentation. Sensors (Basel) 2024;24:2425.
77. Quek TC, Takahashi K, Kang HG, Thakur S, Deshmukh M, Tseng RM, *et al*. Predictive, preventive, and personalized management of retinal fluid via computer-aided detection app for optical coherence tomography scans. EPMA J 2022;13:547-60.
78. Philippi D, Rothaus K, Castelli M. A vision transformer architecture for the automated segmentation of retinal lesions in spectral domain optical coherence tomography images. Sci Rep 2023;13:517.
79. Tan Z, Shi F, Zhou Y, Wang J, Wang M, Peng Y, *et al*. A multi-scale fusion and transformer based registration guided speckle noise reduction for OCT images. IEEE Trans Med Imaging 2024;43:473-88.
80. Liu H, Teng L, Fan L, Sun Y, Li H. A new ultra-wide-field fundus dataset to diabetic retinopathy grading using hybrid preprocessing methods. Comput Biol Med 2023;157:106750.
81. Nguyen TD, Le DT, Bum J, Kim S, Song SJ, Choo H. Retinal disease diagnosis using deep learning on ultra-wide-field fundus images. Diagnostics (Basel) 2024;14:105.
82. Li K, Zhu Q, Wu J, Ding J, Liu B, Zhu X, *et al*. DCT-Net: An effective method to diagnose retinal tears from B-scan ultrasound images. Math Biosci Eng 2024;21:1110-24.
83. Hu D, Li H, Liu H, Oguz I. Domain generalization for retinal vessel segmentation via Hessian-based vector field. Med Image Anal 2024;95:103164.
84. Playout C, Duval R, Boucher MC, Cheriet F. Focused attention in transformers for interpretable classification of retinal images. Med Image Anal 2022;82:102608.
85. Duranton S. ChatGPT — Let The Generative AI Revolution Begin; 2023. Available from: https://www.forbes.com/sites/sylvainduranton/2023/01/07/chatgpt3let-the-generative-ai-revolution-begin/?sh=3c5d94c5af15. [Last accessed on 2024 May 29].
86. Technologies MSTL. How to Use ChatGPT's Advanced Data Analysis Feature. Available from: https://mitsloanedtech.mit.edu/ai/tools/data-analysis/how-to-use-chatgpts-advanced-data-analysis-feature/. [Last accessed on 2024 29 May].
87. Mohammadi SS, Nguyen QD. A user-friendly approach for the diagnosis of diabetic retinopathy using ChatGPT and automated machine learning. Ophthalmol Sci 2024;4:100495.
88. Gopalakrishnan N, Joshi A, Chhablani J, Yadav NK, Reddy NG, Rani PK, *et al*. Recommendations for initial diabetic retinopathy screening of diabetic patients using large language model-based artificial intelligence in real-life case scenarios. Int J Retina Vitreous 2024;10:11.
89. Choudhary A, Gopalakrishnan N, Joshi A, Balakrishnan D, Chhablani J, Yadav NK, *et al*. Recommendations for diabetic macular edema management by retina specialists and large language model-based artificial intelligence platforms. Int J Retina Vitreous 2024;10:22.
90. Ferro Desideri L, Roth J, Zinkernagel M, Anguita R. "Application and accuracy of artificial intelligence-derived large language models in patients with age related macular degeneration". Int J Retina Vitreous 2023;9:71.
91. Potapenko I, Boberg-Ans LC, Stormly Hansen M, Klefter ON, van Dijk EH, Subhi Y. Artificial intelligence-based chatbot patient information on common retinal diseases using ChatGPT. Acta Ophthalmol 2023;101:829-31.
92. Anguita R, Makuloluwa A, Hind J, Wickham L. Large language models in vitreoretinal surgery. Eye (Lond) 2024;38:809-10.
93. You A, Kim JK, Ryu IH, Yoo TK. Application of Generative Adversarial Networks (GAN) for ophthalmology image domains: A survey. Eye Vis (Lond) 2022;9:6.
94. Zhou Y, Wang B, He X, Cui S, Shao L. DR-GAN: Conditional Generative Adversarial Network for fine-grained lesion synthesis on diabetic retinopathy images. IEEE J Biomed Health Inform 2022;26:56-66.
95. Xie Y, Wan Q, Xie H, Xu Y, Wang T, Wang S, *et al*. Fundus image-label Pairs synthesis and retinopathy screening via GANs with class-imbalanced semi-supervised learning. IEEE Trans Med Imaging 2023;42:2714-25.
96. Yang K, Lu Y, Xue L, Yang Y, Chang S, Zhou C. URNet: System for recommending referrals for community screening of diabetic retinopathy based on deep learning. Exp Biol Med (Maywood) 2023;248:909-21.
97. Yoo TK, Kim BY, Jeong HK, Kim HK, Yang D, Ryu IH. Simple code implementation for deep learning-based segmentation to evaluate central serous chorioretinopathy in fundus photography. Transl Vis Sci Technol 2022;11:22.
98. Pham QT, Ahn S, Shin J, Song SJ. Generating future fundus images for early age-related macular degeneration based on generative adversarial networks. Comput Methods Programs Biomed 2022;216:106648.
99. Choi JY, Ryu IH, Kim JK, Lee IS, Yoo TK. Development of a generative deep learning model to improve epiretinal membrane detection in fundus photography. BMC Med Inform Decis Mak 2024;24:25.
100. Burlina P, Paul W, Liu TY, Bressler NM. Detecting anomalies in retinal diseases using generative, discriminative, and self-supervised deep learning. JAMA Ophthalmol 2022;140:185-9.
101. Hou N, Shi J, Ding X, Nie C, Wang C, Wan J. ROP-GAN: An image synthesis method for retinopathy of prematurity based on generative adversarial network. Phys Med Biol 2023;68:205016. [doi: 10.1088/1361-6560/acf3c9].
102. Sun LC, Pao SI, Huang KH, Wei CY, Lin KF, Chen PN. Generative adversarial network-based deep learning approach in classification of retinal conditions with optical coherence tomography images. Graefes Arch Clin Exp Ophthalmol 2023;261:1399-412.
103. Zhao M, Lu Z, Zhu S, Wang X, Feng J. Automatic generation of retinal optical coherence tomography images based on generative adversarial networks. Med Phys 2022;49:7357-67.
104. Moon S, Lee Y, Hwang J, Kim CG, Kim JW, Yoon WT, *et al*. Prediction of anti-vascular endothelial growth factor agent-specific treatment outcomes in neovascular age-related macular degeneration using a generative adversarial network. Sci Rep 2023;13:5639.
105. Zhang Y, Huang K, Li M, Yuan S, Chen Q. Learn single-horizon disease evolution for predictive generation of post-therapeutic neovascular age-related macular degeneration. Comput Methods Programs Biomed 2023;230:107364.
106. Abdelmotaal H, Sharaf M, Soliman W, Wasfi E, Kedwany SM. Bridging the resources gap: Deep learning for fluorescein angiography and optical coherence tomography macular thickness map image translation. BMC Ophthalmol 2022;22:355.
107. Xiang D, Yan S, Guan Y, Cai M, Li Z, Liu H, *et al*. Semi-supervised dual stream segmentation network for fundus lesion segmentation. IEEE Trans Med Imaging 2023;42:713-25.
108. Kwon HJ, Heo J, Park SH, Park SW, Byon I. Accuracy of generative deep learning model for macular anatomy prediction from optical coherence tomography images in macular hole surgery. Sci Rep 2024;14:6913.
109. Odaibo SG. Retina-VAE: Variationally decoding the spectrum of macular disease. arXiv:1907.05195. 2019. Available at https://doi.org/10.48550/arXiv.1907.05195. [Last accessed on 2024 Sept 28].
110. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402-10.

111. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, *et al*. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 2016;57:5200-6.
112. Zhang J, Lin S, Cheng T, Xu Y, Lu L, He J, *et al*. RETFound-enhanced community-based fundus disease screening: Real-world evidence and decision curve analysis. NPJ Digit Med 2024;7:108.
113. Moorfields Eye Hospital L. RETFound - A Foundation Model for Retinal Image of Dataset; 2022. Available from: https://github.com/rmaphoh/RETFound_MAE. [Last accessed on 2024 Sep 28].
114. Rombach RB, Lorenz D, Esser P, Ommer B. Stable Diffusion. 2022. Available from: https://github.com/CompVis/stable-diffusion. [Last accessed on 2024 May 15].
115. Xiao X. Dreambooth on Stable Diffusion; 2022. Available from: https://github.com/XavierXiao/Dreambooth-Stable-Diffusion. [Last accessed on 2024 May 20].
116. Tabuchi H, Engelmann J, Maeda F, Nishikawa R, Nagasawa T, Yamauchi T, *et al*. Using artificial intelligence to improve human performance: efficient retinal disease detection training with synthetic images. Br J Ophthalmol 2024;108:1430-35.
117. Rath S. Segment Anything – A Foundation Model for Image Segmentation; 2023. Available from: https://learnopencv.com/segment-anything/#:~:text=input%20image%20embedding.-,The%20Segment%20Anything%20Dataset, images%20and%201.1%20 billion%20masks. [Last accesse on 2024 Jul 26].
118. Jiang H, Gao M, Liu Z, Tang C, Zhang X, Jiang S, *et al*. GlanceSeg: Real-time microaneurysm lesion segmentation with gaze-map-guided foundation model for early detection of diabetic retinopathy. IEEE J Biomed Health Inform. 2024. Available at https://ieeexplore.ieee.org/document/10472575. [Last accessed on 2024 Sep 28].
119. Rajesh AE, Davidson OQ, Lee CS, Lee AY. Artificial intelligence and diabetic retinopathy: AI framework, prospective studies, head-to-head validation, and cost-effectiveness. Diabetes Care 2023;46:1728-39.
120. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018;1:39.
121. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MY, *et al*. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: A clinical validation study. Lancet Digit Health 2019;1:e35-44.
122. Zhang Y, Shi J, Peng Y, Zhao Z, Zheng Q, Wang Z, *et al*. Artificial intelligence-enabled screening for diabetic retinopathy: A real-world, multicenter and prospective study. BMJ Open Diabetes Res Care 2020;8:e001596.
123. Heydon P, Egan C, Bolter L, Chambers R, Anderson J, Aldington S, *et al*. Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. Br J Ophthalmol 2021;105:723-8.
124. Ipp E, Liljenquist D, Bode B, Shah VN, Silverstein S, Regillo CD, *et al*. Pivotal evaluation of an artificial intelligence system for autonomous detection of referrable and vision-threatening diabetic retinopathy. JAMA Netw Open 2021;4:e2134254.
125. Scheetz J, Koca D, McGuinness M, Holloway E, Tan Z, Zhu Z, *et al*. Real-world artificial intelligence-based opportunistic screening for diabetic retinopathy in endocrinology and indigenous healthcare settings in Australia. Sci Rep 2021;11:15808.
126. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, *et al*. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: A prospective interventional cohort study. Lancet Digit Health 2022;4:e235-44.
127. Yang Y, Pan J, Yuan M, Lai K, Xie H, Ma L, *et al*. Performance of the AIDRScreening system in detecting diabetic retinopathy in the fundus photographs of Chinese patients: A prospective, multicenter, clinical study. Ann Transl Med 2022;10:1088.
128. Lee AY, Yanagihara RT, Lee CS, Blazes M, Jung HC, Chee YE, *et al*. Multicenter, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems. Diabetes Care 2021;44:1168-75.
129. Ruamviboonsuk P, Ruamviboonsuk V, Tiwari R. Recent evidence of economic evaluation of artificial intelligence in ophthalmology. Curr Opin Ophthalmol 2023;34:449-58.
130. Chan SC, Neves AL, Majeed A, Faisal A. Bridging the equity gap towards inclusive artificial intelligence in healthcare diagnostics. BMJ 2024;384:q490.
131. d'Elia A, Gabbay M, Rodgers S, Kierans C, Jones E, Durrani I, *et al*. Artificial intelligence and health inequities in primary care: A systematic scoping review and framework. Fam Med Community Health 2022;10: e001670.
132. Seger E, Ovadya A, Siddarth D, Garfinkel B, Dafoe A, editors. Democratising AI: Multiple Meanings, Goals, and Methods. Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society; 2023. doi: arXiv: 2303.12642.
133. Marwala T. Framework for the Governance of Artificial Intelligence; 2024. Available from: https://medium.com/@tshilidzimarwala/framework-for-the-governance-of-artificial-intelligence-398a2135d345#:~:text=The%20AI%20governance%20model%20conforms,%2C%20accurate%2C%20and%20 secure%20use. [Last accessed on 2024 May 29].