Face2Exp: Combating Data Biases for Facial Expression Recognition

Dan Zeng, Zhiyuan Lin, Xiao Yan, Yuting Liu, Fei Wang, Bo Tang

SUSTech Southern University of Science and Technology

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

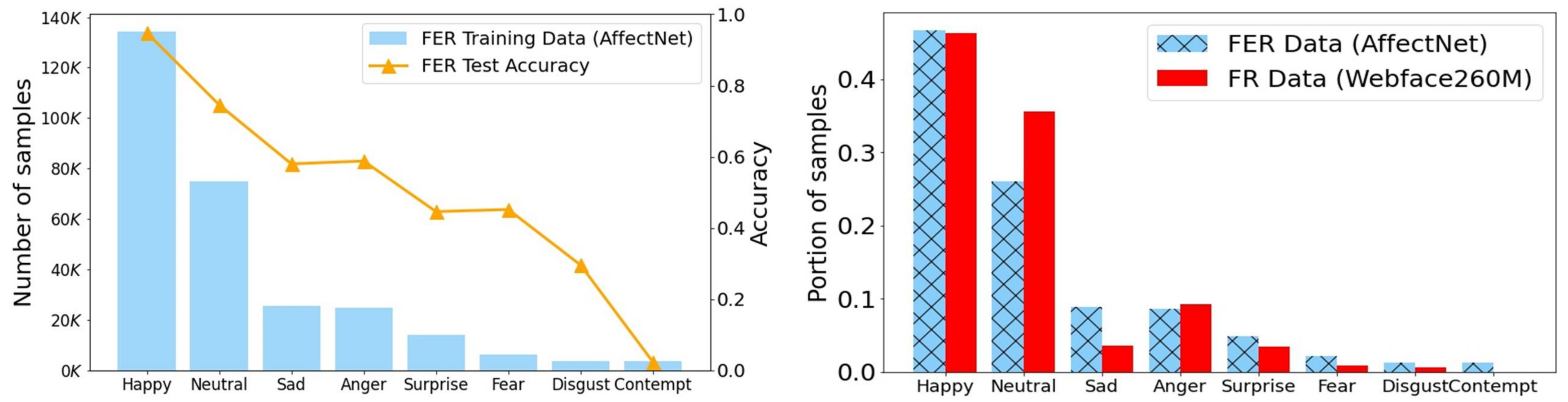
Project website: https://github.com/danzeng1990/Face2Exp.

Introduction

Face2Exp: utilizes large unlabeled FR datasets to enhance FER

Challenges

- ➢ Class imbalance in FER data (a: 1st data bias)
- ➢ Distribution mismatch between FR and FER data (b: 2nd data bias)



(a) (b)

- ➢ The image gap between FR and FER data is large

Each column represent the same facial expression.

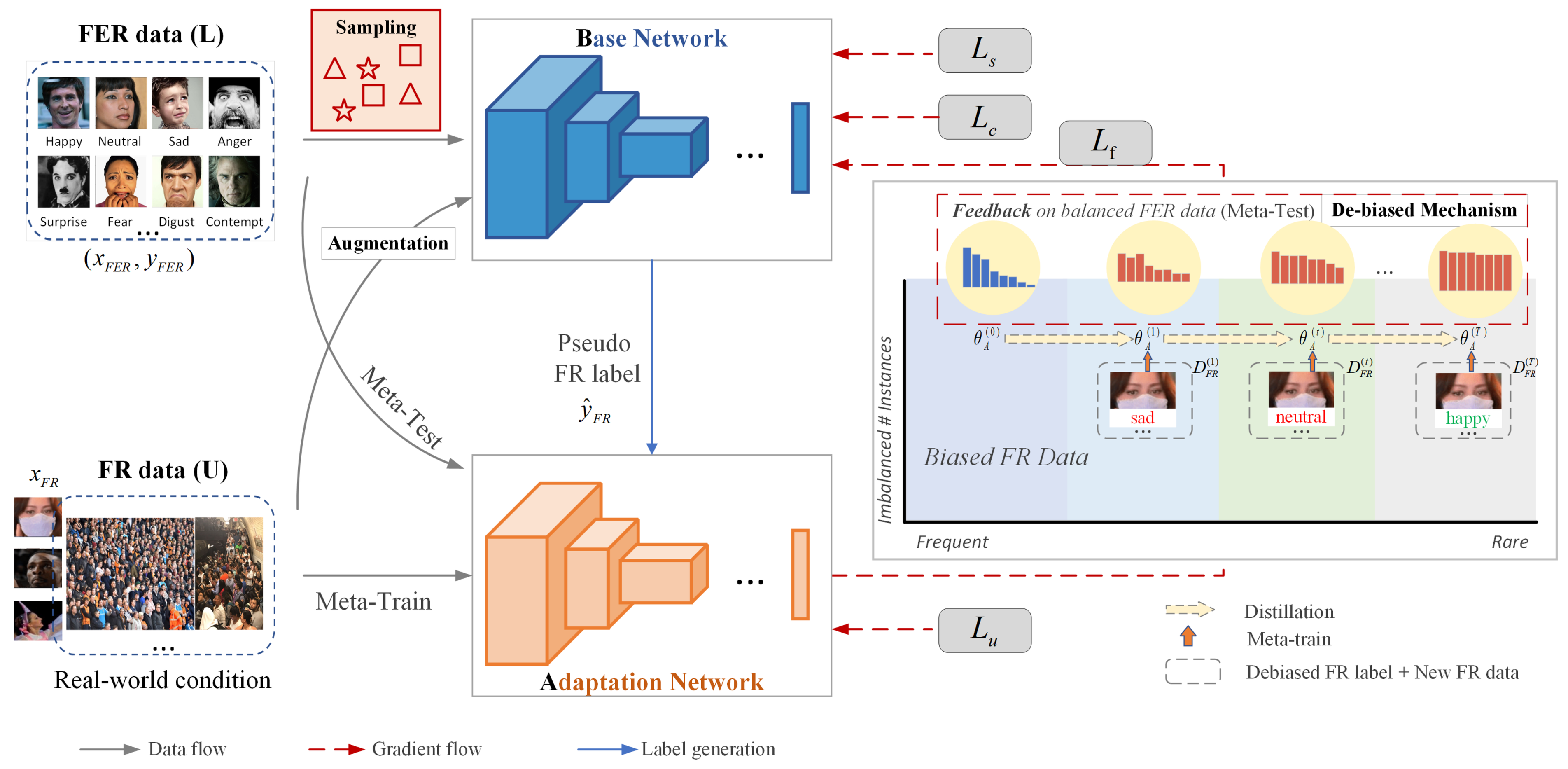
Motivation

- ➢ A base network learns prior expression knowledge from class balanced FER data (to solve (a))
- ➢ The circuit feedback mechanism improves the base network with the cognitive differences from the adaptation network (to solve (b))

Contributions

- ➢ A **general** framework to utilize large unlabeled FR data for other face-related tasks (e.g., gender/race classification, age estimation) that lack high quality data
- ➢ Extract **de-biased** knowledge from auxiliary FR data
- ➢ Obtain comparable results to the state-of-the-art FER methods with **10% labeled** FER data

Method: Meta-Face2Exp



Loss functions

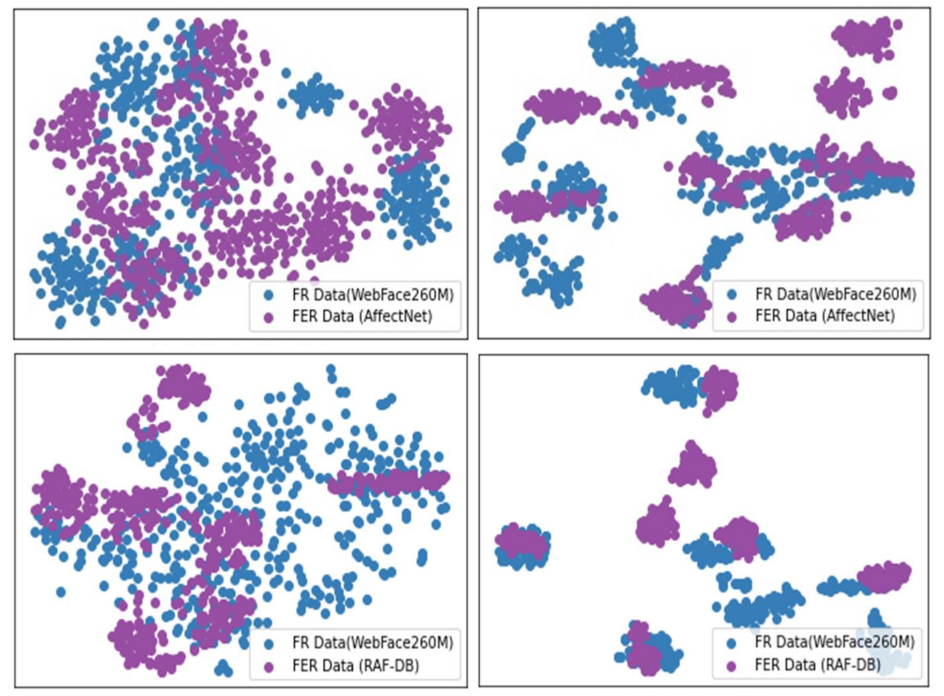
$$\mathcal{L}_u = CE(\hat{y}_{FR}, \mathcal{A}(x_{FR};\theta_{\mathcal{A}})) \qquad \mathcal{L}_s = CE(y_{FER}, \mathcal{B}(x_{FER};\theta_{\mathcal{B}}))$$

$$\mathcal{L}_c = CE(\mathcal{B}(x_{FR};\theta_{\mathcal{B}}), \mathcal{B}(Aug(x_{FR});\theta_{\mathcal{B}})) \qquad \mathcal{L}_f = f \cdot CE(\hat{y}_{FR}, \mathcal{B}(x_{FR};\theta_{\mathcal{B}}))$$

De-biased Mechanism

$$f = \eta_{\mathcal{A}} \cdot \left(\nabla_{\theta_{\mathcal{A}}^{(t+1)}} CE\left(y_{FER}, \mathcal{A}\left(x_{FER};\theta_{\mathcal{A}}^{(t+1)}\right)\right)\right)^T \cdot \nabla_{\theta_{\mathcal{A}}^{(t)}} CE\left(\hat{y}_{FR}, \mathcal{A}\left(x_{FR};\theta_{\mathcal{A}}^{(t)}\right)\right)$$

- ➢ 1st term: the gradients of the *new* adaptation network on debiased FER data
- ➢ 2nd term: the gradients of the *old* adaptation network on biased FR data



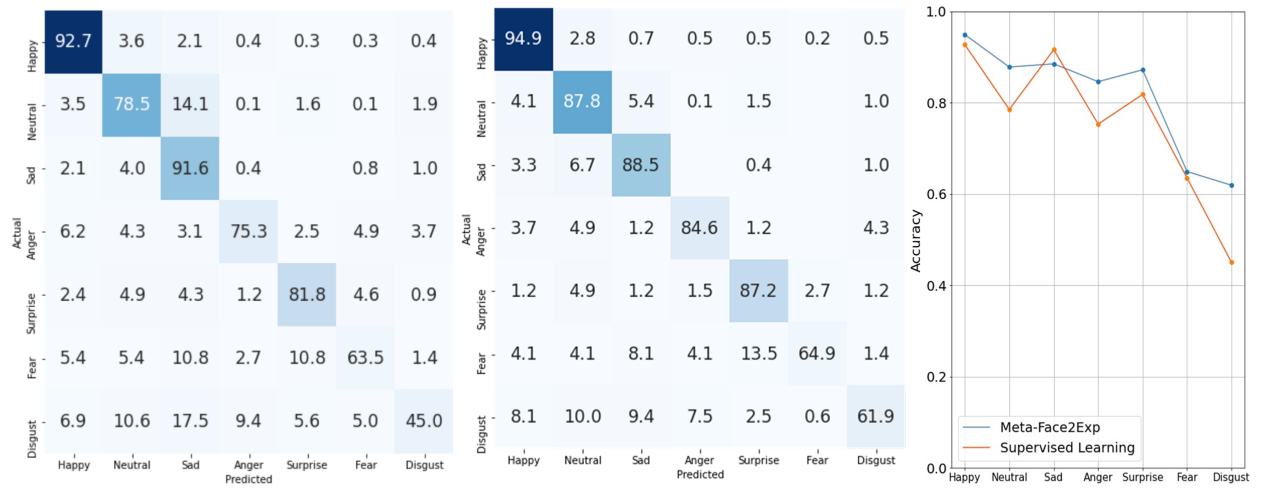
(a) Supervised Learning Model (b) Meta-Face2Exp

- ➢ If two terms have the same/different gradient sign, the base network is updated based on the same/adverse of the current gradients
- ➢ The absolute value of the dot product is the strength of the gradient updates
- ➢ The figure shows that FR and FER data align better with our method

Experiments

Evaluation on class imbalance

➢ De-biased behavior



SL model Meta-Face2Exp Comparison

Meta-Face2Exp yields more balanced accuracy, despite that it is trained with unbalanced FER data. (RAF-DB dataset)

➢ Size of labeled FER data

AffectNet

Models	SL	Meta-Face2Exp		
Data size	100%	1%	5%	10%
Mean Acc	58.37	53.54	61.66	64.23
Std Acc	21.53	14.41	10.69	10.07

RAF-DB

Models	SL	Meta-Face2Exp		
Data size	100%	25%	50%	100%
Mean Acc	84.16	80.87	85.04	88.54
Std Acc	15.48	9.43	9.43	10.00

- Higher mean accuracy with small labeled data (e.g., 5% AffectNet).
- Significantly lower the std accuracy.

➢ Effect of unbalanced FR data

- Meta-Face2Exp is general across different FR datasets (i.e., training our model with Webface260M/ VGGFace2 yields similar accuracy.)

Conclusion

- ➢ Combat two data biases: class imbalance and distribution mismatch
- ➢ Two networks constantly complement each other to extract de-biased knowledge through the circuit feedback paradigm
- ➢ Meta-Face2Exp effectively produces low std and high mean accuracy

Homepage Wechat