**Dataset Description**

**Training dataset**: We collect an incremental dataset based on 300W [1][2], named **JD-landmark**, consisting of LFPW [3], AFW [4], HELEN [5] and IBUG [6], and re-annotate them with the mark-up of Fig.1 (106-key-point style). This dataset, containing about 11,400 faces, is applied as the training dataset (see Fig.2). It is accessible to the participants (with landmark annotations).

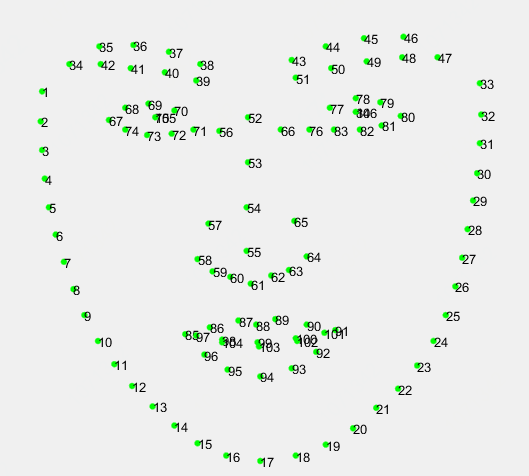


Figure1: The 106-key-point mark-up used for our annotations.

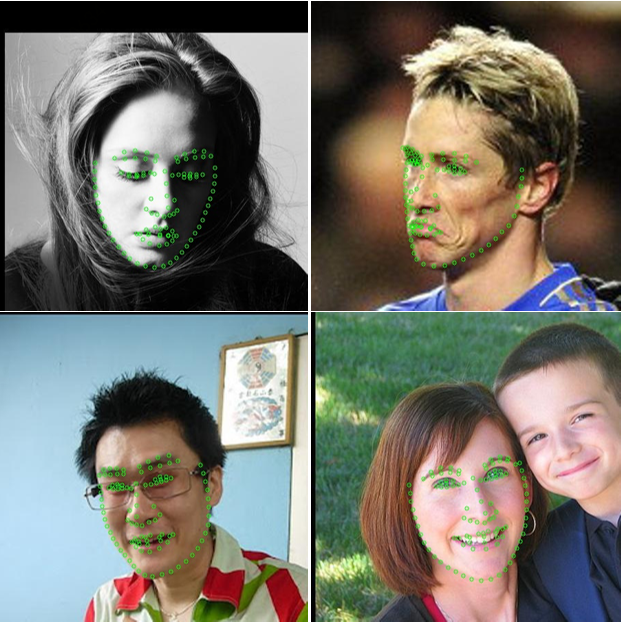


Figure2: Examples of training dataset

**Test dataset 1 (validation):** 2000 webface images (containing large face poses, extreme expressions and occlusions cases) are selected from open source webface database. The participants’ models will be evaluated on this set before the final evaluation. Note that we have appropriately selected or cropped each test image so that it includes only one major face (see Fig. 3).



Figure3: Examples of Test dataset

**Test dataset 2 (final evaluation):** It contains 2000 webface images as well, which is blind to participants throughout the competition. It will be used for the final evaluation.

**Submission Guidelines**

* Participants could send binaries or models with the corresponding runtime environment and the algorithm description to the organizers ([facial\_lmgc\_icme@163.com](mailto:facial_lmgc_icme@163.com)).
* Before April 1st, 2019, the participants can submit the binaries multiple times (but we don’t recommend to submit too much times). The evaluation performance on Test dataset 1 will be returned to the participants via email and we will also report the performance on the website regularly. Besides, the Test dataset 1 will be released on April 1st, so that participants could optimize the algorithm and binaries based on it. During April 1st to April 8th, each team could only submit the binary once, and it will be taken as the final submission and evaluated on Test dataset 2.
* Each binary should accept two parameters as input and be executed like:

“ **./Binary\_filename parameter1 parameter2**”

Here, parameter1 refers to the absolute path for the input file (.jpg) and parameter2 refers to the absolute path for the output file (.txt). At the request of some participants, we will provide the bounding boxes in the test set for choice, which are generated by the same face detector as the training set. If you decide to use the provided bounding box, parameter1 will be the absolute path of the cropped image. Of course, you could use any other face detector you prefer, and then parameter1 refers to the absolute path of the original image and the submitted binary should include both the face detection module and the landmark detection module.

The output file should report the number of key points along with the coordinates for each point (with integer format). The ordering should be the same as the provided annotation files. An example is shown in Fig.4. If multiple faces are detected by the algorithm, please just choose the major face as the output. If no face is detected, the output should be either an empty array or NaN and no txt file should be saved.

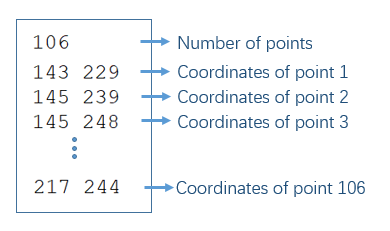


Figure 4: Example of output file

**Evaluation criteria**

Submissions will be evaluated on the area-under-the-curve (AUC) from the cumulative errors distribution (CED) curves. Besides, further statistics from the CED curves such as the failure rate and average normalized mean error (NME) will also be returned to the participants for inclusion in their papers.

The cumulative curve corresponding to the percentage of test images of which the error is less than a threshold ⍺ will be produced. The area-under-the-curve (AUC) is the area under the cumulative distribution curve calculated up to the threshold ⍺, then divided by that threshold. We set the value of ⍺ to be 0.08. Similarly, we consider each image with a point-to-point normalized mean error of α or greater as failure. NME is computed as:

(1)

where “*x*” denotes the ground truth landmarks for a given face, “*y*” denotes the corresponding prediction and “” is computed as . Here, and are the width and height of the enclosing rectangle of the ground truth landmarks. Note that only the successfully detected images will be evaluated for the NME.

**Additional Information**

* The dataset is available for non-commercial research purposes only.
* You agree not to reproduce, duplicate, copy, sell, trade, resell or exploit for any commercial purposes, any portion of the images and any portion of derived data.
* You agree not to further copy, publish or distribute any portion of annotations of the dataset. Except, for internal use at a single site within the same organization it is allowed to make copies of the dataset.
* We reserve the right to terminate your access to the dataset at any time.

**References**

[1] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic.300 Faces in-the-Wild Challenge: The first facial landmark localization Challenge. InInternational Conference on Computer Vision - Workshops (ICCVW), pages 397–403, 2013.

[2] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: Database and results. IVC, 47:3–18, 2016. 3.

[3] Belhumeur, P., Jacobs, D., Kriegman, D., Kumar, N.. ‘Localizing parts of faces using a consensus of exemplars’.  In Computer Vision and Pattern Recognition, CVPR. (2011).

[4] X. Zhu, D. Ramanan.‘Face detection, pose estimation and landmark localization in the wild’, Computer Visionand Pattern Recognition (CVPR) Providence, Rhode Island, June 2012.

[5] Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Boudev, Thomas S. Huang. ‘Interactive Facial Feature Localization’, ECCV2012.

[6] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. A semi-automatic methodology for facial landmark annotation. In CVPR, 2013.