

# 基于注意力惩罚和自适应学习的场景图增强 联合多模态方面情感分析

黄 辰<sup>1,2,3,4</sup>, 刘会杰<sup>1,2,3,4</sup>, 张 龔<sup>1,2,3,4\*</sup>, 杨 超<sup>1,2,3,4</sup>, 宋建华<sup>2,3,4,5</sup>

- (1. 湖北大学计算机学院, 湖北武汉 430062; 2. 智能感知系统与安全教育部重点实验室, 湖北武汉 430062;  
3. 大数据智能分析与行业应用湖北省重点实验室, 湖北武汉 430062;  
4. 湖北省高校人文社科重点研究基地-绩效评价信息管理研究中心, 湖北武汉 430062;  
5. 湖北大学网络空间安全学院, 湖北武汉 430062)

**摘要:** 联合多模态方面情感分析(Joint Multimodal Aspect Sentiment Analysis, JMASA)作为细粒度情感分析领域的重要研究方向,旨在从图像-文本对中联合识别具体的方面术语及其对应的情感极性,近年来受到了越来越多的关注。尽管该任务在社交媒体分析、产品评论挖掘等领域具有重要应用价值,然而,现有方法主要面临两个方面的挑战:一是在利用预训练语言模型融合多模态信息时,模型常对部分无关的视觉或文本标记产生注意力过度信任问题,即分配了不合理的高注意力分数,干扰了对关键情感线索的捕捉;二是现有方法难以显式地建模图像内部对象间的复杂关系,也缺乏有效机制来挖掘图像与文本之间在对象级别的深度语义交互与依赖。为了解决上述问题,本文提出了一种基于注意力惩罚和自适应学习的场景图增强联合多模态方面情感分析方法(Attention Penalty and Adaptive Learning Scene Graph, APALSG),并利用场景图生成(Scene Graph Generation, SGG)来增强联合多模态方面情感分析。具体来说,该方法主要通过专门设计的注意力惩罚策略对超过预设阈值的高注意力分数进行惩罚性衰减,并将衰减的注意力值重新分配给其上下文窗口内的相邻标记。该策略动态调整了模型的注意力分布,有效缓解了对无关信息的过度关注,从而提取出更精准的关键对象特征。此外,设计场景图建模模块,结合图卷积网络(Graph Convolutional Network, GCN)在该场景图上进行消息传播与聚合,获得包含丰富对象间关系上下文信息的视觉表示。最后,还设计了一种自适应学习策略,使模型能够自适应地聚焦于图像-文本对之间与当前方面相关的潜在依赖关系,实现深度的跨模态对齐与融合,并将融合后的多模态特征送入一个分类器,以同时完成方面术语提取和情感分类的联合预测。为全面验证 APALSG 的有效性,在多个公开可用的基准数据集上的实验结果表明,APALSG 在性能上显著优于现有最先进的方法,验证了其有效性。与现有的 JMASA 模型相比,APALSG 在 Twitter-2015、Twitter-2017 和 MACSA 数据集上表现优异,精确率分别提高了 1.46%、2.18% 和 1.19%。

**关键词:** 联合多模态方面情感分析;注意力惩罚;场景图生成;图卷积网络;预训练语言模型

**基金项目:** 湖北省重大攻关项目(JD)(No.2023BAA018);湖北省科技计划重大科技专项(No.2024BAA008);武汉市知识创新专项(No.202311901251001)

**中图分类号:** TP391;TP399

**文献标识码:** A

**文章编号:** 0372-2112(2026)02-0851-11

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20250543

## Attention Penalty and Adaptive Learning Scene Graph for Joint Multimodal Aspect-Based Sentiment Analysis

HUANG Chen<sup>1,2,3,4</sup>, LIU Huijie<sup>1,2,3,4</sup>, ZHANG Yan<sup>1,2,3,4\*</sup>, YANG Chao<sup>1,2,3,4</sup>, SONG Jianhua<sup>2,3,4,5</sup>

(1. School of Computer Science and Information Engineering, Hubei University, Wuhan, Hubei 430062, China;

2. Key Laboratory of Intelligent Sensing System and Security, Ministry of Education, Wuhan, Hubei 430062, China;

3. Hubei Key Laboratory of Big Data Intelligent Analysis and Application, Wuhan, Hubei 430062, China;

4. Hubei Province Project of Key Research Institute of Humanities and Social Sciences at Universities, Wuhan, Hubei 430062, China;

5. School of Cyber Science and Technology, Hubei University, Wuhan, Hubei 430062, China)

**Abstract:** Joint multimodal aspect-based sentiment analysis (JMASA), a crucial research direction in fine-grained sentiment analysis, aims to jointly identify specific aspect terms and their corresponding sentiment polarities from image-text pairs, and has garnered increasing attention in recent years. Although this task holds significant application value in areas such as social media analysis and product review mining, existing methods primarily face two challenges: First, when le-

veraging pre-trained language models to fuse multimodal information, models often exhibit excessive trust in certain irrelevant visual or textual tokens, allocating unnaturally high attention scores, which interferes with capturing key emotional cues; Second, existing methods struggle to explicitly model the complex relationships between objects within an image and lack an effective mechanism to mine deep semantic interactions and dependencies at the object level between images and text. To address these issues, this study proposes a novel scene graph-enhanced method for joint multimodal aspect-based sentiment analysis based on attention penalty and adaptive learning, named APALSG, which utilizes scene graph generation (SGG) to enhance the analysis. Specifically, the method primarily employs a specially designed attention penalty strategy to penalize and attenuate high attention scores that exceed a predefined threshold, redistributing the attenuated attention values to neighboring tokens within their contextual window. This strategy dynamically adjusts the model's attention distribution, effectively mitigating the over-focus on irrelevant information, thereby extracting more precise key object features. Furthermore, a scene graph modeling module is designed, which incorporates graph convolutional networks (GCN) to perform message propagation and aggregation on this scene graph, obtaining visual representations enriched with contextual information about inter-object relationships. Finally, an adaptive learning strategy is also designed, enabling the model to adaptively focus on the potential dependencies between the image-text pair relevant to the current aspect, achieving deep cross-modal alignment and fusion. The fused multimodal features are then fed into a classifier to simultaneously perform joint prediction for aspect term extraction and sentiment classification. To comprehensively validate the effectiveness of APALSG, experimental results on multiple publicly available benchmark datasets demonstrate that APALSG significantly outperforms existing state-of-the-art methods, confirming its efficacy. Compared to existing JMASA models, APALSG shows superior performance on the Twitter-2015, Twitter-2017, and MACSA datasets, improving precision by 1.46%, 2.18%, and 1.19% respectively.

**Keywords:** joint multimodal aspect-based sentiment analysis; attention penalty; scene graph generation; graph convolutional network; pre-trained language models

**Foundation Item(s):** Major Project of Hubei Province (JD) (No.2023BAA018); Major Science and Technology Special Project of Hubei Science and Technology Plan (No.2024BAA008); Wuhan Knowledge Innovation Special Project (No.202311901251001)

## 0 引言

多模态方面级情感分析 (Multimodal Aspect-Based Sentiment Analysis, MABSA) 是人工智能领域的重要研究方向,旨在融合文本、图像、音频等多种模态信息,精确识别特定方面的情感极性,从而克服传统单模态方法的局限,更全面地反映情感的复杂表达方式<sup>[1]</sup>。随着社交媒体和在线教育平台的迅速发展, MABSA 在教育反馈、用户行为分析等方面展现出广阔的应用前景和研究价值<sup>[2]</sup>。该任务包括多模态方面术语提取 (Multimodal Aspect Term Extraction, MATE)、多模态方面情感分类 (Multimodal Aspect Sentiment Classification, MASC) 和联合多模态方面情感分析 (Joint Multimodal Aspect Sentiment Analysis, JMASA)<sup>[3]</sup> 三个子任务。早期研究往往直接融合多模态特征进行情感建模<sup>[4]</sup>, 忽视了在融合过程中模型对部分无关信息产生的注意力过度信任问题。近年来,方面级情感分析 (Aspect-Based Sentiment Analysis, ABSA) 的研究逐步转向基于依赖结构的建模方法<sup>[5]</sup>, 强调通过建模句子中词语之间的语义依赖来提高情感分类的准确性。然而,这类方法主要聚焦于文本内部的语义关系,尚未充分挖掘图像与文本之间对象级依赖所蕴含的情

感信息。

尽管先前关于 JMASA 任务的研究在最新方法上取得了显著进展,但现有方法仍然存在两个关键性局限:一是注意力部分过度信任问题。当前多数 JMASA 方法采用预训练语言模型,这些模型在处理多模态数据时由于多模态幻觉而表现出注意力部分过度信任问题<sup>[6]</sup>。如图 1(a) 所示,对于图像和文本数据,语言模型在训练过程中对草地和逗号分配了较高的注意力分数,而草地和逗号本身不具有明确的情感。这种现象限制了模型对关键对象特征的准确捕捉,进而影响 JMASA 任务的性能。二是难以有效捕捉图像-文本对之间的对象交互。由于缺乏交互,模型难以识别图像-文本之间潜在的依赖关系。简单来说,人类在感知图像和阅读文本时,能够自然地建立两者之间的语义关联。

为解决上述问题,本文提出了一种基于注意力惩罚和自适应学习的场景图增强联合多模态方面情感分析方法 (Attention Penalty and Adaptive Learning Scene Graph, APALSG)。首先,利用预训练语言模型分别对文本和图像进行编码,并提出了一种注意力过度信任惩罚策略。如图 1 所示,使用情感检测器对融合后的多模态特征进行情感预测,完成最终的方面级情感分

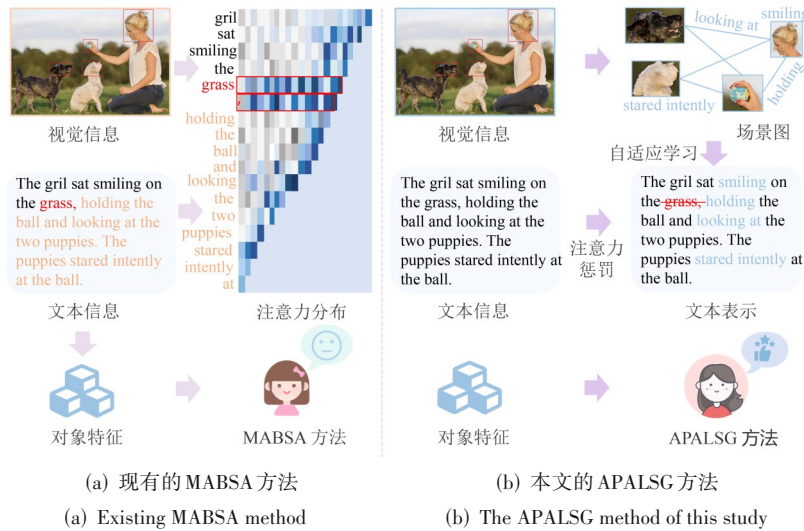


图1 MABSAs方法和APALSs的对比

Figure 1 Intuitive comparison between MABSAs and APALSs

析任务。

针对上述挑战,本文进行深入研究,贡献如下:

(1)提出了一种新的联合多模态方面情感分析模型APALSs,用于MATE、MASC和JMASAs任务。该模型结合图卷积网络对图像内的对象及其关系进行场景图建模,并自适应学习图像-文本对之间潜在的依赖关系。

(2)提出了一种新的注意力惩罚策略,针对语言模型输入中注意力得分过高且不具语义的标记进行惩罚,有效缓解了多模态语言模型中的注意力过度依赖问题。

(3)在三个MABSAs基准数据集上进行广泛的实验,全面分析和验证了APALSs的有效性,并与现有的JMASAs模型进行对比,达到了当前整体的最佳性能。

## 1 相关工作

随着社区多模态数据的快速增长,联合多模态方面情感分析逐渐成为研究热点,并呈现出从粗粒度向细粒度演化的发展趋势<sup>[7]</sup>。粗粒度情感分析主要关注图像-文本对的整体情感判断,而细粒度情感分析则更加侧重于观点性文本的局部信息理解,用户往往通过图文组合来表达更具针对性情感倾向。显然,这一趋势推动了JMASAs方向的深入发展。根据图像视觉特征的特点,以前的方法大致可以分为两类:一类方法通常采用基于Transformer的架构,将整张图像划分为若干视觉块,并与对应文本进行对齐与交互学习<sup>[8]</sup>。这种方法能够建模全局视觉上下文,但对关键区域的识别能力较弱;另一类方法更注重图像中的局部显著区域,通常借助目标检测器如Faster R-CNN或

Mask R-CNN,提取具有高置信度的视觉对象。

近年来,场景图生成(Scene Graph Generation, SGG)在计算机领域引起了广泛关注<sup>[9]</sup>。SGG旨在识别图像内对象之间的关系,并以图结构的形式表示,对象作为节点,关系则以边或关系节点的形式存在。它可以为主下游任务提供有价值的结构信息和语义信息,广泛应用于图像检索<sup>[10]</sup>、视觉常识推理<sup>[11]</sup>和多模态信息提取<sup>[12]</sup>等任务中。为实现更高质量的场景图构建,研究者已提出多种建模方法,例如利用大语言模型理解图像并生成场景图,基于开放词汇构建场景图,面向美学特征的多粒度融合网络。然而,将场景图引入JMASAs任务,以显式建模图像-文本对之间潜在的依赖关系尚未得到充分展开,这就成为了一个突出的研究重点,也为该领域提供了新的见解。

## 2 方法设计与分析

本节详细阐述了APALSs模型。APALSs的整体架构如图2所示,主要有三个核心模块:(1)对象特征提取,提取视觉文本的关键对象特征;(2)场景图建模,提取图像内的对象及其关系,并采用图卷积网络聚合领域信息,获取上下文感知表示;(3)自适应学习,结合场景图自适应学习图像-文本对之间潜在的依赖关系。

### 2.1 模态信息编码

遵循以往的研究工作<sup>[13]</sup>,本文将采用预训练语言模型RoBERTa<sup>[14]</sup>作为文本编码器来提取隐藏的上下文方面表示 $H^t = \{h_1^t, h_2^t, \dots, h_n^t\}$ ,并利用ViT<sup>[15]</sup>作为图像编码器来提取隐藏的上下文视觉表示 $H^v = \{h_1^v, h_2^v, \dots, h_n^v\}$ 。此外,视觉表示需要通过多层感知机

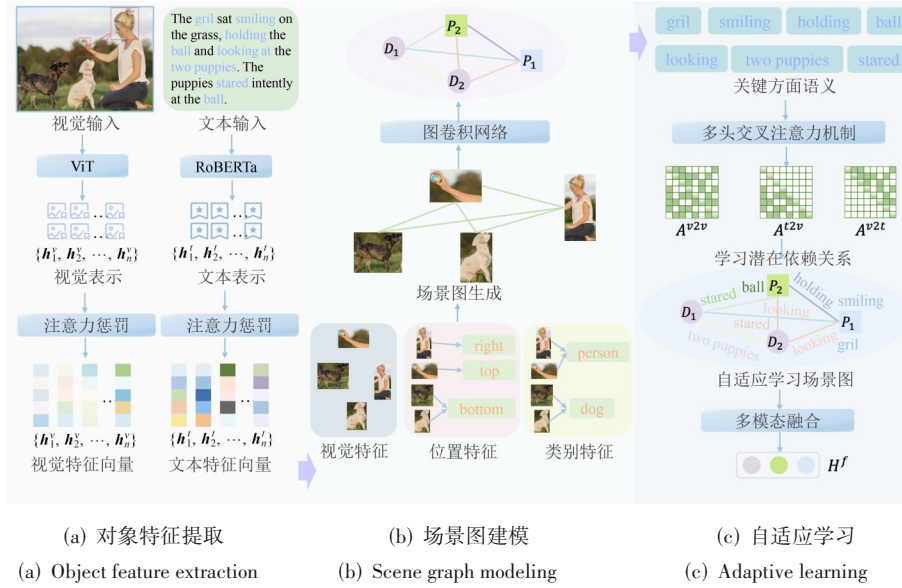


图2 APALSG的整体架构图

Figure 2 The overall architecture diagram of APALSG

(Multi-Layer Perceptron, MLP)实现与方面表示的对齐,  $H^t$ 和 $H^v \in R^{d \times n}$ ,  $d$ 表示隐藏状态维度。

## 2.2 对象特征提取

对象特征提取的目标是利用预训练模型从多模态输入中提取关键对象特征,以便后续场景图建模。对于前一节中得到的上下文视觉表示及上下文方面表示,采用自注意力机制计算注意力分数,计算方式如下所示:

$$[S_t^1, S_t^2, \dots, S_t^n] = \text{attention}(h_t^1, h_t^2, \dots, h_t^n) \quad (1)$$

$$[S_v^1, S_v^2, \dots, S_v^n] = \text{attention}(h_v^1, h_v^2, \dots, h_v^n) \quad (2)$$

其中:  $\text{attention}(\cdot)$ 表示执行自注意力机制来计算自注意力分数;而 $[S_t^1, S_t^2, \dots, S_t^n]$ 和 $[S_v^1, S_v^2, \dots, S_v^n]$ 分别表示图像和方面表示初步计算出的注意力分数。

注意力过度信任惩罚。先前的研究已经证实,预训练语言模型在处理多模态输入时会产生多模态幻觉<sup>[6]</sup>,导致注意力部分过度信任问题。在将预训练语言模型应用于JMASA任务时,该问题同样会复现。为了解决上述问题,本文提出了在注意力计算过程中引入过度信任注意力惩罚策略。具体而言,对于那些表现出过度信任且注意力过度的标记,执行以下操作:

$$S_a^n = (1 - \lambda) \cdot S_b^n, \text{ if } S_b^n > \varphi \quad (3)$$

其中:  $S_b^n$ 表示惩罚前的原始注意力分数;  $\lambda$ 表示注意力惩罚比率值;  $\varphi$ 表示设置的注意力惩罚阈值;  $S_a^n$ 表示执行惩罚操作之后的注意力分数。在惩罚注意力过度的标记后,再通过重新分配其周围标记的注意力值来解决过度信任标记引起的注意力值稀释问题,操作如下:

$$S_a^n = S_b^n + \lambda_a^n \cdot (S_b^n - \varphi), \quad n \sim (p - k, p + k) \quad (4)$$

其中,  $S_b^n$ 表示重新分配前的原始注意力值;  $k$ 表示设置注意力重新分配窗口值;  $\lambda_a^n$ 表示注意力重新分配权重值;  $S_a^n$ 表示重新分配后的注意力值。

本文遵守先前的研究,采用局部窗口  $k$ 对被惩罚标记的注意力进行重新分配,其设计基于语言序列中语义信息具有局部连续性。在方面级情感分析任务中,方面词及其情感修饰通常集中分布于其邻近上下文,而远距离标记往往与当前方面语义相关性较弱。过程如下:

$$[s_{ta}^1, s_{ta}^2, \dots, s_{ta}^n] = \text{penalty}([S_t^1, S_t^2, \dots, S_t^n]) \quad (5)$$

$$[s_{va}^1, s_{va}^2, \dots, s_{va}^n] = \text{penalty}([S_v^1, S_v^2, \dots, S_v^n]) \quad (6)$$

其中:  $\text{penalty}(\cdot)$ 表示执行的注意力过度信任惩罚操作;  $[s_{ta}^1, s_{ta}^2, \dots, s_{ta}^n]$ 和 $[s_{va}^1, s_{va}^2, \dots, s_{va}^n]$ 分别表示图像和方面表示经过惩罚后获得的注意力向量。随后,通过预训练语言模型的编码操作得到视觉特征向量和文本特征向量,具体方法如下:

$$[h_t^1, h_t^2, \dots, h_t^n] = \text{encoder}([s_{ta}^1, s_{ta}^2, \dots, s_{ta}^n]) \quad (7)$$

$$[h_v^1, h_v^2, \dots, h_v^n] = \text{encoder}([s_{va}^1, s_{va}^2, \dots, s_{va}^n]) \quad (8)$$

其中,  $\text{encoder}(\cdot)$ 表示使用模型执行的编码操作。

## 2.3 场景图建模

本文引入SGG技术,提取图像内的对象及其关系,这可以使模型更好地理解图像的高级视觉信息。首先,利用在Visual Genome上预训练的对象检测器Faster R-CNN对每张图像检测一组对象。然后,使用公开的场景图诊断工具包识别每对对象之间的关系。

形式上,对于上一节中得到的图像特征向量,经

过 Faster R-CNN 检测识别之后可以视为对象集表示  $O=\{o_v^i\}$ , 关系集表示为  $R=\{r_v^{ij}\}$ , 其中  $r_v^{ij}$  表示对象  $o_v^i$  和  $o_v^j$  之间的关系。本文定义场景图表示为  $G=\{N,E\}$ , 其中,  $N=O\cup R$  表示包含对象和关系的节点集合,  $E$  表示有向边的集合。值得注意的是, 当关系  $r_v^{ij}$  存在时, 便在构建边集合  $E$  时添加两条有向边  $o_v^i\rightarrow r_v^{ij}$  和  $r_v^{ij}\rightarrow o_v^j$ , 以显式建模对象与关系之间的双向依赖结构。

视觉表示。对于场景图的建模, 本文为每个对象和关系设计了视觉特征、位置特征和类别特征三种特征表示。具体来说, 对象  $c_{o_v^i}^v\in\mathbf{R}^{d_v}$  的视觉特征来自对象检测器区域感兴趣特征的视觉特征, 关系  $c_{r_v^{ij}}^v\in\mathbf{R}^{d_v}$  的视觉特征来自 SGG 模型最终预测层之前的表示。对象和关系的位置特征分别由边界框坐标和联合框坐标转换而来。此外, 对象和关系的类别特征来自对象和关系的类别标签相对应的预训练 Glove 词嵌入<sup>[16]</sup>。然后, 使用一个线性层和 ReLU 激活函数来融合特征:

$$c_{o_v^i}^v = \text{ReLU}\left(\sum_{j\in v,p,g} w_{o_v^i}^j \cdot c_{o_v^i}^j\right) \quad (9)$$

$$c_{r_v^{ij}}^v = \text{ReLU}\left(\sum_{j\in v,p,g} w_{r_v^{ij}}^j \cdot c_{r_v^{ij}}^j\right) \quad (10)$$

其中,  $w_{r_v^{ij}}^v$ ,  $w_{r_v^{ij}}^p$  和  $w_{r_v^{ij}}^g$  是每种特征对应的可训练参数。融合后的特征用于初始化场景图中的节点嵌入。

之后, 采用图卷积网络 (Graph Convolutional Network, GCN) 来聚合领域信息, 获取上下文感知表示<sup>[17]</sup>。第  $L$  层 GCN 中的每个节点根据领域节点的代表进行更新, 具体如下:

$$C_o^L = C_o^{L-1} + \text{ReLU}(A^{ro} \cdot C_r^{L-1} \cdot W_r^L) \quad (11)$$

$$C_r^L = C_r^{L-1} + \text{ReLU}(A^{or} \cdot C_o^{L-1} \cdot W_o^L) \quad (12)$$

其中:  $N_o$  和  $N_r$  分别表示对象和关系的数量;  $A^{ro} \in \mathbf{R}^{N_o \times N_r}$  和  $A^{or} \in \mathbf{R}^{N_r \times N_o}$  分别是对象和关系以及关系到对象的归一化邻接矩阵;  $W^L \in \mathbf{R}^{d \times d}$  是第  $L$  层 GCN 的可训练参数。最后, 可以从第  $L$  层 GCN 的输出中获得对象表示  $H_o = C_o^L$ 。同理可得, 使用上述方法依次获取每个图像内对象的视觉表示, 分别为  $\{H_o^1, H_o^2, \dots, H_o^n\}$ 。

## 2.4 自适应学习

本文优化改进了 2.3 节中的 GCN。首先, 定义注意力矩阵  $\mathbf{AM}$  以及多头注意力函数  $\text{MF}$ , 并据此构建多通道注意力矩阵  $\mathbf{MP}$ , 以匹配多头依赖矩阵  $\mathbf{DM}$ , 具体方式如下:

$$\mathbf{AM}(q, k) = \text{softmax}\left(\frac{qW^Q \cdot (kW^K)^T}{\sqrt{d_k}}\right) \quad (13)$$

$$\mathbf{MF}(q, k) = \sum_{i=1}^N \mathbf{AM}^i(q, k) \quad (14)$$

$$\mathbf{MP} = \mathbf{MF}(H^v, H^v), \mathbf{MF}(H^t, H^v), \mathbf{MF}(H^v, H^t) \quad (15)$$

其中:  $q$  和  $k$  分别代表查询向量和值向量;  $W^Q$  和  $W^K$  是可学习参数;  $d_k$  代表  $K$  的维度;  $\mathbf{MP} \in \mathbf{R}^{N \times M \times L}$  是多通道注意力矩阵;  $L$  是注意力头的数量。

其次, 使用多头依赖矩阵  $\mathbf{DM}$ , 对多通道注意力矩阵  $\mathbf{MP}$  进行掩码操作, 具体实现步骤如下所示:

$$\widehat{A}_{ij} = A_{ij} \cdot M_{ij} \quad (16)$$

$$\mathbf{NM} = \mathbf{DM} \otimes \{A^{v2v}, A^{v2v}, A^{v2t}\} \quad (17)$$

其中:  $A_{ij}$  是原始的邻接矩阵元素;  $M_{ij}$  是掩码矩阵的元素; 掩码矩阵  $\widehat{A}_{ij}$  用于指定哪些依赖关系将被保留。

最后, 将图像隐藏状态  $H^v$ 、文本隐藏状态  $H^t$  以及邻接矩阵  $\mathbf{NM}$  输入到图卷积网络中, 以获得卷积特征  $H^{v2v}$ 、 $H^{v2v}$  和  $H^{v2t}$ , 并且将其取平均值作为最终的多模态融合特征  $H^f$ , 具体如下所示:

$$H^{v2v} = \text{ReLU}(\mathbf{NM}^{v2v} \cdot H^v \cdot W^{v2v}) \quad (18)$$

$$H^{t2v} = \text{ReLU}(\mathbf{NM}^{t2v} \cdot H^t \cdot W^{t2v}) \quad (19)$$

$$H^{v2t} = \text{ReLU}(\mathbf{NM}^{v2t} \cdot H^v \cdot W^{v2t}) \quad (20)$$

$$h^f = \text{mean}(H^{v2v} \cdot H^{t2v} \cdot H^{v2t}) \quad (21)$$

其中: ReLU 表示激活函数; mean 表示对特征求和计算平均值, 作为本文最终的多模态融合特征  $h^f$ 。

## 2.5 情感极性预测

本文使用一个由两层线性神经网络和激活函数组成的解码器进行预测, 并使用损失函数作为训练目标:

$$\widehat{y}_i = \text{softmax}\left(\text{ReLU}(h^f \cdot W_{f1} + a_1) \cdot W_{f2} + a_2\right) \quad (22)$$

$$L_f = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log \widehat{y}_i + (1 - y_i) \cdot \log(1 - \widehat{y}_i) \quad (23)$$

其中:  $W_{f1}$ 、 $W_{f2}$ 、 $a_1$  和  $a_2$  都是可学习参数;  $\widehat{y}_i$  是任务预测结果;  $y_i$  是任务真实标签;  $L_f$  是最终损失。

## 3 实验与分析

在本节中, 本文使用三个公开可用的 MABSA 基准数据集以评估所提出的模型在 MATE、MASC 和 JMASA 任务上的有效性。实验结果旨在回答以下研究问题。

RQ1: 在 MATE、MASC 和 JMASA 任务的表现如何?

RQ2: 主要模块如何影响 APALSG 的性能?

RQ3: 不同的超参数配置如何影响 APALSG 性能?

RQ4: APALSG 在真实数据集上的表现效果如何?

### 3.1 实验设置

数据集本文包括: Twitter-2015<sup>[18]</sup>、Twitter-2017<sup>[18]</sup>、MACSA<sup>[19]</sup>, 相关统计汇总在表 1 中。

评估指标: 本文在 MATE 和 JMASA 任务上使用微

表1 实验数据集的统计数据

Table 1 Statistical data of the experimental dataset

方法	Twitter-2015			Twitter-2017			MACSA		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
积极	928	303	317	1 508	515	493	632	524	532
中性	1 883	670	607	1 638	517	573	1 893	586	613
消极	368	149	113	416	144	168	820	512	240
总数	3 502			2 910			2 460		
单方面	2 159(61.65%)			976(33.54%)			832(33.83%)		
多方面	1 343(38.35%)			1 934(66.46%)			1 628(66.17%)		

观 F1 分数、精确率和召回率来评估 APALSG 的性能。而在 MASC 任务上,遵循先前的研究工作<sup>[20]</sup>,使用准确率 and F1 分数作为评估指标来评估模型性能。

**基线模型:**为了展示 APALSG 模型的有效性,本文将 APALSG 与其他四种类型的 MABSA 方法进行比较。这四种类型方法如下所示。

(1) 文本 ABSA 方法: VistaNet<sup>[21]</sup> 模型使用来自 ResNet 的输入图像的视觉特征。RoBARTa<sup>[22]</sup> 是一个预训练语言模型,具有强大的文本表示能力,并且可以学习两个任意输入之间的对齐进行方面级情感识别。SPAN<sup>[15]</sup> 是一个基于跨度的端到端片段的提取-分类框架,它直接从文本中提取观点并进行分类。D-GCN<sup>[23]</sup> 遵循序列标记范式执行任务,并使用适当的架构对输入词之间的依赖关系进行建模,以实现方面情感分类。

(2) JMASA 方法: UMT-collapse<sup>[24]</sup>、OSCGA-collapse<sup>[25]</sup> 和 RpBERT-collapse<sup>[26]</sup> 都是使用视觉输入来压缩标记。UMT+TomBERT<sup>[18]</sup> 和 OSCGA+TomBERT<sup>[18]</sup> 是两种流水线方法。VLP-MABSA<sup>[27]</sup> 是一种统一的多模态架构,可用于所有预训练和下游任务。CMMT<sup>[28]</sup> 是一种多任务学习框架,用于从图像-文本对中提取方面情感对进行情感分析。AOM<sup>[20]</sup> 是一种面向方面的网络,用于减少视觉和文本内容的噪声。AESAL<sup>[29]</sup> 使用方面增强预训练机制和句法依存图充分学习多模态数据之间的相关性。MIMRL<sup>[30]</sup> 设计了一种基于互信息的解缠多模态表示学习。CAETFN<sup>[31]</sup> 引入一种上下文自适应增强的文本引导融合网络,消除冗余信息,以提升情感分析能力。

(3) MATE 方法: RAN<sup>[32]</sup> 是一种在 MATE 任务上使用对象和文本特征作为输入的新型方法。OSCGA<sup>[33]</sup> 是一种对象感知的神经模型。FITE<sup>[34]</sup> 通过面部表情捕捉视觉情感线索,并取得令人印象深刻的性能。

(4) MASC 方法: ESAFN<sup>[14]</sup> 是一种针对 MASC 的实体敏感注意力和融合网络。TomBERT<sup>[18]</sup> 模拟了视觉和文本表示之间的跨模态交互。CapTrBERT<sup>[35]</sup> 是一种

直接将图像输入到空间中进行转换的双流模型。

**实现细节:**本文在 Windows 系统、Python 版本 3.10、PyTorch 版本 1.12.0 和 NVIDIA RTX 4090 GPU 的环境下实现。在训练过程中,预训练语言模型的注意力头数设置为 12, dropout 设置为 0.1, 批处理大小设置为 4, 隐藏层大小设置为 768。

### 3.2 主要结果分析(RQ1)

在本节中,充分展示了 APALSG 和当前最先进 MABSA 方法在 JMASA、MATE 和 MASC 任务的对比。实验结果表明 APALSG 模型的有效性,实现最优效果。

(1) JMASA 任务表现: JMASA 任务分析结果如表 2 所示。首先, APALSG 优于所有文本 ABSA 方法,这表明利用多模态信息对 ABSA 任务有益。其次,在所有指标上均优于其他的 MABSA 方法。具体而言, APALSG 在 Twitter-2015、Twitter-2017 和 MACSA 数据集上的所有指标均达到了最高水平,精确率分别实现了 1.46%、2.18% 和 1.19% 的显著提升。此外,其他指标(召回率和微观 F1 分数)也显示出相对于 SOTA 模型的改进。这也表明了 APALSG 在 JMASA 任务上的有效性。

(2) MATE 任务表现: 如表 3 所示, APALSG 在 MATE 上的所有指标均出现了提升,实现了整体最优性能。与 AESAL 模型相比,在 Twitter-2015 和 Twitter-2017 数据集上,精确率提高了 2.58% 和 0.86%, 召回率提高了 2.68% 和 0.57%, 微观 F1 分数提高了 3.36% 和 0.56%。这也进一步表明 APALSG 在提取方面术语时的有效性。

(3) MASC 任务表现: 表 4 展示了 APALSG 在 MASC 任务中的优异表现。与 AESAL 模型相比,在 Twitter-2017 数据集上,准确率下降了 0.68%, F1 分数下降了 0.52%。而在 MACSA 数据集上,准确率提升了 0.86%, F1 分数提升了 0.69%。这可能与数据集中图像-文本对的数量有关,因为 Twitter-2017 数据集中的图像-文本对比较少, APALSG 没有充分发挥最大潜力。

### 3.3 消融实验分析(RQ2)

在本节中,为了评估 APALSG 中每个组件的重要性,本研究主要分析了五个变体:(1)排除视觉信息;(2)排除文本信息;(3)无注意力惩罚策略;(4)排除场景图建模;(5)无自适应学习策略。这些修改旨在隔离并评估分析关键组件在 Twitter-2015、Twitter-2017 和 MACSA 数据集上对 JMASA 任务的性能影响。这些变体对模型性能的影响都记录在表 5 中。

APALSG 关键组件在 Twitter-2015、Twitter-2017 和 MACSA 数据集上执行 JMASA 任务的消融实验结果表

表 2 不同方法在 Twitter-2015、Twitter-2017 和 MACSA 数据集上执行 JMASA 任务的结果  
Table 2 Results of performing the JMASA task on the Twitter-2015, Twitter-2017 and MACSA datasets

单位:%  
unit: %

方法	Twitter-2015			Twitter-2017			MACSA		
	精确率	召回率	微观 F1	精确率	召回率	微观 F1	精确率	召回率	微观 F1
VistaNet (AAAI 2019)	54.17	53.30	52.60	58.42	60.10	59.20	52.19	52.40	51.76
RoBERTa* (ARXIV 2019)	61.82	65.30	63.40	65.51	66.83	66.22	53.24	55.62	53.10
SPAN* (ACL 2019)	53.70	53.88	53.73	59.64	61.78	60.57	45.80	46.24	45.72
D-GCN* (COLING 2020)	58.27	58.82	59.40	64.18	64.13	64.10	48.19	50.26	49.12
OSCGA-collapse* (MM 2020)	63.11	63.70	63.24	63.52	63.51	63.55	57.13	57.40	57.43
RpBERT-collapse* (AAAI 2021)	49.22	46.86	48.06	57.04	55.48	56.23	48.72	47.13	48.21
UMT-collapse* (ACL 2020)	61.03	60.42	61.60	60.78	60.03	61.75	56.34	57.13	56.80
UMT+TomBERT* (ACL 2020)	58.43	60.97	59.83	62.37	62.40	62.42	60.35	60.27	61.30
OSCGA+TomBERT* (MM 2020)	61.72	63.42	62.56	63.41	64.02	63.75	63.20	64.24	62.90
VLP-MABSA (ACL 2022)	65.12	68.30	66.57	66.92	69.18	67.98	64.29	64.53	63.84
CMMT (IPM 2022)	64.60	68.72	66.50	67.62	69.42	68.51	63.93	63.87	63.83
AOM (ACL 2023)	67.92	69.30	68.60	68.40	71.02	69.68	65.32	<u>66.07</u>	65.29
AESAL (IJCAI 2024)	<u>68.70</u>	<u>70.40</u>	<u>69.56</u>	<u>69.40</u>	<u>74.82</u>	<u>72.20</u>	<u>66.13</u>	65.90	<u>66.28</u>
MIMRL (TAFFC 2025)	68.27	69.92	69.03	69.11	74.46	71.93	65.82	65.72	65.95
CAETFN(TAFFC 2025)	68.59	70.26	69.40	69.08	74.80	72.05	66.01	66.05	66.10
APALSG(本文)	<b>70.16</b>	<b>71.20</b>	<b>70.39</b>	<b>71.58</b>	<b>76.25</b>	<b>74.36</b>	<b>67.32</b>	<b>67.15</b>	<b>66.50</b>

注:加粗代表效果最优,下划线代表效果次优,\*表示的结果来自文献[20]。

表 3 在 Twitter-2015 和 Twitter-2017 数据集上执行 MATE 任务的结果

单位:%  
unit: %

Table 3 Results of performing the MATE task on the Twitter-2015 and Twitter-2017 datasets

方法	Twitter-2015			Twitter-2017		
	精确率	召回率	微观 F1	精确率	召回率	微观 F1
RAN* (NLPC 2020)	80.52	81.53	81.02	90.70	90.68	90.01
OSCGA* (ARXIV 2021)	81.70	82.11	81.93	90.24	90.70	90.37
FITE (EMNLP 2022)	78.49	73.90	76.52	70.90	68.70	69.24
VLP-MABSA (ACL 2022)	83.60	87.91	85.73	90.74	92.60	91.67
CMMT (IPM 2022)	83.90	88.12	85.92	92.24	93.92	93.11
AOM (ACL 2023)	84.60	87.92	86.27	91.73	92.82	92.28
AESAL (IJCAI 2024)	<u>90.18</u>	<u>90.60</u>	<u>90.36</u>	<u>93.02</u>	<u>96.32</u>	<u>94.70</u>
MIMRL (TAFFC 2025)	89.74	89.79	89.95	92.48	95.70	94.13
CAETFN (TAFFC 2025)	90.05	90.12	90.24	92.80	96.08	94.59
APALSG(本文)	<b>92.76</b>	<b>93.28</b>	<b>93.72</b>	<b>93.88</b>	<b>96.89</b>	<b>95.26</b>

注:加粗代表效果最优,下划线代表效果次优,\*表示的结果来自文献[20]。

明,当移除注意力惩罚时,模型的各项指标均出现下降。其中,微观 F1 分数分别下降 3.91%、2.36% 和 5.21%。这些结果验证了本文设计的注意力惩罚策略对于调整模型注意力分数的有效性,同时有效缓解了过度信任注意力问题。此外,缺少场景图建模和自适应学习也会对模型性能产生较大影响,表明结合场景图自适应学习图像-文本对潜在的依赖关系,可以有效地增强模型在 JMASA 任务上的分析性能。

### 3.4 可解释性分析(RQ3)

为增强注意力惩罚策略的可解释性,本文进一步可视化了引入该策略前后模型在文本序列上的注意

力分布情况。如图 3 所示,在未引入注意力惩罚时,模型对标点符号及背景词分配了较高的注意力权重;而在引入注意力惩罚后,这些语义无关标记的注意力显著降低,与方面相关的情感词及其邻近上下文获得了更高关注度,表明注意力惩罚策略能够对注意力分布进行有效校正。

### 3.5 案例分析(RQ4)

本文在 Twitter-2015 数据集上展现不同方法的预测结果,最终结果如图 4 所示。第一个案例中,尽管 AOM 能够准确检测 boy 的情绪低落,但对 puppy 给出了错误的预测,这是由于缺乏自适应学习机制,未能

表 4 不同方法在 Twitter-2017 和 MACSA 数据集上执行 MASC 任务的结果

单位: %

Table 4 Results of MASC tasks performed by different methods on the Twitter-2017 and MACSA datasets

unit: %

方法	Twitter-2017		MACSA	
	准确率	F1 分数	准确率	F1 分数
ESAFN* (MIPS 2019)	67.82	64.20	78.35	68.73
TomBERT* (TASLP 2020)	70.50	68.00	80.84	71.96
CapTrBERT* (MM 2021)	72.30	70.20	80.30	70.84
VLP-MABSA (ACL 2022)	73.82	71.80	81.21	72.64
CMMT (IPM 2022)	73.80	—	81.70	—
AOM (ACL 2023)	76.42	75.02	82.24	73.46
AESAL (IJCAI 2024)	<b>78.80</b>	<b>75.90</b>	84.42	75.64
MIMRL (TAFFC 2025)	77.48	75.19	84.50	75.70
CAETFN (TAFFC 2025)	77.62	75.30	<u>84.62</u>	<u>75.74</u>
APALSG(本文)	<u>78.12</u>	<u>75.38</u>	<b>85.48</b>	<b>76.43</b>

注:加粗代表效果最优,下划线代表效果次优,\*表示的结果来自文献[20]。

表 5 APALSG 关键组件的消融实验分析

单位: %

Table 5 Analysis of ablation experiments on key components of APALSG

unit: %

方法	Twitter-2015			Twitter-2017			MACSA		
	精确率	召回率	微观 F1	精确率	召回率	微观 F1	精确率	召回率	微观 F1
w/o 视觉信息	67.20	70.35	<b>70.62</b>	70.04	<b>76.38</b>	73.24	66.25	66.42	65.16
w/o 文本信息	66.72	69.16	68.40	67.10	72.21	70.45	62.80	62.93	63.92
w/o 注意力惩罚	65.46	66.54	66.48	66.59	74.30	72.00	63.11	63.04	61.29
w/o 场景图建模	65.30	67.15	67.14	66.10	73.93	72.13	63.20	62.73	62.17
w/o 自适应学习	65.80	70.25	69.64	67.27	75.12	72.82	64.37	65.58	63.58
w/o 关系结构	65.59	67.30	67.72	66.71	74.20	72.39	63.47	62.91	62.38
w/o 随机删除节点	65.62	67.81	68.13	67.04	74.53	72.60	63.82	63.36	62.74
APALSG(本文)	<b>70.16</b>	<b>71.20</b>	70.39	<b>71.58</b>	76.25	<b>74.36</b>	<b>67.32</b>	<b>67.15</b>	<b>66.50</b>

注:加粗代表效果最优。

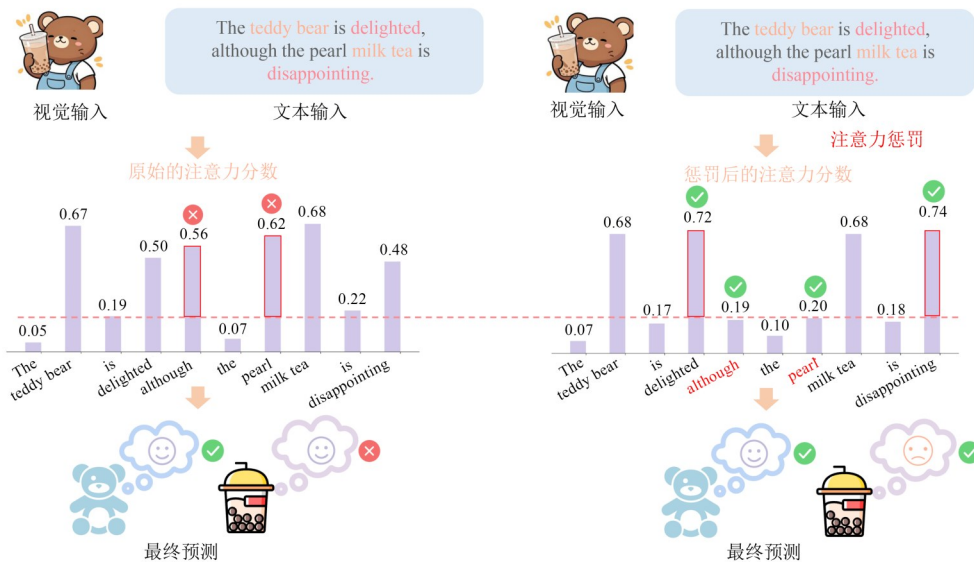


图 3 注意力惩罚的可解释性分析

Figure 3 Interpretability analysis of attention punishment

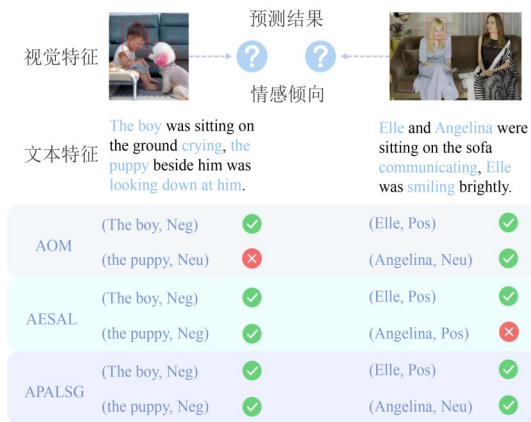


图4 AOM、AESAL和APALSG的结果分析

Figure 4 Analysis of AOM, AESAL and APALSG

充分学习文本中的方面语义信息。AESAL和本文的APALSG均正确预测了所有方面及其对应的情感极性。第二个案例中, AESAL错误地预测 Angelina 的情绪为 Pos, 但正确的情绪为 Neu。这可能是由于缺乏注意力惩罚策略, 导致对相关语义给予过高注意力, 以为和 Elle 一样表现为 Pos。在这两种情况下, APALSG 模型始终表现出优越的性能, 能够正确提取所有方面并对情感进行分类预测, 展示了 APALSG 在 MATE、MASC 和 JMASA 任务中的优越性。

## 4 结论

本文介绍了 APALSG 注意力惩罚和自适应学习的场景图增强联合多模态方面情感分析方法。APALSG 通过专门设计的注意力惩罚策略来调整语言模型的注意力分数, 从而有效缓解了过度信任注意力问题。此外, 结合图卷积网络对图像内对象及其关系进行场景图建模, 并设计了自适应学习策略, 结合场景图自适应学习图像-文本对之间潜在的依赖关系。该方法在 Twitter-2015、Twitter-2017 和 MACSA 数据集上分析 MATE、MASC 和 JMASA 任务, 大量实验评估表明 APALSG 模型有效性。

## 参考文献

- [1] Xu N, Mao W J, Chen G D. Multi-interactive memory network for aspect based multimodal sentiment analysis[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 371-378.
- [2] 张换香, 彭俊杰. 基于方面级情感分析的深度语义挖掘模型[J]. 电子学报, 2024, 52(7): 2307-2319.  
Zhang Huanxiang, Peng Junjie. A deep semantic mining model based on aspect-level sentiment analysis[J]. Acta Electronica Sinica, 2024, 52(7): 2307-2319. (in Chinese)
- [3] Yang X C, Feng S, Wang D L, et al. Few-shot joint multimodal aspect-sentiment analysis based on generative multimodal prompt[PP/OL]. V2. arXiv (2023-05-18) [2025-06-23]. <https://doi.org/10.48550/arXiv.2305.10169>.
- [4] Gao M, Zheng H F, Feng X X, et al. Multimodal fusion using multi-view domains for data heterogeneity in federated learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(16): 16736-16744.
- [5] Chen Q Y, Jin X, Wang Y N, et al. Graph-based unsupervised disentangled representation learning via multimodal large language models[C]//Advances in Neural Information Processing Systems 37. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2024: 103101-103130.
- [6] Huang Q D, Dong X Y, Zhang P, et al. OPERA: Alleviating hallucination in multi-modal large language models via over-trust penalty and retrospection-allocation[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 13418-13427.
- [7] Li Z, Xu B, Zhu C H, et al. CLMLF: A contrastive learning and multi-layer fusion method for multimodal sentiment detection[PP/OL]. V4. arXiv (2022-06-14) [2025-06-23]. <https://doi.org/10.48550/arXiv.2204.05515>.
- [8] Xiao L W, Mao R, Zhao S, et al. Exploring cognitive and aesthetic causality for multimodal aspect-based sentiment analysis[J]. IEEE Transactions on Affective Computing, 2025, 16(4): 3248-3265.
- [9] Xu D F, Zhu Y K, Choy C B, et al. Scene graph generation by iterative message passing[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 3097-3106.
- [10] Yoon S, Kang W Y, Jeon S, et al. Image-to-image retrieval by learning similarity between scene graphs[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 10718-10726.
- [11] Wang Z C, You H X, Li L H, et al. SGEITL: Scene graph enhanced image-text learning for visual commonsense reasoning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(5): 5914-5922.
- [12] Fu Z, Feng J H, Zheng C M, et al. Knowledge-enhanced scene graph generation with multimodal relation alignment (student abstract)[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(11): 12947-12948.
- [13] Wang Z, Liu Y, Yang J N. BERT-based multimodal aspect-level sentiment analysis for social media[C]//Proceedings of the 2022 5th International Conference on Ar-

- tificial Intelligence and Pattern Recognition. New York: ACM, 2023: 187-192.
- [14] Yu J F, Jiang J, Xia R. Entity-sensitive attention and fusion network for entity-level multimodal sentiment classification[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 429-439.
- [15] Hu M H, Peng Y X, Huang Z, et al. Open-domain targeted sentiment analysis via span-based extraction and classification[PP/OL]. V1. arXiv (2019-06-10)[2025-06-23]. <https://doi.org/10.48550/arXiv.1906.03820>.
- [16] Tang K H, Niu Y L, Huang J Q, et al. Unbiased scene graph generation from biased training[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 3713-3722.
- [17] Rakhmatillayevich A M. GLove: Global vectors for word representation[J]. American Journal of Multidisciplinary Bulletin, 2025, 3(5): 359-364.
- [18] Yu J, Jiang J. Adapting BERT for target-oriented multimodal sentiment classification[C]. IJCAI, 2019.
- [19] Yang H, Si Z M, Zhao Y Y, et al. MACSA: A multimodal aspect-category sentiment analysis dataset with multimodal fine-grained aligned annotations[J]. Multimedia Tools and Applications, 2024, 83(34): 81279-81297.
- [20] Zhou R, Guo W Y, Liu X M, et al. AoM: Detecting aspect-oriented information for multimodal aspect-based sentiment analysis[PP/OL]. V1. arXiv (2023-05-31)[2025-06-23]. <https://doi.org/10.48550/arXiv.2306.01004>.
- [21] Truong Q T, Lauw H W. VistaNet: Visual aspect attention network for multimodal sentiment analysis[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 305-312.
- [22] Liu Y H, Ott M, Goyal N, et al. RoBERTa: A robustly optimized BERT pretraining approach[PP/OL]. V1. arXiv (2019-07-26)[2025-06-23]. <https://doi.org/10.48550/arXiv.1907.11692>.
- [23] Chen G M, Tian Y H, Song Y. Joint aspect extraction and sentiment analysis with directional graph convolutional networks[C]//Proceedings of the 28th International Conference on Computational Linguistics. International Committee on Computational Linguistics, 2020: 272-279.
- [24] Yu J F, Jiang J, Yang L, et al. Improving multimodal named entity recognition *via* entity span detection with unified multimodal transformer[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2020: 3342-3352.
- [25] Wu Z W, Zheng C M, Cai Y, et al. Multimodal representation with embedded visual guiding objects for named entity recognition in social media posts[C]//Proceedings of the 28th ACM International Conference on Multimedia. New York: ACM, 2020: 1038-1046.
- [26] Sun L, Wang J Q, Zhang K, et al. RpBERT: A text-image relation propagation-based BERT model for multimodal NER[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(15): 13860-13868.
- [27] Ling Y, Yu J F, Xia R. Vision-language pre-training for multimodal aspect-based sentiment analysis[PP/OL]. V2. arXiv (2022-04-21)[2025-06-23]. <https://doi.org/10.48550/arXiv.2204.07955>.
- [28] Yang L, Na J C, Yu J F. Cross-modal multitask transformer for end-to-end multimodal aspect-based sentiment analysis[J]. Information Processing & Management, 2022, 59(5): 103038.
- [29] Zhu L, Sun H, Gao Q, et al. Joint multimodal aspect sentiment analysis with aspect enhancement and syntactic adaptive learning[C]//Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence. 2024.
- [30] Sun H, Niu Z W, Wang H Y, et al. Multimodal sentiment analysis with mutual information-based disentangled representation learning[J]. IEEE Transactions on Affective Computing, 2025, 16(3): 1606-1617.
- [31] Li J B, Liu R Y, Miao Q G, et al. CAETFN: Context adaptively enhanced text-guided fusion network for multimodal sentiment analysis[J]. IEEE Transactions on Affective Computing, 2025, 16(4): 3122-3138.
- [32] Wu H Q, Cheng S L, Wang J J, et al. Multimodal aspect extraction with region-aware alignment network[M]//Natural language processing and Chinese computing. Cham: Springer International Publishing, 2020: 145-156.
- [33] Yan H, Dai J Q, Ji T, et al. A unified generative framework for aspect-based sentiment analysis[PP/OL]. V1. arXiv (2021-06-08)[2025-06-23]. <https://doi.org/10.48550/arXiv.2106.04300>.
- [34] Yang H, Zhao Y Y, Qin B. Face-sensitive image-to-emotional-text cross-modal translation for multimodal aspect-based sentiment analysis[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2022: 3324-3335.
- [35] Khan Z, Fu Y. Exploiting BERT for multimodal target sentiment classification through input space translation[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM, 2021: 3034-3042.

## 作者简介



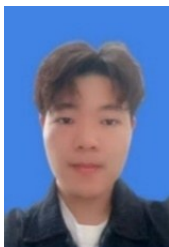
黄 辰 男,1983年8月生,福建龙岩人。  
现为湖北大学计算机学院教授。主要研究方向  
为人工智能、脑机接口。

E-mail: huang@hubu.edu.cn



杨 超 男,1982年9月生,湖北武汉人。  
现为湖北大学计算机学院教授。主要研究方向  
为智能计算、信息安全等。

E-mail: stevenyc@hubu.edu.cn



刘会杰 男,2000年2月生,湖北黄石人。  
现为湖北大学计算机学院硕士研究生。主要研  
究方向为人工智能与脑科学、情感分析。

E-mail: liuhj@stu.hubu.edu.cn



宋建华 女,1973年3月生,湖北襄阳人。  
现为湖北大学网络空间安全学院教授。主要研  
究方向为网络与信息安全。

E-mail: sjhhubu@126.com



张 龔 男,1974年6月生,湖北宜昌人。  
现为湖北大学计算机学院教授。主要研究方向  
为信息安全、大数据分析。中国电子学会会员  
编号:E190197582M。

E-mail: zhangyan@hubu.edu.cn