A 3D FACE MODELING APPROACH FOR IN-THE-WILD FACIAL EXPRESSION RECOGNITION ON IMAGE DATASETS

Son Thai Ly, Nhu-Tai Do, Guee-Sang Lee, Soo-Hyung Kim, and Hyung-Jeong Yang

Dept. of Electronics and Computer Engineering, Chonnam National University, Korea

Introduction

➢ Recent studies mainly focus on learning from 2D images data.

➢ All available 3D databases were constructed in controlled conditions, not in-the-wild environments.

➢ 3D and 2D + 3D information are proved to be useful for FER. There are many papers on in-the-wild 3D face reconstruction.

➢ Therefore, in near future, 3D face could be employed to solve in-the-wild FER problems.

Contribution

➢ For the first time, a novel and competent deep learning approach for fusion 2D and 3D modalities in in-the-wild FER is proposed

➢ Further conduct experiments which combine 2D and 3D facial data in a deep learning manner. this is the first time such a method is reported in the context of FER.

➢ The proposed approach achieves competent recognition accuracy on RAF and SFEW 2.0 dataset.

Pipeline

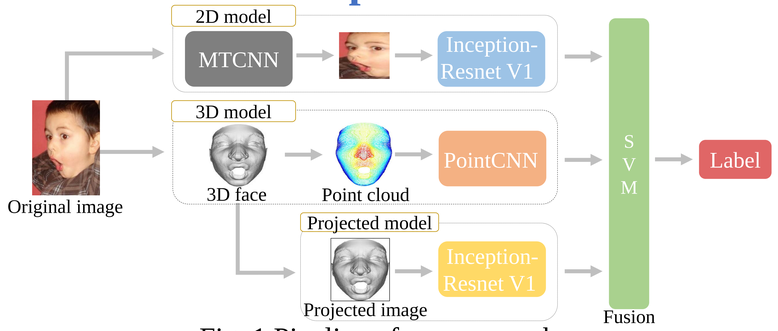

Fig. 1 Pipeline of our approach.

➢ Each model was trained separately and the extracted features are then concatenated and learned by SVMs.

Constructing 3D facial expression data

➢ This study benefits from Tran et al. [1] study for reconstructing the 3D face from the original image dataset.

➢ Once the reconstruction phase is done, the 3D face data are preprocessed as shown below.

Fig. 2 Examples of face reconstruction.

Datasets and training

➢ We evaluate our approach on RAF [2] and SFEW 2.0 [3].

➢ The 2D and projected models were trained with Inception-Resnet V1 network, batch size of 128 for 100 epochs.

➢ The 3D model was trained with PointCNN [4], batch size of 32 for 100 epochs.

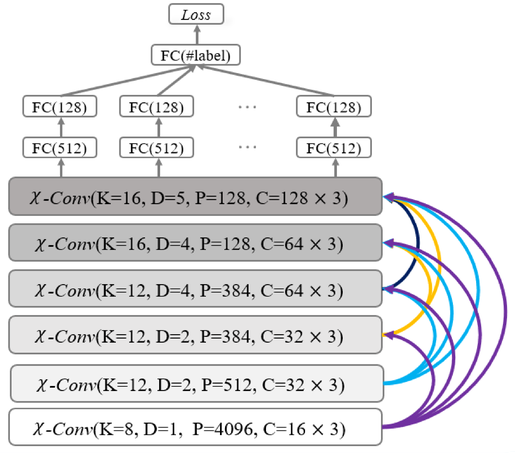

Fig. 3 The set of hyper-parameters for each layer of PointCNN with DenseNet-like links.

Results

Model	RAF	SFEW 2.0
LTNet [5]	86.7	58.2
Cov. Pooling [6]	87	58.1
Transfer learning [7]	80	55.8
DLP-CNN [8]	74.2	51
DSN [9]	84	-
Multimodal fusion [10]	83.8	-
2D model	82.4	52.5
3D model	62.5	38.2
Projected model	55.3	32.4
Fusion 2D + 3D	85	55.2
Fusion 3 models	86.1	56.2

Table 1 Comparison between with state-of-the-art.
All models were trained from scratch.

Discussion and future work

➢This study explores the benefits of 3D facial modeling for in-the-wild FER for the first time.

➢ It is also suggested that the 3D facial expression features could be harvested in many approaches and contributed to improve facial expression recognition performance.

➢ The 3D data preparation phase is time consuming. Therefore, we plan to construct an in-the-wild 3D facial expression database for the sake of academic purpose.

Reference

[1] A. T. Tran et al., CVPR 2018.
[2] S. Li et al. CVPR 2017.
[3] A. Dhall et al., IEEE Multimedia 2012.
[4] Y. Li et al., NeurIPS, 2018
[5] J. Zeng, S. Shan, X. Chen, ECCV 2018
[6] D. Acharya et al., CVPR Workshop 2018
[7] V. Vielzeuf et al., ICMI 2018.
[8] S. Li, W. Deng, J. Du, CVPR 2017
[9] Y. Fan, J. C. Lam, V. O. Li, ICMI 2018
[10] C. Liu, T. Tang, K. Lv, M. Wang, ICMI 2018