史伟, 金宇挺. 情感分析与多模态方法在 Kickstarter 众筹项目中的应用[J]. 智能计算机与应用,2026,16(6):200-205. DOI: 10.20169/j.issn.2095-2163.24102002

情感分析与多模态方法在 Kickstarter 众筹项目中的应用

史 伟[1,2], 金宇挺[2]

(1 浙江海洋大学 经济与管理学院, 浙江 舟山 316000; 2 湖州师范学院 信息工程学院, 浙江 湖州 313000)

摘 要: 在线众筹平台已成为创新项目和创意企业家获取资金的重要渠道。本研究以 Kickstarter 平台为对象,探讨项目描述、评论互动及公众期待程度对众筹成功率的影响。通过结合 BERT 模型与跨模态分析方法,深入解析项目的视觉和文本描述,以及评论中的文字与表情表达。研究重点评估了项目描述与类型的契合度、发起者对评论的情感响应程度,以及情感倾向对资金筹集效果的影响。结果表明,高质量的项目描述和精准的情感表达对吸引支持者和提高融资成功率至关重要。同时,描述中的情感一致性与发起者对评论的适当回应,对众筹项目的成功具有显著的促进作用。

关键词: 多模态; 情感分析; 互联网众筹; BERT

中图分类号: TP391.1 **文献标志码**: A **文章编号**: 2095-2163(2026)06-0200-06

The application of sentiment analysis and multi-modal approach to Kickstarter crowdfunding projects

SHI Wei[1,2], JIN Yuting[2]

(1 School of Economics and Management, Zhejiang Ocean University, Zhoushan 316000, Zhejiang, China;
2 College of Information Engineering, Huzhou Normal University, Huzhou 313000, Zhejiang, China)

Abstract: Online crowdfunding platforms have become essential channels for innovative projects and creative entrepreneurs to secure funding. This study focuses on the Kickstarter platform to explore the impact of project descriptions, comment interactions, and public expectation levels on crowdfunding success rates. By integrating the BERT model with cross-modal analysis methods, the paper delves into the visual and textual descriptions of projects, as well as the textual and emotive expressions in comments. The study evaluates the alignment between project descriptions and types, the emotional responsiveness of creators to comments, and the influence of emotional tendencies on fundraising outcomes. Results indicate that high-quality project descriptions and precise emotional expressions are crucial for attracting supporters and enhancing fundraising success. Moreover, emotional consistency in descriptions and appropriate responses to comments significantly contribute to the success of crowdfunding projects.

Key words: multimodality; sentiment analysis; Internet crowdfunding; BERT

0 引 言

近年来,众筹从新兴融资工具发展为促进创新的重要平台。早期研究集中于众筹如何通过集合小额资金支持企业家和创新者实现创意,随后扩展至其在激进与增量创新、开放性概念及平台创新中的作用。

本研究聚焦于众筹中的情感动态,分析项目叙事和评论中的情感元素对投资决策及众筹表现的影响。尽管社会资本的重要性已被 Cai 等学者[1]加以强调,但对情感因素的研究仍然不足。通过分析 Kickstarter 平台上的项目叙事,本研究探讨多模态手段(文本、图像、表情)如何与投资者产生共鸣并影响资助决策。研究旨在识别关键情感成分及其与投资者情感反应的互动,探索促进情感共鸣与资金支持的叙事策略[2]。研究结果为项目发起者提供优化项目呈现的建议,吸引并维系投资者兴趣;同时为平台管理者提供功能设计优化方向,提升情感互动与筹资效率[3]。此外,本研究结合跨模态 BERT 模型与多模态情感分析方法,识别并解析众筹项目

基金项目: 国家社会科学基金(20BXW013); 浙江省省属高校基本科研业务费专项资金(2023K002)。

作者简介: 史 伟(1981—),男,博士,教授,硕士生导师。主要研究方向:商务智能与情感计算。

通信作者: 金宇挺(1998—),男,硕士研究生。主要研究方向:众筹与情感分析。Email:452724463@qq.com。

收稿日期: 2024-10-20

中的关键情感指标，旨在加深对众筹平台情感交流模式的理解，为项目设计与互动策略优化提供实用建议，从而提升众筹吸引力与成功率，为学术研究与实践提供新视角与专业洞见。

在已有研究中，多模态情感分析通过整合文本、视觉和语音等多种资源来深入理解人类情感。这种组合特征对于提供平行信息和情感行为消歧义至关重要。先前的多模态情感分析（MSA）研究集中于直接相关的情感变化因素，通过融合不同模态来捕捉多模态内容的情感[3]。

多模态研究通过早期、及后期融合在输入和预测阶段整合数据。早期融合整合各模态特征，但可能限制模态内交互并导致同步问题[4]。后期融合将各模态视为独立模型，通过投票或平均法整合结果，忽略了模态间交互[5]。

为克服这些问题，研究转向获取联合表示，如 Graph-MFN 利用图形架构进行记忆使用，MCTN 通过循环转换获得联合表示。这些模型因使用 RNN 而面临长期依赖问题，导致信息丢失[6]。NLP 领域采用基于自注意力的 Transformer 模型解决此问题[7]。在多模态领域，跨模态 Transformer 被用于提升性能，如 MTMM-ES 在多任务学习框架中使用上下文跨模态注意力预测情感和情绪[8]。

在互联网众筹领域，情感传达对吸引投资者和促进资金筹集极为关键。本研究提出假设，探讨众筹项目中图片与文字情感表达的一致性，并利用跨模态 BERT 模型分析其对项目成功的影响。认知一致性理论认为，个体倾向于接受和处理一致性信息，不一致信息可能引发心理不适[9]。在众筹项目中，视觉和文本内容的情感不一致可能引起潜在投资者疑虑，影响其投资决策[10]。研究中假设：

（1）情感一致性的正面影响（H1a）：图片与文字情感表达的一致性可增强众筹项目的吸引力，提高成功概率。例如，积极形象的图片配合乐观描述更能激发投资者共鸣。

（2）情感不一致性的负面影响（H1b）：图片与文字情感表达不一致可能削弱项目吸引力，降低众筹成功率。

本研究探讨众筹项目中发起者对公众评论的情感响应与项目成功之间的关系，以及时间跨度的影响。基于社会交换理论，假设积极互动可增强信任与支持。通过情感分析技术和 BERT 模型量化发起者情感倾向，并评估其对项目支持动态的影响。本研究旨在揭示发起者情感响应的频率与性质如何促进项目成功，深化发起者与支持者互动在众筹中的作用，并为项目管理提供策略性见解[11]。

1　研究背景和数据

1.1　研究背景

Kickstarter 自 2009 年成立以来，已经建立了一个庞大的社区，拥有超过 960 万的资助者，并在 2021 年 11 月之前帮助项目筹集了超过 62 亿美元的资金。Kickstarter 以其公益性质和对创意项目的支持而闻名，不仅为艺术和创意项目提供资金支持，还特别关注社会公益项目和食品及饮料项目的发展。

1.2　数据来源

本研究基于全球知名众筹平台 Kickstarter，选取其于 2022 年 6～12 月区间的食品众筹项目数据。通过开发 Python 爬虫，合法获取项目文本描述和图片资源，并提取基本信息、描述、更新及用户评论。数据清洗与预处理后，以结构化格式存储，用于跨模态情感分析。整个过程严格遵守 Kickstarter 爬虫政策及数据隐私法规，确保合法性与可靠性。

1.3　情感分析方法

本研究利用跨模态 BERT 模型分析众筹项目中图片与文字的情感一致性，并通过情感分析量化发起者对公众评论的情感响应，同时评估众筹前期公众期待对项目成功的影响[12]。通过数据收集、预处理、情感分析、一致性评估、响应量化及相关性分析，构建回归模型验证情感一致性、发起者响应与公众期待对项目成功的影响[13]。此方法将揭示跨模态情感表达在众筹中的作用，为项目管理提供策略性见解如图 1 所示。

2　研究方法

本研究提出了一种跨模态 BERT 模型，能够整合文本和视觉模态信息，以对预训练的 BERT 模型进行微调。在接下来的小节中，第 2.1 节描述了文本与视觉模态的提取。第 2.2 节描述了模态的交互与对齐。第 2.3 节描述了情感分类。

2.1　模态提取

在多模态情感分析流程中，模态特征提取是至关重要的第一步，旨在从文本和视觉模态中提取高质量的特征表示，为后续模块提供基础支持[14]。文本模态的输入为一个序列 $T = \{t_1, t_2, \cdots, t_n\}$，其中 t_i 表示第 i 个词或标记，n 为序列长度。通过预训练的 BERT 模型对文本进行编码，生成上下文感知的嵌

入特征 X_t，其计算公式为：

$$X_t = \mathrm{BERT}(T),\ X_t \in \mathbb{R}^{n\times d} \tag{1}$$

其中，d 表示隐藏层的维度，X_t 能有效捕获文本中的语义信息和上下文依赖。对于视觉模态，输入为图像 $I \in \mathbb{R}^{H\times W\times C}$，这里 H、W、C 分别为图像的高度、宽度和通道数。

通过 ResNet 模型对图像进行特征提取，生成视觉嵌入特征 X_v，其计算公式为：

$$X_v = \mathrm{ResNet}(I),\ X_v \in \mathbb{R}^{m\times d} \tag{2}$$

其中，m 表示分块数量或帧的数量。ResNet 利用深层残差网络结构，有效提取图像的空间和语义信息。文本模态特征 X_t 提供了语言的情感表达，而视觉模态特征 X_v 提供了场景和表情等视觉线索，两者的结合为后续的多模态交互和分类奠定了坚实的基础。

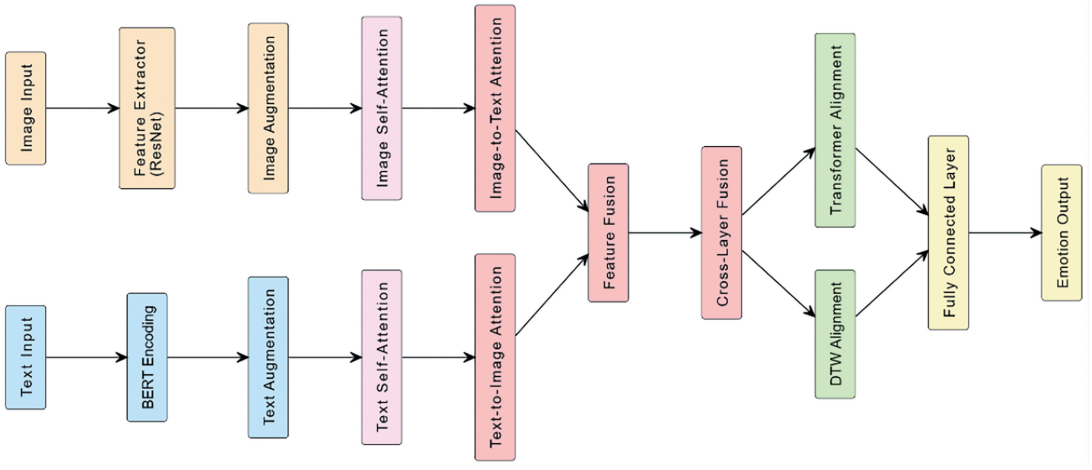


图 1 众筹项目多模态情感分析流程图

Fig. 1 Flowchart of multi-modal emotion analysis for crowdfunding projects

2.2 模态交互与对齐

为了更好地利用文本模态和视觉模态的互补信息，多模态特征处理与融合模块包含 4 个关键步骤：数据增强、自注意力优化、跨模态交互以及模态对齐[15]。这些模块相辅相成，从单模态特征的扩展和优化到跨模态特征的交互与对齐，为多模态情感分析提供了强大的特征基础和融合机制。

2.2.1 数据增强

数据增强模块通过扩展特征空间，提升模型对输入数据扰动的鲁棒性[16]。对于文本模态，采用同义词替换和上下文生成等方法对特征进行增强，生成的文本特征表示为：

$$X_t^{\mathrm{aug}} = X_t + \mathrm{Aug}(T) \tag{3}$$

其中，$\mathrm{Aug}(T)$ 表示文本增强函数。对于视觉模态，利用随机裁剪、旋转、多视角生成等技术扩展图像特征，增强后的视觉特征表示为：

$$X_v^{\mathrm{aug}} = X_v + \mathrm{Aug}(I) \tag{4}$$

通过数据增强，模型能够生成更加丰富的特征表示，减少模态间的不均衡性，并提高后续模块的输入质量。

2.2.2 自注意力优化

自注意力优化模块通过捕捉模态内部的全局依赖关系，有效提升单模态特征的表达能力[17]。对于文本模态，利用多头自注意力机制，计算 Query($\boldsymbol{Q}_t$)、Key($\boldsymbol{K}_t$) 和 Value($\boldsymbol{V}_t$) 矩阵，具体公式如下：

$$\boldsymbol{Q}_t = X_t^{\mathrm{aug}}\boldsymbol{W}_Q,\ \boldsymbol{K}_t = X_t^{\mathrm{aug}}\boldsymbol{W}_k,\ \boldsymbol{V}_t = X_t^{\mathrm{aug}}\boldsymbol{W}_V \tag{5}$$

随后，计算注意力权重：

$$\alpha_t = \mathrm{Softmax}\left(\frac{\boldsymbol{Q}_t\boldsymbol{K}_t^{\mathrm{T}}}{\sqrt{d}}\right) \tag{6}$$

并生成优化后的文本特征：

$$\boldsymbol{X}'_t = \alpha_t\boldsymbol{V}_t \tag{7}$$

视觉模态以相同方式处理，得到优化后的视觉特征 $\boldsymbol{X}'_v$。通过自注意力机制，每种模态的特征表达更加精准，为跨模态交互提供了高质量输入[18]。

2.2.3 跨模态交互

跨模态交互模块通过双向注意力机制实现文本和视觉模态特征的融合。首先，计算文本模态对视觉模态的交互特征：

$$\boldsymbol{X}_{t\to v} = \mathrm{Softmax}\left(\frac{\boldsymbol{Q}_t\boldsymbol{K}_v^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V}_v \tag{8}$$

接着，计算视觉模态对文本模态的交互特征：

$$\boldsymbol{X}_{v\to t} = \mathrm{Softmax}\left(\frac{\boldsymbol{Q}_v\boldsymbol{K}_t^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V}_t \tag{9}$$

并生成优化后的文本特征为：

$$\boldsymbol{X}'_t = \alpha_t\boldsymbol{V}_t \tag{10}$$

视觉模态以相同方式处理，得到优化后的视觉特征 $\boldsymbol{X}'_v$。通过自注意力机制，每种模态的特征表达更加精准，为跨模态交互提供了高质量输入。

2.2.4　模态对齐

模态对齐解决了文本模态和视觉模态之间步长和特征维度不一致的问题。动态时间规整（DTW）利用动态规划，寻找模态特征的最优对齐路径。对此表示为：

$$D[i,j] = |X_t[i] - X_v[j]|^2 + \min(D[i-1,j], D[i,j-1], D[i-1,j-1]) \tag{11}$$

此外，Transformer 对齐通过多头注意力机制动态学习模态间的对齐关系，计算注意力权重为：

$$A_{\text{align}} = \text{Softmax}\left(\frac{\boldsymbol{Q}_t \boldsymbol{K}_v^{\mathrm{T}}}{\sqrt{d}}\right) \tag{12}$$

并生成对齐后的特征：

$$\boldsymbol{X}_{t\to v} = A_{\text{align}} \boldsymbol{V}_v \tag{13}$$

通过模态对齐，文本和视觉模态的特征能够在统一时间维度上有效融合，从而进一步提升特征的兼容性和分类的准确性[19]。

2.3　输出情感分类

情感分类模块是多模态情感分析流程的最后一步，目标是将融合后的多模态特征映射到具体的情感类别上。首先，经过全连接层对融合特征 $\boldsymbol{X}_f$ 进行线性变换，结合激活函数生成情感分类的概率分布：

$$\text{Classifier}(\boldsymbol{X}_f) = \sigma(\boldsymbol{W} \cdot \boldsymbol{X}_f + \boldsymbol{b}) \tag{14}$$

最后，通过取概率分布的最大值索引得到情感分类结果：

$$y = \text{argmax}(\text{Classifier}(\boldsymbol{X}_f)) \tag{15}$$

此模块充分利用融合后的多模态特征，完成了对情感类别的精准分类，为多模态情感分析任务提供最终的预测结果。

3　实验方法

本研究评估了跨模态 BERT 在公共多模态情感分析数据集 CMU-MOSI 和 CMU-MOSEI 上的性能。首先，本研究将介绍有关数据集以及说明在实验中会用到的评估指标[20]。

3.1　数据集

本研究的方法在 2 个数据集上进行了评估：CMU 多模态意见级情绪强度数据集（CMU-MOSI）和 CMU 多模态意见情绪与情绪强度数据集（CMU-MOSEI）。CMU-MOSI 数据集包含 93 个来自 YouTube 电影评论的视频片段，涵盖了 2 199 个话语。每个话语的情感标签由 5 名不同的标注人员在-3 到 3 的连续区间内标注，-3 代表强烈的负面情绪，3 代表强烈的正面情绪。研究将数据集划分为训练集（52 个视频，1 284 个话语）、验证集（10 个视频，229 个话语）和测试集（31 个视频，686 个话语）。

3.2　评价指标

本研究的实验评估与之前工作一致，采用相同的指标[21-22]。情感评分分类使用 7 类精度（Acc7），二元情感分类使用 2 类精度（Acc2）和 $F1$ 分数（$F1$），回归任务使用平均绝对误差（MAE）和相关性（Corr）。除 MAE 外，指标值越高表明模型性能越好。实验结果（见表 1）显示，在 CMU-MOSI 数据集上，本研究提出的跨模态 BERT 模型在性能上优于以往基线模型。

表 1 中，最好的结果以粗体突出显示。h 表示越高越好，l 表示越低越好。同时，T 表示文本、A 表示音频、V 表示视觉。

表 1　CMU-MOSI 数据集上的实验结果

Table 1　Experimental results on the CMU-MOSI dataset

Model	Modality	MAE^l	Acc_2^h	Acc_7^h	F_1^h	Corr^h
EF-LSTM	T+A+V	1.023	75.3	33.7	75.2	0.608
LMF	T+A+V	0.912	76.4	32.8	75.7	0.668
MFM	T+A+V	0.951	78.1	36.2	78.1	0.662
RMFN	T+A+V	0.922	78.4	38.3	78.0	0.681
MulT	T+A+V	0.871	83.0	40.0	82.8	0.698
T-BERT	T	0.784	83.2	41.5	83.2	0.774
本文模型	T+V	**0.775**	**83.9**	**41.9**	**83.7**	**0.781**

4　情感分析与结果

4.1　众筹项目中文字与视觉的的情感

研究对 Kickstarter 平台食品众筹项目的情感分析表明，跨模态 BERT 模型在整合文本与视觉信息方面表现优异，尤其在情感一致性高的项目中预测更准确（见表 2）。研究结果显示，情感积极度与项目成功率正相关，其中“Joy”情感成功率最高，“Fear”最低，验证了文字与图片互补及积极情感表达对提高成功率的关键作用。

表 2 中,加粗代表选用模型在对应情感上的比率最接近实际的占比,以及展示了对应情感的成功率。

表 2 跨模态 BERT 在 Kickstarter 上食品众筹项目的表现

Table 2 BERT′s performance of food crowdfunding projects on Kickstarter %

情感分类	BERT 单模态	跨模态 BERT 多模态	实际占比	项目成功率
Joy	61.1	**63.2**	64.5	79.5
Sadness	22.3	**20.6**	21.3	10.3
Fear	5.5	**4.7**	1.8	1.5
Disgust	2.3	**1.3**	1.1	2.1
Anger	8.8	**10.2**	11.3	6.6

4.2 众筹项目中发起者对评论的情感响应

本研究通过单模态方法分析了发起者对评论的响应时间、响应情感与项目成功率的关系。结果显示,响应时间与项目成功率显著负相关,快速响应(0~6 h 内)与较高成功率(87.6%)相关,而响应时间超 48 h 成功率降至 45.6%(见表 3)。研究表明,发起者快速响应评论对提高项目成功率至关重要。

表 3 众筹项目发起者响应时间与情感比例及项目成功率

Table 3 Response time and emotion ratio of crowdfunding project initiator and project success rate %

平均响应时间/h	正面情感比例	负面情感比例	中性情感比例	项目成功率
0~6	78	5	17	87.6
6~12	73	12	15	82.5
12~24	71	15	14	63.5
24~48	65	23	17	58.1
>48	51	28	21	45.6

5 结束语

本研究通过分析 Kickstarter 平台上 2022 年 6~12 月的食品众筹项目,展示了情感分析在预测众筹项目成功率中的应用价值[23]。跨模态 BERT 模型在整合文本和视觉信息方面表现出色,尤其在情感一致性高的项目中显著提升了预测准确性[24]。快速响应潜在支持者的评论与项目成功率正相关,正面情感表达也正向影响资金筹集可能性[25]。未来研究可进一步探索多模态数据在众筹领域的应用,考虑不同文化背景下情感表达的差异性及其对项目成功率的影响。研究者可开发更复杂的模型,结合机器学习和心理学理论,全面理解支持者的投资决策过程。众筹平台可利用这些研究成果,为项目发起者提供基于数据的策略建议,优化项目描述、视觉呈现和互动性,并开发工具帮助发起者实时监控和分析公众反馈,调整沟通策略[26]。本研究为众筹项目成功提供了新的预测工具,为平台优化和策略制定提供了科学依据,情感分析在众筹领域的应用前景广阔[27]。

参考文献

[1] CAI W, POLZIN F, STAM E. Crowdfunding and social capital: A systematic review using a dynamic perspective[J]. Technological Forecasting and Social Change, 2021, 162: 120412.

[2] 刘忠宝, 雷宇飞. 融合多尺度特征的多模态情感分析模型设计与实验[J]. 实验室研究与探索, 实验室研究与探索, 2024, 43(9): 78-83.

[3] PENGNATE S F, RIGGINS F J. The role of emotion in P2P microfinance funding: A sentiment analysis approach [J]. International Journal of Information Management, 2020, 54: 102138.

[4] AKHTAR M S, CHAUHAN D S, GHOSAL D, et al. Multi-task learning for multi-modal emotion recognition and sentiment analysis [J]. arXiv preprint arXiv, 1905.05812, 2019.

[5] KIM K, PARK S. AOBERT: All-modalities-in-One BERT for multimodal sentiment analysis[J]. Information Fusion, 2023, 92: 37-45.

[6] CAMBRIA E, HOWARD N, HSU J, et al. Sentic blending: Scalable multimodal fusion for the continuous interpretation of semantics and sentics[C]//Proceedings of 2013 IEEE Symposium on Computational Intelligence for Human - like Intelligence (CIHLI). Piscataway, NJ: IEEE, 2013: 108-117.

[7] PORIA S, CHATURVEDI I, CAMBRIA E, et al. Convolutional MKL based multimodal emotion recognition and sentiment analysis [C]//Proceedings of 2016 IEEE 16th International Conference on Data Mining (ICDM). Piscataway, NJ: IEEE, 2016: 439-448.

[8] NOJAVANASGHARI B, GOPINATH D, KOUSHIK J, et al. Deep multimodal fusion for persuasiveness prediction[C]//Proceedings of the 18th ACM International Conference on Multimodal Interaction. New York: ACM, 2016: 284-288.

[9] ZADEH A A B, LIANG P P, PORIA S, et al. Multimodal language analysis in the wild: Cmu-mosei dataset and interpretable dynamic fusion graph[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). ACL, 2018: 2236-2246.

[10] BOGUSZ C I, TEIGLAND R, FELLÄNDER A. Sharing economy: Embracing change with caution[J]. Journal of Business Venturing, 2015, 30(5): 710-725.

[11] PASCHEN J. Choose wisely: Crowdfunding through the stages of the startup life cycle[J]. Business Horizons, 2017, 60(2): 179-188.

[12] CECERE G, LE GUEL F, ROCHELANDET F. Crowdfunding and social influence: An empirical investigation [J]. Applied Economics, 2017, 49(57): 5802-5813.

[13] BURTCH G, HONG Y, LIU D. The role of provision points in online crowdfunding [J]. Journal of Management Information Systems, 2018, 35(1): 117-144.

[14] LIAO C, PALVIA P, LIN H N. The roles of habit and web site quality in e-commerce[J]. International Journal of Information Management, 2006, 26(6): 469-483.

[15] ACHEAMPONG F A, NUNOO-MENSAH H, CHEN W.

Transformer models for text-based emotion detection: A review of BERT-based approaches[J]. Artificial Intelligence Review, 2021, 54(8): 5789-5829.

[16] 王有康, 程春玲. 基于跨模态单向加权的多模态情感分析模型[J]. 计算机科学, 2025, 52(7): 226-232.

[17] 林宜山, 左景, 卢树华. 基于音视频特征优化与跨模态 Transformer 的多模态情感分析[J]. 北京航空航天大学学报, 2026, 52(6): 2219-2228.

[18] 刘佳, 宋泓, 陈大鹏, 等. 非语言信息增强和对比学习的多模态情感分析模型[J]. 电子与信息学报, 2024, 46(8): 3372-3381.

[19] 方旭东, 王兴芬. 基于注意力机制和对比学习的多模态情感分析[J]. 北京信息科技大学学报(自然科学版), 2024, 39(4): 63-70.

[20] 王植, 张珏. 基于跨模态情感联合增强网络的多模态情感分析方法[J]. 甘肃科学学报, 2024, 36(4): 146-152.

[21] 丛子涵, 张思佳. 基于多视图关注网络的图文多模态情感分析模型[J]. 现代电子技术, 2024, 47(12): 157-164.

[22] 刘继明, 张培翔, 刘颖, 等. 多模态的情感分析技术综述[J]. 计算机科学与探索, 2021, 15(7): 1165-1182.

[23] LIU Wei, QIU Jielin, ZHENG Weilong, et al. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition[J]. IEEE Transactions on Cognitive and Developmental Systems, 2021, 14(2): 715-729.

[24] WAGNER J, ANDRE E, LINGENFELSER F, et al. Exploring fusion methods for multimodal emotion recognition with missing data[J]. IEEE Transactions on Affective Computing, 2011, 2(4): 206-218.

[25] KOROMILAS P, GIANNAKOPOULOS T. Deep multimodal emotion recognition on human speech: A review[J]. Applied Sciences, 2021, 11(17): 7962.

[26] YANG K, XU H, GAO K. Cm-bert: Cross-modalbert for text-audio sentiment analysis[C]//Proceedings of the 28th ACM International Conference on Multimedia. New York: ACM, 2020: 521-528.

[27] ZHAO K, ZHOU L, ZHAO X. Multi-modal emotion expression and online charity crowdfunding success[J]. Decision Support Systems, 2022, 163: 113842.