Light-Weight Facial Landmark Prediction Challenge

Chien-Yi Wang
Face Science Team
Microsoft AI R&D Center, Taiwan
chiwa@microsoft.com

Face Science Team at Microsoft Taiwan

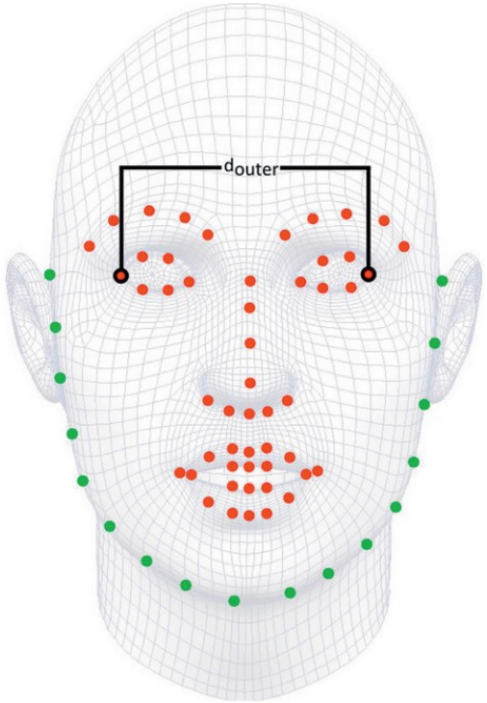
- Topics: Face recognition, Face anti-spoofing, Face synthesis, …
- We are hiring full-time research interns in every semester

Outline

- Introduction
- Dataset
- Challenge Protocol
- Evaluation Metric
- Grading
- Awards

Introduction

- We aim to build a model to predict 68 2D facial landmarks in a single cropped face image with high accuracy, high efficiency, and low computational costs.



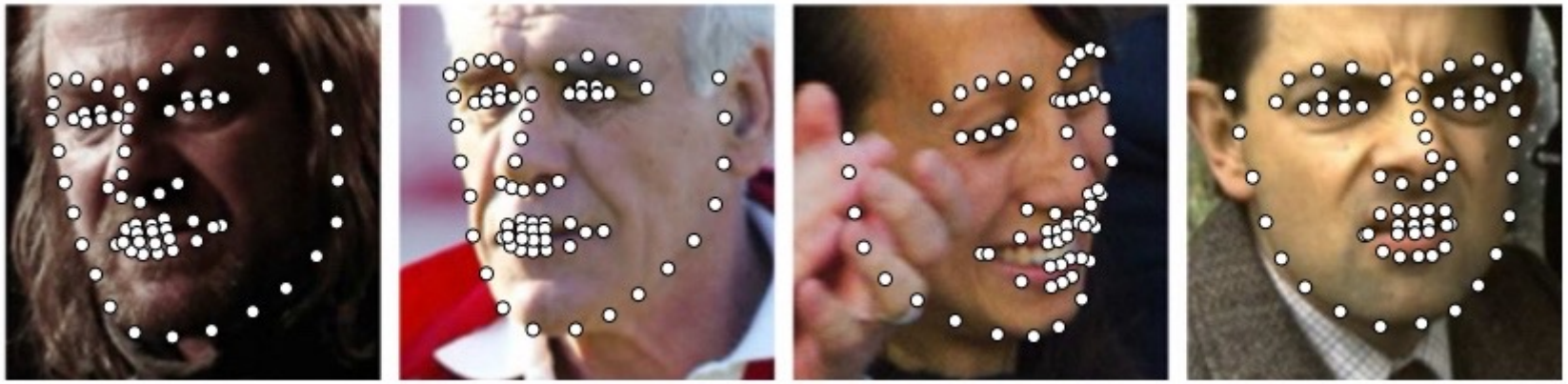


Image Source

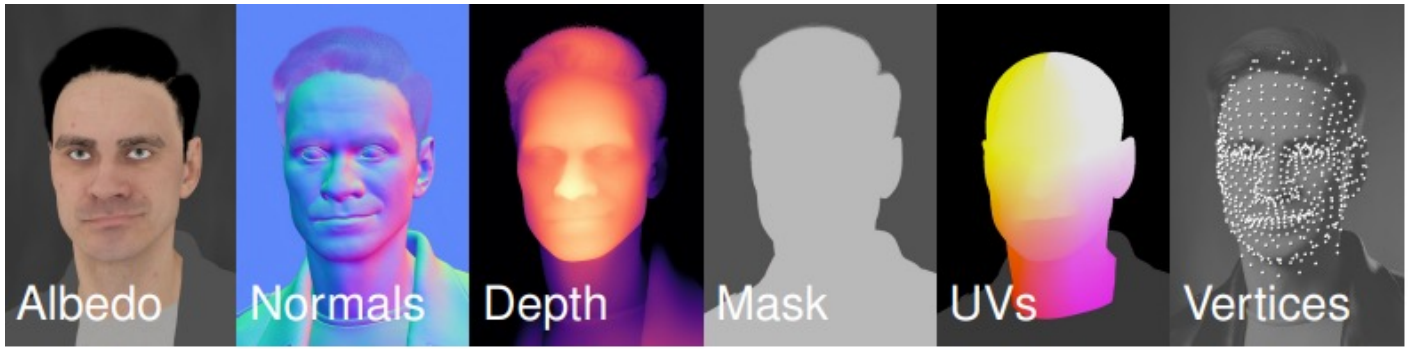
Applications

Facial Motion Retargeting

Talking Head Generation

Training Datasets – Microsoft FaceSynthetics

- 100k diverse synthetic facial images with 68 2D landmark coordinates



Accurate ground truth labels can be generated by our synthetic engine
(only the 68 2D facial landmarks can be used in this challenge)

E. Wood et al., "Fake it till you make it: face analysis in the wild using synthetic data alone", in ICCV 2021

Validation/Testing Datasets – AFLW2000-3D

- The fitted 3D faces of the first 2000 AFLW samples, which can be used for 3D face alignment evaluation
  - The 3D-2D projection is used for the evaluation of 68 2D landmarks in this challenge

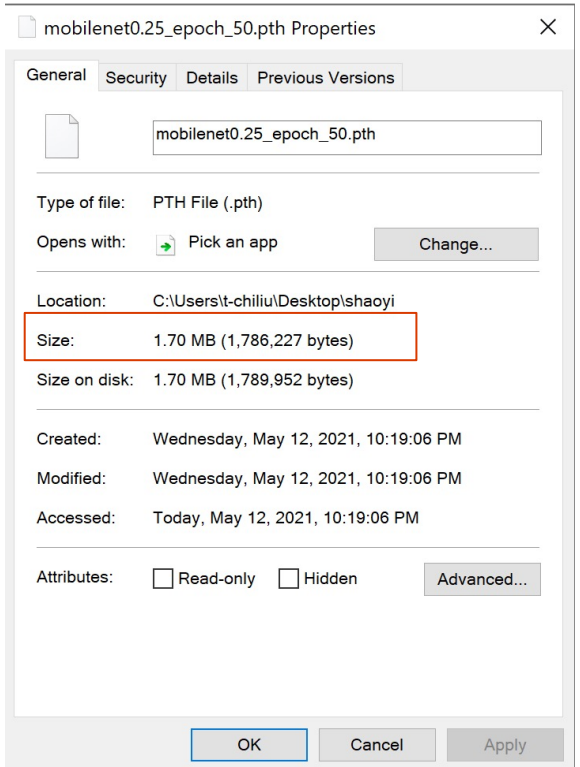
X. Zhu et al., “Face Alignment Across Large Poses: A 3D Solution”, in CVPR 2016

Challenge Protocol

1. We randomly split the AFLW2000 dataset into 10% and 90% as the validation and testing dataset in this challenge
2. You can only train on our training set.
3. The validation set (with labels) can help you validate the results.
4. We will evaluate your method on the testing set (without labels)
5. Face Detector is not allowed during training
6. Please **DO NOT** use external landmark datasets for training your models

Model Constraints

- The upper bound of model size is **15MB**
- The model **should** contain weights (such as neural network)
- We target on **float32** solutions.
  (Float16, int8 or any other quantization methods are **not** allowed.)
- You can **ONLY** train a single model. You **cannot** train an ensembled model.



Evaluation Metric

- Under model constraints, we evaluate the results on the testing dataset with the commonly used **Normalized mean error (NME)**

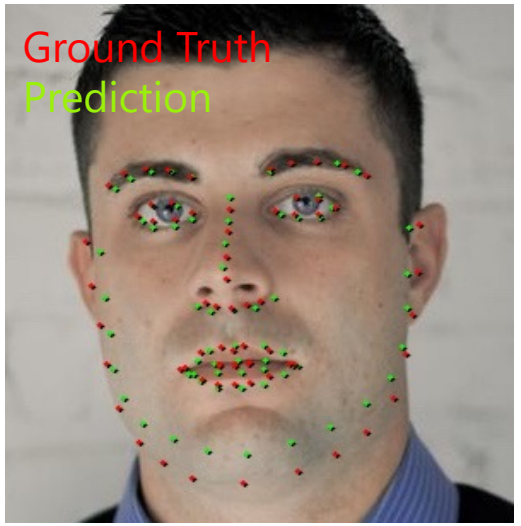
For a matched prediction and the bounding box with width **W** and height **H**, we calculate the NME as follows,

$$NME = \frac{1}{N}\sum_{i=1}^{N}\frac{\sqrt{\Delta x_i^2+\Delta y_i^2}}{d},$$

where N = 68 and d = $\sqrt{W \times H}$

W

H



Dataset format

- Download link : https://drive.google.com/file/d/1hhcsXxGehgf_wf2QJKSuwB7e3xxrTYn9/view?usp=sharing

data/

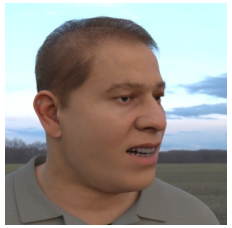
synthetics_train/

annot.pkl

000000.jpg

000001.jpg

……



aflw_val/

annot.pkl

image00013.jpg

……,

Dataset format

- How to load the ground truth pickle file?

```
import pickle
with open('your path', 'rb') as f:
    annot = pickle.load(f)
    X, Y = annot
```

X is a list containing image names.
Y is a list of list, and each element is a list containing 68 ground truth landmark (x,y) coordinates.

Ex. If we print the first element:

```
print(X[0])
print(Y[0])
```

```
image02408.jpg
[(111.59376, 182.81862), (116.13023, 208.07193), (124.482155, 230.69882), (131.35452, 252.12373), (141.70364, 273.31085), (154.61208, 290.53793), (165.997, 301.28796), (182.60072, 310.43527), (206.86557, 311.82864), (233.0994, 302.87238), (254.72496, 288.61755), (271.5397, 273.8399), (284.583, 252.41061), (290.10254, 228.27419), (290.91284, 205.87512), (292.11243, 182.21095), (290.8931, 157.12468), (114.360985, 156.241), (121.548225, 149.99855), (132.0784, 146.11295), (142.52617, 143.1792), (152.90811, 140.92203), (195.95723, 135.28134), (207.4049, 131.23265), (220.66728, 129.4109), (235.43791, 128.7139), (250.00038, 133.73232), (178.41107, 161.37543), (177.87138, 177.33313), (177.42969, 194.58379), (178.1808, 208.63658), (169.74551, 222.30959), (175.17937, 222.24596), (183.7246, 222.6826), (193.31168, 220.80333), (201.22157, 219.13063), (131.96628, 167.29024), (136.33342, 159.93718), (146.66042, 157.46002), (157.50896, 163.65761), (148.9875, 168.21695), (138.85498, 170.50003), (206.21097, 155.15405), (212.75107, 145.31543), (223.15895, 145.14531), (234.53879, 151.05911), (225.13979, 155.05875), (214.31537, 156.92178), (163.45567, 261.2568), (167.67874, 251.91907), (175.46483, 243.55487), (182.99959, 244.69962), (190.2535, 242.16667), (206.98692, 248.48268), (222.47176, 258.56342), (213.36182, 271.5478), (203.61232, 279.2393), (191.35901, 282.38696), (179.65742, 281.66153), (172.34691, 274.74518), (165.565, 261.21948), (175.90991, 254.35864), (184.42104, 252.42915), (194.9466, 252.16464), (220.89227, 259.1861), (201.33879, 271.22598), (190.87549, 273.54205), (181.48886, 272.5461)]
```

How to calculate the NME by yourself

Your prediction

Your ground truth, ex. [[x1,y1],[x2,y2],....]

```
dis = (ldmks - pts68_gt)
dis = np.sqrt(np.sum(np.power(dis, 2), 1))
dis = np.mean(dis)
x = dis / 384
```

$\sqrt{HW}$

NEW!

Schedule

- **Evaluation Server Open**
  2022/06/01 12:00 GMT+8 (中午十二點)
- **Evaluation Server Close**
  2022/06/15 12:00 GMT+8 (中午十二點)
- **Solution File Upload (NTU Cool)**
- 2022/06/15 15:00 GMT+8 (下午三點)
- **Final Leaderboard Release**
- 2022/06/15 20:00 GMT+8 (晚上八點)

- **Oral Presentation**
  2022/06/17 14:20 – 17:20 (TBD)
- **NTU Cool Submission Deadline**
  2022/06/17 23:59 GMT+8 (晚上十二點)
- **NTU Cool Technical Report Deadline (only top 3 teams)**
  2022/06/24 23:59 GMT+8 (晚上十二點)

NEW!

Testing Dataset

- Testing dataset link : https://drive.google.com/file/d/1KNkLYqiZtqef-tt-1PQmirzriwmgHI1-/view?usp=sharing
- There will be 1790 testing images.

NEW!

Evaluation Server

- Our final project challenge is hold on Codalab competition server.
- Link: https://codalab.lisn.upsaclay.fr/competitions/5118?secret_key=19a7d6c1-b907-47fc-a472-1cf6cbf7f853
- Please read all the rules written in the server carefully.
- Maximum submissions: 60
- Maximum submissions per day: 5

NEW!

Evaluation Server -- Submission

- You should first generate the solution file of the 1790 images.
- The filename should be “solution.txt”
- You must report "1+68x2=**137"** values in each row.
  ex. image_name x1 y1 x2 y2 x3 y3 x4 y4 x5 y5 ….
- The numbers are separated with " "(one space)
- The order of the images are not restricted.

```
image03227.jpg 66.4271 129.7086 72.0486 171.3629 56.7890 198.5394 75.9134
image01708.jpg 76.6636 139.5969 97.8672 168.7509 92.8285 209.6318 124.4907
image00809.jpg 116.0135 132.9980 123.1673 175.1914 123.3045 189.5344 103.8
image02637.jpg 43.6828 206.6845 57.5678 240.6826 79.0124 275.0521 94.3887
image03943.jpg 77.0922 160.8842 80.7731 185.0717 87.7598 215.1907 109.9888
image02236.jpg 67.3153 111.8048 67.2469 159.1256 88.7923 192.3095 85.3059
image03852.jpg 81.6116 152.3192 93.9352 170.6814 103.1919 196.3435 122.941
image01572.jpg 254.6684 61.8411 261.1101 110.9259 260.5090 137.3429 259.42
image02622.jpg 48.5405 106.5272 50.9685 145.3769 67.2483 166.7665 76.9288
```

NEW!

Evaluation Server -- Submission

- You need to **compress** it into a zip file. You don't need to put the .txt file into a folder.
  The file **name of the zip is free.** The example is as below.
  ex.
  sample_solution.zip
  |-----solution.txt

- Example File : https://drive.google.com/file/d/1dWwcWSA8vFboXPeDKbOsZANpKQnbkfX3/view?usp=sharing
- The details are on our Codalab Competition page

NEW!

Quantitative Evaluation- Public / Private

- **60%** testing images are public set and **40**% are private set.
- The NME score on the leaderboard are only based on **public set**.
- After closing the evaluation server, we need the team leader to upload the same solution **zip** file on NTU Cool.
- Upload Deadline : **2022/06/15 15:00 GMT+8**

- We will announce the Total score (100% data), Public score (60%), and Private score (40%) on NTU Cool at 2022/06/15 20:00 GMT+8
- **Note** : You should upload the same zip file on the leaderboard, which means the public score we generate should be as same as the score on the leaderboard.

NEW!

Grading (100%)

- Quantitative (50%)
  - Pass baseline score on leaderboard (**10**%) (only Public)
  - Relative Ranking, based on the **total score** (**40**%)

- Presentation (50%) (top-10 teams)
  - Novelty and technical contribution (20%)
  - Completeness of experiments (25%)
    (comparison with different models, ablation studies, visualization, etc)
  - Presentation (5%)

Score	Points	# teams
highest	**40%**	**1**
	38%	**1**
	36%	**1**
	32%	4
	28%	5
	24%	5
Lowest	20%	5

Note:
Only top-10 teams on the final leaderboard (we release on NTU cool) will be chosen for final presentation.
For other teams, the other 50% will based on your report.

NEW!

Report (50%) Format

- Only other teams (not the top-10) should write the report.
  - Novelty and technical contribution (25%)
  - Completeness of experiments (25%)
    (comparison with different models, ablation studies, visualization, etc)

NEW!

Challenge Awards

- Top **3** teams of this challenge will receive cash prizes sponsored by Microsoft AI R&D center
  - 1st: NTD $5k
  - 2nd: NTD $3k
  - 3rd: NTD $2k

- The final ranking is decided by judges based on the **total points** (Quantitative and Presentation) and the content of your work.

NEW!

Code Submission / Report

1. Only the team leader will need to upload your code to **NTU Cool**
2. All your code should be uploaded (including training & testing)
3. **DO NOT** upload the dataset !!
4. You should also upload the **model**, which can generate the score on the leaderboard.
5. You need to upload your presentation or report (either ppt or pdf)
6. **Deadline**: **2022/06/17 23:59 GMT+8** (晚上十二點)

7. (Important!!) For the top 3 teams selected for cash awards, an additional technical report needs to be submitted as well by 6/24 23:59 (晚上十二點). We will open another NTU Cool submission place.
   - Approach (e.g., data preprocessing, model architecture, implementation details, hyperparameter choices, etc.)
   - Experiments
   - Maximum **4** pages (exclude reference)
   - A latex template is provided. https://drive.google.com/drive/folders/12Lb7xAcKtBNGWOnlRo0Hl92fqFHcJJ_l?usp=sharing

NEW!

Code Submission Format

- R07654321/
  - README file (**Important!!!**)
  - Best model file
  - Presentation file / Report file
  - All codes …

- Compress all above files in a zip file named StudentID.zip
  - e.g. R07654321.zip
  - After TAs run “unzip R07654321.zip”, it should generate one directory named “R07654321”

- In **README** file, you need to clearly describe your **environments** and the **steps** to run your code (including training and testing), so that TAs will be able to reproduce your results on leaderboard.
- If TAs cannot reproduce your results, you will receive 0 points in the code part. (unless minor errors)

NEW!

Contact

- 劉致廷
- jackieliu@media.ee.ntu.edu.tw

References

- Synthetic Facial Landmark Datasets:
  - [ICCV 2021] Fake It Till You Make It
- Methods
  - [ECCV 2020] Towards Fast, Accurate and Stable 3D Dense Face Alignment
  - [CVPRW 2019] Accurate 3D Face Reconstruction with Weakly-Supervised Learning
  - [ICCV 2017] How far are we from solving the 2D & 3D Face Alignment problem?