

基于自适应决策边界的命名实体识别

王春雨, 樊媛媛, 梁庆中, 刘超, 李新川

(中国地质大学(武汉) 计算机学院, 湖北 武汉 430074)

摘要: 命名实体识别是自然语言处理的重要任务, 但实际应用中常面临类别分布不平衡的问题, 导致少数类实体难以准确识别, 影响模型整体性能。现有方法多依赖线性或球形决策边界假设, 难以刻画少数类样本在特征空间中的复杂分布, 限制了模型对少数类实体的判别能力。为解决上述问题, 提出一种基于自适应决策边界优化的命名实体识别方法 (ABO-NER)。该方法利用非线性判别函数构建灵活的决策边界, 并针对不同类别特征进行独立建模, 同时引入类内压缩与类间分离约束, 实现决策边界的自适应优化, 从而提升少数类实体的识别能力。在 ACE2005、GENIA、MOVIE、OntoNotes5.0、JNLPBA 和 Few-NERD 6 个公开 NER 数据集上的实验结果表明, 该方法在宏平均 F1 值和微平均 F1 值上均优于现有基线模型, 其中, 在 GENIA、JNLPBA 和 OntoNotes5.0 数据集的极端不平衡设置下, Macro-F1 值分别提升了 3.93%、3.60% 和 2.94%, 验证了该方法在类不平衡场景下的有效性与鲁棒性。

关键词: 命名实体识别; 类别不平衡; 自适应边界优化; 决策边界; 非线性判别函数; 少数类识别

DOI: 10.11907/rjdlk.261064

扫描二维码阅读全文:



中图分类号: TP391.1

文献标识码: A

文章编号: 1672-7800(XXXX)0XX-0001-11

Adaptive Decision Boundary Modeling for Named Entity Recognition

WANG Chunyu, FAN Yuanyuan, LIANG Qingzhong, LIU Chao, LI Xinchuan

(School of Computer Science, China University of Geosciences (Wuhan), Wuhan 430074, China)

Abstract: Named entity recognition is an important task in natural language processing. However, real-world applications often suffer from class imbalance, which makes minority entity types difficult to recognize and consequently affects the overall performance of NER models. Existing methods mainly rely on linear or hyperspherical decision boundary assumptions, which are insufficient to capture the complex feature distributions of minority samples in the representation space, thereby limiting the discriminative capability of models for minority entities. To address this issue, an adaptive boundary optimization-based named entity recognition method (ABO-NER) is proposed. The proposed method constructs flexible decision boundaries using nonlinear discriminative functions and independently models features for different entity classes. Meanwhile, dual constraints of intra-class compactness and inter-class separability are introduced to achieve adaptive optimization of the decision boundaries, thereby improving the recognition performance of minority entities. Experimental results on six public NER datasets (ACE2005, GENIA, MOVIE, OntoNotes5.0, JNLPBA and Few-NERD) demonstrate that our method outperforms all existing baseline models in both Macro-F1 and Micro-F1 scores. Specifically, the Macro-F1 score is increased by 3.93%, 3.60% and 2.94% on the GENIA, JNLPBA and OntoNotes5.0 datasets under extremely imbalanced settings, respectively, which verifies the effectiveness and robustness of our method in class-imbalanced scenarios.

Key Words: named entity recognition; class imbalance; adaptive boundary optimization; decision boundary; nonlinear discriminant function; minority class recognition

收稿日期: 2026-03-13

基金项目: 国家自然科学基金项目 (42571543); 湖北省自然科学基金项目 (2023AFA006)

作者简介: 王春雨 (2001-), 女, 中国地质大学 (武汉) 计算机学院硕士研究生, 研究方向为自然语言处理; 樊媛媛 (1979-), 女, 博士, 中国地质大学 (武汉) 计算机学院副教授、硕士生导师, 研究方向为智能计算、物联网、人工智能; 梁庆中 (1979-), 男, 博士, 中国地质大学 (武汉) 计算机学院副教授、硕士研究生导师, 研究方向为计算机网络、物联网应用, 知识图谱等; 刘超 (1979-), 男, 博士, 中国地质大学 (武汉) 计算机学院副教授、硕士生导师, 研究方向为知识图谱、深度学习与大数据应用; 李新川 (1976-), 男, 博士, 中国地质大学 (武汉) 计算机学院副教授, 研究方向为人工智能、自然语言处理。本文通讯作者: 樊媛媛。

0 引言

近年来,命名实体识别(Named Entity Recognition, NER)作为自然语言处理的核心基础任务,旨在从文本中识别命名实体并归类至预定义类型,是实体链接^[1-2]、关系抽取^[3-4]和知识图谱构建^[5]等任务的重要基础。随着研究的深入,NER的应用边界不断向更复杂的场景拓展,例如情感支持对话中的人物及情绪相关实体识别^[6]、社交媒体文本中的开放领域实体识别^[7]、公平性自然语言处理中的偏置实体识别问题^[8]以及多模态信息抽取^[9]等新兴方向。得益于神经网络的发展,序列标注^[10-11]、超图建模^[12-13]、序列到序列^[14-15]和跨度方法^[16]等技术不断推动NER性能提升。然而,现有方法大多依赖类别分布相对均衡的假设^[17],难以有效应对真实场景中普遍存在的类别不平衡问题。当少数类样本稀缺时,模型往往受多数类主导,导致对少数类实体的识别性能显著下降,从而限制了NER在真实应用场景中的可靠性与泛化能力。

为提升少数类实体识别性能,研究者提出了多种策略。早期方法多借鉴重采样机制^[18],或通过改进损失函数^[19-21]强化少数类贡献,但常伴随过拟合风险或多数类性能下降等问题^[22]。随后,有研究在NER中引入样本检索^[23]、因果推理^[17]、解耦训练方法^[24]等机制,但此类方法往往依赖外部结构或先验假设;另有研究则利用对比学

习^[25]提升类间可分性,但忽视类内一致性,可能导致少数类特征空间被压缩。总体而言,现有方法仍难以根治因分布不均衡引发的决策边界偏移问题。近期研究进一步指出,少数类识别困难不仅源于样本稀缺,更在于不平衡样本学习到的决策边界往往越来越偏向多数类,导致少数类数据样本被忽略^[26]。因此,优化决策边界、调整其形状与位置,正逐渐成为提升少数类判别能力的重要方向。

在决策边界优化研究中,已有一些代表性探索。Zhang等^[27]提出自适应决策边界方法,为每类样本学习球形边界以平衡经验风险与开放空间风险;其后又在DA-ADB框架^[28]中引入距离感知表征学习,实现了特征判别性与边界精度的联合优化。随后,研究逐步突破单一球形假设,发展出更复杂的边界建模。比如:Li等^[29]提出多粒度边界方法;Chen等^[30]通过三元组对比学习自适应校准类内外差异;Liu等^[31]则结合K-center对比学习并引入可调半径机制,使边界随类外样本分布动态扩缩。总体而言,边界优化方法已从球形边界逐渐演进到多粒度与对比学习相结合的方向,有效提升了少数类判别能力。然而,如图1所示,实际场景中的实体分布往往偏离球形假设,使得在球形边界下,不同类别样本容易侵入彼此的判别空间,导致分类混淆。在这种情况下,单一球形边界难以有效区分多数类与少数类;尤其当多数类占据主导空间时,即便进行边界调整,也难以避免少数类被误分类的风险。

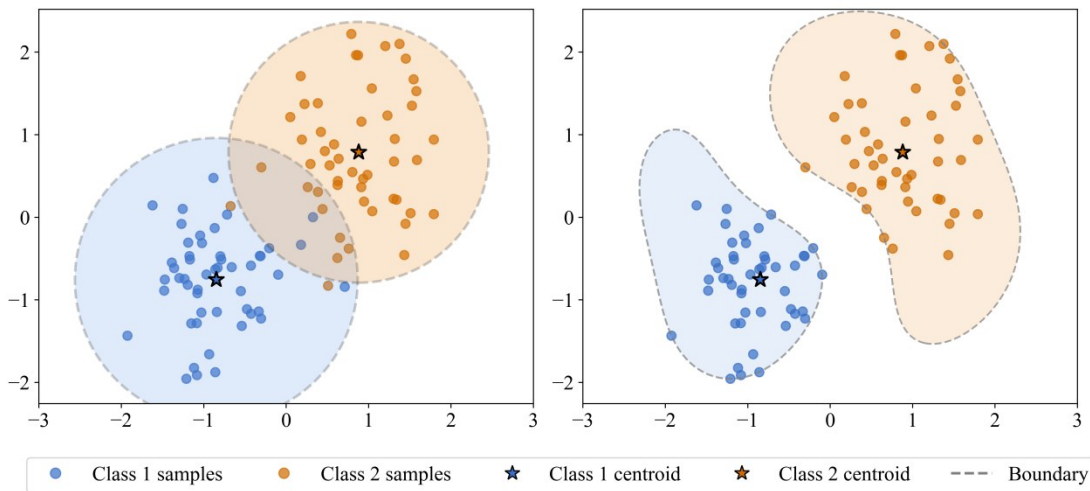


Fig. 1 Comparison of the category distribution and classification effect under different discrimination boundary assumptions

图1 不同判别边界假设下的类别分布与分类效果比较

为了解决该问题,本文提出了一种基于自适应边界优化的命名实体识别方法(ABO-NER)。该方法在特征空间中为各类别构建非线性决策边界,并通过基于边际约束的双边界损失进行优化,使少数类能够形成独立且稳健的判别区域,同时增强类间分离性,降低多数类对边界的干扰。ABO-NER所生成的边界能够更好地贴合实际分布,从而有效缓解少数类与多数类之间的混淆。整体框架包含两个核心模块:实体特征抽取与自适应边界优化。其中,实

体特征抽取模块采用多头双射解编码器生成候选跨度的三维特征矩阵,并结合UNet网络建模相邻跨度间的语义关系,以降低非实体噪声干扰^[32-33];自适应边界建模模块则针对各类别特征进行独立建模,并通过双边界损失实现动态优化。

本文主要贡献如下:①提出了一种自适应边界学习策略(ABO-NER),为各类别独立建模非线性判别边界,并引入基于边际约束的双边界损失,有效增强了少数类的判别

能力;②提出双边界协同优化机制,通过联合约束类内与类间边界,强化少数类的判别区域,从而提升其识别性能;③在多个公开数据集(ACE2005、GENIA、MOVIE、OntoNotes5.0、JNLPBA 和 Few-NERD)上验证了方法的有效性,实验结果显示,ABO-NER 在少数类识别上取得显著提升,同时保持整体性能稳定,为类不平衡文本任务提供可行方案。

1 相关工作

1.1 命名实体识别与类不平衡

命名实体识别是信息抽取的核心任务,其方法发展主要集中于复杂实体(如嵌套和不连续实体)的建模能力提升。序列标注方法通过为每个 token 分配标签(如 BIO)识别连续实体,但在处理嵌套或不连续结构时存在局限^[34]。Tang 等^[35]将 BIO 扩展为 BIOHD,以支持不连续实体。为突破复杂结构限制, Lu 等^[36]提出了基于超图的方法,将嵌套实体映射为超图子图进行分类,并进一步利用神经超图提升性能。与此同时,Seq2Seq 方法提供端到端解决方案,并结合数据增强或提示机制优化^[37-39]。但在长距离依赖下其易受级联错误影响^[40]。基于跨度的方法通过直接预测实体跨度,结合图结构^[41]、注意力机制^[42]及卷积操作^[43],有效增强复杂实体建模能力。例如, Yu 等^[44]利用双仿射注意力增强边界交互, Shen 等^[45]通过边界去噪提升鲁棒性。总体而言,这些方法虽在复杂实体建模上取得显著进展,但在不平衡分布数据中,少数类实体仍难以被有效捕捉,从而限制了整体识别性能。

针对 NER 中的少数类识别困难,研究者借鉴了不平衡学习策略以改善不平衡类别表现。在图像识别和文本分类中,数据重采样和损失加权方法被广泛应用,但在高度不平衡任务中仍难兼顾多数类和少数类性能^[46-49]。为此,解耦训练方法(如 CRT 和 τ -norm)通过将特征学习与分类器训练分离,从而降低类别分布对模型训练的干扰^[24]。监督对比学习通过拉近正样本、推远负样本,学习更具判别性的特征空间^[50]。Li 等^[51]进一步提出均匀目标引导训练,有效提升类不平衡类别性能。在 NER 任务中,少数类识别困难同样引发了研究者们的关注。为应对这一问题,研究者逐渐将不平衡学习策略应用于 NER。Wang 等^[23]通过训练样本检索机制增强少数类表示,缓解样本稀缺问题。Yu 等^[22]结合结构化提示与标签传播促进类别间的知识迁移,提高少数类泛化能力。Nan 等^[17]则从因果推理角度分析类不平衡瓶颈,为少数类识别困难提供理论解释。然而,复杂实体结构与类别不平衡问题叠加,使模型难以兼顾所有类别。因此,如何将不平衡策略与 NER 特有的复杂结构建模深度融合,仍是提升不平衡数据下 NER 性能的关键挑战。

1.2 自适应边界优化

在类别不平衡任务中,传统分类模型往往受到多数类样本的主导作用,使决策边界偏向多数类,从而削弱少数类的判别能力。通过学习更加合理的边界形态,模型能够在稀缺样本条件下更有效地区分少数类,缓解特征分布不均带来的性能瓶颈。近年来,研究者提出了多种边界建模方法以提升分类性能。早期工作中, Zhang 等^[27]提出自适应决策边界方法,并将其用于开放意图分类任务,为每个已知类别学习自适应球形边界,并通过设计损失函数在经验风险与开放空间风险之间取得平衡。随后, Zhang 等^[28]进一步提出 DA-ADB 框架,通过引入距离感知的表征学习,使模型能够同时获得区分性特征和精确边界。随后,研究逐渐突破单一球形假设,发展出更复杂的边界建模策略。Li 等^[29]提出多粒度开放意图分类方法,利用层次化的粒球聚类构建多粒度边界。Chen 等^[30]提出 TCAB 方法,通过三元组对比学习增强判别性,并自适应校准类内与类外样本差异,实现更精确的边界划分。Liu 等^[31]则提出 CLAB 方法,引入 K-center 对比学习以提升特征的均衡性,同时设计可调整半径机制,使边界能随类外样本分布动态扩缩。总体来看,边界优化已从单一球形到多粒度、再到对比学习结合的多样化方向演进,体现出其在应对复杂分布上的持续发展。

然而,这些方法普遍假设特征分布具有紧凑或类球形结构,而真实的 NER 特征分布往往呈现非凸甚至局部稀疏的特征,这使得传统边界难以充分刻画类别之间的差异。为此,本文提出了一种基于样本分布自适应的边界建模方法,能够动态重塑少数类周边的边界形态,从而在复杂不平衡场景下显著提升类不平衡实体的判别能力。

2 本文方法

问题定义:给定一个输入句子 $X = [x_1, x_2, \dots, x_n]$, 其中 X 由 n 个 token 组成, NER 任务的目标是抽取句子 X 中的所有实体,并将它们分类为预定义的类别。每个实体可以表示为一个元组 $s = (x_i, x_j, y_k)$ 。 x_i 和 x_j 分别表示实体的起始和结束边界索引 ($1 \leq i \leq j \leq n$), $y_k \in [1, \dots, C]$ 是实体类型, C 表示实体类型的数量。

本文方法整体框架如图 2 所示。该方法由两个主要组件构成:实体特征抽取(Entity Feature Extraction, EFE),如图 2(a)所示;自适应边界优化(Adaptive Boundary Optimization, ABO),如图 2(b)所示。其中, EFE 通过建模实体跨度之间的局部交互,捕捉丰富的上下文信息,生成高质量的跨度表示。随后, ABO 则在特征空间中动态调整类别边界,通过双边界约束增强类内紧凑性与类间分离性,从而缓解不平衡带来的决策偏移。

2.1 实体特征抽取

EFE 的目标是对句子进行编码,并通过 UNet 卷积神经

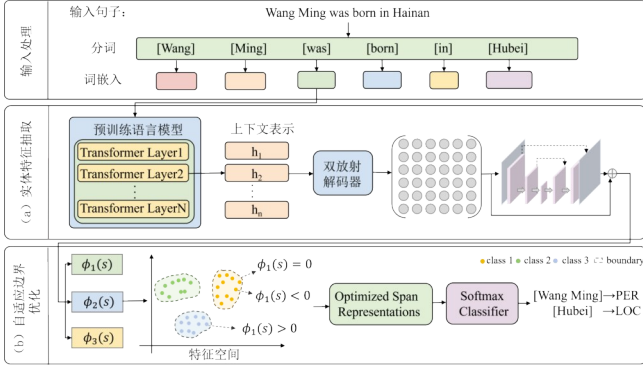


Fig. 2 Overall framework of ABO-NER

图2 ABO-NER整体框架

网络捕捉相邻跨度之间的相关性,从而提取高质量的实体跨度表示。

2.1.1 跨度特征编码

给定一个句子 $X = [x_1, x_2, \dots, x_n]$, EFE 首先使用预训练的语言模型对每个 token x_i 进行向量化,生成 token 级别的特征表示,如式(1)所示。

$$\mathbf{H} = \text{BERT}(X) \quad (1)$$

其中, $\mathbf{H} = [h_1, h_2, \dots, h_n] \in \mathbb{R}^{n \times d}$, d 为嵌入维度, n 表示句子中的 token 数量。考虑到实体的开始和结束上下文差异, EFE 采用两个独立的前馈神经网络获取一个 span 的起始和结束 token 的语义表示向量 $\mathbf{h}_s, \mathbf{h}_e \in \mathbb{R}^{n \times d}$, 如式(2)、式(3)所示。

$$\mathbf{h}_s = \text{LeakyReLU}(\mathbf{H}\mathbf{W}_s) \quad (2)$$

$$\mathbf{h}_e = \text{LeakyReLU}(\mathbf{H}\mathbf{W}_e) \quad (3)$$

其中, $\mathbf{W}_s, \mathbf{W}_e \in \mathbb{R}^{d \times h}$, h 为隐藏层大小, LeakyReLU 指激活函数。最后,使用多头 Biaffine 解码器以创建一个 $n \times n$ 的评分矩阵,如式(4)所示。

$$\mathbf{R} = \text{MHBiaffine}(\mathbf{h}_s, \mathbf{h}_e) \quad (4)$$

其中, r 为双放射模型的特征大小, $\text{MHBiaffine}(\cdot, \cdot)$ 表示多头双放射解码器,通过结合跨度的起始和结束表示,生成评分矩阵 $\mathbf{R} \in \mathbb{R}^{n \times n \times r}$, $\mathbf{R}$ 中的每个单元格 (i, j) 表示跨度的特征向量 $\mathbf{v} \in \mathbb{R}^r$ 。

2.1.2 作用于特征矩阵的 UNet 模型

然而,实体跨度之间往往存在嵌套与重叠等复杂结构,仅依赖起止位置的特征难以充分建模。为此,本文引入 UNet 网络结构,将评分矩阵视为二维特征图,通过卷积操作捕捉局部跨度依赖关系,从而进一步优化并增强跨度特征的表达能力。具体而言,首先将评分矩阵 $\mathbf{R} \in \mathbb{R}^{n \times n \times r}$ 输入到 UNet 的编码器中,通过多层卷积和池化操作逐步提取高层次的特征表示,同时降低特征图的分辨率以扩大感受野。假设经过 k 次池化操作后,特征图的维度变为 $\mathbf{R}_{en} \in \mathbb{R}^{n/2^k \times n/2^k \times d}$, 如式(5)所示。

$$\mathbf{R}_{en} = \text{UNet-Encoder}(\mathbf{R}) \quad (5)$$

其中, d 为编码后的特征维度。解码器部分则通过多层反卷积和上采样操作逐步恢复特征图的空间分辨率,最

终生成与输入尺寸相同的输出 $\mathbf{R}_{de}$ 。解码器的输出表示如式(6)所示。

$$\mathbf{R}_{de} = \text{UNet-Decoder}(\mathbf{R}_{en}) \quad (6)$$

其中, $\mathbf{R}_{de} \in \mathbb{R}^{n \times n \times r}$ 。通过这种设计, UNet 能够有效捕获不同尺度的信息,并生成高质量的跨度表示。

将由 UNet 输出的特征图 $\mathbf{R}_{de}$ 输入到一个分类器中,推理每个跨度的预测标签,如式(7)所示。

$$\mathbf{P} = \sigma(\mathbf{W}_o(\mathbf{R}_{de} + \mathbf{R}) + \mathbf{b}) \quad (7)$$

其中, $\mathbf{W}_o \in \mathbb{R}^{|\mathcal{C}| \times r}$ 和 $\mathbf{b} \in \mathbb{R}^{|\mathcal{C}|}$ 分别表示权重矩阵和偏置项, $\mathbf{P} \in \mathbb{R}^{n \times n \times |\mathcal{C}|}$, σ 表示 Sigmoid 激活函数。该模块采用二元交叉熵损失计算用于训练模型的损失函数,如式(8)所示。

$$L_{BCE} = -\sum_{0 \leq i, j < n} y_{ij} \log(P_{ij}) \quad (8)$$

其中, y_{ij} 表示真实标签,评分矩阵中的每个元素 P_{ij} 表示第 i 个起始位置和第 j 个结束位置之间的跨度得分。

2.2 自适应边界优化

ABO 通过为每个类别构建基于 SIREN 的可学习判别函数,精确刻画类别样本分布形状,以构建不规则决策边界,从而提升少数类实体在不平衡 NER 任务中的辨识能力和鲁棒性。

在获得高质量的实体跨度表示之后,为更细粒度地建模类别边界,本文采用“每类一个判别函数”的设计思路。对于每个类别 $c \in \{1, 2, \dots, C\}$, 单独定义一个可学习的判别函数 $\phi_c(s)$, 其输出值表示样本 s 相对于该类别判别边界的相对位置。每个 $\phi_c(s)$ 都对应一个类别特有的判别边界:若 $\phi_c(s) < 0$, 表示样本 s 属于类别 c , 即 s 落在该类别的边界内部;反之,若 $\phi_c(s) > 0$, 则表示样本 s 不属于该类别,即 s 位于边界外部。因此 $\phi_c(s) = 0$ 的零值等值面构成了类别 c 的决策边界。

为更有效地刻画复杂非线性边界形状,本文采用 SIREN (Sinusoidal Representation Networks) 作为判别函数的建模框架。SIREN 是一种以正弦函数作为激活函数的全连接网络,其形式如式(9)所示。

$$\phi_c(s) = f_c(h; \theta_c) = \mathbf{W}^{(L)} \cdot \sigma(\dots \sigma(\mathbf{W}^{(1)}s + \mathbf{b}^{(1)}) \dots) + \mathbf{b}^{(L)} \quad (9)$$

其中, $\sigma(x) = \sin(w_0 x)$ 为正弦激活函数, w_0 是频率因子, $\theta_c = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}_{l=1}^L$ 是类别 c 对应的模型参数。由于 SIREN 具备捕捉高维细节和复杂边界形状的能力,能够对每类数据分布进行独立建模与适应,进而构建灵活的边界。对于每个样本 s_i , 可得其在所有类别上的判别函数输出向量,如式(10)所示。

$$\boldsymbol{\phi}_i = [\phi_1(s_i), \phi_2(s_i), \dots, \phi_C(s_i)] \in \mathbb{R}^C \quad (10)$$

为进一步增强类别间边界的清晰度,引入一个基于 margin 约束的双边界损失函数,包含正类压缩项与负类排斥项两个部分。设第 i 个样本的真实类别为 y_i , 正类边界约束鼓励其判别输出满足 $\phi_c(s_i) \leq -m$, 即位于所属类别边界内部,并留有 margin 余量,从而提高分类置信度,如式

(11)所示。

$$L_{in} = \frac{1}{N} \sum_i^N (\phi_{c_i}(s_i) + m)^2 \quad (11)$$

负类边界约束鼓励 $\phi_c(s_i \geq +m)$ (当 $c \neq y_i$), 拉开与非目标类别边界的距离, 从而增强类别间的分离性, 如式(12)所示。

$$L_{out} = \frac{1}{N(C-1)} \sum_{c_i \in C} \sum_i^N (\phi_{c_i}(s_i) - m)^2 \quad (12)$$

最终, 整体目标函数为两项损失的加权组合, 如式(13)所示。

$$L_{ABO} = L_{in} + \lambda \cdot L_{out} \quad (13)$$

其中, 超参数 λ 控制类内压缩与类间分离的权重平衡。该损失函数通过引入每类独立的判别函数 $\phi_c(s)$ 和 margin 约束机制, 显式建构了类内收缩、类间分离的边界学习框架, 能够更准确地捕捉少数类样本复杂的分布特性, 从而提升模型对少数类别的判别能力与整体鲁棒性。

2.3 训练与推断

ABO-NER 的优化目标是通过联合优化实体特征抽取(EFE)和自适应边界优化(ABO), 提升模型在类别不平衡场景下的分类性能。具体而言, 总损失函数由两部分组成, 如式(14)所示。

$$L = L_{BCE} + L_{ABO} \quad (14)$$

其中, L_{BCE} 为二元交叉熵损失, 用于优化跨度分类任务; L_{ABO} 为优化自适应边界损失, 用于构建类别特定的非线性决策边界。

在推理阶段, ABO-NER 首先通过 UNet 网络提取融合了相邻跨度语义关系的高质量实体特征。随后, 利用训练好的 SIREN 判别函数, 对每个实体跨度在各类别对应的边界位置进行评估。模型将落在类别边界内部的候选类别筛选出来, 并在这些候选类别中选择预测概率最高的类别作为最终标签。该策略能够有效提升少数类实体在不平衡数据场景下的识别准确率。

3 实验设置

3.1 数据集

为全面评估所提出 ABO-NER 方法在不同数据分布条件下的性能, 本文在 5 个公开 NER 数据集上开展实验, 包括 MOVIE、GENIA^[52]、JNLPBA^[53]、OntoNotes5.0^[54] 和 ACE2005^[44]。此外, 为进一步验证模型在大规模细粒度实体识别任务中的泛化能力, 本文还引入了细粒度长尾数据集 FewNERD^[55]。上述数据集覆盖多个领域, 并包含不同类型的实体结构。其中, ACE200 和 GENIA 属于嵌套型 NER 数据集, 而 MOVIE、OntoNotes5.0、FewNERD 与 JNLPBA 则为平面型 NER 数据集。各数据集统计信息如表 1 所示。

为系统研究模型在类别不平衡场景下的表现, 本文在

Table 1 Basic statistics of the dataset

表 1 数据集基本统计

数据集	领域	类型	句子数	实体数
MOVIE	电影	平面	9 775	21 295
OntoNotes5.0	混合	平面	75 187	81 823
ACE2005	新闻	嵌套	7 291	25 300
GENIA	生物学	嵌套	15 038	46 203
JNLPBA	生物学	平面	16 806	46 748
Few-NERD	混合	平面	131 767	340 387

MOVIE、OntoNotes5.0、ACE2005、GENIA 和 JNLPBA 5 个数据集的训练集上构建不同类别分布的数据版本。具体而言, 采用类不平衡比率 ρ 衡量数据集的类别分布, 其中 ρ 定义为最大类别与最小类别样本数量之比。基于此, 从原始训练集中构建 3 种数据分布设置: 类平衡 ($\rho=1$)、中度不平衡 ($\rho=10$) 和极端不平衡 ($\rho=100$)。在构造过程中, 开发集和测试集保持不变, 并设置一个类别为多数类, 其余类别为少数类, 同时保证所有少数类具有相同的样本数量。进一步地, 为评估模型在真实数据分布条件下的性能, 本文在 OntoNotes5.0、ACE2005 和 Few-NERD 数据集的原始训练集上进行标准 NER 实验, 以验证所提出方法在不同规模和任务场景下的泛化能力。不同不平衡比例下的类别样本分布如表 2 所示。

Table 2 Number of class samples under different ρ

表 2 不同 ρ 值下的类别样本数量

ρ 值	类别	MOVIE	OntoNotes5.0	ACE2005	GENIA	JNLPBA
$\rho=1$	单类别	113	292	200	293	747
	样本数					
	多数类	TITLE:			Protein:	Protein:
$\rho=10$	样本数	650	PER: 1950	PER: 870	1040	2670
	少数类	Other: 65	Other: 195	Other: 87	Other: 104	Other: 267
	样本数					
$\rho=100$	多数类	TITLE:		PER:	Protein:	Protein:
	样本数	1200	PER: 4500	1300	1400	3600
	少数类	Other: 12	Other: 45	Other: 13	Other: 14	Other: 36
	样本数					

3.2 基线模型

为验证 ABO-NER 框架的有效性, 本文选取 12 种模型作为基线模型, 具体分为 3 类。

(1) 结构建模方法类。该方法通过网络结构设计, 优化实体边界建模与跨度级语义交互效果。BARTNER^[15] 采用序列到序列框架, 实现扁平、嵌套与不连续命名实体识别的统一建模; POINTER^[56] 将嵌套命名实体识别转化为句法分析问题; BS^[57] 引入边界平滑策略以缓解标注模糊问题, DS^[54] 则采用融合跨层聚合机制的深度跨度 Transformer 模型; BOPN^[58] 实现面向实体类型感知的边界偏移预测, CNN-NER^[44] 通过对跨度得分矩阵的卷积操作, 建模局部空间交互特征。

(2) 表示学习与对比学习方法类。该方法聚焦于特征表示优化, 以提升类别间的可分性。DiFiNet^[40] 通过边界感知的语义分化策略降低特征噪声; EASAL^[59] 融合主动

学习与伪标签技术,提升低频实体的利用效率;BINDER^[60]构建双编码器对比学习框架,结合动态阈值损失函数,实现对不同结构类型实体的联合建模。

(3)不平衡学习方法类。该方法专门针对命名实体识别中的类别不平衡问题展开研究。LongTailIE^[22]通过标签分组与实例重构策略缓解长尾效应;kNN-NER^[23]借助相似实例检索,增强少数类别的特征表示效果;MProto^[61]采用融合去噪最优传输的多原型网络,解决远距离监督中的噪声问题,是典型的基于原型的学习方法范式。

3.3 实验细节

在实验中,不同数据集使用了与其领域相匹配的预训练模型:ACE2005、MOVIE 和 OntoNotes5.0 采用 RoBERTa-base,而生物医学领域的 GENIA 和 JNLPBA 数据集则采用 BioBERT-base。模型训练的主要超参数如下:双仿射解码器的仿射层隐藏维度设为 200,Dropout 率为 0.4;学习率初始化为 $2e-5$,并采用线性衰减策略;权重衰减系数设为 $1e-2$ 。训练过程中,批大小为 16,实体嵌入维度为 25,实体过滤阈值为 0.5,卷积核大小为 3。所有参数均使用 AdamW 优化器进行优化。

为了保证实验结果的稳定性,每个模型在相同配置下独立运行 3 次,并报告不同随机种子下的平均结果。为确保公平比较,本文研究在 ACE2005、MOVIE、OntoNotes5.0、GENIA、FewNERD 和 JNLPBA 上使用了原始研究中推荐的参数;而对于未在这些数据集上进行过实验的基线方法,则统一采用本文实验中所使用的相同参数进行训练和评估。

4 实验与结果分析

4.1 主实验

ABO-NER 与现有基线模型在嵌套和平面数据集上的实验结果如表 3—表 5 所示,其中粗体和下划线分别表示 F1 得分的最佳和次佳。整体来看,ABO-NER 在不同数据集和各类不平衡比例($\rho=100, 10, 1$)下均表现出稳定且优

异的性能,验证了其在复杂场景中的适应性。

在嵌套数据集上,ABO-NER 的优势尤为突出。以 ACE2005 为例,当 $\rho=100$ 时,其 micro-F1 为 66.84%,较 CNN-NER (62.95%)提升了近 4 个百分点,表明基于决策边界优化的机制能够有效弥补卷积模型在边界模糊类上的不足。在 GENIA 数据集 $\rho=100$ 的场景下,ABO-NER 的 macro-F1 为 54.39%,较 POINTER (50.46%)提升了近 4 个百分点。值得注意的是,随着不平衡程度加剧(ρ 增大),ABO-NER 的性能优势持续放大:在 $\rho=100$ 时,嵌套数据集的平均 macro-F1 提升幅度达到 3%,远超其他模型,进一步验证了多尺度特征融合与非线性边界建模对复杂实体结构的有效适配。

在平面数据集上,ABO-NER 同样保持领先。例如:在 MOVIE 数据集 $\rho=100$ 的条件下,其 macro-F1 (43.29%)和 micro-F1 (55.60%)均优于 CNN-NER (39.04%、51.69%),且 macro-F1 较 MProto 提升了 5.76 个百分点。尽管 MProto 通过原型聚合提升类内一致性,但在全监督场景下,其性能仅优于部分弱基线,明显低于 ABO-NER。即便在相对平衡的场景下,ABO-NER 仍展现出竞争力。例如:在 OntoNotes5.0 数据集 $\rho=1$ 时,其 macro-F1 (63.22%)与 BINDER (63.54%)接近,但 micro-F1 (78.14%)依然更优,说明边界优化机制在平衡分布下同样能够增强特征判别能力。

4.2 原始数据分布实验

为进一步验证模型在真实数据分布下的泛化能力,本文在 ACE2005、OntoNotes5.0 和 FewNERD 数据集的原始训练集上进行了标准 NER 实验,实验结果如表 6 所示,†表示直接引用原论文中的实验结果。

从结果可以看出,ABO-NER 在多个数据集上均取得了更优性能。在 ACE2005 数据集上,ABO-NER 的 macro-F1 和 micro-F1 分别达到了 82.67% 和 89.62%,相比 DiFiNet 和 MProto 等基线模型均取得了一定提升。在 OntoNotes5.0 数据集上,ABO-NER 同样取得更高的 micro-F1 (91.78%),表明该方法在大规模开放域 NER 任务中具有良好的泛化能力。此外,在具有大量细粒度实体类型的 FewNERD 数

Table 3 Results on the nested NER datasets ACE2005 and GENIA

表 3 嵌套数据集 ACE2005 和 GENIA 的结果

(%)

数据集	ACE2005						GENIA					
ρ	100		10		1		100		10		1	
	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro
BARTNER	16.36	49.53	23.32	52.84	33.10	59.38	11.70	43.29	28.47	50.87	33.27	54.19
POINTER	37.26	48.86	48.37	69.48	28.83	71.35	<u>50.46</u>	<u>60.95</u>	52.22	63.22	52.09	64.40
CNN-NER	40.70	62.95	57.61	70.61	<u>65.07</u>	73.94	49.01	60.45	<u>70.64</u>	<u>71.46</u>	<u>70.16</u>	<u>67.37</u>
BS	39.27	60.56	55.77	<u>71.36</u>	61.16	72.14	37.26	56.30	69.34	71.47	69.99	66.47
BOPN	41.67	57.69	<u>58.76</u>	69.13	60.80	68.35	50.39	51.90	56.17	57.86	53.61	54.33
BINDER	36.87	58.21	51.69	65.86	61.51	65.92	46.17	56.45	69.47	70.29	69.42	64.47
DS	<u>44.34</u>	<u>64.88</u>	55.07	69.77	63.13	74.62	41.27	54.88	68.25	68.74	68.87	65.85
DiFiNet	39.12	62.78	55.19	70.40	64.86	<u>75.16</u>	44.61	59.48	69.57	70.59	68.79	66.27
LongTailIE	35.99	53.74	51.83	65.37	58.61	66.06	24.93	48.18	60.90	64.07	61.42	61.54
MProto	40.21	61.17	54.10	69.21	62.35	72.51	47.57	59.20	68.22	70.03	68.54	66.10
ABO-NER	47.04	66.84	59.01	73.05	66.30	75.90	54.39	62.13	71.75	72.68	70.34	68.85

Table 4 Results on the flat NER datasets MOVIE and OntoNotes 5.0

表 4 平面数据集 MOVIE 和 OntoNotes5.0 的结果 (%)

数据集	MOVIE						OntoNotes5.0					
	ρ		100		10		1		100		10	
	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro
BARTNER	07.06	18.39	07.47	19.23	35.25	48.70	30.66	41.97	39.21	51.30	35.72	50.56
POINTER	16.67	43.99	53.49	69.72	38.86	65.54	51.99	68.32	51.31	73.51	62.50	75.68
CNN-NER	39.04	<u>51.69</u>	60.54	68.39	59.61	67.53	56.47	<u>70.45</u>	<u>60.47</u>	73.73	61.36	77.72
BS	<u>41.35</u>	48.26	62.02	68.56	60.74	67.84	52.83	68.50	57.70	<u>75.77</u>	60.07	76.58
BOPN	11.98	19.05	51.94	60.49	54.18	61.98	23.09	39.08	59.55	67.63	57.27	67.06
BINDER	43.19	47.25	63.04	<u>70.32</u>	64.01	<u>69.06</u>	<u>57.75</u>	68.90	58.76	74.16	63.54	76.65
DS	35.71	46.35	<u>63.19</u>	70.02	59.43	64.30	53.48	68.96	59.84	75.90	59.44	77.26
EASAL	22.29	33.38	51.69	62.33	53.06	57.97	45.54	62.19	49.33	63.65	45.54	62.19
DiFiNet	40.00	51.09	62.20	69.94	59.76	67.71	54.39	68.52	58.99	74.57	61.39	77.86
KNN-NER	33.21	39.04	54.30	64.46	48.67	60.30	44.53	56.69	53.31	72.05	55.46	73.68
LongTailIE	34.21	42.92	57.33	66.60	57.60	67.22	52.04	67.32	59.09	75.56	59.99	<u>77.88</u>
MProto	37.53	50.20	58.25	66.56	57.86	65.31	55.10	69.28	59.00	72.01	60.13	76.52
ABO-NER	43.29	55.60	64.94	72.86	<u>61.87</u>	69.82	60.69	73.46	62.20	<u>77.25</u>	<u>63.22</u>	78.14

Table 5 Results on the flat NER dataset JNLPBA

表 5 平面数据集 JNLPBA 的结果 (%)

数据集	JNLPBA					
	ρ		100		10	
	macro	micro	macro	micro	macro	micro
BARTNER	34.52	55.63	35.09	58.07	36.98	63.84
POINTER	46.07	51.16	37.21	62.62	50.76	65.50
CNN-NER	53.75	63.46	66.84	71.06	65.63	68.87
BS	53.86	63.08	66.64	71.10	68.20	72.28
BOPN	40.08	53.91	65.90	70.16	63.43	67.59
BINDER	<u>55.09</u>	60.71	67.09	70.37	64.55	67.80
DS	49.33	59.51	63.47	68.20	63.24	65.82
EASAL	34.52	55.63	50.93	67.15	50.94	63.04
DiFiNet	54.79	63.58	<u>67.17</u>	<u>71.82</u>	64.99	<u>69.31</u>
KNN-NER	53.43	<u>63.78</u>	62.10	68.82	63.26	65.81
LongTailIE	39.31	53.46	60.31	65.48	58.02	61.63
MProto	52.34	62.10	65.15	69.82	64.29	67.53
ABO-NER	58.69	65.37	67.68	72.43	<u>66.37</u>	69.10

据集上, ABO-NER 仍保持稳定性能, 相比 LongTailIE 等方法取得更优结果。这表明, 所提出的自适应决策边界不仅能够缓解类别不平衡问题, 还能够在大规模细粒度实体识别任务中保持稳定表现。

Table 6 Experimental results on ACE2005, OntoNotes5.0, and Few-NERD datasets

表 6 ACE2005, OntoNotes5.0 和 Few-NERD 数据集上的实验结果 (%)

数据集	ACE2005		OntoNotes5.0		Few-NERD	
	macro-F1	micro-F1	macro-F1	micro-F1	macro-F1	micro-F1
DiFiNet	80.67	88.94†	81.57	89.51	77.32	80.43
LongTailIE	75.48	84.37	79.83	87.74	62.60†	67.80†
MProto	78.36	86.35	81.46	90.32	74.80	79.62
ABO-NER	82.67	89.62	83.57	91.78	78.82	81.14

4.3 消融实验

为深入理解 ABO-NER 中各组成模块的作用, 在多个基准数据集上进行了系统的消融实验。具体包括: ①w/o L_{in} : 去除类内压缩项, 仅保留类间分离项; ②w/o L_{out} : 去除类间分离约束, 仅保留压缩项; ③w/o L_{ABO} : 去除基于 margin 的双边界优化损失, 替换为基本交叉熵损失 CE; ④w/o

ABO: 同时去掉判别函数与基于 margin 的双边界优化损失; ⑤MLP: 将判别网络替换为标准 MLP, 以考察 SIREN 判别头的贡献。

表 7—表 9 的消融实验表明, ABO 机制中的判别函数与双边界优化损失在不同不平衡场景下均发挥关键作用, 且协同效应显著。极端不平衡($\rho=100$)时组件移除的性能衰减最为突出。例如, GENIA 数据集上, ABO-NER 的 macro-F1 为 54.39%, 而移除类内压缩(w/o L_{in})后降至 43.75%, 差距达 10.6 个百分点, 凸显 L_{in} 在聚合少数类样本中的重要性; 在 ACE2005 上, 移除类间分离(w/o L_{out})后 macro-F1 从 47.04% 降至 44.84%, 表明 L_{out} 对强化类别边界不可或缺。若将双边界优化替换为 CE 损失(w/o L_{ABO}), 性能同样下降, 例如 MOVIE 数据集 $\rho=100$ 时, macro-F1 较完整模型的要低 1.19%。完全移除 ABO(w/oABO)时的性能普遍垫底, 在 MOVIE 上仅为 39.04%, 验证整体框架的不可替代性。此外, MLP 在 ACE2005 上的 macro-F1 (43.49%) 与完整模型(47.04%)的差距证实了 SIREN 凭借周期激活函数的高频拟合能力。但在平衡场景中存在个别特例, 如在 JNLPBA 上, w/o L_{in} 的 micro-F1 为 69.31%, 略高于 ABO-NER 的 69.10%。这可能是因为在样本分布均衡时, 类间分离需求减弱, 基础分类机制即可维持性能, 但未改变 ABO 机制在多数场景下的优势。

4.4 判别网络的选择讨论

为验证 SIREN 作为判别函数的有效性, 选取 3 种典型非线性网络作为替代方案进行对比。①MLP: 采用 ReLU 激活函数的标准全连接网络, 作为基准; ②FourierMLP: 在输入层引入傅里叶特征映射, 以提升高频特征表达能力; ③RBFNetwork: 依靠高斯核函数实现局部敏感的特征响应, 强调局部邻域的边界刻画能力。通过这些替代方案, 可以全面比较不同函数在边界建模上的表现差异。

实验结果如表 10 所示。在 MOVIE 与 ACE2005 数据集上, 不同判别函数在宏/微平均指标上均呈现出明显差异。

Table 7 Ablation study results on the ACE2005 and GENIA datasets

表 7 在 ACE2005 和 GENIA 数据集上的消融结果

(%)

ρ	GENIA						ACE2005					
	100		10		1		100		10		1	
	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro
MLP	<u>50.24</u>	<u>60.89</u>	70.10	<u>71.91</u>	69.29	67.08	43.49	65.20	57.11	<u>71.22</u>	66.62	74.67
w/o L_{out}	48.89	60.33	70.02	71.23	70.95	<u>67.87</u>	<u>44.84</u>	<u>65.48</u>	<u>58.40</u>	71.08	65.63	75.03
w/o L_{in}	43.75	58.97	<u>70.42</u>	71.46	68.95	65.80	43.19	64.50	55.11	71.11	63.24	<u>75.37</u>
w/o L_{ABO}	49.80	59.96	69.35	71.00	70.08	67.22	42.33	63.96	58.07	71.01	65.98	74.43
w/o ABO	48.81	60.45	69.61	71.46	69.98	67.37	41.53	62.95	55.63	70.61	63.77	73.94
ABO-NER	54.39	62.13	71.75	72.68	<u>70.34</u>	68.85	47.04	66.84	59.01	73.05	<u>66.30</u>	75.90

Table 8 Ablation study results on the MOVIE 和 OntoNotes 5.0 datasets

表 8 在 MOVIE 和 OntoNotes5.0 数据集上的消融结果

(%)

ρ	MOVIE						OntoNotes5.0					
	100		10		1		100		10		1	
	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro
MLP	40.10	51.95	62.60	70.12	61.10	<u>68.63</u>	57.41	72.00	<u>61.77</u>	77.20	61.43	78.13
w/o L_{out}	<u>42.16</u>	<u>54.86</u>	<u>63.39</u>	<u>70.59</u>	<u>61.64</u>	68.62	57.20	72.12	60.15	75.43	60.81	76.00
w/o L_{in}	40.57	52.64	62.06	69.81	60.80	67.68	59.07	72.15	61.10	76.97	<u>62.22</u>	78.27
w/o L_{ABO}	42.10	52.35	61.72	70.35	61.56	68.51	<u>59.23</u>	<u>72.19</u>	61.67	78.09	60.62	76.66
w/o ABO	39.04	51.69	60.54	68.39	59.61	67.53	56.47	70.45	60.47	73.73	61.36	77.72
ABO-NER	43.29	55.60	64.94	72.86	61.87	69.82	60.69	73.46	62.20	<u>77.25</u>	63.22	<u>78.14</u>

Table 9 Ablation study results on the JNLPBA dataset

表 9 在 JNLPBA 数据集上的消融结果

(%)

ρ	JNLPBA					
	100		10		1	
	macro	micro	macro	micro	macro	micro
MLP	53.30	62.77	67.10	71.35	65.99	<u>69.61</u>
w/o L_{out}	53.95	64.54	<u>67.40</u>	71.69	65.05	69.29
w/o L_{in}	<u>56.28</u>	<u>64.61</u>	67.34	<u>71.71</u>	66.40	69.31
w/o L_{ABO}	55.26	64.04	65.65	70.71	65.43	69.85
w/o ABO	53.75	63.46	66.84	71.06	65.63	68.87
ABO-NER	58.69	65.37	67.68	72.43	<u>66.37</u>	69.10

整体而言, ABO-NER 基于 SIREN 的判别函数在多数场景下保持最优或接近最优结果。在 $\rho=100$ 的不平衡场景下,

ABO-NER 的 macro-F1 在 MOVIE 和 ACE2005 上分别达到了 43.29% 和 47.04%, 显著优于 MLP (提升 3.19% 和 3.55%) 以及 RBFNetwork (提升 0.68% 和 4.35%)。这说明正弦激活函数在稀疏样本下仍能保持良好的梯度传播与高频表达能力, 能够有效刻画少数类复杂且不规则的边界分布。在 $\rho=10$ 的中度不平衡下, ABO-NER 依然在宏/微平均指标上整体领先, 例如在 MOVIE 数据集上, 其 micro-F1 达到了 72.86%, 高于 FourierMLP (71.56%) 与 RBFNetwork (71.30%)。值得注意的是, 在 ACE2005 数据集 $\rho=1$ 时, FourierMLP 的 micro-F1 达到了 76.20%, 略高于 ABO-NER 的 75.90%, 虽然频域映射有助于提升整体判别能力, 但其指标波动较大, 对少数类适应性不足。

Table 10 Performance comparison of networks with different discriminant functions

表 10 不同判别函数网络性能比较

(%)

ρ	MOVIE						ACE2005					
	100		10		1		100		10		1	
	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro	macro	micro
MLP	40.10	51.95	62.60	70.12	61.10	68.63	<u>43.49</u>	<u>65.20</u>	57.11	71.22	66.62	74.67
FourierMLP	30.70	47.67	<u>63.95</u>	<u>71.56</u>	<u>61.64</u>	<u>68.94</u>	43.06	64.71	56.70	71.51	67.09	76.20
RBFNetwork	<u>42.61</u>	<u>54.88</u>	62.92	71.30	59.08	67.03	42.69	64.31	<u>57.48</u>	<u>71.61</u>	64.89	74.51
ABO-NER	43.29	55.60	64.94	72.86	61.87	69.82	47.04	66.84	59.01	73.05	<u>66.30</u>	<u>75.90</u>

4.5 参数敏感性分析

在 ABO-NER 模型中, margin 参数 m 用于控制类别判别函数的边界松弛程度。具体而言, 较大的 m 值会在类内收缩约束中强制样本距离其对应类别边界更远, 从而提升分类置信度。同时, 在类间分离约束中, 较大的 m 值也会增强样本与非目标类别的边界间隔。因此, m 在平衡类内紧凑性与类间可分性方面至关重要。

从图 3 和图 4 可以看出, 不同数据集下, ABO-NER 的 Micro-F1 与 m 呈非线性关系。以 ACE2005 数据集为例, 当

$\rho=10$ 时, Micro-F1 在 $m \in [0, 1]$ 区间内波动于 70%~74% 之间, 说明 m 过小或过大均会破坏类内压缩与类间分离之间的平衡。而且极端不平衡场景下, m 的影响尤为显著。例如, 在 MOVIE 数据集中, 当 $\rho=100$ 时 Micro-F1 的波动幅度高达 5%; 而在 $\rho=1$ 时波动幅度收窄至 1% 左右。与此同时, 在 ACE2005 数据集的中度不平衡 ($\rho=10$) 时, 模型在 $m \in [0.3, 0.4]$ 区间内取得最优性能 (Micro-F1 约为 73%); 而在极端不平衡 ($\rho=100$) 下, 最优 m 区间转移至 $m \in [0.8, 1.0]$ 。这一偏移说明, 在严重不平衡场景下需要更大的 m 以强化

类别边界的分离能力。综上所述,margin 参数 m 对 ABO-
NER 在不同类别不平衡场景下的表现影响显著:较小的 m
适用于平衡或轻度不平衡场景,较大的 m 更适配极端不平
衡场景。这表明,合理设置 m 可在稳定多数类性能的同时
提升少数类识别能力,增强模型整体鲁棒性。

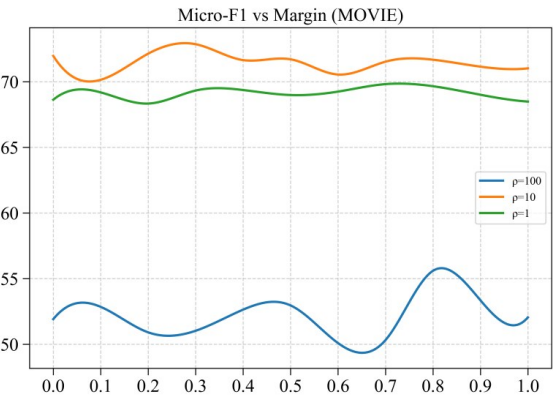


Fig. 3 The impact of Margin (m) on the MOVIE dataset
图3 Margin (m)对MOVIE数据集的影响

4.6 实体向量的可视化分析

为进一步验证本文方法在特征判别性上的优势,在
GENIA 数据集上进行了定性分析。具体地,分别提取 $w/$
 oL_{ABO} 、 $w/oABO$ 以及完整 ABO-NER 模型学习到的实体表
示,并采用 t-SNE 将其映射至二维空间以便可视化。图 5

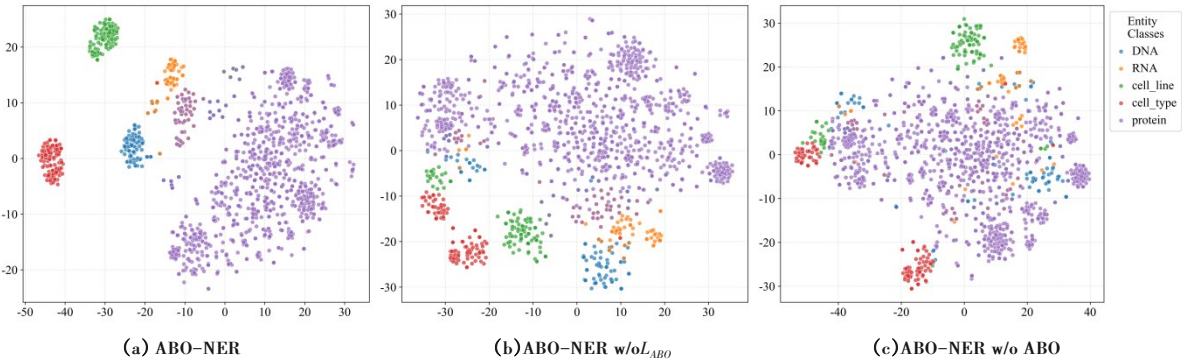


Fig. 5 Spatial distribution of entity features in the GENIA dataset with an imbalance ratio of 100
图5 不平衡比为 100 时 GENIA 数据集的实体特征空间分布

4.7 实体类别性能表现

为了评估 ABO-NER 在不同实体类型上的识别性能,
本文在 GENIA 和 ACE2005 数据集上,对比了 ABO-NER 与
强基线模型 CNN-NER 在不同类别不平衡比例($\rho=100$ 、10、
1)下的表现。

表 11 和表 12 显示,在不同类别不平衡比例下,ABO-
NER 的宏平均 F1 均高于 CNN-NER,且在极端不平衡情况
下优势更为明显。例如,在 GENIA 数据集上,宏平均 F1 达
到了 54.39%,相比 CNN-NER 的 49.01% 提升了 5.38 个百分
点。尤其值得注意的是,ABO-NER 对少数类实体的识别
能力显著增强:当 $\rho=100$ 时,ACE2005 中 VEH 少数类别
的 F1 从 29.58% 提升至 40.76%,GENIA 中 Cell_Type 少数类
别的 F1 从 24.00% 提升至 40.27%。与此同时,多数类表现保

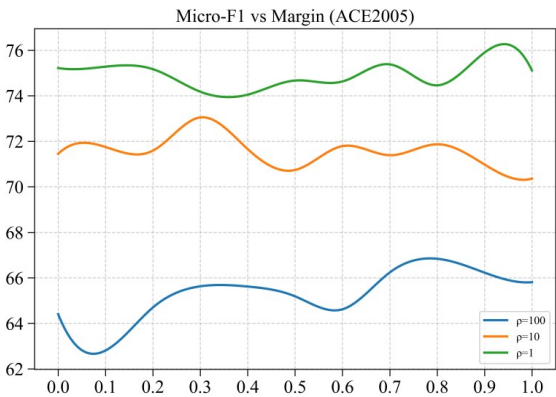


Fig. 4 The impact of Margin (m) on the ACE2005 dataset
图4 Margin (m)对ACE2005数据集的影响

给出了 3 种模型的特征分布情况。可以观察到,在 w/o
ABO 设置下,各类别样本分布高度混杂,类间边界难以区
分;在 w/oL_{ABO} 条件下,虽然类簇初步形成,但不同类别间仍
存在显著重叠,边界模糊不清。相比之下,完整的 ABO-
NER 模型在特征空间中表现出更紧凑的类内聚集和更明
确的类间分离,尤其少数类样本能够形成清晰的独立簇。
这一结果直观地验证了双边界优化与判别函数协同作用
的有效性,进一步说明 ABO-NER 在应对类别不平衡与复
杂边界建模时的优势。

持稳定,如同一设置下 ACE2005 的 PER 类别 F1 达到了
80.73%(CNN-NER 为 78.74%),GENIA 的 Protein 类别 F1
达到了 75.97%(高于 72.64%)。这些结果表明,ABO-NER 能
够有效缓解类别不平衡对少数类识别的影响,同时保持多
数类的高性能。

Table 11 F1 scores for each entity type in the ACE2005 dataset
表 11 ACE2005数据集各实体类型的F1值 (%)

		PER	GPE	LOC	ORG	FAC	VEH	WEA	ALL
$\rho=100$	CNN-NER	78.74	37.37	30.91	14.54	39.42	29.58	54.32	40.70
	ABO-NER	80.73	45.40	33.06	31.09	43.24	40.76	55.00	47.04
$\rho=10$	CNN-NER	81.65	59.23	49.21	59.61	44.04	51.06	58.49	57.61
	ABO-NER	82.55	60.08	52.17	61.36	45.45	54.84	56.60	59.01
$\rho=1$	CNN-NER	80.58	74.36	55.28	70.19	60.30	47.51	67.24	65.07
	ABO-NER	81.79	76.87	51.01	71.44	62.33	48.63	72.00	66.30

Table 12 F1 scores for each entity type in the GENIA dataset

表 12 GENIA 数据集各实体类型的 F1 值 (%)

		Protein	DNA	RNA	Cell_Li ne	Cell_Ty pe	ALL
$\rho=100$	CNN-NER	72.64	32.28	66.30	49.84	24.00	49.01
	ABO-NER	75.97	34.28	68.13	53.29	40.27	54.39
$\rho=10$	CNN-NER	75.79	62.33	79.11	67.64	68.31	70.64
	ABO-NER	76.57	64.47	81.08	68.41	68.22	71.75
$\rho=1$	CNN-NER	65.73	68.22	76.79	69.91	70.13	70.16
	ABO-NER	67.37	69.10	78.00	68.49	68.74	70.34

5 结语

本文针对命名实体识别任务中普遍存在的类别不平衡问题,提出了一种基于自适应边界优化的 NER 方法 (ABO-NER)。该方法通过引入基于 SIREN 的可学习判别函数,精确建模复杂且类别特定的决策边界,并结合类内压缩与类间分离的双边界损失,有效缓解了决策边界向多数类偏移的问题,从而显著提升了少数类实体的识别能力。在多个基准数据集上的实验结果表明,ABO-NER 在不平衡设置下优于现有强基线方法,同时在提升少数类识别性能的同时保持多数类表现稳定。未来研究可进一步探索命名实体识别在复杂交互场景与公平可信自然语言处理中的应用,并结合多源信息与更灵活的边界建模机制,以提升模型在真实不平衡分布环境下的鲁棒性与泛化能力。

参考文献:

- [1] HOU F, WANG R, HE J, et al. Improving entity linking through semantic reinforced entity embeddings[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 6843–6848.
- [2] GU Y, QU X, WANG Z, et al. Read, retrospect, select: an MRC framework to short text entity linking[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 12920–12928.
- [3] LI J, XU K, LI F, et al. MRN: a locally and globally mention-based reasoning network for document-level relation extraction[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, 2021: 1359–1370.
- [4] RATHORE V, BADOLA K, SINGLA P. PARE: a simple and strong baseline for monolingual and multilingual distantly supervised relation extraction[C]//Proceedings of the 60th Annual Meeting of The Association for Computational Linguistics, 2022: 340–354.
- [5] INJY S, MARCO S. Open-CyKG: an open cyber threat intelligence knowledge graph[J]. Knowledge-Based Systems, 2021, 233(1): 1–13.
- [6] ZHENG X, DU Y, QIN X. Multi-level semantics-aware and multi-granularity knowledge-infused model for emotional support conversation[J]. Information Fusion, 2025, 128: 103968.
- [7] MIAO J F, HUANG Y, LIU Y, et al. NoTNER: self-optimizing text reconstruction for open named entity recognition on social media[J]. The Journal of Supercomputing, 2025, 81(13): 1280.
- [8] QIAN Y, HU Y, ZHANG S, et al. Disentangling biased representations: a causal intervention framework for fairer NLP models[C]//Proceedings of the 6th Workshop on Gender Bias in Natural Language Processing (GeBNLP), 2025: 393–402.
- [9] WU Y, LIU J. Multimodal named-entity recognition based on symmetric fusion with contrastive learning[J]. Symmetry, 2026, 18(2): 353.
- [10] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[DB/OL]. <https://arxiv.org/abs/1508.01991>.
- [11] JU M, MIWA M, ANANIADOU S. A neural layered model for nested named entity recognition[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018: 1446–1459.
- [12] OBAJA MUIS A, LU W. Labeling gaps between words: recognizing overlapping mentions with mention separators[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017: 2608–2618.
- [13] WANG B, LU W. Neural segmental hypergraphs for overlapping mention recognition[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 204–214.
- [14] GILLICK D, BRUNK C, VINYALS O, et al. Multilingual language processing from bytes[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016: 1296–1306.
- [15] YAN H, GUI T, DAI J, et al. A unified generative framework for various NER subtasks[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 5808–5822.
- [16] SHEN Y, MA X, TAN Z, et al. Locate and label: a two-stage identifier for nested named entity recognition[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 2782–2794.
- [17] NAN G, ZENG J, QIAO R, et al. Uncovering main causalities for long-tailed information extraction[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021: 9683–9695.
- [18] WANG X, WANG Y. Sentence-level resampling for named entity recognition[C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2022: 2151–2165.
- [19] 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, 2017: 2980–2988.
- [20] CAO K, WEI C, GAIDON A, et al. Learning imbalanced datasets with label-distribution-aware margin loss[C]//Vancouver: 32nd Conference on Neural Information Processing Systems(NeurIPS 2019), 2019.
- [21] LI J, XIAO L. Syrapropa at SemEval-2020 task 11: BERT-based models design for propagandistic technique and span detection[C]//Proceedings of the Fourteenth Workshop on Semantic Evaluation, 2020: 1808–1816.
- [22] YU P, JI H. Shorten the long tail for rare entity and event extraction[C]//Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, 2023: 1339–1350.
- [23] WANG S, LI X, MENG Y, et al. \$k\$-NN-NER: named entity recognition with nearest neighbor search[DB/OL]. <https://arxiv.org/abs/2203.17103>.
- [24] KANG B, XIE S, ROHRBACH M, et al. Decoupling representation and classifier for long-tailed recognition[DB/OL]. <https://arxiv.org/abs/1910.09217>.
- [25] KANG B, LI Y, XIE S, et al. Exploring balanced feature spaces for representation learning[C]//Online: International Conference on Learning Representations, 2020.
- [26] CHEN W, YANG K, YU Z, et al. A survey on imbalanced learning: latest research, applications and future directions[J]. Artificial Intelligence Review, 2024, 57(6): 137.

- [27] ZHANG H, XU H, LIN T E. Deep open intent classification with adaptive decision boundary[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 14374–14382.
- [28] ZHANG H, XU H, ZHAO S, et al. Learning discriminative representations and decision boundaries for open intent detection[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 1611–1623.
- [29] LI Y, OUYANG X, PAN C, et al. Multi-granularity open intent classification via adaptive granular-ball decision boundary[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2025: 24512–24520.
- [30] CHEN G, XU Q, ZHAN C, et al. Improving open intent detection via triplet-contrastive learning and adaptive boundary[J]. IEEE Transactions on Consumer Electronics, 2024, 70(1): 2806–2816.
- [31] LIU X, LI J, MU J, et al. Effective open intent classification with K-center contrastive learning and adjustable decision boundary[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2023: 13291–13299.
- [32] DOZAT T, MANNING C D. Deep biaffine attention for neural dependency parsing[DB/OL]. <https://arxiv.org/abs/1611.01734>.
- [33] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention, 2015: 234–241.
- [34] LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016: 260–270.
- [35] TANG B, HU J, WANG X, et al. Recognizing continuous and discontinuous adverse drug reaction mentions from social media using LSTM-CRF[J]. Wireless Communications and Mobile Computing, 2018(1): 2379208.
- [36] LU W, ROTH D. Joint mention extraction and classification with mention hypergraphs[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015: 857–867.
- [37] STRAKOVÁ J, STRAKA M, HAJIC J. Neural architectures for nested NER through linearization[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 5326–5331.
- [38] ZHANG S, SHEN Y, TAN Z, et al. De-bias for generative extraction in unified NER task[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022: 808–818.
- [39] SHEN Y, TAN Z, WU S, et al. PromptNER: prompt locating and typing for named entity recognition[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2023: 12492–12507.
- [40] CAI Y, LIU Q, GAN Y, et al. Difinet: boundary-aware semantic differentiation and filtration network for nested named entity recognition[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, 2024: 6455–6471.
- [41] WAN J, RU D, ZHANG W, et al. Nested named entity recognition with span-level graphs[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022: 892–903.
- [42] ZHENG Q, WU Y, WANG G, et al. Exploring interactive and contrastive relations for nested named entity recognition[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 2899–2909.
- [43] YAN H, SUN Y, LI X, et al. An embarrassingly easy but strong baseline for nested named entity recognition[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2023: 1442–1452.
- [44] YU J, BOHNET B, POESIO M. Named entity recognition as dependency parsing[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 6470–6476.
- [45] SHEN Y, SONG K, TAN X, et al. Diffusionner: boundary diffusion for named entity recognition[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2023: 3875–3890.
- [46] WANG T, LI Y, KANG B, et al. The devil is in classification: a simple framework for long-tail instance segmentation[C]//Proceedings of the European Conference on Computer Vision, 2020: 728–744.
- [47] AURELIO Y S, DE ALMEIDA G M, DE CASTRO C L, et al. Learning from imbalanced data sets with weighted cross-entropy function[J]. Neural Processing Letters, 2019, 50(2): 1937–1949.
- [48] LAI Z, WANG C, GUNAWAN H, et al. Smoothed adaptive weighting for imbalanced semi-supervised learning: improve reliability against unknown distribution data[C]//Proceedings of the International Conference on Machine Learning, 2022: 11828–11843.
- [49] LUO H, TAN W, NGUYEN N, et al. Re-weighting tokens: a simple and effective active learning strategy for named entity recognition[C]//Proceedings of 2023 Conference on Empirical Methods in Natural Language Processing, 2023: 12725–12734.
- [50] GRAF F, HOFER C, NIETHAMMER M, et al. Dissecting supervised contrastive learning[C]//Proceedings of the International Conference on Machine Learning, 2021: 3821–3830.
- [51] LI T, CAO P, YUAN Y, et al. Targeted supervised contrastive learning for long-tailed recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 6918–6928.
- [52] KIM J D, OHTA T, TATEISI Y, et al. GENIA corpus—a semantically annotated corpus for bio-textmining[J]. Bioinformatics, 2003, 19(suppl_1): i180–i182.
- [53] COLLIER N, OHTA T, TSURUOKA Y, et al. Introduction to the bio-entity recognition task at JNLPBA[C]//Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA/BioNLP), 2004: 73–78.
- [54] ZHU E, LIU Y, LI J. Deep span representations for named entity recognition[C]//Annual Meeting of the Association for Computational Linguistics, 2023: 10565–10582.
- [55] DING N, XU G, CHEN Y, et al. Few-nerd: a few-shot named entity recognition dataset[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021: 3198–3213.
- [56] YANG S, TU K. Bottom-up constituency parsing and nested named entity recognition with pointer networks[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022: 2403–2416.
- [57] ZHU E, LI J. Boundary smoothing for named entity recognition[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022: 7096–7108.
- [58] TANG M, HE Y, XU Y, et al. A boundary offset prediction network for named entity recognition[C]//Proceedings of the the 2023 Conference on Empirical Methods in Natural Language Processing, 2023: 14834–14846.
- [59] LIU Y, HU J, CHEN Z, et al. EASAL: entity-aware subsequence-based active learning for named entity recognition[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2023: 8897–8905.
- [60] ZHANG S, CHENG H, GAO J, et al. Optimizing bi-encoder for named entity recognition via contrastive learning[DB/OL]. <https://arxiv.org/abs/2208.14565>.
- [61] WU S, SHEN Y, TAN Z, et al. MProto: multi-prototype network with denoised optimal transport for distantly supervised named entity recognition[C]//Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, 2023: 2361–2374.