

## 融合局部多视角语言特征和全局特征的对话情感四元组抽取

彭菊红, 张正悦, 丁子胥, 范馨予, 胡长玉, 赵明俊

引用本文

彭菊红, 张正悦, 丁子胥, 范馨予, 胡长玉, 赵明俊. [融合局部多视角语言特征和全局特征的对话情感四元组抽取](#)[J]. 计算机科学, 2026, 53(4): 384-392.

PENG Juhong, ZHANG Zhengyue, DING Zixu, FAN Xinyu, HU Changyu, ZHAO Mingjun. [Multi-view Local Language Feature and Global Feature Fusion for Conversational Aspect-based Sentiment Quadruple Analysis](#) [J]. Computer Science, 2026, 53(4): 384-392.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

### [知识辅助和强化句法驱动的方面级情感分析](#)

Knowledge-assisted and Reinforced Syntax-driven for Aspect-based Sentiment Analysis

计算机科学, 2026, 53(4): 406-414. <https://doi.org/10.11896/jsjcx.250600117>

### [基于去噪扩散模型增强的换电需求数据生成算法](#)

Denoising Diffusion Model-enhanced Algorithm for Battery Swap Demand Data Generation

计算机科学, 2026, 53(4): 163-172. <https://doi.org/10.11896/jsjcx.250600205>

### [KGMamba:基于Kolmogorov-Arnold网络优化图卷积网络和Mamba的基因调控网络预测模型](#)

KGMamba:Gene Regulatory Network Prediction Model Based on Kolmogorov-Arnold Network

Optimizing Graph Convolutional Network and Mamba

计算机科学, 2026, 53(4): 101-111. <https://doi.org/10.11896/jsjcx.250500097>

### [CA-SFTNet:基于空间特征变换和浓缩注意力机制的皮肤病灶分割模型](#)

CA-SFTNet:Skin Lesion Segmentation Model Based on Spatial Feature Transformation and

Concentrated Attention Mechanism

计算机科学, 2026, 53(3): 277-286. <https://doi.org/10.11896/jsjcx.250200049>

### [基于改进YOLO算法的学生行为检测方法](#)

Student Behavior Detection Method Based on Improved YOLO Algorithm

计算机科学, 2026, 53(3): 246-256. <https://doi.org/10.11896/jsjcx.241100165>

# 融合局部多视角语言特征和全局特征的对话情感四元组抽取

彭菊红<sup>1,3</sup> 张正悦<sup>1,3</sup> 丁子胥<sup>1,3</sup> 范馨予<sup>1,3</sup> 胡长玉<sup>1,3</sup> 赵明俊<sup>2</sup>

1 湖北大学人工智能学院 武汉 430062

2 湖北大学计算机学院 武汉 430062

3 湖北大学智能感知系统与安全教育部重点实验室 武汉 430062

(juhongpeng@hubu.edu.cn)

**摘要** 基于对话的方面情感四元组抽取(DiaASQ)是情感分析(ABSA)领域的一个新兴研究方向,其目标旨在从一段对话中识别并提取情感四元组(目标、方面、观点和情感极性)。与传统静态文本的ABSA任务相比,DiaASQ面临以下两大问题:1)对话文本通常较长,目标、方面、观点等情感要素可能分散在多个话语中,难以捕捉长距离依赖关系;2)对话文本结构复杂,通常包含多位发言者和回复关系,信息往往存在跨语句和说话人的情况,回复结构更为复杂。针对上述问题,提出一种融合局部多视角语言特征和全局特征的对话情感四元组抽取(MVLLF-GF)方法。首先,利用多视角语言知识编码器从句法依存关系、语义信息等多个角度对词元进行交互增强,捕捉长距离依赖关系,学习局部特征;其次,使用全局话语编码器从话语层面学习发言者信息和回复关系信息,获取全局特征;再次,使用多粒度融合器对不同层面的特征进行深度整合,增强模型上下文理解能力;最后,使用网格标注的方法实现情感四元组的端到端解码。实验结果表明,在DiaASQ公开中文数据集ZH和英文数据集EN上,与基准模型MVQPN相比,所提模型在Micro F1指标上分别提升了9.13个百分点和6.50个百分点,证明了该方法的有效性。

**关键词:** 对话情感四元组抽取;句法依存关系;注意力机制;语义信息;图卷积网络

**中图分类号** TP391

## Multi-view Local Language Feature and Global Feature Fusion for Conversational Aspect-based Sentiment Quadruple Analysis

PENG Juhong<sup>1,3</sup>, ZHANG Zhengyue<sup>1,3</sup>, DING Zixu<sup>1,3</sup>, FAN Xinyu<sup>1,3</sup>, HU Changyu<sup>1,3</sup> and ZHAO Mingjun<sup>2</sup>

1 School of Artificial Intelligence, HuBei University, Wuhan 430062, China

2 School of Computer Science, HuBei University, Wuhan 430062, China

3 Key Laboratory of Intelligent Perception Systems and Security Ministry of Education(HuBei University), Wuhan 430062, China

**Abstract** Conversational aspect-based sentiment quadruple analysis(DiaASQ) is an emerging research direction in the field of ABSA(Aspect-Based Sentiment Analysis), which aims to identify and extract sentiment quadruples—namely, target, aspect, opinion, and sentiment polarity—from a given dialogue. Compared with traditional ABSA tasks on static texts, DiaASQ faces two major challenges: 1) dialogue texts are often lengthy, with sentiment elements such as targets, aspects, and opinions scattered across multiple utterances, making it difficult to capture long-range dependencies; 2) dialogue structures are more complex, typically involving multiple speakers and reply relationships, where information frequently spans sentences and speakers, leading to intricate interaction patterns. To address these challenges, this paper proposes MVLLF-GF, a model that integrates multi-view local language features with global contextual representations for dialogue-based sentiment quadruple extraction. Specifically, a multi-view linguistic knowledge encoder is employed to enhance token-level interactions from multiple perspectives, including syntactic dependency and semantic information, thereby learning rich local features. A global utterance encoder is then introduced to capture global features by modeling speaker identities and reply relationships at the utterance level. Furthermore, a multi-granularity fusion module is designed to deeply integrate features across different levels, enhancing the model's contextual understanding. Finally, an end-to-end grid tagging mechanism is applied to decode sentiment quadruples. Experimental results on the public DiaASQ Chinese dataset(ZH) and English dataset(EN) demonstrate that the proposed method achieves Micro-F1 improvements of 9.13 percentage points and 6.50 percentage points, respectively, over the baseline model MVQPN, verifying its effectiveness.

到稿日期:2025-09-03 返修日期:2026-01-09

基金项目:国家自然科学基金面上项目(62377009)

This work was supported by the General Program of the National Natural Science Foundation of China(62377009).

通信作者:赵明俊(12669003@qq.com)

**Keywords** Conversational aspect-based sentiment quadruple, Syntactic dependency relation, Attention mechanisms, Semantic information, Graph convolutional network

### 1 引言

方面级情感分析<sup>[1]</sup>是自然语言处理中的重要邻域,旨在检测句子中的情感倾向。具体任务可以分成两大类:第一类是单任务,专注于提取方面词、观点词和判别情感极性;第二类是复合任务,致力于多要素联合抽取,如方面词-观点词配对或方面情感四元组抽取等。如今,关于方面级情感分析的研究已经取得了显著的进步。然而,传统的方面级情感分析大多聚焦于单句或独立文本的情感分析,在动态复杂的对话场景中表现不足,忽略了真实的多轮对话场景中信息跨语句、跨发言人分布的复杂性。在实际的社交网络和客服对话中存在的多轮对话式评论(如电商问答区、用户追评、买家互评等)中,用户的观点往往分散在多轮、多发言人的互动之中,情感信息的表达更加隐晦,结构更为复杂。传统的方面级研究方法难以准确捕捉其中的情感。因此,在对话场景中进行情感分析变得尤为重要。

在此背景下,基于对话的方面情感四元组抽取任务被提出,其目标是在给定的多轮对话中抽取所有的情感四元组(T,A,O,S),即目标词T(对话讨论的主题)、方面词A(目标词所具有的某一项属性)、观点词O(对方面词的态度或评价)和情感极性S,其中情感极性包含积极(Positive)、消极(Negative)、中性(Neutral)。图1给出了多个发言者(A,B,C,D,E)关于手机电量讨论的真实对话,需要从中提取出所有的四元组,如表1所列。方面词、观点词和目标词可能分别在不同的对话中,并且不同组合的四元组情感极性也不尽相同。这不仅要求模型有很好的上下文理解能力,还要求能够整合局部信息和全局信息,精确定位情感要素之间的关联。

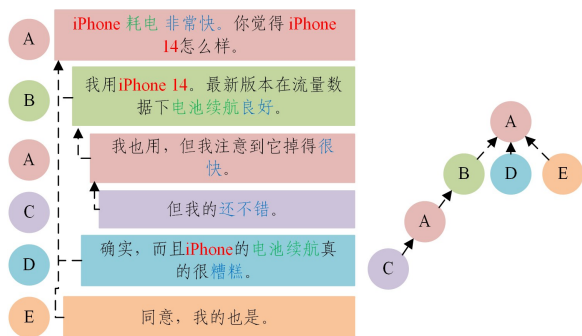


图1 基于对话的四元组实例

Fig.1 Example of conversational aspect-based sentiment quadruple analysis

表1 四元组示例

Table 1 Example of quadruple

Target	Aspect	Opinion	Sentiment
iPhone	耗电	非常快	消极
iPhone 14	电池续航	良好	积极
iPhone 14	电池续航	很快	消极
iPhone 14	电池续航	还不错	中性
iPhone	电池续航	糟糕	消极

现有研究主要包括序列标注方法、多任务学习方法、生成式方法,尽管这些方法都取得了一定的成功,但都存在一定的困难。目前,基于对话的情感四元组抽取任务主要存在以下两个难点。首先,对话较长,目标词、方面词、观点词可能分散在多个话语中,难以捕捉长距离依赖关系。在多轮对话中情感四元组的组成要素——目标词(Target)、方面词(Aspect)、观点词(Opinion)以及情感极性(Sentiment),往往并不在同一句话中,而是分布在对话的不同句子中。例如,在图1中,B在第二句话中提到目标词iPhone 14、方面词“电池续航”以及它的观点词“良好”,此时即可组成一个情感四元组(iPhone14, 电池续航, 良好, 积极),并且他们存在于同一句话中。然而,A在第三句话中提到观点词“很快”,但这句话中并没有明显的目标词和方面词,实际上它的目标词是A第一句话中提到的iPhone14,方面词是B第二句中的“电池续航”,同样组成了一个四元组(iPhone14, 电池续航, 很快, 消极),但他们不存在于同一句话中。这种四元组跨越多个对话的分布,使得模型必须具备强大的长距离信息捕捉与对齐能力。其次,存在对话结构复杂,包含多位发言者和回复关系,发言者之间的信息往往存在跨语句、跨说话人的情况。例如,在图1中,每句话不仅是独立存在的,还存在如图1中右上角的树状结构回复关系。B是对A的回复,A又对B进行了回复,而C同时又对A进行了补充,此时观点词“还不错”对应的目标词应该是B第二句话中提出的“iPhone 14”,方面词则是A第三句话中提到的“电池续航”。他们组成了一个四元组(iPhone14, 电池续航, 还不错, 中性)。这种多发言者、多分支的回复结构(而非简单的序列结构),就需要模型不仅关注词元级(每句话中独立存在的一个个单词)的信息,还需要理解这种回复关系以及对话结构。

为了克服这些难点,本文提出一种融合局部多视角语言特征和全局特征的对话情感四元组抽取。首先,使用图神经网络GCN<sup>[2]</sup>对依存关系、句法依存距离、相对位置距离、语义知识进行学习,融合情感词典从多个视角学习得到局部词元信息;其次,使用图神经网络对对话中的回复关系进行学习,提取全局话语信息;再次,设计了3种注意力得分矩阵,将局部词元信息和全局话语信息进行融合;最后,使用网格标注方法对四元组进行解码,进一步提高了基于对话的方面级情感四元组抽取的性能。本文的贡献总结如下:

1)与以往仅停留在单一层面建模的方法不同,本文从不同层面捕获对话的信息,使用多视角语言知识编码器对词元进行交互增强,捕捉长距离依赖关系,学习局部知识,获得词元级信息。此外,使用全局话语编码器在话语层面学习发言者信息和回复信息,学习全局信息,获得话语级信息,兼顾对话中局部与全局的信息。

2)为了更好地融合局部信息和全局信息,提出多粒度特征融合机制,构建3种注意力矩阵,实现局部信息和全局信息的深度对齐,以及词元级与话语级特征的细粒度交互,提升了情感四元组抽取模型融合不同层面信息的能力。

3)在 DiASQ 任务的公开数据集上进行了充分的对比和消融实验,实验结果表明,本文方法优于其他基线模型。

## 2 相关工作

随着大模型对话系统和社交媒体的广泛应用,情感分析任务需要扩展至对话语境。对话中的情感表达往往涉及多个说话人,且不同的观点元素可能分布在不同的发言之中。对此,Li 等<sup>[3]</sup>首次提出了基于对话的方面级情感四元组抽取 (Conversational Aspect-based Sentiment Quadruple Analysis, DiaASQ)并构建了相应的数据集,旨在从多轮对话中抽取完整的四元组。他们提出了四元组预测网络模型 (MVQPN),重新设计了网格标注方法<sup>[4]</sup>中的标签方案,将原始任务分解为实体边界标注、实体对关系标注和情感极性标注 3 个部分。此外,还设计了发言者掩码、回复掩码和线程掩码,结合掩码多头注意力机制和旋转位置嵌入<sup>[5]</sup>来增强模型的对话感知能力。但其主要建模粒度仍然停留在词元级,未能充分利用话语级信息。为了解决该问题,Cai 等<sup>[6]</sup>在 MVQPN 的基础上进行改进,提出了整体四元组预测网络模型 (Overall-QPN)。该方法尝试将整个对话输入预训练语言模型<sup>[7]</sup> (PLM),直接在编码阶段捕捉全局依赖信息。实验表明,该方法在整体性能上优于 MVQPN,尤其在跨话语四元组的抽取上更具优势。然而,整体建模也带来了新的问题:对话往往较长,容易超出 PLM 的输入限制,导致信息截断和上下文丢失。在此基础上,Lai 等<sup>[8]</sup>提出了另一种改进思路,将 MVQPN 中的 3 种并行注意力机制替换为连续交互机制,并结合  $k$  折交叉训练<sup>[9]</sup>与人工规则约束,增强模型对跨话语依赖关系的捕捉。此外,该方法尝试通过跨语言迁移的方式,将中文数据集训练得到的权重应用于英文任务,取得了较好的效果。除了全局编码,另一类方法尝试显式建模对话中的结构关系。Zhao 等<sup>[10]</sup>提出了多维对话特征交互模型 (MDFI),利用多维对话特征交互机制,结合 ResNet<sup>[11]</sup>网络对发言者、回复关系以及话题线索进行建模,从而增强不同语义单元之间的关联。该方法提升了多轮对话中的特征利用率,但融合策略较为粗糙,容易引入冗余信息,缺乏对结构信息的精细对齐。进一步地,

Li 等<sup>[12]</sup>提出了异构对话三元交互模型 (H2DT),构建了基于发言者与回复者的异构图神经结构,并引入异构注意力机制<sup>[13]</sup>来建模复杂的词元关系,同时设计三元交互评分机制,提升目标-方面-观点三者间的耦合程度。H2DT 在跨发言场景下表现优异,但复杂的图结构和交互机制带来了计算效率和训练稳定性方面的挑战。此外,部分研究尝试使用生成式方法来进行抽取:利用预训练语言模型 T5 将目标词、方面词、观点词以及情感极性四元组以序列化形式直接生成,实现端到端抽取。此类方法避免了复杂的标签设计,有较好的泛化能力,但生成结果受解码策略影响较大,稳定性不足,难以控制生成内容的结构一致性,在多发言者和多轮对话场景中易出现要素搭配错误的问题。

尽管现有研究取得了显著进展,但仍存在两方面的问题。

1)如何准确地提取词元级与话语级信息。现有方法要么侧重于词元级的表示,通过序列标注或局部注意力机制建模词间关系,但难以有效捕捉跨话语、跨发言者的长距离依赖;要么采用整体对话编码的方式,在一定程度上建模了全局上下文,但容易因对话长度过长导致关键信息被忽视,难以准确提取词元级和话语级的信息。2)如何有效地融合不同层面的信息 (词元级与话语级)。部分研究仅通过简单拼接或加权求和的方式融合多层特征,忽略了词元级与话语级之间复杂的对齐关系和交互模式,难以充分发挥两个层面信息的互补优势;一些依赖复杂的图结构或多重交互机制的方法,虽然在一定程度上增强了结构建模能力,但往往会引入大量冗余信息,增加模型复杂度,影响训练稳定性和推理效率。本文针对性地提出融合局部多视角语言特征和全局特征的对话情感四元组抽取,旨在解决这些问题并进一步提高基于对话的方面情感四元组抽取模型的性能。

## 3 MVLLF-GF 模型

MVLLF-GF 模型 (Multi-View Local Language Feature and Global Feature)主要由文本编码、多视角语言知识编码器、全局话语编码器、多粒度融合器、四元组解码 5 个模块构成,其框架图如图 2 所示。

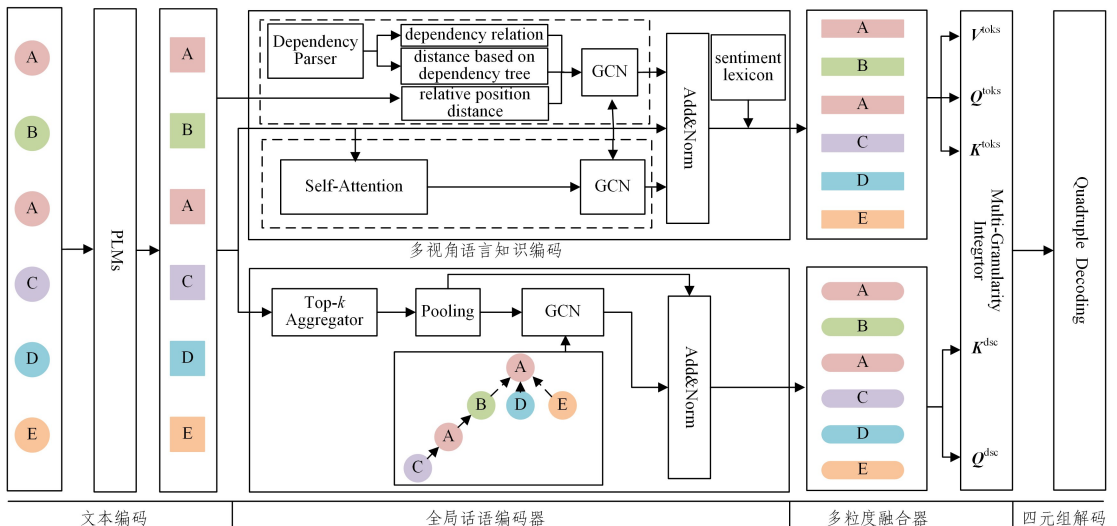


图 2 MVLLF-GF 模型结构

Fig. 2 Structure of MVLLF-GF model

### 3.1 文本编码

本文使用预训练模型(PLMs)对文本特征进行编码。给定一段对话  $D_n = \{u_1, u_2, \dots, u_i\}$ , 对于每句话  $u_i = \{\omega_1, \omega_2, \dots, \omega_i\}$ , 为了获取发言者的信息, 在每句话  $u_i$  后添加发言者的 ID 信息  $x$  (第一个发言者 ID 为 0, 第二个发言者 ID 为 1, 依次向后), 得到  $D_n' = \{u_1', u_2', \dots, u_i'\}$ , 其中  $u_i' = [\text{cls}], u_i, x_i, [\text{cls}]$  是 PLMs 的特殊标记。通过预训练模型, 得到文本的特征表示。

$$\mathbf{H}_n' = \{\mathbf{H}_1', \mathbf{H}_2', \dots, \mathbf{H}_i'\} = \text{PLMs}(D_n') \quad (1)$$

$$\mathbf{H}_i' = \{h_i^{\text{cls}}, h_i^u, h_i^x\} \quad (2)$$

### 3.2 多视角语言知识编码

#### 3.2.1 语言知识增强

DiaASQ 任务需要从对话中抽取方面词和观点词, 要求模型对词元级别的知识与信息进行深入理解。因此, 本文方法通过图神经网络学习多视角语言知识, 提高模型对对话中的每句话的理解能力。

依存句法树被证明可以有效地建立方面词和观点词之间的依存关系。因此, 本文使用句法分析工具<sup>[14]</sup>对每句话进行依存句法分析, 以捕获句子内部的语法关系(如主谓宾结构)。如图 3 所示, Dinner 和 is 存在 nusbj(名词性主语)依存关系, never 和 dinner 之间则不存在依存关系。如果两个词之间存在依存关系, 则依存邻接矩阵为 1, 否则为 0。

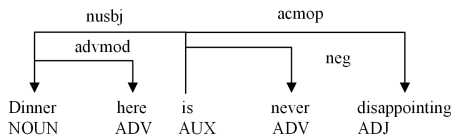


图 3 依存句法分析树

Fig. 3 Dependency parsing tree

$$\mathbf{A}^{\text{syn}} = \begin{cases} 1, & \text{词 } \omega_i \text{ 和 } \omega_j \text{ 存在依存关系} \\ 0, & \text{否则} \end{cases} \quad (3)$$

在自然语言处理的研究中, 句法依存距离<sup>[15]</sup>能够很好地捕捉句子中各个词之间的复杂结构关系。具体来说, 在依存树中如果两个词之间存在句法依存关系, 则它们之间的句法依存距离为 1; 如果它们之间存在多级依存(通过其他词相连), 则句法依存距离为它们在依存树中的最短距离。例如, Dinner 和 here 之间直接相连, 距离为 1, 而 Dinner 和 never 则是通过 is 连接的, 句法依存距离为 2。本文通过句法依存距离来构造句法依存矩阵, 通过依存距离的倒数来表示词语间的关联强度, 距离越近权重越大, 距离越远权重越小, 从而使模型更准确地反映句法结构中词与词的依存紧密程度。

$$\mathbf{A}^{\text{dep}} = \frac{1}{1 + \text{distance}_{\text{dep}}(\omega_i, \omega_j)} \quad (4)$$

其中,  $\text{distance}_{\text{dep}}(\omega_i, \omega_j)$  表示句法依存距离。

在语言知识中, 相对位置距离是指句子中两个词  $\omega_i$  和  $\omega_j$  的相对位置差的绝对值, 即  $|i - j|$ , 是一种重要的语言特征, 用于辅助模型理解句子中词对之间的语义关联和结构关系。本文构造相对位置距离邻接矩阵:

$$\mathbf{A}^{\text{rd}} = \begin{cases} |i - j|, & i \neq j \\ 0, & i = j \end{cases} \quad (5)$$

对于上述构造的 3 种多视角语言特征邻接矩阵, 本文使用图卷积网络建模节点之间的关系。假设图中有  $n$  个节点, 激活函数为  $\sigma$ , 那么第  $l$  层中第  $i$  个节点表示为:

$$\mathbf{h}_i^l = \sigma\left(\sum_{j=1}^n \mathbf{A}_{ij} \mathbf{W}^l \mathbf{h}_j^{l-1} + b^l\right) \quad (6)$$

其中,  $\mathbf{A}_{ij}$  表示邻接矩阵,  $\mathbf{W}^l$  和  $b^l$  分别表示可学习的权重和偏置。将邻接矩阵和文本特征输入 GCN 中得到语言特征输出。

$$\mathbf{H}^{\text{syn}} = \text{GCN}(\mathbf{A}^{\text{syn}}, \mathbf{H}_n') \quad (7)$$

$$\mathbf{H}^{\text{dep}} = \text{GCN}(\mathbf{A}^{\text{dep}}, \mathbf{H}_n') \quad (8)$$

$$\mathbf{H}^{\text{rd}} = \text{GCN}(\mathbf{A}^{\text{rd}}, \mathbf{H}_n') \quad (9)$$

#### 3.2.2 语义知识增强

为进一步加强不同发言之间的语义交互, 提升模型对言语的理解能力, 本文使用自注意力机制<sup>[16]</sup>学习语义知识<sup>[17]</sup>。

$$\mathbf{A}^{\text{sem}} = \text{Atten}(\mathbf{H}_n' \mathbf{W}^Q, \mathbf{H}_n' \mathbf{W}^K) \quad (10)$$

$$\text{Atten}(\mathbf{Q}, \mathbf{K}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right) \quad (11)$$

其中,  $d$  为文本特征  $\mathbf{H}_n$  的维度。与语言特征类似, 语义特征由语义邻接矩阵和文本特征输入 GCN 中得到。

$$\mathbf{H}^{\text{sem}} = \text{GCN}(\mathbf{A}^{\text{sem}}, \mathbf{H}_n') \quad (12)$$

最后, 采用残差连接和归一化<sup>[18]</sup>操作将多视角语言特征、语义特征和文本特征相结合, 得到词元级表示:

$$\mathbf{H}^{\text{tok}} = \text{LN}(\mathbf{H}_n' + \mathbf{H}^{\text{syn}} + \mathbf{H}^{\text{dep}} + \mathbf{H}^{\text{rd}} + \mathbf{H}^{\text{sem}}) \quad (13)$$

#### 3.2.3 情感词典增强

虽然多视角的语言知识和语义知识能够提供句子在语言结构和语义上的特征信息, 但这些特征并不能直接反映词元的情感信息。因此, 本文引入情感词典<sup>[19]</sup>来补充这些情感信息, 对构造的特征予以增强。具体而言, 通过情感词典获得每个词元的情感得分, 其中  $\text{pos\_score}(\omega_i)$  为积极情感得分,  $\text{neg\_score}(\omega_i)$  为消极情感得分, 得分范围为  $[-1, 1]$ , 正数表示积极, 负数表示消极, 0 则表示不在词典中或为中性。本文使用注意力机制, 将情感得分与词元级表示融合, 得到融合词典的词元级表示。

$$s = \begin{cases} \text{pos\_score}(\omega_i), & \text{积极词} \\ 0, & \\ \text{neg\_score}(\omega_i), & \text{消极词} \end{cases} \quad (14)$$

$$\mathbf{A} = \text{sigmoid}(s) \quad (15)$$

$$\mathbf{H}^{\text{toks}} = \mathbf{H}^{\text{tok}} + s \cdot \mathbf{A} \quad (16)$$

### 3.3 全局话语编码

不同于一般文本, 对话文本具有特殊的结构, 多个发言者之间存在回复关系, 同一个人可能有多次发言, 且不同发言之间可能有远距离的依赖关系, 存在回复关系和发言者信息。仅依靠词元表示, 无法很好地捕捉这些结构上的特征。对此, 本文提出了全局话语编码器, 从全局话语层面建模对话的结构化信息(发言者信息、回复关系), 为模型提供跨话语的全局视角。

每句话都是由多个词元构建的,但并不是每个词对 DiaASQ 任务都重要。本文使用 top- $k$  策略<sup>[20]</sup> 选出最关键的  $k$  个词代表整句话,忽略掉一些不重要的词。具体来说,本文对每个词元打分  $s_i^u$ , 选出得分最高的  $k$  的词元进行索引,再进行 Softmax 后加权得到这  $k$  个词元的表示。将其进行最大化化和平均池化,最后将最大化池化特征、平均池化特征以及在文本编码中得到的发言者 ID 信息  $h_i^f$  进行拼接,得到每句话的整体表示  $h_i^u$  以及整个对话表示  $H_n^u$ 。

$$s_i^u = h_i^u \cdot \omega_s + b_s \quad (17)$$

$$\text{id } x_i^u = \text{argmax}(s_i^u, k) \quad (18)$$

$$h_i^k = \text{softmax}(s_i^u [\text{id } x_i^u]) \odot h_i^u [\text{id } x_i^u] \quad (19)$$

$$h_i^p = \text{MLP}(\max(h_i^k) \parallel \text{avg}(h_i^k) \parallel h_i^f) \quad (20)$$

$$H_n^u = \{h_1^p, h_2^p, \dots, h_n^p\} \quad (21)$$

为了明确整个对话的回复结构关系,构建了一个话语邻接矩阵。与多视角语言知识编码器类似,使用 GCN 对话语邻接矩阵进行处理,与对话表示相结合得到话语级表示:

$$A^{\text{disc}} = \begin{cases} 1, & u_i \text{ 与 } u_j \text{ 之间存在回复关系} \\ 0, & \text{否则} \end{cases} \quad (22)$$

$$H^{\text{disc}} = \text{LN}(\text{GCN}(A^{\text{disc}}, H_n^u) + H_n^u) \quad (23)$$

### 3.4 多粒度融合器

通过多视角语言知识编码器获得词元级别的表示,通过全局话语编码器获得话语级别的表示。为了让局部词元和全局话语之间进行信息交互,本文设计了多粒度融合器,为四元组抽取提供全面的表示,如图 4 所示。

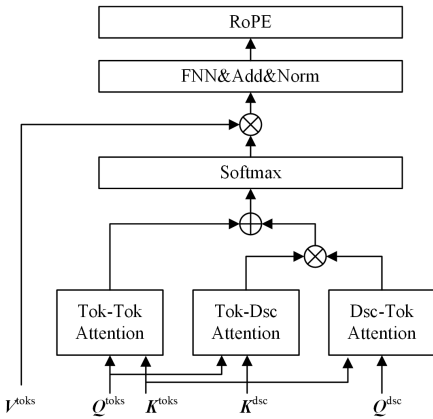


图 4 多粒度融合器

Fig. 4 Multi-granularity interactor

具体来说,构建 3 个注意力得分矩阵:词元-词元注意力得分矩阵 (Tok-Tok Attention)、词元-话语注意力得分矩阵 (Tok-Dsc Attention),以及话语-词元注意力得分矩阵 (Dsc-Tok Attention)。

$$S^{\text{toks-toks}} = \text{Atten}(W_1 Q^{\text{toks}}, W_2 K^{\text{toks}}) \quad (24)$$

$$S^{\text{toks-dsc}} = \text{Atten}(W_1 Q^{\text{toks}}, W_3 K^{\text{dsc}}) \quad (25)$$

$$S^{\text{dsc-toks}} = \text{Atten}(W_4 Q^{\text{dsc}}, W_2 K^{\text{toks}}) \quad (26)$$

其中,  $Q^{\text{toks}}$ ,  $K^{\text{toks}}$  和  $V^{\text{toks}}$  为多视角语言知识编码器的输出,  $Q^{\text{dsc}}$  和  $K^{\text{dsc}}$  为全局话语编码器的输出。本文将 3 种注意力矩阵得分通过式 (27) 进行融合:

$$S = S^{\text{toks-toks}} + S^{\text{toks-dsc}} \cdot S^{\text{dsc-toks}} \quad (27)$$

再将其进行归一化,与  $V^{\text{toks}}$  相乘,得到多粒度融合表示  $H$ 。

$$H = \text{softmax}(S) \cdot V^{\text{toks}} \quad (28)$$

最后,将其通过前馈网络 (FNN)<sup>[21]</sup>、残差连接以及层归一化后,得到最终的特征输出。

$$H_i^f = \text{LN}(\text{FNN}(H) + H) = \{h_1^f, h_2^f, \dots, h_n^f\} \quad (29)$$

### 3.5 四元组解码

受 Li 等的启发,本文使用网格标注的方法来进行解码,如图 5 所示,将四元组预测分解成 3 个子任务。

1) 实体边界标注:确定每个实体(目标 Target、方面词 Aspect、观点词 Opinion)的起止边界。

2) 实体对关系标注:确定实体之间的对应关系(目标-方面词,目标-观点词,方面词-观点词)。

3) 情感极性标注:确定每个四元组对应的情感极性(积极,消极,中性)。

	iPhone	loses	power	very	fast	
	tgt	h2h	t2t			iPhone loses power very fast
		asp	asp	h2h		
			asp		t2t/neg	
				Opi	Opi	
					opi	

图 5 四元组网格标注解码

Fig. 5 Tagging scheme for quadruple extraction

具体来说,创建一个  $n \times n$  的矩阵  $M$ , 矩阵的行和列都对应该文本中的词元。矩阵中的每个单元格  $M_{i,j}$  (其中  $i, j$  表示行和列索引,取值为  $1 \sim n$ ) 用于存储特定的标注信息,用来表示第  $i$  个词和第  $j$  个词之间的语义关系。

实体边界标注:对于目标 (tgt), 遍历文本中的词对。如果词  $w_i$  是目标实体的起始词,词  $w_j$  是目标实体的结束词,那么在矩阵单元格  $M_{i,j}$  中标注“tgt”;方面词 (asp) 和观点词 (opi) 与之类似,分别标注为“asp”“opi”。

实体对关系标注:如果目标实体头部词  $w_{t-\text{head}}$  和方面词实体头部词  $w_{a-\text{head}}$  存在关系,那么在单元格  $M_{t-\text{head}, a-\text{head}}$  中标注“h2h”。与“h2h”类似,使用“t2t”表示目标实体尾部词  $w_{t-\text{tail}}$  和方面词实体尾部词  $w_{a-\text{tail}}$  存在关系。

情感极性标注:在目标实体和观点实体对应的单元格区域内标注情感极性。如果情感为积极,在相应单元格标注“pos”;若为消极,则标注“neg”;若为中性或其他,则标注“neu”。

本文使用多层感知机<sup>[22]</sup> (MLP) 进行标签预测。

$$v_i^f = \text{MLP}(h_i^f) \quad (30)$$

其中,  $\gamma \in \{y^{\text{ent}} \cup y^{\text{pair}} \cup y^{\text{pol}}\}$  包含的 3 类标签分别为实体边界标签  $y^{\text{ent}}$  (目标词 T, 方面词 A, 观点词 B), 实体对关系标签  $y^{\text{pair}}$  (h2h, t2t), 情感极性标签  $y^{\text{pol}}$  (积极, 消极, 中性)。

在自然语言处理中,序列的顺序(对话中的前后话语)

至关重要。传统的位置嵌入(如 Transformer 的正弦和余弦编码)通过固定向量表示位置。但在处理长序列或复杂结构(如对话中的多轮回复)时,无法有效建模位置之间的相对关系。为了增强模型对序列中相对位置信息的感知能力,受 Su 等启发,本文使用旋转位置嵌入(Rotary Position Embedding, RoPE)。与传统的绝对位置编码不同, RoPE 引入旋转矩阵,对嵌入向量进行位置相关的旋转变换,有效捕捉词元之间的相对位置关系,从而提升模型在建模上下文依赖,特别是长距离依赖方面的表现。这对于对话场景中跨语句的结构理解尤为关键。另一方面, RoPE 编码的位置信息与多粒度融合器输出的词元-话语交互特征协同作用,使解码阶段能同时利用局部语义信息与全局对话结构,建立更加一致和连贯的语义关联。

$$\mathbf{u}_i^r = R(\theta, i) \mathbf{v}_i^r \quad (31)$$

其中,  $\theta$  是一个可以学习的参数,  $i$  是位置索引。

对于每个子任务,模型会对最终的第  $i$  个词元表示和第  $j$  个词元表示进行打分(即是否存在某种关系),之后使用 softmax 得到每类标签的概率。

$$\mathbf{S}_{ij}^r = (\mathbf{u}_i^r)^\top \mathbf{u}_j^r \quad (32)$$

$$P_{ij}^r = \text{softmax}(\mathbf{S}_{ij}^r) \quad (33)$$

## 4 实验及结果分析

### 4.1 数据集

本文在中文数据集 ZH 和英文数据集 EN 上评估提出的模型。这两个数据集均包括每个对话语句的发言者与回复关系信息。每个数据集包含 1000 条与电子产品评论相关的对话,平均每个对话包含 7 个语句和 5 位发言者。具体数据如表 2 所列,其中 T-A 表示目标词-方面词匹配个数, T-O 表示目标词-观点词个数, A-O 表示方面词-观点词个数, Total 为所有四元组的总个数。

表 2 数据集统计信息

Table 2 Dataset statistics

数据集	划分	元素对			
		T-A	T-O	A-O	Total
EN	Train	4699	5931	3989	4414
	Dev	60	750	509	555
	Test	592	751	496	545
ZH	Train	4823	6062	4297	4607
	Dev	621	758	538	577
	Test	597	767	523	558

### 4.2 实验参数设置

在对比实验中,本文使用 RobertaLarge<sup>[23]</sup> 和 Chinese-Roberta-wmext-base<sup>[24]</sup> 作英文数据集(EN)和中文数据集(ZH)的预训练模型。Top- $k$  聚合器  $\lambda$  分别设置为 0.5 和 0.8(有利于减少噪声,同时筛选出重要词元并保持语义信息完整性)。Epochs 设置为 35, batch-size 和 dropout 分别设置为 2 和 0.1, 学习率设置为  $1 \times 10^{-4}$ 。本实验设备为 NVIDIA GeForce RTX3090, 实验环境选择 PyTorch。

### 4.3 实验评价指标

在本实验中,由于任务具有复杂度和多维度特性,为了从不同角度全面评估模型在四元组抽取任务中表现,采用了 Micro F1 和 Identification F1 两个评估指标。前者用于衡量模型对整个四元组(包含情感极性)的预测准确度,后者则不区分情感极性,主要关注模型对四元组中的其他元素(目标、方面、观点)的识别能力。此外,本实验还采用 F1 分数评估实体检测和配对检测,包括目标(T)、方面(A)、观点(O)、目标-方面(T-A)、目标-观点(T-O)和方面-观点(A-O)。

### 4.4 基线模型

为了验证模型的有效性,将其与以下基线模型进行对比。

Extract-Classify<sup>[25]</sup>:提出了一种基于 Bert 的双阶段模型,结合了序列标注与条件随机场技术。

SpERT<sup>[26]</sup>:利用 Bert 预训练模型进行基于跨度的联合实体和关系抽取。

ParaPhrase<sup>[27]</sup>:将方面情感四元组预测任务转换为释义生成问题,输入层融入对话分隔符与说话人标识,重构生成模板包含四元组。

Span-ASTE<sup>[28]</sup>:通过考虑目标与观点整体跨度间的交互作用来提升情感一致性,采用双通道跨度剪枝策略优化计算效率;新增“目标跨度检测”模块,检测范围扩展至对话覆盖跨话语要素,将“方面-观点”二元交互升级为“目标-方面-观点”三元交互。

Overall-QPN<sup>[6]</sup>:提出了一种基于整体建模的方法,在编码阶段利用预训练语言模型对整个对话进行编码,以此改进对话层面基于方面的情感四元组分析,通过分段编码+上下文衔接机制保留长对话细节,补充“目标-方面”关联建模层。

IFusionQuad<sup>[29]</sup>:设计了 CloBlock 模块、Biaffine 注意力机制和 InteractiveNet Fusion 模块,融合了对话线程、发言者和回复关系等上下文特征,以增强对四元组抽取的精度与鲁棒性。

H2DT<sup>[12]</sup>:通过构建融合回复关系与发言者角色的异构图,以及引入三元交互机制,从整体对话语境中精确建模各项实体与关系。

DMCA<sup>[30]</sup>:通过构建多尺度滑动窗口与动态层级聚合模块(DHA),高效整合不同语境粒度下的信息,从而提升跨话轮抽取的准确性与鲁棒性。

MDFI<sup>[10]</sup>:通过多维交互网络结构整合发言者、回复关系与话题线程特征,结合双层 ResNet 和反向旋转位置编码,有效提升对话中跨句情感四元组的提取精度。

MVQPN<sup>[3]</sup>:提出了基于网格标注(Grid-tagging)的端到端神经模型,结合多视角交互、自适应位置编码(RoPE)等机制,显著提升跨话轮情感抽取的能力。

### 4.5 对比实验

MVLLF-GF 模型在 ZH 和 EN 两个数据集上的 F1 对比结果如表 3(未使用外部知识模型,仅利用对话本身含有的信息)和表 4(使用外部知识模型,如知识图谱、情感词典)所列。

表3 含外部知识的对比实验结果

Table 3 Comparative experimental results with external knowledge

数据集	模型	Span Match			Pair Extraction			Quadruple	
		T	A	O	T-A	T-O	A-O	Mico	Icen
		(%)							
ZH	Extract-Classify	91.11	75.24	50.06	32.47	26.78	18.90	8.81	9.25
	SpERT	90.69	76.81	54.06	38.05	31.28	21.89	13.00	14.19
	ParaPhrase	—	—	—	37.81	34.42	27.76	23.27	27.98
	Span-ASTE	—	—	—	44.13	34.46	32.21	27.42	30.85
	Overall-QPN	—	—	—	52.86	50.98	53.33	37.77	43.56
	DMCA	<b>92.03</b>	77.07	60.27	56.88	51.70	52.28	42.68	45.36
	MDFI	92.02	74.95	60.20	57.08	<b>53.99</b>	52.27	42.36	45.12
	MVQPN	90.23	76.94	59.35	48.61	43.31	45.44	34.94	37.51
	MVLLF-GF	91.33	<b>78.30</b>	59.87	<b>58.24</b>	51.32	<b>56.30</b>	<b>44.07</b>	<b>47.60</b>
EN	Extract-Classify	88.31	71.71	47.90	34.31	20.94	19.21	11.59	12.80
	SpERT	87.82	74.65	54.17	28.33	21.39	23.64	13.07	13.38
	ParaPhrase	—	—	—	37.22	32.19	30.78	24.54	26.76
	Span-ASTE	—	—	—	42.19	30.44	45.90	26.99	28.34
	Overall-QPN	—	—	—	50.70	49.46	50.31	35.37	39.73
	DMCA	88.11	73.95	63.47	53.08	50.99	52.40	37.96	41.00
	MDFI	88.93	74.15	62.20	52.88	51.04	52.27	37.35	40.48
	MVQPN	88.62	74.71	60.22	47.91	45.58	44.27	33.31	36.80
	MVLLF-GF	<b>89.05</b>	<b>74.74</b>	<b>64.12</b>	<b>54.35</b>	<b>54.13</b>	<b>52.35</b>	<b>39.81</b>	<b>42.58</b>

注:加粗表示最优结果;—表示未进行相关实验。

表4 未含外部知识的对比实验结果

Table 4 Comparative experimental results without external knowledge

数据集	模型	Span Match			Pair Extraction			Quadruple	
		T	A	O	T-A	T-O	A-O	Mico	Icen
		(%)							
ZH	IFusionQuad	91.69	75.90	60.96	54.68	51.81	50.04	41.53	44.56
	H2DT	<b>91.72</b>	76.93	<b>61.87</b>	50.48	48.80	52.40	40.34	42.81
	MVLLF-GF	91.33	<b>78.30</b>	59.87	<b>58.24</b>	51.32	<b>56.30</b>	<b>44.07</b>	<b>47.60</b>
EN	IFusionQuad	88.31	74.23	63.48	52.65	51.82	51.94	35.96	41.49
	H2DT	88.69	73.81	62.61	48.69	48.84	<b>52.47</b>	39.01	42.19
	MVLLF-GF	<b>89.05</b>	<b>74.74</b>	<b>64.12</b>	<b>54.35</b>	<b>54.13</b>	52.35	<b>39.81</b>	<b>42.58</b>

注:加粗表示最优数据。

实验结果表明:相较于 MVQPN,本文模型在 ZH 和 EN 数据集上,F1 分别提高了 9.13 个百分点和 6.5 个百分点。传统方法(如 MVQPN)虽然能捕捉局部依赖,但对长距离信息的对齐能力不足,容易导致跨句情感要素关联失败。而 MVLLF-GF 通过多视角语言知识编码器(结合句法依存、相对位置、语义特征等),增强了词元级的理解能力,能更准确地建立四元组之间的联系。同时,MVLLF-GF 也兼顾了局部词元的信息和全局话语的信息。词元级基线模型(如 SpERT, Span-ASTE)忽略了全局对话结构,只进行整体对话编码(如 Overall-QPN)则容易因输入话语过长而丢失细节,MVLLF-GF 通过引入全局话语编码器,利用发言者信息、回复关系等全局信息,既能把握局部语义精细差别,又能理解全局对话脉络,避免了单一层面建模的局限。相比于 MDFI, H2DT 等依赖图结构的模型,MVLLF-GF 设计了多粒度融合器,实现了词元级信息与话语级信息之间的高效融合,而不是简单拼接或粗糙的特征叠加。这种融合让模型在面对复杂对话(多发言者、多分支回复)时,能更好地区分不同四元组之间的对应关系,从而提升了模型整体的鲁棒性和泛化能力。

此外,与 MVQPN 相比,本文方法在实体检测和配对检测方面均有不同程度的提升。本文模型在两个数据集上相较于其他基线模型均有一定程度的提升,说明通过引入多视角语言特征编码器和全局话语编码器,分别从局部学习词元信息和从全局理解回复关系信息,以及引入多粒度融合器更好地整合两个层面的信息,丰富了模型,从不同层面学习对话文

本的信息,能够有效地提高基于对话的方面情感四元组抽取的准确度。

#### 4.6 消融实验

为了验证不同模块对模型的积极影响,本文进行了消融实验,结果如表 5 所列。

表5 消融实验结果

Table 5 Results of ablation study

模型	ZH		EN	
	Micro F1	Icen F1	Micro F1	Icen F1
	(%)			
Ours	44.07	47.60	39.81	42.58
W/O MLK	41.97	44.58	38.16	41.67
W/O LKE	42.96	46.35	38.66	42.11
W/O SKE	42.52	45.14	38.92	41.17
W/O SL	43.79	47.14	38.51	41.36
W/O GD	43.55	46.72	38.85	41.97
W/O RBR	42.87	45.82	37.93	41.46
W/O SI	42.06	44.84	37.57	40.82
W/O MGI	43.28	46.79	38.24	40.77

多视角语言知识编码层(MLK)通过图神经网络学习局部词元之间的结构关系,并捕捉句法依存关系、语义关系等上下文信息。去除该模块后,模型无法有效理解词元之间的依存关系,特别是在长距离依赖和跨句情感要素提取上,导致上下文信息理解不足,从而影响四元组的提取精度。此外,分别对其中的子模块进行消融实验。多视角语言知识模块(LKE)利用多视角语言知识从不同角度捕获对话中的上下文关系,通过 GCN 对词元之间的结构关系进行建模。去掉该模块

后,模型无法从句法层面和结构层面捕捉方面词与观点词之间的关联,导致对局部信息的理解不足。语义知识增强模块(SKE)通过自注意力机制建模不同发言之间的语义交互,强化了模型的语义理解能力。去除该模块后,模型缺乏对深层语义关系的建模能力,无法有效对齐跨语句的方面词与观点词,从而影响了四元组抽取的精度。情感词典模块(SL)为模型提供了先验的情感极性知识,有助于模型更准确地判断情感极性。若去掉该模块,模型只能依赖上下文推断情感信息,容易造成情感极性预测错误。

全局话语编码层(GD)中编码器的作用是建模发言者信息与回复关系,捕捉跨语句的全局信息。如果去掉该模块,那么模型就无法学习对话层面的整体脉络,尤其在多发言者、多轮次的对话中,无法准确地理解对话回复关系,导致模型对全局依赖的建模能力下降,从而降低了四元组抽取的整体性能。此外,去除其中的回复关系模块(RBR)和发言者信

息模块(SI)后,结果均有不同程度的下降。回复关系模块模型用于建模发言者之间的回复关系,帮助模型理解不同话语间的互动关系。去除该模块,模型将无法准确理解对话的语境关系,导致情感四元组的提取失误。发言者信息模块则提供了每个发言者的ID信息,在多轮对话中尤其重要。去掉该模块,模型将失去对发言者身份和观点来源的理解,无法有效区分不同发言者之间的情感表达,进而影响情感四元组的准确抽取。

去除多粒度融合器(MGI)后,模型无法有效地整合词元级和话语级的信息,导致局部和全局信息的融合不充分,进而影响对复杂对话中四元组的识别与提取,导致整体性能下降。

#### 4.7 案例分析

为了进一步验证 MVLLF-GF 模型的有效性,从 EN 数据集中选取一段对话进行案例分析。MVLLF-GF 与 MVQPQN 的对比结果如表 6 所列。

表 6 案例分析  
Table 6 Case study

对话	正确的四元组(T,A,O,S)	MVQPQN	MVLLF-GF
“[ Angry][Angry][Angry ] The battery of Redmi Note10P is 5000. It is not durable.” “The key is that the heat is too hot” “The heat feels okay,the built-in measurement is about 35 degrees” “Now it may be fine in winter. At that time,I bought it in summer. It was hot when I did anything. It was a few degrees higher than my K30S.” “Tianji 1100 chip is too consuming...” “[ DOGE ] [ Doge ] [ Doge ] Then I am still stupid to wait for the heaven 9000 · I think it is the problem of shrinkage of the Xiaomi battery”	( Redmi Note10P, battery, not durable,neg) ( Redmi Note10P,heat, too hot,neg) ( Xiaomi, battery, shrinkage,neg) ( Tianji 1100,“,too consuming,neg)	( Redmi Note10P, battery, not durable,neg)✓ ( Redmi Note10P,heat, too hot, pos)× ( Xiaomi, battery, shrinkage,neg)✓ ( Redmi Note10P, heat, okay, pos)× ( Redmi Note10P,battery,5000,neg)×	( Redmi Note10P, battery, not durable,neg)✓ ( Redmi Note10P,heat, too hot,neg)✓ ( Xiaomi, battery, shrinkage,neg)✓ ( Xiaomi, battery, shrinkage,neg)✓ ( Redmi Note10P, heat, okay,neg)✓ ( Redmi Note10P, heat, okay, other)×

从表 6 中可以看出,MVLLF-GF 模型相较于 MVPQN 在四元组抽取上展现出更高的准确性与鲁棒性。具体而言,MVPQN 虽能捕捉到部分四元组,但由于其主要依赖序列位置与局部交互建模,容易在多实体和长距离依赖场景下产生错误情感极性,将负向的“too hot”判断为积极 pos;MVLLF-GF 融合了多视角语言特征(依存句法、位置关系及情感先验知识),解决了这个问题,能够准确地判断情感极性。此外,MVPQN 生成无效四元组,如“battery-5000”;MVLLF-GF 通过深度融合句法与语义特征增强了对复杂语义结构的建模能力,并借助全局话语级编码与多粒度融合器提升了跨句一致性,有效避免了 MVPQN 出现的噪声。综合而言,MVLLF-GF 在局部精细建模与全局关系上具有双重优势,能够解决长距离依赖和对话复杂的难题,有效提高基于对话情感四元组抽取的精度。

**结束语** 本文提出了融合局部多视角语言特征和全局特征的对话情感四元组抽取模型。从两个层面理解对话信息:一方面,利用图神经网络学习多视角的语言知识和语义知识局部词元级别特征,获取每个句子的信息;另一方面,通过全局话语编码器,捕获发言者信息和回复关系全局话语特征,学习整段对话的信息。为了融合这两个层面的信息,本文使用多粒度融合器,有效整合了局部信息和全局信息表示,提升了模型的对话理解能力。最后使用网格标注进行四元组解码,从而有效提升了基于对话的方面情感四元组抽取准确度。

然而,本研究仍存在一些不足。首先,尽管多视角语言知识编码器能够有效捕获局部依赖,但其对超长对话中的全局一致性建模仍有限。其次,情感词典的引入在一定程度上提升了情感极性的判别能力,但词典方法依赖先验知识,难以覆盖多领域、多语境的情感表达,存在一定的泛化局限。最后,模型的复杂度相对较高,在实际应用中,推理效率和计算成本仍需进一步优化。

未来的工作将从以下方面展开研究:1)探索更高效的长序列建模机制,例如引入稀疏注意力或层级记忆结构,以提升模型对超长对话的处理能力;2)结合大规模预训练语言模型的生成式能力,将四元组抽取与生成任务进行联合建模,增强模型的通用性与可扩展性;3)引入迁移学习策略,提升模型在多语言、多领域的对话数据上的鲁棒性与泛化能力。

#### 参考文献

- [1] PAPAGEORGIOU H, ANDROUTSOPOULOS I, GALANIS D, et al. Semeval-2015 task 12: aspect based sentiment analysis [C]//Proceedings of the 9th International Workshop on Semantic Evaluation. Piscataway, NJ: IEEE, 2015: 486-495.
- [2] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[J]. arXiv:1609.02907, 2016.
- [3] LI B, FEI H, LI F, et al. Diaasq: A benchmark of conversational aspect-based sentiment quadruple analysis [J]. arXiv: 2211.

- 05705, 2022.
- [4] WU Z, YING C, ZHAO F, et al. Grid tagging scheme for aspect-oriented fine-grained opinion extraction[J]. arXiv:2010.04640, 2020.
- [5] SU J, AHMED M, LU Y, et al. Roformer: Enhanced transformer with rotary position embedding[J]. Neurocomputing, 2024, 568: 127063.
- [6] CAI C, ZHAO Q, XU R, et al. Improving Conversational Aspect-Based Sentiment Quadruple Analysis with Overall Modeling[C]//CCF International Conference on Natural Language Processing and Chinese Computing. Cham: Springer, 2023: 149-161.
- [7] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019:4171-4186.
- [8] LAI Y, FAN S, TONG Z, et al. Conversational aspect-based sentiment quadruple analysis with consecutive multi-view interaction[C]//CCF International Conference on Natural Language Processing and Chinese Computing. Cham: Springer, 2023: 162-173.
- [9] ANGUITA D, GHELARDONI L, GHIO A, et al. The 'K' in K-fold Cross Validation[C]//ESANN. 2012:441-446.
- [10] ZHAO Z, ZHANG L, ZHENG Q, et al. Multi-dimensional feature interaction for Conversational Aspect-Based Quadruple Sentiment Analysis[J]. Neural Processing Letters, 2025, 57(1): 9.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
- [12] LI B, FEI H, LIAO L, et al. Harnessing holistic discourse features and triadic interaction for sentiment quadruple extraction in dialogues[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2024:18462-18470.
- [13] WANG X, JI H, SHI C, et al. Heterogeneous graph attention network[C]//The World Wide Web Conference. 2019:2022-2032.
- [14] SUN K, ZHANG R, MENSAH S, et al. Aspect-level sentiment analysis via convolution over dependency tree[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2019:5679-5688.
- [15] LIU H, XU C, LIANG J. Dependency distance: A new perspective on syntactic patterns in natural languages[J]. Physics of Life Reviews, 2017, 21: 171-193.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 6000-6010.
- [17] CHANG X W, DUAN L G, CHEN J H, et al. A fragment-level extraction method for sentiment triples based on deep fusion of syntactic and semantic features [J]. Computer Science, 2026, 53(2): 322-330.
- [18] BA J L, KIROS J R, HINTON G E. Layer normalization[J]. arXiv:1607.06450, 2016.
- [19] MANNING C D, SURDEANU M, BAUER J, et al. The Stanford CoreNLP natural language processing toolkit [C] // Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics; System Demonstrations. 2014:55-60.
- [20] CHEN M, BEUTEL A, COVINGTON P, et al. Top-k off-policy correction for a REINFORCE recommender system[C]// Proceedings of the twelfth ACM International Conference on Web Search and Data Mining. 2019:456-464.
- [21] BEBIS G, GEORGIOPOULOS M. Feed-forward neural networks[J]. IEEE Potentials, 2002, 13(4): 27-31.
- [22] TOLSTIKHIN I O, HOULSBY N, KOLESNIKOV A, et al. Mlp-mixer: An all-mlp architecture for vision[J]. Advances in Neural Information Processing Systems, 2021, 34: 24261-24272.
- [23] LIU Y, OTT M, GOYAL N, et al. Roberta: A robustly optimized bert pretraining approach[J]. arXiv:1907.11692, 2019.
- [24] CUI Y, CHE W, LIU T, et al. Pre-training with whole word masking for chinese bert[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021, 29: 3504-3514.
- [25] CAI H, XIA R, YU J. Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions[C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. 2021:340-350.
- [26] EBERTS M, ULGES A. Span-based joint entity and relation extraction with transformer pre-training[M]. IOS Press, 2020: 2006-2013.
- [27] ZHANG W, DENG Y, LI X, et al. Aspect sentiment quad prediction as paraphrase generation[J]. arXiv:2110.00796, 2021.
- [28] XU L, CHIA Y K, BING L. Learning span-level interactions for aspect sentiment triplet extraction[J]. arXiv:2107.12214, 2021.
- [29] JIANG H, CHEN X, MIAO D, et al. IFusionQuad: A novel framework for improved aspect-based sentiment quadruple analysis in dialogue contexts with advanced feature integration and contextual CloBlock [J]. Expert Systems with Applications, 2025, 261: 125556.
- [30] LI Y, ZHANG W, LI B, et al. Dynamic multi-scale context aggregation for conversational aspect-based sentiment quadruple analysis[C]//ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2024: 11241-11245.



**PENG Juhong**, born in 1978, Ph.D, associate professor. Her main research interests include signal processing and artificial intelligence methods.



**ZHAO Mingjun**, born in 1974, Ph.D, lecturer. His main research interests include intelligent learning and deep learning.