The Third Micro-Expression Grand Challenge (MEGC2020) Guidelines

The goal of this challenge is to spot macro- and micro-expressions interval in long video sequences. For this challenge, we focus on 98 long videos of CAS(ME)^2 database (300 macro-expressions and 57 micro-expressions) and 147 long videos of SAMM Long Videos dataset (343 macro-expressions and 159 micro-expressions). The details of the databases as follow:

1 CAS(ME)^2 [2] spotting task

In the part A of CAS(ME)^2 database, there are 22 subjects and 98 long videos. The average duration is 148s. The facial movements are classified as macro- and micro-expressions. The video samples may contain multiple macro or micro facial expressions. The onset, apex, offset index for these expressions are given in the excel file. In addition, the eye blinks are labeled with onset and offset time. To download the dataset, please visit:

http://fu.psych.ac.cn/CASME/cas(me)2-en.php

Download and fill in the license agreement form, email to fuxl@psych.ac.cn. (Please use all the video samples in 'rawpic' folder.)

2 SAMM Long Videos [1] spotting task

The original SAMM dataset [1] with 159 micro-expressions. In SAMM Long Videos dataset [3], there are 147 videos. The index of onset, apex and offset frames of micro-movements are outlined in the ground truth excel file. The micro-movements interval is from onset frame to offset frame. In this database, all the micro-movements are labeled. Thus, the spotted frames can indicate not only micro-expression but also other facial movements, such as eye blinks. To download the dataset, please visit:

http://www2.docm.mmu.ac.uk/STAFF/M.Yap/dataset.php

Download and fill in the license agreement form, email to M.Yap@mmu.ac.uk with email subject: SAMM long videos.

References

[1] Adrian K Davison, Cliff Lansley, Nicholas Costen, Kevin Tan, and Moi Hoon Yap. Samm: A spontaneous micro-facial movement dataset. *IEEE Transactions on Affective Computing*, 9(1):116–129, 2018.

[2] Fangbing Qu, Su-Jing Wang, Wen-Jing Yan, He Li, Shuhang Wu, and Xiaolan Fu. Cas (me)ˆ 2: a database for spontaneous macro-expression and micro-expression spotting and recognition. *IEEE Transactions on Affective Computing*, 2017.

[3] Chuin Hong Yap, Connah Kendrick, and Moi Hoon Yap. Samm long videos: A spontaneous facial micro- and macro-expressions dataset. *arXiv preprint arXiv:1911.01519*, 2019.

Result Evaluation Standard

1. True positive in one video definition

The true positive (TP) per interval in one video is first defined based on the intersection between the spotted interval and the ground-truth interval. The spotted interval $W_{spotted}$ is considered as TP if it fits the following condition:

$$\frac{W_{spotted} \cap W_{groundTruth}}{W_{spotted} \cup W_{groundTruth}} \geq k \tag{1}$$

where k is set to 0.5, $W_{groundTruth}$ represents the ground truth of the macro- or micro-expression interval (onset-offset). If the condition is not fulfilled, the spotted interval is regarded as false positive (FP).

2. Result evaluation in one video

Supposing there are m ground truth interval in the video, and n intervals are spotted. According to the overlap evaluation, the TP amount in one video is counted as a ($a \leq m$ and $a \leq n$), therefore FP $= n - a$, FN $= m - a$. The spotting performance in one video can be evaluated by following metrics:

$$Recall = \frac{a}{m},\ Precision = \frac{a}{n} \tag{2}$$

$$F - score = \frac{2TP}{2TP + FP + FN} = \frac{2a}{m + n} \tag{3}$$

Yet, the videos in real life have some complicated situations which influences the evaluation per single video:

- There might be no macro- nor micro-expression in the test video. In this case, $m = 0$, the denominator of recall would be zeros.

- If there is no spotted intervals in the video, the denominator of precision would be zeros since $n = 0$.
- It is impossible to compare two spotting methods when both TP amounts are zero. The metric (recall, precision or F1-score) values both equal to zeros. However, the Method_1 outperforms Method_2, if Method_1 spots less intervals than Method_2.

Thus, to avoid these situations, we propose for single video spotting result evaluation, we just note the amount of TP, FP and FN. Other metrics are not considered for one video.

3. Evaluation for entire database

Supposing in the entire dataset,

- There are V videos including M_1 macro-expressions (MaEs) sequences and M_2 micro-expression (MEs) sequences, where $M_1 = \sum_{i=1}^{V} m_{1i}$ and $M_2 = \sum_{i=1}^{V} m_{2i}$;
- The method spot N_1 MaE intervals and N_2 ME intervals in total, where $N_1 = \sum_{i=1}^{V} n_{1i}$ and $N_2 = \sum_{i=1}^{V} n_{2i}$;
- There are A_1 TPs for MaE and A_2 TPs for ME in total, where $A_1 = \sum_{i=1}^{V} a_{1i}$ and $A_2 = \sum_{i=1}^{V} a_{2i}$.

The dataset could be considered as one long video. The results are firstly evaluated for the MaE spotting and ME spotting separately. Then the overall result for macro- and micro spotting is evaluated. The *recall* and *precision* for entire dataset can be calculated by following formulas:

- for macro-expression:

$$Recall_{MaE\_D} = \frac{A_1}{M_1} \; Precision_{MaE\_D} = \frac{A_1}{N_1} \tag{4}$$

- for micro-expression:

$$Recall_{ME\_D} = \frac{A_2}{M_2} \; Precision_{ME\_D} = \frac{A_2}{N_2} \tag{5}$$

- for overall evaluation:

$$Recall_D = \frac{A_1 + A_2}{M_1 + M_2} \; Precision_D = \frac{A_1 + A_2}{N_1 + N_2} \tag{6}$$

Then, the values of *F1-score* for all these three evaluations are obtained based on:

$$F1 - score = \frac{2 \times (Recall \times Precision)}{Recall + Precision} \tag{7}$$

The champion of the challenge will be the best score for overall results in spotting micro- and macro-expressions.

Baseline Method and Result

The Baseline method and the result will be available in the middle of December.

Submission

For the purpose of result verification and to encourage reproducibility and transparency, all entries must submit the following:

- An evaluation log file (.txt, or .csv)indicating the databases, the video id, the ground truth interval range, and the predicted interval range. This is to ensure that all submissions are fairly and correctly evaluated for comparisons.
- A paper highlighting the contribution of the submission, but not limited to, the method, experimental results and analysis, prepared according to the format stipulated by IEEE FG 2020. For detailed instructions on this, please refer to here. All challenge entries should be accompanied by a paper submission.
- GitHub repository URL containing codes of your implemented method, and all other relevant files such as feature/parameter data. To help publicize our workshop and domain area, please do mention (or add relevant links on) MEGC Workshop 2020 and FG 2020. You may provide this URL in a simple text file while submitting.

For all files except for the paper, please submit in a single zip file and upload to the submission system as supplementary material.

Sample log

Header consists of the database labels ('1' for SAMM Long Videos, '2' for CAS(ME)^2), follow by:
Video_ID GT_onset GT_offset Predicted_onset Predicted_offset Result.
In one video, results are sorted by Predicted_onset.
The submission portal will be open at MEGC2020 Website
Deadline of Challenge: **24 January 2020, 2359 PST (UTC -8)**

Rules

The organizers reserve the right to disqualify submissions with on the basis of

- Challenge results that are likely to be suspicious, i.e. out-of-norm from the distribution of scores from submitters.
- Non-submission of accompanying paper.

1					
006_1	-	-	1000	1050	FP
006_1	2324	2403	2310	2395	TP
006_1	3912	3988	-	-	FN
006_1	-	-	4500	4575	FP
006_1	5343	5424	5349	5360	FP
...					
006_2	-	-	100	150	FP
006_2	180	274	190	250	FP
...					
2					
15_0101	-	-	100	150	FP
15_0102	699	707	700	710	TP
15_0102	-	-	780	789	FP
...					

- Submission of an accompanying paper that has a substantial overlap with any other paper already submitted or published, or to be submitted during the review period

For further enquiries, please contact:

Jingting Li (jingting.li@supelec.fr) - General enquiries about the challenge
Su-Jing Wang (wangsujing@psych.ac.cn) - CAS(ME)^2
Moi Hoon Yap (m.yap@mmu.ac.uk) - SAMM dataset