Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey

6 authors, including:

Some of the authors of this publication are also working on these related projects:

Project   Unsupervised Anomaly Detection in Medical Image View project

Project   Guided M-Net for High-Resolution Biomedical Image Segmentation with Weak Boundaries View project

Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey

Xiao-Qing Zhang[1,2*]   Yan Hu[1,2*]   Zun-Jie Xiao[1]   Jian-Sheng Fang[1]
Risa Higashita[1,3]   Jiang Liu[1,2,4,5]

[1] Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology, Shenzhen 518055, China

[2] Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

[3] Tomey Corporation, Nagoya 4510051, Japan

[4] Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology and Engineering,

Chinese Academy of Sciences, Ningbo 315300, China

[5] Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation,

Southern University of Science and Technology, Shenzhen 518055, China

Abstract: Cataracts are the leading cause of visual impairment and blindness globally. Over the years, researchers have achieved significant progress in developing state-of-the-art machine learning techniques for automatic cataract classification and grading, aiming to prevent cataracts early and improve clinicians′ diagnosis efficiency. This survey provides a comprehensive survey of recent advances in machine learning techniques for cataract classification/grading based on ophthalmic images. We summarize existing literature from two research directions: conventional machine learning methods and deep learning methods. This survey also provides insights into existing works of both merits and limitations. In addition, we discuss several challenges of automatic cataract classification/grading based on machine learning techniques and present possible solutions to these challenges for future research.

Keywords: Cataract, classification and grading, ophthalmic image, machine learning, deep learning.

1 Introduction

According to the World Health Organization (WHO)[1, 2], it is estimated that approximately 2.2 billion people suffer from visual impairment. Cataracts account for about 33% of visual impairment and are the number one cause of blindness (more than 50%) worldwide. Cataract patients can improve life quality and vision through early intervention and cataract surgery, which are efficient methods to reduce the blindness ratio and cataract-blindness burden for society simultaneously.

Clinically, cataracts are the transparency loss of crystalline lens area, which occurs when the protein inside the lens clumps together[3]. They are associated with many factors[4], such as developmental abnormalities, trauma, metabolic disorders, genetics, drug-induced changes, age, etc. Genetics and ageing are two of the most important

factors for cataracts. cataracts can be categorized as age-related cataract, pediatrics cataract (PC), and secondary cataract[3, 4] according to their causes. Depending on the location of the crystalline lens opacity, they can be grouped into nuclear cataract (NC), cortical cataract (CC), and posterior subcapsular cataract (PSC)[5, 6]. NC denotes the gradual clouding and the progressive hardening in the nuclear region. CC is the form of white wedged-shaped and radially oriented opacities and develops from the outside edge of the lens toward the center in a spoke-like fashion[4, 7]. PSC is granular opacities, and its symptom includes small breadcrumbs or sand particles, which are sprinkled beneath the lens capsule[6].

Over the past years, ophthalmologists have used several ophthalmic images to diagnose cataracts based on their experience and clinical training. This manual diagnosis mode is error-prone, time-consuming, subjective, and costly, which is a great challenge in developing countries or rural communities, where experienced clinicians are scarce. To prevent cataracts early and improve the precision and efficiency of cataract diagnosis, researchers have made great efforts in developing computer-aided diagnosis (CAD) techniques for automatic cataract classi-

fication/grading[8] on different ophthalmic images, including conventional machine learning methods and deep learning methods. The conventional machine learning method is a combination of feature extraction and classification/grading. In the feature extraction stage, a variety of image processing methods have been proposed to obtain visual features of cataracts according to different ophthalmic images, such as density-based statistics method, density histogram method, bag-of-features (BOF) method, Gabor wavelet transform, gray level co-occurrence matrix (GLCM), Haar wavelet transform, etc.[6, 9–16] In the classification/grading stage, strong classification methods are applied to recognize different cataract severity levels, e.g., support vector machine (SVM)[17–19]. Over the past ten years, deep learning has achieved great success in various fields, including medical image analysis, which can be viewed as a representation learning approach. It can learn low-level, middle-level, and high-level feature representations from raw data in an end-to-end manner (e.g., ophthalmic images). Various deep neural networks have been utilized to tackle cataract classification/grading tasks like convolutional neural networks (CNNs), attention-based networks, Faster-RCNN, and multilayer perceptron (MLP) neural networks. For instance, Zhang et al.[20] proposed a multi-region fusion attention network to recognize nuclear cataract severity levels.

Previous surveys had summarized cataract types, cataract classification/grading systems, and ophthalmic imaging modalities, respectively[4, 21–26]; however, none had summarized ML techniques based on ophthalmic imaging modalities for automatic cataract classification/grading systematically. To the best of our knowledge, this is the first survey that systematically summarizes recent advances in ML techniques for automatic cataract classi-

fication/grading. This survey mainly focuses on surveying ML techniques in cataract classification/grading, comprised of conventional ML methods and deep learning methods. We surveyed these published papers through Web of Science (WoS), Scopus, and Google Scholar databases. Fig. 1 provides a general organization framework for this survey according to collected papers, our summary, and discussion with experienced ophthalmologists. To understand this survey easily, we also review ophthalmic imaging modalities, cataract grading systems, and commonly-used evaluation measures in brief. Then we introduce ML techniques step by step. We hope that this survey can provide a valuable summary of current research and present potential research directions for ML-based cataract classification/grading in the future.

Abbreviations

CAD: Computer-aided diagnosis; ML: Machine learning; WHO: World Health Organization; NC: Nuclear cataract; CC: Cortical cataract; PSC: Posterior subcapsular cataract; LR: Linear regression; RF: Random forest; SVR: Support vector machine regression; ASM: Active shape model; NLP: Natural language processing; OCT: Optical coherence tomography; LOCS: Lens opacity classification system; OCCCGS: Oxford clinical cataract classification and grading system; SVM: Support vector machine; AI: Artificial intelligence; BOF: Bag-of-features; PCA: Principal component analysis; DWT: Discrete wavelet transform; DCT: Discrete cosine transform; KNN: K-nearest neighbor; ANN: Artificial neural network; BPNN: Back propagation neural network; CNN: Convolutional neural network; RNN: Recurrent neural network; DNN: Deep neural network; MLP: Multilayer perceptron; ROI: Region of interest; AE: Autoencoder;
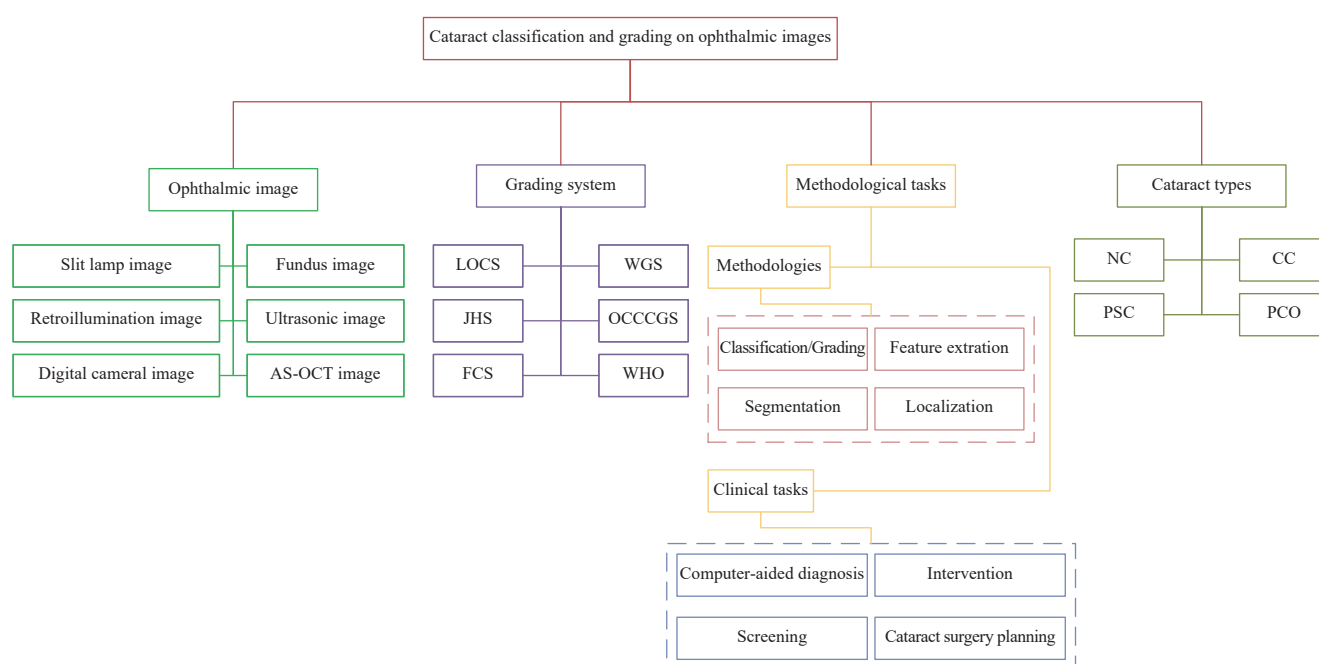

Fig. 1   Overall organization framework of this survey

GAN: Generative adversarial network; CBIR: Content-based image retrieval; AS-OCT: Anterior segment optical coherence tomography; 3D: Three-dimensional; 2D: Two-dimensional; SSL: Semi-supervised learning; AWM: Adaptive window model; GSR: Group sparsity regression; GLCM: Gray level co-occurrence matrix; IMF: Image morphological feature; SWLR: Similarity weighted linear reconstruction; CBAM: Convolutional block attention module; ECA: Efficient channel attention; SE: Squeeze-and-excitation; SPA: Spatial pyramid attention; SRCL: Sparse range-constrained learning; NB：Naive Bayes; EF: Ellipse fitting method; FLD: Fisher linear discriminant; WGS: Wisconsin grading system; SF: semantic feature.

2 Ophthalmic imaging modalities for cataract classification/grading

To our best understanding, this survey introduces six different eye images used for cataract classification/grading for the first time: slit lamp image, retroillumination image, ultrasonic image, fundus image, digital camera image, and anterior segment optical coherence tomography (AS-OCT) image, as shown in Fig. 2. In this section, we will introduce each type of ophthalmic image step by step and then discuss their advantages and disadvantages.
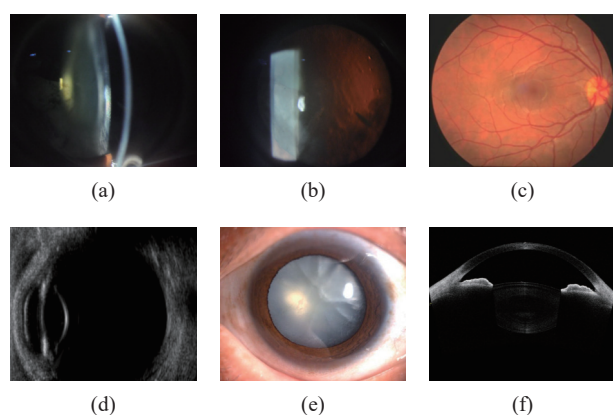

Fig. 2   Six different ophthalmic images: (a) Slit lamp image; (b) Retroillumination image; (c) Ultrasonic image; (d) Fundus image; (e) Digital camera image; (f) Anterior segment optical coherence tomography image.

2.1 Slit lamp image

The slit lamp camera[27, 28] is a high-intensity light source instrument, which is comprised of the corneal microscope and the slit lamp. Slit lamp image can be accessed through a slit lamp camera, which is usually used to examine the anterior segment and posterior segment structure of the human eye eyelid, sclera, conjunctiva, iris, crystalline lens, and cornea. Fig. 3 offers four representative slit lamp images for four different cataract severity levels.

2.2 Retroillumination image

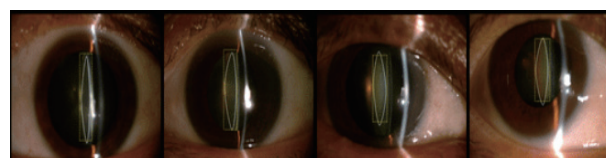Retroillumination image is a non-stereoscopic medical

Fig. 3   Slit lamp images with four severity levels of nuclear cataract

image that is accessed through the crystalline lens camera[29, 30]. It can be used to diagnose CC and PSC in the crystalline lens region. Two types of retroillumination images through the crystalline lens camera can be obtained: an anterior image focused on the iris, which corresponds to the anterior cortex of the lens, and a posterior image focused on 3–5 mm more posteriorly, which intends to image the opacity of PSC.

2.3 Ultrasonic image

In clinical cataract diagnosis, ultrasound image is a commonly-used ophthalmic image modality to evaluate the hardness of the cataract lens objectively[31]. Frequently applied ultrasound imaging techniques are usually developed based on measuring ultrasonic attenuation and sound speed, which may increase the hardness of the cataract lens[32]. High-frequency ultrasound B-mode imaging can be used to monitor local cataract formation, but it cannot measure the lens hardness accurately[33]. In order to make up for the B-scan deficiency, the ultrasound imaging technique built on Nakagami statistical model called ultrasonic Nakagami imaging[34–36] was developed, which can be used to visualize local scatterer concentrations in biological tissues.

2.4 Fundus image

The fundus camera[37, 38] is a unique camera in conjunction with a low power microscope, which is usually used to capture fundus images operated by ophthalmologists or professional operators. Fundus image is a highly specialized form of eye imaging and can capture the eye′s inner lining and the structures of the back of the eye. Fig. 4 shows four fundus images of different cataract severity levels.
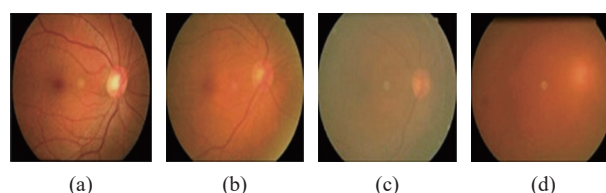

Fig. 4   Four severity levels of cataract on fundus images[39]: (a) Normal; (b) Immature; (c) Mature; (d) Hypermature.

2.5 Digital camera image

Commonly used digital cameras can access digital

camera images like smartphone cameras. Compared with the fundus camera and slit lamp device, the digital camera is easily available and used. Hence, using digital cameras for cataract screening has great potential in the future, especially in developing countries and rural areas, where people have limitations to access expensive ophthalmological equipment and experienced ophthalmologists.

2.6 Anterior segment optical coherence tomography image

Anterior segment optical coherence tomography (AS-OCT)[40] imaging technique is one of optical coherence tomography (OCT) imaging techniques. It can be used to visualize and assess anterior segment ocular features, such as the tear film, cornea, conjunctiva, sclera, rectus muscles, anterior chamber angle structures, and lens[41–44]. AS-OCT image can provide high-resolution visualization of the crystalline lens in vivo in people′s eyes without impacting the tissue, which can help ophthalmologists get different information about the crystalline lens through the circumferential scanning mode. Recent works have suggested that the AS-OCT images can be used to locate the lens region and accurately characterize opacities of different types of cataracts quantitatively[45, 46]. Fig. 5 offers an AS-OCT image, which can quickly help us to understand the structure of the crystalline lens.
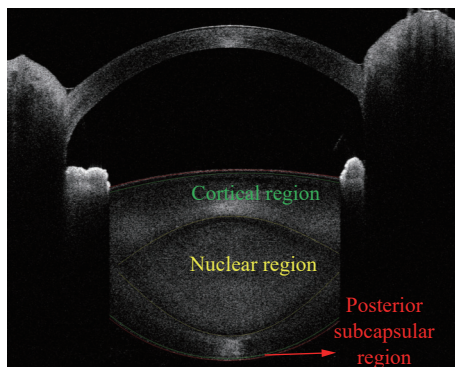

Fig. 5   AS-OCT image. Nuclear region is used for NC diagnosis; cortical region is used for CC diagnosis; posterior subcapsular region is used for PSC diagnosis.

Discussion. Though six different ophthalmic images are used for cataract diagnosis, slit lamp images and fundus images are the most commonly used ophthalmic images for clinical cataract diagnosis and scientific research purposes. This is because existing cataract classification/grading systems are built on them. Slit lamp images can capture the lens region but cannot distinguish the boundaries between nuclear-region, cortical-region, and capsular-region. Hence, it is difficult for clinicians to accurately diagnose different types of cataracts based on slit lamp images. Fundus images only contain opacity information of cataracts and do not contain location in-

formation, which is mainly applied to cataract screening.

Retroillumination images are usually used to diagnose CC and PSC clinically, which have not been widely studied. Digital camera images are ideal ophthalmic images for cataract screening because they can be collected through mobile phones, which are easy and inexpensive for most people. Like fundus images, digital camera images only have opacity information but do not contain location information of different types of cataracts. Ultrasonic images can capture the lens region and evaluate the hardness of the cataract lens, but they cannot distinguish different sub-regions, e.g., cortex region. AS-OCT image is a new ophthalmic image that can distinguish different sub-regions, e.g., cortex and nucleus regions, which are significant for cataract surgery planning and cataract diagnosis. However, there is no cataract classification/grading system built on AS-OCTs; thus, it is urgent to develop a clinical cataract classification/grading system based on AS-OCT images. Moreover, existing automatic AS-OCT image-based cataract classification has rarely been studied.

3 Cataract classification/grading systems

To classify or grade the severity levels of cataracts (lens opacities) accurately and quantitatively, it is crucial and necessary to build standard/gold cataract classification/grading systems for clinical practice and scientific research purposes. This section briefly introduces six existing cataract classification/grading systems.

3.1 Lens opacity classification system

Lens opacity classification system (LOCS) was first introduced in 1988 and has developed from LOCS I to LOCS III[47–49]. LOCS III is widely used for clinical diagnosis and scientific research. In the LOCS III, six representative slit lamp images are used for nuclear cataract classification/grading; five representative retroillumination images are used for cortical cataract classification/grading; five representative retroillumination images are used for grading posterior subcapsular cataract classification/grading, as shown in Fig. 6. Moreover, the severity levels of these three cataract types are graded regularly on a decimal scale by spacing intervals.

3.2 Wisconsin grading system

Wisconsin grading system (WGS) was proposed by the Wisconsin Research Laboratory in 1990[50–52]. It contains four standard photographs for grading cataract severity levels. The grades for cataract severity levels are as follows: Grade 1, as clear or clearer than Standard 1; Grade 2, not as clear as Standard 1 but as clear or clearer than Standard 2; Grade 3, not as clear as Standard 2
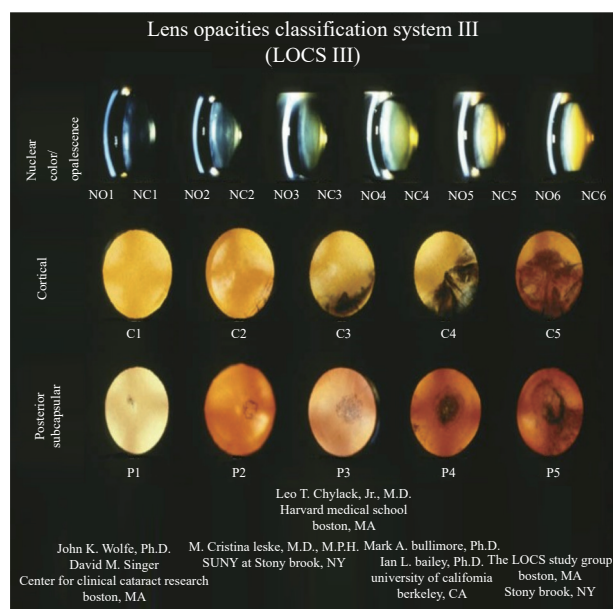
Fig. 6   Lens opacity classification system III

but as clear or clearer than Standard 3; Grade 4, not as clear as Standard 3 but as clear or clearer than Standard 4; Grade 5, at least as severe as Standard 4; and Grades 6–8, cannot be graded due to severe opacities of the lens (please see detailed introduction of Wisconsin grading system in [50]. Wisconsin grading system also uses a decimal grade for cataract grading with 0.1-unit interval space, and the range of the decimal grade is from 0.1 to 4.9.

3.3 Oxford clinical cataract classification and grading system

Oxford clinical cataract classification and grading system (OCCCGS) is also a slit lamp image-based cataract grading system[53, 54]. Although the LOCS III uses photographic transparencies of the lens as cataract grading standards, it adopts standard diagrams and Munsell color samples to grade the severity of cortical, posterior subcapsular, and nuclear cataracts[54]. In the OCCCGS, five standard grading levels are used for evaluating the severity level of cataracts based on cataract features, such as cortical features, nuclear features, morphological features, etc.[53] For example, the severity levels of nuclear cataract are graded as follows: Grade 0: No yellow detectable; Grade 1: Yellow just detectable; Grade 2: Definate-yellow; Grade 3: Orange-yellow; Grade 4: Reddish-brown; Grade 5: Blackish-brown[53].

3.4 Johns Hopkins system

Johns Hopkins system (JHS) was first proposed in the 1980s[55]. It has four standard slit lamp images, which denote the severity level of cataract based on the opalescence of the lens. For nuclear cataract, Grade 1: Opacit-

ies that are definitely present but not thought to reduce visual acuity; Grade 2: Opacities are consistent with visual acuity between 20/20 and 20/30; Grade 3: Opacities are consistent with vision between 20/40 and 20/100; Grade 4: Opacities are consistent with the vision of 20/200 or less.

3.5 WHO cataract grading system

WHO cataract crading system was developed by a group of experts in WHO[56, 57]. The target to develop it is to enable relatively inexperienced observers to grade the most common types of cataracts reliably and efficiently. It uses four severity levels for grading NC, CC, and PSC based on four standard images accordingly.

3.6 Fundus image-based cataract classification system

Xu et al.[58] proposed a fundus image-based cataract classification system (FCS) by observing the blur level. They used five levels to evaluate the blur levels on fundus images: Grade 0: Clear; Grade 1: The small vessel region was blurred; Grade 2: The larger branches of retinal vein or artery were blurred; Grade 3: The optic disc region was blurred; Grade 4: The whole fundus image was blurred.

Discussion. From the six existing cataract classification/grading systems, we can conclude that five cataract classification systems are built on slit lamp images, and one is built on fundus images, which can explain that most existing cataract works based on these two ophthalmic imaging modalities. However, these cataract classification/grading systems are subjective due to the limitations of these two imaging devices. Furthermore, to improve the precision of cataract diagnosis and the efficiency of cataract surgery, it is necessary to develop new and objective cataract classification systems on other ophthalmic image modalities, e.g., AS-OCT images.

4 Datasets

This section introduces ophthalmic image datasets used for cataract classification/grading, which can be grouped into private and public datasets.

4.1 Private datasets

ACHIKO-NC dataset[59]. ACHIKO-NC is the slit lamp lens dataset selected from the SiMES I database, used to grade nuclear cataracts. It comprises 5 378 images with decimal grading scores (0.3 to 5.0). Professional clinicians determine the grading score of each slit lamp image. ACHIKO-NC is a widely used dataset for automatic nuclear cataract grading according to existing works.

ACHIKO-Retro dataset[59]. ACHIKO-Retro is the retro-illumination lens image dataset selected from the SiMES I database, used to grade CC and PSC. Each lens has two eye image types: anterior image and posterior image. The anterior image focuses on the plane centered in the anterior cortex region, and the posterior image focuses on the posterior capsule region. Most previous CC and PSC grading works were conducted on the ACHIKO-Retro dataset.

CC-Cruiser dataset[60]. CC-Cruiser is the slit lamp image dataset collected from Zhongshan Ophthalmic Center (ZOC) of Sun Yat-Sen University, China, which is used for cataract screening. It is comprised of 476 normal images and 410 infantile cataract images.

Multicenter dataset[60]. Multicenter is the slit lamp image dataset, which is comprised of 336 normal images and 421 infantile cataract images. It was collected from four clinical institutions: the Central Hospital of Wuhan, Shenzhen Eye Hospital, Kaifeng Eye Hospital, and the Second Affiliated Hospital of Fujian Medical University, China.

4.2 Public datasets

EyePACS dataset[61]. EyePACS is the fundus image dataset collected from EyePACS, LLC (https://www.eyepacs.com), a free platform for retinopathy screening, used to classify different levels of cataracts. It is made available by California Healthcare Foundation (https://www.kaggle.com/c/diabetic-retinopathy-detection). The dataset comprises 88 702 fundus retinal images in which 1 000 non-cataract images and 1 441 cataract images are provided.

HRF dataset[62]. The high-resolution fundus (HRF) image database is selected from different open-access datasets: structured analysis of the retina (STARE)[63], standard diabetic retinopathy database (DIARE-TDB0)[64], e-ophtha[65], methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology (MESSIDOR) database[66], digital retinal images for vessel extraction (DRIVE) database[67], fundus image registration (FIRE)[68] dataset, digital retinal images for optic nerve segmentation database (DRIONS-DB)[69], Indian diabetic retinopathy image dataset (ID-RiD)[70], available datasets released by Dr. Hossein Rabbani[71], and other Internet sources.

5 Machine learning techniques

This section mainly investigates on ML techniques for cataract classification/grading over the years, which comprises conventional ML methods and deep learning methods.

5.1 Conventional machine learning methods

Over the past years, scholars have developed massive state-of-the-art conventional ML methods to automatically classify/grade cataract severity levels, aiming to assist clinicians in diagnosing cataracts efficiently and accurately. These methods consist of feature extraction and classification/grading, as shown in Fig. 7. Table 1 summarizes conventional ML methods for cataract classification/grading based on different ophthalmic images.
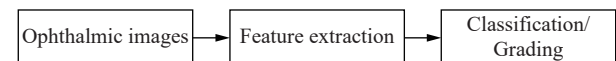

Fig. 7 Flowchart of conventional machine learning based classification and grading

5.1.1 Feature extraction

Considering the characteristics of different imaging techniques and cataract types, we introduce feature extraction methods based on ophthalmic image modalities.

Slit lamp image. The procedure to extract features from slit lamp images is comprised of the lens structure

Table 1 Conventional ML methods for cataract classification/grading based on ophthalmic images

Literature	Method	Image type	Year	Application	Cataract type
[5]	ASM + SVR	Slit lamp image	2009	Grading	NC
[9]	ASM + Ranking	Slit lamp image	2009	Grading	NC
[72]	ASM + SVR	Slit lamp image	2010	Grading	NC
[73]	ASM + Ranking	Slit lamp image	2010	Grading	NC
[74]	ASM + LR	Slit lamp image	2007	Grading	NC
[10]	SF + SWLR	Slit lamp image	2016	Grading	NC
[11]	BOF + GSR	Slit lamp image	2013	Grading	NC
[75]	SVM	Slit lamp image	2016	Grading	NC
[76]	SVM	Slit lamp image	2017	Grading	NC
[77]	SVM	Slit lamp image	2017	Grading	NC
[13]	IGB	Slit lamp image	2014	Grading	NC
[76]	RF	Slit lamp image	2017	Classification	NC

Table 1 (continued) Conventional ML methods for cataract classification/grading based on ophthalmic images

Literature	Method	Image type	Year	Application	Cataract type
[78]	SRCL	Slit lamp image	2018	Grading	NC
[79]	Hough circular transform	Slit lamp image	2019	Classification	NC
[80]	RF	Slit lamp image	2019	Classification	PC
[80]	NB	Slit lamp image	2019	Classification	PC
[6]	Canny + Spatial filter	Retroillumination image	2010	Classification	PSC
[81]	EF + Canny	Retroillumination image	2008	Classification	PSC
[82]	Canny	Retroillumination image	2011	Classification	PSC
[83]	MRF	Retroillumination image	2017	Classification	PSC
[12]	LDA	Retroillumination image	2011	Classification	PSC & CC
[84]	Radial-edge & Region-growing	Retroillumination image	2008	Classification	CC
[85]	PCA	Ultrasonic image	2015	Classification	Cataract
[85]	Bayes	Ultrasonic image	2015	Classification	Cataract
[85]	KNN	Ultrasonic image	2015	Classification	Cataract
[85]	SVM	Ultrasonic image	2015	Classification	Cataract
[85]	FLD	Ultrasonic image	2015	Classification	Cataract
[75]	SVM	Ultrasonic image	2016	Classification	Cataract
[75]	RF	Ultrasonic image	2016	Classification	Cataract
[75]	Bayes network	Ultrasonic image	2016	Classification	Cataract
[86]	PCA + SVM	Ultrasonic image	2014	Classification	Cataract
[87]	Nakagami distribution +CRT	Ultrasonic image	2014	Classification	Cataract
[88]	SVM	Ultrasonic image	2013	Classification	Cataract
[15]	GLCM + KNN	Digital camera image	2015	Classification	Cataract
[89]	GLCM + KNN	Digital camera image	2015	Classification	Cataract
[16]	K-Means	Digital camera image	2016	Classification	Cataract
[14]	IMF	Digital camera image	2016	Classification	Cataract
[17]	SVM	Digital camera image	2016	Classification	Cataract
[90]	DWT + MDA	Fundus image	2015	Classification	Cataract
[18]	Wavelet-SVM	Fundus image	2015	Classification	Cataract
[18]	Texture-SVM	Fundus image	2015	Classification	Cataract
[18]	Stacking	Fundus image	2015	Classification	Cataract
[91]	PCA + Bagging	Fundus image	2015	Classification	Cataract
[91]	PCA + RF	Fundus image	2015	Classification	Cataract
[92]	Multi-feature fusion & Stacking	Fundus image	2019	Classification	Cataract
[91]	PCA + GBDT	Fundus image	2015	Classification	Cataract
[91]	PCA + SVM	Fundus image	2015	Classification	Cataract
[39]	Haar wavelet + Voting	Fundus image	2019	Classification	Cataract
[19]	SVM + GA	Fundus image	2017	Classification	Cataract
[93]	AWM + SVM	Fundus image	2019	Classification	Cataract
[94]	DT	Fundus image	2016	Classification	Cataract
[94]	Bayesian network	Fundus image	2016	Classification	Cataract
[95]	DWT+SVM	Fundus image	2019	Classification	Cataract
[95]	SSL	Fundus image	2019	Classification	Cataract
[96]	RF	AS-OCT image	2021	Classification	NC
[96]	SVM	AS-OCT image	2021	Classification	NC

detection and feature extraction.

Fig. 8 offers a representative slit lamp image-based feature extraction flowchart. Firstly, according to the histogram analysis results of the lens, the foreground of the lens is detected by setting the pixel threshold, and the background of the lens is even based on slit lamp images. Afterward, we analyze the profile on the horizontal median line of the image. The largest cluster on the line is detected as the lens, and the centroid of the cluster is detected as the horizontal coordinate of the lens center. Then, we get the profile on the vertical line through the point. Finally, the center of the lens is estimated, and the lens can be further estimated as an ellipse with the semi-major axis radius estimated from the horizontal and vertical profile.
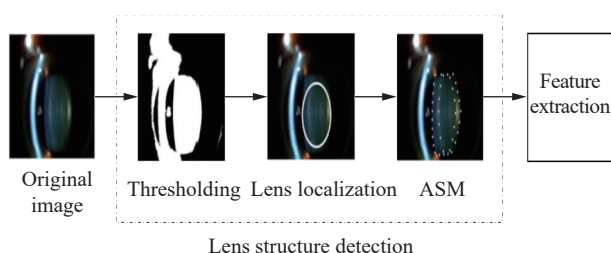
Fig. 8　Procedures for extracting features from slit lamp images for automatic nuclear cataract classification/grading

The lens contour or shape needs to be captured by following the lens location. Researchers commonly used the active shape model (ASM) method for lens contour detection[6, 72, 81] and achieved 95% accuracy in lens structure detection. The ASM can describe the object shape through an iterative refinement procedure to fit an example of the object into a new image based on the statistical models[74]. Based on the detected lens contour, many feature extraction methods have been proposed to extract informative features, such as bag-of-features (BOF), grading protocol-based, semantic reconstruction-based, and statistical texture analysis methods[6, 9–13].

Retroillumination image. It also consists of two stages for feature extraction based on retroillumination images: pupil detection and opacity detection[6, 12, 81, 82], as shown in Fig. 9. Researchers usually use a combination of the Canny edge detection method, the Laplacian method, and the convex hull method to detect the edge pixel in the pupil detection stage. The non-linear least-square fitting method is used to fit an ellipse based on the detected pixels. In the opacity detection stage, the input image is first transformed into the polar coordinate. Based on the polar coordinate, classical image processing methods are applied to detect opacity, such as global threshold, local threshold, edge detection, and region growing. Besides the above methods, Zhang and Li[83] use Watershed and Markov random fields (MRF) to detect the lens opacity. The results showed that the proposed
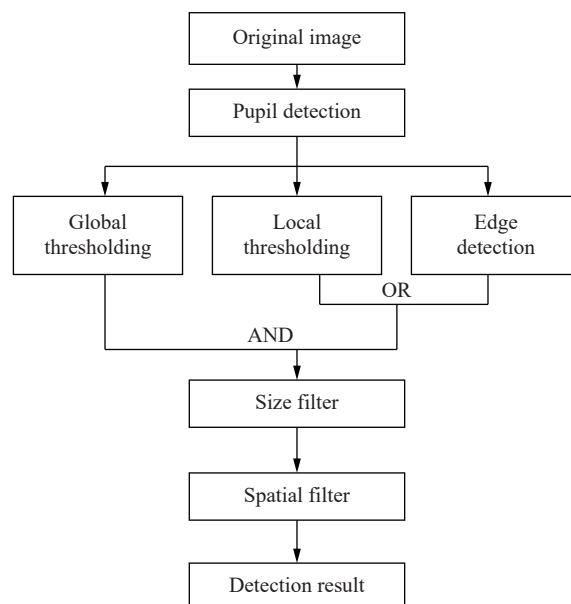

Fig. 9　Procedure for extracting features from retroillumination images for cortical cataract (CC) and posterior subcapsular cataract (PSC) diagnosis

framework got competitive results in PSC detection.

Ultrasound images & digital camera images & AS-OCT images. For ultrasound images, researchers adopt the Fourier transform (FT) method, textural analysis method, and probability density to extract features[75, 86, 87].

The procedure for extracting features from digital camera images is the same as those used for slit lamp images. However, different image processing methods are used to extract features, such as Gabor wavelet transform, gray level co-occurrence matrix (GLCM), morphological image feature, and Gaussian filter[14–16, 89].

The steps to detect the lens region for AS-OCT images are also similar to slit lamp images. Zhang et al.[96] use intensity-based statistics and intensity histogram methods to extract image features from AS-OCT images.

Fundus image. Over the years, researchers have developed various wavelet transform methods to preprocess fundus images for extracting valuable features, as shown in Fig. 10, such as discrete wavelet transform (DWT), discrete cosine transform (DCT), Haar wavelet transform, and top-bottom hat transformation[39, 90, 97].
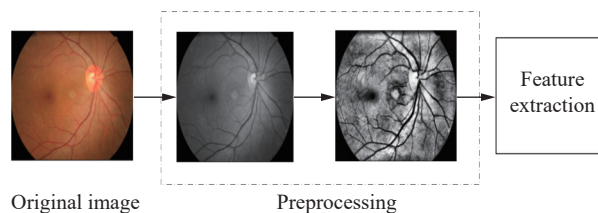

Fig. 10　Procedures to extract features from fundus images for cataract classification or screening

5.1.2 Classification & grading

This section mainly introduces conventional machine learning methods for cataract classification/grading.

Support vector machine. Support vector machine (SVM) is a classical supervised machine learning technique that has been widely used for classification and linear regression tasks. It is a popular and efficient learning method for medical imaging applications. For the cataract grading task, Li et al.[5, 72] utilized support vector machine regression (SVR) to grade the severity level of cataracts and achieved good grading results based on slit lamp images. The SVM classifier is widely used in different ophthalmic images for the cataract classification task. For example, Jiang et al.[76, 77] applied SVM to classify cataract severity levels based on slit lamp images. For other ophthalmic image types, SVM also achieves good results based on extracted features[17−19].

Linear regression. Linear regression (LR) is one of the most well-known ML methods and has been used to address different learning tasks. The concept of LR is still a basis for other advanced techniques, like deep neural networks. Linear functions determine their model in LR, whose parameters are learned from the data by training. Li et al.[74] first studied automatic cataract grading with LR based on slit lamp images and achieved good grading results. Followed by [74], Xu et al.[10, 11] proposed the group sparsity regression (GSR) and similarity weighted linear reconstruction (SWLR) for cataract grading and achieved better grading results.

K-nearest neighbors. K-nearest neighbors (KNN) method is a simple, easy-to-implement supervised machine learning method used for classification and regression tasks. It uses similarity measures to classify new cases based on stored instances. Fuadah et al.[15] used the KNN to detect cataracts on digital camera images and achieved 97.2% accuracy. Caxinha et al.[85] also used KNN for cataract classification based on ultrasonic images collected from the animal model.

Ensemble learning method. Ensemble learning method uses multiple machine learning methods to solve the same problem and usually obtains better classification performance. Yang et al.[18, 39, 92] have used several ensemble learning methods for cataract classification, such as Stacking, Bagging, and Voting. Ensemble learning methods achieved better cataract grading results than single machine learning methods.

Ranking. Ranking denotes a relationship within the list in descending/ascending order. Huang et al.[9, 73] applied the ranking strategy to automatic cataract grading by computing the score of each image from the learned ranking function such as RankBoost and Ranking SVM and achieved competitive performance.

Other machine learning methods. Apart from the above-mentioned conventional ML methods, other advanced ML methods are also proposed for automatic

cataract classification/grading, such as Markov random field (MRF), random forest (RF), Bayesian network, linear discriminant analysis (LDA), K-means, and decision tree (DT), etc. For example, Zhang and Li[83] applied Markov random field (MRF) for automatic CC classification and achieved good results. Song et al.[95] used the semi-supervised learning (SSL) framework for cataract classification based on fundus images and achieved competitive performance.

Furthermore, we can draw the following conclusions:

1) For feature extraction, despite the characteristics of ophthalmic images, various image processing techniques are developed to extract useful features, like edge detection method, wavelet transform method, texture extraction methods, etc. However, no previous work systematically compared these feature extraction methods based on the same ophthalmic images, providing a standard benchmark for other researchers. Furthermore, existing works have not verified the effectiveness of a classical feature extraction method on different ophthalmic images, which is significant for the generalization ability of a feature extraction method and building commonly-used feature extraction baselines on cataract classification/grading tasks.

2) For classification/grading, researchers have made great efforts in developing state-of-the-art ML methods to recognize cataract severity levels on ophthalmic images and demonstrated that ML methods can achieve competitive performance on extracted features. We found that existing research has not made a comparison between ML methods comprehensively based on the same ophthalmic image or different ophthalmic image types; thus, it is necessary to build conventional ML baselines for cataract classification/grading, which can help researchers reproduce previous works and prompt the development of cataract classification/grading tasks.

5.2 Deep learning methods

In recent years, with the rapid development of deep learning techniques, many deep learning methods ranging from the artificial neural network (ANN), multilayer perceptron (MLP) neural network, backpropagation neural network (BPNN), convolutional neural network (CNN), recurrent neural network (RNN), and attention mechanism, to transformer-based methods, have been applied to solve different learning tasks such as image classification and medical image segmentation. In this survey, we mainly pay attention to deep learning methods in cataract classification tasks. Table 2 summaries of deep learning methods for cataract classification/grading based on different ophthalmic images.

5.2.1 Multilayer perceptron neural networks

Multilayer perceptron neural network is one type of artificial neural network composed of multiple hidden layers. Researchers often combined a MLP with hand-engin-

Table 2   Deep learning methods for cataract classification/grading on different ophthalmic images

Literature	Method	Image type	Year	Application	Cataract type
[98]	CNN+RNN	Slit lamp image	2015	Grading	NC
[99]	CNN	Slit lamp image	2017	Classification	PCO
[76]	CNN	Slit lamp image	2017	Classification	PCO
[8]	CNN	Slit lamp image	2017	Classification	Cataract
[100]	CNN	Slit lamp image	2018	Classification	Cataract
[101]	CNN+RNN	Slit lamp image	2018	Classification	PCO
[102]	Faster R-CNN	Slit lamp image	2019	Grading	NC
[103]	CNN	Slit lamp image	2019	Classification	NC
[104]	CNN	Slit lamp image	2019	Classification	NC
[105]	CNN	Slit lamp image	2019	Classification	NC
[106]	CNN	Slit lamp image	2020	Classification	NC
[107]	CNN	Slit lamp image	2021	Classification	Cataract
[60]	R-CNN	Slit lamp image	2021	Classification	PCO
[108]	CNN	Slit lamp image	2021	Classification	cataract
[109]	Wavelet + ANN	Digital camera image	2016	Classification	Cataract
[75]	MLP	Ultrasonic image	2016	Classification	Cataract
[110]	Attention	Ultrasonic image	2020	Classification	Cataract
[111]	CNN	Ultrasonic image	2021	Classification	Cataract
[18]	BPNN	Fundus image	2015	Classification	Cataract
[112]	CNN	Fundus image	2017	Classification	Cataract
[113]	CNN	Fundus image	2017	Classification	Cataract
[92]	CNN	Fundus image	2019	Classification	Cataract
[97]	MLP	Fundus image	2019	Classification	Cataract
[97]	CNN	Fundus image	2019	Classification	Cataract
[112]	CNN	Fundus image	2019	Classification	Cataract
[62]	CNN	Fundus image	2019	Classification	Cataract
[114]	CNN	Fundus image	2020	Classification	Cataract
[115]	CNN	Fundus image	2021	Classification	Cataract
[116]	CNN+RNN	Fundus image	2021	Classification	Cataract
[117]	CNN	Fundus image	2021	Classification	Cataract
[118]	CNN	Fundus image	2022	Classification	Cataract
[119]	CNN	AS-OCT image	2020	Classification	NC
[120]	CNN	AS-OCT image	2021	Classification	NC
[121]	Attention	AS-OCT image	2021	Classification	NC
[20]	Attention	AS-OCT image	2022	Classification	NC
[122]	Attention	AS-OCT image	2022	Classification	NC

eered feature extraction methods for cataract classification to obtain the expected performance. Zhou et al.[97] combined a shallow MLP model with feature extraction methods to classify cataract severity levels automatically. Caixinha et al.[87] utilized the MLP model for cataract classification and achieved 96.7% of accuracy.

Deep neural networks (DNNs) comprise many hidden layers, and they are capable of capturing more informative feature representations from ophthalmic images through comparisons to MLP. DNN is usually used as dense (fully-connected) layers of CNN models. Zhang et al.[92, 113, 114] used DNN models for cataract classification and achieved an accuracy of over 90% on fundus images.

Recently, several works[123–127] have demonstrated that constructing deep network architectures purely on MLPs can get competitive performance on ImageNet classification task with spatial and channel-wise token mixing, to aggregate the spatial information and build dependencies

among visual features. These MLP-based models have promising results in classical computer vision tasks like image classification, semantic segmentation, and image reconstruction. However, to the best of our knowledge, MLP-based models have not been used to tackle ocular diseases tasks, including cataracts based on different ophthalmic images, which can be an emerging research direction for cataract classification/grading in the future.

5.2.2 Convolutional neural networks

Convolutional neural networks have been widely used in the field of ophthalmic image processing and achieved surprisingly promising performance[128–138]. CNN consists of an input layer, multiple convolutional layers, multiple pooling layers, multiple fully-connected layers, and an output layer. The function of convolutional layers is to learn low-level, middle-level, and high-level feature representations from the input images through massive convolution operations in different stages.

For slit lamp images, previous works usually used classical image processing methods to localize the region of interest (ROI) of the lens, and the lens ROI is used as the inputs features of CNN models[99, 102, 139]. For example, Liu et al.[99] proposed a CNN model to detect and grade severity levels of posterior capsular opacification (PCO), which used the ROI of the lens as input of CNN. Xu et al.[102] used original images as input for CNN models to detect cataracts automatically through Faster R-CNN. Long et al.[8] developed an artificial intelligence platform for congenital cataract diagnosis and obtained good diagnosis results. Wu et al.[105] proposed a universal artificial intelligence platform for nuclear cataract management and achieved good results.

For fundus images, Zhang et al.[92, 112–114] achieved competitive classification results with deep CNNs. Zhou et al.[97] proposed the EDST-ResNet model for cataract classification based on fundus images where they used discrete state transition function[140] as the activation function, to improve the interpretability of CNN models. Fig. 11 provides a representative CNN framework for cataract classification on fundus images, which can help audiences know this task easily.

For AS-OCT images, Yin et al.[141–143] proposed CNN-based segmentation frameworks for automatic lens region segmentation based on AS-OCT images, which can help ophthalmologists locate and diagnose different types of cataract efficiently. Zhang et al.[119] proposed a novel CNN model named GraNet for nuclear cataract classification on AS-OCT images but achieved poor results. Xiao et al.[120] used a 3D ResNet architecture for cataract screening based on 3D AS-OCT images. Zhang et al.[122] tested the NC classification performance of state-of-the-art CNNs like ResNet, VGG, and GoogleNet on AS-OCT images, and the results showed that EfficientNet achieved the best performance.

5.2.3 Recurrent neural networks

Recurrent neural networks are typical feedforward neural network architectures where connections between nodes form a directed or undirected graph along a temporal sequence. RNNs are skilled in processing sequential data effectively for various learning tasks[144], e.g., speech recognition. Over the past decades, many RNN variants have been developed to address different tasks, where long short-term memory (LSTM) network is the most representative one. However, researchers have not used pure RNN architecture to classify cataract severity levels yet, which can be a research direction for automatic cataract classification.

5.2.4 Attention mechanisms

Over the past years, attention mechanisms have been proven useful in various fields, such as computer vision[145], natural language processing (NLP)[146], and medical data processing[147]. Generally, the attention mechanism can be taken as an adaptive weighting process in a local-global manner according to feature representations of feature maps. In the field of computer vision, attention can be classified into five categories: channel, spatial, temporal attention, branch attention, and attention combinations such as channel & spatial attention[148]. Each attention category has a different influence on the field of computer vision. Researchers have recently used attention-based CNN models for cataract classification on different ophthalmic images. Zhang et al.[110] proposed a
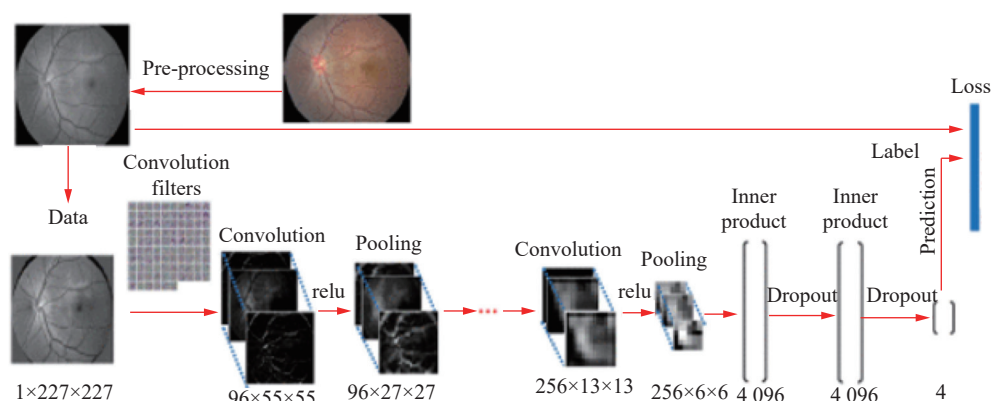

Fig. 11   An example of a CNN model for cataract classification on fundus image[114]

residual attention-based network for cataract detection on ultrasound images. Xiao et al.[121] applied a gated attention network (GCA-Net) to classify nuclear cataract severity levels on AS-OCT images and obtained good performance. Zhang et al.[122] presented a mixed pyramid attention network for AS-OCT image-based nuclear cataract classification in which they constructed the mixed pyramid attention (MPA) block by considering the relative significance of local-global feature representations and different feature representation types, as shown in Fig. 12.
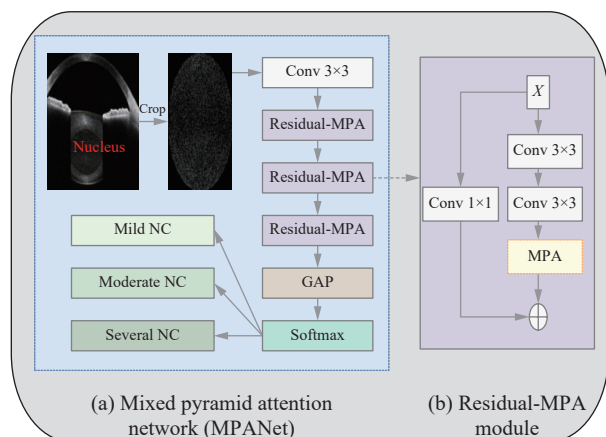
Fig. 12   An example of attention-based CNN architecture for cataract classification on AS-OCT images

Especially self-attention is one representative kind of attention mechanism. Due to its effectiveness in capturing long-range dependencies and generality, it has been playing an increasingly important role in a variety of learning tasks[146, 149, 150]. Massive deep self-attention networks (e.g., transformer) have achieved state-of-the-art performance through mainstream CNNs on visual tasks. Vision transformer (ViT)[151] is a first-pure transformer architecture proposed for image classification and gets promising performance. Recently, researchers also extended the transformer-based models for different medical image analysis tasks[152]. However, no current research work has used transformer-based architectures to recognize cataract severity levels.

5.2.5 Hybrid neural networks

A hybrid neural network indicates that a neural network is composed of two or more than two types of deep neural networks. In recent years, researchers have increasingly used hybrid neural networks to address different learning tasks[153], due to their ability to inherit the advantages from different neural network architectures, such as CNN, MLP, and transformer. Gao et al.[98, 101, 116] proposed a hybrid neural network models for cataract classification/grading by considering the characteristics of RNN and CNN models, as shown in Fig. 13. In the future, we believe that more advanced hybrid neural network models will be designed for cataract classification/grading based on different ophthalmic image modalities.
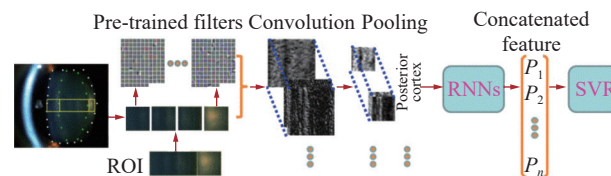
Fig. 13   A hybrid neural network model for nuclear cataract grading on slit lamp images, comprised of a CNN model and an RNN model[98]

Discussion. From Tables 1 and 2, we can conclude as follows:

1) Ophthalmic image perspective. Slit lamp images and fundus images account for most automatic cataract classification/grading works. This is because these two ophthalmic image modalities are easy to access and have clinical gold standards. Except for ultrasound images collected from the animal model and human subjects, five other ophthalmic imaging modalities were used for automatic cataract classification/grading.

2) Cataract classification/grading/perspective. Existing works mainly focused on cataract screening, and most of them achieved over 80% in the accuracy. The number of cataract classification works is more than the number of cataract grading works. Since discrete labels are easy to access and confirm through comparisons to continuous labels.

3) Publication year perspective. Conventional machine learning methods were first used to classify or grade cataracts automatically. With the emergence of deep learning, researchers have gradually constructed advanced deep neural network architectures to automatically predict severity levels of cataracts due to their surprising performance. However, the interpretability of deep learning methods is worse than conventional machine learning methods.

4) Learning paradigm perspective. Most previous machine learning methods belong to supervised learning methods, and only two existing methods used semi-supervised learning methods. Specifically, unsupervised and semi-supervised learning methods have achieved competitive performance in computer vision and NLP, which have not been widely applied to automatic cataract diagnosis.

5) Deep neural network architecture perspective. According to Table 2, we can see that CNNs account for over 50% of deep neural network architectures. Two reasons can explain that: 1) Ophthalmic images are the most commonly-used way for cataract diagnosis by clinicians. 2) Compared with RNN and MLP, CNNs are skilled at processing image data. To enhance the precision of cataract diagnosis, it is better to combine image data with non-imaging data, e.g., age.

6) Performance comparison on private/public datasets. Pratap and Kokil[115] made a comparison of AlexNet, GoogleNet, ResNet, ResNet50, and SqueezeNet

for the cataract classification based on the public Eye-PACS dataset, and results showed all CNN models obtained more than 86% accuracy, and AlexNet got the best performance. Zhang et al.[122] offered the NC classification results of attention-based CNNs (CBAM, ECA, GCA, SPA, and SE) on the private NC dataset. The results showed SE obtained better performance than other strong attention methods. Furthermore, Pratap et al.[62, 102, 114, 115] used pre-trained CNNs for automatic cataract classification on the EyePACS dataset according to transfer learning strategy. They also concluded that training pre-trained CNNs benefited the cataract classification performance more than training CNNs from scratch.

7) Data augmentation techniques. Researchers use the commonly-used method for cataract classification/ grading to augment ophthalmic image data for deep neural networks, such as flipping, cropping, rotation, and noise injection. To further validate or enhance the generalization ability of deep neural network models on cataract classification/grading tasks, other data augmentation techniques should be considered: translation, color space transformations, random erasing, adversarial training, meta-learning, etc.[154]

6 Evaluation measures

This section introduces evaluation measures to assess the performance of cataract classification/grading. In this survey, classification denotes the cataract labels used for learning are discrete, e.g., 1, 2, 3, 4, while grading denotes cataract labels are continuous, such as 0.1, 0.5, 1.0, 1.6, and 3.3.

For cataract classification, accuracy (ACC), *sensitivity* (recall), *specificity*, *precision*, $F1$-*measure*($F1$), and *G-mean* are commonly used to evaluate the classification performance[39, 90, 101].

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{5}$$

$$G\text{-}mean = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}} \tag{6}$$

where TP, FP, TN and FN denote the numbers of true positives, false positives, true negatives, and false negatives, respectively. Other evaluation measures such as receiver operating characteristic curve (ROC), the area under the ROC curve (AUC), and kappa coefficient[39] are also applied to measure the overall performance.

Four evaluation measures are used to evaluate the overall cataract grading performance of methods, titled the exact integral agreement ratio R_0, the percentage of decimal grading errors ≤ 0.5 $R_{e0.5}$, the percentage of decimal grading errors ≤ 1.0 $R_{e1.0}$, and the mean absolute error ε[10, 73, 74, 155, 156].

$$R_0 = \frac{|\lceil G_{gt} \rceil = \lceil G_{pr} \rceil|_0}{N} \tag{7}$$

$$R_{e0.5} = \frac{||G_{gt} - G_{pr}| \leq 0.5|_0}{N} \tag{8}$$

$$R_{e1.0} = \frac{||G_{gt} - G_{pr}| \leq 1.0|_0}{N} \tag{9}$$

$$\varepsilon = \frac{\sum |G_{gt} - G_{pr}|}{N} \tag{10}$$

where G_{gt} and G_{pr} denote the ground-truth grade and the predicted grade. $\lceil . \rceil$ is the ceiling function, $|.|$ is the absolute function, $|.|_0$ a function that counts the number of non-zero values, and N denotes the number of images.

7 Challenges and possible solutions

Although researchers have made significant development in automatic cataract classification/grading over the years, this field still has challenges. This section presents these challenges and gives possible solutions.

7.1 Lacking public cataract datasets

Public cataract datasets are very critical issues for cataract classification/grading. Previous works have made tremendous progress in automatic cataract/grading[11, 86, 90, 157]. There is no public and standard ophthalmology image dataset available except for public challenges datasets and multiple ocular diseases. Hence, it is difficult for researchers to follow previous works because the cataract dataset is unavailable. To this problem, it is necessary and significant to build public and standard ophthalmology image datasets based on standardized medical data collection and storage protocols. Public cataract datasets can be collected from hospitals and clinics with ophthalmologists′ help. This dataset collection mode can ensure the quality and diversity of cataract data and help researchers develop more advanced ML methods.

7.2 Developing standard cataract classification/grading protocols based on new ophthalmic imaging modalities

Most existing works used LOCS III as the clinical gold diagnosis standard to grade/classify cataract severity levels for scientific research purposes and clinical practice. However, the LOCS III is developed based on slit lamp images, which may not work well for ophthalmic images, such as fundus images, AS-OCT images, and ultrasonic images. To solve this problem, researchers have made much effort in constructing new cataract classification/grading standards for other ophthalmic image types. For example, researchers refer to the WHO Cataract Grading System and develop a cataract classification/grading protocol for fundus images[58] based on clinical research and practice, which is widely used in automatic fundus-based cataract classification[92, 113–115].

To address the issue of developing standard cataract classification/grading protocols for new eye images, e.g., AS-OCT images, in this survey, we propose two possible solutions for reference as follows.

1) To develop a cataract grading protocol based on clinical verification. Grulkowski et al.[45, 46] used AS-OCT images to observe the lens opacities of patients based on the LOCS III in clinical practice, and statistical results showed that high correlation between cataract severity levels and clinical features with inter-class and intra-class analysis. The clinical finding may provide clinical support and reference. Hence, it is likely to develop a referenced cataract classification/grading protocol for AS-OCT images based on a clinical verification method like the LOCS III.

2) To build the mapping relationship between two ophthalmic imaging modalities. Lens opacity is an important indicator to measure the severity level of cataracts, which are presented in different ophthalmic images with different forms in clinical research. Therefore, there is potential to construct the mapping relationship between two different ophthalmic imaging modalities to develop a new cataract classification/grading protocol by comparing the lens opacities, e.g., fundus image-based cataract classification system WHO Cataract Grading System.

Furthermore, to verify the effectiveness of new standard cataract grading protocols, we must collect multiple center data from hospitals in different regions.

7.3 How to annotate cataract images accurately

Data annotation is a challenging problem for the medical image analysis field, including cataract image analysis, since it is the significant base for accurate ML-based disease diagnosis. However, clinicians cannot label massive cataract images manually[158, 159] because it is ex-

pensive, time-consuming, and objective. To address these challenges, we offer the following solutions:

1) **Semi-supervised learning.** Song et al.[94, 95] used the semi-supervised learning strategy to recognize cataract severity levels on fundus images, achieving expected performance. It is likely to utilize weakly supervised learning methods to learn useful information from labeled cataract images and let the method automatically label unlabeled cataract images according to the learned information.

2) **Unsupervised learning.** Recent works have shown that deep clustering/unsupervised learning techniques can help researchers acquire labels positively rather than negatively[160–164]. Thus, we can actively apply deep clustering/unsupervised learning methods to label cataract images in the future.

3) **Content-based image retrieval.** The content-based image retrieval (CBIR)[165–168] technique has been widely used for different tasks based on different image features, which can also be utilized to annotate cataract images by comparing the testing images with standard images.

7.4 How to classify/grade cataracts accurately for precise cataract diagnosis

Most of the previous works focused on cataract screening, and few of the previous works considered clinical cataract diagnosis, especially for cataract surgery planning. This is because different cataract severity levels and cataract types should clinically take the corresponding treatments. Hence, it is necessary to develop state-of-the-art methods to accurately classify the severity levels of cataracts, and this survey provides the following research directions.

1) **Clinical prior knowledge injection.** Furthermore, cataracts are associated with various factors[169], e.g., sub-lens regions[10, 12, 45, 46, 85], which can be considered as domain knowledge of cataracts. Thus, we can infuse the domain knowledge into deep networks for automatic cataract classification/grading according to the characteristics of ophthalmic images. For example, Zhang et al.[122] incorporated clinical features in attention-based network design for classification.

2) **Multi-task learning for classification and segmentation.** Over the past decades, multi-task learning techniques have been successfully applied to various fields, including medical image analysis. Xu et al.[102] used the Faster-RCNN framework to detect the lens region and grade nuclear cataract severity levels on slit lamp images and achieved competitive performance. Yin et al.[141, 142] proposed a deep segmentation network framework for automatic lens subregion segmentation based on AS-OCT images, which is a significant basis for cataract diagnosis and cataract surgery planning. Moreover, more multi-task

learning frameworks should be developed for cataract classification and lens segmentation, considering multi-task learning framework usually keeps a good balance between performance and complexity.

3) Transfer learning. Over the years, researchers have used the transfer learning method to improve cataract classification performance[102, 114] with pre-trained CNNs. Large deep neural network models usually perform better than small deep neural network models based on massive data, which previous works have demonstrated. However, it is challenging to collect massive data in the medical field. Thus, it is vital for us to develop transfer learning strategies to take full use of pre-trained parameters for large deep neural network models, to further improve the performance of cataract-related tasks.

4) Multimodality learning. Previous works only used ophthalmic image type for cataract diagnosis[10, 12, 85, 170]. Multimodality learning[171–174] has been utilized to tackle different medical image tasks, which also can be used for automatic cataract classification/grading based on multi-ophthalmic images or the combination of ophthalmic images and non-image data. Multimodality data can be classified into image data and non-image data. However, it is challenging to use multimodal images for automatic cataract classification. Two reasons can explain this: 1) Only slit lamp images and fundus images have standard cataract classification systems, which have no correlation relationship between them; 2) Existing classification systems are subjective, and it is challenging for clinicians to label two levels of cataract severity correctly for different ophthalmic images. Furthermore, we can combine image and non-image data for automatic cataract diagnosis because it is easy to collect non-image data such as age and sex associated with cataracts. Furthermore, recent studies[169, 175] have used non-image data for PC and PCO diagnosis, which demonstrated that it has the potential to use multimodality data to improve the precision of cataract diagnosis

5) Image denoising. Image noise is an important factor in affecting automatic cataract diagnosis on ophthalmic images. Researchers have proposed different methods to remove noise from images based on the characteristics of ophthalmic images, such as Gaussian filter, discrete wavelet transform (DWT), discrete cosine transform (DCT), Haar wavelet transform, and its variants[89, 90, 97]. Additionally, recent research has begun to use the GAN model for medical image denoising and achieved good results.

7.5 Improving the interpretability of deep learning methods

Deep learning methods have been widely used for cataract classification/grading. However, deep learning methods are still considered as a black box without explaining why they make good prediction results or poor

prediction results. Therefore, it is necessary to give reliable explanations for the predicted results based on deep learning methods. Zhang et al.[97] visualized weight distributions of deep convolutional and dense layers to explain the predicted results of deep learning methods. It is likely to analyze the similarities between the weights of convolutional layers to describe the predicted results.

Zhang et al.[142] indicated that infusing domain knowledge into networks can improve the cataract segmentation results based on AS-OCT images and enhance interpretability. It is promising for researchers to combine deep learning methods with different domain knowledge to improve interpretability of deep networks.

Modern neural networks[176] are usually poorly calibrated. Müller et al.[177] also indicated that improving the confidence of predicted probability can explain the predicted results. It is a potential solution for researchers to enhance interpretability by enhancing the confidence levels of predicted probability.

Moreover, researchers have achieved great progress in the interpretability of deep learning methods in the field of computer vision. Zhou et al.[178–181] used the heat maps to localize discriminative regions where CNNs learned. Kornblith et al.[182] proposed the correlation analysis technique to analyze similarities of neural network representations to explain what neural networks had learned. These interpretability techniques are also used for deep learning based cataract classification works to improve interpretability.

7.6 Mobile cataract screening

In the past years, researchers have used digital camera images for cataract screening based on ML techniques and have achieved good results[14–16, 89, 183–185]. Digital camera images can be accessed through mobile phones, which are widely used worldwide. Thus, there is endless potential to embed powerful digital cameras into mobile phones for cataract screening. Although deep learning methods can achieve excellent cataract screening results on digital camera images, their parameters are too many. One possible solution is to develop lightweight network models. Researchers have achieved significant improvement in designing lightweight deep learning methods over the years[186–188], which demonstrated that deep networks kept a good balance between accuracy and speed, which are important for developing cataract screening-based lightweight deep networks.

7.7 How to evaluate the generalization ability of machine learning methods for other eye disease classification tasks

Researchers have proposed massive machine learning techniques for automatic cataract classification/ grading, including conventional machine learning meth-

ods and deep learning methods, and have obtained promising performance based on different eye images. However, only Zhang et al.[122] validated the generalization ability of the proposed attention-based network on cataract and age-related macular degeneration (AMD) classification task, while others focused on the pure cataract classification task. The following reasons can explain that: 1) Due to the characteristics of eye images, it is difficult to validate the effectiveness of a method on different eye disease classification tasks, especially for conventional machine learning methods. 2) Most existing works focused on specific eye disease tasks like cataract and glaucoma rather than universal methods for medical analysis. 3) Appropriate open-access medical datasets are rare. 4) No previous works have systematically concluded existing methods and built standard baselines.

This survey offers the following solutions:

1) More open-access medical datasets are supposed to be released.

2) It is significant to build baselines of different eye disease tasks, which can provide the standard benchmarks for others to follow.

7.8 Others

In addition to the challenges mentioned above, there are other challenges for medical image analysis, including cataracts, such as label noise[189], domain distribution shift across different datasets, and how to train more robust models. However, these challenges have not been studied in the cataract classification/grading field. Taking into account the real diagnosis requirements, it is necessary to address these challenges by following advanced machine learning methods, such as graph convolution method[190], knowledge distillation, adversarial learning, etc.

8 Conclusions

Automatic and objective cataract classification/grading is an important and hot research topic, helping reduce blindness ratio and improve life quality. This survey reviews recent advances in the ML-based cataract classification/grading on ophthalmic imaging modalities and highlights two research fields: conventional ML and deep learning methods. Though ML techniques have made significant progress in cataract classification and grading, there is still room for improvement. First, public cataract ophthalmic images are needed, which can help build ML baselines. Secondly, automatic ophthalmic image annotation has not been studied seriously and is a potential research direction. Thirdly, developing accurate and objective ML methods for clinical cataract diagnosis is effective to help more people improve their vision. It is necessary to enhance the interpretability of deep learning, which can prompt applications in cataract diagnosis. Finally, previous works have used digital camera images for

cataract screening, which is a good idea by devising lightweight ML techniques that address the cataract screening problem for developing countries with limited access to expensive ophthalmic devices.

Acknowledgements

Open Access

References

[1] R. R. A. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, K. Naidoo, K. Pesudovs, S. Resnikoff, A. Silvester, G. A. Stevens, N. Tahhan, T. Y. Wong, H. R. Taylor. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. *The Lancet Global Health*, vol. 5, no. 9, pp. e888–e897, 2017. DOI: 10.1016/S2214-109X(17) 30293-0.

[2] D. Pascolini, S. P. Mariotti. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, vol. 96, no. 5, pp. 614–618, 2012. DOI: 10.1136/bjophthalmol-2011-300539.

[3] P. A. Asbell, I. Dualan, J. Mindel, D. Brocks, M. Ahmad, S. Epstein. Age-related cataract. *The Lancet*, vol. 365, no. 9459, pp. 599–609, 2005. DOI: 10.1016/S0140-6736(05) 17911-2.

[4] Y. C. Liu, M. Wilkins, T. Kim, B. Malyugin, J. S. Mehta. Cataracts. *The Lancet*, vol. 390, no. 10094, pp. 600–612, 2017. DOI: 10.1016/S0140-6736(17)30544-5.

[5] H. Q. Li, J. H. Lim, J. Liu, D. W. K. Wong, N. M. Tan, S.

J. Lu, Z. Zhang, T. Y. Wong. An automatic diagnosis system of nuclear cataract using slit-lamp images. In *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Minneapolis, USA, pp. 3693–3696, 2009. DOI: 10.1109/IEMBS.2009.5334735.

[6]   H. Q. Li, J. H. Lim, J. Liu, D. W. K. Wong, Y. Foo, Y. Sun, T. Y. Wong. Automatic detection of posterior subcapsular cataract opacity for cataract screening. In *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology*, IEEE, Buenos Aires, Argentina, pp. 5359–5362, 2010. DOI: 10.1109/IEMBS.2010.5626467.

[7]   M. Chew, P. P. C. Chiang, Y. F. Zheng, R. Lavanya, R. Y. Wu, S. M. Saw, T. Y. Wong, E. L. Lamoureux. The impact of cataract, cataract types, and cataract grades on vision-specific functioning using rasch analysis. *American journal of Ophthalmology*, vol. 154, no. 1, pp. 29–38.e2, 2012. DOI: 10.1016/j.ajo.2012.01.033.

[8]   E. P. Long, H. T. Lin, Z. Z. Liu, X. H. Wu, L. M. Wang, J. W. Jiang, Y. Y. An, Z. L. Lin, X. Y. Li, J. J. Chen, J. Li, Q. Z. Cao, D. N. Wang, X. Y. Liu, W. R. Chen, Y. Z. Liu. An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nature Biomedical Engineering*, vol. 1, no. 2, Article number 0024, 2017. DOI: 10.1038/s41551-016-0024.

[9]   W. Huang, H. Q. Li, K. L. Chan, J. H. Lim, J. Liu, T. Y. Wong. A computer-aided diagnosis system of nuclear cataract via ranking. In *Proceedings of the 12th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, London, UK, pp. 803–810, 2009. DOI: 10.1007/978-3-642-04271-3_97.

[10] Y. W. Xu, L. X. Duan, D. W. K. Wong, T. Y. Wong, J. Liu. Semantic reconstruction-based nuclear cataract grading from slit-lamp lens images. In *Proceedings of the 19th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, Athens, Greece, pp. 458–466, 2016. DOI: 10.1007/978-3-319-46726-9_53.

[11] Y. W. Xu, X. T. Gao, S. Lin, D. W. K. Wong, J. Liu, D. Xu, C. Y. Cheng, C. Y. Cheung, T. Y. Wong. Automatic grading of nuclear cataracts from slit-lamp lens images using group sparsity regression. In *Proceedings of the 16th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, Nagoya, Japan, pp. 468–475, 2013. DOI: 10.1007/978-3-642-40763-5_58.

[12] X. T. Gao, H. Q. Li, J. H. Lim, T. Y. Wong. Computer-aided cataract detection using enhanced texture features on retro-illumination lens images. In *Proceedings of the 18th IEEE International Conference on Image Processing*, IEEE, Brussels, Belgium, pp. 1565–1568, 2011. DOI: 10.1109/ICIP.2011.6115746.

[13] R. Srivastava, X. T. Gao, F. S. Yin, D. W. Wong, J. Liu, C. Y. Cheung, T. Y. Wong. Automatic nuclear cataract grading using image gradients. *Journal of Medical Imaging*, vol. 1, no. 1, Article number 014502, 2014. DOI: 10.1117/1.JMI.1.1.014502.

[14] A. U. Patwari, M. D. Arif, N. A. Chowdhury, A. Arefin, I. Imam. Detection, categorization, and assessment of eye cataracts using digital image processing. In *Proceedings of the 1st International Conference on Interdisciplinary*

[15] *Research and Development*, Suomi, Thailand, pp. 22.1–22.5, 2011.

[15] Y. N. Fuadah, A. W. Setiawan, T. L. R. Mengko. Performing high accuracy of the system for cataract detection using statistical texture analysis and K-nearest neighbor. In *Proceedings of International Seminar on Intelligent Technology and its Applications*, IEEE, Surabaya, Indonesia, pp. 85–88, 2015. DOI: 10.1109/ISITIA.2015.7219958.

[16] S. Pathak, B. Kumar. A robust automated cataract detection algorithm using diagnostic opinion based parameter thresholding for telemedicine application. *Electronics*, vol. 5, no. 3, Article number 57, 2016. DOI: 10.3390/electronics5030057.

[17] A. A. Khan, M. U. Akram, A. Tariq, F. Tahir, K. Wazir. Automated computer aided detection of cataract. In *Proceedings of the 3rd International Afro-European Conference for Industrial Advancement* , Mrrakesh, Morrocco, pp. 340–349, 2018. DOI: 10.1007/978-3-319-60834-1_34.

[18] J. J. Yang, J. Q. Li, R. F. Shen, Y. Zeng, J. He, J. Bi, Y. Li, Q. Y. Zhang, L. H. Peng, Q. Wang. Exploiting ensemble learning for automatic cataract detection and grading. *Computer Methods and Programs in Biomedicine*, vol. 124, pp. 45–57, 2016. DOI: 10.1016/j.cmpb.2015.10.007.

[19] Z. Q. Qiao, Q. Y. Zhang, Y. Y. Dong, J. J. Yang. Application of SVM based on genetic algorithm in classification of cataract fundus images. In *Proceedings of IEEE International Conference on Imaging Systems and Techniques*, IEEE, Beijing, China, 2017. DOI: 10.1109/IST.2017.8261541.

[20] X. Q. Zhang, Z. J. Xiao, R. Higashita, W. Chen, Y. Hu, J. Yuan, J. Liu. Nuclear cataract classification based on multi-region fusion attention network model. *Journal of Image and Graphics*, vol. 27, no. 3, pp. 948–960, 2022. DOI: 10.11834/jig.210735. (in Chinese)

[21] Z. Zhang, R. Srivastava, H. Y. Liu, X. Y. Chen, L. X. Duan, D. W. K. Wong, C. K. Kwoh, T. Y. Wong, J. Liu. A survey on computer aided diagnosis for ocular diseases. *BMC Medical Informatics and Decision Making*, vol. 14, no. 1, Article number 80, 2014. DOI: 10.1186/1472-6947-14-80.

[22] I. Shaheen, A. Tariq. Survey analysis of automatic detection and grading of cataract using different imaging modalities. *Applications of Intelligent Technologies in Healthcare*, F. Khan, M. A. Jan, M. Alam, Eds., Cham, Germany: Springer, pp. 35–45, 2019. DOI: 10.1007/978-3-319-96139-2_4.

[23] H. I. M. Lopez, J. C. S. Garcia, J. A. D. Mendez. Cataract detection techniques: A review. *IEEE Latin America Transactions*, vol. 14, no. 7, pp. 3074–3079, 2016. DOI: 10.1109/TLA.2016.7587604.

[24] R. Zafar, M. Sharif, M. Yasmin. A survey on the prevalence of cataract and its accompanying risk factors. *Current Medical Imaging*, vol. 14, no. 2, pp. 251–262, 2018. DOI: 10.2174/1573405613666170331103423.

[25] H. E. Gali, R. Sella, N. A. Afshari. Cataract grading systems: A review of past and present. *Current Opinion in Ophthalmology*, vol. 30, no. 1, pp. 13–18, 2019. DOI: 10.1097/ICU.0000000000000542.

[26] J. H. L. Goh, Z. W. Lim, X. L. Fang, A. Anees, S.

Nusinovici, T. H. Rim, C. Y. Cheng, Y. C. Tham. Artificial intelligence for cataract detection and management. *Asia-Pacific Journal of Ophthalmology*, vol. 9, no. 2, pp. 88–95, 2020. DOI: 10.1097/01.APO.0000656988.16221.04.

[27] A. F. Fercher, H. C. Li, C. K. Hitzenberger. Slit lamp laser Doppler interferometer. *Lasers in Surgery and Medicine*, vol. 13, no. 4, pp. 447–452, 1993. DOI: 10.1002/lsm.1900130409.

[28] S. R. Waltman, H. E. Kaufman. A new objective slit lamp fluorophotometer. *Investigative Ophthalmology*, vol. 9, no. 4, pp. 247–249, 1970.

[29] M. A. Vivino, A. Mahurkar, B. Trus, M. L. Lopez, M. Datiles. Quantitative analysis of retroillumination images. *Eye*, vol. 9, no. 1, pp. 77–84, 1995. DOI: 10.1038/eye.1995.12.

[30] A. Gershenzon, L. D. Robman. New software for lens retro-illumination digital image analysis. *Australian and New Zealand Journal of Ophthalmology*, vol. 27, no. 3–4, pp. 170–172, 1999. DOI: 10.1046/j.1440-1606.1999.00201.x.

[31] C. C. Huang, R. M. Chen, P. H. Tsui, Q. F. Zhou, M. S. Humayun, K. K. Shung. Measurements of attenuation coefficient for evaluating the hardness of a cataract lens by a high-frequency ultrasonic needle transducer. *Physics in Medicine & Biology*, vol. 54, no. 19, pp. 5981–5994, 2009. DOI: 10.1088/0031-9155/54/19/021.

[32] C. C. Huang, H. Ameri, C. DeBoer, A. P. Rowley, X. C. Xu, L. Sun, S. H. Wang, M. S. Humayun, K. K. Shung. Evaluation of lens hardness in cataract surgery using high-frequency ultrasonic parameters *in vitro*. *Ultrasound in Medicine & Biology*, vol. 33, no. 10, pp. 1609–1616, 2007. DOI: 10.1016/j.ultrasmedbio.2007.05.002.

[33] P. H. Tsui, C. C. Chang. Imaging local scatterer concentrations by the Nakagami statistical model. *Ultrasound in Medicine & Biology*, vol. 33, no. 4, pp. 608–619, 2007. DOI: 10.1016/j.ultrasmedbio.2006.10.005.

[34] P. H. Tsui, C. K. Yeh, Y. Y. Liao, C. C. Chang, W. H. Kuo, K. J. Chang, C. N. Chen. Ultrasonic Nakagami imaging: A strategy to visualize the scatterer properties of benign and malignant breast tumors. *Ultrasound in Medicine & Biology*, vol. 36, no. 2, pp. 209–217, 2010. DOI: 10.1016/j.ultrasmedbio.2009.10.006.

[35] P. H. Tsui, C. C. Huang, C. C. Chang, S. H. Wang, K. K. Shung. Feasibility study of using high-frequency ultrasonic Nakagami imaging for characterizing the cataract lens *in vitro*. *Physics in Medicine & Biology*, vol. 52, no. 21, pp. 6413–6425, 2007. DOI: 10.1088/0031-9155/52/21/005.

[36] P. H. Tsui, M. C. Ho, D. I. Tai, Y. H. Lin, C. Y. Wang, H. Y. Ma. Acoustic structure quantification by using ultrasound Nakagami imaging for assessing liver fibrosis. *Scientific Reports*, vol. 6, Article number 33075, 2016. DOI: 10.1038/srep33075.

[37] A. Plesch, U. Klingbeil, J. Bille. Digital laser scanning fundus camera. *Applied Optics*, vol. 26, no. 8, pp. 1480–1486, 1987. DOI: 10.1364/AO.26.001480.

[38] O. Pomerantzeff, R. H. Webb, F. C. Delori. Image formation in fundus cameras. *Investigative Ophthalmology & Visual Science*, vol. 18, no. 6, pp. 630–637, 1979.

[39] L. C. Cao, H. Q. Li, Y. J. Zhang, L. Zhang, L. Xu. Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. *Information Fusion*, vol. 53, pp. 196–208, 2020. DOI: 10.1016/j.inffus.2019.06.022.

[40] M. Ang, M. Baskaran, R. M. Werkmeister, J. Chua, D. Schmidl, V. A. dos Santos, G. Garhöfer, J. S. Mehta, L. Schmetterer. Anterior segment optical coherence tomography. *Progress in Retinal and Eye Research*, vol. 66, pp. 132–156, 2018. DOI: 10.1016/j.preteyeres.2018.04.002.

[41] R. M. Werkmeister, S. Sapeta, D. Schmidl, G. Garhöfer, G. Schmidinger, V. A. dos Santos, G. C. Aschinger, I. Baumgartner, N. Pircher, F. Schwarzhans, A. Pantalon, H. Dua, L. Schmetterer. Ultrahigh-resolution OCT imaging of the human cornea. *Biomedical Optics Express*, vol. 8, no. 2, pp. 1221–1239, 2017. DOI: 10.1364/BOE.8.001221.

[42] N. Hirnschall, S. Amir-Asgari, S. Maedel, O. Findl. Predicting the postoperative intraocular lens position using continuous intraoperative optical coherence tomography measurements. *Investigative Ophthalmology & Visual Science*, vol. 54, no. 8, pp. 5196–5203, 2013. DOI: 10.1167/iovs.13-11991.

[43] N. Yamazaki, A. Kobayashi, H. Yokogawa, Y. Ishibashi, Y. Oikawa, M. Tokoro, K. Sugiyama. In vivo imaging of radial keratoneuritis in patients with *acanthamoeba* keratitis by anterior-segment optical coherence tomography. *Ophthalmology*, vol. 121, no. 11, pp. 2153–2158, 2014. DOI: 10.1016/j.ophtha.2014.04.043.

[44] N. Hirnschall, T. Buehren, F. Bajramovic, M. Trost, T. Teuber, O. Findl. Prediction of postoperative intraocular lens tilt using swept-source optical coherence tomography. *Journal of Cataract & Refractive Surgery*, vol. 43, no. 6, pp. 732–736, 2017. DOI: 10.1016/j.jcrs.2017.01.026.

[45] I. Grulkowski, S. Manzanera, L. Cwiklinski, J. Mompeán, A. de Castro, J. M. Marin, P. Artal. Volumetric macro- and micro-scale assessment of crystalline lens opacities in cataract patients using long-depth-range swept source optical coherence tomography. *Biomedical Optics Express*, vol. 9, no. 8, pp. 3821–3833, 2018. DOI: 10.1364/BOE.9.003821.

[46] D. Pawliczek, C. Dalke, H. Fuchs, V. Gailus-Durner, M. H. de Angelis, J. Graw, O. V. Amarie. Spectral domain-optical coherence tomography (SD-OCT) as a monitoring tool for alterations in mouse lenses. *Experimental Eye Research*, vol. 190, Article number 107871, 2020. DOI: 10.1016/j.exer.2019.107871.

[47] L. T. Chylack Jr, J. K. Wolfe, D. M. Singer, M. C. Leske, M. A. Bullimore, I. L. Bailey, J. Friend, D. McCarthy, S. Y. Wu. The lens opacities classification system III. *Archives of Ophthalmology*, vol. 111, no. 6, pp. 831–836, 1993. DOI: 10.1001/archopht.1993.01090060119035.

[48] L. T. Chylack Jr, M. C. Leske, R. Sperduto, P. Khu, D. McCarthy. Lens opacities classification system. *Archives of Ophthalmology*, vol. 106, no. 3, pp. 330–334, 1988. DOI: 10.1001/archopht.1988.01060130356020.

[49] L. T. Chylack, M. C. Leske, D. McCarthy, P. Khu, T. Kashiwagi, R. Sperduto. Lens opacities classification system II (LOCS II). *Archives of Ophthalmology*, vol. 107, no. 7, pp. 991–997, 1989. DOI: 10.1001/archopht.1989.01070020053028.

[50] B. E. K. Klein, R. Klein, K. L. P. Linton, Y. L. Magli, M. W. Neider. Assessment of cataracts from photographs in the beaver dam eye study. *Ophthalmology*, vol. 97, no. 11, pp. 1428–1433,1990.DOI:10.1016/S0161-6420(90)32391-6.

[51] The Age-Related Eye Disease Study Research Group. The age-related eye disease study (AREDS) system for classifying cataracts from photographs: AREDS Report No. 4. *American Journal of Ophthalmology*, vol. 131, no. 2, pp. 167–175, 2001. DOI: 10.1016/S0002-9394(00)00732-7.

[52] S. H. Fan, C. R. Dyer, L. Hubbard, B. Klein. An automatic system for classification of nuclear sclerosis from slit-lamp photographs. In *Proceedings of the 6th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, Montreal, Canada, pp. 592-601, 2003. DOI: 10.1007/978-3-540-39899-8_73.

[53] J. M. Sparrow, A. J. Bron, N. A. P. Brown, W. Ayliffe, A. R. Hill. The oxford clinical cataract classification and grading system. *International Ophthalmology*, vol. 9, no. 4, pp. 207–225, 1986. DOI: 10.1007/BF00137534.

[54] A. B. Hall, J. R. Thompson, J. S. Deane, A. R. Rosenthal. LOCS III versus the oxford clinical cataract classification and grading system for the assessment of nuclear, cortical and posterior subcapsular cataract. *Ophthalmic Epidemiology*, vol. 4, no. 4, pp. 179–194, 1997. DOI: 10.3109/09286589709059192.

[55] S. K. West, F. Rosenthal, H. S. Newland, H. R. Taylor. Use of photographic techniques to grade nuclear cataracts. *Investigative Ophthalmology* & *Visual Science*, vol. 29, no. 1, pp. 73–77, 1988.

[56] B. Thylefors, L. T. Chylack Jr, K. Konyama, K. Sasaki, R. Sperduto, H. R. Taylor, S. West. A simplified cataract grading system the WHO Cataract Grading Group. *Ophthalmic Epidemiology*, vol. 9, no. 2, pp. 83–95, 2002. DOI: 10.1076/opep.9.2.83.1523.

[57] WHO Programme for the Prevention of Blindness, WHO Cataract Grading Group. A Simplified Cataract Grading System, WHO/PBL/01.81. World Health Organization, Switzerland, 2001.

[58] L. Xu, C. W. Yang, H. Yang, S. Wang, Y. Y. Shi, X. D. Song. The study of predicting the visual acuity after phacoemulsification according to the blur level of fundus photography. *Ophthalmology in China*, vol. 19, no. 2, pp. 81–83, 2010. (in Chinese)

[59] J. Liu, D. W. K. Wong, Z. Zhang, B. H. Lee, X. T. Gao, F. S. Yin, J. L. Zhang, M. T. Htoo. Integrating research, clinical practice and translation: The Singapore experience. In *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Osaka, Japan, pp. 7148–7151, 2013. DOI: 10.1109/EMBC.2013.6611206.

[60] J. W. Jiang, S. T. Lei, M. M. Zhu, R. Y. Li, J. Y. Yue, J. J. Chen, Z. W. Li, J. M. Gong, D. R. Lin, X. H. Wu, Z. L. Lin, H. T. Lin. Improving the generalizability of infantile cataracts detection via deep learning-based lens partition strategy and multicenter datasets. *Frontiers in Medicine*, vol. 8, Article number 470, 2021. DOI: 10.3389/FMED.2021.664023.

[61] J. Cuadros, G. Bresnick. EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. *Journal of Diabetes Science and Technology*, vol. 3, no. 3, pp. 509–516, 2009. DOI: 10.1177/193229680900300315.

[62] T. Pratap, P. Kokil. Computer-aided diagnosis of cataract using deep transfer learning. *Biomedical Signal Processing and Control*, vol. 53, Article number 101533, 2019. DOI: 10.1016/j.bspc.2019.04.010.

[63] A. D. Hoover, V. Kouznetsova, M. Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000. DOI: 10.1109/42.845178.

[64] T. Kauppi, V. Kalesnykiene, J. K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, J. Pietilä. DIARET-DB0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms. Lappeenranta University of Technology, Finland, 2006. [Online], Available: https://www.it.lut.fi/project/imageret/diaretdb0/.

[65] E. Decenciére, G. Cazuguel, X. Zhang, G. Thibault, J. C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ, A. Chabouis. Teleophta: Machine learning and image processing methods for teleophthalmology. *IRBM*, vol. 34, no. 2, pp. 196–203, 2013. DOI: 10.1016/j.irbm.2013.01.010.

[66] E. Decenciére, X. W. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, J. C. Klein. Feedback on a publicly distributed image database: The messidor database. *Image Analysis* & *Stereology*, vol. 33, no. 3, pp. 231–234, 2014. DOI: 10.5566/ias.1155.

[67] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, B. Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004. DOI: 10.1109/TMI.2004.825627.

[68] C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, A. A. Argyros. Fire: Fundus image registration dataset. *Journal for Modeling in Ophthalmology*, vol. 1, no. 4, pp. 16–28, 2017.

[69] E. J. Carmona, M. Rincón, J. García-Feijoó, J. M. Martínez-de-la-Casa. Identification of the optic nerve head with genetic algorithms. *Artificial Intelligence in Medicine*, vol. 43, no. 3, pp. 243–259, 2008. DOI: 10.1016/j.artmed.2008.04.005.

[70] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, F. Meriaudeau. Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. *Data*, vol. 3, no. 3, Article number 25, 2018. DOI: 10.3390/data3030025.

[71] T. Mahmudi, R. Kafieh, H. Rabbani, A. M. Dehnavi, M. Akhlagi. Comparison of macular OCTs in right and left eyes of normal people. In *Proceedings of SPIE 9038, Medical Imaging 2014: Biomedical Applications in Molecular, Structural, and Functional Imaging*, SPIE, San Diego, USA, pp. 472–477, 2014. DOI: 10.1117/12.2044046.

[72] H. Q. Li, J. H. Lim, J. Liu, P. Mitchell, A. G. Tan, J. J. Wang, T. Y. Wong. A computer-aided diagnosis system of nuclear cataract. *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1690–1698, 2010. DOI: 10.1109/TBME.2010.2041454.

[73] W. Huang, K. L. Chan, H. Q. Li, J. H. Lim, J. Liu, T. Y. Wong. A computer assisted method for nuclear cataract grading from slit-lamp images using ranking. *IEEE Transactions on Medical Imaging*, vol. 30, no. 1, pp. 94–107, 2011. DOI: 10.1109/TMI.2010.2062197.

[74] H. Q. Li, J. H. Lim, J. Liu, T. Y. Wong. Towards automatic grading of nuclear cataract. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Lyon, France, pp. 4961–4964, 2007. DOI: 10.1109/IEMBS.2007.4353454.

[75] M. Caixinha, J. Amaro, M. Santos, F, Perdigão, M. Gomes, J. Santos. *In-vivo* automatic nuclear cataract detection and classification in an animal model by ultrasounds. *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2326–2335, 2016. DOI: 10.1109/TBME.2016.2527787.

[76] J. W. Jiang, X. Y. Liu, K. Zhang, E. P. Long, L. M. Wang, W. T. Li, L. Liu, S. Wang, M. M. Zhu, J. T. Cui, Z. Z. Liu, Z. L. Lin, X. Y. Li, J. J. Chen, Q. Z. Cao, J. Li, X. H. Wu, D. N. Wang, J. H. Wang, H. T. Lin. Automatic diagnosis of imbalanced ophthalmic images using a cost-sensitive deep convolutional neural network. *Bio-Medical Engineering OnLine*, vol. 16, no. 1, Article number 132, 2017. DOI: 10.1186/s12938-017-0420-1.

[77] L. M. Wang, K. Zhang, X. Y. Liu, E. P. Long, J. W. Jiang, Y. Y. An, J. Zhang, Z. Z. Liu, Z. L. Lin, X. Y. Li, J. J. Chen, Q. Z. Cao, J. Li, X. H. Wu, D. N. Wang, W. T. Li, H. T. Lin. Comparative analysis of image classification methods for automatic diagnosis of ophthalmic images. *Scientific Reports*, vol. 7, Article number 41545, 2017. DOI: 10.1038/srep41545.

[78] J. Cheng. Sparse range-constrained learning and its application for medical image grading. *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2729–2738, 2018. DOI: 10.1109/TMI.2018.2851607.

[79] A. B. Jagadale, S. S. Sonavane, D. V. Jadav. Computer aided system for early detection of nuclear cataract using circle Hough transform. In *Proceedings of the 3rd International Conference on Trends in Electronics and Informatics*, IEEE, Tirunelveli, India, pp. 1009–1012, 2019. DOI: 10.1109/ICOEI.2019.8862595.

[80] K. Zhang, X. Y. Liu, J. W. Jiang, W. T. Li, S. Wang, L. Liu, X. J. Zhou, L. M. Wang. Prediction of postoperative complications of pediatric cataract patients using data mining. *Journal of Translational Medicine*, vol. 17, no. 1, Article number 2, 2019. DOI: 10.1186/s12967-018-1758-2.

[81] H. Q. Li, L. Ko, J. H. Lim, J. Liu, D. W. K. Wong, T. Y. Wong, Y. Sun. Automatic opacity detection in retro-illumination images for cortical cataract diagnosis. In *Proceedings of IEEE International Conference on Multimedia and Expo*, IEEE, Hannover, Germany, pp. 553–556, 2008. DOI: 10.1109/ICME.2008.4607494.

[82] Y. C. Chow, X. T. Gao, H. Q. Li, J. H. Lim, Y. Sun, T. Y. Wong. Automatic detection of cortical and PSC cataracts using texture and intensity analysis on retro-illumination lens images. In *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Boston, USA, pp. 5044–5047, 2011. DOI: 10.1109/IEMBS.2011.6091249.

[83] W. J. Zhang, H. Q. Li. Lens opacity detection for serious posterior subcapsular cataract. *Medical & Biological Engineering & Computing*, vol. 55, no. 5, pp. 769–779, 2017. DOI: 10.1007/s11517-016-1554-1.

[84] H. Q. Li, L. Ko, J. H. Lim, J. Liu, D. W. K. Wong, T. Y. Wong. Image based diagnosis of cortical cataract. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Vancouver, Canada, pp. 3904–3907, 2008. DOI: 10.1109/IEMBS.2008.4650063.

[85] M. Caixinha, E. Velte, M. Santos, F. Perdigão, J. Amaro, M. Gomes, J. Santos. Automatic cataract classification based on ultrasound technique using machine learning: A comparative study. *Physics Procedia*, vol. 70, pp. 1221–1224, 2015. DOI: 10.1016/j.phpro.2015.08.263.

[86] M. Caixinha, E. Velte, M. Santos, J. B. Santos. New approach for objective cataract classification based on ultrasound techniques using multiclass SVM classifiers. In *Proceedings of IEEE International Ultrasonics Symposium*, IEEE, Chicago, USA, pp. 2402–2405, 2014. DOI: 10.1109/ULTSYM.2014.0599.

[87] M. Caixinha, D. A. Jesus, E. Velte, M. J. Santos, J. B. Santos. Using ultrasound backscattering signals and nakagami statistical distribution to assess regional cataract hardness. *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 12, pp. 2921–2929, 2014. DOI: 10.1109/TBME.2014.2335739.

[88] D. Jesus, E. Velte, M. Caixinha, M. Santos, J. Santos. Using of the ultrasound frequency dependent attenuation and Nakagami distribution for cataract evaluation. In *Proceedings of the 3rd IEEE Portuguese Meeting in Bioengineering*, IEEE, Braga, Portugal, 2013. DOI: 10.1109/ENBENG.2013.6518388.

[89] Y. N. Fuadah, A. W. Setiawan, T. L. R. Mengko, Budiman. Mobile cataract detection using optimal combination of statistical texture analysis. In *Proceedings of the 4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, IEEE, Bandung, Indonesia, pp. 232–236, 2015. DOI: 10.1109/ICICI-BME.2015.7401368.

[90] L. Y. Guo, J. J. Yang, L. H. Peng, J. Q. Li, Q. F. Liang. A computer-aided healthcare system for cataract classification and grading based on fundus image analysis. *Computers in Industry*, vol. 69, pp. 72–80, 2015. DOI: 10.1016/j.compind.2014.09.005.

[91] W. M. Fan, R. F. Shen, Q. Q. Zhang, J. J. Yang, J. Q. Li. Principal component analysis based cataract grading and classification. In *Proceedings of the 17th International Conference on E-health Networking, Application & Services*, IEEE, Boston, USA, pp. 459–462, 2015. DOI: 10.1109/HealthCom.2015.7454545.

[92] H. Y. Zhang, K. Niu, Y. M. Xiong, W. H. Yang, Z. Q. He, H. X. Song. Automatic cataract grading methods based on deep learning. *Computer Methods and Programs in Biomedicine*, vol. 182, Article number 104978, 2019. DOI: 10.1016/j.cmpb.2019.07.006.

[93] C. S. Huo, F. Akhtar, P. Z. Li. A novel grading method of cataract based on AWM. In *Proceedings of the IEEE 43rd Annual Computer Software and Applications Conference*, IEEE, Milwaukee, USA, pp. 368–373, 2019. DOI: 10.1109/COMPSAC.2019.10234.

[94] W. A. Song, P. Wang, X. D. Zhang, Q. Wang. Semi-supervised learning based on cataract classification and grading. In *Proceedings of the 40th IEEE Annual Com-*

puter Software and Applications Conference, IEEE, Atlanta, USA, pp. 641–646, 2016. DOI: 10.1109/COMPSAC.2016.227.

[95] W. A. Song, Y. Cao, Z. Q. Qiao, Q. Wang, J. J. Yang. An improved semi-supervised learning method on cataract fundus image classification. In Proceedings of the 43rd IEEE Annual Computer Software and Applications Conference, IEEE, Milwaukee, USA, pp. 362–367, 2019. DOI: 10.1109/COMPSAC.2019.10233.

[96] X. Q. Zhang, J. S. Fang, Z. J. Xiao, R. Higashita, W. Chen, J. Yuan, J. Liu. Classification algorithm of nuclear cataract based on anterior segment coherence tomography image. Computer Science, vol. 49, no. 3, pp. 204–210, 2022. DOI: 10.11896/jsjkx.201100085. (in Chinese)

[97] Y. Zhou, G. Q. Li, H. Q. Li. Automatic cataract classification using deep neural network with discrete state transition. IEEE Transactions on Medical Imaging, vol. 39, no. 2, pp. 436–446, 2020. DOI: 10.1109/TMI.2019.2928229.

[98] X. T. Gao, S. Lin, T. Y. Wong. Automatic feature learning to grade nuclear cataracts based on deep learning. IEEE Transactions on Biomedical Engineering, vol. 62, no. 11, pp. 2693–2701, 2015. DOI: 10.1109/TBME.2015.2444389.

[99] X. Y. Liu, J. W. Jiang, K. Zhang, E. P. Long, J. T. Cui, M. M. Zhu, Y. Y. An, J. Zhang, Z. Z. Liu, Z. L. Lin, X. Y. Li, J. J. Chen, Q. Z. Cao, J. Li, X. H. Wu, D. N. Wang, H. T. Lin. Localization and diagnosis framework for pediatric cataracts based on slit-lamp images using deep features of a convolutional neural network. PLoS One, vol. 12, no. 3, Article number 0168606, 2017. DOI: 10.1371/journal.pone.0168606.

[100] K. Zhang, X. Y. Liu, F. Liu, L. He, L. Zhang, Y. H. Yang, W. T. Li, S. Wang, L. Liu, Z. Z. Liu, X. H. Wu, H. T. Lin. An interpretable and expandable deep learning diagnostic system for multiple ocular diseases: Qualitative study. Journal of Medical Internet Research, vol. 20, no. 11, Article number e11144, 2018. DOI: 10.2196/11144.

[101] J. W. Jiang, X. Y. Liu, L. Liu, S. Wang, E. P. Long, H. Q. Yang, F. Q. Yuan, D. Y. Yu, K. Zhang, L. M. Wang, Z. Z. Liu, D. N. Wang, C. Z. Xi, Z. L. Lin, X. H. Wu, J. T. Cui, M. M. Zhu, H. T. Lin. Predicting the progression of ophthalmic disease based on slit-lamp images using a deep temporal sequence network. PLoS One, vol. 13, no. 7, Article number 0201142, 2018. DOI: 10.1371/journal.pone.0201142.

[102] C. X. Xu, X. J. Zhu, W. W. He, Y. Lu, X. X. He, Z. J. Shang, J. Wu, K. K. Zhang, Y. L. Zhang, X. F. Rong, Z. N. Zhao, L. Cai, A. Y. Ding, X. R. Li. Fully deep learning for slit-lamp photo based nuclear cataract grading. In Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention, Springer, Shenzhen, China, pp. 513–521, 2019. DOI: 10.1007/978-3-030-32251-9_56.

[103] D. S. J. Ting, M. Ang, J. S. Mehta, D. S. W. Ting. Artificial intelligence-assisted telemedicine platform for cataract screening and management: A potential model of care for global eye health. British Journal of Ophthalmology, vol. 103, no. 11, pp. 1537–1538, 2019. DOI: 10.1136/bjophthalmol-2019-315025.

[104] D. Kim, T. J. Jun, Y. Eom, C. Kim, D. Kim. Tourna-

ment based ranking CNN for the cataract grading. In Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, Berlin, Germany, pp. 1630–1636, 2019. DOI: 10.1109/EMBC.2019.8856636.

[105] X. H. Wu, Y. L. Huang, Z. Z. Liu, W. Y. Lai, E. P. Long, K. Zhang, J. W. Jiang, D. R. Lin, K. X. Chen, T. Y. Yu, D. X. Wu, C. Li, Y. Y. Chen, M. J. Zou, C. Chen, Y. Zhu, C. Guo, X. Y. Zhang, R. X. Wang, Y. H. Yang, Y. F. Xiang, L. J. Chen, C. X. Liu, J. H. Xiong, Z. Y. Ge, D. D. Wang, G. H. Xu, S. L. Du, C. Xiao, J. H. Wu, K. Zhu, D. Y. Nie, F. Xu, J. Lv, W. R. Chen, Y. Z. Liu, H. T. Lin. Universal artificial intelligence platform for collaborative management of cataracts. British Journal of Ophthalmology, vol. 103, no. 11, pp. 1553–1560, 2019. DOI: 10.1136/bjophthalmol-2019-314729.

[106] S. M. Hu, X. T. Wang, H. Wu, X. Z. Luan, P. Qi, Y. Lin, X. D. He, W. He. Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images. IEEE Access, vol. 8, pp. 174169–174178, 2020. DOI: 10.1109/ACCESS.2020.3025346.

[107] J. W. Jiang, L. M. Wang, H. R. Fu, E. P. Long, Y. B. Sun, R. Y. Li, Z. W. Li, M. M. Zhu, Z. Z. Liu, J. J. Chen, Z. L. Lin, X. H. Wu, D. N. Wang, X. Y. Liu, H. T. Lin. Automatic classification of heterogeneous slit-illumination images using an ensemble of cost-sensitive convolutional neural networks. Annals of Translational Medicine, vol. 9, no. 7, Article number 550, 2021. DOI: 10.21037/atm-20-6635.

[108] S. M. Hu, X. Z. Luan, H. Wu, X. T. Wang, C. H. Yan, J. Y. Wang, G. T. Liu, W. He. ACCV: Automatic classification algorithm of cataract video based on deep learning. BioMedical Engineering OnLine, vol. 20, no. 1, Article number 78, 2021. DOI: 10.1186/s12938-021-00906-3.

[109] H. R. M. Tawfik, R. A. K. Birry, A. A. Saad. Early recognition and grading of cataract using a combined log Gabor/discrete wavelet transform with ANN and SVM. International Journal of Computer and Information Engineering, vol. 12, no. 12, pp. 1038–1043, 2018. DOI: 10.5281/zenodo.2022731.

[110] X. F. Zhang, J. C. Lv, H. Zheng, Y. S. Sang. Attention-based multi-model ensemble for automatic cataract detection in B-scan eye ultrasound images. In Proceedings of International Joint Conference on Neural Networks, IEEE, Glasgow, UK, 2020. DOI: 10.1109/IJCNN48605.2020.9207696.

[111] H. J. Wu, J. C. Lv, J. Wang. Automatic cataract detection with multi-task learning. In Proceedings of International Joint Conference on Neural Networks, IEEE, Shenzhen, China, 2021. DOI: 10.1109/IJCNN52387.2021.9533424.

[112] L. L. Zhang, J. Q. Li, I. Zhang, H. Han, B. Liu, J. J. Yang, Q. Wang. Automatic cataract detection and grading using deep convolutional neural network. In Proceedings of the 14th IEEE International Conference on Networking, Sensing and Control, IEEE, Calabria, Italy, pp. 60–65, 2017. DOI: 10.1109/ICNSC.2017.8000068.

[113] Y. Y. Dong, Q. Y. Zhang, Z. Q. Qiao, J. J. Yang. Classification of cataract fundus image based on deep learning. In Proceedings of IEEE International Conference on Imaging Systems and Techniques, IEEE, Beijing, China, 2017. DOI: 10.1109/IST.2017.8261463.

[114] X. Xu, L. L. Zhang, J. Q. Li, Y. Guan, L. Zhang. A hybrid global-local representation CNN model for automatic cataract grading. *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 556–567, 2020. DOI: 10.1109/JBHI.2019.2914690.

[115] T. Pratap, P. Kokil. Efficient network selection for computer-aided cataract diagnosis under noisy environment. *Computer Methods and Programs in Biomedicine*, vol. 200, Article number 105927, 2021. DOI: 10.1016/j.cmpb.2021.105927.

[116] A. Imran, J. Q. Li, Y. Pei, F. Akhtar, T. Mahmood, L. Zhang. Fundus image-based cataract classification using a hybrid convolutional and recurrent neural network. *The Visual Computer*, vol. 37, no. 8, pp. 2407–2417, 2021. DOI: 10.1007/s00371-020-01994-3.

[117] M. S. Junayed, M. B. Islam, A. Sadeghzadeh, S. Rahman. CataractNet: An automated cataract detection system using deep learning for fundus images. *IEEE Access*, vol. 9, pp. 128799–128808, 2021. DOI: 10.1109/ACCESS.2021.3112938.

[118] Y. C. Tham, J. H. L. Goh, A. Anees, X. F. Lei, T. H. Rim, M. L. Chee, Y. X. Wang, J. B. Jonas, S. Thakur, Z. L. Teo, N. Cheung, H. Hamzah, G. S. W. Tan, R. Husain, C. Sabanayagam, J. J. Wang, Q. Y. Chen, Z. Y. Lu, T. D. Keenan, E. Y. Chew, A. G. Tan, P. Mitchell, R. S. M. Goh, X. X. Xu, Y. Liu, T. Y. Wong, C. Y. Cheng. Detecting visually significant cataract using retinal photograph-based deep learning. *Nature Aging*, vol. 2, no. 3, pp. 264–271, 2022. DOI: 10.1038/s43587-022-00171-6.

[119] X. Q. Zhang, Z. J. Xiao, R. Higashita, W. Chen, J. Yuan, J. S. Fang, Y. Hu, J. Liu. A novel deep learning method for nuclear cataract classification based on anterior segment optical coherence tomography images. In *Proceedings of IEEE International Conference on Systems*, IEEE, Toronto, Canada, pp. 662–668, 2020. DOI: 10.1109/SMC42975.2020.9283218.

[120] Z. Xiao, X. Zhang, R. Higashita, W. Chen, J. Yuan, J. Liu. A 3D CNN-based multi-task learning for cataract screening and left and right eye classification on 3D AS-OCT images. In *Proceedings of the 3rd International Conference on Intelligent Medicine and Health*, ACM, Macau, China, pp. 1–7, 2021. DOI: 10.1145/3484377.3484378.

[121] Z. J. Xiao, X. Q. Zhang, R. Higashita, Y. Hu, J. Yuan, W. Chen, J. Liu. Gated channel attention network for cataract classification on AS-OCT image. In *Proceedings of the 28th International Conference on Neural Information Processing*, Springer, Sanur, Indonesia, pp. 357–368, 2021. DOI: 10.1007/978-3-030-92238-2_30.

[122] X. Q. Zhang, Z. J. Xiao, X. L. Li, X. Wu, H. X. Sun, J. Yuan, R. Higashita, J. Liu. Mixed pyramid attention network for nuclear cataract classification based on anterior segment OCT images. *Health Information Science and Systems*, vol. 10, no. 1, Article number 3, 2022. DOI: 10.1007/s13755-022-00170-2.

[123] I. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. H. Zhai, T. Unterthiner, J. Yung, A. Steiner, D. Keysers, J. Uszkoreit, M. Lucic, A. Dosovitskiy. MLP-mixer: An all-MLP architecture for vision. In *Proceedings of the 35th Conference on Neural Information Processing Systems*, 2021.

[124] D. Z. Lian, Z. H. Yu, X. Sun, S. H. Gao. AS-MLP: An axial shifted MLP architecture for vision. [Online], Available: https://arxiv.org/abs/2107.08391, 2021.

[125] Y. Mansour, K. Lin, R. Heckel. Image-to-image MLP-mixer for image reconstruction. [Online], Available: https://arxiv.org/abs/2202.02018, 2021.

[126] H. Touvron, P. Bojanowski, M. Caron, M. Cord, A. El-Nouby, E. Grave, G. Izacard, A. Joulin, G. Synnaeve, J. Verbeek, H. Jégou. ResMLP: Feedforward networks for image classification with data-efficient training. [Online], Available: https://arxiv.org/abs/2105.03404, 2021.

[127] H. J. Zheng, P. C. He, W. Z. Chen, M. Y. Zhou. Mixing and shifting: Exploiting global and local dependencies in vision MLPs. [Online], Available: https://arxiv.org/abs/2202.06510, 2022.

[128] C. Szegedy, W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich. Going deeper with convolutions. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Boston, USA, pp. 1–9, 2015. DOI: 10.1109/CVPR.2015.7298594.

[129] A. Krizhevsky, I. Sutskever, G. E. Hinton. ImageNet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Curran Associates Inc, Lake Tahoe, USA, pp. 1097–1105, 2012.

[130] H. Chen, Q. Dou, D. Ni, J. Z. Cheng, J. Qin, S. L. Li, P. A. Heng. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Munich, Germany, pp. 507–514, 2015. DOI: 10.1007/978-3-319-24553-9_62.

[131] H. Z. Fu, Y. W. Xu, S. Lin, D. W. K. Wong, J. Liu. DeepVessel: Retinal vessel segmentation via deep learning and conditional random field. In *Proceedings of the 19th International Conference on Medical Image Computing and Computer-assisted Intervention*, Springer, Athens, Greece, pp. 132–139, 2016. DOI: 10.1007/978-3-319-46723-8_16.

[132] M. D. Abràmoff, Y. Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, M. Niemeijer. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigative Ophthalmology & Visual Science*, vol. 57, no. 13, pp. 5200–5206, 2016. DOI: 10.1167/iovs.16-19964.

[133] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, C. I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. DOI: 10.1016/j.media.2017.07.005.

[134] A. Mansoor, J. J. Cerrolaza, R. Idrees, E. Biggs, M. A. Alsharid, R. A. Avery, M. G. Linguraru. Deep learning guided partitioned shape model for anterior visual pathway segmentation. *IEEE Transactions on Medical Imaging*, vol. 35, no. 8, pp. 1856–1865, 2016. DOI: 10.1109/TMI.2016.2535222.

[135] S. N. Xie, R. Girshick, P. Dollár, Z. W. Tu, K. M. He. Aggregated residual transformations for deep neural networks. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Honolulu, USA, pp. 1492–1500, 2017. DOI: 10.1109/CVPR.2017.634.

[136] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Las Vegas, USA, pp. 2818–2826, 2016. DOI: 10.1109/CVPR.2016. 308.

[137] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Honolulu, USA, pp. 2261–2269, 2017. DOI: 10.1109/CVPR.2017.243.

[138] X. Y. Zhang, X. Y. Zhou, M. X. Lin, J. Sun. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, USA, pp. 6848–6856, 2018. DOI: 10.1109/CVPR.2018.00716.

[139] P. S. Grewal, F. Oloumi, U. Rubin, M. T. S. Tennant. Deep learning in ophthalmology: A review. *Canadian Journal of Ophthalmology*, vol. 53, no. 4, pp. 309–313, 2018. DOI: 10.1016/j.jcjo.2018.04.019.

[140] L. Deng, P. Jiao, J. Pei, Z. Z. Wu, G. Q. Li. GXNOR-Net: Training deep neural networks with ternary weights and activations without full-precision memory under a unified discretization framework. *Neural Networks*, vol. 100, pp. 49–58, 2018. DOI: 10.1016/j.neunet.2018.01.010.

[141] P. S. Yin, M. K. Tan, H. Q. Min, Y. W. Xu, G. H. Xu, Q. Y. Wu, Y. F. Tong, H. Risa, J. Liu. Automatic segmentation of cortex and nucleus in anterior segment OCT images. In *Proceedings of the 1st International Workshop on Computational Pathology and Ophthalmic Medical Image Analysis*, Springer, Granada, Spain, pp. 269–276, 2018. DOI: 10.1007/978-3-030-00949-6_32.

[142] S. H. Zhang, Y. G. Yan, P. S. Yin, Z. Qiu, W. Zhao, G. P. Cao, W. Chen, J. Yuan, R. Higashita, Q. Y. Wu, M. K. Tan, J. Liu. Guided M-Net for high-resolution biomedical image segmentation with weak boundaries. In *Proceedings of the 6th International Workshop on Ophthalmic Medical Image Analysis*, Springer, Shenzhen, China, pp. 43–51, 2019. DOI: 10.1007/978-3-030-32956-3_6.

[143] G. P. Cao, W. Zhao, R. Higashita, J. Liu, W. Chen, J. Yuan, Y. B. Zhang, M. Yang. An efficient lens structures segmentation method on AS-OCT images. In *Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, IEEE, Montreal, Canada, pp. 1646–1649, 2020. DOI: 10.1109/EMBC44109.2020.9175944.

[144] Y. Hu, X. Q. Zhang, L. Xu, F. X. He, Z. Tian, W. She, W. Liu. Harmonic loss function for sensor-based human activity recognition based on LSTM recurrent neural networks. *IEEE Access*, vol. 8, pp. 135617–135627, 2020. DOI: 10.1109/ACCESS.2020.3003162.

[145] M. Choi, H. Kim, B. Han, N. Xu, K. M. Lee. Channel attention is all you need for video frame interpolation. *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 7, pp. 10663–10671, 2020. DOI: 10.1609/aaai.v34i07.6693.

[146] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, USA, pp. 6000–6010, 2017.

[147] Z. J. Zhang, H. Z. Fu, H. Dai, J. B. Shen, Y. W. Pang, L. Shao. ET-Net: A generic edge-attention guidance network for medical image segmentation. In *Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention*, Springer, Shenzhen, China, pp. 442–450, 2019. DOI: 10.1007/978-3-030-32239-7_49.

[148] M. H. Guo, T. X. Xu, J. J. Liu, Z. N. Liu, P. T. Jiang, T. J. Mu, S. H. Zhang, R. R. Martin, M. M. Cheng, S. M. Hu. Attention mechanisms in computer vision: A survey. *Computational Visual Media*, to be published. DOI: 10.1007/s41095-022-0271-y.

[149] X. L. Wang, R. Girshick, A. Gupta, K. M. He. Non-local neural networks. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, USA, pp. 7794–7803, 2018. DOI: 10.1109/CVPR.2018.00813.

[150] W. H. Yu, M. Luo, P. Zhou, C. Y. Si, Y. C. Zhou, X. C. Wang, J. S. Feng, S. C. Yan. MetaFormer is actually what you need for vision. [Online], Available: https://arxiv.org/abs/2111.11418, 2021.

[151] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. H. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of the 9th International Conference on Learning Representations*, 2021.

[152] K. L. He, C. Gan, Z. Y. Li, I. Rekik, Z. H. Yin, W. Ji, Y. Gao, Q. Wang, J. F. Zhang, D. G. Shen. Transformers in medical image analysis: A review. [Online], Available: https://arxiv.org/abs/2202.12165, 2022.

[153] A. Gulati, J. Qin, C. C. Chiu, N. Parmar, Y. Zhang, J. H. Yu, W. Han, S. B. Wang, Z. D. Zhang, Y. H. Wu, R. M. Pang. Conformer: Convolution-augmented transformer for speech recognition. In *Proceedings of the 21st Annual Conference of the International Speech Communication Association*, Shanghai, China, pp. 5036–5040, 2020.

[154] C. Shorten, T. M. Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of Big Data*, vol. 6, no. 1, Article number 60, 2019. DOI: 10.1186/s40537-019-0197-0.

[155] C. Y. I. Cheung, H. Q. Li, E. L. Lamoureux, P. Mitchell, J. J. Wang, A. G. Tan, L. K. Johari, J. Liu, J. H. Lim, T. Aung, T. Y. Wong. Validity of a new computer-aided diagnosis imaging program to quantify nuclear cataract from slit-lamp photographs. *Investigative Ophthalmology & Visual Science*, vol. 52, no. 3, pp. 1314–1319, 2011. DOI: 10.1167/iovs.10-5427.

[156] H. Q. Li, J. H. Lim, J. Liu, D. W. K. Wong, T. Y. Wong. Computer aided diagnosis of nuclear cataract. In *Proceedings of the 3rd IEEE Conference on Industrial Electronics and Applications*, IEEE, Singapore, pp. 1841–1844, 2008. DOI: 10.1109/ICIEA.2008.4582838.

[157] A. W. P. Foong, S. M. Saw, J. L. Loo, S. Shen, S. C. Loon, M. Rosman, T. Aung, D. T. H. Tan, E. S. Tai, T. Y. Wong. Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES). *Ophthalmic Epidemiology*, vol. 14, no. 1, pp. 25–35, 2007. DOI: 10.1080/09286580600878844.

[158] S. Resnikoff, V. C. Lansingh, L. Washburn, W. Felch, T.

M. Gauthier, H. R. Taylor, K. Eckert, D. Parke, P. Wiedemann. Estimated number of ophthalmologists worldwide (international council of ophthalmology update): Will we meet the needs? *British Journal of Ophthalmology*, vol. 104, no. 4, pp. 588–592, 2020. DOI: 10.1136/bjophthalmol-2019-314336.

[159] S. Resnikoff, W. Felch, T. M. Gauthier, B. Spivey. The number of ophthalmologists in practice and training worldwide: A growing gap despite more than 200 000 practitioners. *British Journal of Ophthalmology*, vol. 96, no. 6, pp. 783–787, 2012. DOI: 10.1136/bjophthalmol-2011-301378.

[160] S. Ravi, H. Larochelle. Optimization as a model for few-shot learning. In *Proceedings of the 5th International Conference on Learning Representations*, Toulon, France, 2016.

[161] N. Grira, M. Crucianu, N. Boujemaa. Unsupervised and semi-supervised clustering: A brief survey. *A review of Machine Learning Techniques for Processing Multimedia Content*, 2004.

[162] S. Fogel, H. Averbuch-Elor, D. Cohen-Or, J. Goldberger. Clustering-driven deep embedding with pairwise constraints. *IEEE Computer Graphics and Applications*, vol. 39, no. 4, pp. 16–27, 2019. DOI: 10.1109/MCG.2018.2881524.

[163] X. F. Guo, X. W. Liu, E. Zhu, J. P. Yin. Deep clustering with convolutional autoencoders. In *Proceedings of the 24th International Conference on Neural Information Processing*, Springer, Guangzhou, China, pp. 373–382, 2017. DOI: 10.1007/978-3-319-70096-0_39.

[164] Y. C. Hsu, Z. Kira. Neural network-based clustering using pairwise constraints. [Online], Available: https://arxiv.org/abs/1511.06321, 2015

[165] B. Ramamurthy, K. R. Chandran. Content based image retrieval for medical images using canny edge detection algorithm. *International Journal of Computer Applications*, vol. 17, no. 6, pp. 32–37, 2011. DOI: 10.5120/2222-2831.

[166] J. Sivakamasundari, G. Kavitha, V. Natarajan, S. Ramakrishnan. Proposal of a content based retinal image retrieval system using kirsch template based edge detection. In *Proceedings of International Conference on Informatics, Electronics & Vision*, IEEE, Dhaka, Bangladesh, 2014. DOI: 10.1109/ICIEV.2014.6850744.

[167] Y. F. Fathabad, M. A. Balafar. Content based image retrieval for medical images. *International Journal on Technical and Physical Problems of Engineering*, vol. 4, no. 12, pp. 177–182, 2012.

[168] J. S. Fang, Y. W. Xu, X. Q. Zhang, Y. Hu, J. Liu. Attention-based saliency hashing for ophthalmic image retrieval. In *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine*, IEEE, Seoul, Korea, pp. 990–995, 2020. DOI: 10.1109/BIBM49941.2020.9313536.

[169] D. R. Lin, J. J. Chen, Z. L. Lin, X. Y. Li, K. Zhang, X. H. Wu, Z. Z. Liu, J. L. Huang, J. Li, Y. Zhu, C. Chen, L. Q. Zhao, Y. F. Xiang, C. Guo, L. M. Wang, Y. Z. Liu, W. R. Chen, H. T. Lin. A practical model for the identification of congenital cataracts using machine learning. *EBioMedicine*, vol. 51, Article number 102621, 2020. DOI: 10.1016/j.ebiom.2019.102621.

[170] M. M. Yang, J. J. Yang, Q. Y. Zhang, Y. Niu, J. Q. Li. Classification of retinal image for automatic cataract detection. In *Proceedings of the 15th IEEE International Conference on e-Health Networking, Applications and Services*, IEEE, Lisbon, Portugal, pp. 674–679, 2013. DOI: 10.1109/HealthCom.2013.6720761.

[171] D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, A. Madabhushi. *Deep Learning In Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Cham, Germany: Springer, 2017. DOI: 10.1007/978-3-030-00889-5.

[172] J. Q. Ma, Z. Zhao, J. L. Chen, A. Li, L. C. Hong, E. H. Chi. SNR: Sub-network routing for flexible parameter sharing in multi-task learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Hawaii, USA, pp. 216–223, 2019. DOI: 10.1609/aaai.v33i01.3301216.

[173] Z. Guo, X. Li, H. Huang, N. Guo, Q. Z. Li. Deep learning-based image segmentation on multimodal medical imaging. *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 162–169, 2019. DOI: 10.1109/TRPMS.2018.2890359.

[174] B. Cheng, M. X. Liu, H. I. Suk, D. G. Shen, D. Q. Zhang, A. S. D. N. Initiative. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain Imaging and Behavior*, vol. 9, no. 4, pp. 913–926, 2015. DOI: 10.1007/s11682-015-9356-x.

[175] S. F. Mohammadi, M. Sabbaghi, H. Z-Mehrjardi, H. Hashemi, S. Alizadeh, M. Majdi, F. Taee. Using artificial intelligence to predict the risk for posterior capsule opacification after phacoemulsification. *Journal of Cataract & Refractive Surgery*, vol. 38, no. 3, pp. 403–408, 2012. DOI: 10.1016/j.jcrs.2011.09.036.

[176] C. Guo, G. Pleiss, Y. Sun, K. Q. Weinberger. On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, pp. 1321–1330, 2017.

[177] R. Müller, S. Kornblith, G. Hinton. When does label smoothing help? In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vancouver, Canada, Article number 422, 2019.

[178] B. L. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba. Learning deep features for discriminative localization. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Las Vegas, USA, pp. 2921–2929, 2016. DOI: 10.1109/CVPR.2016.319.

[179] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Venice, Italy, pp. 618–626, 2017. DOI: 10.1109/ICCV.2017.74.

[180] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi. A survey of methods for explaining black box models. *ACM Computing Surveys*, vol. 51, no. 5, Article number 93, 2018. DOI: 10.1145/3236009.

[181] J. S. Lu, C. M. Xiong, D. Parikh, R. Socher. Knowing when to look: Adaptive attention via a visual sentinel for image captioning. In *Proceedings of IEEE Conference on*

Computer Vision and Pattern Recognition, IEEE, Honolulu, USA, pp. 3242–3250, 2017. DOI: 10.1109/CVPR. 2017.345.

[182] S. Kornblith, M. Norouzi, H. Lee, G. E. Hinton. Similarity of neural network representations revisited. In *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, USA, pp. 3519–3529, 2019.

[183] R. Supriyanti, Y. Ramadhani. The achievement of various shapes of specular reflections for cataract screening system based on digital images. In *Proceedings of International Conference on Biomedical Engineering and Technology*, Kualalumpur, Malaysia, pp. 75–79, 2011.

[184] R. Supriyanti, Y. Ramadhani. Consideration of iris characteristic for improving cataract screening techniques based on digital image. In *Proceedings of the 2nd International Conference on Biomedical Engineering and Technology*, Hong Kong, China, pp. 130–133, 2012.

[185] J. Rana, S. M. Galib. Cataract detection using smartphone. In *Proceedings of the 3rd International Conference on Electrical Information and Communication Technology*, IEEE, Khulna, Bangladesh, 2017. DOI: 10.1109/EICT.2017.8275136.

[186] Y. Cheng, D. Wang, P. Zhou, T. Zhang. A survey of model compression and acceleration for deep neural networks. [Online], Available: https://arxiv.org/abs/1710.09282, 2017.

[187] Y. A. Jiang, S. Q. Wang, V. Valls, B. J. Ko, W. H. Lee, K. K. Leung, L. Tassiulas. Model pruning enables efficient federated learning on edge devices. In *Proceedings of the 34th Conference on Neural Information Processing Systems*, Vancouver, Canada, 2019.

[188] Y. H. He, X. Y. Zhang, J. Sun. Channel pruning for accelerating very deep neural networks. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Venice, Italy, pp. 1398–1406, 2017. DOI: 10.1109/ICCV.2017.155.

[189] H. Song, M. Kim, D. Park, Y. Shin, J. G. Lee. Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, to be published. DOI: 10.1109/TNNLS.2022.3152527.

[190] J. X. Zhong, N. N. Li, W. J. Kong, S. Liu, T. H. Li, G. Li. Graph convolutional label noise cleaner: Train a plug-and-play action classifier for anomaly detection. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Long Beach, USA, pp. 1237–1246, 2019. DOI: 10.1109/CVPR.2019.00133.
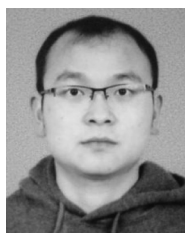
Xiao-Qing Zhang received the B. Sc. degree in water conservancy and hydropower engineering from South China Agricultural University, China in 2016, the M. Sc. degree in computer engineering from Zhengzhou University. Currently, he is a Ph. D. degree candidate with Department of Computer Science and Engineering, Southern University of Science and Technology, China.

His research interests include deep learning, interpretability, and medical image processing.

E-mail: 11930927@mail.sustech.edu.cn

ORCID iD: 0000-0003-1518-4781

Yan Hu received the B. Sc. and M. Sc. degrees in computer science from Northeast Normal University, China in 2008 and 2011, respectively, and the Ph. D. degree in optics from University of Tokyo, Japan in 2016. Currently, she is a research assistant professor in Department of Computer Science and Engineering, Southern University of Science and Technology, China.

Her research interests include surgical assistance, intraoperative navigation, and medical image processing.

E-mail: huy3@sustech.edu.cn

Zun-Jie Xiao received the B. Sc. degree in bioinformatics from Dalian University of Technology, China in 2019. Currently, he is a Ph. D. degree candidate with Department of Computer Science and Engineering, Southern University of Science and Technology, China.

His research interests include deep learning and medical image processing.

E-mail: 11930387@mail.sustech.edu.cn
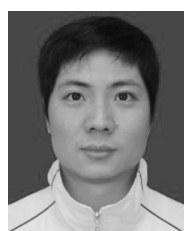
Jian-Sheng Fang received the B. Sc. degree in computer science and technology from Harbin Institute of Technology, China, the M. Sc. degree in computer application from Sun Yat-sen University, China. He is a Ph. D. degree candidate with Department of Computer Science and Engineering, Southern University of Science and Technology, China.

His research interests include deep learning and image retrieval.

E-mail: 11949039@mail.sustech.edu.cn

Risa Higashita received the Ph. D. degree in biomedical engineering from Nagoya University, Japan in 2004. Currently, she is a visiting professor in Department of Computer Science and Engineering, Southern University of Science and Technology, China.

Her research interests include medical image processing and ophthalmology.

E-mail: lisahigashita@gmail.com

Jiang Liu received the B. Sc. degree in computer science from University of Science and Technology of China, China in 1988, the M. Sc. and Ph. D. degrees from National University of Singapore, Singapore in 1992 and 2004, respectively. He founded the Intelligent Medical Imaging Research Team which was once the world's largest ophthalmic medical image processing team, focusing on ophthalmic artificial intelligence research. Currently, he is a professor in Department of Computer Science and Engineering, Southern University of Science and Technology, China.

His research interests include artificial intelligence, eye-brain research, precision medicine, and surgical robots.

E-mail: liuj@sustech.edu.cn (Corresponding author)

ORCID iD: 0000-0001-6281-6505