Emotion challenge fact sheet

iCV

February 2017

1 Team details

- Team name

  CBSR-CASIA

- Team leader name

  Jianzhu, Guo

- Team leader address, phone number and email

  Address: No.95 Zhongguancun East Road, Haidian District, Beijing 100190, China
  Phone: 86-15600118066
  Email: guojianzhu1994@foxmail.com

- Rest of the team members

  Shuai, Zhou (equal contribution)
  Jilin, Wu (equal contribution)
  Jun, Wan (corresponding author)
  Xiangyu, Zhu
  Zhen, Lei

- Team website URL (if any)

  `http://www.cbsr.ia.ac.cn/users/jwan/research.html`

- Affiliation

  National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences (CASIA).

2 Contribution details

- Title of the contribution

  Multi-modality Network with Visual and Geometry based information for Micro Emotion Recognition

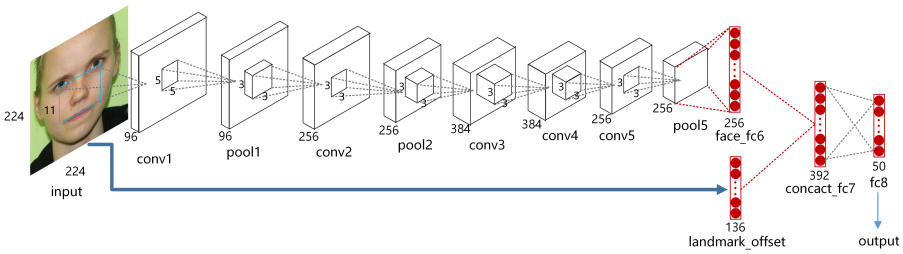

Figure 1: Architecture of the network architecture

- Final score

  It is not available now.

- General method description

  We fisrt extract the each ID's lamdmark feature, and subsract it's mean as the geometrical information, then we get each ID's landmark offset of $R^{136\times1}$. And we use an modified AlexNet to extract the $R^{256\times1}$ feature of the 224×224 image on the second-last full connected layer. Then we concact the two features, and connect to a full connected layer. We use Hinge Loss to train the net, the ground truth is a single label indicating the complementary and dominant emotion.

- References

  [1] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." neural information processing systems (2012): 1097-1105. [2] Kazemi, Vahid, and Josephine Sullivan. "One millisecond face alignment with an ensemble of regression trees." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.

- Representative image / diagram of the method

- Describe data preprocessing techniques applied (if any)

  Face detection and face landmark extraction using dlib and face crop and weakly aligned using the eyes center and the upper flip. We crop the image to 224×224 size.

3 Face Landmarks Detection

3.1 Features / Data representation

Describe features used or data representation model FOR FACE LANDMARKS DETECTION (if any)

3.2 Dimensionality reduction

Dimensionality reduction technique applied FOR FACE LANDMARKS DETECTION (if any)

3.3 Compositional model

Compositional model used, i.e. pictorial structure FOR FACE LANDMARKS DETECTION (if any)

3.4 Learning strategy

Learning strategy applied FOR FACE LANDMARKS DETECTION (if any)

3.5 Other techniques

Other technique/strategy used not included in previous items FOR FACE LANDMARKS DETECTION (if any)

3.6 Method complexity

Method complexity FOR FACE LANDMARKS DETECTION

4 Dominant emotion recognition

4.1 Features / Data representation

Describe features used or data representation model FOR DOMINANT EMOTION RECOGNITION (if any)

4.2 Dimensionality reduction

Dimensionality reduction technique applied FOR DOMINANT EMOTION RECOGNITION (if any)

4.3 Compositional model

Compositional model used, i.e. pictorial structure FOR DOMINANT EMOTION RECOGNITION (if any)

4.4 Learning strategy

Learning strategy applied FOR DOMINANT EMOTION RECOGNITION (if any)

4.5 Other techniques

Other technique/strategy used not included in previous items FOR DOMINANT EMOTION RECOGNITION (if any)

4.6 Method complexity

Method complexity FOR DOMINANT EMOTION RECOGNITION

5 Complementary emotion recognition

5.1 Features / Data representation

Describe features used or data representation model FOR COMPLEMENTARY EMOTION RECOGNITION (if any)

5.2 Dimensionality reduction

Dimensionality reduction technique applied FOR COMPLEMENTARY EMOTION RECOGNITION (if any)

5.3 Compositional model

Compositional model used, i.e. pictorial structure FOR COMPLEMENTARY EMOTION RECOGNITION (if any)

5.4 Learning strategy

Learning strategy applied FOR COMPLEMENTARY EMOTION RECOGNITION (if any)

5.5 Other techniques

Other technique/strategy used not included in previous items FOR COMPLEMENTARY EMOTION RECOGNITION (if any)

5.6 Method complexity

Method complexity FOR COMPLEMENTARY EMOTION RECOGNITION

6 Joint dominant and complementary emotion recognition

6.1 Features / Data representation

Describe features used or data representation model FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION (if any)

Landmark offset and AlextNet feature extraction.

6.2 Dimensionality reduction

Dimensionality reduction technique applied FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION (if any)

6.3 Compositional model

Compositional model used, i.e. pictorial structure FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION (if any)

6.4 Learning strategy

Learning strategy applied FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION (if any)

6.5 Other techniques

Other technique/strategy used not included in previous items FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION (if any)

6.6 Method complexity

Method complexity FOR JOINT DOMINANT AND COMPLEMENTARY EMOTION RECOGNITION

7 Global Method Description

- Total method complexity: all stages
- Which pre-trained or external methods have been used (for any stage, if any)
- Which additional data has been used in addition to the provided training and validation data (at any stage, if any)
- Qualitative advantages of the proposed solution

  If takes advantage of the geometry feature and texture pattern.

- Results of the comparison to other approaches (if any)
- Novelty degree of the solution and if is has been previously published

8 Other details

- Language and implementation details (including platform, memory, parallelization requirements)

  We use Python with OpenCV and Dlib package and Caffe deep learning framework.

- Detailed list of prerequisites for compilation
- Human effort required for implementation, training and validation?
- Training/testing expended time?

  Training takes about an hour on Titan-X GPU, validation of 7000 images takes about half minute.

- General comments and impressions of the challenge?

  The task is hard, and it's hard to find effective texture feature to do this challenge well.