

基于深度学习的实体关系抽取研究综述

夏振涛^{1,3}, 曲维光¹, 顾彦慧¹, 周俊生¹, 李斌²

(1.南京师范大学 计算机科学与技术学院, 江苏省 南京市 210023;

2.南京师范大学 文学院, 江苏省 南京市 210097;

3.南京擎盾信息科技有限公司, 江苏省 南京市 210000)

摘要

作为信息抽取的一项核心子任务, 实体关系抽取对于知识图谱、智能问答、语义搜索等自然语言处理应用都十分重要。关系抽取在于从非结构化文本中自动地识别实体之间具有的某种语义关系。该文聚焦句子级别的关系抽取研究, 介绍用于关系抽取的主要数据集并对现有的技术作了阐述, 主要分为: 有监督的关系抽取、远程监督的关系抽取和实体关系联合抽取。我们对比用于该任务的各种模型, 分析它们的贡献与缺陷。最后介绍中文实体关系抽取的研究现状和方法。

关键词: 关系抽取; 有监督方法; 远程监督方法; 实体关系联合抽取

Review of Entity Relation Extraction based on deep learning

XIA zhentao^{1,3}, QU Weiguang¹,

GU Yanhui¹, ZHOU Junsheng¹, LI Bin²

(1.School of Computer Science and Technology, Nanjing Normal University,
Nanjing, Jiangsu 210023, China;

2.School of Chinese Language and

Literature, Nanjing Normal University, Nanjing, Jiangsu 210097, China;

3.Aegis Data, Nanjing, Jiangsu 210000, China)

Abstract

As a core task of Information Extraction, Entity Relation Extraction plays an important role in many Natural Language Processing applications such as knowledge graph, intelligent question answering system and semantic search. Relation extraction tasks aim to find the semantic relation between a pair of entity mentions from unstructured texts. This paper focuses on the sentence-level relation extraction, introduces the main datasets for this task, and expounds the current status of relation extraction technology which can be divided into: supervised relation extraction, distant supervision relation extraction and joint extraction of entities and relations. We compare the various models for this task and analyze their contributions and defects. Finally, the research status and methods of Chinese entity relation extraction are introduced.

Keywords: relation extraction, supervised method, distant supervision method, joint extraction of entities and relations

©2020 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

基金项目: 国家自然科学基金“汉语抽象意义表示关键技术研究”(61772278); 江苏省高校哲学社会科学基金“面向机器学习的汉语复句语料建设研究”(2019JSA0220); 国家社会科学基金“中文抽象语义库的构建及自动分析研究”(18BYY127)。

1 引言

信息抽取是自然语言处理的一项重要任务，它的基本目的之一是从原始的非结构化文本中提取有意义的结构化信息，以用于智能问答、检索等自然语言处理应用。信息抽取本身是一项庞大的任务，包括命名实体识别、关系抽取、事件抽取等子任务。本文，我们聚焦关系抽取任务。

本文将关系定义为两个或多个实体之间的某种联系，关系抽取就是自动识别实体之间具有的某种语义关系。本文所研究的是二元关系抽取，关注两个实体之间的语义关系，得到(subject:主语, relation:关系, object:宾语)三元组，其中subject和object表示两个实体，relation表示实体之间的语义关系。例如，给出如下的句子：刘翔，1983年7月13日出生于上海，中国男子田径队110米栏运动员。我们可以抽取出实体关系三元组实例(刘翔, 出生地, 上海)。

这里我们主要研究基于深度学习的关系抽取方法，包括：有监督关系抽取、远程监督关系抽取和实体关系的联合抽取。基于传统方法的关系抽取研究综述可参考文献 (Pawar et al., 2017)。此外我们还探讨中文关系抽取的研究现状、方法等。

本文的结构如下：第1节介绍用于关系抽取的数据集；第2,3,4节分别探讨有监督关系抽取、远程监督关系抽取和实体关系联合抽取相关方法研究及其优缺点；第5节探讨中文实体关系抽取的研究现状和方法；第6节对本文进行总结并展望关系抽取未来的研究方向。

2 数据集

2.1 有监督数据集

有监督的数据集需要人工标注，意味着数据包含高质量的实体关系三元组，几乎没有噪音。但是人工标注数据集耗时耗力，因此这些数据集通常规模较小。下面介绍的两个数据集中，每个句子中的命名实体进行了标记并且实体对表达的关系可以被预测。

ACE 2005数据集：自动内容抽取数据集(ACE)包含599个与新闻和电子邮件相关的文档，并划分为7种主要的关系类型。其中，6种主要关系类型包含足够的实例，每种关系类型平均有700个实例用于训练和测试。

SemEval-2010 Task 8 数据集：该数据集 (Hendrickx et al., 2009)包含10717个样本，其中8000个用于训练，2717个用于测试。数据集中包含9种有序关系类型。关系的方向性实际上使关系的数量加倍，因为只有在顺序也是正确的情况下，才认为实体对被正确标记。最终数据集有19种关系类别(其他类型为 $2*9+1$ 种)。

2.2 远程监督数据集

为了避免手工构建用于关系抽取数据集的繁重任务，远程监督方法 (Mintz et al., 2009)将文档与已知的知识库对齐，用于自动生成大量训练数据。这种方法基于距离监督假设：如果知识库中的实体对之间存在关系，那么每个包含该实体对的文档都会表达该关系。下面介绍的数据集是基于远程监督方法构造的。

New York Time(NYT)数据集：远程监督假设是一个强假设并且会导致错误标签问题，即提到两个实体的句子不一定表达他们在知识库中的关系。为了缓解这一问题，降低噪声，可以将问题建模为一个多示例学习问题，放宽远程监督假设 (Riedel et al., 2010)我们在后续章节详细介绍多示例学习方法。该数据集通过将Freebase与纽约时报语料库(NYT)的关系对齐而形成。数据集包含53种可能的关系类别，包括一个特殊的关系类别NA(表示实体之间没有关系)。训练数据包含522611个句子，281270个实体对和18252个关系。

3 有监督的关系抽取方法研究

有监督的关系抽取方法被认为是一个多分类问题，通常分为基于特征工程的方法、基于核函数的方法和基于神经网络的方法。基于特征工程的方法，许多研究的主要工作是如何能抽取复杂的特征。基于核函数的方法，直接以原始字符串作为处理对象，计算两个对象的核函数。利用神经网络可以自动从文本中获取表征关系的特征，完成关系抽取。本节中，我们主要介绍神经网络用于关系抽取的方法。

3.1 基于神经网络的方法

利用神经网络进行关系抽取, 主要分为三个步骤: 特征表示、神经网络的构建与学习和模型分类。其中, 特征表示是将文本表示成分布式特征信息, 比如将文本中的单词映射为分布式词向量 (Collobert et al., 2008; Mikolov et al., 2013(a); Pennington et al., 2014; Mikolov et al., 2013(b)), 这样就避免了传统特征工程中特征向量的稀疏性问题。我们把现有的研究方法总结为如下三类: 融合词汇特征的方法、基于最短依存路径的方法和基于注意力机制的方法。

3.1.1 融合词汇特征的神经网络模型

MV-RNN (Socher et al., 2012)是最早的利用神经网络进行关系抽取的方法,该方法沿着成分解析树建立递归神经网络学习树结构中单词或短语的语义向量表示。树中每个节点由矩阵-向量(MV)表示, 向量用来表示成分的内在含义, 矩阵用来捕获它如何改变相邻单词或短语的含义。对于关系分类任务, 该方法首先在解析树中找到要对其关系进行分类的两个实体之间的路径。然后选择路径的最高节点, 并使用该节点的向量作为特征对关系进行分类, 同时他们在模型中加入词性特征、命名实体标签特征和WordNet语义类别特征。该方法在SemEval-2010数据集上F1值达到82.4%。

MV-RNN包含特征学习过程, 这种方法依赖递归过程中使用的语法树, 句法分析中的错误会抑制学习高质量特征的能力。因此, 现有方法利用卷积神经网络(CNN) (Zeng et al., 2014; Kim, 2014; Collobert et al., 2011)提取词汇和句子级层次特征, 用于关系抽取任务。该方法不需要复杂的句法或语义预处理, 模型的输入是一个带有两个标记实体名词的句子。然后通过词嵌入将单词映射为分布式词向量, 分别提取词汇级特征和句子级特征, 并将两种特征拼接形成最终的特征向量。其中, 将带标记实体及其上下文词语所对应的词向量和WordNet中语义类别特征拼接作为词汇级特征向量,用最大池化卷积神经网络(CNN)自动提取句子级特征表示。更进一步, CR-CNN模型 (Santos et al., 2015)利用新的卷积神经网络处理关系分类任务,对于给定的输入文本, 模型将其转换成分布式表示, 然后通过卷积层构造句子的特征表示 r , 不同于Zeng使用softmax计算得分, 该方法模型通过类别权重矩阵和 r 执行点积操作得到每个关系类别的得分, 并且采用排名损失函数进行训练。该工作只使用词向量作为特征输入, 没有使用任何其他外部资源。

3.1.2 基于最短依存路径的神经网络模型

上述方法在关系抽取任务是有效的, 但是, 当主语和宾语之间的距离较长时, 往往会受到其他不相关信息的干扰。考虑到句法特征在关系识别中起到至关重要的作用, 许多研究通过神经网络从最短依存路径中学习关系表示。该方法基于最短路径假设: 如果 e_1 和 e_2 是句子中的两个名词实体, 我们假设 e_1 和 e_2 之间的最短路径, 描述了它们之间的关系。这是因为,(1)如果 e_1 和 e_2 是属于同一个谓词的论元, 那么它们的最短路径应该通过该谓词; (2)如果 e_1 和 e_2 属于不同的谓词论元结构, 它们的最短路径将通过一系列谓词, 并且任何连续的谓词都共享一个公共论元。路径上谓词的顺序表明了该关系的主语和宾语的正确分配。

基于最短路径假设, 通过卷积神经网络可以从最短依存路径学习关系表示 (Xu et al., 2015(a)),该模型采用主语到宾语的最短路径作为输入, 依存路径上的每个节点通过词嵌入矩阵生成局部特征向量, 然后将这些特征向量组合, 用CNN网络获取全局特征向量。同时, 为了处理关系的方向性(即在关系中正确区分主语和宾语), 该方法在模型中引入负采样策略, 采用宾语到主语的最短依存路径为负样本, 研究表明负采样策略显著提升了模型的性能。

同年, 基于SDP-LSTM神经网络方法 (Xu et al., 2015(b))用于关系分类, 该方法利用两个实体之间的最短路径(SDP), 以LSTM为单元的多通道(四种信息: 单词、词性、语法关系、WordNet语义特征)循环神经网络沿SDP提取多样化特征, 最大池化层从每条路径中的LSTM节点收集信息; 为了缓解过拟合, 模型采用了dropout正则化策略。

然而, 最短依存路径(SDP)是一种特殊的结构, 在这个结构中, 每两个相邻单词由依存关系分隔开。由于卷积神经网络和循环神经网络的局限性, 以往的研究将依存关系等同于单词或词性等句法特征。循环卷积神经网络(RCNN) (Cai et al., 2016)利用基于LSTM的双通道循环神经网络对SDP中的全局模式进行编码, 并利用卷积层捕获依存关系中每两个相邻单词的局部特征。基于RCNN模型, 他们提出一种双向循环卷积神经网络(BRCNN), 可以同时学习沿着SDP前向和后向的双向信息表示, 增强了实体间关系方向分类的能力。

神经网络模型直接在解析树上操作很难并行化，计算效率很低。基于主语和宾语之间的最短路模型虽然在计算上更高效，但这种假设有很大的局限性，会丢掉一些关键信息。受到图卷积神经网络 (Marcheggiani et al., 2017; Kipf et al., 2016) 的启发，一种适用于关系抽取的图卷积神经网络的扩展方法 (Zhang et al., 2018) 利用高效的图卷积运算对输入语句的依存结构进行编码，然后提取以实体为中心的表示，实现关系预测。此外基于注意力引导的图卷积网络方法 (AGGCNs) (Guo et al., 2019) 提出用于关系抽取。该模型由M个相同的块组成，每个块包含注意力指导层、密集连接层和线性组合层。模型直接把整棵依存树作为输入，每个块以表示图的节点嵌入和邻接矩阵作为输入，在注意力引导层中，先将原始的图转化为邻接矩阵，然后通过多头注意力转化为全连接的基于注意力指导的邻接矩阵，矩阵中的每个元素对应相应节点之间边的权重，从而捕捉到领域的信息。密集连接层得到的矩阵被送入N个单独的密集连接层，产生新的表示。最后，应用线性组合将N个紧密连接的层的输出组合成隐藏的表示。经过上述注意力引导的图卷积模型，得到所有tokens的表征，然后将句子的表征和实体表征合并，运用前馈神经网络得到最终表示，最后用逻辑回归分类器预测关系。

3.1.3 基于注意力机制的神经网络模型

通过神经网络对依存树建模能提高关系抽取的性能，但这类方法还是需要依赖词汇资源如WordNet，或自然语言处理工具来提取特征。并且，对于关系重要的信息可以出现在句子的任意位置。为了解决这个问题，基于注意力机制的双向长短期记忆网络 (AttBLSTM) (Zhou et al., 2016) 用来捕获句子中重要的语义信息。该方法的注意力机制计算公式为：

$$M = \tanh(H) \quad (1)$$

$$\alpha = \text{softmax}(W^T M) \quad (2)$$

$$r = H\alpha^T \quad (3)$$

其中，H为双向LSTM输出层组成的矩阵，w是学习的参数，r是输出向量的加权求和。最终的句子可以表示为

$$h^* = \tanh(r) \quad (4)$$

模型用一个softmax分类器得到关系类别。实验表明，该方法不使用任何外部资源，可以得到很好的性能。

现有的基于注意力机制的关系抽取方法并没有充分利用实体信息，而实体信息可能是关系分类的最关键特征。因此，一种融合潜在实体类型的实体注意力机制模型 (Lee et al., 2019) 用于关系抽取。在该方法中，为了捕获句子上下文信息，利用了自注意力机制 (self-attention) (Tan et al., 2018; Vaswani et al., 2017) 获得单词表示，并利用双向LSTM构建了神经网络模型。为了充分利用句子中实体对的信息，模型融合了实体相对位置特征和实体潜在种类特征，最终句子的表示可以通过注意力机制得到：

$$u_i = \tanh(W^H [h_i, p_i^{e_1}, p_i^{e_2}] + W^E [h_{e_1}, t_1, h_{e_2}, t_2]) \quad (5)$$

$$\alpha_i = \frac{\exp(v^T u_i)}{\sum_{j=1}^n \exp(v^T u_j)} \quad (6)$$

$$z = \sum_{i=1}^n \alpha_i h_i \quad (7)$$

其中， h_i 为第i位置双向LSTM的输出， $h_{e_1}; h_{e_2}$ 分别为实体 $e_1; e_2$ 位置双向LSTM的输出， $p_i^{e_1}; p_i^{e_2}$ 分别为对应于第i个单词相对于句子中的第一个实体 e_1 和第二个实体 e_2 的位置。该方法还使用实体类型信息提升性能，用于实体类型没有标注，他们使用主题聚类方法得到实体的潜在种类，通过注意力机制得到实体类型表示t。

3.2 小结

本章详细讨论了有监督方法的关系抽取，给出了现有的一些经典方法。目前，神经网络用于有监督的关系抽取成为主流方法，我们详细讨论了，神经网络方法中融合词汇特征的方法、基于最短依存路径的方法和基于注意力机制的方法。表1给出现有方法用于有监督关系抽取在SemEval-2010数据集上的性能对比。从图中可以看出，基于最短依存路径的方法总体上达到了较好的性能，原因在于利用最短依存路径对句子建模，可以丰富句子的全局语义信息，显著帮助关系分类。而融合词汇的方法，只能利用句子的局部特征，关系分类性能较低。基于注意力机制的方法不采用任何外部特征资源，利用注意力机制的可解释性自动挖掘出句子中的重要语义信息。

方法分类	模型	特征	F1值
融合词汇特征方法	MV-RNN (Socher et al., 2012)	Word embeddings+POS,NER, WordNet	82.4%
	CNN (Zeng et al., 2014)	Word embeddings + word position embeddings, WordNet	82.7%
	CR-CNN (Santos et al., 2015)	Word embeddings + word position embeddings	84.1%
基于最短依存路径方法	depLCNN (Xu et al., 2015(a))	WordNet, words around nominals	85.6%
	SDP-LSTM (Xu et al., 2015(b))	Word Embeddings, POS embeddings, WordNet embeddings, grammar relation embeddings	83.7%
	BRCNN (Cai et al., 2016)	Word embeddings+POS, NER, WordNet embeddings	86.3%
	C-GCN (Zhang et al., 2018)	Word embeddings	84.8%
	AGGCNs (Guo et al., 2019)	Word embeddings	85.7%
基于注意力机制方法	AttBLSTM (Zhou et al., 2016)	Word embeddings, position embeddings	84.0%
	BiLSTM with Entity-aware attention (Lee et al., 2019)	Word emneddings, Latent entity Typing	85.2%

Table 1: 有监督关系抽取方法在数据集SemEval-2010对比

4 基于远程监督的关系抽取方法研究

有监督的关系抽取需要依赖人工标注的数据集，限制了该方法的适用领域。因此，远程监督方法 (Mintz et al., 2009)将文档与已知的知识库对齐，用于自动生成大量训练数据。然而，远程监督假设是一个强假设并且会导致错误标签问题，即提到两个实体的句子不一定表达他们在知识库中的关系。因此可以将远程监督关系抽取任务作为一个多示例学习问题来放宽假设 (Riedel et al., 2010)。在用于关系抽取的多示例学习中，知识库(KB)中的每个实体对标记一个句子包。包中的所有句子都包含实体对的提及，但它们不一定包含直接关系。多示例学习是对包标签预测，而不是为每个句子预测关系标签。它假定，如果实体对存在关系，则包中至少有一个示例反映给定实体对的关系。

4.1 基于卷积神经网络的远程监督方法

Riedel的方法抽取文本特征时，需要依赖自然语言处理工具，会造成错误传播问题。因此，分段卷积网络(PCNN) (Zeng et al., 2015)用来提取特征，并利用多示例学习方法缓解数据

噪音问题。在多示例训练中把目标函数定义在包上，首先对包中的每个示例分别预测，得到相应的关系概率，然后选取概率最大的示例标签作为包的标签，并利用包的标签更新网络参数。

PCNN在远程监督数据集上得到了不错的效果，但这种方法仍然有缺陷。首先，PCNN将远程监督关系抽取看作一个单标签学习问题，并为每个实体对选择一个关系标签，而忽略了同一个实体对可能存在多个关系的事实。针对这个问题，利用两个不同的损失函数，(Jiang et al., 2016)处理多标签分类问题。此外，PCNN基于Riedel提出的假设来生成标记数据，根据这一假设，PCNN在训练和预测中只选择每个实体对可能的句子。然而，选择一个句子会丢失包含其他句子中的信息。对于这个问题，假设“两个实体间的关系可以自动从提到这两个实体的所有句子中显示表达或隐式推断”(Jiang et al., 2016)，在使用卷积神经网络自动提取每个句子的特征后，他们使用跨句最大池化来选择不同句子的特征，然后将最重要的特征聚合为每个实体对的表示。由于结果表示有不同句子的特征组成，因此该方法充分利用这些句子中包含的所有可用信息。此外，利用句子级别的注意力机制(Lin et al., 2016)来自动捕获不同句子的重要程度，过滤噪声句子。

上述方法使用的神经网络，模型大多是相对较浅的卷积神经网络，通常只涉及一个卷积层和一个全连接层，而且不清楚更深的模型结构是否能够从噪声数据中提取信号。一种基于残差学习的卷积神经网络。(Huang et al., 2017)用于关系抽取，他们将词嵌入和位置嵌入合并到一个深度残差网络中，通过恒等映射到卷积层中。实验表明，该方法利用9层带残差学习的卷积神经网络可以显著提升远程监督关系抽取性能。

4.2 基于注意力机制的远程监督方法

现有方法在选择有效示例和缺乏实体背景知识方面存在缺陷，一个基于PCNN的句子级注意力机制模型(APCNNs)(Ji et al., 2017)用来选择有效示例，该模型充分利用了知识库中的监督信息。他们从Freebase和Wikipedia页面中提取实体描述来补充背景知识。对于一个包，模型首先使用PCNNs提取每个句子的特征向量 v_{sen} 。受到TransE模型的启发，在TransE模型中，用 $e_1 + r \approx e_2$ 对一个三元组 $r(e_1, e_2)$ 建模，在APCNNs中，用 $(e_1 - e_2)$ 表示句子中 e_1 和 e_2 之间的关系。然后，模型通过一个隐含层用串联 $[v_{sen}; e_1 - e_2]$ 的方式计算每个句子的注意力权重。最后，所有句子特征向量的加权求和就是包的特征。此外，为了将更多的背景知识融入到模型中，该方法使用卷积神经网络来提取实体描述的特征向量。

使用注意力机制，是一种通过学习多个示例的权值分布来选择有效示例的方法。但是，基于深度神经网络的远程监督学习中存在两个重要的表示学习问题：(1)在一个示例中，目标实体对上下文表示学习问题；(2)多个示例的有效示例选择表示学习。在先前的研究工作中，通常采用1-D向量的单词级和句子级注意力机制。1-D注意力向量的缺陷是，它只关注句子中一个或少量的方面，或一个或少量的示例。其结果是不同语义方面的句子，或者不同的多个有效句子被忽略。受结构化自注意力句子嵌入(Lin et al., 2017)的启发，一种新的基于双向LSTM的多层结构化自注意力机制模型(MLSSA)(Du et al., 2018)用于缓解上述两个问题。针对第一个问题，他们提出一个基于二维矩阵的单词级注意力机制，该机制包含多个向量，每个向量都聚焦于句子的不同方面，从而更好地学习上下文表示。针对第二个问题，他们提出一种用于多示例学习的二维句子级注意力机制，其中包含多个向量，每个向量都集中在不同的有效示例上，以更好地选择句子。

4.3 融合知识库的方法

为了缓解远程监督中错误标注问题，许多研究利用现有知识库添加信息。首先，一种无标签的远程监督方法(Wang et al., 2018)在距离假设不充分的条件下，不使用关系标签，只利用知识库(KG)的先验知识直接、柔和的监督分类器的学习。除了关系示例外，知识库中还包括其他相关信息，比如关系的别名，现有的关系抽取模型通常忽略这些可用的信息。一种远程监督关系抽取方法—RESIDE(Vashishth et al., 2018)利用知识库中附加的边信息改进关系抽取。具体的，它使用实体类型和关系别名信息在预测关系时施加软约束，使用图卷积神经网络从文本中编码语法信息，即使在有限额外信息可用时也能提高性能。

远程监督可以自动标注足够数量的训练数据；然而，这些数据通常只覆盖关系的有限部分。许多关系都是长尾关系，数据仍然不足。目前的远程监督模型忽略了长尾关系问题，难以从纯文本中提取出全面的信息。受在尾部的数据和在顶部的数据之间丰富的语义关联的启发，一种用于长尾不平衡数据的远程监督关系提取方法(Zhang et al., 2019)利用分布顶部数据丰富的类

的知识来提高尾部数据贫乏类的性能。首先，他们提出利用知识图嵌入的类标签间的隐式关系知识，利用图卷积网络学习显式关系知识。其次，通过粗到细的知识感知注意机制，将关联知识集成到关联抽取模型中。

4.4 小结

在本节中，我们详细讨论了一些用于远程监督关系抽取的经典方法，下面表2给出这些方法在NYT数据集上抽取关系示例前100(TOP-100)、前200(Top-200)、前300(TOP-300)前500(Top-500)的对比结果,这里使用Precision@N(P@N)为评估指标。远程监督关系抽取只需要手动标注少量的关系实例，适用于没有标注语料库的关系抽取，但其实现过程在数据集中引入了噪声，使得该方法的性能低于有监督的关系抽取方法。许多后续的工作都试图利用选择性注意力机制、融合知识库等方法来处理噪声和放宽远程监督假设，通过去噪进一步提高性能。远程监督关系抽取方法利用的是弱标注数据，一般的神经网络方法都是以数据驱动模型，但纯数据驱动模型并不能充分挖掘数据中的潜在信息，从表中我们可以发现，融合知识库的方法相对有着较好的性能，这类方法能够更好地将结构化知识融入神经网络模型中，用知识指导模型。融合知识库的方法，不仅仅是关系抽取任务，在其他自然语言处理任务中也具有重要意义。

方法分类	模型	Top-100	Top-200	Top-300	Top-500
基于卷积神经网络方法	PCNN (Zeng et al., 2015)	72.3	69.7	64.1	-
	PCNN+MIL (Zeng et al., 2015)	86.0	80.0	-	69.0
	PCNN+ATT (Lin et al., 2016)	76.2	73.1	67.4	-
	MIMLCNN (Jiang et al., 2016)	69.0	64.0	59.0	53.0
	ResCNN-9 (Huang et al., 2017)	79.0	69.0	61.0	-
基于注意力机制的方法	APCNN+D (Ji et al., 2017)	87.0	83.0	-	74.0
	MLSSA (Du et al., 2018)	90.0	81.5	77.0	-
融合知识库的方法	LFDS (Wang et al., 2018)	90.0	88.0	-	83.0
	RESIDE (Vashishth et al., 2018)	84.0	78.5	75.6	-

Table 2: 远程监督关系抽取方法在数据集NYT数据集对比

5 实体和关系联合抽取方法研究

实体和关系联合抽取在于从非结构化文本中同时进行实体识别和关系抽取。传统方法以流水线方式处理抽取实体关系三元组任务，即首先提取实体，然后识别他们之间的关系。这个独立的框架使任务易于处理，并且每个组件都可以更加灵活。但它忽略了这两个子任务之间的相关性，在这种方式下，每个子任务都是独立的模型。这样，实体识别的结果可能会影响关系分类的性能，导致错误传递。与流水线方法不同，联合学习框架能利用单个模型提取实体和关系，能够有效地集成实体和关系的信息。

5.1 基于共享参数的联合抽取方法

最早的联合框架模型 (Li et al., 2014)利用结构化感知机和集束搜索方式同时提取实体及其关系。该框架采用了一种基于半马尔科夫链思想的分段解码器，克服了传统的基于字符的标注方式。此外，考虑到不精确的搜索，他们提出一些新的有效的全局特征作为约束来捕捉实体和关系之间的相互依赖性。

在有监督的关系分类任务中，可以用基于LSTM的神经网络来表示实体之间的关系，但這些方法只使用有限的语言学结构，并且不能对实体和关系联合建模。因此，一种端到端的模型 (Miwa et al., 2016) 用于提取词序列和依存树结构上实体之间的关系。该方法通过使用双向顺序(从左到右和从右到左)和双向树形结(自底向上和自顶向下)的LSTM-RNNs在单个模型中对实体和关系联合建模。与传统的增量端到端关系抽取模型不同，该模型在训练中做了两个改进：实体预训练(对实体模型进行预训练)和scheduled sampling(以一定的概率用正确的标签替换不可靠的预测标签)，这些改进缓解了早期训练阶段实体识别性能低的问题，并且允许实体信息进一步帮助下游关系分类。Miwa提出的模型是局部训练的，没有考虑到增量决策之间的结构性对应。因此，一个全局优化的神经网络模型 (Zhang et al., 2017) 用于端到端关系抽取，为了更好地学习上下文表示，提出了新的LSTM特征。Miwa的方法依靠外部句法解析器获取句法信息，这对于关系抽取至关重要，但解析错误可能导致树LSTM的编码不准确，从而降低关系抽取性能。在Zhang的方法中，使用双仿射注意力解析器 (Dozat et al., 2016) 的LSTM隐藏层来增强输入表示。由于解析器是预训练的，它包含了关于每个单词的丰富语法信息，但不显式地表示解析决策，从而避免了由于不正确的解析而导致的问题。此外，Miwa的方法在预测实体边界或做出关系分类决策时没有明确学习片段的表示，在Zhang的方法中，采用LSTM-Minus方式，将一个片段建模为最后一个和第一个LSTM隐藏层向量之差。

为了不使用任何依存树信息，Katiyar等人 (Katiyar et al., 2017; Zheng et al., 2017) 用多层双向LSTM对句子建模，对实体识别和关系抽取都看作是序列化标注任务。在关系抽取中，对于每个词，使用指针网络找到当前词和相关之前词的关系类型。同样，Zheng把实体和关系抽取任务看作是序列标注任务，与Katiyar方法不同，他们提出一种新的序列化标注方式，用双向LSTM和单向LSTM分别编码和解码，输出层同时对实体和关系标注，完成关系抽取任务。

5.2 基于全局优化的联合抽取方法

上述联合学习的方法是通过共享参数的方式实现，这样的好处是不需要在两个子任务上附加约束。但是由于独立的子模型解码器，子模型之间的联系没有得到充分利用。一个联合最小化风险训练的方法用于实体和关系联合抽取 (Sun et al., 2018)。在这个模型中，把实体识别作为序列化标注任务，关系抽取作为分类任务。两个任务的模型之间共享参数，并且优化一个全局损失函数，弥补了训练和测试之间的差异。通过共享参数联合学习，一方面，在实体类型和关系类型判定时没有显示交互，一些复杂的解码算法可以同时判断实体边界和类型，但是通过在ACE05数据集上的实验发现 (Sun et al., 2019)，边界的识别正确率很高，相对的类型判定就低一些。因此，将实体关系联合抽取分为两个子任务，分别是实体范围检测 (Entity Span Detection) 和实体关系类型推导 (Entity Relation Type Deduction)。在实体范围检测中使用序列标注的方法，在实体关系类型推导中使用一种基于图卷积网络的联合模型，同时两个模型进行联合训练。

在句子中，实体三元组会存在重叠问题，Zeng等人 (Zeng et al., 2018) 首先在神经网络建模中利用拷贝机制 (Gu et al., 2016; He et al., 2017) 解决重叠问题。他们把重叠问题分为三种类型：Normal, EntityPairOverlap(SEO), SingleEntityOverlap(SEO)。他们的方法基于sequence-to-sequence模型 (Dong et al., 2016)，在模型中有两个主要部分：编码器(endcoder)和解码器(decoder)。编码器首先把句子转换成固定长度的向量。然后，解码器读取这个向量并生成三元组。

实验表明Zeng的方法对训练数据依赖性强，不能提取多个单词实体的情况。一个层次化的强化学习框架 (Takanobu et al., 2019) 通过一个高层强化学习过程识别关系指示词，用低层的强化学习过程识别实体。高层的过程在某个特定位置检测关系指示词，如果确定了某个关系，将触发低层过程识别该关系对应的实体。当低层任务完成后，高层强化学习过程继续搜索句子中下一个关系。Li等人 (Li et al., 2019) 将实体关系抽取任务转换为多轮问答问题，即将实体和关系的提取转换为从上下文确定答案的任务。这种方法提供了一个比较好的捕捉标签层次依赖的方法。但这中间方法计算效率低，因为它需要在单个句子中扫描所有实体模板问题和相关的关系模板问题。

此外一种用于实体识别和关系提取的端到端联合模型 ((GraphRel)(Fu et al., 2019) 通过关系加权GCN来考虑命名实体和关系之间的交互来解决实体重叠问题。GraphRel通过堆叠Bi-LSTM句子编码器和GCN依赖树编码器来学习自动提取每个单词的隐藏特征。然后GraphRel标记实体提及单词并预测连接提及的关系三元组，这是第一阶段预测。为了在考虑到三元组之间

的相互作用的情况下进行预测，该模型在GraphRel第二阶段添加了一个新颖的关系加权GCN。第一阶段GraphRel接收到实体损失和关系损失，沿着依赖关系链接提取节点隐藏特征，同时建立具有关系加权边的新全连接图。然后，通过对中间图进行操作，第二阶段GCN在最终分类每个边之前有效地考虑实体之间的相互作用以及（可能重叠的）关系。

5.3 小结

在本节中，我们详细讨论了一些用于实体关系联合抽取的经典方法，表3给出实体和关系联合抽取方法在不同数据集上的结果对比。许多较早的实体关系抽取都采用了流水线框架，流水线框架具有集成不同数据源和学习算法的灵活性，但其缺点也很明显。首先，他们受到错误传播的严重影响，实体提取阶段的错误会传播到关系分类阶段。其次忽略了实体提取和关系分类的相关性。第三，流水线框架导致计算效率低下。在实体提取阶段后，将每个实体对传递到关系分类模型，以识别它们之间的关系。由于大多数实体对没有关系，这中两阶段的方式是低效的。在本节中讨论的联合建模技术采用了实体识别和关系识别任务之间的双向信息流，很好的解决了流水线方式的缺陷。从实际的角度来看，联合抽取方法非常重要，因为良好的实体抽取性能是实现良好关系抽取性能的必要条件。

datasets	model	Entity			Relation		
		P	R	F1	P	R	F1
ACE04	(Li et al., 2014)	0.835	0.762	0.797	0.608	0.361	0.453
	(Miwa et al., 2016)	0.808	0.829	0.818	0.487	0.481	0.484
	(Katiyar et al., 2017)	0.812	0.781	0.796	0.502	0.488	0.493
	(Li et al., 2019)	0.844	0.829	0.836	0.501	0.487	0.494
ACE05	(Li et al., 2014)	0.852	0.769	0.808	0.654	0.398	0.495
	(Miwa et al., 2016)	0.852	0.769	0.808	0.572	0.540	0.556
	(Zhang et al., 2017)	-	-	0.836	-	-	0.575
	(Katiyar et al., 2017)	0.840	0.813	0.831	0.605	0.553	0.578
	(Sun et al., 2018)	0.839	0.832	0.836	0.649	0.551	0.596
	(Sun et al., 2019)	0.861	0.824	0.842	0.681	0.523	0.591
CONLL04	(Zhang et al., 2017)	-	-	0.856	-	-	0.678
NYT	(Zheng et al., 2017)	0.59	0.479	0.529	0.597	0.451	0.514
	(Sun et al., 2018)	-	-	-	0.652	0.406	0.500
	(Zeng et al., 2018)	-	-	-	0.610	0.566	0.587
	(Takanobu et al., 2019)	-	-	-	0.714	0.586	0.644
	(Fu et al., 2019)	-	-	-	0.639	0.600	0.619

Table 3: 实体关系联合抽取在不同数据集的结果

6 中文实体关系抽取研究现状

在中文研究方面，由于标注语料的短缺，关系抽取的研究相对于英文数据集上的研究较少。本文聚焦神经网络的中文实体关系抽取研究，传统方法可参考(武文雅 et al., 2018)的工作，远程监督方法可参考(白龙 et al., 2019)的工作。

6.1 中文实体关系抽取数据集

COAE2016: 该数据集来源于第八届中文倾向性分析评测(COAE2016)的面向知识抽取的关系分类任务。该数据集包含988句训练集、483句测试集，以及10种关系类型(人物的出生日期, 人物的出生地, 人物的毕业院校, 人物的配偶, 人物的子女, 组织机构的高管, 组织机构的员工数, 组织机构的创始人, 组织机构的成立时间, 组织机构的总部地点)

ACE2005: ACE 2005数据集收集自新闻专线、广播和网络日志。关系分为6大类和18个小类，包含8023个关系事实和18个关系子类型。

DuIE: 2019年，中国计算机学会、中国中文信息学会联合百度公司举办的语言与智能技术竞赛开放了基于百度百科和百度信息流的大规模中文信息抽取数据集(Li et al., 2019)。该数据集

包含49个关系种类数和458184个关系实例数。

6.2 基于神经网络方法研究现状

为了在中文语料中获得更丰富的高级特征，高层语义注意力机制的分段卷积神经网络模型(武文雅 et al., 2019)用于中文关系抽取。在模型的向量表示中，添加了HowNet中的上位词向量特征。该方法在ACE2005数据集上的实验结果F1值达到73.94%，COAE2016数据集上F1值达到78.41%。由于中文句式和语法结构复杂，并且汉语有更多歧义，会影响中文实体关系分类的效果，一种基于多特征自注意力的实体关系抽取方法(李卫疆 et al., 2019)，充分考虑词汇、句法、语义和位置特征，使用基于自注意力的双向长短期记忆网络来进行关系预测。

此外，利用多粒度信息和外部知识进行中文关系抽取(Li et al., 2019)。具体的，多粒度信息主要包含三部分：字向量、词向量和词义向量。除了会用到字向量和词向量外，还会用到词义向量。使用HowNet作为外部知识库，对于给定词，通过检索HowNet可以获得词的所有词义信息。然后通过SAT模型(Niu et al., 2017)将每个词义转换为实值向量。将字向量、词向量和词义向量通过Lattice LSTM编码层，最后经过关系分类层，将编码层得出的隐藏层状态作为输入，经过注意力计算，进行关系分类。通过在不同领域的三个数据集(Chinese SanWen, ACE 2005 Chinese corpus and FinRE)上进行的实验表明，他们的模型具有显著的优越性，在ACE2005数据集上F1达到78.71%。

基于DuIE中文信息抽取数据集，一个端到端的框架(Liu et al., 2019)用于关系抽取，该框架首先在具有关系提及层的原始文本中捕获关系提及，然后进行实体标记，其目的是使用给定的关系提及对相应的三元组实体进行解码。此方法在验证集下的f1值达到84.8%。

6.3 小结

目前基于神经网络的中文实体关系抽取在公共数据集上的研究还较少，现有方法为了提取中文的丰富语义特征，都融合了知识库资源。2019年，中国计算机学会、中国中文信息学会联合百度公司举办的语言与智能技术竞赛开放了基于百度百科和百度信息流的大规模中文信息抽取数据集DuIE，有利推动了中文关系抽取研究的发展，中文关系抽取方法将不断涌现、性能不断提高。表3给出中文数据集上的实验结果对比。

数据集	方法	F1值
COAE2016	(武文雅 et al., 2019)	78.41%
	(李卫疆 et al., 2019)	81.49%
ACE2005	(武文雅 et al., 2019)	73.94%
	(Li et al., 2019)	78.71%
DuIE	(Liu et al., 2019)	84.8%

Table 4: 关系抽取在中文数据集上的结果

7 总结与展望

关系抽取作为信息抽取不可或缺的部分，是知识图谱、文本内容理解的重要支撑技术之一。根据领域的划分，可分为限定域关系抽取和开放域关系抽取。本文详细讨论了限定域关系抽取的三大类方法：有监督方法、远程监督方法和实体关系联合抽取方法。根据本文的论述，前沿的关系抽取技术在英文数据集ACE2004、ACE2005、SemEval-2010和NYT-10做了许多工作，在中文数据集上相对较少。

本文通过对现有关系抽取研究方法的总结，提出以下关系抽取未来的研究路线：

(1)前沿的关系抽取技术在主流英文数据集ACE2004、ACE2005、SemEval-2010和NYT-10做了许多工作。NYT-10数据集是自动构建的，通过将Freebase知识库与纽约时报语料库(NYT)的关系对齐而形成，此数据集没有手动注释，存在着数据噪声的问题。SemEval-2010数据集通过引入手动注释达到了相对较高的质量，但数据规模依然太小。未来工作可以开发出高质量的基于中文的关系抽取数据集，并且不断提升关系抽取技术在中文数据集上的性能。

(2)由于大量的关系事实都是通过多个句子来表达，句子级的关系抽取受到了不可避免的限制，因此，未来关系抽取的研究方向会从句子级推广到篇章级，通过读取和推理一个文档中的多个句子，能够有效提升关系抽取性能。

(3)关系抽取是一项复杂的任务，无论是有监督的数据，还是远程监督数据，纯数据驱动模型是远远不够的。如何从现有数据中挖掘和学习有用信息，以及如何将结构化知识、语言知识、领域知识融合进关系抽取模型中，这是两个重要的课题对于语言理解有着重要意义。

(4)多模态学习在关系抽取任务中的应用和研究。在互联网上存在着多种形式的的数据，如自然语言、图片、结构化文本，这每一种数据形式可以称为一种模态。在构造数据集时，可以融入这些多模态数据。通过利用多模态之间的互补性，剔除模态间的冗余性，从而可以学习到更好的特征表示。

参考文献

- Kurt D. Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge and Jamie Taylor. 2008. Freebase: A collaboratively created graph database for structuring human knowledge. *International Conference on Management of Data*, pages 1247-1250.
- 白龙, 靳小龙, 席鹏弼. 2019. 基于远程监督的关系抽取研究综述. *中文信息学报*, 33(10):10-17.
- Rui Cai, Xiaodong Zhang and Houfeng Wang. 2016. Bidirectional recurrent convolutional neural network for relation classification. *The 54th Annual Meeting of the Association for Computational Linguistics*, pages 756-765.
- Ronan Collobert, Jason Weston, Michael Karlen, Koray Kavukcuoglu and Pavel P. Kuksa. 2011. language processing (almost) from scratch. *Journal of Machine Learning Research*, pages 2493-2537.
- Ronan Collobert and Jason Weston. 2008. A Unified Architecture for Natural Language Processing. *The Twenty-Fifth International Conference on Machine Learning*, pages 160-167.
- Li Dong and Mirella Lapata. 2016. Language to logical form with neural attention. *The 54th Annual Meeting of the Association for Computational Linguistics*, pages 7-12.
- Timothy Dozat and Christopher D. Manning. 2016. Deep biaffine attention for neural dependency parsing. *The 5th International Conference on Learning Representations*, pages 24-26.
- Jinhua Du, Jingguang Han, Andy Way and Dadong Wan. 2018. Multi-level structured self-attentions for distantly supervised relation extraction. *The 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2216-2225.
- Tsu-Jui Fu, Peng-Hsuan Li and Wei-Yun Ma. 2019. GraphRel: Modeling Text as Relational Graphs for Joint Entity and Relation Extraction. *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 1409-1418.
- Jiatao Gu, Zhengdong Lu, Hang Li and Victor O. K. Li. 2016. GraphRel: Modeling Text as Relational Graphs for Joint Entity and Relation Extraction. *the 55th Annual Meeting of the Association for Computational Linguistics*, pages 199-208.
- Zhijiang Guo, Yan Zhang and Wei Lu. 2019. Attention Guided Graph Convolutional Networks for Relation Extraction. *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 241-251.
- Shizhu He, Cao Liu, Kang Liu and Jun Zhao. 2017. Generating natural answers by incorporating copying and retrieving mechanisms in sequence-to-sequence learning. *The 55th Annual Meeting of the Association for Computational Linguistics*, pages 199-208.
- Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Sebastian Padó, Marco Pennacchiotti, Lorenza Romano, Stan Szpakowicz. 2009. SemEval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. *The 5th International Workshop on Semantic Evaluation*, pages 15-16.
- Yi Yao Huang and William Yang Wang. 2017. Deep residual learning for weakly-supervised relation extraction. *The 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1803-1807.

- Guoliang Ji, Kang Liu, Shizhu He and Jun Zhao. 2017. Distant supervision for relation extraction with sentence-level attention and entity descriptions. *The 31th AAAI Conference on Artificial Intelligence*, pages 3060-3066.
- Xiaotian Jiang, Quan Wang, Peng Li and Bin Wang. 2016. Relation extraction with multi-instance multi-label convolutional neural networks. *The 26th International Conference on Computational Linguistics: Technical Papers*, pages 1471-1480.
- Kambhatla N. 2016. Combining lexical, syntactic and semantic features with Maximum Entropy models for extracting relations. *The Meeting on Association for Computational Linguistics*, pages 22-28.
- Arzoo Katiyar and Claire Cardie. 2017. Going out on a limb: Joint extraction of entity mentions and relations without dependency trees. *The 55th Annual Meeting of the Association for Computational Linguistics*, pages 917-928.
- Yoon Kim. 2014. Convolutional neural networks for sentence classification. *The 2014 Conference on Empirical Methods in Natural Language Processing*, pages 25-29.
- Thomas N. Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *The 5th International Conference on Learning Representations*, pages 24-26.
- Jooheon Lee, Sangwoo Seo and Yong Suk Choi. 2019. Semantic Relation Classification via Bidirectional LSTM Networks with Entity-Aware Attention Using Latent Entity Typing. *Symmetry*, pages 785.
- Qi Li and Heng Ji. 2014. Incremental joint extraction of entity mentions and relations. *The 52nd Annual Meeting of the Association for Computational Linguistics*, pages 402-412.
- Shuangjie Li, Wei He, Yabing Shi, Wenbin Jiang, Haijin Liang, Ye Jiang, Yang Zhang, Yajuan Lyu and Yong Zhu. 2019. DuIE: A Large-Scale Chinese Dataset for Information Extraction. *The CCF International Conference on Natural Language Processing and Chinese Computing*, pages 791-800.
- Xiaoya Li, Fan Yin, Zijun Sun, Xiayu Li, Arianna Yuan, Duo Chai, Mingxin Zhou and Jiwei Li. 2019. Entity-Relation Extraction as Multi-Turn Question Answering. *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 1340-1350.
- Ziran Li, Ning Ding, Zhiyuan Liu, Haitao Zheng and Ying Shen. 2019. chinese Relation Extraction with Multi-Grained Information and External Linguistic Knowledge. *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 4377-4386.
- Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan and Maosong Sun. 2016. Neural relation extraction with selective attention over instances. *The 54th Annual Meeting of the Association for Computational Linguistics*, pages 2124-2133.
- Zhouhan Lin, Minwei Feng, Cícero Nogueira dos Santos, Mo Yu, Bing Xiang, Bowen Zhou and Yoshua Bengio. 2017. A structured self-attentive sentence embedding. *The 5th International Conference on Learning Representations*, pages 24-26.
- 李卫疆, 李涛, 漆芳. 2019. 基于多特征自注意力BLSTM的中文实体关系抽取. *中文信息学报*, 33(10):47-56.
- Zhenhua Liu, Tianyi Wang, Wei Dai, Zehui Dai and Guangpeng Zhang. 2019. A Relation Proposal Network for End-to-End Information Extraction. *The Natural Language Processing and Chinese Computing*, pages 782-790.
- Diego Marcheggiani and Ivan Titov. 2017. Encoding sentences with graph convolutional networks for semantic role labeling. *The 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1506-1515.
- Tomas Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean. 2013(b). Efficient estimation of word representations in vector space. *The 1st International Conference on Learning Representations*, pages 2-4.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado and Jeffrey Dean. 2013(a). Distributed Representations of Words and Phrases and their Compositionality. *The 27th Annual Conference on Neural Information Processing Systems*, pages 3111-3119.

- Melanie Reiplinger, Michael Wiegand and Dietrich Klakow. 2009. Distant supervision for relation extraction without labeled data. *The 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 2-7.
- Makoto Miwa and Mohit Bansal. 2016. End-to-end relation extraction using lstms on sequences and tree structures. *The 54th Annual Meeting of the Association for Computational Linguistics*, pages 7-12.
- Yilin Niu, Ruobing Xie, Zhiyuan Liu and Maosong Sun. 2017. Improved Word Representation Learning with Sememes. *The 55th Annual Meeting of the Association for Computational Linguistics*, pages 2049-2058.
- Sachin Pawar, Girish K. Palshikar and Pushpak Bhattacharyya. 2017. Relation Extraction : A Survey. *arXiv:1712.05191v1 [cs.CL]* .
- Jeffrey Pennington, Richard Socher and Christopher D. Manning. 2014. Glove: Global vectors for word representation. *The 2014 conference on empirical methods in natural language processing*, pages 1532-1534.
- Sebastian Riedel, Limin Yao and Andrew McCallum. 2010. Modeling Relations and Their Mentions without Labeled Text. *European Conference on Machine Learning & Knowledge Discovery in Databases*, pages 20-24.
- Bryan Rink and Sanda M. Harabagiu. 2010. UTD: Classifying semantic relations by combining lexical and semantic resources. *The 5th International Workshop on Semantic Evaluation*, pages 15-16.
- Cícero Nogueira dos Santos, Bing Xiang and Bowen Zhou. 2015. Classifying relations by ranking with convolutional neural networks. *The 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing*, pages 626-634.
- Richard Socher, Brody Huval, Christopher D. Manning and Andrew Y. Ng. 2012. Semantic compositionality through recursive matrix-vector spaces. *The 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, pages 1201-1211.
- Changzhi Sun, Yeyun Gong, Yuanbin Wu, Ming Gong, Daxin Jiang, Man Lan, Shiliang Sun and Nan Duan. 2019. Joint Type Inference on Entities and Relations via Graph Convolutional. *The 57th Annual Meeting of the Association for Computational Linguistics*, pages 1361-1370.
- Changzhi Sun, Yuanbin Wu, Man Lan, Shiliang Sun, Wenting Wang, Kuang-Chih Lee and Kewen Wu. 2018. Extracting Entities and Relations with Joint Minimum Risk Training. *The 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2256-2265.
- Ryuichi Takanobu, Tianyang Zhang, Jiexi Liu and Minlie Huang. 2019. A Hierarchical Framework for Relation Extraction with Reinforcement Learning. *The Thirty-Third AAAI Conference on Artificial Intelligence*, pages 7072-7079.
- Zhixing Tan, Mingxuan Wang, Jun Xie, Yidong Chen and Xiaodong Shi. 2018. Deep semantic role labeling with self-attention. *The Thirty-Second AAAI Conference on Artificial Intelligence*, pages 4929-4936.
- Shikhar Vashishth, Rishabh Joshi, Sai Suman Prayaga, Chiranjib Bhattacharyya and Partha P. Talukdar. 2018. Improving distantly-supervised neural relation extraction using side information. *The 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1257-1266.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, pages 5998-6008.
- Oriol Vinyals, Meire Fortunato and Navdeep Jaitly. 2015. Pointer networks. *Advances in Neural Information Processing Systems*, pages 2692-2700.
- Guanying Wang, Wen Zhang, Ruoxu Wang, Yalin Zhou, Xi Chen, Wei Zhang, Hai Zhu and HuaJun Chen. 2018. Label-free distant supervision for relation extraction via knowledge graph embedding. *The 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2246-2255.
- 武文雅, 陈钰枫, 徐金安. 2019. 基于高层语义注意力机制的中文实体关系抽取. *广西师范大学学报*, 33(01):32-41.

- 武文雅, 陈钰枫, 徐金安. 2018. 中文实体关系抽取研究综述. *计算机与现代化*, 000(008):25-31.
- Jason Weston, Antoine Bordes, Oksana Yakhnenko and Nicolas Usunier. 2013. Connecting language and knowledge bases with embedding models for relation extraction. *The 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1366-1371.
- Kun Xu, Yansong Feng, Songfang Huang and Dongyan Zhao. 2015(a). Semantic relation classification via convolutional neural networks with simple negative sampling. *The 2015 Conference on Empirical Methods in Natural Language Processing*, pages 17-21.
- Yan Xu, Lili Mou, Ge Li, Yunchuan Chen, Hao Peng and Zhi Jin. 2015(b). Classifying relations via long short term memory networks along shortest dependency paths. *The 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1785-1794.
- Dmitry Zelenko, Chinatsu Aone and Anthony Richardella. 2003. Kernel Methods for Relation Extraction. *Journal of Machine Learning Research*, pages 1083-1106.
- Daojian Zeng, Kang Liu, Yubo Chen and Jun Zhao. 2015. Distant supervision for relation extraction via piecewise convolutional neural networks. *The 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1753-1762.
- Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou and Jun Zhao. 2014. Relation classification via convolutional deep neural network. *The 2015 Conference on Empirical Methods in Natural Language Processing*, pages 23-29.
- Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu and Jun Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. *The 56th Annual Meeting of the Association for Computational Linguistics*, pages 506-514.
- Meishan Zhang, Yue Zhang and Guohong Fu. 2017. End-to-end neural relation extraction with global optimization. *The 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1730-1740.
- Ningyu Zhang, Shumin Deng, Zhanlin Sun, Guanying Wang, Xi Chen, Wei Zhang and Huajun Chen. 2019. Long-tail Relation Extraction via Knowledge Graph Embeddings and Convolution Networks. *The 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3016-3025.
- Yuhao Zhang, Peng Qi and Christopher D. Manning. 2018. Graph convolution over pruned dependency trees improves relation extraction. *The 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2205-2215.
- Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou and Bo Xu. 2017. Joint extraction of entities and relations based on a novel tagging scheme. *The 55th Annual Meeting of the Association for Computational Linguistics*, pages 1227-1236.
- Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao and Bo Xu. 2016. Attention-based bidirectional long short-term memory networks for relation classification. *The 54th Annual Meeting of the Association for Computational Linguistics*, pages 207-212.