

基于深度学习的事件抽取研究综述*

王浩畅¹, 周郴莲¹, Marius Gabriel PETRESCU²

¹(东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

²(Universitatea Petrol-gaze din Ploiesti, Bucharest 100680, Romania)

通信作者: 周郴莲, E-mail: chenlian_zhou@163.com



摘要: 事件抽取是从非结构化的自然语言文本中自动抽取用户感兴趣的事件信息, 并以结构化的形式表示出来。事件抽取是自然语言处理与理解中的重要方向, 在政府公共事务管理、金融业务、生物医学等不同领域有着很高的应用价值。根据对人工标注数据的依赖程度, 目前基于深度学习的事件抽取方法主要分为两类: 有监督和远程监督学习方法。对当前深度学习中事件抽取技术进行了全面的综述。围绕有监督中 CNN、RNN、GAN、GCN 与远程监督等方法, 系统地总结了近几年的研究情况, 并对不同的深度学习模型的性能进行了详细对比与分析。最后, 对事件抽取面临的挑战进行了分析, 针对研究趋势进行了展望。

关键词: 事件抽取; 有监督学习; 深度学习; 远程监督; 信息抽取

中图法分类号: TP18

中文引用格式: 王浩畅, 周郴莲, Marius Gabriel PETRESCU. 基于深度学习的事件抽取研究综述. 软件学报, 2023, 34(8): 3905–3923. <http://www.jos.org.cn/1000-9825/6645.htm>

英文引用格式: Wang HC, Zhou CL, Petrescu MG. Survey on Event Extraction Based on Deep Learning. Ruan Jian Xue Bao/Journal of Software, 2023, 34(8): 3905–3923 (in Chinese). <http://www.jos.org.cn/1000-9825/6645.htm>

Survey on Event Extraction Based on Deep Learning

WANG Hao-Chang¹, ZHOU Chen-Lian¹, Marius Gabriel PETRESCU²

¹(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

²(Universitatea Petrol-gaze din Ploiesti, Bucharest 100680, Romania)

Abstract: Event extraction is to automatically extract event information in which users are interested from unstructured natural language texts and express it in a structured form. Event extraction is an important direction in natural language processing and understanding and is of high application value in different fields, such as government management of public affairs, financial business, and biomedicine. According to the degree of dependence on manually labeled data, the current event extraction methods based on deep learning are mainly divided into two categories: supervised learning and distantly-supervised learning. This article provides a comprehensive overview of current event extraction techniques in deep learning. Focusing on supervised methods such as CNN, RNN, GAN, GCN, and distant supervision, this study systematically summarizes the research in recent years. Additionally, the performance of different deep learning models is compared and analyzed in detail. Finally, the challenges facing event extraction are analyzed, and the research trends are forecasted.

Key words: event extraction; supervised learning; deep learning; distant supervision; information extraction

随着云计算与大数据时代的迅速推进, 计算机已经是人们平时获取信息最重要的途径。从各种数据形式中获取最有用的、潜在的信息已成为人们关注的重点方向, 信息抽取技术应运而生。信息抽取就是从海量的文本、图片和视频等数据里面自动抽取用户需要的结构化信息的过程。事件抽取作为信息抽取技术的主要分支之一, 同时

* 基金项目: 国家自然科学基金 (61402099, 61702093)

收稿时间: 2020-12-22; 修改时间: 2021-06-28, 2021-08-18; 采用时间: 2022-01-13; jos 在线出版时间: 2022-05-24

CNKI 网络首发时间: 2023-03-17

还是该方向最有挑战性的任务之一. 事件抽取一直吸引着许多研究机构和学者, 如消息理解会议 (message understanding conference, MUC)^[1]和自动内容抽取 (automatic content extraction, ACE) 会议^[2]就把事件抽取作为典型任务.

事件抽取任务研究是从非结构化的自然语言文本中自动抽取用户感兴趣的事件信息并以结构化的形式表示^[3], 融合了来自计算机科学、语言学、数据挖掘、人工智能和知识建模等多个领域的知识和经验, 对人们了解社会有着深远的影响. 事件抽取在不同领域中具有许多应用, 例如结构化事件能够直接扩充知识库并进行逻辑推理. 事件检测与监控一直是政府公共事务管理的重点, 实时了解社会事件的爆发和演变有助于对其迅速做出反应并采取措施. 在金融业务领域, 事件抽取可以帮助公司快速发现产品的市场响应并推断信号以执行风险分析、评估等操作. 在生物医学领域, 事件抽取能够识别生物分子 (例如基因或蛋白质) 状态的变化, 以及它们之间的相互作用. 事件抽取在应用需求的推动下展开, 由人工标注数据的依赖情况可以将目前基于深度学习的事件抽取方法主要分为两类: 有监督和远程监督学习方法. 本文首先简单介绍深度学习中事件抽取的发展, 再从事件抽取研究的方法, 对其发展状况和技术推进两个维度全面阐述事件抽取的工作, 然后概述了深度学习中事件抽取的数据集及对评价指标进行分析, 最后讨论了事件抽取中面临的挑战及研究趋势, 并对其进行了总结和展望.

1 事件抽取的发展

20 世纪 80 年代末, 事件抽取的研究开始蓬勃发展, 其中耶鲁大学、MUC 会议、ACE 会议和文本分析会议 (text analysis conference, TAC)^[4]的相关测评都推动着事件抽取技术的快速发展, 情况如表 1 所示.

表 1 事件抽取有关的描述

机构/会议	事件抽取任务描述	应用领域/语言	成果
耶鲁大学	开创了事件类型识别的研究, 同时也进行了事件理解的研究	新闻热点话题信息分类	信息抽取系统的设计与实现
MUC会议	定义了各种概念、模型的规范以及完整的评估系统, 提出了采用模板抽取整个事件场景	新闻报道	为事件抽取研究奠定了坚实的基础
ACE会议	将事件抽取描述为一项任务, 即从文本中识别事件并抽取事件论元	不受特定领域与方案的限制	事件抽取的准确性和多样性取得了重大进展
TAC-KBP	将事件抽取描述为自动抽取文本中特定类型的事件属性的过程	中文, 英文, 西班牙语	提供了大量的语料库

MUC 会议每两年举办一次, 主要是美国国防部高级研究计划局 (defense advanced research projects agency, DARPA)^[1]赞助的. 它从 1987–1997 年总共举办了 7 届, 积极推动了事件抽取研究方向的确立与后续发展的方向.

ACE 会议是 MUC 的延伸, 在大量的应用需求下, 1999 年美国国家标准与技术研究院 (national institute of standards and technology, NIST)^[2]组织的 ACE 评测会议开始发展起来. ACE 从 2000–2007 年共举办了 8 届, 这是事件抽取领域最有影响力的评测会议.

TAC 会议是 NIST 在 2008 年成立的, 自 2009 年开始 ACE 就成为 TAC 中的一个子任务. 随着云计算和大数据时代的到来, 数据呈爆炸式增长, 上述测评会议所发布的依靠人工标注方式获得的语料库已经无法满足需求. 2014 年 TAC 增加了知识库生成 (knowledge base population, KBP)^[4]评测任务, 同时也增加了事件抽取的任务. 如今, 事件抽取已成为 TAC-KBP 公开评测的主要任务, 可以从大型文本语料库中自动抽取事件信息, 完成对知识库中不足论元^[5]的补充.

传统基于特征的方法是利用人工构建事件候选触发词与论元, 这会导致模型的扩展性和移植性较差. 而机器学习方法在特征提取过程容易出现误差传播问题, 极大影响事件抽取模型的性能. 随着深度学习的崛起, 研究者们逐渐将深度学习方法引入事件抽取任务中, 大量基于有监督的卷积神经网络 (convolutional neural networks, CNN)^[6]、递归神经网络 (recurrent neural networks, RNN)^[7]、生成对抗网络 (generative adversarial networks, GAN)^[8]、图卷积网络 (graph convolutional networks, GCN)^[9]与远程监督方法的事件抽取模型被提出.

2 事件抽取研究的方法

近年来,深度学习技术已广泛用于复杂结构的建模,并验证了对许多 NLP 任务都有效,例如机器翻译^[10]、关系抽取^[11]和情感分析^[12]等.双向长期短期记忆 (bi-directional long short-term memory, Bi-LSTM) 模型^[13]是一种双向 RNN,可以捕获前后上下文中每个单词的信息,同时利用其信息对单词表示进行建模. CNN 是另一种有效的模型,可以提取事件语义表示同时捕获其结构特征.

事件抽取在应用需求的推动下展开,由人工标注数据的依赖程度可以将目前基于深度学习的事件抽取方法主要分为两类:有监督和远程监督学习方法.有监督学习在训练过程中使用人工标注的数据集,而远程监督的学习方法通过对齐远程知识库自动对语料进行标注来获取带标记语料数据,具体比较情况如表 2 所示.

表 2 有监督与远程监督学习方法对比

方法	主要思想	优势	劣势
有监督	对实体和文本通过特征工程或表征学习来提取句法或语义信息	人工标注语料,具有较高的准确率	标注成本高、效率低,导致训练数据缺乏且移植性较差
远程监督	将大量的无标注语料与现有的知识库进行对齐,从而快速地构建大规模监督语料	构建语料效率高、成本低	知识库不充分,不能保证语料的质量,且数据严重不均衡

2.1 基于深度学习的有监督事件抽取

在有监督中运用深度学习方法进行事件抽取,已经成为这几年事件抽取的热门研究方向.表 3 整理了深度学习框架下有监督事件抽取的经典模型^[14-28].其中,模型主要是改进 CNN、RNN 输入特征或网络结构,比如添加不同特征、结合多种注意力机制和引入依存树等挖掘更深层次事件语义信息来提升其性能.下面对相关模型进行深入研究和分析.

表 3 有监督事件抽取的经典模型

方法	模型名称	创新	优势	劣势	提出年份及机构
CNN	DMCNN ^[14]	设计了一个动态多池卷积神经网络	自动抽取词汇级和句子级信息	单独预测所有候选论元,捕获远程依赖的效率低	2015年 ACL
	CNN ^[15]	率先将卷积神经网络用于检测事件	自动学习及抽取有效特征	单独预测所有候选论元,没有考虑它们的相互作用	2015年 ACL
	NC-CNN ^[16]	率先使用非连续 CNN	捕获预测触发词所必须的远程和非连续的依赖	捕获不必要与嘈杂信息	2016年 ACL
	HMEAE ^[17]	提供事件论元角色的概念层有效的归纳偏差	捕获论元的概念相关性	忽略了事件论元的相关性,捕获远程依赖的效率低	2019年 EMNLP
	Bi-LSTM-CNN ^[18]	不依赖任何语言,不需要特征工程	捕获特定上下文的序列信息和语义信息	忽略了事件的相互依赖性,产生错误传播	2016年 ACL
RNN	JRNN ^[19]	率先用神经网络进行联合事件抽取	同时识别事件触发词和论元角色	忽略了事件的依赖性及论元的相关性,容易出现错误传播	2016年 NAACL
	BDLSTM-TNNs ^[20]	设计了一个双向动态多池长短期记忆张量神经网络	自动抽取,捕获论元间的相互作用	忽略了事件的相互依赖性,会出现事件歧义	2016年 CCL
	dbRNN ^[21]	设计了一个依赖桥递归神经网络	明确了事件的语法信息	忽略了事件论元的相关性	2018年 AAAI
	DAG-GRU ^[22]	语法信息与注意力机制结合	捕获上下文和句法信息	语法链接会出现干扰,降低了语言信息的可用性和可靠性	2018年 EMNLP
	HBTNGMA ^[23]	设计了一个门控多层注意力机制	自动抽取并动态集成句子级和文档级信息	忽略了事件的依赖性,会出现错误传播的问题	2018年 EMNLP
	TLNN ^[24]	设计了一个触发感知 Lattice 网络	解决触发不匹配和多义触发	忽略了事件的依赖性	2019年 EMNLP

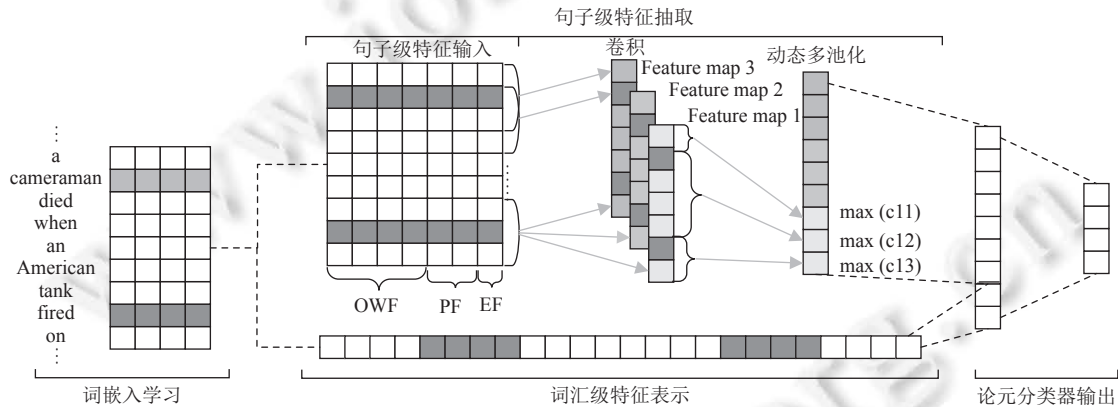
表 3 有监督事件抽取的经典模型 (续)

方法	模型名称	创新	优势	劣势	提出年份及机构
GAN	SELF ^[25]	设计了一种自调节学习方法	避免错误及虚假的特征信息	抽取全局特征时,会受到噪声的负面影响	2018年 ACL
	AEM ^[26]	率先将对抗训练用于开放域事件抽取	可以从长文本生成论元	会产生无关事件,处理较长文章时模型准确率下降	2019年 EMNLP
GCN	GCN-ED ^[27]	率先把语法信息集成到神经网络事件检测	图卷积网络捕获句法信息	破坏了相邻词表示的多样性	2018年 AAAI
	MOGANED ^[28]	率先将GAT应用于事件检测的研究	用一阶句法图和高阶句法图来显式地模拟候选触发词的多阶表示	触发词预测不准确,事件类型存在歧义	2019年 EMNLP

2.1.1 基于 CNN 模型的事件抽取方法

CNN 模型能够自动抽取有效特征进行事件抽取与识别任务,自动使用最大池化层获取整个句子的表示形式,捕获最有价值的信息。

Chen 等人^[14]和 Nguyen 等人^[15]在 2015 年率先尝试使用 CNN^[29]自动抽取有效特征进行事件抽取和识别任务,由此避免了特征提取中的错误传播以及使用复杂的 NLP 工具包的问题。其中,Chen 等人^[14]将事件类型特征用于论元角色分类任务,提出了一种基于动态多池卷积神经网络 (DMCNN),如图 1 所示。他们都在 Li 等人^[30]事件类型识别结果的基础上进行了对比实验,其中 F1 值分别提高了 1.6% 和 1.5%,验证了事件类型特征的有效性。

图 1 DMCNN 事件抽取模型论元分类阶段体系结构^[14]

2016 年, Nguyen 等人^[16]在 Lei 等人^[31]的非连续 n-gram 的基础上,提出了一种基于卷积神经网络的 skip-grams 事件检测模型 (NC-CNN)。与 Lei 等人工作的不同之处在于,他们通过位置嵌入模拟句子中触发候选词的相对距离。实验结果表明,他们在 2015 年事件识别^[15]F1 值为 69.0% 的基础上提高了 2.3%,验证了使用非连续 n-gram 机制进行事件检测优于传统的 CNN 模型。Zhang 等人^[32]则结合文本特征和视觉模型提出了一种多模态的事件抽取方法。在 ACE 2005 数据集上,该方法进行事件触发标签和论元标签的实验的 F1 值分别提高了 7.1% 和 8.5%。

与前面工作不同的是, Liu 等人^[33]将事件检测和词义消歧当作两个相似的任务。他们利用两个单独的 CNN 模型对其上下文建模,匹配两项任务中特征相似的神经表示。通过实验对比了多种深度学习模型,验证了其模型的效果,而且他们的方法还不依赖于手工注释的事件实体提及。

除了上述基于单一模型的深度学习应用在事件抽取中的应用,还有一些基于混合模型的方法。2019 年, Lin 等人^[34]利用错误预测的数据重新构造模型训练的损失函数,提出了一种基于 CNN 与 RNN 联合的代价敏感正则化方法 (CR-INS)。通过实验证明,他们的方法可以提高神经网络事件检测模型的性能。Wang 等人^[17]发现事件抽取中的方法大部分都是将每个论元角色单独分类,而忽略了不同论元之间的概念相关性。他们提出了一种分层模块化

事件论元抽取模型 (HMEAE). 在 ACE 2005 数据集上的实验表明, 该方法比 DMCNN 模型的 $F1$ 值提高了 5.8%.

由于生物学领域的事件结构较复杂, 深度学习应用相对较少, 但是近几年也取得了一定的进步. 2018 年, Li 等人^[35]为了解决生物学事件抽取中复杂触发词的问题, 提出了一种并行多池卷积神经网络 (PMCNN), 可以捕获句子的语义特征. 实验结果表明, 在 MLEE 数据集^[36]上进行触发词识别的任务 $F1$ 值达到了 59.65%, 比以前的工作提高了 2.24%. 2020 年, Zhu 等人^[37]和 Ramponi 等人^[38]都结合深度学习进行了生物学事件抽取的研究. Zhu 等人^[37]忽略了生物学事件的交互性, 将其分解为简单子任务, 提出了一种混合深度神经网络的事件组合策略方法. 在 MLEE 数据集上进行事件触发词识别任务的实验 $F1$ 值为 60.05%, 在其他数据集上的 $F1$ 值也较高, 证明了其方法的有效性和泛化能力较强. 而 Ramponi 等人^[38]率先结合 CNN 架构提出了一种跨域边缘检测的生物学事件抽取模型. 在 5 个生物学语料库上进行实验, 均取得不错的效果, 证明了其模型的鲁棒性和泛化性较好.

2.1.2 基于 RNN 模型的事件抽取方法

Goller 等人^[39]在 1996 年就提出了用 RNN 模型来处理可变的输入长度序列. 然而, 梯度消失或者梯度爆炸这些典型的问题经常出现在训练过程中. RNN 模型的一些变体, 如 LSTM^[40]和门控循环单元 (gated recurrent unit, GRU)^[41]在内的大多数循环神经网络架构都是使用门控控制信息流, 解决传统 RNN 出现的问题.

2016 年, Nguyen 等人^[19]提出了一种双向递归神经网络的联合框架 (JRNN). 他们利用两个 RNN 构建句子的局部特征和全局特征, 同时进行事件识别和论元角色分类的任务, 如图 2 所示. 在 ACE 2005 数据集上的实验表明, 模型在论元角色识别中的 $F1$ 值比 DMCNN 模型提高了 1.9%, 证明了该模型可以高效地学习多事件句子特征. Duan 等人^[42]提出了一种文档级递归神经网络模型 (DLRNN), 可以自动学习跨句子特征. 通过实验证明, 该模型在句子级和跨句子级的事件检测都优于其他模型, 能够很好地捕获文档级信息. 现有研究大部分都是单独抽取事件论元, 而不考虑论元间的相互作用, Chen 等人^[20]提出了一种动态多池的双向长短期记忆网络 (BDLSTM-TNNs). 他们通过 LSTM 的上下文感知词表示模型捕获文本中单词的语义, 然后利用神经网络的张量层探索候选论元间的相互作用同时预测所有候选论元. 该模型在事件论元分类上获得了较好的性能, 而且比 DMCNN 模型的 $F1$ 值提高了 0.6%, 揭示了该模型能够自动学习并捕获单词语义信息.

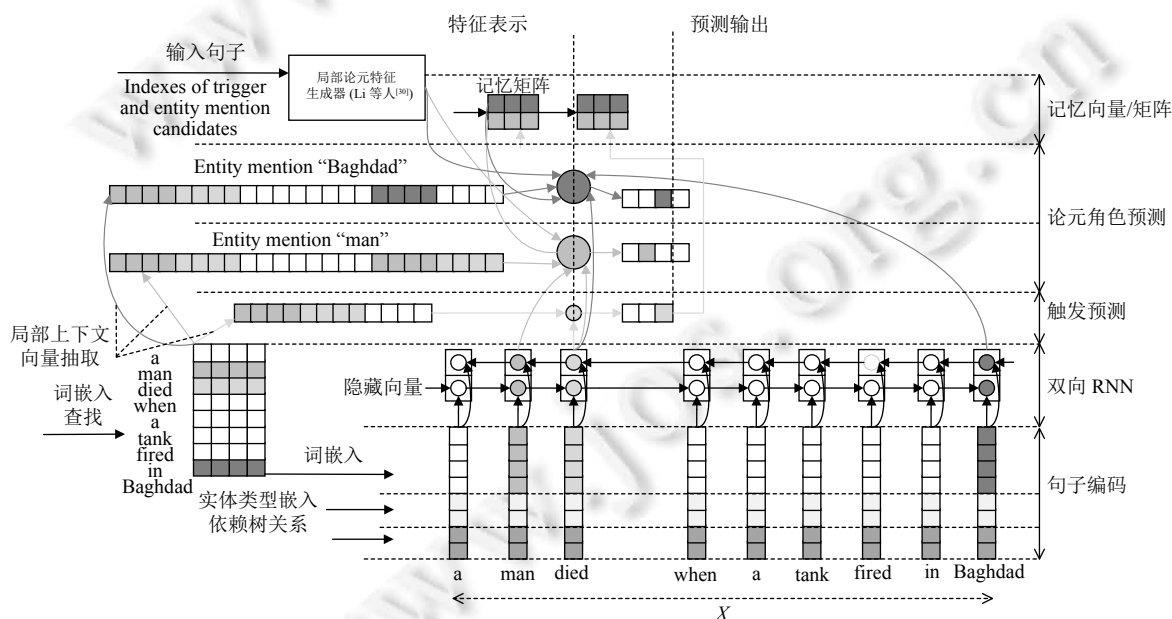


图 2 JRNN 模型^[19]

2018 年, Sha 等人^[21]提出了一种基于依赖桥递归神经网络的事件抽取模型 (dbRNN). 该模型不同于传统事件抽取方法需要依赖词法句法特征, 而是通过依赖关系桥对每个单词建模来表达与语法相关的信息. 实验结果表明,

该模型在事件触发词分类和论元分类中 $F1$ 值比 JRNN 模型分别提高了 1.6% 和 3.3%。Nguyen 等人^[43]通过深度学习中的共享隐藏表示,对实体提及、事件触发与论元进行联合预测。他们提出了一个基于深度学习的模型 (EMD-Joint3EE),采用双向 RNN 为句子中的单词引入共享的隐藏表示。实验结果表明,该模型优于基于特征的方法,与其他深度学习模型对比,在事件论元识别和论元角色分类中 $F1$ 值分别提高了 2.7% 和 2.0%。Orr 等人^[22]利用具有依存关系的文本信息双向捕获上下文信息和序列信息,把 GRU 模型推广到图结构上,提出了一个 DAG-GRU 的模型。在 ACE 2005 数据集的实验表明,它比其他神经网络模型性能更好。

以往的很多研究,都很难实现复杂上下文的事件语义表示,Wu 等人^[44]提出了一种基于注意力机制的递归神经网络 (RNN-ARG),可以捕获语义规律。而 Ding 等人^[45]提出了一种基于多层注意力和深度上下文的词表示事件抽取方法。该模型集成了 RNN 和注意力机制在事件抽取中的效果,可以获得更深层的语义与上下文信息,这使得模型的精度和召回率都有所提高。Zhang 等人^[46]提出了一种基于注意力机制 Bi-LSTM 模型,利用模型中的关系类型嵌入进行事件检测。实验表明,该模型在触发词识别和类型分类中 $F1$ 值分别达到了 76.1% 和 73.9%,说明它在 Bi-LSTM 神经网络之前集成注意力机制,能够更好地捕获句子的关键信息。

Liu 等人^[47]针对数据稀缺和单语言歧义的问题,提出了一种门控多语言注意力机制 (GMLATT) 的事件检测框架,这是率先在事件检测任务中引入注意力机制建模多语言数据的工作。该方法在实验结果中优于其他传统的深度学习方法,其中事件类型分类中 $F1$ 值提高了 0.7%。Chen 等人^[23]利用分层的偏置标注网络检测一个句子中的多个事件,提出了一种基于门控多层注意力机制 (HBTNGMA) 的偏置标注网络。与 JRNN 模型相比事件类型识别的 $F1$ 值提高了 4.0%,实验结果表明,该模型可以同时检测一个句子中的多个事件并动态集成上下文信息。Zhao 等人^[48]提出了一种文件嵌入增强双向 RNN 模型 (DEEB-RNN),通过基于有监督分层注意力机制 RNN 模型进行事件检测,然后用其结果增强另一个双向 RNN 模型识别事件触发与句子类型。在 ACE 数据集上,该模型在事件类型分类的任务中 $F1$ 值达到了 74.0%。

2019 年,Chen 等人^[49]为了减少其他句子噪声带来的负面影响,提出了一种多头注意力分割机制的方法进行事件检测。实验结果表明,该方法在触发词识别和事件类型分类的 $F1$ 值分别是 76.8% 和 74.2%,分别提高了 0.9% 和 0.8%。Mehta 等人^[50]则提出了一种分解式双线性多层注意力机制 (FBMA)。该方法使用多层注意力的模型构建句子和文档表示信息,通过 3 个事件数据集的实验证明了其模型具有较好的性能。黄细凤^[51]提出了一种基于动态掩蔽注意力机制的事件抽取模型 (Dy MAN)。在 ACE 2005 数据集上进行了实验,对于多事件抽取任务,与之前最好的模型 JRNN 相比,Dy MAN 模型在触发词分类任务上取得了 9.8% 的提升。

表 4 将注意力机制按其结构分 3 种类别,分别是单层自注意力、多层注意力和多头注意力机制。无论是基于 CNN、LSTM 还是 GRU 衍生出的深度学习事件抽取模型,均可以通过引入不同注意力机制或其组合来提高性能,也可以有效缓解监督学习中的噪声问题。

表 4 注意力机制相关方法统计

类别	模型名称
Self-attention	RNN-ARG ^[44]
	Re+Bi-LSTM+Att1 ^[46]
	Dy MAN ^[51]
Multi-level attention	RNN+Attention ^[45]
	GMLATT ^[47]
	HBTNGMA ^[23]
	DEEB-RNN ^[48]
	FBMA ^[50]
Multi-head attention	Bi-LSTM+Attention ^[49]

2.1.3 基于 GAN 模型的事件抽取方法

GAN 是基于神经网络的生成模型,主要是由生成器和判别器这两个部分构成。其中,生成器能够学习真实的

事件特征分布情况, 有利于自身特征的生成以进行抽取的操作, 同时还能骗过判别器. 判别器则能够让接收到的事件特征实现真假判别.

2018年, Hong等人^[25]提出了一种基于GAN生成虚假特征的自调节学习方法(SELF), 采用了Bi-LSTM和GAN联合的循环网络来消除虚假特征. 在ACE 2005数据集上的实验表明, 该方法在事件触发词识别和事件分类中F1值分别为77.0%和73.0%, 进一步证明了GAN方法的有效性和较强的泛化性能. Zhang等人^[52]也提出了一种生成对抗网络的逆强化学习方法, 通过动态机制学习实体和事件的特征. 实验结果表明, 该方法能够正确处理错误与混乱事件类别, 与dbRNN模型在事件触发词识别和角色识别的F1值相比都有所提升.

2019年, Liu等人^[53]提出了一种对抗模仿学习的事件抽取方法. 该方法不依赖NLP工具包构建知识表示, 可以从原始语句获取知识进行事件检测, 在ACE 2005数据集上的实验结果表明, 在事件类型上与DLRNN模型相比F1值提高了4.3%. 而Wang等人^[26]发现贝叶斯图模型, 如果依赖Gibbs采样进行论元推理, 需要较长时间才能收敛, 因此他们提出了一种对抗神经事件模型(AEM), 如图3所示. 该模型使用Dirichlet分布对事件建模, 利用生成器网络捕获潜在的事件信息. 这是率先将对抗训练用于开放域事件抽取的研究. 实验结果表明, 与其他深度学习方法相比, 该模型的事件抽取性能更好, 特别是在长文本数据集中F1提高了15%; 与现有的贝叶斯图模型相比, 该模型需要的训练时间更短.

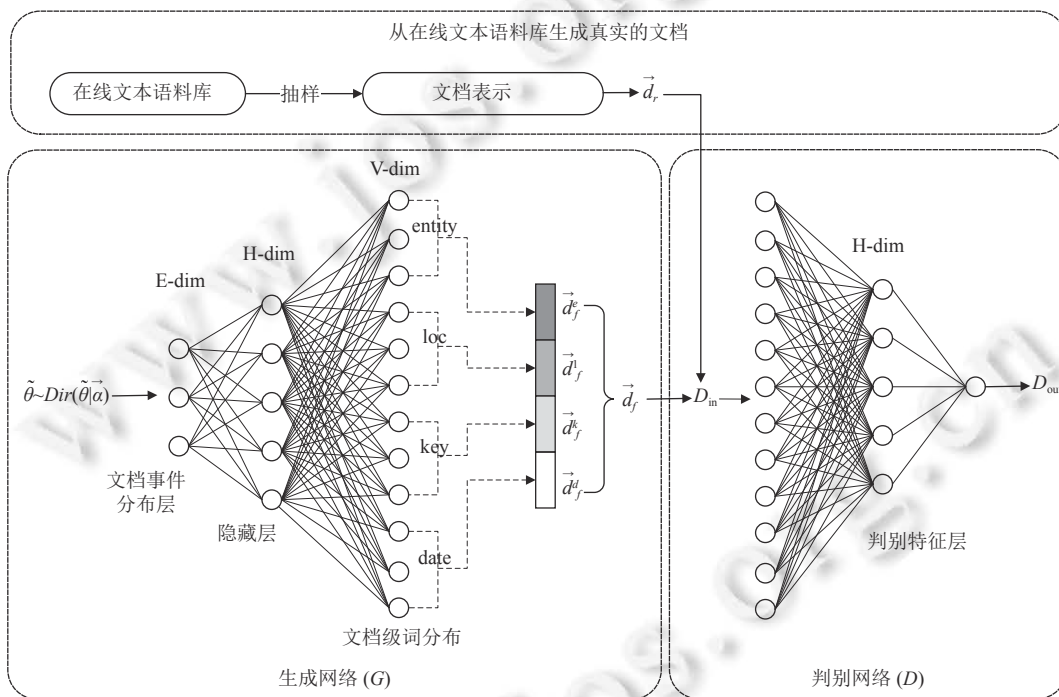


图3 AEM框架^[26]

2.1.4 基于GCN模型的事件抽取方法

GCN能够依靠图谱理论在拓扑图上完成卷积操作, 在事件抽取中, 根据单词间的句法连通性构建句子图引入句子结构信息.

2018年, Nguyen等人^[27]提出了一种依赖树卷积神经网络, 他们依靠实体提及聚合卷积向量. 该方法不同于传统的神经网络模型只考虑句子的顺序表示, 而是依靠实体提及聚合卷积向量把语法集成到GCN进行事件检测. 由于现有研究捕获长距离的依存关系效率低, 需要大量的人力资源, 而且还不能充分建模事件间的关联, Liu等人^[54]针对这个问题提出了一种多个事件联合抽取的框架(JMEE). 他们引入了句法捷径弧增强信息流和注意力图卷积来建模图信息, 可以联合多个事件触发词和论元进行事件抽取. 实验表明, 在事件触发词分类和角色识别上F1值

分别比 dbRNN 模型提高了 1.8% 和 1.6%。

由于现有研究仅使用依存树中的一阶句法关系识别触发词, Yan 等人^[28]提出了一种多阶图网络注意力的事件检测方法 (MOGANED)。该方法在一阶句法图的基础上, 引入高阶句法图建模候选触发词的多阶表示, 在 ACE 2005 数据集上进行实验, 事件类型识别的 $F1$ 值为 75.7%, 比 JMEE 模型提高了 2%。

2020 年, Cui 等人^[55]针对现有依赖树图卷积网络的研究忽略了句法分析树关系结构的问题, 提出了一种关系感知图卷积网络 (RA-GCN)。实验表明, 其在事件类型识别上 $F1$ 值比 Yan 等人^[28]的 MOGANED 模型提高了 1.9%, 说明了该模型可以利用句法关系标签专门为单词间的关系建模。这是率先同时使用 GCN 的句法依赖结构和关系标签进行事件抽取的研究。

2.2 基于远程监督的事件抽取

远程监督 (distant supervision, DS) 可以为事件抽取自动生成大规模标注数据^[56]。为了解决数据标注问题, Chen 等人^[57]利用远程监督方法结合 Freebase^[58]构建了一个基于维基百科的事件数据集。他们通过 Freebase 找出事件关键论元, 自动检测事件和触发词, 利用 FrameNet^[59]过滤噪声。实验结果表明, 该模型能够学到合理权重来缓解远程监督中噪声问题, 同时还充分挖掘有用信息。而 Zeng 等人^[60]则从维基百科和 Freebase 中自动生成训练事件抽取的数据, 将事件抽取训练实例从数千个扩展到数十万个。该模型集成了远程监督的知识库, 自动从未标记文本中标注事件数据, 并开发了一个基于 Bi-LSTM 和 CRF 的联合神经网络模型。实验结果证明, 该模型可以与已有的数据相结合自动抽取事件, 还可以进行多类型的事件检测。Keith 等人^[61]利用 2016 年全年收集的警察死亡语料库, 提出了一个基于特征逻辑回归和卷积神经网络分类器的远程监督模型。通过实验结果可以看出, 该模型与现有的抽取模型的 $F1$ 值相比有所提高, 说明其方法可以更好地进行事件抽取, 但是仍存在人工标注导致的错误传播的问题。Rao 等人^[62]提出了一种抽象意义表示 (AMR) 的方法识别生物医学文本分子事件。他们对事件结构的 AMR 子图进行假设验证, 在给定 AMR 的情况下利用远程监督神经网络模型, 识别事件子图。通过对其在 BioNLP 共享任务的 GENIA 事件抽取子任务^[63]上实验表明, 仅以蛋白质为论元的简单事件下 $F1$ 值达到了 94.74%, 而复杂事件下降到了 74.18%, 这是由于 AMR 存在错误解析导致的。

2.3 对事件抽取中深度学习方法的探索

FrameNet (FN) 中的帧由一个词法单元和一组帧元素组成, 分别扮演与 ACE 事件触发词和论元相似的角色, 缓解了 ACE 数据集事件类型稀疏的问题。Liu 等人^[64]提出了一种基于概率软逻辑的全局推理方法检测 FN 事件。他们还构建了 Event-FN 数据集, 缓解了数据稀疏的问题。通过实验证明, 该方法可以实现相关事件类型之间的信息共享。Wadden 等人^[65]把全局上下文合并到 IE 框架中, 提出了一种基于上下文的跨实体、关系和事件抽取框架 (DYGIE++)。该方法在 ACE 2005 数据集上, 事件触发检测和论元角色分类的 $F1$ 值分别为 76.5% 和 52.5%。他们对其他数据集也进行了验证, 都优于对比的模型。

Yang 等人^[66]对事件、实体及篇章内不同事件的依赖关系进行建模, 提出了一种完全端到端学习的模型。实验结果表明, 该模型在事件类型分类和论元识别的任务中 $F1$ 值都提高了 1.0%。Han 等人^[67]利用端到端的方法, 提出了一种基于共享表示与结构化预测的联合事件和时间关系抽取模型。他们运用两阶段学习方法, 首先允许事件和关系模块共享相同的上下文嵌入与神经表示, 其次利用结构化的推理和学习方法共同分配事件和事件关系标签, 避免了常规管道系统中错误传播的问题。对其进行实验的 $F1$ 值分别提高了 10% 和 6.8%, 说明该模型对于端到端事件和时间关系抽取是有效的。

在事件抽取的过程中, 总是有数据不平衡的问题出现, 而且训练数据稀少也会影响模型的训练效果。针对这些问题, Zhang 等人^[68]提出了一种基于迁移学习的神经网络框架 (JointTransition), 采用从左到右的递增阅读顺序捕获实体和事件提及的依赖结构。在 ACE 2005 数据集上的实验表明, 事件触发词分类的任务中 $F1$ 值达到了 73.8%, 证明了该方法的有效性。Lu 等人^[69]提出了一种基于蒸馏学习和知识泛化的 Δ 表示学习方法。实验结果表明, 在 ACE 2005 数据集上 $F1$ 值提高了 0.7%, 在 TAC-KBP 2017 数据集上 $F1$ 值提高了 1.53%。Deng 等人^[70]提出了一种基于动态内存的原型网络 (DMB-PN), 包括了事件触发词识别和 few-shot 事件分类两个阶段。他们还定义“few-shot 事件检测”新问题, 也创建了新的数据集 FewEvent。实验结果表明, DMB-PN 不仅比其他基准模型更好地处理

数据稀缺的问题,而且在事件类型多样和数据极少的情况下模型的性能较好. Deng 等人^[71]在 2021 年提出了一种基于本体嵌入的 ED 模型 (OntoED),同时构建了一个新的数据集 OntoEvent. 实验结果表明,在事件角色识别和事件类型分类的任务中 $F1$ 值比 JMEE 模型分别提高了 15.32% 和 6.85%,证明了其方法鲁棒性较好.

另外,一些工作还通过结合预训练模型提高事件抽取的性能. Yang 等人^[72]针对手工创建的数据费力且数量有限的问题,提出了一种基于语言生成预训练的事件抽取模型 (PLMEE). 他们为了解决训练数据不足的问题,采用原型网络自动生成标注数据. 在 ACE 2005 数据集上的实验表明,事件类型分类和论元分类的任务中 $F1$ 值分别为 81.1% 和 58.9%. 2020 年, Du 等人^[73]则针对错误传播的问题,提出了一种基于问答任务的事件抽取模型. 他们在预训练 BERT 的基础上将事件抽取转换为问答任务,并以此为模型在事件触发词识别和论元分类任务上的 $F1$ 值分别提高了 0.39% 和 0.81%.

Gangal 等人^[74]针对 RAMS 数据集在事件论元抽取中准确率存在较大差距的问题,提出了一种基于预训练 BERT 的事件论元抽取方法 (BERTing RAMS). 实验结果表明,该方法具有较好的跨句准确性. Zhang 等人^[75]为了减少候选论元数量的问题,提出了一种两步隐式事件论元检测方法,将问题分解为两个子问题:论元头词检测和头跨度扩展,其中编码模块采用了预训练 BERT 进行上下文编码. 在 RAMS 数据集上的实验结果表明,该模型获比其他对比模型的性能更好.

在生物医学上识别触发词相关联的嵌套结构化事件时准确率不高, Huang 等人^[76]提出了一种基于层次知识图的生物医学事件抽取方法 (GEANet-SciBERT). 该方法在预训练语言模型 SciBERT^[77]的基础上,加入了一种新的图神经网络模型 GEANet 作为补充. 在 BioNLP 2011 GENIA 事件抽取任务中,该方法在所有事件和复杂事件上 $F1$ 值分别提高 1.41% 和 3.19%.

2.4 对中文事件抽取方法的研究

中文事件抽取不仅存在比英文更加严重的数据稀缺问题,而且也存在方法层面和语言特性层面的问题. 中文语言词语间还没有显式间隔,在进行分词时会出现比英文更加明显的错误和误差. 中文语言的复杂性和灵活性让相同语义的词语、短语和句子有更多表达方式,即同一类型事件触发词可以使用更多词语表达.

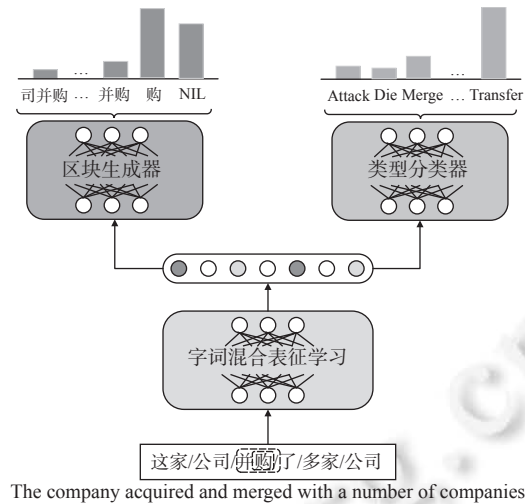
根据现有的中文事件抽取相关研究,从文本粒度方面可以将其分为两类:句子级事件抽取和篇章级事件抽取. 其中,句子级事件抽取主要是利用句子内部获取特征,即识别句子的中文触发词,并判断实体在事件中所扮演的角色. 而篇章级事件抽取主要包含了跨句、跨文档抽取的特征信息.

2.4.1 基于句子级的中文事件抽取方法

句子级事件抽取主要是识别句子中每个词可能提及的单个事件,以及判断句子实体在该事件中扮演的角色.

以往的研究中,一直存在单词和事件触发词不匹配的问题,2018 年 Lin 等人^[78]提出了一种基于区块提议网络的模型 (NPNs),如图 4 所示. NPNs 是以字符方式的范例进行事件检测,首先学习每个字符的混合表示,在字符和单词中获取结构信息和语义信息;然后用基于表示学习的方法进行中文事件触发词识别同时进行事件分类. 在 ACE 2005 和 TAC-KBP 2017 数据集上的实验表明,触发检测上 $F1$ 值分别提高了 2.3% 和 3.43%,类型识别上 $F1$ 值分别提高了 1.6% 和 1.51%. 为了解决论元角色重叠的问题,2020 年 Xu 等人^[79]在提出了一种中文联合事件抽取框架 (JMCEE). 该框架主要是从预训练的 BERT 语言模型共享特征表示的事件触发和论元特征,然后联合进行预测. 在中文 ACE 2005 数据集上的实验表明,与 NPNs 相比,在事件触发识别和分类中 $F1$ 值分别提高了 14.5% 和 8.1%.

Liu 等人^[80]提出了一种基于触发检测动态记忆网络 (TD-DMN). 他们利用了动态记忆网络的多跳机制更改捕获的上下文信息,将事件触发检测转换为一个问答问题. 通过在 ACE 2005 数据集上进行实验证明,多跳机制的引入可以使模型性能提高. 在中文事件触发词抽取中存在语义信息和部分事件实例无法有效抽取的问题,李培峰等人^[81]提出了一种基于语义的中文事件触发词抽取联合学习模型. 在 ACE 2005 中文数据集进行的实验表明,该方法模型的性能效果较好. Feng 等人^[18]采用 CNN 与 Bi-LSTM 结合的混合神经网络,捕获特定上下文的序列信息和短语信息. 他们的方法不需要特征工程参与,是一种与语言无关的模型. 实验表明,在英文事件识别中 $F1$ 值达到了 73.4%,在中文和西班牙语中平均提高了 3.0%. Zeng 等人^[82]提出了一个基于卷积 Bi-LSTM 的中文事件抽取模型,将 LSTM 和 CNN 结合起来捕获词级和句子级信息.

图 4 NPNs 总体框架^[78]

事件检测一般是在原始文本中定位触发词后,将其分类到正确的事件类型。但是,涉及到没有自然分隔符的语言(如中文)时,会出现触发词不匹配和定位不准确的问题, Ding 等人^[24]提出了一种基于触发感知的 Lattice 神经网络模型(TLNN)。实验结果表明,在 ACE 2005 和 KBP 2017 数据集上均优于其他深度学习模型,证明了该模型可以提高触发词定位的准确性和事件分类的准确性。2019 年, Subburathinam 等人^[83]提出了一种跨语言迁移学习的关系与事件抽取模型。他们利用跨语言关系和事件相关语言通用性,结合图卷积网络将所有事件实体、触发词和上下文表示为复杂结构化的多语言公共空间。该方法在英文、中文和阿拉伯语中的实验表明,其在事件论元角色识别上 $F1$ 值达到了 63.1%,验证了跨语言迁移学习方法的可行性。

2.4.2 基于篇章级的中文事件抽取方法

篇章级事件抽取以文本中描述的主要事件为中心,用简洁、结构化的形式呈现给用户。篇章级事件抽取方法还额外考虑了跨句子或跨文档的特征辅助任务实现。例如,相同/类似事件对应的论元一致性,同一文档内的相同单词触发的事件类型一致性特征,相同类型的实体参与的事件一致性特征。

2018 年, Yang 等人^[84]针对现有研究中数据稀缺的问题,以及事件抽取都是从一个句子中抽取事件的情形,提出了一个篇章级中文金融事件抽取系统(DCFEE),能够自动生成大规模的标签数据,同时可以从整个文档中抽取事件。实验表明,该方法不需要人工标注数据,可以自动从文档中抽取中文金融事件。2019 年,仲伟峰等人^[85]提出了一种基于联合标注和全局推理的篇章级事件抽取方法,主要是利用注意力机制的序列标注模型联合抽取句子级事件的触发词和实体,然后利用多层感知机判断实体在事件中扮演的角色。在 ACE2005 数据集上,验证了模型的有效性。2020 年,朱培培等人^[86]针对中文事件检测方法获得的局部上下文的信息不足以解决事件触发词的歧义性的问题,提出了一种基于篇章信息和 Bi-GRU 的中文事件检测方法。在 ACE2005 数据集上的实验结果表明,该方法比当前最好的模型的 $F1$ 值提高了 1.5%。

2.5 数据集

事件抽取中常用的数据集主要包括 ACE 2005、TAC-KBP、BioNLP-ST、GENIA、MLEE 和 CEC^[87],其中 ACE 2005 是应用最广泛的数据集。

ACE 数据集是语言数据联盟(linguistic data consortium, LDC)发布的,由实体、关系和事件注释组成的各种类型的数据^[2]。ACE 2005 事件数据集定义了 8 个事件类型和 33 个子类型,每个事件子类型对应一组论元角色,是来自媒体、广播新闻等不同方面的英文、中文和阿拉伯语事件。表 5 描述了相关的事件类型和其子类型。

表 5 ACE 2005 中的事件类型和子类型

编号	事件类型	编号	事件子类型
1	Life	1-5	Be-Born, Marry, Divorce, Injure, Die
2	Movement	6	Transport
3	Contact	7-8	Meet, Phone-Write
4	Conflict	9-10	Attack, Demonstrate
5	Business	11-14	Merge-org, Declare-Bankruptcy, Start-Org, End-Org
6	Transaction	15-16	Transfer-Money, Transfer-Ownership
7	Personnel	17-20	Elect, Start-Position, End-Position, Nominate
8	Justice	21-33	Arrest-Jail, Release-Parole, Trial-Hearing, Charge-Indict, Execute, Pardon, Fine, Convict, Acquit, Sentence, Sue, Extradite, appeal

TAC-KBP 数据集已在 2015 年 TAC 会议事件检测评估中发布^[5]. TAC-KBP 中的事件类型和子类型是根据 ACE 语料库定义的, 包括 9 个事件类型和 38 个子类型. TAC-KBP 2015 为英文语料库, 但 TAC-KBP 2016 为所有任务添加了中文和西班牙语数据.

生物学领域的事件抽取数据集, 常见的是 BioNLP-ST、GENIA、MLEE 和 PC 等. BioNLP-ST 是从生物学领域科学文献的生物分子事件抽取中获得的, 包括 GE、CG、PC、GRO、GRN 和 BB^[88]等任务. GENIA 事件数据集是为 GENIA 项目编写并标注的生物学文献集合事件. 而 MLEE 数据集是分子到器官系统的多个生物组织水平的事件. PC 数据集则是与生物分子途径模型有关的反应事件.

除了上述常见的数据集外, 还有一些其他领域的数据集, 如中文突发事件语料库 CEC (Chinese emergency corpus)^[87]是由上海大学语义智能实验室所构建, 选取了地震、火灾、交通事故、恐怖袭击和食物中毒这 5 类突发事件的新闻报道进行标注, 合计有 332 篇. 与 ACE 数据集相比, CEC 较小, 但它在事件和事件论元的注释方面更全面. TERQAS 研讨会建立了一个名为 TimeBANK 数据集, 主要用于突发新闻事件抽取^[89]. 还有军事情报领域的 MUC 数据集^[90]、丁效等人的音乐领域事件抽取数据集^[91], 以及 Yang 的中文金融事件抽取数据集^[92,93].

2.6 模型性能及评价

事件抽取常采用准确率 (precision, P)、召回率 (recall, R) 和 $F1$ 值 ($F1$ -measure, $F1$) 这 3 项作为基本评价指标. 其中, P 是指系统中抽取出的正确个数占抽取总数的比例, 用来衡量抽取准确程度; R 是指系统中正确抽取的个数占所有正确总数的比例, 用来衡量抽取全面程度; $F1$ 值是 P 和 R 的加权平均值, 作为系统性能的总体评价. 事件抽取所采用的这 3 个评价指标的具体公式如下:

$$P = \frac{TP}{TP + FP},$$

$$R = \frac{TP}{TP + FN},$$

$$F1 = \frac{2 \times P \times R}{P + R}.$$

事件抽取模型的性能通常通过 $F1$ 值来综合判断, $F1$ 值越大, 模型抽取的效果越好. 表 6 对比了深度学习模型在 ACE 2005 英文数据集实验的 $F1$ 值.

- 触发词检测 (trigger detection): 如果触发偏移量正确 (即, 触发词在文本中相应位置), 则可以正确检测到该触发词的正确匹配.
- 类型识别 (type identification): 如果触发偏移量和事件类型都与参考触发及其事件类型匹配, 则可以正确识别事件类型.
- 论元检测 (argument detection): 如果论元的偏移量与任何参考论元匹配, 则可以正确检测到论元提及 (即正确识别事件中的论元).
- 角色识别 (role identification): 论元角色的事件类型, 偏移量和角色与任何参考论元提及匹配时, 将正确识别

论元角色.

从表 6 中 $F1$ 值的结果, 我们可以注意到并非所有模型都对触发检测、类型识别、论元检测和角色识别这 4 个任务进行了实验, 因为其中一些模型仅需要实现某些子任务. 表中对深度学习的方法进行实验得到的最高的 $F1$ 值为 GAN 中的触发检测 77.0% 和 GCN 中的类型识别 77.6%、论元检测 68.4%、角色识别 60.3%, 均高于 CNN 与 RNN 的深度学习方法中的 76.8%、74.2%、68.4% 和 59.3%. 这一结果验证了随着深度学习技术的发展, 其对事件抽取性能的提升效果较好. 由 CNN 和 RNN 相关模型的实验结果对比可知, RNN 中对于 LSTM、GRU 等方法的使用以及结合 Attention 机制效果明显比传统的 CNN 模型更好, 特别是事件类型识别和论元检测任务的提升更明显. 由 CNN、RNN、GCN、GAN 和其他深度学习方法的实验结果对比可知, 在事件抽取任务中引入生成对抗网络和图卷积网络模型的性能提升更高. 由此可知, 表 6 中的数据说明模型的性能不仅受事件抽取任务的影响, 还取决于所使用的深度学习方法, 并且采用不同方法进行实验得到的事件抽取性能往往有较大差距.

表 6 ACE 2005 英文数据集上的模型性能 (%)

方法	模型	评测值 $F1$			
		Trigger	Even type	Argument	Argument role
CNN	DMCNN ^[14]	73.5	69.1	59.1	53.5
	HMEAE ^[17]	N/A	N/A	N/A	59.3
	S-CNNs ^[92]	74.8	69.1	58.6	53.1
RNN	JRNN ^[19]	71.9	69.3	62.8	55.4
	BDLSTM-TNNs ^[20]	72.2	68.9	60.0	54.1
	DLRNN ^[42]	N/A	70.5	N/A	N/A
	dbRNN ^[21]	N/A	71.9	67.7	58.7
	DEEB-RNN ^[48]	N/A	74.0	N/A	N/A
	Re+Bi-LSTM+Att1 ^[46]	76.1	73.9	N/A	N/A
	RNN-BiGRU ^[45]	74.9	71.2	64.8	56.6
	GMLATT ^[47]	74.1	72.4	N/A	N/A
	HBTNGMA ^[23]	N/A	73.3	N/A	N/A
	EMD-Joint3EE ^[43]	72.5	69.8	59.9	52.1
GAN	Bi-LSTM+multi-head att ^[49]	76.8	74.2	N/A	N/A
GCN	SELF Bi-LSTM+GAN ^[25]	77.0	73.0	N/A	N/A
	JMEE ^[54]	75.9	73.7	68.4	60.3
	MOGANED ^[28]	N/A	75.7	N/A	N/A
其他	RA-GCN ^[55]	N/A	77.6	N/A	N/A
	JOINTEVENTENTITY ^[66]	71.0	68.7	50.6	48.4
	BERT_QA ^[73]	75.82	72.39	55.29	53.31

接下来, 对中文事件抽取在 ACE 2005 数据集上的实验结果, 也进行了对比, 具体如表 7 所示.

表 7 ACE 2005 数据集在中文事件抽取模型的对比 (%)

模型	Trigger			Even type		
	P	R	$F1$	P	R	$F1$
Rich-C ^[94]	62.2	71.9	66.7	58.9	68.1	63.2
C-BiLSTM ^[86]	68.1	69.2	68.7	61.6	64.7	63.1
NPNs ^[78]	64.8	73.8	69.0	60.9	69.3	64.8
TLNN ^[82]	67.34	74.68	70.82	64.45	71.47	67.78
Bi-GRU-doc ^[24]	81.3	72.1	76.4	76.3	67.7	71.7

在表 7 中, 中文事件抽取模型主要是围绕 Chen 等人^[94]提出的 Rich-C 作为 baseline 展开的, 其中 Rich-C 是通过联合学习事件触发特征和事件类型一致性特征来进行抽取的. 从表 7 的实验结果可以得出, 在事件触发词检测和事件类型识别中 Bi-GRU-doc 模型的 $F1$ 值分别是 76.4% 和 71.7%, 比 Rich-C 的高了 9.7% 和 8.5%, 从这里可以看出基于篇章级的神经网络模型, 利用句子上下文信息对中文事件抽取是有效的. TLNN 模型的 R 值也比 Rich-C 模型分别高了 2.78% 和 3.37%, 证明了该模型可以提高触发词定位的准确性和事件分类的准确性. C-BiLSTM 和 NPNs 在事件触发词和事件类型识别的 P 、 R 、 $F1$ 都高于 Rich-C 的, 这说明了深度学习模型在中文事件抽取上是可行的. 由此可知, 表 7 中的数据说明了中文事件模型的性能不仅受词语的语义影响, 还取决于对深度学习方法的运用, 基于中文篇章级和字符信息融合的方法, 有利于解决中文事件触发词和类型稀疏的问题.

3 事件抽取面临的挑战及研究趋势

事件抽取作为信息抽取中的重要任务之一, 能够检测句子提到的某些事件, 也可以对事件类型进行分类与识别事件论元. 它为知识图谱、推荐系统、信息检索等任务提供基础的数据支持, 同时, 事件抽取在语义分析、篇章理解、自动问答等领域也具有重要意义. 然而, 人工智能的发展对事件抽取的准确性提出了更高的要求, 深度学习虽已成为事件抽取的关键技术, 但在领域自适应性和召回率方面仍有很大的提升空间.

(1) 结合深度学习技术进展的事件抽取研究. 事件抽取是在事件触发和论元识别的基础上发展起来的. 它在某种程度上取决于事件类型、触发词识别和论元检测的效果, 但是这些基础技术准确率不高. 在深度学习技术被大规模使用后, 事件抽取的效果得到了很大的提升. 2015 年至今, 该领域的研究热点集中于 CNN、RNN、GAN、GCN 以及 Attention 机制、少样本学习、预训练技术等方法的研究, 还有联合多种网络来进行事件抽取. 在事件抽取中, 如何更有效地结合最新的深度学习技术进行抽取的研究是未来的一个重要趋势.

(2) 段落级和篇章级事件抽取的研究. 近几年的事件抽取工作主要聚集在利用深度学习方法进行句子级事件抽取, 而实际应用时同一事件经常出现在不同句子中, 这时需要通过整个篇章确定事件的具体情况. 同时, 还存在抽取的事件信息不完整的问题, 在日常应用中, 我们经常要了解事件发生的全过程. 现有基于深度学习的跨句子级事件抽取模型性能较差, 但是段落甚至篇章级的事件抽取要求深度模型具有更复杂的推理能力、更高的准确率以及更好的灵活性. 未来利用深度学习技术融合多个句子进行段落和篇章级事件抽取是一个重要的研究方向.

(3) 面向特定领域事件抽取系统的设计与研究. 基于特征或传统机器学习的事件抽取方法, 已经覆盖了多数可能的输入和特征, 而基于深度学习的模型往往依赖网络的复杂程度带来对隐含事件信息挖掘性能的提升. 面向特定领域事件抽取系统的领域与深度学习技术更好地融合, 进行可扩展性与可移植性的进一步提升是将来的关键研究内容.

(4) 跨语言、跨领域的事件抽取的研究. 目前, 事件抽取的水平还限制在对独立语言、单一领域的事件文本处理上, 跨语言、跨领域的研究尚处于探索阶段. 例如, 在中文事件抽取的相关研究主要集中在生物医学、微博、新闻、紧急情况等方面, 其他领域和跨语言事件抽取技术的研究很少. 随着深度学习技术的进一步完善, 跨语言、跨领域的事件抽取必将成为研究热点.

(5) 事件抽取中的深度迁移学习的研究. 在事件抽取中, 由于触发词特征和数据集注释的不同, 可能无法在其他文本上很好地进行研究. 尽管有一些迁移学习技术已应用在事件抽取的研究中, 但涉及比较少. 针对深度迁移学习技术的进一步研究, 有利于我们开发一个健壮的识别器来识别不同领域的事件类型, 探索事件抽取任务中的少样本、零样本学习, 解决领域不匹配和跨域不匹配的问题.

(6) 基于远程监督的事件抽取的深入研究. 为了缓解远程监督中经常会出现的错误标签问题, 研究者们分别结合多示例学习、注意力机制、噪声建模等方法提出了多种模型, 但如何建立更有效的方法缓解远程监督中错误标签的影响仍是事件抽取中研究的重点问题.

4 结论及展望

在本文中, 对深度学习中事件抽取技术进行了全面最新的概述. 事件抽取发展至今, 在研究内容上逐渐由单一

领域向多领域、跨领域发展,事件类型的定义方式表示为由人工预先定义转变为事件类型自动发现、挖掘;在研究方法上,深度学习的方法在事件抽取任务的性能上表现良好,基于深度学习的框架日益成为主流,在此基础上结合远程监督、强化学习、少样本学习和零资源学习的思想等可以为事件抽取性能的提升提供新思路.特别是,中文事件抽取的任务复杂,将深度学习的方法应用在其中,是今后需要努力的一个方向.

References:

- [1] Grishman R, Sundheim B. Message understanding conference-6: A brief history. In: Proc. of the 16th Conf. on Computational Linguistics. Copenhagen: Association for Computational Linguistics, 1996. 466–471. [doi: [10.3115/992628.992709](https://doi.org/10.3115/992628.992709)]
- [2] Missingham R. Access to Australian Government information: A decade of change 1997–2007. *Government Information Quarterly*, 2008, 25(1): 25–37. [doi: [10.1016/j.giq.2007.07.001](https://doi.org/10.1016/j.giq.2007.07.001)]
- [3] Ahn D. The stages of event extraction. In: Proc. of the 2006 Workshop on Annotating and Reasoning about Time and Events. Sydney: Association for Computational Linguistics, 2006. 1–8.
- [4] Malakasiotis P. AUEB at TAC 2009. In: Proc. of the 2nd Text Analysis Conf. Gaithersburg: NIST, 2009. 1–6.
- [5] Mitamura T, Liu ZZ, Hovy EH. Overview of TAC KBP 2015 event nugget track. In: Proc. of the 2015 Text Analysis Conf. Gaithersburg: NIST, 2015. 1–11.
- [6] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: Arbib MA, ed. *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1995. 255–258. [doi: [10.5555/303568.303704](https://doi.org/10.5555/303568.303704)]
- [7] Kombrink S, Mikolov T, Karafiát M, Burget L. Recurrent neural network based language modeling in meeting recognition. In: Proc. of the 12th Annual Conf. of the Int'l Speech Communication Association. Florence: ISCA, 2011. 2877–2880.
- [8] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proc. of the 27th Int'l Conf. on Neural Information Processing Systems. Montreal: MIT Press, 2014. 2672–2680. [doi: [10.5555/2969033.2969125](https://doi.org/10.5555/2969033.2969125)]
- [9] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: Proc. of the 5th Int'l Conf. on Learning Representations. Toulon: OpenReview.net, 2017. 1–14.
- [10] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: Proc. of the 3rd Int'l Conf. on Learning Representations. San Diego: ICLR, 2015. 1–15.
- [11] Zeng DJ, Liu K, Lai SW, Zhou GY, Zhao J. Relation classification via convolutional deep neural network. In: Proc. of the 25th Int'l Conf. on Computational Linguistics. Dublin: Dublin City University and Association for Computational Linguistics, 2014. 2335–2344.
- [12] Tang DY, Qin B, Liu T. Document modeling with gated recurrent neural network for sentiment classification. In: Proc. of the 2015 Conf. on Empirical Methods in Natural Language Processing. Lisbon: Association for Computational Linguistics, 2015. 1422–1432. [doi: [10.18653/v1/D15-1167](https://doi.org/10.18653/v1/D15-1167)]
- [13] Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Trans. on Signal Processing*, 1997, 45(11): 2673–2681. [doi: [10.1109/78.650093](https://doi.org/10.1109/78.650093)]
- [14] Chen YB, Xu LH, Liu K, Zeng DJ, Zhao J. Event extraction via dynamic multi-pooling convolutional neural networks. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing. Beijing: Association for Computational Linguistics, 2015. 167–176. [doi: [10.3115/v1/P15-1017](https://doi.org/10.3115/v1/P15-1017)]
- [15] Nguyen TH, Grishman R. Event detection and domain adaptation with convolutional neural networks. In: Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing. Beijing: Association for Computational Linguistics, 2015. 365–371. [doi: [10.3115/v1/P15-2060](https://doi.org/10.3115/v1/P15-2060)]
- [16] Nguyen TH, Grishman R. Modeling skip-grams for event detection with convolutional neural networks. In: Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing. Austin: Association for Computational Linguistics, 2016. 886–891. [doi: [10.18653/v1/D16-1085](https://doi.org/10.18653/v1/D16-1085)]
- [17] Wang XZ, Wang ZQ, Han X, Liu ZY, Li JZ, Li P, Sun MS, Zhou J, Ren X. HMEAE: Hierarchical modular event argument extraction. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 5777–5783. [doi: [10.18653/v1/D19-1584](https://doi.org/10.18653/v1/D19-1584)]
- [18] Feng XC, Huang LF, Tang DY, Ji H, Qin B, Liu T. A language-independent neural network for event detection. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin: Association for Computational Linguistics, 2016. 66–71. [doi: [10.18653/v1/P16-2011](https://doi.org/10.18653/v1/P16-2011)]
- [19] Nguyen TH, Cho K, Grishman R. Joint event extraction via recurrent neural networks. In: Proc. of the 2016 Conf. of the North American

- Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego: Association for Computational Linguistics, 2016. 300–309. [doi: [10.18653/v1/N16-1034](https://doi.org/10.18653/v1/N16-1034)]
- [20] Chen YB, Liu SL, He SZ, Liu K, Zhao J. Event extraction via bidirectional long short-term memory tensor neural networks. In: Proc. of the 15th China National Conf. on Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Yantai: Springer, 2016. 190–203. [doi: [10.1007/978-3-319-47674-2_17](https://doi.org/10.1007/978-3-319-47674-2_17)]
- [21] Sha L, Qian F, Chang BB, Sui ZF. Jointly extracting event triggers and arguments by dependency-bridge RNN and tensor-based argument interaction. In: Proc. of the 32nd AAAI Conf. on Artificial Intelligence (AAAI2018), the 30th Innovative Applications of Artificial Intelligence (IAAI2018), and the 8th AAAI Symp. on Educational Advances in Artificial Intelligence (EAAI2018). New Orleans: AAAI, 2018. 5916–5923.
- [22] Orr JW, Tadepalli P, Fern X. Event detection with neural networks: A rigorous empirical evaluation. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018. 999–1004. [doi: [10.18653/v1/D18-1122](https://doi.org/10.18653/v1/D18-1122)]
- [23] Chen YB, Yang H, Liu K, Zhao J, Jia YT. Collective event detection via a hierarchical and bias tagging networks with gated multi-level attention mechanisms. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018. 1267–1276. [doi: [10.18653/v1/D18-1158](https://doi.org/10.18653/v1/D18-1158)]
- [24] Ding N, Li ZR, Liu ZY, Zheng HT, Lin ZB. Event detection with trigger-aware lattice neural network. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 347–356. [doi: [10.18653/v1/D19-1033](https://doi.org/10.18653/v1/D19-1033)]
- [25] Hong Y, Zhou WX, Zhang JL, Zhou GD, Zhu QM. Self-regulation: Employing a generative adversarial network to improve event detection. In: Proc. of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics, 2018. 515–526. [doi: [10.18653/v1/P18-1048](https://doi.org/10.18653/v1/P18-1048)]
- [26] Wang R, Zhou DY, He YL. Open event extraction from online text using a generative adversarial network. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 282–291. [doi: [10.18653/v1/D19-1027](https://doi.org/10.18653/v1/D19-1027)]
- [27] Nguyen TH, Grishman R. Graph convolutional networks with argument-aware pooling for event detection. In: Proc. of the 32nd AAAI Conf. on Artificial Intelligence and the 30th Innovative Applications of Artificial Intelligence Conf. and the 8th AAAI Symp. on Educational Advances in Artificial Intelligence. New Orleans: AAAI, 2018. 5900–5907. [doi: [10.5555/3504035.3504759](https://doi.org/10.5555/3504035.3504759)]
- [28] Yan HR, Jin XL, Meng XB, Guo JF, Cheng XQ. Event detection with multi-order graph convolution and aggregated attention. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 5766–5770. [doi: [10.18653/v1/D19-1582](https://doi.org/10.18653/v1/D19-1582)]
- [29] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa R. Natural language processing (almost) from scratch. *The Journal of Machine Learning Research*, 2011, 12: 2493–2537. [doi: [10.5555/1953048.2078186](https://doi.org/10.5555/1953048.2078186)]
- [30] Li Q, Ji H, Huang L. Joint event extraction via structured prediction with global features. In: Proc. of the 51st Annual Meeting of the Association for Computational Linguistics. Sofia: Association for Computational Linguistics, 2013. 73–82.
- [31] Lei T, Barzilay R, Jaakkola T. Molding CNNs for text: Non-linear, non-consecutive convolutions. In: Proc. of the 2015 Conf. on Empirical Methods in Natural Language Processing. Lisbon: Association for Computational Linguistics, 2015. 1565–1575. [doi: [10.18653/v1/D15-1180](https://doi.org/10.18653/v1/D15-1180)]
- [32] Zhang TT, Whitehead S, Zhang HW, Li HZ, Ellis J, Huang LF, Liu W, Ji H, Chang SF. Improving event extraction via multimodal integration. In: Proc. of the 25th ACM Int'l Conf. on Multimedia. Mountain View: ACM, 2017. 270–278. [doi: [10.1145/3123266.3123294](https://doi.org/10.1145/3123266.3123294)]
- [33] Liu WY, Nguyen TH. Similar but not the same: Word sense disambiguation improves event detection via neural representation matching. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018. 4822–4828. [doi: [10.18653/v1/D18-1517](https://doi.org/10.18653/v1/D18-1517)]
- [34] Lin HY, Lu YJ, Han XP, Sun L. Cost-sensitive regularization for label confusion-aware event detection. In: Proc. of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: Association for Computational Linguistics, 2019. 5278–5283. [doi: [10.18653/v1/P19-1521](https://doi.org/10.18653/v1/P19-1521)]
- [35] Li LS, Liu Y, Qin MY. Extracting biomedical events with parallel multi-pooling convolutional neural networks. *IEEE/ACM Trans. on Computational Biology and Bioinformatics*, 2020, 17(2): 599–607. [doi: [10.1109/TCBB.2018.2868078](https://doi.org/10.1109/TCBB.2018.2868078)]
- [36] Pyysalo S, Ohta T, Miwa M, Cho HC, Tsujii J, Ananiadou S. Event extraction across multiple levels of biological organization. *Bioinformatics*, 2012, 28(18): i575–i581. [doi: [10.1093/bioinformatics/bts407](https://doi.org/10.1093/bioinformatics/bts407)]

- [37] Zhu LX, Zheng HR. Biomedical event extraction with a novel combination strategy based on hybrid deep neural networks. *BMC Bioinformatics*, 2020, 21(1): 47. [doi: [10.1186/s12859-020-3376-2](https://doi.org/10.1186/s12859-020-3376-2)]
- [38] Ramponi A, Plank B, Lombardo R. Cross-domain evaluation of edge detection for biomedical event extraction. In: *Proc. of the 12th Language Resources and Evaluation Conf. Marseille: European Language Resources Association*, 2020. 1982–1989.
- [39] Goller C, Kuchler A. Learning task-dependent distributed representations by backpropagation through structure. In: *Proc. of the 1996 Int'l Conf. on Neural Networks. Washington: IEEE*, 1996. 347–352. [doi: [10.1109/ICNN.1996.548916](https://doi.org/10.1109/ICNN.1996.548916)]
- [40] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997, 9(8): 1735–1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)]
- [41] Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics*, 2014. 1724–1734. [doi: [10.3115/v1/D14-1179](https://doi.org/10.3115/v1/D14-1179)]
- [42] Duan SY, He RF, Zhao WL. Exploiting document level information to improve event detection via recurrent neural networks. In: *Proc. of the 8th Int'l Joint Conf. on Natural Language Processing. Taipei: Asian Federation of Natural Language Processing*, 2017. 352–361.
- [43] Nguyen TM, Nguyen TH. One for all: Neural joint modeling of entities and events. *Proc. of the AAAI Conf. on Artificial Intelligence*, 2019, 33(1): 6851–6858. [doi: [10.1609/aaai.v33i01.33016851](https://doi.org/10.1609/aaai.v33i01.33016851)]
- [44] Wu WT, Zhu XX, Tao JM, Li PF. Event detection via recurrent neural network and argument prediction. In: *Proc. of the 7th CCF Int'l Conf. on Natural Language Processing and Chinese Computing. Hohhot: Springer*, 2018. 235–245. [doi: [10.1007/978-3-319-99501-4_20](https://doi.org/10.1007/978-3-319-99501-4_20)]
- [45] Ding RX, Li ZJ. Event extraction with deep contextualized word representation and multi-attention layer. In: *Proc. of the 14th Int'l Conf. on Advanced Data Mining and Applications. Nanjing: Springer*, 2018. 189–201. [doi: [10.1007/978-3-030-05090-0_17](https://doi.org/10.1007/978-3-030-05090-0_17)]
- [46] Zhang JL, Zhou Wx, Hong Y, Yao JM, Zhang M. Using entity relation to improve event detection via attention mechanism. In: *Proc. of the 7th CCF Int'l Conf. on Natural Language Processing and Chinese Computing. Hohhot: Springer*, 2018. 171–183. [doi: [10.1007/978-3-319-99495-6_15](https://doi.org/10.1007/978-3-319-99495-6_15)]
- [47] Liu W, Yang ZY, Liu ZT. Chinese event recognition via ensemble model. In: *Proc. of the 25th Int'l Conf. on Neural Information Processing. Siem Reap: Springer*, 2018. 255–264. [doi: [10.1007/978-3-030-04221-9_23](https://doi.org/10.1007/978-3-030-04221-9_23)]
- [48] Zhao Y, Jin XL, Wang YZ, Cheng XQ. Document embedding enhanced event detection with hierarchical and supervised attention. In: *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics*, 2018. 414–419. [doi: [10.18653/v1/P18-2066](https://doi.org/10.18653/v1/P18-2066)]
- [49] Chen JL, Hong Y, Zhang JL, Yao JM. Using mention segmentation to improve event detection with multi-head attention. In: *Proc. of the 2019 Int'l Conf. on Asian Language Processing. Shanghai: IEEE*, 2019. 367–372. [doi: [10.1109/IALP48816.2019.9037683](https://doi.org/10.1109/IALP48816.2019.9037683)]
- [50] Mehta S, Islam MR, Rangwala H, Ramakrishnan N. Event detection using hierarchical multi-aspect attention. In: *Proc. of the 2019 World Wide Web Conf. San Francisco: ACM*, 2019. 3079–3085. [doi: [10.1145/3308558.3313659](https://doi.org/10.1145/3308558.3313659)]
- [51] Huang XF. Event extraction based on dynamic masked attention. *Application Research of Computers*, 2020, 37(7): 1964–1968 (in Chinese with English abstract). [doi: [10.19734/j.issn.1001-3695.2018.12.0927](https://doi.org/10.19734/j.issn.1001-3695.2018.12.0927)]
- [52] Zhang TT, Ji H. Event extraction with generative adversarial imitation learning. arXiv:1804.07881, 2018.
- [53] Liu J, Chen YB, Liu K. Exploiting the ground-truth: An adversarial imitation based knowledge distillation approach for event detection. *Proc. of the 2019 AAAI Conf. on Artificial Intelligence*, 2019, 33(1): 6754–6761. [doi: [10.1609/aaai.v33i01.33016754](https://doi.org/10.1609/aaai.v33i01.33016754)]
- [54] Liu X, Luo ZC, Huang HY. Jointly multiple events extraction via attention-based graph information aggregation. In: *Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics*, 2018. 1247–1256. [doi: [10.18653/v1/D18-1156](https://doi.org/10.18653/v1/D18-1156)]
- [55] Cui SY, Yu BW, Liu TW, Zhang ZY, Wang XB, Shi JQ. Edge-enhanced graph convolution networks for event detection with syntactic relation. arXiv:2002.10757, 2020.
- [56] Zeng DJ, Liu K, Chen YB, Zhao J. Distant supervision for relation extraction via piecewise convolutional neural networks. In: *Proc. of the 2015 Conf. on Empirical Methods in Natural Language Processing. Lisbon: Association for Computational Linguistics*, 2015. 1753–1762. [doi: [10.18653/v1/D15-1203](https://doi.org/10.18653/v1/D15-1203)]
- [57] Chen YB, Liu SL, Zhang X, Liu K, Zhao J. Automatically labeled data generation for large scale event extraction. In: *Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver: Association for Computational Linguistics*, 2017. 409–419. [doi: [10.18653/v1/P17-1038](https://doi.org/10.18653/v1/P17-1038)]
- [58] Bollacker K, Evans C, Paritosh P, Sturge T, Taylor J. Freebase: A collaboratively created graph database for structuring human knowledge. In: *Proc. of the 2008 ACM SIGMOD Int'l Conf. on Management of Data. Vancouver: ACM*, 2008. 1247–1250. [doi: [10.1145/1376616.1376746](https://doi.org/10.1145/1376616.1376746)]
- [59] Baker CF, Fillmore CJ, Lowe JB. The Berkeley FrameNet project. In: *Proc. of the 36th Annual Meeting of the Association for*

- Computational Linguistics and the 17th Int'l Conf. on Computational Linguistics. Montreal: Association for Computational Linguistics, 1998. 86–90. [doi: [10.3115/980845.980860](https://doi.org/10.3115/980845.980860)]
- [60] Zeng Y, Feng YS, Ma R, Wang Z, Yan R, Shi CD, Zhao DY. Scale up event extraction learning via automatic training data generation. In: Proc. of the 32nd AAAI Conf. on Artificial Intelligence and the 30th Innovative Applications of Artificial Intelligence Conf. and the 8th AAAI Symp. on Educational Advances in Artificial Intelligence. New Orleans: AAAI, 2018. 742. [doi: [10.5555/3504035.3504777](https://doi.org/10.5555/3504035.3504777)]
- [61] Keith KA, Handler A, Pinkham M, Magliozzi C, McDuffie J, O'Connor B. Identifying civilians killed by police with distantly supervised entity-event extraction. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. Copenhagen: Association for Computational Linguistics, 2017. 1547–1557. [doi: [10.18653/v1/D17-1163](https://doi.org/10.18653/v1/D17-1163)]
- [62] Rao S, Marcu D, Knight K, Daumé III H. Biomedical event extraction using abstract meaning representation. In: Proc. of the 2017 BioNLP. Vancouver: Association for Computational Linguistics, 2017. 126–135. [doi: [10.18653/v1/W17-2315](https://doi.org/10.18653/v1/W17-2315)]
- [63] Kim JD, Wang Y, Takagi T, Yonezawa A. Overview of genia event task in BioNLP shared task 2011. In: Proc. of the 2011 BioNLP Shared Task Workshop. Portland: Association for Computational Linguistics, 2011. 7–15.
- [64] Liu SL, Chen YB, He SZ, Liu K, Zhao J. Leveraging framenet to improve automatic event detection. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin: Association for Computational Linguistics, 2016. 2134–2143. [doi: [10.18653/v1/P16-1201](https://doi.org/10.18653/v1/P16-1201)]
- [65] Wadden D, Wennberg U, Luan Y, Hajishirzi H. Entity, relation, and event extraction with contextualized span representations. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 5784–5789. [doi: [10.18653/v1/D19-1585](https://doi.org/10.18653/v1/D19-1585)]
- [66] Yang BS, Mitchell TM. Joint extraction of events and entities within a document context. In: Proc. of the 2016 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego: Association for Computational Linguistics, 2016. 289–299. [doi: [10.18653/v1/N16-1033](https://doi.org/10.18653/v1/N16-1033)]
- [67] Han RJ, Ning Q, Peng NY. Joint event and temporal relation extraction with shared representations and structured prediction. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 434–444. [doi: [10.18653/v1/D19-1041](https://doi.org/10.18653/v1/D19-1041)]
- [68] Zhang JC, Qin YX, Zhang Y, Liu MC, Ji DH. Extracting entities and events as a single task using a transition-based neural model. In: Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence. Macao: IJCAI.org, 2019. 5422–5428. [doi: [10.24963/ijcai.2019/753](https://doi.org/10.24963/ijcai.2019/753)]
- [69] Lu YJ, Lin HY, Han XP, Sun L. Distilling discrimination and generalization knowledge for event detection via delta-representation learning. In: Proc. of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: Association for Computational Linguistics, 2019. 4366–4376. [doi: [10.18653/v1/P19-1429](https://doi.org/10.18653/v1/P19-1429)]
- [70] Deng SM, Zhang NY, Kang JJ, Zhang YC, Zhang W, Chen HJ. Meta-learning with dynamic-memory-based prototypical network for few-shot event detection. In: Proc. of the 13th Int'l Conf. on Web Search and Data Mining. Houston: ACM, 2020. 151–159. [doi: [10.1145/3336191.3371796](https://doi.org/10.1145/3336191.3371796)]
- [71] Deng SM, Zhang NY, Li LQ, Chen H, Tou HX, Chen MS, Huang F, Chen HJ. OntoED: Low-resource event detection with ontology embedding. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing. Association for Computational Linguistics, 2021. 2828–2839. [doi: [10.18653/v1/2021.acl-long.220](https://doi.org/10.18653/v1/2021.acl-long.220)]
- [72] Yang S, Feng DW, Qiao LB, Kan ZG, Li DS. Exploring pre-trained language models for event extraction and generation. In: Proc. of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: Association for Computational Linguistics, 2019. 5284–5294. [doi: [10.18653/v1/P19-1522](https://doi.org/10.18653/v1/P19-1522)]
- [73] Du XY, Cardie C. Event extraction by answering (almost) natural questions. In: Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2020. 671–683. [doi: [10.18653/v1/2020.emnlp-main.49](https://doi.org/10.18653/v1/2020.emnlp-main.49)]
- [74] Gangal V, Hovy E. BERTERING RAMS: What and how much does BERT already know about event arguments?—A study on the RAMS dataset. In: Proc. of the 3rd BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP. Association for Computational Linguistics, 2020. 1–10. [doi: [10.18653/v1/2020.blackboxnlp-1.1](https://doi.org/10.18653/v1/2020.blackboxnlp-1.1)]
- [75] Zhang ZS, Kong X, Liu ZZ, Ma XZ, Hovy E. A two-step approach for implicit event argument detection. In: Proc. of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2020. 7479–7485. [doi: [10.18653/v1/2020.acl-main.667](https://doi.org/10.18653/v1/2020.acl-main.667)]
- [76] Huang KH, Yang M, Peng NY. Biomedical event extraction with hierarchical knowledge graphs. In: Proc. of the 2020 Findings of the Association for Computational Linguistics. Association for Computational Linguistics, 2020. 1277–1285. [doi: [10.18653/v1/2020.findings-emnlp.114](https://doi.org/10.18653/v1/2020.findings-emnlp.114)]
- [77] Beltagy I, Lo K, Cohan A. SciBERT: A pretrained language model for scientific text. In: Proc. of the 2019 Conf. on Empirical Methods in

- Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 3615–3620. [doi: [10.18653/v1/D19-1371](https://doi.org/10.18653/v1/D19-1371)]
- [78] Lin HY, Lu YJ, Han XP, Sun L. Nugget proposal networks for Chinese event detection. In: Proc. of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: Association for Computational Linguistics, 2018. 1565–1574. [doi: [10.18653/v1/P18-1145](https://doi.org/10.18653/v1/P18-1145)]
- [79] Xu N, Xie HH, Zhao DY. A novel joint framework for multiple Chinese events extraction. In: Proc. of the 19th China National Conf. on Chinese Computational Linguistics. Springer, 2020. 174–183. [doi: [10.1007/978-3-030-63031-7_13](https://doi.org/10.1007/978-3-030-63031-7_13)]
- [80] Liu SB, Cheng R, Yu XM, Cheng XQ. Exploiting contextual information via dynamic memory network for event detection. In: Proc. of the 2018 Conf. on Empirical Methods in Natural Language Processing. Brussels: Association for Computational Linguistics, 2018. 1030–1035. [doi: [10.18653/v1/D18-1127](https://doi.org/10.18653/v1/D18-1127)]
- [81] Li PF, Zhou GD, Zhu QM. Semantics-based joint model of Chinese event trigger extraction. Ruan Jian Xue Bao/Journal of Software, 2016, 27(2): 280–294 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4833.htm> [doi: [10.13328/j.cnki.jos.004833](https://doi.org/10.13328/j.cnki.jos.004833)]
- [82] Zeng Y, Yang HH, Feng YS, Wang Z, Zhao DY. A convolution BiLSTM neural network model for Chinese event extraction. In: Proc. of the 5th CCF Conf. on Natural Language Processing and Chinese Computing, NLPCC 2016, and the 24th Int'l Conf. on Computer Processing of Oriental Languages. Kunming: Springer, 2016. 275–287. [doi: [10.1007/978-3-319-50496-4_23](https://doi.org/10.1007/978-3-319-50496-4_23)]
- [83] Subburathinam A, Lu D, Ji H, May J, Chang SF, Sil A, Voss C. Cross-lingual structure transfer for relation and event extraction. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019. 313–325. [doi: [10.18653/v1/D19-1030](https://doi.org/10.18653/v1/D19-1030)]
- [84] Yang H, Chen YB, Liu K, Xiao Y, Zhao J. DCFEE: A document-level Chinese financial event extraction system based on automatically labeled training data. In: Proc. of the 2018 ACL System Demonstrations. Melbourne: Association for Computational Linguistics, 2018. 50–55. [doi: [10.18653/v1/P18-4009](https://doi.org/10.18653/v1/P18-4009)]
- [85] Zhong WF, Yang H, Chen YB, Liu K, Zhao J. Document-level event extraction based on joint labeling and global reasoning. Journal of Chinese Information Processing, 2019, 33(9): 88–95, 106 (in Chinese with English abstract). [doi: [10.3969/j.issn.1003-0077.2019.09.011](https://doi.org/10.3969/j.issn.1003-0077.2019.09.011)]
- [86] Zhu PP, Wang ZQ, Li SS, Wang HL. Chinese event detection based on document information and Bi-GRU. Computer Science, 2020, 47(12): 233–238 (in Chinese with English abstract). [doi: [10.11896/j.sjcx.191100031](https://doi.org/10.11896/j.sjcx.191100031)]
- [87] Meng HJ. Research on event extraction technology in the field of unexpected events [MS. Thesis]. Shanghai: Shanghai University, 2015 (in Chinese with English abstract).
- [88] Nédellec C, Bossy R, Kim JD, Kim JJ, Ohta T, Pyysalo S, Zweigenbaum P. Overview of BioNLP shared task 2013. In: Proc. of the 2013 BioNLP Shared Task Workshop. Sofia: Association for Computational Linguistics, 2013. 1–7.
- [89] Pustejovsky J, Hanks P, Sauri R, See A, Gaizauskas R, Setzer A, Radev DR, Sundheim B, Day D, Ferro L, Lazo M. The timebank corpus. Corpus Linguistics, 2003, 2003: 647–656.
- [90] Sundheim BM, Chinchor NA. Survey of the message understanding conferences. In: Proc. of the 1993 Workshop on Human Language Technology. Plainsboro: Association for Computational Linguistics, 1993. 56–60. [doi: [10.3115/1075671.1075684](https://doi.org/10.3115/1075671.1075684)]
- [91] Ding X, Song F, Qin B, Liu T. Research on typical event extraction method in the field of music. Journal of Chinese Information Processing, 2011, 25(2): 15–20 (in Chinese with English abstract). [doi: [10.3969/j.issn.1003-0077.2011.02.003](https://doi.org/10.3969/j.issn.1003-0077.2011.02.003)]
- [92] Yang H. Research on the methods for document-level event extraction from chinese unstructured texts [MS. Thesis]. Harbin: Harbin University of Science and Technology, 2019 (in Chinese with English abstract).
- [93] Zhang ZK, Xu WR, Chen QQ. Joint event extraction based on skip-window convolutional neural networks. In: Proc. of the 5th CCF Conf. on Natural Language Processing and Chinese Computing, NLPCC 2016, and the 24th Int'l Conf. on Computer Processing of Oriental Languages. Kunming: Springer, 2016. 324–334. [doi: [10.1007/978-3-319-50496-4_27](https://doi.org/10.1007/978-3-319-50496-4_27)]
- [94] Chen C, Ng V. Joint modeling for Chinese event extraction with rich linguistic features. In: Proc. of the 2012 COLING. Mumbai: The COLING 2012 Organizing Committee, 2012. 529–544.

附中文参考文献:

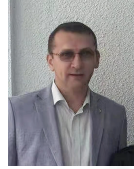
- [51] 黄细凤. 基于动态掩蔽注意力机制的事件抽取. 计算机应用研究, 2020, 37(7): 1964–1968. [doi: [10.19734/j.issn.1001-3695.2018.12.0927](https://doi.org/10.19734/j.issn.1001-3695.2018.12.0927)]
- [81] 李培峰, 周国栋, 朱巧明. 基于语义的中文事件触发词抽取联合模型. 软件学报, 2016, 27(2): 280–294. <http://www.jos.org.cn/1000-9825/4833.htm> [doi: [10.13328/j.cnki.jos.004833](https://doi.org/10.13328/j.cnki.jos.004833)]
- [85] 仲伟峰, 杨航, 陈玉博, 刘康, 赵军. 基于联合标注和全局推理的篇章级事件抽取. 中文信息学报, 2019, 33(9): 88–95, 106. [doi: [10.11896/j.sjcx.191100031](https://doi.org/10.11896/j.sjcx.191100031)]

[3969/j.issn.1003-0077.2019.09.011](http://www.jos.org.cn/3969/j.issn.1003-0077.2019.09.011)]

- [86] 朱培培, 王中卿, 李寿山, 王红玲. 基于篇章信息和Bi-GRU的中文事件检测. 计算机科学, 2020, 47(12): 233-238. [doi: [10.11896/j.issn.1003-0077.2019.09.011](https://doi.org/10.11896/j.issn.1003-0077.2019.09.011)]
- [87] 孟环建. 突发事件领域事件抽取技术的研究 [硕士学位论文]. 上海: 上海大学, 2015.
- [91] 丁效, 宋凡, 秦兵, 刘挺. 音乐领域典型事件抽取方法研究. 中文信息学报, 2011, 25(2): 15-20. [doi: [10.3969/j.issn.1003-0077.2011.02.003](https://doi.org/10.3969/j.issn.1003-0077.2011.02.003)]
- [92] 杨航. 面向非结构化中文文本的篇章级事件抽取研究 [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2019.



王浩畅(1974-), 女, 博士, 教授, 主要研究领域为人工智能, 自然语言处理, 数据挖掘, 生物信息学.



Marius Gabriel PETRESCU(1966-), 男, 博士, 教授, 主要研究领域为人工智能.



周柳莲(1995-), 女, 硕士, 主要研究领域为自然语言处理.

www.jos.org.cn

www.jos.org.cn