MDPI

Systematic Review

Denoising of Optical Coherence Tomography Images in Ophthalmology Using Deep Learning: A Systematic Review

Hanya Ahmed [1,*,†], Qianni Zhang [1,†], Robert Donnan [2,†] and Akram Alomainy [1,†]

1   Department of Electronic Engineering and Computer Science, Queen Mary University of London,
    London E1 4NS, UK
2   Department of Engineering and Materials Science, Queen Mary University of London, London E1 4NS, UK
*   Correspondence: h.t.ahmed@qmul.ac.uk
†   These authors contributed equally to this work.

Abstract: Imaging from optical coherence tomography (OCT) is widely used for detecting retinal diseases, localization of intra-retinal boundaries, etc. It is, however, degraded by speckle noise. Deep learning models can aid with denoising, allowing clinicians to clearly diagnose retinal diseases. Deep learning models can be considered as an end-to-end framework. We selected denoising studies that used deep learning models with retinal OCT imagery. Each study was quality-assessed through image quality metrics (including the peak signal-to-noise ratio—*PSNR*, contrast-to-noise ratio—*CNR*, and structural similarity index metric—*SSIM*). Meta-analysis could not be performed due to heterogeneity in the methods of the studies and measurements of their performance. Multiple databases (including Medline via PubMed, Google Scholar, Scopus, Embase) and a repository (ArXiv) were screened for publications published after 2010, without any limitation on language. From the 95 potential studies identified, a total of 41 were evaluated thoroughly. Fifty-four of these studies were excluded after full text assessment depending on whether deep learning (DL) was utilized or the dataset and results were not effectively explained. Numerous types of OCT images are mentioned in this review consisting of public retinal image datasets utilized purposefully for denoising OCT images ($n = 37$) and the Optic Nerve Head (ONH) ($n = 4$). A wide range of image quality metrics was used; *PSNR* and *SNR* that ranged between 8 and 156 dB. The minority of studies ($n = 8$) showed a low risk of bias in all domains. Studies utilizing ONH images produced either a *PSNR* or *SNR* value varying from 8.1 to 25.7 dB, and that of public retinal datasets was 26.4 to 158.6 dB. Further analysis on denoising models was not possible due to discrepancies in reporting that did not allow useful pooling. An increasing number of studies have investigated denoising retinal OCT images using deep learning, with a range of architectures being implemented. The reported increase in image quality metrics seems promising, while study and reporting quality are currently low.

Keywords: deep learning; ophthalmology; image processing; optical coherence tomography

1. Introduction

Optical coherence tomography (OCT) stands at the forefront of modern medical imaging techniques, harnessing the power of low-coherence infrared light to delve deep into biological structures with unprecedented clarity and precision [1]. This revolutionary technology affords longer exposure times due to its inherent biological safety, presenting a stark departure from the ionizing radiation associated with conventional X-rays. Moreover, when juxtaposed with Magnetic Resonance Imaging (MRI) and Computerized Tomography (CT), OCT emerges as a cost-effective alternative, democratizing access to high-quality diagnostic imaging. However, amidst the brilliance of OCT lies a challenge inherent to all imaging modalities—noise. Inevitably introduced during the imaging process, noise mingles with the signal emanating from the object under scrutiny, influencing the resultant intensity observed by the detecting pixel [2]. Of particular concern is speckle noise,

a byproduct of low coherence in irradiance, which casts a shadow over the signal-to-noise ratio, obscuring critical details within the imagery. Within the realm of ophthalmology, OCT serves as a versatile tool, facilitating the acquisition of cross-sectional and volumetric images that illuminate the intricate landscape of biological tissues and retinal structures. These images serve as invaluable aids in diagnosing a myriad of ocular diseases, ranging from diabetic retinopathy (DR) [3] to age-related macular degeneration (AMD) [4], guiding clinicians toward tailored treatment strategies.

Meanwhile, in the realm of computer science and Artificial Intelligence (AI), Machine Learning (ML) emerges as a beacon of innovation. By mining insights from past data, ML algorithms prognosticate future trends, obviating the need for explicit programming or human intervention [5]. At its core, ML epitomizes the essence of pattern recognition, endowing computers with the capacity to glean insights from vast datasets with unparalleled efficiency.

Within the ML landscape, deep learning (DL) commands center stage, propelled by leaps and bounds in computational prowess and the proliferation of "big data". Convolutional Neural Networks (CNNs) epitomize this evolution, revolutionizing DL with their ability to extract features, classify images, and recognize patterns at breakneck speeds. By emulating the intricate workings of the human brain through the deployment of filters and intricate layers, CNNs herald a new era of computational efficiency.

The intersection of DL and OCT heralds a realm of boundless possibilities, marked by advancements in volumetric data handling, heightened sensitivity, and specificity in detecting structural alterations and the tantalizing prospect of denoising retinal images to unprecedented levels of clarity. This synthesis prompts a critical examination, as we delve into recent studies illuminating the applications of DL to OCT imagery, evaluating their impact on image quality assessment.

Moreover, as we stand on the precipice of a new era in medical imaging and computational innovation, it behooves us to explore the clinical ramifications of integrating these cutting-edge computational techniques into the fabric of healthcare delivery. Through a judicious examination of recent computational innovations and their potential clinical applications, we chart a course toward enhanced diagnostic accuracy, streamlined treatment pathways, and, ultimately, improved patient outcomes. Thus, this review not only serves as a testament to the symbiotic relationship between technology and healthcare but also as a compass guiding future research endeavors and clinical initiatives.

2. Overview of Optical Coherence Tomography

Optical coherence tomography (OCT) stands as a maturing imaging technology, offering resolution ranging from millimeters to sub-millimeters and boasting a penetration depth comparable to that achieved in human skin [6]. This innovative technique predominantly employs low-coherence infrared light to safely delve into biological tissues, affording longer exposure times in contrast to X-rays. Figure 1 illustrates the conceptual system configuration of Michelson interferometry, the foundation upon which OCT operates. Within this setup, the interferometric probe beam, formed by recombined reflected beams at the beam splitter, is directed toward the surface under examination, with a detector poised to capture the backscatter emanating from this surface [7].

Central to the functioning of OCT is the notion of coherence, a defining characteristic of light wherein all rays maintain a consistent and calculable phase over a defined period. However, the utilization of low coherence, while advantageous in probing biological tissues, introduces an unintended consequence—the introduction of noise altering pixel intensity and distorting the resulting image. This phenomenon manifests as artifacts, leading to a compromised signal-to-noise ratio within OCT images [7]; while various forms of noise may afflict the imaging process, speckle noise emerges as the predominant type encountered in OCT imagery [8].

At the heart of OCT lies the Michelson interferometer, serving as its primary setup. Here, an optical probe directs low-coherent light toward the sample, penetrating its surface

and awaiting the rebound of reflected light. Subsequently, this reflected light is channeled to the interferometer via an optical fiber for meticulous analysis, as depicted in Figure 1 [9].
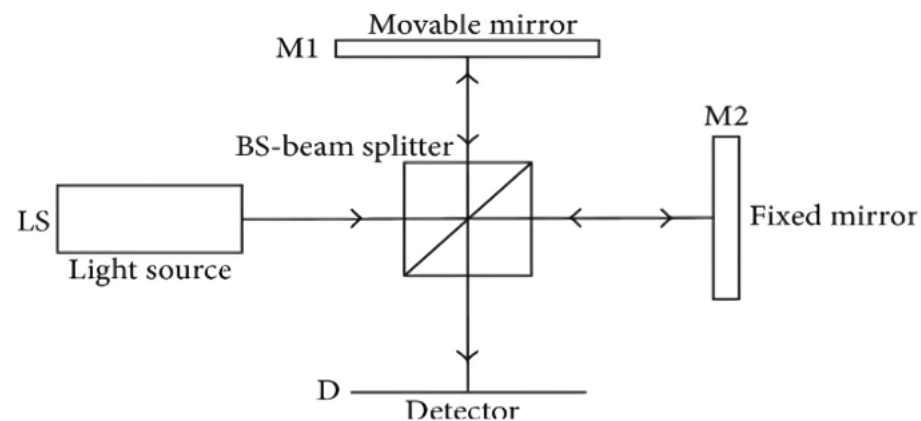

Figure 1. A system diagram of the principle of Michelson interferometry used in OCT [9]. The signal from a collimated light source (LS) is partitioned into two orthogonal beam paths by a beam splitter (BS); one collinear and the other normal to the ls pointing. The normal beam is reflected off a movable mirror (M1), the collinear off a fixed mirror (M2). The reflected beams are recombined at the bs and the co-propagating superposition is recorded as an interferogram at the detector (D). Image [9].

3. Material and Methods

This paper presents a systematic review focusing on studies concerning diagnostic accuracy. The reporting of this study adheres to the PRISMA-DTA guidelines [10].

3.1. Eligibility Criteria

This systematic review is designed to address specific queries utilizing the PICO framework, which encompasses Population, Intervention, Control, and Outcomes. Our investigation centers on several key questions. Firstly, we aim to explore the various implementations and accuracy outcomes associated with employing deep learning (DL) techniques for denoising retinal optical coherence tomography (OCT) imagery. Secondly, we seek to evaluate the effectiveness of these DL-based denoising models in mitigating noise and enhancing the quality of retinal OCT images. Thirdly, we will assess the performance of DL-based denoising models against other denoising methods or untreated OCT images used as benchmarks. Finally, our primary objective is to evaluate the effectiveness of DL-based denoising models by analyzing improvements in image quality through the application of widely used image metrics. Inclusion criteria for studies are as follows:

- Population (P): Studies focusing on the utilization of DL models with retinal imagery obtained from either clinical or research settings.
- Intervention and Control (I, C): Studies employing DL-based models for tasks such as image denoising, speckle reduction, or super-resolution, compared with a reference test.
- Outcomes (O): Studies reporting any estimate of image quality metrics (such as *PSNR*, *CNR*, *SSIM*) applied at either the image or pixel level.

Our exclusion criteria encompass studies that do not meet specific standards, including a lack of clear explanation regarding the utilized DL model, absence of an effective comparative analysis with other state-of-the-art denoising methods, and classification as reviews rather than original research contributions. By meticulously delineating these inclusion and exclusion criteria, our goal is to ensure the robustness and reliability of the systematic review findings. This approach facilitates comprehensive insights into the utility and efficacy of DL in enhancing the quality of retinal OCT imagery.

3.2. Search Methods for Identifying Studies

An electronic search was conducted within the following electronic databases: Google Scholar, Scopus, Medline (via PubMed), Embase, and ArXiv, covering entries up to 16 November 2023. Results were filtered to include publications from 2010 onward, as deep learning (DL) for computer vision and image analysis gained prominence following its development in 2012 by Krizhevsky et al. [11]. Language restrictions were not applied. Customized keywords were utilized for each database. Refer to Table 1 for the specific search queries employed. Moreover, additional studies were identified through screening conference proceedings and journal articles. Furthermore, manual cross-referencing of the bibliographies of included papers was performed to ensure comprehensive coverage.

Table 1. The results of the electronic search in multiple databases.

Database	Keywords	Results	Date
Google Scholar	Image denoising AND deep learning AND "optical coherence tomography"	4220	16 November 2023
Medline	Image denoising AND deep learning AND "optical coherence tomography"	32	16 November 2023
Scopus	TITLE-ABS-KEY (image AND denoising AND deep AND learning AND "optical coherence tomography")	68	16 November 2023
Embase	("image denoising"/exp OR "image denoising" OR (("image"/exp OR image) AND ("denoising"/exp OR denoising))) AND ("deep learning"/exp OR "deep learning" OR (deep AND ("learning"/exp OR learning))) AND ("optical coherence tomography"/exp OR "optical coherence tomography")	37	16 November 2023
ArXiv	Image denoising AND deep learning AND "optical coherence tomography"	9	16 November 2023

3.3. Study Selection

To efficiently manage citations, BibTeX was employed as a tool. Initial screening involved the removal of duplicate entries based on titles and abstracts. Subsequently, a thorough evaluation of articles was conducted to identify eligible studies in accordance with the predetermined inclusion and exclusion criteria. This meticulous process ensured the selection of relevant and appropriate studies for the systematic review. Additionally, any discrepancies or uncertainties during the screening process were resolved through discussion among the research team, ensuring consistency and accuracy in study selection.

3.4. Data Collection and Extraction

Data collection was conducted independently from the included studies and meticulously revised to address any discrepancies or disagreements. Comprehensive information was extracted, encompassing various data items vital for analysis. These included bibliographic details such as authors' names and publication years, details regarding the data modality and type of dataset utilized, hardware specifications, and dataset size (including train/validation/test sets, if provided). Moreover, inclusion and exclusion criteria at the image level, if available, were noted, along with the specified objective of the study (e.g., image denoising, speckle reduction, super resolution). Information regarding pre-processing techniques, data augmentation strategies, and the deep learning (DL) approach employed, including the neural network (NN) architecture utilized, was also recorded. Additionally, details regarding the loss function employed and the image quality metrics used for evaluation were documented. The resulting findings from each study were thoroughly examined. In cases where an article compared multiple NN architectures, the most accurate one was reported to ensure clarity and consistency in the analysis. This comprehensive approach

to data extraction facilitated a robust and thorough examination of the included studies, contributing to the reliability and validity of the systematic review findings.

3.5. Risk of Bias and Applicability

This review focuses specifically on the denoising aspect, and to assess the risk of bias in the included studies, we adapted and employed the QUADAS-2 tool. This modified tool encompasses four main domains addressing the risk of bias: the data selection, index test, reference standard, and flow-and-timing. Additionally, it evaluates three domains regarding the applicability of the study to patient selection, the index test, and the reference standard.

Within the "data selection" domain, we scrutinized papers with vague data-split strategies and limited information on the dataset, which could potentially lead to data leakage, indicating a high risk of bias. Moving on to the "index test" domain, we assessed indicators such as the lack of description of the model and the absence of details regarding test recreation and reproducibility. The "flow-and-timing" domain was evaluated based on indicators such as the implementation of multiple reference standards (i.e., state-of-the-art denoisers) within the same article and the appropriateness of intervals between the index test and reference standard. Finally, within the "reference standard" domain, we considered indicators like inadequate information on reference standard definition and the utilization of only one reference test.

In cases where concerns arose regarding the relevance of the studies, certain factors were meticulously reviewed. These included the dataset used, the procedure employed for creating clean data, the specific deep learning (DL) model utilized, and its performance concerning image quality metrics. Table 2 outlines the key questions utilized in our assessment process, providing a structured framework for evaluating the risk of bias across the included studies. Through this comprehensive approach, we aimed to ensure the rigor and reliability of our review findings.

Table 2. Modified leading questions of QUADAS-2 for critical appraisal.

Domain	Leading Questions
Data Selection	1—Are any data imbalances addressed in the article? 2—Was the dataset split explained correctly for training, validation, and testing? 3—Did the study collect sufficient noisy–clean image pairs?
Index test	1—Was the methodology sufficiently explained for reproducibility? 2—Were the results of deep learning models explained without knowledge of state-of-the-art denoisers? 3—Did the study apply any image quality metrics, sensitivity or robustness analysis on their model?
Flow and Timing	1—Was the full dataset utilized in the analysis? 2—Did each image have a reference clean image? 3—Were the reference clean images produced similarly? 4—Did the model show a sufficient interval between the reference and index test?
Reference Standard	1—Were state-of-the-art results of denoisers mentioned and utilized for interpretation? 2—Did the study describe the noisy–clean image procedure and minimize bias? 3—Were limitations, biases, and generalization issues reported sufficiently?

3.6. Data Synthesis and Analysis

Due to the diverse array of study designs and image quality measures utilized, our quantitative synthesis was primarily confined to examining the outcomes related to image denoising and speckle reduction. Given the substantial variability in the image quality metrics employed for quantifying denoising and super resolution, the scope for direct comparison was somewhat constrained, while a minority of studies reported metrics such as contrast-to-noise ratio (*CNR*), equivalent number of looks (*ENL*), and the structural

similarity index measure (*SSIM*), the predominant metrics presented were peak signal-to-noise ratio (*PSNR*) or signal-to-noise ratio (*SNR*). Consequently, conducting a comparative quantitative analysis across studies was limited to reporting any or all evaluation metrics mentioned, calculated by:

$$PSNR = 10 \log(\frac{L^2}{MSE}), \tag{1}$$

$$SNR = 10 \log(\frac{P_{signal}}{P_{noise}}), \tag{2}$$

L denotes the maximum possible pixel value, and *MSE* is the mean squared error of the image. P_{signal} and P_{noise} are the mean and standard deviation of pixel values, respectively. Next, the structural similarity index (*SSIM*) is a well-known image quality metric that focuses on perceived similarity. The *SSIM* focuses on texture, quality degradation and visible structures. The *SSIM* is defined as

$$SSIM = \frac{(2\sigma_{nc} + c_2)(2\mu_n\mu_c + c_1)}{(\mu_n^2 + \mu_c^2 + c_1)(\sigma_n^2 + \sigma_c^2 + c_2)} \tag{3}$$

μ_n, μ_c, and σ_n, σ_c are the mean value and standard variation in noisy (n)–clean (c) image pairs, respectively. Lastly, contrast-to-noise ratio (*CNR*) utilizes ROIs of background and signal areas for speckle repression with respect to both areas. Equivalent number of looks (*ENL*) is a metric assessing the smoothing of the predicted image. It does not require a reference image since it utilizes selected ROIs of background and signal. *ENL* is defined as

$$ENL = \frac{\mu_b^2}{\sigma_s^2} \tag{4}$$

σ_s is the standard deviation of the signal representation, and μ_b is the mean value for background representation. *CNR* is calculated through

$$CNR = 10 \log(\frac{\mu_s - \mu_b}{\sqrt{\sigma_b^2 + \sigma_s^2}}) \tag{5}$$

μ_s and σ_s are the mean value and standard deviation of the signal representation, respectively. For background representation, μ_b and σ_b are the mean value and standard deviation.

4. Results

4.1. Study Selection and Study Characteristics

Out of the initial pool of 4399 studies identified, a rigorous evaluation was conducted for 41 studies based on the criteria outlined in Table 2, utilizing their full texts. Subsequently, 54 studies were excluded following a thorough assessment of their full texts. The reasons for exclusion were carefully documented and categorized, with detailed explanations provided in Table S1. Ultimately, after meticulous manual screening, a total of 41 studies were deemed eligible for inclusion in our review. Notably, the number of studies included per year exhibited an upward trend over the observation period, as depicted in Figure 2. This trend underscores the increasing interest and attention devoted to the subject matter over time, highlighting the evolving landscape of research in this field.

The studies examined in this review were categorized into two main groups based on the type of optical coherence tomography (OCT) images utilized within the field of ophthalmology: Optic Nerve Head (ONH) images and retinal images. A summary of these studies is presented in Tables 3 and 4, respectively, highlighting the utilization of various deep learning (DL) models for tasks such as image denoising, speckle reduction, and super resolution. Specifically, retinal image datasets utilized in this review were predominantly public datasets purposely employed for denoising OCT images (*n* = 37). Notable datasets included

DUKE [12], Topcon [13], OPTIMA [14], Cirrcus [15], and Heindberg [16]. Conversely, ONH images (*n* = 4) primarily comprised private datasets created by researchers.
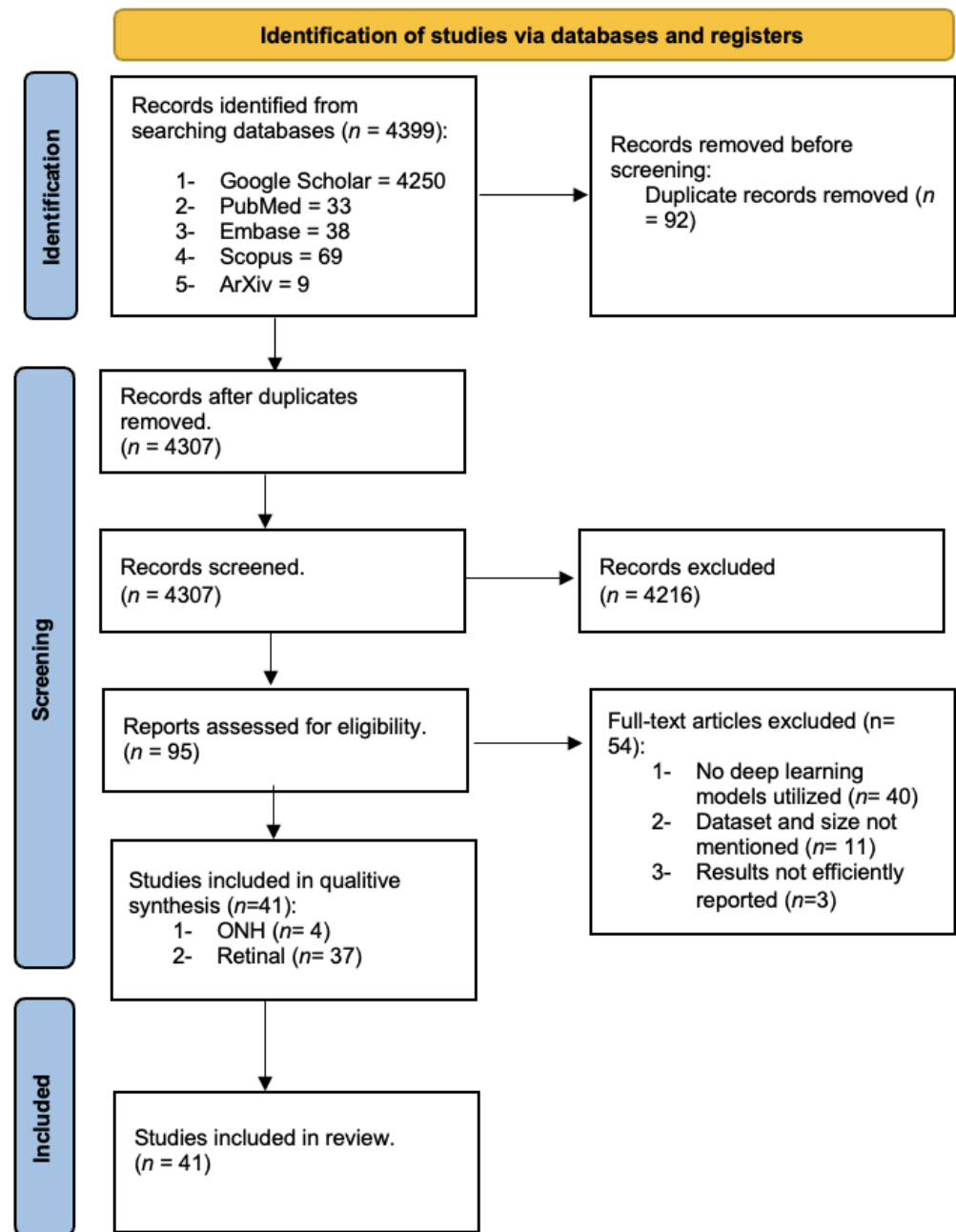

Figure 2. Flowchart of the search following the PRISMA guidelines.

The majority of studies (*n* = 39) incorporated multiple reference tests to evaluate their proposed methods against previous state-of-the-art denoisers. These reference tests encompassed both traditional programming and DL denoising models. Specifically, 26 studies implemented DL models and traditional programming as reference tests, while 9 studies utilized traditional programming-based state-of-the-art denoisers such as BM3D and NLM. However, five studies did not specify any established reference test.

Regarding the choice of dataset for denoising, the DUKE dataset (*n* = 22) was the most frequently utilized, followed by Topcon (*n* = 6), with ONH images being the least employed (*n* = 4). Various DL models were deployed and integrated into hybrid frame-

works, as depicted in Figure 3, which illustrates the distribution of studies implementing each DL model. Notably, 44% of the studies implemented a hybrid generative adversarial network (GAN), with the super-resolution GAN (SR-GAN) being the most utilized (*n* = 3) alongside the conditional GAN (cGAN) (*n* = 3). Additionally, 13% of the studies employed a traditional U-Net model (*n* = 6) for denoising, with only one hybrid framework reported.

Figure 3. Deep learning models approaches involved for denoising OCT images (*n* = studies).

Evaluation of denoising, speckle reduction, and studies on super resolution primarily relied on metrics such as *PSNR* and *SNR* as the major image quality indicators. Furthermore, other widely used image quality metrics included the *SSIM* (*n* = 27), *CNR* (*n* = 24), and *ENL* (*n* = 16). These metrics are comprehensively displayed in Tables 3 and 4, providing insight into the methodologies and outcomes of the reviewed studies.

Table 3. Summary of findings in the selected studies utilizing Optic Nerve Head (ONH) datasets for image denoising, speckle reduction, and super-resolution.

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Devalla, 2019 [17]	24,832 B-Scans (23,280/1552)	Augmentation: Rotation, flip, elastic deformation Pre-Processing: NA	U-Net with residual blocks	NA	MAE	SNR = 8.14 dB CNR = 7.63 dB MSSIM = 0.65
Cheong, 2021 [18]	2628 B-Scans (2328/300)	Augmentation: rotation, translation, flip, and scaling Pre-Processing: NA	Hybrid model (framework containing U-Net followed by a choice of ResNet, ResXNet, or EfficientNet)	NA	Shadow + content + style	PSNR = 11.1% AGM = 57.2% CNR = 154% SSIM = 187%
Tian, 2020 [19]	350 HQ scans (315/35)	Augmentation: flip, rotation, cropping	RMA-GAN	SRCNN, EDSR, ESRGAN	Content + perceptual + adversarial + MS-SSIM + TV	PSNR = 25.7 dB SSIM = 0.77
Hu, 2020 [20]	2500 B-Scans	Augmentation: NA	MSUN with self-fusion	NA	L1 + MSE	PSNR = 10.1 dB SSIM = 0.57
Akter, 2020 [21]	157 B-Scans (100/57)	Augmentation: NA Pre-processing: contrast adjustment, sharpening filter, manually removed noise using Fiji: ImageJ	U-Net	WIN5-RB, Autoencoder, DnCNN, Dense-UNet	MSE	PSNR = 29.8 dB SSIM = 0.90 MSE = 0.005 MAE = 0.03
Halupka, 2018 [22]	55,080 B-Scans (40,711/5587/ 8780)	Augmentation: flip, rotation, cropping	GAN	WGAN, BM3D, DD-CDWT	MSE + VGG + adversarial	PSNR = 32.3 dB SSIM = 0.78 MSE = 40.3 MS-SSIM = 0.92

Keywords: MSE—Mean squared error; MAE—Mean absolute error; SNR—Signal-to-noise ratio; PSNR—Peak signal-to-noise ratio; CNR—Contrast-to-noise ratio; SSIM—Structural similarity index measure; MS-SSIM—Multi-scale structural similarity index measure; RMA-GAN—Realistic mixed attention GAN; SRCNN—Super-resolution CNN; EDSR—Enhanced deep super-resolution network; ESRGAN—Enhanced super-resolution GAN; MSUN—Multi-scale U-Net; WGAN—Wasserstein GAN.

Table 4. Summary of findings in the selected studies utilizing retinal datasets for image denoising, speckle reduction, and super-resolution.

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Wei, 2018 [23]	Duke [12] 26 B-Scans (22/4)	Pre-processing: KNN for clustering Augmentation: scaling	DnCNN	BM3D, NLM, BM3D-SAPCA, LPG-PCA, Low Rank, FFDnet	NA	PSNR = 28.2 dB CNR = 3.9 dB MSR = 6.2
Chen, 2020 [24]	36 B-Scans (25/11)	Pre-processing: aligning, averaging, thresholding, adding speckle noise	DN-GAN	MSBTD, SBSDI, BM3D, K-SVD, Tikhonov, SRResNet, GAN-SRResNet, DCSRN, GAN-U-Net	L1 + perceptual	PSNR = 27.9 dB SSIM = 0.9 FBE = 3.6
Gour, 2020 [25]	Duke [12] and Topcon [13], 23 B-Scans	NA DnCNN	Adaptive Median Filtering, wavelet thresholding, Tikhonov, BM3D, K-SVD, MSBTD, Anisotropic diffusion, STAT, Bayesian, Isotropic diffusion, SE-CNN	MSE	PSNR = 27.5 dB SSIM = 0.68	
Hassan, 2021 [26]	10,000 B-Scans (8000/2000)	Pre-processing: added speckle noise	D-GAN	Wavelet, Bilateral, NLM, BM3D	Euclidean + perceptual + adversarial	PSNR = 35.4 dB MSE = 0.19
Ma, 2018 [15]	Duke [12] and Topcon [13] 521 B-Scans (512/9)	Pre-processing: registration, alignment and enhancing contrast Augmentation: flip, scaling, rotation, non-rigid transformation	cGAN	NLM, BM3D, STROLLR, K-SVD, MAP, DnCNN, ResNet	MSE + L1 + edge	SNR = 60.1 dB CNR = 10.0 dB ENL = 126.9 dB EPI = 1.0
Guo, 2020 [27]	Duke [12], A2A SD-OCT, 90 B-Scans (10/80)	NA	Nonlocal GAN	NLM, BM3D, K-SVD, BM4D, GCBD, GAN-MSE, DnCNN, GAN-WDP, DeGAN	Binary cross-entropy	SNR = 40.1 dB ENL = 981.3 dB CNR = 7.4 dB
Qiu, 2020 [28]	47 B-scans (37/10)	Pre-processing: averaged and registered the B-scans to create denoised image pairs	DnCNN	NLM, BM3D	Perceptually sensitive (SSIM loss)	PSNR = 26.4 dB SSIM = 0.71 MSE = 89.6 MS-SSIM = 0.91

Table 4. *Cont.*

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Huang, 2021 [13]	OCT2017 [29], 84,500 B-scans (83,416/32/968)	Augmentation: crop	AC-SRResNet	BM3D, U-Net, SRResNet	L1	SNR = 41.8 dB CNR = 44.6 dB EPI = 0.72
Halupka, 2018 [22]	69 OCT volumes (51/7/11)	Pre-processing: averaged and registered the B-scans to create denoised image pairs	GAN	BM3D, DD-CDWT, CNN-WGAN	Adversarial + MSE + perceptual	PSNR = 32.3 dB SSIM = 0.78 MS-SSIM = 0.92 MSE = 40.3
Qiu, 2021 [30]	Duke [12], 52 groups of 50 B-scans each (37/15)	NA	P2PGAN-N2N	Median, NLM, BM3D	Adversarial + L1	SNR = 35.5 dB SSIM = 0.81 CNR = 4.0 dB ENL = 260.3 dB R = 0.94
Abassi, 2019 [31]	Duke [12], 28 B-Scans (10, 18)	Augmentation: flip, rotate, crop	MIFCN	KSVD, BM3D, SAIST, PG-GMM, BM4D, SSR	MSE	PSNR = 27.4 dB CNR = 3.8 dB ENL = 2750.8dB
Shi, 2019 [32]	Topcon [13] and Cirrus [15], 11 groups of 256 B-Scans (2/9)	NA	DeSpec-Net	NLM, BM3D, STROLLR, K-SVD, MAP, Intra-volume compounding, DnCNN	L1	SNR = 40.2 dB CNR = 9.7 dB ENL = 166.2 dB EPI = 0.91
Huang, 2020 [33]	Duke [12], 26 B-scans (10/16)	Pre-processing: registering and averaging images, removing any over smoothed images	DRGAN	Median, Bilateral, NLM, Wavelet, BM3D, SNR-GAN, NWSR, edge-sensitive cGAN, HDCycleGAN, Nonlocal GAN, SiameseGAN	Adversarial + reconstruction + cycle-consistency + novel noise	PSNR = 24.4 dB SSIM = 0.58 CNR = 3.2 dB EPI = 0.98 MSR = 4.8 ENL = 317.4 dB
Yu, 2018 [34]	Duke [12], 15 B-scans (8/3/4)	Pre-processing: crop, removing unaligned images	DN-GAN	BM3D, BM3DPCA, LPGPCA, FFDNET	MSE + adversarial	PSNR = 31.0 dB CNR = 3.3 dB MSR = 3.7

Table 4. *Cont.*

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Tajmirr-iahi, 2021 [35]	Topcon [13], 240 B-Scans (200/40)	Augmentation: rotation, shift, flip, and crop	Autoencoder	GT-SC-GMM, BM3D, MSBTD, Tikhonov	MSE	SNR = 108.8 dB CNR = 82.2 dB ENL = 58.4 dB TP = 0.79 EP = 0.98 CT = 4.68
Sengupta, 2021 [36]	Duke [12], 1600 B-Scans (1400/200)	Pre-processing: crop	EdgeWaveNet	NLM, DeBlur-GAN, RDNSR-GAN, RED-GAN	L1 + adversarial + Sobel edge	PSNR = 22.8 dB SSIM = 0.61
Mehdi-zadeh, 2021 [37]	71 B-scans (51/20)	Augmentation: created patches	DnCNN	NA	L2 + L1 + perceptual + VGG	PSNR = 33.6 dB PSI = 0.23 JNB = 13.9 S3 = 0.26
Cai, 2018 [38]	Topcon [13], 256 B-scans (246/10)	Pre-processing: averaged and registered the B-scans to create denoised image pairs	ResNet	Median, NLM, BM3D	MSE	PSNR = 34.8 dB SSIM = 0.52
Zhou, 2022 [39]	5000 B-scans (4500/480/20)	Pre-processing: crop	Transformer-IP2	BM3D, PNLM, NCDF, OBNLM, DnCNN, CNN-NLM, Neighbor2Neighbor	Neighbor- 2Neighnor + PNLM	SNR = 154.6 dB CNR = 7.9 dB ENL = 13,160.3 dB
Anoop, 2021 [40]	Duke [12] and Optima [14], 2720 B-scans (2176/544)	Pre-processing: noise distribution is found for each image, patches and denoised image pairs were created	DenseNet121	CAD, OBNLM, TVG, Wavelet, K-SVD, DnCNN	Cross-entropy	PSNR = 31.0 dB SSIM = 0.91
Fu, 2021 [41]	Duke [12], 21 B-scans (16/5)	Pre-processing: registering and averaging images, removed any over smoothed images	ADGAN	Wavelet, NLM, BM3D, NWSR, HDCycleGAN	Adversarial + cycle-consistency	PSNR = 27.6 dB SSIM = 0.62 CNR = 3.1 dB ENL = 530.8 dB
Wang, 2021 [42]	Topcon [13] and Cirrus [15], 1920 B-scans (512/1408)	Pre-processing: creating denoised image pairs from [23]	Capsule cGAN	BM3D, K-SVD, NLM, MAP, STROLLR, DnCNN, ResNet, Cycle-GAN, cGAN	L1 + Adversarial + SSIM	SNR = 59.0 dB CNR = 11.4 dB ENL = 417.2 dB EPI = 1.0

Table 4. *Cont.*

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Zhou, 2022 [43]	Topcon [13] and Cirrus [15], 1920 B-scans (512/1408)	Pre-processing: registering and averaging images	Cycle-GAN with mini-cGAN	NLM, BM3D, STROLLR, K-SVD, MAP, DnCNN, DPDNN, NAGAN with mini-cGAN	L1 + MSE	SNR = 20.9 dB CNR = 12.5 dB SSI = 0.09 EPI = 0.99
Wu, 2021 [44]	3737 B-scans, (3537/200)	Pre-processing: crop and contrast enhancement	cGAN	Cycle-GAN, DnCNN, BM3D, DCWT, NLM, MPE, cGAN, EGAN, SR	Adversarial + cycle-consistency + structural consistency + regularization	SNR = 35.0 dB CNR = 7.2 dB EPI = 0.92 CRSB = 0.14
Das, 2020 [45]	Duke [12] 45 B-scans and 384 OCT volumes, (2000/17)	Pre-processing: crop	SRGAN	SBSDI, SSR, NWSR, SRGAN	Adversarial + cycle-consistency + identity mapping	PSNR = 39.2 dB CNR = 4.7 dB
Huang, 2019 [46]	Duke [12], 26 B-scans (10/16)	Pre-processing: crop	SDSR-OCT	BM3D + Bicubic, NWSR, SRCNN	Pixel + perceptual + GAN	PSNR = 28.1 dB CNR = 4.6 dB ENL = 537.5 dB EPI = 0.95
Ge, 2022 [47]	Duke [12], 10 B-scans	Pre-processing: clear images are obtained by registering and averaging and crop	Self2Self-OCT	BM3D, NWSR, DnCNN, DIP, TSI	Background noise attenuation + self-prediction	PSNR = 24.8 dB SSIM = 0.99
Ma, 2022 [48]	Duke [12], 26 B-scans (10/16)	NA	DSGAN	MIFCN, Edge-sensitive cGAN, SDSR-OCT	Adversarial + SSIM + MSE	PSNR = 28.1 dB SSIM = 0.95 CNR = 3.7 dB
Xie, 2022 [49]	Duke [12], 26 B-scans (22/4)	NA	GAN	K-SVD, BM3D, wGAN, cGAN, SDSR, HDcycleGAN, DRGAN	Adversarial + cycle-consistency + perceptual	PSNR = 27.6 dB EPI = 1.0 CNR = 3.1 dB ENL = 73.8 dB MSR = 5.1 SSIM = 0.65
Xie, 2023 [50]	Duke [12], 26 B-scans (10/16)	NA	MGAN	NLM, BM3D, DnCNN, MIFCN, SDSR-OCT	Adversarial + pixel-level error + BCE+ SSIM	PSNR = 28.1 dB SSIM = 0.95 EPI = 0.99 CNR = 3.6 dB

Table 4. *Cont.*

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Ahmed, 2022 [51]	Duke [12], 18 B-scans (10/8)	Pre-processing: clean images are obtained by BM3D, BM3DDEB, Weiner and HWT	DenseNet with AG	BM3D, NLM	MSE + pixel difference	PSNR = 23.5 dB CNR = 7.7 dB ENL = 585.5 dB
Ahmed, 2022 [52]	Duke [12] and dentistry, 28 B-scans (18/12)	NA	Autoencoder with MFSK and AG	BM3D, NLM, DnCNN, GAN	MSE + pixel difference	PSNR = 26.9 dB CNR = 7.0 dB ENL = 213.7 dB SSIM = 0.68
Zhou, 2023 [53]	OCT2017 [29] and OCTID [54], 5620 B-scans (5000/ 600/20)	Pre-processing: crop	Transformer-based NLM	N2N, DRGAN, Den-mimic-net, Contourlet, BM3D, INLSM, NLM, OBNLM, PNLM	MSE + gradient	CNR = 15.7 dB SNR = 51.1 dB ENL = 23,787 dB
Kande, 2020 [55]	Duke [12], 28 B-Scans (10/18)	NA	SiameseGAN	MSBTD, MIFCN, Shared Encoder, WGAN U-Net, WGAN ResNet	MS-SSIM + perceptual	PSNR = 28.3 dB SSIM = 0.83 MSR = 4.2 CNR = 2.6 dB TP = 0.68 EP = 0.66
Qiu, 2020 [56]	Duke [12], 52 groups of 50 B-scans each (37/15)	Pre-processing: crop	DBPN	BM3D, Bicubic, NWSR, U-Net	MSE	PSNR = 31.3 dB RMSE = 0.027 MS-SSIM = 0.92
Zhou, 2021 [57]	Topcon [13] and Cirrus [15], 521 B-scans (512/9)	Pre-processing: registering and averaging images Augmentation: flip, scaling, rotation, non-rigid transformation	GAN with HRNet	NLM, STROLLR, DnCNN, DPDNN, Edge-cGAN, mini-cGAN	L1 + MSE + Adversarial	SNR = 40.4 dB CNR = 11.2 dB SSI = 0.09 EPI = 0.96

Table 4. *Cont.*

Paper	Data Size	Pre-Processing and Augmentation	Model	Compared to	Loss Function	Image Quality Metrics
Ahmed, 2022 [58]	Duke [12], 18 B-scans (10/8)	NA	DnCNN	BM3D, Weiner, NLM	CNR + pixel difference	PSNR = 29.6 dB CNR = 11.5 dB ENL = 1196.6 dB

Keywords: MSE—Mean squared error; MAE—Mean absolute error; SNR—Signal-to-noise ratio; PSNR—Peak signal-to-noise ratio; CNR—Contrast-to-noise ratio; SSIM—Structural similarity index measure; MS-SSIM—Multi-scale structural similarity index measure; ENL—Equivalent number of looks; EPI—Edge preservation index; RMA-GAN—Realistic mixed attention GAN; SRGAN—Super-resolution GAN; EGAN—Enhanced GAN; SRCNN—Super-resolution CNN; SSR—Self Super-resolution; SRResNet—Super-resolution ResNet; WGAN—Wasserstein GAN; DeGAN—Denoising GAN; BCE—Binary Cross Entropy; MGAN—Multi-task generative adversarial network; HWT—Hyperanalytic Wavelet Transform; AG—Attention Gate; MKSF—Multi-kernel speckle filtering block; N2N—Neighbor2Neighbor; DRGAN—Disentangled representation generative adversarial network; NLM—Nonlocal means.

4.2. Risk of Bias and Applicability

Every study included in the review underwent a thorough assessment of risk of bias, with the results meticulously documented in Table S2. Out of the studies evaluated, 10 (27%) were identified as having a low risk of bias across all four domains. Notably, the domain presenting the most challenges was the "index test", with only 22 studies (59.5%) categorized as having a low risk of bias in this area. This finding underscores the importance of critically evaluating the methodology and execution of the index test within each study to ensure the reliability and validity of the findings. Through this rigorous risk-of-bias assessment, we aimed to provide a comprehensive evaluation of the methodological robustness of the included studies, thereby enhancing the credibility and trustworthiness of our review outcomes.

4.3. Findings of the Studies

When focusing on peak signal-to-noise ratio (*PSNR*) and signal-to-noise ratio (*SNR*) as metrics for denoising images, a significant variation in deep learning (DL) techniques across different types of optical coherence tomography (OCT) images was observed. Specifically, for studies utilizing Optic Nerve Head (ONH) images, *PSNR* or *SNR* values ranged from 8.1 to 25.7 dB, while, for retinal datasets, these metrics ranged from 26.4 to 158.6 dB.

The majority of studies predominantly applied public retinal datasets such as Duke [12], Topcon [13], and Cirrus [15], primarily due to their extensive availability and widespread use. However, it is noteworthy that these datasets typically do not provide "clean" versus "noisy" image pairs required for denoising tasks. Consequently, researchers resorted to generating clean images through alternative means. For instance, clean images were generated by averaging multiple B-scans or employing traditional programming techniques. In some cases, unsupervised DL techniques were also utilized to generate clean images. These approaches aimed to provide a reliable basis for assessing the effectiveness of DL-based denoising methods, despite the absence of explicitly labeled clean and noisy image pairs within the datasets.

5. Discussion

Numerous DL tasks consisting of detection, segmentation and classification are challenging in ophthalmology since OCT is the main imaging technique in that specific field. OCT introduces speckle noise that adds artifacts, thereby impairing image quality and confounding correct clinical interpretation [59]. Thus, DL has provided solutions for denoising and speckle reduction in OCT images to overcome this problem. In this systematic review, many studies were compiled and analyzed to assess the application of DL for image denoising of OCT images in ophthalmology. Over the past ten years, multiple researchers have shown a developing and promising body of evidence supporting DL for this task. Even so, there was a limited supply of quality studies for comparison across traditional programming and DL. A number of findings require more detailed discussion. The challenge of addressing numerous deep learning (DL) tasks in ophthalmology, including detection, segmentation, and classification, is compounded by the prevalent use of optical coherence tomography (OCT) as the primary imaging technique in this field. OCT introduces speckle noise, which can adversely affect image quality, thereby complicating accurate clinical interpretation. DL techniques have emerged as promising solutions for denoising and reducing speckle artifacts in OCT images, with this systematic review analyzing a multitude of studies to assess the application of DL for image denoising in ophthalmology. Over the past decade, researchers have generated a growing body of evidence supporting the efficacy of DL for this purpose. As shown in Figure 3, the main DL models implemented were the GAN, DnCNN, Autoencoder, U-Net, Noise2Noise, and transformer, etc.

Generative adversarial networks (GANs) exhibit promise in synthesizing synthetic OCT images with reduced noise levels. By training a generator network to generate realistic images and a discriminator network to differentiate between real and generated images, GANs have demonstrated their capability to denoise OCT scans effectively while preserving

clinically relevant features. However, GANs can be challenging to train and prone to mode collapse, where the generator produces limited variations in images, potentially limiting their diversity and generalizability in clinical settings. In contrast, Denoising Convolutional Neural Networks (DnCNNs) utilize deep convolutional layers to understand the underlying structure of noisy OCT images and produce clean counterparts. Through iterative training on paired noisy–clean image datasets, DnCNNs efficiently enhance image clarity. However, DnCNNs may struggle with complex noise patterns and require large amounts of labeled data for training, which may be resource-intensive and time-consuming to acquire. The U-Net architecture, characterized by its symmetric encoder–decoder structure with skip connections, has showcased remarkable performance in semantic segmentation tasks, including the denoising of OCT images. By integrating contextual information from various spatial scales, U-Net effectively preserves anatomical structures. Nonetheless, U-Net architectures may suffer from memory inefficiency and computational overhead, particularly when handling high-resolution OCT images. Autoencoders, consisting of an encoder and decoder network, learn to reconstruct input data from a compressed representation. Trained on noisy OCT images, autoencoders can encode essential features suitable for clinical interpretation. However, autoencoders may struggle with capturing complex image structures and may require careful tuning of hyperparameters to achieve optimal performance. Transformers, initially designed for natural language processing tasks, have recently been applied in image processing, including denoising. Leveraging self-attention mechanisms to capture long-range dependencies, transformers effectively preserve spatial information in OCT images. However, transformers may suffer from scalability issues when applied to large-scale image datasets due to their computational complexity and memory requirements. The Noise2Noise approach involves training deep learning models directly on pairs of noisy images, eliminating the need for clean reference images during training. By exploiting inherent redundancies in noisy data, Noise2Noise denoises OCT images without access to ground truth clean images, making it suitable for real-world clinical applications where clean reference images may not be readily available. Nevertheless, Noise2Noise may struggle with highly variable noise patterns and may require careful regularization techniques to prevent overfitting to noise artifacts.

These innovative DL techniques have the potential to revolutionize clinical workflows in ophthalmology by automating the denoising process and improving the efficiency and accuracy of OCT image interpretation. By providing clinicians with high-quality, denoised images, these techniques can facilitate more confident diagnoses and treatment decisions, ultimately leading to improved patient outcomes. However, the scarcity of high-quality studies for comparison across traditional programming and DL remains limited, necessitating a closer examination of key findings.

Firstly, a significant proportion of studies (62%) conducted their DL model training and testing on public datasets, validating their reported image quality metrics and comparing them with other state-of-the-art DL models (reference tests). These studies demonstrated relatively high *PSNR* or *SNR* values for denoised images, indicating substantial improvement in image quality attributable to DL models. Given the considerable noise inherent in OCT images, DL techniques hold significant promise for enhancing diagnosis, segmentation, and detection tasks.

Secondly, only studies utilizing the Duke [12] public dataset had access to clean versus noisy image pairs for training purposes. Other studies, utilizing datasets like Topcon [12] and Cirrus [15], lacked clean images and thus resorted to creating their own algorithms or applying traditional programming techniques to obtain "ground truths" for their B-scans. This variance in ground truth generation methods complicates cross-study comparisons and underscores the importance of clearly outlining and validating strategies for creating clean images to ensure the production of robust models.

Thirdly, there was considerable variability in both the reporting and conduct of the studies, posing challenges for comprehensive quantitative and qualitative analyses across the board. Given the wide array of available image quality metrics beyond *PSNR* or *SNR*

(such as *CNR*, *ENL*, and *SSIM*), it was observed that studies typically reported only two or three metrics. This lack of uniformity in reporting was further exacerbated by the implementation of multiple and varied reference tests. For instance, several studies utilized a diverse range of reference tests, resulting in a vague impact assessment on either *PSNR* or *SNR*. Consequently, there is a pressing need for a standardized structure for image denoising studies, encompassing specific reference tests (including the appropriate balance of traditional programming and deep learning models) and a predetermined set of specified image quality metrics (such as *PSNR*, *SSIM*, *CNR*, and *ENL*). Consequently, the feasibility of conducting a meta-analysis was limited, as only a small subset of studies provided similar image quality metrics beyond *PSNR* or *SNR*. This situation underscores the significant consequences stemming from the lack of minimum standards and tools available in the field of image denoising, where diagnostic studies are currently limited to utilizing only two tools (STRAD-AI and QUADS-2). Since the studies included in this review solely reported either *PSNR* or *SNR* of the denoised images in comparison to multiple reference tests, it remains uncertain whether DL provides significant assistance in image denoising and speckle reduction. Ideally, the impact of DL for denoising OCT images in ophthalmology should be demonstrated in practice-based settings and validated by its ability to improve further objectives such as detection [60], classification [61], and segmentation [62], which the majority of included studies did not consider.

Fourthly, a significant proportion (74%) of papers were excluded from the review due to the absence of DL methodologies, instead focusing on improving traditional programming methods such as wavelets and shearlets [63], NLM [64], and BM3D [65]. This suggests that DL has not been extensively investigated for OCT image denoising. The remaining papers were excluded due to inadequate descriptions of datasets and their utilization, as well as a lack of reference testing to demonstrate their impact on *PSNR* and *SNR*.

Fifthly, and perhaps most importantly, none of the studies incorporated input from clinicians regarding the produced results. This is particularly concerning considering that the primary purpose of denoising OCT images is to assist clinicians in diagnosing various retinal diseases. Ideally, clinicians should provide feedback on the reported denoised images, assessing whether useful data has been removed or added that could significantly impact diagnosis accuracy.

Lastly, it is essential to emphasize that OCT is the primary instrument in ophthalmology for capturing multiple types of images (including ONH and retinal images) crucial for detecting retinal diseases. This systematic review has revealed both strengths and limitations, with a systematic and extensive assessment of studies conducted to compare DL for image denoising, speckle reduction, and super resolution for OCT images in ophthalmology; while there was a restricted timeline justified by the focus on DL, a substantial and diverse body of evidence has been presented, validating the necessity for such limitations. Additionally, the absence of a meta-analysis was attributed to the mixed nature of reporting and the lack of quality in comparative result analysis.

Therefore, there has been a growing number of studies investigating the denoising of OCT images in ophthalmology using deep learning, with various computational architectures being explored. Among the reported metrics of image quality, the peak signal-to-noise ratio (*PSNR*) has emerged as a reliable metric for intercomparison, with values spanning from 8.1 to 25.7 dB for ONH images and 26.4 to 158.6 dB for retinal datasets. Moving forward, it is imperative for future studies to clearly outline reference tests and datasets, relying on a common, extensive, and clinically meaningful outcome basis to drive progress in the field.

6. Conclusions

In summary, the landscape of research in denoising Optical Coherence Tomography (OCT) images within ophthalmology has seen a notable surge, with a diverse array of studies delving into the application of deep learning techniques. This exploration has encompassed various computational architectures, reflecting a dynamic and evolving field

seeking optimal solutions for image enhancement. Among the array of metrics used to evaluate image quality, the peak signal-to-noise ratio (*PSNR*) has emerged as a consistent benchmark for intercomparison, offering insights into the effectiveness of denoising methodologies. Notably, *PSNR* values have exhibited substantial ranges, underscoring the complexity and variability inherent in OCT image denoising efforts. Looking ahead, it is essential for future investigations to prioritize transparency and standardization in methodologies, particularly in outlining reference tests and datasets. By establishing a common foundation grounded in clinically meaningful outcomes, researchers can foster more robust advancements and ensure the translation of findings into tangible benefits for clinical practice and patient care.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jimaging10040086/s1, Table S1: List of excluded studies and the reasons; Table S2: Risk-of-bias assesment of papers reviewed.

Author Contributions: H.A. and R.D. wrote the main manuscript text, and H.A. prepared Figures 1–3. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

MSE—Mean squared error; SNR—Signal-to-noise ratio; PSNR—Peak signal-to-noise ratio; CNR—Contrast-to-noise ratio; SSIM—Structural similarity index measure; ENL—Equivalent number of looks; BCE—Binary Cross Entropy; MGAN—Multi-task generative adversarial network; HWT—Hyperanalytic Wavelet Transform; AG—Attention Gate; MKSF—Multi-kernel speckle filtering block; N2N—Neighbor2Neighbor; DRGAN—Disentangled representation generative adversarial network; NLM—Nonlocal means; EPI—Edge preservation index; RMA-GAN—Realistic mixed attention GAN; SRGAN—Super-resolution GAN; EGAN—Enhanced GAN; SRCNN—Super-resolution CNN; SSR—Self Super-resolution; SRResNet—Super-resolution ResNet; WGAN—Wasserstein GAN; DeGAN—Denoising GAN; BCE—Binary Cross Entropy; MGAN—Multi-task generative adversarial network. MSUN—Multi-scale U-Net.

References

1. Aumann, S.; Donner, S.; Fischer, J.; Müller, F. Optical Coherence Tomography (OCT): Principle and Technical Realization. In *High Resolution Imaging in Microscopy and Ophthalmology*; Springer International Publishing: Cham, Switzerland, 2019; pp. 59–85.
2. Huang, S.; Tang, C.; Xu, M.; Qiu, Y.; Lei, Z. BM3D-based total variation algorithm for speckle removal with structure-preserving in OCT images. *Appl. Opt.* **2019**, *58*, 62336243. [CrossRef] [PubMed]
3. Tey, K.Y.; Teo, K.; Tan, A.C.S.; Devarajan, K.; Tan, B.; Tan, J.; Schmetterer, L.; Ang, M. Optical coherence tomography angiography in diabetic retinopathy: A review of current applications. *Eye Vis.* **2019**, *6*, 110. [CrossRef] [PubMed]
4. Müller, P.L.; Liefers, B.; Treis, T.; Rodrigues, F.G.; Olvera-Barrios, A.; Paul, B.; Dhingra, N.; Lotery, A.; Bailey, C.; Taylor, P.; et al. Reliability of Retinal Pathology Quantification in Age-Related Macular Degeneration: Implications for Clinical Trials and Machine Learning Applications. *Transl. Vis. Sci. Technol.* **2021**, *10*, 44 [CrossRef] [PubMed]
5. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85117.
6. Colston, B.W.; Sathyam, U.S.; DaSilva, L.B.; Everett, M.J.; Stroeve, P.; Otis, L.L. Dental OCT. *Opt. Express* **1998**, *3*, 230238. [CrossRef] [PubMed]
7. Lai, Y.-C.; Lin, J.-Y.; Yao, C.-Y.; Lyu, D.-Y.; Lee, S.-Y. Interactive OCT-Based Tooth Scan and Reconstruction. *Sensors* **2019**, *19*, 4234. [CrossRef] [PubMed]
8. Wong, A.; Mishra, A.; Bizheva, K.; Clausi, D.A. General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery. *Opt. Express* **2010**, *18*, 8338–8352. [CrossRef]
9. Novacam Technologies, Inc. How Low-Coherence Interferometry (LCI) Works. 2021. Available online: https://www.novacam.com/technology/how-lci-works/ (accessed on 17 November 2023).
10. McInnes, M.D.F.; Moher, D.; Thombs, B.D.; McGrath, T.A.; Bossuyt, P.M.; Clifford, T.; Cohen, J.F.; Deeks, J.J.; Gatsonis, C.; Hooft, L.; et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *J. Am. Med. Assoc.* **2018**, *319*, 388396. [CrossRef]

11. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst. (NIPS)* **2012**, *25*, 10971105. [CrossRef]

12. Fang, L.; Li, S.; Nie, Q.; Izatt, J.A.; Toth, C.A.; Farsiu, S. Sparsity based denoising of spectral domain optical coherence tomography images. *Biomed. Opt. Express* **2012**, *3*, 927942. [CrossRef]

13. Huang, Y.; Zhang, N.; Hao, Q. Real-time noise reduction based on ground truth free deep learning for optical coherence tomography. *Biomed. Opt. Express* **2021**, *12*, 20272040. [CrossRef] [PubMed]

14. Optima—Cyst Segmentation Challenge. 2015. Available online: https://optima.meduniwien.ac.at/optima-segmentation-challenge-1/ (accessed on 17 November 2023).

15. Ma, Y.; Chen, X.; Zhu, W.; Cheng, X.; Xiang, D.; Shi, F. Speckle noise reduction in optical coherence tomography images based on edge-sensitive cgan.*Biomed. Opt. Exp.* **2018**, *9*, 5129–5146. [CrossRef] [PubMed]

16. Mokhtari, M.; Kamasi, Z.G.; Rabbani, H. Automatic detection of hyperreflective foci in optical coherence tomography B-scans using morphological component analysis. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 1497–1500.

17. Devalla, S.K.; Subramanian, G.; Pham, T.H.; Wang, X.; Perera, S.; Tun, T.A.; Aung, T.; Schmetterer, L.; Thiery, A.H.; Girard, M.J.A. A Deep Learning Approach to Denoise Optical Coherence Tomography Images of the Optic Nerve Head. *Sci. Rep.* **2019**, *9*, 14454. [CrossRef] [PubMed]

18. Cheong, H.; Devalla, S.K.; Chuangsuwanich, T.; Tun, T.A.; Wang, X.; Aung, T.; Schmetterer, L.; Buist, M.L.; Boote, C.; Thiery, A.H.; et al. OCT-GAN: Single step shadow and noise removal from optical coherence tomography images of the human optic nerve head. *Biomed. Opt. Express* **2021**, *12*, 14821498. [CrossRef] [PubMed]

19. Tian, C.; Yang, J.; Li, P.; Zhang, S.; Mi, S. Retinal fundus image superresolution generated by optical coherence tomography based on a realistic mixed attention GAN. *Med. Phys.* **2022**, *49*, 318531998. [CrossRef] [PubMed]

20. Hu, D.; Malone, J.D.; Atay, Y.; Tao, Y.K.; Oguz, I. Retinal OCT Denoising with Pseudo-Multimodal Fusion Network. In *OMIA 2020: Ophthalmic Medical Image Analysis*; Springer: Cham, Switzerland, 2020; pp. 125–135.

21. Akter, N.; Perry, S.; Fletcher, J.; Simunovic, M.; Roy, M. Automated Artifacts and Noise Removal from Optical Coherence Tomography Images Using Deep Learning Technique. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, ACT, Australia, 1–4 December 2020; pp. 2536–2542.

22. Halupka, K.; Lee, B.A.M.; Lucy, K.; Rai, R.; Ishikawa, H.; Wollstein, G.; Schuman, J.; Garnavi, R. Retinal optical coherence tomography image enhancement via deep learning. *Biomed. Opt. Express* **2018**, *9*, 62056221. [CrossRef] [PubMed]

23. Wei, X.; Liu, X.; Yu, A.; Fu, T.; Liu, D. Clustering-Oriented Multiple Convolutional Neural Networks for Optical Coherence Tomography Image Denoising. In Proceedings of the 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 13–15 October 2018; pp. 1–5.

24. Chen, Z.; Zheng, Z.; Shen, H.; Zheng, Z.; Dai, P.; Ouyang, P. DN-GAN: Denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images. *Biomed. Signal Process. Control* **2020**, *55*, 101632. [CrossRef]

25. Gour, N.; Khanna, P. Speckle denoising in optical coherence tomography images using residual deep convolutional neural network. *Multimed. Tools Appl.* **2020**, *79*, 1567915695. [CrossRef]

26. Hasan, M.J.; Alom, M.S.; Fatema, U.; Wahid, M.F. Deep Learning Based Retinal OCT Image Denoising using Generative Adversarial Network. In Proceedings of the International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), Rajshahi, Bangladesh, 8–9 July 2021; pp. 1–6.

27. Guo, A.; Fang, L.; Qi, M.; Li, S. Unsupervised Denoising of Optical Coherence Tomography Images with Nonlocal-Generative Adversarial Network. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 112. [CrossRef]

28. Qiu, B.; Huang, Z.; Liu, X.; Meng, X.; You, Y.; Liu, G.; Yang, K.; Maier, A.; Ren, Q.; Lu, Y. Noise reduction in optical coherence tomography images using a deep neural network with perceptually sensitive loss function. *Biomed. Opt. Express* **2020**, *11*, 817830. [CrossRef]

29. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **2018**, *172*, 1122–1131. [CrossRef]

30. Qiu, B.; Zeng, S.; Meng, X.; Jiang, Z.; You, Y.; Geng, M.; Li, Z.; Hu, Y.; Huang, Z.; Zhou, C.; et al. Comparative study of deep neural networks with unsupervised Noise2Noise strategy for noise reduction of optical coherence tomography images. *J. Biophotonics* **2021**, *14*, e202100151. [CrossRef] [PubMed]

31. Abbasi, A.; Monadjemi, A.; Fang, L.; Rabbanni, H.; Zhang, Y. Three-dimensional optical coherence tomography image denoising through multi-input fully convolutional networks. *Comput. Biol. Med.* **2019**, *108*, 18. [CrossRef]

32. Shi, F.; Cai, N.; Gu, Y.; Hu, D.; Ma, Y.; Chen, Y.; Chen, X. DeSpecNet: A CNN-based method for speckle reduction in retinal optical coherence tomography images. *Phys. Med. Biol.* **2019**, *64*, 175010. [CrossRef]

33. Huang, Y.; Xia, W.; Lu, Z.; Liu, Y.; Chen, H.; Zhou, J.; Fang, L.; Zhang, Y. Noise-Powered Disentangled Representation for Unsupervised Speckle Reduction of Optical Coherence Tomography Images. *IEEE Trans. Med. Imaging* **2021**, *40*, 26002614. [CrossRef] [PubMed]

34. Yu, A.; Liu, X.; Wei, X.; Fu, T.; Liu, D. Generative Adversarial Networks with Dense Connection for Optical Coherence Tomography Images Denoising. In Proceedings of the 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 13–15 October 2018; pp. 1–5.

35. Tajmirriahi, M.; Kafieh, R.; Amini, Z.; Rabbani, H. A Lightweight Mimic Convolutional Auto-Encoder for Denoising Retinal Optical Coherence Tomography Images. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 18. [CrossRef]

36. Sengupta, S.; Singh, A.; Lakshminarayanan, V. EdgeWaveNet: Edge aware residual wavelet GAN for OCT image denoising. In Proceedings of the Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications, Online, 15–20 February 2021; Volume 11601, pp. 110–115.

37. Mehdizadeh, M.; MacNish, C.; Xiao, D.; Alonso-Caneiro, D.; Kugelman, J.; Bennamoun, M. Deep feature loss to denoise OCT images using deep neural networks. *Biomed. Opt.* **2021**, *26*, 046003. [CrossRef]

38. Cai, N.; Shi, F.; Hu, D.; Chen, Y. A ResNet-based universal method for speckle reduction in optical coherence tomography images. In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), Washington, DC, USA, 4–7 April 2018.

39. Zhou, Q.; Wen, M.; Ding, M.; Zhang, X. Unsupervised despeckling of optical coherence tomography images by combining cross-scale CNN with an intra-patch and inter-patch-based transformer. *Opt. Express* **2022**, *30*, 1880018820. [CrossRef]

40. Anoop, B.N.; Kalmady, K.S.; Udathu, A.; Siddharth, V.; Girish, G.N.; Kothari, A.R.; Rajan, J. A cascaded convolutional neural network architecture for despeckling OCT images. *Biomed. Signal Process. Control* **2021**, *66*, 102463. [CrossRef]

41. Fu, Z.; Yu, X.; Ge, C.; Aziz, M.Z.; Liu, L. ADGAN: An Asymmetric Despeckling Generative Adversarial Network for Unpaired OCT Image Speckle Noise Reduction. In Proceedings of the IEEE 6th Optoelectronics Global Conference (OGC), Shenzhen, China, 15–18 September 2021; pp. 212–216.

42. Wang, M.; Zhu, W.; Yu, K.; Chen, Z.; Shi, F.; Chen, X. Semi-Supervised Capsule cGAN for Speckle Noise Reduction in Retinal OCT Images. *IEEE Trans. Med. Imaging* **2021**, *40*, 11681183. [CrossRef]

43. Zhou, Y.; Yu, K.; Wang, M.; Ma, Y.; Peng, Y.; Chen, Z.; Zhu, W.; Shi, F.; Chen, X. Speckle Noise Reduction for OCT Images Based on Image Style Transfer and Conditional GAN. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 139150. [CrossRef] [PubMed]

44. Wu, M.; Chen, W.; Chen, Q.; Park, H. Noise Reduction for SD-OCT Using a Structure-Preserving Domain Transfer Approach. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 34603472. [CrossRef] [PubMed]

45. Das, V.; Dandapat, S.; Bora, P.K. Unsupervised Super-Resolution of OCT Images Using Generative Adversarial Network for Improved Age-Related Macular Degeneration Diagnosis. *IEEE Sens. J.* **2020**, *20*, 87468756. [CrossRef]

46. Huang, Y.; Lu, Z.; Shao, Z.; Ran, M.; Zhou, J.; Fang, L.; Zhang, Y. Simultaneous denoising and super-resolution of optical coherence tomography images based on generative adversarial network. *Opt. Express* **2019**, *27*, 1228912307. [CrossRef] [PubMed]

47. Ge, C.; Yu, X.; Li, M.; Mo, J. Self-Supervised Denoising of single OCT image with Self2Self-OCT Network. In Proceedings of the IEEE 7th Optoelectronics Global Conference (OGC), Shenzhen, China, 6–11 December 2022; pp. 200–204.

48. Ma, Z.; Xie, Q.; Fan, F.; Zhu, J. DSGAN: A generative model for speckle noise reduction in retinal optical coherence tomography images. In Proceedings of the SPIE 12320, Optics in Health Care and Biomedical Optics XII, Online, 5–12 December 2022; p. 123201H.

49. Xie, K.; Luo, M.; Chen, H.; Yang, M.; He, Y.; Liao, P.; Zhang, Y. Speckle denoising of optical coherence tomography image using residual encoder–decoder CycleGAN. *Signal Image Video Process.* **2023**, *17*, 1521–1533. [CrossRef]

50. Xie, Q.; Ma, Z.; Zhu, L.; Fan, F.; Meng, X.; Gao, X.; Zhu, J. Multi-task generative adversarial network for retinal optical coherence tomography image denoising. *Phys. Med. Biol.* **2022**, *68*, 045002. [CrossRef] [PubMed]

51. Ahmed, H.; Zhang, Q.; Donnan, R.; Alomainy, A. Framework of Unsupervised based Denoising for Optical Coherence Tomography. In Proceedings of the International Conference on Biomedical Signal and Image Processing (ICBIP), Suzhou, China, 19–21 August 2022; pp. 19–24.

52. Ahmed, H.; Zhang, Q.; Donnan, R.; Alomainy, A. Attention Based Speckle Reduction for Optical Coherence Tomography in Ophthalmology and Dentistry. In Proceedings of the International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 5–7 November 2022; pp. 1–6.

53. Zhou, Q.; Wen, M.; Yu, B.; Lou, C.; Ding, M.; Zhang, X. Self-supervised transformer based non-local means despeckling of optical coherence tomography images. *Biomed. Signal Process. Control* **2023**, *80*, 104348. [CrossRef]

54. Gholami, P.; Roy, P.; Parthasarathy, M.K.; Lakshminarayanan, V. OCTID: Optical coherence tomography image database. *Comput. Electr. Eng.* **2020**, *81*, 106532. [CrossRef]

55. Kande, N.A.; Dakhane, R.; Dukkipati, A.; Yalavarthy, P.K. SiameseGAN: A Generative Model for Denoising of Spectral Domain Optical Coherence Tomography Images. *IEEE Trans. Med. Imaging* **2021**, *40*, 180192. [CrossRef]

56. Qiu, B.; You, Y.; Huang, Z.; Meng, X.; Jiang, Z.; Zhou, C.; Liu, G.; Yang, K.; Ren, Q.; Lu, Y. N2NSR-OCT: Simultaneous denoising and super-resolution in optical coherence tomography images using semisupervised deep learning. *J. Biophotonics* **2020**, *14*, e202000282. [CrossRef]

57. Zhou, Y.; Li, J.; Wang, M.; Zhu, W.; Peng, Y.; Chen, Z.; Wang, L.; Wang, T.; Yao, C.; Wang, T.; et al. High-Resolution Hierarchical Adversarial Learning for OCT Speckle Noise Reduction. In *MICCAI 2021: Medical Image Computing and Computer Assisted Intervention*; Springer: Cham, Switzerland, 2021; pp. 372–381.

58. Ahmed, H.; Zhang, Q.; Donnan, R.; Alomainy, A. Unsupervised Region-Based Denoising for Optical Coherence Tomography Framework. In Proceedings of the International Conference on Computational Intelligence and Applications (ICCIA), Nanjing, China, 24–26 June 2022; pp. 267–273.

59. Hao, S.; Hao, G. Research on OCT Image Processing Based on Deep Learning. In Proceedings of the International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 17–19 July 2020; pp. 208–212.

60. Abbas, Q.; Qureshi, I.; Yan, J.; Shaheed, K. Machine Learning Methods for Diagnosis of Eye-Related Diseases: A Systematic Review Study Based on Ophthalmic Imaging Modalities. *Arch. Comput. Methods Eng.* **2022**, *29*, 3861–3918. [CrossRef]
61. Syed, A.M.; Faizan, M.; Akbar, M.U.; Fatima, J. Automated Techniques for Detection and Classification of Diabetic Macular Edema: A Review. *Asian J. Eng. Sci. Technol.* **2016**, 39–43.
62. Stankiewicz, A.; Marciniak, T.; Dabrowski, A.; Stopa, M.; Marciniak, E.; Obara, B. Segmentation of Preretinal Space in Optical Coherence Tomography Images Using Deep Neural Networks. *Sensors* **2021**, *21*, 7521. [CrossRef] [PubMed]
63. Jian, Z.; Yu, L.; Rao, B.; Tromberg, B.J.; Chen, Z. Three-dimensional speckle suppression in optical coherence tomography based on the curvelet transform. *Opt. Express* **2010**, *18*, 1024103. [CrossRef]
64. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; pp. 60–65.
65. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Trans. Image Process.* **2007**, *16*, 20802095. [CrossRef]