Biomedical Signal Processing and Control 70 (2021) 103030

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

CDC-Net: Cascaded decoupled convolutional network for lesion-assisted detection and grading of retinopathy using optical coherence tomography (OCT) scans

Bilal Hassan [a,*], Shiyin Qin [a,b], Taimur Hassan [c,d], Muhammad Usman Akram [c], Ramsha Ahmed [e], Naoufel Werghi [d]

[a] School of Automation Science and Electrical Engineering, Beihang University (BUAA), Beijing, 100191, China
[b] School of Electrical Engineering and Intelligentization, Dongguan University of Technology, Dongguan, 523808, China
[c] Department of Computer and Software Engineering, National University of Sciences and Technology, Islamabad, 44000, Pakistan
[d] Center for Cyber-Physical Systems, Khalifa University of Science and Technology, Abu Dhabi, 127788, United Arab Emirates
[e] School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing, 100083, China

ARTICLE INFO

Keywords:
Optical coherence tomography (OCT)
Deep learning
Lesion segmentation
Retinopathy grading
Retinal analysis

ABSTRACT

Retinopathy refers to any injury in the retinal region of the eye that can lead to distorted vision or even blindness. The segmentation of retinal lesions or biomarkers is crucial for the precise classification and grading of retinopathy. Optical coherence tomography imaging is the widely used eye examination tool by ophthalmologists due to its comprehensive visualization of the retinal lesions, which can assist in the prompt treatment of retinal conditions. However, due to vast clinical optical coherence tomography applications and the incidence of ocular syndromes, the number of scans collected daily outweighs ophthalmologists' capacity to interpret these in a meaningful way. Many studies have been proposed previously to address this issue using optical coherence tomography scans. However, to the best of our knowledge, no framework can perform joint segmentation of the retinal lesions and retinopathy grading. In this paper, we propose a novel framework to address this shortcoming. We propose a new cascaded decoupled convolutional network comprising two separate modules that work together to perform lesion-assisted grading of retinopathy according to the clinical standards. We thoroughly evaluated the proposed framework using 26841 multi-vendor scans spanned over four publicly accessible datasets. The obtained results validate the efficacy of the proposed scheme over other state-of-the-art frameworks, where it achieved the mean Dice score of 0.820 (3.66% improvement) in segmentation of retinal lesions and 98.89% accuracy in the grading of retinopathy with 98.34% true positive rate and 99.17% true negative rate.

1. Introduction

Humans perceive visual details through their eyes. The human eye consists of three layers, where retina is the innermost layer responsible for creating vision [1]. It consists of two segments, the macular and the peripheral. The macular region (also known as macula) is essential for forming vivid central vision, while the peripheral retina produces lateral vision. The light focused by the biconvex lens falls on the retina, thereby stimulating rods and cones. The visual information is then transmitted to the brain via the optic nerve for interpretation [1].

Many retinal complications are due to diabetes [1]. In most cases, elevated blood sugar causes retinal blood vessels to rupture, resulting in blood leakage and other fluid deposits (lesions) in the retinal substructures. This medical condition is known as diabetic retinopathy (DR), which is the leading cause of blindness worldwide [2–4]. The four commonly associated fluids in the etiology of DR and macular degenerations are intraretinal fluid (IRF), subretinal fluid (SRF), pigment epithelial detachment (PED), and drusen/reticular pseudodrusen (RPD) [5]. Besides, retinal fluid (IRF, SRF) leakage due to DR can cause macular thickening or macular edema (ME). In non-diabetic subjects, ME can still occur for several reasons, including cataract surgery, macular degeneration, uveitis, retinal vein, or arterial blockages [6]. Macular degeneration (also known as age-related macular degeneration (AMD)) is another retinal syndrome that often affects older people. AMD generally manifests into two stages, namely dry or non-exudative AMD and wet or exudative AMD. RPD is formed in the early stages of AMD, and vision is distorted and blurred due to the atrophy of

* Corresponding author.
E-mail address: bilalhassan@buaa.edu.cn (B. Hassan).

https://doi.org/10.1016/j.bspc.2021.103030
Received 24 April 2021; Received in revised form 2 July 2021; Accepted 30 July 2021
Available online 17 August 2021

Table 1
Description of retinal lesions investigated in this study.

Lesions	Description
IRF	IRF corresponds to the adjacent fluid-filled cavities comprising the columns of retinal tissue. These cavities can appear as discrete, low-reflective cystoid spaces, so IRF is often referred to as intraretinal cystoid fluid.
SRF	SRF involves the accumulation of translucent or lipid-rich lesions in the subretinal zone (between the neurosensory retina and the underlying retinal pigment epithelium (RPE)).
PED	PED is relevant to AMD syndrome and reflects the disruption of the RPE together with the overlying retina of the remaining Bruch's membrane due to blood leakage and fluid deposition. In this study, we have considered two types of PED, fibrovascular and serous.
RPD	Drusen is morphologically classified as hard tissue, characterized as slight hyalin formations with delineated boundaries that can also be considered as age-related low-risk alterations; whereas soft drusen is an accumulation of amorphous or granular material that is known to be precursor to AMD. In comparison to drusen, which is found below the RPE, the RPD lesion is observed superficial to the RPE.

the retinal pigment epithelium (RPE). The further development of this disease can lead to wet macular degeneration, where abnormal blood vessels in the choroid block the retina and start to leak fluid deposits (PED) [7,8]. This retinal complication is also known as neovascular AMD or choroidal neovascularization (CNV) [9]. Table 1 describes the four retinal manifestations (IRF, SRF, PED, and RPD) based on the morphology, volume, or reflectivity patterns.

Many retinal examination techniques can adequately visualize the retinal anatomical symptoms in diabetic and non-diabetic patients. The two widely used clinical practice techniques are fundus, and optical coherence tomography (OCT) imaging due to their non-invasive nature and ability to visualize retinal pathology more accurately [10–13]. However, since OCT imaging can intuitively depict the biological tissues and substructures, it is more commonly used for various medical applications [14–16]. Further, spectral-domain OCT (SD-OCT) imaging offers better sensitivity and higher acquisition speed, allowing the collection of three-dimensional (3-D) scans with improved pixel density. The resultant 3-D retinal volume consists of sequential 2-D imageries called OCT brightness scan (B-scan). These B-scans are defined in the horizontal-axis plane, with each B-scan in a separate vertical position. Besides, each B-scan is composed of a series of axial lines, termed as A-scans. Fig. 1 shows the standard SD-OCT nomenclature along with various SD-OCT based retinal lesions.

In addition to this, OCT provides exhaustive imaging traits of the disease phenotype, which could be used for timely intervention and treatment of retinal disorders [17,18]. ME and AMD are both progressive retinal conditions that can manifest into intermediate or advanced stages reflected by the presence of underlying retinal lesions. Hence, considering OCT imaging modality along with the underlying retinal morphological changes, many researchers have proposed grading systems to categorize the ME and AMD diseases [9,19–22]. Broadly, as observed in Table 2, both ME and AMD diseases can be classified in two main types following the clinical rules defined by the early treatment diabetic retinopathy study (ETDRS) [22] and age-related eye disease study (AREDS) [9]. An automated tool for screening these ocular syndromes and segmentation of underlying lesions is a critical need. Such methods greatly help determine the severity of the condition, allowing physicians to recognize more precise diagnoses and treatments.

1.1. Related works

In the past decade, many studies have been presented in the context of OCT-based retinal image analysis. These studies are either focused on OCT imaging's clinical perspective for visualizing retinal diseases or focused on utilizing OCT imaging to segment and extract retinal features such as layers, fluids, and lesions. Also, several OCT-based automated retinal diagnosis systems have been proposed in the past decade.

Table 2
Clinically adopted strategy for retinopathy grading based on findings from the OCT imaging only.

Diseases	Types	Description
ME	CME	Increased retinal thickness due to presence of IRF, irregularity in retinal substructures, reduced intraretinal reflectivity, and flattening of the foveal depression.
	SRD	Any of the above, associated with SRF at the inferior retinal layers.
AMD	Dry-AMD	When only RPD are formed on the inferior retina, beneath the macula, leading to degradation or degeneration over time.
	Wet-AMD	When RPD along with fibrovascular PED, SRF or IRF are observed within the OCT scan.

CME = Cystoid Macular Edema, SRD = Serous Retinal Detachment.

1.1.1. Clinical evaluations

Clinicians have thoroughly adapted OCT imaging for the screening of retinal subjects and have consistently emphasized the relevance of OCT in the evaluation of the retina. Clinically, OCT imagery acts as an initial examination procedure for screening ME subjects [23] or for the ME severity analysis [22] as it presents the early and objective visualization of ME pathology. Similarly, the authors in [24] defined quantitative biomarkers to detect the presence of intermediate AMD using OCT scans of older adults. Moreover, studies focused on comparing the OCT imaging-based findings with other retinal examination techniques are also proposed in the literature. In [25], a comparison for grading AMD and CNV using OCT imaging, fundus fluorescein angiography, and fundus photography has been made, where OCT is concluded to be a useful tool to pick even the slightest early variations of CNV and AMD.

1.1.2. Retinal layers and lesions segmentation

The segmentation of retinal layers and retinal lesions is crucial to the treatment and prognosis of retinal pathologies. Manual delineation of retinal layers and various lesions associated with the ocular syndromes is a very time-consuming and tedious process. Consequently, several researchers have proposed fully and semi-automated frameworks as prospective solutions. In [26], a graph theory and dynamic programming (GTDP) based model is proposed to segment up to eight retinal layers in the OCT scans. The automated segmentation and extraction of the retinal layers are not limited to healthy eyes, but several researchers have presented algorithms for similar tasks using OCT scans of various retinal pathologies [27,28]. Further, an improved version of [26] for the extraction of the retinal layer, and retinal fluid was introduced in [29]. They evaluated the algorithm using 610 OCT scans of 10 severe ME subjects.

Besides, GTDP framework [26] is also used in the literature to segment the retinal layers in 15 AMD patients and six control subjects [30]. Subsequently, a 3-D representation of retinal morphology was created in [30] to investigate retinal hypoxia and oxygen distribution across the retina. In addition to this, an approach based on the random forest was proposed in [31] to automatically extract retinal layers and fluids from central serous retinopathy (CSR) affected scans. They validated the framework on a local dataset comprising 48 OCT scan [31]. Similarly, an automated framework was proposed in [32] for extracting retinal layers from the optic nerve head scans. In another method [33], to identify the IRF lesions in OCT scans of AMD and ME subjects, authors extracted 312 different features and performed the classification of IRF lesions using support vector machines (SVM), linear discriminant classifier, and Parzen window. Further, a framework based on neutrophil sets

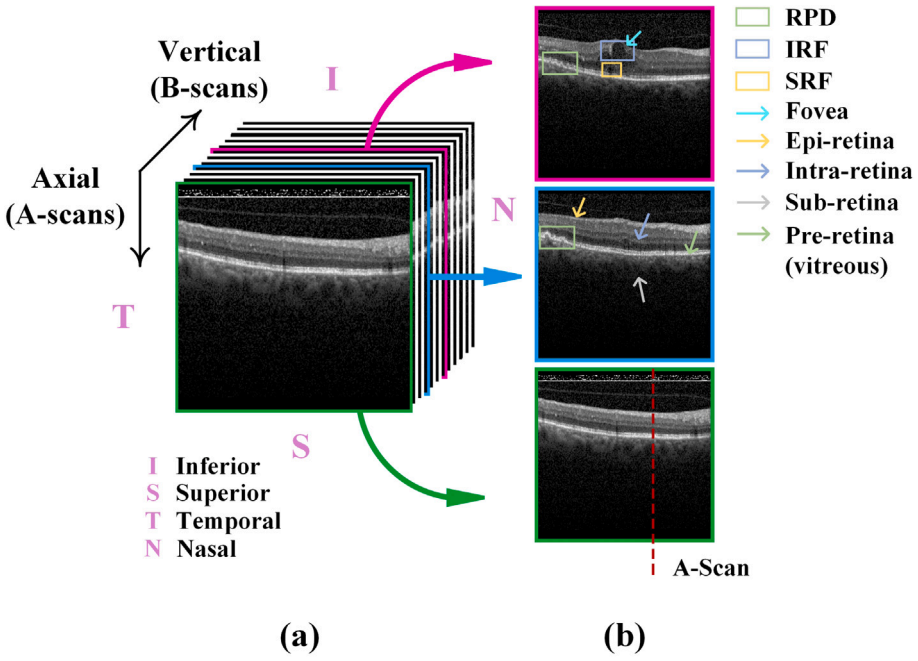

Fig. 1. OCT imaging nomenclature. (a) OCT volumetric scan consisting of multiple B-scans. (b) Visualization of OCT B-scans with various lesions, where drusen/reticular pseudodrusen (RPD) is observed in the temporal retina (blue and magenta B-scans), intraretinal fluid (IRF) and subretinal fluid (SRF) can be seen in temporal retina (magenta B-scan).

and graphing algorithms is presented in [34] to automatically segment IRF using OCT scans of ME subjects.

Much recently, the field of retinal imaging has surged from the applications of deep learning, in particular for the extraction of retinal layers and segmentation of lesions [35]. Compared with traditional methods, deep learning methods have many advantages, such as the robustness and reliability for retinal information extraction [36,37]. In [38], a deep learning framework is developed to segment the retinal layers and fluid regions. The authors in [39] presented a convolutional neural network (CNN) based model to detect and quantify SRF and IRF. Their proposed method is evaluated using 1200 multi-vendor OCT volumes corresponding to three groups of patients (retinal vein occlusion (RVO), ME, and AMD) [39]. Similarly, using multi-vendor OCT retinal scans, a fully convolutional network (FCN) is proposed in [40] to segment the intraretinal cysts automatically. In general, deep learning approaches have shown impressive performance for the extraction and segmentation of retinal information. However, such methods rely heavily on the large cohort of annotated data and substantial training time to achieve better results.

1.1.3. Retinal diagnosis

In addition to extraction of retinal information, such as layers and lesions, many computer-aided diagnostics (CAD) tools have also been proposed by researchers in the literature to predict various retinal pathologies [41–43]. The motivation towards developing these CAD systems is to mass screen the retinal subjects across the globe, especially in rural settings where consultation with an expert ophthalmologist is not easily available. Besides, such systems can be operated locally by the technicians to filter out the retinal subjects so that only critical patients are referred to the doctor for rapid examination. Some researchers have developed CAD systems to screen normal and abnormal retina [44,45], while others have proposed methods that can classify different retinal abnormalities [46,47].

Furthermore, transfer learning-based frameworks are also proposed with promising results for screening different retinal pathologies [48]. An automatic diagnosis system for analyzing CME and macular hole disease in the retina is proposed in [49]. They validated the system using retinal scans of 18 patients. In another approach [50], an automatic system is introduced to screen AMD and ME subjects. In [51], a framework based on the multi-scale histogram of gradients (HOG) algorithm with SVM classifier is proposed to screen the AMD, ME, and healthy eyes. Furthermore, the authors in [8] presented a deep learning-based automated diagnosis of healthy, dry AMD, wet AMD, and ME subjects using OCT imagery. In [52], a deep retinal screening system to classify healthy, AMD, CNV, and ME eyes is proposed. They evaluated the algorithm for 1000 OCT scans containing 250 scans against each class. Similarly, in [53], a surrogate-assisted CNN-based framework is proposed to classify the AMD, ME, and normal OCT scans.

After conducting a comprehensive literature review, we observed that existing studies had been presented with a focus either towards the segmentation of retinal features or towards the diagnosis of retinal pathologies. In this research, we aim to recognize and segment the retinal anomalies towards achieving the lesion-assisted grading of retinal pathologies. In the past, the CSR, RVO, AMD, and ME pathologies have been investigated to recognize the associated fluid lesions [31,39]. Similarly, researchers have highlighted the significance of OCT imaging-based fluid biomarkers for retinal diagnosis, where authors in [33] identified the presence of IRF in ME and AMD subjects. However, the segmented lesions are not intuitively utilized afterward for the grading of underlying retinal pathology. At the core of this work, we presume that an adequately trained lesion-assisted deep learning framework can assist ophthalmologists through the automated delineation of various biomarkers and grading of underlying retinal conditions.

1.2. Key contributions

In this research, we propose a cascaded decoupled convolutional network (CDC-Net) for the segmentation of retinal lesions towards analysis and grading of retinopathy. The main contributions of this research are as follows:

1. The proposed CDC-Net is a hybrid model composed of two decoupled CNN modules (CDC-m1 and CDC-m2) to provide an in-depth interpretation of retinal pathology in terms of underlying lesions using OCT scans.

Table 3
List of abbreviations.

Acronym	Definition	Acronym	Definition	Acronym	Definition
AMD	Age-related Macular Degeneration	CNV	Choroidal Neovascularization	PED	Pigment Epithelial Detachment
A-scan	Axial Scan	Conv	Convolution	ReLU	Rectified Linear Units
AUC	Area Under the Curve	CSR	Central Serous Retinopathy	ROC	Receiver Operating Characteristic
BN	Batch Normalization	DR	Diabetic Retinopathy	RPD	Drusen/Reticular Pseudodrusen
B-scan	Brightness Scan	ILM	Inner Limiting Membrane	RPE	Retinal Pigment Epithelium
CDC-m1	CDC-Net Module 1	IRF	Intraretinal Fluid	SOTA	State-of-the-Art
CDC-m2	CDC-Net Module 2	ME	Macular Edema	SRD	Serous Retinal Detachment
CDC-Net	Cascaded Decoupled Convolutional Network	NLM	Non-local Means	SRF	Subretinal Fluid
CME	Cystoid Macular Edema	OCT	Optical Coherence Tomography	TConv	Transposed Convolution

Table 4
List of symbols.

Symbol	Description	Symbol	Description
Gr	Normalized Gradient	$\mathcal{L}_{\mathcal{CE}}$	Cross-entropy Loss
I_{OCT}	Raw OCT B-scan	$f(\varphi)$	Softmax Function
W_{DL}	Dark-to-Light Window	IoU	Intersection-over-Union
W_{LD}	Light-to-Dark Window	DSC	Dice Coefficient
F	Feature Maps	TPR	True Positive Rate
K	Receptive Field	TNR	True Negative Rate
R_A	ReLU Activation	ACC	Accuracy
$\mathcal{L}_{\mathcal{D}}$	Dice Loss Function	T_p	True Positives
$\mathcal{P}$	Network Segmented Labels	T_n	True Negatives
$\mathcal{G}$	Ground Truth Labels	F_p	False Positives
$\mathcal{W}_c$	Class Weighing Factor	F_n	False Negatives

2. CDC-Net is the single lesion-assisted retinal grading framework, where CDC-m1 localizes and segments the four retinal manifestations, namely IRF, SRF, PED, and RPD from the candidate scan. The extracted retinal lesions information guides the CDC-m2 to grade the retina accordingly (ME into CME or SRD and AMD into dry-AMD or wet-AMD).
3. A novel normalized gradient-based preprocessing mechanism is introduced to extract the chorioretinal region in the OCT scans, allowing the model to deal with different artifacts, vendor annotations, and inter-scanner variations. The lesion segmentation performance of the CDC-Net and other state-of-the-art (SOTA) frameworks is determined with and without the preprocessing steps. Considering the mean Dice score, the improvement rate of preprocessing is observed between 13.36% and 26.44%.
4. A prediction of the candidate OCT volume (eye-level grading) based on a novel run-length method by detecting the longest uninterrupted connectivity at the B-scan level, thereby grading the underlying pathology.
5. A rigorous evaluation of the CDC-Net for segmenting retinal lesions, and also for the classification and grading of retinopathy using 26 841 multi-vendor OCT scans from four publicly accessible datasets. The proposed framework outperformed other SOTA solutions by achieving 3.66% higher lesions segmentation accuracy in terms of mean Dice score and better classification accuracy in grading retinopathy across all four datasets.

The rest of the paper is organized as follows: Section 2 presents the proposed methodology, Section 3 describes the experimental setup, followed by the results in Section 4. Section 5 presents a detailed discussion about CDC-Net, and finally Section 6 concludes the paper. In addition, Tables 3 and 4 enlist the important acronyms and symbols used in this research.

2. Proposed methodology

CDC-Net combines two cascaded decoupled CNN modules, namely CDC-m1 and CDC-m2, dedicated for the segmentation of lesions and retinopathy grading, respectively. First of all, the input scan is preprocessed through normalized gradients to extract the retina and suppress the background content. Afterward, the scan is passed to the CDC-m1 module, which segments the retinal lesions in the candidate scan. The segmented lesions are then overlaid on the preprocessed scan, which is passed on to the CDC-m2 module to perform lesion-assisted classification and grading according to the clinical protocols as listed in Table 2. In the same way, all scans in the OCT volume are processed, and then the longest uninterrupted connectivity is examined to generate the eye or volume-level grading. CDC-Net has been evaluated on 26 841 scans from four publicly available datasets acquired using different OCT devices. Fig. 2 presents the high-level block diagram of the proposed CDC-Net framework.

2.1. Preprocessing

Preprocessing is the initial phase in the CDC-Net framework, where each OCT volume is processed at the B-scan level to eliminate various OCT artifacts as shown in Fig. 3. In the preprocessing stage, the raw OCT B-scan (I_{OCT}) is first denoised by applying the non-local means (NLM) filtering [54], followed by different morphological operations. Next, the chorioretinal region in the scan is extracted through the normalized gradients approach [29]. Finally, the inner limiting membrane (ILM) and the choroid layers are delineated on the scan to obtain the final preprocessed scan. The incorporated preprocessing technique to isolate the chorioretinal region helps overcome the erroneous lesion segmentation because of the noisy artifacts. The normalized gradients Gr_{an} and Gr_{bn} are expressed in Eqs. (1)–(4):

$$Gr_{a_{n,m}} = \left(\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} I_{OCT_{i,j}} W_{DL_{n-i,m-j}} \right), \tag{1}$$

$$Gr_{b_{n,m}} = \left(\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} I_{OCT_{i,j}} W_{LD_{n-i,m-j}} \right), \tag{2}$$

$$Gr_{an} = \frac{(Gr_a - min(Gr_a))}{(max(Gr_a) - min(Gr_a))}, \tag{3}$$

$$Gr_{bn} = \frac{(Gr_b - min(Gr_b))}{(max(Gr_b) - min(Gr_b))}, \tag{4}$$

where $Gr_{a_{n,m}}$ is the resultant pixel for the dark-to-light gradient, $Gr_{b_{n,m}}$ is a resultant pixel for light-to-dark gradient, $W_{DL} = [+1, -1]^T$ is the dark-to-light window, $W_{LD} = [-1, +1]^T$ is the light-to-dark window, M and N are the respective width and height of W_{DL} and W_{LD}, whereas m and n denote the current pixel coordinates. Both windows are kept vertically aligned here to pick horizontal layers transition. After extracting the normalized gradients (Gr_{an} and Gr_{bn}), ILM and the choroid layers are extracted by the binarization of Gr_{an} and Gr_{bn} using hysteresis thresholding approach [17], followed by an iterative technique to select the brightest top and bottom pixels from Gr_{an} and Gr_{bn}, respectively. The extracted layers are utilized in generating a retinal mask to isolate the candidate retina. Fig. 4 shows the various preprocessing steps involved to obtain the final preprocessed scan with cropped retina, delineated ILM (green), and choroid (purple) layers.

2.2. CDC-Net architecture

CDC-Net is a hybrid CNN-based framework that extracts and identifies retinal lesions for the grading of retinopathy. The formation

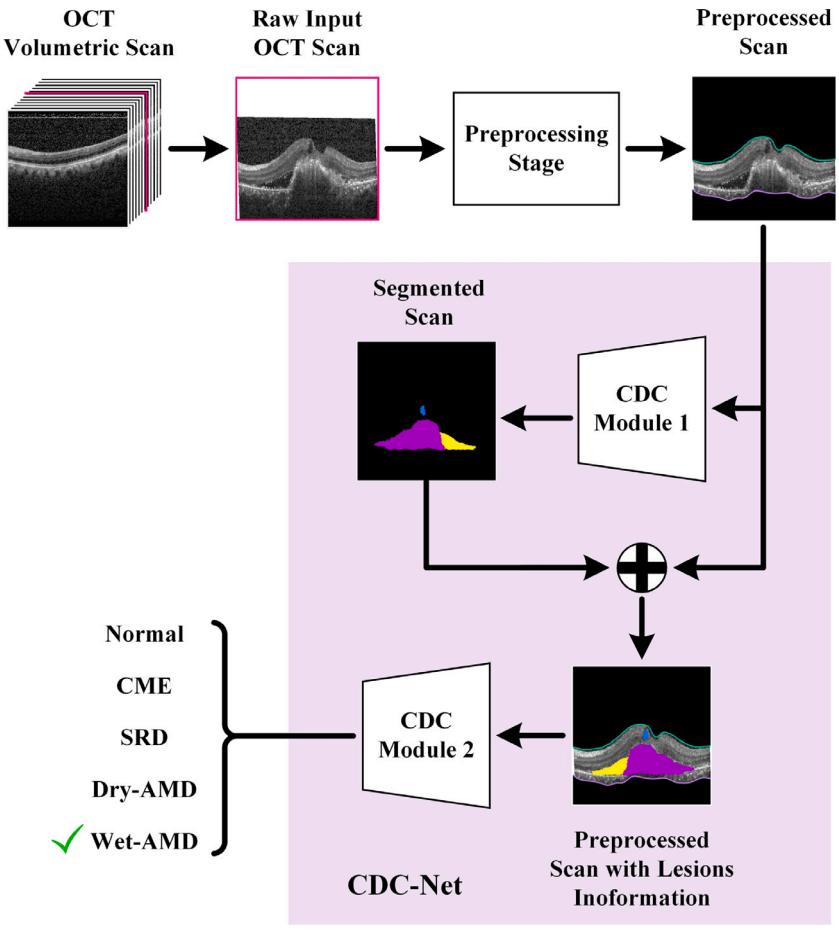

Fig. 2. Proposed CDC-Net block diagram.

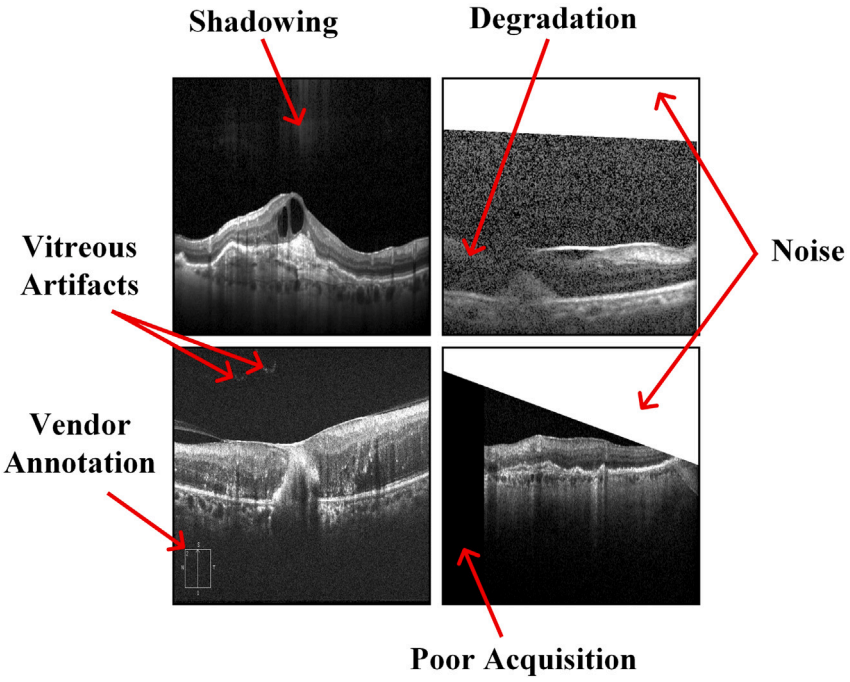

Fig. 3. Visualization of various OCT artifacts that degrade the scan quality and affect both the lesions segmentation and retinopathy grading performance.

of retinal lesions in different regions causes different types of visual impairment. CDC-Net is a combination of two cascaded decoupled CNN modules, as shown in Fig. 2. These modules are named CDC-m1 and CDC-m2. CDC-m1 is a convolutional feature encoder–decoder

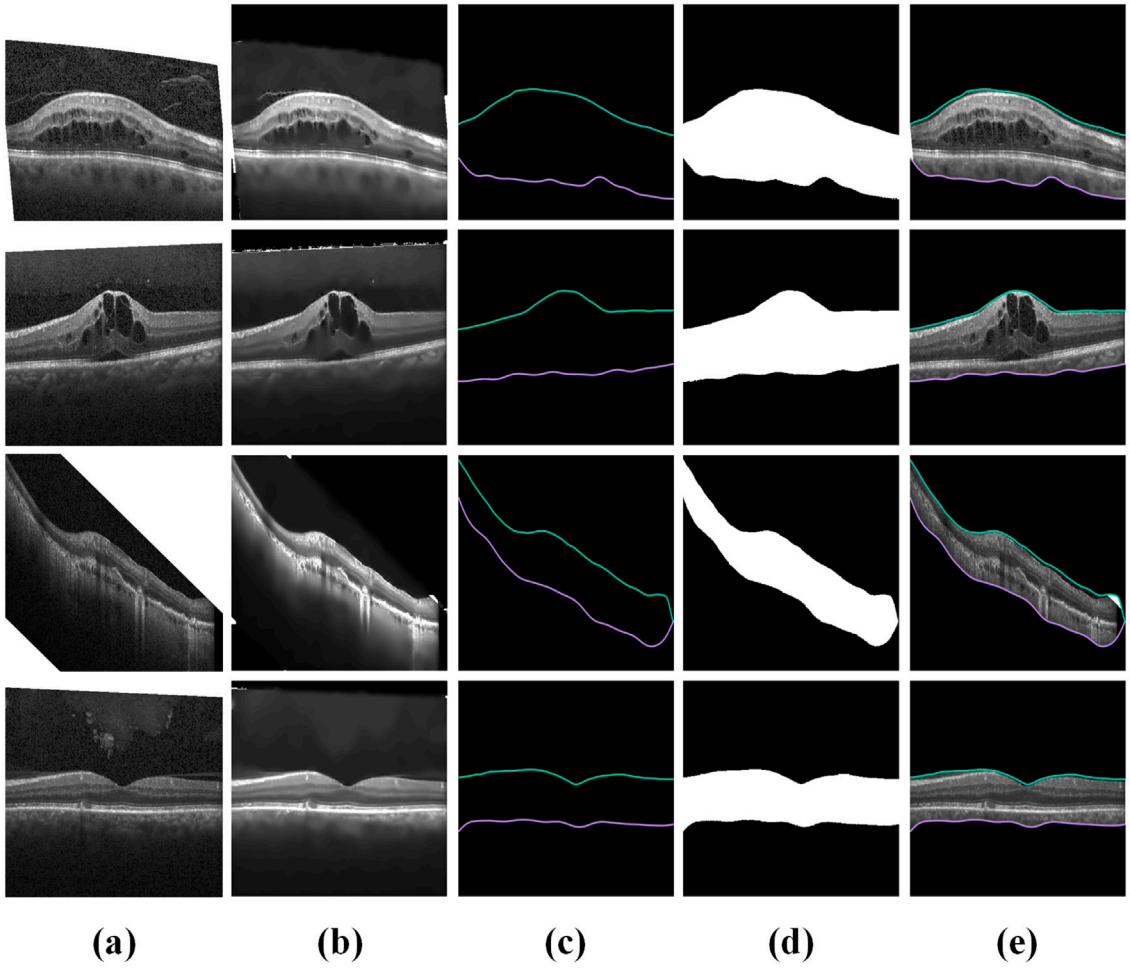

Fig. 4. Randomly selected OCT scans at different phases of preprocessing stage. (a) Raw OCT scans. (b) Denoised scans with NLM filtering. (c) Extracted ILM (green) and choroid (purple) layers using normalized gradients. (d) Chorioretinal masks. (e) Final preprocessed scans with isolated chorioretinal region and delineated ILM choroid layers.

architecture used to segment the retinal lesions by generating a lesions probability map. CDC-m2 takes the lesions mapped scan as an input to perform the lesion-assisted grading of the candidate retina. The detailed description of these two modules is presented next:

2.2.1. CDC-m1 for segmentation of retinal lesions

CDC-m1 consists of two blocks (feature encoder and feature decoder) that jointly perform the segmentation and extraction of retinal lesions, as shown in Fig. 5. Feature encoder block decomposes the candidate input OCT scan while retaining the lesion region through trained kernel weights. Afterward, CDC-m1 up-samples the decomposed scan to the original resolution in which only lesion pixels are preserved. Unlike in conventional segmentation architectures, the CDC-m1 encoder computes the most coherent channel's gradient at each depth to extract the finer lesion details, which are directly passed on to the respective decoder depth. Besides, the transposed convolution is performed at each decoder depth for un-pooling feature maps rather than separate convolution and up-sampling operations. This approach involves lesser memory consumption than the feature maps concatenation method in traditional architectures and offers a better spatial representation of lesion segments. Also, CDC-m1 computes contours-based features instead of measuring edge-based features, which helps in better preserving the morphology of the lesions during the downsampling process.

- **Feature encoder:** Each depth of the CDC-m1 encoder involves sequential convolution (Conv), batch normalization (BN), and rectified linear units (ReLU) operations, as depicted by the blue arrows in Fig. 5. To generate the subsequent feature maps of an input image I_{OCT}, each Conv operation is performed using a 3×3 sized receptive field K with same padding and single stride as expressed in Eqs. (5), (6):

$$F[i,j] = (I_{OCT} * K)[i,j], \tag{5}$$

$$F[i,j] = \sum_{n}\sum_{m} K[n,m] I_{OCT}[i-n, j-m], \tag{6}$$

here $F[i,j]$ is the resultant feature maps having indexes of rows and columns marked with i and j respectively. Moreover, the region corresponding to lesions in the input OCT scan is preserved due to Conv weight adjustment during the training phase. Next, these weights are normalized using the ReLU activation function to truncate the negative values and ensure that only pixels corresponding to the feature map's lesions are carried forward. The ReLU activation is applied to feature maps as expressed in Eq. (7):

$$R_A(F) = \begin{cases} F, & \text{if } F > 1 \\ 0, & \text{otherwise} \end{cases} \quad \forall F = \{n, m \mid n \in i, m \in j\}, \tag{7}$$

here, $R_A(F)$ is the ReLU function applied to F. The height and width of the feature maps are denoted by i and j, respectively. Furthermore, the feature maps after each encoder depth are reduced through a 2×2 max-pooling operation with stride two, as represented by the red arrows in Fig. 5. In addition to this, the lesion features from the most coherent channel at each depth of the encoder are passed to the corresponding decoder end. The feature maps at the last encoder depth are reduced to 1/32 size of the input scan's resolution.

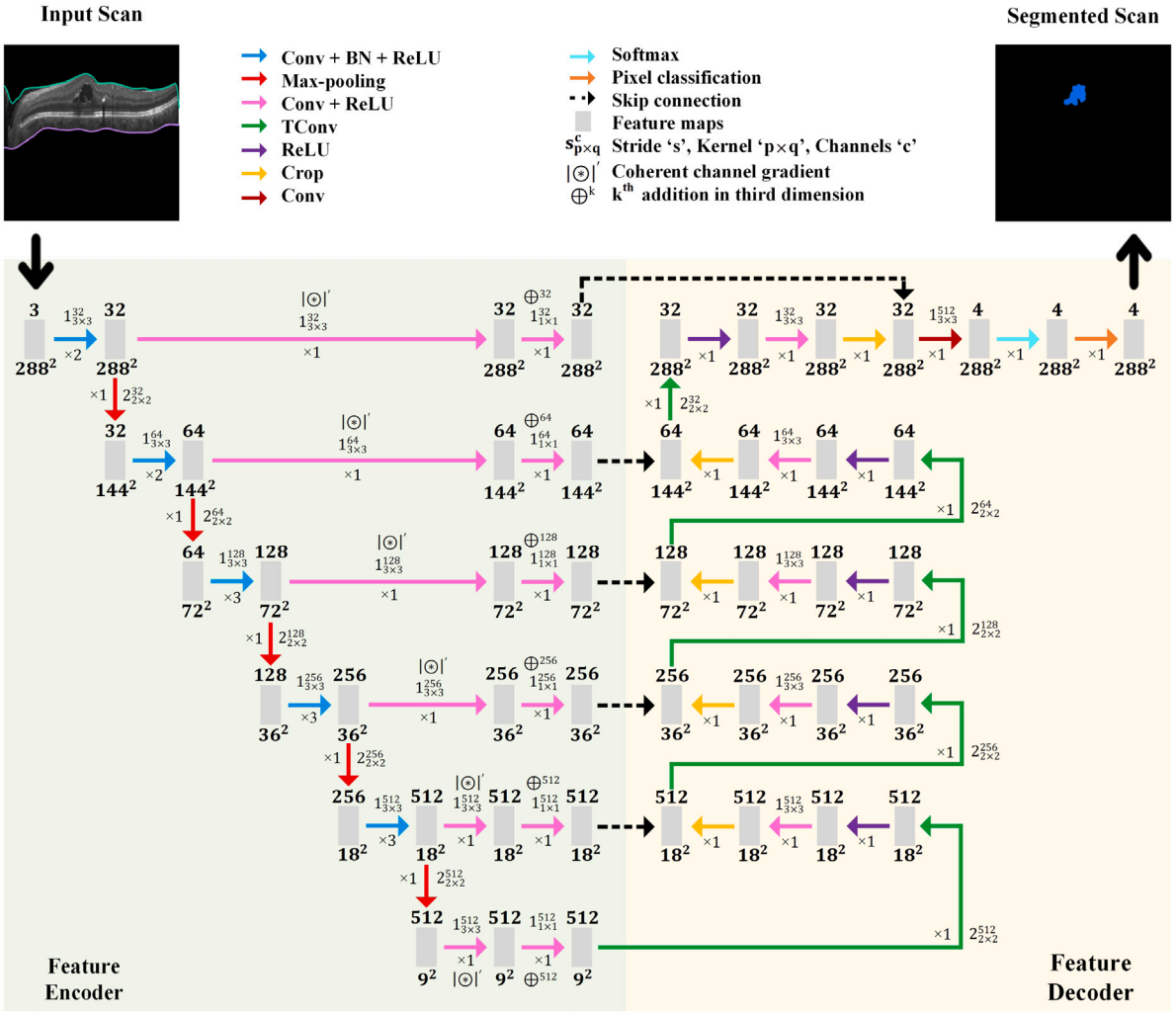

Fig. 5. Depth-wise architecture of CDC-m1 comprising of feature encoder (green) and feature decoder (yellow) blocks. Different color arrows represent various operations performed at each depth of the network. Dotted black arrow depict the skip connections between the encoder side and corresponding decoder depth.

- **Feature decoder:** The feature decoder block in CDC-m1 restores the feature maps to the original resolution in a step-wise manner. At each depth of the decoder block, 2 × 2 transposed convolution (TConv) operation with stride two is performed to upsample the feature maps by a factor of two, as shown in Fig. 6. TConv provides better and most versatile upsampling of abstract features [55], which is different from the computationally expensive upsampling techniques (such as bilinear interpolation or max-unpooling) that are usually used in typical decoder settings. Besides, finer lesion segments from the encoder side are added at the respective decoder depth to retain the retinal lesions geometry effectively. Softmax and pixel classification layers are used at the end of CDC-m1 to assign the label against each pixel having the highest class probability. Finally, various morphological enhancements are performed to refine the segmented retinal lesions to remove the stray pixels and small blobs.

2.2.2. *CDC-m2 for grading of retinopathy*

The segmented retinal lesions through CDC-m1 are embedded into the preprocessed scan, which is then passed on to the CDC-m2 to perform the lesion-assisted classification and retinopathy grading. The OCT scan containing retinal lesions information provides an intuitive way of analyzing the disease severity. For instance, an AMD-affected scan can be further graded easily as exudative or non-exudative based on the underlying retinal lesions or biomarkers. In this sense, CDC-Net is a unique framework as it diagnoses the disease and utilizes the segmented lesions to properly grade retinopathy as per the clinical standards as shown in Table 2.

CDC-m2 is a lesion-assisted directed acyclic graph convolutional network that contains 61 layers in total with integrated inception and residual blocks stacked in a parallel configuration as shown in Fig. 7. The inception block in CDC-m2 is inspired by the fundamental component of GoogLeNet [56] architecture, while the idea of a residual block is taken from the ResNet [57] model. The purpose of stacking these two blocks in a parallel configuration is to go deeper and wider simultaneously. These blocks jointly work to extract multi-scale features of retinal lesions using different receptive fields (convolutional kernels) to promote the grading performance in CDC-m2. The detailed explanation of these blocks is as follows:

- **Inception block:** The inception module in CDC-m2 ensures dimensionality reduction and projection while maintaining the sparse representation of the input feature maps. In the inception module, Conv operations with multiple receptive fields (1 × 1, 3 × 3, and 5 × 5) are performed to learn more local and global retinal lesions information disseminated in the OCT scan. However, since 3 × 3 and 5 × 5 sized Conv operations are usually expensive and increase the computational cost, an additional 1 × 1 Conv is applied to achieve dimensionality reduction. Besides, ReLU activation is applied after each Conv operation, enabling the network to handle considerable variations in terms of information and location embedded in the scan.

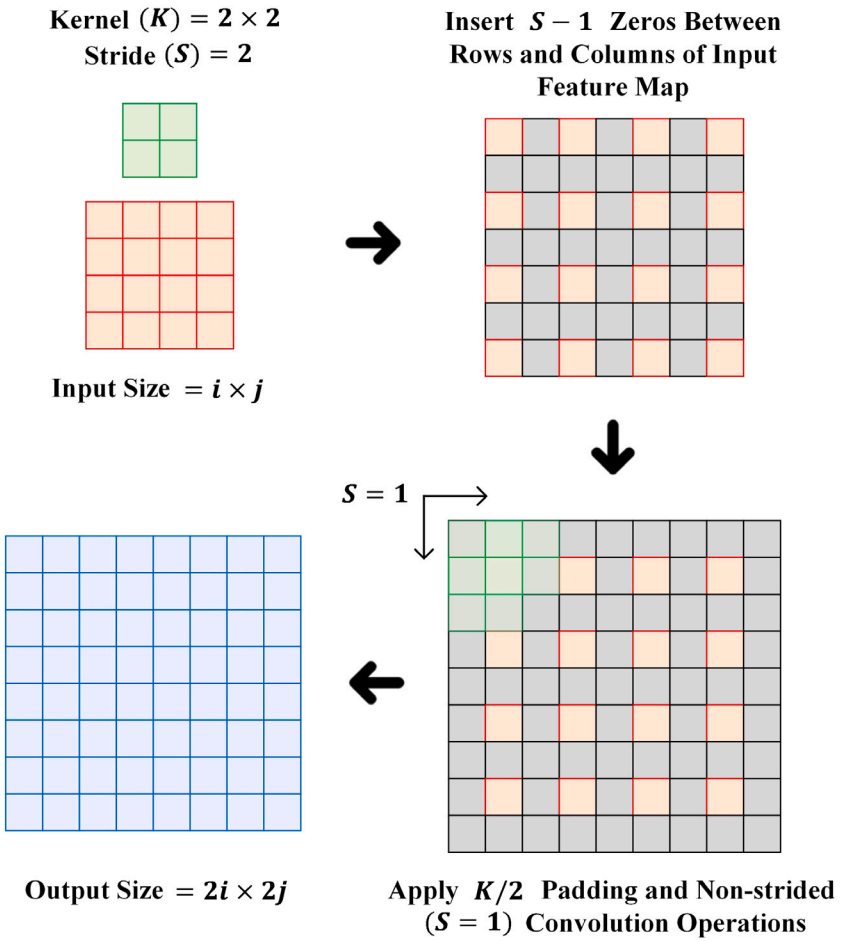

Fig. 6. Illustration of 2 × 2 transposed convolution operation with stride two applied to set of feature maps at each decoder depth to increase the spatial resolution by a factor of two.

- **Residual block:** Generally, in CNNs, the vanishing gradients make the training process more challenging [55]. One solution is to take the activations from one layer and transmit them to another via residual learning, allowing the network to learn deeper and high-resolution features [55]. Following this technique, we integrated residual block in the CDC-m2 architecture, which contains one 3 × 3 Conv and two 1 × 1 Conv with a skip connection between the input and output of the block.

In addition, 2 × 2 max-pooling with stride two is used in the CDC-m2 model to half the resolution of the feature maps. Further, the inception and residual blocks are configured at the higher layers, when the feature maps are reduced to 18 × 18 representation, to optimize the memory utilization during the training phase. Finally, global average pooling with a dropout of 40 percent is used before the fully connected layers in order to minimize the overfitting problem.

The proposed CDC-Net framework differentiates between healthy and retinopathy OCT scans. Further, it can grade the ME disease as CME or SRD and AMD as dry (non-exudative) or wet (exudative) by analyzing the presence of segmented retinal lesions. The layer-wise configurations and the learnables of the CDC-Net framework are listed in Table 5. Here, it should be noted that the number of layers and size is determined after exhaustive experimentation on various publicly accessible datasets.

Table 5
Summary of operational layers in CDC-Net model.

Layers	CDC-m1		CDC-m2	
	Number of operations	Params (million)	Number of operations	Params (million)
Input	1	0	1	0
Convolution	31	17.07	18	11.30
Batch normalization	13	0.01	12	0.01
Rectified linear unit	35	0	17	0
Max-pooling	5	0	5	0
Transposed convolution	5	1.75	–	–
Crop	5	0	–	–
Softmax	1	0	1	0
Addition	6	0	2	0
Concatenation	–	–	1	0
Global average pooling	–	–	1	0
Dropout	–	–	1	0
Fully connected	–	–	1	0.01
Classification	1	0	1	0
Total	103	18.83	61	11.32

2.2.3. Loss function

In this research, two different loss functions are used to train both CDC-Net modules, as the task of these two modules is different from each other. The focus of CDC-m1 is towards the segmentation of retinal lesions, where the major challenge is to overcome the imbalance between the number of background class pixels (which are significantly higher) and the remaining pixels of retinal lesion classes. In such cases (especially in medical image segmentation), the Dice loss function ($\mathcal{L}_{\mathcal{D}}$) and its variants are usually used by many researchers due to its substantial resistance to class imbalance [58]. Therefore, $\mathcal{L}_{\mathcal{D}}$ is also employed in CDC-m1 to ensure that the segmentation performance reflects the overall situation instead of mainly relying on the dominant

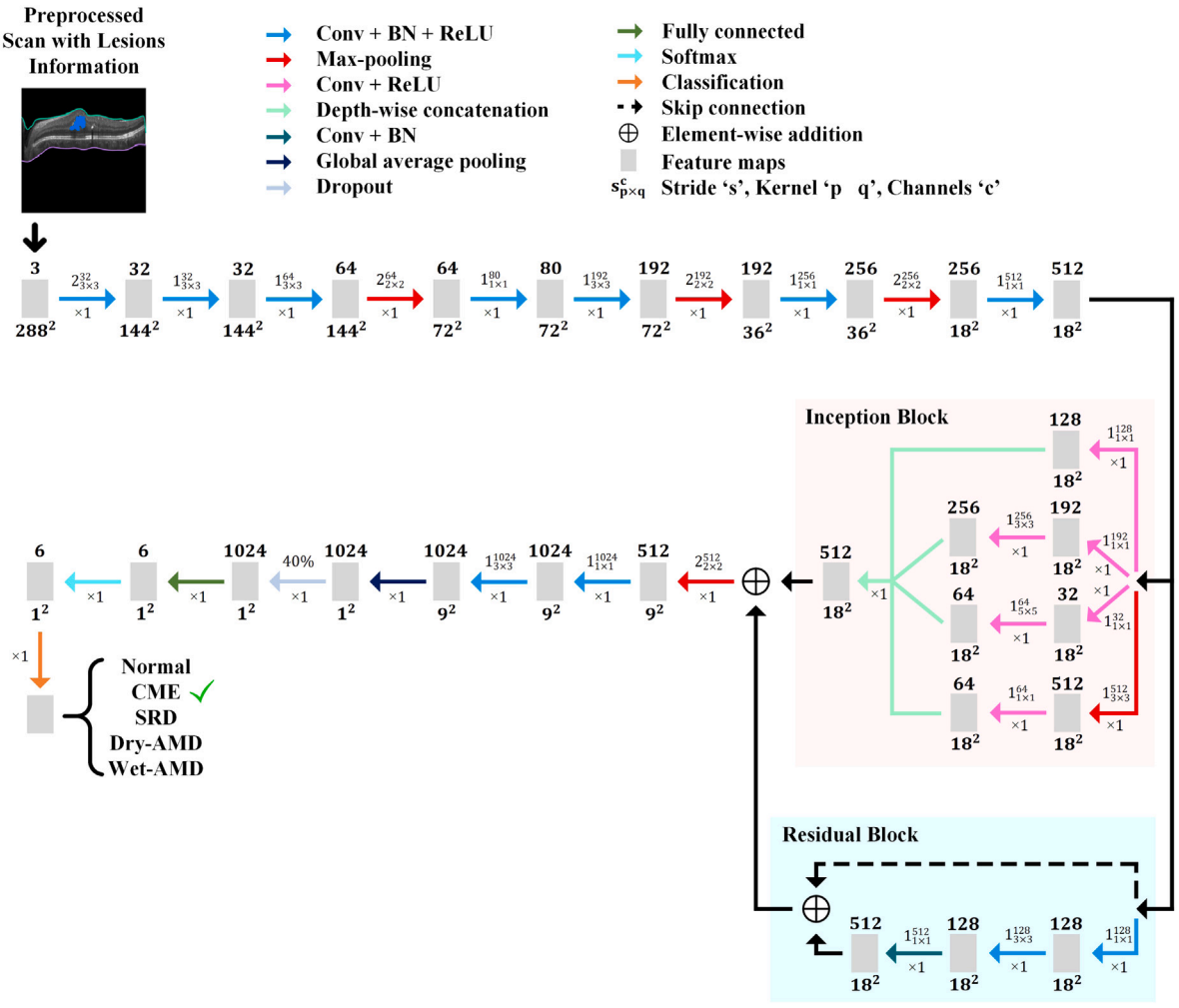

Fig. 7. Layer-wise architecture of CDC-m2 with integrated inception (pink) and residual (blue) blocks.

class (background). $\mathcal{L}_{\mathcal{D}}$ is measured as expressed in Eq. (8):

$$\mathcal{L}_{\mathcal{D}} = 1 - \frac{2\sum_{c=1}^{C}\mathcal{W}_c\sum_{e=1}^{E}(\mathcal{P}_{c,e}\mathcal{G}_{c,e})}{\sum_{c=1}^{C}\mathcal{W}_c\sum_{e=1}^{E}(\mathcal{P}_{c,e}^2+\mathcal{G}_{c,e}^2)}, \tag{8}$$

where $\mathcal{P}$ and $\mathcal{G}$ represent the segmented labels in the network predicted image and corresponding ground truth image, respectively. C is the total classes, E denotes the number of elements in the first two dimensions of $\mathcal{P}$. The weighting factor for each class is denoted by $\mathcal{W}_c$, which regulates the per class contribution to the total loss. $\mathcal{W}_c$ is an essential parameter since it tends to offset the effect of more significant regions on the Dice score and facilitates the learning process of the network to segment smaller regions. $\mathcal{W}_c$ is computed as expressed in Eq. (9):

$$\mathcal{W}_c = \frac{1}{\left(\sum_{e=1}^{E}\mathcal{P}_{ce}\right)^2}, \tag{9}$$

As the CDC-m2 is meant to perform the grading of retinopathy in mutually exclusive categories. We use the categorical cross-entropy loss ($\mathcal{L}_{\mathcal{CE}}$) to train it to output a probability (φ_c) over the C retinopathy classes for each OCT scan. $\mathcal{L}_{\mathcal{CE}}$ is computed as expressed in Eqs. (10), (11):

$$\mathcal{L}_{\mathcal{CE}} = -\sum_{c=1}^{C}\mathcal{T}_c\log\left(f(\varphi)_c\right), \tag{10}$$

$$f(\varphi)_c = \frac{\exp\left(\varphi_c\right)}{\sum_{c=1}^{C}\exp\varphi_c}, \tag{11}$$

$f(\varphi)_c$ here is the softmax function, and $\mathcal{T}$ is the target vector containing the one-hot term that is against the positive class, while the remaining terms (corresponding to other retinopathy classes) in $\mathcal{T}$ are zero. Therefore, $\mathcal{L}_{\mathcal{CE}}$ can further be simplified by discarding the summation elements as expressed in Eq. (12):

$$\mathcal{L}_{\mathcal{CE}} = -\log\left(\frac{\exp\left(\varphi_+\right)}{\sum_{c=1}^{C}\exp\varphi_c}\right), \tag{12}$$

where, φ_+ is the confidence score by the CDC-m2 against the positive class.

2.3. Eye-level grading

The proposed CDC-Net framework incorporates a simple yet efficient technique to generate the eye-level grading after processing all individual B-scans in a single OCT volume. This technique is developed after analyzing the pattern of graded abnormalities within the OCT volume. We observed retinal lesions' continuous presence in the consecutive B-scans leading to a similar retinopathy classification pattern. Therefore, the eye-level grading technique in CDC-Net is based on the longest uninterrupted connectivity of graded scans across the OCT volume, as shown in Fig. 8. Since the longest uninterrupted connectivity in Fig. 8 is of Dry-AMD, the complete volume (eye-level) is graded as Dry-AMD positive.

Table 6
Dataset details.

Datasets	Scans *(Subjects)*				Total scans	Training scans	Validation scans	Testing scans	Link
	Normal	Drusen (Dry-AMD)	ME	CNV (Wet-AMD)					
Duke-1 [24]	11 500 *(115)*	26 900 *(269)*	–	–	38 400	15 000	400	23 000	
Duke-2 [29]	–	–	610 *(10)*	–	610	–	500	110	
Duke-3 [51]	1407 *(15)*	723 *(15)*	1101 *(15)*	–	3231	–	500	2731	
Zhang [52]	51 390 *(3719)*	8867 *(881)*	11 599 *(876)*	37 456 *(968)*	109 312	107 712	600	1000	
Total	64 297 *(3849)*	36 490 *(1165)*	13 310 *(901)*	37 456 *(968)*	151 553	122 712	2000	26 841	–

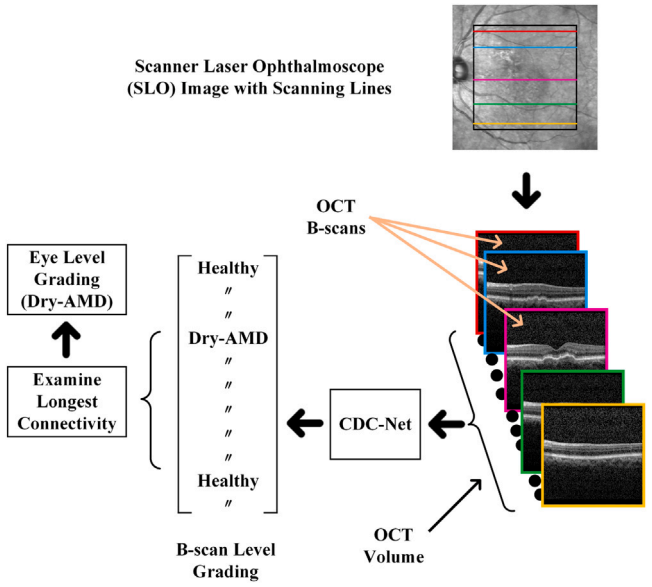

Fig. 8. Eye-level grading strategy adopted in CDC-Net framework.

3. Experimental setup

3.1. Data

In this research, four different datasets are used for the validation purpose of the CDC-Net model. The first three datasets (Duke-1 [24], Duke-2 [29], and Duke-3 [51]) are collected and arranged by the vision and image processing (VIP) laboratory of Duke University. These datasets contain a total of 42 241 OCT scans from 439 subjects, covering three different retinal pathologies (normal, drusen or dry-AMD, and ME). Besides, the scans in the Duke-1 dataset are acquired using the Bioptigen device, while the Spectralis machine is used to capture the scans in the Duke-2 and Duke-3 datasets. The final dataset (Zhang [52]) is provided by the Zhang lab at the University of California, San Diego. This dataset is imaged by the Spectralis OCT machine and contains 151 553 scans of four different retinal conditions (normal, drusen or dry-AMD, ME, CNV or wet-AMD). Moreover, OCT scans from two datasets (Duke-1 and Zhang) are used for training the network, while the remaining two datasets (Duke-2 and Duke-3) are not exposed to the network during the training process and are only used for validation purposes. Furthermore, all necessary retinal lesion annotations and cross-validations of retinopathy are arranged locally through three expert clinicians. The detailed summary of datasets with split ratio for training and validation of the network is presented in Table 6.

3.2. System and software

The proposed CDC-Net framework is implemented with MATLAB R2020a software, using the computer vision and deep learning toolboxes, installed on a 64-bit Windows operating system. The machine is configured as Intel i7-7500U @ 2.90 GHz processor with NVIDIA GTX 1080 Ti graphics card and 16 GB DDR4 memory.

Table 7
CDC-Net training hyper-parameters.

Parameters	Values
Training OCT scans	122 712
Validation OCT scans	2000
Optimizer	SGDM
Momentum	0.9
Initial learning rate	0.001 (CDC-m1), 0.0001 (CDC-m2)
Weight decay	0.005 (CDC-m1), 0.0001 (CDC-m2)
Mini-batch Size	128
Total epochs	80
Iterations per epoch	958
Total iterations	76 640
Validation frequency	10

3.3. Training details

The training for both CDC-Net modules (CDC-m1 and CDC-m2) is performed using 122 712 scans taken from two datasets (15 000 from Duke-1 and 107 712 from Zhang). Stochastic gradient descent with momentum (SGDM) is utilized as an optimizer to update both modules' network parameters. The mini-batch size of 128 is specified for training the network, and the training process lasted for 80 epochs with 958 iterations in each epoch. Further, 2000 OCT scans from all four datasets (400 from Duke-1, 500 each from Duke-2 and Duke-3, and 600 from Zhang) are used for validation purposes. The validation frequency is set to 10 epochs, which enabled the network to validate eight times on the unseen data during the training process. Table 7 shows the hyper-parameter details for both the CDC-Net modules, which are empirically tuned following systematic experiments on various publicly accessible datasets. The training performance curves of CDC-Net are shown in Fig. 9, where we can observe that the validation accuracy for both the modules started to converge around the 50th epoch.

3.4. Performance metrics

In this research, different metrics have been used to evaluate the CDC-Net model extensively and compare its efficacy with other SOTA solutions. We started with intersection-over-union (IoU) and Dice coefficient (DSC) to measure the retinal lesion segmentation performance. These metrics are computed as expressed in Eqs. (13), (14):

$$IoU = \frac{\text{Area}(\mathcal{P} \cap \mathcal{G})}{\text{Area}(\mathcal{P} \cup \mathcal{G})}, \tag{13}$$

$$DSC = \frac{2 \times IoU}{1 + IoU}, \tag{14}$$

$\mathcal{P}$ and $\mathcal{G}$ here are the segmented scan by the model and corresponding ground truth scan. Besides, we used the true positive rate (TPR), true negative rate (TNR), and accuracy (ACC) metrics to measure the performance of the CDC-Net model in terms of correct detection of retinal lesions in the candidate OCT scan and for grading of retinopathy. These metrics are computed as expressed in Eqs. (15)–(17):

$$TPR = \frac{T_p}{T_p + F_n}, \tag{15}$$

$$TNR = \frac{T_n}{T_n + F_p}, \tag{16}$$

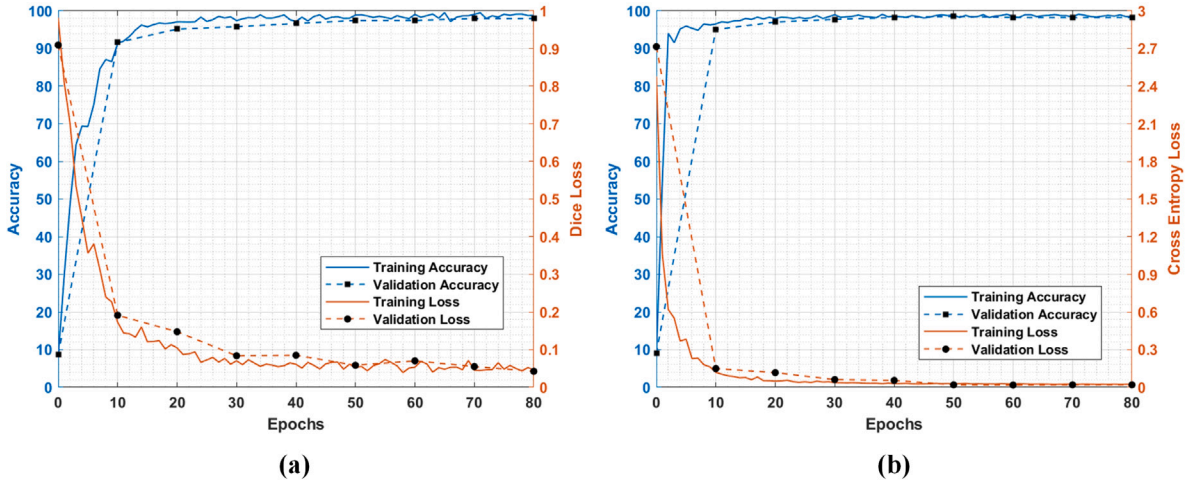

Fig. 9. CDC-Net training performance curves. (a) CDC-m1. (b) CDC-m2.

Table 8
Performance comparison of CDC-Net with SOTA algorithms in terms of lesion segmentation. The best and second-best results are indicated in bold and underline, respectively.

Lesions	FCN-32s [59]		FCN-8s [59]		U-net [60]		SegNet [61]		DeepLabv3+ [62]		CDC-Net	
	IoU	*DSC*	*IoU*	*DSC*	*IoU*	*DSC*	*IoU*	*DSC*	*IoU*	*DSC*	*IoU*	*DSC*
IRF	0.434	0.605	0.510	0.675	0.585	0.738	0.634	0.776	0.686	0.814	**0.693**	**0.819**
SRF	0.259	0.411	0.325	0.490	0.592	0.744	0.548	0.708	0.472	0.642	**0.631**	**0.774**
PED	0.576	0.731	0.699	0.823	0.762	0.865	**0.818**	**0.900**	0.790	0.883	0.789	0.882
RPD	0.335	0.502	0.506	0.672	0.523	0.687	0.618	0.764	**0.705**	**0.827**	0.673	0.805
Overall	0.401	0.562	0.510	0.665	0.616	0.758	0.654	0.787	0.663	0.791	**0.697**	**0.820**

$$ACC = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}, \tag{17}$$

where T_p, T_n, F_p, F_n denote the true positives, true negatives, false positives and false negatives, respectively. Moreover, the retinopathy diagnosis and grading are further assessed by plotting the receiver operating characteristic (ROC) curves and confusion matrices.

4. Results

CDC-Net is validated in terms of precise segmentation of lesions and lesion-assisted detection of retinal pathologies using four publicly available datasets. Besides, we compare the lesions segmentation performance of the CDC-Net with SOTA models like FCN [59], U-net [60], SegNet [61], DeepLabv3+ [62] where we incorporate the CDC-Net preprocessing mechanism in all the models to draw a fair comparison.

4.1. Evaluation for lesion segmentation

We first present the evaluation results for the segmentation of four retinal lesions investigated in this study. Table 8 compares the segmentation performance of CDC-Net with other SOTA frameworks using mean *IoU* and *DSC* metrics averaged over the complete test dataset. Here, we can observe that CDC-Net demonstrates better segmentation performance for IRF and SRF considering mean *DSC*, where it exceeds the second-best framework by a margin of 0.61% and 4.03%, respectively. Whereas for PED and RPD segmentation, CDC-Net lags behind the best results by 2.04% and 2.73%, respectively. In terms of overall segmentation performance across all the lesions, CDC-Net outperforms other frameworks and achieved 5.12% and 3.66% higher *IoU* and *DSC* scores in comparison to the second-best performing method [62].

The segmentation performance of all SOTA algorithms using randomly selected OCT scans from four datasets is shown in Fig. 10. Here, we can notice that CDC-Net's segmented retinal lesions are more precise in morphology and delineation, with fewer false positives and false negatives. Besides, CDC-Net better identified the finer lesions with similar texture in the scan, which are often excluded or falsely predicted by its competitors. This is because CDC-Net retains the contextual information in OCT scans while segmenting the retinal lesions, and the best features are merged during the encoding process.

Table 9
Comparison of retinal fluid segmentation on Duke-2 [29] dataset. Bold marks the best performance while the second-best performance is underlined.

Methods	Mean *IoU*			Mean *DSC*		
	Grader-1	Grader-2	Combined	Grader-1	Grader-2	Combined
[1]	0.285	0.373	0.329	0.444	0.543	0.495
[35]	–	–	0.478	–	–	0.647
[34]	–	–	0.404	0.624	0.527	0.575
[29]	–	–	0.361	–	–	0.530
CDC-Net	0.438	0.627	**0.532**	0.609	0.771	**0.690**
Inter-grader		0.408			0.581	

Next, we present the segmentation performance using the Duke-2 dataset exclusively. This dataset contains 610 OCT scans of 10 ME subjects along with the fluid (merged IRF and SRF) annotations provided by two separate graders. As the annotations provided by the organizers correspond to only one class (fluid), we merged the predicted retinal lesions of CDC-Net to draw the comparison. Table 9 shows the segmentation performance of CDC-Net in comparison to the other SOTA algorithms. Here considering the mean *DSC* metric, we can observe that CDC-Net outperforms the second-best method [35] by a neat margin of 6.64% and inter-grader agreement by 18.76%. It is pertinent to mention here that, unlike the remaining frameworks, CDC-Net is trained to segment four different retinal lesions, yet it achieved the best results, evidencing the efficacy of CDC-Net in correctly ruling out the non-existing lesions in the OCT scans.

Further, the segmentation performance of CDC-Net in comparison to other SOTA algorithms using randomly selected OCT scans from the Duke-2 dataset is shown in Fig. 11. The mean *IoU* and *DSC* values in Fig. 11(d–f) represent the average of segmentation scores achieved

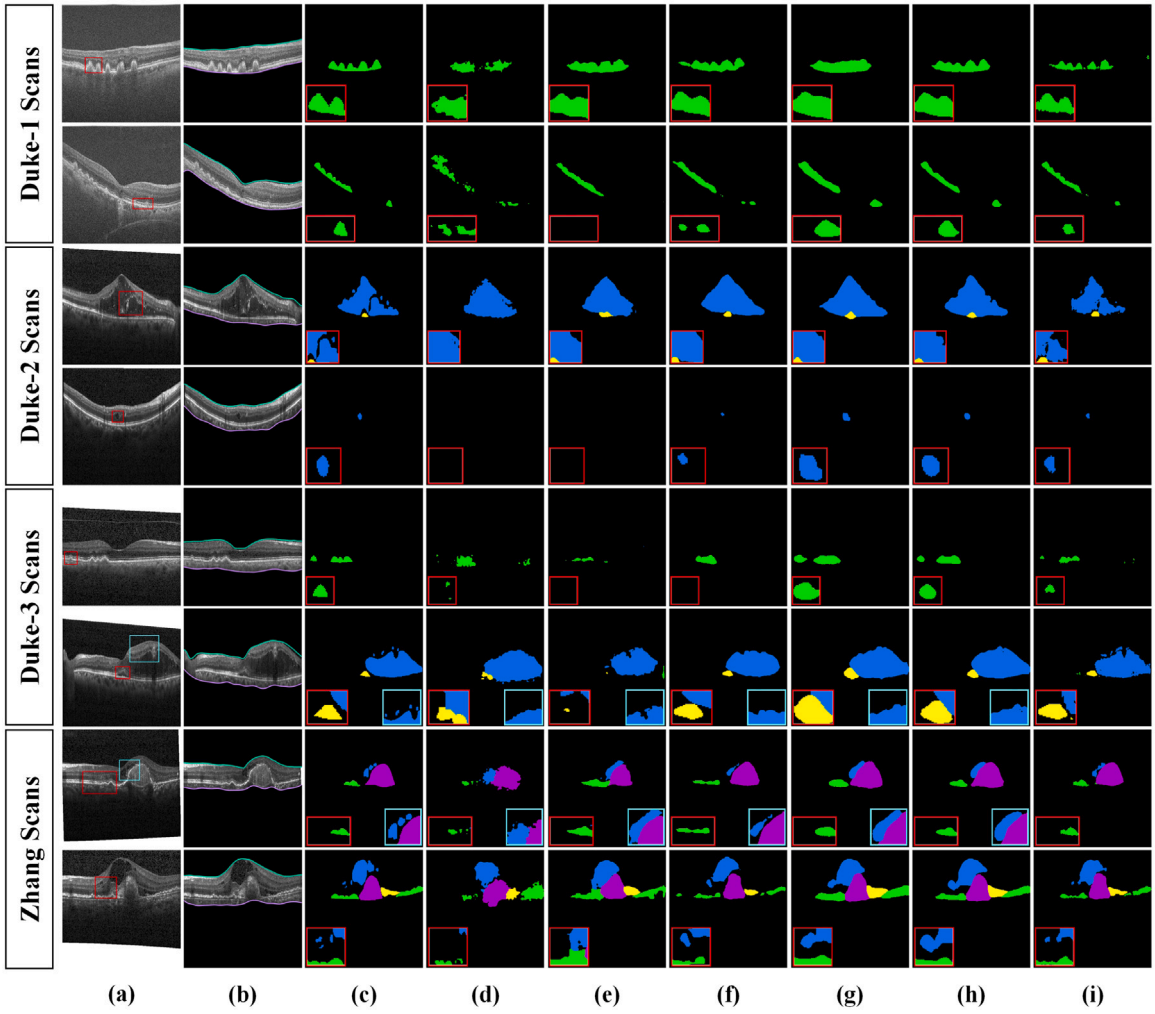

Fig. 10. Lesion segmentation performance comparison between CDC-Net and other SOTA solutions using four public datasets (Duke-1 [24], Duke-2 [29], Duke-3 [51], Zhang [52]). (a) Raw OCT scans. (b) Preprocessed scans. (c) Ground truth labels (IRF (blue), SRF (yellow), PED (magenta), RPD (green)). (d) FCN-32s [59]. (e) FCN-8s [59]. (f) U-net [60]. (g) SegNet [61]. (h) DeepLabv3+ [62]. (i) CDC-Net.

against the annotations of both the graders (as shown in Fig. 11(b–c)). Here, we can observe that CDC-Net consistently demonstrated superior performance in comparison to other methods (see Fig. 11(f)). Moreover, we demonstrate the results of our proposed framework separately with respect to the annotations from both the graders as shown in Fig. 12. Here we can observe that CDC-Net's performance is more precise against the grader two annotations (see Fig. 12(c)), containing lesser false positive (blue) and false negative pixels (red) in each case.

Furthermore, we compared the retinal lesion segmentation performance of the CDC-Net model and other SOTA frameworks with and without incorporating the preprocessing mechanism as shown in Table 10. Here, we can observe that each model's performance declined considerably using the raw scans (without preprocessing). This is due to the presence of different artifacts and vendor annotations in the original OCT scans, as shown in Fig. 3, which affects the network's segmentation performance and accuracy. On the other hand, by integrating an exclusive preprocessing stage as proposed in this study, a considerable performance gain (between 13.36% to 26.44%) is observed for the lesion segmentation task as shown in Table 10. This also justifies using a certain preprocessing mechanism to suppress the background contents in the OCT scans.

Table 10
Performance gain in lesion segmentation by incorporating preprocessing in all frameworks.

Methods	w/preprocessing		w/o preprocessing		Improvement (%)	
	IoU	*DSC*	*IoU*	*DSC*	*IoU*	*DSC*
FCN-32s [59]	0.401	0.562	0.286	0.444	40.34	26.44
FCN-8s [59]	0.510	0.665	0.371	0.542	37.34	22.79
U-net [60]	0.616	0.758	0.458	0.628	34.51	20.66
SegNet [61]	0.654	0.787	0.493	0.661	32.53	19.09
DeepLabv3+ [62]	0.663	0.791	0.521	0.685	27.31	15.50
CDC-Net	0.697	0.820	0.567	0.723	23.00	13.36

4.2. Evaluation for lesion detection

Next, we analyzed the performance of CDC-Net in detecting retinal lesions using a complete test dataset containing 26 841 multi-vendor OCT scans. First, in Table 11 we report the lesion detection performance of CDC-Net in comparison with other SOTA algorithms using the TPR, TNR, and ACC metrics. For each type of retinal lesion, these metrics are computed separately, and the scans that did not contain the relevant lesion were ignored in the calculation. Table 11 also depicts the mean value (averaged across the four retinal lesions) of each metric for all the frameworks. For the ACC metric, we can see that CDC-Net demonstrates superior detection performance for IRF and SRF lesions, where it exceeds the second-best method by a margin of 1.24% and 4.44%, respectively. Moreover, CDC-Net comes second in the accurate detection of RPD and third for accurate detection of PED, lagging behind the best method by a margin of 1.69% and 0.61%,

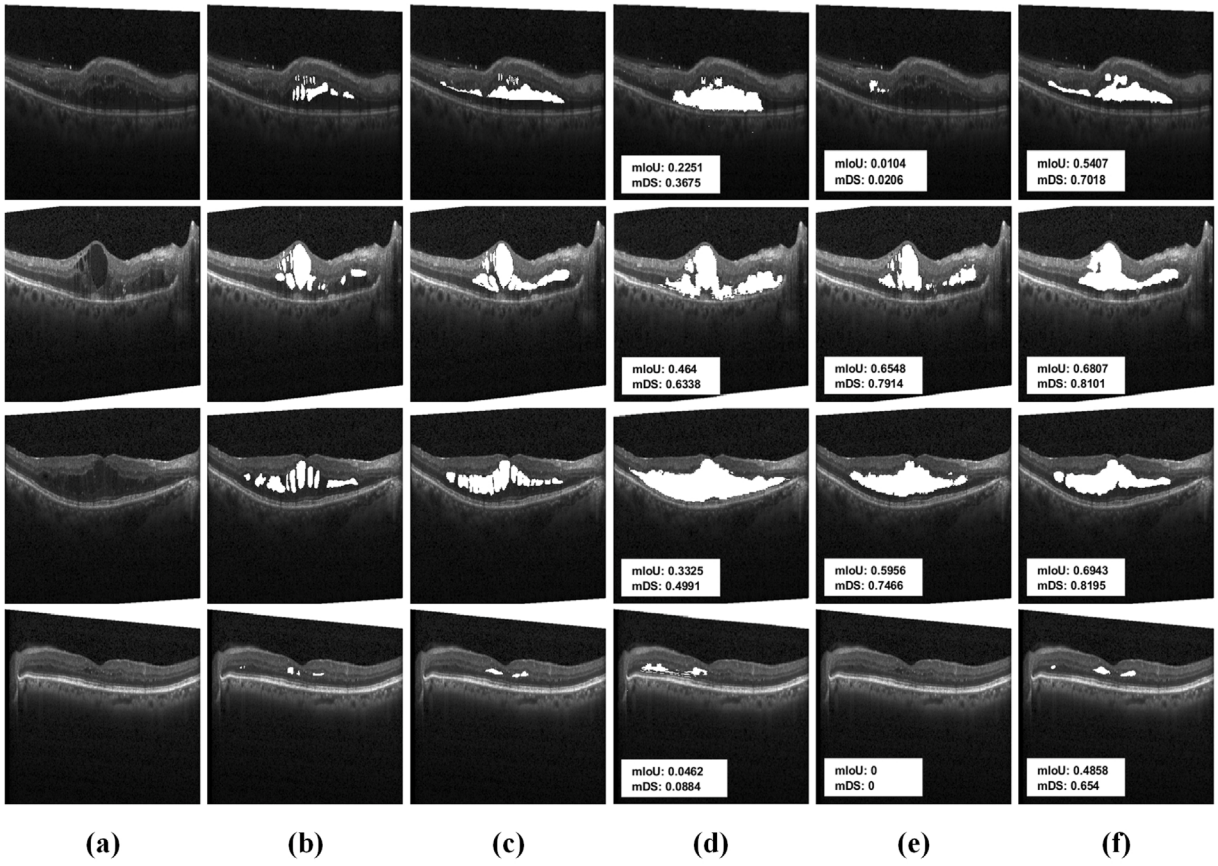

Fig. 11. Comparison of fluid segmentation performance on Duke-2 [29] dataset. (a) Raw OCT scans. Ground truth labels by (b) Grader-1 and (c) Grader-2. (d) Structure tensors method [1]. (e) Kernel regression method [29]. (f) CDC-Net (mIoU and mDS represent the mean values computed against the labels of both Graders).

Table 11
Performance comparison of CDC-Net with SOTA algorithms in terms of correct lesion detection in OCT scans. The best and second-best performance is indicated in bold and underline, respectively.

Lesions	FCN-32s [59]			FCN-8s [59]			U-net [60]			SegNet [61]			DeepLabv3+ [62]			CDC-Net		
	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR	ACC	TPR	TNR	ACC
IRF	0.690	0.727	0.724	0.787	0.829	0.826	0.872	0.923	0.919	0.894	0.952	0.947	0.947	0.968	0.967	**0.948**	**0.983**	**0.979**
SRF	0.628	0.662	0.660	0.768	0.790	0.788	0.907	0.923	0.922	0.883	0.904	0.902	0.827	0.855	0.852	**0.939**	**0.965**	**0.963**
PED	0.898	0.920	0.918	0.910	0.923	0.922	0.911	0.951	0.948	**0.968**	**0.989**	**0.987**	0.968	0.985	0.983	0.965	0.982	0.981
RPD	0.679	0.698	0.696	0.771	0.810	0.807	0.874	0.898	0.896	0.881	0.906	0.904	**0.913**	**0.965**	**0.961**	0.908	0.949	0.945
Mean	0.724	0.752	0.749	0.809	0.838	0.836	0.891	0.924	0.921	0.906	0.938	0.935	0.914	0.943	0.941	**0.940**	**0.969**	**0.967**

respectively. In terms of mean ACC for detection of all retinal lesions, CDC-Net outperforms the second-best method [62] by a margin of 2.76%. Similarly, the proposed framework outperforms the second-best results by achieving the mean TPR and TNR of 0.940 (2.84% high) and 0.969 (2.75% high), respectively.

In Fig. 13, we report some qualitative results to demonstrate the detection performance of CDC-Net using OCT scans of various retinal pathologies. Here, we can appreciate the efficacy of CDC-Net in precisely detecting and segmenting the retinal lesions. For instance, in the first row of Fig. 13, we can see that CDC-Net is able to detect even the minor IRF segments (such as pseudocyst) from ME-affected OCT scans. Similarly, in the second row of Fig. 13, we can observe the dry-AMD affected scans where the proposed framework can efficiently detect the RPD lesions. Moreover, CDC-Net's performance for detecting SRF and PED lesions from CNV affected scans can be visualized in the third and the last rows of Fig. 13.

To further evaluate the retinal lesion detection performance of CDC-Net, we altered the pixel classification threshold between 0 and 1, with a stride of 0.01 to plot the ROC curve for each lesion class as shown in Fig. 14(a). We can observe that the proposed framework achieved the minimum area under the curve (AUC) value of 0.975 for SRF. Nevertheless, considering that both SRF and IRF have highly similar intensity values and features in the OCT scan, CDC-Net's performance in distinguishing these regions is quite impressive. Furthermore, the AUC values higher than 0.975 suggest that CDC-Net has skillfully detected each lesion class's presence in the OCT scans.

4.3. Evaluation for retinopathy classification

In another series of experiments, we demonstrate the performance of CDC-Net in terms of retinopathy classification and grading. Table 12 presents the dataset-wise and combined performance of CDC-Net in lesion-assisted detection of retinal pathologies. Here, we can observe that CDC-Net performed consistently well in comparison to the other SOTA frameworks. Considering the ACC metric, we can see that CDC-Net outperforms the second-best framework by a margin of 0.16%, 1.88%, 4.65%, and 2.95% using the Duke-1, Duke-2, Duke-3, and Zhang datasets, respectively. Moreover, our proposed framework achieved 98.89% ACC with 98.34% TPR and 99.17% TNR in accurate retinopathy classification using combined datasets. Fig. 15(a) shows the

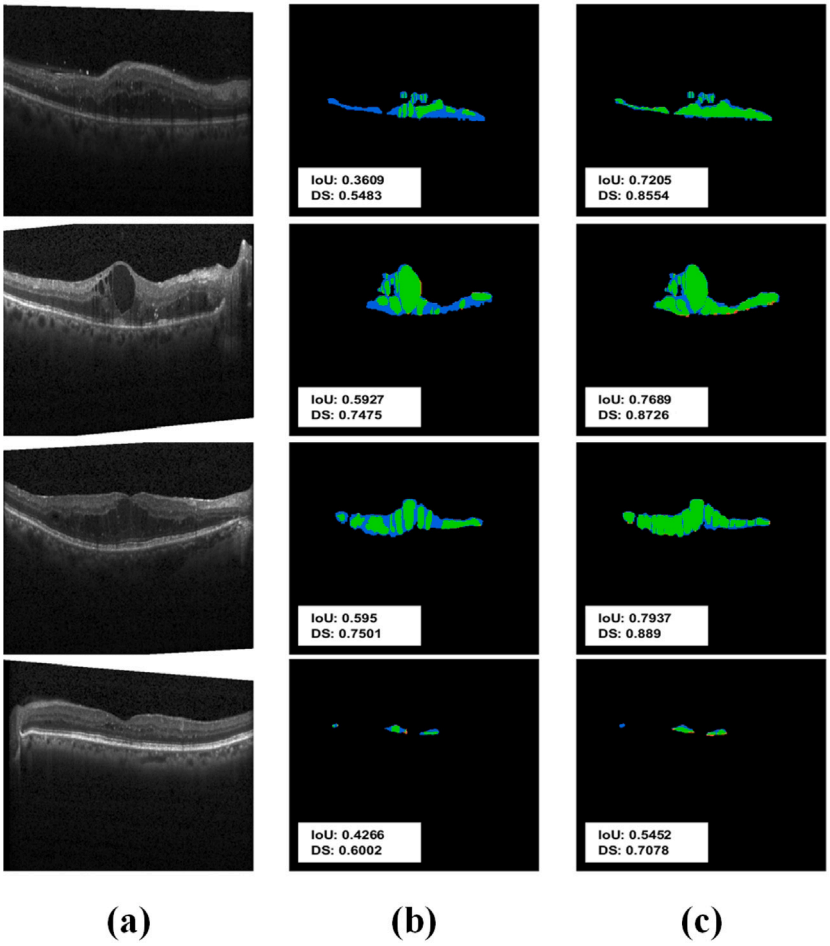

Fig. 12. Comparison of fluid segmentation performance on Duke-2 [29] dataset. (a) Raw OCT scans. Overlap of CDC-Net predicted labels with (b) Grader-1 and (c) Grader-2 annotations (green depicts the overlapped region, blue and red show the false positives and false negatives, respectively. IoU and DS mark the intersection over union and Dice score values).

confusion matrix of CDC-Net for classification of retinopathy, where we can observe that the proposed framework faced the most confusion in discriminating the healthy and AMD classes. Specifically, 128 healthy scans are misclassified as AMD, and 268 AMD scans are wrongly classified as healthy. On the other hand, the trained CDC-Net model produced 54 misclassifications for the ME class with 51 false negatives (17 predicted as healthy and 34 predicted as AMD) and three false positives (3 AMD scans predicted as ME).

Besides, we plotted the ROC curves as depicted in Fig. 14(b) for further validation, which shows that the minimum AUC score achieved by CDC-Net is 0.998 for classification of ME in comparison to other pathologies. Similarly, with an AUC score of 0.999, the best performance is observed for diagnosing healthy versus abnormal retinal pathologies (retinopathy). The obtained high AUC values (greater than 0.99) show that the proposed CDC-Net framework is highly insensitive to the classification thresholds, where the decisions are firmly based on the underlying segmented retinal lesions.

In addition to retinopathy classification, the proposed CDC-Net framework can also grade the candidate OCT scan according to clinical standards (ME into CME or SRF, AMD into dry or wet AMD). The confusion matrix of CDC-Net in terms of retinopathy grading is shown in Fig. 15(b), where we can observe that the proposed framework achieved an overall TPR of 98.20% in precise grading of OCT scans. Furthermore, we performed a separate experiment to evaluate the CDC-Net model for direct grading of retinal pathologies without incorporating the lesion-assisted mechanism. The confusion matrix depicting the direct grading by the CDC-Net model is shown in Fig. 15(c), where we can observe a decline of 11.70% in performance due to frequent misclassifications between SRD and CME, as well as dry and wet AMD affected OCT scans. These findings indicate that retinal lesions' prior segmentation holds significant importance in the accurate grading of retinopathy.

Table 12
Comparison of retinopathy grading on all datasets. The best and second-best results are highlighted in bold and underline, respectively. The results presented are in percentages.

Datasets	Metrics	CDC-Net	[41]	[42]	[1]	[51]	[52]	[43]
Duke-1 [24]	TPR	97.93	98.22	**99.26**	–	–	–	–
	TNR	**98.82**	90.43	95.65	–	–	–	–
	ACC	**98.33**	95.88	98.17	–	–	–	–
Duke-2 [29]	TPR	**96.53**	–	–	95.46	–	–	–
	TNR	**98.23**	–	–	96.17	–	–	–
	ACC	**97.80**	–	–	95.99	–	–	–
Duke-3 [51]	TPR	**100**	–	–	–	**100**	–	98.80
	TNR	**100**	–	–	–	86.67	–	–
	ACC	**100**	–	–	–	95.55	–	96.00
Zhang [52]	TPR	**98.90**	–	–	–	–	97.80	94.50
	TNR	**99.63**	–	–	–	–	97.40	–
	ACC	**99.45**	–	–	–	–	96.60	96.00
Combined	TPR	**98.34**	–	–	–	–	–	–
	TNR	**99.17**	–	–	–	–	–	–
	ACC	**98.89**	–	–	–	–	–	–

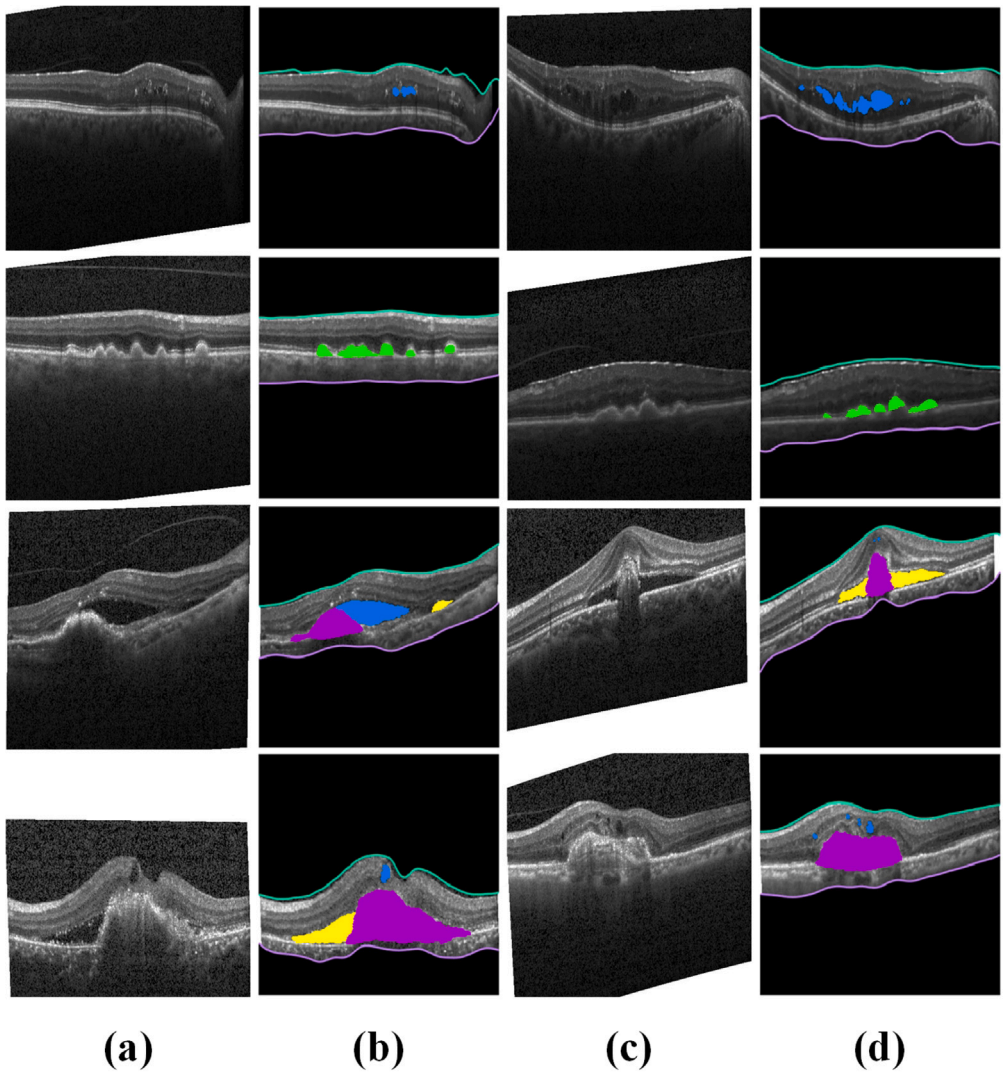

Fig. 13. (a,c) Raw OCT scans. (b,d) Visualization of lesions detected by the CDC-Net model.

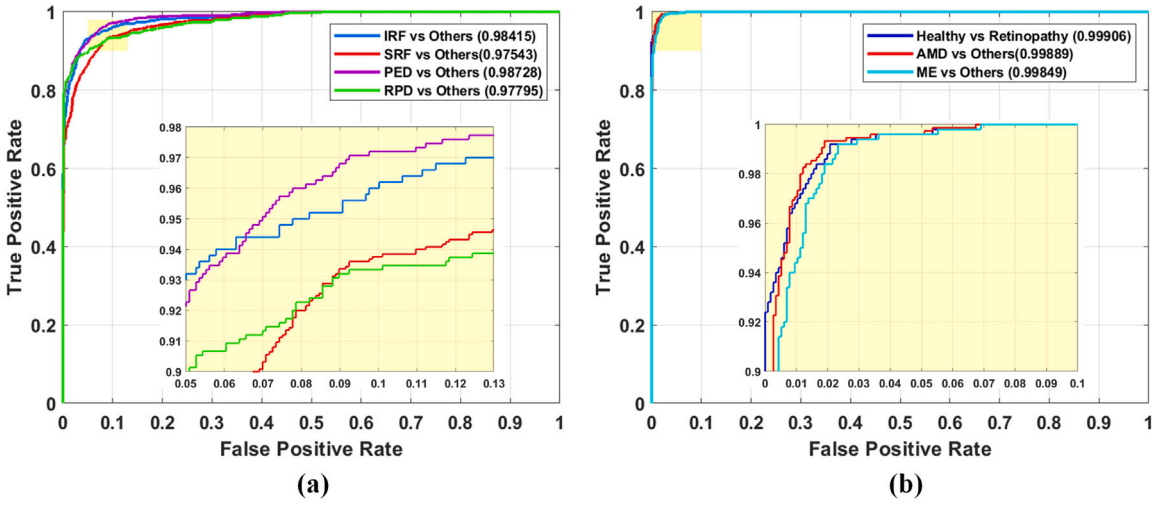

Fig. 14. CDC-Net ROC performance curves: (a) Detection of retinal lesions by CDC-m1. (b) Classification of retinopathy by CDC-m2.

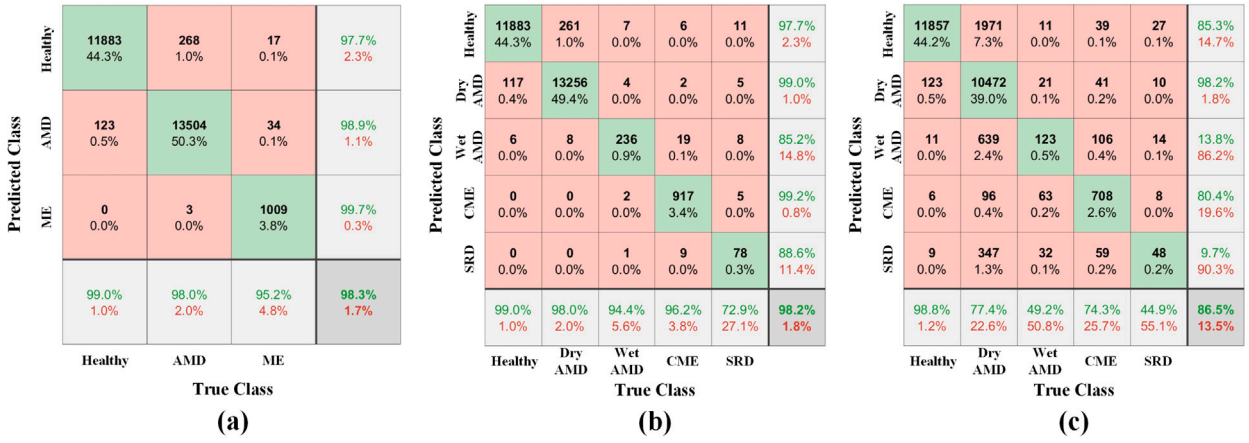

Fig. 15. Confusion matrix showcasing performance of the CDC-Net framework on complete test dataset. (a) Classification of retinopathy. (b) Lesion-assisted grading of retinopathy. (c) Direct grading without incorporating the segmented lesions information.

5. Discussion

This paper proposes a novel hybrid architecture called CDC-Net, which integrates two decoupled convolutional networks that jointly work towards the lesion-assisted classification and grading of retinopathy using multi-vendor OCT scans. Retinal lesions play a crucial role in diagnosing and evaluating the severity of underlying ocular syndromes. To the best of our knowledge, the proposed framework is the first of its kind to segment four different retinal lesions (IRF, SRF, RPD, and PED) and to utilize the segmented lesions for the grading of retinopathy. Besides, CDC-Net is thoroughly evaluated using 26 841 OCT scans over various experiments, and the performance is compared with existing SOTA solutions in terms of different evaluation metrics.

Previously researchers have proposed solutions for automated segmentation of retinal lesions such as IRF and SRF [11,13,14], but they did not use these biomarkers to classify and grade the candidate retina. Lesion-aware CNN framework to classify the retinopathy is previously proposed in [15], but their method lacks generalizability as it was not evaluated on multi-vendor OCT scans. Also, the authors in [15] did not grade the severity of the disease based on the segmented retina lesions. In contrast, the proposed framework's performance is assessed against various datasets separately, where the demonstrated results validate that CDC-Net consistently outperformed various existing SOTA solutions.

In addition, the preprocessing mechanism introduced in this study greatly enhances the lesion segmentation performance of CDC-Net and pre-trained models [59–62], as shown in Table 10. This is because the raw OCT scans contain various unwanted artifacts, as shown in Fig. 3. Also, OCT scans are usually in grayscale, where the intensities and background texture (vitreous and sclera regions) appear similar to certain lesions (such as IRF and SRF). Therefore, it is important to crop the retinal region before feeding the segmentation network, and without preprocessing, any models can easily misclassify these regions. We also molded the CDC-Net framework for the direct grading of retinopathy instead of incorporating the lesion-assisted mechanism, as shown in Fig. 15(c). However, as shown in the figure, this leads to an increase in the number of misclassifications between healthy and dry-AMD, SRD and CME, and dry-AMD and wet-AMD. Although the direct retinopathy grading scheme is computationally efficient, its performance is far inferior to the lesion-assisted method. This indicates the significance of retinal lesions in the classification and management of ocular syndromes with different degrees of progression.

6. Conclusion

This paper proposes an automated framework dubbed CDC-Net for segmenting the retinal lesions and retinopathy grading using OCT scans. Besides, CDC-Net provides in-depth and intuitive retinopathy grading attributed to the segmented lesions. CDC-Net can be regarded as a generic model that overlooks the noisy artifacts and variations in multi-vendor OCT scans. The CDC-Net performance is validated on 26 841 scans from four publicly accessible datasets and contrasted to several existing SOTA solutions using various evaluation metrics. CDC-Net achieved mean IoU and DSC scores of 0.697 and 0.820 for segmenting the retinal lesions, and accuracy of 98.89% with 98.34% TPR and 99.17% TNR for grading retinopathy. The proposed CDC-Net framework is currently limited to the classification and grading of ME and AMD retinal pathologies. However, it could be extended in future work to classify more pathologies (such as CSR, glaucoma) and also provide their severity analysis. Furthermore, exploring the correlation of other imaging modalities, such as clinical findings in fundus photography, can further enhance the intuitive grading of retinopathy syndrome.

CRediT authorship contribution statement

Bilal Hassan: Conceptualization, Methodology, Software, Data curation, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Shiyin Qin:** Supervision, Writing - review & editing, Project administration. **Taimur Hassan:** Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Muhammad Usman Akram:** Supervision, Writing - review & editing, Project administration. **Ramsha Ahmed:** Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Naoufel Werghi:** Supervision, Writing - review & editing, Project administration.

Declaration of competing interest

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

Acknowledgments

We are immensely grateful to our three clinical collaborators for providing the OCT annotations used in this study. We are also thankful to the authors in [24,29,51,52] for making the datasets publicly available, which enabled us to conduct this research.

References

[1] B. Hassan, G. Raja, T. Hassan, M.U. Akram, Structure tensor based automated detection of macular edema and central serous retinopathy using optical coherence tomography images, J. Opt. Soc. Amer. A 33 (4) (2016) 455–463, http://dx.doi.org/10.1364/JOSAA.33.000455.

[2] C. Sabanayagam, R. Banu, M.L. Chee, R. Lee, Y.X. Wang, G. Tan, et al., Incidence and progression of diabetic retinopathy: a systematic review, Lancet Diabetes Endocrinol. 7 (2) (2019) 140–149, http://dx.doi.org/10.1016/S2213-8587(18)30128-1.

[3] B. Hassan, R. Ahmed, B. Li, A. Noor, Z.U. Hassan, A comprehensive study capturing vision loss burden in Pakistan (1990-2025): Findings from the global burden of disease (GBD) 2017 study, PLoS One 14 (5) (2019) e0216492, http://dx.doi.org/10.1371/journal.pone.0216492.

[4] S. Das, K. Kharbanda, M. Suchetha, R. Raman, E. Dhas, Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy, Biomed. Signal Process. Control 68 (2021) 102600, http://dx.doi.org/10.1016/j.bspc.2021.102600.

[5] T. Hassan, M.U. Akram, N. Werghi, Exploiting the transferability of deep learning systems across multi-modal retinal scans for extracting retinopathy lesions, in: 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2020, pp. 577–581, http://dx.doi.org/10.1109/BIBE50027.2020.00099.

[6] T. Hassan, M.U. Akram, B. Hassan, A.M. Syed, S.A. Bazaz, Automated segmentation of subretinal layers for the detection of macular edema, Appl. Opt. 55 (3) (2016) 454–461, http://dx.doi.org/10.1364/AO.55.000454.

[7] B. Hassan, S. Qin, R. Ahmed, Seadnet: Deep learning driven segmentation and extraction of macular fluids in 3D retinal OCT scans, in: 2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), IEEE, 2020, pp. 1–6, http://dx.doi.org/10.1109/ISSPIT51521.2020.9408988.

[8] S. Kaymak, A. Serener, Automated age-related macular degeneration and diabetic macular edema detection on oct images using deep learning, in: 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP), IEEE, 2018, pp. 265–269, http://dx.doi.org/10.1109/ICCP.2018.8516635.

[9] A.-R.E.D.S.R. Group, et al., The age-related eye disease study (AREDS): design implications areds report no. 1, Control. Clin. Trials 20 (6) (1999) 573–600, http://dx.doi.org/10.1016/s0197-2456(99)00031-8.

[10] T. Hassan, M.U. Akram, B. Hassan, A. Nasim, S.A. Bazaz, Review of OCT and fundus images for detection of macular edema, in: 2015 IEEE International Conference on Imaging Systems and Techniques (IST), IEEE, 2015, pp. 1–4, http://dx.doi.org/10.1109/IST.2015.7294517.

[11] B. Hassan, T. Hassan, B. Li, R. Ahmed, O. Hassan, Deep ensemble learning based objective grading of macular edema by extracting clinically significant findings from fused retinal imaging modalities, Sensors 19 (13) (2019) 2970, http://dx.doi.org/10.3390/s19132970.

[12] D.R. Nayak, D. Das, B. Majhi, S.V. Bhandary, U.R. Acharya, Ecnet: An evolutionary convolutional network for automated glaucoma detection using fundus images, Biomed. Signal Process. Control 67 (2021) 102559, http://dx.doi.org/10.1016/j.bspc.2021.102559.

[13] B. Hassan, R. Ahmed, B. Li, O. Hassan, T. Hassan, Automated retinal edema detection from fundus and optical coherence tomography scans, in: 2019 5th International Conference on Control, Automation and Robotics (ICCAR), IEEE, 2019, pp. 325–330, http://dx.doi.org/10.1109/ICCAR.2019.8813311.

[14] A. Thomas, P. Harikrishnan, A.K. Krishna, P. Palanisamy, V.P. Gopi, A novel multiscale convolutional neural network based age-related macular degeneration detection using OCT images, Biomed. Signal Process. Control 67 (2021) 102538, http://dx.doi.org/10.1016/j.bspc.2021.102538.

[15] V. Das, S. Dandapat, P.K. Bora, Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images, Biomed. Signal Process. Control 54 (2019) 101605, http://dx.doi.org/10.1016/j.bspc.2019.101605.

[16] J. Xie, S. Chen, N. Wang, L. Wang, E. Bo, L. Liu, Automatic differentiation of nonkeratinized stratified squamous epithelia and columnar epithelia through feature structure extraction using OCT, Biomed. Signal Process. Control 60 (2020) 101919, http://dx.doi.org/10.1016/j.bspc.2020.101919.

[17] T. Hassan, M.U. Akram, A. Shaukat, S.G. Khawaja, B. Hassan, Structure tensor graph searches based fully automated grading and 3D profiling of maculopathy from retinal OCT images, IEEE Access 6 (2018) 44644–44658, http://dx.doi.org/10.1109/ACCESS.2018.2862626.

[18] B. Hassan, T. Hassan, Fully automated detection, grading and 3D modeling of maculopathy from OCT volumes, in: 2019 2nd International Conference on Communication, Computing and Digital Systems (C-CODE), IEEE, 2019, pp. 252–257, http://dx.doi.org/10.1109/C-CODE.2019.8680996.

[19] T. Otani, S. Kishi, Y. Maruyama, Patterns of diabetic macular edema with optical coherence tomography, Am. J. Ophthalmol. 127 (6) (1999) 688–693, http://dx.doi.org/10.1016/S0002-9394(99)00033-1.

[20] B. Hassan, S. Qin, R. Ahmed, T. Hassan, A.H. Taguri, S. Hashmi, N. Werghi, Deep learning based joint segmentation and characterization of multi-class retinal fluid lesions on OCT scans for clinical use in anti-vegf therapy, Comput. Biol. Med. (2021).

[21] J. de Moura, G. Samagaio, J. Novo, P. Almuina, M.I. Fernández, M. Ortega, Joint diabetic macular edema segmentation and characterization in oct images, J. Digit. Imaging 33 (5) (2020) 1335–1351, http://dx.doi.org/10.1007/s10278-020-00360-y.

[22] B.L. Sikorski, G. Malukiewicz, J. Stafiej, H. Lesiewska-Junk, D. Raczynska, The diagnostic function of OCT in diabetic maculopathy, Mediat. Inflamm. 2013 (2013) http://dx.doi.org/10.1155/2013/434560.

[23] M. Dysli, R. Rückert, M.R. Munk, Differentiation of underlying pathologies of macular edema using spectral domain optical coherence tomography (SD-OCT), Ocul. Immunol. Inflamm. 27 (3) (2019) 474–483, http://dx.doi.org/10.1080/09273948.2019.1603313.

[24] S. Farsiu, S.J. Chiu, R.V. O'Connell, F.A. Folgar, E. Yuan, J.A. Izatt, C.A. Toth, et al., Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography, Ophthalmology 121 (1) (2014) 162–172, http://dx.doi.org/10.1016/j.ophtha.2013.07.013.

[25] N.F. Mokwa, T. Ristau, P.A. Keane, B. Kirchhof, S.R. Sadda, S. Liakopoulos, Grading of age-related macular degeneration: comparison between color fundus photography, fluorescein angiography, and spectral domain optical coherence tomography, J. Ophthalmol. 2013 (2013) 385915, http://dx.doi.org/10.1155/2013/385915.

[26] S.J. Chiu, X.T. Li, P. Nicholas, C.A. Toth, J.A. Izatt, S. Farsiu, Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation, Opt. Express 18 (18) (2010) 19413–19428, http://dx.doi.org/10.1364/OE.18.019413.

[27] Y. Wiseman, E. Fredj, Contour extraction of compressed JPEG images, J. Graph. Tools 6 (3) (2001) 37–43, http://dx.doi.org/10.1080/10867651.2001.10487544.

[28] Y. Huang, R.P. Danis, J.W. Pak, S. Luo, J. White, X. Zhang, A. Narkar, A. Domalpally, Development of a semi-automatic segmentation method for retinal oct images tested in patients with diabetic macular edema, PLoS One 8 (12) (2013) e82922, http://dx.doi.org/10.1371/journal.pone.0082922.

[29] S.J. Chiu, M.J. Allingham, P.S. Mettu, S.W. Cousins, J.A. Izatt, S. Farsiu, Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema, Biomed. Opt. Express 6 (4) (2015) 1172–1194, http://dx.doi.org/10.1364/BOE.6.001172.

[30] K.J. McHugh, D. Li, J.C. Wang, L. Kwark, J. Loo, V. Macha, S. Farsiu, L.A. Kim, M. Saint-Geniez, Computational modeling of retinal hypoxia and photoreceptor degeneration in patients with age-related macular degeneration, PLoS One 14 (6) (2019) e0216215, http://dx.doi.org/10.1371/journal.pone.0216215.

[31] D. Xiang, G. Chen, F. Shi, W. Zhu, Q. Liu, S. Yuan, X. Chen, Automatic retinal layer segmentation of OCT images with central serous retinopathy, IEEE J. Biomed. Health Inf. 23 (1) (2018) 283–295, http://dx.doi.org/10.1109/JBHI.2018.2803063.

[32] D. Kaba, Y. Wang, C. Wang, X. Liu, H. Zhu, A. Salazar-Gonzalez, Y. Li, Retina layer segmentation using kernel graph cuts and continuous max-flow, Opt. Express 23 (6) (2015) 7366–7384, http://dx.doi.org/10.1364/OE.23.007366.

[33] P.L. Vidal, J. De Moura, J. Novo, M.G. Penedo, M. Ortega, Intraretinal fluid identification via enhanced maps using optical coherence tomography images, Biomed. Opt. Express 9 (10) (2018) 4730–4754, http://dx.doi.org/10.1364/BOE.9.004730.

[34] A. Rashno, D.D. Koozekanani, P.M. Drayna, B. Nazari, S. Sadri, H. Rabbani, K.K. Parhi, Fully automated segmentation of fluid/cyst regions in optical coherence tomography images with diabetic macular edema using neutrosophic sets and graph algorithms, IEEE Trans. Biomed. Eng. 65 (5) (2017) 989–1001, http://dx.doi.org/10.1109/TBME.2017.2734058.

[35] T. Hassan, M.U. Akram, M.F. Masood, U. Yasin, Deep structure tensor graph search framework for automated extraction and characterization of retinal layers and fluid pathology in retinal SD-oct scans, Comput. Biol. Med. 105 (2019) 112–124, http://dx.doi.org/10.1016/j.compbiomed.2018.12.015.

[36] L. Fang, D. Cunefare, C. Wang, R.H. Guymer, S. Li, S. Farsiu, Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search, Biomed. Opt. Express 8 (5) (2017) 2732–2744, http://dx.doi.org/10.1364/BOE.8.002732.

[37] C.S. Lee, A.J. Tyring, N.P. Deruyter, Y. Wu, A. Rokem, A.Y. Lee, Deep-learning based, automated segmentation of macular edema in optical coherence tomography, Biomed. Opt. Express 8 (7) (2017) 3440–3448, http://dx.doi.org/10.1364/BOE.8.003440.

[38] A.G. Roy, S. Conjeti, S.P.K. Karri, D. Sheet, A. Katouzian, C. Wachinger, N. Navab, Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks, Biomed. Opt. Express 8 (8) (2017) 3627–3642, http://dx.doi.org/10.1364/BOE.8.003627.

[39] T. Schlegl, S.M. Waldstein, H. Bogunovic, F. Endstraßer, A. Sadeghipour, A.-M. Philip, D. Podkowinski, B.S. Gerendas, G. Langs, U. Schmidt-Erfurth, Fully automated detection and quantification of macular fluid in oct using deep learning, Ophthalmology 125 (4) (2018) 549–558, http://dx.doi.org/10.1016/j.ophtha.2017.10.031.

[40] G. Girish, B. Thakur, S.R. Chowdhury, A.R. Kothari, J. Rajan, Segmentation of intra-retinal cysts from optical coherence tomography images using a fully convolutional neural network model, IEEE J. Biomed. Health Inf. 23 (1) (2018) 296–304, http://dx.doi.org/10.1109/JBHI.2018.2810379.

[41] J. Qiu, Y. Sun, Self-supervised iterative refinement learning for macular oct volumetric data classification, Comput. Biol. Med. 111 (2019) 103327, http://dx.doi.org/10.1016/j.compbiomed.2019.103327.

[42] Y. Sun, H. Zhang, X. Yao, Automatic diagnosis of macular diseases from OCT volume based on its two-dimensional feature map and convolutional neural network with attention mechanism, J. Biomed. Opt. 25 (9) (2020) 096004, http://dx.doi.org/10.1117/1.JBO.25.9.096004.

[43] A. Butola, D.K. Prasad, A. Ahmad, V. Dubey, D. Qaiser, A. Srivastava, P. Senthilkumaran, B.S. Ahluwalia, D.S. Mehta, Deep learning architecture "lightoct" for diagnostic decision support using optical coherence tomography images of biological samples, Biomed. Opt. Express 11 (9) (2020) 5017–5031, http://dx.doi.org/10.1364/BOE.395487.

[44] B. Hassan, G. Raja, Fully automated assessment of macular edema using optical coherence tomography (OCT) images, in: 2016 International Conference on Intelligent Systems Engineering (ICISE), IEEE, 2016, pp. 5–9, http://dx.doi.org/10.1109/INTELSE.2016.7475153.

[45] R.M. Kamble, G.C. Chan, O. Perdomo, M. Kokare, F.A. Gonzalez, H. Müller, F. Mériaudeau, Automated diabetic macular edema (DME) analysis using fine tuning with inception-resnet-v2 on OCT images, in: 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, 2018, pp. 442–446, http://dx.doi.org/10.1109/IECBES.2018.8626616.

[46] S. Kuwayama, Y. Ayatsuka, D. Yanagisono, T. Uta, H. Usui, A. Kato, N. Takase, Y. Ogura, T. Yasukawa, Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images, J. Ophthalmol. 2019 (2019) 6319581, http://dx.doi.org/10.1155/2019/6319581.

[47] B. Hassan, S. Qin, R. Ahmed, RRI-Net: Classification of multi-class retinal diseases with deep recurrent residual inception network using OCT scans, in: 2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), IEEE, 2020, pp. 1–6, http://dx.doi.org/10.1109/ISSPIT51521.2020.9408820.

[48] F. Li, H. Chen, Z. Liu, X. Zhang, Z. Wu, Fully automated detection of retinal disorders by image-based deep learning, Graefe's Arch. Clin. Exp. Ophthalmol. 257 (3) (2019) 495–505, http://dx.doi.org/10.1007/s00417-018-04224-8.

[49] L. Zhang, W. Zhu, F. Shi, H. Chen, X. Chen, Automated segmentation of intraretinal cystoid macular edema for retinal 3D OCT images with macular hole, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), IEEE, 2015, pp. 1494–1497, http://dx.doi.org/10.1109/ISBI.2015.7164160.

[50] J. Sugmk, S. Kiattisin, A. Leelasantitham, Automated classification between age-related macular degeneration and diabetic macular edema in OCT image using image segmentation, in: The 7th 2014 Biomedical Engineering International Conference, IEEE, 2014, pp. 1–4, http://dx.doi.org/10.1109/BMEiCON.2014.7017441.

[51] P.P. Srinivasan, L.A. Kim, P.S. Mettu, S.W. Cousins, G.M. Comer, J.A. Izatt, S. Farsiu, Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images, Biomed. Opt. Express 5 (10) (2014) 3568–3577, http://dx.doi.org/10.1364/BOE.5.003568.

[52] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (5) (2018) 1122–1131, http://dx.doi.org/10.1016/j.cell.2018.02.010.

[53] Y. Rong, D. Xiang, W. Zhu, K. Yu, F. Shi, Z. Fan, X. Chen, Surrogate-assisted retinal OCT image classification based on convolutional neural networks, IEEE J. Biomed. Health Inf. 23 (1) (2018) 253–263, http://dx.doi.org/10.1109/JBHI.2018.2795545.

[54] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2, IEEE, 2005, pp. 60–65, http://dx.doi.org/10.1109/CVPR.2005.38.

[55] B. Hassan, S. Qin, T. Hassan, R. Ahmed, N. Werghi, Joint segmentation and quantification of chorioretinal biomarkers in optical coherence tomography scans: A deep learning approach, IEEE Trans. Instrum. Meas. 70 (2021) 1–17, http://dx.doi.org/10.1109/TIM.2021.3077988.

[56] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015, pp. 1–9, http://dx.doi.org/10.1109/CVPR.2015.7298594.

[57] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2016, pp. 770–778, http://dx.doi.org/10.1109/CVPR.2016.90.

[58] Y. Zhao, A. Gafita, B. Vollnberg, G. Tetteh, F. Haupt, A. Afshar-Oromieh, et al., Deep neural network for automatic characterization of lesions on 68 ga-PSMA-11 PET/CT, Eur. J. Nucl. Med. Mol. Imaging 47 (3) (2020) 603–613, http://dx.doi.org/10.1007/s00259-019-04606-y.

[59] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015, pp. 3431–3440, http://dx.doi.org/10.1109/CVPR.2015.7298965.

[60] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241, http://dx.doi.org/10.1007/978-3-319-24574-4_28.

[61] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495, http://dx.doi.org/10.1109/TPAMI.2016.2644615.

[62] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), Springer, 2018, pp. 801–818, http://dx.doi.org/10.1007/978-3-030-01234-2_49.