

Exploiting Granger Causality for Dynamic Multimodal Emotion Recognition in Video Learning Context

Table

Table 1 Comparative Results with Other Models

Method	Model	Accuracy(%)	Recall(%)	F1-score
Feature Extraction Method	LSTM [1]	67.59	67.89	0.67
	CNN-LSTM [2]	73.23	73.14	0.73
	Transformer [3]	79.52	80.03	0.79
Static Fusion Method	CMHA [4]	83.97	84.01	0.84
	CLA-TCN [5]	81.52	81.03	0.81
	PCAF [6]	86.69	86.62	0.86
Ours		96.79	96.80	0.97

Table 2 Comparative Results with Other Models

Setting	Accuracy(%)	Recall(%)	F1-score	
$(f_e f_p, f_p f_e, f_i f_v, f_v f_i)$	93.26	93.53	0.93	
$(f_e f_i, f_i f_e, f_p f_v, f_v f_p)$	94.68	94.42	0.95	
Ours		96.79	96.80	0.97

Table 3 Verification of the Effectiveness of Granger Causality and Dynamic Multimodal Attention Fusion

Exp.	P	V	G	D	C	Accuracy(%)	Recall(%)	F1 Score	Params(MB)
1	✓	-	-	-	-	63.57	63.54	0.64	0.467
2	-	✓	-	-	-	77.03	76.96	0.77	0.450
3	✓	✓	-	-	-	80.39	80.27	0.80	0.79
4	✓	✓	✓	-	-	88.82	88.80	0.89	1.675
5	✓	✓	✓	-	✓	95.44	95.45	0.95	10.726
6	✓	✓	✓	✓	-	96.79	96.80	0.97	4.646

Table 4 Comparison with other models on HCI dataset

Method	Accuracy(%)		F1-score	
	Arousal	Valence	Arousal	Valence
LSTM [1]	66.21	67.90	0.66	0.67
CNN-LSTM [2]	70.92	71.90	0.69	0.70
Transformer [3]	79.83	80.18	0.78	0.79
CMHA [4]	80.44	81.46	0.79	0.80
CLA-TCN [5]	79.17	79.78	0.78	0.78
PCAF [6]	85.34	86.62	0.84	0.85
Ours	93.73	94.35	0.91	0.93

References

- [1] Alhagry S, Fahmy AA, El-Khoribi RA. Emotion recognition based on EEG using LSTM recurrent neural network. *International Journal of Advanced Computer Science and Applications*. 2017;8(10).
- [2] Chakravarthi B, Ng SC, Ezilarasan M, Leung MF. EEG-based emotion recognition using hybrid CNN and LSTM classification. *Frontiers in Computational Neuroscience*. 2022;16:1019776.
- [3] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is All You Need. *arXiv*. 2017;.
- [4] Zheng J, Zhang S, Wang Z, Wang X, Zeng Z. Multi-channel weight-sharing autoencoder based on cascade multi-head attention for multimodal emotion recognition. *IEEE Transactions on Multimedia*. 2022;25:2213–2225.
- [5] Ye H, Zhou Y, Tao X. A Method of Multimodal Emotion Recognition in Video Learning Based on Knowledge Enhancement. *Computer Systems Science & Engineering*. 2023;47(2).
- [6] Zhang Y, Tao X, Ai H, Chen T, Gan Y. Multimodal Emotion Recognition by Fusing Video Semantic in MOOC Learning Scenarios. *arXiv preprint arXiv:240407484*. 2024;.