被引用論文の表現も利用した引用文生成

有田朗人 [1] 杉山 弘晃 [2]　堂坂浩二 [3] 田中陸斗 [3] 平博順 [1]
[1] 大阪工業大学大学院 [2]NTT コミュニケーション科学基礎研究所 [3] 秋田県立大学
m21a02@oit.ac.jp　h.sugi@ieee.org
{dohsaka,B20P025}@akita-pu.ac.jp

概要

本研究では，論文執筆者支援に必要な技術の一つとして引用文自動生成の精度向上に取り組んだ．高精度な引用文生成のためには，引用元と引用先の論文の関係性，引用する際の文脈といった情報や，引用文生成時の適切な文の表現パターンの学習が必要である．しかし，深層学習を利用した引用文生成では，学習時の入力データサイズの制限から引用文直前の 1 文を利用したり，論文のアブストラクトが利用されることが多かった．そのため，本研究では，引用文と類似した被引用論文中の文も引用文生成の学習に使用することで，入力データ量を抑えつつ，高精度な引用文生成モデルの学習を試みた．

1　はじめに

出版される科学技術論文数は増加の一途を辿っている．近年，arXiv [1] を始めとするプレプリントサーバの利用も進み，論文執筆時に引用すべき文献の量が一段と増加している．しかし，1 人の研究者が論文調査にかける時間は限られるため，論文執筆時に，引用すべき論文が抜け落ちてしまう危険が，年々増加している可能性がある．我々は，これまで論文執筆支援の枠組みについて考察し，論文支援に必要な技術がいくつかのフェーズに分かれることを明らかにしている [2]．本研究ではそのフェーズの一つである引用文生成技術について取り上げる．

引用文生成技術は，それまでの文脈を与える論文テキストと引用先論文のテキストを入力とし，文脈にあった適切な引用文を自動生成する技術である．従来，深層学習を用いた引用文生成手法がいくつか提案されている．例えば，Xing らは，sequence to sequence モデルの一種である，Pointer-Generator Networks [3] を利用した手法 [4] を提案している．Pointer-Generator Networks は，デコーダ部分で新たに単語を生成するだけでなく，学習に使用される引用元テキスト中の未知語も利用することで引用文生成精度を向上させようとしている．

また，Jia-Yan らは，引用文中で複数の論文を引用する場合にも対応するため，複数箇所の入力情報をコンパクトに扱うことができる Fusion-in-Decoder [5] を利用している．さらに，引用文生成精度を高めるため，論文引用の際の「引用意図」を自動判定し，引用意図カテゴリ [6] も学習・識別の入力とすることで引用文生成を試みている [7]．

高精度な引用文生成のためには，引用元と引用先の論文の関係性，引用する際の文脈といった情報や，引用文生成時の適切な文の表現パターンの学習が必要である．しかし，深層学習を利用した引用文生成では，学習時に入力される 1 データに対する情報は，それほど大規模なものは使用できないため，従来の研究では，引用の文脈情報として，引用文直前の 1 文を利用したり，論文のアブストラクトが利用されることが多かった．また，被引用論文の情報としても，被引用論文の本文全てを利用することは困難で，被引用論文のアブストラクトが利用されることが多かった．

しかし，論文のアブストラクトを学習の入力とした場合，引用文で実際に使用される具体的な表現が情報として含まれにくく，高精度な引用文生成が実現しにくいと考えられる．そのため，本研究では，引用文と類似した被引用論文中の文も引用文生成の学習に使用することで，入力データ量を抑えつつ，高精度な引用文生成モデルの学習を試みた．

2　提案手法

本研究では，従来の複数論文を引用する場合にも対応した引用文生成モデル [7] をベースとし，より高精度な引用文生成が実現するため，被引用論文中の文の表現も引用文生成モデルの学習に使用する．本研究の提案手法の概要を図 1 に示す．

提案手法では，引用文の直前の文，引用意図カテ

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

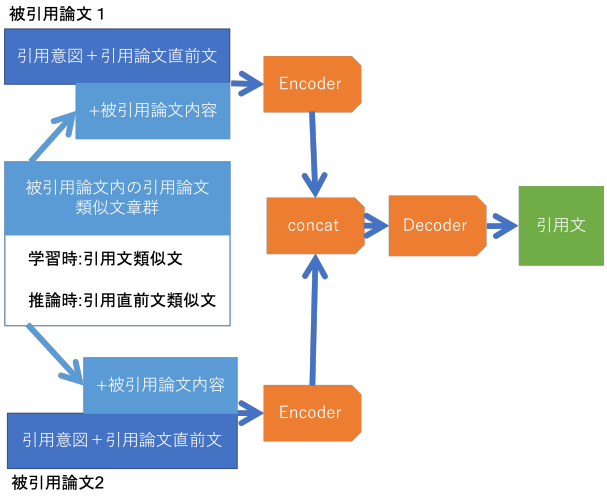

図 1　提案手法

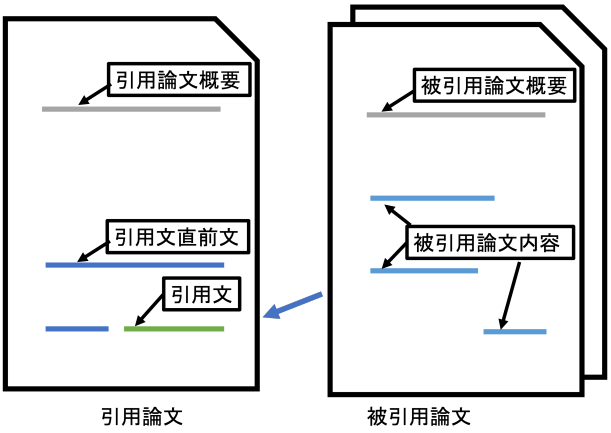

図 2　本研究で扱う用語について

ゴリ，被引用論文中の，引用文に文類似度が近い文集合を入力とし，正解の引用文を生成させるような学習を Fusion-in-Decoder [5] を用いて行う．

Jia-Yan らの手法 [7] では，被引用論文の情報として，被引用論文のアブストラクトが用いられていたのに比べ，提案手法では，被引用論文中の文の情報も学習することで，引用文の文の表現がより的確なものになることが期待される．

なお，本研究で使用する「引用文」，「引用論文概要」等の用語の関係を図 2 に示す．

3　評価実験

3.1　評価用データ

評価用データには，Xing らが作成した citation generation データセット [4] を用いた．このデータセットは，ACL Anthology Network(AAN) コーパス [8] 中の 1000 件の論文の引用文に対し，引用文献情報と引用文正解をアノテーションしたデータセットである．1000 件の論文のうち，600 件が訓練データ，400 件がテストデータである．

ただし，我々が確認したところ，一部のテストデータが訓練データに含まれていることが確認されたため，本研究では，重複した 103 件のデータを訓練データから除外して実験を行った．

3.2　実験設定

引用直前文は，引用文直前の 2 文を使用した．

被引用論文の文については，被引用論文の概要を除く論文本文全体のテキストから引用文に類似性の高い文を抽出して使用した．SentenceBERT[9] を用いて得られた文ベクトルについて，引用文とコサイン類似度で類似度 0.6 以上の上位 6 文を使用した．類似度 0.6 以上の文が 6 文ない場合は，類似度上位 3 文を学習時に使用した．テスト時は，引用文直前文と類似度の高い上位 6 文を入力に使用した．これらの類似文は，学習時，テスト時ともに文を連結して入力とした．

引用意図カテゴリは，生成する引用文の傾向を引用意図カテゴリで狭めることで，より意図する知識が抽出されること，引用意図カテゴリごとにおける特徴的な単語がより多く生成され，精度が上がるのではないかと考え利用した．め，あらかじめ引用文に対して Cohan らの引用意図推測モデル [6] を利用し，引用意図カテゴリを擬似的に付与したものを用いた．

実験では，学習，識別モデルとして Fusion-in-Decoder [5] を用い，引用文生成モデルの事前学習済みモデルに，T5[10] を用いた．

3.3　実験結果

実験結果を表 1 に示す．まず，被引用論文の概要と本文どちらが精度に貢献しているかを検証するため，引用側の情報を引用論文概要にして，被引用論文の情報をアブストラクト，論文本文の類似文に変更した場合の実験を行った．ROUGE-1 の評価で 0.15 ポイントとわずかに，被引用論文の内容を入力した時が高かった．しかし，ROUGE-2，ROUGE-L の評価ではそれぞれ 0.06 ポイント，1.1 ポイント低い結果となった．

このことから，被引用文の内容と被引用論文の概要にはほとんど違いがない，もしくは，引用論文の概要と被引用論文の概要，引用論文の概要と被引用

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

表 1 実験結果

入力データ	ROUGE-1	ROUGE-2	ROUGE-L
引用論文の概要, 被引用論文の概要	20.87	2.60	15.40
引用論文の概要, 被引用論文本文中の類似文	21.02	2.54	14.30
引用論文の直前文, 被引用論文の概要	19.44	2.14	14.11
引用論文の直前文, 被引用論文本文中の類似文	**22.08**	**3.43**	**16.52**

図 3 提案手法による引用文生成の例

引用文直前文 (入力):

Such an extraction-based definition of summarization has also been quite common in most existing general summarization work #OTHEREFR. By definition in order to generate an impact summary of a paper, we must look at how other papers cite the paper, use this information to infer the impact of the paper, and select sentences from the original paper that can reflect the inferred impact.

被引用論文中の類似文 (入力):

We believe that similar improvements can be achieved on other discourse annotation tasks in the scientific literature domain. For such information access and retrieval purposes, the relevance of a citation within a paper is often crucialIn particular,we plan to investigate the use of scientific attribution information for the citation function classification task.To demonstrate the use of automatic scientific attribution classification, we studied its util- ity for one well known discourse annotation taskIn bibli- ographic information retrieval, anchor text, i.e., the context of a citation can be used to charac- terise (index) the cited paper using terms out- side of that paperIn particular, we plan to investigate the use of scientific at- tribution information for the citation function classification task.

引用文 (正解):

This is because in citations, the discussion of the paper cited is usually mixed with the content of the paper citing it, and sometimes also with discussion about other papers cited #REFR.

提案手法による引用文生成の結果:

This paper, use the citation-based approach of #REFR to select sentences from the original paper that reflect the impact of the paper.

論文の内容それぞれの関係においては大きく違いがないということが考えられる．

次に，引用論文の直前文を固定し，被引用論文の概要と内容どちらが精度に貢献しているかを検証した．ROUGE-1，ROUGE-2，ROUGE-L スコアは，本文を使用した方がそれぞれ 2.64 ポイント，1.29 ポイント，2.41 ポイントとすべて向上する結果となった．この結果から，引用文の内容を取り入れることで精度が上がるのではないかと考えられる．ただし，引用文の本文の情報が加わったから精度が上がったのではなく，引用文直前文がセットで入力されがことで精度が向上してのではないかと考えている．

図 3 に本提案手法による引用文生成の結果の例を示す．引用文直前文をキーとし検索をかけたことで，今回の例においては"citation","paper"という共通キーワードが入力データにおいて複数回登場していることから，重要な単語としての引用文の生成で利用されやすくなったことが要因で，精度が向上したと考えられる．

4 おわりに

本研究では，引用論文の引用文直前文と，被引用論文の内容から引用論文の引用文の直前文と類似している文を新たに学習データに含め，組み合わせることで，引用文生成の精度向上を実現した．今後は，ROUGE 以外の人手等の評価，より大規模な引用文データセットによる評価なども行っていきたいと考えている．

謝辞

本研究の遂行にあたり，ご助言・ご協力をいただきました，NTT コミュニケーション科学基礎研究所成松宏美氏，電気通信大学南泰浩教授，工学院大学大和淳司教授，名古屋大学東中竜一郎教授，農研機構菊井玄一郎チーム長に感謝いたします．

参考文献

[1] Gerry McKiernan. arxiv.org: the los alamos national laboratory e-print server. **International Journal on Grey Literature**, Vol. 1, pp. 127–138, 2000.

[2] Hiromi Narimatsu, Kohei Koyama, Kohji Dohsaka, Ryuichiro Higashinaka, Yasuhiro Minami, and Hirotoshi Taira. Task definition and integration for scientific-

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

document writing support. In **Proceedings of the Second Workshop on Scholarly Document Processing**, pp. 18–26, Online, June 2021. Association for Computational Linguistics.

[3] A. See, Peter J. Liu, and Christopher D. Manning. Get to the point: Summarization with pointer-generator networks. In **ACL**, 2017.

[4] Xinyu Xing, Xiaosheng Fan, and Xiaojun Wan. Automatic generation of citation texts in scholarly papers: A pilot study. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics**, pp. 6181–6190, Online, July 2020. Association for Computational Linguistics.

[5] Gautier Izacard and Edouard Grave. Leveraging passage retrieval with generative models for open domain question answering. In **EACL**, 2021.

[6] Arman Cohan, Waleed Ammar, Madeleine van Zuylen, and Field Cady. Structural scaffolds for citation intent classification in scientific publications. In **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)**, pp. 3586–3596, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[7] Jia-Yan Wu, Alexander Te-Wei Shieh, Shih Ju Hsu, and Yun-Nung Chen. Towards generating citation sentences for multiple references with intent control. **ArXiv**, Vol. abs/2112.01332, , 2021.

[8] Dragomir R. Radev, Pradeep Muthukrishnan, Vahed Qazvinian, and Amjad Abu-Jbara. The acl anthology network corpus. **Language Resources and Evaluation**, pp. 1–26, 2013.

[9] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. **ArXiv**, Vol. abs/1908.10084, , 2019.

[10] Colin Raffel, Noam M. Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. **ArXiv**, Vol. abs/1910.10683, , 2020.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).