The Second Micro-Expression Grand Challenge (MEGC2019) Spotting Challenge Guidelines

December 18, 2018

The goal of this challenge is to spot micro-movements interval in long video sequences. For this challenge, we focus on 57 micro-expressions from 87 long videos of CAS(ME)^2 database and 159 micro-movements from 79 long videos of SAMM database. 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 87 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.

2 Micro-facial movement spotting in long videos of SAMM database [1]

In SAMM database, there are 32 subjects and each has 7 videos. The average length of videos is 35.3s. For this challenge, we focus on 79 videos, each contains one/multiple micro-movements, with a total of 159 micro-movements. 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.

Result Evaluation Standard

1. True positive in one video definition

Supposing there are m micro-expressions in the video, and n intervals are spotted. The result of this spotted interval $W_{spotted}$ is considered as true positive (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 micro-expression interval (inset-offset).

If not, the spotted interval is regarded as false positive (FP).

2. Result evaluation in one video

Supposing the number of TP in one video is a ($a \leq m$ and $a \leq n$), then FP $= n - a$, FN $= m - a$, the values of following metrics could be calculated:

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

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

Yet, in the real life, there would be following situations for single video:

- The test video do not have micro-expression sequences, thus, $m = 0$, the denominator of recall would be zeros.
- The spotting method does not spot any intervals. The denominator of precision would be zeros since $n = 0$.
- If there are two spotting methods, Method_1 spots p intervals and Method_2 spots q intervals, and $p \leq q$. Supposing for both methods, the number of true positive is 0, thus the metric (recall, precision or F1-score) values both equal to zeros. However, in fact, the Method_1 performs better than Method_2.

Considering these real situations, we propose for single video, we just note the result of TP, FP and FN without calculation of other metrics.

3. Evaluation for entire database

Supposing in the entire database, there are V videos and M micro-expression sequences, and the method spot N intervals in total. The database could be considered as one long video, thus, the metrics for entire database can be calculated by following formulas:

$$Recall_D = \frac{\sum_{i=1}^{V} a_i}{\sum_{i=1}^{V} m_i} = \frac{A}{M} \tag{4}$$

$$Precision_D = \frac{\sum_{i=1}^{V} a_i}{\sum_{i=1}^{V} n_i} = \frac{A}{N} \tag{5}$$

$$F1 - score_D = \frac{2 \times (Recall_D \times Precision_D)}{Recall_D + Precision_D} \tag{6}$$

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 2019. 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 2019 and FG 2019. 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, '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.

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
...					

The submission portal will be open at MEGC2019 Website
Deadline of Challenge: **27 January 2019, 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.
- 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)$\hat{2}$
Moi Hoon Yap (m.yap@mmu.ac.uk) - SAMM dataset