Guidelines for Utilizing Human Behavior Experiment Data

These guidelines are intended for interpreting and utilizing the image index and similarity assessment results data collected through an experiment involving 121 clinical physicians evaluating the similarity of Chest X-ray images.

This data will be officially released following the conference presentation of this paper and will be freely available for use by researchers in related fields for non-commercial purposes.

To use this data, please cite this paper as the reference.

1. Chest X-ray Data

We utilized the CheXpert 1.0 dataset for the human behavior experiment. Only the experimental results data interpreted by the physicians are provided; the CheXpert images are not included and must be downloaded separately.

The img\_list.csv file (located in the <make\_image\_numpy\_file> folder) provides the original image filenames from the CheXpert 1.0 dataset (second column) and their corresponding indices used in our experiment (first column). All experimental data references images by these indices.

1. Interpreting Behavioral Experiment Data

Each of the 121 subjects was assigned an anonymized participant number from 1 to 121.

Subjects were assigned to Group A and Group B as follows:

* Group A: Participants 1 to 62
* Group B: Participants 63 to 121

The behavioral measurement data is provided in CSV files named after the subject number, located in the <Group\_A\_human\_data\_main> and <Group\_B\_human\_data\_main> folders under the <1.subject\_measurement\_data> folder.

Each CSV file consists of 18 columns. The following describes the meaning of each column:

* Column 1: Indicates the display order of the triplet presented to the subject. While the triplet combinations were identical for all subjects, the display order was randomized.
* Column 2: Indicates the unique number of the triplet. The same triplet number presents the same image set to all subjects.
* Column 3: Indicates the participant's group. Group A is assigned triplet numbers 11-510, and Group B is assigned triplet numbers 521-1020.
* Column 4: Index of the image 1 presented in the triplet for the subject.
* Column 5: Index of the image 2 presented in the triplet for the subject.
* Column 6: Index of the image 3 presented in the triplet for the subject.
* Column 7: The time taken by the subject to perceive the triplet and move the balls associated with each image to make the final similarity judgment, measured in seconds.
* Column 8: The time taken to input the similarity judgment and confidence level for each image pair, measured in seconds.
* Column 9: Final pixel position of the ball associated with the image 1 on the subject’s display device.
* Column 10: Final pixel position of the ball associated with the image 2 on the subject’s display device.
* Column 11: Final pixel position of the ball associated with the image 3 on the subject’s display device.
* Column 12: Screen size of the subject's display device in pixels.
* Column 13: Confidence in the similarity judgment between image 1 and image 2 (2: Very uncertain, 4: Uncertain, 6: Neutral, 8: Fairly certain, 10: Very certain).
* Column 14: Confidence in the similarity judgment between image 2 and image 3 (2: Very uncertain, 4: Uncertain, 6: Neutral, 8: Fairly certain, 10: Very certain).
* Column 15: Confidence in the similarity judgment between image 3 and image 1 (2: Very uncertain, 4: Uncertain, 6: Neutral, 8: Fairly certain, 10: Very certain).
* Column 16: The distance between the ball associated with image 1 and the ball associated with image 2 as reported by the subject (smaller numbers indicate higher similarity).
* Column 17: The distance between the ball associated with image 2 and the ball associated with image 3 as reported by the subject (smaller numbers indicate higher similarity).
* Column 18: The distance between the ball associated with image 3 and the ball associated with image 1 as reported by the subject (smaller numbers indicate higher similarity).