

基于深度学习的实体关系抽取方法文献综述

易战军 曾锐

宜春学院人工智能与信息工程学院 江西宜春 336000

摘要：实体关系抽取作为自然语言处理（NLP）领域信息抽取的核心任务，在知识图谱构建、智能问答等场景中发挥关键作用，其效率与准确性直接影响海量非结构化文本的价值挖掘。深度学习凭借自主特征学习能力成为实体关系抽取主流技术，递归神经网络（RNN）及其变体（LSTM、Bi-LSTM）通过时序建模解决长距离依赖问题，但存在参数训练复杂或训练周期长等不足；卷积神经网络（CNN）利用卷积与池化操作高效提取局部语义特征，结合位置特征等改进进一步提升性能，却在长句建模中存在局限；图神经网络（GNN）通过图结构与节点交互优化全局关系建模，在文档级实体关系抽取中优势显著，其优化聚焦依存树剪枝、多模态图表示等方向；联合学习通过参数共享或序列标注实现端到端建模，部分方法针对特定领域优化特征融合，提升抽取鲁棒性；大语言模型（LLM）借助大规模预训练与微调增强语义理解和泛化能力，在低资源、零样本场景及专业术语处理中表现突出，参数高效微调等技术进一步优化其性能。

关键词：实体关系抽取；RNN；LSTM；CNN；GNN；LLM

引言

实体关系抽取通过识别文本中的实体及其语义关系，在知识图谱构建、智能问答系统、个性化推荐等应用场景中发挥着关键作用。近年来，深度学习技术的突破性进展为实体关系抽取提供了全新的解决方案，显著提升了信息抽取的准确性和处理效率。本文系统梳理了基于深度学习的实体关系抽取方法，深入分析了其研究现状、核心技术、面临的挑战以及未来的发展方向。

1 基于深度学习的实体关系抽取研究进展

随着深度学习技术的兴起，基于神经网络的方法逐渐成为关系抽取的主流。与传统方法不同，深度学习模型具有自主学习能力，无需依赖人工进行特征提取。神经网络模型通过输入低维、连续的词向量，输出高维向量来表示语义，省去了人工特征选择步骤，避免了因特征选取不当而导致的效果偏差。基于深度学习的实体关系抽取方法已形成多个研究方向，主要包括循环神经网络模型（RNN）、卷积神经网络模型（CNN）、图神经网络模型（GNN）、联合学习方法以及大语言模型（LLM）等，为实体关系抽取研究提供了丰富的技术手段。

1.1 递归神经网络（RNN）在实体关系抽取中的应用

最早用于实体关系抽取的神经网络模型是递归神经网络（RNN）。在自然语言处理领域，Socher 等人^[1]开创性地将

矩阵 - 递归神经网络（MV-RNN）应用于句法依存分析任务中，该模型在于通过句法分析将文本的线性顺序转换为树结构，为每个树节点分别赋予向量和矩阵形式的表征，这种设计有效克服了传统词向量模型在长距离语义关联捕捉方面的局限性。实验结果表明，该模型在 SemEval-2010 task 8 数据库上实现了 82.2% 的 F 值抽取效果。然而，该模型有缺陷：首先，解析树的每个节点都需要独立训练参数；其次，句法分析误差或句子复杂度过高都会显著影响性能。递归神经网络（RNN）模型在实际应用中面临梯度爆炸问题，尤其是在处理需要长时间标注的输入序列时表现更为突出。这种现象导致靠前实体对后续实体的感知能力下降，最终影响分类性能。为解决这一问题，长短时记忆网络（LSTM）应运而生，LSTM 通过引入“门”的概念，利用遗忘门、输入门和输出门的控制机制，成功克服了传统 RNN 的缺陷，能够有效学习长距离依赖关系。在 LSTM 的基础上，研究者们进一步进行了改进和创新。Ningthoujam 等人^[2]提出了一种基于实体间最短依赖路径的 LSTM 模型，通过消除噪声干扰显著提升了实体识别的准确率，在 I2B2 数据集上取得了优异的实验效果。

1.2 卷积神经网络（CNN）在实体关系抽取中的应用

卷积神经网络（CNN）最初在图像处理领域得到了广泛应用。随后，研究者们将其成功扩展到自然语言处理任务

中。在实体关系抽取领域, Liu 等人^[3]在该任务中首次引入卷积神经网络 (CNN) 方法, 通过整合词汇特征, 并借助词典对输入词汇进行编码处理, 成功地将语义知识融入神经网络模型, 在 ACE2005 数据集的实验过程中, 该方法使 F1 值提升了 9%。Zeng 等人^[4]在此基础上进一步改进, 直接采用 CNN 模型提取局部语义特征, 并在句子级别特征抽取中引入位置特征 (position features, PF), 取得了显著的效果。Nguyen 等人^[5]对 CNN 结构进行了创新性改造, 提出了一种多卷积窗口的 CNN 模型, 它能够学习多层次粒度, 摆脱对外部词汇特征的依赖性, 从而丰富了输入句子的表示能力。与 Zeng 等人的方法相比, 其性能提升了 0.1%。在实体关系抽取领域。

1.3 图神经网络 (GNN) 在实体关系抽取中的应用

图神经网络 (GNN) 在文本数据处理中展现出独特优势, 其通过构建图结构, 利用节点间的信息交互机制, 能够有效识别实体间的关系。相较于传统的卷积神经网络 (CNN) 和循环神经网络 (RNN), 在全局关系建模上 GNN 具有显著优势, 特别是在文档级实体关系抽取任务中很强势。当前, GNN 技术的优化主要集中在三个关键方向: 依存树剪枝策略、多模态图表示学习以及动态边权重计算。这些技术创新不断推动着句法结构信息在深度语义理解中的应用广度和深度。在 GNN 的实际应用中, 图卷积神经网络 (GCN) 因其强大的关系建模能力而备受关注。Song 等人^[6]的研究表明, 通过构建依赖森林结构, 可以有效解决解析树噪声问题, 实现句法信息的精准提取。然而, 在实际应用中仍存在一定的局限性, 虽然能够有效结合词特征和句法特征, 但对实体间结构以及层次结构的关注仍有待提升。针对这一挑战, 研究者们提出了多种创新性解决方案。Fu 等人^[7]则采用了一种基于图卷积网络的堆叠式编码器方法, 创新性地将 BiLSTM 句子编码器与 GCN 依赖树编码器相结合, 实现了隐藏特征的自动提取, 在关系抽取任务中取得了显著效果。

1.4 联合学习在实体关系抽取中的应用

联合学习主要可分为两类: 一类为参数共享方法, 另一类为序列标注方法。其中, 前者采用参数共享的方式处理实体和关系, 而后者则通过序列标注方式处理实体及其关系。在实体识别与关系抽取领域研究者们提出了多种创新性方法。Miwa 等人^[8]提出了一种端到端神经网络模型, 该模型创新性地结合了双向 LSTM 和树型结构 (tree-LSTM) 模

型, 能够有效捕捉文本特征和依赖树结构。具体而言, 双向 LSTM 子网络负责执行实体识别任务, 而 tree-LSTM 子网络则专注于实体关系抽取, 这种模块化设计使得模型能够同时处理实体和关系信息。针对实体关系冗余问题, Zheng 等人^[9]提出了一种新型标注方法, 成功克服了传统参数共享方法的固有缺陷。该方法通过在神经网络模型中引入偏置损失函数, 显著提升了关系抽取的效果。具体来说, 偏置损失函数通过加强有效实体对的关系表示, 有效抑制了无效实体对的干扰, 从而实现了更精准的关系抽取。这一改进显著提升了模型的性能表现。

1.5 大语言模型 (LLM) 在实体关系抽取中的应用

通过预训练和微调语义信息, 大语言模型 (LLM) 的发展提升了实体关系抽取的准确性。大语言模型不仅拓展了 LLM 在自然语言处理中的潜力, 更为处理海量非结构化数据提供了新的解决方案。张西硕等^[10]的研究表明, GPT-4 等模型通过预训练获得的通用语言表征, 能够在无需特定领域标注数据的情况下识别实体间的隐含语义关系。其提出的 GPT-RE 模型通过将关系抽取任务重构为生成式问题, 在 FewRel 数据集上实现了接近传统监督模型 78% 的性能水平。此外, Wang 等人^[11]提出了一种基于双大语言模型协作的关系抽取方法, 该方法先提取短语, 再利用模型保存的语义相似度, 并通过迭代循环进行聚类, 从而显著提升了关系抽取的准确性。

2 结论与展望

近年来, 基于深度学习的实体关系抽取研究在模型架构创新方面取得了显著进展。研究已从早期依赖手工特征的统计模型, 逐步演进为具备自主语义学习能力的深度学习神经网络体系。具体而言, 循环神经网络 (RNN) 及其变体通过时序建模有效解决了序列标注问题; 卷积神经网络 (CNN) 实现了高效的特征提取; 基于参数共享的联合学习方法成功打破了传统流水线模式的性能瓶颈。预训练语言模型与图神经网络的深度融合进一步推动语义表征能力的质变, 尤其是在零样本场景下, 大语言模型 (LLM) 展现出了强大推理能力为低资源条件下的关系抽取开辟了全新研究路径。然而, 当前研究仍面临三大主要挑战: 首先, 模型对复杂语言现象的适应性不足, 嵌套实体、跨句指代及隐含关系等场景的识别准确率仍有较大提升空间; 其次, 各领域自适应能力存在明显局限, 现有多数模型在医学、法律等专业术语密集的领

域表现显著下降；最后，计算效率与可解释性之间的矛盾日益突出，特别是大语言模型的黑箱特性导致其难以在关键应用场景中部署。这些瓶颈问题凸显出当前研究在语言学理论指导、跨模态知识融合以及轻量化架构设计等方面仍存在明显空白。

参考文献：

[1] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursive matrix–vector spaces[C]//Proc of Empirical Methods in Natural Language Processing. 2012:1201–1211.

[2] NINGTHOUJAM D, YADAV S, BHATTACHARYYA P, et al. Relation extraction between the clinical entities based on the shortest dependency path based lstm[J]. arXiv:1903.09941, 2019.

[3] Liu Chunyang, Sun Wenbo, Chao Wenlan, et al. Convolution neural network for relation extraction[C]//Advanced Data Mining and Applications. 2013:231–242.

[4] ZENG D, LIU K, LAI S, et al. Relation classification via convolutional deep neural network[C]//Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers, 2014:2335–2344.

[5] NGUYEN T H, GRISHMAN R. Relation extraction: perspective from convolutional neural networks[C]//Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing, 2015:39–48.

[6] SONG L, ZHANG Y, GILDEA D, et al. Leveraging dependency forest for neural medical relation extraction[J]. arXiv:1911.04123, 2019.

[7] FU T J, LI P H, MA W Y. GraphRel: modeling text as relational graphs for joint entity and relation extraction[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019:1409–1418.

[8] Miwa M, Bansal M. End-to-end relation extraction using LSTMs on sequences and tree structures[C]//Proc of Meeting of the Association for Computational Linguistics. 2016:1105–1116.

[9] Zheng Suncong, Wang Feng, Bao Hongyun, et al. Joint extraction of entities and relations based on a novel tagging scheme[C]//Proc of Meeting of the Association for Computational Linguistics. 2017:1227–1236.

[10] 张西硕, 柳林, 王海龙, 等. 知识图谱中实体关系抽取方法研究[J]. 计算机科学与探索, 2024, 18(03):574–596.

[11] WANG J X, ZHANG L L, LEE W S, et al. When phrases meet probabilities: enabling open relation extraction with cooperating large language models[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics. Stroudsburg:ACL, 2024:13130–13147.

作者简介: 易战军(1977.08—),男,汉族,江西宜春人,硕士,主要研究方向为自然语言处理,双师型高级。

曾锐(1991.06—),男,汉族,江西万载人,硕士研究生,主要研究方向图像处理,助教。

基金: 江西省教育厅科技项目(GJJ180852)基于深度学习框架的实体关系抽取方法模型研究。