Tutorial

ACCV 2018

2 – 6 December 2018 Perth Western Australia

Facial Micro-Expression Analysis – A Computer Vision Challenge

V. Challenges & Future Avenues

JOHN SEE Multimedia University, Malaysia

ANH CAT LE NGO TrustingSocial

SZE-TENG LIONG Feng Chia University, Taiwan

1st Micro-Expression Grand Challenge 2018 @ FG 2018, Xi’an, China

So, we decided that we should meet up and have a “real-world” look at each other’s expressions…

1st Micro-Expression Grand Challenge 2018 @ FG 2018, Xi'an, China

- 5 "Objective classes" (grouped by Facial AU) instead of emotion classes
- Cross-database protocols
  - **Holdout Database Evaluation (HDE)**
    - Train on one dataset, Test on the other. Swap, repeat. (WAR, UAR)
  - **Composite Database Evaluation (CDE)**
    - Combine both datasets, evaluate by LOSO (F1-score)

Class	CASME II	SAMM	Composite
I	25	24	49
II	15	13	28
III	99	20	119
IV	26	8	34
V	20	3	23
Total	185	68	253

Facial Micro-Expressions Grand Challenge 2018 Summary

Moi Hoon Yap[1], John See[2], Xiaopeng Hong[3], Su-Jing Wang[4]
[1] School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, UK
[2] Faculty of Computing and Informatics, Multimedia University, Malaysia
[3] Center for Machine Vision and Signal Analysis, University of Oulu, Finland
[4] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing, 100101, China

***Abstract*— This paper summarises the Facial Micro-Expression Grand Challenge (MEGC 2018) held in conjunction with the 13th IEEE Conference on Automatic Face and Gesture Recognition (FG) 2018. In this workshop, we aim to stimulate new ideas and techniques for facial micro-expression analysis by proposing a new cross-database challenge. Two state-of-the-art datasets, CASME II and SAMM, are used to validate the performance of existing and new algorithms. Also, the challenge advocates the recognition of micro-expressions based on AU-centric objective classes rather than emotional classes. We present a summary and analysis of the baseline results using LBP-TOP, HOOF and 3DHOG, together with results from the challenge submissions.**

inconsistencies adds further justification for the introduction of new classes based on AUs only [3].

This challenge aims to stimulate the micro-expressions researchers in developing new techniques for the AU-centric objective classes. A summary of the objective classes are as illustrated in Table I. A single composite database for this experiment has a total of 253 micro-expressions.

TABLE I

THE TOTAL NUMBER OF MOVEMENTS ASSIGNED TO THE NEW OBJECTIVE CLASSES FOR CASME II AND SAMM.

1st Micro-Expression Grand Challenge 2018 @ FG 2018, Xi'an, China

THE RESULTS OF HOLDOUT-DATABASE EVALUATION (TASK A).

Method	WAR			UAR		
	@SAMM	@CASME II	Average	@SAMM	@CASME II	Average
LBP-TOP	0.338	0.232	0.285	0.327	0.316	0.322
3DHOG	0.353	0.373	0.363	0.269	0.187	0.228
HOOF	0.441	0.265	0.353	0.349	**0.346**	0.348
Peng et al.	**0.544**	**0.578**	**0.561**	**0.440**	0.337	**0.389**
Khor et al.	0.485	0.384	0.435	0.382	0.322	0.352

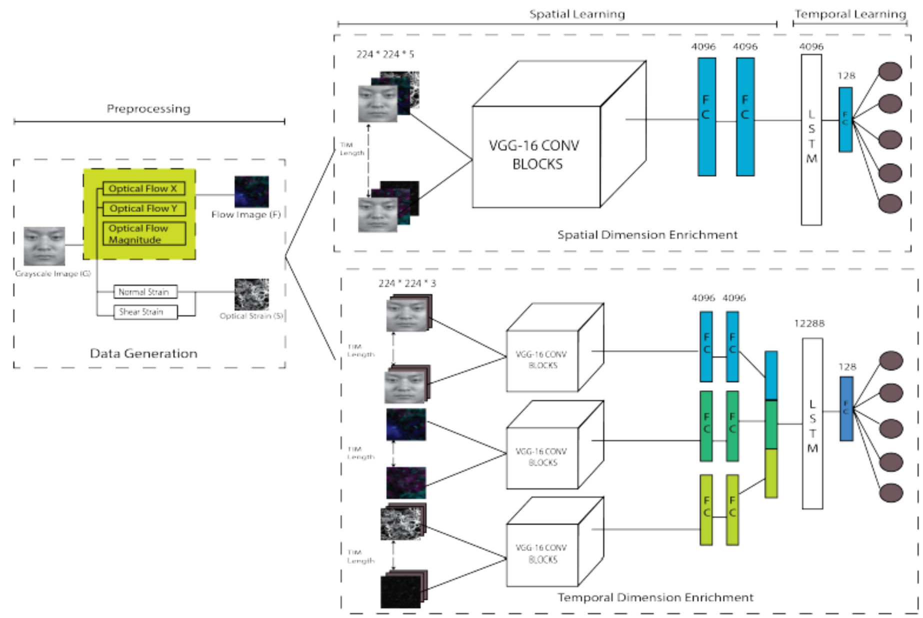
THE RESULTS OF COMPOSITE DATABASE EVALUATION (TASK B) BASED ON LOSO CROSS VALIDATION.

Method	F1-Score	Weighted F1-score
LBP-TOP	0.400	0.524
3DHOG	0.271	0.436
HOOF	0.404	0.527
Peng at al.	**0.639**	**0.733**
Merghani et al.	0.454	0.579
Khor et al.	0.393	0.523

- UAR results are very close
- 6 papers accepted (50%) – 3 challenge, 3 non-challenge

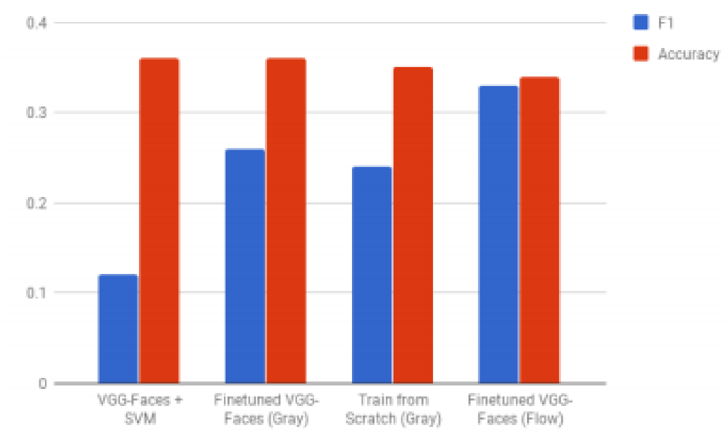
Enriched Long-term Recurrent Convolutional Network (ELRCN)

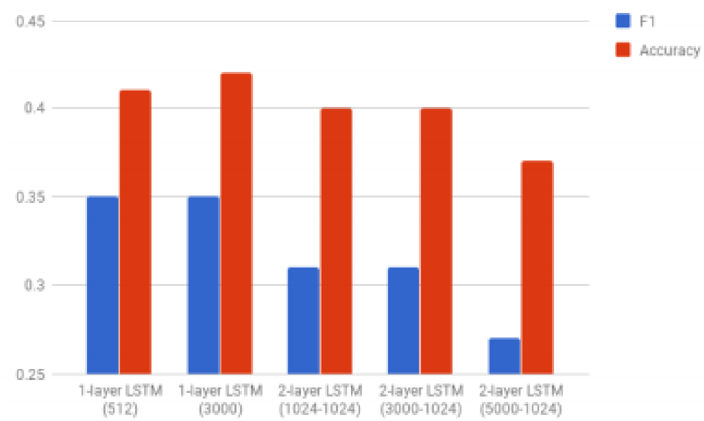
- 2 ways of enriching a CNN-LSTM pairing
  - **Spatially**: Gray, OF, OS images stacked along CNN channel, feats. passed to LSTM
  - **Temporally**: Separate CNN streams for Gray, OF and OS images , late fusion after FC, feats passed to LSTM



Khor, H. Q. et al. (2018). **Enriched Long-term Recurrent Convolutional Network for Facial Micro-Expression Recognition**. In *FG 2018*, pp. 667-674

Enriched Long-term Recurrent Convolutional Network (ELRCN)





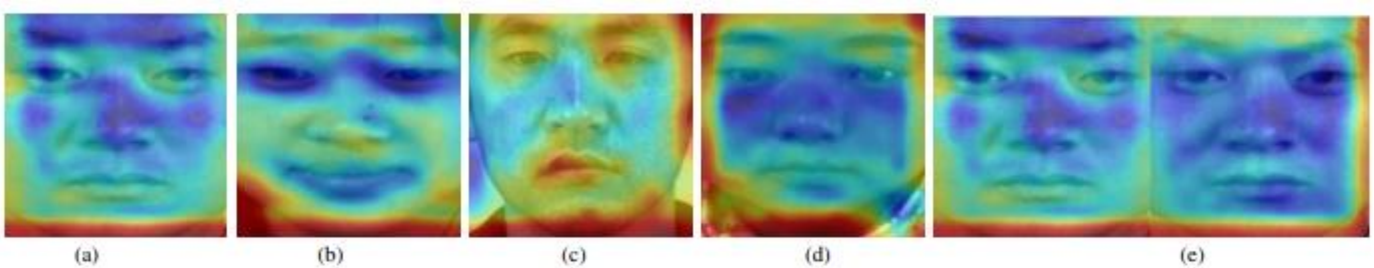


Figure 9. Grad-CAM visualizations of the ELRCN-TE model: *from left:* (a) Subject 1 (CASME II), Happiness; (b) Subject 2 (CASME II), Others; (c) Subject 13 (SAMM), Class III; (d) Subject 5 (CASME II), Objective class III, CDE protocol (e) Subject 1 (CASME II), Happiness, Comparison between single domain and cross domain experiments.

Khor, H. Q. et al. (2018). **Enriched Long-term Recurrent Convolutional Network for Facial Micro-Expression Recognition**. In *FG 2018*, pp. 667-674

Transfer learning of macro-trained deep models

- Train deep models on macro-expression apex samples ➔ Transfer learning on micro-expression apex samples
  - **ResNet10** pre-trained on 4 macro-exp. datasets using apex frames

    CK+ (852 images)

    Oulu CASIA NIR & VIS (1200 images)

    Jaffe (151 images)

    MUGFE (8228 images)

    TOTAL: 10,431 images → oversample to 5,000 images/expression
  - Fine-tuning on micro-exp datasets using apex frames → oversample to 200 images/expression
  - Assumption: That apex information is available!

	Hold-database Evaluation (HDE)	
	WAR	***UAR***
LBP-TOP	0.285	0.332
HOOF	0.353	0.348
HOG3D	0.363	0.228
Our method	**0.561**	**0.389**

TABLE VII. RECOGNITION ACCURACY AND F1 SCORE OF OUR METHODSON THE COMPOSITE DATASETS IN CDE

	Leave-One-subject-Out (LOSO)	
	Accuracy (%)	***F1 score***
Our method	**74.70**	**0.64**

Peng, M. et al. (2018). **From Macro to Micro Expression Recognition: Deep Learning on Small Datasets Using Transfer Learning** . In *FG 2018,* pp. 657-661

1st Micro-Expression Grand Challenge 2018 @ FG 2018, Xi'an, China

Insights:

- Cross-database task is challenging
  - Leveraging macro-expression samples seem to work reasonably well
  - Lack of data -> LSTMs not suitable
- Efforts underway to create a new large-scale database
- We need more people to work on this area!

Challenges

DATABASES

Subjectivity in humans

- Certain emotions (e.g. happiness) are easier to elicit compared to others (e.g. fear, sadness, anger)
- Some people are more "poker-faced" than others – they hide their emotions well!

Sample distribution

- Bias learning → Imbalanced distribution of samples per emotion, samples per subject

Creative strategies for inducement

- Complementary info from body region[1], or heart rate from skin variations[2]

[1] Song et al. (2013). **Learning a sparse codebook of facial and body micro expressions for emotion recognition**," Proc of ACM Int. Conf. Multimodal Interaction
[2] Gupta et al. (2018). **Exploring the feasibility of face video based instantaneous heart-rate for micro-expression spotting**," CVPR Workshops.

Challenges

DATABASES

Subject diversity

- Most datasets contain a majority of subjects from one particular country or ethnicity

Environment and setting

- Real-world scenarios are much needed: Job interviews, criminal interrogation, patient assessment etc. (but many cannot pass ethic committees!)
- How about “two truths and a lie” game?

Challenges

SPOTTING

Landmark detection

- Room for improvement in existing methods. ME requires very stable detection / robust against noise to capture minute changes in facial muscles.

Threshold or classify?

- Most existing works employ rule-based strategies ➔ Not robust and adaptable!
- Per-frame classification of ME occurrence ➔ Rigid and noisy!

Onset and offset detection

- Current works do not consider detecting the start and end frames, which could be useful to trim ME sequences before classification

Challenges

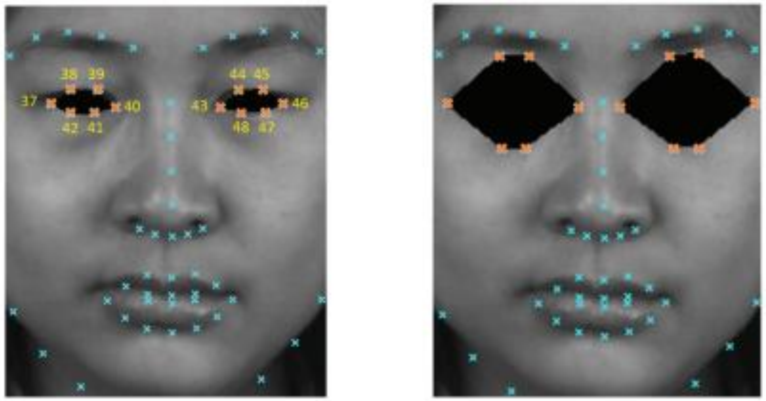
RECOGNITION

Block Selection

- Block-based methods of extracting features are quite popular
- Assignment of weights to blocks with key information ➔ New: Learn which blocks are discriminative[3]

Eyes: To Keep or Not To Keep?

- Some[4] works mask out eye regions to avoid eye blink motions, some[5] think otherwise



[3] Zong et al. (2018). **Learning from hierarchical spatiotemporal descriptors for micro-expression recognition.** IEEE T-MM
[4] Liong et al. (2016). **Automatic Micro-expression Recognition from Long Video using a Single Spotted Apex**," ACCV Workshops.
[5] Duan et al. (2016). **Recognizing spontaneous micro-expression from eye region**," Neurocomputing.

Challenges

RECOGNITION

Feature crafting / learning

- Most crafted features circa 2014-2016 are still holding strong results – shape, motion
- DL getting popular 2016 onwards – pushing the limits

Cross-DB recognition

- Realistic setting (multi-environment)
- How to generalise across domains[6]?
- MEGC[7] leads this effort

[6] Zong et al. (2018). **Domain regeneration for cross-database micro-expression recognition**. IEEE T-IP.
[7] Yap et al. (2018). **Facial Micro-Expressions Grand Challenge 2018 Summary,**," IEEE FG.

Experiment-related issues

Evaluation Protocol

- Use LOSO cross-validation[8] instead of LOVO cross-validation (some works still do this! ☹)
  - LOVO exposes the training to samples belonging to the test sample subject

Performance Metrics

- Use F1-score instead of Accuracy
  - Accuracy tends to be bias in imbalanced datasets or heavily skewed data
  - Use unweighted metrics that give equal emphasis to rare classes

Class Labels

- A few works consider fewer number of classes than it should be ➔ problem benchmarking!
- Emotion classes vs. Objective classes[9]

[8] Le Ngo et al. (2014). **Spontaneous subtle expression recognition: imbalanced databases and solutions**. ACCV
[9] Davison et al. (2017). **Objective classes for micro-facial expression recognition.** Arxiv.

2nd MEGC @ IEEE FG 2019

14th IEEE International Conference on
Automatic Face and Gesture Recognition
FG2019

https://facial-micro-expressiongc.github.io/MEGC2019/

2 Challenges:

- Cross-DB Recognition
- Spotting

Submission of papers: Challenge and non-challenge papers

Important Dates

- Submission Deadline: 27 January 2019
- Notification: 12 February 2019
- Camera-Ready: 15 February 2019

Acknowledgement

Many thanks to the our sponsors:

- Multimedia University, Malaysia
- Ministry of Higher Education, Fundamental Research Grant Scheme
- Belt and Road Initiative Research Scholar Exchange

our team throughout the years (2014-Current):

Anh Cat Le Ngo

Yandan Wang

Raphael C.W. Phan

Sze-Teng Liong

Yee-Hui Oh

Huai-Qian Khor

and collaborators:

Moi-Hoon Yap
Manchester Metro. Uni

Xiaopeng Hong
Univ of Oulu, Finland

Su-Jing Wang
Institute of Psychology, CAS

The Not Face

рахмат
danke
謝謝
ngiyabonga
شكراً جزيلاً
Баярлалаа
spas
teşekkür ederim
mersi
kia ora
barka
welalin
tack
mahalo
tapadh leat
спасибо
faafetai lava
vinaka
dank je
misaotra
paldies
grazzi
спасибі
blagodaram
matondo
kiitos
dankie
nanni
nandri
dhanyavad
mauruuru
köszönöm
thank you
aciū
gracias
хвала
asante
manana
obrigada
bedankt
bayarlalaa
gràçie
hvala
dankon
akun
murakoze
tenki
enkosi
djiere dieuf
tau
mochchakkeram
дякую
mamnun
dziękuję
chokrane
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
go raibh maith agat
sobodi
dėkuji
sagolun
sukriya
kop khun krap
obrigado
mèsi
didi madloba
ありがとう
arigatô
takk
dakujem
trugarez
najis tuke
terima kasih
tanemirt
grazie
rahmet
kam sah hamnida
rahmat
ευχαριστώ
diolch
dhanyavadagalu
shukriya
merce
мерси
তোমাকে ধন্যবাদ
감사합니다
xièxie
merci