基于FLAT的农业病虫害命名实体识别

任义[1],沈洁[*1],袁帅[1]
1.沈阳建筑大学，计算机科学与工程学院，辽宁沈阳 110168
renyi@sjzu.edu.cn,1007404126@qq.com,yuanshuai@sjzu.edu.cn

摘要

针对传统命名实体识别方法中词嵌入无法表征一词多义及字词融合的模型存在特征提取不够准确的问题，本文提出了一种基于FLAT的交互式特征融合模型，该模型首先通过外部词典匹配获得字、词向量，经过BERT预训练后，通过设计的交互式特征融合模块充分挖掘字词间的依赖关系。另外，引入对抗训练提升模型的鲁棒性。其次，采用了特殊的相对位置编码将数据输入到自注意力机制，最后通过CRF得到全局最优序列。本文模型在农业病虫害数据集上识别的准确率、召回率、F1值分别达到了93.76%、92.14%和92.94%。

关键词： 命名实体识别； 农业病虫害； 对抗训练； 特征融合； 自注意力机制

Named Entity Recognition of Agricultural Pests and Diseases based on FLAT

Yi Ren[1],Jie Shen[*1],Shuai Yuan[1]
1.School Of Computer Science And Engineering, Shenyang Jianzhu University
Shenyang, Liaoning 110168, China
renyi@sjzu.edu.cn,1007404126@qq.com,yuanshuai@sjzu.edu.cn

Abstract

Aiming at the problem that feature extraction is not accurate enough in the traditional named entity recognition method in which word embedding cannot represent the polysemy of a word and word fusion model, this paper proposes an interactive feature fusion model based on FLAT. The model first obtains word and word vector through external dictionary matching. After BERT pre-training, The interactive feature fusion module is designed to fully explore the dependency relationship between words. In addition, adversarial training is introduced to improve the robustness of the model. Secondly, a special relative position encoding is used to input the data into the self-attention mechanism, and finally the globally optimal sequence is obtained by CRF. The identification accuracy, recall rate and F1 value of the model in the agricultural pest and disease data set reached 93.76%, 92.14% and 92.94%, respectively.

Keywords: Named entity recognition , Agricultural pests and diseases , Adversarial training , Feature fusion , Self-attention mechanisms

1 引言

*通讯作者
国家自然科学基金（62073227）；辽宁省教育厅基金（LJKZ0581，LJKZ0584）

目前我国农业的发展受到多种因素制约，除了水涝干旱等自然灾害外，农业病虫害是农业生产最为常见的问题之一。面对海量的非结构化的农业病虫害相关文本数据，人们无法快速准确获取到农业病虫害的防治信息。为了更好地解决我国农业生产实践中遇到的病虫害问题，信息化防治农业病虫害成为提高农业生产效率、增加农业产量的重要手段。农业病虫害命名实体识别任务能够帮助准确识别和分类出相关命名实体。

命名实体识别(王颖洁 et al., 2023)又称实体抽取，主要分为基于规则、机器学习和深度学习的方法。基于规则的方法需要大量领域专家来构建规则，人工及时间成本太高，可迁移性差(Xu K et al., 2019)。基于机器学习的方法主要包括隐马尔科夫模型(Morwal S, 2012)、支持向量机(Isozaki H and Kazawa H, 2002)、最大熵模型(Saha S K et al., 2009)和条件随机场(Lafferty J et al., 2001)。但是，基于机器学习的方法依赖人工制定的特征模板，不具备领域通用性。近年来，随着深度学习的发展，基于深度学习的命名实体识别方法，因其具有能够从数据中自主学习特征而无需人为设定，逐步成为中文命名实体识别的主流模型。刘新亮(2021)等提出一种基于BERT-CRF模型的命名实体识别方法，完成对生鲜蛋供应链领域的命名实体识别。郭知鑫(2021)等提出基于BERT-BiLSTM-CRF的实体识别模型，对法律文本中的案件实体进行智能识别。郭军成(2021)等基于BiLSTM-CRF模型融入BERT层，实现中文简历命名实体识别。由于上述方法未考虑到词向量特征对实体识别效果的影响，Y.Zhang(2018)首次提出了Lattice-LSTM模型，该模型利用字向量和词向量信息，在公开数据集上取得了较好的结果。基于Lattice的模型难以充分利用GPU的并行计算，推理速度通常较慢，X.Li(2020)提出FLAT模型，它将Lattice结构转换为平面结构，在保留Lattice原有信息的基础上提高了并行计算的能力。在以往的研究中，字词融合大多采用不同特征表示向量（如字符向量、词向量）的拼接或累加的方式提取信息，这就造成了不同特征表示之间的相互依赖关系被忽略。

近年来，基于通用领域的命名实体识别已经相对成熟，由于缺少公开标注的数据集，针对农业病虫害领域的命名实体识别研究仍处于探索阶段，现阶段国内外只有少数学者针对农业领域开展了一定研究。李想(2017)等、张剑(2018)等提出基于条件随机场的方法，对农作物、病虫害、农药进行实体识别。郑泳智(2021)等将BERT模型和BiLSTM-CRF模型相结合，实现对农业病虫害领域的命名实体识别。目前，极少数的研究将字词融合运用到农业病虫害领域的命名实体识别中。基于上述问题，本文提出了一种基于FLAT的交互式特征融合模型，对字级嵌入和词级嵌入特征向量进行交互学习，并加入对抗训练以提升模型的鲁棒性。

2 数据处理与标注

2.1 数据获取

本文通过数据获取、数据标注两个步骤，建立农业病虫害领域的数据集。农业病虫害领域命名实体识别目前还没有公开可使用的数据集，本文节选了百度百科有关农业病虫害的信息作为文本语料初始数据。通过数据清洗、去噪、去冗等预处理，保证数据的可靠性。接下来就是标注标签工作，由于是自己定义的标签类别，所以需要人工手动标注，而实体的标注需要大量特定领域的知识，从而又增加了标注的难度。另外，本研究通过查阅资料和咨询专家的方法，基于现有研究本文进一步将疾病类别划分为更细粒度的实体，分别为“病害”和“虫害”。此外，与农作物相关的实体，如防治药剂和为害症状等也被考虑在内。经过数据清洗后的文字约6万字，包含1476个句子，其中有1968个农业病虫害实体。

2.2 数据标注

本文预先定义好的命名实体类别包括病害、虫害、防治药剂、防治方法、为害症状、为害地区、作物，实体定义和样例如表1。通过使用Brat标注工具(Mahanazuddin S et al., 2021)，对获取的语料进行人工标注。本文采用BIO规则对语言序列进行标注，其中B（Begin）描述句子中每个命名实体的开始位置；I（Internal）描述命名实体除起始位置外的其他部分，O（Other）用来描述句子中其他非预先定义好的实体。标注中为了更好地识别命名实体的类别信息，本文将类别信息与BIO规则进行融合，类别信息如病害实体用（-Disease）表示、为害地区实体用（-Area）表示等。以句子“水稻云形病主要发生在长江流域”为例，其序列标注如图1所示。

实体	定义	样例
病害	农业病害名称	水稻纹枯病、赤霉病
虫害	农业虫害名称	褐飞虱、白粉虱
防治药剂	防治农业病虫害的药剂学名、俗名、生物防治药剂名	井冈霉素、多菌灵
防治方法	防治农业病虫害的农业防治方法、生物防治方法等	加强肥水管理、检疫、水旱轮作
为害症状	农业病虫害危害作物的特征	失水青枯、枯萎霉烂
为害地区	发生农业病虫害的地区	长江流域、江苏、海南
作物	农作物名称及品种	水稻、小麦、临稻6号、辽粳326

表 1: 实体定义和样例

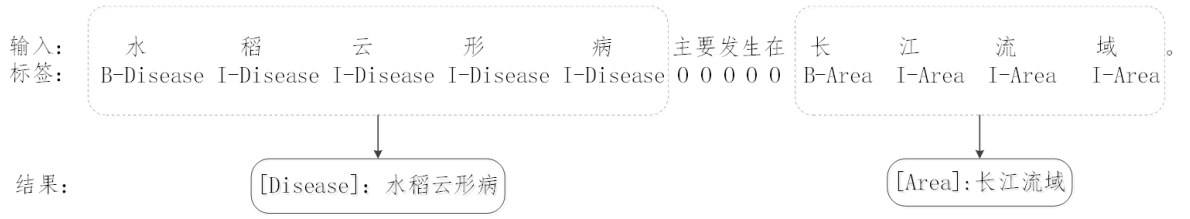

图 1: 语料序列标注示例

3 本文模型

本文提出了一种基于交互式特征融合与对抗训练的农业病虫害命名实体识别模型。对于输入序列$X = \{x_1, x_2, x_3, ..., x_n\}$,将$x_i$与外部词典匹配得到句子中的潜在词汇向量$d_i$，$x_i$通过BERT-WWM层生成具有丰富信息的字向量$c_i$，$c_i$与词典特征$d_i$拼接得到向量$w_i$，$w_i = c_i \oplus d_i$。随后，$w_i$与$d_i$进行交互式特征融合，接着对融合之后的向量进行对抗训练。使用相对位置编码将信息输入到自注意力机制，为了提升模型的性能，引入残差连接、归一化和前馈神经网络，最后通过线性层和条件随机场得到全局最优的标注结果。整体模型架构如图2所示。

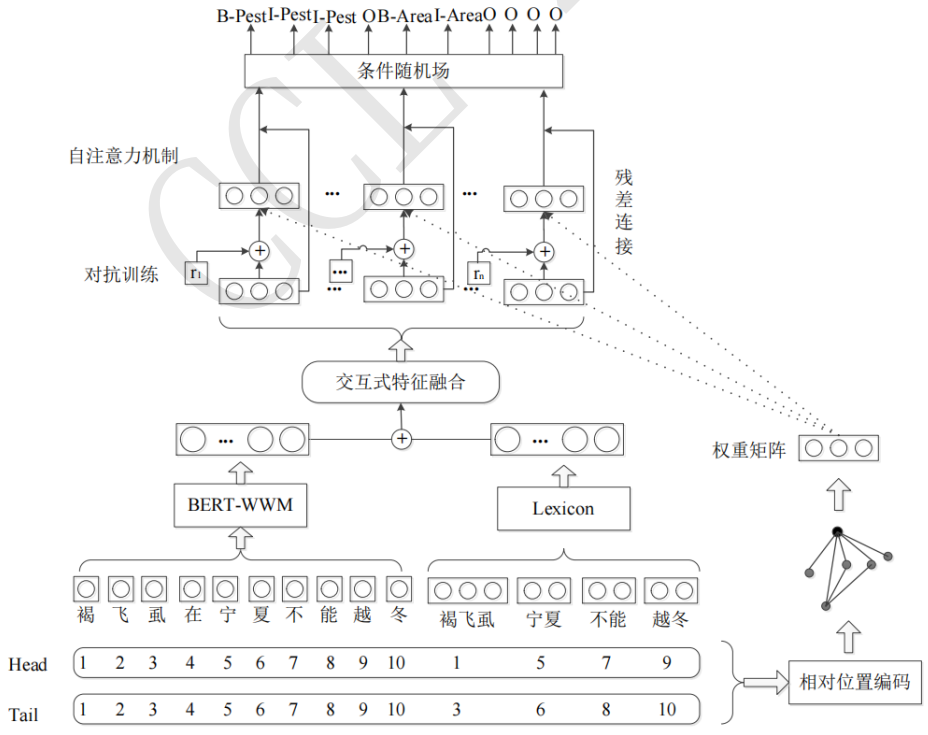

图 2: 总体模型架构

第二十二届中国计算语言学大会论文集，第410页-第419页，哈尔滨，中国，2023年8月3日至5日。

3.1 交互式特征融合

为了解决特征融合时不同特征之间的相互依赖关系被忽略的问题，而传统的字词融合只是通过向量的简单累加来达到融合的目的，本文提出一种交互式特征融合机制，如图3所示。

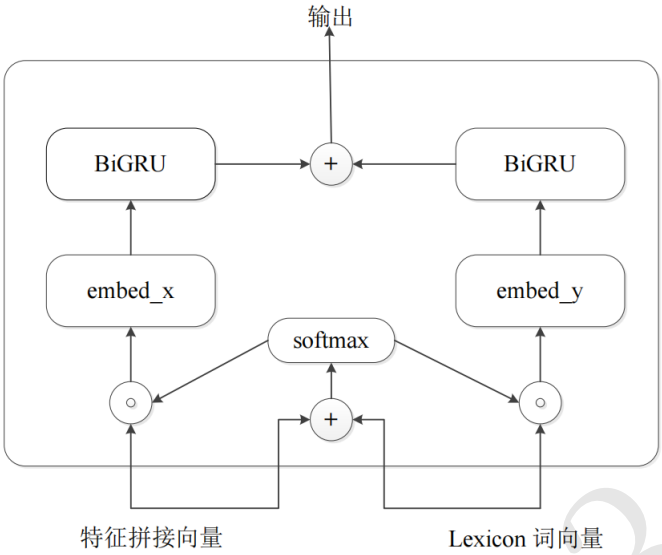

图 3: 交互式特征融合结构图

该机制通过一种交互的方式将不同特征进行充分融合。上述提及的d_i和w_i先经过简单融合，经过softmax函数筛选特征，之后再与w_i按元素相乘，这样通过以w_i特征向量为主体融合字词特征向量的部分重要特征得到$embed_x$，同理可得$embed_y$，其数学原理可以概括为下式(1)-(2)。

$$embed_x = w_i \odot \text{softmax}(w_i + d_i) \tag{1}$$

$$embed_y = d_i \odot \text{softmax}(w_i + d_i) \tag{2}$$

本文引入softmax激活函数更新不同特征向量的权重，从而提高有用信息的比重，丰富输入信息的特征表示。式中，$\odot$表示哈达玛积。

将特征交互得到的$embed_x$和$embed_y$分别经过BiGRU(Bidirectional Gated Recurrent Units)训练后进行融合，进一步丰富输入序列的语义表示，最终实现了整个交互式特征融合，其过程可概括为式(3)。

$$output = \text{BiGRU}(embed_x) + \text{BiGRU}(embed_y) \tag{3}$$

通过BiGRU层特征提取后，可以更加充分捕获上下文之间的关系。GRU工作原理的详细计算公式如式(4)-(7)所示。

$$r_t = \sigma(W_{ir}x_t + b_{ir} + W_{hr}h_{(t-1)}) + b_{hr}) \tag{4}$$

$$z_t = \sigma(W_{iz}x_t + b_{iz} + W_{hz}h_{(t-1)}) + b_{hz}) \tag{5}$$

$$n_t = \tanh(W_{in}x_t + b_{in} + r_t \odot (W_{hn}h_{(t-1)} + b_{hn}) \tag{6}$$

$$h_t = (1 - z_t) * n_t + z_t \odot h_{(t-1)} \tag{7}$$

式中，x_t表示t时刻的输入信息，$h_{(t-1)}$表示$t-1$时刻的隐藏状态，h_t表示t时刻的隐藏状态，W与b分别为权重矩阵与偏置项，σ为sigmoid非线性变换函数，tanh为激活函数，$\odot$为哈达玛积，r_t为重置门，z_t为更新门，n_t为候选隐藏状态。更新门与重置门通过sigmoid函数将值压缩在0到1之间。重置门用于决定前一时刻的状态是否需要保留，更新门则显示了保留前一时刻状态的程度。候选隐藏状态包含了t时刻的输入x_t的信息和对$t-1$时刻的隐藏状态$h_{(t-1)}$选择性的

保留了信息。

简单的GRU不能充分利用文本的上下文信息，本文设计使用BiGRU网络模型来提取序列信息的关键特征。BiGRU由两层GRU组成，分别是前向GRU与后向GRU，将他们各自得到的隐藏层状态拼接得到最终的隐藏层状态，详情见如下公式(8)-(10)。

$$\overrightarrow{h_t} = \overrightarrow{\mathrm{GRU}}(x_t) \tag{8}$$

$$\overleftarrow{h_t} = \overleftarrow{\mathrm{GRU}}(x_t) \tag{9}$$

$$h_t = (\overrightarrow{h_t}, \overleftarrow{h_t}) \tag{10}$$

式中，$\overrightarrow{h_t}$、$\overleftarrow{h_t}$分别为前、后向GRU的隐藏层状态，$\overrightarrow{\mathrm{GRU}}(x_t)$为前向GRU的计算过程，$\overleftarrow{\mathrm{GRU}}(x_t)$为后向GRU的计算过程，隐藏状态$h_t$为序列信息的特征表示。

3.2 对抗训练

对抗训练最早应用于计算机视觉领域，指在训练样本中添加一些可能导致误分类的微小扰动，并使神经网络适应这种改变，现已逐渐引入到自然语言处理领域中。本文利用对抗训练在保留原始数据的基础上加入对抗样本来发挥部分数据增强的作用，提高模型识别边界模糊实体的能力，进而提升模型在农业病虫害命名实体识别时的性能及鲁棒性。

对抗训练的数学原理可以概括为式(11)。

$$\min_{\theta} E_{(x,y)\sim D}\left[\max_{r_{adv}\in\Omega} L\left(\theta, x + r_{adv}, y\right)\right] \tag{11}$$

它可以看成由2部分组成，分别是内部扰动最大化和外部误差损失最小化。其中，r_{adv}表示在输入样本中添加的扰动，Ω表示扰动的空间范围，L表示损失函数，θ表示模型参数，x表示输入样本，y表示样本的标签，D表示输入样本的空间分布。

常用的对抗训练算法有FGM(Fast Gradient Method)、PGD(Projected Gradient Descent)和FreeLB(Free Large-Batch)，由于本文模型的计算量较大，本文选用训练速度较快的FGM对抗训练算法，其扰动r_{adv}的计算方法如式(12)-(13)所示。

$$r_{adv} = \epsilon \cdot \left(\frac{g}{\|g\|_2}\right) \tag{12}$$

$$g = \bigtriangledown_x(L(x, y, \theta)) \tag{13}$$

式中，ϵ为超参数的小有界范数，$\|g\|_2$表示梯度的$L2$范数，g为损失函数关于x的梯度。得到对抗样本X_{adv}如式(14)。

$$X_{adv} = x + r_{adv} \tag{14}$$

对抗样本会模仿标签中数据集的自然误差，使模型更能容忍模型参数波动带来的变化，从而提高模型的鲁棒性。在生成对抗样本之后，将交互式特征融合之后的向量与对抗样本一起送入自注意力机制训练。

3.3 相对位置编码

H.Yan(2019)指出，位置和方向信息在命名实体识别任务中非常重要。对于lattice中的两个span如x_i和x_j，本文使用4种相对距离来表示x_i和x_j之间的关系，其计算公式如式(15)-(18)所示。

$$d_{ij}^{(hh)} = head[i] - head[j] \tag{15}$$

$$d_{ij}^{(ht)} = head[i] - tail[j] \tag{16}$$

$$d_{ij}^{(th)} = tail[i] - head[j] \tag{17}$$

$$d_{ij}^{(tt)} = tail[i] - tail[j] \tag{18}$$

式中，$d_{ij}^{(hh)}$表示x_i的$head$与x_j的$head$之间的距离，$d_{ij}^{(ht)}$表示x_i的$head$与x_j的$tail$之间的距离，$d_{ij}^{(th)}$表示x_i的$tail$与x_j的$head$之间的距离，$d_{ij}^{(tt)}$表示x_i的$tail$与x_j的$tail$之间的距离。举例来说，如图2，假设“越冬”与“褐飞虱”分别为x_i与x_j，则$d_{ij}^{(ht)} = head[i] - tail[j] = head[\prime越\prime] - tail[\prime虱\prime] = 6$，同理可得$d_{ij}^{(hh)}$=8，$d_{ij}^{(th)}$= 9，$d_{ij}^{(tt)}$= 7。

最终的相对位置编码通过这4种相对距离经过简单的非线性变换得到，具体计算公式如式(19)。

$$R_{ij} = \text{ReLU}(W_r(p_{d_{ij}^{(hh)}} \oplus p_{d_{ij}^{(th)}} \oplus p_{d_{ij}^{(ht)}} \oplus p_{d_{ij}^{(tt)}})) \tag{19}$$

式中，W_r为可学习参数，$\oplus$表示拼接运算，p_d的计算与原始的Transformer相同，计算公式如下。

$$p_d^{(2k)} = \sin(\frac{d}{10000^{2k/d_{model}}}) \tag{20}$$

$$p_d^{(2k+1)} = \cos(\frac{d}{10000^{2k/d_{model}}}) \tag{21}$$

式中，d表示$d_{ij}^{(hh)}$、$d_{ij}^{(ht)}$、$d_{ij}^{(th)}$、$d_{ij}^{(tt)}$，k表示位置编码的维度索引。

本文使用改进后的自注意力机制，用新的注意力打分函数代替原本的缩放点积模型，公式如下。

$$\text{Attention}(A, V) = \text{softmax}(A)V \tag{22}$$

$$A_{ij} = W_q^T E_{x_i}^T E_{x_j} W_{k,E} + W_q^T E_{x_i}^T R_{ij} W_{k,R} + u^T E_{x_j} W_{k,E} + v^T R_{ij} W_{k,R} \tag{23}$$

$$[Q, K, V] = E_x[W_q, W_k, W_v] \tag{24}$$

式中，Q表示查询，K表示键，V表示值，d_{head}是多头注意力机制中每个头的维度，E是Embedding层，W、u、v为可学习参数，R_{ij}表示相对位置编码。

4 实验结果与分析

为验证本文模型的有效性，对自建的数据集按照训练集、测试集、验证集为6:2:2比例进行划分，验证集用于模型训练及优化，三个数据集无重叠交叉，因此测试集的训练结果可以作为模型性能的评价指标。

4.1 实验设置

实验采用Ubuntu操作系统，运行环境为RTX3080 GPU，内存为10G。模型所用的优化算法为SGD(Tian Yingjie et al., 2023)，为了缓解过拟合的问题，引入Dropout机制，模型参数设置如表2所示。

超参数设置	数值
隐藏单元数	8
多头注意力机制$head$个数	8
输入维度	160
学习率	6e-4
epoch	20
batch size	2
Dropout	0.5

表 2: 实验参数设置

4.2 评价指标

本文采用准确率（Precision，P）、召回率（Recall，R）以及准确率和召回率的调和平均数F1值（F1-score，F1）作为评价指标，其数值越高代表模型效果越好，各评价指标的计算公式如式(25)-(27)所示。

$$P = \frac{TP}{TP + FP} \times 100\% \tag{25}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{26}$$

$$F1 = \frac{2PR}{P + R} \times 100\% \tag{27}$$

式中，TP表示准确识别出的实体个数，FP表示错误识别出的实体个数，FN表示数据集中存在但未被识别出的实体个数。

4.3 对比实验

本文采用的主要对比模型如下：

BERT-BiLSTM-CRF：该模型是各领域命名实体识别基于字符的主流模型，主要由BERT-base-chinese预训练模型获取字向量，后输入到 BiLSTM 中对句子上下文进行建模，然后使用CRF进行解码。

BERT-IDCNN-CRF：与上述模型类似，其中IDCNN相比BiLSTM能够实现并行运算。

BERT-ADV-BiLSTM-CRF：在BERT-BiLSTM-CRF模型的基础上引入对抗训练，增强模型的泛化性。

BERT-FLAT：基于Transformer设计了一种巧妙位置编码来融合Lattice结构，可以无损引入词汇信息。

LatticeLSTM：加入了字词融合，避免因分词错误导致实体识别错误，是基于词汇增强的中文实体识别方法。

Radical-BiLSTM-CRF：采用BiLSTM-CRF神经网络，该网络同时利用字符级和部首级radical-level表示。

CNN-BiRNN-CRF：采用五笔画输入法获得笔画级表示，并将其与预训练的字符嵌入相连接，以探索字符的形态和语义信息，CNN用于提取N-gram特征。

LRCNN：改进LatticeLSTM，用CNN代替LSTM以提升性能，使用rethinking机制，通过高层特征的语义来优化词向量权重。

NFLAT：NFLAT对FLAT进行了解耦，第一阶段使用InterFormer融合词的边界和语义信息，第二阶段使用Transformer对上下文进行词汇信息编码，最后使用条件随机场作为解码器来预测序列标签，达到了SOTA。

本文在自建的农业病虫害数据集上设置五组模型分别进行实验，对比实验结果如表3所示。在模型训练中，本文的模型具有更加优异的表现。

模型	准确率/%	召回率/%	F1值/%
BERT-BiLSTM-CRF	83.95	71.53	77.24
BERT-IDCNN-CRF	80.47	72.24	76.13
BERT-ADV-BiLSTM-CRF	81.82	77.58	79.64
BERT-FLAT	88.65	91.09	89.85
本文模型	93.76	92.14	92.94

表 3: 对比实验结果

由表3可知，基于BERT-BiLSTM-CRF的识别效果优于BERT-IDCNN-CRF，这是因为IDCNN只能获取局部特征，而BiLSTM能够获取上下文全局特征。基于BERT-ADV-BiLSTM-CRF的识别效果较BERT-BiLSTM-CRF又有了较大的提升，表明了对抗训练的引入可以提高模型对实体边界的识别能力，可以证明对抗训练在命名实体识别任务中的有效性。

实验表明BERT-FLAT模型对于农业病虫害的识别效果相较其它模型有质的进步，所以本文在BERT-FLAT模型的基础上进行改进，本文模型的准确率、召回率、F1值分别提升了5.11%、1.05%和3.09%。

为了避免本文模型在自建的农业病虫害数据集上的实体识别结果具有偶然性以及证明模型在通用领域仍具有较好的表现，因此，在SIGHAN Bakeoff 2006的MSRA数据集上也进行了实验，MSRA数据集是微软亚洲研究院提供的较为权威的命名实体识别数据集，其规模如表4所示。

种类	训练集（K）	测试集（K）
句子	46.4	4.4
字符	2169.9	172.6
实体	74.8	6.2

表 4: MSRA数据集介绍

MSRA数据集包含三种实体，分别是人名、机构名、地名。MSRA数据集命名实体识别的实验结果如表5所列。Radical-BiLSTM-CRF模型同时利用字符级和部首级表示来提取特征，F1值为90.95%；CNN-BiRNN-CRF模型融入笔画特征以探索字符的形态和语义信息，F1值为91.67%；Lattice-LSTM在字向量的基础上引入词向量信息，且不会受到分词错误的影响，F1值为93.18%；LR-CNN在Lattice-LSTM的基础上使用rethinking机制来优化词向量权重同时提高了运行效率，F1值为93.71%；NFLAT去除self-attention的冗余计算，提高模型的性能和效率，F1值达到了94.55%。实验结果表示，本文模型在MSRA数据集上的表现更好，将F1值提升了0.97%。

模型	准确率/%	召回率/%	F1值/%
Radical-BiLSTM-CRF(Dong C et al., 2016)	91.28	90.62	90.95
CNN-BiRNN-CRF(Yang F et al., 2018)	92.04	91.31	91.67
Lattice-LSTM	93.57	92.79	93.18
LR-CNN(Tao Gui et al., 2019)	94.50	92.93	93.71
NFLAT(Wu S et al., 2022)	94.92	94.19	94.55
本文模型	95.39	95.65	95.52

表 5: MSRA数据集命名实体识别实验结果

4.4 消融实验

为了证明本文提出的模型中交互式特征融合与对抗训练这两部分的有效性，表6列出了模型在去除这两部分之后在自建农业病虫害数据集的性能。

模型	准确率/%	召回率/%	F1值/%
本文模型	93.76	92.14	92.94
BERT-FLAT-交互式特征融合	92.95	92.40	92.67
BERT-FLAT	88.65	91.09	89.85

表 6: 消融实验结果

由表6可知，在去除对抗训练和交互式特征融合后，模型的性能均有所下降：

（1）在去除对抗训练后，综合指标F1值下降了0.27%，这是由模型识别边界模糊实体的能力变弱、泛化性变差引起的。

（2）在（1）的基础上再去除交互式特征融合后，综合指标F1值下降了2.82%，这是因为特征融合变成了不同特征的简单累加，完全忽略了它们之间的相互依赖关系，语义特征无法得到充分表示。基于（1）、（2）所述，可以证明本文模型中交互式特征融合与对抗训练的有效性。

5 结语

由于农业领域缺乏公开的语料库，本文首先构建了农业病虫害命名实体识别的数据集。对于农业领域中文命名实体识别任务，考虑到用字词融合的方法来获取丰富的语义表示，在FLAT的基础上加入交互式特征融合模块以充分提取字词间的依赖关系，此外加入对抗训练来提升模型的鲁棒性和泛化性。下一步研究重点将集中在以下两方面：一是将自建的农业数据集进一步扩充以及修正或增强存在的噪音误差，以提升模型的识别效果。二是在扩充规模的数据集进行显著性验证。三是继续研究特征融合模块，引入更加丰富的特征信息。

参考文献

Dong C, Zhang J and Zong C. 2016. Natural Language Understanding and Intelligent Applications: 5th CCF Conference on Natural Language Processing and Chinese Computing, NLPCC 2016, and 24th International Conference on Computer Processing of Oriental Languages. *Character-based LSTM-CRF with radical-level features for Chinese named entity recognition*, ICCPOL 2016, Kunming, China, December 2—6, 2016, Proceedings 24. Springer International Publishing, 2016: 239–250.

Isozaki H and Kazawa H. 2002. International Conference on Computational Linguistics-volume. *Efficient Support Vector Classifiers for Named Entity Recognition*, 1–7.

Lafferty J, McCallum A and Pereira F. C. 2001. ICML. *Conditional random fields: Probabilistic models for segmenting and labeling sequence data*, 282–289.

Lin Y, Shen S and Liu Z. 2016. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). *Neural relation extraction with selective attention over instances*, 2124–2133.

Mahanazuddin S, Shaymaa A L S and Shorabuddin S. 2021. Studies in health technology and informatics. *DeIDNER corpus: Annotation of Clinical Discharge summary notes for named entity recognition USING BRAT TOOL*, 281–432.

Morwal S . 2012. *Named Entity Recognition using Hidden Markov Model (HMM).* Int.j.nat.lang.comput,1(4):15–23

Saha S K, Sarkar S and Mitra P. 2009. Journal of Biomedical Informatics. *Feature selection techniques for maximum entropy based biomedical named entity recognition*, 42(5):905–911.

Tao Gui, Ruotian Ma, Qi Zhang, Lujun Zhao, Yu-Gang Jiang, and Xuanjing Huang. 2019. In Proceedings of the 28th International Joint Conference on Artificial Intelligence. *Cnn-based chinese ner with lexicon rethinking*, AAAI Press, 4982–4988.

Tian Yingjie, Zhang Yuqi and Zhang Haibin. 2023. Mathematics. *Recent Advances in Stochastic Gradient Descent in Deep Learning*, 11(3).

Wu S , Song X and Feng Z. 2022. arXiv. *NFLAT: Non-Flat-Lattice Transformer for Chinese Named Entity Recognition*,2205.05832.

Xiaonan L, Hang Y A N and Xipeng QIU. 2020. Association for Computational Linguistics. *FLAT: Chinese NER Using Flat-Lattice Transformer*, 6836–6842.

Xu K, Yang Z G and Kang P P. 2019. Computers in Biology and Medicine. *Document-level attention-based BiLSTM-CRF incorporating disease dictionary for disease named entity recognition*, 108(22):122–132.

Yan H, Deng B and Li X. 2019. arXiv preprint arXiv. *TENER: adapting transformer encoder for named entity recognition*,1911.04474.

Yang F, Zhang J and Liu G. 2018. Natural Language Processing and Chinese Computing: 7th CCF International Conference. *Five-stroke based CNN-BiRNN-CRF network for Chinese named entity recognition,NLPCC 2018, Hohhot, China, August 26—30, 2018, Proceedings, Part I 7.* Springer International Publishing, 2018: 184–195.

ZHANG Y and YANG J. 2018. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics，Association for Computational Linguistics. *Chinese NER using lattice LSTM*, 1554–1564.

郭军成,万刚,胡欣杰. 2021. 计算机应用. 基于*BERT*的中文简历命名实体识别, 41(S01):5.

郭知鑫,邓小龙. 2021. 北京邮电大学学报. 基于*BERT-BiLSTM-CRF*的法律案件实体智能识别方法, 44(4):129.

李想,魏小红,贾璐. 2017. 农业机械学报. 基于条件随机场的农作物病虫害及农药命名实体识别, 48(S1):178–185.

刘新亮,张梦琪,谷情. 2021. 农业机械学报. 基于*BERT-CRF*模型的生鲜蛋供应链命名实体识别, 52(S01):7.

王颖洁,张程烨,白凤波,汪祖民,季长清. 2023. 计算机科学与探索. 中文命名实体识别研究综述, 17(02):324–341.

张剑,吴青,羊昕旖. 2018. 计算机与现代化. 基于条件随机场的农业命名实体识别, 2018(1): 123-126.

郑泳智,吴惠,朱定局. 2021. 计算机与数字工程. 基于荔枝和龙眼病虫害知识图谱的问答系统, 49(12):2618–2622.