

文章编号: 2095-2163(2021)05-0082-06

中图分类号: TP391

文献标志码: A

基于深度学习的韩国语文本情感分类

赵天锐

(战略支援部队信息工程大学 洛阳校区, 河南 洛阳 471000)

摘要: 文本情感分析(又称意见挖掘),是对带有情感倾向的文本进行分析、处理、归纳和推理的过程。本文提出将卷积神经网络(CNN)和双向长短时记忆网络(Bi-LSTM)相结合,作为提取文本特征的方式,而后添加自注意力(Self-Attention)机制形成情感分析模型。通过在建的NAVER电影评论数据库中进行比较实验,证明本文模型的分​​类准确率可达90.32%,较SVM、CNN、LSTM、Bi-LSTM等模型有较大的性能提升。该方法不仅可以较好地完成韩语短文本情感分析,对其它非通用语种和领域的情感分析任务也具备参考和借鉴意义。

关键词: 韩语; 情感分析; FastText; CNN; Bi-LSTM; Self-Attention

Sentiment analysis of Korean text based on deep learning

ZHAO Tianrui

(Luoyang Campus, Information Engineering University of PLA Strategic Support Forces, Luoyang Henan, 471001, China)

[Abstract] Text sentiment analysis, also known as opinion mining, is the process of analyzing, processing, summarizing and reasoning texts with emotional tendencies. The current process of using deep learning for sentiment analysis can be roughly divided into training vector representation in corpus and using vector representation for text classification. This paper proposes to combine CNN and Bi-LSTM as a way to extract text features, and then add a self-attention mechanism to form an sentiment analysis model. Through comparative experiments in the self-built NAVER movie review database, it is proved that the classification accuracy of this model can reach 90.32%, which has a greater performance improvement than SVM, CNN, LSTM, Bi-LSTM. This method can not only complete the sentiment analysis of Korean short texts, but also has reference significance for sentiment analysis tasks in other non-universal languages.

[Key words] Korean; sentiment analysis; FastText; CNN; Bi-LSTM; Self-Attention

0 引言

情感分析本质上属于文本分类任务,是以文本中包含的情感倾向为标准,对文本进行分类。随着计算机和互联网的不断发展,对海量文本信息进行处理、分析后得出信息中的情感分布成为可能。对社交媒体、新闻评论等进行情感分析后得到的结论,也可以进一步应用在舆情分析、观点挖掘等领域当中。

传统情感分析方法分为基于情感词典的方法和基于机器学习的方法,两种方法各有所长。而随着深度学习技术的持续进步,提出了将深度学习和词向量相结合,引入情感分析领域的研究方法。目前在自然语言处理领域,词向量是具有代表性的语义分布式表示方式,其中具有代表性的是 Word2Vec^[1]和 FastText^[2]。但是,目前大部分相关研究都是基于英文、中文等通用语。而对韩语而言,无论是韩国国内学者还是国际学者,对韩语情感分析的研究成果较少。因此,本文利用网络爬虫,爬取韩语影评语料形成语料库,使用 FastText 方法训练词向量,而后

使用多个深度学习模型进行对比实验,寻找适合韩语的情感分析模型。进而提出融合卷积神经网络和双向长短时记忆网络,并加入 Self-Attention 机制的韩语情感分析模型,并在自建的韩语数据集上证明了该模型的有效性。

1 相关研究

情感分析最早是由 Pang 等^[3]于 2002 年提出的概念。近来,深度学习因在自然语言处理任务中逐渐显露出优越的性能^[4],受到越来越多的关注。由于基于深度学习的方法可以使模型从数据中自行学习语言特征,该方法与传统机器学习算法相比通常能取得更好的效果。由于文本的数据结构较为稀疏且长度较大,为了能实现对文本的处理,需要较好的文本特征提取方式,而其中影响力最大的就是词嵌入(Word Embedding)。Milokov 等提出的 Word2Vec 是词嵌入的代表性方法,该方法可以利用上下文信息得出各单词之间的相关性,最终以高维向量的形式将这种相关性展现出来。因为词嵌入的出现,之

作者简介: 赵天锐(1997-),男,硕士研究生,主要研究方向:情感分析、计算语言学。

收稿日期: 2021-01-05

前被认为缺乏实用性的神经网络算法再次受到关注,并且这种算法也被应用于情感分析任务,收到了较好的效果。文献[5]中提出了可以捕捉文本序列信息的循环神经网络(RNN)。而针对 RNN 学习长期依赖困难的问题,学界相继提出了长短期记忆网络(LSTM)和双向长短期记忆网络(Bi-LSTM)。这些模型通过在隐藏层增加 3 种“门”结构,控制神经元的状态,也提高了情感分析的准确度。此外,在图像处理方面展现出优越性能的卷积神经网络(CNN)也被用来解决情感分类任务^[6]。在 Attention 注意力机制提出后,Bahdanau 等首先将其应用到机器翻译领域。实验结果证明效果高于传统的神经网络模型。于是,很多专家学者将其应用在文本分类任务中,也收到了较好的验证效果。尤其是引入 Attention 机制的 BiLSTM 模型,可以达到很高的准确率。

上述研究方法大多已应用于通用语的情感分析,韩国国内和国际学者以韩语文本为研究对象也进行了一定的探索。文献[7]中对韩语文本进行分词,训练词向量,而后运用 CNN 进行文本分类;文献[8]则在进行词嵌入后使用 RNN 进行分类。此外,结合韩语本身的语言特征,文献[9]完成了以音节为单位,对文本进行预处理的研究;文献[10]则以音素为单位,研究文本的预处理方案;文献[11]中是先通过比较试验,选出了单词、音节、音素中最适合进行韩语文本预处理的单位,而后提出了双向堆叠的 Bi-LSTM 模型,在电影评论数据集中达到约 88.95% 的准确率。而 Lee^[12]以购物评论为对象进行分析,结合不同用户群体评论中的关键字特征,提高了神经网络模型的效率。

2 融入自注意力机制的 Bi-LSTM + CNN 模型结构

2.1 文本预处理和向量化表示

2.1.1 文本预处理

实验所使用的语料通过爬虫程序从‘Naver 영화’上抓取。抓取后需对原始语料进行清洗,语料的清洗和降噪会对模型分类的准确率产生影响。韩语文本虽然存在分写法,但和英文不同,文本中的空格符不能直接当做分词的依据。因此,需要对文本进行分词操作,本文采用的分词工具是基于 Python 语言编写的韩语分词标注工具:Konlpy。

2.1.2 文本的向量化表示

文本向量化是指将文本的分词结果用词向量进

行表示。本文采用的向量化表示方法是 Facebook 提出的一种高效的浅层网络——FastText。FastText 和 Word2Vec 中的 CBOW 原理相似,整个模型有输入层、隐藏层和输出层。两个模型的输入都是用向量表示的单词,隐藏层都是多个词向量的叠加平均,输出都是特定的 Target。FastText 和 CBOW 区别在于:

(1)前者输入的是多个单词及其 n-gram 特征,而后者输入的是目标单词的上下文;

(2)前者输入的特征经过嵌入层,后者输入的特征经过独热编码;

(3)前者输出的是对应文档类型的标签,而后者输出的是目标词汇。FastText 利用 h-softmax 的分类功能,通过对分类树所有叶子节点的遍历,寻找概率值最大的标签。本文采用 Python 语言编写 Gensim 库中的 FastText 模型,对构建的语料库中文本进行 10 000 次迭代后,得到由该模型训练的词向量。

借助 FastText 输出的词向量,可以更清楚地看到单词间语义的亲疏关系。衡量词义远近关系的标准是余弦相似度,余弦相似度越大则单词间的联系越紧密。在影评语料中与“나쁘다”关联度最高的 10 个词见表 1。

表 1 词向量语义相似度
Tab. 1 Word Vector Semantic Similarity

单词	余弦相似度
다들다	0.475 757 032 632 827 8
다짜기다	0.405 774 623 155 593 9
일하다	0.396 124 243 736 267 1
지급하다	0.375 405 132 770 538 3
입다	0.368 692 517 280 578 6
다러워기다	0.359 301 805 496 215 8
찍다	0.355 432 569 980 621 3
노답다	0.354 354 232 549 667 4
깨분하다	0.352 682 709 693 908 7
프랑하다	0.343 380 123 376 846 3

2.2 卷积神经网络

卷积神经网络最初应用于计算机视觉领域中的图像分类、目标检测、图像语义分割等任务,是一种包含卷积计算的前馈神经网络。后经实验发现,其用于文本分类中也可以得到较好的效果。经典的卷积神经网络包括输入层、卷积层、池化层、全连接层和输出层。其中,卷积层对输入的特征进行特征提取;池化层对提取过后的特征进行下采样以减少运算量,提高运行效率,得到局部最优值;池化层后是全连接层和输出层;为了防止出现过拟合,还要加入 Dropout 操作,最终通过分类器进行分类。卷积神经

网络结构如图 1 所示。

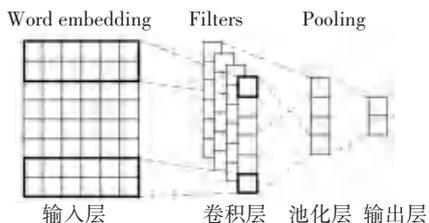


图 1 CNN 结构图

Fig. 1 CNN Structure

本文将 CNN 的卷积核大小设置为 (3, 4, 5), 每个大小相同的卷积核的通道数设置为 2。这样 CNN 可以分别提取词汇在原句中的 3-gram、4-gram 和 5-gram 特征。

2.3 双向长短时记忆网络

1997 年 Hochreiter 等针对循环神经网络 (RNN) 的梯度消失和梯度爆炸等问题, 首次提出了长短时记忆网络 (LSTM)。与原来的 RNN 相比, LSTM 加入了“门”结构来控制信息的传递, 较好地解决了文本的长距离依赖问题。LSTM 主要包含输入门 (i_t)、遗忘门 (f_t)、输出门 (o_t) 与一个记忆单元 (C_t)。具体结构如图 2 所示。

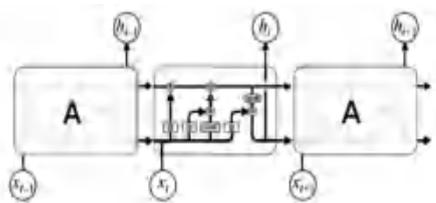


图 2 LSTM 内部结构图

Fig. 2 LSTM Structure

最初提出的 LSTM 模型在处理自然语言时, 只考虑了上文的语义信息而忽略了下文信息, 为了弥补这个缺陷, 可以使用双向长短时记忆网络 (Bi-LSTM) 模型。Bi-LSTM 通过构建对齐的双层 LSTM 模型, 同时进行自前向后和自后向前传播模型。Bi-LSTM 模型包含了输入层、前向传递层、后向传递层和输出层。本文使用的 Bi-LSTM 结构如图 3 所示:

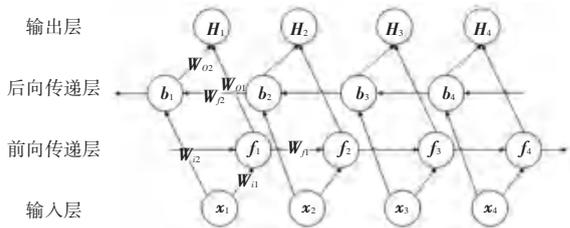


图 3 Bi-LSTM 模型结构

Fig. 3 Bi-LSTM Structure

的特征, 将预处理过的每条短文本表示为向量形式后, 作为 Bi-LSTM 网络的输入。假设: 输入内容 $X = \{x_1, x_2, \dots, x_t\}$, X 是每条短文本中所有单词向量组成的句向量, 每个 $x_i (i = 1, 2, \dots, t)$ 是单个词向量。前向传递层和后向传递层分别用于学习上下文的文本信息, 同时连接到输出层。 W_{i1}, W_{i2} 分别为输入门到前向单元和后向单元的权值矩阵, W_{f1}, W_{f2} 分别为前向和后向传递层隐含单元遗忘门的权值矩阵。 W_{o1}, W_{o2} 分别为前向和后向传递层到输出门的权值矩阵。本文中的 Bi-LSTM 接受词向量作为输入, 最终输出值由前向传递层和后向传递层共同学习到的语义信息组成。如将“이 영화는 좋다”这句话输入到模型当中, 分词后得到{“이”, “영화”, “좋다”}, 而后 3 个单词的词向量会输入网络, 前向传递层会得到 3 个向量 $\{h_{f0}, h_{f1}, h_{f2}\}$, 后向传递层也会得到 3 个向量 $\{h_{b0}, h_{b1}, h_{b2}\}$ 。将前后传递层的最后一项输出进行拼接, 可以得到 $[h_{f2}, h_{b2}]$, 将该向量作为输入接入全连接层, 可以进行文本情感分类。

2.4 自注意力机制

注意力机制的出现最早是为了解决计算机视觉领域的问题, Google Mind 在循环神经网络的基础上, 添加了注意力机制处理图像分类任务。随后 Bahdanau 等^[13]将 Attention 机制引入自然语言处理领域, 实现了在机器翻译任务中翻译和对齐同时进行。2017 年, Google 机器翻译团队大量使用了自注意力 (Self-Attention) 机制, 并在不同的自然语言处理任务上进行实验, 取得了良好的效果。自注意力机制的基本结构如图 4 所示。

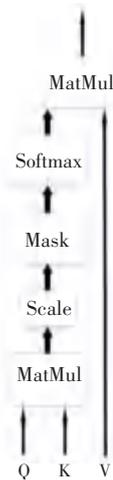


图 4 Self-Attention 基本结构

Fig. 4 Self-Attention Structure

Google 团队在发布的论文中提出放缩点积

为充分考虑文本的上下文信息, 准确提取文本

Attention 机制(scaled dot-Product attention),其实质是在原本使用点积进行相似度计算的 Attention 基础上,除以一个标准尺度 $\sqrt{d_k}$,起到调节作用。这可以防止 Q 和 K 的点积结果过大,然后通过 softmax 操作将注意力得分归一化为概率分布。注意力机制的得分计算如式(1):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (1)$$

2.5 融入自注意力机制的 Bi-LSTM + CNN 模型

本文采用融入自注意力机制的 Bi-LSTM + CNN 模型,作为韩国语情感分析模型。模型整体架构如图 5 所示。

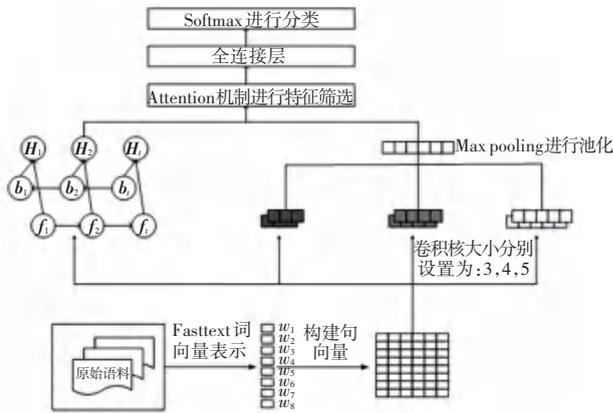


图5 融入自注意力机制的 Bi-LSTM + CNN 模型结构

Fig. 5 ATT-Bi-LSTM+CNN Structure

本文建构的神经网络模型,利用 Bi-LSTM 和 CNN 分别对输入的词向量进行特征提取并进行组合,而后通过词语关联度的注意力机制对组合后的特征进行进一步的优化,将优化后的特征作为全连接层的输入,以完成情感分类任务。该模型的优势在于将 Bi-LSTM 和 CNN 的优势进行有机结合:前者关注与文本语境相关的特征,后者关注局部特征,从而对文本信息的捕捉更加全面。这种网络对传统单一神经网络存在的问题,如:忽略词汇上下文语义信息,梯度消失和梯度离散等能起到一定的缓解作用。注意力机制的引入让神经网络在进行分类时,能给予情感关键词更高的权重,有助于情感分类。

3 实验

本文在预处理过的文本上使用 FastText 模型训练词向量,而后将所得词向量输入 Bi-LSTM + CNN 模型,进行迭代训练后输出预测的情感标签。实验

环境见表 2。

表 2 实验环境

Tab. 2 Lab Environment

硬件环境	CPU: Intel Core i7 GPU: NVIDIA Telsa T4
软件环境	Windows10 64bit Python 3.7.7 Pytorch 1.2.0

3.1 实验数据

实验所用数据是通过爬虫获取的韩国影评网站“Naver 영화”上的相关评论。网站采用满分为 10 分的评分机制,用户可以对电影进行评论并打出分数。在抓取评论后,结合评论内容和分数对评论进行情感标注。最终选取积极和消极的评论各 10 000 条组成实验数据集。部分数据样例见表 3。

表 3 实验数据样例

Tab. 3 Sample Experimental Data

评论内容	分数
영화보는대내 감동전수림에있네요~최고!	10
태절은 배우의 커리어 연기가 왜 어우어 귀구려	1
헐리웃 해개로 가는 한국영화 산업의 좋은 본보기	1
한국영화의 길을 걸구는 영화	0
여우 집노릇 이게 무슨 필적어니?	0
특별히게 비틀고 도인드만 주나고 좋은 영화가 아니다.	0

3.2 评价标准

本文采用准确率 (Precision)、召回率 (Recall) 和 $F1 - score$ 3 种通用的国际标准,对各种实验模型进行评估。其中,准确率指被正确分类的样本占总体样本的比例;召回率是被正确分类的样本占此类样本的比例; $F1 - score$ 是当 $a = 1$ 时的准确率和召回率的调和均值。3 种评价标准的计算公式见表 4。公式中各项数据的含义见表 5。

表 4 实验评价标准计算公式

Tab. 4 Experimental Evaluation Standard Calculation

评价指标	计算公式
准确率 (Precision)	$\frac{TP}{TP + FP}$
召回率 (Recall)	$\frac{TP}{TP + FN}$
$F1 - score$	$\frac{2 * Precision * Recall}{Precision + Recall}$

表 5 混淆矩阵

Tab. 5 Confusion Matrix

真实类别	预测类别	
	积极	消极
积极	TP	FN
消极	FP	TN

3.3 模型参数

实验参数的设置要经过多轮实验调整才能得到良好的效果。在本文进行的实验中,首先在数据集上使用 FastText 训练词向量,然后利用训练的词向量进行情感分类。经过多次对比实验,结合数据集的实际状况,最终设置的 FastText 模型参数见表 6;在得到词向量后要利用其进行情感分类,本文提出了融入自注意力机制的 Bi-LSTM + CNN 模型,最终设置的模型参数分别见表 7 和表 8。

表 6 FastText 模型参数设置

Tab. 6 FastText Model Parameters

参数	参数含义	参数值
size	词向量维度	100
window	滑动窗口大小	5
iter	迭代次数	10 000
min_n	最小字符长度	2
max_n	最大字符长度	5

表 7 CNN 参数设置

Tab. 7 CNN Model Parameters

参数含义	参数值
卷积核尺寸	3,4,5
滑动窗口大小	128
激活函数	Relu
池化	Max Pooling
Dropout	0.5
迭代周期	1000

表 8 Bi-LSTM 参数设置

Tab. 8 Bi-LSTM Model Parameters

参数	参数值
层数	2
学习率	1e-3
批处理个数	128
Dropout	0.5
损失函数	CrossEntropy
全连接层激活函数	Relu
迭代周期	1 000

3.4 实验结果与分析

为证明模型的有效性,在自建数据集上将经典的情感分类算法和本文提出的融入自注意力机制的 Bi-LSTM + CNN 模型进行对比实验。为控制变量,实验采用的词向量皆为在数据集中通过 FastText 模型训练得到的词向量。

(1) CNN 模型: CNN 模型注重局部特征,在自建韩国语文本数据集上取得的效果较差;

(2) LSTM 模型: LSTM 与传统的循环神经网络相比,加入了“门”机制,可以部分解决梯度爆炸和梯度离散问题;

(3) Bi-LSTM 模型: Bi-LSTM 模型可以提取文本上下文相关的全局特征,相较于单向的 LSTM 模型分类效果有所提升;

(4) ATT-CNN 模型: 应用注意力机制的卷积神经网络可以缩短训练时间,同时通过注意力机制弥补了捕捉信息的不足;

(5) ATT-Bi-LSTM 模型: 在 Bi-LSTM 模型中加入了 Self-Attention 机制。先利用 Bi-LSTM 网络提取文本序列的信息,而后借助 Self-Attention 机制对关键词给予更多权重;

(6) ATT-Bi-LSTM+CNN 模型: 本文提出的模型,先使用 Bi-LSTM 和 CNN 网络分别提取文本的上下文和局部特征,而后经过 Self-Attention 优化进行情感分析的算法。

本文算法与其它算法在自建数据集上的表现见表 9。

表 9 不同模型评价标准对比

Tab. 9 Comparison of Different Models

模型	准确率/%	召回率/%	F1 - score
CNN	83.14	82.52	0.828 2
LSTM	83.72	83.51	0.836 1
Bi-LSTM	86.27	84.78	0.850 2
ATT-CNN	85.21	84.16	0.846 8
ATT-Bi-LSTM	87.36	85.55	0.859 0
ATT-Bi-LSTM+CNN	0.32	87.63	0.889 5

通过表(9)的数据,可以得到如下结论:

(1) 通过对比 LSTM 和 Bi-LSTM 的数据,双向模型更好地考虑上下文的信息,在自建数据集上的表现较单向模型更优;

(2) 分别对比 CNN 与 ATT-CNN、Bi-LSTM 和 ATT-Bi-LSTM 的数据可知,在原有的模型中增加注意力机制,可以有效地实现对文本中重点信息的关注,从而提升模型的表现。CNN 加入注意力机制后准确率、召回率和 F1 - score 的值分别提升了 2.08%、1.64% 和 0.0186; Bi-LSTM 则分别提升了 1.09%、0.77% 和 0.008 8。

(3) 通过对比 ATT-CNN、ATT-Bi-LSTM 和 ATT-Bi-LSTM+CNN 的效果可知,相比于单一的 CNN 和 Bi-LSTM,两者提取特征的组合更能在上下

文语境信息和局部特征两方面取得平衡。本文提出的模型在数据集上取得了较好的效果, 准确率、召回率和 $F1 - score$ 均高于其它模型, 证明了本文提出算法的有效性。

4 结束语

本文提出了融入自注意力机制的 Bi-LSTM + CNN 韩语短文本情感分析模型, 使用 FastText 模型训练词向量。综合利用 CNN 和 Bi-LSTM 关注局部特征和提取上下文语境信息的优势。为证明模型有效性, 本文设计了对比实验, 结果显示本文提出的模型在自建数据集中取得了优于其他算法的效果, 证明了模型的有效性。

然而本文在处理文本时, 只考虑了上下文信息, 并未充分利用韩语的词性、音素等语言学特征, 因此未来可以考虑在分类时加入此类特征, 观察其对分类结果的影响。

参考文献

- [1] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient Estimation of Word Representations in Vector Space [C]//Proceedings of the International Conference on Learning Representations.2013: 1-12.
- [2] JOULIN A, GRAVE E, BOJANOWSKI P, T Mikolov, et al. Bag of Tricks for Efficient Text Classification [C]//Proceedings of the European Chapter of the Association for Computational Linguistics. 2017: 427-431.
- [3] PANG B, LEE L, VAITHYANATHAN S. Thumbs up: sentiment classification using machine learning techniques [C]//Proceedings of the ACL - 02 Conference on Empirical Methods in Natural Language Processing, 2002: 79-86.
- [4] KIM Y. Convolutional Neural Networks for Sentence Classification [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014: 1746-1751.
- [5] SOCHER R, LIN C Y, NG A Y, et al. Parsing Natural Scenes and Natural Language with Recursive Neural Networks [C]// Proceedings of the 28th International Conference on Machine Learning, ICML 2011, Bellevue, Washington, USA, June 28 - July 2, 2011: 129-136.
- [6] KIM Y. Convolutional Neural Networks for Sentence Classification [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL. 2014: 1746-1751.
- [7] Dowoo, Kim, Myoung - Wan, et al. Categorization of Korean News Articles Based on Convolutional Neural Network Using Doc2Vec and Word2Vec [J]. Journal of KIISE, 2017, 44(7): 742-747.
- [8] KIM J M, LEE J H. Text Document Classification Based on Recurrent Neural Network Using Word2vec [J]. Journal of Korean institute of intelligent systems, 2017, 27(6): 560-565.
- [9] KIM S D, PARK S B, PARK S Y, et al. A Syllable Kernel based Sentiment Classification for Movie Reviews [J]. Journal of Korean Institute of Intelligent Systems, 2010, 20(2): 202-207.
- [10] 정호주, 김도현, 최용석. 한글 음소 단위 Trigram-bigram 기반의 오타피드백 이터레이팅 기법 [J], 한국정보과학회 학술발표논문집, 2015, 6.
- [11] OH Y, KIM M, KIM W. Korean Movie-review Sentiment Analysis Using Parallel Stacked Bidirectional LSTM Model [J]. Journal of KIISE, 2019, 46(1): 45-49.
- [12] 이광호. 딥러닝을 활용한 소극물-고려문벌 리액션 영상 분석 [D]. 국민대학교, 2019.
- [13] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate [J]. Computer ence, 2014.
- [14] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn [C]// Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
- [15] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2818-2826.
- [16] BAGAUDTINOV T, ALAHI A, FLEURET F, et al. Social scene understanding: End-to-end multi-person action localization and collective activity recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4315-4324.
- [17] WU J, WANG L, WANG L, et al. Learning actor relation graphs for group activity recognition [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 9964-9974.
- [18] you need [J]. arXiv preprint arXiv:1706.03762, 2017.

(上接第 81 页)

- [6] LI X, CHOOCHUAH M. Sbgar: Semantics based group activity recognition [C]//Proceedings of the IEEE international conference on computer vision. 2017: 2876-2885.
- [7] TSUNODA T, KOMORI Y, MATSUGU M, et al. Football action recognition using hierarchical lstm [C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 99-107.
- [8] RAMANATHAN V, HUANG J, ABU-EL-HAJJA S, et al. Detecting events and key actors in multi-person videos [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 3043-3053.
- [9] QI M, WANG Y, QIN J, et al. StagNet: an attentive semantic RNN for group activity and individual action recognition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(2): 549-565.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all