

农业病虫害知识问答意图识别与槽位填充联合模型研究

郭旭超¹ 郝霞² 姚晓闯³ 李林¹

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 山东农业大学信息科学与工程学院, 泰安 271018;
3. 中国农业大学土地科学与技术学院, 北京 100083)

摘要: 农业病虫害领域的意图识别和槽位填充研究仍处于起步阶段, 除语料严重匮乏外, 还面临任务相互独立、忽略彼此相关性和未充分利用意图嵌入信息等问题。为此, 提出了一种基于意图嵌入信息和槽位门控机制的意图识别与槽-位填充联合模型(AgIG-IDSF)。首先, 该模型在共享编码模块引入了注意力机制用于丰富上下文语义特征; 其次, 提出了一种融合意图嵌入表示和槽位门控机制的意图-槽位交互方法用以增强意图信息指导槽位填充任务的能力, 进而提高模型的整体识别性能。在包含22个意图类别、10个槽位类别和11 976条标注样本的自构建语料上进行了实验。结果表明, 在该语料上 AgIG-IDSF 模型的意图识别准确率为94.41%, 槽位填充 F1 值为94.01%, 整体识别准确率达88.07%, 显著优于包含双向关联模型在内的多种基准模型, 表明了该模型在识别农业病虫害意图与槽位方面的有效性。此外, 在公共数据集上的实验结果还表明了该模型具有一定的泛化能力。

关键词: 农业病虫害; 意图识别; 槽位填充; 意图嵌入表示; 门控机制; 注意力机制

中图分类号: TP391

文献标识码: A

文章编号: 1000-1298(2023)01-0205-11

OSID:



Joint Intent Detection and Slot Filling of Knowledge Question Answering for Agricultural Diseases and Pests

GUO Xuchao¹ HAO Xia² YAO Xiaochuang³ LI Lin¹

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China
3. College of Land Science and Technology, China Agricultural University, Beijing 100083, China)

Abstract: Joint intent detection and slot filling plays an important part in natural language understanding for knowledge answering, but it is still in its infancy in the field of agricultural diseases and pests. In addition to lack of corpus, it also faces several challenges such as task independence, mutual correlations ignoring, and intent embedded information neglect. To address the above questions, a novel joint intent detection and slot filling model based on intent embedding and slot-gated mechanism, named AgIG-IDSF was proposed. Firstly, the attention mechanism, which could extract the context-aware features, was introduced into shared encoder to further enrich the contextual semantic features. Secondly, an intent-slot interaction method with intent embedding and the slot-gated mechanism was designed to enhance the ability of intent detection to guide the slot filling task. Finally, the comparative experiments from various aspects were conducted on a self-constructed corpus named AGIS, which mainly contained 22 intent categories, 10 slot categories, and 11 976 annotated samples. The experimental results showed that AgIG-IDSF achieved the intent detection accuracy of 94.41%, slot filling F1-score of 94.01%, and overall semantic accuracy of 88.07% on the self-constructed corpus, which were significantly better than a variety of benchmark models, including bidirectional mutual models. It demonstrated the effectiveness of AgIG-IDSF in jointly identifying the intent and slots in the field of agricultural diseases and pests. In addition, the experimental results on public datasets, i.e., ATIS and SNIPS also showed that the model had a certain generalization ability.

Key words: agricultural diseases and pests; intent detection; slot filling; intent embedded information; slot-gated mechanism; attention mechanism

收稿日期: 2022-02-16 修回日期: 2022-04-11

基金项目: 国家科技创新 2030-新一代人工智能重大项目(2021ZD0113702)

作者简介: 郭旭超(1992—), 男, 博士生, 主要从事自然语言处理与知识图谱研究, E-mail: gxe@cau.edu.cn

通信作者: 李林(1963—), 女, 教授, 博士生导师, 主要从事大数据管理与挖掘研究, E-mail: lilincau@126.com

0 引言

意图识别与槽位填充是自然语言理解的两大核心子任务^[1],是进行知识问答的前提。前者一般可被视为句子级的文本分类任务,旨在将给定问句划分为预先设定好的意图类别;而后者可被视为字符级序列标注任务,即将长度为 n 的问句序列 $\mathbf{X} = [x_1 \ x_2 \ \cdots \ x_n]$ 标记为槽位标签序列 $\mathbf{Y} = [y_1 \ y_2 \ \cdots \ y_n]$ ^[2]。以问句“水稻恶苗病的俗称叫什么?”为例,本文旨在通过上述任务分别识别出该问句的语义槽“水稻恶苗病”和意图“别称”,从而构成三元组(水稻恶苗病,别称,?),其中“?”表示根据语义槽(头实体)“水稻恶苗病”和意图(关系)“别称”从知识图谱中查找并反馈给用户的答案(尾实体)。由此可见,意图识别与槽位填充任务是正确解析问句语义的关键,在基于知识图谱的农业病虫害问答系统中发挥重要作用。

目前,针对意图识别与槽位填充的方法总体上可分为流水线方法和联合抽取方法。流水线方法将意图识别与槽位填充分别进行建模^[3-8]。而目前农业病虫害领域主要采用流水线方法实现意图识别和槽位填充任务^[9-12]。此类方法容易导致错误传播问题,并且容易忽略两者之间的相关性。

为此,近年来提出许多基于深度学习的联合抽取方法,此类方法能够对上述任务进行联合建模,可进一步细分为参数共享、单向关联和双向关联模型^[13]。参数共享模型采用共享参数隐式地学习意图识别与槽位填充任务之间的相关性^[14]。例如,一些学者分别将门控循环网络(Gated recurrent unit, GRU)、注意力机制、RNN 以及 BERT 相结合,提出了多种面向意图识别与槽位填充的隐式联合抽取模型^[15-17]。虽然上述方法的整体识别性能相对于流水线方法有所提升,但仅通过联合优化损失函数隐式地利用两者之间的相关性,可解释性较弱^[13]。

单向关联模型采用意图或槽位信息显式地指导槽位填充/意图识别任务。例如文献[18]提出了一种语义槽门控机制(Slot-gated mechanism, Slot - Gated),建立了意图上下文语义信息到槽位填充任务的单向关联,从而提高了模型整体性能。文献[19]提出了一种结合单词级意图标签的栈式传播框架(StackPropagation),该框架直接利用意图信息作为槽位填充任务的输入,以便获取意图语义知识。此外,该模型还为栈式传播框架执行了字符级意图识别机制以进一步缓解错误传播问题。文献[20]在文献[17]基础上通过引入意图-槽位注意力机

制,提出了 JointIDSF 模型,该模型通过“软”意图标签嵌入显式地将意图上下文信息集成到槽位填充任务中。在面向越南语的 PhoATIS 数据集上全局 F1 值达到 86.25%,比 JointBERT + CRF 提升 0.7 个百分点。

双向关联模型旨在对两任务之间的双向关联关系进行建模。例如,文献[21-22]在文献[18]基础上分别提出了两种基于双向门控机制的联合抽取模型 DCNN - Gate 和 Biass - Gate,建立了意图识别与槽位填充之间的双向关联。文献[23]则提出了一种意图识别-槽位填充双向关联网络(Slot filling-intent detection, SF - ID),构建了两任务之间的直接关联,增强彼此交互能力,但需考虑任务顺序。为解决上述问题,文献[24]提出了一种协同互作 Transformer 机制,该机制同时考虑两任务之间的交叉影响,在 ATIS 和 SNIPS 数据集上分别取得 87.4% 和 90.30% 的整体识别精度。

农业病虫害领域的意图识别与槽位填充仍存在以下问题:该领域的意图识别与槽位填充研究仍处于初级阶段,因此还有任务相互独立、忽略彼此相关性问题亟待解决;相应地,已有模型并未充分考虑意图嵌入表示信息,导致意图识别指导槽位填充任务能力有限;就数据而言,该领域公开可用的语料仍处于空白,直接增加了识别农业病虫害意图和语义槽的成本和难度。

因此,为了能更好地解决农业病虫害领域的意图识别与槽位填充问题,本文首先在领域专家指导下收集并标注首个面向农业病虫害的意图识别与槽位填充数据集,然后针对现有模型存在的问题提出一种基于意图嵌入表示和门控机制的意图识别与槽位填充联合模型 AgIG - IDSF。

1 领域语料收集与构建

1.1 类别设置

由于本文研究包含意图识别和槽位填充两个子任务,因此需要预先确定意图和槽位两种标签类型。对于槽位,在已有工作基础^[25-26]上初步确定了虫害、病害、作物、药剂、病原、周期/时期、部位和品种等 8 种槽位类型。另外,为了确保完整性,本文还额外添加了“症状”和“其他”,最终确定了 10 种槽位类别;对于意图,本文在对收集语料进行分析基础上,结合文献[27]确定了病害、虫害、危害作物等 21 种常用意图。此外,考虑到数据集的完备性,本文同样设置了“其他”类别,用于表示与农业病虫害相关的其他未知意图,以便进一步细分。最终得到如表 1 所示的 22 种意图类别。

表 1 意图类别

Tab. 1 Details of intent categories

编号	类别	英文简写	编号	类别	英文简写
1	危害作物	HaCrop	12	防治时期	Conperd
2	别名	Nick	13	症状	Symptom
3	病原/传播媒介	Pathogeny	14	防治方法	Method
4	危害部位	Haloc	15	危害特点	DamaCha
5	分布区域	Location	16	形态特征	Shape
6	作物品种	Varieties	17	生活习性	Habit
7	危害时期	Haperd	18	简介	Intro
8	防治药剂	Drug	19	发病规律	Rule
9	属科	Family	20	病害	Disease
10	虫态/周期	Period	21	虫害	Pest
11	包含	Include	22	其他	Other

1.2 语料收集与标注

本文所需语料主要来源于中国农技推广信息平台 (<http://njtg.nercita.org.cn/user/index.shtml>)。由于用户来自全国各地,因此形成的问答语料更具普遍性和广泛性,更加适用于实际应用场景。鉴于此,本文在该平台上爬取了 2019 年 4 月 21 日至 2021 年 5 月 15 日之间的农业病虫害相关问答数据作为原始语料,通过解析、筛选和去噪等操作对语料进行了预处理,最终得到 11 976 条原始样本。

语料详细标注过程如图 1 所示。由图 1 可知,给定原始语料,标注对象包含槽位与意图两方面:前

者表示参与问答的实体(即槽位),为了加快标注速度,对于所有待标注样本,本文首先采用文献[25]所提 CNNs-BiLSTM-Attention 模型预测每个字符的标签,然后通过人工校验和修正,得到最终的槽位标注序列;本文主要通过人工方式对每条样本的意图类别进行标注。最终得到完整标注结果。最终得到包含 10 个槽位类型、22 个意图类别和 11 976 条训练样本的中文农业病虫害意图识别与槽位填充标注数据集(Intent selection and slot filling corpus for agricultural diseases and pests, AGIS)。为了使数据集的标签分布尽可能均匀,本文按照比例 8:1:1 采用分层抽样法对数据集进行划分,其中训练集、验证集、测试集样本数分别为 9 576、1 199、1 199 条。部分样本如表 2 所示。数据集各类意图与槽位数据分布如图 2 所示。

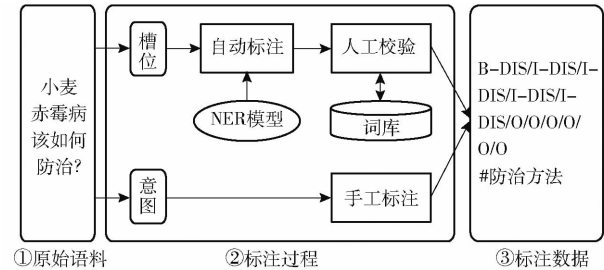


图 1 意图识别与槽位填充语料标注过程

Fig. 1 Labeling process of corpus for intent detection and slot filling

表 2 中文农业病虫害意图识别与槽位填充数据集样例

Tab. 2 Samples of intent detection and slot filling for Chinese agricultural diseases and pests

序号	问句	槽位标签	意图标签
1	包菜的黑斑病的发病原因是什么?	O/O/O/B-DIS/I-DIS/I-DIS/O/O/O/O/O/O/O/O/O/	发病原因
2	玉米瘤黑穗病主要危害玉米的哪些部位?	B-DIS/I-DIS/I-DIS/I-DIS/I-DIS/I-DIS/O/O/O/O/O/O/O/O/O/ O/O/O/O/	危害部位
3	芝麻的主要病害有哪些?	B-CRO/I-CRO/O/O/O/O/O/O/O/O/O/	病害

注:CRO 表示作物,DIS 表示病害。

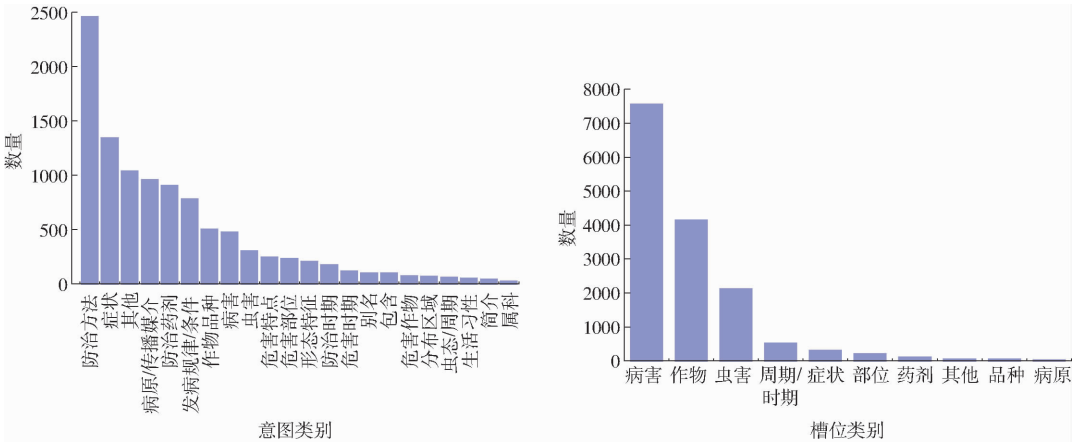


图 2 AGIS 数据集意图与槽位数据分布图

Fig. 2 Data distributions for intents and slots in AGIS

2 AgIG - IDSF 模型

本节首先给出意图识别与槽位填充的形式化定

义,然后描述本文所提模型 AgIG - IDSF 的实现细节。如图 3 所示,本文所提模型主要由共享编码模块、意图识别模块、意图-槽位交互模块以及槽位填

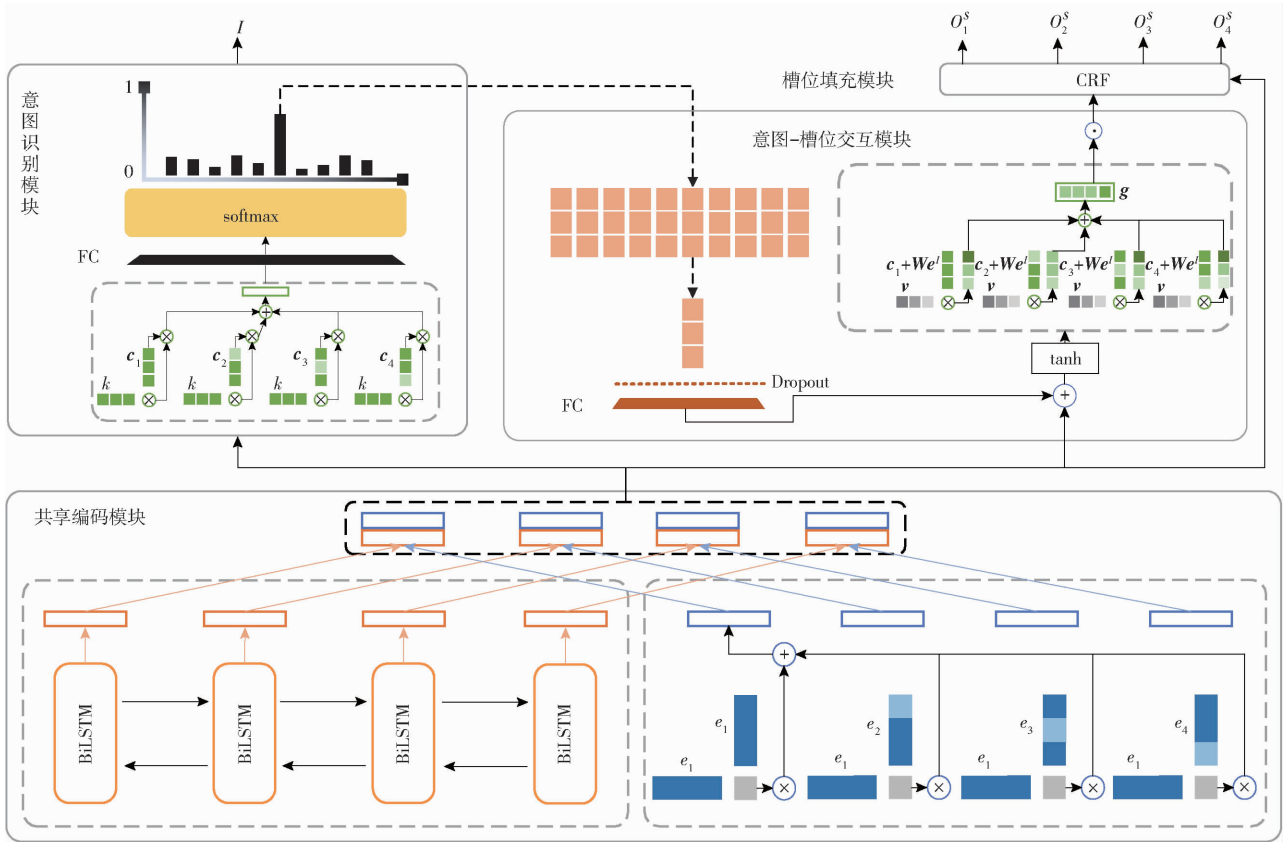


图 3 AgIG - IDSF 模型整体架构图

Fig. 3 Overall framework of AgIG - IDSF model

充模块构成。

2.1 问题定义

给定长度为 n 的农业病虫害问句 X , 农业病虫害意图识别与槽位填充任务可被分解为: ① 句子级意图识别, 一般将其定义为文本分类问题, 预测出相应的意图类别 $O_i^l \in \{O_1^l, O_2^l, \dots, O_m^l\}$ 。② 字符级槽位填充, 为了避免中文分词带来的错误传播问题和未知词问题 (Out of vocabulary, OOV), 本文依然采用字符作为基本输入单元。将其定义为字符级序列标注问题, 即从所有可能的槽集合 $S = \{s_1, s_2, \dots, s_t\}$ 中为每个字符 x_i 预测出对应的语义槽标签。其中, t 为预先定义好的槽位标签个数。最终输出语义槽标签序列 $O^s = [O_1^s \ O_2^s \ \dots \ O_n^s]$ 。

2.2 共享编码模块

该模块旨在学习输入序列上下文语义特征, 以便作为意图识别和槽位填充后续子模块的输入。对于给定输入 X , 经过嵌入层得到的上下文嵌入表示为 $E = [e_1 \ e_2 \ \dots \ e_n]$, 为了能够在不采用 BERT 等大规模预训练模型的情况下充分利用输入序列丰富的语义特征, 在借鉴文献 [19] 基础上, 引

入自注意力编码器, 该编码器主要包含 BiLSTM 和自注意力机制两部分。BiLSTM 用于提取文本序列双向时序特征表示 $H = [h_1 \ h_2 \ \dots \ h_n]$, 其中 $h_i \in \mathbf{R}^d$ 表示第 i 个字符前向和后向上下文表示的级联, 即 $h_i = [\vec{h}_i; \overleftarrow{h}_i]$ 。自注意力机制 [28] 则主要用于学习长距离上下文依赖信息。本文首先基于

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{2d}} \right) V \tag{1}$$

采用权重矩阵 $W_Q, W_K, W_V \in \mathbf{R}^{n \times d}$ 分别将嵌入表示 E 映射为维度为 d 的 Q, K 和 V 。然后基于

$$C = [H; A] \tag{2}$$

得到自注意力机制模块的输出 $A \in \mathbf{R}^{n \times 2d}$ 。

最后, 通过级联操作将双向时序特征表示 H 和注意力上下文表示 A 在特征维度上进行拼接, 得到最终的共享上下文语义 $C = [c_1 \ c_2 \ \dots \ c_n] \in \mathbf{R}^{n \times 2d}$ 。

2.3 意图识别模块

意图识别模块旨在将长度可变的嵌入表示序列 C 聚合为大小固定的意图上下文向量 c^l 。因为注意力机制能够动态调整每个字符级向量的权重, 进而

实现动态聚合,缓解语义损失问题,因此本文选用注意力机制聚合文本序列的共享上下文语义 $\mathbf{C}$,其计算过程为

$$\alpha_i = \text{softmax}(\mathbf{W}\mathbf{c}_i + \mathbf{b}) \quad (3)$$

$$\mathbf{c}^l = \sum_{i=1}^n \alpha_i \mathbf{c}_i \quad (4)$$

其中 $\mathbf{W} \in \mathbf{R}^{2d}$ $\mathbf{c}_i \in \mathbf{R}^{2d}$ $\mathbf{c}^l \in \mathbf{R}^{2d}$ $\sum_{i=1}^n \alpha_i = 1$

式中 $\mathbf{W}$ ——可训练权重

$\mathbf{c}_i$ ——第 i 个字符的共享上下文语义

$\mathbf{b}$ ——偏置

α_i ——第 i 个字符归一化权重,取 0 ~ 1

$\mathbf{c}^l$ ——意图上下文语义表示向量,为 n 个字符的上下文语义表示的加权和

所有意图标签相对于问句 X 的概率分布为

$$\mathbf{p}^l = \sigma(\mathbf{W}^l \mathbf{c}^l + \mathbf{b}^l) \quad (5)$$

其中 $\mathbf{W}^l \in \mathbf{R}^{m \times 2d}$ $\mathbf{b}^l \in \mathbf{R}^m$ $\mathbf{p}^l \in \mathbf{R}^m$ $\sum_{i=1}^m p_i^l = 1$

式中 $\mathbf{W}^l$ ——可训练权重

$\mathbf{b}^l$ ——偏置 m ——意图标签数量

$\mathbf{p}^l$ ——意图标签概率分布(权重分布),则第 i 个意图标签对应的概率 $p_i^l \in \{p_1^l, p_2^l, \dots, p_m^l\}$, p_i^l 值越大,表示问句 X 对应意图为 i 的可能性越大

σ ——激活函数,本文采用 softmax

本文选择最大概率对应的意图标签作为最终预测的意图类别,即

$$\bar{y}^l = \text{argmax}(\mathbf{p}^l) \quad (6)$$

采用交叉损失函数计算意图识别的损失 L_l ,公式为

$$L_l = - \sum_{i=1}^m (\bar{y}^l = y^l) \lg p_i^l \quad (7)$$

$\bar{y}^l = y^l$ 是指,当预测意图标签与真值标签相同时,取值为 1,否则为 0。

2.4 意图-槽位交互模块

该模型主要用于对意图与槽位的相关性进行建模,使得槽位上下文表示中能够融合意图信息,从而实现意图信息指导槽位填充任务。与文献[18,22]采用意图上下文向量 $\mathbf{c}^l$ 作为输入不同,本文从另一个角度出发,通过引入意图嵌入表示矩阵 $\mathbf{E}^l \in \mathbf{R}^{m \times d_l}$,采用意图标签概率分布中最大概率对应的意图嵌入表示向量 $\mathbf{e}^l \in \mathbf{R}^{d_l}$ 作为该模块的输入,其优点在于该向量中包含丰富的意图特征,这有助于提高意图指导槽位填充任务的能力,进而提高模型的性能。

相应地,为了充分利用意图嵌入表示向量和字符级上下文语义特征,对意图和槽位之间的关系进

行建模,本文采用意图嵌入表示向量 $\mathbf{e}^l$ 和共享上下文表示 $\mathbf{C}$ 作为门控机制的输入,用于生成联合权重向量 $\mathbf{g}$ 。 $\mathbf{g}$ 越大,说明意图与槽位的相关性越高,在槽位填充任务中越重要。 $\mathbf{g}$ 的计算公式为

$$\mathbf{g} = \sum \mathbf{v} \tanh(\mathbf{c}_i + \mathbf{W}\mathbf{e}^l) \quad (8)$$

其中

$$\mathbf{v} \in \mathbf{R}^{2d} \quad \mathbf{W} \in \mathbf{R}^{d_l \times 2d}$$

式中 $\mathbf{v}$ ——可训练权重

2.5 槽位填充模块

本文采用条件随机场(CRF)^[23-24]对相邻字符之间的槽位标签依赖进行建模,即

$$\mathbf{C}'_s = \mathbf{C} + \mathbf{C}\mathbf{g} \quad (9)$$

首先采用联合权重向量 $\mathbf{g}$ 对字符级上下文表示 $\mathbf{C}$ 进行加权求和,得到融合意图信息的槽位上下文语义 $\mathbf{C}'_s$,最后采用全连接层进行两次变换,即

$$\mathbf{p}^s = \mathbf{W}_{s2}(\tanh(\mathbf{W}_{s1}\mathbf{C}'_s + \mathbf{b}_{s1})) + \mathbf{b}_{s2} \quad (10)$$

其中 $\mathbf{W}_{s1} \in \mathbf{R}^{2d \times d}$ $\mathbf{W}_{s2} \in \mathbf{R}^{d \times t}$ $\mathbf{b}_{s1} \in \mathbf{R}^d$ $\mathbf{b}_{s2} \in \mathbf{R}^t$

式中 $\mathbf{W}_{s1}$ 、 $\mathbf{W}_{s2}$ ——可训练权重

$\mathbf{b}_{s1}$ 、 $\mathbf{b}_{s2}$ ——偏置

得到最终的槽位上下文语义表示 $\mathbf{p}^s \in \mathbf{R}^{n \times t}$ 。

对句子 X ,一个可能的槽位标签序列 $\mathbf{O}^s = (O_1^s, O_2^s, \dots, O_n^s)$ 的得分可定义为

$$s(X, \mathbf{O}^s) = \sum_{i=1}^n P_{i, o_i^s} + \sum_{i=1}^n T_{o_{i-1}^s, o_i^s} \quad (11)$$

式中 P_{i, o_i^s} ——第 i 个字符预测为标签 O_i^s 的概率,等同于 p^s

$T_{o_{i-1}^s, o_i^s}$ ——从标签 O_{i-1}^s 转移到标签 O_i^s 的概率

则标签序列 $\mathbf{O}^s$ 的条件概率为

$$P_o(\mathbf{O}^s | X) = \frac{\exp(s(X, \mathbf{O}^s))}{\sum_{\mathbf{O}_i^s \in \mathbf{O}_{\text{all}}^s} \exp(s(X, \mathbf{O}_i^s))} \quad (12)$$

式中 $\mathbf{O}_{\text{all}}^s$ ——句子 X 所有可能的标签序列

本文通过最小化负似然对数来计算损失值。最后,本文采用维特比(Viterbi)算法对得分最高的标签序列进行解码,得到最终的预测标签

$$\begin{aligned} L_s &= -\lg(P_o(\mathbf{O}^s | X)) = \\ &= -\lg \frac{\exp(s(X, \mathbf{O}^s))}{\sum_{\mathbf{O}_i^s \in \mathbf{O}_{\text{all}}^s} \exp(s(X, \mathbf{O}_i^s))} = \\ &= \lg \sum_{\mathbf{O}_i^s \in \mathbf{O}_{\text{all}}^s} \exp(s(X, \mathbf{O}_i^s)) - s(X, \mathbf{O}^s) \end{aligned} \quad (13)$$

最后,对两任务进行联合优化,最终的目标函数定义为两损失函数的加权和

$$L_{\text{oss}} = \varphi L_l + (1 - \varphi) L_s \quad (14)$$

φ 为加权因子,介于 0 ~ 1 之间,根据文献[19],本文将其设置为 0.5。

3 实验

3.1 实验环境与参数设置

本文选择第 1 节所构建的 AGIS 数据集作为基准数据集,以评价本文所提模型对农业病虫害意图识别与槽位填充任务的识别效果,同时在公共数据集 ATIS 和 SNIPS 上进行实验,以验证该模型的泛化性。其中,ATIS 主要包含了航班、城市等 6 种航空旅行信息、21 种意图类别、120 种槽位类别和 5 871 条样本;而 SNIPS 涵盖了天气、餐厅和娱乐等多个领域,共包含 7 个意图类别、72 个槽位类别和 15 884 条样本。根据文献[29],环境配置如表 3 所示。在超参数设置方面,根据模型调试结果,AGIS 和 SNIPS 数据集上无需设置 Dropout,而在 ATIS 数据集上对应的 Dropout 设置为 0.1,以缓解过拟合问题。其他超参数设置如表 4 所示。除第 3.3.5 节外,对于中文数据集,所有模型均采用维度为 300 的预训练向量^[30]作为初始的字符级嵌入表示;英文数据集则采用斯坦福大学(<https://github.com/stanfordnlp/GloVe>)预训练的词嵌入向量。

表 3 实验环境配置
Tab.3 Experimental configuration

实验环境	型号及版本
操作系统	Windows 11
深度学习框架	Pytorch 1.7
编程语言	Python 3.7.10
GPU 加速库	CUAD 11.1
中央处理器	Intel Core i7-9700K CPU 3.60GHz
图形处理器	NVIDIA GeForce GTX 1080 Ti 11GB

表 4 超参数设置
Tab.4 Settings for hyper-parameters

超参数	数值/类型
字符嵌入维度	300
BiLSTM 嵌入维度	256
意图嵌入表示维度	256
注意力输出维度	256
学习率	0.001
最大迭代次数	100
批量大小	32
优化器	RAAdam

3.2 评价指标及对比模型

为了定量评价本文所提模型的有效性,本文采用 F1 分数评价模型的槽位填充性能,采用准确率 (Accuracy, Acc) 评价模型的意图识别性能,同时采用句子级准确率评价模型整体识别性能,即意图和槽位标签均识别正确的问句占总样本的百分比。

为了充分说明本文所提模型的有效性,本文采

用以下 5 种主流模型进行对比实验:① Slot - Gated^[18]、② StackPropagation^[19]、③ SF - ID^[23]、④DCA - NET^[24]、⑤BiAss - Gate^[22]。其中,①和②属于单向关联模型,③、④和⑤属于双向关联模型。此外,本文还引入了 JointBERT^[17]、JointIDSF^[20],以深入探究本文所提模型在采用 BERT 预训练模型时的性能。

3.3 实验结果及分析

3.3.1 实验结果

图 4 给出了各模型在 AGIS 数据集上的实验结果。与其他对比模型相比,本文所提模型无论在槽位填充和意图识别单任务上还是在整体识别精度上均得到较大提升,这一方面是因为 AgIG - IDSF 模型采用注意力机制有助于丰富上下文语义特征,另一方面是因为所引入意图嵌入表示包含丰富的意图信息,有助于提高指导槽位填充任务的能力(详情可参考 3.3.3 节)。虽然 Slot - Gated 模型同样利用门控机制指导槽位填充任务,但由于对意图嵌入表示信息的获取能力有限,导致在 AGIS 数据集上的性能不如其他模型。虽然 StackPropagation 采用字符级意图标签投票机制在一定程度上缓解了错误传播问题,但同样无法充分利用意图嵌入表示信息,导致对槽位填充任务的指导能力有限,因此限制了模型性能。SF - ID 虽然能够在一定程度上利用意图识别与槽位填充两任务之间的互作关系,但需考虑建模顺序,导致上述模型的识别结果稍低。主要表现为该模型的整体识别精度与 AgIG - IDSF 模型相差 1.84 个百分点。而 BiAss - Gate 进一步改进了 Slot - Gated 模型,具备双向关联能力,使得整体识别精度进一步提高,达到 87.82%。DCA - NET 是近几年提出的较先进的双向关联模型,该模型所提的协同 Transformer 包含两个彼此交互的注意力机制,打破了交互顺序限制,显式地建立了彼此的双向关联。因此在 AGIS 集上取得优于其他对比模型的识别结果。而本文所提模型槽位填充 F1 值和全局识别精度更高,相较于 DCA - NET 分别提高 0.24、0.16 个百分点。因此,上述实验结果不仅表明了本

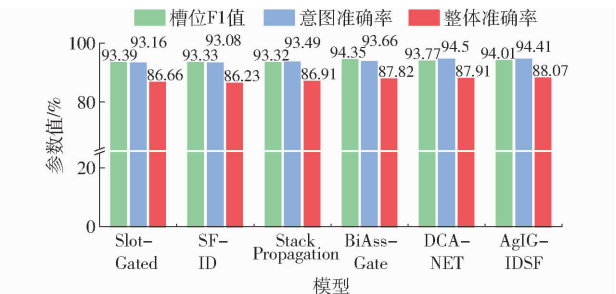


图 4 各模型在 AGIS 数据集上的实验结果

Fig.4 Experimental results for each model on AGIS

文所提模型在农业病虫害意图识别与槽位填充任务上的有效性,而且其性能优于双向关联模型。

为了进一步说明本文所提模型的识别细节,如图 5 所示,本文以 AGIS 数据集为例,分别统计了 AgIG - IDSF、DCA - NET、BiAss - Gate、StackPropagation、SF - ID 和 Slot - Gated 对该数据集各类意图和槽位的识别详情。由图 5a 可知,对每类意图,所有模型在防治方法、病害和防治药剂等意图类别上的识别精度普遍较高,这是因为上述意图类别样本占比较大,训练相对充分。而由于简介、分布区域和别名等意图类别样本量少,导致各模型在上述意图上的识别精度普遍偏低。此外,通过对预测结果进行统计后发现,诸如“危害时期”意图与“虫态/周期”意图容易在上下文语义上混淆,因此降低了彼此的识别精度。而本文所提模型不仅能够有效识别那些较易识别的意图,而且进一步提高了别名、简介和防治时期等较难识别意图的精度。例如,相较于基准模型 DAC - NET,本文所提模型针对上述意图的准确率分别提高 9.09、24.24、5.79 个百分点。对每类槽位,如图 5b 所示,本文所提模型不仅能够有效识别病害、作物和药剂等常见槽位类型,而且对症状等难以识别槽位类型仍保持有较高的识别性能,例如相较于 DCA - NET,其 F1 值提

高 7.5 个百分点,这得益于本文所提的融合意图嵌入表示的门控机制能够使槽位上下文语义表示充分融合意图信息进而提高模型的槽位填充性能。

其次,如图 6 所示,本文同样以 AGIS 数据集为例,可视化了 AgIG - IDSF、DCA - NET、BiAss - Gate、StackPropagation、SF - ID 和 Slot - Gated 模型的意图分类混淆矩阵。由图可知,大多数预测意图分布于对角线上,表现出对各类意图识别的有效性。此外,部分模型将“其他”意图归类为危害时期、防治时期和防治药剂等,对预测结果进行分析后认为包含“其他”意图的问句语义容易与包含上述意图的样本语义混淆,导致识别错误;相反,各模型在属科、别名和发病规律/条件等意图类别上错误预测数量较少,这是因为包含上述意图的问句具有明显的语义特征。比如问句“小麦赤霉病的发病规律是什么?小麦赤霉病还有其他名称吗?”,模型很容易将其意图分别识别为“发病规律/条件”和“别称”。而与大多数模型相比,本文所提模型在多数意图类别上取得较高的识别精度,虽然略逊于 DCA - NET 模型,但由图 4 可知,本文所提模型的整体识别精度优于 DCA - NET 模型。

3.3.2 泛化性分析

为了进一步探究本文所提模型的泛化性,将 AgIG - IDSF 模型及对比模型在 ATIS 和 SNIPS 两种公共数据集上进行了对比实验。如表 5 所示,本文所提模型在 ATIS 和 SNIPS 数据集上取得最大整体识别精度,分别为 87.57% 和 90.43%,高出 BiAss - Gate 模型 0.34、0.14 个百分点,表现出最优性能。而其他模型性能则呈现出不同程度的差异性。比如在 SNIPS 数据集上,Slot - Gated 模型的槽位填充 F1 分数、意图识别准确度和整体识别精度分别与 AgIG - IDSF 模型的相应值相差 1.41、5.11、8.23 个百分点,可能的原因有:中英文不同语言本身差异性;垂直领域语料的专业性;由于缺乏充足的意图嵌入表示信息,槽位上下文语义融合意图信息存在限制。虽然 StackPropagation 模型的意图标签投票机制在一定程度上缓解了上述问题,但仍然存在意图特征缺失问题^[24,31],其识别精确度比 AgIG - IDSF 略低。另外,由实验结果可知 AgIG - IDSF 在两种数据集上的性能明显优于双向关联模型 DCA - NET。因此,上述实验结果不仅再次表明了该模型优于双向关联模型,而且还验证了该模型还具有一定泛化能力和稳定性。

3.3.3 消融实验

为验证本文所提模型各模块的有效性,本文以

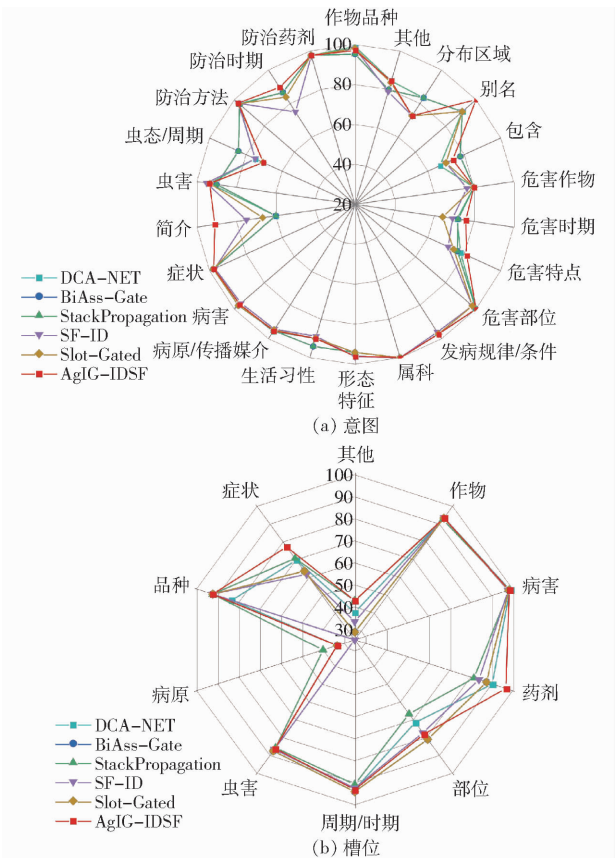


图 5 各模型对 AGIS 各类意图和槽位识别细节
Fig.5 Detailed results of intents and slots on AGIS

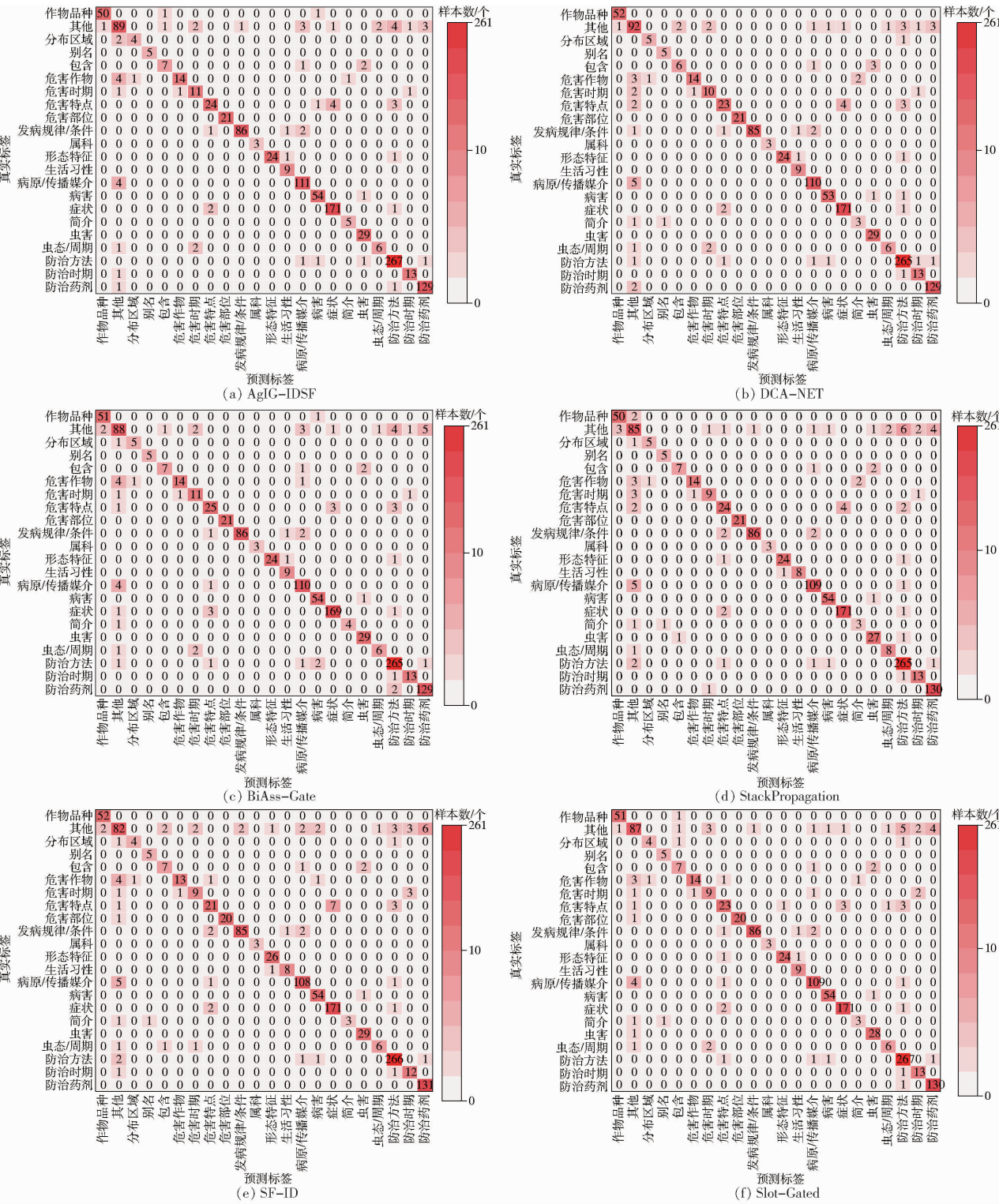


图6 各模型对 AGIS 意图识别混淆矩阵

Fig.6 Confusion matrixes for each model on AGIS

表5 各模型在 ATIS 和 SNIPS 数据集上的识别结果

Tab.5 Experimental results for each model on ATIS and SNIPS

模型	ATIS			SNIPS		
	槽位 F1 值	意图准确率	整体准确率	槽位 F1 值	意图准确率	整体准确率
Slot - Gated	95.67	95.97	85.33	94.80	93.60	82.20
SF - ID	95.12	96.97	85.22	93.62	98.29	84.43
StackPropagation	95.33	96.30	86.00	94.71	98.00	87.71
DCA - NET	95.82	97.65	86.56	95.68	98.71	90.00
BiAss - Gate	95.62	97.42	87.23	96.35	98.14	90.29
AgIG - IDSF	95.97	97.31	87.57	96.21	98.71	90.43

AgIG-IDSF 为基准模型设置了 4 个对照组,分别以 AGIS、ATIS 和 SNIPS 数据集进行实验,结果如表 6 所示。如实验 1,当 AgIG-IDSF 模型采用注意力机制时,该模型在 3 种数据集上的整体识别精度高于实验 2 不采用注意力机制的 AgIG-IDSF 模型的相应值,分别提高 0.91、0.34、0.29 个百分点,这在一定程度上说明了在共享编码模块采用注意力机制的有效性。实验 3 直接采用式(4)所得意图上下文语义

义作为门控机制的输入,虽然在 SNIPS 数据集上取得 88.43% 的整体识别精度,但在 AGIS 和 ATIS 数据集上的整体识别精度明显偏低,仅为 46.29% 和 28.56%,泛化性表现较差,同时反映了本文所提模型引入意图嵌入表示的合理性和必要性。相应地,与实验 4 采用级联槽位代替门控机制相比,本文所提 AgIG-IDSF 模型在 3 种数据集上的整体识别精度更优,表明了门控机制的有效性。

表 6 AgIG-IDSF 在 3 种数据集上的消融实验结果										
Tab.6 Ablation experiments for AgIG-IDSF on three datasets										
实验	模型	AGIS			ATIS			SNIPS		
		槽位 F1	意图准	整体准	槽位 F1	意图准	整体准	槽位 F1	意图准	整体准
		值	确率	确率	值	确率	确率	值	确率	确率
1	AgIG-IDSF	94.01	94.11	88.07	95.97	97.31	87.57	96.21	98.71	90.43
2	-注意力机制	93.73	93.58	87.16	95.75	97.31	87.23	95.87	98.29	90.14
3	+意图上下文语义	51.41	92.99	46.29	59.63	96.98	28.56	95.48	97.86	88.43
4	+Cascade	94.19	93.74	87.49	98.51	98.10	87.35	95.54	97.57	88.86

3.3.4 意图识别与槽位填充

为了探究意图识别和槽位填充任务对全局识别精度的影响程度,如图 7 所示,以 AGIS 数据集为例,采用混淆矩阵可视化了两任务句子级识别结果。其中,横坐标表示意图标签样本数,纵坐标表示槽位标签样本数,均分为正确和错误识别两类。通过分析图 7 可得如下结论:各模型可识别出大部分样本对应的意图与槽位标签,均正确识别数超过 1 000 条;意图识别或槽位填充单任务预测错误标签数较多,例如在图 7b 中,意图标签识别正确但槽位标签识别错误的样本数为 79 条,而槽位标签识别

正确但意图标签识别错误的样本数为 53 条,说明在给定样本中单任务识别错误率对模型整体识别性能影响较大;并且与意图识别相比,槽位填充任务对模型整体识别精度影响更大,因为在意图正确识别情况下,上述模型的槽位填充错误率更高。而与其他模型相比,本文所提模型不仅提高了同时识别正确样本的数量,而且也降低了同时错误样本数量,分别为1 056 条和 13 条,因此提高了模型的整体识别精度。由上述可知,相对于意图识别任务,槽位填充任务对模型整体识别性能的影响更大。

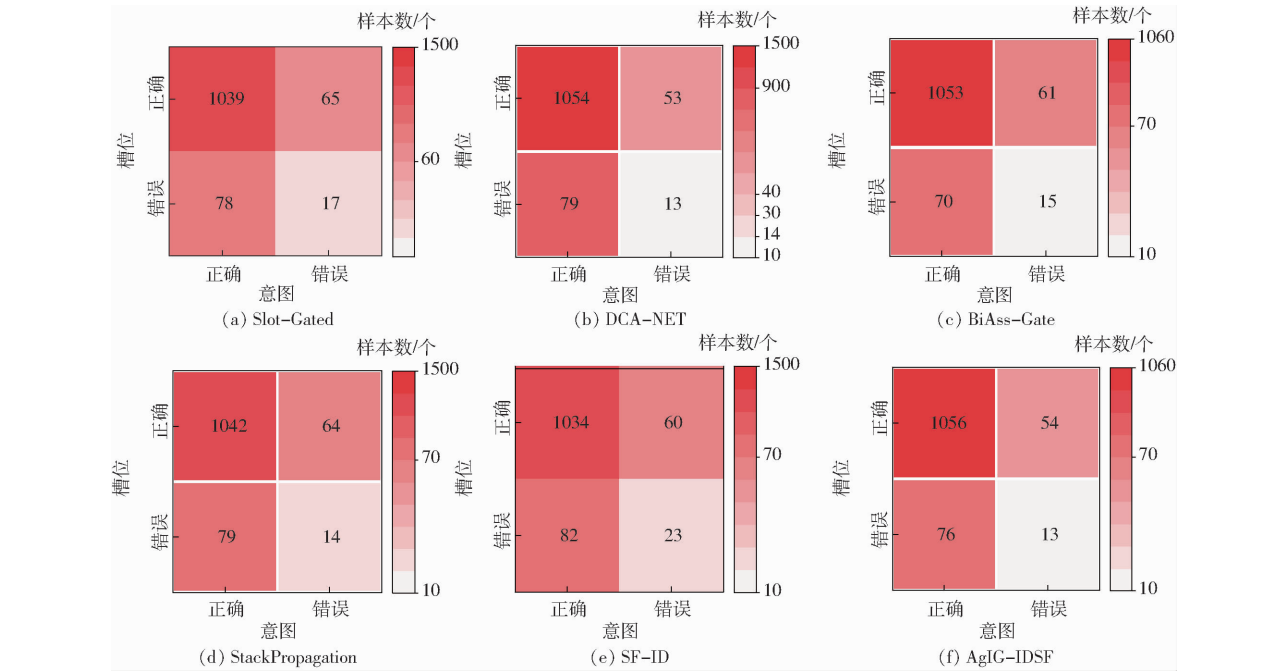


图 7 句子级意图识别与槽位填充混淆矩阵

Fig.7 Sentence-level confusion matrix of intent detection and slot filling for AGIS dataset

3.3.5 BERT 有效性

以上研究均采用基于传统 word2vec 或 glove 预训练的词嵌入,但此类词嵌入是静态的,无法根据动态生成上下文相关的高质量文本表示。近年来,BERT 等预训练模型的提出极大促进了各项 NLP 任务的性能。因此,为探究本文所提模型在采用预训练模型条件下的有效性,采用 BERT 动态生成包含丰富语义的文本表示,并与 JointBERT 和 JointIDSF 在 3 种数据集上进行了对比实验,结果如表 7 所示。JointBERT 在 3 种基准数据集上的整体识别准确率分别为 88.82%、87.34% 和 88.90%,显然优于图 4 和表 5 所示多数模型的相应值,说明 BERT 有助于提高模型的意图识别与槽位填充性能。基于此,

JointIDSF 模型在 3 种数据集上的识别结果略有上升,这是因为该模型引入意图槽位注意力机制,能够使意图根据其重要性与句子中的每个字符动态融合,进而促使意图信息指导槽位填充任务。然而,由于该模型所提意图槽位注意力机制强依赖于预测的意图概率矩阵,未引入意图嵌入信息,因此对整体识别性能提升能力有限。而本文所提 AgIG-IDSF 模型在 3 种数据集上均取得最优整体识别准确率,分别为 89.15%、87.57% 和 92.29%,比 JointIDSF 模型分别提升 0.09、0.22、2.29 个百分点。总之,上述结果说明了 BERT 的有效性,还表明了 AgIG-IDSF 模型在采用 BERT 情况下能够进一步提高农业病虫害意图识别与槽位填充性能。

表 7 基于 BERT 的模型在各数据集上的识别结果
Tab.7 Experimental results for BERT-based models on three datasets %

模型	AGIS			ATIS			SNIPS		
	槽位 F1 值	意图准确率	整体准确率	槽位 F1 值	意图准确率	整体准确率	槽位 F1 值	意图准确率	整体准确率
JointBERT	94.36	94.58	88.82	95.77	97.09	87.34	95.20	98.70	88.90
JointIDSF	93.09	95.70	89.06	95.52	97.64	87.35	95.74	98.29	90.00
AgIG-IDSF	94.27	94.82	89.15	95.52	97.54	87.57	96.77	98.14	92.29

4 结束语

本文主要针对农业病虫害领域意图识别与槽位填充存在的语料有限、现有模型未充分利用意图嵌入表示信息等问题,构建了首个农业病虫害领域的意图识别与槽位填充数据集,并提出了面向该领域的基于意图嵌入信息和门控机制的意图识别与槽位填充联合模型 AgIG-IDSF。该模型在采用注意力

机制进一步丰富上下文语义基础上,引入了意图嵌入表示信息,并采用门控机制对上述向量进行融合,从而生成了包含丰富意图信息的槽位上下文语义。实验结果表明,本文所提 AgIG-IDSF 模型能够更好地利用意图信息指导槽位填充任务,与其他模型相比在农业病虫害的意图识别与槽位填充任务上性能更优,且具有一定的泛化性和稳定性,为构建基于知识图谱的农业病虫害问答系统提供了技术支撑。

参 考 文 献

[1] QIN L, WEI F, NI M, et al. Multi-domain spoken language understanding using domain-and task-aware parameterization[J]. Transactions on Asian and Low-Resource Language Information Processing, 2022, 21(4): 1-17.

[2] SAHA T, PRIYA N, SAHA S, et al. A transformer based multi-task model for domain classification, intent detection and slot-filling[C]//2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021: 1-8.

[3] LIN T E, XU H. Deep unknown intent detection with margin loss[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 5491-5496.

[4] WANG Y, HUANG J, HE T, et al. Dialogue intent classification with character-CNN-BGRU networks[J]. Multimedia Tools and Applications, 2020, 79(7): 4553-4572.

[5] XIA C, ZHANG C, YAN X, et al. Zero-shot user intent detection via capsule neural networks[C]//Conference on Empirical Methods in Natural Language Processing, 2019.

[6] LIU Z, WINATA G I, XU P, et al. Coach: a coarse-to-fine approach for cross-domain slot filling[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 19.

[7] GONG Y, LUO X, ZHU Y, et al. Deep cascade multi-task learning for slot filling in online shopping assistant[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019: 6465-6472.

[8] ZHANG Jianguo, BUI T, YOON S, et al. Few-shot intent detection via contrastive pre-training and fine-tuning[C]//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, 2021: 1906-1912.

[9] 郑冰智,吴惠萍,朱定局,等.基于荔枝和龙眼病虫害知识图谱的问答系统[J].计算机与数字工程,2021,49(12):2618-2622.

ZHENG Yongzhi, WU Huilin, ZHU Dingju et al. Question and answer system based on the knowledge graphs of litchi and

- longan diseases and insect pests[J]. Computer & Digital Engineering, 2021, 49(12):2618–2622. (in Chinese)
- [10] 吴赛赛. 基于知识图谱的作物病虫害智能问答系统设计与实现[D]. 北京:中国农业科学院,2021.
WU Saisai. Design and implementation of intelligent question and answering system for crop diseases and pests based on knowledge graph[D]. Beijing: Chinese Academy of Agricultural Sciences, 2021. (in Chinese)
- [11] 张博凯,李想. 基于知识图谱的 Android 端农技智能问答系统研究[J]. 农业机械学报,2021,52(增刊):164–171.
ZHANG Bokai, LI Xiang. Design of agricultural question answering system based on knowledge graph[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(Supp.):164–171. (in Chinese)
- [12] 周晗. 基于知识图谱的水稻病虫害问答研究[D]. 北京:中国农业大学,2021.
ZHOU Han. Research for rice diseases and insect pests question answering based on knowledge graph[D]. Beijing: China Agricultural University, 2021. (in Chinese)
- [13] YANG P,JI D,AI C, et al. AISE: attending to intent and slots explicitly for better spoken language understanding[J]. Knowledge-Based Systems,2021,211:106537.
- [14] MA Ziyu, SUN Bin, LI Shutao. A two-stage selective fusion framework for joint intent detection and slot filling[C]//IEEE Transactions on Neural Networks and Learning Systems,2022:1–12.
- [15] ZHANG X, WANG H. A joint model of intent determination and slot filling for spoken language understanding[C]//Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, 2016: 2993–2999.
- [16] LIU B, LANE I. Attention-based recurrent neural network models for joint intent detection and slot filling[J]. arXiv preprint arXiv:1609.01454, 2016.
- [17] CHEN Q, ZHUO Z, WANG W. Bert for joint intent classification and slot filling[J]. arXiv preprint arXiv:1902.10909, 2019.
- [18] GOO C W, GAO G, HSU Y K, et al. Slot-gated modeling for joint slot filling and intent prediction[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018: 753–757.
- [19] QIN L, CHE W, LI Y, et al. A stack-propagation framework with token-level intent detection for spoken language understanding[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), 2019: 2078–2087.
- [20] DAO M H, TRUONG T H, NGUYEN D Q. Intent detection and slot filling for Vietnamese[J]. arXiv preprint arXiv:2104.02021, 2021.
- [21] SUN Chengai, LÜ Liangyu, LIU Tailu, et al. A joint model based on interactive gate mechanism for spoken language understanding[J]. Applied Intelligence,2022,52(6):6057–6064.
- [22] 王丽花,杨文忠,姚苗,等. 意图识别与语义槽填充的双向关联模型[J]. 计算机工程与应用,2021,57(3):196–202.
WANG Lihua, YANG Wenzhong, YAO Miao, et al. Bidirectional association model for intent detection and slot filling[J]. Computer Engineering and Applications, 2021, 57(3):196–202. (in Chinese)
- [23] HAIHONG E, NIU P, CHEN Z, et al. A novel bi-directional interrelated model for joint intent detection and slot filling[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 5467–5471.
- [24] QIN L, LIU T, CHE W, et al. A co-interactive transformer for joint slot filling and intent detection[C]//ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 8193–8197.
- [25] GUO X, ZHOU H, SU J, et al. Chinese agricultural diseases and pests named entity recognition with multi-scale local context features and self-attention mechanism[J]. Computers and Electronics in Agriculture, 2020, 179: 105830.
- [26] 郭旭超,唐詹,刁磊,等. 基于部首嵌入和注意力机制的病虫害命名实体识别[J]. 农业机械学报,2020,51(增刊2):335–343.
GUO Xuchao,TANG Zhan,DIAO Lei, et al. Recognition of Chinese agricultural diseases and pests named entity with joint radical embedding and self-attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(Supp.2):335–343. (in Chinese)
- [27] 吴赛赛,周爱莲,谢能付,等. 基于深度学习的作物病虫害可视化知识图谱构建[J]. 农业工程学报,2020,36(24):177–185.
WU Saisai, ZHOU Ailian, XIE Nengfu, et al. Construction of visualization domain-specific knowledge graph of crop diseases and pests based on deep learning[J]. Transactions of the CSAE, 2020, 36(24):177–185. (in Chinese)
- [28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems (NIPS 2017), 2017.
- [29] 刘莉,陶红燕,方静,等. 基于近红外高光谱的梨叶片炭疽病与黑斑病识别[J]. 农业机械学报,2022,53(2):221–230.
LIU Li, TAO Hongyan, FANG Jing, et al. Identifying anthracnose and black spot of pear leaves on near-infrared hyperspectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(2):221–230. (in Chinese)
- [30] LI S, ZHAO Z, HU R, et al. Analogical reasoning on Chinese morphological and semantic relations[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2018: 138–143.
- [31] QIN L, XU X, CHE W, et al. AGIF: an adaptive graph-interactive framework for joint multiple intent detection and slot filling[C]//Findings of the Association for Computational Linguistics: EMNLP 2020, 2020: 1807–1816.