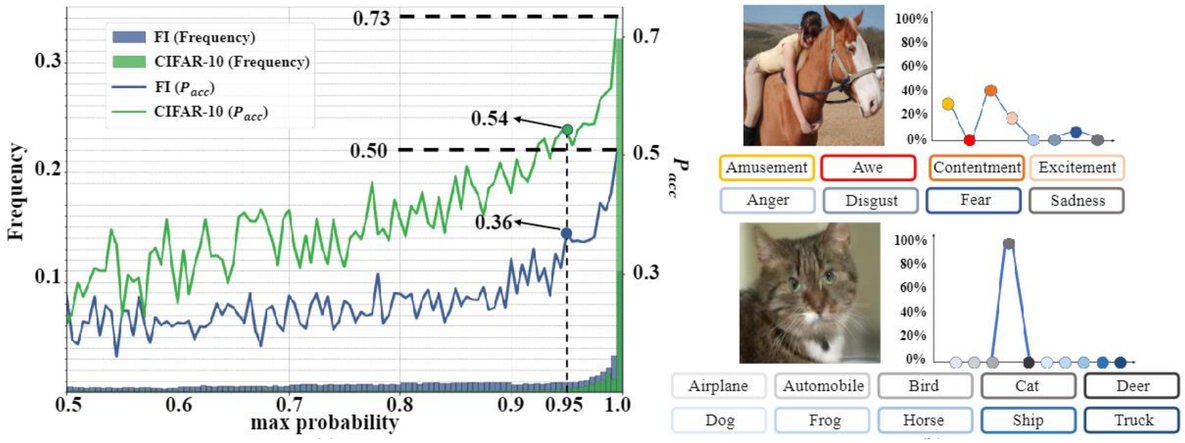
S^2-VER: Semi-Supervised Visual Emotion Recognition

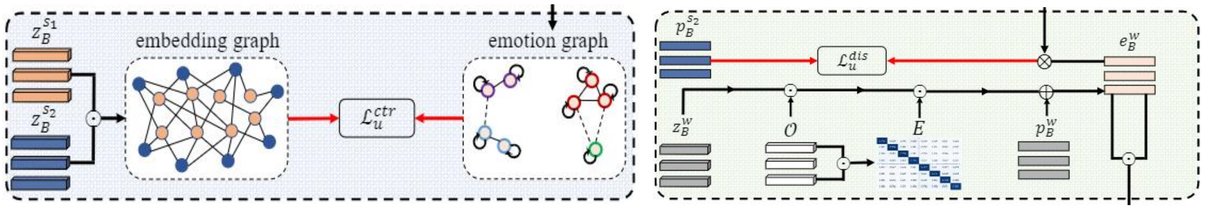
Guoli Jia, Jufeng Yang

南开大学
Nankai University

Semi-Supervised Visual Emotion Recognition



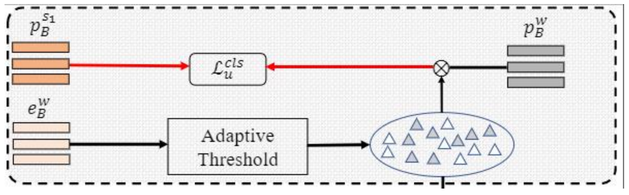
Reliable Emotion Label Learning



Learning emotional discriminative representation

Generating smoothing label

Ambiguity-Aware Adaptive Threshold



Addressing the ambiguity by Adjusting the threshold for each sample, based on the polarity.

Code is publicly available: https://github.com/exped1230/S2-VER

exped1230@gmail.com, yangjufeng@nankai.edu.cn

Method	FI			SE30K8			LDL		
	80	800	1600	80	400	800	80	800	1600
Yang *et al.* [42]	19.9±0.36	25.4±0.37	30.1±0.33	19.8±0.41	22.7±0.32	26.0±0.55	21.4±0.26	26.5±0.29	32.3±0.41
RCA [41]	18.4±0.33	25.9±0.39	31.4±0.17	18.6±0.29	21.9±0.33	26.5±0.33	23.8±0.48	29.2±0.19	33.2±0.21
WSCNet [40]	20.2±0.37	27.5±0.41	31.2±0.39	18.4±0.25	23.2±0.28	27.4±0.36	22.3±0.46	29.2±0.31	35.2±0.45
PDANet [50]	21.4±0.26	26.6±0.22	33.2±0.31	20.6±0.18	23.4±0.25	27.7±0.45	23.5±0.29	30.5±0.30	33.5±0.33
π-Model [16]	22.9±0.54	28.3±0.36	31.7±0.21	22.1±1.71	23.7±0.69	26.9±0.31	24.3±0.61	31.8±0.59	34.4±0.27
Pseu-Lab [17]	22.9±0.48	31.3±0.43	33.5±0.31	23.4±1.10	25.9±0.50	27.6±0.16	24.2±0.49	32.3±0.44	35.8±0.13
VAT [21]	23.6±0.78	31.5±0.77	35.1±0.37	24.4±0.69	27.2±0.39	28.9±0.25	26.3±0.58	34.5±0.49	38.9±0.36
Mean-Tea [32]	23.8±0.51	29.3±0.48	33.9±0.33	24.3±0.67	26.7±0.53	28.2±0.22	26.6±0.54	33.8±0.42	38.6±0.20
MixMatch [4]	26.3±1.53	35.1±0.74	38.0±0.32	26.6±0.87	28.3±0.62	29.6±0.40	28.1±0.78	34.2±0.52	38.9±0.23
ReMix [3]	29.7±0.68	35.4±0.53	38.3±0.42	26.4±1.10	29.9±0.98	31.9±0.63	29.1±0.67	35.3±0.54	39.2±0.35
UDA [36]	28.5±0.87	37.7±0.56	40.3±0.38	27.3±0.89	29.6±0.64	32.2±0.37	30.7±0.76	40.9±0.58	43.4±0.47
FixMatch [30]	28.2±0.78	37.4±0.51	42.2±0.29	29.7±0.70	32.2±0.57	32.7±0.46	32.4±0.84	39.4±0.45	43.2±0.24
FlexMatch [47]	29.7±0.90	38.2±0.49	42.9±0.17	28.5±1.03	33.2±0.60	33.9±0.26	33.2±0.93	41.3±0.71	46.7±0.42
CoMatch [18]	36.7±0.87	43.5±0.39	47.9±0.26	29.9±0.65	32.5±0.47	35.3±0.26	**38.1**±0.53	42.1±0.31	45.3±0.27
Ours	**39.1**±0.66	**46.9**±0.46	**51.8**±0.21	**30.1**±0.73	**33.3**±0.62	**36.2**±0.49	37.9±0.80	**43.6**±0.47	**47.4**±0.43

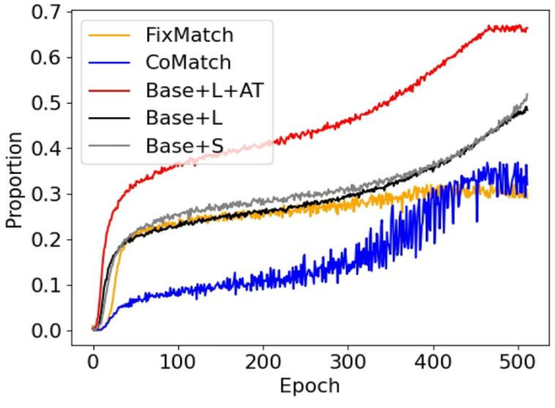
Comparison with SOTA methods

Method	Emotion-6			UnBiasedEmo		
	20%	50%	100%	20%	50%	100%
FixMatch [30]	46.6	48.3	49.0	65.6	69.2	71.5
FlexMatch [47]	48.1	50.1	51.2	67.2	71.1	73.4
CoMatch [18]	50.2	51.2	52.4	68.5	70.8	73.8
Ours	**51.7**	**53.5**	**54.0**	**70.8**	**76.7**	**78.7**

- SSL algorithm is beneficial for small-scale datasets
- S^2-VER performs better than other methods

M_{acc}	150	300	450	Avg
FixMatch	44.9	59.7	61.8	55.5
FlexMatch	36.6	44.2	47.4	42.7
FlexMatch(W)	31.4	37.4	40.6	36.5
CoMatch	48.1	67.4	66.9	60.8
Base + S	59.8	70.7	67.8	66.1
Base + L	**58.9**	**72.3**	**69.3**	**66.8**

Accuracy of Pseudo labels reaching threshold



Proportion reaching threshold