doi：10. 12068/j. issn. 1005－3026. 2026. 20260034

基于双图神经网络的对话情感四元组抽取方法

张 引，吕久鸿，赵玉丽
（东北大学 软件学院，辽宁 沈阳 110169）

摘 要：针对情感要素跨话语分散、长距离依赖突出以及结构信息利用不足的问题，提出1种基于双图神经网络的对话情感四元组抽取方法．该方法采用1套层级框架，其由上下文表示、结构感知交互、多视图特征融合和结构化解码推理四部分构成．上下文表示层结合预训练语言模型的多层语义信息与话语距离编码，增强对对话上下文的表征能力；结构感知交互层联合建模说话者关系、话语距离关系和句法依存关系，以刻画跨话语关联与句内结构约束；多视图特征融合层通过协同注意力与门控残差机制整合序列语义特征和图结构特征；结构化解码推理层通过候选实体筛选与关系推断完成情感四元组识别．在DiaASQ数据集上的实验结果表明，该方法能够有效提升复杂对话场景下情感四元组抽取性能，整体表现优于对比基线模型．

关 键 词：情感四元组抽取；双图神经网络；多视图特征融合；结构化解码；跨话语依赖
中图分类号：TP 391.1 **文献标志码**：A **文章编号**：1005-3026(2026)07-0162-11

Dialogue Sentiment Quadruple Extraction Method Based on Dual-Graph Neural Networks

ZHANG Yin, *LYU Jiu-hong*, *ZHAO Yu-li*
(School of Software, Northeastern University, Shenyang 110169, China. Corresponding author: ZHANG Yin, E-mail: zhangyin@mail.neu.edu.cn)

Abstract: To address the problems of cross-utterance dispersion of sentiment elements, prominent long-distance dependencies, and insufficient utilization of structural information, a dialogue sentiment quadruple extraction method was proposed based on dual-graph neural networks. The proposed method employs a hierarchical framework consisting of four modules: context representation, structure-aware interaction, multi-view feature fusion, and structured decoding inference. In the context representation layer, multi-level semantic information derived from a pre-trained language model was combined with utterance distance encoding to enhance dialogue context representation. The structure-aware interaction layer jointly modeled speaker relations, utterance distance relations, and syntactic dependency relations to capture cross-utterance associations and intra-sentence structural constraints. Sequential semantic features and graph structural features were integrated through co-attention and gated residual mechanisms in the multi-view feature fusion layer. Sentiment quadruple identification was accomplished through candidate entity filtering and relation inference in the structured decoding inference layer. Experimental results on the DiaASQ dataset demonstrate that the proposed method effectively improves sentiment quadruple extraction performance in complex dialogue scenarios and outperforms baseline models.

Key words: sentiment quadruple extraction; dual-graph neural network; multi-view feature fusion; structured decoding; cross-utterance dependency

随着社交媒体与智能客服系统的发展，文本方面级情感分析的研究对象逐渐由单句评论扩展至多轮对话．多轮对话中说话人交替频繁、上下文跨度较大，目标实体、评价方面、观点词与情感极性等要素常分散于不同话轮，产生明显的跨话语依赖．传统方面级情感分析方法主要面向单

收稿日期：2026-03-24
基金项目：辽宁省自然科学基金资助项目(2022-MS-124)；中央高校基本科研业务费专项资金资助项目(N25LPY032)．
作者简介：张 引(1985—)，男，辽宁沈阳人，东北大学副教授．
通信作者：张 引，E-mail：zhangyin@mail.neu.edu.cn.

句语境,通常依赖局部上下文完成方面词与观点表达的匹配,难以处理对话场景中由说话人交互、回复关系和长距离上下文共同造成的关系断裂问题.对话式方面级情感四元组分析(conversational aspect-based sentiment quadruple analysis, DiaASQ)正是在这一背景下提出,其目标是从多轮对话中抽取由目标实体、评价方面、观点词和情感极性构成的结构化四元组[1].

现有相关研究可分为三类:第一类是面向对话情绪识别的粗粒度方法,如DialogueGCN将图卷积网络引入对话情绪识别任务,通过建模话语间关系增强上下文推理能力[2];DialogueCRN从认知推理角度刻画情绪状态的动态变化[3].这类方法证明了对话结构信息对情感理解的重要性,但输出通常是话语级情绪类别,难以描述说话者围绕特定目标实体和评价方面形成的细粒度情感关系.第二类是面向句子级的情感要素抽取方法,如Xu等[4]提出基于位置感知标注的情感三元组抽取方法;Zhang等[5]进一步将任务扩展为评价对象、评价方面、观点词和情感极性的四元组预测框架.这类方法为情感四元组建模提供了基础,但多依赖单句内部的局部语义关系,难以直接适应多轮对话中要素跨话语分布的情况.第三类是面向DiaASQ任务的对话级方法:Li等[1]构建DiaASQ数据集并将四元组抽取扩展到多轮对话场景;Xiao等[6]通过上下文融合编码增强跨话语表示能力;Li等[7]利用动态多尺度上下文聚合整合不同范围的对话线索;Huang等[8]从多粒度表示交互角度提升细粒度情感要素定位能力;Jiang等[9]融合线程、说话人身份和回复关系等特征以改进对话结构建模;Lai等[10]通过连续多视图交互增强跨话语特征传播.

尽管上述方法已取得一定进展,但仍存在以下不足.首先,部分基于序列编码或上下文融合的方法主要依赖预训练语言模型学习隐式语义关联,对说话人关系、话语距离关系和句法依存关系等显式结构约束利用不足.预训练语言模型能够有效捕捉上下文语义,但其注意力分布并不直接等价于任务所需的对话结构关系,在长对话中仍可能出现情感线索扩散和关系匹配偏移.其次,已有图结构方法通常将多类异质关系混合到同一图空间传播,容易造成不同结构信号之间的相互干扰;同时,部分方法更重视对话级宏观关系,对句内依存结构等微观约束建模不足,影响目标实体与观点词之间的精确关联.最后,在解码阶段,长对话会产生较大的候选组合空间,若缺乏有效的候选筛选与高阶关系建模,容易增加错误匹配风险;低阶关系交互方式也难以充分刻画头实体、尾实体与关系标签之间的复杂依赖[9,11-12].

针对上述问题,本文提出1种基于双图神经网络的对话情感四元组抽取方法.该方法并非以图神经网络替代预训练语言模型,而是在预训练语言模型获得上下文语义表示的基础上,引入图神经网络对显式结构信息进行补充建模.本方法分别刻画句内句法依存关系和对话级交互关系,通过多视图融合机制整合序列语义特征和图结构特征,并结合结构化解码推理完成情感四元组识别.在DiaASQ中文子数据集上的实验结果表明,该方法在完整四元组匹配性能上优于对比模型,消融实验进一步验证了不同结构建模模块的有效性.

1　图神经网络技术

图神经网络(graph neural networks, GNN)是一类面向图结构数据的深度学习方法,能够通过节点间的消息传递与邻域聚合学习结构化表示.在自然语言处理中,词语之间的句法依存关系、话语之间的回复关系、说话人交互关系以及话语距离关系均可抽象为图结构,因此GNN可作为序列语义模型之外的结构建模工具.虽然Transformer及其预训练模型已在文本表示学习中取得良好效果,但其优势主要体现在基于自注意力机制的上下文语义建模能力上.对于DiaASQ这类结构化抽取任务而言,模型不仅需要理解语义内容,还需要判断目标实体、评价方面、观点词和情感极性之间是否构成有效组合.当单纯依赖序列语义表示时,模型难以显式区分不同说话人的交互关系、跨话语距离关系以及句内依存约束.GNN的作用在于引入结构归纳偏置,使模型能够沿预定义或学习得到的结构边传播信息,从而提升跨话语关联建模和四元组关系推理能力.

图卷积网络[13](graph convolutional network, GCN)和图注意力网络[14](graph attention network, GAT)是其中具有代表性的两类模型.二者在建模侧重上具有互补性:GCN更适合刻画相对稳定的句法依存结构[15],GAT更适合处理对话中说话人交互、话语距离等动态关系.基于这种互补性,本文在预训练语言模型语义表示的基础上引入GNN,以增强模型对对话结构信息和细粒度情感关系的建模能力.

2 基于双图神经网络的对话情感四元组抽取方法

2.1 问题定义

DiaASQ任务旨在从多轮对话中抽取完整的情感四元组$q=(t,a,o,s)$,其4个元素依次为目标实体(target)、评价方面(aspect)、观点词(opinion)和情感极性(sentiment).与传统单句情感分析不同,DiaASQ面向具有回复结构的多说话人对话场景,四元组要素可能分布于不同话轮,需要结合上下文语义、话语关系和说话人交互信息进行联合判断.因此,该任务并非单句内部的局部识别问题,而是面向多轮对话的结构化抽取问题.

设输入对话经预训练语言模型编码后得到上下文表示矩阵$\boldsymbol{H}\in\mathbf{R}^{N\times d}$,其中$N$表示文本单元数量,$d$表示维度.为刻画对话结构关系,构建图结构$G=(V,E,\boldsymbol{A})$,其中$V$表示节点集合,$E$表示结构连接关系,$\boldsymbol{A}$为邻接矩阵.DiaASQ的目标是学习从语义表示与结构信息到四元组集合的映射函数,该函数如式(1)所示:

$$F_\theta:(\boldsymbol{H},G)\mapsto Q. \tag{1}$$

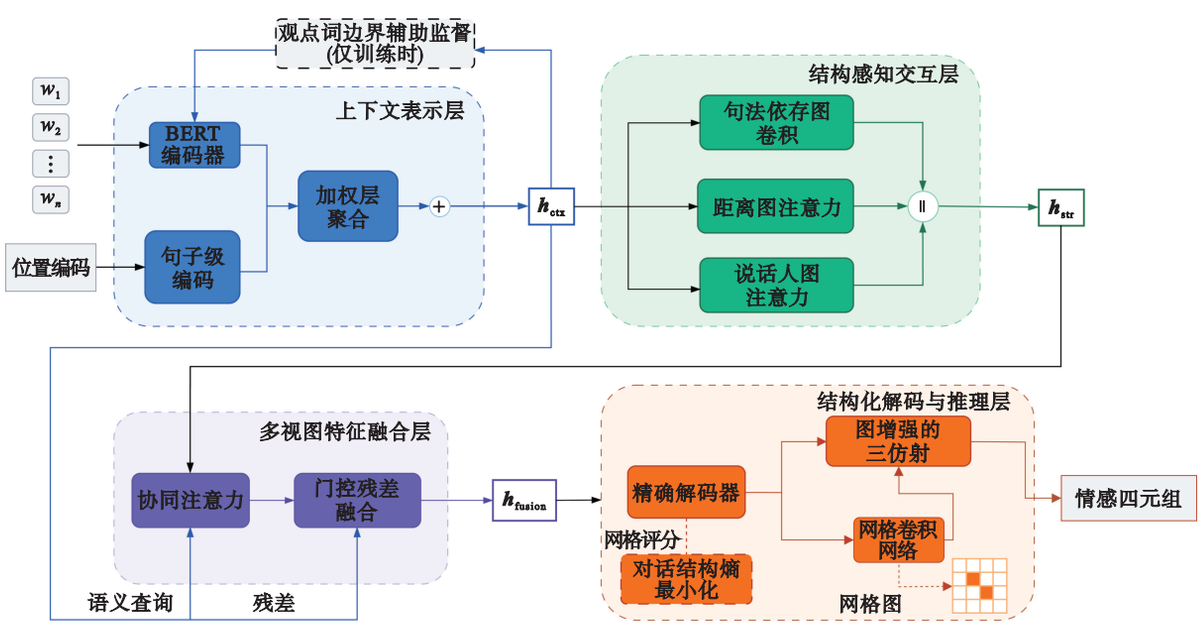
其中:F_θ为从语义表示与结构信息到四元组集合的映射函数,θ为模型参数;$Q=\{(t_k,a_k,o_k,s_k)\}_{k=1}^{K}$表示对话中的情感四元组集合,$K$为四元组数量.该过程通常涉及实体边界识别、实体关系配对和情感极性判定等子任务.

真实对话中,情感要素常分散于不同说话人的回复中,简单线性拼接上下文容易引入噪声并破坏对话逻辑.因此,DiaASQ任务要求模型同时具备全局上下文建模、结构关系建模和细粒度边界识别能力.

2.2 网络架构

本文方法的整体网络结构如图1所示.针对DiaASQ任务中情感要素跨话轮分散、说话人交互复杂以及四元组关系组合空间较大的问题,模型按照信息流处理过程划分为4个层次:上下文表示层(contextual representation layer, CRL)、结构感知交互层(structure-aware interaction layer, SIL)、多视图特征融合层(multi-view feature fusion layer, MVFL)以及结构化解码与推理层(structured decoding and inference layer, SDIL).4个层次依次完成语义编码、结构建模、异质特征融合和情感四元组生成.

CRL以多轮对话文本为输入,利用预训练语言模型提取上下文相关的词级表示,并引入话语位置信息以刻画对话中的顺序关系.考虑到不同层语义表示在边界识别和语义抽象方面具有互补性,本层进一步通过加权层聚合获得上下文表示$\boldsymbol{h}_{\text{ctx}}$,作为后续结构建模与融合的基础.同时,训练阶段从该表示中引出观点词边界辅助监督,用于增强模型对观点词边界的感知能力,但该分支不参与推理阶段的最终解码.



注:BERT(bidirectional encoder representations from Transformer)为基于Transformer的双向编码器表征.

图1 面向DiaASQ任务的双图神经网络整体框架

Fig. 1 Overall framework of dual-graph neural networks for DiaASQ

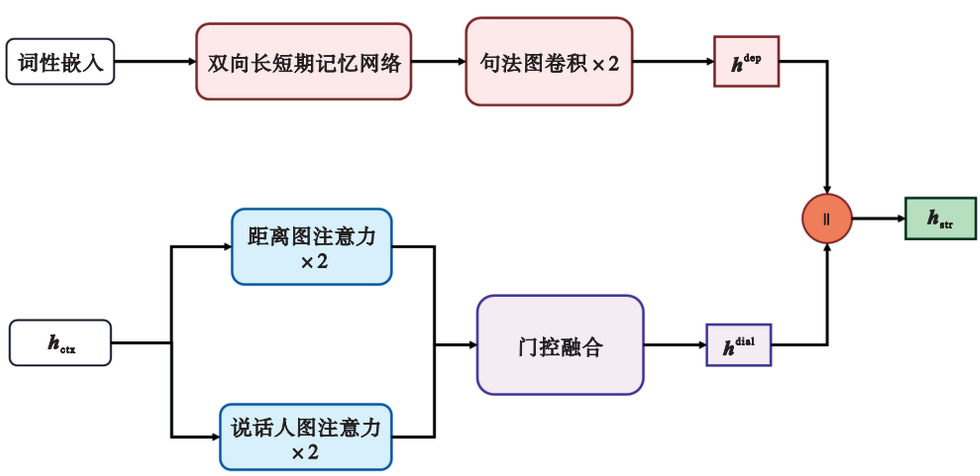
SIL层在$\boldsymbol{h}_{\mathrm{ctx}}$的基础上引入多视角结构信息.该层分别构建句内句法依存分支、话语距离分支和说话人交互分支,其中句法依存分支用于捕捉句内局部依赖关系,话语距离分支用于建模不同话轮之间的位置关联,说话人交互分支用于表示对话参与者之间的交互关系.3个结构分支的输出经过拼接整合,形成结构感知表示$\boldsymbol{h}_{\mathrm{str}}$.由此,模型能够同时利用句内细粒度依赖和对话级交互结构,为跨话题情感要素关联提供结构支撑.

MVFL层用于融合上下文语义表示与结构感知表示.由于不同结构分支所刻画的信息粒度存在差异,直接拼接或叠加容易引入冗余噪声,因此本层采用协同注意力机制,将语义表示为查询,对结构信息进行选择性对齐;随后通过门控残差融合机制保留原始语义信息并整合结构特征,得到融合表示$\boldsymbol{h}_{\mathrm{fusion}}$.该过程既促进了结构信息的有效利用,也缓解了图传播过程中局部边界信息被削弱的问题.

SDIL层基于h_{fusion}完成情感四元组的结构化预测.该层首先通过精确解码器配合网格标注识别候选实体边界,并结合引入的对话结构熵最小化(dialogue structural entropy minimization, DSEM)模块对候选实体空间进行约束.DSEM的名称源于对话结构熵最小化思想[11],其作用是在对话图上划分语义相关的候选子簇,保留簇内依赖并减少跨话语无效配对.随后利用网格卷积神经网络(grid convolutional neural network, Grid CNN)捕捉网格空间中的局部模式,并通过图增强的三仿射(GNN-enhanced triaffine)机制对候选实体之间的高阶关系进行建模,最终输出由目标实体、评价方面、观点词和情感极性组成的情感四元组.整体来看,本方法并非简单堆叠多个功能模块,而是以对话上下文表示为基础,逐步引入句内句法结构、话语距离和说话人交互信息,并通过多视图特征融合与结构化解码完成情感四元组预测.

2.3　对话与句法双图结构建模

DiaASQ任务中的结构依赖同时存在于对话层面和句法层面.目标实体、评价方面与观点词可能分散在不同话语中,其关联受到说话人身份、话语距离和回复上下文的共同影响,属于对话层面的宏观依赖;实体边界、观点词定位以及局部修饰关系则依赖句子内部的依存路径,属于词级层面的微观约束.两类结构信息在节点粒度、连接依据和传播方式上存在差异,若统一纳入单一图空间进行消息传递,则身份一致性、时序邻近性和句法支配关系等不同性质的结构信号容易相互干扰.本文所称的双图神经网络结构是指对话结构图与句法依存图这两个不同粒度的结构空间.其中,对话结构图进一步由说话人视图和距离视图刻画跨话语关联,句法依存图用于建模句内词语之间的依存约束.整体框架如图2所示.



注:Ⅱ表示拼接操作.

图2　对话与句法双图结构建模框架

Fig. 2　Framework of dialogue and syntax dual-graph structure modeling

对话中的跨话语关联并非由单一因素决定.同一说话人的表达通常在多个话轮中保持一定的语义连续性,体现说话人身份的一致性;相邻话语往往具有更紧密的上下文联系,随着话轮间隔增大关联逐渐减弱,体现时序邻近关系.二者对应的结构含义不同,直接在同一拓扑中混合传播会削弱各自的建模作用.为此,本文将对话结构图拆分为说话人视图和距离视图2个子空间.

说话人视图通过邻接矩阵连接同一说话人的话语节点,其定义如式(2)所示:

$$\boldsymbol{A}_{ij}^{\mathrm{spk}}=\mathbb{I}\left[s_i=s_j\right]. \tag{2}$$

其中:$\boldsymbol{A}_{ij}^{\mathrm{spk}}$表示说话人视图邻接矩阵中第$i$个话语节点与第$j$个话语节点之间的连接关系;$s_i$为第$i$个话语节点的说话人标识;$\mathbb{I}[\cdot]$为指示函数.当$s_i=s_j$时,$\boldsymbol{A}_{ij}^{\mathrm{spk}}=1$,表示2个话语节点由同一说话人产生并建立连接;否则$\boldsymbol{A}_{ij}^{\mathrm{spk}}=0$.该矩阵使同一说话人的话语能够跨越时序间隔直接建立连接,从而在保持说话人身份一致性的基础上维持语义连贯.距离视图以指数衰减函数描述话语间的时序邻近关系,其权重的邻接矩阵定义如式(3)所示:

$$\boldsymbol{A}_{ij}^{\mathrm{dis}}=\exp\left(-\frac{|i-j|}{\tau}\right). \tag{3}$$

其中:$\boldsymbol{A}_{ij}^{\mathrm{dis}}$表示距离视图邻接矩阵中第$i$个话语节点与第$j$个话语节点之间的时序邻近性权重;$|i-j|$表示二者在话语序列中的相对距离;$\tau$为控制衰减速度的温度参数.说话人视图和距离视图分别在图注意力框架下完成特征传播,二者的邻域定义相互独立.由于不同对话样本中说话人一致性和时序邻近性的作用程度并不固定,模型进一步采用门控机制自适应调节两类结构表示的贡献,计算过程如式(4)和式(5)所示:

$$g=\sigma\left(\boldsymbol{W}\left[\boldsymbol{h}^{\mathrm{spk}}\|\boldsymbol{h}^{\mathrm{dis}}\right]+b\right), \tag{4}$$

$$\boldsymbol{h}^{\mathrm{dial}}=g\odot\boldsymbol{h}^{\mathrm{spk}}+\left(1-g\right)\odot\boldsymbol{h}^{\mathrm{dis}}. \tag{5}$$

其中:$\boldsymbol{h}^{\mathrm{spk}}$和$\boldsymbol{h}^{\mathrm{dis}}$分别为说话人视图和距离视图的输出;$g$为门控权重;$\boldsymbol{h}^{\mathrm{dial}}$为对话级结构表示,用于刻画跨话语范围内的宏观结构依赖,并在后续阶段与句法分支的输出共同参与结构特征融合;$\boldsymbol{W}$和b分别为门控机制中的可学习的权重矩阵和偏置量;$\sigma(\cdot)$为非线性激活函数;$\odot$为哈达玛积.

对话结构图能够描述话语之间的连接关系,但难以刻画句子内部词语之间的修饰、支配和边界约束.在四元组抽取过程中,目标实体边界、观点词定位以及短语结构判断都依赖句内词级关系.对话文本中常见省略、口语化和不规范表达,实体与修饰词之间也未必在线性序列上相邻,仅依靠上下文窗口难以稳定捕捉此类局部结构关系.为补充句内微观结构信息,本文引入句内句法依存GCN.预训练模型通常采用子词切分,而句法依存树以完整词为节点,首先结合词性嵌入和双向长短期记忆(bidirectional long short-term memory,BiLSTM)网络构造词级表示,再通过2层GCN沿依存边进行特征传播,整个过程表述如式(6)所示:

$$\boldsymbol{H}^{(l+1)}=\sigma\left(\tilde{\boldsymbol{A}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}+b^{(l)}\right). \tag{6}$$

其中:$\tilde{\boldsymbol{A}}$为加入自环并归一化后的依存邻接矩阵;$\boldsymbol{H}^{(l)}$为第l层的输入特征;$\boldsymbol{W}^{(l)}$和$b^{(l)}$分别为可学习的权重矩阵和偏置项;$\sigma(\cdot)$为非线性激活函数,用于对邻接聚合后的节点进行非线性映射.2层图卷积能够覆盖直接依存关系及部分二阶依存路径,同时避免因层数过深造成的节点表示趋同.

模型将对话结构图得到的宏观结构表示$\boldsymbol{h}^{\mathrm{dial}}$与句法依存图得到的微观结构表示$\boldsymbol{h}^{\mathrm{dep}}$进行整合,得到结构感知表示$\boldsymbol{h}_{\mathrm{str}}$,并送入后续多视图特征融合模块处理.通过将说话人关系、话语距离和句法依存路径分别建模,模型能够在同一结构建模阶段同时保留跨话语依赖和句内词级约束,为后续四元组关系识别提供结构化表示.

2.4 多视图融合机制

CRL输出细粒度语义表示,SIL输出融合对话拓扑与句法依存的结构表示.二者来源不同,直接拼接难以实现语义信息与结构信息的有效对齐,同时还可能造成特征维度膨胀;简单加权平均又难以区分不同词元对结构信息的需求.另一方面,图神经网络在多层传播后容易使邻近节点表示趋同,削弱实体识别所需的局部语义差异.因此,本文在语义表示与结构表示之间引入多视图特征融合机制.

该机制以CRL输出的语义表示作为Query,以SIL输出的结构表示作为Key和Value,通过协同注意力检索与当前词元相关的结构上下文.考虑到不同词元在四元组抽取中的作用不同,结构信息不作统一注入,而是根据当前语义状态进行选择性聚合.为进一步控制结构噪声,协同注意力之后引入门控残差融合,门控系数由结构上下文与原始语义表示共同计算,其形式如式(7)和式(8)所示:

$$g_{\mathrm{fusion}}=\sigma\left(\boldsymbol{W}_{\mathrm{fusion}}\left[c_{\mathrm{str}};\boldsymbol{h}_{\mathrm{ctx}}\right]+b_{\mathrm{fusion}}\right), \tag{7}$$

$$\boldsymbol{h}_{\mathrm{fusion}}=\mathrm{LayerNorm}\left(g_{\mathrm{fusion}}\odot c_{\mathrm{str}}+\left(1-g_{\mathrm{fusion}}\right)\odot\boldsymbol{h}_{\mathrm{ctx}}\right). \tag{8}$$

其中:g_{fusion}为结构信息与原始语义之间的融合权重;$\boldsymbol{W}_{\mathrm{fusion}}$和$b_{\mathrm{fusion}}$分别为门控残差融合中的可学习的权重矩阵和偏置项;$\boldsymbol{h}_{\mathrm{ctx}}$表示来自CRL层的输出语义;$c_{\mathrm{str}}$表示当前词元相关的结构上下文;LayerNorm(·)用于对融合后的表示进行归一化处理.通过该过程,融合表示既引入跨话语结构的上下文,又保留原始语义中的局部边界信息,从而为后续实体识别与关系分

类提供输入.

2.5　解耦式结构化解码与两阶段训练

结构化解码层负责将融合表示映射为离散情感四元组.若直接对所有实体候选进行端到端联合解码,候选对数量会随序列长度近似呈$O(N^2)$增长,其中大量配对并不构成有效语义关联,容易引入跨话语噪声.实体识别与边界确定主要依赖局部形态和句法线索,而关系判别与极性预测则更依赖跨话语语义交互,二者的优化目标并不完全一致.为降低候选空间规模和任务耦合带来的影响,本文采用如图3所示的解耦式结构化解码流程,将解码过程划分为实体候选生成、候选连接约束、关系类型推理与情感极性分类等阶段.首先通过网格标注确定待抽取实体,其次通过对话结构熵最小化机制压缩候选连接空间,最后利用三仿射关系推理器完成关系类型推理,并最终得到情感极性预测结果,体现了解耦式结构化解码的设计思想.

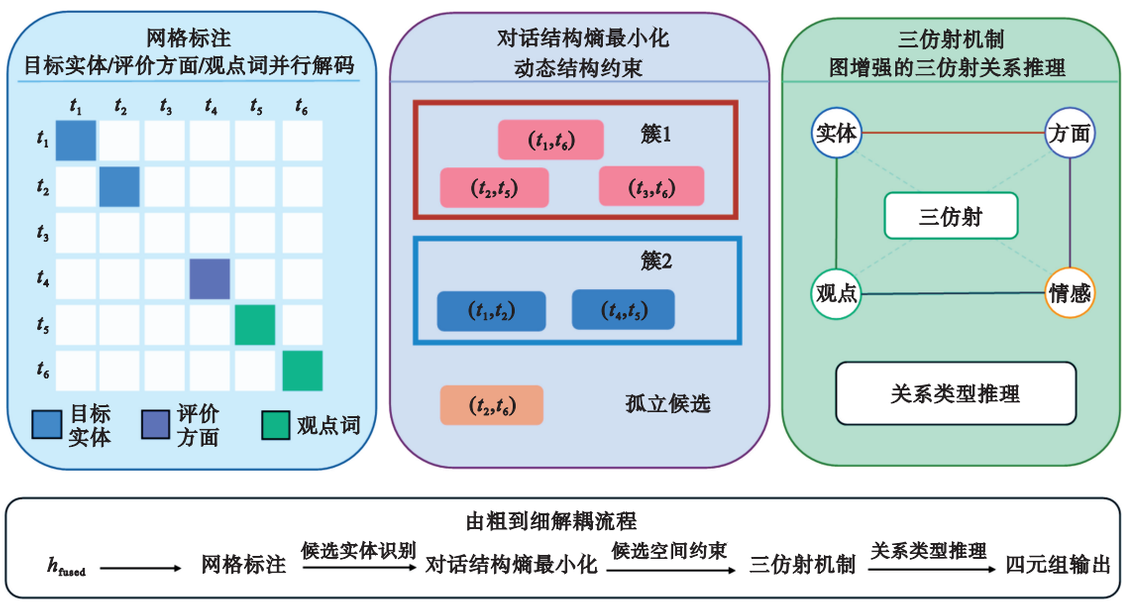


图3　解耦式结构化解码流程示意图

Fig. 3　Schematic diagram of decoupled structured decoding process

在实体候选生成阶段,将边界识别转化为二维跨度网格上的分类问题.设第i个位置与第j个位置构成候选跨度$s_{i,j}$,其评分函数定义如式(9)所示:

$$\phi(i,j)=\left(\boldsymbol{h}_i^{\text{head}}\right)^{\mathrm{T}}\boldsymbol{U}\boldsymbol{h}_j^{\text{tail}}+\left(\boldsymbol{h}_i^{\text{head}}\right)^{\mathrm{T}}\boldsymbol{T}\boldsymbol{h}_j^{\text{tail}}+b. \tag{9}$$

其中:$\phi(i,j)$为候选跨度$s_{i,j}$的类别评分函数;$\boldsymbol{h}_i^{\text{head}}$和$\boldsymbol{h}_j^{\text{tail}}$分别为第$i$个位置作为候选实体起始边界、第$j$个位置作为候选实体结束边界时的向量表示;$\boldsymbol{U}$为共享双线性权重矩阵,用于建模起止边界之间类别无关的匹配关系;$\boldsymbol{T}$为类别相关的高阶参数张量,用于刻画不同实体类别下边界组合的交互差异.通过该评分函数,二维网格中的每个单元均对应1个由起始位置和结束位置确定的候选跨度,模型能够在跨度级别完成实体边界与实体类别的联合判别,从而缓解一维序列标注中重叠边界相互覆盖的问题.网格卷积网络在网格空间内聚合相邻跨度的评分信息,用于增强局部连续的真实边界区域,并抑制孤立噪声单元.为进一步保留观点词边界信息,在实体解码前引入观点词边界辅助监督,以BIOES序列标注体系对观点词边界提供监督.B,I,O,E,S分别表示观点词的起始位置、内部位置、非观点词位置、结束位置和单独成词位置,用于减少图传播过程中局部边界特征被平滑的问题.

候选实体生成后,本方法引入基于对话结构熵最小化约束,而不是直接对候选实体进行全连接配对.设对话图分区方案为$\mathcal{P}$,其优化目标如式(10)所示:

$$\min_{\mathcal{P}}\left[\mathcal{H}_1(G)-\sum_{i=1}^{K}\frac{V_i}{2m}\mathcal{H}_1(C_i)\right]. \tag{10}$$

其中:$\mathcal{P}$表示对话图G的语义分区方案;K为语义簇数量;C_i表示第i个语义簇;$\mathcal{H}_1(G)$表示对话图在未划分状态下的一维结构熵,用于度量整体图结构的平均编码长度;$\mathcal{H}_1(C_i)$表示语义簇C_i内部子图的一维结构熵,用于刻画簇内结构的不确定性;V_i为语义簇C_i的簇内节点度或边权的累加结果;m为图G的总边权.该目标函数通过最小化平均编码长度确定语义簇边界,使语义相关的长距离依赖保留在同一簇内,同时减少跨话语候选连接,将后续关系推理限制在更紧凑的候选空间内.

在关系类型推理阶段,简单拼接或点积难以充分刻画目标实体、评价方面和观点词等情感要素之间的非对称依赖关系.本文采用三仿射关系推理机制,将候选头实体、候选尾实体与关系通道置于统一交互空间.对于候选头实体表示$\boldsymbol{h}_{\mathrm{h}}$与候选尾实体表示$\boldsymbol{h}_{\mathrm{t}}$,第$r$个关系通道的交互得分定义如式(11)所示:

$$s_r^{\mathrm{tri}}=\boldsymbol{h}_{\mathrm{h}}^{\mathrm{T}}\boldsymbol{W}_r\boldsymbol{h}_{\mathrm{t}}+[\boldsymbol{h}_{\mathrm{h}};\boldsymbol{h}_{\mathrm{t}}]^{\mathrm{T}}w_r+b_r. \tag{11}$$

其中:s_r^{tri}表示候选头实体与候选尾实体在第r个关系通道上的交互得分;$\boldsymbol{h}_{\mathrm{h}}$和$\boldsymbol{h}_{\mathrm{t}}$分别表示候选头实体和候选尾实体的表示向量;$\boldsymbol{W}_r$表示关系通道$r$对应的双线性切片矩阵,用于刻画该关系类型下头尾实体之间的非对称交互;$[\boldsymbol{h}_{\mathrm{h}};\ \boldsymbol{h}_{\mathrm{t}}]$表示头尾实体表示在特征维度上的拼接;$w_r$和$b_r$分别为第$r$个关系通道对应的线性项参数和偏置项.所有关系通道的双线性切片共同构成三维参数张量,使模型能够在头实体、尾实体与关系类型之间建立高阶交互.不同关系类型对应独立参数空间,使头尾实体之间的交互方式能够随关系类型变化;线性项用于保留实体自身语义,避免关系交互项完全覆盖个体特征.

由于实体边界检测主要依赖局部边界线索,而关系推理则需要建立在较可靠的实体候选之上,二者的优化目标和输入条件并不相同,因此本方法采用两阶段训练策略.

预热阶段截断关系分类与情感极性模块的梯度流,仅优化实体边界检测与观点词边界辅助监督,任务损失函数如式(12)所示:

$$\mathcal{L}_{\mathrm{warm}}=\mathcal{L}_{\mathrm{ent}}+\mathcal{L}_{\mathrm{op}}. \tag{12}$$

其中:$\mathcal{L}_{\mathrm{warm}}$为预热阶段损失函数;$\mathcal{L}_{\mathrm{ent}}$为实体边界检测损失;$\mathcal{L}_{\mathrm{op}}$为观点词边界辅助监督的序列标注损失.该阶段用于优先获得较稳定的候选实体集合,以减少错误候选对下游关系分类器的影响.

随后进入全局微调阶段,解除梯度截断并联合优化完整流水线,目标函数如式(13)所示:

$$\mathcal{L}_{\mathrm{global}}=\mathcal{L}_{\mathrm{ent}}+\mathcal{L}_{\mathrm{op}}+\lambda_{\mathrm{rel}}\mathcal{L}_{\mathrm{rel}}+\mathcal{L}_{\mathrm{pol}}. \tag{13}$$

其中:$\mathcal{L}_{\mathrm{global}}$为全局阶段任务损失函数;$\mathcal{L}_{\mathrm{rel}}$为三仿射关系分类损失;$\mathcal{L}_{\mathrm{pol}}$为情感极性预测损失;$\lambda_{\mathrm{rel}}$为关系分支损失权重.在全局微调阶段,本方法进一步引入动态采样机制,逐步降低真实实体标注在模型输入中的比例,并增加预测实体结果的参与程度,以缓解训练阶段与推理阶段输入分布不一致的现象.针对候选关系空间中正负样本不平衡的情况,关系推理损失采用Focal Loss[16],可以降低大量易分类负样本对参数更新的影响,使训练过程更多关注低频正样本和边界模糊样本.

3 实验及结果分析

3.1 实验数据准备

DiaASQ数据集是面向对话文本情感四元组抽取的基准数据集,语料来源于微博平台电子产品领域,包含1 000个对话片段和7 452条语句[1].数据集提供对话标识、分词序列、说话人映射、回复索引及情感四元组标注,本文实验使用其中文版本,并按照8∶1∶1的比例划分训练集、验证集与测试集.该数据集具有典型的对话级细粒度情感分析特征:每组对话平均涉及5位说话人,回复关系形成树状线程结构,约20%评价对象与观点词跨语句分布,要求模型同时处理长距离上下文、说话人交互与跨话语情感要素关联.其标注者间Cohen's Kappa系数为0.86,说明标注结果具有较高一致性.

数据预处理围绕模型输入格式展开.原始文本首先经BERT分词器进行子词级切分,并完成词级索引到子词级词元索引的映射;随后将四元组标注转化为包含实体边界、关系类型与情感极性的张量表示,以适配网格标注与结构化解码过程.同时,根据数据中的说话人信息、回复索引和话语顺序构建说话人邻接矩阵、线程掩码和话语距离矩阵,并结合句法依存分析生成词级依存结构.变长对话序列经填充后组成固定维度批次,作为模型输入.

3.2 实验设置与复杂度分析

当前模型基于PyTorch框架构建,BERT基座采用chinese-roberta-wwm-ext,输入序列最大长度设为512,超出部分截断,不足部分零填充.优化算法采用AdamW并配合分层学习率策略:BERT层学习率为1.5×10^{-5},其余模块为1×10^{-4}.损失函数由加权交叉熵、Focal Loss及权重为1.0的观点词边界辅助任务损失组合构成,关系判定子任务损失权重设为5.0.训练批大小为1,最大训练轮数为50.DSEM语义剪枝阈值p综合剪枝粒度与计算开销后确定为0.35.主要超参数配置见表1.

表1 主要超参数配置

Table 1 Main parameter configuration

超参数	设定值
学习率(BERT层/其他)	$1.5\times10^{-5}/1\times10^{-4}$
隐藏层维度/投影维度	768/256
GNN堆叠层数	2
协同注意力头数	4
关系损失权重	5.0
语义剪枝阈值	0.35

实验构建多层次评价体系,从实体边界、二元关系、三元组识别和完整四元组抽取等角度评估模型性能,并采用$F1$值作为主要指标.Target,Aspect与Opinion分别衡量目标实体、评价方面和观点词的跨度抽取效果,预测结果需与标注结果在左右边界及要素类型上保持一致;TA,TO与AO分别衡量目标实体–评价方面、目标实体–观点词和评价方面–观点词之间的二元关系抽取效果,要求参与配对的2个要素跨度均识别正确,且二者之间的关系匹配正确;Iden $F1$表示不区分情感极性时的四元组识别性能,用于反映模型对目标实体、评价方面和观点词组合关系的抽取能力;Micro $F1$表示完整情感四元组抽取的整体性能,要求目标实体、评价方面、观点词和情感极性均预测正确;Intra $F1$用于评估同一语句内部四元组的抽取效果,以反映方法对话语内情感要素组合的识别能力.上述指标共同覆盖了从单一情感要素识别到完整四元组预测的不同评价层次.

核心指标采用$F1_{\text{micro}}$,在全局样本层面统计各类别的TP(真阳性),FP(假阳性)与FN(假阴性)后计算综合得分,其定义如式(14)所示:

$$F1_{\text{micro}}=\frac{2\sum_c \text{TP}_c}{2\sum_c \text{TP}_c+\sum_c \text{FP}_c+\sum_c \text{FN}_c}. \quad (14)$$

其中,c表示类别.

该指标对高频类别更敏感,在四元组严格匹配的场景下能够直接反映模型对主流关系模式的整体识别能力.

本方法的计算开销主要来自上下文编码、结构建模、二维网格标注和候选实体对关系分类.设输入序列长度为S,隐藏层维度为d,对话中的话语数为U,图神经网络层数为L_{G}.上下文编码阶段主要由BERT完成,其自注意力计算随序列长度呈平方级增长.结构建模阶段,距离视图和说话人视图在话语级图上进行带掩码的信息传播,复杂度约为$O(L_{\text{G}}U^2d)$;句法依存分支在词级依存图上传播,其复杂度与依存边数量相关.二维网格标注阶段需要在词元对空间上预测实体边界及相关标签,复杂度约为$O(S^2d_{\text{g}}C_{\text{g}})$,其中$d_{\text{g}}$为网格内部投影维度,$C_{\text{g}}$为网格标签类别数.关系推理阶段,设经对话结构熵最小化约束后的候选实体对数量为P,关系类型数为R,三仿射投影维度为d_{r},则关系分类复杂度约为$O(PRd_{\text{r}}^2)$.本文实验基于NVIDIA RTX 4090 GPU开展,完整训练时间约为4 h,显存占用峰值约为12 GB.

3.3　消融实验

为评估各组件对模型性能的独立贡献,本文以完整模型为基准设置8组消融实验,每次移除1个模块,并在完整四元组微平均(Micro $F1$)、不区分情感极性的识别(Iden $F1$)和语句内四元组(Intra $F1$)三项指标上记录性能变化,结果见表2.

表2　模型消融实验结果

Table 2　Ablation experiment results of model

模型	$F1$/%		
	Micro	Iden	Intra
full model	43.51	46.69	46.39
no-distGAT	40.18	43.44	44.08
no-speakerGAT	42.25	44.96	45.84
no–GCN	41.82	44.61	45.61
no-co-attention	41.90	45.14	45.75
no-dsem	42.06	44.30	46.08
no-two-step	42.55	44.96	46.71
no-prec-decoder	41.71	44.76	44.98
no-triaffine	41.10	43.84	44.16

注:no表示移除.

完整模型在Micro $F1$和Iden $F1$上分别达到43.51%和46.69%,均高于各消融变体,说明各模块的组合能够提升完整四元组匹配和实体边界识别效果.虽然no-two-step的Intra $F1$略高于完整模型,但其Micro $F1$和Iden $F1$均出现下降,表明单一关系指标的提升并未转化为整体四元组抽取性能的提升.

从结构建模模块看,移除距离图注意力分支后,Micro $F1$降至40.18%,为所有变体中降幅最大的结果,Iden $F1$和Intra $F1$也分别降至43.44%和44.08%.这说明话语距离与回复结构对跨话语情感要素的连接具有关键作用.移除说话人图注意力分支后,Micro $F1$降至42.25%,表明说话人身份一致性有助于恢复同一说话人在不同轮次中的态度连续性,但其影响弱于语句距离与线程结构.移除依存句法GCN后,Micro $F1$降至41.82%,Iden $F1$降至44.61%,说明句法依存关系能够为实体边界和句内结构判断提供词级约束.

从特征融合与候选约束模块看,以向量拼接替代协同注意力后,Micro $F1$降至41.90%.这说明语义表示与结构表示之间需要显式对齐,简单拼接难以根据当前词元状态选择有效结构信息.移除对话结构熵最小化约束后,Micro $F1$降至42.06%,Iden $F1$降至44.30%,说明结构熵最小化约束能够减少跨话语的无效候选对,压缩后续关

系推理空间.将两阶段训练方式退化为端到端联合解码后,Micro $F1$ 降至42.55%,Iden $F1$ 降至44.96%.由于关系推理依赖实体边界检测生成的候选结果,训练初期直接联合优化容易使边界误差传递至关系分类模块,因此两阶段训练有助于提高整体解码稳定性.

从解码与关系推理模块看,移除精确解码器后,Micro $F1$ 降至41.71%,Iden $F1$ 和Intra $F1$ 分别降至44.76%和44.98%.该结果表明,二维跨度网格和局部网格修正能够提升实体边界定位质量,减少错误跨度进入后续关系推理.移除三仿射机制后,Micro $F1$ 降至41.10%,Iden $F1$ 和Intra $F1$ 分别降至43.84%和44.16%,整体降幅仅次于no-distGAT.三仿射机制将头实体表示、尾实体表示和候选关系类别同时纳入打分过程,为每个候选实体对计算不同关系类别下的匹配得分.移除该机制后,模型难以充分区分实体对的方向性差异以及关系类型与情感极性的组合关系,因此整体性能下降较为明显.

综合来看,距离图注意力分支和三仿射机制对完整四元组匹配的影响最明显,分别对应跨话语依赖建模和关系类型推理2个环节;句法依存GCN与精确解码器主要影响实体边界和句内结构判断;协同注意力、对话结构熵最小化约束与两阶段训练则分别从特征融合、候选约束和训练稳定性方面改善整体性能.消融结果表明,各模块并非独立提升单项指标,而是在结构建模、特征融合和关系推理过程中形成互补.

3.4 与其他模型对比实验

为评估本文模型在DiaASQ任务上的性能,本节将其与9种基线模型进行比较.对比方法涵盖流水线式序列标注、生成式抽取框架和对话级结构建模等范式,实验结果如表3所示.

CRF-Extract[17]与SpERT[18]属于流水线式抽取方法,主要在局部文本范围内完成实体识别与关系推理,其Micro $F1$ 分别为8.81%与14.19%.这类方法缺少对多轮对话中说话人关系、回复结构和跨话语依赖的建模,且不同阶段之间容易产生误差传递,因此在完整四元组抽取上的表现较欠佳.ParaPhrase[5]将四元组抽取转化为自回归生成问题,Micro $F1$ 提升至23.27%,一定程度上缓解了传统训练方式的误差级联问题,但缺少显式结构约束和边界对齐机制,在长对话中仍容易出现要素错配.Span-ASTE[19]引入跨度级特征交互与候选剪枝策略,Micro $F1$ 达到27.42%,多词实体的边界识别能力有所改善,但其整体仍面向句子级抽取场景,对跨话语关系的处理能力有限.

DiaASQ[1]是面向对话级情感四元组抽取的代表性模型,通过多视图交互网络融合线程、说话人和回复关系信息,Micro $F1$ 达到34.94%,较句子级方法有明显提升.H2DT[20]进一步构建回复-说话人异构图,并利用图卷积增强跨话语建模,Micro $F1$ 提升至38.17%.MDFI[21]通过多维特征交互,Micro $F1$ 达到42.36%,在对话级方法中表现较好.Triple GNNs[22]同时引入句法依存GCN与图注意力结构,Micro $F1$ 为41.77%,说明句内句法约束能够对跨话语建模形成补充.SEMDia[11]采用结构熵最小化方法对对话进行语义子簇划分,Micro $F1$ 达到42.44%,表明候选空间压缩有助于减少无效配对,提高关系推理效果.

表3 模型性能对比

Table 3 Model performance comparison %

模型	$F1$							
	T	A	O	TA	TO	AO	Micro	Iden
CRF-Extract	91.11	75.24	50.06	32.47	27.68	18.90	8.81	9.25
SpERT	90.69	76.81	54.06	38.05	31.28	21.89	14.19	13.00
ParaPhrase	—	—	—	37.81	34.32	27.26	23.27	27.98
Span-ASTE	—	—	—	44.13	34.46	32.21	27.42	30.85
DiaASQ	90.23	76.94	59.35	48.61	43.31	45.44	34.94	37.51
H2DT	91.72	76.93	61.87	50.48	48.80	52.40	38.17	40.34
MDFI	92.02	74.95	60.20	57.08	53.99	52.27	42.36	45.12
Triple GNNs	88.72	75.36	59.25	56.75	50.86	52.60	41.77	45.24
SEMDia	91.90	77.25	64.70	56.14	51.78	56.22	42.44	44.61
本文方法	92.00	79.04	59.72	58.05	51.42	55.03	43.51	46.69

注:T,A,O分别表示目标实体、评价方面和观点词的跨度匹配$F1$.

本文方法在Micro $F1$ 和Iden $F1$ 上分别达到43.51%和46.69%,均为表中最佳结果.与本文Micro $F1$ 最接近的SEMDia相比,本文方法分别在Micro $F1$ 和Iden $F1$ 上提高1.07和2.08个百分

点;与 MDFI 相比,分别提高 1.15 和 1.57 个百分点.SEMDia 在候选空间压缩方面具有优势,但其关系解码主要依赖较简单的实体交互方式,对头尾实体间非对称关系的表达能力有限.本文方法在候选约束基础上进一步引入三仿射机制,对头实体、尾实体和关系通道进行联合建模,因此在完整四元组匹配上取得更高结果.MDFI 强调多维特征交互,但对语义表示与结构表示之间的对齐处理较弱.本文方法通过协同注意力和门控残差融合对两类表示进行选择性整合,使结构信息能够更有效地服务于实体识别与关系判别.

从细粒度指标看,本文方法在评价方面的 $F1$ 指标上达到 79.04%,为所有方法中的最高值,说明对话结构解耦建模与句法分支对评价方面的识别具有一定帮助;在 TA 和 AO 指标上分别达到 58.05% 和 55.03%,处于较高水平,表明该方法能够较好地建模目标实体、评价方面与观点词之间的组合关系.综合各项指标可以看出,本文方法的优势主要来自对话结构建模、句法约束、候选空间压缩和高阶关系推理之间的协同作用.

4　结　　语

针对多轮对话情感四元组抽取中跨话语依赖建模不足、结构信号混合干扰以及候选关系空间较大等问题,本文提出 1 种基于双图神经网络的对话级情感四元组抽取方法.该方法将说话人关系与话语距离关系解耦建模,并引入依存句法图卷积分支以补充句内词级约束;在特征融合阶段,通过协同注意力对齐语义表示与结构表示;在解码阶段,利用对话结构熵最小化模块压缩候选空间,并通过三仿射解码器建模头实体、尾实体与关系类型之间的高阶交互.在 DiaASQ 基准数据集上的实验结果表明,本方法在 Micro $F1$ 和 Iden $F1$ 上均优于对比方法,消融实验也验证了各模块的有效性.

本文方法仍存在一定局限,受对话文本中情感表达分散、语句口语化和标注边界复杂等因素影响,整体 Micro $F1$ 仍为 43.51%,说明对话级情感四元组抽取仍具有较高难度.其中,实体与观点词边界识别仍是影响性能的重要因素;同时,对话结构熵最小化模块在压缩候选空间时可能过滤部分低频或跨簇四元组,从而影响召回率.

后续研究可进一步改进复杂边界识别与候选约束策略,引入更细粒度的词法、句法和上下文边界特征,并结合对比学习、数据增强或困难样本挖掘等方法缓解正负样本不均衡问题.此外,本文实验主要基于中文 DiaASQ 数据集,尚未验证该方法在英文及其他非中文对话语料上的适用性.不同语言在分词方式、句法依存结构和情感表达习惯上存在差异,跨语言迁移能力、多语言预训练表示的利用以及更长对话序列下的泛化性能仍有待进一步研究.

参考文献:

[1] Li B B, Fei H, Li F, et al. DiaASQ: a benchmark of conversational aspect-based sentiment quadruple analysis [C]//Findings of the Association for Computational Linguistics: ACL 2023. Toronto: Association for Computational Linguistics, 2023: 13449−13467.

[2] Ghosal D, Majumder N, Poria S, et al. DialogueGCN: a graph convolutional neural network for emotion recognition in conversation [C]// 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: Association for Computational Linguistics, 2019: 154−164.

[3] Hu D, Wei L W, Huai X Y. DialogueCRN: contextual reasoning networks for emotion recognition in conversations [C]// 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). Online: Association for Computational Linguistics, 2021: 7042−7052.

[4] Xu L, Li H, Lu W, et al. Position-aware tagging for aspect sentiment triplet extraction [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Online: Association for Computational Linguistics, 2020: 2339−2349.

[5] Zhang W X, Deng Y, Li X, et al. Aspect sentiment quad prediction as paraphrase generation [C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Online and Punta Cana: Association for Computational Linguistics, 2021: 9209−9219.

[6] Xiao X S, Chen J W, Li Q E, et al. Enhancing conversational aspect-based sentiment quadruple analysis with context fusion encoding method [C]//Natural Language Processing and Chinese Computing. Cham: Springer, 2023: 185−196.

[7] Li Y Q, Zhang W Y, Li B 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). Seoul: IEEE, 2024: 11241−11245.

[8] Huang P J, Xiao X S, Xu Y H, et al. DMIN: a discourse-specific multi-granularity integration network for conversational aspect-based sentiment quadruple analysis [C]//Findings of the Association for Computational Linguistics: ACL 2024. Bangkok, Thailand: Association for Computational Linguistics, 2024: 16326−16338.

[9] Jiang H Y, Chen X L, Miao D Q, 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.

[10] Lai Y Q, Fan S X, Tong Z L, et al. Conversational aspect-based sentiment quadruple analysis with consecutive multi-view interaction [C]//Natural Language Processing and Chinese Computing. Cham: Springer, 2023: 162-173.
[11] Peng K, Cao C, Peng H, et al. Dialogues aspect-based sentiment quadruple extraction via structural entropy minimization partitioning [C]//Proceedings of the 34th ACM International Conference on Information and Knowledge Management. New York: Association for Computing Machinery, 2025: 2326-2335.
[12] Tong Z L, Wei W, Qu X Y, et al. Multi-level association refinement network for dialogue aspect-based sentiment quadruple analysis [C]//Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vienna : Association for Computational Linguistics, 2025: 14035-14057.
[13] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [EB/OL]. (2016-09-09) [2025-09-10]. https://arxiv.org/abs/1609.02907.
[14] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[EB/OL].(2017-10-30) [2025-06-13]. https://doi.org/10.48550/arXiv.1710.10903.
[15] Marcheggiani D, Titov I. Encoding sentences with graph convolutional networks for semantic role labeling [C]// Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen: Association for Computational Linguistics, 2017: 1506-1515.
[16] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017: 2980-2988.
[17] Cai H J, Xia R, Yu J F. 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 (Volume 1: Long Papers). Online: Association for Computational Linguistics, 2021: 340-350.
[18] Eberts M, Ulges A. Span-based joint entity and relation extraction with Transformer pre-training[C]// ECAI 2020: 24th European Conference on Artificial Intelligence. Amsterdam: IOS Press, 2020: 2006-2013.
[19] Xu L, Chia Y K, Bing L D. Learning span-level interactions for aspect sentiment triplet extraction [C]// 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). Online: Association for Computational Linguistics, 2021: 4755-4766.
[20] Li B B, Fei H, Liao L Z, et al. Harnessing holistic discourse features and triadic interaction for sentiment quadruple extraction in dialogues[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver, 2024: 18462-18470.
[21] Zhao Z Y, Zhang L, Zheng Q S, et al. Multi-dimensional feature interaction for conversational aspect-based quadruple sentiment analysis [J]. *Neural Processing Letters*, 2025, 57(1): 9.
[22] Li B B, Li Y Q, Jia S Y, et al. Triple GNNs: introducing syntactic and semantic information for conversational aspect-based quadruple sentiment analysis[C]//2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD). Tianjin: IEEE, 2024: 998-1003.

(上接第 161 页)

[4] Dou X Y, Chen X H, Liang D Q, et al. A time-delay overlapping modulation-based high spectral efficiency and secure DCSK system[J]. *IEEE Access*, 2021, 9: 122685-122695.
[5] Liu Z F, Zhang L, Wu Z Q. Reliable and secure pre-coding OFDM-DCSK design for practical cognitive radio systems with the carrier frequency offset[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 6(1): 189-200.
[6] Boullé N, Dallas V, Nakatsukasa Y, et al. Classification of chaotic time series with deep learning [J]. *Physica D: Nonlinear Phenomena*, 2020, 403: 132261.
[7] Jia B W, Wu H Y, Guo K Y. Chaos theory meets deep learning: a new approach to time series forecasting [J]. *Expert Systems with Applications*, 2024, 255: 124533.
[8] Aricioğlu B, Uzun S, Kaçar S. Deep learning based classification of time series of Chen and Rössler chaotic systems over their graphic images [J]. *Physica D: Nonlinear Phenomena*, 2022, 435: 133306.
[9] Uzun S, Kaçar S, Arıcıoğlu B. Deep learning based classification of time series of chaotic systems over graphic images[J]. *Multimedia Tools and Applications*, 2024, 83(3): 8413-8437.
[10] Pourafzal A, Fereidunian A, Safarihamid K. Chaotic time series recognition: a deep learning model inspired by complex systems characteristics [J]. *International Journal of Engineering* , 2023, 36(1): 1-9.
[11] Akmeşe Ö F, Emin B, Alaca Y, et al. The time series classification of discrete-time chaotic systems using deep learning approaches [J]. *Mathematics*, 2024, 12(19): 3052.
[12] Tan M X, Le Q. EfficientNetV2: smaller models and faster training[C]// International Conference on Machine Learning. San Diego: PMLR, 2021: 10096-10106.
[13] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
[14] Gupta S, Akin B. Accelerator-aware neural network design using AutoML [PP/OL]. *arXiv* (2020-03-05) [2025-12-28]. https://doi.org/10.48550/arXiv.2003.02838.
[15] Xia Y X, Tse C K, Lau F C M. Performance of differential chaos-shift-keying digital communication systems over a multipath fading channel with delay spread [J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2004, 51(12): 680-684.
[16] Ramezani-Kebrya A, Antonakopoulos K, Cevher V, et al. On the generalization of stochastic gradient descent with momentum [J]. *Journal of Machine Learning Research* , 2024, 25(22): 1-56.