

# 基于卷积神经网络与 SVM 分类器的隐喻识别<sup>\*</sup>

黄孝喜 李晗雨 王荣波 王小华 谌志群

(杭州电子科技大学认知与智能计算研究所 杭州 310018)

**摘要:**【目的】针对中英文的隐喻数据集,提出一种基于卷积神经网络与 SVM 分类器的隐喻识别方法。【方法】将实验数据向量化,结合词性特征和关键词特征作为卷积神经网络的输入,通过卷积层和池化层提取特征,应用 SVM 进行分类。针对卷积神经网络的池化层中特征采样的不完全性,提出将 MaxPooling 与 MeanPooling 组合在一起的改进方法。【结果】相对于直接使用卷积神经网络,利用本文方法进行隐喻识别的准确率在英文动宾语料、英文形容词-名词词组语料和中文隐喻语料分别提高 4.12%、0.84% 和 4.50%。【局限】中文分词不准确,影响词向量模型训练;卷积神经网络的层数过少,影响特征的完整性。【结论】根据中英文数据集上隐喻识别的结果分析,该方法在两个数据集上都取得了良好效果。

**关键词:** 隐喻识别 卷积神经网络 支持向量机 特征提取

**分类号:** TP391

**DOI:** 10.11925/infotech.2096-3467.2018.0114

## 1 引言

隐喻不仅仅是一种修辞手段,更是人类的重要认知方式,如“他的眼睛是星星”和“天上亮的是星星”等句子中的“星星”意义不同。第一句中的“星星”使用隐喻表示“眼睛是明亮的”;第二句中的“星星”指肉眼可见的宇宙中的天体。隐喻是从源域到目标域的认知映射,如“时间就是金钱”中,时间是源域,金钱是目标域。虽然在某些研究中,识别这些映射(概念隐喻)本身就是分析主题,但是更重要的是在文本(语言隐喻)中检测它们的表现形式。

隐喻研究在自然语言处理(NLP)中起到重要作用,如机器翻译、信息检索和问答系统等。作为自然语言中的一种普遍现象和认知方式,隐喻识别和解释的研究不仅吸引了语言学家,而且也引起了认知科学家的关注。其中隐喻识别主要是在语料中自动地识别出含

有隐喻的短语或句子。根据隐喻内容的语法特征,可以将隐喻分为三类:名词隐喻、动词隐喻和形容词隐喻。

(1) 名词隐喻,名词在句子中作为主语、宾语等成分,最基本的形式是通过动词“是”相连,如“他的眼睛是星星”和“香蕉的味道是淡淡的甜味”中“是”可以作为喻词,“甲是乙”表示通过乙来描述甲。第一句为隐喻表达,喻词为“是”,本体为眼睛,喻体为星星,眼睛与星星语义相似值较小。第二句具有相同的结构,但并不是隐喻表达,“是”作为谓语,并不是喻词,主语与宾语是“味道”和“甜味”,具有相似的语义,则并不是本体与喻体。

(2) 动词隐喻,动词作用于名词,如“I kill a process.”中“kill”通常作用于有生命的物体,而“process”是无生命的、抽象的,“杀死进程”解释不通,是隐喻表达。

(3) 形容词隐喻,形容词作用于名词,如“Bright

通讯作者: 李晗雨, ORCID: 0000-0003-0106-8552, E-mail: lhu3492216@163.com。

<sup>\*</sup>本文系教育部人文社会科学研究规划基金项目“融合深度神经网络模型的汉语隐喻计算研究”(项目编号: 18YJA740016)和教育部人文社会科学研究青年基金项目“基于语义相关性的汉语组块切分模型研究”(项目编号: 12YJCZH201)的研究成果之一。

painting/bright idea”和“Heavy table/heavy feeling”是形容词-名词词组, 形容词通常形容事物的物理特性, 但也可通过尺寸或重量等感官特征描述抽象特征, 当形容词修饰抽象特征时, 词组为隐喻表达。“painting”和“table”是具体事物, 则“bright painting”和“heavy table”是字面表达; “idea”和“feeling”是抽象的感官特征, 则“bright idea”和“heavy feeling”是隐喻表达。

近年来, 卷积神经网络被用于解决自然语言处理的各项任务, 并取得较好的效果。但迄今为止, 还没有应用到隐喻识别中。隐喻识别是隐喻机器处理的基本问题, 本文旨在提出一种新颖的结合卷积神经网络和SVM分类器的隐喻识别方法, 在中英文数据上都取得了较好的效果。

## 2 相关工作

传统的隐喻识别任务包括基于语义的隐喻识别方法和基于统计的隐喻识别方法。Wilks<sup>[1]</sup>认为隐喻会导致语义上的异常中断, 则在选择优先模型的基础上增加隐喻识别, 并加入带有辅助理解的情景知识, 通过语义选择优先的异常中断触发进行隐喻识别。Fass<sup>[2]</sup>在选择优先中断的基础上, 提出修正语义学, 并给出一个隐喻解释性语言的系统 Met\*, 但是 Met\*系统是手工构造的语料库, 在实际应用上受到限制。Neuman 等<sup>[3]</sup>根据词语在 WordNet 中的类别、搭配信息和名词的抽象度以及常用搭配等信息, 针对不同类别的隐喻, 设计对应的识别算法, 其中名词隐喻识别算法的准确率相对较高。Shutova 等<sup>[4]</sup>结合概念隐喻理论和语义优先理论, 提出对动词和名词进行聚类识别隐喻, 以英国国家语料库(BNC)中部分已经标注的动宾结构的隐喻作为种子集合, 将动词聚类获得源域概念集合, 名词聚类获得目标域概念集合, 并将动词与名词聚类构成一个映射关系。聚类使用谱聚类方法, 按照优先强度方法进行筛选, 去除优先度弱的词语, 最终得到目标域和源域的集合。这种方法与基于 WordNet 的隐喻识别方法有一定的差异, 利用聚类算法能形成更大的动词集合, 适用于更大范围的隐喻表达, 但是这种方法仅适用于“V+N”形式的隐喻表达。Hovy 等<sup>[5]</sup>使用树核的方法实现隐喻识别, 通过句法依存关系(本体与喻体)建立隐喻树, 利用统计模型找出树的差异点, 进而找出隐喻。选取依存树的语义和词

义作为特征, 通过 SVM 和 CRF 进行训练, 发现 SVM 模型在该算法中有更好的分类效果。Rai 等<sup>[6]</sup>应用条件随机场(CRF)进行隐喻识别, 并结合概念特征、上下文特征、句法特征和情感特征, 证明不同的特征对隐喻识别的影响, 在这种方法中, 情感特征对隐喻识别并没有显著影响。Tsvetkov 等<sup>[7]</sup>提出使用随机森林分类器识别不同语言(西班牙语、波斯语和俄语)的隐喻, 该方法使用独立语言(而不是词汇或特定语言)的概念特征, 加入抽象性和可成像性特征, 并通过空间向量进行潜在语义分析找出词语间的关系特征, 通过随机森林分类器进行隐喻识别。

近年, 较多研究基于深度学习方法对文本进行分析, Kalchbrenner 等<sup>[8]</sup>将卷积神经网络应用于自然语言处理, 设计一个动态卷积神经网络(DCNN)模型, 以处理不同长度的文本。Graves 等<sup>[9-10]</sup>提出 BLSTM, 采用双向 LSTM 提取文本上下文的双向特征, 进行分类及识别任务, 区别于 LSTM 只能考虑上文信息而不考虑下文信息, 根据建模目标对象的上下文信息, 设计 BLSTM。也有部分研究将深度学习应用到隐喻的机器处理上, Dinh 等<sup>[11]</sup>使用词向量结合多层感知器(完全连接的前馈神经网络)的架构, 包含一个输入层、多个隐藏层和一个输出层, 并且添加窗口大小为 5 的上下文特征, 对隐喻识别有较好的效果, 但是并没有使用不同语言以及应用更先进的神经网络进行实验。Bizzoni 等<sup>[12]</sup>应用基本神经网络模型, 结合预先训练的词向量模型, 并且在不同数量的数据集和不同的词向量模型上, 对名词与形容词形式的隐喻词组进行隐喻识别, 但没用应用复杂的神经网络模型; Rei 等<sup>[13]</sup>应用监督的语义相似性网络进行隐喻识别, 提出的框架优于前馈神经网络的识别结果, 网络中的源域和目标域之间的门控功能, 通过 Sigmoid 函数充当过滤器, 选出有用信息从而使隐喻识别效果更好, 但是语义相似性网络在处理隐喻句子上存在缺陷。

汉语隐喻研究发展相对较晚, 例如, 王治敏等<sup>[14]</sup>和徐扬<sup>[15]</sup>利用最大熵模型对名词性隐喻进行识别; 李斌等<sup>[16]</sup>通过最大熵模型结合 CRF 解决明喻问题, 并识别出本体和喻体及其相似点; 黄孝喜<sup>[17]</sup>利用依存句法分析得出隐喻句的各种依存模式, 提出一种模式匹配算法, 实现隐喻识别。

综上所述, 目前相关研究主要是基于机器学习和

基于简单的神经网络进行隐喻识别,而基于深度学习的中文隐喻识别的研究尚少,与之相比,本文方法使用卷积神经网络,没有复杂结构,仅仅使用输入层、卷积层和池化层,对不同语言、不同格式的隐喻表达进行特征提取,如短语和句子格式的隐喻表达。笔者使用非静态的词向量,而不像 Kim<sup>[18]</sup>在 CNN 模型中引入词性特征,仅使用简单的静态词向量和微调超参数,解决一词多义现象。最后根据卷积神经网络提取的特征,通过 SVM 分类器进行隐喻识别。

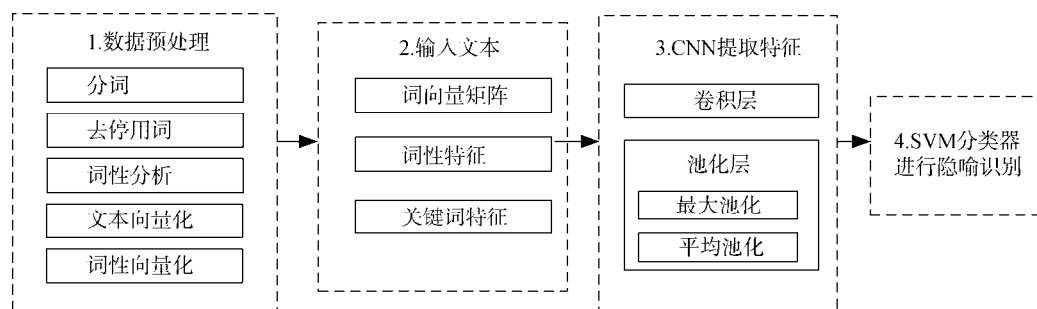


图 1 基于 SVM 与 CNN 的隐喻识别的整体框架

### 3.1 词向量训练

使用神经网络对文本进行处理时,需要将文本转换为计算机可识别的向量表示形式, Mikolov 等<sup>[19-20]</sup>基于 NNLP 提出 Word2Vec 模型,中文词向量模型通过搜狗新闻数据,而英文词向量模型通过 Word2Vec 官方数据进行训练, Word2Vec 包含 C-BOW(Continuous Bag-of-Words) 和 Skip-Grams 两种结构构建词向量,本文应用 Skip-Gram 训练的词向量模型,神经网络的层数为 100,网络层数量意味着输入数据量,也能更高地提升准确率,输入语料为完成分词的数据,对数据进行两次迭代操作:第一次是统计词频构建内部词典数据结构,第二次是进行神经网络训练,这两次步骤是分开进行的。训练窗口大小为 5,词向量维度为 200。Skip-Gram 模型的计算如公式(1)<sup>[20]</sup>所示。

$$p(W_a/W_i) = \frac{e^{U_a V_i}}{\sum_j e^{U_j V_i}} \quad (1)$$

其中,  $V_i$  是 Embedding 层矩阵里的列向量,也被称为  $W_i$  的输入向量,  $U_j$  是 Softmax 层矩阵里的行向量,也被称为  $W_i$  的输出向量。因此, Skip-Gram 模型的本质是计算输入词语的输入向量与目标词的输出向量之间的余弦相似度,并进行 Softmax 归一化。

## 3 研究方法

基于 CNN 与 SVM 进行隐喻识别,对实验语料进行分词、词性分析和去除停用词等数据预处理,利用 Skip-Grams 框架对实验语料进行词向量和 Word/Pos 向量训练,构建文本词向量矩阵,并结合词性特征及关键词特征作为神经网络的输入,通过 CNN 的卷积层和池化层对输入句子进行特征提取,利用 SVM 分类器进行隐喻识别,具体实验框架如图 1 所示。

### 3.2 输入文本

(1) 关键词特征,提取句子的关键词,作为句子的局部特征,同句子的向量矩阵进行拼接,共同作为神经网络的输入。

(2) 词性特征,由于词语存在一词多义的现象,在不同语境中存在不同语义,提出词语与词性的拼接形式训练词性向量,实现词义消歧。给定短文本句子  $S=(Word1, Word2, \dots, Wordn)$ ,  $n$  表示文本句子的长度,  $Word_i$  表示第  $i$  个单词,对应的词性标注为  $P=(Pos_1, Pos_2, \dots, Pos_n)$ ,  $Pos_i$  为第  $i$  个单词对应的词性标注,使用 Word/Pos 形式的输入(如“太阳/ns 好像/v 给/p 浓雾/n 罩住/n 一般/a 变成/v 了/ul 一块/m 圆形/n 水磨/n 玻璃/n 暗淡无光/i”),通过预先训练的 Word/Pos 向量模型映射后, Word/Pos 转换为  $n$  维的向量  $wp_i$ ,则长度为  $m$  的句子为  $m \times n$  维矩阵,作为神经网络的输入。

### 3.3 神经网络结构

卷积神经网络是一个前馈神经网络,一般是对图片进行处理,对文本处理有一定的改动,本文是在经典的卷积神经网络的基础上,即结合卷积层和池化层建立网络,提取特征后再应用 SVM 进行隐喻识别。卷积神经网络结构如图 2 所示。

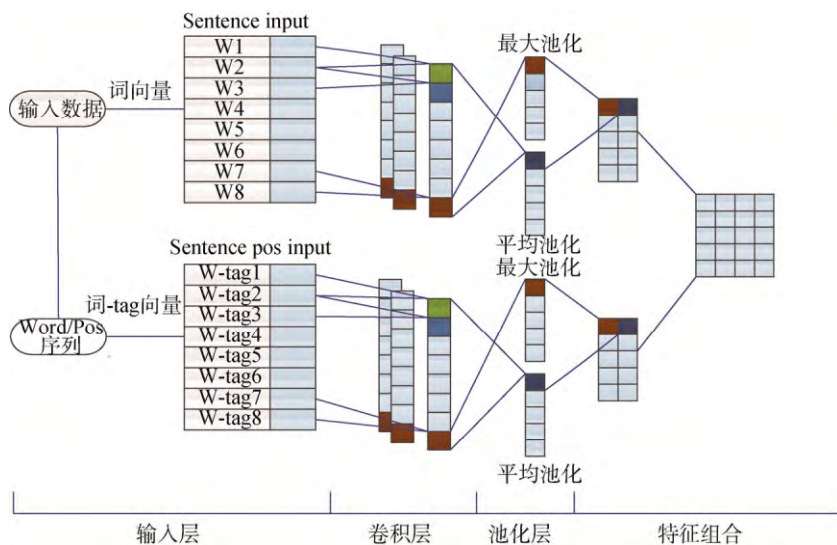


图2 卷积神经网络结构

输入层是以矩阵的形式输入句子，要求每个输入都要相同，则要将句子长度一定，添加零使矩阵大小相同，作为网络的输入。

一般由卷积核共享权值。卷积核一般是随机初始化的，在网络的训练过程中卷积核将学习得到合理的权值，直接好处是减少网络各层之间的连接，减少过拟合的风险。卷积神经网络一般有以下约束<sup>[21]</sup>。

(1) 特征提取，一般是由卷积层作为特征提取层，通过滤波器(卷积核)提取特征，通常选取  $h \times n$  维大小的滤波器在文本矩阵上滑动获取卷积特征， $n$  为词向量的维度， $h$  为每次卷积划过多少个相邻的词。卷积核的大小是为了获取文本的上下文特征，一般卷积神经网络处理文本时，设置不同大小的卷积核，则可以提取更多的特征，但是，卷积核过多会影响网络训练的效率，网络的分类效果可能并没有提高。考虑到网络的训练效率，本文使用 4 个滤波器，主要大小为  $2 \times n$ ,  $3 \times n$ ,  $4 \times n$ ,  $5 \times n$ ，其中  $n$  为词向量的维度。

(2) 特征映射，网络的每个计算层由多个特征映射组成，每个特征映射是一个平面，平面上所有的神经元的权值相等，特征映射结构采用激活函数，常用的激活函数有 Tanh、ReLU 等，Tanh 和 ReLU 函数对比如图3所示，Tanh 是双曲正切函数，当其作为激活函数时，特征映射具有位移不变性；ReLU 具有收敛快、计算简单的优点，但是有使梯度变为 0 的缺点。本文使用 Tanh 函数作为激活函数。

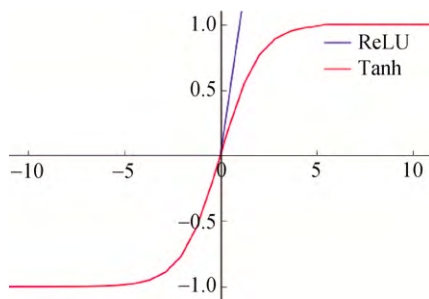


图3 Tanh 和 ReLU 对比

(3) 子抽样层(池化层)，在卷积层后的一个实现局部特征平均和子抽样的计算层，将卷积层得到的特征图进行聚合统计，降低计算的数据量，有效防止过拟合现象。本文的池化层使用两种：MaxPooling 和 MeanPooling，即对卷积层输出的特征集合进行取最大值和平均值的操作，得到文本的最突出的整体特征和平均特征，然后将 MaxPooling 和 MeanPooling 求得的特征值进行级联，作为输入到分类器的特征。

### 3.4 SVM 分类器

支持向量机(SVM)<sup>[22]</sup>是一种二分类模型，基本模型是定义在特征空间上的间隔最大的线性分类器，支持向量机还使用核技巧(Kernel Trick)，使它成为实质上的非线性分类器。核技巧的基本思想是通过一个非线性变换将输入空间对应于一个特征空间，使得在输入空间  $R^n$  中的超平面模型对应于特征空间  $H$  中的超平面模型。核函数如公式(2)<sup>[22]</sup>所示。



$$K(x, z) = \varphi(x) \cdot \varphi(z) \quad (2)$$

其中,  $K(x, z)$  为核函数,  $\varphi(x)$  为映射函数,  $\varphi(x) \cdot \varphi(z)$  为其内积。本文采用高斯核函数(RBF)作为核函数, 如公式(3)<sup>[22]</sup>所示。

$$K(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right) \quad (3)$$

RBF 核函数可将 SVM 中的输入样本映射到高维特征空间上, 解决线性不可分问题。

## 4 实验及结果分析

### 4.1 数据集

Tsvetkov 等<sup>[7]</sup>建立一个英文隐喻识别模型, 通过双语词典的转换方法可识别俄语和葡萄牙语的隐喻, 构建动词(SVO)和形容词-名词词组(AN)的隐喻语料。形容词-名词词组(AN)语料是一个包含 884 个隐喻形容词-名词短语对和 884 对具有字面意义的训练集(TSV-TRAIN)以及一个包含 100 个字面和 100 个隐喻对的测试集。SVO 语料是一个包含 3 737 条手工标注的句子作为训练集, 以及一个包含 111 条隐喻句, 111 条非隐喻句的测试集。其中, 表 1 显示句子中隐喻动词与主语和宾语的关系。表 2 显示 TSV-TRAIN 中形容词名词短语的一部分。

表 1 TSV 中动词隐喻的主语-动词或动词-宾语关系

Verb Noun	Class	Relation
See development	Metaphorical	VO
Live dream	Metaphorical	VO
Envy eat	Metaphorical	SV
Break window	Literal	VO
Boy cry	Literal	SV
Paint dry	Literal	SV

表 2 TSV-TRAIN 中的形容词-名词短语

Metaphorical	Literal
bright smile	blue fence
bushy eyebrows	blinding light
cautious smile	biting dog
dark history	bright sun
deep faith	bright light
desolate beauty	burning tree
economic battle	burning arm
fading memory	dark face
faint impression	dirty hands

笔者使用 TSV-TRAIN 和 SVO 测试集的数据进行英文隐喻识别的实验, 每个数据集的 90% 作为实验的训练语料, 10% 作为测试语料。

实验的中文隐喻语料主要来自《文学比喻词典》和《读者》语料库, 从中选出 2 000 条句子, 其中包含 1 000 条隐喻句子, 1 000 条非隐喻句子, 将每个句子标注为隐喻或非隐喻, 句子形式如“太阳像面火镜”和“他的眼睛是星星”。其中, 训练数据和测试数据比例与英文实验数据相同。

### 4.2 结果与分析

使用迁移学习思想, 通过预先训练的词向量模型得到 200 维的词向量, 在训练词向量过程中, 可能得到语料库中不包含的词, 对其赋予一个 200 维的零向量。实验可能存在不拟合的结果, 使用十次交叉验证法解决此问题。为了证明本文提出的方法对英文数据有效性, 设计 3 组对比实验, 结果如表 3 所示。

表 3 英文语料隐喻识别

实验	准确率
CNN-sentence-eng	81.80%
CNN_SVM -sentence-eng	87.23%
CNN-Word/Pos-eng	86.00%
CNN_SVM -Word/Pos-eng	90.12%
CNN-AN-eng	86.36%
CNN_SVM -AN-eng	87.20%
Rei 等 <sup>[13]</sup>	83.00%

(1) 第一组实验, 使用 TSV 的动词隐喻数据集, 以独立句子作为输入。CNN-sentence-eng 使用卷积神经网络对动词隐喻进行识别; CNN\_SVM-sentence-eng 使用卷积神经网络与 SVM 进行隐喻识别。

(2) 在第一组实验基础上, 第二组实验添加词性特征。CNN-Word/Pos-eng 和 CNN\_SVM-Word/Pos-eng 分别表示直接使用卷积神经网络和使用 CNN-SVM 方法。

(3) 第三组实验, 使用形容词-名词短语(AN)隐喻数据集, 并不是以独立句子作为输入, 而是以短语组合作为输入。CNN\_SVM-AN-eng 为本文方法, 与直接使用卷积神经网络的隐喻识别方法(CNN-AN-eng)以及 Rei 等<sup>[13]</sup>的实验结果进行对比。

从表 3 可看出, 卷积神经网络与 SVM 对隐喻的识

别准确率高于直接使用 CNN 的结果,通过第二组实验,发现添加词性特征后,使用 CNN 或使用 CNN-SVM 的准确率都有提高,但是 CNN-SVM 的结果更好,提高了约 4.12%的准确率。由于卷积神经网络处理文本时,要通过卷积核滑动提取特征,则网络的输入要更大一些,而 TSV 的形容词-名词短语数据集由长度为 2 的词组组成,为了更符合卷积神经网络的输入,第三组实验使用 4 个词组作为一次输入进行训练,在扩大卷积神经网络输入大小的同时,使提取的特征更加多样全面。本文方法在对 AN 形式的隐喻进行识别时,比 Rei 等的实验准确率提高了约 4.2%。

从表 4 中可以看出,本实验对中文数据集同样适用, CNN-SVM 方法的隐喻识别准确率高于直接使用 CNN 的结果。但是中文实验并没有添加词性特征,因此仅仅证明 CNN-SVM 方法的有效性。

表 4 中文语料隐喻识别

实验	准确率
CNN-sentence-ch	72.5%
CNN_SVM -sentence-ch	77.00%

从表 3 和表 4 可以看出,本文方法使用 SVM 分类器,对比实验使用 Softmax 分类器,这是主要区别。而英文隐喻识别的准确率比中文高,这是由于对中文语料分词准确率低于英文语料,影响了中文词性分析和文本词向量训练的准确率,进而影响了中文隐喻识别的效果。

## 5 结 语

本文介绍用于隐喻识别的 CNN-SVM 方法,通过卷积神经网络提取隐喻句子的特征,并应用 SVM 进行分类,实现隐喻分类。此外,提出词与词性结合的方法,有效解决了一词多义现象,并改进卷积神经网络的池化层,使提取的特征更加准确全面。实验对比词向量和 Word/Pos 结合向量分别在 CNN 与 CNN-SVM 上的结果变化,发现 Word/Pos 结合向量在 CNN-SVM 模型上的隐喻识别结果有较大提升。另外, CNN 在中英文语料上的隐喻识别都有较好效果。

在未来工作中,将在文本中加入情感特征,进一步考察情感对隐喻的影响。此外,针对中文语料,在分词方法上要有所改进,以提高实验结果。

## 参考文献:

- [1] Wilks Y. A Preferential, Pattern-seeking, Semantics for Natural Language Inference[A]// Words and Intelligence I. Text, Speech and Language Technology[M]. Springer, 2007.
- [2] Fass D. Met\*: A Method for Discriminating Metonymy and Metaphor by Computer[J]. Computational Linguistics, 1991, 17(1): 49-90.
- [3] Neuman Y, Assaf D, Cohen Y, et al. Metaphor Identification in Large Texts Corpora[J]. PLoS One, 2013, 8(4): e62343.
- [4] Shutova E, Sun L, Korgonen A. Metaphor Identification Using Verb and Nouns Clustering[C]// Proceedings of the 23rd International Conference on Computational Linguistics. 2010.
- [5] Hovy D, Srivastava S, Kumar S, et al. Identifying Metaphorical Word Use with Tree Kernels[C]// Proceedings of the 1st Workshop on Metaphor in NLP. 2013.
- [6] Rai S, Chakraverty S, Tayal D K. Supervised Metaphor Detection Using Conditional Random Fields[C]// Proceedings of the 4th Workshop on Metaphor in NLP. 2016.
- [7] Tsvetkov Y, Boytsov L, Gershman A, et al. Metaphor Detection with Cross-Lingual Model Transfer[C]// Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. 2014.
- [8] Kalchbrenner N, Grefenstette E, Blunsom P. A Convolutional Neural Network for Modelling Sentences[OL]. arXiv Preprint, arXiv: 1404.2188.
- [9] Graves A, Fernández S, Schmidhuber J. Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition[A]// Artificial Neural Networks: Formal Models and Their Applications[M]. Springer, 2005.
- [10] Graves A, Schmidhuber J. Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures[J]. Neural Networks, 2005, 18:602-610.
- [11] Dinh E L D, Gurevych I. Token-Level Metaphor Detection Using Neural Networks[C]// Proceedings of the 4th Workshop on Metaphor in NLP. 2016.
- [12] Bizzoni Y, Chatzikyriakidis S, Ghanimifard M. "Deep" Learning: Detecting Metaphoricity in Adjective-Noun Pairs[C]// Proceedings of the Workshop on Stylistic Variation. 2017: 43-52.
- [13] Rei M, Bulat L, Kiela D, et al. Grasping the Finer Point: A Supervised Similarity Network for Metaphor Detection[C]// Proceedings of EMNLP. 2017: 1537-1546.
- [14] 王治敏, 王厚峰, 俞士汶. 基于机器学习方法的汉语名词隐喻识别[J]. 高技术通讯, 2006, 17(6): 575-580. (Wang Zhimin, Wang Houfeng, Yu Shiwen. Chinese Nominal Metaphor Recognition Based on Machine Learning[J]. Chinese High Technology Letter, 2006, 17(6): 575-580.)

- [15] 徐扬. 基于最大熵模型的汉语隐喻现象识别[J]. 计算机工程 and 科学, 2007, 29(4): 95-103. (Xu Yang. Recognition of the Chinese Metaphor Phenomena Based on the Maximum Entropy Model[J]. Computer Engineering and Science, 2007, 29(4): 95-103.)
- [16] 李斌, 于丽丽, 石民, 等. “像”的明喻计算[J]. 中文信息学报, 2008, 22(6): 27-32. (Li Bin, Yu Lili, Shi Min, et al. Computation of Chinese Simile with “Xiang”[J]. Journal of Chinese Information Processing, 2008, 22(6): 27-32.)
- [17] 黄孝喜. 隐喻机器理解的若干关键问题研究[D]. 杭州: 浙江大学, 2009. (Huang Xiaoxi. Research on Some Key Issues of Metaphor Computation[D]. Hangzhou: Zhejiang University, 2009.)
- [18] Kim Y. Convolutional Neural Networks for Sentence Classification [OL]. arXiv Preprint, arXiv: 1408.5882.
- [19] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations in Vector Space[OL]. arXiv Preprint, arXiv: 1301.3781.
- [20] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and Their Compositionality[A]// Advances in Neural Information Processing Systems[M]. Springer, 2013.
- [21] LéCun Y, Bottou L, Bengio Y, et al. Gradient-based Learning Applied to Document Recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [22] 李航. 统计学习方法[M]. 第 1 版. 北京: 清华大学出版社, 2016: 95-123. (Li Hang. Statistical Learning Method[M]. The 1st Edition. Beijing: Tsinghua University Publishing House, 2016: 95-123.)

### 作者贡献声明:

李晗雨, 黄孝喜: 提出研究思路, 采集分析数据;  
李晗雨, 黄孝喜, 王荣波: 进行实验;  
黄孝喜, 李晗雨, 王小华: 起草及修改论文;  
王荣波, 黄孝喜, 湛志群: 修改论文;  
黄孝喜: 论文最终版本修订。

### 利益冲突声明:

所有作者声明不存在利益冲突关系。

### 支撑数据:

支撑数据由作者自存储, E-mail: lhu3492216@163.com。

- [1] 李晗雨. M\_Result.txt. 中文隐喻句子数据.
- [2] 李晗雨. NM\_Result.txt. 中文非隐喻句子数据.
- [3] 李晗雨. met.txt. 英文形容词-名词短语隐喻数据.
- [4] 李晗雨. nonmet.txt. 英文形容词-名词短语非隐喻数据.
- [5] 李晗雨. txt\_svo\_m.txt. 英文动词(SVO)隐喻句子数据.
- [6] 李晗雨. txt\_svo\_nm.txt. 英文动词(SVO)非隐喻句子数据.
- [7] 李晗雨. txt\_pos\_m11.txt. 英文动词(SVO)隐喻句子词性分析数据.
- [8] 李晗雨. txt\_pos\_nm11.txt. 英文动词(SVO)非隐喻句子词性分析数据.
- [9] 李晗雨. 文学比喻词典.xls. 文学比喻词典语料库.
- [10] 李晗雨. duzhe.txt. 《读者》20 年电子版语料库.

收稿日期: 2018-01-29  
收修改稿日期: 2018-06-04

## Recognizing Metaphor with Convolution Neural Network and SVM

Huang Xiaoxi Li Hanyu Wang Rongbo Wang Xiaohua Chen Zhiquan  
(Institute of Cognitive and Intelligent Computing, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract: [Objective]** This paper presents a new method to recognize metaphor, from the Chinese and English datasets. **[Methods]** First, we mapped the experimental dataset to vector space, which was also input to a convolutional neural network along with the property and keyword features. Then, we extracted the needed features with the help of convolutional and pooled layers, as well as classified them using SVM. Finally, we combined the Max-Pooling and Mean-Pooling to improve the extracted features' accuracy. **[Results]** Compared with the traditional models, our method increased the accuracy of extracted features from the corpus of English verb-object, English adjective-noun and Chinese metaphor by 4.12%, 0.84% and 4.50% respectively. **[Limitations]** The Chinese word segmentation affects the training of word vector model. We need to add more layers to the convolutional neural networks. **[Conclusions]** The proposed method could effectively identify metaphor from Chinese and English corpus.

**Keywords:** Metaphor Recognition Convolution Neural Network Support Vector Machines Feature Extraction