Manuscript ID : 00001-89308

Source ID : 00000067

JOURNAL OF COMPUTER ENGINEERING & TECHNOLOGY

Volume 7, Issue 1, January -June 2024, Pages 33-41, Page Count - 9

Synergy of Graph-Based Sentence Selection and Transformer Fusion Techniques For Enhanced Text Summarization Performance

Yugandhar Reddy Gogireddy [(1)*] Adithya Nandan Bandaru [(2)] Venkata Sumanth [(3)]

[(1)] Department of Computational Intelligence, School of Computing, SRM Institute of Science and Technology Chennai Main Campus (Deemed to be University), Kattankulathur, India.
[(2)] Department of Computational Intelligence, School of Computing, SRM Institute of Science and Technology Chennai Main Campus (Deemed to be University), Kattankulathur, India.
[(3)] Department of Computational Intelligence, School of Computin, SRM Institute of Science and Technology Chennai Main Campus (Deemed to be University), Kattankulathur, India.

Abstract

This paper presents a new method to improve text summarization by combining the strengths of Graph Neural Networks with Transformer-based models. Text summarization is pivotal in natural language processing; it deals with the compression of large documents into smaller summaries while retaining essential information. The work proposed here focuses on enhancing the summarization process through the segmentation of articles into Elementary Discourse Units and further selecting important ones for the summary. In this paper, we propose a multi-stage methodology involving the application of GNN for the selection of salient EDUs from input articles. The process, therefore, aims at arriving at the most important pieces of information requisite for summarization. Following that, we fuse the selected EDUs using BART (Bidirectional and Auto-Regressive Transformers), which is a recent Transformerbased architecture for text generation. In evaluating the performance of our method, experiments are carried out on the CNN/Daily Mail dataset since it is a widely used benchmark for text summarization. We compare our method against the baseline model in whose respect we assess its performance using ROUGE scores, a measure that judges the quality of summaries against human-written references. The experimental results show that the proposed approach can outperform the baseline methods with ROUGE scores.The precision, recall, and F1 scores achieved using our method are higher than those obtained using other baseline methods, implying that our method is suitable for generating informative and coherent summaries. Our study suggests that the performance of GNN, in combination with Transformer-based models for text fusion, holds promise in yielding an effective methodology for text summarization. We present a novel method for improving text summarization, a fundamental task in natural language processing. Summarization is aimed at distilling important information from a given huge amount of text or document into shorter, more comprehensive summaries while retaining the main points. Our approach optimizes the process by segmenting the articles into Elementary Discourse Units (EDUs) and selecting the most important ones for summarization. In the first stage of our approach, we apply Graph Neural Networks (GNN) to identify the most important EDUs from the input articles. This step is a way to bring out information that should be captured for summarization purposes. Next, we use BART to fuse these selected EDUs, following a common Transformer architecture famous for text generation tasks. We validated our method by performing experiments on the CNN/Daily Mail dataset, benchmarked widely in the field. We compared the proposed model against baseline methods for a performance measure based on ROUGE scores, primarily indicating similarity between summaries and human-generated references. As observed, our approach has outperformed the baseline models in terms of precision, recall, and F1 scores. It can be concluded, therefore, that our method yields more informative and coherent summaries. This study establishes that GNNs convey content selection capability that works best with mapping information of EDUs to their source text, and fusing information using Transformer-based models in text summarization presents promising advances.

Author Keywords

Text summarization, Graph Neural Networks (GNN), Transformer-based models, BART (Bidirectional and Auto-Regressive Transformers), Natural language processing (NLP), Elementary Discourse Units (EDUs), ROUGE scores, Document summarization, Neural network architectures, Information extraction.

ISSN Print: 2347-3908
Source Type: Journals
Publication Language: English
Abbreviated Journal Title: JCET
Publisher Name: IAEME Publication
Major Subject: Physical Sciences
Subject area: Artificial Intelligence

ISSN Online: 2347-3916
Document Type: Journal Article
DOI:
Access Type: Open Access
Resource Licence: CC BY-NC
Subject Area classification: Computer Science
Source: SCOPEDATABASE

Reference

[1] Vs, Vibashan, et al. "Image fusion transformer." 2022 IEEE International conference on image processing (ICIP). IEEE, 2022.
[2] Roy, Swalpa Kumar, et al. "Multimodal fusion transformer for remote sensing image classification." IEEE Transactions on Geoscience and Remote Sensing (2023).
[3] Ma, Jiayi, et al. "SwinFusion: Cross-domain long-range learning for general image fusion via swin transformer." IEEE/CAA Journal of Automatica Sinica 9.7 (2022): 1200-1217.
[4] Song, Y., Dai, Y., Liu, W., Liu, Y., Liu, X., Yu, Q., ... & Li, M. (2024). DesTrans: A medical image fusion method based on transformer and improved DenseNet. Computers in Biology and Medicine, 174, 108463.
[5] Ma, Kai & Tian, Miao & Tan, Yongjian & Xie, Xuejing & Qiu, Qinjun. (2022). What is this article about? Generative summarization with the BERT model in the geosciences domain. Earth Science Informatics. 15. 1-16. 10.1007/s12145-021-00695-2.
[6] Challagundla, Bhavith Chandra, Yugandhar Reddy Gogireddy, and Chakradhar Reddy Peddavenkatagari. "Efficient CAPTCHA Image Recognition Using Convolutional Neural Networks and Long Short-Term Memory Networks." International Journal of Scientific Research in Engineering and Management (IJSREM) (2024).
[7] Qu, Linhao, et al. "TransFuse: A unified transformer-based image fusion framework using self-supervised learning." arXiv preprint arXiv:2201.07451 (2022).
[8] Tang, Wei, and Fazhi He. "FATFusion: A functional–anatomical transformer for medical image fusion." Information Processing & Management 61.4 (2024): 103687.
[9] Zhang, Yundong, Huiye Liu, and Qiang Hu. "Transfuse: Fusing transformers and cnns for medical image segmentation." Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. Springer International Publishing, 2021.
[10] Tang, Wei, et al. "MATR: Multimodal medical image fusion via multiscale adaptive transformer." IEEE Transactions on Image Processing 31 (2022): 5134-5149.
[11] Tang, Wei, Fazhi He, and Yu Liu. "TCCFusion: An infrared and visible image fusion method based on transformer and cross correlation." Pattern Recognition 137 (2023): 109295.
[12] Yang, X., Huo, H., Wang, R., Li, C., Liu, X., & Li, J. (2023). DGLT-Fusion: A decoupled global–local infrared and visible image fusion transformer. Infrared Physics & Technology, 128, 104522.
[13] Li, Ming, et al. "MFVT: an anomaly traffic detection method merging feature fusion network and vision transformer architecture." EURASIP Journal on Wireless Communications and Networking 2022.1 (2022): 39.
[14] Odusami, Modupe, Rytis Maskeliūnas, and Robertas Damaševičius. "Pixel-level fusion approach with vision transformer for early detection of Alzheimer's disease." Electronics 12.5 (2023): 1218.
[15] Imfeld, Moritz, et al. "Transformer fusion with optimal transport." arXiv preprint arXiv:2310.05719 (2023).
[16] Zhang, Jun, et al. "Transformer based conditional GAN for multimodal image fusion." IEEE Transactions on Multimedia (2023).
[17] Huy, Pham Canh, et al. "Short-term electricity load forecasting based on temporal fusion transformer model." Ieee Access 10 (2022): 106296-106304.
[18] Siriwardhana, Shamane, et al. "Multimodal emotion recognition with transformer-based self supervised feature fusion." Ieee Access 8 (2020): 176274-176285.
[19] Rao, Dongyu, Tianyang Xu, and Xiao-Jun Wu. "TGFuse: An infrared and visible image fusion approach based on transformer and generative adversarial network." IEEE Transactions on Image Processing (2023).
[20] Khan, Aisha Urooj, et al. "Mmft-bert: Multimodal fusion transformer with bert encodings for visual question answering." arXiv preprint arXiv:2010.14095 (2020).
[21] Zheng, Guowei, et al. "A transformer-based multi-features fusion model for prediction of conversion in mild cognitive impairment." Methods 204 (2022): 241-248.
[22] Wang, Tao, et al. "O-Net: a novel framework with deep fusion of CNN and transformer for simultaneous segmentation and classification." Frontiers in neuroscience 16 (2022): 876065.
[23] Mustafa, H.T., Shamsolmoali, P. and Lee, I.H., 2024. TGF: Multiscale transformer graph attention network for multi-sensor image fusion. Expert Systems with Applications, 238, p.121789.

[24] Zhang, H., Zou, Y., Yang, X., & Yang, H. (2022). A temporal fusion transformer for short-term freeway traffic speed multistep prediction. Neurocomputing, 500, 329-340.
[25] Choi, Jae-Ho, Ki-Bong Kang, and Kyung-Tae Kim. "Fusion-Vital: Video-RF Fusion Transformer for Advanced Remote Physiological Measurement." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 2. 2024.
[26] Shvetsova, N., Chen, B., Rouditchenko, A., Thomas, S., Kingsbury, B., Feris, R. S., ... & Kuehne, H. (2022). Everything at once-multi-modal fusion transformer for video retrieval. In Proceedings of the ieee/cvf conference on computer vision and pattern recognition (pp. 20020-20029).
[27] Hu, J. F., Huang, T. Z., Deng, L. J., Dou, H. X., Hong, D., & Vivone, G. (2022). Fusformer: A transformer-based fusion network for hyperspectral image super-resolution. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.
[28] Liu, Jing, et al. "Image-text fusion transformer network for sarcasm detection." Multimedia Tools and Applications 83.14 (2024): 41895-41909.
[29] Alaba, Simegnew Yihunie, and John E. Ball. "Transformer-Based Optimized Multimodal Fusion for 3D Object Detection in Autonomous Driving." IEEE Access (2024).
[30] Jia, Sen, Zhichao Min, and Xiyou Fu. "Multiscale spatial–spectral transformer network for hyperspectral and multispectral image fusion." Information Fusion 96 (2023): 117-129.
[31] Liu, Yanyu, et al. "An Improved Hybrid Network with a Transformer Module for Medical Image Fusion." IEEE Journal of Biomedical and Health Informatics (2023).
[32] Park, Seonghyun, An Gia Vien, and Chul Lee. "Cross-modal transformers for infrared and visible image fusion." IEEE Transactions on Circuits and Systems for Video Technology (2023).
[33] Nagrani, Arsha, et al. "Attention bottlenecks for multimodal fusion." Advances in neural information processing systems 34 (2021): 14200-14213.
[34] Yi, Shi, et al. "TCPMFNet: An infrared and visible image fusion network with composite auto encoder and transformer–convolutional parallel mixed fusion strategy." Infrared Physics & Technology 127 (2022): 104405.