Robust Emotion Recognition in Context Debiasing

Dingkang Yang, Kun Yang, Mingcheng Li, Shunli Wang, Shuaibing Wang, Lihua Zhang

Academy for Engineering and Technology, Fudan University, Shanghai, China

Background & Motivation

Context-Aware Emotion Recognition (CAER)

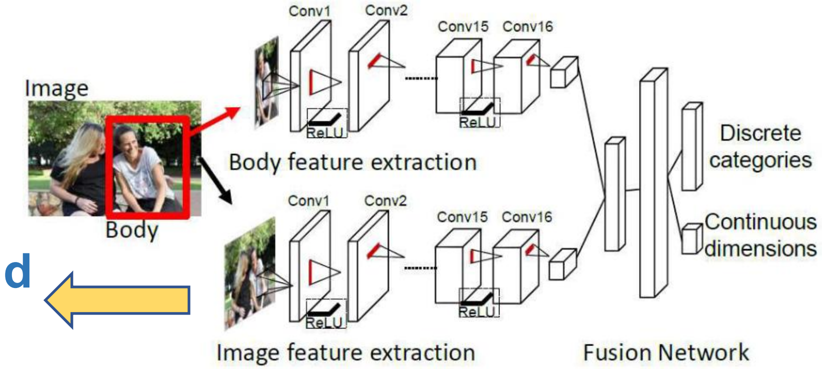
Input image

Person information

Context information

Emotion recognition

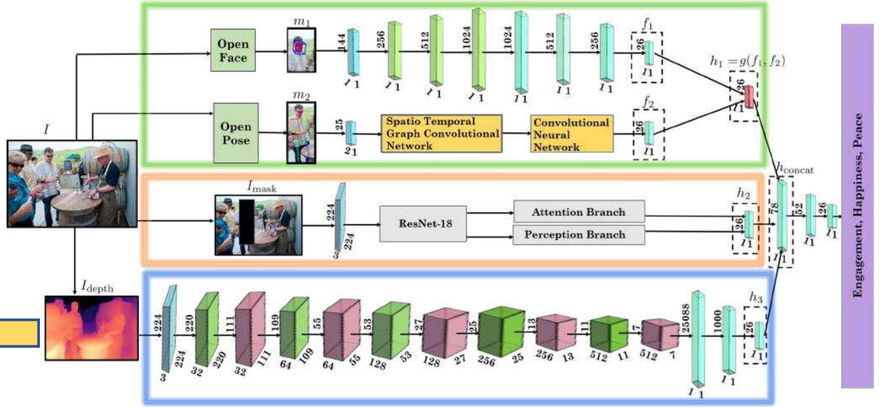
Background context



EMOT-Net

Kosti R, Alvarez J M, Recasens A, et al. Emotion recognition in context. In CVPR 2017: 1667-1675.

Scene and socio-dynamic contexts

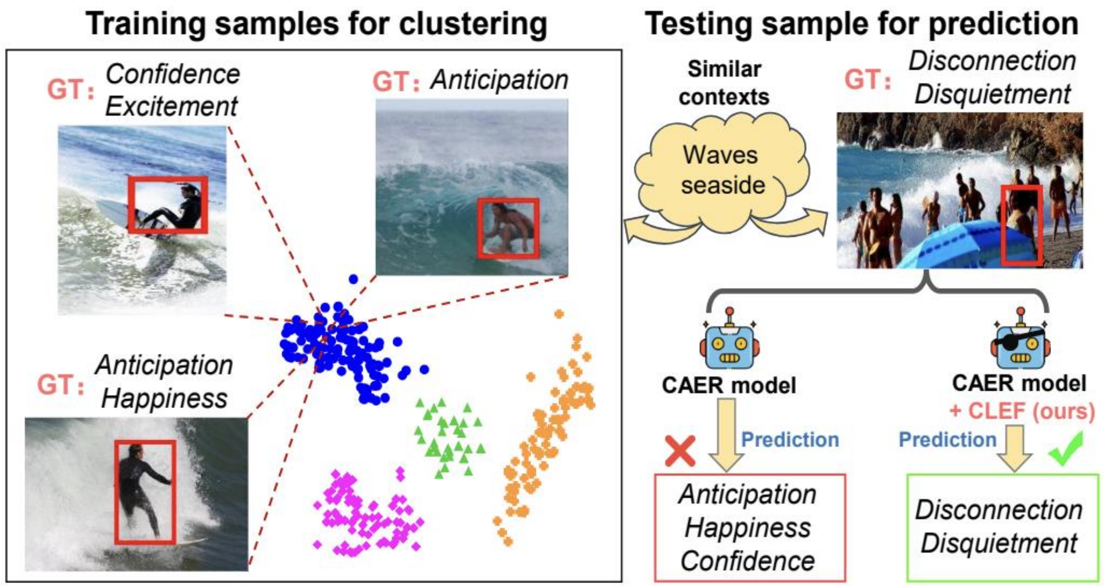


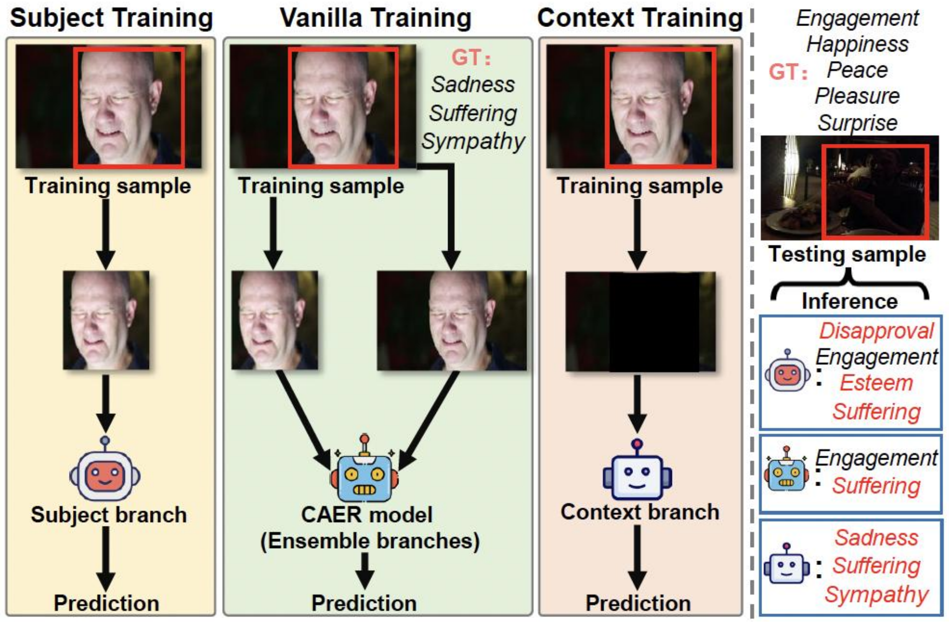
EmotiCon

Mittal T, Guhan P, Bhattacharya U, et al. Emoticon: Context-aware multimodal emotion recognition using frege's principle. In CVPR 2020: 14234-14243.

Background & Motivation

The Context Bias in the CAER Task

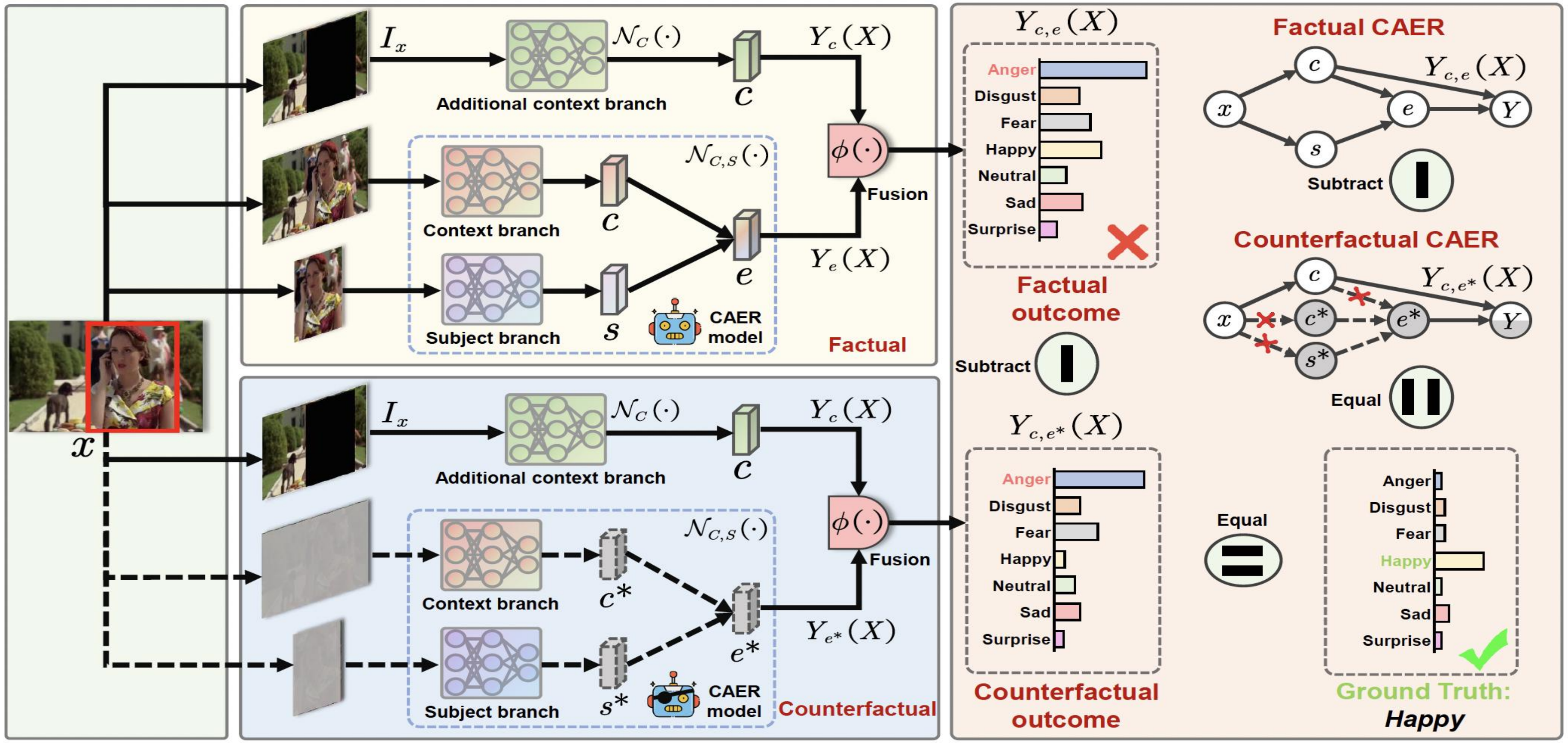




- Context-specific semantics easily yield spurious shortcuts with emotion labels during training to confound the model, giving erroneous results.
- Conversely, our CLEF effectively corrects biased predictions.
- The indirect effect of the good context prior follows ensemble branches, narrowing the emotion candidate space.
- The bad direct effect follows the context branch, causing pure bias.

Methodology

CVPR SEATTLE, WA JUNE 17-21, 2024



In addition to the vanilla CAER model, we introduce an additional context branch in a non-intrusive manner to capture the pure context bias as the direct context effect. By comparing factual and counterfactual outcomes, our framework effectively mitigates the interference of the harmful bias and achieves debiased emotion inference.

Experiments

Methods	mAP (%)
HLCR [7]	30.02
TEKG [5]	31.36
RRLA [24]	32.41
VRD [14]	35.16
SIB-Net [25]	35.41
MCA [56]	37.73
EMOT-Net [19]	27.93
EMOT-Net + **CLEF**	**31.67** (↑ 3.74)
CAER-Net [20]	23.85
CAER-Net + **CLEF**	**27.44** (↑ 3.59)
GNN-CNN [65]	28.16
GNN-CNN + **CLEF**	**32.18** (↑ 4.02)
CD-Net [53]	28.87
CD-Net + **CLEF**	**32.51** (↑ 3.64)
EmotiCon [32]	35.28
EmotiCon + **CLEF**	**38.05** (↑ 2.77)

Quantitative results on EMOTIC.

Methods	Accuracy (%)
Fine-tuned VGGNet [43]	64.85
Fine-tuned ResNet [13]	68.46
SIB-Net [25]	74.56
MCA [56]	79.57
GRERN [11]	81.31
RRLA [24]	84.82
VRD [14]	90.49
EMOT-Net [19]	74.51
EMOT-Net + **CLEF**	**77.03** (↑ 2.52)
CAER-Net [20]	73.47
CAER-Net + **CLEF**	**75.86** (↑ 2.39)
GNN-CNN [65]	77.21
GNN-CNN + **CLEF**	**79.53** (↑ 2.32)
CD-Net [53]	85.33
CD-Net + **CLEF**	**88.41** (↑ 3.08)
EmotiCon [32]	88.65
EmotiCon + **CLEF**	**90.62** (↑ 1.97)

Quantitative results on CAER-S.

Experiments

EMOTIC Dataset

	Testing Image	Ground Truth	Vanilla Method	w/ CLEF
(a)		Disconnection Disquietment Doubt/Confusion Engagement	Affection Happiness Peace Sympathy	Disconnection Disquietment Doubt/Confusion Engagement
(b)		Anticipation Doubt/Confusion Engagement	Anticipation Confidence Disapproval Disconnection Embarrassment	Anticipation Doubt/Confusion Engagement Suffering
(c)		Confidence Excitement Sensitivity Yearning	Pain Suffering Sensitivity	Confidence Excitement Sensitivity Yearning

CAER-S Dataset

	Testing Image	Ground Truth	Vanilla Method	w/ CLEF
(d)		Anger	Neutral	Anger
(e)		Happy	Sad	Happy
(f)		Disgust	Happy	Disgust

Qualitative results of the vanilla and CLEF-based baseline on EMOTIC and CAER-S datasets.

Conclusion

CVPR
SEATTLE, WA JUNE 17-21, 2024

- ✓ We are the first to embrace counterfactual thinking to investigate causal effects in the CAER task and reveal that the context bias as the adverse direct causal effect misleads the models to produce spurious prediction shortcuts.

- ✓ We devise CLEF, a model-agnostic CAER debiasing framework that facilitates existing methods to capture valuable causal relationships and mitigate the harmful bias in context semantics through counterfactual inference. CLEF can be readily adapted to state-of-the-art (SOTA) methods with different structures, bringing consistent and significant performance gains.

- ✓ Extensive experiments are conducted on several largescale CAER datasets. Comprehensive analyses show the broad applicability and effectiveness of our framework.

Thanks!