

# Facial expression comparison dataset

This dataset consists of face image triplets along with annotations that specify which two faces in the triplet form the most similar pair in terms of facial expression. To the best of our knowledge, there is no existing large scale face dataset with such expression comparison annotations.

## Dataset content:

The dataset is released as CSV files. Each line in a CSV file corresponds to one data sample, which consists of three face images and annotations (by multiple human annotators) that indicate which two faces in the triplet form the most similar pair in terms of expression.

Each face image is specified using an image url and a face bounding box (top-left and bottom-right coordinates). The users need to download the images themselves using the provided urls. The face regions can be cropped from the downloaded images using the provided bounding boxes.

**Note:** Some of the images in this dataset may not be publicly available (if an owner removes their images from public domain) when a user is trying to download them. In such cases, depending on the hosting website and the tool used to download, a dummy image may get downloaded instead of the original image. We recommend users to run a face detection algorithm on the downloaded images and verify that a face is indeed present at the location specified by the provided bounding box.

Each annotation is an integer from the set  $\{1, 2, 3\}$ . A value of 1 means the expressions on second and third faces in the triplet are visually more similar to each other when compared to the expression on the first face. A value of 2 means the expressions on the first and third faces in the triplet are visually more similar to each other when compared to the expression on the second face. A value of 3 means the expressions on the first and second faces in the triplet are visually more similar to each other when compared to the expression on the third face. Each human annotator who participated in the annotation process has a unique id. Most of the samples in this dataset were annotated by six human raters. However, there are few samples which have more than six annotations.

Each triplet in the dataset is also categorized into one of the following three types: one-class triplet, two-class-triplet and three-class triplet. Please see the next section for the definition of these types.

Each line in the CSV files has the following entries:

- URL of image1 (string)
- Top-left column of the face bounding box in image1 normalized by width (float)
- Bottom-right column of the face bounding box in image1 normalized by width (float)
- Top-left row of the face bounding box in image1 normalized by height (float)
- Bottom-right row of the face bounding box in image1 normalized by height (float)
- URL of image2 (string)
- Top-left column of the face bounding box in image2 normalized by width (float)
- Bottom-right column of the face bounding box in image2 normalized by width (float)
- Top-left row of the face bounding box in image2 normalized by height (float)
- Bottom-right row of the face bounding box in image2 normalized by height (float)
- URL of image3 (string)
- Top-left column of the face bounding box in image3 normalized by width (float)
- Bottom-right column of the face bounding box in image3 normalized by width (float)
- Top-left row of the face bounding box in image3 normalized by height (float)
- Bottom-right row of the face bounding box in image3 normalized by height (float)
- Triplet\_type (string) - A string indicating the variation of expressions in the triplet.
- Annotator1\_id (string) - This is just a string of random numbers that can be used to search for all the samples in the dataset annotated by a particular annotator.
- Annotation1 (integer)
- Annotator2\_id (string)
- Annotation2 (integer)
- ....

## How was this dataset collected?

The triplets in this dataset were generated by sampling images from an internal face dataset in which each image has one or more of the following emotion labels: *Amusement, Anger, Awe, Boredom, Concentration, Confusion, Contemplation, Contempt, Contentment, Desire, Disappointment, Disgust, Distress, Doubt, Ecstasy, Elation, Embarrassment, Fear, Interest, Love, Neutral, Pain, Pride, Realization, Relief, Sadness, Shame, Surprise, Sympathy, Triumph*. To reduce the effect of category-bias in the triplet dataset, we sampled the images such that all the above categories are well represented.

Each triplet was shown to human annotators and each annotator was asked to pick two images from the triplet which are visually the most similar. Annotators were instructed to focus only on the expressions and ignore other factors such as identity, gender, ethnicity, pose, age, etc.

This dataset specifically focuses on three types of triplets:

- One-class triplet: All the three images in the triplet share a common label.
- Two-class triplet: Only two images in the triplet share a label and the third one does not share any label with the other two.
- Three-class triplet: None of the images in the triplet share a label.

**Note:** The images in the face dataset (from which we sampled the triplets) were not labeled exhaustively, i.e., an image does not have all the labels that are applicable to it. The current categorization into one-class, two-class and three-class triplets is done based on the existing labels. Hence, a triplet currently classified as a two/three-class triplet may not actually be a two/three-class triplet if the images had been labeled exhaustively.

## Dataset statistics:

This dataset is divided into two partitions:

- TRAIN: faceexp-comparison-data-train-public.csv
  - For training expression models
- TEST: faceexp-comparison-data-test-public.csv
  - For evaluation of the trained models

The number of triplets and the number of faces in each partition are shown in the table below:

Partition	Triplet type				Faces
	One-class	Two-class	Three-class	All	
Training set	152,674	150,234	146,235	449,143	130,516
Test set	17,059	17,107	16,894	51,060	25,427
Full dataset	169,733	167,341	163,129	500,203	155,943

## Citation:

Please cite the below paper if you use this dataset in your work.

R Vemulapalli, A Agarwala, "[A Compact Embedding for Facial Expression Similarity](#)", CoRR, abs/1811.11283, 2018.

## Contact:

Please email Raviteja Vemulapalli ([ravitejavemu@google.com](mailto:ravitejavemu@google.com)) or Aseem Agarwala ([aseemaa@google.com](mailto:aseemaa@google.com)) for any questions regarding the dataset.

## Acknowledgements:

We thank Gautam Prasad, Ting Liu, Brendan Jou, Alan Cowen, Florian Schroff and Hartwig Adam from Google for their support and suggestions during the data collection process.