

基于知识图谱的水产养殖病害诊断技术研究

陆光豪, 李海涛, 赵瑞金

(青岛科技大学 信息科学技术学院, 山东 青岛 266061)

摘要: 水产养殖病害是影响水产养殖效益的重要因素, 由于水产养殖病害文本数据杂乱无章, 无法快速准确定位疾病原因, 从而耽误诊断和治疗时机, 导致水产养殖质量和产量下降; 为解决上述问题, 深入知识图谱的工作原理和模型特征, 采用知识图谱技术完成水产养殖病害诊断总体方案设计, 建立水产病害语料库, 引入 H-BIO 标注策略, 完成标注方案设计、改进 BiLSTM 模型构建, 进行实体关系抽取和水产病害模型训练, 完成水产养殖病害知识图谱可视化设计, 并进行水产病害联合抽取实验; 实验结果表明: 基于知识图谱的改进 BiLSTM 模型在实体关系抽取方面效果较好、可靠性较高, 有效提高了水产病害联合抽取准确率, 构建了水产养殖病害可视化知识图谱, 能够辅助作业人员快速准确进行水产病害诊断和治疗, 对提升水产养殖生产效益具有十分重要的作用。

关键词: 水产养殖; 病害诊断; 知识图谱; H-BIO 标注; BiLSTM 模型

Research on Diagnosis Technology of Aquaculture Diseases Based on Knowledge Graph

LU Guanghao, LI Haitao, ZHAO Ruijin

(College of Information Science and Technology, Qindao University of Science and Technology, Qingdao 266061, China)

Abstract: Diseases in aquaculture are an important factor of affecting the efficiency of aquaculture. Due to the disorderly text data of aquaculture diseases, it is difficult to quickly and accurately locate the causes of diseases, which delays diagnosis and treatment, leading to a decrease in the quality and yield of aquaculture. To solve the above problems, the working principles and model features of knowledge graph are deeply studied, knowledge graph technology is used to implement the overall design of aquaculture disease diagnosis, establish a corpus of aquaculture diseases, introduce the H-BIO annotation strategy, complete the annotation scheme design, improve the construction of the BiLSTM model, extract the entity relationships and train aquaculture disease models, complete the visualization design of aquaculture disease knowledge graphs, and conduct the experiments on joint extraction of aquaculture diseases. Experimental results show that the improved BiLSTM model based on knowledge graph has good performance and high reliability in entity relationship extraction, effectively improving the accuracy of joint extraction of aquatic diseases. The visual knowledge graph of aquatic disease is constructed, which can help operators quickly and accurately diagnose and treat aquatic diseases. It plays a very important role in improving the efficiency of aquaculture.

Keywords: aquaculture; disease diagnosis; knowledge graph; H-BIO annotation; BiLSTM model

0 引言

随着国民经济的快速发展, 我国水产养殖规模也逐年扩大, 而水产养殖的病害问题一直是困扰水产养殖效益的关键。但是由于水产养殖数量大、种类多, 产生水产养殖病害的原因也千奇百怪, 作业人员无法快速诊断和治疗, 造成水产养殖质量和产量上不去, 有时甚至会出现数百亿的亏损^[1]。因此对复杂的水产养殖病害进行诊断研究具有十分重要的意义。

国内外许多学者开始研究医学知识图谱, 文献 [2] 通过构建概念图的方式从而研究形成了一个医学知识图谱, 文献 [3] 应用知识图谱的结构框架构建了基于中文电子病

历知识图谱。还有学者针对知识图谱的语料库进行优化, 文献 [4] 对疾病特征进行深层次挖掘, 构建了开放式中文症状语料库, 文献 [5] 提出了基于 CNN 的文本分类模型, 使用预训练模型的词嵌入向量作为输入, 获取了更全面的特征。但由于水产病害文本句式特征复杂, 专有名词过长且频度副词多, 使得以往的模型不能全面地学习文本特征, 因此上述研究无法直接应用在水产病害的诊断中。文献 [6] 在提出了 BP-MLL 算法, 通过神经网络来解决多标签文本分类问题。文献 [7] 运用交叉熵损失函数对 BP-MLL 算法进行改进, 但是神经网络模型都较为简单, 无法表达文本的完整性。因此, 后来又出现了许多基于 CNN、RNN

收稿日期: 2024-03-07; 修回日期: 2024-04-24。

基金项目: 山东省重点研发计划(科技示范工程)课题(2021SFGC0701), 青岛市海洋科技创新专项(22-3-3-hygg-3-hy)。

作者简介: 陆光豪(1997-), 男, 硕士。

通讯作者: 李海涛(1978-), 男, 博士, 副教授。

引用格式: 陆光豪, 李海涛, 赵瑞金. 基于知识图谱的水产养殖病害诊断技术研究[J]. 计算机测量与控制, 2024, 32(9): 101-107.

和 Transformer 的深度学习文本分类模型，用于解决上述问题。

本文以提升水产养殖病害诊断效果为目标，深入研究知识图谱技术，将其应用在水产养殖病害诊断中，通过构建水产养殖病害语料库，将不同类型的病害数据转化为图文数据，整合水产养殖病害信息，实现病害信息资源的继承和共享，帮助作业人员快速准确查询自己所需的数据信息，辅助完成水产养殖病害诊断和治疗，从而提高水产养殖病害诊断效率，提升水产养殖生产效益，同时对实现渔业可持续发展具有一定的借鉴意义。

1 知识图谱构建总体方案

知识图谱是一种多实体关系网，反映不同文本实体之间相互关联情况^[8]，主要是通过对杂乱无序的文本数据实体及其相互关系进行数据处理、提取，将其转化为结构化的文本数据，可以帮助作业人员快速、准确查询水产养殖病害信息，达到高效辅助疾病诊断和治疗的目的。

水产病害诊断系统知识图谱总体结构分为 4 层：语料库构建层、模型训练层、技术模块层和页面展示层，语料库构建层为整个知识图谱提供数据来源，经过数据处理后进入模型训练层进行计算、优化，最后通过技术模块实现在终端的可视化展示。

图 1 为基于知识图谱的水产养殖病害诊断系统总体方案。

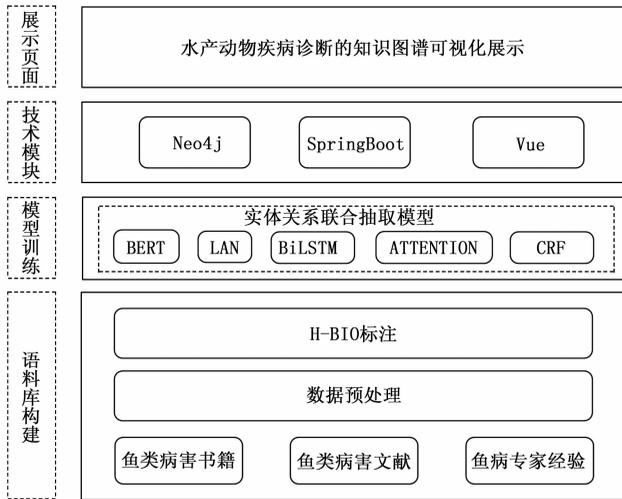


图 1 水产养殖病害知识图谱构建总体方案

语料库构建层是基于知识图谱的水产养殖病害诊断系统的基础，数据来源较为广泛，包括水产养殖病害图文书籍、网络文献以及专家经验等，语料库构建是通过不同的方法收集水产养殖病害信息，通过数据预处理后筛选出有用数据，再利用 H-BIO 标注方法进行标注，形成结构化数据。

模型训练层是知识图谱的核心，利用 BERT、LAN、BiLSTM 等模型对语料库构建层传输的结构化数据进行实体关系联合抽取，从而得到不同类型的水产病害实体关系。

页面展示层是基于知识图谱的水产养殖病害诊断系统的终端，通过开源数据库 Neo4j 存储抽取的三元组数据，利用 SpringBoot 和 Vue 模块完成知识图谱可视化设计，方便作业人员通过可视化界面快速掌握水产病害种类和防治方法，提高诊断效率。

2 水产病害语料库建立

2.1 标注方案设计

本文构建的水产病害语料库数据主要从互联网、百度百科、网络文献等来源获取，数据量为 43 万字符。由于从不同途径搜集的水产病害数据中含有大量无用信息，不进行有效筛选会造成水产病害语料库冗杂，降低信息搜索效率。同时水产病害数据存在类型多、格式不规范等问题，无法直接用于知识图谱构建。因此需要采取合适的标注策略，对语料库数据文本进行标注，标注方案如图 2 所示。

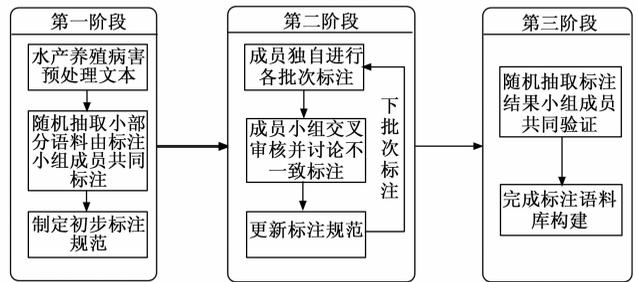


图 2 语料库数据文本标注方案

由图 2 可知，水产病害语料库标注方案主要分为 3 个阶段^[9]：

第一阶段主要完成水产病害语料库的预处理，由标注小组成员随机抽取部分语料数据文本进行共同标注，制定初步标注规范。

第二阶段主要由标注小组成员分批次各自完成各自分配的标注任务，再进行交叉审核，对不一致的标注内容进行讨论，更新标注规范后继续交叉审核，直至完成全部标注工作。

第三阶段为水产病害文本数据标注验证环节，标注小组成员随机抽取标注结果进行验证，验证有效后将水产病害语料数据进行存储，完成语料数据库的构建。

2.2 语料库预处理

语料库数据预处理主要是通过文本格式转换、语料清洗、预料标注和文本表示学习等操作，将语料库数据转换为当前需要的文本格式，去除其中的干扰字、错别字等，利用合适的标准规则对水产病害语料数据进行人工标注，从而得到标准的水产病害语料数据库^[10]。语料库预处理流程如图 3 所示。

通过文本格式转换、语料清洗、预料标注和文本表示学习等操作，将语料库数据转换为当前需要的文本格式，去除其中的干扰字、错别字等，利用合适的标准规则对水产病害语料数据进行人工标注，从而得到标准的水产病害语料数据库。

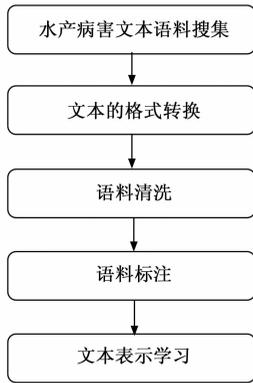


图 3 语料库预处理流程

通过对语料数据库中的病害实体进行分类, 定义 6 类实体关系, 分别为引发、感染、流行温度、症状和防治, 具体实体对应关系如表 1 所示^[11]。

表 1 具体实体对应关系

实体 1	关系	实体 2
病害	引发	病原体
	感染	易感鱼群
	流行温度	温度
	症状	患病症状
	防治	防治药物

以实体 1 是草鱼出血病为例, 若实体 2 是草鱼出血病病毒, 说明实体 1 和实体 2 是引发关系。

若实体 2 是草鱼, 草鱼易感染出血病, 说明实体 1 和实体 2 是感染关系。

若实体 2 是流行温度, 草鱼出血病的易感温度为流行温度, 说明实体 1 和实体 2 是流行温度关系。

若实体 2 是全身出血, 草鱼出血病症状为全身出血, 说明实体 1 和实体 2 是症状关系。

若实体 2 是草鱼出血症疫苗和氯化钠, 草鱼出血病需疫苗和氯化钠治疗, 说明实体 1 和实体 2 是防治关系。

2.3 标注结果分析

水产病害语料数据库中数据文本受各种语法和语义的影响, 分词难度大、准确率低, 因此本文采取 H-BIO 标准策略进行文本标注, 从而以标签形式表示水产病害与其他实体之间的相互关系, 水产病害实体关系标签如表 2 所示^[12]。

表 2 水产病害实体关系标签

标签种类	示例	数量
非实体	O	1
关系实体	H-1, I-H-1, HB-YF-2, HI-YF-2, HB-GR-2, HI-GR-2, HB-WD-2, HI-WD-2, HB-BS-2, HI-BS-2, HB-FZ-2, HI-FZ-2	12

本文针对水产病害疾病和引发、感染、流行温度、症

状和防治这 6 类实体及相互关系进行标注, 采用参数调优方法, 对标注长度、标签等内容进行多轮迭代标注, 通过对比实验标签与实际标签情况, 计算实验标签标准准确率, 从而验证标注方案的准确性, 语料库标注结果如表 3 所示^[14]。

表 3 语料库标注结果验证

实体关系	第一次/%	第二次/%	第三次/%	最终/%
引发关系	0.825	0.868	0.882	0.922
感染关系	0.763	0.806	0.824	0.830
流行温度关系	0.816	0.832	0.841	0.844
症状关系	0.741	0.790	0.820	0.831
防治关系	0.790	0.822	0.836	0.839
综合 F 值	0.787	0.823	0.841	0.853

由表 3 可知: 经过 3 轮循环标注, 标注结果的准确性逐渐增加, 最终标注一致性达到 0.853, 达到了 80% 以上, 说明本文建立的语料库标注结果有效, 可以用于水产病害的实体关系抽取模型中。

3 水产病害模型训练

3.1 改进 BiLSTM 模型构建

水产养殖病害语料库文本数据容易出现大量实体重叠的情况, 而传统的实体识别模型具有一定的领域局限性, 只能对特定领域实体进行识别, 无法覆盖水产医学全领域, 因此本文采用改进型 BiLSTM 模型对文本数据进行实体识别。

本文研究的改进型 BiLSTM 模型主要由 BERT 模型、LAN 层、BiLSTM 模型、Attention 模型和 CRF 模型这 5 个部分组成^[15], 改进型 BiLSTM 模型结构如图 4 所示。

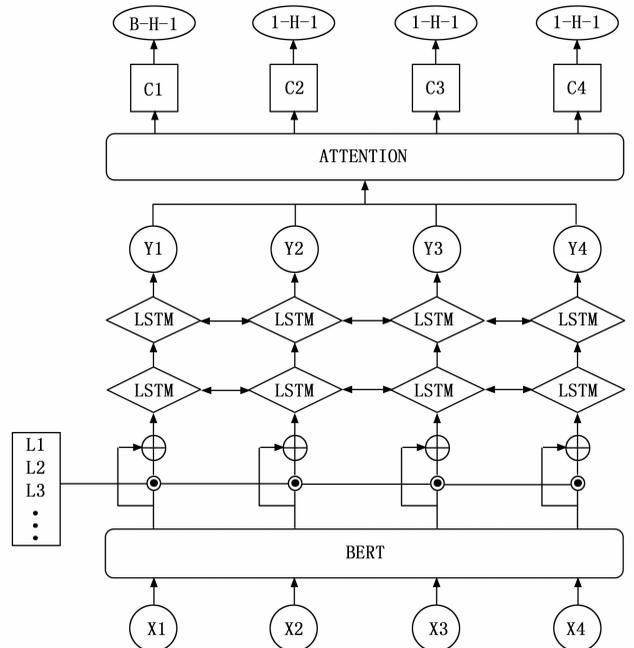


图 4 改进型 BiLSTM 模型结构示意图

1) BERT 模型:

BERT 模型为预训练模型, 主要由字向量、句子向量和位置向量组成, 能够双向表示输入层的文本数据上下文相互关系, 为双向自注意力表示模型。由于 BERT 模型无法记录时序信息, 因此增加位置向量进行定位, 提高关系抽取准确性^[16]。图 5 为水产病害的 BERT 模型输入示意图。

2) LAN 层:

LAN 层主要是进行标签信息嵌入, 而标签信息包含有水产病害实体之间相互关系的信息, 因此在水产病害语料库文本句子中嵌入设定的标签信息, 可以帮助作业人员快速获取带有标签信息的有效数据^[17]。

经过 BERT 模型预训练后的数据文本序列与标签映射到词嵌入空间, 进而得到文本矩阵 $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ 和标签矩阵 $\mathbf{K} = \{K_1, K_2, \dots, K_k\}$ 。将对 BERT 模型输出的向量与对应的标签进行映射, 在得到 BERT 模型的全连接矩阵和标签矩阵后, 将这两个矩阵进行拼接, \mathbf{K} 个标签对应的是一个 k 行 k 列的矩阵, 矩阵的每一行对应一个标签, 通过缩放点积注意力实现标签嵌入的计算, 计算公式为:

$$Q = \frac{\mathbf{X} * \mathbf{K}^T}{\sqrt{K}} \quad (4)$$

3) BiLSTM 模型:

水产病害语料库文本数据含有长依赖信息, 同时还存在大量专有名词等反向依赖信息, 因此本文采取 BiLSTM 模型, 通过两个前向和后向的 LSTM 层叠加, 有效获取距离较长的不同实体双向依赖关系^[18]。图 6 为水产病害的 BiLSTM 模型结构。

假设词嵌入层的特征用 w_i 表示, w_i 传入 LSTM 层前向和后向后得到的上文表示和下文表示计算公式为^[19]:

$$C_f(w_i) = f(W^{(f)} C_f(w_{i-1})) \quad (5)$$

$$C_b(w_i) = f[W^{(b)} C_b(w_{i+1})] \quad (6)$$

$$x_i = [C_f(w_i), C_b(w_i)] \quad (7)$$

其中: $C_f(w_i)$ 为 w_i 传入前向 LSTM 层后得到的上文表示, $C_b(w_i)$ 为 w_i 传入后向 LSTM 层后得到的下文表示, x_i 表示上文表示 $C_f(w_i)$ 和下文表示 $C_b(w_i)$ 的合集。

最后通过 \tanh 激活函数计算得到潜在语义向量 y_i ^[20]:

$$y_i = \tanh(Wx_i + b) \quad (8)$$

其中: W 为词嵌入层的特征合集, b 为激活函数拟合系数。

4) Attention 模型:

Attention 模型又称注意力机制模型, 通过计算水产病害语料库不同文本字的权重值及归一化处理, 使文本字一致性的权重高于非一致性的权重, 提高实体关系提取的准确性^[21]。水产病害的 Attention 模型计算过程如图 7 所示。

Attention 模型训练的第一步是利用不同的查询方法和关键字匹配计算方法计算关键字权重值, 第二步是利用类 SoftMax 函数进行归一化处理, 计算公式为^[22]:

$$a_i = \text{Softmax}(\text{Sim}_i) = \frac{e^{\text{Sim}_i}}{\sum_{j=1}^L e^{\text{Sim}_j}} \quad (9)$$

第三步是将 a_i 的计算结果进行加权求和, 得到 Attention 模型结果^[23]:

$$\text{Attention}(\text{Query}, \text{Source}) = \sum_{i=1}^L a_i \cdot \text{Value}_i \quad (10)$$

5) CRF 模型:

CRF 模型主要是在 Attention 模型输出的标签中添加相应约束条件, 从而确保输出结果符合条件。相应的约束条件在语料库数据训练时会自动被学习, 再传输至下一个模型中。本文研究的水产病害语料库中, 最后输出的结果都以标签“HB-”或“B-”或“O”开头, 如果出现“HI-”开头, 代表该标签不符合规范。同一个实体的关系标签必须保持一致, 否则会造成混乱, 例如“HB-lable1-2 HI-lable2-2HI-lable3-2 HI-...”, lable1, lable2, lable3 应该属于相同的关系标签。为了避免输出不合法的标签序列, 在 Attention 模型输出的标签中添加 CRF 模型来进行约束限制, 从而大大减少输出的结果中非法标签的数量。

假设输入序列为 $\mathbf{X} = (X_1, X_2, \dots, X_n)$, Attention 模型输出的概率矩阵为 \mathbf{P} , 输出标注序列为 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$, 则输出的预测标签结果为:

$$\mathbf{S}(\mathbf{X}, \mathbf{y}) = \sum_{i=0}^n \mathbf{A}_{y_i, y_{(i+1)}} + \sum_{i=1}^n \mathbf{P}_{i, y_i} \quad (11)$$

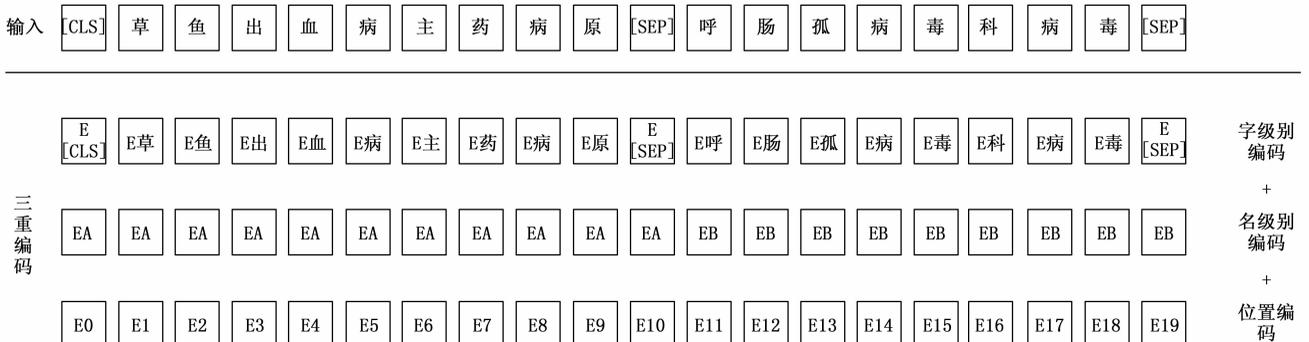


图 5 水产病害的 BERT 模型输入示意图

的症状有时也会表现出相同的症状。因此，通过知识图谱可以将水产动物所患疾病种类、表现症状等都一一对应，养殖户或者动物医生则根据病鱼的体表症状以及其他的并发症状，在知识图谱上推断出可能患有的水产动物疾病，从而快速定位水产病害类型，避免对水产动物疾病产生错误的判断。例如一个水产动物可能患有几种不同的水产病害，而各种发病的症状又是相同，如果按照经验来判断这个水产动物患的什么病，会导致因为经验不足分辨不清具体病原体，而造成错误的判断，影响救治水产动物的最佳时期。可视化知识图谱可以帮助养殖户根据其他的并发症或者信息对这些疾病进行区分判断，通过进一步地观察、检查最终对所患疾病进行确认，从而选择合适准确的治疗方式进行诊断治疗，提高救治效率。

由图 10 可知：不同实体类型用不同颜色进行区分，两节点间的信息表上不同实体之间的语义关系，本文研究的实体类型和颜色的对应关系如表 4 所示。

表 4 实体类型和颜色的对应关系

实体类型	颜色
病害名称	黄色
易感鱼群	蓝色
流行温度	粉色
患病症状	橘色
防治药物	绿色

通过知识图谱可以将水产动物所患疾病、患病症状、防治药物等信息一一对应，辅助作业人员根据水产动物患病症状推断患病情况，从而科学采取有效防治手段进行治疗。

可视化知识图谱支持图片显示功能，可以显示出水产疾病的症状，帮助作业人员精准判断；同时通过爬虫技术可以实现简单问题解答功能，辅助完成水产病害诊断和治疗。

5 实验与分析

为验证 H-BIO 标注策略和实体关系抽取模型的有效性，本文将构建的水产病害语料库 43 万数据按 8 : 2 比例划分训练集和测试集，通过精确率 P 、召回率 R 和调和平均数 F_1 来测量抽取结果。各评价指标的计算公式为^[13]：

$$P = \frac{S}{Q} \times 100\% \quad (1)$$

$$R = \frac{S}{F} \times 100\% \quad (2)$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \times 100\% \quad (3)$$

其中： S 表示正确抽取到的水产动物疾病关系的系数； Q 表示抽取到的水产动物疾病关系的系数； F_1 表示水产动物疾病关系的系数。

进行实体关系抽取模型训练时，需对模型参数进行优化设置，通过测试得到最优模型参数值，如表 5 所示。

在模型参数值不变的情况下，本文通过将改进型 BiLSTM 模型与其他传统模型 (Bi LSTM+CRF 模型、BERT +BiGRU+CRF 模型、BERT + BiLSTM+CRF 模型) 进行对比试验^[26]，试验结果如表 6 所示。

表 5 最优模型参数值

模型参数	参数值
学习率	1×10^{-5}
批处理	128
迭代次数	200
衰减度	0.5
标签嵌入层数	6
BiLSTM 层数	2

表 6 不同模型试验结果

模型	精确率 P /%	召回率 R /%	F_1 值 /%
BiLSTM+CRF 模型	83.31	83.64	83.47
BERT+BiGRU+CRF 模型	85.64	85.82	85.60
BERT+BiLSTM+CRF 模型	86.12	86.58	86.22
本模型	92.61	92.20	92.86

由表 6 可知：运用 Word2Vec 作为输入的 BiLSTM+CRF 模型，通过提取特征，模型准确率、召回率、 F_1 值分别为 83.31%、83.64% 和 83.47%。Word2Vec 方法是通过假设词的位置信息不重要，将词映射为词向量，但在实际的水产医学语料库中，不同上下文语境中相同词表达的含义可能不同，因此实体的位置信息不能忽略^[27]。而 BiLSTM+CRF 模型没考虑位置信息对水产医学实体的影响，导致模型识别效果不够好。BERT+BiLSTM+CRF 模型在 BiLSTM+CRF 模型的基础上引入 BERT 模型，有效解决了 BiLSTM+CRF 模型存在的一词多义问题，其模型准确率、召回率、 F_1 值分别提高了 2.37%、1.83% 和 2.10%。BERT+BiLSTM+ATT+CRF 模型则同时引入 BERT 模型和 Attention 模型，Attention 模型有效解决了水产病害语料库中因长序列语句而出现实体标注不一致问题，但在水产医学命名实体识别任务中添加 Attention 机制与 BERT 中的 self-Attention 机制作用类似^[28]，因此与 BERT+BiLSTM+CRF 模型相比，模型准确率、召回率、 F_1 值分别提高了 0.12%、0.22% 和 0.17%。针对上述模型没有充分考虑嵌套实体识别的问题，本文设计的实体关系抽取模型增加了 LAN 层和注意力机制模型， F_1 值高达 92.86%，精确率、召回率、 F_1 值分别提升了 6.97%、6.38%、6.26%，大大提高了关系抽取准确性，验证了改进型 BiLSTM 模型的有效性，具有更好的识别效果，水产医学命名实体中嵌套实体识别准确率有了一定提高。

改进 BiLSTM 模型有效解决了水产病害文本因一文多类而造成数据集不足的问题，当水产病害出现多种状况后，能够快速识别水产动物名称、发病部位、病原体、水产动

物疾病、药品名称、临床症状 6 类实体, 确定病害与引发、感染、流行温度、症状及防治之间的关系, 从而能够帮助工作人员快速识别水产病害特征, 提出相应的治疗方案, 快速完成水产病害诊治。

6 结束语

为解决水产病害数据杂乱、诊断难度大、治疗不及时等问题, 本文设计研究了基于知识图谱的水产养殖病害诊断技术, 完成了知识图谱构建总体方案设计, 通过优化标注方案, 完成水产病害语料库预处理, 确定病害、病原体、易感鱼群、流行温度、患病症状和防治药物这 6 种实体以及不同实体之间对应的关系。采取 H-BIO 标准策略完成水产病害语料文本数据的标注, 分析 BERT 模型、BiLSTM 模型、Attention 模型和 CRF 模型计算方法, 建立改进型 BiLSTM 模型, 完成知识图谱可视化设计, 最后进行实验验证, 实验结果表明: 基于知识图谱的水产养殖病害诊断技术具有较好的文本识别效果, 病害诊断准确率较高, 能够帮助作业人员快速准确定位病害种类, 辅助提供治疗方法, 能有效提高水产养殖质量和产量, 提升水产养殖效益。

参考文献:

[1] 李悦. 农作物病虫害知识图谱构建研究 [D]. 北京: 中国农业科学院, 2021.

[2] 袁光辉. 基于电子病历的知识图谱构建研究 [D]. 大连: 大连理工大学, 2021.

[3] 姜京池. 基于医学知识图谱的疾病诊断与健康预测模型研究 [D]. 哈尔滨: 哈尔滨工业大学, 2019.

[4] HE B, DONG B, GUAN Y, et al. Building a comprehensive syntactic and semantic corpus of chinese clinical texts [J]. *Journal of Biomedical Informatics*, 2017, 69 (7): 203-217.

[5] 贺琳, 张雨, 巴韩飞. 基于注意力机制和深度学习模型的外来海洋生物命名实体识别 [J]. *大连海洋大学学报*, 2021, 36 (3): 503-509.

[6] 孙娟娟, 于红, 冯艳红, 等. 基于深度学习的渔业领域命名实体识别 [J]. *大连海洋大学学报*, 2018, 33 (2): 265-269.

[7] KIM S, LIU H, YEGANOVA L, et al. Extracting drug-drug interactions from liter at reusing a rich feature-based linear kernel approach [J]. *Journal of Biomedical Informatics*, 2015, 55 (2): 23-30.

[8] 王智广, 文红英, 鲁强, 等. 地质领域开放式实体关系联合抽取 [J]. *计算机工程与设计*, 2021, 42 (4): 996-1005.

[9] 徐进. 基于深度学习的实体关系联合抽取 [D]. 成都: 电子科技大学, 2021.

[10] 张军莲, 张一帆, 汪鸣泉, 等. 基于图卷积神经网络的中文实体关系联合抽取 [J]. *计算机工程*, 2021, 47 (12): 103-111.

[11] 史冰莹, 李佳琦, 张磊. 基于 CNN 的农作物病虫害图像识别模型 [J]. *计算机系统应用*, 2020, 29 (6): 89-96.

[12] 孙弋, 梁兵涛. 基于 BERT 和多头注意力的中文命名实体识别方法 [J]. *重庆邮电大学学报 (自然科学版)*, 2021: 1

-10.

[13] 马建红, 魏字默, 陈亚萌. 基于信息融合标注的实体及关系联合抽取方法 [J]. *计算机应用与软件*, 2021, 38 (7): 8-10.

[14] DAI D, XIAO X, LUY Y, et al. Joint extraction of entities and overlapping relations using position-attentive sequence labeling [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33 (10): 6300-6308.

[15] 高大林. 融合注意力机制与图神经网络的实体关系联合抽取 [D]. 徐州: 中国矿业大学, 2021.

[16] 杨锦锋, 关毅, 何彬, 等. 中文电子病历命名实体和实体关系语料库构建 [J]. *Journal of Software*, 2016, 27 (11): 2725-2746.

[17] 姜鑫, 张思佳. 面向水产领域疾病的语料库构建 [J]. *现代畜牧科技*, 2021 (11): 38-39.

[18] 孙红, 陈强越. 融合 BERT 词嵌入和注意力机制的中文文本分类 [J]. *小型微型计算机系统*, 2021, 7 (14): 1-6.

[19] 胡玉琦, 李婧, 常艳鹏, 等. 引入注意力机制的 Bi GRU-CNN 情感分类模型 [J]. *小型微型计算机系统*, 2020, 41 (8): 1602-1607.

[20] LIU N, WANG Q, REN J. Label-embedding bi-directional attentive model for multi-label text classification [J]. *Neural Processing Letters*, 2021, 53 (9): 1-15.

[21] 姜鑫. 水产动物疾病诊断的知识图谱构建 [D]. 大连: 大连海洋大学, 2022.

[22] 赵明, 杜亚茹, 杜会芳, 等. 植物领域知识图谱构建中本体非分类关系提取方法 [J]. *农业机械学报*, 2016, 47 (9): 278-284.

[23] AI-KHATIB K, HOU Y, WACHSMUTH H, et al. End-to-end argumentation knowledge graph construction [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34 (5): 7367-7374.

[24] 赵鹏飞, 赵春江, 吴华瑞, 等. 基于注意力机制的农业文本命名实体识别 [J]. *农业机械学报*, 2021, 52 (1): 185-192.

[25] 任媛, 于红, 杨鹤, 等. 融合注意力机制与 BERT+BiLSTM+CRF 模型的渔业标准定量指标识别 [J]. *农业工程学报*, 2021, 37 (10): 135-141.

[26] 刘巨升. 面向水产动物疾病诊断及防治的命名实体识别研究 [D]. 大连: 大连海洋大学, 2022.

[27] 杨鹤, 于红, 刘巨升, 等. 基于 BERT+BiLSTM+CRF 深度学习模型和多元组合数据增广的渔业标准命名实体识别 [J]. *大连海洋大学学报*, 2021, 36 (4): 661-669.

[28] 杨鹤, 于红, 刘巨升, 等. 基于 BERT+BiLSTM+CRF 深度学习模型和多元组合数据增广的渔业标准命名实体识别 [J]. *大连海洋大学学报*, 2021, 36 (4): 661-669.

[29] NOZZA D, MANCHANDA P, FERSINI E, et al. Learning to adapt with word embeddings: domain adaptation of named entity recognition systems [J]. *Information Processing & Management*, 2021, 58 (3): 221-230.