

# 用于方面级情感分析的多信息融合图卷积网络<sup>①</sup>



高玮军, 张玉莹, 焦成寅

(兰州理工大学 计算机与通信学院, 兰州 730050)

通信作者: 张玉莹, E-mail: [376502886@qq.com](mailto:376502886@qq.com)

**摘要:** 近年来, 方面级情感分析利用图神经网络挖掘依赖句法信息逐渐成为趋势, 但现有方法大多未考虑不同关系类型对内容词的影响, 难以区分关键的关联词. 此外, 多视角信息的相互补充对捕捉情感特征起重要作用, 但在过去的研究中融合机制常被忽视. 为解决这些问题, 提出一种多源信息融合图卷积网络 (multi-source information graph convolutional network, MSI-GCN) 有效捕获和集成三视角信息. 首先, 设计了一个双通道信息提取模块 SSD-GCN (syntax-semantics dual graph convolutional network), 由类型嵌入的句法增强图卷积网络 (TES-GCN) 和语义图卷积网络 (SEM-GCN) 组成. TES-GCN 通过引入类型嵌入层, 使用句法模块学习不同类型的权重来增强句法信息. SEM-GCN 对自注意矩阵进行编码, 捕获语义信息, 并引入正交正则化来增强语义关联. 其次, 嵌入外部知识图表示丰富词汇特征. 最后, 引入局部门控-全局卷积网络, 充分利用视角之间的互补性, 对其进行有效融合. 本文在 4 个公开数据集上对提出的方法进行了评估, 准确率和 *Macro-F1* 值相比于基线模型均有所提升.

**关键词:** 方面级情感分析; 多源信息融合; 图卷积网络; 外部知识; 句法类型嵌入

引用格式: 高玮军, 张玉莹, 焦成寅. 用于方面级情感分析的多信息融合图卷积网络. 计算机系统应用, 2025, 34(8): 14-24. <http://www.c-s-a.org.cn/1003-3254/9909.html>

## Multi-information Fusion GCN for Aspect-based Sentiment Analysis

GAO Wei-Jun, ZHANG Yu-Ying, JIAO Cheng-Yin

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract:** In recent years, aspect-based sentiment analysis utilizing graph neural networks to exploit syntactic information has become increasingly popular. However, existing methods often overlook the impact of different types of relations on content words, making it challenging to distinguish key relational terms. In addition, the mutual supplementation of multi-perspective information plays a crucial role in capturing sentiment features, but integration mechanisms have often been neglected in past research. To address these issues, a multi-source information graph convolutional network (MSI-GCN) is proposed to effectively capture and integrate three perspectives of information. First, a dual-channel information extraction module, syntax-semantics dual graph convolutional network (SSD-GCN), is designed, comprising a type-enhanced syntax graph convolutional network (TES-GCN) and a semantic graph convolutional network (SEM-GCN). TES-GCN enhances syntactic information by incorporating a type embedding layer and learning different weights through the syntactic module. SEM-GCN encodes self-attention matrices to capture semantic information and employs orthogonal regularization to strengthen semantic associations. Second, external knowledge graph embeddings enrich vocabulary features. Finally, a local-global convolutional network is introduced to leverage the complementarity between perspectives for effective integration. The proposed method is evaluated on four public datasets, showing improvements in accuracy and *Macro-F1* scores compared to baseline models.

① 基金项目: 国家自然科学基金 (6172059)

收稿时间: 2024-12-17; 修改时间: 2025-01-24, 2025-02-12; 采用时间: 2025-02-18; csa 在线出版时间: 2025-06-24

CNKI 网络首发时间: 2025-06-25

**Key words:** aspect-based sentiment analysis (ABSA); multi-source information fusion; graph convolutional network (GCN); external knowledge; syntactic type embedding

方面级情感分析 (aspect-based sentiment analysis, ABSA) 是一种细粒度的情感分析任务, 其目标是对具有主观个人情感的文本进行整理以及分析, 识别并理解与特定方面或主题相关的情感倾向. 用户生成的内容通常涉及多个方面, 每个方面可能对应不同的情感倾向. 以 SemEval2014 中的某条评论为例 (如图 1 所示), 若选择“price”作为方面词, 其对应的意见词为“reasonable”, 则情感极性为积极; 而若选择“service”作为方面词, 其意见词为“poor”, 则情感极性为消极; 两个方面拥有两种不同的情感极性, 分别为积极和消极.

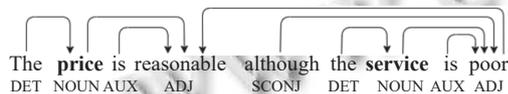


图 1 依存树示例

在 ABSA 中, 一些研究人员明确利用句子的句法结构来建立方面词与其相关情感词之间的连接, 一般通过依存解析获得的依存树来提供全面的语法信息<sup>[1,2]</sup>, 其中的依赖关系表示一般是使用循环神经网络 (RNN) 进行编码<sup>[3]</sup>. 近年来, 图神经网络 (GNN), 例如图卷积网络 (GCN) 和图注意力网络 (GAT), 已被用来学习依赖树的表示<sup>[4,5]</sup>. 虽然利用句法知识对 ABSA 任务有益, 但仍存在缺陷. 基于依赖的方法忽视了某些观点和方面之间的依赖关系, 依存解析结果可能不准确. GCN 的表达能力和图关系权重得分密切相关<sup>[6]</sup>, 而当前方法大多将依赖关系视为相同的权重分数, 忽略了不同关系类型之间的区别, 进而限制了模型对依赖关系的理解. 此外, 在处理复杂和非正式的在线评论时, 由于缺乏语法敏感性可能难以解析, 模型理解困难. 此前研究人员通过将外部知识整合到模型中, 从而利用情感和常识知识来提高模型理解能力<sup>[7,8]</sup>. 通过整合知识图谱, 模型可以利用具有类似嵌入的同义表达来增强对罕见和晦涩单词 (特别是晦涩情感词) 的理解, 从而提高模型准确性. 然而, 知识图谱的集成是一个复杂的过程, 由于图计算随着句子长度的增加而呈指数增长<sup>[9]</sup>, 大多方法无法将知识图谱“平滑”地合并到模型中. 并且对于多个路径的表示, 之前的大多研究<sup>[7-10]</sup>只是对其进行连接操作. Li 等<sup>[4]</sup>虽采用相互 Bi-Affine 变换来融合两个不

同视角, 但仍缺少对不同特征间互补性的关注.

针对以上问题, 本文主要贡献如下.

(1) 提出了一种多源信息融合模型 MSI-GCN, 能够有效提取句法和语义信息, 同时整合外部知识图谱表示, 增强对罕见和晦涩单词的理解以提高情感分类的准确性.

(2) 设计了 SSD-GCN 模块, 包括提取增强句法信息的 TES-GCN 和提取语义信息的 SEM-GCN. TES-GCN 引入了类型嵌入层, 用于获取依赖关系类型的表示, 并利用句法模块对不同类型赋予权重, 克服了传统图模型中的均匀加权问题.

(3) 提出一种多信息的融合模块, 引入局部门控-全局卷积网络, 有效利用 3 个视角表示之间的互补性融合句法信息、语义信息和外部知识.

(4) 在 4 个基准数据集上评估了该方法. 实验结果显示, 与经典基线模型及主流图卷积网络相比, MSI-GCN 在准确率和 Macro-F1 指标上均取得了显著提升. 同时, 通过消融实验进一步验证了各模块的有效性.

## 1 相关工作

本节介绍方面级情感分析研究的整体发展趋势及图神经网络和外部知识在方面级情感分析中的具体应用.

### 1.1 方面级情感分析方法

ABSA 作为细粒度情感分析任务, 通常以分类问题形式进行处理. 在早期的研究中, 大多利用传统的机器学习方法, 需要大量的人力和时间成本. 随着深度学习兴起, 传统方法逐渐被取代, 深度学习逐渐成为主流方法. 循环神经网络 (RNN)、长短期记忆网络 (LSTM) 和门控循环神经网络 (GRU) 等序列模型表现出强大的表示能力. Tang 等<sup>[11]</sup>提出 TD-LSTM 模型, 将句子围绕目标分为两部分, 并利用 Bi-LSTM 捕捉目标左右上下文的表征. 在此基础上, Zhang 等<sup>[12]</sup>提出性能优于 LSTM 的双向 GRU. 同时, 许多研究人员引入注意力机制, 并根据任务需求进行创新<sup>[13,14]</sup>. Chen 等<sup>[13]</sup>为捕捉相隔较远的情感特征, 引入了一个分层注意力网络. Ma 等<sup>[14]</sup>为提高对方面词的理解能力, 提出了交互式注意力网络 (IAN) 交互式地学习上下文和方面目标并分别生成

表示. 一些研究人员也将卷积神经网络 (CNN) 引入方面级情感分析任务中<sup>[15,16]</sup>. 然而, 这些方法在提取方面及其上下文之间的情感依赖特征方面仍有提升空间.

### 1.2 基于图神经网络的方法

近年来, 图神经网络 (GNN) 在 ABSA 上表现出强劲的趋势, 在提取句法信息中发挥了巨大作用<sup>[5,17-21]</sup>. Wang 等<sup>[5]</sup>对传统依赖解析树进行重组和简化, 并基于此提出关系图注意力网络 (R-GAT) 来编码新的树结构, 增强了句法信息. Yuan 等<sup>[19]</sup>提出了句法图注意力网络 (SGAN), 克服基于依赖解析树构建图的局限性. 为解决严重依赖依存树的方法在准确建模方面的挑战, Liang 等<sup>[17]</sup>利用双图注意力网络对每个方面与其上下文的情感关系进行模拟学习, 更精确提取句法信息, 提升了情感分类的准确性. Qiu 等<sup>[18]</sup>提出一种子句级的关系图注意力网络, 将给定的句子分割成从句, 从句级交互获得两个方面句法之间的关系. Zhao 等<sup>[20]</sup>提出一种具有多种权重机制的图卷积网络, 在提取句法信息时设计约束机制以减少与方面术语无关的词语产生的噪声. 张文轩等<sup>[21]</sup>利用依存树与注意力机制构建了情感与句法信息的双图卷积网络更准确地提取特定方面的情感依赖关系. 然而, 这些方法仍将所有句法依赖关系视为均等, 忽视了句法类型的重要性.

### 1.3 基于外部知识的方法

情感和常识知识对于模型理解自然语言至关重要, 结合外部知识可以显著增强 ABSA 模型性能<sup>[10,22-26]</sup>. 为充分利用外部知识辅助, Zhang 等<sup>[10]</sup>将外部知识融入词表示中. Ma 等<sup>[22]</sup>将常识信息融入 LSTM 中, 并将其有效应用于 ABSA 任务. 随着图神经网络的流行, 逐渐开始利用 GCN 或 GAT 来传播外部知识特征. Zhou 等<sup>[26]</sup>通过整合句法和常识知识来增强上下文嵌入, 利用 GCN 进行全面的节点融合. Liang 等<sup>[7]</sup>将 SenticNet 的情感知识分数与句法依存树相结合, 以提高邻接矩阵中情感词的权重, 从而通过方面词及其在 GCN 中的编码来识别关键情感信息. 通过主题图注意力机制, 李阳等<sup>[24]</sup>将方面词与意见词的外部知识进行融合, 以此预测基于情感词典生成的候选标签. Zhong 等<sup>[8]</sup>参考 AR-BERT<sup>[25]</sup>的知识图谱, 将知识嵌入与上下文嵌入结合起来, 使用 GCN 在图节点之间传播和融合信息.

## 2 模型概述

本文提出的模型 (MSI-GCN) 主要由 5 个部分组成: a) 输入层和编码层; b) 双通道图卷积网络, 包括 TES-GCN 和 SEM-GCN; c) 外部知识表示; d) 多信息融合模块; e) 输出层. 具体结构如图 2 所示.

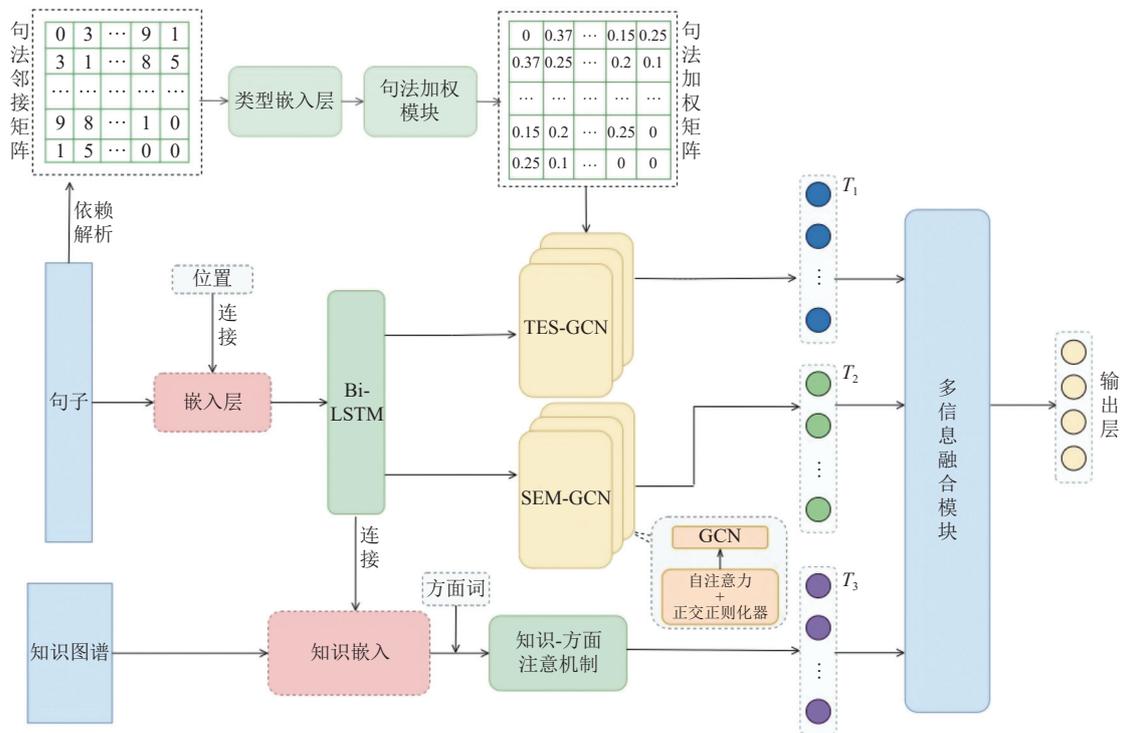


图 2 MSI-GCN 模型结构示意图

## 2.1 输入和编码层

给定句子-方面对 $(S, A)$ , 其中 $S = \{x_1, x_2, \dots, x_{\tau+1}, x_{\tau+2}, \dots, x_n\}$ ,  $A = \{x_{\tau+1}, x_{\tau+2}, \dots, x_n\}$ , 我们利用 GloVe<sup>[27]</sup>表示 $S$ 和 $A$ 中的每个单词. 利用映射矩阵 $E \in \mathbb{R}^{|V| \times d_x}$ 将每个单词 $x_i$ 转换为低维空间内的紧凑向量, 其中,  $|V|$ 表示词汇量大小,  $d_x$ 表示词嵌入的维度. 单词与目标方面位置的接近程度能够反映其重要性, 因此我们将相对位置嵌入整合到词向量中. 对于 BERT<sup>[28]</sup>编码, 输入格式为“[CLS]句子[SEP]方面[SEP]”. 句子 $S$ 和方面 $A$ 的嵌入可以表示为 $O^S = \{o_1^s, o_2^s, \dots, o_{\tau+1}^s, o_{\tau+2}^s, \dots, o_{\tau+m}^s, \dots, o_n^s\}$ 和 $O^A = \{o_1^a, o_2^a, \dots, o_m^a\}$ .

在获得嵌入后, 我们通过使用两个独立的双向长短期记忆网络 (BiLSTM) 将词向量序列转换为上下文级表示. 模型不仅理解了整个句子序列 ( $O^S$ ) 和方面词序列 ( $O^A$ ) 的上下文, 同时也能够初步理解序列中单词的顺序与句子之间的关系. 句子中每个单词和方面单词的隐藏表示, 即 $h_i^s$ 和 $h_j^a$ 可以通过以下等式获得:

$$h_i^s = \left[ \overrightarrow{LSTM}(o_i^s); \overleftarrow{LSTM}(o_i^s) \right], i \in \{1, n\} \quad (1)$$

$$h_j^a = \left[ \overrightarrow{LSTM}(o_j^a); \overleftarrow{LSTM}(o_j^a) \right], j \in \{1, m\} \quad (2)$$

句子 $S$ 的输出表示为 $H^S = \{h_1^s, h_2^s, \dots, h_i^s, \dots, h_n^s\}$ , 方面词 $A$ 的输出表示为 $H^A = \{h_1^a, h_2^a, \dots, h_j^a, \dots, h_m^a\}$ .  $H^S$ 用作第 2.2 节中的初始特征矩阵,  $H^A$ 用于第 2.3 节中知识方面注意力机制的阐述.

## 2.2 双通道图卷积网络构建

SSD-GCN 由基于类型嵌入句法增强的 TES-GCN 和基于自注意力机制的 SEM-GCN 组成. 最终, 输出句法表示  $T_1$  和语义表示  $T_2$ .

### 2.2.1 TES-GCN

TES-GCN 的核心是构建增强句法依赖图, 具体做法是通过依赖类型嵌入并赋予句法图权重来实现.

首先根据单词之间的依存关系建立了句法无向图, 其中每个词对应一个节点, 边表示它们的依赖关系. 为每个节点添加自循环, 将其类型标记为“self”, 并构建了从 1 开始编号的依赖类型索引. 句法邻接矩阵 $A_{ij}^{\text{tes}}$ 表示如下:

$$A_{ij}^{\text{tes}} = \begin{cases} T, & \text{if } w_i \text{ and } w_j \text{ have dependency relation} \\ 0, & \text{others} \end{cases} \quad (3)$$

经过类型嵌入层后, 生成句法类型嵌入  $H_{\text{type}} \in \mathbb{R}^{d_t \times n \times n}$ . 其中,  $d_t$  是类型嵌入的维数. 句法加权矩阵由

句法加权模块产生, 公式为:

$$A^{\text{tes}} = \sigma(W_s H_{\text{type}} + b_s) \quad (4)$$

将句法加权矩阵输入句法 GCN 层中, 通过聚合邻接矩阵的邻节点信息来更新节点表示, 权重表示邻接的重要性. 句法 GCN 层将句法相关词整合到词表示中, 词在 GCN 层中的隐藏状态表示为:

$$H_{\text{tes}}^{(l+1)} = \text{ReLU}(L^{\text{tes}} H_{\text{tes}}^{(l)} W_{\text{tes}}^{(l)}) \quad (5)$$

其中,  $W_{\text{tes}}^{(l)}$  是 GCN 第  $l$  层学习的参数,  $L^{\text{tes}}$  是拉普拉斯矩阵, 定义为:

$$L^{\text{tes}} = \tilde{D}^{-\frac{1}{2}} \tilde{A}^{\text{tes}} \tilde{D}^{\frac{1}{2}} \quad (6)$$

其中,  $\tilde{D}$  是对角矩阵,  $\tilde{D}_{ii} = \sum_j A_{ij}^{\text{tes}}$ ,  $A^{\text{tes}} = \tilde{A}^{\text{tes}} + I$ , 其中  $I$  是单位矩阵. 初始隐藏状态  $H_{(0)}^{\text{tes}}$  设置为  $H_c$ . 第  $l$  层 GCN 从  $l$  跳邻居处捕获信息.

TES-GCN 的最终输出表示为  $H_{\text{tes}}^{(l+1)}$ , 其中对非方面词使用零向量屏蔽机制, 仅保留隐藏的方面词表示. MSA-GCN 的第 1 个视角表示  $T_1$  通过平均池化获得, 公式如下:

$$H_{\text{tes-zero}}^{l+1} = [0, h_{\tau+1}^{(l+1)\text{tes}}, h_{\tau+2}^{(l+1)\text{tes}}, \dots, h_{\tau+m}^{(l+1)\text{tes}}] \quad (7)$$

$$T_1 = \text{AvgPooling}(h_{\tau+1}^{(l+1)\text{tes}}, h_{\tau+2}^{(l+1)\text{tes}}, \dots, h_{\tau+m}^{(l+1)\text{tes}}) \quad (8)$$

### 2.2.2 SEM-GCN

SEM-GCN 模块能够从句子中捕获语义信息, 其核心是结合自注意力机制构建词级自相关邻接矩阵, 可以将其视为表示词之间语义关联的带权无向图. 由于不依赖句法信息, SEM-GCN 在捕获单词之间的语义信息方面表现出更大的灵活性.

自注意力<sup>[29]</sup>能够高效地并行计算词汇间的注意力分数. SEM-GCN 通过自注意力层计算得到注意力分数矩阵  $A_{\text{sem}} \in \mathbb{R}^{n \times n}$ , 并将  $A_{\text{sem}}$  作为 SEM-GCN 的邻接矩阵, 其表达式为:

$$A_{\text{sem}} = \text{Softmax} \left( \frac{QW^Q \times (KW^K)^T}{\sqrt{d}} \right) \quad (9)$$

其中,  $Q$  和  $K$  为 SEM-GCN 模块中前一层的特征表示,  $W^Q$  和  $W^K$  是可训练的权重矩阵,  $d$  表示输入节点特征的维数. 需要注意的是, 仅使用单个自注意力头来生成句子的注意力分数矩阵. 由于词汇间的注意力分布重叠较小, 通过正交正则化器强制所有注意力分数向量的正交性可以提升语义表示性能. 正交正则化器的公

式表示如下:

$$R_O = \|A^{\text{sem}}(A^{\text{sem}})^T - I\|_F \quad (10)$$

其中,  $I$ 是单位矩阵, 下标  $F$ 表示弗罗贝尼乌斯范数.  $A^{\text{sem}}$  ( $A^{\text{sem}})^T$ 的非对角元素被最小化, 确保 $A^{\text{sem}}$ 矩阵能够保持正交.

最终, 采用与 TES-GCN 相同的处理操作, 如式 (11)-式 (13) 所示:

$$H_{\text{sem}}^{(l+1)} = \text{ReLU}(L^{\text{sem}} H_{\text{sem}}^l W_{\text{sem}}^{(l)}) \quad (11)$$

$$H_{\text{sem-zero}}^{l+1} = [0, 0, h_{\tau+1}^{(l+1)\text{sem}}, h_{\tau+2}^{(l+1)\text{sem}}, \dots, h_{\tau+m}^{(l+1)\text{sem}}] \quad (12)$$

$$T_2 = \text{AvgPooling}(h_{\tau+1}^{(l+1)\text{sem}}, h_{\tau+2}^{(l+1)\text{sem}}, \dots, h_{\tau+m}^{(l+1)\text{sem}}) \quad (13)$$

### 2.3 外部知识表示

ABSA 任务面临一个普遍问题, 即许多单词缺乏

独特的含义, 导致出现多义词或同义词, 难以准确识别情感极性. 为了解决这个问题, 我们使用 WordNet<sup>[30]</sup>作为外部知识图谱来丰富词汇特征, 并将它们作为知识源嵌入编码器. WordNet 是一个英语词汇数据库, 其中每个单词都与其同义词集合相关联, 以此形成了广泛的语义网络. 其中的单词和同义词集通过各种关系类型进行连接, 如同义词、反义词、上位词、下位词、单义和蕴含等. 使用 WordNet 作为先验知识可以客观地理解句子. 以“rambutan”为例, 与“fruit”或“food”等类似词连接, 有助于模型理解不常见词“rambutan”的含义. 此外, 由于它们在知识图谱中的关联, 它们的嵌入在嵌入空间中具有相似的表示, 如图 3 所示. 从而利用常见同义词来理解不常见和晦涩的单词含义, 特别是情感单词, 更容易理解句子的含义.

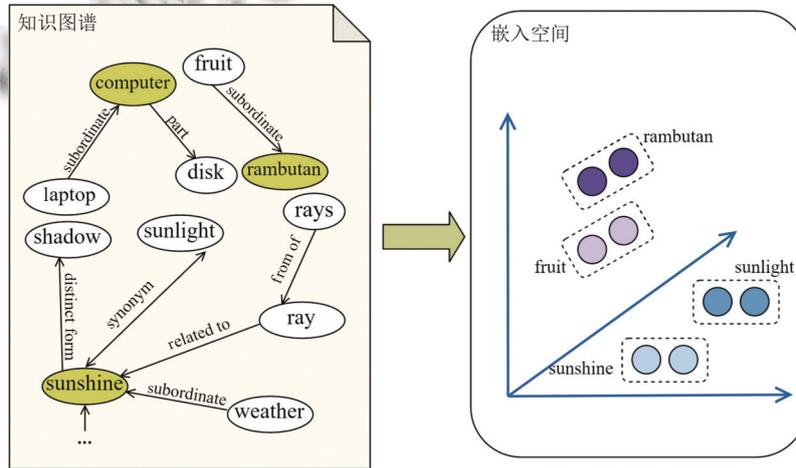


图3 知识图谱示例

参考 Islam 等<sup>[25]</sup>的研究, 引用了一种高效的方法来处静态外部知识图. 其核心思想是从知识图谱中提取知识嵌入, 并与上下文和方面建立联系. DistMult<sup>[31]</sup>和 ANALOGY<sup>[32]</sup>是两种常用的知识图嵌入方法, 它们通过不同的方式捕捉图谱中实体与关系之间的语义信息. DistMult 在处理已知关系时表现优秀, 能够快速且高效地计算实体间的关系. ANALOGY 更擅长从类比中推理出新的、未显现的关系, 能够捕捉知识图谱中的隐性结构. 结合这两种方法, 可以在处理显性和隐性关系的同时, 提升模型的能力. 本文利用这两种基于语义匹配的模型来将这些嵌入与文本中的上下文和方面建立联系. 具体来说, 首先获得知识嵌入, 分别表示为  $K^S$  和  $K^A$ , 再将其与 Bi-LSTM 的输出进行连接. 将上下

文和方面词的知识嵌入表示分别称为  $(KH)^S$  和  $(KH)^A$ , 如式 (14) 和式 (15) 所示, 其中, “;”表示连接.

$$(KH)^S = (K^S; H^S) \quad (14)$$

$$(KH)^A = (K^A; H^A) \quad (15)$$

为有效收集与方面词相关的上下文词汇, 采用了知识-方面注意机制, 如式 (16) 和式 (17) 所示:

$$\beta_t = \sum_{i=1}^n [(KH)_i^S]^T (KH)_i^A = \sum_{i=\tau+1}^{\tau+m} [(KH)_i^S]^T (KH)_i^A \quad (16)$$

$$\alpha_t = \frac{\exp(\beta_t)}{\sum_{i=1}^n \exp(\beta_i)} \quad (17)$$

$$T^{KH} = \sum_{i=1}^n \alpha_t (KH)_i^S \quad (18)$$

最后,屏蔽非方面词,应用平均池化来获得知识视角的表示  $T_3$ ,如式(19)、式(20)所示:

$$T_{zero}^{KH} = [0, 0, t_{\tau+1}^{kh}, t_{\tau+2}^{kh}, \dots, t_{\tau+m}^{kh}, \dots, 0] \quad (19)$$

$$T_3 = AvgPooling(t_{\tau+1}^{kh}, t_{\tau+2}^{kh}, \dots, t_{\tau+m}^{kh}) \quad (20)$$

## 2.4 多信息融合模块

多信息融合模块旨在开发一种从3个角度融合表示的方法.由于这3个视角的表示来自不同领域,因此该模块的主要目标是使模型能够从数据中获取最好的数据,并剔除最差的数据.与以前简单地合并多个路径的表示不同,我们采用局部门控全局卷积网络来融合不同视角的表示,让表述可以相互补充.多信息融合模块如图4所示,输入为  $\{T_1, T_2, T_3\}$ .

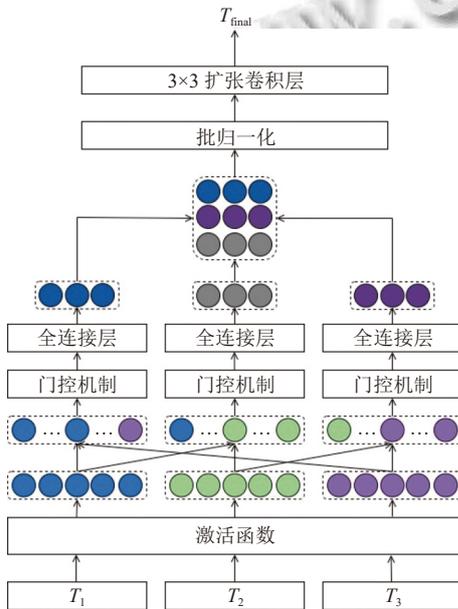


图4 多信息融合模块结构图

最初,  $\{T_1, T_2, T_3\}$  先分别通过激活函数.在局部门控过程中,从  $[T_1; T_2]$ 、 $[T_1; T_3]$  和  $[T_2; T_3]$  交叉应用特征,“;”表示向量串联运算符.交叉融合的局部特征分别输入门控机制中,这3个门控层的参数通过反向传播独立训练,并通过滤除噪声和调整门控控制机制中的权重,产生局部门控表示.局部交叉和门控机制表示如下:

$$g_{12} = \sigma(W_{12}^g [T_1; T_2]) \quad (21)$$

$$T_{12}^g = g_{12}T_1 + (1 - g_{12})T_2 \quad (22)$$

$$g_{13} = \sigma(W_{13}^g [T_1; T_3]) \quad (23)$$

$$T_{13}^g = g_{13}T_1 + (1 - g_{13})T_3 \quad (24)$$

$$g_{23} = \sigma(W_{23}^g [T_2; T_3]) \quad (25)$$

$$T_{23}^g = g_{23}T_2 + (1 - g_{23})T_3 \quad (26)$$

其中,符号  $\sigma$  表示 Sigmoid 函数,  $W_{12}^g$ 、 $W_{13}^g$  和  $W_{23}^g$  是可训练参数.

数据输入全连接层以协调维度,分别表示为  $T_{12}$ 、 $T_{13}$  和  $T_{23}$ .最后,为了充分利用这3个视角表示之间的互补性,使用扩张卷积进行全局融合.扩张卷积在不改变图像输出特征图的大小的情况下扩展了感受野,使得每个卷积输出包含更广泛的信息(即全局信息融合),如式(27)所示:

$$T_{final} = DilatedConvolution\left(BatchNorm\left(\left[T_{12}^g, T_{13}^g, T_{23}^g\right]^T\right)\right) \quad (27)$$

其中,  $\left[T_{12}^g, T_{13}^g, T_{23}^g\right]^T$  是通过列拼接获取的情感特征.为了增强数据分布的均匀性,使用批归一化.随后,输入  $3 \times 3$  扩张卷积层,选择性地融合特征,从而在不添加参数的情况下扩展了感受野,促进了全局融合.

## 2.5 情感输出层

$T_{final}$  作为最终预测结果,通过 *Softmax* 函数进行分类输出,生成情感标签的概率分布  $p$ ,如式(28)所示:

$$p = Softmax(W_o T_{final} + b_o) \quad (28)$$

其中,  $W_o$  和  $b_o$  分别表示可训练权重和偏置项.模型损失函数采用交叉熵损失函数,如式(29)所示:

$$Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C \hat{y}^{ij} \log(y^{ij}) \quad (29)$$

其中,  $N$  表示训练样本的数量,  $C$  表示情感类别的数量.  $y^{ij}$  和  $\hat{y}^{ij}$  分别是预测标签和真实标签.

## 3 实验

### 3.1 数据集介绍

为评估 MSI-GCN 的性能,本文使用4个基准数据集进行评估,分别是 Twitter 以及由 SemEval 官方发布的 ASC 基准数据集 Lap14、Rest14 和 Rest16<sup>[33]</sup>.这4个数据集的统计数据如表1所示.

### 3.2 实验参数设置与评价指标

本文实验使用300维的 GloVe 作为初始词嵌入,位置嵌入和知识嵌入维数分别设置为100和600. Bi-

LSTM的维数设置为300,批量大小设置为16,dropout设为0.3,以防止过拟合。L2正则化调整系数 $\lambda$ 范围为 $10^{-5}$ – $10^{-3}$ 。本文使用Adam优化器进行训练,并将学习率设置为 $10^{-3}$ 以优化模型参数。利用Spacy解析树获得的依赖关系和类型来构建句法图A,依赖类型的嵌入大小设置为300并随机初始化。经过研究,将GCN层的深度设置为2。

表1 数据集样本统计表

数据集	类别	积极	中立	消极
Twitter	Train	1561	3127	1560
	Test	173	346	173
Lap14	Train	994	464	870
	Test	341	169	128
Rest14	Train	2164	637	807
	Test	728	196	196
Rest16	Train	1620	88	709
	Test	597	38	190

模型的有效性通过准确率 (Accuracy, *Acc*) 以及 *Macro-F1* 指标进行评估,如式(30)–式(33)所示:

$$Precision = \frac{TP}{TP + FP} \quad (30)$$

$$Recall = \frac{TP}{TP + FN} \quad (31)$$

$$Macro-F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (32)$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (33)$$

### 3.3 对比实验

为了评估所提出模型的有效性,我们在4个公共数据集上将其性能与经典基线模型及主流的图卷积神经网络方法进行了比较。

(1) TD-LSTM<sup>[11]</sup>: 利用两个LSTM网络分别获取左侧和右侧上下文的隐藏向量,然后将这些隐藏向量结合生成方面表示。

(2) IAN<sup>[14]</sup>: 采用交互注意力网络以动态学习上下文与方面词间的关系。

(3) ASGCN<sup>[34]</sup>: 利用依赖树通过GCN获取节点句法信息,并结合注意力机制进行情感分类。

(4) Dual-GCN<sup>[4]</sup>: 利用两个独立的GCN分别提取语义信息和句法信息,然后通过双向仿射函数和新开发的正则化函数将其融合。

(5) EEGCN<sup>[35]</sup>: 结合BiLSTM和Transformer提

取上下文信息为处理长依赖文本提供显著优势,然后通过依赖图构建上下文依赖信息,进一步增强上下文表示。

(6) SEGCN<sup>[36]</sup>: 通过建立依赖图,利用方面与上下文词之间的距离来设置上下文权重,引入情感知识来构建情感词和方面词之间的关系。

(7) Sentic-GCN<sup>[8]</sup>: 将句法依赖树与SenticNet中的情感分数相结合来获得情感增强邻域矩阵,并通过GCN将情感信息传播到相邻节点。

(8) SK-GCN<sup>[26]</sup>: 利用GCN对句法依存图与外部知识进行建模,并将提取特征进行拼接。

### 3.4 结果与分析

通过表2的实验结果可以看出,MSI-GCN在4个公开数据集上的性能表现优于其他模型,其准确率和*Macro-F1*值均有一定提升。从结果中,我们发现了以下几点。首先,与基于注意力机制和单纯基于语义信息的模型相比(例如TD-LSTM、IAN),基于GCN的模型在捕捉上下文和特定方面之间的依赖关系方面表现优越,将更多信息(例如句法、情感特征)合并到模型中,因此取得了更好的结果。其次,MSI-GCN包含一个句法加权模块,能够自动学习句法依赖关系的权重,使其可以根据依赖类型的重要性分配权重信息。与ASGCN、EEGCN、SEGCN和Sentic-GCN相比,克服了为所有依赖类型分配相同权重的问题。此外,MSI-GCN模型通过多信息融合模块,利用局部门控-全局卷积融合机制进行多视角表示,优于依赖语法或外部知识的模型(例如Dual-GCN、EEGCN、SEGCN和SK-GCN),能够更好地完成情感分析。

在Lap14数据集上,MSI-GCN性能略逊于SEGCN。这一结果差异可能是由于Rest16数据集的样本量较小,限制了外部知识图的利用。但值得注意的是,MSI-GCN在Twitter数据集上明显优于其他模型,这可能与Twitter数据集的非正式特征有关。我们的模型因对外部知识图的充分利用,从而促成了出色的结果。此外,通过将MSI-GCN模型与BERT模型集成,进一步提高了性能并在4个数据集上取得了最高分。这表明本文充分利用了BERT的优势,将其与MSI-GCN相结合,实现了有效的增益。

### 3.5 消融实验

我们设计了消融实验来定量评估MSI-GCN中各个模块的影响,实验结果见表3。

表2 MSI-GCN 与基准模型的性能对比 (%)

编码方式	模型	Twitter		Lap14		Rest14		Rest16	
		Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1
GloVe	TD-LSTM	70.80	69.00	78.00	66.73	71.83	68.43	82.16	54.21
	IAN	72.50	70.81	72.05	67.38	79.26	70.09	84.74	55.21
	ASGCN	72.15	70.40	75.55	71.05	80.77	72.02	88.99	67.48
	Dual-GCN	75.92	74.29	78.48	74.74	84.27	78.08	—	—
	EEGCNN	72.44	70.67	77.83	75.21	81.70	73.63	88.83	68.91
	SEGCN	76.51	75.32	79.43	<b>76.49</b>	84.72	77.51	—	—
	Sentic-GCN	—	—	77.90	74.71	84.03	75.38	90.88	74.82
	SK-GCN	71.97	70.22	77.62	73.84	81.53	72.90	85.17	68.08
	MSI-GCN (Ours)	<b>78.92</b>	<b>77.68</b>	<b>79.52</b>	76.35	<b>84.75</b>	<b>77.52</b>	<b>90.93</b>	<b>75.91</b>
BERT	Dual-GCN-BERT	77.40	76.02	81.80	78.10	87.13	81.16	—	—
	SEGCN-BERT	77.40	76.02	81.01	77.96	87.31	81.09	—	—
	Sentic-GCN-BERT	—	—	82.12	79.05	86.92	81.03	91.97	79.56
	MSI-GCN-BERT (Ours)	<b>78.72</b>	<b>78.04</b>	<b>82.94</b>	<b>79.72</b>	<b>87.94</b>	<b>82.17</b>	<b>92.48</b>	<b>80.24</b>

表3 消融实验结果 (%)

模型	Twitter		Lap14		Rest14		Rest16	
	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1	Acc	Macro-F1
w/o PI	77.86	76.27	78.94	75.41	83.64	76.69	89.36	74.55
w/o Sem	76.87	75.70	77.72	73.75	82.83	76.59	88.74	72.54
w/o Tes	77.82	76.33	77.91	75.14	83.95	77.96	89.98	70.63
w/o Type	77.54	76.05	77.55	74.94	83.37	77.03	89.15	69.09
w/o KG	78.09	74.97	78.58	75.03	83.37	76.54	88.64	73.12
w/o FFL	76.67	74.84	77.36	73.74	82.92	76.44	87.62	71.74
MSI-GCN	78.92	77.68	79.52	76.35	84.75	77.52	90.93	74.82

w/o PI: 在构建邻接矩阵时去除相对位置信息, 确保 GCN 输出的邻接矩阵中的每个节点仍然不知道其与句子中方面词的相对位置。

w/o Sem: 去掉基于自注意力的语义信息, 根据控制变量法, 利用 Bi-LSTM 的输出作为 MSI-GCN 中的  $T_2$ 。

w/o Type: 将所有依赖关系类型的权重设置为固定值 1, 而不是从嵌入中学习。

w/o Tes: 去掉了句法图卷积网络, 其余部分类似于 MSI-GCN。

w/o KG: 去除了知识图嵌入。由于在标准 MSI-GCN 中, 知识嵌入与 Bi-LSTM 的隐藏输出相结合, 所以将 Bi-LSTM 的输出作为  $T_3$ 。

w/o FFL: 去掉局部门控全局卷积网络, 简单地将这 3 个视角的表示连接起来。

从表 3 结果可以观察到删除相对位置信息后模型性能下降, 这表明当一个词利用 GCN 聚合相邻节点时, 没有差异的贡献值会导致噪声增加。同样, 将依赖类型视为相同也使得模型性能下降, 这验证了句法加权模块在学习依赖关系权重方面的有效性。此外, 删除语义模块和句法模块都显著降低了性能, 这表明语义

信息和句法信息对情感分析的重要性。去掉外部知识图的引用使模型性能下降表明了外部知识图的重要作用, 有效利用知识图谱中的概念级知识可以使模型更准确地理解原始单词和口语表达。最后, 使用简单的串联而不是局部门控全局卷积网络也会导致性能下降, 因为简单的串联会失去大部分信息的相关性, 比没有 SemGCN 或 TesGCN 的模型还要严重。这表明这 3 个视角的表示在某种程度上是互补的, 充分验证了多信息融合模块的可靠性。

### 3.6 GCN 层数影响

为研究双通道中不同 GCN 层数对模型性能的影响, 我们在 Rest14、Lap14 和 Twitter 数据集上对 GCN 层数从 1-7 进行测试, 结果如图 5 和图 6 所示。

根据实验结果, MSI-GCN 模型在 2 层 GCN 时表现最佳。当层数过少时, 难以充分聚合深层节点的特征信息, 导致模型性能较差。而层数过多时, 节点间的信息会变得过于复杂, 容易转化为噪声。此外, 层数过多会增加模型的计算复杂度, 而将层数限制为两层可以在保持一定性能的同时降低计算成本, 特别是在处理大规模数据时更为明显。层数过多还可能导致模型不

稳定和梯度爆炸等问题. 限制层数有助于控制模型复杂度, 避免过拟合, 提高模型对新数据的泛化能力.

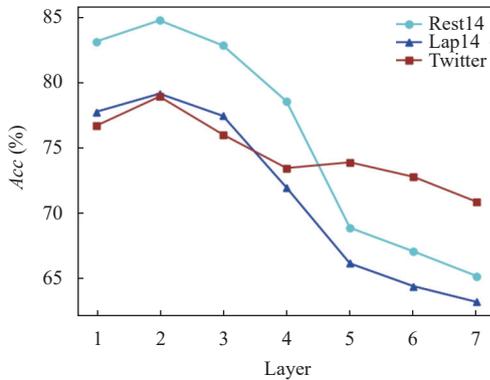


图5 GCN网络层数与准确率的关系

### 3.7 案例分析

通过案例研究, 评估 MSI-GCN 模型在实际应用中的性能提升情况, 具体如表 4 所示. 我们将 MSI-GCN 与 IAN、ASGCN 和 Sentic-GCN 进行比较, 从数据集选取典型案例进行进一步分析, 并探讨这些案例在实际应用的结果.

从实验结果中可以观察到, 基于上下文的方法 IAN 在预测结果方面表现较差, 而基于语义信息的模型则

表现更好. 在文本 2 中, ASGCN 未能捕捉到从方面词汇中的远程否定语义, 因此预测结果为中性. 从文本 3 的结果可以发现, 通过引入外部知识, Sentic-GCN 和 MSI-GCN 能够更准确地捕捉情感信息, 从而准确预测中性情感极性. 在文本 4 中, “wish”语气的假设影响了模型对整体语义的捕捉. 尽管基于依存关系或情感信息的语法模型能够更好地捕捉情感信息, 但它们仍然受到句子综合语义上下文的影响, 在情感分析预测中存在一定困难. 然而, MSI-GCN 模型利用了这 3 种视角特征的互补性, 并增强了语法信息, 使得模型能够准确预测情感极性.

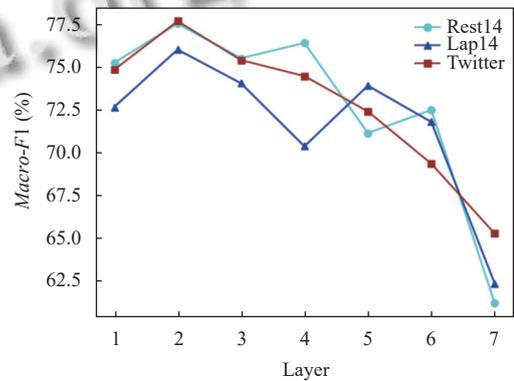


图6 GCN网络层数与 Macro-F1 的关系

表 4 案例分析

方面词	句子	IAN	ASGCN	Sentic-GCN	MSI-GCN	Label
{food, service}	Great food but the service was dreadful!	(N, N)	(P, N)	(P, N)	(P, N)	(P, N)
{i5}	Not as fact as I would have expect for an i5.	P	O	N	N	N
{food, prices}	How can they survive serving mediocre food at exorbitant prices?	(N, N)	(N, N)	(O, N)	(O, N)	(O, N)
{webcam}	I wish it had a webcam though, then it would be perfect.	P	P	O	N	N

为检验 MSI-GCN 的有效性, 我们对句子中单词的情感贡献进行了可视化分析. 图 7 是对 Lap14 数据集的句子“Not as fact as I would have expected for an i5”中各单词情感贡献的可视化, 深度与情感贡献度成正比. 其中, “only Sem”表示仅保留语义通道; “only Sy”表示仅保留语法通道; “only KG”表示知识图谱视角. 通过结果发现, 语义通道和知识图谱视角更强调程度副词和情感词, 从而使 MSI-GCN 能够有效理解句子的语义.

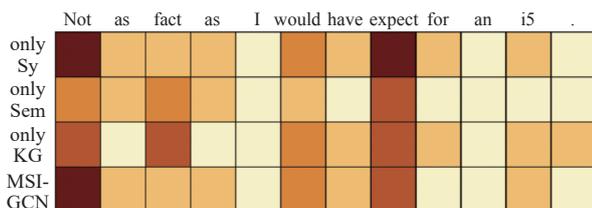


图7 情感贡献可视化

当仅依赖语义图进行特征提取时, 由于缺乏远程句法依赖信息, 模型无法准确捕捉方面词汇“i5”与意见词“Not”之间的远距离语法关系. 而在通过句法图进行特征学习时, 模型能够有效建立方面词“i5”与“Not”之间的联系, 但存在干扰词“expect”, 给该词赋予相对较大的权重. 然而, MSI-GCN 则能够有效利用这 3 种视角特征的互补性, 从而有效地连接方面词汇和意见词, 并自适应地减少对分散词汇“expect”的关注, 最终实现情感极性的正确分类, 这验证了 MSI-GCN 的有效性.

## 4 结束语

本文提出了一种用于多信息融合的图卷积神经网络. 首先, 通过依存类型嵌入加权语法依存矩阵, 以增强语法信息, 并同时捕捉语义特征以编码自注意力矩

阵. 引入正交正则化以增强语义相关性, 并嵌入知识图谱以利用基于知识的注意力过滤噪声. 最后, 利用局部门控全局卷积网络融合3种视角特征, 通过发挥特征的互补性, 获得更准确的情感倾向. 本文方法在4个标准数据集上进行了验证, 在准确率和 *Macro-F1* 值方面均取得了良好结果. 在未来, 一方面我们将考虑探索更多知识库或使用大规模预训练语言模型, 以提供更强的语义表示能力, 进一步提升性能. 另一方面, 由于采用了双通道 GCN, 模型学习时间有所增加, 我们也将把工作重点放在探索如何结合知识蒸馏或加速算法, 在融入大模型研究的同时, 兼顾效率提升, 以提高学习速度, 实现性能与效率的平衡发展.

### 参考文献

- 1 Dong L, Wei FR, Tan CQ, *et al.* Adaptive recursive neural network for target-dependent twitter sentiment classification. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Baltimore: ACL, 2014. 49–54. [doi: [10.3115/v1/p14-2009](https://doi.org/10.3115/v1/p14-2009)]
- 2 Huang BX, Carley K. Syntax-aware aspect level sentiment classification with graph attention networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong: ACL, 2019. 5469–5477. [doi: [10.18653/v1/d19-1549](https://doi.org/10.18653/v1/d19-1549)]
- 3 Wang WY, Pan SJ, Dahlmeier D, *et al.* Recursive neural conditional random fields for aspect-based sentiment analysis. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: ACL, 2016. 616–626. [doi: [10.18653/v1/d16-1059](https://doi.org/10.18653/v1/d16-1059)]
- 4 Li RF, Chen H, Feng FX, *et al.* Dual graph convolutional networks for aspect-based sentiment analysis. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). ACL, 2021. 6319–6329. [doi: [10.18653/v1/2021.acl-long.494](https://doi.org/10.18653/v1/2021.acl-long.494)]
- 5 Wang K, Shen WZ, Yang YY, *et al.* Relational graph attention network for aspect-based sentiment analysis. arXiv:2004.12362, 2020.
- 6 Wu LF, Chen Y, Shen K, *et al.* Graph neural networks for natural language processing: A survey. Foundations and Trends® in Machine Learning, 2023, 16(2): 119–328. [doi: [10.1561/22000000096](https://doi.org/10.1561/22000000096)]
- 7 Liang B, Su H, Gui L, *et al.* Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. Knowledge-based Systems, 2022, 235: 107643. [doi: [10.1016/j.knosys.2021.107643](https://doi.org/10.1016/j.knosys.2021.107643)]
- 8 Zhong QH, Ding L, Liu JH, *et al.* Knowledge graph augmented network towards multiview representation learning for aspect-based sentiment analysis. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(10): 10098–10111. [doi: [10.1109/tkde.2023.3250499](https://doi.org/10.1109/tkde.2023.3250499)]
- 9 Yasunaga M, Ren HY, Bosselut A, *et al.* QA-GNN: Reasoning with language models and knowledge graphs for question answering. arXiv:2104.06378, 2021.
- 10 Zhang Z, Zhou ZL, Wang YN. SSEGCN: Syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis. Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Seattle: ACL, 2022. 4916–4925. [doi: [10.18653/v1/2022.naacl-main.362](https://doi.org/10.18653/v1/2022.naacl-main.362)]
- 11 Tang DY, Qin B, Feng XC, *et al.* Effective LSTMs for target-dependent sentiment classification. arXiv:1512.01100, 2015.
- 12 Zhang MS, Zhang Y, Vo DT. Gated neural networks for targeted sentiment analysis. Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix: AAAI Press, 2016. 3097–3093. [doi: [10.1609/aaai.v30i1.10380](https://doi.org/10.1609/aaai.v30i1.10380)]
- 13 Chen P, Sun ZQ, Bing LD, *et al.* Recurrent attention network on memory for aspect sentiment analysis. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen: ACL, 2017. 452–461. [doi: [10.18653/v1/d17-1047](https://doi.org/10.18653/v1/d17-1047)]
- 14 Ma DH, Li SJ, Zhang XD, *et al.* Interactive attention networks for aspect-level sentiment classification. arXiv: 1709.00893, 2017.
- 15 Lu Q, Zhu ZF, Zhang GY, *et al.* Aspect-gated graph convolutional networks for aspect-based sentiment analysis. Applied Intelligence, 2021, 51(7): 4408–4419. [doi: [10.1007/s10489-020-02095-3](https://doi.org/10.1007/s10489-020-02095-3)]
- 16 Zhou J, Jin SQ, Huang XL. ADeCNN: An improved model for aspect-level sentiment analysis based on deformable CNN and attention. IEEE Access, 2020, 8: 132970–132979. [doi: [10.1109/access.2020.3010802](https://doi.org/10.1109/access.2020.3010802)]
- 17 Liang S, Wei W, Mao XL, *et al.* BiSyn-GAT+: Bi-syntactic aware graph attention network for aspect-based sentiment analysis. Proceedings of the 2022 Findings of the Association

- for Computational Linguistics: ACL 2022. Dublin: ACL, 2022. 1835–1848. [doi: [10.18653/v1/2022.findings-acl.144](https://doi.org/10.18653/v1/2022.findings-acl.144)]
- 18 Qiu ZX, Chen KH, Xue Y, *et al.* Modeling inter-aspect relations with clause and contrastive learning for aspect-based sentiment analysis. *IEEE Transactions on Computational Social Systems*, 2024, 11(2): 2833–2842. [doi: [10.1109/tcss.2023.3302331](https://doi.org/10.1109/tcss.2023.3302331)]
- 19 Yuan L, Wang J, Yu LC, *et al.* Syntactic graph attention network for aspect-level sentiment analysis. *IEEE Transactions on Artificial Intelligence*, 2024, 5(1): 140–153. [doi: [10.1109/tai.2022.3227535](https://doi.org/10.1109/tai.2022.3227535)]
- 20 Zhao ZG, Tang MW, Tang W, *et al.* Graph convolutional network with multiple weight mechanisms for aspect-based sentiment analysis. *Neurocomputing*, 2022, 500: 124–134. [doi: [10.1016/j.neucom.2022.05.045](https://doi.org/10.1016/j.neucom.2022.05.045)]
- 21 张文轩, 殷雁君, 智敏. 用于方面级情感分析的情感增强双图卷积网络. *计算机科学与探索*, 2024, 18(1): 217–230. [doi: [10.3778/j.issn.1673-9418.2209033](https://doi.org/10.3778/j.issn.1673-9418.2209033)]
- 22 Ma YK, Peng HY, Cambria E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM. *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2018. 5876–5883. [doi: [10.1609/aaai.v32i1.12048](https://doi.org/10.1609/aaai.v32i1.12048)]
- 23 Zhang K, Liu Q, Qian H, *et al.* EATN: An efficient adaptive transfer network for aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(1): 377–389. [doi: [10.1109/tkde.2021.3075238](https://doi.org/10.1109/tkde.2021.3075238)]
- 24 李阳, 唐积强, 朱俊武, 等. 基于 prompt 和知识增强的方面级情感分析. *计算机学报*, 2023, 46(1): 77–83. [doi: [10.11896/jsjcx.220300279](https://doi.org/10.11896/jsjcx.220300279)]
- 25 Islam SM, Bhattacharya S. AR-BERT: Aspect-relation enhanced aspect-level sentiment classification with multi-modal explanations. *arXiv:2108.11656v2*, 2022.
- 26 Zhou J, Huang JX, Hu QV, *et al.* SK-GCN: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification. *Knowledge-based Systems*, 2020, 205: 106292. [doi: [10.1016/j.knosys.2020.106292](https://doi.org/10.1016/j.knosys.2020.106292)]
- 27 Pennington J, Socher R, Manning C. GloVe: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha: ACL, 2014. 1532–1543. [doi: [10.3115/v1/d14-1162](https://doi.org/10.3115/v1/d14-1162)]
- 28 Devlin J, Chang MW, Lee K, *et al.* BERT: Pre-training of deep bidirectional Transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis: ACL, 2019. 4171–4186. [doi: [10.18653/v1/n19-1423](https://doi.org/10.18653/v1/n19-1423)]
- 29 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. *Proceedings of the 31st Conference on Neural Information Processing Systems*. Long Beach: NIPS, 2017. 1–15. [doi: [10.48550/arxiv.1706.03762](https://doi.org/10.48550/arxiv.1706.03762)]
- 30 Miller GA. WordNet: A lexical database for English. *Communications of the ACM*, 1995, 38(11): 39–41. [doi: [10.1145/219717.219748](https://doi.org/10.1145/219717.219748)]
- 31 Yang BS, Yih WT, He XD, *et al.* Embedding entities and relations for learning and inference in knowledge bases. *Proceedings of the 3rd International Conference on Learning Representations*. San Diego: OpenReview.net, 2014. 1–12. [doi: [10.48550/arxiv.1412.6575](https://doi.org/10.48550/arxiv.1412.6575)]
- 32 Liu H, Wu Y, Yang Y. Analogical inference for multi-relational embeddings. *Proceedings of the 34th International Conference on Machine Learning*. Sydney: JMLR.org, 2017. 2168–2178.
- 33 Pontiki M, Galanis D, Papageorgiou H, *et al.* Semeval-2016 task 5: Aspect based sentiment analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation*. San Diego: ACL, 2016. 19–30. [doi: [10.18653/v1/s16-1002](https://doi.org/10.18653/v1/s16-1002)]
- 34 Zhang C, Li QC, Song DW. Aspect-based sentiment classification with aspect-specific graph convolutional networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong: ACL, 2019. 4568–4578. [doi: [10.18653/v1/d19-1464](https://doi.org/10.18653/v1/d19-1464)]
- 35 Li C, Zhang JD, Tang HL, *et al.* Advancing aspect-based sentiment analysis through deep learning models. *Proceedings of the 20th International Conference on Advanced Data Mining and Applications*. Sydney: Springer, 2024. 228–242. [doi: [10.1007/978-981-96-0847-8\\_16](https://doi.org/10.1007/978-981-96-0847-8_16)]
- 36 Zheng YQ, Li X, Nie JY. Store, share and transfer: Learning and updating sentiment knowledge for aspect-based sentiment analysis. *Information Sciences*, 2023, 635: 151–168. [doi: [10.1016/j.ins.2023.03.102](https://doi.org/10.1016/j.ins.2023.03.102)]

(校对责编: 王欣欣)