

RJafroc documentation

Dev P. Chakraborty

2019-08-04

Contents

1	Prerequisites	5
2	Introduction	7
3	FROC data format	9
3.1	Introduction	9
3.2	An actual FROC dataset	10
3.3	The FROC Excel data file	13
4	TBA ROI paradigm	17
4.1	References	17

Chapter 1

Prerequisites

TBA

Chapter 2

Introduction

- This is the book describing the **RJafroc** packages.
- The name of the book is RJafrocBook
- Modality and treatment are used interchangeably.
- Reader is a generic radiologist, or a computer aided detection algorithm, or any algorithmic “reader”
- TBA

Chapter 3

FROC data format

3.1 Introduction

- In the free-response ROC (___FROC___) paradigm (Bunch et al., 1978) the observer’s task is to indicate (i.e., **mark** the location of) and **rate** (i.e., assign an ordered label - or confidence level - representing the degree of suspicion) regions in the image that are perceived as suspicious for presence of disease. Accordingly, FROC data consists of **mark-rating pairs**, where each mark indicates a region ¹ that was considered suspicious for presence of a localized lesion and the rating is the corresponding confidence level. The number of mark-rating pairs on any particular case is a-priori unpredictable. It is a non-negative random integer (i.e., 0, 1, 2, ...) that depends on the case, the reader and the modality. The relatively unstructured nature of FROC data makes FROC paradigm data more difficult to analyze than ROC paradigm data ².
- By adopting a proximity criterion, each mark is classified by the investigator as a lesion localization (LL) - if it is close to a real lesion - or a non-lesion localization (NL) otherwise.
- The rating can be an integer or quasi- continuous (e.g., 0 – 100), or a floating point value, as long as higher numbers represent greater confidence in presence of one or more lesions in the ROI ³.

¹In order to avoid confusion with the ROI-paradigm, I do not like to use the term ROI to describe the marks made by the observer.

²Recall that the ROC paradigm always yields a single rating per case.

³The directionality of the rating is not a limitation. If lower values correspond to increased confidence level, it is only necessary to transform the observed rating by subtracting it from a constant value. The constant value can be chosen arbitrarily, typically as the maximum of all observed ratings, thereby ensuring that the transformed value is always non-negative.

- Region-level-normal ratings are stored in the NL field and region-level-abnormal ratings are stored in the LL field.

3.2 An actual FROC dataset

An actual FROC dataset (Zanca et al., 2009) is included as `dataset04`, which has the following `dataset` structure:

```
str(dataset04)
#> List of 8
#> $ NL          : num [1:5, 1:4, 1:200, 1:7] -Inf -Inf 1 -Inf -Inf ...
#> $ LL          : num [1:5, 1:4, 1:100, 1:3] 4 5 4 5 4 3 5 4 4 3 ...
#> $ lesionNum   : int [1:100] 1 1 1 1 1 1 1 1 1 1 ...
#> $ lesionID    : num [1:100, 1:3] 1 1 1 1 1 1 1 1 1 1 ...
#> $ lesionWeight: num [1:100, 1:3] 1 1 1 1 1 1 1 1 1 1 ...
#> $ dataType    : chr "FROC"
#> $ modalityID  : Named chr [1:5] "1" "2" "3" "4" ...
#> ..- attr(*, "names")= chr [1:5] "1" "2" "3" "4" ...
#> $ readerID    : Named chr [1:4] "1" "3" "4" "5"
#> ..- attr(*, "names")= chr [1:4] "1" "3" "4" "5"
```

Examination of the output reveals that:

- The `dataset` structure is a list with 8 members.
- This is a 5-treatment 4-reader dataset (the lengths of the first and second dimensions, respectively, of the NL and LL arrays). The names of the treatments are in the `modalityID` array:

```
attributes(dataset04$modalityID)
#> $names
#> [1] "1" "2" "3" "4" "5"
```

For example, the name of the second treatment is "2".

- The names of the readers are in the `readerID` array:

```
attributes(dataset04$readerID)
#> $names
#> [1] "1" "3" "4" "5"
```

For example, the name of the second reader is "3". Apparently reader "2" “dropped out” of the study.

3.2.1 Numbers of non-diseased and diseased cases

```
length(dataset04$NL[1,1,,1])
#> [1] 200
length(dataset04$LL[1,1,,1])
#> [1] 100
```

- The third dimension of the NL array is the total number of **all** cases, i.e., 200, and the third dimension of the LL array, i.e., 100, is the total number of diseased cases.
- Subtracting the number of diseased cases from the number of all cases yields the number of non-diseased cases.
- Therefore, in this dataset, there are 100 diseased cases and 100 non-diseased cases.

3.2.2 Why dimension the NL array for the total number of cases?

- Because, in addition to LLs, NLs are possible on diseased cases.
- Only LLs are possible on diseased cases.
- Only NLs are possible on non-diseased cases.
- The missing values are filled in with `-Inf`.

3.2.3 Ratings on a non-diseased case

- For treatment 1, reader 1 and case 1 (the first non-diseased case), the NL ratings are:

```
dataset04$NL[1,1,1,]
#> [1] -Inf -Inf -Inf -Inf -Inf -Inf -Inf
```

3.2.4 The meaning of a negative infinity rating

- Obviously, a real rating cannot be negative infinity ⁴. This value is reserved for **missing ratings**, and more generally, **missing marks** ⁵. For

⁴If an observer is so highly confident in the *absence* of a localized lesion, he will simply *not mark* the location in question; if he did, then, logically, he should mark *all* areas in the image that are definitely not lesions; in the FROC paradigm only regions with a reasonable degree of suspicion are marked. The radiologist only wishes to draw attention to regions that are reasonably suspicious; the definition of “reasonable” is determined by clinical considerations.

⁵Since there is a one-to-one correspondence between marks and ratings.

example, since all values in the above code chunk are negative infinities, this means this treatment-reader-case combination did not yield any mark-rating pairs. This possibility, alluded to above, is only possible with FROC data. All other paradigms (ROC, LROC and ROI) yield at least one rating per case.

- The length of the fourth dimension of the NL array is determined by that treatment-reader-case combination yielding the maximum number of NLs. Consider the following chunk:

```
for (i in 1:5)
  for (j in 1:4)
    for (k in 1:200)
      if (all(dataset04$NL[i,j,k,] != -Inf))
        cat(i, j, k, all(dataset04$NL[i,j,k,] != -Inf), "\n")
#> 5 4 192 TRUE
```

- This shows that the fourth dimension of the NL array has to be of length 7 because *one, and only reader*, specifically reader “4”, made 7 NL marks on a diseased case in treatment “5”!

3.2.5 Ratings on a non-diseased case

Unlike non-diseased cases, diseased cases can have both NL and LL ratings.

- For treatment 1, reader 1, case 51 (the 1st diseased case) the NL ratings are:

```
dataset04$NL[1,1,51,]
#> [1] -Inf -Inf -Inf -Inf -Inf -Inf -Inf
dataset04$lesionNum[1]
#> [1] 1
dataset04$LL[1,1,1,]
#> [1] 4 -Inf -Inf
mean(is.finite(dataset04$LL))
#> [1] 0.3043333
```

. There are only two finite values because this case has two ROI-level-abnormal regions, and 2 plus 2 makes for the assumed 4-regions per case. The corresponding \$lesionNum field is 1.

```
mean(is.finite(dataset04$NL[, , 1:50,]))
#> [1] 0.05942857
dataset04$NL[1,1,51,]
```

```
#> [1] -Inf -Inf -Inf -Inf -Inf -Inf -Inf
dataset04$lesionNum[1]
#> [1] 1
dataset04$LL[1,1,1,]
#> [1] 4 -Inf -Inf
mean(is.finite(dataset04$LL))
#> [1] 0.3043333
```

```
mean(is.finite(dataset04$NL[,1:50,]))
#> [1] 0.05942857
dataset04$NL[1,1,51,]
#> [1] -Inf -Inf -Inf -Inf -Inf -Inf -Inf
dataset04$lesionNum[1]
#> [1] 1
dataset04$LL[1,1,1,]
#> [1] 4 -Inf -Inf
mean(is.finite(dataset04$LL))
#> [1] 0.3043333
```

- The ratings of the 2 ROI-level-abnormal ROIs on this case are 4. The mean rating over all ROI-level-abnormal ROIs is 3.6785323.

```
mean(is.finite(dataset04$NL[,1:50,]))
#> [1] 0.05942857
dataset04$NL[1,1,51,]
#> [1] -Inf -Inf -Inf -Inf -Inf -Inf -Inf
dataset04$lesionNum[1]
#> [1] 1
dataset04$LL[1,1,1,]
#> [1] 4 -Inf -Inf
mean(is.finite(dataset04$LL))
#> [1] 0.3043333
```

3.3 The FROC Excel data file

An Excel file in JAFROC format containing simulated ROI data corresponding to `dataset04`, is included with the distribution. The first command (below) finds the location of the file and the second command reads it and saves it to a dataset object `ds`.

```
fileName <- system.file(
  "extdata", "includedFrocData.xlsx", package = "RJafroc", mustWork = TRUE)
```

```
ds <- DfReadDataFile(fileName)
ds$dataType
#> [1] "FROC"
```

The `DfReadDataFile` function automatically recognizes that this is an *ROI* dataset. Its structure is similar to the JAFROC format Excel file, with some important differences, noted below. It contains three worksheets:

The figure displays three screenshots of an Excel spreadsheet titled 'IncludedFrocData.xlsx'. The first two screenshots show the 'Data' worksheet, which contains a table with columns A (CaseID), B (LesionID), C (Weight), D (Truth), and E (Truth). The first 50 rows represent normal cases (CaseID 1-50), and the next 40 rows represent abnormal cases (CaseID 51-90). The third screenshot shows the 'Truth' worksheet, which contains a table with columns A (CaseID), B (LesionID), and C (Truth). The first 50 rows represent normal cases (CaseID 1-50), and the next 40 rows represent abnormal cases (CaseID 51-90).

- The **Truth** worksheet - this indicates which cases are diseased and which are non-diseased and the number of ROI-level-abnormal region on each case.
 - There are 50 normal cases (labeled 1-50) under column **CaseID** and 40 abnormal cases (labeled 51-90).
 - The **LesionID** field for each normal case (e.g., **CaseID** = 1) is zero and there is one row per case. For abnormal cases, this field has

a variable number of entries, ranging from 1 to 4. As an example, there are two rows for **CaseID** = 51 in the Excel file: one with **LesionID** = 2 and one with **LesionID** = 3.

- The **Weights** field is always zero (this field is not used in ROI analysis).

ReaderID	ModalityID	CaseID	FP_Rating
1	1	1	0.949868
2	1	1	-0.05825
3	1	1	-0.776378
4	1	1	0.012073
5	1	1	-0.641182
6	1	1	-0.149436
7	1	1	-0.669429
8	1	1	-1.15841
9	1	1	0.584815
10	1	1	0.799213
11	1	1	-1.45574
12	1	1	0.55311
13	1	1	0.409021
14	1	1	0.409462
15	1	1	-0.034937
16	1	1	1.04011
17	1	1	0.244483
18	1	1	-0.674879

ReaderID	ModalityID	CaseID	TP_Rating
1	1	1	0.949868
2	1	1	-0.05825
3	1	1	-0.776378
4	1	1	0.012073
5	1	1	-0.641182
6	1	1	-0.149436
7	1	1	-0.669429
8	1	1	-1.15841
9	1	1	0.584815
10	1	1	0.799213
11	1	1	-1.45574
12	1	1	0.55311
13	1	1	0.409021
14	1	1	0.409462
15	1	1	-0.034937
16	1	1	1.04011
17	1	1	0.244483
18	1	1	-0.674879

- The FP (or NL) worksheet - this lists the ratings of ROI-level-normal regions.
 - For **ReaderID** = 1, **ModalityID** = 1 and **CaseID** = 1 there are 4 rows, corresponding to the 4 ROI-level-normal regions in this case. The corresponding ratings are . The pattern repeats for other treatments and readers, but the rating are, of course, different.
 - Each **CaseID** is represented in the FP worksheet (a rare exception could occur if a case-level abnormal case has 4 abnormal regions).

ReaderID	ModalityID	CaseID	LesionID	TP_Rating
1	1	1	51	2
2	1	1	51	3
3	1	1	52	1
4	1	1	52	2
5	1	1	52	4
6	1	1	53	1
7	1	1	53	2
8	1	1	89	2
9	1	1	89	4
10	1	1	90	1
11	1	1	90	2
12	2	1	51	2
13	2	1	51	3
14	2	1	52	1
15	2	1	52	2
16	2	1	52	4
17	2	1	53	1
18	2	1	53	2

- The TP (or LL) worksheet - this lists the ratings of ROI-level-abnormal regions.
 - Because normal cases generate TPs, one does not find any entry with **CaseID** = 1-50 in the TP worksheet.
 - The lowest **CaseID** in the TP worksheet is 51, which corresponds to the first abnormal case.
 - There are two entries for this case, corresponding to the two ROI-level-abnormal regions present in this case. Recall that corresponding to this **CaseID** in the **Truth** worksheet there were two entries with **LesionID** = 2 and 3. These must match the **LesionID**'s listed for this case in the TP worksheet. Complementing these two entries, in the **FP** worksheet for **CaseID** = 51, there are 2 entries corresponding to the two ROI-level-normal regions in this case.
 - One should be satisfied that for each abnormal case the sum of the number of entries in the TP and FP worksheets is always 4.

Chapter 4

TBA ROI paradigm

- One can think of the ROI paradigm as similar to the FROC paradigm, but with localization accuracy restricted to belonging to a region (one cannot distinguish multiple lesions within a region). The ROIs are defined prior to the study and made known to all observers participating in the study. Unlike the FROC paradigm, a **rating is required for every ROI**.

4.1 References

Bibliography

- Bunch, P. C., Hamilton, J. F., Sanderson, G. K., and Simmons, A. H. (1978). A free-response approach to the measurement and characterization of radiographic-observer performance. *J of Appl Photogr. Eng.*, 4:166–171.
- Zanca, F., Jacobs, J., Van Ongeval, C., Claus, F., Celis, V., Geniets, C., Provost, V., Pauwels, H., Marchal, G., and Bosmans, H. (2009). Evaluation of clinical image processing algorithms used in digital mammography. *Medical Physics*, 36(3):765–775.