DR Dataset Analysis

1.) Diaretdb0:

a.) No. of Classes : 5 (Classes are based upon types of lesions in images)

b.) Images in each class :

Total Images = 130
Images with Microaneurysms = 106
Images with Haemorrhages = 80
Images with Hard Exudates = 71
Images with Soft Exudates = 41
Images with Neovascularization = 20

c.) Type of task : Segmentation/classification purposes

d.) Folder structure :

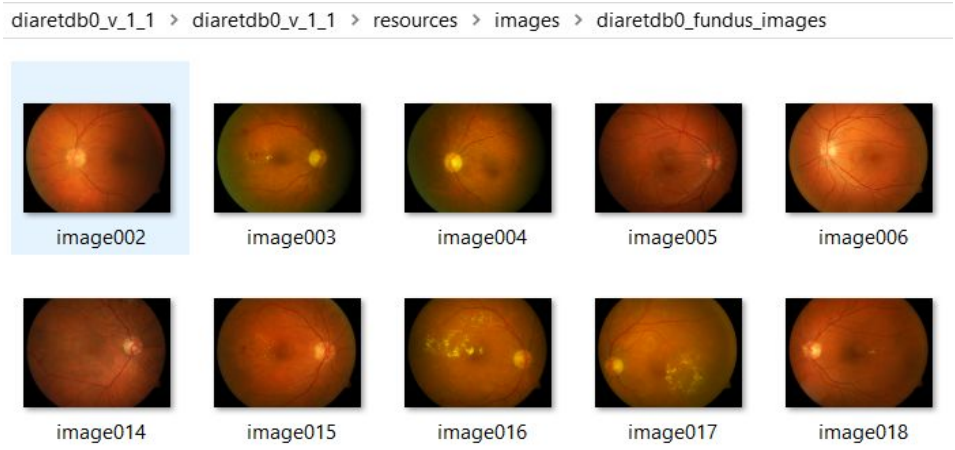


Fig 1 : Destination of the main folder that includes images.

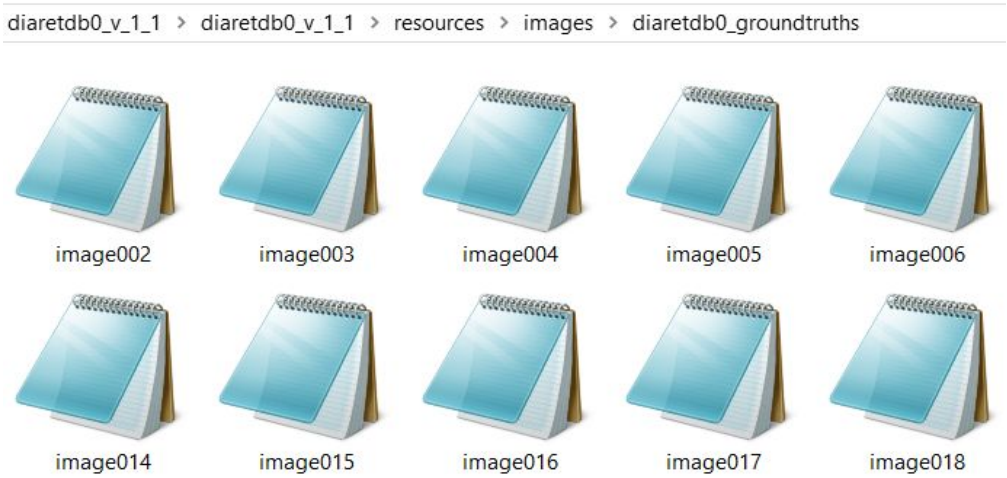


Fig 2 : Destination of the main folder that includes ground truths.

e.) Annotation format and example :

Annotation format = redsmalldots n/a hardexudates n/a neovascularisation

f.) Image with its ground truth :

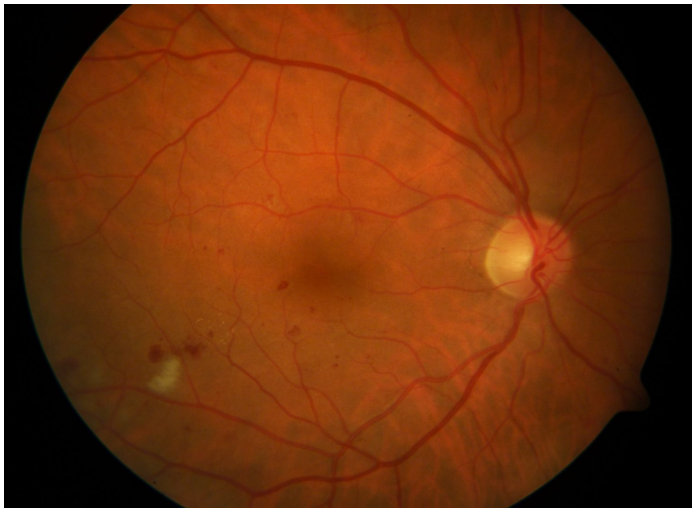


Fig 3 : Sample fundus image from DiaretDB0

Fig 4 : Ground truth for image in Fig 3

Dataset LINK : https://www.it.lut.fi/project/imageret/diaretdb0/

2.) Diaretdb1:

a.) No. of Classes : 4 (Classes are based upon types of lesions in images)

b.) Images in each class :

Total images = 89
Images with Microaneurysms = 72
Images with Haemorrhages = 51
Images with Hard Exudates = 47
Images with Soft Exudates = 36

c.) Type of task : Segmentation/classification purposes

d.) Folder structure :

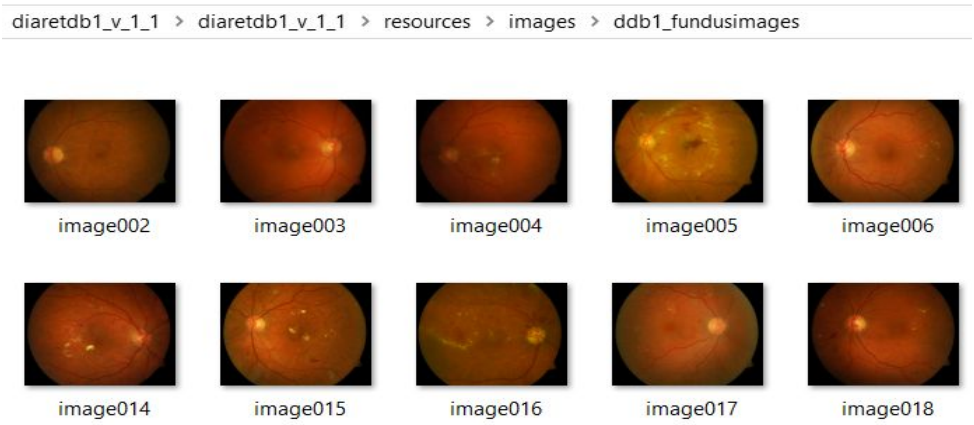


Fig 5 : Destination of the main folder that includes images.

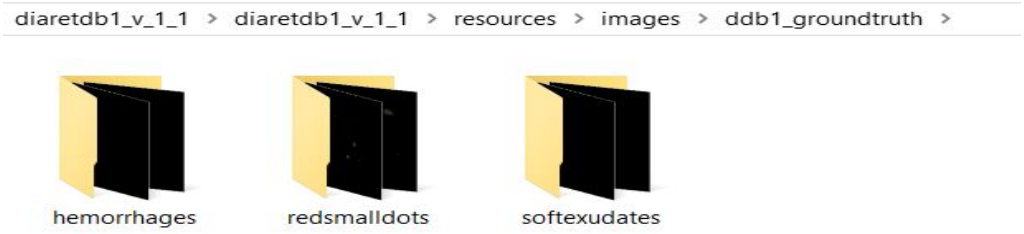


Fig 6 : Destination of the main folder that includes annotation files.

e.) **Annotation format and example :** Annotation is provided in the form of masks created according to different lesion classes present in each image.

f.) **Image with its ground truth :**

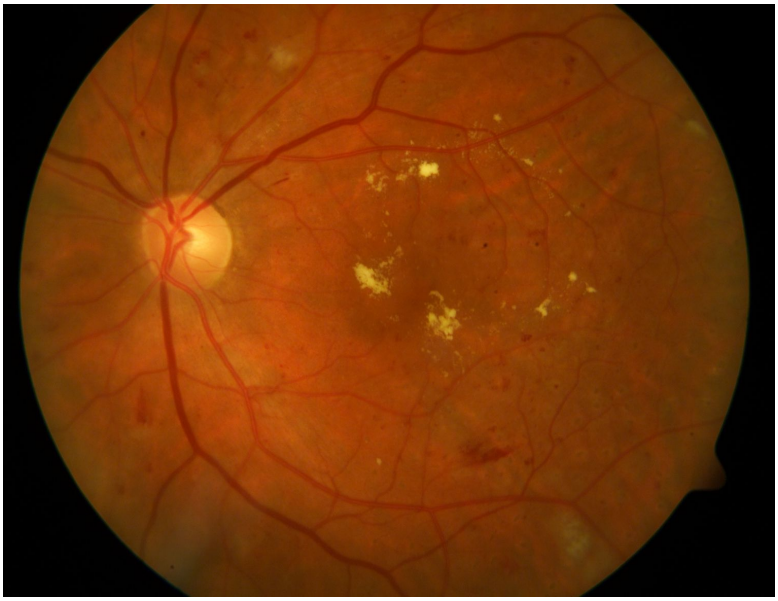
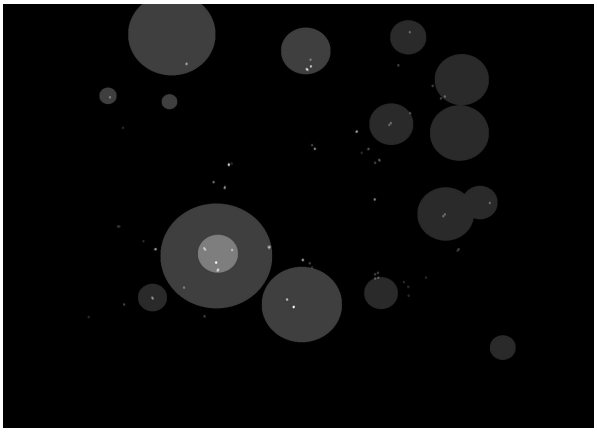
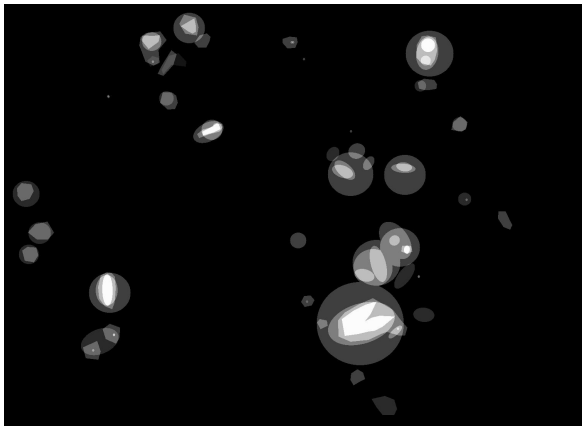


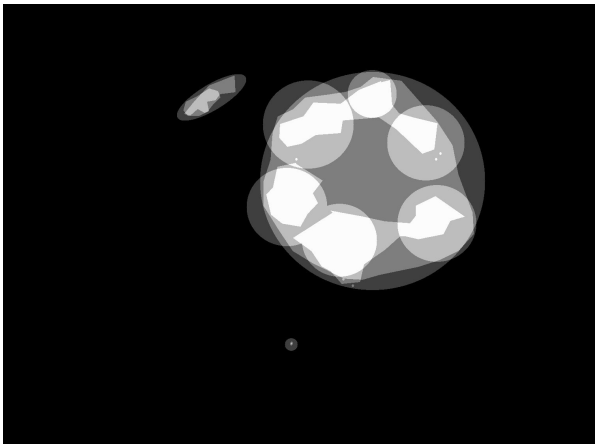
Fig 7 : Sample fundus image from DiaretDB1



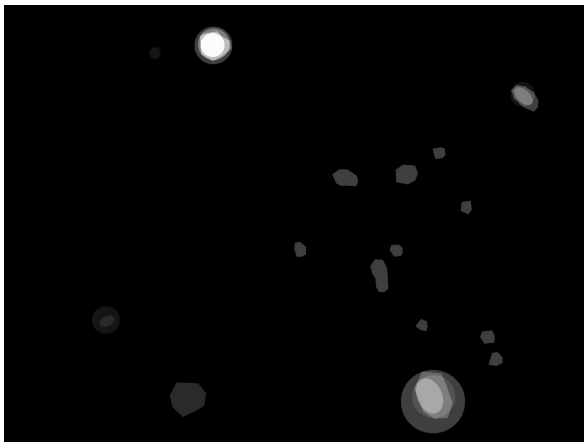
(a)



(b)



(c)



(d)

Fig 8 : ‘a’ describes the mask for microaneurysm, ‘b’ for haemorrhages, ‘c’ for hard exudates, ‘d’ for soft exudates.

Dataset LINK : https://www.it.lut.fi/project/imageret/diaretdb1/

3.) Messidor:

a.) No. of classes : 4

b.) Images in each class :

Total Images = 1200

Based of Retinopathy grading :

No. of Grade 0 Images = 546

No. of Grade 1 Images = 153

No. of Grade 2 Images = 247
No. of Grade 3 Images = 254

Based on Risk of Macular Edema:

Images having Risk factor 0 = 974
Images having Risk factor 1 = 75
Images having Risk factor 2 = 151

c.) Type of task : Classification purposes

d.) Folder Structure :

Messidor dataset is available in 3 groups each having 400 images. Each group has been divided into 4 folders each having 100 images along with its annotation file.

Here as seen, Base11 is one of the 4 folders from Group 1 of the dataset. It has 100 images and an annotation file for those images.

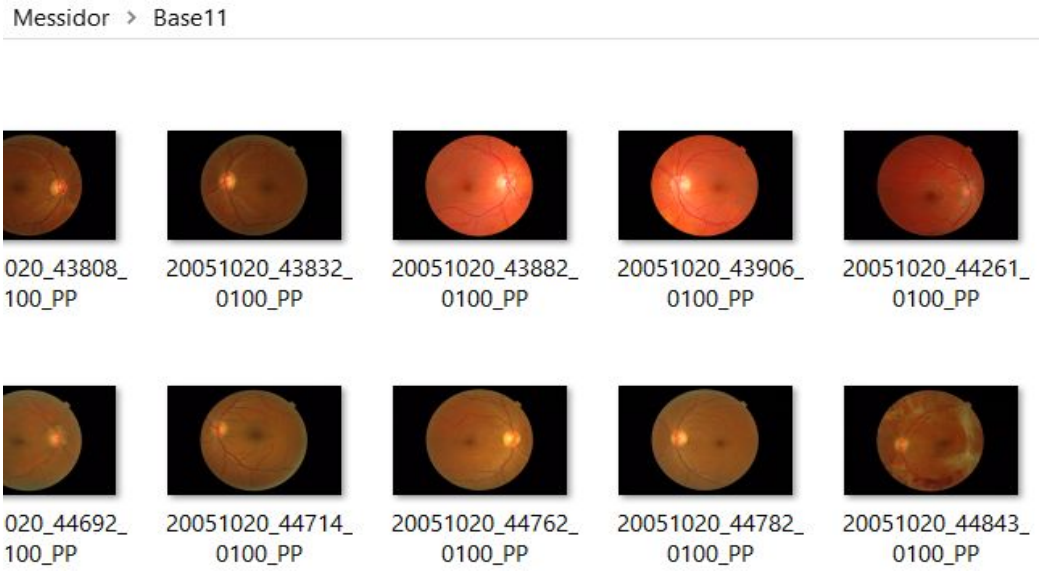


Fig 9 : Destination of the main folder that includes images in Messidor

e.) Annotation format and example :

An annotation excel file is provided for each folder.

Format:

[Image_001.tiff Dept. Retinopathy-Grade Risk of Edema]

f.) Image with its ground truth :

Fig 10 : Sample fundus image from Messidor

Image name	Ophthalmologic department	Retinopathy grade	Risk of macular edema
20051020_45137_0100_PP.tif	Service Ophtalmologie Lariboisière	3	2

Fig 11 : Ground truth describing the image in Fig 10.

Dataset LINK : http://www.adcis.net/en/third-party/messidor/

4.) Messidor-2

a.) No. of classes: 5

b.) Images in each class:

Total Images = 1748

Based on Retinopathy Grading:

No. of Grade 0 Images = 1017
No. of Grade 1 Images = 270
No. of Grade 2 Images = 347
No. of Grade 3 Images = 75
No. of Grade 4 Images = 35

Based on Risk of Macular Edema:

Images having Macular Edema = 151
Images not having Macular Edema = 1593

c.) Type of task : Classification purposes

d.) Folder Structure :

Messidor-2 is divided into 4 zipped groups that can be downloaded separately.

e.) Annotation format and example:

image_id	adjudicated_dr_grade	adjudicated_dme	adjudicated_gradable
20051020_43808_0100_PP.png	0	0	1
20051020_43832_0100_PP.png	1	0	1
20051020_43882_0100_PP.png	1	0	1
20051020_43906_0100_PP.png	2	1	1
20051020_44261_0100_PP.png	0	0	1
20051020_44284_0100_PP.png	0	0	1
20051020_44338_0100_PP.png	0	0	1
20051020_44349_0100_PP.png	2	0	1
20051020_44400_0100_PP.png	0	0	1
20051020_44431_0100_PP.png	0	0	1
20051020_44598_0100_PP.png	2	0	1
20051020_44636_0100_PP.png	2	0	1
20051020_44692_0100_PP.png	0	0	1
20051020_44714_0100_PP.png	0	0	1

Fig 12 : Example of Annotation file

Dataset LINK : http://www.adcis.net/en/third-party/messidor2/

5.) DRIONS_DB

a.) **No. of classes :** Images are not grouped into any classes. This dataset is basically used for optic disk detection.

b.) **Total Images :** 110 (All images are optic disk centered.

c.) **Type of task :** Used for optic disk detection

d.) Folder Structure :

Fig 13 : Destination of the main folder that includes images in Drions_DB.

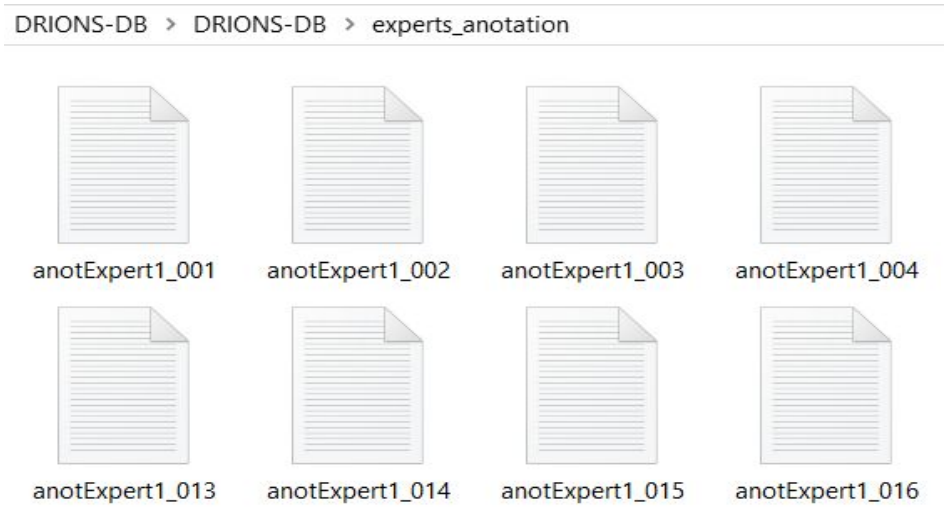


Fig 14 : Destination of the main folder that includes annotation files for Drions_DB .

e.) Annotation format and example :

The annotations for each image is a text file having 36 lines of text that denotes 36 points in the 2d plane of the image. These points correspond to detecting the optic disk in the retinal image.

Example:

1.) 328, 245

2.) 312, 231
3.) 365, 257
.
.
.
35.) 278, 180
36.) 245, 198

f.) Image with Ground Truth :

anotExpert1_037 - Notepad

File Edit Format View Help

```
348.566 , 217.938
345.812 , 226.116
342.923 , 234.300
338.135 , 241.611
332.897 , 248.945
324.875 , 253.019
316.675 , 256.114
308.217 , 257.953
299.640 , 259.817
290.862 , 259.091
282.521 , 256.345
275.013 , 251.966
268.099 , 246.899
261.391 , 241.405
254.498 , 235.373
250.494 , 227.199
246.953 , 218.601
245.870 , 209.311
246.209 , 199.889
247.994 , 190.513
252.269 , 181.961
257.820 , 174.220
264.495 , 167.426
272.858 , 162.923
281.509 , 159.497
290.338 , 156.554
299.640 , 156.866
308.823 , 157.233
317.912 , 159.108
325.647 , 164.265
332.561 , 170.077
338.119 , 177.023
343.200 , 184.162
346.593 , 192.221
350.046 , 200.423
349.581 , 209.311
```

(a) **(b)**

Fig 14 : ‘a’ describes the fundus image and ‘b’ describes the annotation for ‘a’

Dataset LINK : http://www.ia.uned.es/~ejcarmona/DRIONS-DB.html

6.) HRF (High Resolution Fundus Imaging)

a.) No. of classes : 3 (Here the total images are divided into 3 classes based on diabetic retinopathy, glaucoma, and normal image.

b.) Images in each class :

Total Images = 45

Images consisting Diabetic retinopathy = 15

Images consisting Glaucoma = 15

Normal Images = 15

c.) Type of Task : Segmentation/Classification purposes

i.) Here the classification is to be done on the basis of the 3 classes listed above.

ii.) Segmentation of lesions and optic disc is possible in all the images.

d.) Folder Structure:

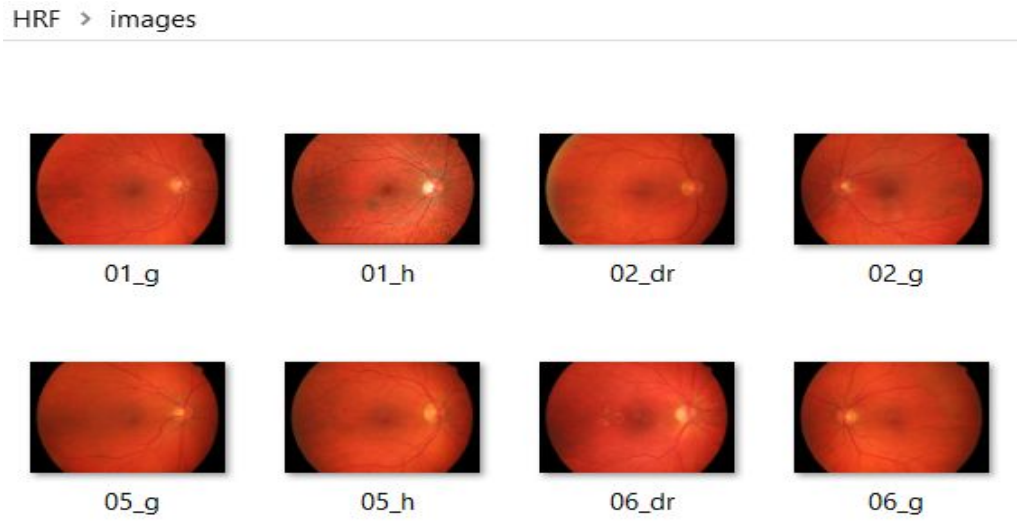


Fig 15 : Destination of the main folder that includes images in HRF dataset.

e.) Annotation Format :

There isn't any seperate annotation given in the HRF dataset. But the images have their classes labeled in their name itself.

Example :

"01_dr" = 1st image classified as having DR.

"01_g" = 1st image classified as having Glaucoma.

"01_h" = 1st image classified as normal retina.

f.) Image with Ground Truth:

Fig 16 : Retinal Fundus image having retinopathy with its vessel flow

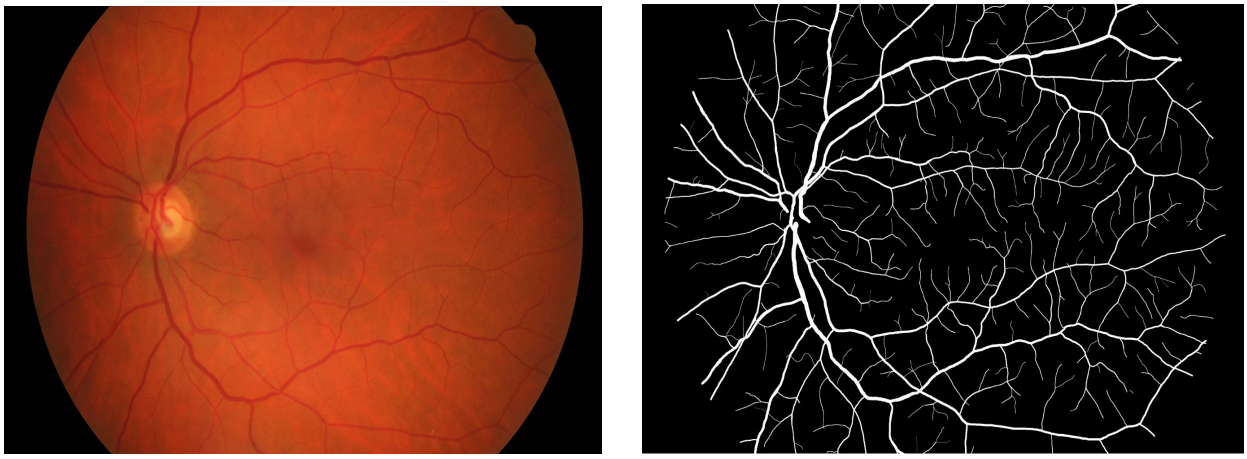


Fig 17 : Retinal Fundus image having Glaucoma with its vessel flow

Fig 18 : Retinal Fundus image that is normal with its vessel flow.

Dataset LINK : https://www5.cs.fau.de/research/data/fundus-images/

7.) INSPIRE-STEREO

a.) No. of classes : There are no classes used to distribute this data.

b.) Total Images : 30

c.) Type of task : Segmentation purposes.

d.) Folder Structure :

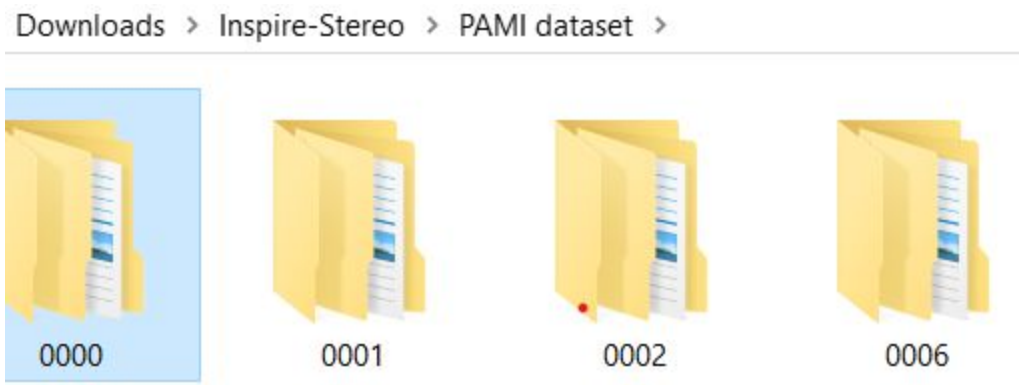


Fig 19 : Destination of the main folder that includes images in INSPIRE-STEREO.

e.) **Annotation format :** There isn't any annotation given in this dataset. There are fundus images focussed around optic disk and a depth reference of the optic disk for each retinal image. This depth reference is taken into consideration during Optical Coherence Tomography for determining the cross sectional region of the macula and optic disc.

f.) **Image with depth reference:**

(a) **(b)**

Fig 20 : 'a' describes the optic focussed retinal image and 'b' is the depth reference image of 'a'.

Dataset LINK : https://medicine.uiowa.edu/eye/inspire-datasets

8.) **INSPIRE-AVR**

a.) **No. of classes :** Images aren't segmented based on any classes.

b.) **Total Images :** 40

c.) **Type of task :** Segmentation purposes

d.) **Folder structure :**

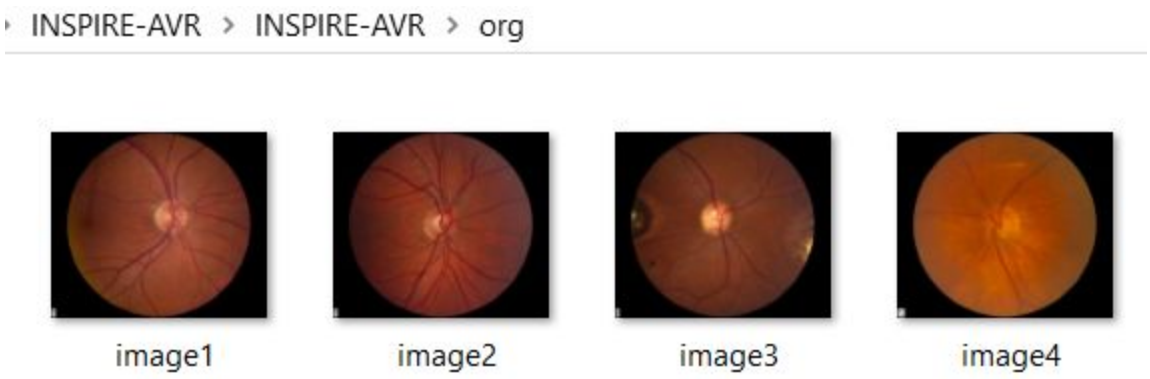


Fig 21 : Destination of the main folder that includes images in INSPIRE-AVR.

e.) **Annotation format :** This dataset is annotated with an AVR (Arterio Venous Ratio). 2 simultaneous AVR are calculated for each image and provided in the Ground Truth file.

f.) **Image with its AVR :**

Fig 22 : Optic centered image taken from INSPIRE-AVR

Annotation :

Image	AVR-obs1	AVR-obs2
image19.jpg	0.67	0.63

Dataset LINK : https://medicine.uiowa.edu/eye/inspire-datasets

9.) HEI-MED (Hamilton Eye Institute Macular Edema Dataset)

a.) No. of classes :

i.) The classes in this dataset aren't constant but the common things that the dataset is divided upon are exudates and other lesions that include cotton wool spots, drusens, microaneurysms, and other fluid occurring regions.

ii.) Basically the most common classes include 3 names(exudates, microaneurysms, drusens).

iii.) There are a lot of other classes describing the kind of retinopathy(Proliferative or not), extent of macular edema, pigmentations, fluid leakages, etc

b.) Total Images = 169 (All have their separate ground truth file)

c.) Type of task : Majorly classification purposes**.**

d.) Folder Structure :

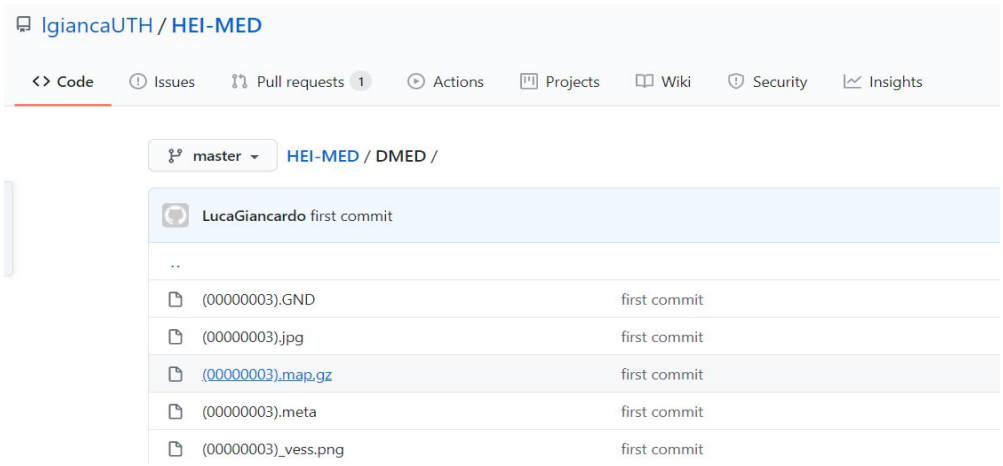


Fig 23 : Destination of the main folder that includes images in HEI-MED.

e.) Annotation format : There is a separate annotation(ground truth file) provided for each image. Each image consists of all the classes in that image and each class is listed down in a different line.

Example :

```
3
MicroAneurysm 0
MicroAneurysm 0
MicroAneurysm 0
0
38
9
NPDR
7: NPDR Severe + CSME
6: NPDR Severe - CSME
5: NPDR Moderate + CSME
4: NPDR Moderate - CSME
3: NPDR Mild/Minimal + CSME
2: NPDR Mild/Minimal - CSME
1: Absent
0: Unknown
PDR
5: PDR + CSME
4: PDR - CSME
3: PDR + HRC + CSME
2: PDR + HRC - CSME
1: Absent
0: Unknown
```

Fig 24 : Image describing a sample annotation file.

Each line firstly defines the number of lesions of a particular type followed by the name of the lesion upto that number. Here 3 in the 1st line defines 3 Microaneurysms followed by 38 in the 6th line that defines 38 different classes and their names.

f.) Image with its ground truth :

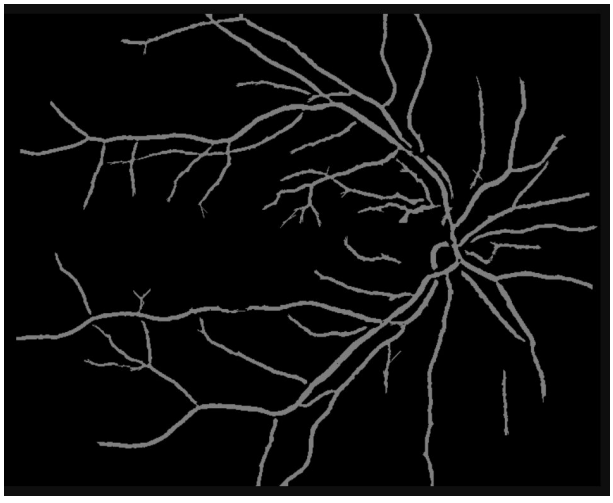
(a)

```
5
Other Bright Lesion 1
Other Bright Lesion 1
Other Bright Lesion 1
Other Bright Lesion 1
Other Bright Lesion 1
0
0
0
NA
NA
NA
NA
NA
NA
NA
```

(b)

```
~ImageName:~(00000168).jpg
~PatientGender~F
~PatientRace~CAUCASIAN
~PatientDOB~10/21/1954
~QualityValue~0.839159
~DiabetesType~TYPE II
~ONrow~923
~ONcol~1625
```

(c)



(d)

Fig 25 : ‘a’ contains a sample fundus image, ‘b’ is the annotation file for ‘a’, ‘c’ describes the patient’s metadata, ‘d’ is the vessel flow of ‘a’.

Dataset LINK : https://github.com/IgiancaUTH/HEI-MED

10.) STARE Dataset

a.) **No. of classes :** 39 classes (Here the experts have annotated the data into these many classes defining each and every problem causing in the retina.

b.) **Total Images =** 400

c.) **Type of task :** Classification and Segmentation purposes.

d.) **Annotation format :** Here there are many kinds of annotations available.

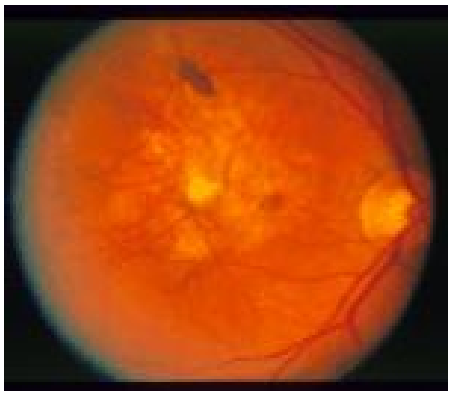
Example :

6	Drusen	5: Large, soft, many (20) 4: Large, soft, few (13) 3: Fine, many (11) 2: Fine, few (19) 1: Absent (307) 0: Unknown (27)
7	Preretinal Hemmorhage	2: Present anywhere (9) 1: Absent (364) 0: Unknown (24)
8	Subretinal Hemmorhage	2: Present anywhere (31) 1: Absent (343) 0: Unknown (23)
9	Microaneurism or Dot Hemmorhage	3: Many anywhere (32) 2: Few anywhere (40) 1: Absent (298) 0: Unknown (27)
10	VH	2: Present anywhere (3) 1: Absent (370) 0: Unknown (24)
11	Small or Medium blot Hemmorhage	3: High density, not regional (57) 2: Low density, not regional (66) 1: Absent (251)

Fig 26 : In STARE, the annotations are denoted as shown in the figure.

A total of 39 annotations are available along with their classification in terms of severity where all images are segmented based on the lesions.

e.) Image with ground truth :



(a)

im0105 13 14 Choroidal Neovascularization AND Age Related Macular Degeneration

(b)

Fig 27 : ‘a’ is a sample image and ‘b’ is the annotation given to ‘a’ in the dataset.

Dataset LINK : https://cecas.clemson.edu/~ahoover/stare/

11.) DRIVE Dataset

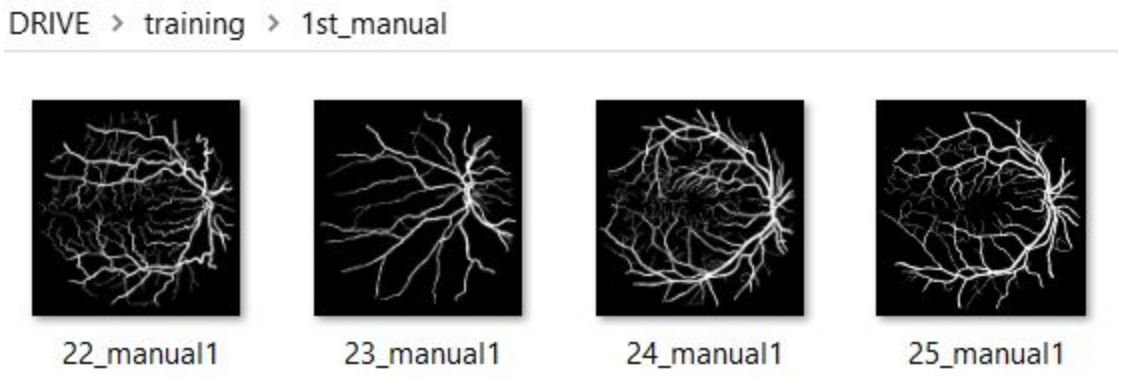
a.) **No. of classes :** There is no distribution of classes in this dataset. This dataset is used for blood vessel extraction

b.) **Total Images =** 40

c.) **Type of task :** Extraction of blood vessels

d.) **Folder Structure:**

(a)

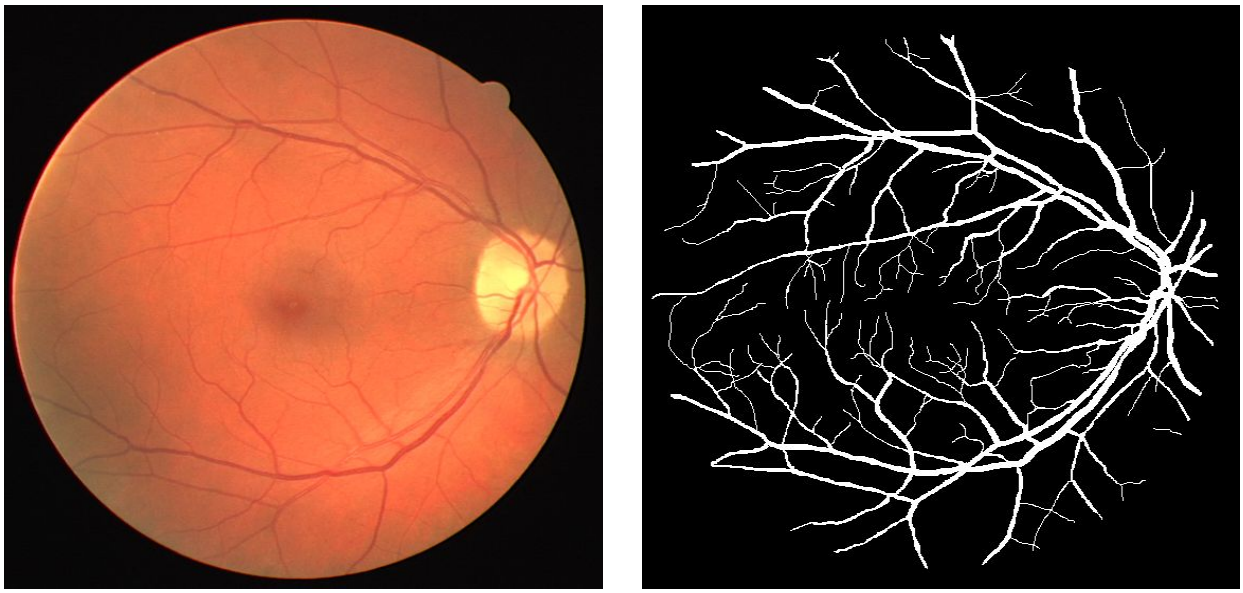


(b)

Fig 28 : ‘a’ and ‘b’ show the folder locations of the image data and their annotations respectively.

e.) **Annotation format :** There are no annotations in this dataset. The purpose of DRIVE is to extract out the blood vessels of the retinal images.

f.) **Image with its vessel flow :**



(a) **(b)**

Fig 29 : ‘a’ is a sample image and ‘b’ shows the vessel extraction of ‘a’

Dataset LINK : https://drive.grand-challenge.org/

12.) EyePACS

a.) **No. of classes :** 5

b.) **Total Images =** more than 35000 (In the Training directory)

c.) **Type of task :** Detection/Grading of Diabetic Retinopathy

d.) **Folder structure :**

The dataset is very huge (>82 GB) and is divided and zipped into folders as shown in the Fig 30. Images from each folder have to be extracted and their annotations/grades are provided in the “labels.csv” file.

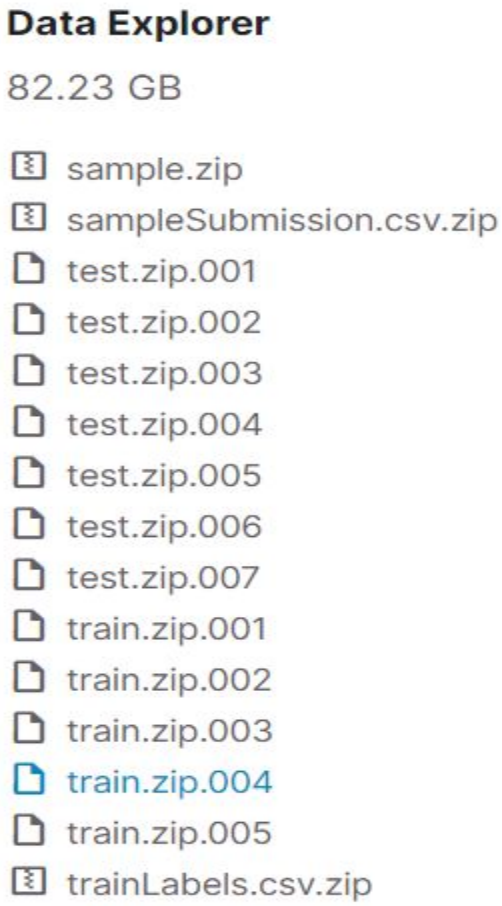


Fig 30 : This image is a sample of the folder structure of the EyePACS database.

e.) Annotation Format :

Here all the images in the dataset are graded according to the severity of Diabetic Retinopathy ranging from 0 to 4 (5 classes). There is a csv file consisting of image name along with its DR grade (level).

Example : [1_left.jpeg 3]

Here image name is given as patient-id_(left/right).jpeg followed by DR grade.

f.) Image sample with Annotation :

(a)

image	level
15_right	2

(b)

Fig 31 : ‘a’ is a fundus retinal image and ‘b’ shows the DR grading for ‘a’.

Dataset LINK : https://www.kaggle.com/c/diabetic-retinopathy-detection/overview

13.) E-Ophtha-EX

a.) No. of classes : The images are not divided into any classes. The dataset is used for Segmentation/Extraction of Exudates.

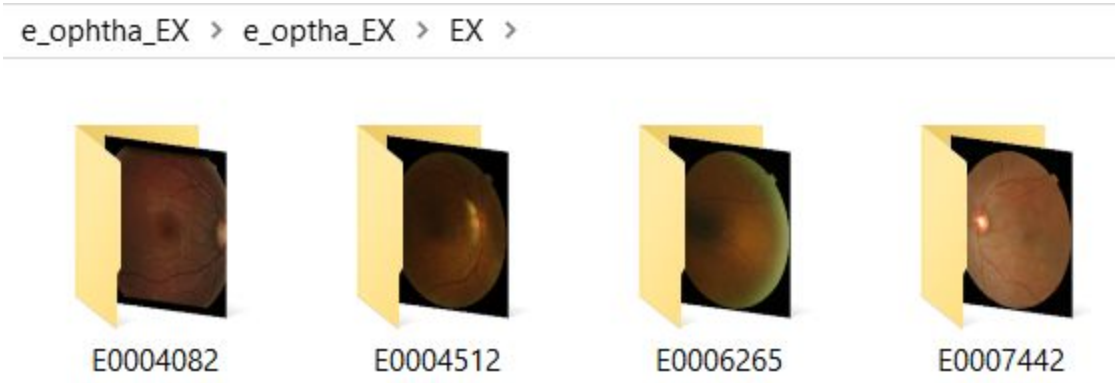
b.) Total Images = 82

i.) Images having Exudates = 47

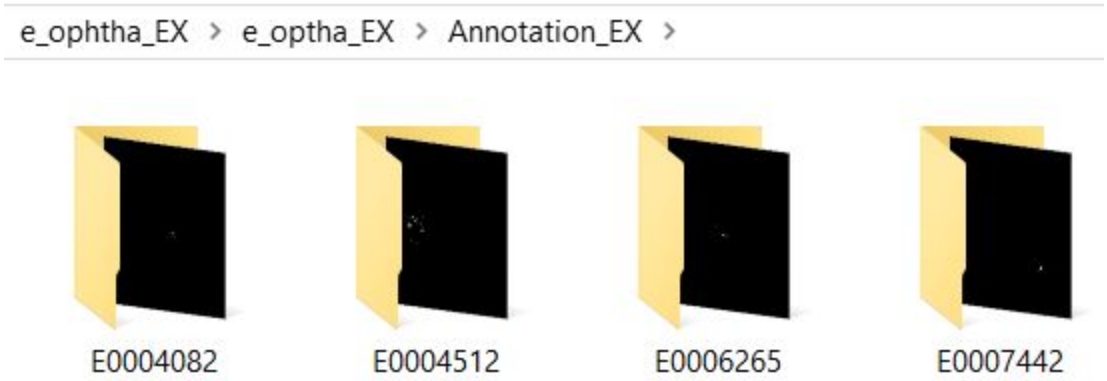
ii.) Normal Image = 35

c.) Type of task : Segmentation of Exudates

d.) Folder Structure :



(a)



(b)

Fig 32 : ‘a’ and ‘b’ both show the folder locations of the image data and their annotations respectively.

There is one more folder named “healthy” that consists of normal images having no lesions.

e.) Annotation format :

This dataset is used for Extraction of Exudates and so the annotations are images that contain the masks for the regions having exudates.

f.) Image sample with its annotation :

(a) **(b)**

Fig 33 : ‘a’ is a sample fundus containing exudates and ‘b’ denotes the mask of the exudates in ‘a’.

Dataset LINK : http://www.adcis.net/en/third-party/e-ophtha/

14.) E-Ophtha-MA

a.) **No. of classes :** The images are not divided into any classes. The dataset is used for Segmentation/Extraction of Microaneurysms.

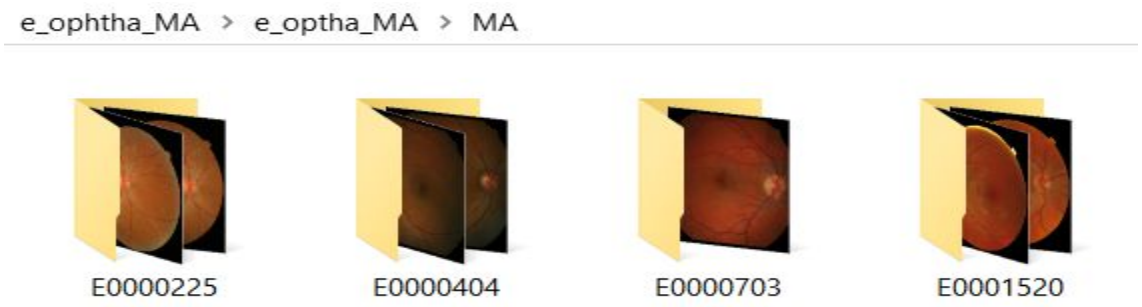
b.) **Total Image =** 381

i.) Images having Microaneurysms = 148

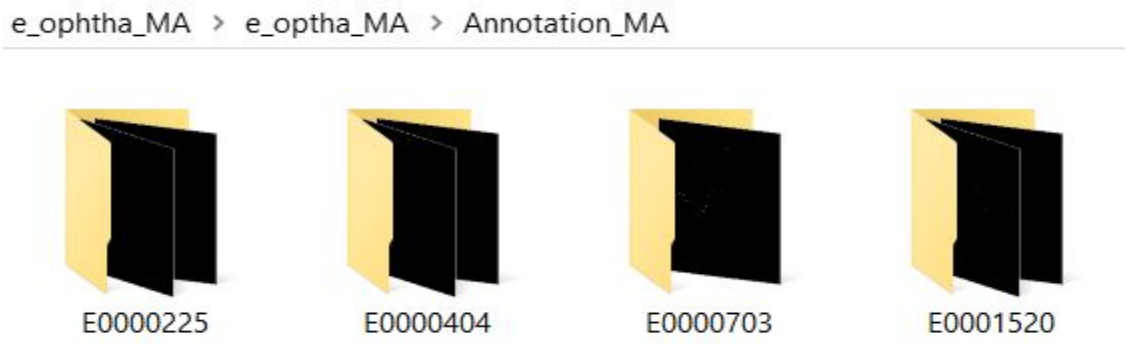
ii.) Normal Image = 233

c.) **Type of task :** Segmentation of Microaneurysms

d.) **Folder Structure :**



(a)



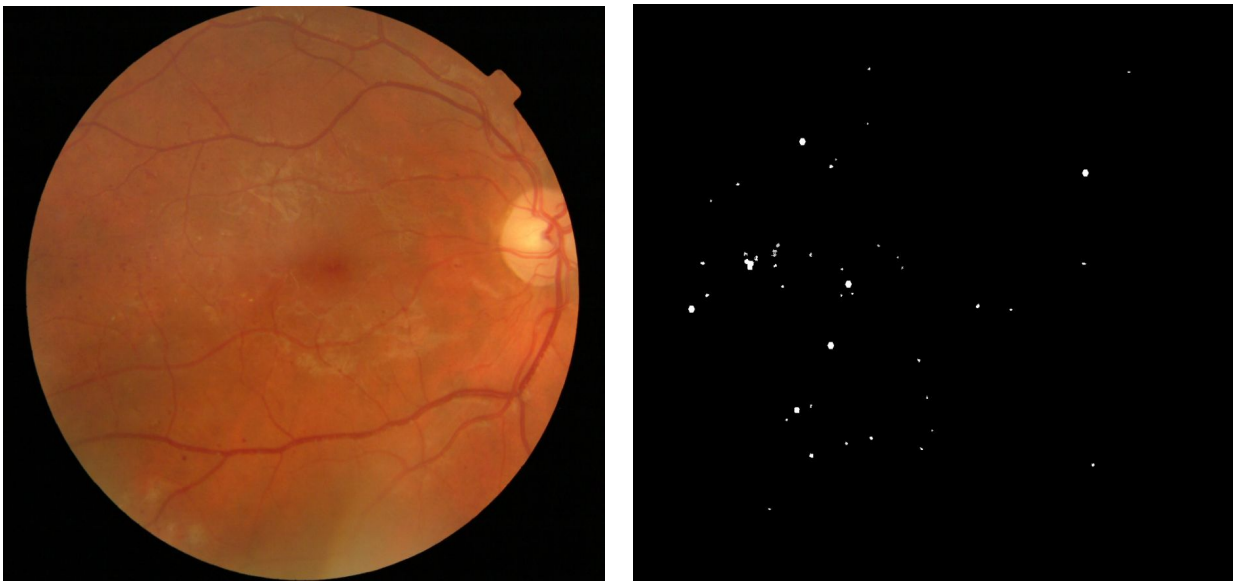
(b)

Fig 34 : ‘a’ and ‘b’ both show the folder locations of the image data along with their annotations.

e.) Annotation format :

This dataset is used for Extraction of Microaneurysms and so annotations are the images that contain the masks for the regions having Microaneurysms.

f.) Image sample with its annotation :



(a) **(b)**

Fig 35 : ‘a’ is a fundus image containing microaneurysms and ‘b’ denotes the mask of microaneurysms in ‘a’.

Dataset LINK : http://www.adcis.net/en/third-party/e-ophtha/