Computerized Medical Imaging and Graphics 108 (2023) 102269

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Unraveling the complexity of Optical Coherence Tomography image segmentation using machine and deep learning techniques: A review

Mehmood Nawaz [a], Adilet Uvaliyev [a], Khadija Bibi [a], Hao Wei [a], Sai Mu Dalike Abaxi [a], Anum Masood [b], Peilun Shi [a], Ho-Pui Ho [a], Wu Yuan [a,*]

[a] Department of Biomedical Engineering, The Chinese University of Hong Kong, Hong Kong Special Administrative Region of China
[b] Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Index terms:
Optical coherence tomography
Image segmentation
Deep learning
Convolution neural network

ABSTRACT

Optical Coherence Tomography (OCT) is an emerging technology that provides three-dimensional images of the microanatomy of biological tissue in-vivo and at micrometer-scale resolution. OCT imaging has been widely used to diagnose and manage various medical diseases, such as macular degeneration, glaucoma, and coronary artery disease. Despite its wide range of applications, the segmentation of OCT images remains difficult due to the complexity of tissue structures and the presence of artifacts. In recent years, different approaches have been used for OCT image segmentation, such as intensity-based, region-based, and deep learning-based methods. This paper reviews the major advances in state-of-the-art OCT image segmentation techniques. It provides an overview of the advantages and limitations of each method and presents the most relevant research works related to OCT image segmentation. It also provides an overview of existing datasets and discusses potential clinical applications. Additionally, this review gives an in-depth analysis of machine learning and deep learning approaches for OCT image segmentation. It outlines challenges and opportunities for further research in this field.

1. Introduction

Image segmentation is a fundamental technique that is utilized in several fields, including medical imaging, autonomous vehicles, robotics, video analytics, computer vision, etc. In the medical area, manual segmentation of medical images is time-consuming and requires expertise. To resolve this issue, researchers proposed many types of segmentation techniques. For example, Fig. 1 shows the OCT image segmentation using U-net deep learning model. This article analyzed different types of OCT (Optical Coherence Tomography) image segmentation techniques. OCT is a non-invasive imaging technique that generates cross-sectional and volumetric images with high resolution. OCT images are generated by the magnitude and time delay of back-reflected light (Fujimoto and Drexler, 2015a). In OCT, A-scans are measurements of properties of back-scattered light against depth or axial direction. The B-Scans or cross-sectional images are obtained by moving incident light in the transverse direction and acquiring A-scans sequentially. The volumetric images are constructed by obtaining B-Scans sequentially. The light from a source is split to sample and reference arm with known optical path length to calculate the magnitude and time delay of back-scattered light (Aumann et al., 2019). The first OCT system was introduced in 1991 (Huang et al., 1991), where the length of the optical path in the reference arm had to change for each scan. Later, the spectral domain OCT methods that work without mechanical scan of the reference arm and have high-speed OCT imaging acquisition (Choma et al., 2003; Leitgeb et al., 2003) were introduced. Ultrasound and microscopy imaging techniques have some common characteristics with OCT (Fujimoto and Drexler, 2015b). Ultrasound has a low resolution of merely 0.1–1 millimeter (mm) but has a high penetration depth. On the other hand, microscopy has a one-micrometer resolution or higher but has a low penetration depth of around one hundred micrometers. In terms of these parameters, OCT resides between ultrasound and microscopy.

OCT imaging is widely used in ophthalmology, where an optical beam travels inside the eye with low optical attenuation. This makes OCT suitable for imaging of the retina and the interior part of an eye. The resolution and contrast of retinal images obtained with OCT are better than other non-contact imaging methods (Drexler and Fujimoto, 2015). It is estimated that over 20 million eye patients benefit from OCT imaging every year (Swanson, 2015). It is used to diagnose eye diseases such as macular edema (Hee et al., 1995) and glaucoma (Bowd et al., 2001; Lu et al., 2023). Besides ophthalmology, it has applications in various medical fields (shown in Fig. 2) like dermatology to study

* Corresponding author.
E-mail address: wyuan@cuhk.edu.hk (W. Yuan).

https://doi.org/10.1016/j.compmedimag.2023.102269
Received 3 May 2023; Received in revised form 30 June 2023; Accepted 3 July 2023
Available online 14 July 2023

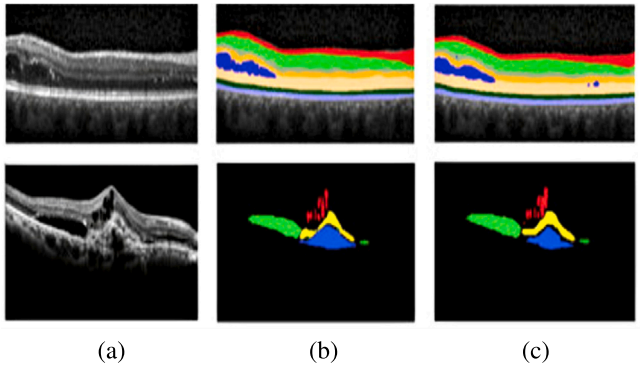


Fig. 1. An example of OCT image segmentation using deep learning technique (Schmidt-Erfurth et al., 2018). (a) shows the original OCT image, (b) shows the multiple-layer segmentation using U-net, and (c) shows the ground truth.

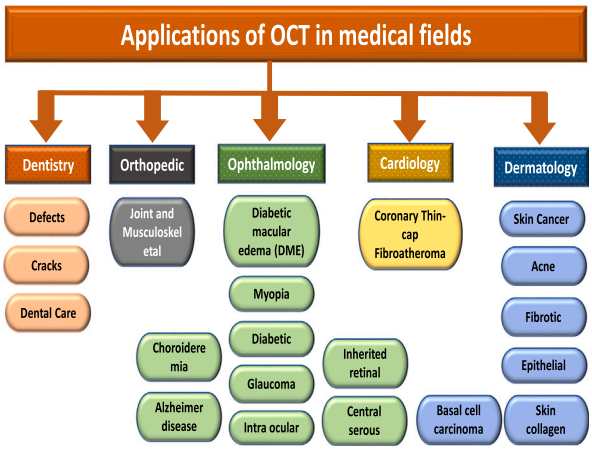


Fig. 2. Applications of Optical Coherence Tomography (OCT) imaging technique in different medical modalities.

skin diseases, such as non-melanoma skin cancer (Themstrup et al., 2014), acne lesions (Manfredini et al., 2017) and contact dermatitis (Boone et al., 2015; López-Varela et al., 2023). Unlike confocal laser microscopy or multi-photon tomography, OCT has a high penetration depth in tissues. This is one of the advantages of OCT in dermatology, as it allows visualization down to the depth of the skin layers (Holmes and Welzel, 2015). In cardiology, OCT imaging is used for lesion detection (Barlis et al., 2008; van Soest et al., 2009). OCT is also employed to visualize different oral and dental structures with high resolution to help the diagnosis of dental caries (Amaechi et al., 2001; Wilder-Smith et al., 2008) and oral cancer (Wilder-Smith et al., 2004).

Some of the earlier OCT segmentation techniques were based on traditional machine learning algorithms, as shown in Table 1. Machine Learning in OCT Medical Images is a branch of Artificial Intelligence (AI) that uses algorithms to learn from data and make predictions or decisions without being explicitly programmed (Mahesh, 2020). It is used to automatically analyze medical images, such as Optical Coherence Tomography (OCT), to detect and diagnose diseases, as well as to predict outcomes and suggests further interventions. Machine Learning techniques are used to identify patterns in OCT images, such as spots, cracks, and shadows, so that healthcare professionals can make accurate medical decisions.

Huang et al. (2018) proposed a support vector machine (SVM) based technique for the segmentation of plaque components from intravascular OCT images. Lang et al. (2015) segmented microcystic macular edema from OCT images using a Random Forest Classifier (RFC), as shown in Fig. 4. This classifier can be trained to identify

Table 1
Overview of traditional machine learning-based OCT segmentation methods.

Ref	Year	Algorithm	Segmentation
Liu et al. (2023)	2023	Support vector machine	Skin tissue OCT
Abbas et al. (2022)	2022	Advances in TML	Blood vessel
Huang et al. (2018)	2018	Support vector machine	Blood vessel
Guo et al. (2019)	2018	Support vector machine	Blood vessel
Amrute et al. (2017)	2017	K-means	Blood vessel
Lang et al. (2015)	2014	Random forest	Lesion areas in retina
Lang et al. (2013)	2013	Random forest	Retinal layer

various features of OCT images, such as intensity, size, shape, and texture. This can be used to detect and classify various tissue types in the image, such as healthy tissue, tumor, cysts, and other lesions. The classifier can also be used to detect changes in the OCT images over time, such as new lesions or changes in the tissue structure. Another type of OCT segmentation technique is based on deep learning. Deep learning is a branch of machine learning which enables computers to learn from observation and to perform human-like tasks using bio-inspired neural networks with multiple layers (Kelleher, 2019). Neural networks consist of neurons that are connected in a sequential way to form a neural network. Each neuron is considered a component that processes information.

One of the challenges in building a deep learning model for OCT image segmentation is the difficulty of obtaining labeled data. Acquiring accurate ground truth is laborious and requires expert knowledge. Another problem is the lack of homogeneity in the dataset as different OCT devices output different images, and models trained on images from a single source might not work on images from multiple sources (Ran and Cheung, 2021). Also, there is an insufficient amount of generally accessible OCT datasets, and there might be problems with overfitting (Yanagihara et al., 2020). Deep learning models are computationally expensive and can take a long time to train, making it difficult to deploy them in a clinical setting (Ching et al., 2018). Therefore, we can overcome on these challenges using transfer learning, data augmentation, high-speed GPU, optimized hyper-parameters, etc.

Li et al. (2020a) developed a fully supervised deep-learning method for the segmentation of retinal layers from OCT images. This method requires a large amount of labeled data for training, which is expensive. This limits the use of deep learning in OCT to a few data sets with sufficient manual annotations. However, obtaining labeled images is a time-consuming and expert-dependent process. To tackle this problem, techniques based on semi-supervised learning using labeled and unlabeled data for model training were developed. Liu et al. (2018) designed a semi-supervised method with adversarial training for retinal fluid and layer segmentation in OCT scans. Other than semi-supervised learning, non-supervised (weakly supervised) techniques that use annotations other than pixel-level annotations were also developed. Xing et al. (2021) proposed a weakly supervised technique that uses only image-level labels to segment retinal detachment from SD-OCT (spectral domain) images. A complete OCT segmentation process is shown in Fig. 3.

Contributions: The main contributions of our review article are summarized as:

- **Summary of important OCT segmentation techniques:** Review comprises of traditional machine learning and deep learning methods for OCT image segmentation.
- **Detailed assessment and challenges of OCT image segmentation techniques:** We carried out an extensive study of the current OCT image segmentation techniques, their challenges, and how to overcome these challenges.
- **Resources for new researchers:** The quantitative and qualitative analysis of these approaches help the new researchers to better understand the basics of OCT image segmentation techniques.

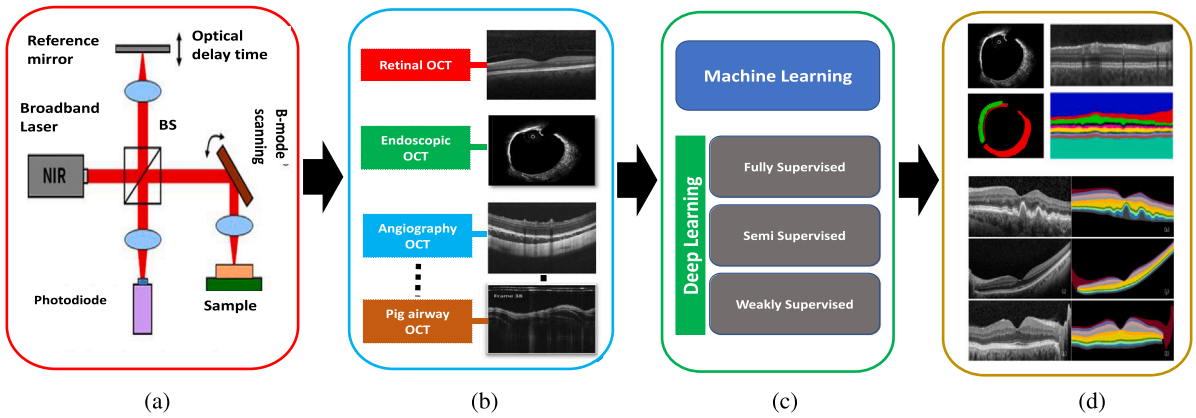


Fig. 3. Flow chart of OCT image segmentation. (a) shows the OCT image acquisition system, (b) shows the different types of OCT images, (c) shows the image segmentation methods including machine learning and deep learning, and (d) shows the segmentation results of OCT images.

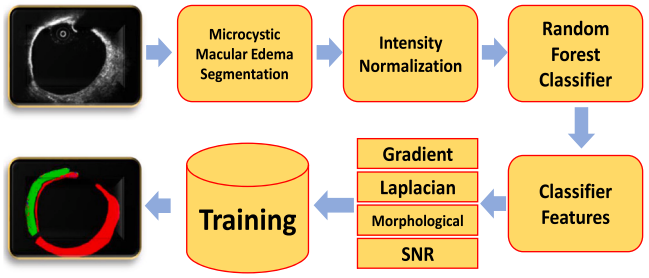


Fig. 4. Shows the traditional machine learning technique for OCT image segmentation (Moccia et al., 2018).

- **Possible directions for future:** We explored new ideas for future direction, including real-time OCT segmentation, 3D OCT segmentation, etc.

The rest of the paper is structured as follows: in Section 2, we reviewed traditional machine learning and deep learning-based OCT segmentation methods. Section 3 introduces datasets, evaluation metrics, and labeling techniques. Section 4 explains the performance comparisons of different OCT segmentation methods. Discussion of the reviewed methods is presented in Section 5. Section 6, and VII discuss applications of OCT methods in medicine and future work, respectively.

2. Related works

In this section, we reviewed state-of-the-art OCT image segmentation methods. These methods are based on traditional machine learning, fully supervised deep learning, semi-supervised deep learning, and weakly supervised deep learning approaches, as shown in Table 2.

2.1. Traditional machine learning techniques for OCT images

In this section, we have reviewed traditional machine learning OCT segmentation methods, as shown in Table 1. Huang et al. (2018) built a segmentation method based on a support vector machine (SVM) for intravascular OCT images. This technique segments three types of plaque components: lipid-containing regions, fibrous tissue, and calcification. Attenuation coefficient and image texture features calculated from the gray level co-occurrence matrix are used as features to build the SVM model. A dataset for training and testing was obtained from Nanjing Drum Tower Hospital consisting of 28 scans obtained from 11 people. This technique has achieved 83% pixel accuracy in OCT image segmentation. Similarly, Guo et al. (2019) developed Least Squares Support Vector Machine-based model for the segmentation of plaque components from intra-coronary OCT images. A custom dataset of 9 scans obtained from one patient is used. It has achieved 80% accuracy and 89.6% recall in OCT image segmentation results.

In another study, a segmentation method based on a Random Forest (RF) classifier is developed for the segmentation of microcystic macular edema from OCT images (Lang et al., 2015). In this technique, each pixel has classified as either background or lesion region. They used intensity and location-based features to train this model. A dataset obtained by the spectral scanner from 9 microcystic macular edema subjects was used for training and validation. Performance evaluation shows that the model achieved 85% precision, 79% recall, and 80% F-measure scores. Similarly, Lang et al. (2013) proposed RFC (Random Forest classifier) based retinal layers segmentation technique. They use the location of each pixel and its relationship with neighbor pixels to train the RFC model, as shown in Fig. 4. The training dataset comprised of 10 OCT volumes obtained with a spectral scanner. This method has achieved 79% on the dice coefficient metric score.

In another study (Amrute et al., 2017), the K-means clustering-based OCT image segmentation method is developed. The clustering approach is used to segment pixels based on their intensity into groups and then segment the object of interest. This method is validated by 658 annotated OCT scans obtained from public hospitals. It has achieved a 92% F-measure score in segmentation results.

2.2. Deep learning techniques for OCT images

2.2.1. Fully supervised deep learning techniques

This section explains how convolutional neural networks (CNN) use convolution operations to segment the OCT images. An overview of deep learning techniques is shown in Table 2. Schlegl et al. (2018) developed a method to segment macular fluid with the help of a convolutional neural network using encoder–decoder architecture. It used 1200 OCT volumes obtained from Vienna Reading Center. This method has achieved 91% precision and 84% recall for intraretinal fluid segmentation. However, for subretinal fluid segmentation, it has achieved 61% precision and 81% recall. In another study (Lee et al., 2017), a U-Net architecture-based segmentation method is proposed to segment the macular fluid OCT images. This network uses a modified U-Net architecture. It is trained on 1289 annotated OCT scans that were obtained from the Ophthalmology Department of the University of Washington. It has achieved a 91.1% Dice coefficient score in segmentation. Li et al. (2020a) proposed a method with a convolutional neural network called DeepRetina to segment retinal layers. This network also has an encoder–decoder structure like Lee et al. (2017). They use modified Xception architecture as the encoder backbone in this network. DeepRetina has achieved 90% IoU and 92% sensitivity score on the publicly available Duke dataset (Chiu et al., 2015).

Table 2
Summary of deep learning methods for OCT segmentation. It contains the last five years research articles on OCT image segmentation.

Ref.	Year	Network type	Group of network	Group of applications
Chen et al. (2023)	2023	Fully supervised	Deep network with rough fuzzy discretization	OCT fundus image segmentation, Retinal fluid
Garcia-Marin et al. (2023)	2023	Fully supervised	Patch-based CNN	Corneal segmentation of AS-OCT images
Lu et al. (2023)	2023	Semi-supervised	Boundary-enhanced semi-supervised	Surgical, ROI, Retinal fluid
López-Varela et al. (2023)	2023	Fully supervised	CNN based fully automatic segmentation	Choriocapillaris flow voids in OCTA images
Zhang et al. (2023)	2023	Fully supervised	CNN following a U-Net-like design	Surgical navigation, ROI, Retinal fluid
Chen et al. (2023)	2023	Fully supervised	Rough fuzzy discretization (RFDDN)	Retinal fluid
Ren et al. (2023)	2023	Fully supervised	Deep learning with level-set method	Retinal fluid
Schlegl et al. (2018)	2023	Weakly supervised	TSSK-Net with biomarker localization	Retinal fluid
Venkatesh et al. (2022)	2022	Fully supervised	CNN with 2 MacTel structure	Retinal fluid
Yang et al. (2022)	2022	Fully supervised	DCU-net: a cascade U-net	Retinal fluid
Liu et al. (2022)	2022	Semi supervised	Encoder and 3 decoder with modified Unet	Retinal fluid
He et al. (2022)	2022	Weakly supervised	Unet Netwrok	Retinal fluid
Wang et al. (2021)	2021	Weakly supervised	CycleGAN Network	Lesion areas
Soltanian-Zadeh et al. (2021)	2021	Weakly supervised	FCN Network	Ganglion cells
Xing et al. (2021)	2021	Weakly supervised	DeepLab Framework	Lesion areas
Wang et al. (2020)	2020	Weakly supervised	DeepLab v3 Framework	Lesion areas
Li et al. (2020a)	2020	Fully supervised	CNN with encoder–decoder structure	Retinal layers
Heisler et al. (2020)	2020	Semi supervised	GAN Network	Retinal layers
Liu et al. (2019)	2019	Fully supervised	ResNet-101 Framework	Choroidal vessels
Dos Santos et al. (2019)	2019	Fully supervised	FCN Framework	Cornea
Sedai et al. (2019)	2019	Semi supervised	Dense-Unet Framework	Retinal layers
Schlegl et al. (2018)	2018	Fully supervised	CNN with encoder–decoder structure	Retinal fluid
Liu et al. (2018)	2018	Semi supervised	Segmentation and discriminator network with modified Unet	Retinal layers and fluid
Lee et al. (2017)	2017	Fully supervised	Modified Unet Framework	Retinal fluid
Roy et al. (2017)	2017	Fully supervised	FCN Framework	Retinal layers and fluid

Roy et al. (2017) proposed a fully convolutional network to segment the retinal layer and fluid. It employs an extended path of convolutional blocks (decoders) for semantic segmentation after learning a hierarchy of contextual information using a compact path of convolutional blocks (encoders). In Fully Convolutional Networks (FCN), all the convolutional layers are replaced by fully connected layers. In a study (Liu et al., 2019), a fully convolutional network is proposed for the segmentation of choroidal vessels. This method uses ResNet-101 as a backbone network, as shown in Fig. 5. Additionally, a refinement network module which is called RefineNet (Lin et al., 2017) is used to enable the higher resolution prediction. It is trained on 40 OCT images acquired from 10 patients. RefineNet achieved an 84% dice coefficient score on segmentation, which is very close to the grader performance. Dos Santos et al. (2019) proposed a method for cornea segmentation from OCT scans. It is a modified network with a fully convolutional U-shape architecture that is called CorneaNet. It used 280 OCT volumes obtained with an ultra-high resolution OCT machine. CorneaNet achieved a 99.6% dice coefficient score in segmentation results.

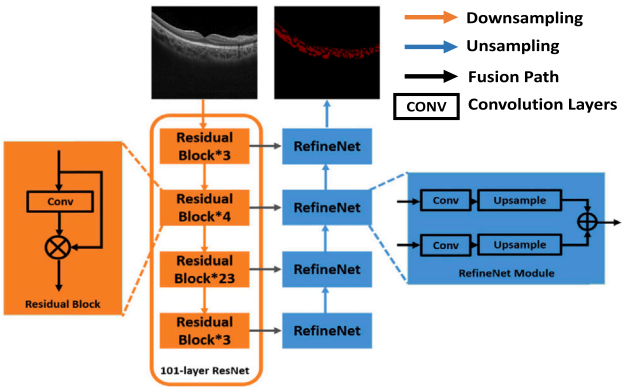


Fig. 5. An example of fully supervised deep learning technique for OCT image segmentation (Liu et al., 2019).

2.2.2. *Semi-supervised deep learning techniques for OCT images*

Liu et al. (2022) introduced a self-ensembling semi-supervised method for retinal fluid segmentation. This method is based on the student–teacher model. These models comprised one encoder and three decoders with modified U-Net architecture, as shown in Fig. 6. These three decoders contain an output probability map, contour map, and distance map that are used to compute the final segmentation map. The final segmentation map successfully distinguishes the close fluid regions. This method also contains the Selective Kernel Module (SKM) (Li et al., 2019b), which allows receptive fields to change according to the fluid scale. The weighted dice, contour, regression, and consistency loss functions are used to train the student model. It has achieved a 76% Mean dice score in the RETOUCH dataset. Sedai et al. (2019) proposed another semi-supervised method for retinal layers segmentation. This method is based on Dense-Unet, which consists of two networks (Li et al., 2018) architecture. The first network is used for soft labeling of unlabeled OCT images and the second network is used for the final segmentation of OCT images. The first network was trained using labeled data with cross-entropy loss and the second network is trained using soft labels with an uncertainty map. It used 700 labeled OCT scans obtained with the Cirrus device. This method has achieved an 82% mean dice coefficient score in the segmentation results.

Heisler et al. (2020) proposed a semi-supervised method with adversarial learning for peripapillary retinal layer segmentation. This method has a GAN (Generative Adversarial Network) based architecture. It is trained on 326 OCT volumes obtained with a custom OCT machine. It has achieved a 93% mean dice coefficient score in the segmentation results. In another study, Liu et al. (2018) proposed a semi-supervised method with adversarial training for retinal layers and fluid areas. This method contains segmentation and a discriminator network. The segmentation network is made by adding Global Convolution Layer (GCL) (Peng et al., 2017) and Boundary Refinement Layers (BRL) (Peng et al., 2017) in the encoder block of the U-Net architecture. The

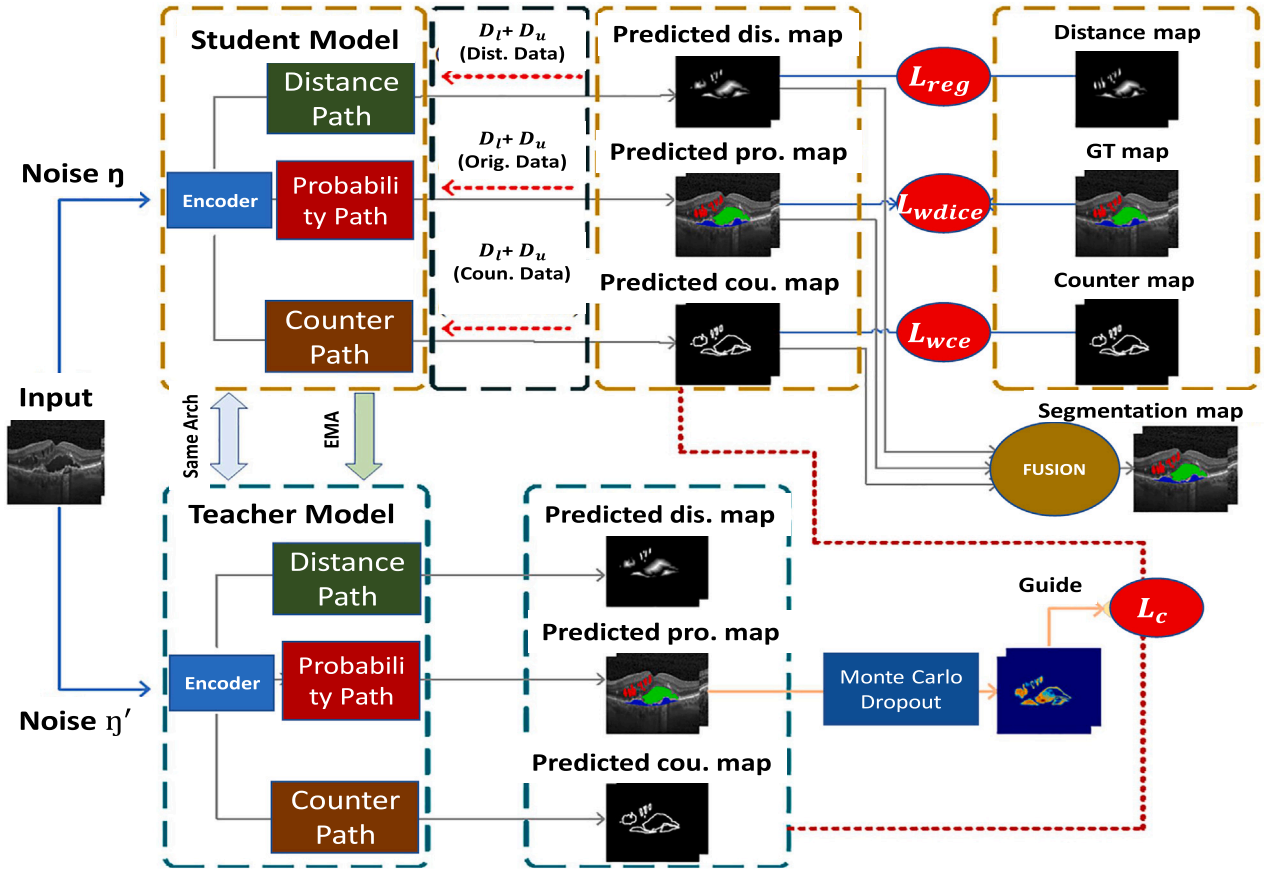


Fig. 6. Semi-supervised deep learning technique for OCT images segmentation (Liu et al., 2022).

discriminator network is the same as the segmentation network except for the GCL and BRL layers that are substituted with the convolution layer in the encoder and decoder blocks, respectively. The discriminator network used the Leaky ReLU activation function instead of the ReLU function in the convolutional layers. The segmentation network is trained by minimizing the total loss function, including weighted cross entropy, weighted dice, semi-supervised loss, and adversarial loss. The discriminator network is trained on discriminator loss, which differentiates the ground truth maps from the segmentation probability map. This method has achieved a 90% mean dice coefficient score on the Duke dataset, and a 93% mean dice coefficient score on the AROI dataset.

2.2.3. *Weakly-supervised deep learning techniques for OCT images*

He et al. (2022) proposed a self-supervised method for retinal fluid segmentation by using point annotations. It applies the contrastive learning tasks technique to learn visual features from unlabeled images. This network is trained by using cross-entropy loss and contrastive learning loss functions. This self-supervised method has achieved a 71% dice coefficient score on the publicly available AI Challenger dataset. Xing et al. (2021) proposed a weakly supervised technique for lesion area segmentation. This method has two training stages. In the first training stage, a classification model with convolutional neural network architecture is trained with image-level annotations. In this stage, a modified Class Activation Mapping (CAM) technique (Zhou et al., 2016) is used to highlight the lesion areas in the final map. In the second training stage, a DeepLab model is used for the segmentation network, which is trained with pseudo labels. This method used 23 OCT volumes of patients with neurosensory retinal detachment. The deepLab model has acquired a 94% dice coefficient score and demonstrated comparable performance to the pixel-level-based annotated data.

Wang et al. (2020) developed a weakly supervised segmentation technique for lesion area segmentation in OCT images. Similar to the previous technique, this method generates pixel-level pseudo labels and then uses them to train the supervised segmentation module. This method is trained on 23 OCT volumes obtained from patients with neurosensory retinal detachment with the help of the DeepLab v3 segmentation network. This method achieved a 92.4% dice coefficient score in segmentation results. Wang et al. (2021) proposed a weakly supervised method for lesion area segmentation in retinal OCT images. This network uses CycleGAN architecture that was trained to translate images with lesions into normal images, as shown in Fig. 7. The lesion regions from the OCT scan are segmented by subtracting the original image from the reconstructed image. This technique obtained an 82% dice coefficient score on a publicly available dataset (Kermany et al., 2018). Soltanian-Zadeh et al. (2021) proposed a weakly supervised technique for ganglion cell segmentation from adaptive optics OCT images. This network has a fully convolutional architecture. It uses human click points for individual ganglion cells in the training stage. It used the Indiana University dataset for training, which was obtained with an OCT machine. This technique achieved 88% recall, 87% precision, and 87% F1 score in segmentation.

3. Datasets and evaluation

The development of deep learning in medical image analysis is only achievable with the public availability of datasets. Hence, in this section, we provide an overview of the OCT datasets, presented in Table 3, which are broadly used for segmentation tasks. We provide details on the accessibility, diseases, tissues scanned, resolutions, and download links to facilitate research requirements.

Table 3 mainly includes 19 representative OCT datasets in this field, where most (13 of 19, 68%) are scanned at the retina, followed by

Table 3
The overview of widely used OCT dataset for segmentation task. The link to each data set is given in the last column of this table.

Ref	Name	Year	Region	Number	Resolution	Diseases	Annotations	Access	Link
Melinščak et al. (2021b)	AROI	2021	Retina	25*128	1024 × 512	AMD	Layer, lesion	OA	Link
Li et al. (2020b)	OCTA-500	2020	Retina	361,600	400 × 640	AMD/Healthy	2D vessel	OA	Link
Farsiu et al. (2014)	AMD-OCT	2014	Retina	384*100	1001 × 1001	AMD/Healthy	3 layers	OA	Link
Chiu et al. (2012)	AMD-OCT2	2012	Retina	20*100	1000 × 100	AMD	3 layers	OA	Link
Chiu et al. (2015)	DME-OCT	2015	Retina	11*61	496 × 768	DME	Layers/lesion	OA	Link
Yang et al. (2021)	Esophagus-OCT	2021	Esophagus	30*784	512 × 512	Barrett's esophagus/Healthy	Esophageal epithelium	OA	Link
Bogunović et al. (2019)	AI Challenger	2019	Retina	49*128	512 × 512	DME/AMD	Lesion	AoR	Link
Bogunović et al. (2019)	RETOUCH	2019	Retina	46*128+24*49	512 × 1024	Fluids	Lesion	AoR	Link
Oguz et al. (2015)	OPTIMA	2015	Retina	22*128+8*49	512 × 1024	Fluids	Lesion	Not	Link
Rashno et al. (2017)	UMN	2017	Retina	725	512 × 1024	Fluids/DME	Lesion	Private	Link
Tian et al. (2016)	Miami-OCT	2016	Retina	610	768 × 496	Diabetic retinopathy	5 layers	OA	Link
Gawlik et al. (2018)	ONH-OCT	2018	Retina	416	384 × 496 × 145	Glaucoma/healthy	5 tissues	OA	Link
Masood et al. (2019)	Choroid-OCT	2019	Retina	525	760 × 456	Healthy	2 layers	AoR	Link
Lee et al. (2017)	CNN-OCT	2017	Retina	1289	432 × 32	Macular edema	Intra-Retinal Fluid (IRF)	AoR	Link
Dos Santos et al. (2019)	Cornea-OCT	2019	Cornea	20,160	512 × 1024	Keratoconic/ healthy	Epithelium, Bowman's layer and stroma	Private	None
Yuan et al. (2022a)	Rat-colon	2021	Colon	10,000	–	Healthy	Layers, muscles	Private	None
Yuan et al. (2022a)	Sheep-airway	2021	Airway	11	–	Healthy	Layers, muscles	Private	None
Li et al. (2019c)	Guinea Pigs-OCT	2019	Esophagus	235	512 × 168	Healthy	Esophageal Layers	Private	None
Pfister et al. (2019)	Dermal-OCT	2019	Skin	100*512	512 × 304	Dermal fillers	Healthy	Private	None

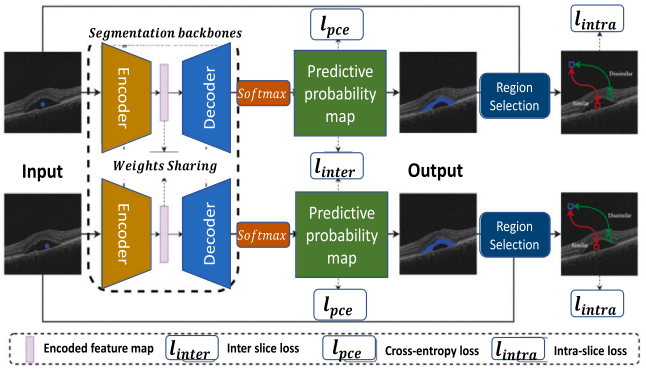


Fig. 7. Weakly-supervised deep learning technique for OCT images segmentation (He et al., 2022).

the skin, esophagus, cornea, colon, and airway. It makes sense due to the large-scale use of OCT systems in ophthalmology clinics. We will initially discuss several popular OCT datasets scanned at the retina, providing additional information such as the institution where it was conducted, the patient cohort, the targeted diseases, the equipment used, the acquisition process, and the scanning parameters.

3.1. Retinal public datasets

Retinal public datasets are sets of data collected from various sources and made available to the public for academic and research purposes. These datasets provide a valuable resource for researchers and clinicians to study the complexity of retinal diseases and to develop new methods for diagnosis and treatment. The Ophthalmological diseases represented in OCT datasets include AMD, DME, DR, glaucoma, fluids caused by one or both of them, and also healthy ones. These diseases can cause structural abnormalities, such as distortions, thinning, and breakage, which are the targets of related segmentation tasks. In general, retinal layer (6/16;37.5%) and lesion (6/16;37.5%) segmentation are mainstream, compared to the vessel, choroid, and other tissues segmentation.

The most prominent retinal public datasets include the Eye Disease Database (EDDB) (Camara et al., 2022), the Ophthalmic Image Analysis Society (OIAS) Image Dataset (Yang et al., 2022), the Retinal OCT Image Dataset (ROCID) (Venkatesh et al., 2022) from the University of Iowa, the National Institutes of Health Age-related Eye Disease Study (AREDS) (Yang et al., 2022) and the EyePACS Digital Dataset. The Eye Disease Dataset (EDDB) contains the largest collection of retinal images, with over 18,000 image sets from over 4,000 patients. The OIAS image dataset provides image sets of various retinal diseases such as age-related macular degeneration, diabetic retinopathy, and glaucoma. The ROCID contains over 1,200 OCT images and associated clinical data from a variety of retinal diseases. The AREDS dataset contains a large set of images from the Age-related Eye Disease Study. Finally, the EyePACS Digital Database (Venkatesh et al., 2022) contains several thousand images from patients who have been diagnosed with various retinal diseases.

3.2. Other publicly available datasets

The Vision and Image Processing (VIP) Laboratory (Farsiu et al., 2014; Srinivasan et al., 2014; Chiu et al., 2012, 2015; Yang et al., 2021) from Duke University released a public dataset AROI (Annotated Retinal Optical Images) (Melinščak et al., 2021a). This dataset contains 25 annotated images of AMD (age-related macular degeneration) patients. These macular SD-OCT volumes were recorded using a Zeiss Cirrus HD OCT-4000 device. Each OCT volume consists of 128 B-scans with a resolution of 1024 × 512 pixels, and the size of each pixel is 1.96 × 11.74 μm. This dataset has three types of annotations: Pigment Epithelial Detachment (PED), subretinal fluid and subretinal hyperreflective material, and Intra-Retinal Fluid (IRF). To calculate the intra-observer error, an expert has annotated 75 B-scans twice. During a second annotation, the expert did not have access to the previous annotations.

The VIP Laboratory released (Farsiu et al., 2014) another dataset called the A2 A SD-OCT dataset. This is collected by four A2 A SD-OCT clinics (Devers Eye Institute, Duke Eye Center, Emory Eye Center, and National Eye Institute). This dataset includes Age-Related Eye Disease Study 2 (AREDS2) and Ancillary SD-OCT (A2 A SD-OCT) data. In these datasets, the subjects should meet the following inclusion criteria: 50 to 85 years old, with both eyes or one eye exhibiting moderate AMD with large drusen (125 mm above) and advanced AMD in the fellow eye, and no history of vitreoretinal surgery or eye disease that may affect vision. For all participants, 0° and 90° rectangular volumes centered at the fovea (defined as volumes collected with the fast scan direction

oriented, horizontally and vertically) were recorded for both eyes using 1000 A-scans per B-scan and 100 B-scans per volume. In the A2 A SD-OCT dataset, they studied 119 of the 122 control participants without AMD and 314 of the 345 participants with AMD that had no eye illness at the baseline. Based on these, each subject's eligible eyes were randomly selected as study eyes.

In Chiu et al. (2012), the SD-OCT dataset used for cross-validation included volume scans obtained from 45 patients: 15 normal patients, 15 dry AMD patients, and 15 DME patients. All SD-OCT volumes were acquired using Spectralis Device at Duke, Harvard, and Michigan. The Duke Enterprise collected the data from the patients that came for retinal practices in the Duke Eye Center (Chiu et al., 2015) and used the DME (ICD-9 362.07) billing code associated with their visit. An ophthalmologist has identified six clinically imaged patients with severe DME pathology and varying image quality using a standard Spectralis Device with a 61-line plethysmographic protocol. The photographer determines the average of the B-scans, and each averaged B-scan ranges from 9 to 21 raw images.

The human subjects study dataset was conducted with institutional review board approval from Duke University Health System (Pro0090173) and the University of North Carolina (UNC; 17-3037). These subjects (Yang et al., 2021) were recruited from patients undergoing routine care endoscopy at UNC Healthcare. Initially, they recruited 54 patients, and 30 were successfully obtained using OCT (784 B-scans in total). Thirty subjects were studied, of which six had non-dysplastic Barrett's Esophagus (BE) and one had low-grade dysplastic BE. The captured OCT data were manually cropped around regions of interest (ROI; esophagus) and independently labeled by three graders experienced in evaluating OCT images. The annotation of the most experienced grader (Grader #1) served as the gold standard, and the annotations of the other two graders were used to test for inter-rater variability. Table 3 shows the statistics of publicly available datasets.

3.3. Private datasets

Private datasets (Kermany et al., 2018) are usually collected and distributed by private companies, such as eye care providers (Lee and Maa, 2023) or medical imaging companies (Guadalupe Mori et al., 2023). These companies may collect data from patients over a period of time, and then make the data available for researchers to use for their studies. The data may include a variety of different types of OCT images, including images of the anterior chamber, the vitreous, the retina, and the choroid. These images can provide valuable insight into the structure and functioning of the eye and allow researchers to better understand and diagnose eye diseases. There are the following private datasets, which are accessible on request.

3.3.1. The UMN OCT dataset

It is a large collection of optical coherence tomography (OCT) scans from the University of Minnesota (Cramer, 2023). It is a private dataset. It contains over 18,000 OCT scans from a variety of sources, including retinal, choroidal, and anterior segment images. Each scan is accompanied by detailed metadata, including age, gender, and diagnosis. This data can be used to study a variety of ophthalmologic conditions and diseases, including glaucoma, age-related macular degeneration, diabetic retinopathy, and more. The dataset also contains 3D volumetric reconstructions of OCT scans, allowing for a more detailed study of retinal layers and other structures.

3.3.2. The STARE (structured analysis of the retina) dataset

It contains over 400 images of retinal scans from 20 different diseases, including diabetic retinopathy, glaucoma, macular degeneration, and other eye diseases (Almustafa et al., 2023). Each image is labeled with a disease name and the annotations include the presence or absence of features such as hard exudates, microaneurysms, and cotton wool spots. The dataset also includes additional metadata such as patient age, gender, and diagnosis. This dataset has been used to develop a variety of image analysis algorithms for automated retinal image analysis and has made significant contributions to the field of medical image analysis.

3.3.3. The stanford online OCT dataset

It includes over 25,000 images from 2000 patients who have been diagnosed with a variety of retinal diseases, including age-related macular degeneration and diabetic retinopathy (Chinn et al., 2023). The images have been carefully annotated by experienced ophthalmologists and are organized into distinct folders, including OCT images, OCT segmentation masks, and patient metadata.

3.3.4. The MESSIDOR (multi-ethnic study of susceptibility and risk of age-related macular degeneration) dataset

It includes over 1000 participants from five different ethnic groups, including African-Americans, Latinos, Asians, Caucasians, and Native Americans (Jacoba et al., 2023). The dataset includes both demographic information as well as medical data from eye exams, including visual acuity and macular thickness measurements. The dataset was collected for a study to identify genetic and environmental factors that can increase the risk of AMD, and to create a risk score to predict the likelihood of developing the disease.

3.3.5. The SEED (spectral-domain eye OCT) dataset

It is a collection of high-resolution OCT images of the human eye obtained from a spectral-domain OCT system (Fenner et al., 2023). It contains over 40,000 retinal images of both healthy and diseased eyes, with a wide range of pathology. The images have been collected from multiple sites around the world and have been anonymized to protect patient privacy. The dataset is divided into two parts: the public dataset, which contains 33,000 images, and the private dataset, which contains an additional 7500 images. The images are accompanied by a dataset of manual segmentations and annotations created by experienced physicians.

3.3.6. BIT and ABI-lab dataset

In addition to the multiple publicly available retina OCT datasets, the Biophotonics Imaging Technology Lab (BIT) at Johns Hopkins University and Advanced Biophotonics Imaging Laboratory (ABILab) at the Chinese University of Hong Kong have various in-house datasets of endoscopic OCT images (Yuan et al., 2016, 2022a; Park et al., 2017; Kermany et al., 2018). In contrast to retinal OCT, endoscopic OCT moves along specific lengths along luminal organs inside the body of test subjects. The OCT catheter is sent through the organ of interest using a plastic sheath. At each imaging depth, the catheter rotates to capture a-lines at various angles around the entire circumference. The collation of all a-lines at each step would be a b-frame showing the cross section at that imaging depth. Then the catheter would be pulled back in fixed steps (pitch) along the organ to image b-frames at each step (Yuan et al., 2017).

(i) Rat datasets: Rat colon images were collected using 800 nm ultra-high resolution endoscopic spectral domain OCT (Park et al., 2017; Yuan et al., 2022a). A circumferential OCT catheter was used to acquire an image in 3 cm of rat colon (*in vivo*) with pitch size 20 μm. Every centimeter of the rat colon image includes 500 corresponding images. They combined 10,000 raw images from 20 scans. In this dataset, the layers of interest are colonic mucusa (CM), sub mucusa (SM), and muscularis externa (ME). During manual labeling, these layers can be distinguished by the human eye due to the ultra-high resolution OCT imaging.

(ii) Sheep datasets: Sheep small airway (fifth generation airways, with diameter 1 mm or less) OCT scans were obtained *in vivo* using 800 nm ultra-high resolution SD-OCT (Yuan et al., 2017, 2022b). The region of interest is the airway walls consisting of 4 layers: epithelium (EP), basement membrane (BM), smooth muscle (SM), and adventitia

Table 4
Overview of labeling tools, which support images, videos, and point clouds data.

Tool name	Supported format	Type
Amazon SageMaker (Das et al., 2020)	Text, images, video, 3D cloud points	Automatic
Label Studio (Tkachenko et al., 2020)	Text, images, video, multi-domain	Automatic
Labelbox (Labelbox, 2019)	Images, video, text, audio, geospatial	Automatic
ITK-SNAP (Yushkevich et al., 2016)	Images, video	Manual
Tagtog (Cejuela et al., 2014)	Text	Automatic/manual
Playment (Magistretti, 2017)	Images, text, video, audio, sensor fusion, geo-local	Automatic
Dataturk (Pizenberg et al., 2018)	Text, images ,video	Manual
LightTag (Perry, 2021)	Text	Automatic/manual
Superannotate (Panferov et al., 2020)	Text, images, video, audio, lidar	Automatic/manual
CVAT (Sekachev et al., 2020)	Image, video	Automatic

(AD), along with two other classes cartilage (CT) and Blood Vessels (BV). These are also visible to the naked eye for manual annotating.

(iii) Esophagus dataset: Endoscopic OCT esophageal images were taken (in vivo) from five guinea pigs (male, Hilltop, Scottsdale, PA) utilizing an 800 nm endoscopic OCT system (with an OCT endoscope of a 1.3-mm outer diameter) (Li et al., 2019c,a). Before being inserted into the esophagus, the endoscope was cleaned. The Johns Hopkins University Animal Care and Use Committee (ACUC)-approved the animal handling procedure. It contains 235 OCT images from two guinea pigs with EOE, one control, and the hand segmentation layer maps for each of these images.

3.4. OCT dataset labeling

In the last decade, numerous artificial intelligence (AI) based image labeling tools have been developed to label images and videos, as shown in Table 4. Some of them are specially used to label medical images like Hitachi semantic segmentation editor (Mandrioli, 2019), CVAT, ITK-SNAP, etc. The Hitachi semantic segmentation editor (Mandrioli, 2019) is used to label the endoscopic OCT images. The polygon drawing tool is used to plot points along the boundary of any region, and the bounded area would be labeled as the selected class. Since all classes are directly connected, the same boundary line is always shared between two classes with no space. Another classical and widely used labeling tool in radiology images (CT, MRI) is called ITK-SNAP (Yushkevich et al., 2006).

3.5. OCT dataset analysis

We have reviewed different OCT datasets that are used in OCT segmentation. Each OCT dataset has different structures and features due to different parameters setting of OCT devices on different animals. The rat colon dataset has the most straightforward structure of three classes of interest in fixed order stretching along the entire circumference. The third class, ME, is usually of a much lower proportion. So, this dataset suffers from class imbalance. Sheep airways are more challenging due to the possibility of cartilage and blood vessels appearing inside other layers (adventitia). In addition, these two classes do not stretch along the entire circumference but appear as isolated patches in the b-frames of the airway cross-section.

The smooth muscle class follows a structure that runs along the circumference, but it is often discontinuous. Consequently, of the six classes of interest, three classes are discontinuous and three classes are ongoing. This leads to an even more pronounced class imbalance in the sheep airway dataset. Most of the dog airway annotations only contain the smooth muscle class as labeled. This makes it the most difficult dataset for endoscopic OCT segmentation, due to its severe class imbalance coupled with the irregular shape of the b-frame images.

Table 5
Typical layout of a confusion matrix.

Ground Truth/Prediction	1	2	3
1	G_1P_1	G_1P_2	G_1P_3
2	G_2P_1	G_2P_2	G_2P_3
3	G_3P_1	G_3P_2	G_3P_3

4. Performance comparison

In this section, we discussed the quantitative and qualitative analyses of the state-of-the-art OCT segmentation methods. It includes widely used evaluation metrics, parameter settings, and performance comparisons.

4.1. Evaluation metrics

Most of the evaluation metrics for image segmentation are calculated by a confusion matrix (Moccia et al., 2018), also known as a contingency table. This table comprises the total number of ground truth pixels for each class and the total number of predicted pixels for their corresponding classes. An overview of the structure of the confusion matrix can be seen in Table 5. Where $G_yP'_y$ refers to the ground truth (G_y) of the pixel and the predicted pixels(P_y).

The content of each cell in the confusion matrix would be the number of pixels that correspond to that condition. For instance, the cell labeled G_1P_2 tabulates the number of pixels that were labeled as class 1 but were predicted as class 2. When the row and column indices align, i.e., the cells on the diagonal of the matrix, the predictions that match the ground truth are referred to as true positive (TP) predictions.

For each row, the cells that are not present in the diagram are false negative (FN) predictions since the model has outputted a different class compared to the ground truth class. Each column, on the other hand, shows the predicted labels of being in a particular class. If there is a deviation from the diagonal with respect to a column, this would be a false positive (FP) prediction, as the pixel was not supposed to be classified into that class.

4.1.1. Dice coefficient

The Dice coefficient is a measure of similarity between two sets of data and is commonly used in medical image analysis to evaluate the performance of algorithms for segmenting organs or other structures in OCT, computed tomography (CT), and magnetic resonance imaging (MRI) scans. The Dice coefficient is calculated by measuring the overlap between the two sets of data and is typically expressed as a fraction between 0 and 1, where 1 indicates a perfect match. *Fx* measures can

be considered as providing certain weights for the true positives of intersecting over the union. When $x = 1$, it is known as the dice score.

$$F1score(Dice) = \frac{2 \times TP}{2 \times TP + FN + FP} \quad (1)$$

4.1.2. Pixel accuracy

Pixel accuracy refers to the proportion of correctly labeled pixels out of all possible pixels (n). This is simply calculated as:

$$Accuracy = \frac{TP}{n} \quad (2)$$

The proper distribution of each class is essential for obtaining accurate pixel accuracy. When one class occupies a disproportionately large amount of pixels, pixel accuracy results may be unreliable.

4.1.3. Intersect over union (IoU)

It is also known as the Jaccard score. It is calculated for each class separately. It is the intersection between ground truth and prediction labels (TP) divided by the combination of True Positive, False Negative, and False Positive. It is calculated as:

$$IoU = \frac{TP}{TP + FN + FP} \quad (3)$$

4.1.4. Precision

Precision specifically looks at how many of the positive predictions of a class are truly correct.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

4.1.5. Recall

The recall specifically looks at how many of the positive ground truths of a class are correctly predicted to be of that class.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

4.1.6. Area under ROC curve

The receiver operative characteristic curve is commonly used for single-class analysis. It involves changing the decision threshold for that class to plot out the true positive rate (Recall) and false positive rate. It uses the probability of correctly classifying randomly selected pixels when any chosen fixed threshold is applied to the overall input data (Hanley and McNeil, 1982).

$$FalsePositiveRate = \frac{FP}{FP + TN} \quad (6)$$

So, the area under the ROC curve (AUROC) should be equal to 1, which is only possible when the true positive rate = 1 and the false positive rate = 0. The worst case would result in AUROC = 0.5, in which any pixel would be randomly assigned to either belong to a class or not.

4.1.7. Mean absolute error (MAE)

It is the average absolute difference between the class index of predictions and ground truths. It can be useful when the index number carries certain information within itself.

$$MAE = \frac{\sum |y' - y|}{n} \quad (7)$$

4.1.8. Mean square error (MSE)

It is a measure of the average of the squares of the errors or deviations from the predicted value. It is used to measure the accuracy of a regression model and is the most commonly used loss function for regression problems. It is calculated by finding the average of the squares of the differences between the predicted values and the actual values.

$$MSE = \frac{\sum (y' - y)^2}{n} \quad (8)$$

4.1.9. Hausdorff distance

Hausdorff distance is another metric that has been seeing increased usage in medical image segmentation (Karimi and Salcudean, 2019). Unlike the earlier metrics that look at every pixel, Hausdorff distance only considers the boundary pixels of every patch of pixels belonging to the same class. It calculates the distance between every corresponding point within one or two boundaries. Although by definition, Hausdorff distance only considers the maximum distance between two points. The average Hausdorff distance is also used to account for all boundary pixels of a patch.

$$HD(Gt, Pd) = mean_{p_{pd} \in Pd} min_{p_{gt} \in Gt} \| p_{gt} - p_{pd} \|^2 \quad (9)$$

Where p_{pd} is predicted pixels and p_{gt} is ground truth pixels during segmentation.

4.1.10. Parameters

The number of parameters is another way to compare the effectiveness of different deep learning networks. One parameter count refers to one learnable weight in the network that will be updated throughout the training process. The larger the number of parameters, the more space is required to save the completed model, and more computational resource is needed for processing the inputs.

4.2. Performance evaluation

In this section, we compared the performance of different OCT segmentation models in terms of precision, recall, F-measure, dice coefficient, mean absolute error, etc. Dos Santos et al. (2019) technique has the highest performance in terms of precision and recall, as shown in Table 6. The highest performance is due to the easiest segmentation task compared to other methods relatively. This method segmented different layers in the cornea, which have quite predictable structures compared to lesion segmentation. Similarly, the fully supervised deep learning methods (Li et al., 2020a; Schlegl et al., 2018) have similar precision but different recall values. These methods contain two sub-techniques. The first technique segments the retinal layers, which are easy to segment, while the second technique segments retinal fluid, which is difficult. In terms of F1-score, weakly-supervised learning methods (He et al., 2022; Xing et al., 2021; Wang et al., 2020, 2021) have good results. These methods (Xing et al., 2021; Wang et al., 2020) include Deeplab architecture and have the highest performance in terms of precision, recall, and F-measure. This implies the advantage of using DeepLab for segmentation tasks such as lesion segmentation. The machine learning algorithms (Guo et al., 2019) with support vector machine have higher recall than random forest algorithm (Lang et al., 2015). It is likely that the Support Vector Machine (SVM) algorithm is more suitable for this task than the random forest algorithm.

5. Discussion

5.1. Traditional machine learning techniques

The machine learning methods have a few limitations, such as pre-feature extraction for training a model and lack of high-level features during segmentation. These methods have different applications that require pre-processing steps to segment the region of interest in OCT images. In Huang et al. (2018), they extract the first features from the gray-level co-occurrence matrix and store them using the RFC algorithm. So, this illustrates that there are some complicated steps needed before training an algorithm. The other limitation of these methods is processing time. As shown in Table 6, these methods are generally slower in segmentation tasks as compared to the deep learning methods. On the other hand, they do not need annotated data for training, which is the advantage of machine learning algorithms in segmentation applications. As in Amrute et al. (2017), K-means clustering is used to segment the region of interest in the image by grouping pixels based on their intensity and then applying a post-processing segmentation task. In general, machine learning models have a good performance on different datasets, as shown in Table 6.

Table 6
Performance of segmentation methods.

Ref.	Method	Dataset	Precision	Recall	F1-score	Inference time	IoU	MAE
Huang et al. (2018)	Machine learning	Custom	0.82	0.81	0.83	2-3 min	0.82	– –
Guo et al. (2019)	Machine learning	Custom	0.91	0.90	0.90	– –	0.89	– –
Lang et al. (2015)	Machine learning	Custom	0.85	0.79	0.84	– –	0.91	– –
Lang et al. (2013)	Machine learning	Custom	0.82	0.82	0.79	2.4 s	0.78	1.21 ± 1.45
Amrute et al. (2017)	Machine learning	Custom	0.93	0.92	0.92	<3 s	0.92	– –
Li et al. (2020a)	Fully supervised	Duke	0.91	0.92	0.90	0.52 s	0.90	– –
Liu et al. (2019)	Fully supervised	Custom	– –	– –	– –	0.72s	– –	– –
Schlegl et al. (2018)	Fully supervised	Custom	0.76	0.82	0.80	0.61 s (Spectralis), 0.35 (Cirrus scans)	0.91	– –
Dos Santos et al. (2019)	Fully supervised	Custom	0.98	0.97	0.97	0.0249 s	0.93	– –
Lee et al. (2017)	Fully supervised	Custom	0.91	0.92	0.91	13.2 s	0.91	– –
Liu et al. (2022)	Semi-supervised	RETOUCH	0.73	0.74	0.76	– –	– –	– –
Sedai et al. (2019)	Semi-supervised	Custom	0.83	0.80	0.82	– –	0.80	– –
Liu et al. (2018)	Semi-supervised	Duke	0.91	0.90	0.90	0.1 s	– –	– –
Heisler et al. (2020)	Semi-supervised	Custom	0.94	0.93	0.93	– –	0.93	– –
He et al. (2022)	Weakly-supervised	AI Challenger	0.75	0.71	0.72	0.042 s	– –	– –
Xing et al. (2021)	Weakly-supervised	Custom	0.93	0.91	0.94	– –	0.92	– –
Wang et al. (2020)	Weakly-supervised	Custom	0.93	0.90	0.92	– –	0.90	– –
Wang et al. (2021)	Weakly-supervised	Publicly available Kermany et al. (2018)	– –	– –	0.82	0.039 s	0.84	– –
Soltanian-Zadeh et al. (2021)	Weakly-supervised	Custom	0.87	0.88	0.87	– –	0.85	– –

5.2. Fully supervised deep learning techniques

These methods have demonstrated their effectiveness in various tasks by using deep learning networks in OCT segmentation applications, as shown in Figs. 8 and 9. One of the difficulties of training these networks is the class imbalance problem. In some OCT applications, the region of interest is tiny, like lesion areas in OCT images. Segmenting this small area is a challenging task in deep learning models. To solve this problem, a focal-loss function technique is used in this model (Li et al., 2020a). According to the survey on small area segmentation in OCT images (Johnson and Khoshgoftaar, 2019), the choice of the backbone network is very important. For example, in Li et al. (2020a), models trained on different backbones, such as ResNet101, Inception-V3 and modified Xception that have quite different results. One disadvantage of fully supervised methods is training time. They take a long training time to get better results on a large number of images. For instance, in Lee et al. (2017), it took around seven days to train a network. Generally, fully supervised approaches demonstrate low latency and produce high precision, recall, and F-measure scores in OCT image segmentation, as shown in Table 6.

5.3. Semi-supervised deep learning techniques

These methods have good accuracy in the segmentation of OCT images as compared to the machine learning algorithms as shown in Table 6. In Liu et al. (2022), the semi-supervised based segmentation results achieved a mean dice coefficient score of 76%, compared to 73% without the use of a semi-supervised learning network in the RETOUCH dataset. Same in Sedai et al. (2019), the segmentation results with the semi-supervised network have achieved 82% mean dice coefficient score and 80% without the semi-supervised network by using a custom dataset. In Heisler et al. (2020), the model has achieved a 93.04% mean dice coefficient score with a semi-supervised based network and 92.85% without a semi-supervised network. This difference demonstrates that semi-supervised learning-based models have good segmentation accuracy. One of the limitations of these methods is their segmentation capability. They are not efficient in segmenting the small multi-regions in OCT images. It is not easy to generalize these segmentation models into a global OCT segmentation model. The other problem is that these models are trained for different datasets, which makes it difficult to compare their performance on the same dataset.

5.4. Weakly-supervised deep learning techniques

These models segment the region of interest in OCT images without pixel-level annotation, as shown in Table 6. These models take less processing time and perform well for different applications. One of the difficulties of these methods is their complicated structure. They usually have pre-processing and post-processing steps. For example, in Xing et al. (2021), Ruiwen Xing et al. proposed a weakly supervised segmentation model that takes three steps to train: training classification, class activation maps to output saliency maps, and level set method. The limitation of these models is a customized task. It is hard to use these models for multiple tasks. For example, this technique (He et al., 2022) relies on the fact that fluid and background regions should have similar intensities. So, the same intensity areas are not common in all OCT images.

6. Future direction

AI models with a large scale of data and parameters have recently come to light (Qiu et al., 2023; Yuan et al., 2022b; Thiboutot et al., 2022; Chen et al., 2020; Yuan et al., 2020; Luo et al., 2023a; Shahid et al., 2022; Nawaz et al., 2023; Arsalan et al., 2022; Shi et al., 2023). These models have been pre-trained and have shown great results in different tasks (Nawaz and Yan, 2020b; Nawaz et al., 2022a,b; Nawaz and Yan, 2020a; Nawaz et al., 2019; Abaxi et al., 2023,?; Luo et al., 2023b). Chat-GPT is one example of this, making people think about the possibilities large AI models can bring to our lives. In healthcare informatics, large AI models are creating new strategies. Due to deep learning, there is now a lot more data in the biomedical and healthcare industry, which is essential for the development and improvement of AI models for health-related topics.

The explainability of deep learning models is important for clinical applications. To have stable segmentation models, the researcher should focus on machine learning and deep learning models to extract high-level features from different medical modalities. It can be seen that traditional segmentation algorithms such as level set (Xing et al., 2021) are combining deep learning methods with conventional machine learning methods to get good results. Further research could be done to integrate deep learning methods with more traditional segmentation methods.

To reduce the class imbalance problem in the future, we can incorporate relevant loss functions into the networks, such as focal loss.

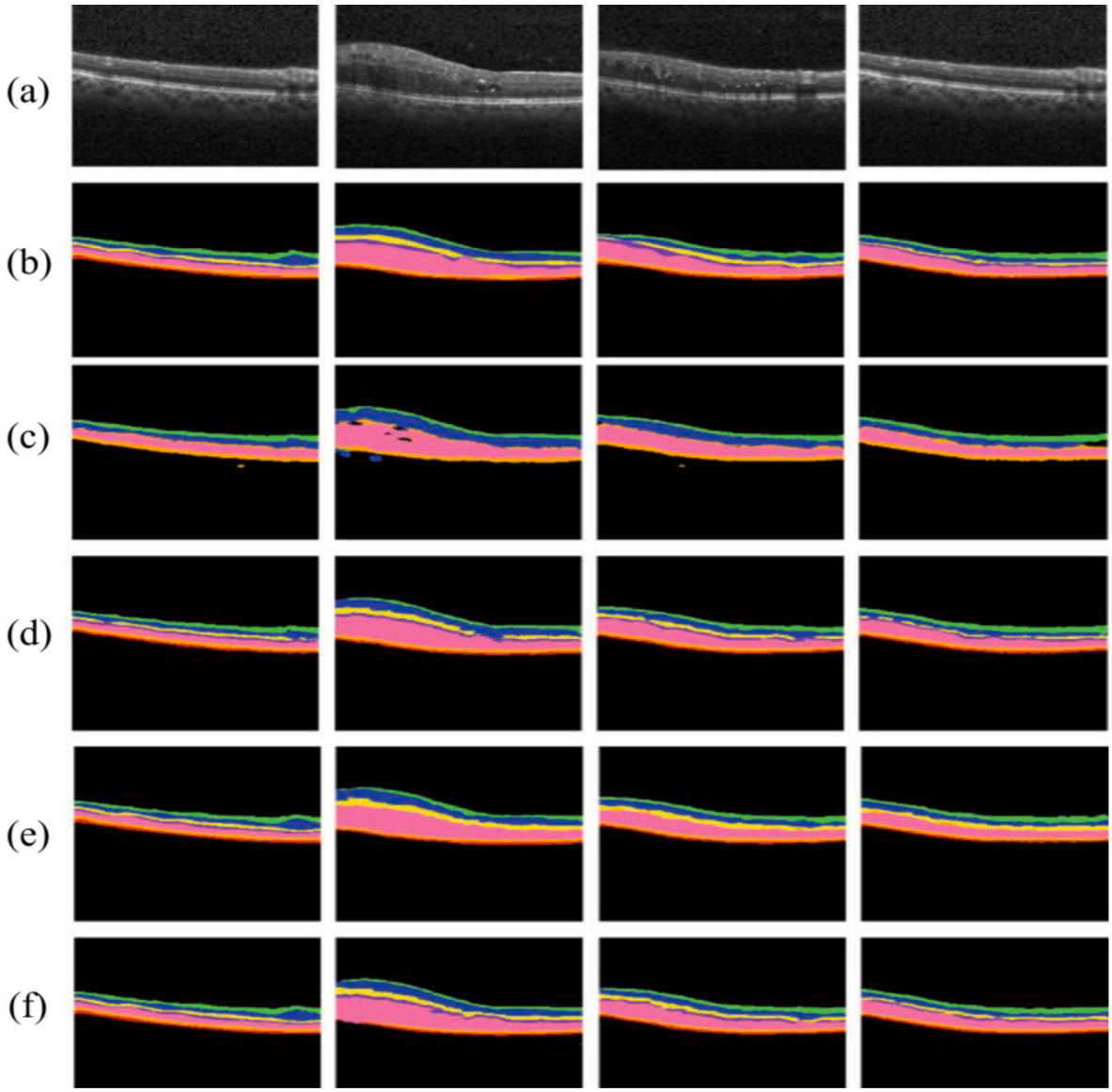


Fig. 8. Results of Deep learning based OCT images techniques. (a) shows the original OCT images, (b) shows the ground truth, (c) shows the result of SegNet technique (Badrinarayanan et al., 2017), (d) shows the results of U-Net technique (Ronneberger et al., 2015), (e) shows the results of LROA-S technique (Devalla et al., 2018), and (f) shows the results of LROA-U technique (Lazaridis et al., 2020). The 10-layers segmented are the nerve fiber layer (NFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), outer limiting membrane (OLM), photoreceptor inner segment (IS), photoreceptor outer segment (OS), and pigment epithelium layer (PEL).

We can modify the current techniques to make them generalize for a specific task. One of the new techniques in image segmentation is reinforcement learning (Minaee et al., 2021). So, in the future, we can emphasize this area to make a global deep learning segmentation method. To build robust models, techniques must be tested on various datasets to ensure reproducibility. Currently, techniques are built on private datasets. So, there is a need to have more publicly available datasets for segmentation.

Another interesting and valuable direction is to consider the credibility of labels. Due to the variety and complexity of diseases, different experts may have different annotations at the same regions or pixels, where we cannot tell right from wrong (or no right and wrong). Breaking the ideal assumption about annotations may introduce more ideas into the methods.

7. Conclusion

In this paper, we reviewed state-of-the-art OCT segmentation techniques. An in-depth review of both conventional machine learning and recent deep learning approaches was conducted to analyze the accuracy of Optical Coherence Tomography (OCT) image segmentation. Machine and Deep learning have demonstrated promise in improving the precision of segmentation in OCT images, and numerous studies have been conducted to examine the efficacy of different OCT segmentation algorithms. Additionally, several datasets have been created to evaluate the potential of the machine and deep learning for OCT image segmentation.

The application of machine learning and deep learning techniques to OCT image segmentation is still relatively new, and numerous challenges must be overcome before it can be adopted in clinical practice. This includes the need for more diverse datasets, more efficient training algorithms, and more reliable segmentation algorithms. This paper summarizes and analyzes recent methods in OCT image segmentation and proposes ideas for further exploration of the field. We hope that our research will provide a useful starting point for new researchers, allowing them to develop more advanced OCT segmentation methods.

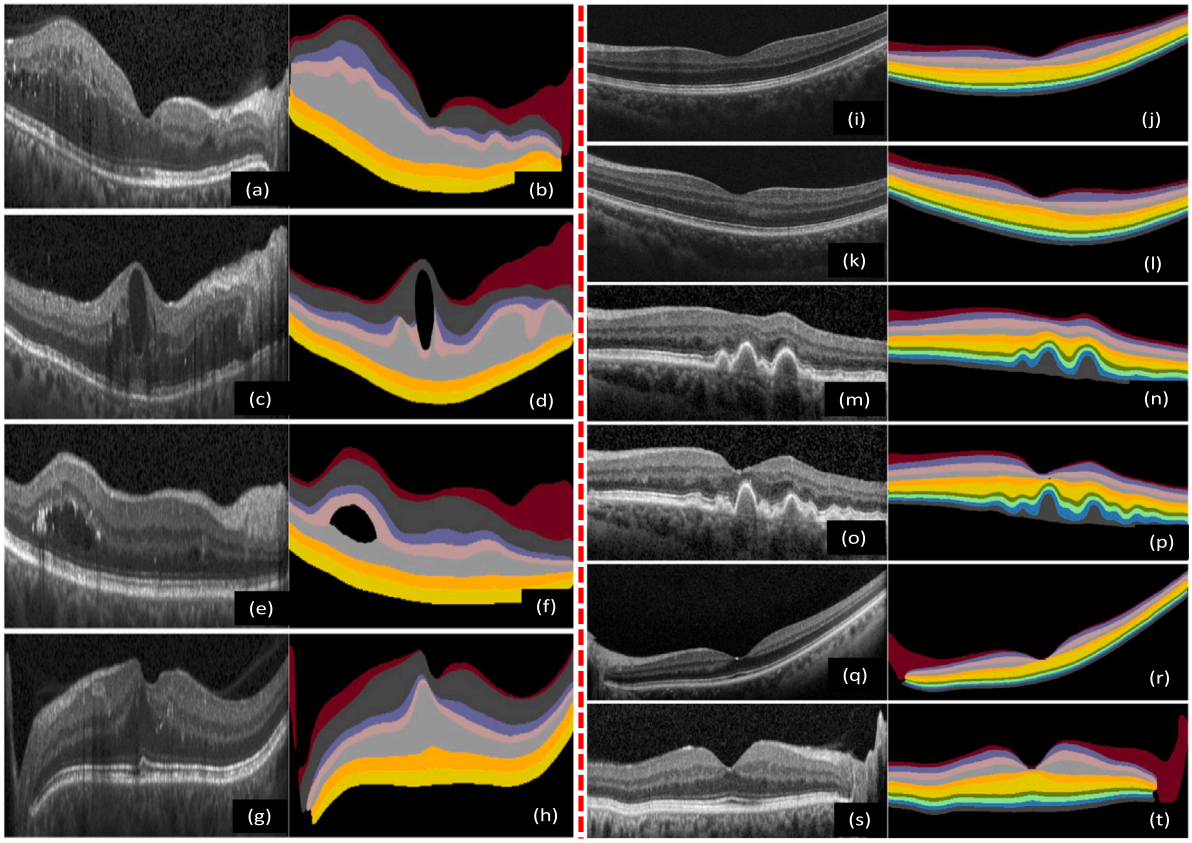


Fig. 9. Left Image: Retinal 7-layers segmentation results (Moradi et al., 2023). Where *a*, *c*, *e*, and *g* show the original OCT images from the Duke data set while *b*, *d*, *f*, and *h* are the corresponding segmentation results. **Right Image:** Retinal 10-layers segmentation results. Where *i*, *k*, *m*, *o*, *q* and *s* show the original OCT images from the Duke data set while *j*, *l*, *n*, *p*, *r* and *t* are the corresponding segmentation results.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

I want to convey my sincere gratitude to Mr. Hing Lam Chang and Mr. Minqing Zhang for using their abilities to help me to enhance this manuscript. This work is supported in part by the Research Grants Council (RGC) of Hong Kong SAR (ECS24211020, GRF14203821, GRF14216222), the Innovation and Technology Fund (ITF) of Hong Kong SAR (ITS/240/21), and the Science, Technology and Innovation Commission (STIC) of Shenzhen Municipality, China (SGDX20220530111005039).

References

Abaxi, S.M.D., Nawaz, M., Shi, P., Wei, H., Abbasi, S.A., Yuan, W., 2023. LiDeOCTNet: A lightweight OCT-aware framework for segmentation of irregularly layered tissue structures. TechRxiv.

Abbas, Q., Qureshi, I., Yan, J., Shaheed, K., 2022. Machine learning methods for diagnosis of eye-related diseases: a systematic review study based on ophthalmic imaging modalities. Arch. Comput. Methods Eng. 29 (6), 3861–3918.

Almustafa, K.M., Sharma, A.K., Bhardwaj, S., 2023. STARC: Deep learning Algorithms' modelling for STructured analysis of retina classification. Biomed. Signal Process. Control 80, 104357.

Amaechi, B., Higham, S., Podoleanu, A.G., Rogers, J.A., Jackson, D.A., 2001. Use of optical coherence tomography for assessment of dental caries: quantitative procedure. J. Oral. Rehabil. 28 (12), 1092–1093.

Amrute, J.M., Athanasiou, L., Rikhtegar, F., Jose, M., Camarero, T.G., Edelman, E.R., 2017. Automated segmentation of bioresorbable vascular scaffold struts in intracoronary optical coherence tomography images. In: 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering. BIBE, IEEE, pp. 297–302.

Arsalan, M., Khan, T.M., Naqvi, S.S., Nawaz, M., Razzak, I., 2022. Prompt deep lightweight vessel segmentation network (PLVS-Net). IEEE/ACM Trans. Comput. Biol. Bioinform. 20 (2), 1363–1371.

Aumann, S., Donner, S., Fischer, J., Müller, F., 2019. Optical coherence tomography (OCT): principle and technical realization. High Resolut. Imaging Microsc. Ophthalmol. 59–85.

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39 (12), 2481–2495.

Barlis, P., Serruys, P.W., Gonzalo, N., van der Giessen, W.J., de Jaegere, P.J., Regar, E., 2008. Assessment of culprit and remote coronary narrowings using optical coherence tomography with long-term outcomes. Am. J. Cardiol. 102 (4), 391–395.

Bogunović, H., Venhuizen, F., Klimscha, S., Apostolopoulos, S., Bab-Hadiashar, A., Bagci, U., Beg, M.F., Bekalo, L., Chen, Q., Ciller, C., et al., 2019. RETOUCH: the retinal OCT fluid detection and segmentation benchmark and challenge. IEEE Trans. Med. Imaging 38 (8), 1858–1874.

Boone, M.A., Jemec, G.B., Del Marmol, V., 2015. Differentiating allergic and irritant contact dermatitis by high-definition optical coherence tomography: a pilot study. Arch. Dermatol. Res. 307 (1), 11–22.

Bowd, C., Zangwill, L.M., Berry, C.C., Blumenthal, E.Z., Vasile, C., Sanchez-Galeana, C., Bosworth, C.F., Sample, P.A., Weinreb, R.N., 2001. Detecting early glaucoma by assessment of retinal nerve fiber layer thickness and visual function. Invest. Ophthalmol. Vis. Sci. 42 (9), 1993–2003.

Camara, J., Rezende, R., Pires, I.M., Cunha, A., 2022. Retinal glaucoma public datasets: what do we have and what is missing? J. Clin. Med. 11 (13), 3850.

Cejuela, J.M., McQuilton, P., Ponting, L., Marygold, S.J., Stefancsik, R., Millburn, G.H., Rost, B., FlyBase Consortium, et al., 2014. Tagtog: interactive and text-mining-assisted annotation of gene mentions in PLOS full-text articles. Database 2014.

Chen, D., Yuan, W., Park, H.-C., Li, X., 2020. In vivo assessment of vascular-targeted photodynamic therapy effects on tumor microvasculature using ultrahigh-resolution functional optical coherence tomography. Biomed. Opt. Express 11 (8), 4316–4325.

Chen, Q., Zeng, L., Lin, C., 2023. A deep network embedded with rough fuzzy discretization for OCT fundus image segmentation. Sci. Rep. 13 (1), 328.

Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., Kalinin, A.A., Do, B.T., Way, G.P., Ferrero, E., Agapow, P.-M., Zietz, M., Hoffman, M.M., et al., 2018. Opportunities and obstacles for deep learning in biology and medicine. J. R. Soc. Interface 15 (141), 20170387.

Chinn, E., Arora, R., Arnaout, R., Arnaout, R., 2023. ENRICHing medical imaging training sets enables more efficient machine learning. J. Am. Med. Inform. Assoc. ocad055.

Chiu, S.J., Allingham, M.J., Mettu, P.S., Cousins, S.W., Izatt, J.A., Farsiu, S., 2015. Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. Biomed. Opt. Express 6 (4), 1172–1194.

Chiu, S.J., Izatt, J.A., O'Connell, R.V., Winter, K.P., Toth, C.A., Farsiu, S., 2012. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. Invest. Ophthalmol. Vis. Sci. 53 (1), 53–61.

Choma, M.A., Sarunic, M.V., Yang, C., Izatt, J.A., 2003. Sensitivity advantage of swept source and fourier domain optical coherence tomography. Opt. Express 11 (18), 2183–2189.

Cramer, D.E., 2023. Applications of generative adversarial networks in single image datasets. Sch. Horiz. Univ. Minn. Morris Undergrad. J. 10 (1), 2.

Das, P., Ivkin, N., Bansal, T., Rouesnel, L., Gautier, P., Karnin, Z., Dirac, L., Ramakrishnan, L., Perunicic, A., Shcherbatyi, I., et al., 2020. Amazon SageMaker Autopilot: a white box AutoML solution at scale. In: Proceedings of the Fourth International Workshop on Data Management for End-To-End Machine Learning. pp. 1–7.

Devalla, S.K., Renukanand, P.K., Sreedhar, B.K., Subramanian, G., Zhang, L., Perera, S., Mari, J.-M., Chin, K.S., Tun, T.A., Strouthidis, N.G., et al., 2018. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomed. Opt. Express 9 (7), 3244–3265.

Dos Santos, V.A., Schmetterer, L., Stegmann, H., Pfister, M., Messner, A., Schmidinger, G., Garhofer, G., Werkmeister, R.M., 2019. CorneaNet: fast segmentation of cornea OCT scans of healthy and keratoconic eyes using deep learning. Biomed. Opt. Express 10 (2), 622–641.

Drexler, W., Fujimoto, J.G., 2015. Retinal optical coherence tomography imaging. Opt. Coherence Tomogr. 1685.

Farsiu, S., Chiu, S.J., O'Connell, R.V., Folgar, F.A., Yuan, E., Izatt, J.A., Toth, C.A., Age-Related Eye Disease Study 2 Ancillary Spectral Domain Optical Coherence Tomography Study Group, et al., 2014. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. Ophthalmology 121 (1), 162–172.

Fenner, B.J., Li, H., Gan, A.T., Song, Y.S., Tham, Y.C., Jonas, J.B., Wang, Y.X., Cheng, C.Y., Wong, T.Y., Teo, K.Y., et al., 2023. Genetic variability of complement factor H has ethnicity-specific associations with choroidal thickness. Invest. Ophthalmol. Vis. Sci. 64 (2), 10.

Fujimoto, J.G., Drexler, W., 2015a. Introduction to OCT. In: Drexler, W., Fujimoto, J.G. (Eds.), Optical Coherence Tomography: Technology and Applications. Springer International Publishing, Cham, pp. 3–64. http://dx.doi.org/10.1007/978-3-319-06419-2_1.

Fujimoto, J.G., Drexler, W., 2015b. Introduction to OCT. Opt. Coherence Tomogr. 3.

Garcia-Marin, Y.F., Alonso-Caneiro, D., Fisher, D., Vincent, S.J., Collins, M.J., 2023. Patch-based CNN for corneal segmentation of AS-OCT images: effect of the number of classes and image quality upon performance. Comput. Biol. Med. 152, 106342.

Gawlik, K., Hausser, F., Paul, F., Brandt, A.U., Kadas, E.M., 2018. Active contour method for ilm segmentation in ONH volume scans in retinal OCT. Biomed. Opt. Express 9 (12), 6497–6518.

Guadalupe Mori, V.H., Ogosi Auqui, J.A., Rosales Huamani, J.A., Arenas Ñiquin, J.L., 2023. Industry 4.0: implementation of technologies in medical manufacturing companies. In: Intelligent Sustainable Systems: Selected Papers of WorldS4 2022, Vol. 1. Springer, pp. 619–626.

Guo, X., Tang, D., Molony, D., Yang, C., Samady, H., Zheng, J., Mintz, G.S., Maehara, A., Wang, L., Pei, X., et al., 2019. A machine learning-based method for intracoronary oct segmentation and vulnerable coronary plaque cap thickness quantification. Int. J. Comput. Methods 16 (03), 1842008.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143 (1), 29–36.

He, X., Fang, L., Tan, M., Chen, X., 2022. Intra-and inter-slice contrastive learning for point supervised OCT fluid segmentation. IEEE Trans. Image Process. 31, 1870–1881.

Hee, M.R., Puliafito, C.A., Wong, C., Duker, J.S., Reichel, E., Rutledge, B., Schuman, J.S., Swanson, E.A., Fujimoto, J.G., 1995. Quantitative assessment of macular edema with optical coherence tomography. Arch. Ophthalmol. 113 (8), 1019–1029.

Heisler, M., Bhalla, M., Lo, J., Mammo, Z., Lee, S., Ju, M.J., Beg, M.F., Sarunic, M.V., 2020. Semi-supervised deep learning based 3D analysis of the peripapillary region. Biomed. Opt. Express 11 (7), 3843–3856.

Holmes, J., Welzel, J., 2015. OCT in dermatology. Opt. Coherence Tomogr. 2189.

Huang, Y., He, C., Wang, J., Miao, Y., Zhu, T., Zhou, P., Li, Z., 2018. Intravascular optical coherence tomography image segmentation based on support vector machine algorithm. MCB Mol. Cell. Biomech. 15 (2), 117–125.

Huang, D., Swanson, E.A., Lin, C.P., Schuman, J.S., Stinson, W.G., Chang, W., Hee, M.R., Flotte, T., Gregory, K., Puliafito, C.A., et al., 1991. Optical coherence tomography. Science 254 (5035), 1178–1181.

Jacoba, C.M.P., Doan, D., Salongcay, R.P., Aquino, L.A.C., Silva, J.P.Y., Salva, C.M.G., Zhang, D., Alog, G.P., Zhang, K., Locaylocay, K.L.R.B., et al., 2023. Performance of automated machine learning for diabetic retinopathy image classification from multi-field handheld retinal images. Ophthalmol. Retin.

Johnson, J.M., Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. J. Big Data 6 (1), 1–54.

Karimi, D., Salcudean, S.E., 2019. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. IEEE Trans. Med. Imaging 39 (2), 499–513.

Kelleher, J.D., 2019. Deep Learning. In: MIT Press Essential Knowledge Series, The MIT Press, [Online]. Available: https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2234376&site=ehost-live&scope=site.

Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F., et al., 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172 (5), 1122–1131.

Labelbox, W., 2019. Labelbox: image annotation tools, which supported images, video, text, audio, and geospatial formats.

Lang, A., Carass, A., Sotirchos, E., Calabresi, P., Prince, J.L., 2013. Segmentation of retinal oct images using a random forest classifier. In: Medical Imaging 2013: Image Processing, Vol. 8669. SPIE, pp. 199–205.

Lang, A., Carass, A., Swingle, E.K., Al-Louzi, O., Bhargava, P., Saidha, S., Ying, H.S., Calabresi, P.A., Prince, J.L., 2015. Automatic segmentation of microcystic macular edema in OCT. Biomed. Opt. Express 6 (1), 155–169.

Lazaridis, G., Xu, M., Afgeh, S.S., Montesano, G., Garway-Heath, D., 2020. Bio-inspired attentive segmentation of retinal OCT imaging. In: International Workshop on Ophthalmic Medical Image Analysis. Springer, pp. 1–10.

Lee, D., Maa, A., 2023. Ethical and legal considerations in eye telehealth programs. In: Ocular Telehealth. Elsevier, pp. 185–198.

Lee, C.S., Tyring, A.J., Deruyter, N.P., Wu, Y., Rokem, A., Lee, A.Y., 2017. Deep-learning based, automated segmentation of macular edema in optical coherence tomography. Biomed. Opt. Express 8 (7), 3440–3448.

Leitgeb, R., Hitzenberger, C., Fercher, A.F., 2003. Performance of fourier domain vs. time domain optical coherence tomography. Opt. Express 11 (8), 889–894.

Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.-W., Heng, P.-A., 2018. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans. Med. Imaging 37 (12), 2663–2674.

Li, Q., Li, S., He, Z., Guan, H., Chen, R., Xu, Y., Wang, T., Qi, S., Mei, J., Wang, W., 2020a. DeepRetina: layer segmentation of retina in OCT images using deep learning. Transl. Vis. Sci. Technol. 9 (2), 61.

Li, K., Liang, W., Mavadia-Shukla, J., Park, H.-C., Li, D., Yuan, W., Wan, S., Li, X., 2019a. Super-achromatic optical coherence tomography capsule for ultrahigh-resolution imaging of esophagus. J. Biophotonics 12 (3), e201800205.

Li, X., Wang, W., Hu, X., Yang, J., 2019b. Selective kernel networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 510–519.

Li, D., Wu, J., He, Y., Yao, X., Yuan, W., Chen, D., Park, H.-C., Yu, S., Prince, J.L., Li, X., 2019c. Parallel deep neural networks for endoscopic OCT image segmentation. Biomed. Opt. Express 10 (3), 1126–1135.

Li, M., Zhang, Y., Ji, Z., Xie, K., Yuan, S., Liu, Q., Chen, Q., 2020b. Ipn-v2 and octa-500: methodology and dataset for retinal image segmentation. arXiv preprint arXiv:2012.07261.

Lin, G., Milan, A., Shen, C., Reid, I., 2017. Refinenet: multi-path refinement networks for high-resolution semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1925–1934.

Liu, X., Bi, L., Xu, Y., Feng, D., Kim, J., Xu, X., 2019. Robust deep learning method for choroidal vessel segmentation on swept source optical coherence tomography images. Biomed. Opt. Express 10 (4), 1601–1612.

Liu, X., Cao, J., Fu, T., Pan, Z., Hu, W., Zhang, K., Liu, J., 2018. Semi-supervised automatic segmentation of layer and fluid region in retinal optical coherence tomography images using adversarial learning. IEEE Access 7, 3046–3061.

Liu, X., Ouellette, S., Jamgochian, M., Liu, Y., Rao, B., 2023. One-class machine learning classification of skin tissue based on manually scanned optical coherence tomography imaging. Sci. Rep. 13 (1), 867.

Liu, X., Wang, S., Cao, J., Zhang, Y., Wang, M., 2022. Uncertainty-guided self-ensembling model for semi-supervised segmentation of multiclass retinal fluid in optical coherence tomography images. Int. J. Imaging Syst. Technol. 32 (1), 369–386.

López-Varela, E., de Moura, J., Novo, J., Fernández-Vigo, J.I., Moreno-Morillo, F.J., Ortega, M., 2023. Fully automatic segmentation and monitoring of choriocapillaris flow voids in OCTA images. Comput. Med. Imaging Graph. 104, 102172.

Lu, Y., Shen, Y., Xing, X., Ye, C., Meng, M.Q.-H., 2023. Boundary-enhanced semi-supervised retinal layer segmentation in optical coherence tomography images using fewer labels. Comput. Med. Imaging Graph. 105, 102199.

Luo, W., Abbasi, S.A., Zhu, S., Li, X., Ho, H.-P., Yuan, W., 2023a. Electrically switchable and tunable infrared light modulator based on functional graphene metasurface. Nanophotonics 12 (9), 1797–1807.

Luo, W., Li, X., Abbasi, S.A., Zhu, S., Ho, H.-P., Yuan, W., 2023b. Analysis of the D-shaped PCF-based SPR sensor using resonance electron relaxation and fourier domain method. Opt. Lasers Eng. 166, 107588.

Magistretti, B., 2017. Playment Raises 1.6 Million to Improve AI Training Through Crowdsourced Data Tagging. VentureBeat.

Mahesh, B., 2020. Machine learning algorithms-a review. Int. J. Sci. Res. (IJSR) 9, 381–386 [Internet].

Mandrioli, D., 2019. Hitachi semantic segmentation editor. [Online]. Available: https://github.com/Hitachi-Automotive-And-Industry-Lab/semantic-segmentation-editor.

Manfredini, M., Greco, M., Farnetani, F., Ciardo, S., De Carvalho, N., Mandel, V.D., Starace, M., Pellacani, G., 2017. Acne: morphologic and vascular study of lesions and surrounding skin by means of optical coherence tomography. J. Eur. Acad. Dermatol. Venereol. 31 (9), 1541–1546.

Masood, S., Fang, R., Li, P., Li, H., Sheng, B., Mathavan, A., Wang, X., Yang, P., Wu, Q., Qin, J., et al., 2019. Automatic choroid layer segmentation from optical coherence tomography images using deep learning. Sci. Rep. 9 (1), 3058.

Melinščak, M., Radmilović, M., Vatavuk, Z., Lončarić, S., 2021a. Annotated retinal optical coherence tomography images (AROI) database for joint retinal layer and fluid segmentation. Automatika 62 (3–4), 375–385.

Melinščak, M., Radmilovič, M., Vatavuk, Z., Lončarić, S., 2021b. Aroi: Annotated retinal OCT images database. In: 2021 44th International Convention on Information, Communication and Electronic Technology. MIPRO, IEEE, pp. 371–376.

Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D., 2021. Image segmentation using deep learning: A survey. IEEE Trans. Pattern Anal. Mach. Intell.

Moccia, S., De Momi, E., El Hadji, S., Mattos, L.S., 2018. Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics. Comput. Methods Programs Biomed. 158, 71–91.

Moradi, M., Chen, Y., Du, X., Seddon, J.M., 2023. Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans. Comput. Biol. Med. 106512.

Nawaz, M., Chan, R.W., Malik, A., Khan, T., Cao, P., 2022a. Hand gestures classification using electrical impedance tomography images. IEEE Sens. J.

Nawaz, M., Khan, S., Qureshi, R., Yan, H., 2019. Clustering based one-to-one hypergraph matching with a large number of feature points. Signal Process., Image Commun. 74, 289–298.

Nawaz, M., Qureshi, R., Teevno, M.A., Shahid, A.R., 2022b. Object detection and segmentation by composition of fast fuzzy C-mean clustering based maps. J. Ambient Intell. Humaniz. Comput. 1–16.

Nawaz, M., Yan, H., 2020a. Saliency detection using deep features and affinity-based robust background subtraction. IEEE Trans. Multimed. 23, 2902–2916.

Nawaz, M., Yan, H., 2020b. Saliency detection via multiple-morphological and superpixel based fast fuzzy C-mean clustering network. Expert Syst. Appl. 161, 113654.

Nawaz, M., Zuojun, W., Khanc, S., Irfan, M., Wonge, E.C., Chane, R., Cao, P., 2023. Cross modality generative learning framework for anatomical transitive Magnetic Resonance Imaging (MRI) from Electrical Impedance Tomography (EIT) image. Comput. Med. Imaging Graph.

Oguz, I., Zhang, L., Abramoff, M.D., Sonka, M., 2015. Graph-based retinal fluid segmentation from OCT images. In: Proceeding Optima Challenge-MICCAI.

Panferov, V., Tailakov, D., Donets, A., 2020. Recognition of rocks lithology on the images of core samples. In: 2020 Science And Artificial Intelligence Conference. SAI Ence, IEEE, pp. 54–57.

Park, H.-C., Mavadia-Shukla, J., Yuan, W., Alemohammad, M., Li, X., 2017. Broadband rotary joint for high-speed ultrahigh-resolution endoscopic OCT imaging at 800 nm. Opt. Lett. 42 (23), 4978–4981.

Peng, C., Zhang, X., Yu, G., Luo, G., Sun, J., 2017. Large kernel matters–improve semantic segmentation by global convolutional network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4353–4361.

Perry, T., 2021. Lighttag: text annotation platform. arXiv preprint arXiv:2109.02320.

Pfister, M., Schützenberger, K., Pfeiffenberger, U., Messner, A., Chen, Z., Dos Santos, V.A., Puchner, S., Garhöfer, G., Schmetterer, L., Gröschl, M., et al., 2019. Automated segmentation of dermal fillers in OCT images of mice using convolutional neural networks. Biomed. Opt. Express 10 (3), 1315–1328.

Pizenberg, M., Carlier, A., Faure, E., Charvillat, V., 2018. Web-based configurable image annotations. In: Proceedings of the 26th ACM international conference on Multimedia. pp. 1368–1371.

Qiu, J., Li, L., Sun, J., Peng, J., Shi, P., Zhang, R., Dong, Y., Lam, K., Lo, F.P.W., Xiao, B., Yuan, W., Xu, D., Lo, B., 2023. Large AI models in health informatics: applications, challenges, and the future. arXiv:2303.11568.

Ran, A., Cheung, C.Y., 2021. Deep learning-based optical coherence tomography and optical coherence tomography angiography image analysis: An updated summary. Asia-Pac. J. Ophthalmol. 10 (3), 253–260.

Rashno, A., Koozekanani, D.D., Drayna, P.M., Nazari, B., Sadri, S., Rabbani, H., Parhi, K.K., 2017. Fully automated segmentation of fluid/cyst regions in optical coherence tomography images with diabetic macular edema using neutrosophic sets and graph algorithms. IEEE Trans. Biomed. Eng. 65 (5), 989–1001.

Ren, X., Zhao, Y., Fan, J., Wu, H., Chen, Q., Kubo, T., 2023. Semantic segmentation of superficial layer in intracoronary optical coherence tomography based on cropping-merging and deep learning. Infrared Phys. Technol. 104542.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C., Navab, N., 2017. Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed. Opt. Express 8 (8), 3627–3642.

Schlegl, T., Waldstein, S.M., Bogunovic, H., Endstraßer, F., Sadeghipour, A., Philip, A.-M., Podkowinski, D., Gerendas, B.S., Langs, G., Schmidt-Erfurth, U., 2018. Fully automated detection and quantification of macular fluid in OCT using deep learning. Ophthalmology 125 (4), 549–558.

Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B.S., Waldstein, S.M., Bogunović, H., 2018. Artificial intelligence in retina. Prog. Retin. Eye Res. 67, 1–29.

Sedai, S., Antony, B., Rai, R., Jones, K., Ishikawa, H., Schuman, J., Gadi, W., Garnavi, R., 2019. Uncertainty guided semi-supervised segmentation of retinal layers in OCT images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 282–290.

Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., Kruchinin, D., Zankevich, A., Sidnev, D., 2020. Opencv/cvat: V1. 1.0. Zenodo.

Shahid, A.R., Nawaz, M., Fan, X., Yan, H., 2022. View-adaptive graph neural network for action recognition. IEEE Trans. Cogn. Dev. Syst.

Shi, P., Qiu, J., Abaxi, S.M.D., Wei, H., Lo, F.P.-W., Yuan, W., 2023. Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation. Diagnostics 13 (11), 1947.

Soltanian-Zadeh, S., Kurokawa, K., Liu, Z., Zhang, F., Saeedi, O., Hammer, D.X., Miller, D.T., Farsiu, S., 2021. Weakly supervised individual ganglion cell segmentation from adaptive optics OCT images for glaucomatous damage assessment. Optica 8 (5), 642–651.

Srinivasan, P.P., Kim, L.A., Mettu, P.S., Cousins, S.W., Comer, G.M., Izatt, J.A., Farsiu, S., 2014. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. Biomed. Opt. Express 5 (10), 3568–3577.

Swanson, E.A., 2015. Oct technology transfer and the OCT market. Opt. Coherence Tomogr. 2529.

Themstrup, L., Banzhaf, C., Mogensen, M., Jemec, G., 2014. Optical coherence tomography imaging of non-melanoma skin cancer undergoing photodynamic therapy reveals subclinical residual lesions. Photodiagnosis Photodyn. Ther. 11 (1), 7–12.

Thiboutot, J., Yuan, W., Park, H.-c., Li, D., Loube, J., Mitzner, W., Yarmus, L., Li, X., Brown, R.H., 2022. Visualization and validation of the microstructures in the airway wall in vivo using diffractive optical coherence tomography. Academic Radiol. 29 (11), 1623–1630.

Tian, J., Varga, B., Tatrai, E., Fanni, P., Somfai, G.M., Smiddy, W.E., Debuc, D.C., 2016. Performance evaluation of automated segmentation software on optical coherence tomography volume data. J. Biophotonics 9 (5), 478–489.

Tkachenko, M., Malyuk, M., Shevchenko, N., Holmanyuk, A., Liubimov, N., 2020. Label studio: Data labeling software. Open source software available from https://github.com/heartexlabs/label-studio.

van Soest, G., Goderie, T., Gonzalo, N., Koljenović, S., van Leenders, G., Regar, E., Serruys, P., van der Steen, A., 2009. Imaging atherosclerotic plaque composition with intracoronary optical coherence tomography. Neth. Heart J. 17 (11), 448–450.

Venkatesh, R., Reddy, N.G., Mishra, P., Agrawal, S., Mutalik, D., Yadav, N.K., Chhablani, J., 2022. Spectral domain OCT features in type 2 macular telangiectasia (type 2 MacTel): its relevance with clinical staging and visual acuity. Int. J. Retin. Vitreous 8 (1), 26.

Wang, J., Li, W., Chen, Y., Fang, W., Kong, W., He, Y., Shi, G., 2021. Weakly supervised anomaly segmentation in retinal OCT images using an adversarial learning approach. Biomed. Opt. Express 12 (8), 4713–4729.

Wang, T., Niu, S., Dong, J., Chen, Y., 2020. Weakly supervised retinal detachment segmentation using deep feature propagation learning in SD-OCT images. In: International Workshop on Ophthalmic Medical Image Analysis. Springer, pp. 146–154.

Wilder-Smith, P., Jung, W.-G., Brenner, M., Osann, K., Beydoun, H., Messadi, D., Chen, Z., 2004. In vivo optical coherence tomography for the diagnosis of oral malignancy. Lasers Surg. Med. Off. J. Am. Soc. Laser Med. Surg. 35 (4), 269–275.

Wilder-Smith, P., Otis, L., Zhang, J., Chen, Z., 2008. Dental oct. In: Optical Coherence Tomography. Springer, pp. 1151–1182.

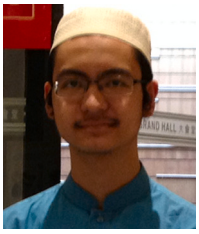
Xing, R., Niu, S., Gao, X., Liu, T., Fan, W., Chen, Y., 2021. Weakly supervised serous retinal detachment segmentation in SD-OCT images by two-stage learning. Biomed. Opt. Express 12 (4), 2312–2327.
Yanagihara, R.T., Lee, C.S., Ting, D.S.W., Lee, A.Y., 2020. Methodological challenges of deep learning in optical coherence tomography for retinal diseases: a review. Transl. Vis. Sci. Technol. 9 (2), 11.
Yang, X., Li, Z., Guo, Y., Zhou, D., 2022. DCU-net: a deformable convolutional neural network based on cascade U-net for retinal vessel segmentation. Multimedia Tools Appl. 81 (11), 15593–15607.
Yang, Z., Soltanian-Zadeh, S., Chu, K.K., Zhang, H., Moussa, L., Watts, A.E., Shaheen, N.J., Wax, A., Farsiu, S., 2021. Connectivity-based deep learning approach for segmentation of the epithelium in in vivo human esophageal oct images. Biomed. Opt. Express 12 (10), 6326–6340.
Yuan, W., Brown, R., Mitzner, W., Yarmus, L., Li, X., 2017. Super-achromatic monolithic microprobe for ultrahigh-resolution endoscopic optical coherence tomography at 800 nm. Nat. Commun. 8 (1), 1–9.
Yuan, W., Chen, D., Sarabia-Estrada, R., Guerrero-Cázares, H., Li, D., Quiñones Hinojosa, A., Li, X., 2020. Theranostic OCT microneedle for fast ultrahigh-resolution deep-brain imaging and efficient laser ablation in vivo. Sci. Adv. 6 (15), eaaz9664.
Yuan, W., Feng, Y., Chen, D., Gharibani, P., Chen, J.D., Yu, H., Li, X., 2022a. In vivo assessment of inflammatory bowel disease in rats with ultrahigh-resolution colonoscopic OCT. Biomed. Opt. Express 13 (4), 2091–2102.
Yuan, W., Mavadia-Shukla, J., Xi, J., Liang, W., Yu, X., Yu, S., Li, X., 2016. Optimal operational conditions for supercontinuum-based ultrahigh-resolution endoscopic OCT imaging. Opt. Lett. 41 (2), 250–253.
Yuan, W., Thiboutot, J., Park, H.-c., Li, A., Loube, J., Mitzner, W., Yarmus, L., Brown, R.H., Li, X., 2022b. Direct visualization and quantitative imaging of small airway anatomy in vivo using deep learning assisted diffractive OCT. IEEE Trans. Biomed. Eng.
Yushkevich, P.A., Gao, Y., Gerig, G., 2016. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBC, IEEE, pp. 3342–3345.
Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31 (3), 1116–1128.
Zhang, H., Yang, J., Tang, X., Guo, B., Zhang, J., Wang, X., Zhang, A., 2023. Learning oct segmentation from a single label. In: Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXVII, Vol. 12367. SPIE, pp. 150–153.
Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016. Learning deep features for discriminative localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2921–2929.

Mehmood Nawaz received a Ph.D. degree from the City University of Hong Kong and an M.Sc. degree from Shanghai Jiao Tong University, China. His BS degree is from Baha-Uddin-Zakariya University, Multan, Pakistan. He is currently a postdoctoral fellow in the Department of Biomedical Engineering at The Chinese University of Hong Kong. His research interests medical image segmentation, including graph matching, image segmentation, image reconstruction, shape matching, and pattern recognition, etc.

Adilet Uvaliyev received his undergraduate degree from The University of Hong Kong. He is currently working on machine learning and deep learning techniques for OCT image segmentation at the Chinese University of Hong Kong. His research interests are machine learning, deep learning, medical image analysis, etc.

Khadija Bibi received the B.Sc degree in Bio-Technology from Baha-uddin-Zakariya (B.Z.U) University, Multan, Pakistan. She is currently working on Medical Images Segmentation using deep-learning models at the Chinese University of Hong Kong. Her research interests include medical image segmentation, droplet segmentation, PCR data analysis, image reconstruction, shape matching, pattern recognition, etc.

Hao Wei received his Master's and Bachelor's degree from Central South University, China. Currently, he is a Ph.D. candidate in the Department of Biomedical Engineering, The Chinese University of Hong Kong. His research interests mainly include medical image segmentation, classification, and domain generalization, etc.



Sai Mu Dalike Abaxi received B.Eng. in Computer Science from the University of Hong Kong. He is currently working as a Research Assistant in the Department of Biomedical Engineering the Chinese University of Hong Kong. His research interests include deep learning in medical image analysis, segmentation, and quantification, etc.

Anum Masood received her Ph.D. degree in Computer Science and Engineering from Shanghai Jiao Tong University, Shanghai, China in 2019. Her B.Sc. and M.Sc. degrees in Computer Science are from the COMSATS University Islamabad, Islamabad, Pakistan. She worked as a Lecturer with the Department of Computer Science, COMSATS University Islamabad, from 2014 to 2020. Currently, she is a post-doctoral researcher at the Norwegian University of Science and Technology, Norway, and affiliated with PET Centre, St. Olav's Hospital, Trondheim, Norway. Her research interests include medical image analysis, automated cancer detection, machine learning, and image processing.

Peilun Shi received his B.Eng in Mechanical Engineering at the University of Birmingham in 2021. And he received M.Res in Medical Robotics and Image-Guided Intervention from Hamlyn Centre at Imperial College London in 2022. He joined the Chinese University of Hong Kong in 2022. His research interests include artificial intelligence in medical image analysis and health care.

Aaron H.P. Ho received his BEng and Ph.D. in Electrical and Electronic Engineering in the University of Nottingham in 1986 and 1990 respectively. His Ph.D. thesis was on the interdiffusion of semiconductor superlattices. During 1990–1992, he was a post-doctoral research fellow at the University of Leeds, working on the growth and characterization of ferromagnetic superlattices. He then returned to the University of Nottingham to participate in an industrial research project on laser ultrasound evaluation of engineering ceramics using laser interferometers. In 1994, he joined the Fiber Optics Components Operation of Hewlett-Packard as a senior process engineer in the semiconductor laser device fabrication division. He returned to Hong Kong to take up a lecturing position in the Department of Physics and Materials Science at the City University of Hong Kong in 1996. He joined The Chinese University of Hong Kong in 2002.

Wu Yuan joined the Department of Biomedical Engineering at the Chinese University of Hong Kong as an Assistant Professor in January 2020. He received the Ph.D. degree in Electronic Engineering from the Chinese University of Hong Kong in December 2008. He was a research associate at the Department of Biomedical Engineering, Johns Hopkins University, with a joint appointment to the Department of Neurologic Surgery, Mayo Clinic.