README

label.json stores exported label information in json format, including: basic information, BI-RADS label, cropping, tumor masking, and tissue segmentation. An example is shown below:



The essential entry is ‘case’. Each case has a unique id, tumor type diagnosis (benign or malignant) and one or more ultrasound images (captured from different angles).

“**birads**” field is created by users who have access to the specific dataset. The meaning of encoded number is listed below:

**SHAPE** =

-1: UNKNOWN

0: Oval

1: Round

2: Irregular

**ORIENTATION** =

-1: UNKNOWN

0: Parallel

1: Not Parallel

**MARGIN** =

-1: UNKNOWN

0: Circumscribed

*Not circumscribed*

1: Indistinct

2: Angular

3: Microlobulated

4: Spiculated

**ECHO PATTERN** =

-1: UNKNOWN

0: Anechoic

1: Hyperechoic

2: Complex cystic and solid

3: Hypoechoic

4: Isoechoic

5: Heterogeneous

**POSTERIOR FEATURE** =

-1: UNKNOWN

0: No posterior features

1: Enhancement

2: Shadowing

3: Combined pattern

**CALCIFICATION** =

-1: UNKNOWN

0: Calcifications in a mass

1: Calcifications outside of a mass

2: Intraductal calcifications

3: NONE

**BI-RADS ASSESSMENT** =

-1: UNKNOWN

0: Category 0: Incomplete - Need Additional Imaging

1: Category 1: Negative

2: Category 2: Benign

3: Category 3: Probably Benign

*Category 4: Suspicious*

4A: Category 4A: Low suspicion for malignancy

4B: Category 4B: Moderate suspicion for malignancy

4C: Category 4C: High suspicion for malignancy

5: Category 5: Highly suggestive of malignancy

6: Category 6: Known biopsy-proven malignancy

“**masking**” field includes 4 sub-fields: cropping, tumor, tissue, and creator.

“cropping” is 2 anchor points needed to crop an image. It includes 4 values: [x1, x2, x3, x4]. (x1, x2) is the start cropping point in the left-top and (x3, x4) is the end cropping point in the right-bottom.

“tumor” stores a curve enclosing the tumor region. It consists of a series of points.

“tissue” stores 3 or 4 curves across the image horizontally from left to right.