基于中文信息与越南语句法指导的越南语事件检测

陈龙[1,2]，**郭军军**[1,2]，**张亚飞***[1,2]，**高盛祥**[1,2]，**余正涛**[1,2]
1. 昆明理工大学，信息工程与自动化学院，昆明，650500
2. 昆明理工大学，云南省人工智能重点实验室，昆明，650500
2722383436@qq.com， guojjgb@163.com， zyfeimail@163.com，
gaoshengxiang.yn@foxmail.com， ztyu@hotmail.com

摘要

当前基于深度学习的事件检测模型都依赖足够数量的标注数据，而标注数据的稀缺及事件类型歧义为越南语事件检测带来了极大的挑战。根据“表达相同观点但语言不同的句子通常有相同或相似的语义成分”这一多语言一致性特征，本文提出了一种基于中文信息与越南语句法指导的越南语事件检测框架。首先通过共享编码器策略和交叉注意力网络将中文信息融入到越南语中，然后使用图卷积网络融入越南语依存句法信息，最后在中文事件类型指导下实现越南语事件检测。实验结果表明，在中文信息和越南语句法的指导下越南语事件检测取得了较好的效果。

关键词： 事件检测；越南语；中文信息；图卷积网络

Vietnamese event detection based on Chinese information and Vietnamese syntax guidance

Long Chen[1,2]**,Junjun Guo**[1,2]**,,Yafei Zhang***[1,2]
Shengxiang Gao[1,2],Zhengtao Yu[1,2]
1. Faculty of Information Engineering and Automation,
Kunming University of Science and Technology Kunming 650500, China
2. Yunnan Key Laboratory of Artificial Intelligence,
Kunming University of Science and Technology Kunming 650500, China
2722383436@qq.com， guojjgb@163.com， zyfeimail@163.com，
gaoshengxiang.yn@foxmail.com， ztyu@hotmail.com

Abstract

Current event detection models based on deep learning rely on a sufficient amount of labeled data and only focus on specific in-formation such as trigger words. However, the scarcity of annotation data for Vietnamese news events and the ambiguity of event types have brought great challenges to Vietnamese event detection. According to the fea-ture that sentences expressing the same viewpoint but different languages usually have the same or similar semantic components, this paper proposes a Vietnamese event detection framework that combines Chinese information and Vietnamese syntax. First, use the shared encoder strategy and cross-attention network to integrate Chinese semantic information into Vietnamese, then use graph convolutional network to obtain Vietnamese representation based on Vietnamese dependency syntactic information. Finally, the Vietnamese semantic representation based on Chinese event type

*通信作者：张亚飞，email地址:zyfeimail@163.com
基金项目：国家自然科学基金（61762056，61972186，61732005，61761026）；国家重点研发计划（2018YFC0830105, 2018YFC0830101, 2018YFC0830100）；云南高科技人才项目（201606）；云南省重大科技专项计划（202002AD080001-5），云南省基础研究计划（202001AS070014，2018FB104）。

information is extracted through the event type perception network to realize Vietnamese news event detection. Experimental results show that under the guidance of Chinese information and Vietnamese syntax, Vietnamese event detection has achieved good results.

Keywords: Event Detection , Vietnamese , Chinese information , Graph convolutional network

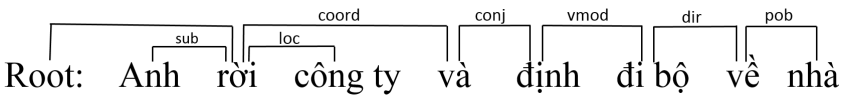
1 引言

事件检测（Event Detection，ED）是事件抽取中一个关键任务，可以从海量的文本中快速、准确地获取事件信息并对其进行分类。目前基于深度学习的事件检测模型都依赖足够数量的标注数据，并且通过触发词等特定的线索进行事件分类。对于一些标记资源丰富的语言，如中文、英文等，(Li et al., 2013; Liu et al., 2019; Chen et al., 2020)事件检测模型基于大规模标记语料进行有监督训练且取得了较好的效果。而现有越南语标记语料非常稀缺,造成其事件检测效果较差。同时，由于句子中局部信息可能会触发多个事件类型，例如，在“他离开了公司打算步行回家。”这句话中，触发词“离开”表达的是交通事件或结束位置事件，但是结合“步行”信息，就能准确判断为交通事件。由此可见，由于触发词表达的局限性，易引起事件类型歧义。综上所述，受标注数据数量的限制与事件类型歧义的影响，使得越南语事件检测面临极大的挑战。对于标注数据稀缺语言的事件检测，Chen等人 (Chen and Ji, 2009)是使用源语言检测器在并行文档上标注事件触发词，(Zhu et al., 2014; Faruqui and Kumar, 2015; Zou et al., 2018)使用机器翻译来获取额外的触发词标注数据，以扩充目标语言训练数据的规模。这些方法受机器翻译性能以及越南语音节分词的特性的影响，不能直接应用于越南语触发词的标注。例如，给定中-越语句:

V: Một chiếc xe tăng đã **nổ súng** vào một khách sạn.
C: 一辆坦克向酒店**开火**。
Type: 冲突

在中文语句C中，包含了“冲突”事件的触发词“开火(Ngọn lửa）”，而越南语的每一个音节常常是一个有意义的单位，这些音节也是构成多音节词的基础，并且相同含义的多音节词在不同的语句中构成也不同，因此中文触发词“开火（Ngọn lửa）”通过翻译的方式不能在越南语句子V中准确匹配到触发词。事件检测依赖于充分理解事件句的语义信息，根据McDonald等人 (McDonald et al., 2013)提出不同语言中表达相同想法的句子通常有相同或相似的语义成分这一特征，同时结合Liu等人 (Liu et al., 2019)通过目标事件类型计算句子的表示实现无触发词的事件检测思想，本文提出使用中文语句以及所标注的中文事件类型信息共同指导越南语事件检测的思想，通过共享编码器和双语交叉注意力机制使中文信息深入融入到越南语中，使得越南语可获得中文语句所表达的事件信息，最后在中文事件类型的指导下对越南语检测，从而减少对越南语数据标注信息的依赖。

此外，在越南语事件检测中，由于事件类型歧义影响为越南语事件检测带来挑战。针对局部信息所引发多个事件而造成事件类型歧义问题，近年来，(Ji and Grishman, 2008; Liao and Grishman, 2010)也使用文档级特征信息来解决歧义，对于句子中没有关注到的重要信息，Chen等人 (Chen et al., 2015)使用上下文信息保存到句子级序列建模中，以保留更多关键信息。(Yang and Mitchell , 2016; Katherine et al., 2017; Liu et al., 2016)利用句子特征信息进行建模，Nguyen等人 (Nguyen and Grishman, 2016)通过引入记忆向量和矩阵在标记句子的过程中存储预测信息。然而，基于特征的序列建模方法在获取非常长的相关性方面效率低下，并且没有充分地提取关键信息之间的关联。因此，本文利用越南语依存句法信息，使用图卷积网络(Graph GCN)增强越南语中信息之间的关联，例如，给定越南语语句:



在越南语句中，从“rời（离开）”到“đi bộ（步行）”中间要经过三次跳跃，但是根据依存句法表示的快捷弧，只需要两次跳跃便得到这两个相互依赖的信息，从而加强了信息之间的依赖

关系。与序列建模方法相比，采用图卷积网络对快捷弧进行建模，以图中相邻节点的代表向量学习每个节点的句法上下文表示，可实现越南语特征的有效提取，从而提高越南语事件检测的准确性。

本文提出融合中文信息与越南语句法的越南语事件检测框架，首先通过共享编码器策略和交叉注意力网络（Cross attention network）使越南语获得中文语义表征，然后使用图卷积网络使越南语融入依存句法信息，最后通过设计事件类型感知网络实现对越南语事件信息表征，以实现在中文信息和越南语句法信息指导下的越南语事件检测。

综上所述，本文的主要贡献如下：

(1) 本文提出了融合中文信息和越南语句法的越南语事件检测模型，通过中文语句信息与越南语句法信息的融入，实现越南语事件检测。

(2) 设计了中越双语信息融合网络与事件类型感知网络，以实现基于中文信息指导的越南语语义表示，从而减少越南语事件检测对越南语数据标注信息的依赖。

(3) 利用越南语依存句法信息，通过句法图卷积网络提取越南语句法特征并将其融入越南语语句中，消除事件类型歧义，提高事件检测的准确性。

2 相关工作

事件检测是自然语言处理中的一个热点问题，近年来受到了广泛的关注。事件检测旨在检测非结构化文本中的事件信息并对其进行分类。根据事件检测任务不同可分为两类：单语训练和多语言联合训练。

目前事件检测的方法研究主要集中在单语。Li等人 (Li et al., 2013)将常见的事件检测任务建模为触发词分类，预测给定句子中的每个单词是否是事件触发器以及它触发的事件类型。当前，随着深度学习的发展，Nguyen等人 (Nguyen and Grishman , 2015)通过使用神经网络来自动学习任务的特征，以达到对事件的检测; Liu等人 (Liu et al., 2016)使用概率软逻辑模型以逻辑形式编码全局信息，并利用FrameNet中的事件信息来引导训练;Nguyen等人 (Nguyen and Grishman, 2016)通过引入卷积网络（CNN）的连续模型来改进非卷积模型，从而获得更好的性能;Katherine等人 (Katherine et al., 2017)构建事件之间相互依赖关系实现对事件的检测;Liu等人 (Liu et al., 2017)利用了额外的参数信息和FrameNet，通过监督注意机制显式地利用论点信息进行事件检测; Nguyen等人 (Nguyen and Grishman , 2018)通过卷积神经网络从语句的句法层面进行特征提取，利用结构化信息来提高触发词的识别，进而提高事件检测。Liu等人 (Liu et al., 2018)通过图卷积神经网络从语句的句法层面进行深层次特征提取，以此提高语句的理解和触发词的识别。Chen等人 (Chen et al., 2020)提出为模型提供带有漂白语句（实体用通用的方式指代）的模型，通过将文本中的实体用指代的方式表示。使模型能够提取封闭本体下的事件并推广到未知的事件类型；Du等人 (Du and Cardie , 2020)通过将事件公式化为问题解答（QA）任务来引入事件提取的新范式，以端到端的方式提取事件参数。但它们的性能受到特定语言中标记数据量的限制，需要大规模标注数据进行训练。

对于低资源语言而言，由于注释的复杂性和高成本使得训练数据严重不足，而多语言事件检测尝试在不同语言之间传递知识以解决这一困难。 (Zhu et al., 2014; McDonald et al., 2013)使用机器翻译来获取额外的标记数据，以扩充目标语言训练数据的规模。Hsi等人 (Hsi et al., 2016)使用语言相关和语言独立功能的组合在多种语言上进行训练，尤其关注目标域训练数据量非常有限的情况。Guo等人 (Guo et al., 2015)使用源语言的训练数据，针对目标语言训练分布式表征实现跨语言转移。Mayhew等人 (Mayhew et al., 2017)提出使用一种或多种高资源语言中可用的带注释的数据辅助目标语言训练，以增强目标语言模型性能。 (Ni et al., 2017; Subburathinam and Lu et al., 2019)提出将训练好的源语言模型迁移到目标语言，并通过对齐的方式以达到对目标语言的适应以解决目标语言稀缺问题，并达到预期的效果;Liu等人 (Liu and Chen et al., 2018)通过其他语言提供的大量信息，利用多语言数据传达的补充信息来解决单语中存在的数据稀缺和歧义问题。Subburathinam等人 (Subburathinam and Lu et al., 2019)通过训练一个公共空间，将实体、触发词等信息融入进去，并用源语言的触发词等注释信息训练事件抽取器，应用于目标语言。Liu等人 (Liu and Chen et al., 2019)设计一种上下文相关的翻译方法来构建不同语言之间的词汇映射，同时利用共享句法顺利的方法处理词序差异问题，进而实现不同语言间的知识转移。

综上所述，对于相似语系，应用翻译和迁移的方法可以取得较好效果。然而，受机器翻译

性能和越南语音节特性影响，传统模型在越南语事件检测上存在不足。因此，根据表达相同观点但语言不同的句子通常有相同或相似的语义成分这一特征，提出一种基于中文信息与越南语句法指导的越南语事件检测框架，利用中文语句以及所标注的中文事件类型信息辅助越南语事件检测，从而减少越南语事件检测对越南语数据标注信息的依赖。

3 方法

针对越南语事件检测任务，本文提出了一个融合中文信息和越南语句法的越南语事件检测模型，该模型结合了中文信息与越南语依存句法信息，并将其融入到越南语语句中，以指导越南语事件检测。模型的体系结构如图 1所示：

本文模型有三部分组成：双语信息融合模块，图卷积模块和事件检测器。1.双语信息融合模块：主要由共享编码器网络和交叉注意力网络两个网络模块组成。1）共享编码器网络首先通过编码器对中文语句进行编码并获取到中文编码隐层向量表示和中文句子级向量表示，然后越南语词级向量与中文句子级向量融合，再通过共享编码器策略获取越南语隐层向量和越南语句子级向量。2）交叉注意力网络将获得中文隐层向量与越南语隐层向量进行联合学习，得到融合中文词级信息的越南语向量表示。2.句法图卷积模块将越南语向量表示与越南语依存句法信息进行联合学习，得到融合依存句法信息的向量表示。3. 最后通过事件检测器中的事件类型感知网络实现基于中文事件类型信息的越南语语义表示，以完成越南语事件检测。

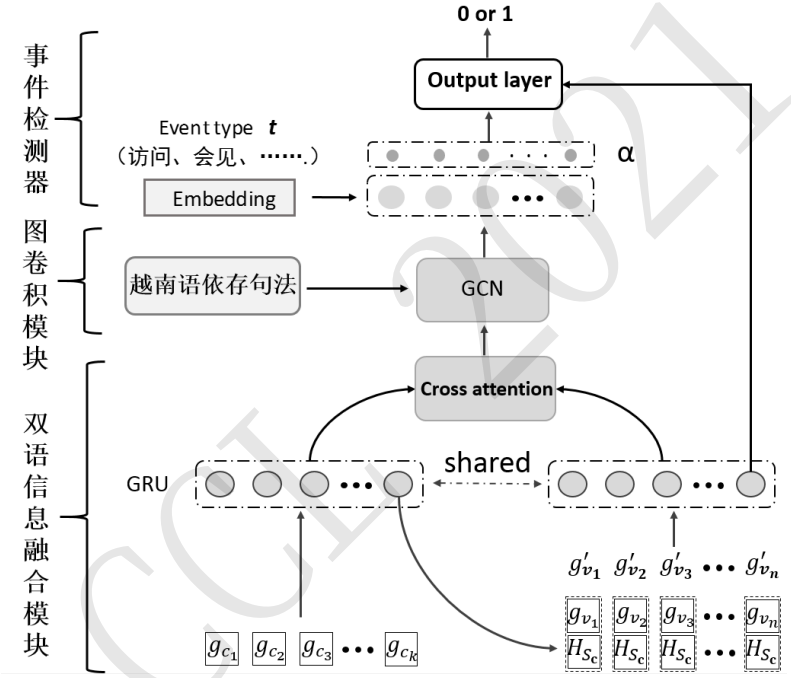


图 1. 模型结构图

3.1 双语信息融合模块

本文实验模型使用编码器对中越双语语句进行向量化表示在此基础上，本文提出共享编码器策略，然后通过交叉注意力网络获得最终的越南语向量表示。

共享编码器策略.当编码器在对汉语、越南语进行编码训练的过程时，为了将汉越语言训练过程中的语义距离减小，我们将编码器生成的参数进行共享。首先通过GRU网络读取输入的汉语词向量c_i与实体向量e_{c_i}所构成的$g_{c_i} = [c_i; e_{c_i}]$，获得中文语句隐层$h_i^{(c)}$:

$$h_i^{(c)} = GRU(g_{c_i}; h_{i-1}^{(c)}) \tag{1}$$

由中文语句子输入，可将GRU编码器最后的输出作为句子级向量表示H_{S_c}:

$$H_{s_c} = h_k^{(c)} \tag{2}$$

通过中文语义信息的融入，可使越南语能关注到更多事件信息。因此，越南语词向量v_j和实体向量e_{v_j}所构成$g_{v_j} = \left[v_j; e_{v_j}\right]$,再与中文句子级向量$H_{S_c}$构成$g'_{v_j} = [g_{v_j}; H_{S_c}]$,通过共享编码器策略得到越南语隐层$h_j^{(v)}$:

$$h_j^{(v)} = GRU(g_{v_j}'; h_{j-1}^{(v)}) \tag{3}$$

由越南语句子输入，可将GRU编码器最后的输出作为句子级向量表示H_{S_v}:

$$H_{s_v} = h_n^{(v)} \tag{4}$$

交叉注意力网络. 在中文和越南语之间使用交叉注意力网络，该网络允许越南语词级隐层状态通过关注中文词级隐层状态来表示，从而得到这种语言的跨语言特征表示，以此达到语义上对齐两种语言的目的。如图 2所示:

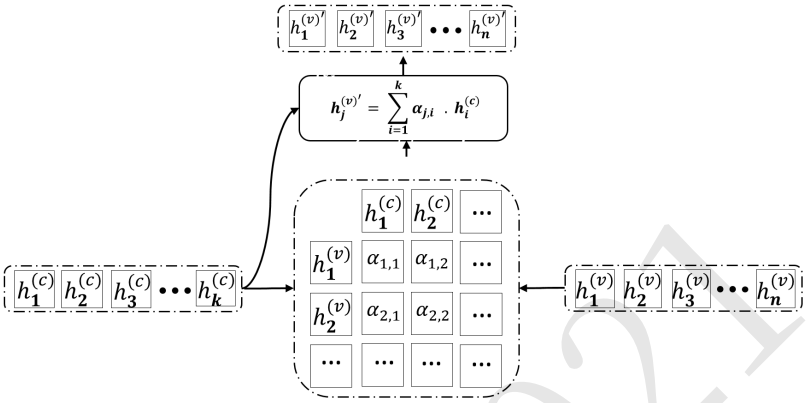


图 2. 交叉注意力网络

由共享编码器策略得到越南语向量$H_V = \left\{h_1^{(v)}, h_2^{(v)}, h_3^{(v)} ... h_n^{(v)}\right\}$，通过中文每一个特征表示$\left\{h_i^{(c)}\right\}_{i=1}^{k}$的加权值来表示越南语第$j$个词的特征表示$h_j^{(v)'}$:

$$h_j^{(v)'} = \sum_{i-1}^{k} \alpha_{j.i} \cdot h_i^{(c)} \tag{5}$$

注意力权重$\alpha_{i,j}$是通过相应的匹配分数$m_{j,i}$上计算$Softmax$函数而获得的。其中匹配分数又是基于特征向量$h_i^{(c)}$和$h_j^{(v)}$的双线性乘积来计算的:

$$\alpha_{j,i} = \frac{e^{m_{j,i}}}{\sum_{x=1}^{n} e^{m_{j,x}}} \tag{6}$$

$$m_{j,i} = tanh(h_j^{(v)} W h_i^{(c)} + b) \tag{7}$$

其中，$W \in R^{n \times n}$和$b \in R$是其训练的注意力参数。通过交叉注意力网络，将中文语句放在越南语语句的上下文嵌入中，可以进一步使越南语关注到中文事件信息。

3.2 句法图卷积模块

依存关系构建.为了对关键信息的有效聚合，本文采用 (Vu et al., 2018)越南语开源依存句法工具来构建依存句法关系。由无向图$\zeta = (\gamma, \varepsilon)$作为越南语句子$S_V$的句法分析树，其中$\gamma = \nu_1, \nu_2, \nu_3 ... \nu_n (|\gamma| = n)$和$\varepsilon$分别是节点集和边集。在$\gamma$中，每个$\nu_i$是表示单词$w_i$和$S_V$ 的节点，每个边$(\nu_i, \nu_j) \in \varepsilon$是来自单词$w_i$到单词$w_j$的有向句法弧，类型标签为$K(w_i, w_j)$。此外，为了让信息朝相反的方向流动，本文还添加了带有类型标签$K'(w_i, w_j)$的反向边(ν_j, ν_i)和自循环，即任何$\nu_i \in \gamma$的(ν_i, ν_i)。所以最终得到标签$K(w_i, w_j)$的三种类型表示为:

$$K(w_i, w_j) = \begin{cases} along & (\nu_i, \nu_j) \in \varepsilon \\ rev & i! = j \ \ and \ \ (\nu_j, \nu_i) \in \varepsilon \\ loop & i == j \end{cases} \tag{8}$$

句法图卷积模块被设计用来捕获句法依存之间的关系，通过依存关系中边的类型标签构建邻接矩阵，应用经过公式 5表示的越南语词级表征$h_j^{(v)'}$作为网络的输入，初始化网络第一层的节点表示$\bar{h}_{v_j}^0$；在句法图卷积网络模块的第k层，可以通过以下方法计算节点$v \in \gamma$的图卷积向量$\bar{h}_{v_j}^{k+1}$：

$$\bar{h}_{v_j}^{k+1} = f(\sum_{u \in N(v)} (W_{k(u,v)}^{(k)} \bar{h}_{v_i}^k + b_{k(u,v)}^{(k)})) \tag{9}$$

其中$K(u,v)$表示边(u,v)的类型标签所构建的邻接矩阵；$W_{K(u,v)}^{(k)}$和$b_{K(u,v)}^{(k)}$分别是某个类型标签$K(u,v)$的权值矩阵和偏差；$N(v)$是v的邻域集，包括v（由于自循环）；f是非线性激活函数。

3.3 事件检测器

事件触发词是这项任务的重要线索。例如，死亡事件通常由“死”、“去世”等词触发。但是，这个信息隐藏在我们的任务中，因为带注释的触发词没有标注，为了对隐藏的触发词进行建模，本文在方法中引入事件感知网络。首先将中越语句标记为0或1，如表 1所示，假设有三个预定义的事件类型（用T_1、T_2和T_3表示），而根据中文语句判断包含T_1事件，然后将中越双语语句转换为三个二分类实例：

实例	标签
[(C,V) , T_1]	1
[(C,V) , T_2]	0
[(C,V) , T_3]	0

表 1. 中越双语语句二分类实例

根据汉语语句所标记的目标事件类型得到嵌入向量t_1进行打分，以此感知触发的事件类型。在下面的等式中，特别是通过注意来计算第k-h个隐状态的分数，以使目标事件类型的触发词比其他词获得更高的分数：

$$\alpha^k = \frac{exp(\bar{h}_{v_k} \cdot t_1^T)}{\sum_i exp(\bar{h}_{v_j} \cdot t_1^T)} \tag{10}$$

最后，可得经过分数评估之后越南语句子表示S_{att}：

$$S_{att} = \alpha^T \overline{H}_v \tag{11}$$

其中$\alpha = [\alpha^1, \alpha^2, \alpha^3 ... \alpha^n]$是注意力的矢量，$\overline{H}_v = [\bar{h}_{v_1}, \bar{h}_{v_2}, \bar{h}_{v_3} \ldots \bar{h}_{v_n}]$是越南语语句向量矩阵。

根据所得越南语句子的表示，最终输出O连接到两个组件：v_{att}和$v_{\mathbf{global}}$。如图 3所示：

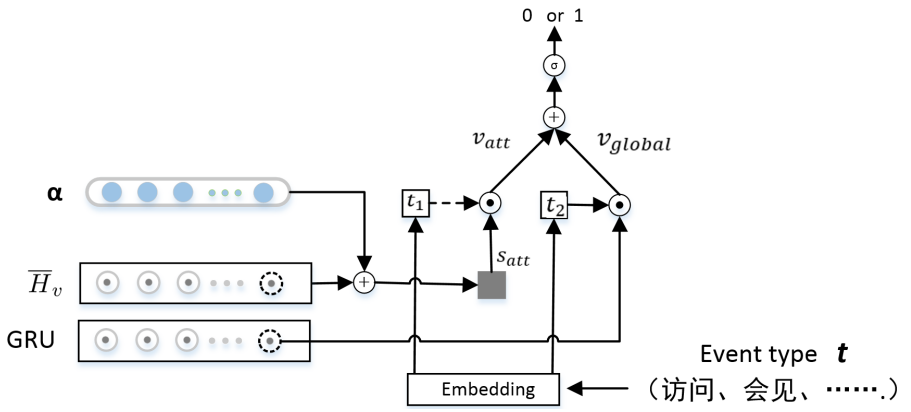


图 3. 事件检测器模块

一方面，v_{att}是由s_{att}和t_1点积计算所得，它被设计用来捕捉越南语本地特征。另一方面，越南语通过共享编码器策略得到越南语句子向量表示H_{S_v}，因此$v_{\mathbf{global}} = H_{s_v} \cdot t_2^T$是捕捉越南语句子的整体特征信息。最后，O是被定义$v_{att}$和$v_{\mathbf{global}}$的加权和：

$$O = \sigma(\lambda \cdot v_{att} + (1-\lambda) \cdot v_{global}) \tag{12}$$

这里σ是$sigmoid$函数，$\lambda \in [0,1]$是一个用于权衡v_{att}和$v_{\mathbf{global}}$的超参数。

3.4 偏置损失函数

由于每个训练样本都是一个[(C , V) , T]，根据给定的句对是否传递一个T类型的事件使其标签是1或0。例如，我们总共有7个目标事件类型，如果由中文句子标注一个事件类型，那么它将被转换成6个负样本和1个正样本,因此负样本数量比正样本多，于是我们通过一个偏置损失函数来加强正样本影响。

$$J(\theta) = \frac{1}{T}\sum_{i=1}^{T}(o(x^i) - y^i)^2(1 + y^i \cdot \beta) + \delta \left\|\theta\right\|^2 \tag{13}$$

其中x是由汉越句对和中文标注的事件类型组成的一对，$y \in [0,1]$，θ是我们模型的参数，$\delta > 0$是L2规范化项的权重。$(1 + y^i \cdot \beta)$是偏差项。具体而言，该项的值对于负样本（y^i为0）为1，对于阳性样本（y^i为1）为1+β，其中$\beta \geqslant 0$。

3.5 训练

本文通过随机梯度下降（SGD）来训练该模型，正则化由L_2实现。对于x，模型给它分配了一个标签$\tilde{y}$, 根据下式：

$$\tilde{y} = \begin{cases} 0 & o(x) < 0.5 \\ 1 & otherwise \end{cases} \tag{14}$$

这里的是一对[(C , V) , T]，是模型对于的输出，$\tilde{y}$是最终的预测结果。

4 实验

4.1 实验数据

本文实验的数据集是通过爬取汉越双语新闻网站的新闻文本和维基百科中的汉语-越南语翻译文章获得。首先对所获得的文本数据分句处理，然后经过手动对齐，获得19K个汉越可比语料对，并从中选取18K个汉越可比语料对作为训练集，选取1124个汉越可比语料作为测试集，并根据事件检测任务中通用的ACE2005数据集的格式对其进行标注[1]。本文构建的语料中划分了7种事件类型和1种非事件类型，如表 2所示：

事件类型	事件触发词
访问（chuyến thăm）	拜访,出访,考察,探访...
会见（Gặp）	接待,见面,接见,会谈...
合作（Tiếp xúc）	联合,交流,合作,合同...
经济（Thuộc kinh tế）	衰退,下降,上升...
换届（Thay đổi）	推举,选举,推选,投票选....
贸易（Giao dịch）	出口,进口,转让...
冲突（Cuộc xung đột）	争端,对抗,侵犯,冲突...

表 2. 事件类型及触发词

1：https://github.com/yitiaoyu-bot/data

4.2 评价指标

本文使用精确值（P）、召回率（R）和来评估结果。
精确值P（Precision）：正确预测的事件在总预测为事件中所占的比例。
召回率R（Recall）：正确预测的事件在全部实际为事件的总数中的比例。
F1-measure: $\frac{2\times P\times R}{P+R}$

4.3 参数设置

模型的参数设置如下：使用300维作为汉越双语语句的单词嵌入，使用50维作为汉越双语语句的实体嵌入，每一批次大小为100。L2设置为1e-5。偏差项中的β为1.0。此外，本文通过实验研究公式 12中参数λ对模型性能影响，如下图 4所示，当λ为0.3时，模型性能最优。

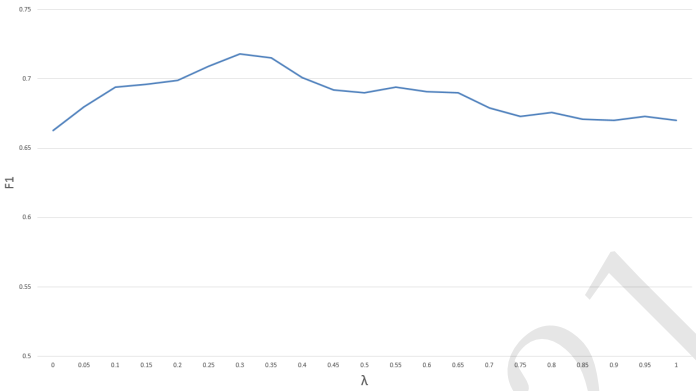


图 4. 设置不同λ值的实验结果

5 实验结果及分析

5.1 模型图卷积层数对模型性能的影响

模型	评价指标		
	层数	P(%)	R(%)	F1(%)
GCN	1	77.5	66.9	71.8
	2	74.3	67.0	70.4
	3	73.8	66.8	70.1

表 3. 图卷积层数实验结果

图卷积层数实验如表 3所示，当图卷积的层数为1时模型达到最佳效果，随着层数的增加性能均有所下降。因此，在后续的实验中，模型均采用一层图卷积。

5.2 参数对比实验

模型	评价指标		
	P(%)	R(%)	F1(%)
100dim	73.0	64.0	68.2
200dim	73.5	65.2	69.1
300dim	77.5	66.9	71.8
512dim	75.9	66.3	70.8

表 4. Embedding维度实验结果

为验证本模型的嵌入向量维度对性能的影响,如表 4所示，当嵌入维度为100和200维度时，此时模型性能都所下降。由于嵌入维度比较低，其语意信息表达能力不足，从而影响模型性能。当嵌入维度比较高的时候，此时会模型出现过拟合趋势，从而使模型性能开始下降。所以，最后模型嵌入选定为300维为最优。

5.3 对比实验

为验证本文方法对越南语事件检测的效果，共选择四个基线方法进行试验。

LSTM-ED：该方法是通过共享LSTM编码器的方式和事件检测模块联合进行事件检测。

CNN-ED：该方法是利用卷积神经网络对句法特征进行提取和融入，以完成事件检测。

DPCNN：该方法使用翻译的方式进行触发词标记，进而训练事件检测模型。

TBNNAM： 该模型是基于目标事件类型计算句子的表示实现无触发词的事件检测。

模型	评价指标		
	P(%)	R(%)	F1(%)
LSTM-ED	72.1	67.0	69.4
CNN-ED	72.7	65.9	69.1
DPCNN	73.6	64.2	68.5
TBNNAM	65.2	60.0	62.4
Ours	77.5	66.9	71.8

表 5. 对比实验结果

表 5显示了事件检测的F1值结果。第一部分LSTM-ED、CNN-ED模型是融入中文信息和中文事件类型信息的事件检测模型，第二部分DPCNN、TBNNAM是传统单语事件检测模型，第三部分Ours是我们的实验模型。通过对比实验可知，本文模型的F1值均超过其他基线模型。

1）对比DPCNN和TBNNAM模型分析得出，虽然使用翻译的方式扩充越南语语料在DPCNN和TBNNAM模型进行训练取得一定效果，但是本模型在中文语句的辅助下越南语事件检测能取得更好的效果。本文模型通过中文语义信息的融入，充分利用中文标注的事件类型信息辅助越南语事件检测，从而提升了越南语事件检测的性能。

2) 对比LSTM-ED和CNN-ED模型，结果表明捕获语句的特征信息有助于提高模型的事件检测性能。本文模型使用共享GRU编码器，使得参数更少因此更容易收敛，比共享LSTM编码器取得了更好的效果。本文模型与CNN-ED对比表明，使用图卷积模块可以捕捉到CNN未能捕捉到的越南语句法关系特征，通过融入越南语句法关系特征有效提高了事件检测的准确性。

权重可视化：事件感知网络中注意力向量权重α权重可视化。如图 5所示，对于例句：

V: Việt Nam sẵn sàng nỗ lực vì sự kinh tế phát triển, thương mại của hai nước và cùng nhau thúc đẩy trao đổi song phương trong lĩnh vực này.

C:越方愿为两国经贸发展做出努力，共同推动双方在这个领域的交流。

Type:合作

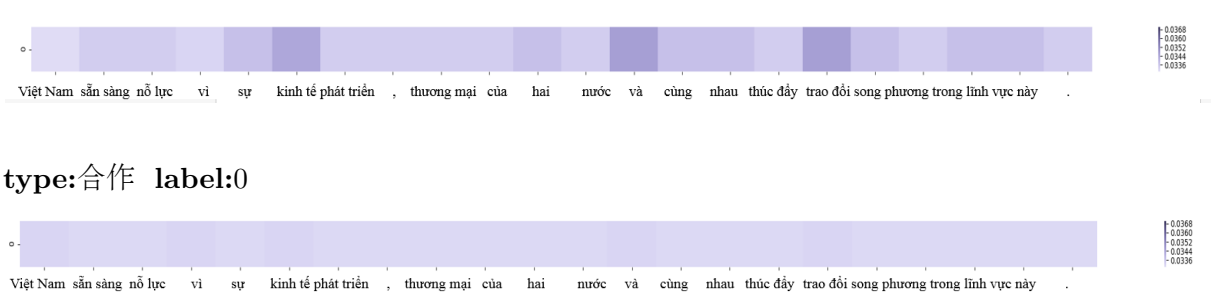


图 5. 模型学习样本注意力权重α的可视化

图 5展示了模型学习到的注意力向量α的例子。通过事件类型感知网络，例句由“合作”事件成功捕捉到“kinh tế（经贸）”、“và（和）”和“trao đổi（交流）”这三个特性，并将其分配了较大的注意力分数。对于负样本，由于没有关键线索，模型给每个单词分配了几乎相等的注意力分数。

5.4 消融实验

为验证本模型的中文语义信息、句法图卷积模块、以及交叉注意力机制的有效性，如下表所示（特表说明“（-）”表示未使用该网络结构）：

模型	评价指标		
	P(%)	R(%)	F1(%)
(-)Chinese-sen	66.8	58.2	62.3
(-)GCN	73.8	63.1	68.0
(-)cr-attention	75.2	64.5	69.4
OUR	77.5	66.9	71.8

表 6. 消融实验结果

消融实验结果如表 6所示，本文模型的P 、R 、F1 值均超过其他基线模型。当模型未融入中文信息时，模型P 、R 、F1均有所下降，因此可以证明中文语义信息的融入，可以使模型更好的辅助越南语进行事件检测，这对模型检测越南语事件的性能提升是非常重要的，也是模型中必不可少的一部分。同时对比交叉注意力机制，由于没有进行对中文语句深层次信息融合，模型性能也有所下降，说明交叉注意力机制对于模型提升也有帮助。当模型未使用句法图卷积模块时，模型的P 、R 、F1 均在下降，因此可以证明句法图卷积模块可以有效的对句子深层次的特征进行提取和融入，可以更好的辅助模型对越南语事件信息的准确识别。

5.5 数据质量对比实验

为验证数据质量对本模型的性能的影响，本文将输入的可比语料句对按照比例随机打乱（0.05、0.1和0.15的比例）生成一定比例的噪声数据。其实验结果如表 7所示：

噪声数据比例	评价指标		
	P(%)	R(%)	F1(%)
0	77.5	66.9	71.8
0.05	75.6	64.7	69.7
0.10	74.3	67.5	70.7
0.15	73.4	65.0	69.0

表 7. 数据质量对比实验结果

由表 7实验结果可知，当以0.05的比例打乱数据时，R 值降低，说明少量的的噪声数据影响模型对事件句的识别，使得模型总体性能下降。当以0.1的比例打乱数据时，R值上升而P值下降，说明由于训练数据噪声的影响，使得模型的泛化能力得到了提高，进而对事件句的判断更加准确，但对非事件句造成了较大误判。当以0.15的比例打乱数据时，模型的性能大幅度下降，这是由于噪声数据比例的不断增加，其对模型的影响不断叠加，从而影响模型对事件句的判断，使得模型性能下降。

6 总结

本文提出了一个融合中文信息与越南语句法的越南语事件检测模型。该模型结合了中文的语义信息和越南语句法信息，通过融入中文语句的语义信息，使得中文标注的事件类型可以指导越南语事件检测，从而解决越南语数据标注稀缺的问题。同时，本文模型利用图卷积网络对

越南语句法特征进行提取并将其融入越南语中，提高了越南语事件检测的准确性。通过对比实验，证明了加入中文信息与越南语句法信息可以有效的提升越南语事件检测效果，其实验结果证明了本文模型的有效性。

参考文献

Qi Li and Heng Ji and Liang Huang. 2013. Sofia. Joint event extraction via structured prediction with global features. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics.*Volume 1 (Long Papers): 73–82.

Shulin Liu and Yang Li and Feng Zhang and Tao Yang and Xinpeng Zhou. 2019. Minneapolis. Event Detection without Triggers. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational linguistics : Human Language Technologies.* Volume 1 (Long and Short Papers): 735–744.

Zhu Zhu and Shoushan Li and Guodong Zhou and Rui Xia. 2014. Baltimore. Bilingual Event Extraction: a Case Study on Trigger Type De-termination. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics.* Volume 2 (Short Papers): 842–847.

Zheng Chen and Heng Ji. 2009. Boulder, Colorado. Can one language bootstrap the other: A case study on event extraction. *Proceedings of the NAACL HLT 2009 Workshop on Semi-supervised Learning for Natural Language Processing*: pages: 66–74.

Manaal Faruqui and Shankar Kumar. 2015. Denver, Colorado. Multilingual open relation extraction using cross-lingual projection. *The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.* Papers: 1351-1356.

Bowei Zou and Zengzhuang Xu and Yu Hong and Guodong Zhou. 2018. New Mexico. Adversarial feature adaptation for cross-lingual relation classification. *In Proceedings of the 27th International Conference on Computational Linguistics.* Papers: 437–448.

Ryan T. McDonald and Joakim Nivre and Yvonne Quirmbach-Brundage and Yoav Goldberg et al. 2013. Sofia. Universal Dependency Annotation for Multilingual Parsing. *Pro-ceedings of the 51st Annual Meeting of the Association for Computational Linguistics.* Short Papers: 92–97.

Bishan Yang and Tom M. Mitchell. 2016. San Diego California. Joint extraction of events and entities within a document context. *In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics.* Papers: 289–299.

Katherine A. Keith, Abram Handler, Michael Pinkham, et al. 2017. Copenhagen, Denmark. Identifying civilians killed by police with distantly super-vised entity-event extraction. *Proceedings of the 2017 Conference on Empirical Methods in Natural.* Papers: 1547–1557.

Shulin Liu and Yubo Chen and Shizhu He and Kang Liu and Jun Zhao. 2016. Berlin, Germany. Leveraging frame to improve automatic event detection. *In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics.* Papers: 2134–2143.

Thien Huu Nguyen and Kyunghyun Cho and Ralph Grishman. 2016. San Diego California. Joint event extraction via recurrent neural networks. *In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.* Papers: 300–309.

Shasha Liao and Ralph Grishman. 2010. Uppsala, Sweden. Using document level cross-event inference to improve event extraction. *In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics.* Papers: 789–797.

Heng Ji and Ralph Grishman. 2008. Columbus,Ohio. Refining event extraction through cross-document inference. *In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics.* Papers: 254–262.

Thien Huu Nguyen and Ralph Grishman. 2016. Austin, Texas. Modeling skip-grams for event detection with convolutional neural networks. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics.* Papers: 886–891.

Shulin Liu and Kang Liu and Shizhu He and Jun Zhao. 2016. Phoenix, Arizona. A probabilistic soft logic based approach to exploiting latent and global information in event classification. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence.* Papers: 2993–2999.

Shulin Liu and Yubo Chen and Kang Liu and Jun Zhao. 2017. Vancouver, Canada. Exploiting argument information to improve event detection via supervised attention mechanisms. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics.* Papers: 1789–1798.

Thien Huu Nguyen and Ralph Grishman. 2015. Beijing. Event detection and domain adaptation with convolutional neural networks. *In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing.* Volume 2 (Short Papers): 365–371.

Andrew Hsi and Yiming Yang and Jaime G. Carbonell and Ruochen Xu. 2016. Leveraging multilingual training for limited resource event extraction. *The COLING 2016 Organizing Committee. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics.* Papers: 1201–1210.

Jiang Guo and Wanxiang Che and David Yarowsky and Haifeng Wang and Ting Liu. 2015. Cross-lingual dependency parsing based on distributed representations. *Association for Computational Linguistics. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing.* Volume 1(Long Papers): 1234–1244.

Stephen Mayhew and Chen-Tse Tsai and Dan Roth. 2017. Cheap translation for cross-lingual named entity recognition. *In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.* Papers: 2536–2545.

Jian Ni and Georgiana Dinu and Radu Florian. 2017. Weakly supervised cross-lingual named entity recognition via effective annotation and representation projection. *In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics.* Volume 1(Long Papers):1470–1480.

Thanh Vu, Dat Quoc Nguyen, Dai Quoc Nguyen, et al. 2018. Vn CoreNLP: A Vietnamese Natural Language Processing Toolkit. *In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations.* Papers:56-60.

Yunmo Chen and Tongfei Chen and Seth Ebner and Aaron Steven White and Benjamin Van Durme. 2020. Reading the Manual: Event Extraction as Definition Comprehension. *Proceedings of the Fourth Workshop on Structured Prediction for NLP .* Papers:74–83.

Thien Huu Nguyen and Ralph Grishman. 2018. Graph Convolutional Networks with Argument-Aware Pooling for Event Detection. *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence,(AAAI-18), the 30th innovative Applications of Artificial Intelligence(IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18) .* Papers: 5900–5907.

Du xinya, and Claire Cardie. 2020. Event Extraction by Answering (Almost) Natural Questions. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing .* Papers:671–683.

Xiao Liu 、Zhunchen Luo and Heyan Huang 2018. Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium.* Papers:1247–1256.

Jian Liu 、Yubo Chen、Kang Liu and Jun Zhao. 2018. Event Detection via Gated Multilingual Attention Mechanism. *Jian Liu and Yubo Chen and Kang Liu and Jun Zhao.2018. Event Detection via Gated Multilingual Attention Mechanism. Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence,(AAAI-18), the 30th innovative Applications of Artificial Intelligence(IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18).* Papers:4865–4872.

Ananya Subburathinam and Di Lu and Heng Ji and Jonathan May et al. 2019. Cross-lingual Structure Transfer for Relation and Event Extraction. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP.* Papers:313–325.

Jian Liu and Yubo Chen and Kang Liu and Jun Zhao. 2019. Neural Cross-Lingual Event Detection with Minimal Parallel Resources. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP.* Papers:738—748.

Yubo Chen and Liheng Xu and Kang Liu and Daojian Zeng and Jun Zhao. 2015. Beijing. Event extraction via dynamic multi-pooling convolutional neural networks. *In Pro-ceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Confer-ence on Natural Language Processing of the Asian Federation of Natural Language Process-ing.* Papers: 167–176.