






Data Descriptor

Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research

Prasanna Porwal ^{1,*} , Samiksha Pachade ¹ , Ravi Kamble ¹ , Manesh Kokare ¹ ,
Girish Deshmukh ², Vivek Sahasrabudde ³ and Fabrice Meriaudeau ⁴ 

¹ Center of Excellence in Signal and Image Processing, Department of Electronics and Telecommunication Engineering, Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded 431606, India; 2017pec601@sggs.ac.in (S.P.); kambleravi@sggs.ac.in (R.K.); mbkokare@sggs.ac.in (M.K.)

² Eye Clinic, Sushrusha Hospital, Nanded 431601, India; girishnanded@gmail.com

³ Department of Ophthalmology, Shankarrao Chavan Government Medical College, Nanded 431606, India; smvivek195@gmail.com

⁴ Centre for Intelligent Signal and Imaging Research, Department of Electrical & Electronic Engineering, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Malaysia; fabrice.meriaudeau@utp.edu.my

* Correspondence: porwalprasanna@sggs.ac.in

Received: 5 June 2018; Accepted: 6 July 2018; Published: 10 July 2018



Abstract: Diabetic Retinopathy is the most prevalent cause of avoidable vision impairment, mainly affecting the working-age population in the world. Recent research has given a better understanding of the requirement in clinical eye care practice to identify better and cheaper ways of identification, management, diagnosis and treatment of retinal disease. The importance of diabetic retinopathy screening programs and difficulty in achieving reliable early diagnosis of diabetic retinopathy at a reasonable cost needs attention to develop computer-aided diagnosis tool. Computer-aided disease diagnosis in retinal image analysis could ease mass screening of populations with diabetes mellitus and help clinicians in utilizing their time more efficiently. The recent technological advances in computing power, communication systems, and machine learning techniques provide opportunities to the biomedical engineers and computer scientists to meet the requirements of clinical practice. Diverse and representative retinal image sets are essential for developing and testing digital screening programs and the automated algorithms at their core. To the best of our knowledge, IDRiD (Indian Diabetic Retinopathy Image Dataset), is the first database representative of an Indian population. It constitutes typical diabetic retinopathy lesions and normal retinal structures annotated at a pixel level. The dataset provides information on the disease severity of diabetic retinopathy, and diabetic macular edema for each image. This makes it perfect for development and evaluation of image analysis algorithms for early detection of diabetic retinopathy.

Dataset: 10.21227/H25W98

Dataset License: CC-BY 4.0

Keywords: retinal fundus images; diabetic retinopathy; diabetic macular edema

1. Summary

Diabetic Retinopathy (DR) is the result of microvascular retinal changes triggered by diabetes and it is the most common leading cause of preventable blindness in the working-age population in the world [1,2]. Whereas, Diabetic Macular Edema (DME) is a complication associated with DR,

characterized by accumulation of fluid or retinal thickening that can occur at any stage of DR [3,4]. International Council of Ophthalmology (ICO) report [5] indicate that 1 out of 3 individuals affected with diabetes had some form of DR and also highlighted that 1 in 10 had vision-threatening DR. In India it is the sixth common cause of blindness [6].

DR is referred as a clinical diagnosis, depicted by the presence (see Figure 1) of one or more several retinal lesions like microaneurysms, hemorrhages, hard exudates and soft exudates [7].

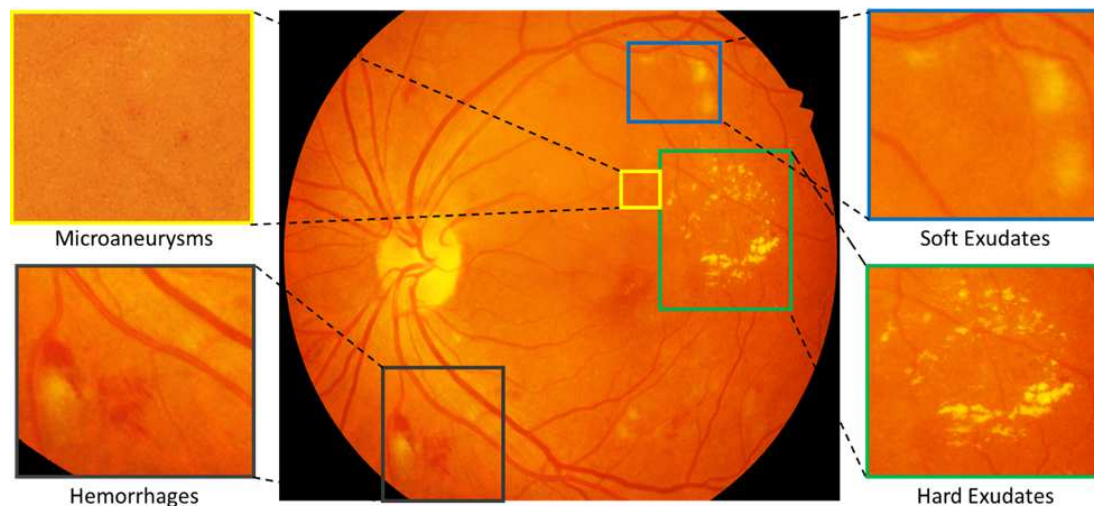


Figure 1. Color fundus photograph containing different retinal lesions associated with diabetic retinopathy. Enlarged parts illustrating presence of Microaneurysms, Soft Exudates, Hemorrhages and Hard Exudates.

Early diagnosis and treatment of DR can prevent vision loss [8]. Hence, diabetic patients are referred to do a regular biannual or annual follow-up and frequent consultation for the screening of their retina [9]. The elimination of preventable visual impairment is mainly dependent on the pool of expert clinicians and basic health care infrastructure essential for the treatment of the eye [10,11]. In the Indian subcontinent, against national eye care experts: population ratio of 1:107,000, in various regions this ratio is 1:9000 whereas in some other parts there is only one eye care expert for 608,000 population [12,13]. Due to the large number of people that require a continuous follow-up and shortage of ophthalmologists, management of DR needs attention to develop computer-aided diagnosis tool [14,15]. The recent technological advances in computing power, communication systems, and machine learning techniques provide opportunities to the biomedical engineers and computer scientists to meet the requirements of clinical practice [16,17]. The raw images with ground truths facilitates the scientific community for development, validation, comparison and aid in the further improvement of DR lesion detection algorithms used in clinical application [18]. Precise pixel level annotation of abnormalities associated with DR like microaneurysms, soft exudates, hard exudates and hemorrhages is invaluable resource for performance evaluation of individual lesion segmentation techniques. Whereas, the reliable information about disease severity level of DR, and DME is useful in development and evaluation of image analysis and retrieval algorithms for early detection of the disease [19].

This dataset was available as a part of “Diabetic Retinopathy: Segmentation and Grading Challenge (<http://biomedicalimaging.org/2018/challenges/>)” organized in conjunction with IEEE International Symposium on Biomedical Imaging (ISBI-2018), Washington D.C. The data challenge was hosted on Grand Challenges in Biomedical Imaging Platform [20]. Information about specifications and data accessibility is provided in the Table 1.

Table 1. Specifications Table.

Subject area	Biomedical Imaging, Ophthalmology
More specific subject area	Retinal image analysis for detection of DR and DME
Type of data	Image, CSV
How data was acquired	Retinal Fundus Camera. Model: Kowa VX-10 α
Data format	Raw and Manual Annotations
Experimental factors	Mydriasis with one drop of tropicamide at 0.5% concentration
Experimental features	Retinal image of humans affected by diabetes was captured with 39 mm distance between lenses and examined eye using non-invasive fundus camera having xenon flash lamp.
Data source location	Eye Clinic, Sushrusha Hospital Building, Nanded, (M.S.), India
Data accessibility	https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid

2. Data Description

The IDRiD dataset, is a new publicly available retinal fundus image database consisting of 516 images categorized in two parts:

- Retinal images with the signs of DR and/or DME.
- Normal retinal images (without signs of DR and/or DME).

The dataset provides ground truths associated with the signs of Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) and normal retinal structures given below and described as follows:

- Pixel level annotations of typical diabetic retinopathy lesions and optic disc.
- Image level disease severity grading of diabetic retinopathy, and diabetic macular edema.
- Optic Disc and Fovea center co-ordinates.

2.1. Pixel Level Annotated Data

This dataset consists of 81 color fundus images with signs of DR and 164 without signs of DR. Precise pixel level annotation as shown in Figure 2 of abnormalities associated with DR like microaneurysms (MA), soft exudates (SE), hard exudates (EX) and hemorrhages (HE) is provided as a binary mask for performance evaluation of individual lesion segmentation techniques. It includes color fundus images (.jpg files) and separate binary masks for each lesion type (.tif files). Along with the lesion masks, it also consist of optic disc (OD) mask for all 81 images (see example in Figure 6).

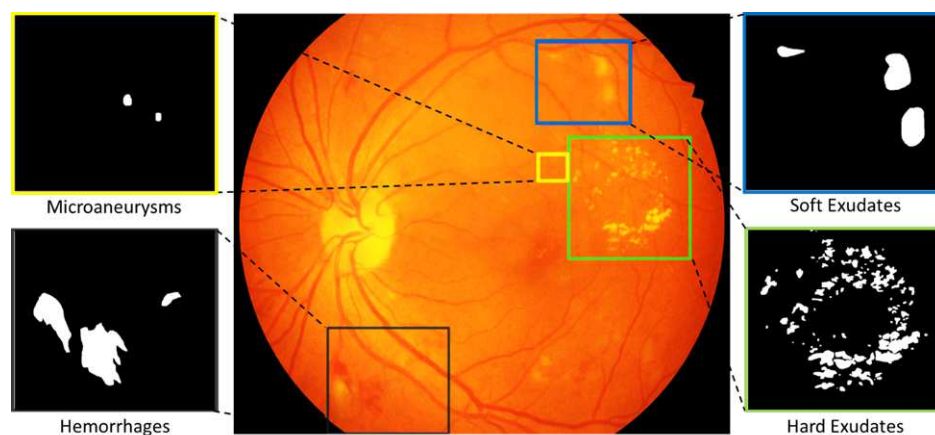


Figure 2. Color fundus photograph containing different retinal lesions associated with diabetic retinopathy. Enlarged parts illustrating sample annotations of Microaneurysms, Soft Exudates, Hemorrhages and Hard Exudates.

2.2. Image Level Disease Grading

The medical experts graded the full set of 516 images with a variety of pathological conditions of DR and DME. The dataset is divided into training and testing set comprising of 413 (80%) and 103 (20%) images respectively by maintaining appropriate mixture of disease stratification. Similarly, the expert labels of DR and DME severity level for the dataset are provided in two CSV files *a.IDRiD_DiseaseGrading_TrainingLabels.CSV* and *b.IDRiD_DiseaseGrading_TestingLabels.CSV*.

Figure 3 illustrates the information available in both CSV files with each column description given as follows:

- A. *Image No*: Name (serial number) of deidentified and renamed patient image.
- B. *DR Grade*: DR severity level in range 0 (No apparent DR) to 4 (Severe DR).
- C. *Risk of DME*: Macular edema severity level in range 0 (No DME) to 2 (Severe DME).

	A	B	C
1	Image No	DR Grade	Risk of DME
2	IDRiD_001	3	1
3	IDRiD_002	2	2
4	IDRiD_003	4	2
5	IDRiD_004	2	1
6	IDRiD_005	1	0

Figure 3. Sample DR and DME expert labels in CSV file.

2.3. Optic Disc and Fovea Center Location

Along with the annotations presented above, the dataset provides center pixel locations of optic disc [OD_x, OD_y] and fovea [F_x, F_y] for all 516 images as shown in Figure 4.

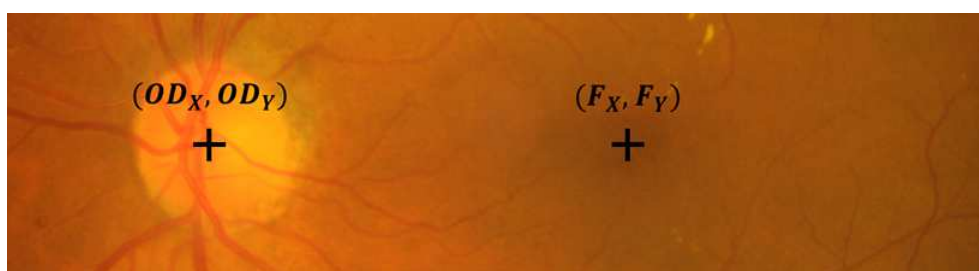


Figure 4. Sample cropped image from the IDRiD database illustrating the OD and fovea center location.

The dataset is divided into training and testing set comprising of 413 (80%) and 103 (20%) images respectively. Center pixel markups for OD and Fovea are made available in two separate folders. One folder consist of OD center markup files *IDRiD_OD_Center_Training_Set_Markups.CSV* and *IDRiD_OD_Center_Testing_Set_Markups.CSV*; and other folder consist of fovea center markup files *IDRiD_Fovea_Center_Training_Set_Markups.CSV* and *IDRiD_Fovea_Center_Testing_Set_Markups.CSV*. Each CSV file consist of three columns representing image no, X co-ordinate and Y co-ordinate. Where, X and Y co-ordinates are of center pixel location of OD/Fovea in the image.

Table 2 summarizes the available data with its description, quantity of data and different file types.

Table 2. List of data available in the created dataset.

Data	Description	Quantity	Data Type	File Format
Color Fundus Images of Retina	Raw Data	516	Image	jpg
Disease Severity Grading of DR and DME	Image level grading	516	Tabular	CSV
Center co-ordinates of OD and Fovea	Manual center co-ordinates	516	Tabular	CSV
Binary Masks of different lesions	Precise pixel level manual annotation	81	Image	tif
Binary Masks of optic disc	Precise pixel level manual annotation	81	Image	tif

3. Experimental Design, Materials and Methods

3.1. Ethics Statement

Informed consent was received from the patients of this study. Appropriate care has been taken for privacy protection of patients as per the guidelines from the local ethics committee and ethics of the clinical practices and medical research. The dataset has also received approval from the local research ethics committee of Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded (M.S.), India.

Details regarding the data acquisition and annotation is as follows:

3.2. Data Acquisition

Retinal fundus imaging is non-invasive and painless mean to screen retina [9,21]. The fundus images in IDRiD database were acquired from an Eye Clinic located in Nanded, (M.S.), India. Retinal images of humans affected by diabetes were captured with 39 mm distance between lenses and examined eye using non-invasive fundus camera having xenon flash lamp. The details of pretreatment of samples and camera specifications are as follows:

- *Pretreatment of Samples:* All the subjects in the dataset had undergone mydriasis prior to capturing of images. Mydriasis is process of pupil dilation which was done with one drop of tropicamide at 0.5% concentration.
- *Fundus Camera Specifications:* Images were acquired using a Kowa VX-10 α digital fundus camera with 50° field of view (FOV). The images have resolution of 4288 \times 2848 pixels and are stored in jpg file format. The size of each image is about 800 KB.
- *Data Quality:* The dataset is formed by extracting 516 images from the thousands of examinations done during the period 2009–2017. Experts verified that all images are of adequate quality, clinically relevant, that no image is duplicated and that a reasonable mixture of disease stratification representative of diabetic retinopathy (DR) and diabetic macular edema (DME) is present.

3.3. Annotation of Images

This dataset provides three type of annotations, namely pixel level annotations of lesions, image level DR and DME grading and center markups for OD and Fovea. Details of the ground truths for each of the three types is explained as follows:

- **Pixel Level Annotation:** Initially, all observers were trained by expert ophthalmologists for the identification of individual lesion. An image processing expert chose 81 images with contextual data comprising soft exudates, hard exudates, microaneurysms, and hemorrhages. The pixel level annotation is done by a master's student using special software developed by ADCIS [22] specifically for annotation purposes. Figure 5 shows the sample image from the database and manually drawn contours. Later the markings on each of these images were reviewed by two retinal specialists, and they were finalized when the necessary consensus was reached. The final

groundtruth images for all lesions and optic disc are shown in Figure 6. Similar pixel level lesion annotations are available in the E-Optha dataset [23].

- **DR and DME Grading:** The medical experts graded full set of 516 images with variety of pathological conditions of DR and DME. The diabetic retinal images were classified into separate groups ranging from 0 (No apparent DR) to 4 (Severe DR) according to the International Clinical Diabetic Retinopathy Scale [24], similar to existing Kaggle DR Dataset [25]. The risk of macular edema can be determined by the presence of exudates [26], severity grading of DME is done based on occurrences of hard exudates near to macula center as per the definitions provided by Messidor database [27].
- **Optic Disc and Fovea Center Location Markup:** The OD and fovea center markups are done by a master's and PhD student. The final center co-ordinates are obtained by computing average of two locations. The averaged markups were further verified by a retinal expert.

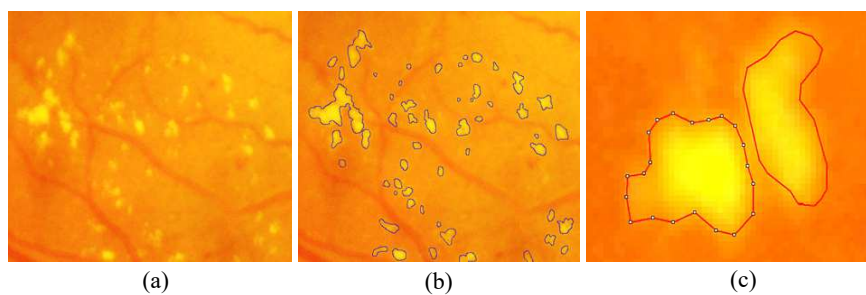


Figure 5. Enlarged part of fundus image containing hard exudates from our database: (a) presence of hard exudates; (b) the manual mark-ups for hard exudates; and (c) markup contour from annotator display.

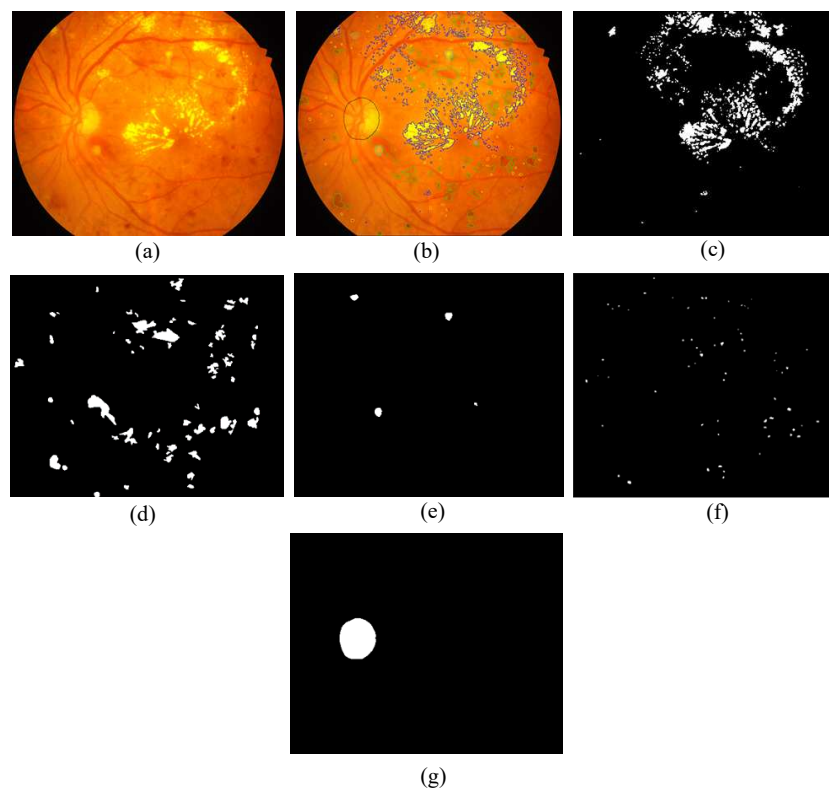


Figure 6. Retinal photograph and different annotations: (a) sample fundus image from the presented dataset; (b) groundtruth markups of annotator; (c–g) sample groundtruths of hard exudates, hemorrhages, soft exudates, microaneurysms and optic disc respectively.

Author Contributions: Conceptualization, P.P. and M.K.; Methodology, P.P., S.P., R.K., G.D. and V.S.; Resources, G.D.; Data Curation, S.P. and P.P.; Disease Grading, G.D. and V.S.; Software for OD and Fovea center markups, R.K.; Annotation Validation, G.D. and V.S.; Investigation, M.K. and F.M.; Writing—Original Draft Preparation, P.P.; Writing—Review & Editing, M.K. and F.M.; Visualization, P.P. and R.K.; Supervision, M.K. and F.M.; Project Administration, M.K. and F.M.

Acknowledgments: We would like to thank the Technical Education Quality Improvement Program (TEQIP-II), the World Bank project, for supporting the project and providing a state of the art Center of Excellence in Signal and Image Processing research lab. We express our gratitude towards Board of Management of Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded for sponsoring the organization of the challenge based on the IDRiD dataset.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Reichel, E.; Salz, D. Diabetic retinopathy screening. In *Managing Diabetic Eye Disease in Clinical Practice*; Springer: Berlin, Germany, 2015; pp. 25–38.
- International Diabetes Federation (IDF). *IDF Diabetes Atlas*; IDF: Brussels, Belgium, 2017.
- Bandello, F.; Parodi, M.B.; Lanzetta, P.; Loewenstein, A.; Massin, P.; Menchini, F.; Veritti, D. Diabetic macular edema. In *Macular Edema*; Karger Publishers: Basel, Switzerland, 2010; Volume 47, pp. 73–110.
- Ciulla, T.A.; Amador, A.G.; Zinman, B. Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies. *Diabetes Care* **2003**, *26*, 2653–2664. [[CrossRef](#)] [[PubMed](#)]
- International Council of Ophthalmology (ICO). *Guidelines for Diabetic Eye Care*, 2nd ed.; International Council of Ophthalmology (ICO): San Francisco, CA, USA, 2017.
- Bourne, R.R.; Stevens, G.A.; White, R.A.; Smith, J.L.; Flaxman, S.R.; Price, H.; Jonas, J.B.; Keeffe, J.; Leasher, J.; Naidoo, K.; et al. Causes of vision loss worldwide, 1990–2010: A systematic analysis. *Lancet Glob. Health* **2013**, *1*, e339–e349. [[CrossRef](#)]
- Wong, T.Y.; Cheung, C.M.G.; Larsen, M.; Sharma, S.; Simó, R. Diabetic Retinopathy. *Nat. Rev. Disease Prim.* **2016**, *2*, 16012. [[CrossRef](#)] [[PubMed](#)]
- Abràmoff, M.D.; Garvin, M.K.; Sonka, M. Retinal imaging and image analysis. *IEEE Rev. Biomed. Eng.* **2010**, *3*, 169–208. [[CrossRef](#)] [[PubMed](#)]
- Jelinek, H.; Cree, M.J. *Automated Image Detection of Retinal Pathology*; CRC Press: Boca Raton, FL, USA, 2009.
- Jones, S.; Edwards, R. Diabetic retinopathy screening: A systematic review of the economic evidence. *Diabet. Med.* **2010**, *27*, 249–256. [[CrossRef](#)] [[PubMed](#)]
- Lin, S.; Ramulu, P.; Lamoureux, E.L.; Sabanayagam, C. Addressing risk factors, screening, and preventative treatment for diabetic retinopathy in developing countries: A review. *Clin. Exp. Ophthalmol.* **2016**, *44*, 300–320. [[CrossRef](#)] [[PubMed](#)]
- Raman, R.; Gella, L.; Srinivasan, S.; Sharma, T. Diabetic retinopathy: An epidemic at home and around the world. *Indian J. Ophthalmol.* **2016**, *64*, 69–75. [[CrossRef](#)] [[PubMed](#)]
- Porwal, P.; Pachade, S.; Kokare, M.; Deshmukh, G.; Sahasrabuddhe, V. Automatic Retinal Image Analysis for the Detection of Diabetic Retinopathy. In *Biomedical Signal and Image Processing in Patient Care*; IGI Global: Hershey, PA, USA, 2018; pp. 146–161.
- Ting, D.S.W.; Cheung, G.C.M.; Wong, T.Y. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: A review. *Clin. Exp. Ophthalmol.* **2016**, *44*, 260–277. [[CrossRef](#)] [[PubMed](#)]
- Walter, T.; Klein, J.C.; Massin, P.; Erginay, A. A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina. *IEEE Trans. Med. Imaging* **2002**, *21*, 1236–1243. [[CrossRef](#)] [[PubMed](#)]
- Shortliffe, E.H.; Blois, M.S. The computer meets medicine and biology: Emergence of a discipline. In *Biomedical Informatics*; Springer: New York, NY, USA, 2006; pp. 3–45.
- Patton, N.; Aslam, T.M.; MacGillivray, T.; Deary, I.J.; Dhillon, B.; Eikelboom, R.H.; Yogesana, K.; Constable, I.J. Retinal image analysis: Concepts, applications and potential. *Prog. Retin. Eye Res.* **2006**, *25*, 99–127. [[CrossRef](#)] [[PubMed](#)]

18. Trucco, E.; Ruggeri, A.; Karnowski, T.; Giancardo, L.; Chaum, E.; Hubschman, J.P.; Al-Diri, B.; Cheung, C.Y.; Wong, D.; Abramoff, M.; et al. Validating retinal fundus image analysis algorithms: Issues and a proposal. *Investig. Ophthalmol. Vis. Sci.* **2013**, *54*, 3546–3559. [[CrossRef](#)] [[PubMed](#)]
19. Porwal, P.; Pachade, S.; Kamble, R.; Kokare, M.; Deshmukh, G.; Sahasrabuddhe, V.; Meriaudeau, F. Indian Diabetic Retinopathy Image Dataset (IDRI). *IEEE DataPort* **2018**, doi:10.21227/H25W98. [[CrossRef](#)]
20. Porwal, P.; Pachade, S.; Kamble, R.; Kokare, M.; Deshmukh, G.; Sahasrabuddhe, V.; MacGillivray, T.; Sidibé, D.; Giancardo, L.; Quéllec, G.; et al. Diabetic Retinopathy Segmentation and Grading Challenge. IEEE ISBI Challenge. 2018. Available online: <https://idrid.grand-challenge.org/> (accessed on 2 July 2018).
21. Patton, N.; Aslam, T.; MacGillivray, T.; Pattie, A.; Deary, I.J.; Dhillon, B. Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: A rationale based on homology between cerebral and retinal microvasculatures. *J. Anat.* **2005**, *206*, 319–348. [[CrossRef](#)] [[PubMed](#)]
22. Advanced Concepts in Imaging Software (ADCIS). Aphelion Image Annotator. ADCIS France. 2018. Available online: <http://www.adcis.net/en/Image-Processing-And-Analysis-Software-And-Custom-Engineering-Developments.html> (accessed on 2 July 2018).
23. Decencièrre, E.; Cazuguel, G.; Zhang, X.; Thibault, G.; Klein, J.C.; Meyer, F.; Marcotegui, B.; Quéllec, G.; Lamard, M.; Danno, R.; et al. TeleOphta: Machine learning and image processing methods for teleophthalmology. *Irbm* **2013**, *34*, 196–203. [[CrossRef](#)]
24. Wu, L.; Fernandez-Loaiza, P.; Sauma, J.; Hernandez-Bogantes, E.; Masis, M. Classification of diabetic retinopathy and diabetic macular edema. *World J. Diabetes* **2013**, *4*, 290. [[CrossRef](#)] [[PubMed](#)]
25. Cuadros, J.; Bresnick, G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. *J. Diabetes Sci. Technol.* **2009**, *3*, 509–516. [[CrossRef](#)] [[PubMed](#)]
26. Giancardo, L.; Meriaudeau, F.; Karnowski, T.P.; Li, Y.; Garg, S.; Tobin, K.W., Jr.; Chaum, E. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Med. Image Anal.* **2012**, *16*, 216–226. [[CrossRef](#)] [[PubMed](#)]
27. Decencièrre, E.; Zhang, X.; Cazuguel, G.; Lay, B.; Cochener, B.; Trone, C.; Gain, P.; Ordonez, R.; Massin, P.; Erginay, A.; et al. Feedback on a publicly distributed image database: The Messidor database. *Image Anal. Stereol.* **2014**, *33*, 231–234. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).