Evaluation of Bicycle data

Evaluation of annotation dataset

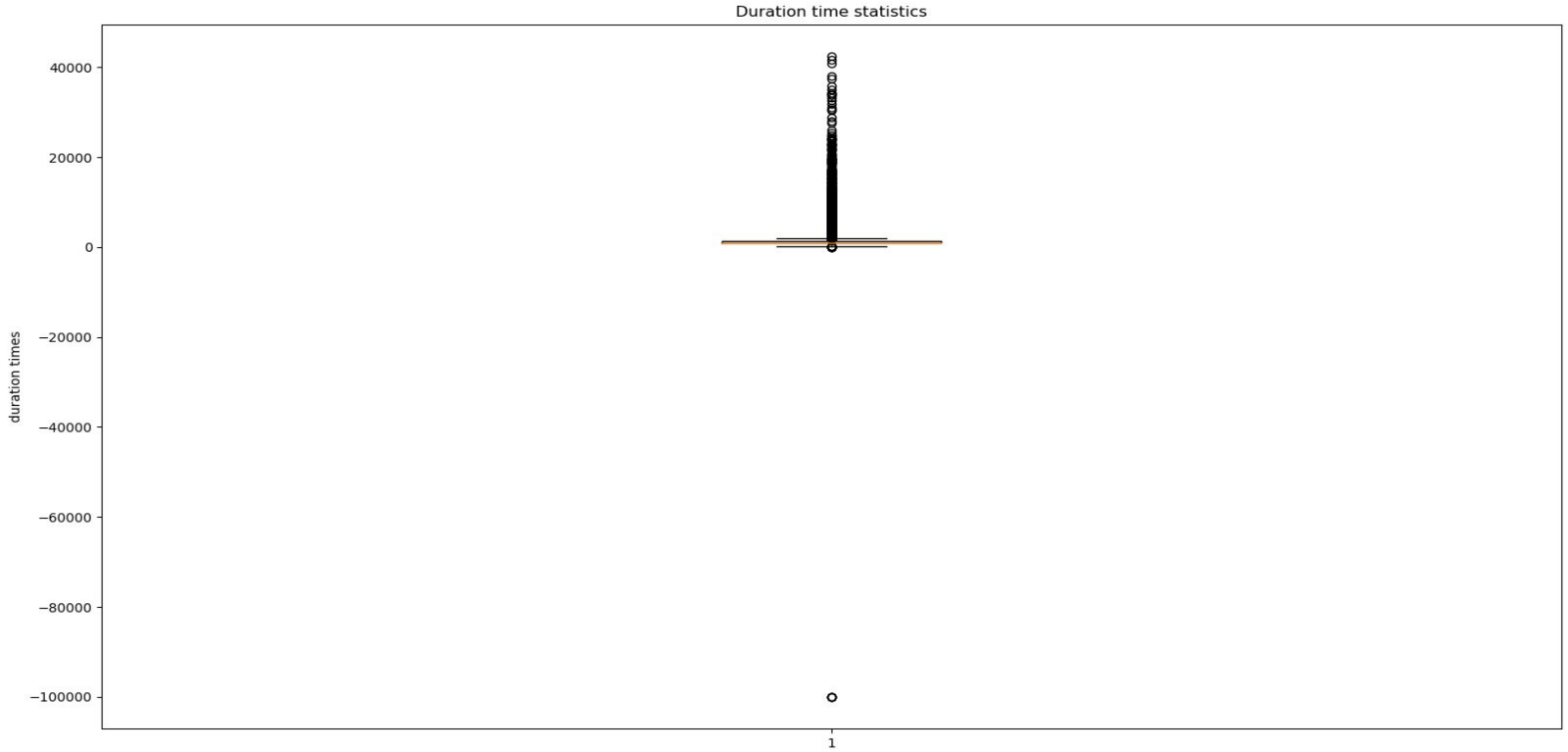
1.1. The total number of anotaters in the given dataset are 22.

1.2 The mean, minimum and maximum duration times are 1284.33, -99999 and 42398 ms respectively.

1.3 There are different work packages sizes that are assigned to different annotators. The range of various work package sizes ranges from 1 to 5.

1.4 There are in total 68 in the dataset. Highly disagree responses comes when there are equal number of positive and negative responses in each task in the dataset.

Box plot for duration time statistics



Corrupt responses and can't solve responses

- The number of cannot solve response are 17 and the number of corrupt data responses are 4.
- In each of questions, each annotator has utilized this option and did not provide any answer to the actual question.
- When asked about weather there is a bicycle in the corresponding image, the annotation was not able to provide the answer and have chosen these options.

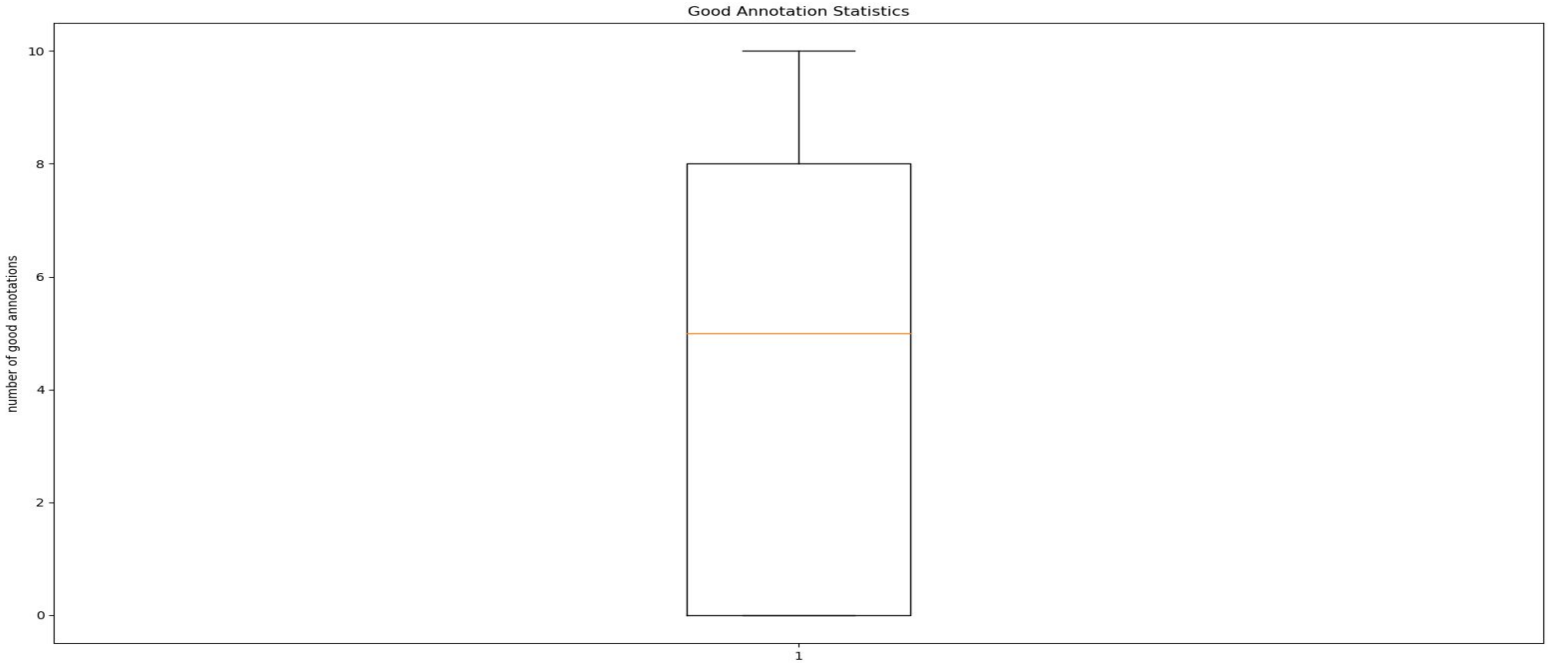
Evaluation of reference data

- There are in total 9087 images. In these images according to reference data, 4586 images contains a bicycle and remaining 4501 images does not contain a bicycle.
- This means around 50.46 percent of the images contains bicycle and 49.53 percent does not contain bicycle in the images.
- **These statistics reveal that, the given reference dataset is a well balanced dataset.**

Good Annotators

- The following statistics describes the overall good annotation statistics:
  - count 9087
  - mean 4.060086
  - std 3.813204
  - min 0.000000
  - 25% 0.000000
  - 50% 5.000000
  - 75% 8.000000
  - max 10.000000

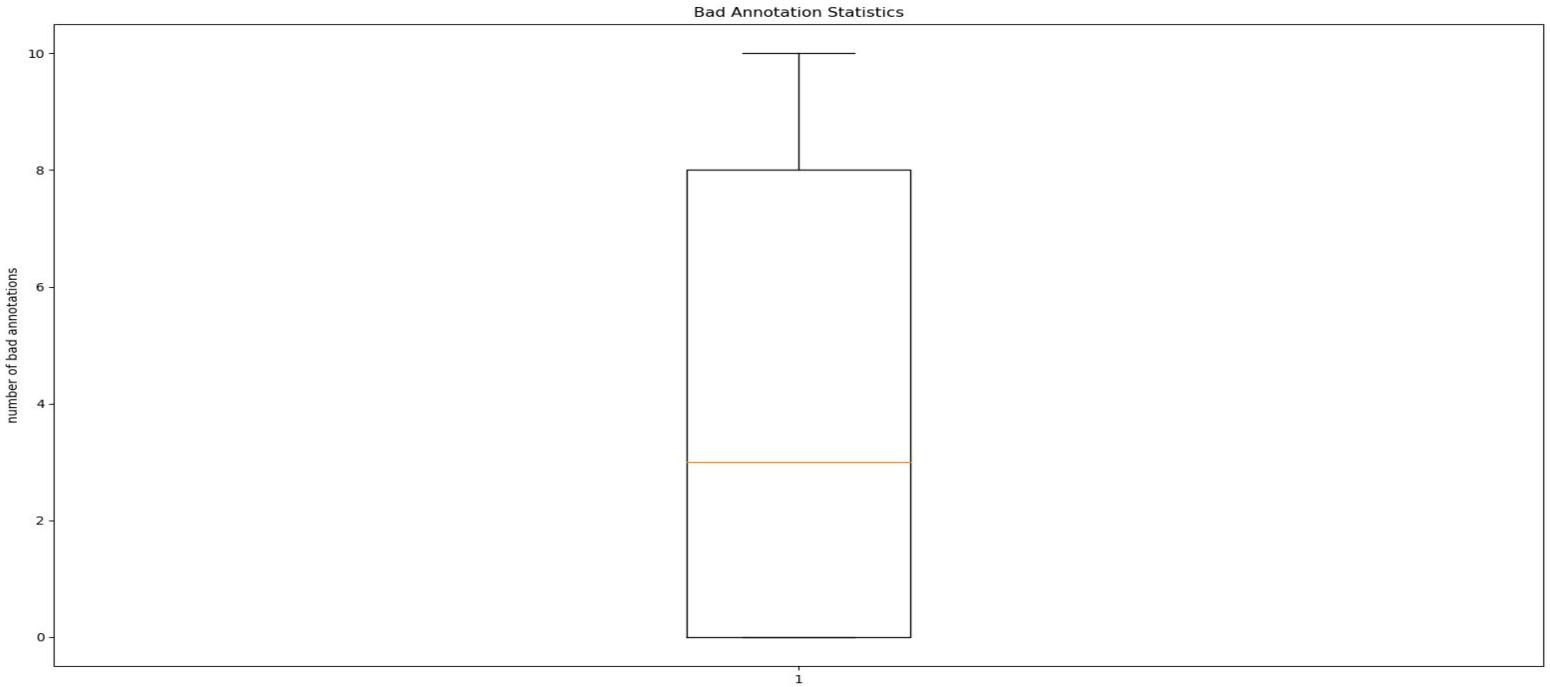
Good annotation box plot



Bad Annotation Statistics

- The following statistics describes the overall good annotation statistics:
  - Count: 9087
  - mean 3.915704
  - std 3.808875
  - min 0.000000
  - 25% 0.000000
  - 50% 3.000000
  - 75% 8.000000
  - max 10.000000
  -

Bad annotation box plot



Interpretation of the good and bad annotations

- The median value in case of good annotation is 5 and for bad annotation is 3.
- This indicates that the distribution of number of bad annotators in the dataset is skewed right and good annotation is symmetric.

Thank You!!