

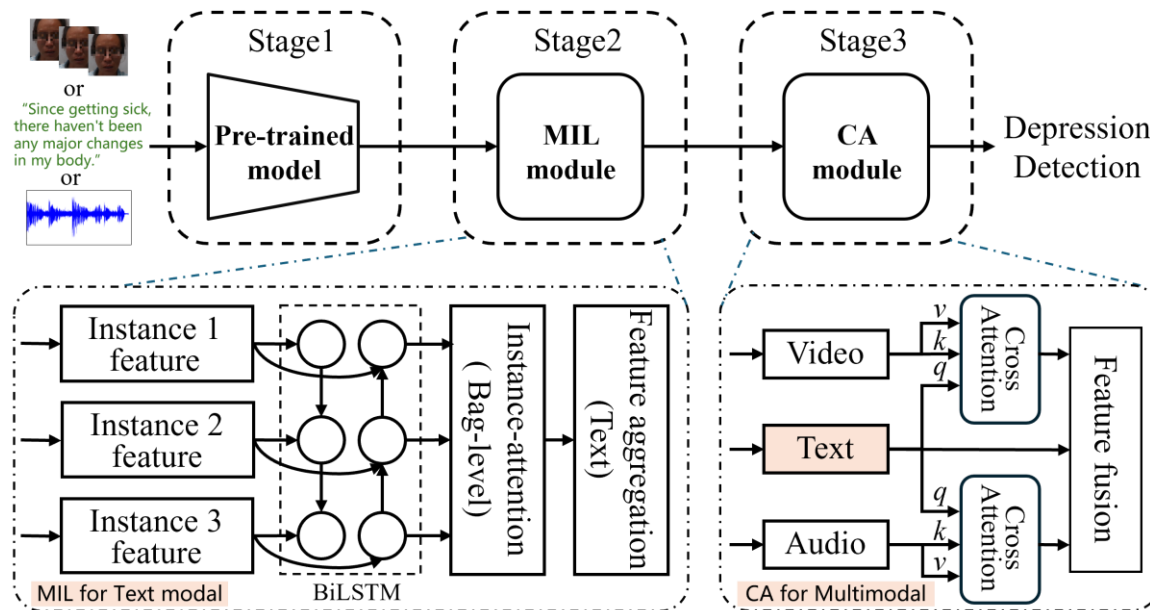
# MILCAnet: A dominant feature attention framework for enhanced multimodal data analysis in depression detection

**Qian RONG, Cheng SONG, Yaru ZHANG, Yao YU,  
Ping LIANG, Chuan PANG, Jie YU, Shuai DING**

Frontiers of Computer Science, DOI: [10.1007/s11704-025-41443-9](https://doi.org/10.1007/s11704-025-41443-9)

# Problems & Ideas

- Problems of key feature extraction and multimodal fusion:
  - There is redundancy in the feature extraction of long-temporal, and the extraction of key features is unclear.
  - The existing feature fusion methods for multimodal data are interfered with by the features of inferior modalities.
- Ideas: a depression detection model combining key feature extraction for long-temporal data and feature fusion dominated by dominant modalities.



An overview of the proposed method (i.e., MILCAnet).

# Main Contributions

- Contributions:
  - We propose an intra-modal dominant feature attention fusion module (MIL module) that effectively suppresses redundant information by enhancing the attention to intra-modal key temporal data;
  - We propose an inter-modal dominant feature attention fusion module (CA module) that enhances the multimodal data feature fusion performance and task suitability through inter-modal dominant modal attention;
  - The proposed method is validated in a non-contact depression detection task based on multimodal data fusion, which improves the performance of depression detection.

Table2 Comparison of the MILCAnet with SOTA methods

Method	Modal	Accuracy	Precision	Recall	F1_score
TCNN	A. T	0.7812	0.7836	0.7812	0.7941
MULT	A. V. T	0.8281	0.8325	0.8281	0.8285
HEALNet	A. V. T	0.7968	0.8013	0.7969	0.7973
MILCAnet	A. T	0.8125	0.8209	0.8125	0.8091
MILCAnet	A. V. T	<b>0.8437</b>	<b>0.8460</b>	<b>0.8438</b>	<b>0.8441</b>