



# Group Emotion Recognition with Individual Facial Emotion CNNs and Global Image based CNNs

Lianzhi Tan, Kaipeng Zhang, Kai Wang, Xiaoxing Zeng, Xiaojiang Peng, Yu Qiao  
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

## Introduction

### Aim:

Classify a group perceived emotion in an image as positive, neutral or negative.

### Motivation:

- Individual facial emotion is the most important cue for group emotion recognition
- Group emotion is relevant to the global scene, e.g. an image is likely to be positive taken from a wedding party while negative from a funeral

### Contributions:

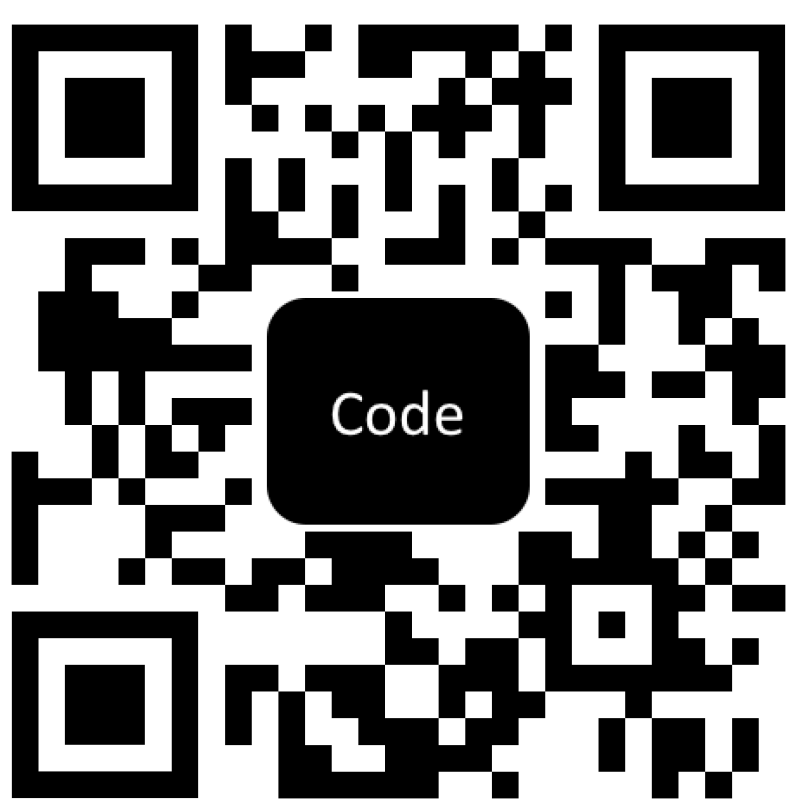
1. An aligned facial emotion CNN
2. An non-aligned facial emotion CNN
3. A global image based CNN
4. An extra dataset searched using train/val sets
5. Winner of the GREco task in the EmotiW17 challenge

## Submissions

- ▷ **1:** Aligned (3.0) + Global ResNet101 (0.5) + Non-aligned (0.5)
- ▷ **2:** Aligned (3.0) + Global VGG19 (0.5) + Non-aligned (0.5)
- ▷ **3:** Aligned (3.0, train on trainset) + Global ResNet101 (0.5) + Non-aligned (0.5)
- ▷ **4:** Aligned (3.0, train on trainset) + Global VGG19 (0.5) + Non-aligned (0.5)
- ▷ **5:** Aligned (3.2) + Global ResNet101 (0.8, extra data) + Non-aligned (0.6)
- ▷ **6:** Aligned (3.2) + Global VGG19 (0.8, extra data) + Non-aligned (0.6)
- ▷ **7:** Aligned (3.2, train on trainset) + Global ResNet101 (0.8, extra data) + Non-aligned (0.6)

Runs	Validation		Test		
	Overall	Positive	Neutral	Negative	Overall
1	82.7	82.0	57.6	76.7	79.9
2	81.7	82.3	57.0	74.7	79.1
3	-	85.2	63.6	67.6	78.9
4	-	85.2	64.84	66.6	78.8
5	83.7	83.9	58.8	76.4	<b>80.9</b>
6	83.8	83.9	58.2	74.0	79.8
7	-	85.5	59.4	70.0	79.1

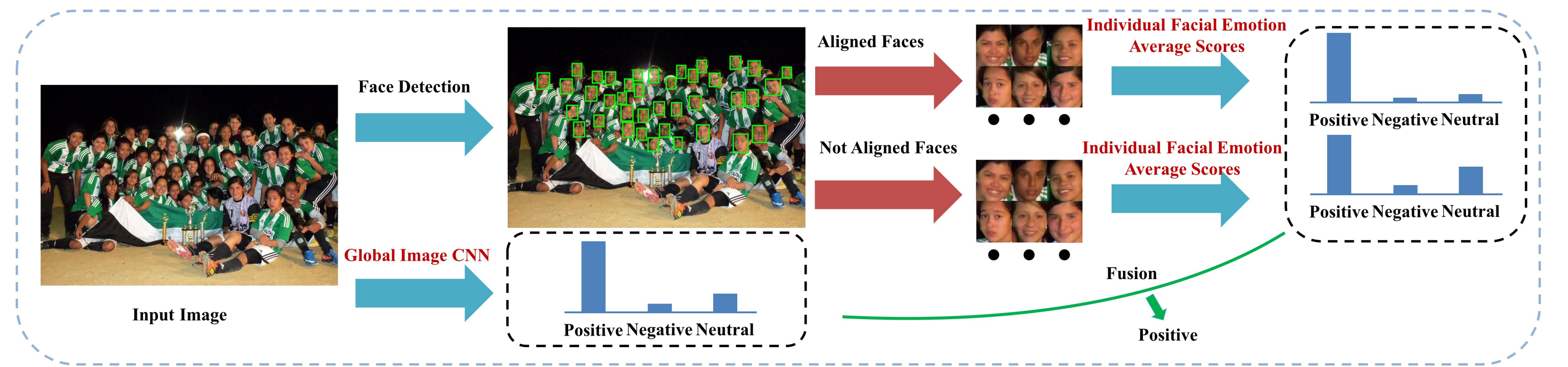
## Code and Extra Data



▷ **Code release:** <https://github.com/pengxj/GroupEmotionRecognition>

▷ **Extra data release:** <https://pan.baidu.com/s/1gf3tdzt>

## Framework



**Figure 1:** Our method contains two kinds of CNNs, namely the individual facial emotion CNNs and the global image based CNNs as indicated by red texts. The final prediction is made by averaging all the scores of CNNs from all faces and the image.

## Individual Facial Emotion CNNs

**Face detection:** We use MTCNN [2] to detect and align faces in the images

**Aligned Facial Model Architecture:**

- ResNet64 [4]

**Training:**

- With large-margin softmax (L-Softmax) [3], we pretrain it on Web-face, and then finetune on ExpW, finally finetune on EmotiW2017

**Non-aligned Facial Model Architecture:**

- ResNet34

**Training:**

- Pretrain on FERPlus, and then finetune on EmotiW2017

**Evaluation on validation set:**

	Aligned(96 × 112)		Non-aligned (48 × 48)
	Softmax	L-Softmax	Softmax
rm small faces	73.8	79.7	70.73
keep all faces	74.1	80.19	69.97

## Global image based CNNs

**Architecture:** VGG19, BN-Inception, ResNet101

**Training:** VGG19 is pretrained on Places dataset, and the others on ImageNet

**Evaluation on validation set:**

	VGG19		BN-Inception		ResNet101	
	Softmax	L-Softmax	Softmax	L-Softmax	Softmax	L-Softmax
+extra data	-	73.2	-	67.3	<b>74.7</b>	73.2
-	67.2	70.6	60.54	65.3	71.04	-

## Score combination

- ▷ **Aligned facial model:** ResNet64 with L-Softmax, and keep all faces
- ▷ **Global image model:** ResNet101 with Softmax
- ▷ **Non-aligned facial model:** ResNet34 with Softmax, and the top 3 faces for testing

Model ( score weight)	Acc. on validation set (%)
Aligned	80.2
Aligned (3.0) + Global (0.5)	81.0
Aligned (3.0) + Global (0.5) + Non-aligned (0.5)	82.7
Aligned (1.0) + Global (1.0) + Non-aligned (1.0)	78.4

## References

- [1] Abhinav. Dhall. From individual to group-level emotion recognition: EmotiW 5.0. In *ICMI'17*.
- [2] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 2016.
- [3] Weiyang Liu. Large-margin softmax loss for convolutional neural networks. In *ICML*, 2016.
- [4] Weiyang Liu. Sphreface: Deep hypersphere embedding for face recognition. In *ICCV*, 2017.