Journal of Computer Applications
计算机应用,2021,41(3):699 - 705
ISSN 1001-9081
CODEN JYIIDU
2021-03-10
http://www.joca.cn

文章编号:1001-9081(2021)03-0699-07
DOI:10.11772/j.issn.1001-9081.2020060837

基于优势演员–评论家算法的强化自动摘要模型

杜嘻嘻,程　华*,房一泉
(华东理工大学 信息科学与工程学院,上海 200237)
(*通信作者电子邮箱hcheng@ecust.edu.cn)

摘　要: 针对长文本自动摘要任务中抽取式模型摘要较为冗余,而生成式摘要模型时常有关键信息丢失、摘要不准确和生成内容重复等问题,提出一种面向长文本的基于优势演员–评论家算法的强化自动摘要模型(A2C-RLAS)。首先,用基于卷积神经网络(CNN)和循环神经网络(RNN)的混合神经网络的抽取器(extractor)来提取原文关键句;然后,用基于拷贝机制和注意力机制的重写器(rewriter)来精炼关键句;最后,使用强化学习的优势演员–评论家(A2C)算法训练整个网络,把重写摘要和参考摘要的语义相似性(BERTScore值)作为奖励(reward)来指导抽取过程,从而提高抽取器提取句子的质量。在CNN/Daily Mail数据集上的实验结果表明,与基于强化学习的抽取式摘要(Refresh)模型、基于循环神经网络的抽取式摘要序列模型(SummaRuNNer)和分布语义奖励(DSR)模型等模型相比,A2C-RLAS的最终摘要内容更加准确、语言更加流畅,冗余的内容有效减少,且A2C-RLAS的ROUGE和BERTScore指标均有提升。相较于Refresh模型和SummaRuNNer模型,A2C-RLAS模型的ROUGE-L值分别提高了6.3%和10.2%;相较于DSR模型,A2C-RLAS模型的F1值提高了30.5%。

关键词: 自动摘要模型;抽取式摘要模型;生成式摘要模型;编码器–解码器;强化学习;优势演员–评论家算法
中图分类号: TP391.1　　**文献标志码:** A

Reinforced automatic summarization model based on advantage actor-critic algorithm

DU Xixi, CHENG Hua*, FANG Yiquan
(*Institute of Information Science and Engineering, East China University of Science and Technology, Shanghai* 200237, *China*)

www.joca.cn

Abstract: The extractive summary model is relatively redundant and the abstractive summary model often loses key information and has inaccurate summary and repeated generated content in long text automatic summarization task. In order to solve these problems, a Reinforced Automatic Summarization model based on Advantage Actor-Critic algorithm (A2C-RLAS) for long text was proposed. Firstly, the key sentences of the original text were extracted by the extractor based on the hybrid neural network of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Then, the key sentences were refined by the rewriter based on the copy mechanism and the attention mechanism. Finally, the Advantage Actor-Critic (A2C) algorithm in reinforcement learning was used to train the entire network, and the semantic similarity between the rewritten summary and the reference summary (BERTScore (Evaluating Text Generation with Bidirectional Encoder Representations from Transformers) value) was used as a reward to guide the extraction process, so as to improve the quality of sentences extracted by the extractor. The experimental results on CNN/Daily Mail dataset show that, compared with models such as Reinforcement Learning-based Extractive Summarization (Refresh) model, a Recurrent Neural Network based sequence model for extractive summarization (SummaRuNNer) and Distributional Semantics Reward (DSR) model, the A2C-RLAS has the final summary with content more accurate, language more fluent and redundant content effectively reduced, at the same time, A2C-RLAS has both the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) and BERTScore indicators improved. Compared to the Refresh model and the SummaRuNNer model, the ROUGE-L value of the A2C-RLAS model is increased by 6.3% and 10.2% respectively; compared with the DSR model, the F1 value of the A2C-RLAS model is increased by 30.5%.

Key words: automatic summary model; extractive summary model; abstractive summary model; encoder-decoder; reinforcement learning; Advantage Actor-Critic (A2C) algorithm

0　引言

互联网上的信息量正以指数级的速度增长,如何从海量的文本信息中准确、快速地获取有用信息,是当今人们最迫切的需求[1]。文本自动摘要技术对给定源文本信息进行概括与总结,并产生简洁、流畅且保留关键信息的短文本,这不仅可以缓解搜索引擎检索导致的冗余问题,而且可以有效解决信息过载的缺陷[2]。

文本自动摘要可分为抽取式方法(Extractive)和生成式方

收稿日期: 2020-06-17;**修回日期:** 2020-10-08;**录用日期:** 2020-10-27。　**基金项目:** 赛尔网络下一代互联网技术创新项目(NGII20170520)。
作者简介: 杜嘻嘻(1994—),女,安徽亳州人,硕士研究生,主要研究方向:自然语言处理;　程华(1975—),男,安徽黄山人,副研究员,博士,主要研究方向:智能信号处理、信息安全;　房一泉(1975—),女,上海人,高级实验师,硕士,主要研究方向:信息安全。

法(Abstractive)。抽取式方法抽取源文本关键句组成摘要,如TextRank算法[3]将句子作为节点,用相邻边的权重大小表示节点的语义相似度,选取得分最高的k个节点作为摘要。生成式方法是近年来流行的方法,在理解原文的基础上,逐词生成摘要,可得到语义更精确的摘要,通常可用机器学习或深度学习实现,如基于注意力机制的序列到序列(sequence to sequence, seq2seq)模型[4]、利用卷积神经网络(Convolutional Neural Network, CNN)对源文本编解码,并引入注意力机制,在DUC数据集上取得了不错的效果。

目前,研究者们深入研究了自动文本摘要领域,提出了许多新颖的模型。Sharma等[5]提出实体驱动的生成式摘要模型(System for ENtity-drivEn Coherent Abstractive summarization model, SENECA),利用实体信息产生包含有用信息的连贯摘要。Zhang等[6]提出通过抽取空缺句的预训练生成式摘要模型(Pre-training with Extracted GAp-Sentences for abstractive Summarization model, PEGASUS),利用新的预训练目标(Gap Sentences Generation, GSG)用于摘要总结、空缺句子生成,在所有12个下游数据集上都达到了最先进的性能。Saito等[7]提出具有重要标记的条件摘要模型(Conditional summarization model with Important Tokens, CIT),该模型将显著性模型提取的标记序列显式地提供给seq2seq模型作为附加输入文本。

然而,针对长文本的自动摘要,若采用生成式方法,解码时间较长,摘要不准确,其原因是过长的源文本导致编码器不能准确提取输入文本的语义信息,且生成过程缺少关键信息的指导。如文献[8]的生成式摘要方法采用强化学习框架,输入整篇文档,生成的摘要不反映文章核心内容,摘要无法达到与源文本内容等价的效果。若采用抽取式方法,抽取原文中整句作为摘要,其包含较多非摘要内容,得到的摘要信息冗长、不够精简且摘要句间不连贯。

针对上述问题,长文本自动摘要可先抽取源文本关键句,再用生成式方法重写摘要,以解决摘要的冗余与连贯问题。生成式方法依靠抽取的结果,因此抽取结果的好坏直接影响最后摘要的质量。如果抽取得不好,甚至抽取错误将导致摘要不准确、偏离文章中心。对于该问题,本文提出了面向长文本的基于优势演员-评论家算法[9]的强化自动摘要模型(Reinforced Automatic Summarization model based on Advantage Actor-Critic algorithm, A2C-RLAS),在强化学习框架下用A2C算法的多次迭代来改进抽取过程,其中用BERTScore (evaluating text generation with Bidirectional Encoder Representations from Transformers)[10]得分较好地评价重写摘要和参考摘要的语义相似性,反馈抽取质量的改善效果。在CNN/Daily Mail[11]数据集上的实验结果表明,与基于强化学习的抽取式摘要(Reinforcement learning-based extractive summarization, Refresh)模型、基于循环神经网络的抽取式摘要序列模型(a Recurrent Neural Network based sequence model for extractive Summarization, SummaRuNNer)和分布语义奖励(Distributional Semantics Reward, DSR)模型等相比,A2C-RLAS模型在ROUGE(Recall-Oriented Understudy for Gisting Evaluation)[12]和BERTScore指标上均有提升,其中相较于Refresh模型和SummaRuNNer模型,A2C-RLAS模型的ROUGE-L值分别提高了6.3%和10.2%;相较于DSR模型,A2C-RLAS模型的F1值提高了30.5%。本文模型摘要语义也更加丰富、语言也更加简洁流畅。

1 问题描述

在长文本自动摘要中,已有的抽取式方法抽取原文句子组成摘要,摘要的长度往往是参考摘要长度的几倍,例如基于encoder-decoder的抽取式模型Refresh[13]对数据集CNN/Daily Mail中测试集样本进行抽取,结果如表1,参考摘要往往比抽取摘要短,平均每句参考摘要13.5个词,平均每句抽取摘要27.8个词,平均抽取摘要句的长度是平均参考摘要句长度的2倍多,表明直接抽取整句作为摘要的方法将导致摘要信息冗长、不够精简。

表1 摘要的统计

Tab. 1 Summary statistics

摘要	测试样本数	平均源文本长度	平均源文本句子数	平均摘要句子数	平均摘要长度
参考摘要	11 490	706.92	31.65	3.72	50.17
抽取摘要	11 490	706.92	31.65	4.67	129.75

例1为该模型抽取的摘要样例,其中抽取摘要包含的关键信息用下划线标出,抽取摘要包含10个短句,参考摘要仅包含2个短句,抽取摘要的长度是参考摘要长度的3~4倍,而且抽取摘要中同一人名重复出现,句间不够连贯。因此摘要不应该仅仅来自于抽取,要得到精炼、简洁和流畅的摘要需要进一步重写。

例1

参考摘要:ben powers joined the cast of “good times” for its sixth and final season. he played thelma ’ s husband keith anderson.

抽取摘要:<u>ben powers joined the cast of “good times” for its sixth and final season</u> from 1978 to 1979 season , playing keith , a professional football player . his character and thelma are wed on the third episode that year , but he injures his leg while walking out of the church , straining their relationship . <u>ben powers</u> , who <u>played thelma ’ s husband keith anderson</u> on the final season of the classic cbs sitcom “good times” , has died . he was 64 .

生成式摘要允许摘要中包含新的词语或短语,灵活性高、更连贯,但是在面临长文本摘要时,输入整篇文档影响了生成过程,导致解码缓慢,通常会出现关键信息丢失、摘要不准确的问题。如文献[8]采用生成式方法,提出分布语义奖励(Distributional Semantics Reward, DSR)模型,实验结果表明,输入是整篇长文档时,输出摘要无法涵盖输入长文本中的全部关键信息。但生成式摘要对短文本生成效果好,可以生成准确流畅的摘要。例2为DSR模型生成的摘要样例,模型生成摘要与参考摘要几乎一致,且摘要准确流畅、可读性强。因此对于长文本,通过重写原文关键句,不仅可以保留原文中最重要的信息,而且可达到更加简洁流畅的摘要目标。

例2

源文本:european finance ministers on saturday urged swedes to vote “yes” to adopting the euro , saying entry into the currency bloc would be good for sweden and the rest of europe .

参考摘要:european finance ministers urge swedes to vote yes to euro

生成摘要:eu finance ministers urge swedes to vote yes to euro

在摘要生成过程中,关键句经重写后得到摘要,重点在于如何抽取出关键句,抽取结果影响生成的摘要质量,抽取出非关键句将导致摘要无法概括文章核心内容。如文献[14]通过抽取任务中估计的显著性来约束生成任务中学习的注意力,以增强它们的一致性,重新生成后减少了抽取式摘要包含的冗余信息,但没有多次迭代的过程,导致抽取出非关键句,以至于摘要没有准确把握原文核心思想。可以采用强化学习来解决,通过A2C算法多次迭代改进抽取过程,使抽取器获得更准确的句子用于重写。

2 基于A2C算法的强化自动摘要模型

在强化学习框架下,本文提出面向长文本的基于A2C算法的强化自动摘要模型(A2C-RLAS),如图1所示。基于抽取器(extractor)和重写器(rewriter)过程得到新的摘要,作为A2C算法的actor,critic基于actor抽取关键句的行为评判该行为的得分,评价器(evaluator)通过BERTScore值评价重写摘要与参考摘要的语义相似性,作为奖励,actor则根据critic的评分修改选行为的概率,最终改进抽取过程。

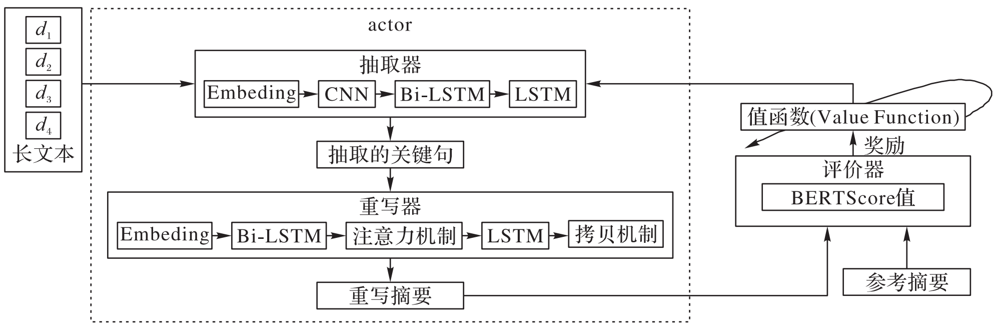

图1　A2C-RLAS模型

Fig. 1　A2C-RLAS model

2.1　抽取器

抽取器基于encoder-decoder框架,如图2所示。分层文档编码器读取输入序列并生成结合全局上下文语义信息的句向量;解码器结合指针机制获取原文中每个句子被抽取的概率,同时抽取出若干关键句,组成短文本。

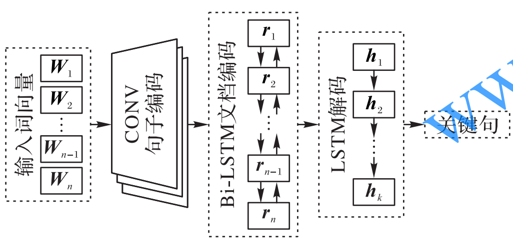

图2　抽取器模型

Fig. 2　Extractor model

2.1.1　分层文档编码器

由于抽取器要抽取源文本整句,因此模型输入需要句向量表示源文本中的每一个句子。采用文献[15]提出的时间卷积模型得到源文本中每个句子的句向量。具体地,对于源文本中每一个句子,由句中每个词的词向量得到句子矩阵,通过卷积、非线性激活和最大池化操作,卷积网络输出第i个句子的句向量$\boldsymbol{r}_i$。

为了进一步整合文档的全局上下文信息并捕捉句子间长期语义依赖,将句向量$\boldsymbol{r}_i$送入双向长短期记忆(Long Short-Term Memory, LSTM)网络[16]进行训练,生成包含上下文语义信息的句向量$\boldsymbol{h}_i$,进而使用$\boldsymbol{h}_i$来表示源文本中第i个句子。

2.1.2　解码器

得到句向量$\boldsymbol{h}_i$后,使用基于LSTM的指针网络[17]作为解码器,进行关键句的选择。对于每个抽取时间步t,源文本中每个句子被抽取的概率是:

$$u_i^t=\begin{cases}\boldsymbol{v}_p^{\mathrm{T}}\tanh(\boldsymbol{W}_{p1}\boldsymbol{h}_i+\boldsymbol{W}_{p2}\boldsymbol{e}_t), & i_t\neq i_k, \forall k<t\\ -\infty, & 其他\end{cases}\tag{1}$$

$$P(i_t|i_1,i_2,\cdots,i_{t-1})=\mathrm{softmax}(u^t)\tag{2}$$

其中$\boldsymbol{e}_t$是指针网络在抽取时间步t时的输出,具体如下:

$$a_i^t=\boldsymbol{v}_g^{\mathrm{T}}\tanh(\boldsymbol{W}_{g1}\boldsymbol{h}_i+\boldsymbol{W}_{g2}\boldsymbol{z}_t)\tag{3}$$

$$\alpha^t=\mathrm{softmax}(a^t)\tag{4}$$

$$\boldsymbol{e}_t=\sum_i\alpha_i^t\boldsymbol{W}_{g1}\boldsymbol{h}_i\tag{5}$$

其中$\boldsymbol{z}_t$是单层LSTM的输出,$\boldsymbol{W}$和$\boldsymbol{v}$均为可学习参数。抽取模型可看作一个二分类问题,即对源文本中每个句子分为摘要句和非摘要句。

2.2　重写器

为得到简短、流畅的摘要,抽取器提取出关键句构成短文本后,重写器对短文本进行生成式改写。重写器采用标准的基于注意力机制的encoder-decoder模型[18],输入短文本$x=\{x_1,x_2,\cdots,x_M\}$,输出摘要序列$y=\{y_1,y_2,\cdots,y_N\}$。为考虑当前词的上下文状态,编码器采用双向长短期记忆网络(Bi-directional Long Short-Term Memory network, Bi-LSTM)。引入注意力机制,使解码时可以选择性地读取编码后的语义向量的信息。

针对未登录词(Out-Of-Vocabulary, OOV)问题,采用拷贝机制[19]解决。在解码端设置指针开关,指针开关为拷贝模式P,则复制原文单词;指针开关为生成模式G,则从词表生成一个新词。其模型结构如图3所示。

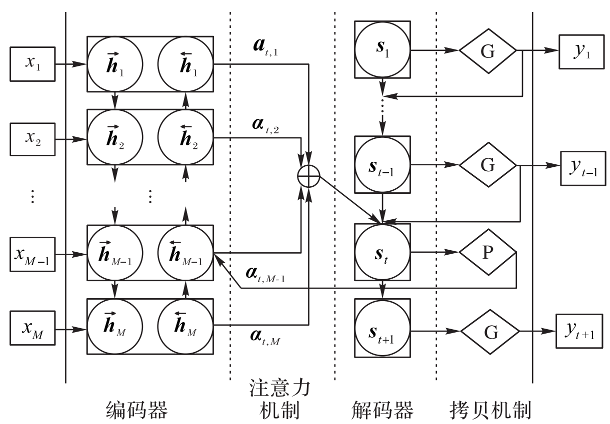

图3　重写器模型

Fig. 3　Rewriter model

www.joca.cn

2.3 强化学习

在长文本摘要过程中，需要对抽取过程进行多次迭代，采用强化学习中 Advantage Actor-Critic（A2C）算法。相较于 REINFORCE 算法[20]，A2C 算法增加评论家网络，可解决高方差问题，从而有效地提高抽取的关键句的质量。

由于传统的摘要评价指标 ROUGE 只能计算候选摘要与参考摘要之间的 n 元短语重叠度，无法捕捉摘要的语义关系。而本文方法要重写关键句，显然 ROUGE 不再适用于重写摘要的评价。因此，采用评价器（evaluator）计算重写摘要和参考摘要的 BERTScore 值，作为强化学习中基于语义的奖励，以提高抽取器抽取句子的质量。

2.3.1 基于语义的奖励

BERTScore[10] 利用来自 BERT（Bidirectional Encoder Representations from Transformers）的预训练上下文嵌入，并通过余弦相似性匹配参考摘要和候选摘要中的单词，被定义为：

$$R_{\mathrm{BERT}}=\frac{\sum_{m_i\in m}\mathrm{idf}(m_i)\max_{\hat{m}_j\in\hat{m}}\boldsymbol{M}_i^{\mathrm{T}}\hat{\boldsymbol{M}}_j}{\sum_{m_i\in m}\mathrm{idf}(m_i)} \tag{6}$$

$$P_{\mathrm{BERT}}=\frac{\sum_{m_j\in m}\mathrm{idf}(m_j)\max_{\hat{m}_i\in\hat{m}}\boldsymbol{M}_j^{\mathrm{T}}\hat{\boldsymbol{M}}_i}{\sum_{m_j\in m}\mathrm{idf}(m_j)} \tag{7}$$

$$F_{\mathrm{BERT}}=2\frac{R_{\mathrm{BERT}}\cdot P_{\mathrm{BERT}}}{R_{\mathrm{BERT}}+P_{\mathrm{BERT}}} \tag{8}$$

其中：$\boldsymbol{M}$ 和 $\hat{\boldsymbol{M}}$ 表示参考摘要中词 m 和候选摘要中词 $\hat{m}$ 来自 BERT 的上下文词嵌入；函数 idf(·)用于计算逆文档频率。

2.3.2 A2C 算法

采用强化学习 A2C（Advantage Actor-Critic）算法，actor 是指策略函数 $\pi_\theta(s,a)$。在每一个抽取时间步 t，抽取网络的当前状态是已抽取出的关键句 d_i 的集合，actor 通过观察当前状态 $c_t=(D,d_{i_{t-1}})$，根据策略 $\pi_\theta(s,a)$ 采样一个动作 $i_t\sim\pi_\theta(c_t,i)$，从原文 D 中抽取下一个句子 d_i，critic 网络评价当前重写后摘要句 $g(d_{i_t})$ 与参考摘要 b_t 对应句子的 BERTScore 值作为奖励：

$$r(t+1)=\mathrm{BERTScore}_{\mathrm{F}_1}(g(d_{i_t}),b_t) \tag{9}$$

而一次摘要过程为一个回合，对满足策略 $\pi_\theta(s,a)$ 的实际采样行为的评价通过反算、折扣后可得到整个摘要的总回报 R_t：

$$R_t=\sum_{t=1}^{N}\gamma^t r(t+1) \tag{10}$$

其中：N 表示抽取的关键句数目；γ 是折扣因子，为一个常数。

R_t 作为动作值函数 $Q^{\pi_\theta}(c,i)$ 的估计，A2C 算法选取仅基于状态的状态值函数 $V^{\pi_\theta}(c)$ 作为基准（baseline）函数，得到 Advantage 函数：

$$A^{\pi_\theta}(c,i)=Q^{\pi_\theta}(c,i)-V^{\pi_\theta}(c) \tag{11}$$

Advantage 函数记录了在状态 c_t 时采取行为 i_t 会比在状态 c_t 多出的价值，这与策略改善的目标是一致的，策略目标函数梯度可表示为：

$$\nabla_\theta J(\theta)=E_{\pi_\theta}[\nabla_\theta\log\pi_\theta(c,i)A^{\pi_\theta}(c,i)] \tag{12}$$

由 Advantage 函数得到的策略梯度有正有负，为正时鼓励抽取该句子，为负时不会鼓励抽取该句子。

此时，actor 的损失函数为：

$$L_\pi=-\log\pi_\theta(c,i)\cdot A^{\pi_\theta}(c,i) \tag{13}$$

而 critic 可以直接使用均方误差 MSE 作为损失函数。在每步更新中，策略函数的参数 θ 更新如下：

$$\theta\leftarrow\theta+\alpha\nabla_\theta\log\pi_\theta(c,i)A^{\pi_\theta}(c,i) \tag{14}$$

www.joca.cn

3 实验与结果分析

3.1 数据集

实验评估使用 CNN/Daily Mail 数据集[11]，由美国有线新闻网（Cable News Network，CNN）和每日邮报网新闻（Daily Mail）组成，如表2。该数据集包含287 227对训练集、13 368对验证集和11 490对测试集。本文用训练集来训练模型，保存多个训练效果较好的模型用验证集来挑选出最优模型，在测试集上测试得出最终结果。

表2 CNN/Daily Mail 数据集

Tab. 2 CNN/Daily Mail dataset

数据集	样本数	平均源文本长度	平均源文本句子数	平均摘要长度	平均摘要句子数
CNN/Daily Mail(train/val/test)	312 085(287 227/13 368/11 490)	766	29.74	53	3.72

3.2 评价指标

自动摘要评价采用通用的摘要评测准则 ROUGE[12]。ROUGE 指标计算参考摘要与候选摘要之间的 N 元短语重叠度来衡量摘要的质量。ROUGE 评价指标主要由 ROUGE-N（N 可取1、2、3等）和 ROUGE-L 组成。

本文采用 ROUGE-1（unigram）和 ROUGE-2（bigram）以衡量摘要信息的丰富性，ROUGE-L 代表参考摘要和候选摘要的最长公共子序列，用于衡量摘要内容的流畅性。

3.3 参数设置

对于所有的实验，基于 word2vec[21] 训练词向量，词向量的维数为128，LSTM 的隐藏单元向量维数为256，单词表使用30 000个单词。在训练阶段使用 Adam 优化器[22]训练模型，学习速率设置为0.001，动量参数设置为 $\beta_1=0.9$，$\beta_2=0.999$ 和 $\varepsilon=10^{-8}$，折扣因子设置为0.95，batch-size 设置为32，解码时使用 Beam 大小为5的集束搜索（Beam search）。

3.4 对比模型

为验证 A2C-RLAS 方法生成摘要的性能，选取目前表现良好的抽取式模型和生成式模型进行对比验证。各对比模型的详情如下：

1）Lead-3：直接抽取源文本的前三句组成文章摘要。

2）SUMO（Structured sUmmarization MOdel）[23]：将抽取式文本摘要看作树归纳问题，利用结构化注意（structured attention），基于迭代结构改进方法来学习文本表示。

3）Refresh[13]：将单文档摘要问题抽象为句子排序任务，提出一种通过强化学习对 ROUGE 评价标准进行全局优化的抽取式文本自动摘要模型。

4）Pointer-generator+coverage[24]：基于 encoder-decoder 框架，通过指针从原文复制单词以解决 OOV 问题，并引入 coverage 机制，避免生成重复。

5）SummaRuNNer[25]：基于门控循环单元-循环神经网络

(Gate Recurrent Unit- Recurrent Neural Network, GRU-RNN)的抽取式模型，提出了一个新颖的生成式训练机制，可以不需要在训练阶段引入抽取式标签。

6) BERT-ext+abs[26]：将 BERT 预训练方法应用于解决文本自动摘要问题，提出一种两阶段自动摘要模型。

3.5　对比实验结果

采用传统的评测准则 ROUGE 评价对比实验结果，如表 3，分为抽取式模型、生成式模型和本文模型三组，指标最高的分数用加粗数字标出。由表 3 可以看出，本文模型在 ROUGE-1 和 ROUGE-L 评分上是所有模型中最好的，表明本文模型的整体表现最好。相较于抽取式模型 SUMO，本文模型在 ROUGE-1 和 ROUGE-L 指标上分别提高了 0.33 和 1.69 个百分点；相较于生成式模型 Pointer-generator+coverage，本文模型在 ROUGE-1、ROUGE-2 和 ROUGE-L 指标上分别提高了 1.80、0.93 和 2.51 个百分点，表明本文模型抽取出的关键句内容更加全面。与 BERT-ext+abs 模型相比，本文模型在 ROUGE-1、ROUGE-2、ROUGE-L 指标上分别提升了 1.19、0.34 和 1.06 个百分点，表明基于强化学习的 A2C-RLAS 模型摘要的信息性和流畅性更好。但在 ROUGE-2 评分上 A2C-RLAS 没有达到最高水平，相较于 SUMO 模型差了 0.19 个百分点。R-AVG 评分是 ROUGE-1、ROUGE-2 和 ROUGE-L 的平均值，本文模型的 R-AVG 评分是 32.81，取得了相较于基线模型的最好效果。综合上述对比实验可知，本文模型性能取得了一定的提升，达到了最好效果。

表 3　不同模型的 ROUGE 评价指标对比　　单位：%

Tab. 3　Comparison of ROUGE evaluation indicator of different models　　unit：%

模型		ROUGE 分数 1	ROUGE 分数 2	ROUGE 分数 L	R-AVG
抽取式模型	Lead-3	40.34	17.70	36.57	31.54
	Refresh	40.00	18.20	36.60	31.60
	SUMO	41.00	**18.40**	37.20	32.20
生成式模型	SummaRuNNer	39.60	16.20	35.30	30.37
	Pointer-generator+coverage	39.53	17.28	36.38	31.06
	BERT-ext+abs	40.14	17.87	37.83	31.95
本文模型	A2C-RLAS	**41.33**	18.21	**38.89**	**32.81**

不采用传统的评测准则 ROUGE，增加以 BERTScore 为评测指标的对比实验，结果如表 4，指标最高的分数用加粗数字标出。文献[8]在以交叉熵损失(cross-ENTropy loss, XENT)训练模型和以 ROUGE 作为奖励训练模型的基础上，提出 DSR 模型，引入分布式语义奖励，生成语义正确的短语。文献[27]在指针生成(Pointer-Generator, PG)网络和基于覆盖机制的指针生成网络(Pointer-Generator network with the coverage mechanism, PG+Cov.)的基础上，引入新的注意力机制，提出基于递减注意力机制的指针生成模型(Pointer-Generator model with the diminishing attention mechanism, PG+DimAttn)，利用子模块函数的特性提高神经抽象摘要的覆盖率。

在精确率 Precision、召回率 Recall 和 F1 值上，本文模型超过了各对比实验模型。相较于 DSR 模型，本文模型在 Precision(P)、Recall(R)和 F1 值(F1)上分别提高了 20.46、19.94 和 20.39 个百分点，在 ROUGE-L 指标上提高了 7.93 个百分点。与 PG+DimAttn 模型相比，本文模型的 F1 值提高了 0.05 个百分点，在 ROUGE-L 指标上提高了 1.95 个百分点。综合 ROUGE 指标和 BERTScore 指标结果表明，基于语义的奖励方法能够提高摘要语义信息，摘要信息性和流畅性也得到进一步提升。

表 4　不同模型的 BERTScore 和 ROUGE-L 评价指标对比　单位：%

Tab. 4　Comparison of BERTScore and ROUGE-L evaluation indicators of different models　　unit：%

模型		BERTScorede 得分 P	BERTScorede 得分 R	BERTScorede 得分 F1	ROUGE-L
文献[8]模型	XENT	62.18	63.79	62.77	29.46
	ROUGE	59.31	61.33	60.11	33.89
	ROUGE+XENT	61.07	62.17	61.38	33.07
	DSR+ROUGE	65.92	64.56	65.01	33.61
	DSR+XENT	66.06	67.63	66.64	31.28
	DSR	66.27	67.98	66.93	30.96
文献[27]模型	PG	—	—	86.72	33.63
	PG+Cov.	—	—	87.24	35.92
	PG+DimAttn	—	—	87.27	36.94
本文模型	A2C-RLAS	**86.73**	**87.92**	**87.32**	**38.89**

3.6　消融实验

本节对 A2C-RLAS 模型中的重写器(rewriter)和强化学习(RL)进行消融分析，结果如表 5 所示，指标最高的分数用加粗数字标出。在 ROUGE-1、ROUGE-2 和 ROUGE-L 指标上，A2C-RLAS 相较于 A2C-RLAS(no RL)分别提高了 2.95、2.09 和 2.85 个百分点，相较于 A2C-RLAS(no rewriter)分别提高了 0.16、0.05 和 1.13 个百分点，相较于 A2C-RLAS(no RL, no rewriter)分别提高了 1.16、0.10 和 2.48 个百分点。在 BERTScore 的 Precision、Recall 和 F1 值指标上，A2C-RLAS 相较于 A2C-RLAS(no RL)分别提高了 5.49、5.29 和 5.39 个百分点，相较于 A2C-RLAS(no rewriter)分别提高了 0.32、0.57 和 0.44 个百分点，相较于 A2C-RLAS(no RL, no rewriter)分别提高了 2.41、1.71 和 2.07 个百分点。

表 5　不同模型的 ROUGE 评价指标对比　　单位：%

Tab. 5　Comparison of ROUGE evaluation indicator of different models　　unit：%

模型	ROUGE 分数 1	ROUGE 分数 2	ROUGE 分数 L	BERTScorede 得分 P	BERTScorede 得分 R	BERTScorede 得分 F1
A2C-RLAS(no RL)	38.38	16.12	36.04	81.24	82.63	81.93
A2C-RLAS(no rewriter)	41.17	18.16	37.76	86.41	87.35	86.88
A2C-RLAS(no rewriter, no RL)	40.17	18.11	36.41	84.32	86.21	85.25
A2C-RLAS	**41.33**	**18.21**	**38.89**	**86.73**	**87.92**	**87.32**

上述性能提升的原因是不含重写器的模型直接抽取原文关键句组成摘要，而原文整句往往过长，没有重写关键句导致摘要包含过多冗余信息，并且参考摘要往往是重写过的，比较精简，导致计算参考摘要和抽取摘要的精确匹配时 ROUGE 值较低，相应的 BERTScore 值也较低。不含强化学习的模型的抽取过程没有多次迭代，导致抽取出非关键句，经重写后无法与参考摘要相匹配，ROUGE 评分低，此时参考摘要和模型生成摘要的语义相似度也较低。上述实验验证了重写器(rewriter)和强化学习(RL)存在的必要性。

3.7　模型摘要示例

表 6 展示了本文模型和其他对比模型的摘要样例。从摘要内容可以发现，A2C-RLAS 模型摘要更加全面、高度贴近参考摘要，表明该模型经过强化学习的 A2C 算法多次迭代后，抽取器成

www.joca.cn

功抽取出全部关键句，且关键句经重写后，去除了冗余信息，保留核心内容且简洁流畅。经计算，A2C-RLAS模型在BERTScore的Precision、Recall和F1值的评价指标上有明显优势，表明本文模型摘要语义几乎和参考摘要语义一致，达到了摘要目的。

综合表4、5和表6可知，A2C-RLAS充分发挥了抽取式方法的优势，摘要内容紧扣文章中心；重写器充分发挥了生成式方法的优势，得到精简流畅的摘要，且没有句内重复及OOV词，对原文内容有较好的覆盖，总体表现良好。

表6 不同模型生成的摘要示例

Tab. 6 Examples of summaries generated by different models

模型	摘要内容	BERTScore
参考摘要	shadow chancellor refused to rule out a five-point rise in corporation tax . he was also evasive about the 40p rate of income tax during interview . businesses are already concerned about labour's plans not to cut tax . the main rate has been brought down by coalition from 28 to 21 per cent .	
Lead-3	the shadow chancellor refused to rule out a five-point rise in corporation tax . ed balls yesterday opened the door to a tax raid on businesses , which critics warn would endanger the economic recovery . in a tv interview , the shadow chancellor refused to rule out a five-point rise in the main rate of corporation tax if labour form the next government .	{BERTScore： Precision=88.2140, Recall=87.5591, F1=87.8853}
Refresh	the shadow chancellor refused to rule out a five-point rise in corporation tax . in a tv interview , the shadow chancellor refused to rule out a five-point rise in the main rate of corporation tax if labour form the next government . he was also evasive about the 40p rate of income tax , into which millions of people have been dragged in recent years , saying he was unsure how quickly the threshold could be raised . businesses are already concerned about labour's stated plans not to cut corporation tax -- paid by all large firms on their profits -- as the tories have vowed to if they win office .	{BERTScore： Precision=87.2914, Recall=91.8925, F1=89.5329}
Pointer-generator+coverage	shadow chancellor refused to rule out a five-point rise in corporation tax . he was also evasive about the 40p rate of income tax . businesses are already concerned about labour's stated plans not to cut corporation tax – paid by all large firms on their profits .	{BERTScore： Precision=93.5651, Recall=93.3751, F1=93.4700}
A2C-RLAS	shadow chancellor refused to rule out a five-point rise in corporation tax . he was also evasive about the 40p rate of income tax , into which millions have been dragged in recent years . businesses are already concerned about labour's plans not to cut corporation . main rate has been brought down under coalition from 28 % to 21 % .	{BERTScore： Precision=95.1275, Recall=97.1508, F1=96.1179}

www.joca.cn

4 结语

本文提出了面向长文本的基于A2C算法的强化自动摘要模型A2C-RLAS，该模型基于CNN和RNN混合神经网络的抽取器从原文中提取出关键信息，指导基于拷贝机制和注意力机制的重写器的重写，使用强化学习连接两个网络，结合基于语义的奖励方式训练整个模型。实验结果表明，A2C-RLAS模型在长文本自动摘要任务中，生成的摘要内容更准确，包含更多原文关键信息，摘要语言更加流畅，且有效避免了生成内容的重复。

此外，本文模型可以考虑进一步改进：首先，词嵌入可以使用预训练模型BERT；然后，为考察生成式摘要的“事实性”，可在模型中融入事实性的知识。最后，可考察本文模型对不同场景下的新数据集的适用性，比如学术论文摘要。

参考文献（References）

[1] 黄波，刘传才. 基于加权TextRank的中文自动文本摘要[J]. 计算机应用研究，2020，37(2)：407-410.(HUANG B, LIU C C. Chinese automatic text summarization based on weighted TextRank [J]. Application Research of Computers, 2020, 37(2): 407-410.)

[2] GAMBHIR M, GUPTA V. Recent automatic text summarization techniques: a survey[J]. Artificial Intelligence Review, 2017, 47(1): 1-66.

[3] ERKAN G, RADEV D R. LexRank: graph-based lexical centrality as salience in text summarization [J]. Journal of Artificial Intelligence Research, 2004, 22(1): 457-479.

[4] RUSH A M, CHOPRA S, WESTON J. A neural attention model for abstractive sentence summarization [C]// Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2015: 379-389.

[5] SHARMA E, HUANG L, HU Z, et al. An entity-driven framework for abstractive summarization [C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing/the 9th International Joint Conference on Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2019: 3280-3291.

[6] ZHANG J, ZHAO Y, SALEH M, et al. PEGASUS: pre-training with extracted gap-sentences for abstractive summarization [C]// Proceedings of the 37th International Conference on Machine Learning. La Jolla, CA: International Machine Learning Society, 2020:11328-11339.

[7] SAITO I, NISHIDA K, NISHIDA K, et al. Abstractive summarization with combination of pre-trained sequence-to-sequence and saliency models [EB/OL]. [2020-04-12]. https://arxiv. org/pdf/2003. 13028v1. pdf.
[8] LI S, LEI D, QIN P, et al. Deep reinforcement learning with distributional semantic rewards for abstractive summarization [C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing/ the 9th International Joint Conference on Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2019: 6038-644.
[9] CHENG H, CAI J, FANG Y. RL-Gen: a character-level text generation framework with reinforcement learning in domain generation algorithm case [C]// Proceedings of the 26th International Conference on Neural Information Processing, CCIS 1143. Cham: Springer, 2019: 690-697.
[10] ZHANG T, KISHORE V, WU F, et al. BERTScore: evaluating text generation with BERT [EB/OL]. [2020-03-07]. https://arxiv. org/pdf/1904. 09675v3. pdf.
[11] HERMANN K M, KOČISKÝ T, GREFENSTETTE E, et al. Teaching machines to read and comprehend [C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 1693-1701.
[12] LIN C Y. ROUGE: a package for automatic evaluation of summaries [C]// Text Summarization Branches Out: Proceedings of the ACL-04 Workshop. Stroudsburg, PA: Association for Computational Linguistics, 2004: 74-81.
[13] NARAYAN S, COHEN S B, LAPATA M. Ranking sentences for extractive summarization with reinforcement learning [C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg, PA: Association for Computational Linguistics, 2018: 1747-1759.
[14] CHEN Y, MA Y, MAO X, et al. Abstractive summarization with the aid of extractive summarization [C]// Proceedings of the 2018 Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data, LNCS 10987. Cham: Springer, 2018: 3-15.
[15] KIM Y. Convolutional neural networks for sentence classification [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: Association for Computational Linguistics, 2004: 1746-1751.
[16] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [EB/OL]. [2018-09-07]. https://arxiv. org/pdf/1412. 3555v1. pdf.
[17] VINYALS O, FORTUNATO M, JAITLY N. Pointer networks [C]// Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 2692-2700.
[18] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [EB/OL]. [2018-09-20]. https://arxiv. org/pdf/1409. 0473v7. pdf.
[19] GU J, LU Z, LI H, et al. Incorporating copying mechanism in sequence-to-sequence learning [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2016: 1631-1640.
[20] WILLIAMS R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning [J]. Machine Learning, 1992, 8(3/4): 229-256.
[21] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [EB/OL]. [2018-07-08]. https://arxiv. org/pdf/1301. 3781v3. pdf.
[22] KINGMA D P, BA J L. Adam: a method for stochastic optimization [EB/OL]. [2018-07-21]. https://arxiv. org/pdf/1412. 6980v9. pdf.
[23] LIU Y, TITOV I, LAPATA M. Single document summarization as tree induction [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg, PA: Association for Computational Linguistics, 2019: 1745-1755.
[24] SEE A, LIU P J, MANNING C D. Get to the point: summarization with pointer-generator networks [C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2017: 1073-1083.
[25] NALLAPATI R, ZHAI F, ZHOU B. SummaRuNNer: a recurrent neural network based sequence model for extractive summarization of document [C]// Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2017: 3075-3081.
[26] BAE S, KIM T, KIM J, et al. Summary level training of sentence rewriting for abstractive summarization [C]// Proceedings of the 2nd Workshop on New Frontiers in Summarization. Stroudsburg, PA: Association for Computational Linguistics, 2019: 10-20.
[27] HAN S, LIN X, JOTY S. Resurrecting submodularity in neural abstractive summarization [EB/OL]. [2020-10-10]. https://arxiv. org/pdf/1911. 03014v3. pdf.

www.joca.cn

This work is partially supported by the CERNET Innovation Project (NGII20170520).

DU Xixi, born in 1994, M. S. candidate. Her research interests include natural language processing.

CHENG Hua, born in 1975, Ph. D., research associate. His research interests include intelligent signal processing, information safety.

FANG Yiquan, born in 1975, M. S., senior experimentalist. Her research interests include information safety.