

基于门控图卷积与动态依存池化的事件论元抽取

王士浩 王中卿 李寿山 周国栋

苏州大学计算机科学与技术学院 江苏 苏州 215006

(shwang10@stu.suda.edu.cn)

摘要 事件论元抽取是事件抽取任务中一个极具挑战性的子任务。该任务旨在抽取事件中的论元及论元扮演的角色。研究发现,句子的语义特征和依存句法特征对事件论元抽取都有着非常重要的作用,现有的方法往往未考虑如何将两种特征有效地融合起来。因此,提出一种基于门控图卷积与动态依存池化的事件论元抽取模型。该方法使用 BERT 抽取出句子的语义特征;然后通过依存句法树设计两个相同的图卷积网络,抽取句子的依存句法特征,其中一个图卷积的输出会通过激活函数作为门控单元;接着,语义特征和依存句法特征通过门控单元后相加融合。此外,还设计了一个动态依存池化层对融合后的特征进行池化。在 ACE2005 数据集上的实验结果表明,该模型可以有效地提升事件论元抽取效果。

关键词: 事件论元抽取;图卷积;依存句法特征;语义特征;门控机制

中图法分类号 TP391

Event Argument Extraction Using Gated Graph Convolution and Dynamic Dependency Pooling

WANG Shi-hao, WANG Zhong-qing, LI Shou-shan and ZHOU Guo-dong

School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006, China

Abstract Event argument extraction is a very challenging subtask of event extraction. This task aims to extract the arguments in the event and the role they played. It is found that the semantic features and dependency features of sentences play a very important role in event argument extraction, and the existing methods often don't consider how to integrate them effectively. Therefore, this paper proposes an event argument extraction model using gated graph convolution and dynamic dependency pooling. This method uses BERT to extract the semantic features of sentences, and then two same graph convolution networks are used to extract the dependency features of sentences based on the dependency tree. The output of one graph convolution is used as the gating unit through the activation function. Then semantic features and dependency features are added and fused through gate unit. In addition, a dynamic dependency pooling layer is designed to pool the fused features. The experiment results on ACE2005 dataset show that the proposed model can effectively improve the performance of event argument extraction.

Keywords Event argument extraction, Graph convolution, Dependency features, Semantic features, Gate mechanism

1 引言

事件抽取任务旨在从文本中抽取出标志事件发生的触发词(Trigger)和对应事件的各个论元(Argument)。触发词是表示事件发生的词,一般为动词或名词,少数情况下为形容词。论元是事件的参与者,主要由实体、时间、数值构成。事件抽取任务一般分为两个部分:第一部分是事件检测任务(Event Detection),旨在识别句子中的触发词并对其类型进行分类;第二部分是事件论元抽取(Event Argument Extraction),旨在识别作为论元的实体并对其论元角色进行分类。如图1中例句所示,事件检测程序首先会识别出句子中的“died”和“fired”为触发词,并将其分类为Die(死亡)事件和Attack(攻击)事件。接着,事件论元抽取程序会抽取出每个事件的论元并判断其论元角色。“Baghdad”“cameraman”“American tank”在Die(死亡)事件

中论元角色分别是Place(地点)、Victim(受害者)、Instrument(工具)。本文主要研究的是第二个子任务——事件论元抽取。

在事件论元抽取任务上,语义特征和依存句法特征对模型性能提升都有很大的帮助。在基于语义特征的模型上,常使用卷积神经网络、循环神经网络、Transformer等编码器计算单词在句子中的语义。卷积神经网络在处理文本时,由于卷积核大小的限制,模型易限于局部的特点,从而忽略了距离较远词的信息。循环神经网络在处理较长的句子时,无法捕捉词之间长期依赖关系,并且可能产生梯度消失和梯度爆炸的问题。Transformer通过多头自注意力机制将句子中词与词之间的距离变为了1,解决了卷积神经网络只关注于局部的特点和循环神经网络无法捕捉长期的依赖关系等问题。

依存句法分析是自然语言处理的基础任务之一,反映了句子中词与词之间的依存关系。依存句法特征对事件论元抽

取有着十分重要的帮助。在一个事件中,触发词和论元之间往往有着直接或间接的依存关系。图 1 给出了例句 1 的依存关系,我们可以看到触发词“died”和 3 个论元之间都有依存关系,“Baghdad”是其复合名词修饰(nmod),“cameraman”是

其名词主语(nsubj),“tank”是“fired”的名词主语(nsubj),而“fired”又是“died”的状语从句修饰词(advcl)。对候选论元进行编码时,融合与其存在密切联系的信息,可以帮助模型判断候选论元的角色。

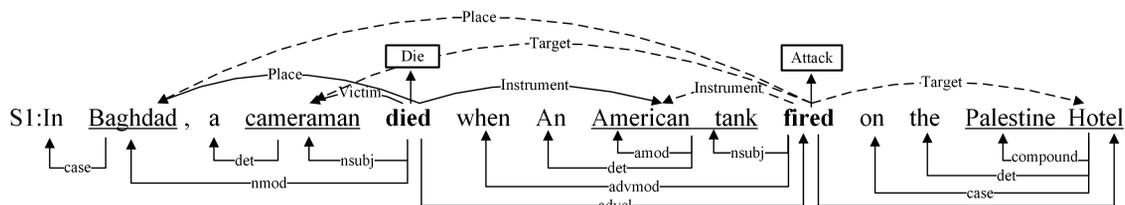


图 1 例句 S1 的触发词、论元以及依存关系

Fig. 1 Triggers, arguments and dependency of example S1

为了更好地将语义特征和依存句法特征融合,本文提出了基于门控图卷积和动态依存池化的事件论元抽取模型。根据所使用的特征范围,我们将特征分为了句级特征和词级特征。首先使用 BERT^[1] 获取句子的语义特征;然后根据句子的依存句法树构造邻接图,使用两个图卷积^[2] 获取句子的依存句法特征。受到文献^[3] 的启发,我们只对其中一个图卷积使用 Sigmoid 激活函数,作为门控单元,控制语义特征和依存句法特征的流量。因为只有一个图卷积使用了激活函数,这样的设计降低了模型梯度消失的风险。接着,分别让语义特征和依存句法特征通过门控单元,以残差结构将语义特征与依存句法特征相加融合。在池化方式上,对论元进行分类时最重要的词是触发词、论元,以及和论元有直接依存关系的词。例如在例句 1 中,当触发词为“fired”、候选论元为“cameraman”时,我们难以判断“cameraman”是受害者还是攻击者。通过候选论元有直接依存关系的“died”,我们可以很容易地判断“cameraman”的论元角色是受害者。因此,我们提出动态依存池化,将触发词、候选论元以及和候选论元有直接依存关系的词的隐层向量进行最大池化,获得最终的句级特征。随后,从语义特征中提取出触发词和候选论元以及两者的上下词的隐层向量,将其拼接在一起作为词级特征。最后,将句级特征和词级特征进行拼接,通过全连接层和 softmax 激活函数进行论元角色分类。

本文第 2 节介绍了事件论元抽取的相关工作;第 3 节阐述基于门控图卷积与动态依存池化的事件论元抽取模型;第 4 节介绍实验设置并给出实验结果和分析;最后总结全文。

2 相关工作

事件论元抽取作为事件抽取的子任务,一直是研究中的难点。事件抽取任务分为事件检测和事件论元抽取任务。其中,事件论元抽取任务输入依赖于事件检测任务输出。近年来,对事件抽取的研究层出不穷。根据所使用的特征范围,可以将模型分为句子级事件抽取和篇章级事件抽取。根据学习方法的不同,可以将模型分为基于流水线模型的事件抽取方法和基于联合学习模型的事件抽取方法。

传统的事件抽取模型需要进行大量的特征工程,人工构建候选触发词和候选论元的特征,然后使用基于统计的分类模型对候选触发词和候选论元进行分类。Ahn 等^[4] 首先对事件抽取进行了定义,将事件抽取任务当作多分类任务,使用词级特征、句级特征以及外部知识对触发词和论元进行分类。

Liao 等^[5] 使用了篇章级的跨事件信息来提升事件抽取模型效果。Hong 等^[6] 认为实体类型一致性是事件抽取的关键信息,提出了跨实体推理事件抽取模型。McClosky 等^[7] 将事件抽取任务看成依存分析问题,将事件触发词与论元之间的关系树看成依存关系树,以此提升模型性能。Li 等^[8] 提出了基于结构化感知器的联合模型,同时抽取触发词和论元。该工作中还设计了如触发词和论元的词性、语法、语义等局部特征和能够进行触发词和论元交互的全局特征帮助模型进行分类。在中文数据集上, Li 等^[9] 提出了一种篇章级联合模型,为了更好地表示在一个篇章中触发词与论元间的关系,在联合模型中设计了多种限制。

随着神经网络的飞速发展,越来越多的研究者们也开始使用基于神经网络的模型进行事件抽取。Chen 等^[10] 率先使用神经网络的方法进行事件抽取,同时设计了动态多池化层保留多事件句子更多的信息,验证了神经网络模型的有效性。在 Li 等^[8] 和 Chen 等^[10] 工作的基础上, Nguyen 等^[11] 吸取两者的优点,提出一种基于 RNN 的联合学习模型,既可以通过神经网络自动提取特征,也可以避免管道模型的错误传递。为了有效利用论元角色概念上的联系, Wang 等^[12] 提出了层次模块化事件论元抽取方法。该模型利用了论元角色和上位概念之间的联系,初始化多个概念模块计算每个模块的得分,最后将有联系的模块得分相加,辅助论元角色分类。

目前,依存句法特征大多通过图卷积神经网络进行编码。Negyun 和 Grishman 等^[13] 提出了一种基于依存关系的图卷积和实体池化的模型用于事件检测。Yan 等^[14] 根据依存关系构造了多个不同阶次的图,同时使用了注意力机制将不同的图融合到一起。Liu^[15] 设计了一种联合学习事件抽取框架,通过图卷积网络进行事件抽取。与上述 3 个模型不同, Sha 等^[16] 将依存关系加入 LSTM 中,使模型能够捕捉句法信息,在论元分类阶段,加入了论元之间的联系特征,以提升模型效果。

3 基于门控图卷积和动态依存池化的事件论元抽取模型

本部分将详细介绍基于门控图卷积和动态依存池化的事件论元抽取模型。如图 2 所示,本文所提出的模型可分为 3 个部分:(1)句级特征抽取层;(2)词级特征抽取层;(3)论元角色分类层。

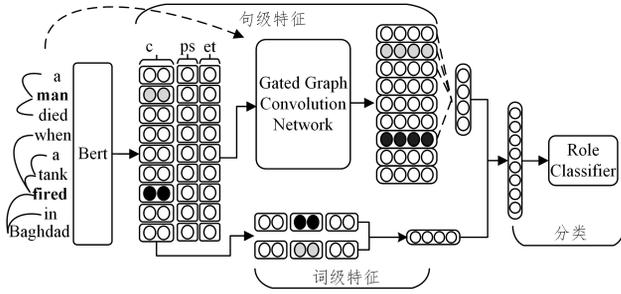


图2 基于门控图卷积与动态依存池化的事件论元抽取模型

Fig. 2 Event argument extraction model using gated graph convolution and dynamic dependency pooling

3.1 句级特征抽取

3.1.1 语义特征抽取

本文使用预训练模型 BERT 对句子进行编码,得到句子的语义特征。BERT 模型使用多层 Transformer 的架构,通过多头自注意力机制,每个单词都会与所有词进行注意力计算,使句子中每个词之间的距离变为 1,克服了循环神经网络中的长期依赖问题。句子通过 BERT 编码后,每一个词都获得了全局的语义特征。给定句子 $X = \{w_1, w_2, \dots, w_i, \dots, w_n\}$,其中 n 表示句子的长度, w_i 表示句子中的触发词, w_n 表示句子中的候选论元。将句子输入 BERT 模型中,得到句子的语义特征表示 $X = \{c_1, c_2, \dots, c_i, \dots, c_n\}$ 。

3.1.2 门控图卷积

对于一个长度为 n 的句子,我们使用 stanfordCoreNLP¹⁾ 工具进行依存句法分析。本文将句子中每一个词作为一个节点,根据依存关系构造该句子的邻接矩阵 $G = \{V, E\}$ 。 $V = \{v_1, v_2, \dots, v_n\}$ 包含了 n 个节点,对应句子中每个词, E 为每个节点之间的边。参考了 Kipf 等^[2]的工作,我们首先对每一个节点创建一个自环 (v_i, v_i) ,如果节点 v_i 和 v_j 之间有依存关系,我们将添加边 (v_i, v_j) ,为了保证信息反向流通,我们添加边 (v_j, v_i) 。图卷积是一种在图上进行学习的神经网络方法,图中的每一个节点都会因为邻居结点或者更远结点的影响而改变自身的状态,关系越亲近,影响越大。图卷积的公式如式(1)所示:

$$h_i^{(k)} = \sigma \left(\frac{G_i^T (h_i^{(k-1)} W + b)}{\sum G_i} \right) \quad (1)$$

其中, G_i 是图的邻接矩阵, h_i 是输入的特征, W 是权重参数, b 是偏移量, σ 是激活函数, k 是图卷积的层数。

门控卷积的输入如图 2 所示。第一个是通过 BERT 获得的语义特征 c ,第二个是句子中每个词与触发词和候选论元的位置向量 ps ,本文给每一个相对位置都初始化了一个随机的隐层向量。第三个是当前事件类型的特征向量 et ,本文给每个事件类型都初始化了一个隐层向量。将这三者进行拼接,最终得到门控卷积的输入。门控图卷积的结构如图 3 所示。首先将上一层得到的隐层向量 X 进行两次图卷积操作,得到句子的依存句法特征。然后对其中一个依存句法特征使用 Sigmoid 函数,将其作为门控单元,控制依存句法特征和语义特征的流量,我们记为 σ 。接着将依存句法特征与 σ 进行点

乘,将语义特征与 $(1-\sigma)$ 进行点乘。依存句法特征和语义特征通过门控单元后相加融合。具体公式见式(2):

$$\begin{cases} H = X \otimes (1-\sigma) + \text{GraphConv}_1(X) \otimes \sigma \\ \sigma = \text{Sigmoid}(\text{GraphConv}_2(X)) \end{cases} \quad (2)$$

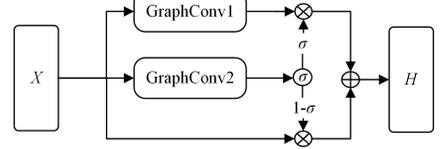


图3 门控图卷积网络

Fig. 3 Gated graph convolution network

3.1.3 动态依存池化

经过特征融合后,我们得到隐层向量 $H = \{h_1, h_2, \dots, h_n\}$,接着我们需要对隐层向量进行池化操作。动态依存池化将候选论元、触发词以及和候选论元有直接依存关系的词进行最大池化。具体见式(3):

$$S = \max\{\{h_a, h_t\} \cup \{h_i\}\}, (v_i, v_a) \text{ or } (v_a, v_i) \in E; \\ i, a, t \in [1, n] \quad (3)$$

其中, t 表示触发词的位置, a 表示候选论元的位置, v 代表节点, E 表示图的边。

3.2 词级特征抽取

词级特征是事件抽取的重要特征,在处理多义词时,我们可以通过当前词的上下词来判断单词的含义,提升模型效果。句子通过 BERT 模型后得到 $X = \{c_1, c_2, \dots, c_i, \dots, c_n\}$; $t, a \in [1, n]$,其中 c_t 表示触发词的向量, c_a 表示候选论元的向量。我们将设置一个大小为 3 的窗口,分别以触发词和候选论元为中心,计算窗口中向量的和。最后将代表触发词词级特征的向量 L_t 和代表候选论元词级特征的向量 L_a 拼接起来作为词级特征。具体见式(4):

$$\begin{cases} L_t = \text{sum}(c_{t-1}, c_t, c_{t+1}) \\ L_a = \text{sum}(c_{a-1}, c_a, c_{a+1}) \\ L = L_t \oplus L_a \end{cases} \quad (4)$$

3.3 论元角色分类

在论元角色分类阶段,我们将获得的词级特征 L 和句级特征 S 进行拼接,得到向量 H' 。接着,将隐层向量 H' 进行线性变换,使其映射到标签空间,具体如式(5)所示:

$$o = W \times H' + b \quad (5)$$

其中,权重矩阵 $W \in R^{K \times d}$, K 是论元角色的个数, d 是向量的维度, b 是偏移量。接着,通过 Softmax 函数得到各个论元角色的概率。候选论元 w_a 属于论元角色 r 的概率计算见式(6):

$$P(r | w_a, \theta) = \frac{e^{\theta_r}}{\sum_{i=1}^K e^{\theta_i}} \quad (6)$$

3.4 训练与优化

在训练中,本文使用自适应矩估计 (Adaptive Moment Estimation, Adam)^[17] 算法更新网络参数,并采用交叉熵误差作为损失函数,如式(7)所示:

$$L(\theta) = - \sum_i \log p(r_i | w_i, \theta) \quad (7)$$

在分类层之前,本文使用了 Dropout 机制随机丢弃一部分网络,防止过拟合。

¹⁾ <https://stanfordnlp.github.io/CoreNLP/>

4 实验

4.1 实验数据集

本实验使用 ACE2005 进行事件论元抽取。ACE2005 广泛用于事件抽取任务,其定义了 8 种事件类型(又分为 33 种子事件类型)和 35 种论元角色。该语料共包含了 599 个文档,在数据集划分上,我们采取了与文献[10-12]相同的划分方式,使用 40 篇新闻类型文档作为测试集,在剩下的文档中随机选取 30 篇作为验证集,529 篇作为训练集。实验主要评测模型在论元角色分类上的结果,当一个论元的事件类型、位置和论元角色都正确时,则论元角色分类正确。与以往工作相同,本文使用准确率、召回率、F1 值作为评价指标。

4.2 实验参数设置

实验所采用的各个参数如表 1 所列。

表 1 模型超参数设置

Table 1 Model hyperparameter setting

Hyperparameter	Value
Learning Rate	2×10^{-5}
Batch Size	8
Bert Dimension	768
Graph Hidden Layer Dimension	918
Event Type Embedding Dimension	50
Position Embedding Dimension	50
Dropout Probability	0.5

4.3 实验结果与分析

本文关注事件论元抽取任务,事件检测模型使用 Chen 等^[10]提出的 DMCNN 模型。在前人研究的基础上,本文将选取下列模型作为基准系统进行对比分析。

(1)DMCNN^[10]:由 Chen 等人提出,是第一个使用神经网络方法进行事件抽取的模型。该模型设计了一个动态多池化卷积神经网络保留句子中更多的信息,改善模型效果。

(2)RBPB^[18]:由 Sha 等人提出,该模型利用论元之间的相关性提升模型效果。

(3)JRNN^[11]:由 Nguyen 等人提出,采用联合模型,使用双向 RNN 进行事件抽取,同时设计了记忆矩阵来捕获论元角色和事件类型之间的关系。

(4)dbRNN^[16]:由 Sha 和 Qian 等人提出,该模型将依存关系加入双向 LSTM 中,在每一个时间步根据边来计算隐层向量,同时使用一个张量来表示论元之间的关系。

(5)HMEAE^[12]:由 Wang 等人提出,该模型利用了不同论元角色概念上的联系来提升论元角色分类的效果。

(6)GGCN:本文所提出的事件论元抽取模型,使用门控图卷积与动态依存池化来抽取论元。

表 2 列出了我们的模型和选取的基准模型在论元分类上的结果。

表 2 实验结果

Table 2 Experiment results

Model	Argument Role Classification		
	P	R	F
DMCNN	62.2	46.9	53.5
RBPB	54.1	53.5	53.8
JRNN	54.2	56.7	55.4
dbRNN	66.2	52.8	58.7
HMEAE	62.2	56.6	59.3
GGCN(ours)	57.4	63.5	60.3

(单位:%)

通过 GGCN 与选取的 5 个基准模型对比可以看出,本文提出的模型 F1 值达到了 60.3%,比基准模型 HMEAE 在 F1 值上高出 1%。GGCN 模型在事件论元抽取性能上的提升主要来源于两个方面。

(1)通过门控机制将依存句法特征和语义特征融合。依存句法特征和语义特征能够帮助模型提升性能,通过门控机制将依存句法特征和语义特征进行动态融合,降低了模型梯度消失的风险。

(2)动态依存池化的加入。动态依存池化保证了候选论元提取特征的多样性,又避免了池化时对论元分类无关词的影响。

4.4 门控图卷积对事件论元抽取的影响

为了验证门控图卷积对事件论元抽取结果的影响,本部分设置了使用不同特征以及不同融合方式的实验。

(1)Model_Sem:在原有模型基础上,句级特征只使用 BERT 抽取出来的语义特征。

(2)Model_Dep:在原有模型基础上,句级特征只使用单个图卷积抽取出来的依存句法特征。

(3)Model_Connect:模型在获取语义特征与依存句法特征后,只使用简单的拼接将两种特征融合一起。

表 3 列出了上述模型在事件论元抽取的结果。通过分析我们可以得到如下两个结论。

(1)通过对比 Model_Dep 和 Model_Sem,可以发现 Model_Dep 的 F1 值到达了 59.2%,比 Model_Sem 高出了 1.5%。这说明了通过图卷积模型提取的依存句法特征能够有效提升事件论元抽取性能。

(2)GGCN 模型获得了最高的 F1 值 60.3%,比 Model_Sem 高出 2.6%,比 Model_Dep 高出 1.1%,比 Model_Connect 高出了 0.6%。Model_Connect 也比 Model_Sem 和 Model_Dep 分别高出 2% 和 0.5%。这表明语义特征和依存句法特征的融合可以有效提升论元抽取结果。Model_Connect 只是将依存句法特征和语义特征进行简单的拼接融合,因此提升效果有限。门控图卷积通过门控机制将语义特征和依存句法特征动态结合,比简单的拼接能更好地防止梯度消失,提升模型效果。

表 3 使用不同特征和融合方式的实验结果

Table 3 Experiment results using different features and fusion methods

Model	Argument Role Classification		
	P	R	F
Model_Sem	57.1	58.3	57.7
Model_Dep	55.9	63.0	59.2
Model_Connect	58.9	60.6	59.7
GGCN	57.4	63.5	60.3

4.5 不同池化机制对事件论元抽取的影响

为了探究动态依存池化对事件论元抽取效果的影响,我们在门控图卷积模型上使用了多个不同的池化机制,并在 ACE2005 语料上进行进行实验。

(1)最大池化:将所有隐层向量进行最大池化。

(2)动态多池化:将事件触发词和候选论元分成 3 个部分,对每个部分的隐层向量进行最大池化,最后将 3 个部分得到的向量拼接在一起。

(3)动态依存池化:本文所提出的池化机制,将触发词、候选论元以及和候选论元有直接依存关系的词的隐层

向量进行最大池化。

表 4 列出了不同池化机制的对结果的影响,从中可以得出如下结论。

(1)最大池化结果最差,因为最大池化只关注最强的特征,抽取的特征过于单一,在处理含有多个事件的句子时无法根据触发词和论元的变化而抽取不同的特征。如例 1 中含有 Die 类型事件和 Attack 类型事件,“cameraman”在两个事件中的论元角色分别是 Victim(受害者)和 Target(目标)。无论判断“cameraman”在哪一个事件中的角色,最大池化都会抽取到相同的结果,导致效果较差。

(2)动态多池化将句子划分成多个部分进行最大池化,保留了更多的信息,因此比最大池化的效果更好。但是动态多池化没有舍弃掉那些对分类无关紧要的特征,保留了一部分噪音,因此其实验结果比动态依存池化低一些。如例 1 中,我们判断“cameraman”在 Die 类型事件中的论元角色时,句子中的“when”“on the Palestine Hotel”等内容并不会对“cameraman”的论元角色分类提供有价值的信息,甚至会产生一定的噪音。

(3)我们提出的动态依存池化获得了最好的结果,F1 值达到了 60.3%,这是因为动态依存池化只使用了触发词、候选论元以及和候选论元有直接依存关系的词的隐层向量,这样在对含句子进行论元抽取时既保证了特征的多样性,又可以有效地去除掉对论元分类无关紧要的特征。如例 1 中当触发词为“fired”时,论元“cameraman”的论元角色难以确定,但是通过和“cameraman”有着依存关系的词“died”可以很容易地判定“cameraman”的论元角色是受害者,避免其他词的干扰。

表 4 使用不同池化机制的实验结果

Table 4 Experiment results using different pooling mechanisms

Model	Argument Role Classification (单位:%)		
	P	R	F
最大池化	54.7	60.3	57.4
动态多池化	56.7	62.9	59.6
动态依存池化	57.4	63.5	60.3

结束语 本文提出了基于门控图卷积和动态依存池化的事件论元抽取模型。该模型首先使用 BERT 抽取出句子的语义特征,然后使用依存句法树和图卷积抽取句子的依存句法特征,并通过门控机制将两种特征进行融合。在池化操作上,本文提出了动态依存池化,降低了池化后隐层向量中的噪音,提高了论元抽取效果。在 ACE2005 语料上进行的实验表明,本文提出的模型能够有效提升论元分类的结果。

本文工作只考虑了句子级的事件论元抽取。然而在实际应用中,一个事件的论元可能分布在多个句子中。在未来的工作中,我们将进一步探究篇章级的事件论元抽取。

参考文献

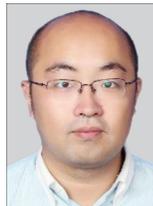
- [1] DEVLIN J, CHANG M W, LEE K, et al. Bert: pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of NAACL-HLT. 2019:4171-4186.
- [2] KIPF T N, WELLING M. Semi-Supervised Classification with Graph Convolutional Networks [C] // Proceedings of ICLR. 2017.
- [3] SU J L. Reading comprehension question answering model based

on CNN; DGCNN [EB/OL] (2018-07-28). <https://spaces.ac.cn/archives/-5409>.

- [4] DAVID A. The stages of event extraction [C]// Proceedings of the Workshop on Annotating and Reasoning about Time and Events. 2006:1-8.
- [5] LIAO S, GRISHMAN R. Using document level cross-event inference to improve event extraction [C]// Proceedings of ACL. 2010:789-797.
- [6] HONG Y, ZHANG J F, MA B, et al. Using cross-entity inference to improve event extraction [C] // Proceedings of ACL-HLT. 2011:1127-1136.
- [7] MCCLOSKEY D, SURDEANU M, CHRISTOPHER D M. Event extraction as dependency parsing [C] // Proceedings of ACL-HLT. 2011:1626-1635.
- [8] LI Q, JI H, HUANG L. Joint event extraction via structured prediction with global features [C]// Proceedings of ACL. 2013:73-82.
- [9] LI P, ZHU Q, ZHOU G. Joint modeling of argument identification and role determination in Chinese event extraction with discourse-level information [C]// Proceedings of IJCAI. 2013.
- [10] CHEN Y B, XU L H, LIU K, et al. Event extraction via dynamic multi-pooling convolutional neural networks [C]// Proceedings of ACL. 2015:409-419.
- [11] NGUYEN T H, CHO K, GRISHMAN R. Joint event extraction via recurrent neural networks [C] // Proceedings of NAACL-HLT. 2016:300-309.
- [12] WANG X Z, WANG Z Q, HAN X, et al. HMEAE: Hierarchical Modular Event Argument Extraction [C] // Proceedings of EMNLP. 2019:5781-5787.
- [13] NGUYEN T, GRISHMAN R. Graph convolutional networks with argument-aware pooling for event detection [C]// Proceedings of AACL. 2018:5900-5907.
- [14] YAN H, JIN X L, MEMG X B, et al. Event Detection with Multi-Order Graph Convolution and Aggregated Attention [C]// Proceedings of EMNLP. 2019:5766-5770.
- [15] LIU X, ZHUNCHEN L, HEYAN H. Jointly multiple events extraction via attentionbased graph information aggregation [C]// Proceedings of EMNLP. 2018:1247-1256.
- [16] SHA L, QIAN F, CHANG B B, et al. Jointly extracting event triggers and arguments by dependency-bridge RNN and tensor-based argument interaction [C]// Proceedings of AACL. 2018:5916-5923.
- [17] KINGM D, BA J. Adam: A Method for Stochastic Optimization [C]// Proceedings of ICLR. 2015.
- [18] SHA L, LIU J, LIN C Y, et al. RBPB: Regularization-based pattern balancing method for event extraction [C]// Proceedings of ACL. 2016:1224-1234.



WANG Shi-hao, born in 1997, postgraduate, is a member of China Computer Federation. His main research interests include natural language processing and so on.



WANG Zhong-qing, born in 1987, Ph.D., associate professor. His main research interests include natural language processing and so on.