文章编号:1671-6833(2026)05-0077-08

基于三重提示和对比学习的中文医疗命名实体识别

刘 纳[1,2], 季 喆[1,2], 吴克东[1,2], 刘 磊[1,2]

(1.北方民族大学 计算机科学与工程学院,宁夏 银川 750021;2.北方民族大学 图形图像智能处理国家民委重点实验室,宁夏 银川 750021)

摘 要:针对中文医疗小样本命名实体识别任务中预训练知识表征不足、边界模糊及类别混淆等问题,提出一种基于三重提示引导的多阶段识别方法。首先,构建三重提示学习机制,通过边界定位、类型判别与语境验证的分步推理范式,结合离散模板的可解释性与连续向量的可优化性,增强模型对实体结构的解析能力;其次,设计双向门控网络,动态调控提示信息与文本表征的特征交互,强化复杂术语的语义捕获效果;最后,引入对比学习优化特征空间分布,提升实体类内聚合度与类间区分度。实验结果表明:该模型在 IMCS-V2-NER、cMedQANER 和 CCKS2019 3 个医疗命名实体识别数据集上 *F*1 值分别达到 91.86%、84.06%和 88.93%。证明该方法能够稳定提升实体识别性能,并在低资源场景下表现出较好的泛化能力。

关键词:小样本; 命名实体识别; 提示学习; 特征交互; 对比学习

中图分类号: TP391 **文献标志码**:A **doi**:10.13705/j.issn.1671-6833.2026.05.002

随着医疗信息化的发展,全球医疗系统积累了海量临床数据,非结构化文本占据显著比例,其中蕴含大量医学实体信息,如药物特征与治疗方案等。高精度的实体识别可将非结构化医疗文本有效结构化,不仅为病理机制研究提供多维数据支撑,也能为临床决策提供高效的智能诊断。因此,研究非结构化医疗数据中的实体识别具有重要的价值。

中文医疗命名实体识别(chinese medical named entity recognition, CMNER)[1]旨在从海量中文医疗文本中自动提取疾病、症状等关键实体,实现非结构化数据的结构化表达,为医疗信息抽取与知识图谱等下游任务提供高质量数据支持。

在方法的选择上,主流命名实体识别(named entity recognization, NER)[2]方法多基于深度学习模型,因其无须依赖烦琐的特征工程,且具备挖掘深层次、抽象特征的能力,在研究中得到了广泛应用。2024 年,林令德等[3]提出多层动态融合方法,从预训练模型中提取多层隐状态并动态整合,结合条件随机场(conditional random field, CRF)序列标注,显著增强语义表征与模型泛化能力。同年,Yang 等[4]针对中文电子病历语义模糊与结构复杂的问题,通过 BERT 提取动态词向量,多头注意力增强语义表示,CRF 进行全局最优解码,显著增强实体识别能力。2025 年,Zhang 等[5]提出基于机器阅读理解的医学 NER 方法,通过双向长短期记忆网络(bidirectional long short-term memory network, BiLSTM)和双向注意力机制获取全局语义表示,以索引精准定位命名实体边界,有效提升了模型对医疗实体的识别能力。

现阶段命名实体识别(NER)主流方案以深度学习为主流,但该类方法在中文医疗文本处理中存在显著短板:一是医疗术语体系复杂、中文文本多义问题难以处理;二是模型训练需要消耗大量人工标注资源。小样本学习场景下,提示学习范式[6]能够有效缓解上述两类痛点,为相关研究提供新方向。Brown 等[7]首次提出提示学习概念,并通过 GPT-3 验证了其在多任务迁移中的有效性。

目前,提示学习需解决的核心问题在于提示模板的设计[8],早期的提示学习方法主要是人工构建提示模板来引导模型获取知识。Schick 等[9]提出模式利用训练的方法,将输入样本转换成填空式短语以增强语言对任务的理解。Cui 等[10]将 NER 重构为序列到序列任务,通过枚举类型候选跨度并依据

收稿日期:2026-02-23;**修订日期**:2026-04-24

基金项目:国家自然科学基金资助项目(62162001);宁夏重点研发计划引才专项项目(2024BEH04020)

作者简介:刘纳(1986—),女,宁夏银川人,北方民族大学讲师,博士,主要从事数据挖掘与自然语言处理研究,E-mail:Liuna@nun.edu.cn。

引用本文:刘纳,季喆,吴克东,等. 基于三重提示和对比学习的中文医疗命名实体识别[J]. 郑州大学学报(工学版),2026,47(5):77-84.[Liu Na, Ji Zhe, Wu Kedong, et al. Chinese medical named entity recognition based on triple hint and contrast learning[J]. Journal of Zhengzhou University (Engineering Science),2026,47(5):77-84.]

模板填充得分进行判定,从而完成实体识别。

离散提示[11]通常以自然语言片段构建模板,在离散空间中进行搜索,具备较为良好的语义直观性和可解释性,能够有效引导模型理解任务意图。Jiang 等[12]提出结合 BiLSTM 与离散提示的自监督模型,通过构建输入与复述指令之间的相似性信号,在无标注数据下实现实体识别。

以上方法依赖人工设计,泛化能力有限,难以适应复杂任务场景。为克服这些问题,连续提示被提出。提示表示映射为可训练的向量,通过优化策略自动学习任务相关表达。Mai 等[13]以问答式连续提示将实体识别改写为掩码预测,并以字符级枚举提升边界精度,显著强化小样本表现。但连续提示可解释性较弱,难以对提示行为进行显式调控。针对以上问题,Liu 等[14]提出了多源知识融合的深度提示微调框架。该方法将上下文知识、标签先验与语义结构编码注入软提示嵌入,提升模型对实体边界与类别的识别能力。

基于以上内容,本文提出了一种融合连续与离散提示的三重提示引导的中文医疗命名实体识别方法 PLC-GT,构建三重提示引导的混合提示学习机制,兼顾离散可解释性与连续可优化性,采用分阶段策略增强实体边界与语义建模;引入双向门控网络对文本与提示动态融合加权;结合对比学习聚合同类、分离异类实体表示,提升识别与泛化性能。

1 中文医疗命名实体识别模型

1.1 任务定义

CMNER[15]旨在从医疗文本中自动识别并分类医学相关的实体,将每个单词或字符映射到相关的实体类型。其本质就是一个 token 序列预测问题,即针对输入的文本序列 $X=(x_1,x_2,\cdots,x_t)$,模型需要输入对应的标签序列 $Y=(y_1,y_2,\cdots,y_t)$,其中 Y 取值取自预定义的医学类别集合 Z,如疾病、药物、治疗方法等。在小样本学习场景下,数据集 D 被划分为支持集、查询集以及验证集。传统 N-way K-shot 采样的方式为每个类别固定采样 K 个实例,医疗数据类型分布不均衡,固定采样可能导致某些类别数据不足,影响模型的训练能力。本文采用贪心采样策略[16],即对每个类别随机抽样 $K\sim 2K$ 个样本,使样本数量在一定范围内浮动,从而缓解数据不均衡问题,提高模型的识别能力。

1.2 模型整体框架图

本文针对小样本中文医疗命名实体识别任务,提出了一种融合三重提示学习模板、BGN 与对比学习[17](contrastive learning, CL)的模型框架,整体模型框架图如图 1 所示。首先,特征选择器对文本与提示特征进行筛选,输入 MCBERT 预训练模型获得深层语义表征;其次,引入双向门控网络(BGN),通过门控机制对文本与提示信息进行动态融合,依据语义表征自适应调节二者贡献;最后,结合对比学习拉近同类、分离异类实体表示,提升判别能力。

1.2.1 提示增强

本研究采用任务特定的提示策略,通过引入类特定词显式编码标签语义,增强类别区分能力,缓解标注数据不足导致的偏差[16]。

具体而言,首先为每个预定义的类别标签分配一个独特的类特定词,例如,将疾病标签“dis”映射为“disease”。这些类特定词保留了类别名称的语义信息,模型能够在学习过程中利用其内在含义来

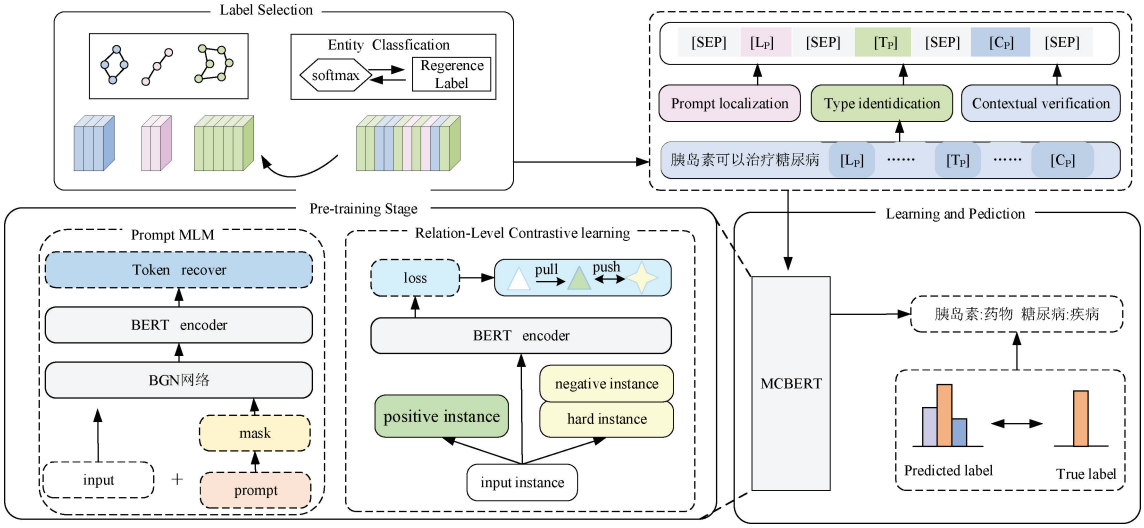

图 1 本文整体模型框架图

Figure 1 Overall model framework of this study

增强类别区别能力,减少对有限标注数据的依赖。在小样本任务中,所有类特定词按照类别顺序拼接,形成提示符,进一步附加至原始输入序列 $S=\{x_1,x_2,\cdots,x_t\}$ 的末尾,以构造类特定词和文本的扩展序列 $S^*=\{x_1,x_2,\cdots,x_t,w_1,w_2,\cdots,w_n,w_{n+1}\}$。相比于仅依赖有限的训练数据,该方法在输入中显式引入了类别相关的先验信息,使得模型能够在训练过程中自适应地捕捉类别语义特征。

1.2.2　三重提示模板

针对医疗文本中存在复杂术语、语义模糊等问题时,本文提出一种基于离散-连续混合三重提示模板,层层递进地引导模型完成实体识别任务。第一阶段,借助二分类提示与编码器结构,明确实体在文本中的边界,实现实体定位;第二阶段,通过多种类型提示,识别实体所属的具体类别,细化分类结果;第三阶段,结合上下文提示和解码器层,判断实体在当前语境中的语义,确保分类的准确性。

在实体定位阶段,通过构造显示的实体定位提示模板,将实体定位转化为显式的二分类问题,极大提高了实体定位的准确率。

给定医疗文本序列 $S=\{x_1,x_2,\cdots,x_t\}$,任意候选实体片段 $m=x_{a:b}\subseteq S$,提示序列输入形式为"$Z=$[CLS]m 是否为实体?",将输入序列 Z 送入预训练模型编辑器 $f_\theta(\cdot)$,获取对应的上下文语义表示,通过 Tanh 非线性激活函数和 Sigmoid 函数相结合来计算候选实体的概率,如公式(1)所示:

$$P(m\mid S)=\sigma(\boldsymbol{W}_e^{\mathrm{T}}\mathrm{Tanh}\,f_\theta(Z)+\boldsymbol{b}_e)。\tag{1}$$

式中:$\boldsymbol{W}_e$、$\boldsymbol{b}_e$ 为待优化参数,为优化实体存在性判断,采用二元交叉熵损失函数,如式(2)[18]所示:

$$L=-(y\log P(m\mid S)+(1-y)\log(1-y)P(m\mid S))。\tag{2}$$

式中:真实标签 $y\in\{0,1\}$,表示片段是否为真实实体,实体定位的输出为候选片段的二分类判别。

实体类型识别阶段基于已抽取的候选实体片段,确定其细粒度实体类型。为此,本研究构建提示增强机制,将预定义类别转化为掩码提示,使模型依托上下文完成实体类型推断。

模型通过对 MASK 位置的上下文向量进行线性映射,计算该实体类型概率分布,如式(3)所示:

$$P(c\mid m,S)=\frac{\exp(\boldsymbol{w}_c^{\mathrm{T}}h+\boldsymbol{b}_c)}{\sum_{c'\in c}\exp(\boldsymbol{w}_{c'}^{\mathrm{T}}h+\boldsymbol{b}_c)}。\tag{3}$$

式中:c 表示所有实体类别集合;$\boldsymbol{w}_c$ 为类别权重矩阵;$\boldsymbol{b}_c$ 为待优化参数。使用 Softmax 运算为每个候选类别预测概率分数,为优化类型识别能力,采用负对数似然函数作为训练目标,如式(4)所示:

$$L=-\log P(c''\mid m,S)。\tag{4}$$

式中:$c''\in C$ 为实体片段 m 的真实类别标签。通过优化该损失函数,使模型在所有候选类别中分配给真实类别的预测概率达到最大。

鉴于医疗文本的语义强依赖上下文信息,本文设计语境验证提示机制:在给定上下文条件下,对候选实体片段与目标类别的匹配性进行显式验证,从而判断该片段在当前语境中是否应被归入该类别,并降低脱离语境导致的误判。

将 MASK 位置的上下文向量与实体嵌入进行深度交互融合,构建设计类别-语境匹配向量来计算二者的匹配概率,如式(5)所示:

$$P(y\mid c,m,S)=\sigma(\boldsymbol{V}^{\mathrm{T}}(L*e_C)+\boldsymbol{b}_v)。\tag{5}$$

式中:e_C 表示类别 C 的可学习嵌入;$\boldsymbol{V}$、$\boldsymbol{b}_v$ 表示匹配参数,其中最高概率为实体所属类型。

1.3　双向门控网络

本研究提出 BGN 网络架构,由两条独立 GRU 分支分别编码文本与提示序列来动态筛选关键信息,通过可学习融合权重自适应整合双通路输出,增强三重提示的适应性与实体语义表示能力。

门控信号生成是基于当前输入与历史状态的联合信息生成调控门,如式(6)和式(7)所示:

$$Z_t=\sigma(\boldsymbol{W}_z[h_{t-1};x_t]+b_z);\tag{6}$$

$$r_t=\sigma(\boldsymbol{W}_r[h_{t-1};x_t]+b_z)。\tag{7}$$

式中:$[\cdot;\cdot]$表示向量拼接;$\boldsymbol{W}_z$、$\boldsymbol{W}_r$ 表示门控权重矩阵;$b_z\in R^{d_h}$ 为偏置项。前一时刻 h_{t-1} 与候选状态$\widetilde{h}_t$进行加权组合得到最终隐状态 h_t[19],如式(8)所示:

$$h_t=(1-z_t)h_{t-1}+z_t\widetilde{h}_t。\tag{8}$$

在双通路编码过程中,模型分别获得文本特征状态 $h_t^{(w)}$ 和提示特征状态 $h_t^{(p)}$。为实现二者特征融合,设计可学习的动态加权融合,如式(9)所示:

$$G_t=\boldsymbol{\alpha}h_t^{(w)}+(\mathbf{1}-\boldsymbol{\alpha})h_t^{(p)}。\tag{9}$$

式中:$\boldsymbol{\alpha}\in[0,1]^{d_h}$ 表示自适应权重向量。

1.4　对比学习

为实现不同实体类别的有效分离,本文采用对比学习训练实体特征提取器,在训练阶段利用对比学习损失函数来拉近正样本距离,疏远负样本距离。

首先,利用实体在不同语义路径下的表示差异强化,对每个实体的最终表示 h_i^{final} 和平均表示 h_i^{avg} 构成正样本对,如式(10)所示:

$$Q=(h_i^{\mathrm{final}},h_i^{\mathrm{avg}}\mid i\in B)。\tag{10}$$

为提升模型对相近类别间细粒度差异的区分能

力,设计困难负样本挖掘策略。负样本对通过困难样本采样策略进行生成,在同一批次中筛选出混淆度较高的实体对作为负样本,增加模型对边界判别的敏感度。若 $y_i \neq y_j$ 且两者相似度高于阈值 γ,则构成负样本对,如式(11)所示:

$$N = (h_i^{\text{final}}, h_j^{\text{final}}) \mid (i,j), y_i \neq y_j \wedge \text{sim}(h_i, h_j > \gamma)。 \quad (11)$$

通过温度缩放的对比学习损失函数调整模型对正负样本对的敏感程度,提升语义空间区分度,损失函数如式(12)[20]所示:

$$L = -\frac{1}{|Q|}\sum_{\boldsymbol{h}_i,\boldsymbol{h}_j \in p} \log \frac{\exp(\boldsymbol{h}_i^{\mathrm{T}}\boldsymbol{h}_j/\tau)}{\sum_{k=1}^{N}\exp(\boldsymbol{h}_i^{\mathrm{T}}\boldsymbol{h}_k/\tau)}。 \quad (12)$$

式中:Q 表示生成样本对的集合;$\boldsymbol{h}_i$ 和 $\boldsymbol{h}_j$ 表示两个样本的特征表示;τ 表示温度系数。

2 实验分析

2.1 数据集

本实验采用了 IMCS-V2-NER[21]、cMedQANER[22]和 CCKS2019[23]3 个中文医疗命名实体识别数据集来验证 PLC-GT 模型的有效性。此外,在通用领域命名实体识别数据集 CLUENER2020[24]上验证 PLC-GT 模型的泛化性,数据集详细信息如下所示。

(1)IMCS-V2-NER[21]:该数据集来源于智能诊疗医患对话,包括药品名称等 5 种医学实体,划分为训练集 2 472 条、验证集 833 条和测试集 811 条。

(2)cMedQANER[22]:该数据集来源于中文医疗问答系统,包括疾病等 11 种医学实体,样本划分为训练集 1673 条、验证集 215 条和测试集 175 条。

(3)CCKS2019[23]:该数据集来源于全国知识图谱与语义计算大会所提供的电子病历信息,包括药物等 6 种医学实体,样本划分为训练集 1 000 条、验证集 379 条和测试集 379 条。

(4)CLUENER2020[24]:该数据集是在清华大学开源文本分类数据集 THUCNER 中部分数据进行命名实体标注,包括公司等 10 种实体,样本划分为训练集 10 748 条、验证集 1 343 条和测试集 1 345 条。

2.2 实验环境与参数设置

本实验所使用的深度学习框架为 PyTorch1.8.1,实验系统为 Ubuntu 20.04,开发环境为 python3.8,Cuda 版本为 12.4,GPU 为 RTX4090(24 GB)。

在参数设置上,隐藏层维度设置为 768 维,学习率为 1e-4,学习样本大小 *Batch_size* = 16,温度系数 *Tau* = 0.1,序列最大长度 *Max_length* = 128 以及优化器使用 AdamW。训练阶段迭代次数为 10 000 次,测试阶段设置为 5 000 次迭代。

2.3 实验结果与分析

为了评估 PLC-GT 模型在命名实体识别任务中的有效性表现,选取 7 个小样本命名实体识别经典模型(NNshot[25]、StructShot[25]、TemplateNER[10]、LabelBERT[26]、PromptNER[27]、LightNER[28]、ProML[29])为基线进行对比。实验以精确率(*Precision*)、召回率(*Recall*)及 *F*1 值作为评价指标。

2.3.1 医疗领域数据集的性能对比

本文模型在中文医疗领域的 IMCS-V2-NER、cMedQANER 和 CCKS2019 3 个医疗数据集上进行实验,实验结果如表 1 所示,由表 1 可知,本研究提出的模型在 3 个中文医疗命名实体识别数据集上和两种贪心策略设置下的实验结果均优于现有的主流基线方法。其中,在 IMCS-V2-NER 和 cMedQANER 数据集上,5~10 shot 相对于 1~2 shot 的设置中,*F*1 值分别增加了 1.36 百分点和 1.46 百分点。三重提示机制通过多源异构方式融合类型语义与语境先验,联合关注实体类别表征与边界结构,强化模型对上下文与标签边界的感知,随着样本数量增加,语境先验不断积累,模型对类型间区分度的判别更充分,实体边界更加清晰,从而更易区分不同语境下的实体类型。

具体而言,在 cMedQANER 数据集上,ProML 在低资源设定下均优于基线,ProML 采用引导式提示,以少量提示样本激活预训练模型,降低模型对大规模数据的依赖。然而,医疗文本中实体随语境变化的语义差异,二者在复杂场景的抽取能力受限,缺少显式语义约束,易出现过拟合和识别错误。

2.3.2 通用领域数据集的性能对比

为进一步验证 PLC-GT 模型在中文通用领域的泛化能力,本文在 CLUENER2020 数据集上进行对比实验,实验结果如表 2 所示。由表 2 可知,本文提出的模型展现出优越的性能,整体表现显著优于基线模型。在 5-way 1~2shot 和 5-way 5~10shot 设置下,PLC-GT 模型的 *F*1 值分别为 76.63% 和 83.21%,与 LabelBERT 模型进行对比,分别提升了 1.75 百分点和 6.01 百分点;相较于仅采用连续提示机制的 LightNER 模型,分别提升了 0.66 百分点和 2.75 百分点,验证了该模型在小样本场景下的泛化能力。三重提示模板有效结合了可学习向量的灵活表达与显式类别语义的先验引导,在提升模型判别能力的同时增强不同语境下泛化能力。

2.4　消融实验

2.4.1　模块消融

为评估各模块的贡献,本文以 *F*1 值为评价指标,实验结果如图 2 所示。

在移除提示学习模块(w/o Prompt)后,PLC-GT 模型在两种样本设置下,分别降低了 2.13 和 2.31 百分点。结果表明,提示学习模块在建模全局上下文语义和丰富文本特征表示方面发挥了关键作用,极大提升实体识别准确性。移除双向门控网络(w/o BAN)后,模型在 3 个数据集上的 *F*1 均出现不同程度下降,双向门控网络通过对文本和提示特征进行动态融合,在信息筛选与权重调控方面起到了积极的作用。当对比学习模块(w/o CL)被移除时,模型性能显著下降,在 cMedQANER 数据集中,5-way 1~2shot 设置下,*F*1 值降幅最大,降低了 8.11 百分点。表明该模块在实体类型判别中具有关键作用。通过构建正负样本对,对比学习有效增强了类别间的语义区分能力,缓解了标签不平衡带来的性能。

表 1　医疗数据集实验结果

Table 1　Experimental results of medical data set　%

数据集	模型	5-way 1~2shot			5-way 5~10shot		
		Precision	*Recall*	*F*1	*Precision*	*Recall*	*F*1
IMCS-V2-NER	NNshot	78.11	85.63	81.70	84.17	81.33	82.72
	structShot	83.90	81.61	81.99	79.63	86.96	83.12
	TemplateNER	82.22	88.51	85.26	86.10	90.79	88.38
	LabelBERT	83.82	90.62	87.14	87.20	91.19	90.55
	PromptNER	78.40	88.92	83.33	82.27	90.07	85.99
	LightNER	85.67	89.88	87.72	87.34	92.81	90.58
	ProML	81.76	91.07	86.16	84.61	90.93	88.12
	PLC-GT(本文)	**87.91**	**91.89**	**89.86**	**88.91**	**93.89**	**91.86**
cMedQANER	NNshot	69.28	72.91	71.05	69.96	80.36	74.80
	structShot	66.51	83.66	74.11	73.10	81.70	77.16
	TemplateNER	73.21	82.07	78.60	77.91	80.61	80.98
	LabelBERT	78.09	84.41	80.82	80.66	84.43	83.76
	PromptNER	70.14	83.22	76.12	75.35	84.42	79.62
	LightNER	77.71	81.19	82.22	80.70	83.08	82.92
	ProML	77.63	81.03	79.29	80.33	83.69	81.97
	PLC-GT(本文)	**79.66**	**85.75**	**82.60**	**81.90**	**85.76**	**84.06**
CCKS2019	NNshot	63.63	61.96	74.35	76.35	80.79	73.33
	structShot	70.26	80.85	75.19	75.43	83.35	85.00
	TemplateNER	81.93	82.56	82.74	83.71	84.01	82.80
	LabelBERT	80.80	86.21	84.11	83.99	87.98	85.94
	PromptNER	73.23	79.10	76.05	77.01	84.90	80.76
	LightNER	79.91	83.82	85.05	85.14	90.77	87.81
	ProML	75.16	78.42	76.76	77.17	85.22	80.09
	PLC-GT(本文)	**82.22**	**88.51**	**85.56**	**85.97**	**91.28**	**88.93**

表 2　通用领域数据集实验结果

Table 2　Experimental results of universal domain data set　%

数据集	模型	5-way 1~2shot			5-way 5~10shot		
		Precision	*Recall*	*F*1	*Precision*	*Recall*	*F*1
CLUENER2020	NNshot	61.80	67.71	64.62	66.71	75.43	70.80
	structShot	69.89	75.24	72.47	71.15	83.33	76.76
	TemplateNER	70.11	78.83	74.43	77.47	84.11	78.06
	LabelBERT	71.61	77.21	74.88	78.54	85.95	77.20
	PromptNER	70.96	76.89	73.52	74.94	81.33	76.63
	LightNER	72.01	79.94	75.97	76.41	83.08	80.54
	ProML	72.18	75.91	74.00	75.16	85.95	80.18
	PLC-GT(本文)	**72.45**	**80.77**	**76.63**	**79.39**	**87.21**	**83.21**

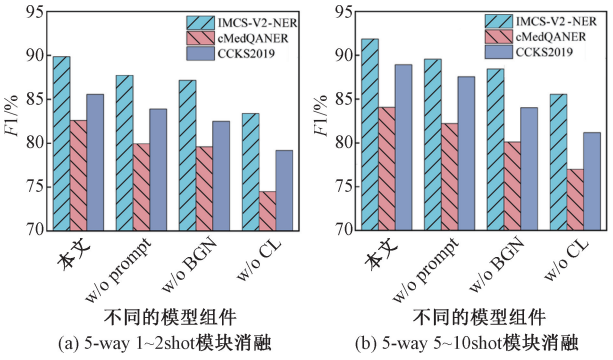

图 2 模块消融实验结果

Figure 2 Experimental results of modular ablation

2.4.2 提示消融

为评估提示模板对模型性能的影响,本文设计并测试了 5 种提示模板,提示模板如表 3 所示,消融实验如图 3 所示。

表 3 提示模板

Table 3 Prompt templates

提示模板	示例
P_1	这是一个<MASK>实体。
P_2	<候选实体>是<MASK>实体。
P_3	<候选实体>是<MASK>实体吗?
P_4	<候选实体>是一个实体吗? <候选实体>是<MASK>实体吗? <候选实体>是一个实体吗?
P_5	<候选实体>是<MASK>实体吗? 在上下文中,<候选实体>属于<MASK>吗?

图 3 实验结果表明,不同模板对模型性能影响显著。相较结构简单、难以建立候选实体与类别语义关联的 P_1、P_2,疑问句 P_3 能更明确任务目标,有效提升识别准确性。P_4 通过二重提示将 NER 拆分为“实体定位-类型分类”两步,缓解边界模糊。在此基础上,P_5 引入语境验证提示,增强类型语义聚合与推断能力,降低模板同质化对泛化能力限制。

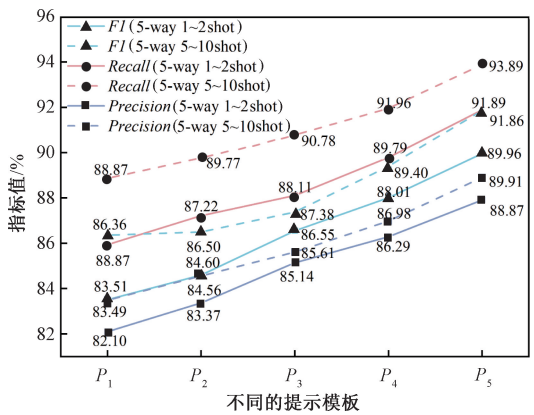

图 3 提示模板消融实验结果

Figure 3 Prompt template ablation experimental results

从提示结构看,P_1 ~ P_4 为非三重提示,单样本仅依赖单一模板与单条语义信息,面对多义实体或强上下文依赖候选实体时易出现漏检并产生边界模糊。而 P_5 为三重提示,在同一候选实体上引入多源、多层次语义引导,提供冗余且互补信息,强化实体与类型的语义关联,提升了上下文感知和推理能力。

3 结论

针对中文医疗领域中实体边界模糊与一词多义问题,本文提出三重提示引导的命名实体识别模型 PLC-GT。该模型融合可解释离散与可优化连续提示生成提示向量以弥补小样本场景下的语义不足。通过双向门控网络动态筛选文本与提示信息中的关键信息,强化实体语义表达;同时引入对比学习优化实体表征,显著提升类别区分能力。

尽管 PLC-GT 模型在平坦数据集的命名实体识别中表现优异,但在处理多层嵌套或跨句实体时仍存在局限。多层嵌套实体伴随着边界模糊与语义高度重叠,使传统方法难以准确表示其层次关系。未来研究将重点探究面向嵌套实体的提示设计与对比学习策略,提升在多层嵌套医疗文本中的适应性。

参考文献:

[1] Kang Hui, Xiao Jingwu, Zhang Yunpeng, et al. A research toward Chinese named entity recognition based on transfer learning[J]. International Journal of Computational Intelligence Systems, 2023, 16: 56.

[2] Liu Xin, Xu Hongzhen, Liu Aihua, et al. Geological named entity recognition based on MacBERT and R-Drop [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(3): 89-95. [刘昕, 徐洪珍, 刘爱华, 等. 基于 MacBERT 和 R-Drop 的地质命名实体识别[J]. 郑州大学学报(工学版), 2024, 45(3): 89-95.]

[3] Lin Lingde, Liu Na, Xu Zhenshun, et al. Chinese medical named entity recognition based on multi-layer dynamic fusion [J]. Computer Engineering and Applications, 2024, 60(15): 161-169. [林令德, 刘纳, 徐贞顺, 等. 基于多层动态融合的中文医疗命名实体识别[J]. 计算机工程与应用, 2024, 60(15): 161-169.]

[4] Yang Haotian, WANG Li, Yang Yanpeng, et al. Named entity recognition in electronic medical records incorporating pre-trained and multi-head attention[J]. IAENG International Journal of Computer Science, 2024, 51(4): 401-408.

[5] Zhang Ziqi, Zheng Xiangwei, Zhang Jinsong. Machine

reading comprehension based named entity recognition for medical text[J]. Multimedia Tools and Applications, 2025, 84(28): 33431-33451.

[6] Cui Jinman, Li Dongmei, Tian Xuan, et al. Survey on prompt learning[J]. Computer Engineering and Applications, 2024, 60(23): 1-27. [崔金满, 李冬梅, 田萱, 等. 提示学习研究综述[J]. 计算机工程与应用, 2024, 60(23): 1-27.]

[7] Brown T B, Mann B, Ryder N, et al. Language models are few-shot learners[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 1877-1901.

[8] Zheng Guofeng, Liu Na, Li Chen, et al. Chinese medical named entity recognition based on prompt tuning and contrastive learning[J]. Computer Engineering and Applications, 2026,62(1):231-242. [郑国风, 刘纳, 李晨, 等. 基于提示调优与对比学习的中文医疗命名实体识别方法[J]. 计算机工程与应用, 2026,62(1):231-242.]

[9] Schick T, Schütze H. Exploiting cloze-questions for few-shot text classification and natural language inference [C]//Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Stroudsburg: ACL, 2021: 255-269.

[10] Cui Leyang, Wu Yu, Liu Jian, et al. Template-based named entity recognition using BART[C]//Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Stroudsburg: ACL, 2021: 1835-1845.

[11] Petroni F, Rocktäschel T, Riedel S, et al. Language models as knowledge bases? [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). Stroudsburg: ACL, 2019: 2463-2473.

[12] Jiang Cong, Xu Qingyang, Song Yong, et al. Discrete sequence rearrangement based self-supervised Chinese named entity recognition for robot instruction parsing[J]. Intelligence & Robotics, 2023, 3(3): 337-354.

[13] Mai Chengcheng, Chen Yu, Gong Ziyu, et al. PromptCNER: a segmentation-based method for few-shot Chinese NER with prompt-tuning[J]. ACM Transactions on Asian and Low-Resource Language Information Processing, 2025, 24(9): 1-24.

[14] Liu Jiang, Fei Hao, Li Fei, et al. TKDP: threefold knowledge-enriched deep prompt tuning for few-shot named entity recognition[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(11): 6397-6409.

[15] Chen Yuxiang, Dan Zhiping, Gao Zhun, et al. Chinese medical continual named entity recognition based on continual learning and knowledge distillation[C]//Proceedings of the 2024 IEEE International Conference on High Performance Computing and Communications (HPCC). Piscataway: IEEE, 2024: 885-893.

[16] Chen Junqi, Ding Zhaoyun, Jin Guang, et al. The optimal sampling strategy for few-shot named entity recognition based with prototypical network[C]//Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE, 2024: 1-8.

[17] Wang Chengyu, Zhao Shan, Yan Tianwei, et al. Hierarchical label-enhanced contrastive learning for Chinese NER[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(6): 10504-10514.

[18] Qin Xiao, Tan Sijing, Chun Xin, et al. Simplify prompt template and optimum label-select for few-shot NER [C]//Applied intelligence. Singapore: Springer Nature Singapore, 2025: 370-381.

[19] Bhushan R C, Donthi R K, Chilukuri Y, et al. Biomedical named entity recognition using improved green anaconda-assisted Bi-GRU-based hierarchical ResNet model [J]. BMC Bioinformatics, 2025, 26: 34.

[20] Huang Yucheng, He Kai, Wang Yige, et al. COPNER: contrastive learning with prompt guiding for few-shot named entity recognition[C]//Proceedings of the 29th International Conference on Computational Linguistics. Stroudsburg: ACL, 2515-2527.

[21] Zhu Wei, Wang Xiaoling, Zheng Huanran, et al. PromptCBLUE: a Chinese prompt tuning benchmark for the medical domain[PP/OL]. V1. arXiv (2023-10-22)[2025-12-24]. https://doi.org/10.48550/arXiv.2310.14151.

[22] Liu Feifei, Liu Mingtong, Li Meiting, et al. Automatic knowledge extraction from Chinese electronic medical records and rheumatoid arthritis knowledge graph construction[J]. Quantitative Imaging in Medicine and Surgery, 2023, 13(6): 3873-3890.

[23] Wang Haitao, He Zhengqiu, Zhu Tong, et al. CCKS2019 shared task on inter-personal relationship extraction[PP/OL]. V1. arXiv (2019-08-29)[2025-12-24]. https: doi.org/10.48550/arXiv.1908.11337.

[24] Xu Liang, tong Yu, Dong Qianqian, et al. CLUENER2020: fine-grained named entity recognition dataset and benchmark for Chinese [PP/OL]. V1. arXiv (2020-01-20)[2025-12-24]. https://doi.org/10.48550/arXiv.2001.04351.

[25] Yang Yi, Katiyar A. Simple and effective few-shot named entity recognition with structured nearest neighbor learning [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Stroudsburg: ACL, 2020: 6365-6375.
[26] Ma Jie, Ballesteros M, Doss S, et al. Label semantics for few shot named entity recognition[C]//Proceedings of the Findings of the Association for Computational Linguistics. Stroudsburg: ACL, 2022: 1956-1971.
[27] Shen Yongliang, Tan Zeqi, Wu Shuhui, et al. PromptNER: prompt locating and typing for named entity recognition[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg: ACL, 2023: 12492-12507.
[28] Chen Xiang, Li Lei, Deng Shumin, et al. LightNER: a lightweight tuning paradigm for low-resource NER *via* pluggable prompting[C]//Proceedings of the 29th International Conference on Computational Linguistics. Stroudsburg: ACL, 2022: 2374-2387.
[29] Chen Yanru, Zheng Yanan, Yang Zhilin. Prompt-based metric learning for few-shot NER[C]//Proceedings of the Findings of the Association for Computational Linguistics: ACL 2023. Stroudsburg: ACL, 2023: 7199-7212.

Chinese Medical Named Entity Recognition Based on Triple Hint and Contrast Learning

LIU Na[1,2], JI Zhe[1,2], WU Kedong[1,2], LIU Lei[1,2]

(1. School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China; 2. State Key Laboratory of Intelligent Processing of Graphics and Images, North Minzu University, Yinchuan 750021, China)

Abstract: To address the issues of insufficient pretrained knowledge representation, ambiguous boundaries, and category confusion in Chinese medical few-shot named entity recognition, a multi-stage recognition approach guided by triple prompts was proposed in this study. Firstly, a triple-prompt learning mechanism was constructed, in which a stepwise reasoning paradigm, comprising boundary localization, type discrimination, and contextual verification was adopted. By integrating the interpretability of discrete templates with the optimizability of continuous vectors, the capability of the model to parse entity structures was enhanced. Secondly, a bidirectional gating network was designed to dynamically regulate feature interactions between prompt information and textual representations, thereby strengthening semantic capture for complex terminology. Finally, contrastive learning was introduced to optimize the distribution of the feature space, improving intra-class compactness and inter-class separability. Experimental results indicated that $F1$ scores of 91.86%, 84.06%, and 88.93% were achieved on the IMCS-V2-NER, cMedQANER, and CCKS2019 datasets, respectively. These findings demonstrated that the proposed method consistently improved entity recognition performance and exhibited strong generalization capability under low-resource settings.

Keywords: small sample; named entity recognition; cue learning; feature interaction; contrast learning