Learn From All: Erasing Attention Consistency for Noisy Label Facial Expression Recognition

Yuhang Zhang, Chengrui Wang, Xu Ling, Weihong Deng

Noisy Label Facial Expression Recognition

Previous:
Select small-loss samples as clean samples for training, which needs to know the noise rate.
Ours:
Suppress the model from remembering noisy labels in the feature level without knowing the exact noise rate.

SCN only focus on a part of the feature that can be considered related to the noisy labels to remember the noisy labels.

Our EAC forces the model to focus on similar parts before and after the flip to prevent the model from remembering noisy labels.

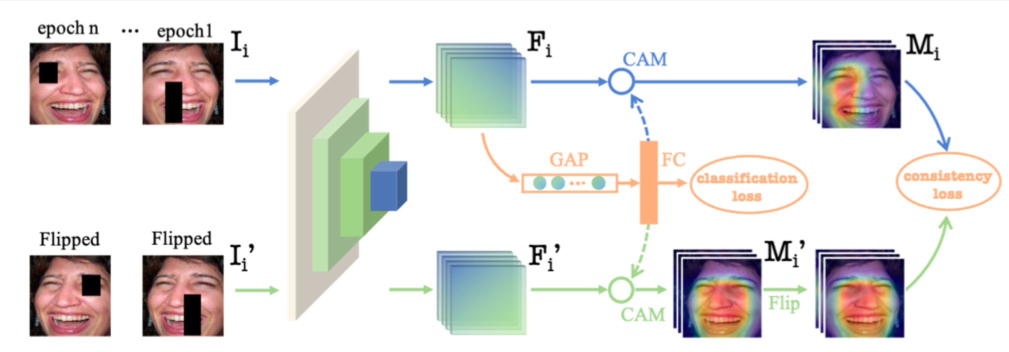
Motivation

Model focus on parts of the feature to remember noisy labels, we need to guide the model to focus on the whole feature.

The semantic meaning is the same of the original images and there flipped ones.

Acquire codes and pre-trained models through scanning the QR code in the upper right.

Simple Technique: Erasing Attention Consistency

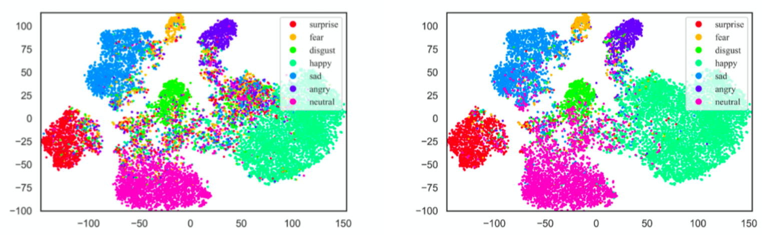


1. Imbalanced framework: Only compute classification loss using original images
2. Flip attention consistency: Prevent model from remembering noisy labels of original images through attention map consistency with the flipped ones.
3. Random Erasing: Guide the model to focus on the whole feature instead of parts of the feature and regularize the model from overfitting the flipped attention maps.

Learning with noisy labels in Facial Expression Recognition

Method	Noise(%)	RAF-DB(%)	FERPlus(%)	AffectNet(%)
Baseline	10	81.01	83.29	57.24
SCN (CVPR20)	10	82.15	84.99	58.60
RUL (NeurIPS21)	10	86.17	86.93	60.54
EAC (Ours)	10	88.02	87.03	61.11
Baseline	20	77.98	82.34	55.89
SCN (CVPR20)	20	79.79	83.35	57.51
RUL (NeurIPS21)	20	84.32	85.05	59.01
EAC (Ours)	20	86.05	86.07	60.29
Baseline	30	75.50	79.77	52.16
SCN (CVPR20)	30	77.45	82.20	54.60
RUL (NeurIPS21)	30	82.06	83.90	56.93
EAC (Ours)	30	84.42	85.44	58.91

Feature Visualization



(a). EAC learned features labeled with the noisy labels (b). EAC learned features labeled with the latent truth

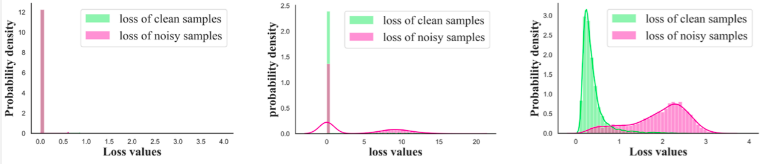
The learned features of our method labeled with noisy labels and latent truth.

Notice that noisy samples are pushed to the boundary by our method when learning with noisy labels in the left image.

The right image shows that our method learns features related to the latent truth even training with noisy labels.

Attention Maps and Loss Values

Our method predicts consistent attention maps on the flipped images.



(a). Baseline (b). SCN (c). EAC

Loss of clean and noisy samples after training for 60 epochs.