RAF-DB Facial Expression Recognition

Fan Yixuan

Department of Electronic Engineering, Tsinghua University

June 6, 2022

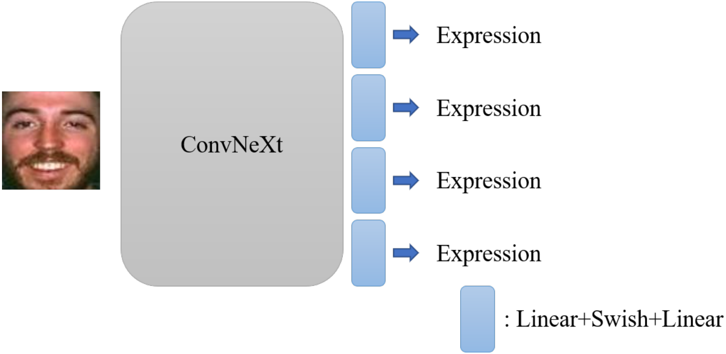


Figure 1: The distribution of the training samples from the seven expressions is sparse in the high-dimensional space, which leads to overfitting. Thus, the feature vector used for classification is divided into several sub-vectors and the corresponding classification heads are **ensembled**. If **channel of CNN** can be viewed as an equivalent structure of **token of Transformer**, it is similar to DeiT. It is also very similar to dropout, but shows superior performance on this dataset (+0.59%).

Experimental Results: **Accuracy 90.97%**

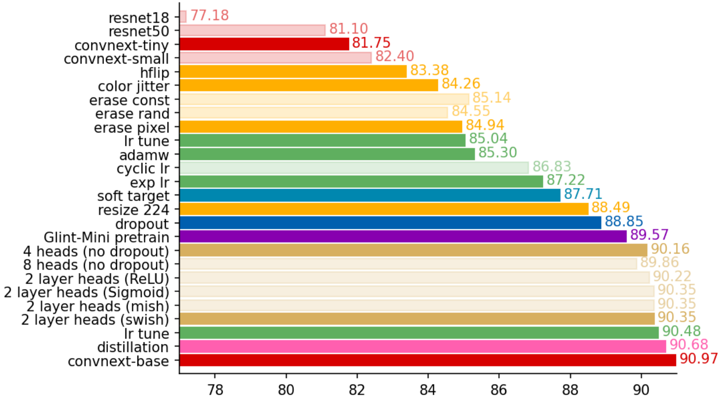


Figure 2: We achieve a performance that is competitive with SOTA (while using less external data) by stacking general tricks for image recognition. Structurally, we use a modern CNN without any attention-based modules and **propose a plug-and-play multi-head trick**.

Compare to SOTA

Method		Acc	Comment
RAN [8]	TIP 2020	86.90	
SCN [7]	CVPR 2020	87.03	
DACL [1]	WACV 2021	87.78	
KTN [2]	TIP 2021	88.07	
PSR [6]	CVPR 2020	88.98	
DMUE [4]	CVPR 2021	89.42	
FDRL [3]	CVPR 2021	89.47	
ARM [5]	arXiv 2021	90.42	Wrong approach with information disclosure.
TransFER [9]	ICCV 2021	90.91	Stacking IR50, 'Local CNN' and 8 layers of ViT, no code available.
POSTER [10]	arXiv 2022.4	92.05	Need a well-trained keypoint detector, no code available.
Ours		90.97	

Confusion Matrix

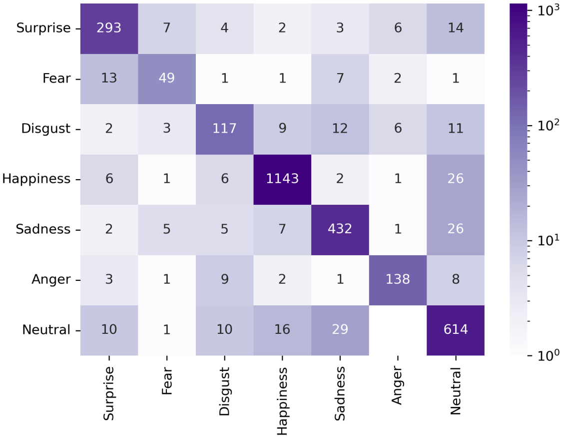
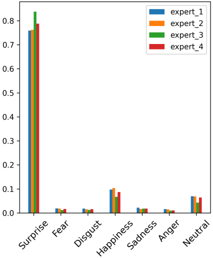
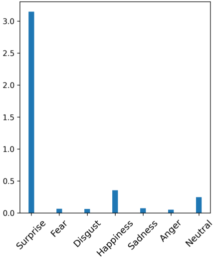


Figure 3: The vertical axis is the true label, and the horizontal axis is the predicted label.

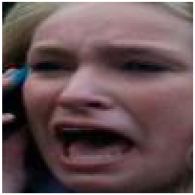
Mistakes

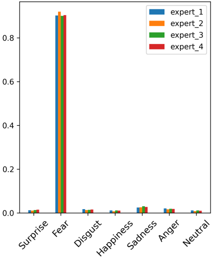
Happiness

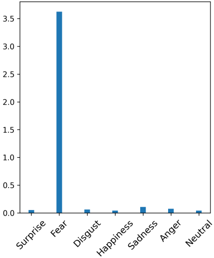


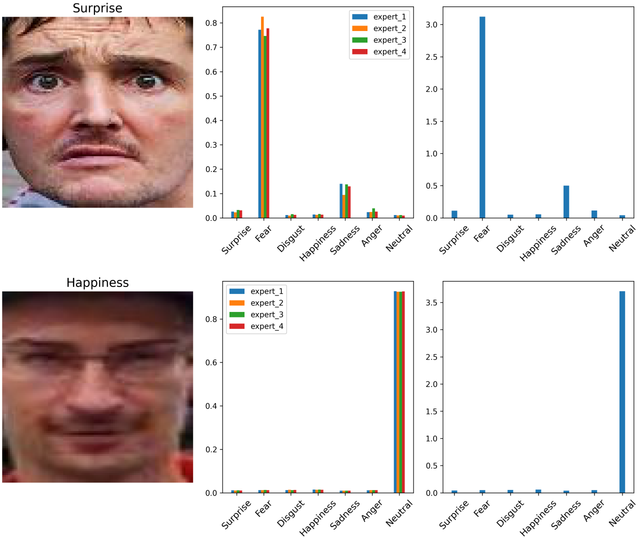


Sadness

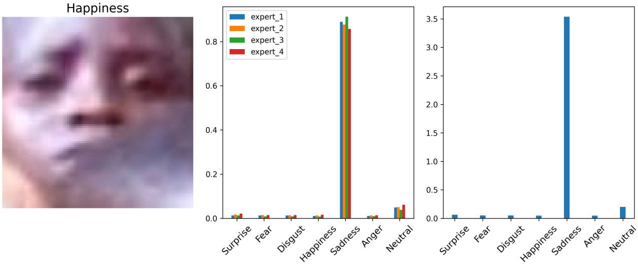


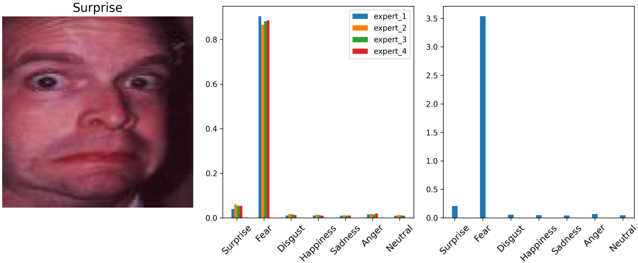





Surprise
Happiness
expert_1
expert_2
expert_3
expert_4
Surprise
Fear
Disgust
Happiness
Sadness
Anger
Neutral

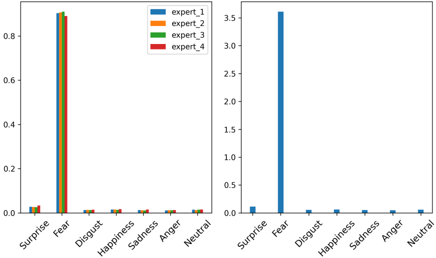
Mistakes


Happiness
expert_1
expert_2
expert_3
expert_4
0.0
0.2
0.4
0.6
0.8
Surprise
Fear
Disgust
Happiness
Sadness
Anger
Neutral
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Surprise
Fear
Disgust
Happiness
Sadness
Anger
Neutral

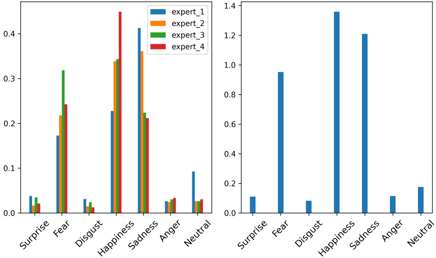

Surprise
expert_1
expert_2
expert_3
expert_4
0.0
0.2
0.4
0.6
0.8
Surprise
Fear
Disgust
Happiness
Sadness
Anger
Neutral
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Surprise
Fear
Disgust
Happiness
Sadness
Anger
Neutral

Mistakes

Surprise

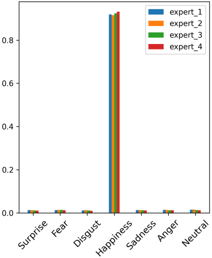


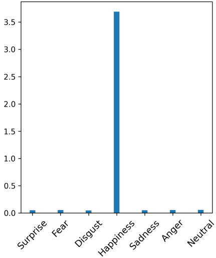
Sadness



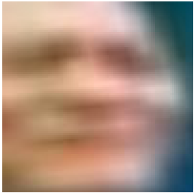
Mistakes

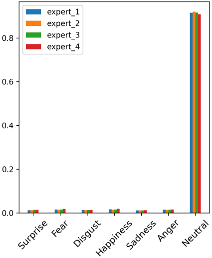
Sadness

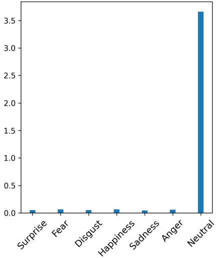




Happiness







[1] Amir Hossein Farzaneh and Xiaojun Qi.
Facial expression recognition in the wild via deep attentive center loss.
In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2402–2411, 2021.

[2] Hangyu Li, Nannan Wang, Xinpeng Ding, Xi Yang, and Xinbo Gao.
Adaptively learning facial expression representation via cf labels and distillation.
IEEE Transactions on Image Processing, 30:2016–2028, 2021.

[3] Delian Ruan, Yan Yan, Shenqi Lai, Zhenhua Chai, Chunhua Shen, and Hanzi Wang.
Feature decomposition and reconstruction learning for effective facial expression recognition.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7660–7669, 2021.

[4] Jiahui She, Yibo Hu, Hailin Shi, Jun Wang, Qiu Shen, and Tao Mei.
Dive into ambiguity: latent distribution mining and pairwise uncertainty estimation for facial expression recognition.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6248–6257, 2021.

[5] Jiawei Shi, Songhao Zhu, and Zhiwei Liang.
Learning to amend facial expression representation via de-albino and affinity.
arXiv preprint arXiv:2103.10189, 2021.

[6] Thanh-Hung Vo, Guee-Sang Lee, Hyung-Jeong Yang, and Soo-Hyung Kim.
Pyramid with super resolution for in-the-wild facial expression recognition.
IEEE Access, 8:131988–132001, 2020.

[7] Kai Wang, Xiaojiang Peng, Jianfei Yang, Shijian Lu, and Yu Qiao.
Suppressing uncertainties for large-scale facial expression recognition.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6897–6906, 2020.

[8] Kai Wang, Xiaojiang Peng, Jianfei Yang, Debin Meng, and Yu Qiao.
Region attention networks for pose and occlusion robust facial expression recognition.
IEEE Transactions on Image Processing, 29:4057–4069, 2020.

[9] Fanglei Xue, Qiangchang Wang, and Guodong Guo.
Transfer: Learning relation-aware facial expression representations with transformers.
In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 3601–3610, October 2021.

[10] Ce Zheng, Matias Mendieta, and Chen Chen.
Poster: A pyramid cross-fusion transformer network for facial expression recognition.
arXiv preprint arXiv:2204.04083, 2022.

Thanks!