

Retraction

Retracted: Research on Intelligent English Translation Method Based on the Improved Attention Mechanism Model

Scientific Programming

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

Copyright © 2023 Scientific Programming. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] R. Wang, "Research on Intelligent English Translation Method Based on the Improved Attention Mechanism Model," *Scientific Programming*, vol. 2021, Article ID 9667255, 8 pages, 2021.

Research Article

Research on Intelligent English Translation Method Based on the Improved Attention Mechanism Model

Rong Wang 

School of Foreign Languages, Shaanxi Xueqian Normal University, Xi'an 710100, Shaanxi, China

Correspondence should be addressed to Rong Wang; rongw@mail.bnu.edu.cn

Received 14 October 2021; Revised 2 November 2021; Accepted 9 November 2021; Published 23 November 2021

Academic Editor: Bai Yuan Ding

Copyright © 2021 Rong Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of neural machine algorithms for English translation is a hot topic in the current research. English translation using the traditional sequential neural framework, which is too poor at capturing long-distance information, has its own major limitations. However, the current improved frameworks, such as recurrent neural network translation, are not satisfactory either. In this paper, we establish an attention coding and decoding model to address the shortcomings of traditional machine translation algorithms, combine the attention mechanism with a neural network framework, and implement the whole English translation system based on TensorFlow, thus improving the translation accuracy. The experimental test results show that the BLUE values of the algorithm model built in this paper are improved to different degrees compared with the traditional machine learning algorithms, which proves that the performance of the proposed algorithm model is significantly improved compared with the traditional model.

1. Introduction

Natural language is an important vehicle for knowledge and information dissemination, as well as an outward expression of human civilization and wisdom [1]. NLP, a cross-cutting field spanning computer linguistics and artificial intelligence, explores how computers can understand and process complex human language and, on this basis, achieves a true sense of human-computer interaction [2]. NLP covers a wide range of research areas, including machine translation, information extraction, text summary generation, question and answer systems, sentiment analysis, reading comprehension, opinion analysis, data mining, and many other areas that are currently very compelling [3]. As the Internet continues to evolve and the era of artificial intelligence approaches, the amount of data generated on the Internet is growing exponentially and how to efficiently process and extract useful data has become an urgent problem for major companies [4].

Language is the main tool for cultural exchange, but there is a huge language gap between the mother tongues of different countries, which undoubtedly brings many

obstacles to the cultural exchange of people in the world [5, 6]. Therefore, the demand for human translators is increasing, and this has led to the high price of human translators, which is difficult for the general public, to afford such an expensive price. The automated nature of machine translation makes translation between languages easy and efficient, which can undoubtedly contribute to a wide range of communication between countries around the world. According to investor opinion's forecast, with the increasing globalization of the economy, machine translation technology will play an increasingly important role in it and the market size of machine translation will reach \$1.5 billion in 2024 [7, 8].

Deep learning, a new machine learning method, is capable of automatically learning abstract features and establishing mapping relationships between input and output signals. Due to its powerful feature learning and representation capabilities, the application of deep neural networks to both speech recognition and image processing has yielded results that far exceed traditional methods, which have brought more dynamism to these fields [9]. With the recent success of deep learning in signal variable processing, more and more

researchers have started to work on applying neural networks to symbolic variables and have made similar progress, a classic example being the application of deep learning to reading comprehension [10], which likewise provides a new way of thinking for machine translation research.

Neural machine translation (NMT), based on deep learning, is capable of automatically learning abstract features and establishing relationships between source and target utterances and has recently obtained far better performance than SMT in various tasks of machine translation [11–14]. The encoder-decoder model is currently one of the most widely used models in NMT [10, 15], which first maps the source utterance into a distributed representation at the encoder side (source side) and then uses an attention mechanism at the decoder side (target side) to generate a target word cyclically [12]. Most encoders in previous studies have used recurrent neural networks (RNNs) to sequentially to encode the source-side utterances [13], but this has some shortcomings. Most notably, although RNNs can learn relationships between words, they still cannot solve the long-term dependency problem and are very slow to train due to the model structure. Self-attention networks (SANs) can not only learn relationships between words like RNNs [14] but also capture relationships between words by explicitly focusing on all words regardless of the distance between them and solve the problem of slow computation of RNNs.

Therefore, machine translation models based on SANs are currently doing the most excellent performance in various machine translation tasks. In this paper, we propose an English-Chinese translation model with improved attention mechanism to address the shortcomings of the traditional decoder-translator algorithm framework. The core idea of the algorithm is to combine the attention mechanism with neural networks and use deep learning methods to train the local attention of the translation model. Therefore, it can improve the translation system's ability to connect the context and then effectively improve the translation quality.

2. Related Work

Banchs and Costa-jussà [15] spent decades perfecting his invention; the machine they constructed not only failed to achieve the desired results in practice but also translated sentences with very poor readability, which made many scientists at the time to consider the invention useless. The conjecture of machine translation was therefore shelved until the emergence of electronic computers, and British and American engineers Booth and Weaver thought that automatic translation could be carried out with the help of computers [16].

With the introduction of statistical machine translation in 1990, the research on machine translation gradually began to go into a boom period and thus also entered a period of rapid development [17]. Unlike the rule-based or instance-based machine translation in the past, statistical machine translation has a very good mathematical foundation and can effectively utilize large-scale corpus, which also results in significant improvement in the performance of translation. This time also saw the emergence of many studies related to

statistical machine translation, such as word-based SMT, phrase-based SMT, and grammar-based SMT [18].

And, with the proposal of neural machine translation in 2013 [19], the research of machine translation has thus entered a brand new era. Taking neural network as the basic unit to construct the corresponding machine translation model, using the ability that neural network can automatically learn abstract features and establish the mapping relationship between input signal and output signal, with the support of large-scale corpus, the performance of neural machine translation has reached an unprecedented height and the performance on various machine translation tasks is far more than that of statistical machine translation [20], which becomes the dominant research approach. With the introduction of Transformer model by Google in 2017 [21], it has pushed the quality of machine translation to another peak, the accuracy rate has reached an astonishing 85%, and most sentences can be translated more accurately, which can basically meet people's daily needs, and the field of machine translation has thus shown a prosperous scene.

The Institute of Linguistics, Chinese Academy of Social Sciences, Tsinghua University, Harbin Institute of Technology, Northeastern University, Hong Kong University of Science and Technology, and University of Macau are all conducting research on machine translation [22, 23]. In the past few years, we have obtained a series of self-developed machine translation systems, such as “THUMT” of Tsinghua University, “Multi-language Machine Translation System” of Chinese Academy of Sciences, and “Chinese-Portuguese-English Translation System” of the University of Macau [24, 25], and a series of excellent achievements. In addition, Internet companies such as Baidu, Netease, Tencent, KDXunfei, and Sogou are also increasing their research on machine translation, which has given birth to translation software such as Baidu Translator, Sogou Translator, and Youdao Translator and real-time voice translation tools such as Xunfei Translator [26, 27].

3. Basic Modeling Research

3.1. Encoder-Decoder Architecture Model. The encoder encodes the input data information of the neural network into a fixed-length piece of data. The decoder takes the data encoded by the encoder and decodes it in reverse and then outputs it as a translated sentence. This is the idea underlying the sequence model [28], as shown in Figure 1.

3.2. Recurrent Neural Networks. The encoder-decoder framework is part of a neural network, and running the framework requires building a proper neural network model. Among the many neural network models, recurrent neural networks are the most widely used. RNNs are deformed models of feedforward neural networks, whose main feature is that they can handle data sequences of different lengths. Figure 2 shows the structure of recurrent neural network algorithm, which has the recursive property that the state of each time has a greater relationship with the previous activation state [29].

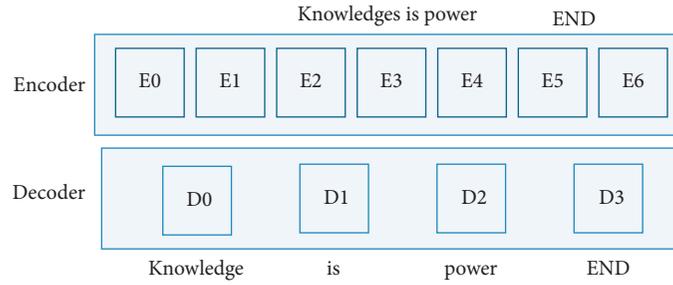


FIGURE 1: Encoder-decoder translation flowchart.

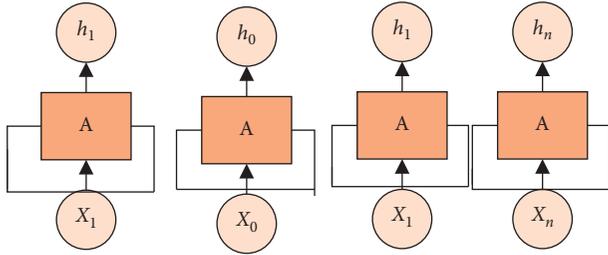


FIGURE 2: Recurrent neural network algorithm structure.

Although the input data of the theoretical recurrent neural network can be of infinite length, in the experimental state, if the data length is infinite, it will cause the problem of excessive gradient of the neural network, so the neural network may not be able to capture the correlation between the chapter contexts, resulting in the degradation of translation quality [30].

3.3. Recurrent Neural Networks with Attention Mechanism. Due to the various drawbacks of recurrent neural networks, the current mainstream approach is to combine recurrent neural networks with encoder-decoders, which can reduce the gradient problem caused by long data sequences [31].

Since the recurrent neural network operation mechanism is a left-right sequential operation process, this will significantly limit the parallel operation capability of the model itself and the sequential operation of data will also cause the problem of data module loss. The above problem will be improved by using the attention mechanism, which can change the data distance to 1 at any position in the translated data, so that it does not depend on the effect of the previous sequential operation on the current operation and the system will have better parallelism [32].

The mathematical principle of the attention mechanism is that the input data of the neural network is first weighted and imported into the encoder, which then imports the data to the decoder and the decoder queries the data weights in the decoding process as the inverse input data and then realizes the weighted average of the data of each state. The simplified implementation flow of the attention mechanism is shown in Figure 3 [11].

3.4. Model Construction. To better highlight the advantages of the attention mechanism, the model framework in this

paper is based on a recurrent neural network model incorporating the attention mechanism for building an encoder-decoder framework to implement the translation task. The neuronal connectivity part of the model is implemented using the attention mechanism, which can bring out the advantages of the attention mechanism better.

3.5. Model Framework. Figure 4 shows the attention mechanism model constructed in this paper, the overall structure of which consists of an encoder and a decoder. Among them, the encoder consists of a header attention single-layer structure and a preceding network single-layer structure and the number of blocks of the whole encoder is N_c . The decoder structure is similar to the encoder and also consists of N_c block structures, except that there is no head attention layer.

This neural network uses a differential network connection approach, the distinctive feature of which is that the network enters the normalization hierarchy (Add and Norm) for data processing [33].

3.6. Construction of the Attention Mechanism Module. The attention mechanism module is mainly divided into the encoder module and the decoder module. The input part of the encoder module is the whole data sequence, and the input matrices are set to Q , K , and V . The attention mechanism function can be regarded as a mapping relationship, and the function equation is [15]

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

Another important structure in the encoder module is the head attention single-layer structure, the model of which is shown in Figure 5.

4. Experiment and Analysis

4.1. Extraction of Feature Parameters. In order to further improve the system computing efficiency and reduce the interference of data unrelated to the speech signal, the relevant information data should be unified and thus find the parameter features and then realize the subsequent calculation. Figure 6 shows the structure for extracting feature parameters.

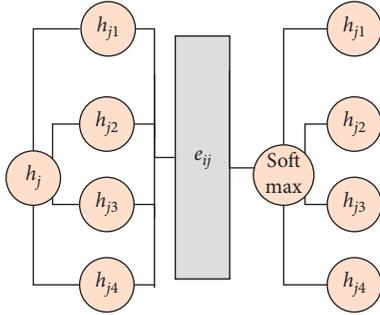


FIGURE 3: Sketch of attention mechanism.

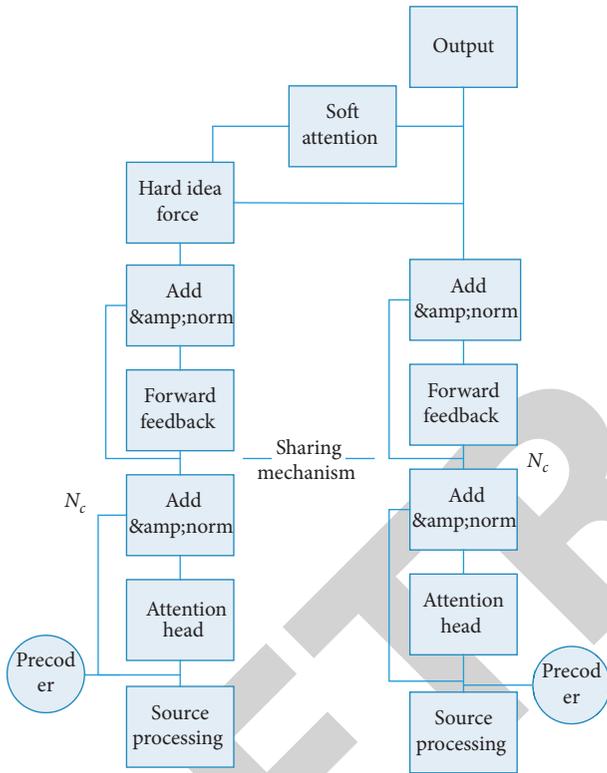


FIGURE 4: Model framework structure.

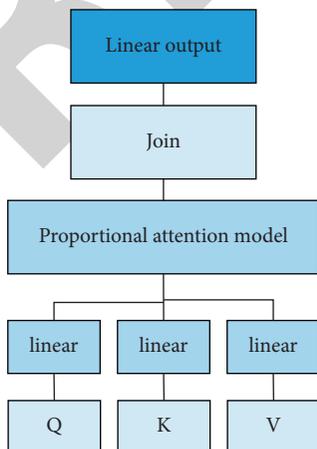


FIGURE 5: Structure schematic.

Nonperiodic continuous time signals use the Fourier transform to calculate the continuous map of the signal, but the actual control system is used to derive continuous signal discrete sampling values, so as to calculate the signal map by discrete sampling values. Fast Fourier transform [2] is as follows:

$$X[K] = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)nk}, \quad k = 0, 1, 2, \dots, N, \quad (2)$$

where $x[n]$ is the discrete speech sequence obtained by sampling and $X[K]$ is the k -point reset sequence. The discrete speech sequences are transformed into Mel frequency scale:

$$\text{Mel}(f) = 2579 \lg \left(1 + \frac{f}{700} \right), \quad (3)$$

where $\text{Mel}(f)$ is the Mel frequency and f is the actual frequency.

The discrete cosine transform DTC is carried out on the filtered output to derive the feature parameter extraction result P of the speech signal $w(n)$, which is calculated as

$$P = \sum_{n=1}^N F(l)w(n)\cos(\pi n(M + 0.5)). \quad (4)$$

After a speech signal generates a spectrogram to be weighted, windowed, and framed, each short-time analysis window is able to obtain spectral information by fast Fourier transform and then the MFCC two-dimensional map is derived using Mel filtering [14, 34].

4.2. Experimental Environment Setup. The experimental dataset used in this paper is the IWSLT2018 corpus data collection, which has a small amount of data. The experimental environment uses Python to program the attention mechanism neural network, and the parameters of the experimental setup environment are shown in Table 1.

4.3. Experimental Tests and Result Analysis. The experiments were performed as follows:

- (1) Processing the corpus and cutting long sentences into words
- (2) Numbering the words and storing them as files, and then, storing the files to the PC
- (3) Normalizing the text, completing the sentences that are not long enough, and intercepting the sentences that are too long
- (4) Training the processed sentences, and then, evaluating the BLUE values [12, 13]

In the experimental tests, the RNN (recurrent neural network) translation model, LSTM translation model, and neural network translation model incorporating the attention mechanism was evaluated using the comparison test method. The test results are shown in Table 2.

Table 2 shows the test results of the three models with the default corpus. It can be seen that the basic RNN model has

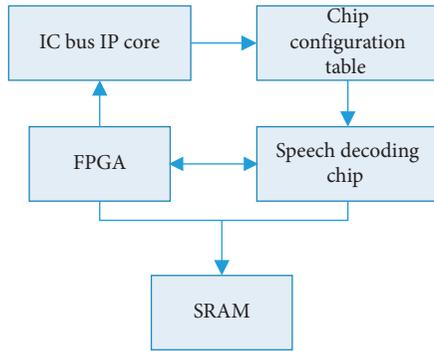


FIGURE 6: Structure for extracting feature parameters.

TABLE 1: Environmental parameters of the experimental setup.

Parameters	Content
Corpus data size	25000
Word dimension	512
Layers of neural network	2
Deep learning framework	TensorFlow
Hardware configuration	i7-9700K, 128 GB

TABLE 2: Test results of the 3 models.

Model	BLUE value
RNN	15.88
LSTM	21.03
ATT + RNN	24.33

the lowest BLUE value and the model built in this paper has the highest BLUE value, which also indicates that the proposed model does have some improvement on translation quality.

In the second comparison experiment, the sentences of the corpus are clustered according to their lengths and then the translation models are tested with sentences of different lengths to verify the long sentence translation ability of the translation models; the experimental results of different models are shown in Table 3.

Table 3 shows the performance comparison of machine translation models with different network structures under the English-German machine translation task. The first column of the table indicates the different kinds of translation models, which include RNN and its variant models such as LSTM state and GRU [27]. All three different RNN structures were modeled as well as experimented, and the one with the highest BELU value among the three was taken for presentation in the table. “Attn” indicates the attention mechanism, while “+Attn” indicates the addition of the attention mechanism to the model. As for “+IntHeads (8),” it indicates replacing the multiheaded attention network in the Transformer model with an interactive multiheaded attention network with 8 attention heads, while “+IntHeads (16)” replaces the multiheaded attention network in the Transformer model with an interactive multiheaded attention network with 16 heads, noting that the number of heads increases, but the parameters of the model hardly increase.

In order to compare the performance of these translation models more visually, the data in Table 3 are visualized as a histogram, as shown in Figure 7.

As can be seen from Figure 7, the Transformer model has the highest score both in the development set and in the test set, and as a machine translation model built using only the self-attention mechanism, its performance does outperform the machine translation models based on RNN and CNN. Comparing the first to the third columns of Table 3, we can find that the performance of the translation model based on RNN or its variants is average without using the attention mechanism, while the performance of the model achieves a more obvious improvement after adding the attention mechanism. The fourth row of the table shows the translation models based on CNN and attention mechanism. Due to the characteristics of CNN, its ability to obtain local information is greatly enhanced, so its performance is greatly improved compared with the first three models, which also confirm the importance of local information to the performance of machine translation.

In addition, by comparing the last three histograms in each dataset, it can be found that the performance of the model is significantly improved after replacing the multi-headed attention network in the Transformer model with interactive multiheaded attention network, except for the test set Test1. This also validates the effectiveness of the interactive multiheaded attention network, which enables the attention heads to share the learned feature representations by improving the performance of the model. It is also found that the more the number of attention heads, the better the interaction between attention heads in the interactive multiheaded attention network and thus the better the performance of the model.

For example, compared with CNN-based machine translation models, the self-attention mechanism focuses on all words in a sentence at the same time, which makes the Transformer model unable to learn local information well, so if certain measures can be taken to enhance the ability of the Transformer model, the performance of the Transformer model will be further improved to acquire local information.

In order to fully demonstrate the advantages of the designed model, the syntax-based and phrase-based intelligent recognition models are used in the experimental process to realize the comparison experiment. The number of control points in each system is recorded, and the distribution of control points is analyzed. The distribution of nodal points can describe the semantic and contextual relevance of English translation, and the dense distribution of nodal points indicates that the system has a high accuracy of English translation recognition. Figures 8 and 9 both show the distribution of the system’s recognition nodes. The compact distribution of nodes in Figure 8 indicates that the system has higher recognition performance and more accurate calibration results, and the problem of incoherence in English translation has been solved; the loose distribution of nodes in the syntactic and phrase-based recognition system in Figure 9 and the compact distribution of nodes in the 1st, 4th, and 5th experiments indicate that the system has higher calibration accuracy. But the translation result warning

TABLE 3: Translation model performance comparison.

Translation model	Dev	Test1	Test2	Test3	Average
RNN	16.6	15.9	17.2	16.3	16.5
RNN + Attn	19.0	19.5	20.9	19.4	19.8
CNN + Attn	28.1	26.1	29.1	25.6	27.0
Transformer	31.8	30.8	33.6	29.3	31.2
Transformer IntHeads (8)	32.7	30.7	34.8	30.7	32.1
Transformer IntHeads (16)	33.0	30.9	35.2	31.1	32.4

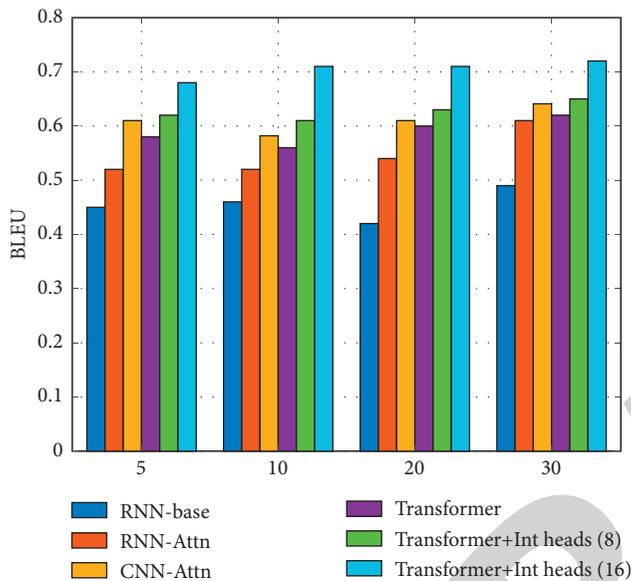


FIGURE 7: Performance of different machine translation models.

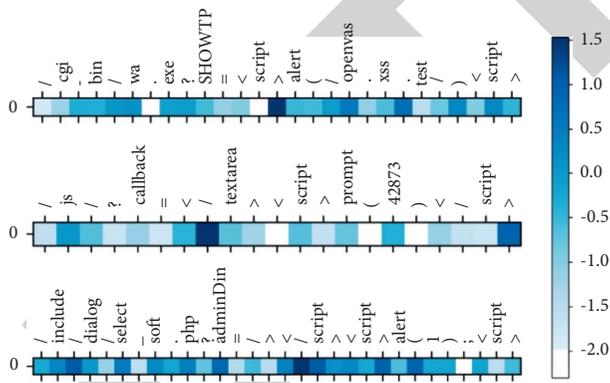


FIGURE 8: Attention weight distribution in word translation.

coherence is relatively poor. Secondly, this system alternates between loose and compact distribution of nodal control points, which indicates that it is not stable [15].

The above description shows that the intelligent recognition model for English translation designed in the paper has high proofreading accuracy and can identify the problem of contextual incoherence in English translation results, giving translation results that are consistent with contextual coherence and rationality.

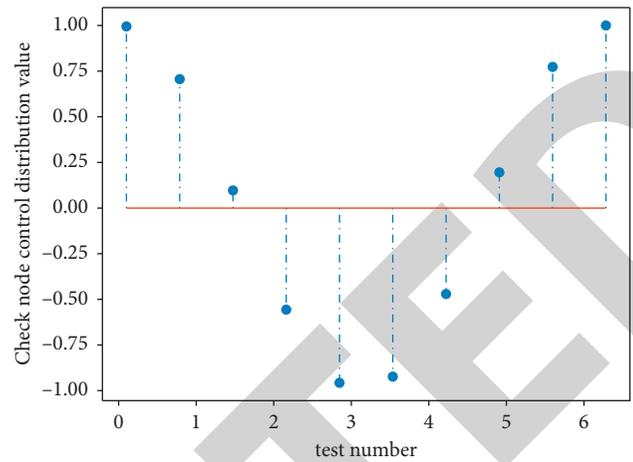


FIGURE 9: Intelligent proofreading control point distribution.

5. Conclusions

To address the shortcomings of the traditional encoding-decoding algorithm translation model, which suffers from inaccurate translation and semantic mutilation, this paper proposes to combine the attention mechanism and recurrent neural network model to build an English-Chinese translation model with improved attention mechanism and use TensorFlow to train the translation model at the same time.

Data Availability

The datasets used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Scientific Research Fund Project of Shaanxi Xueqian Normal University in 2020 named “Research on the Applicability of Online Translation to English Historical Novel Translation in the Age of Artificial Intelligence” under Grant No. 2020YBRS14.

References

- [1] W. Gong, H. Chen, and Z. Zhang, “A Novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-svm and multichannel data fusion,” *Sensors*, vol. 19, no. 7, Article ID 1693, 2019.
- [2] T. Nishihara, A. Tamura, T. Ninomiya, Y. Omote, and H. Nakayama, “Supervised visual attention for multimodal neural machine translation,” in *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 4304–4314, Barcelona, Spain, December 2020.

- [3] S. Feng, S. Liu, N. Yang, M. Li, M. Zhou, and K. Q. Zhu, "Improving attention modeling with implicit distortion and fertility for machine translation," in *Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics*, pp. 3082–3092 Technical Papers, Osaka, Japan, December 2016.
- [4] P. Koehn, *Statistical Machine translation*, Cambridge University Press, Cambridge, UK, 2009.
- [5] J. Bao, D. Tang, N. Duan et al., "Table-to-text: describing table region with natural language," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, LA, USA, February 2018.
- [6] X. Wang, Z. Tu, and M. Zhang, "Incorporating statistical machine translation word knowledge into neural machine translation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 12, pp. 2255–2266, 2018.
- [7] R. L. Smith and Z. Fang, "Techniques, applications and future prospects of diamond anvil cells for studying supercritical water systems," *The Journal of Supercritical Fluids*, vol. 47, no. 3, pp. 431–446, 2009.
- [8] X. M. Yu, W. Z. Feng, H. Wang, Q. Chu, and Q. Chen, "An attention mechanism and multi-granularity-based Bi-LSTM model for Chinese Q&A system," *Soft Computing*, vol. 24, no. 8, pp. 5831–5845, 2020.
- [9] H. Ryo, N. Takahiro, and M. Toshiyuki, "Translation and inclination control for intelligent tension Pole based on mode decoupling method," *Ieej Journal of Industry Applications*, vol. 7, no. 2, pp. 158–165, 2018.
- [10] G. R. Mirams, Y. Cui, A. Sher et al., "Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk," *Cardiovascular Research*, vol. 91, no. 1, pp. 53–61, 2011.
- [11] C. Shi, D. Lin, and T. Ishida, "Agent metaphor for machine translation mediated communication," in *Proceedings of the 2013 international conference on Intelligent user interfaces*, South California, U.S.A, March 19–22 2013.
- [12] T. Xie, C. Zhang, Z. Zhang, and K. Yang, "Utilizing active sensor nodes in smart environments for optimal communication coverage," *IEEE Access*, vol. 7, pp. 11338–11348, 2018.
- [13] S. Chhaniyara, C. Brunskill, B. Yeomans et al., "Terrain trafficability analysis and soil mechanical property identification for planetary rovers: a survey," *Journal of Terra-mechanics*, vol. 49, no. 2, pp. 115–128, 2012.
- [14] Y. Gao, Y. Li, R. Y. Lau, Y. Xu, and M. Bashar, "Finding semantically valid and relevant topics by association-based topic selection model," *ACM Transactions on Information Systems*, vol. 9, no. 1, pp. 3.1–3.22, 2018.
- [15] R. E. Banchs and M. R. Costa-jussà, "A vector-space dynamic feature for phrase-based statistical machine translation," *Journal of Intelligent Information Systems*, vol. 37, no. 2, 2010.
- [16] C. Karthik, K. Suresh, K. Valarmathi, and R. Jacob Rajesh, "Model-based control for moisture in paper making process," *Advances in Intelligent Systems and Computing*, vol. 324, pp. 257–264, 2015.
- [17] Z. Zhang, C. Zhang, M. Li, and T. Xie, "Target positioning based on particle centroid drift in large-scale WSNs," *IEEE Access*, vol. 8, pp. 127709–127719, 2020.
- [18] L. Jin, "Research on pronunciation accuracy detection of English Chinese consecutive interpretation in English intelligent speech translation terminal," *International Journal of Speech Technology*, no. 5, pp. 1–8, 2021.
- [19] T. B. Adji, B. Baharudin, and N. B. Zamin, "Annotated disjunct in link grammar for machine translation," in *Proceedings of the International Conference on Intelligent & Advanced Systems*, pp. 205–208, IEEE, Kuala Lumpur, Malaysia, November 2007.
- [20] X. Li, L. Liu, Z. Tu, S. Shi, and M. Meng, "Target foresight based attention for neural machine translation," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 1380–1390 (Long Papers), New Orleans, Louisiana, June 2018.
- [21] Z. Zhang, W. Li, and Q. Su, "Automatic translating between ancient Chinese and contemporary Chinese with limited aligned corpora," in *Proceedings of the CCF International Conference on NLP and Chinese Computing*, pp. 157–167, Springer, Dunhuang, China, October 2019.
- [22] Y. Yin, J. Su, H. Wen, J. Zeng, Y. Liu, and Y. Chen, "POS tag-enhanced coarse-to-fine attention for neural machine translation," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 18, no. 4, pp. 1–14, 2019.
- [23] Y. X. Li, Q. Wang, and Y. J. Zhang, "A Novel information technology based college physical education model[J]," *Advances in Intelligent Systems and Computing*, vol. 191, pp. 53–56, 2013.
- [24] A. Gelbukh, "Lecture Notes in Computer Science] Computational Linguistics and Intelligent Text Processing," *English-Arabic Statistical Machine Translation: State of the Art*, vol. 9041, pp. 520–533, 2015, (Chapter 39).
- [25] C. Zhang, T. Xie, K. Yang et al., "Positioning optimisation based on particle quality prediction in wireless sensor networks," *IET Networks*, vol. 8, no. 2, pp. 107–113, 2019.
- [26] Z. Shi, X. Zhang, J. Tu, and Z. Yang, "An efficient and improved particle swarm optimization algorithm for swarm robots system," *Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013*, vol. 212, no. 212, pp. 329–337, 2013.
- [27] J. Mizera-Pietraszko and P. Pichappan, "Study of the Establishment of a Reliable English-Chinese Machine Translation System Based on Artificial Intelligence," *Advances in Intelligent Systems and Computing*, vol. 613, pp. 13–23, 2018, (Chapter 2).
- [28] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 5998–6008, Long Beach, CA, USA, December 2017.
- [29] D. Wu, C. Zhang, L. Ji, R. Ran, H. Wu, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, 2021.
- [30] W. Qiu, Y. J. Shu, and Y. J. Xu, "Research on Chinese multi-documents automatic summarizations method based on improved TextRank algorithm and seq2seq," in *Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing*, pp. 196–201, Harbin, China, January 2021.
- [31] S. Zhao and Z. Zhang, "Attention-via-attention neural machine translation," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, LA, USA, February 2018.
- [32] S. R. Indurthi, I. Chung, and S. Kim, "Look harder: a neural machine translation model with hard attention," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 3037–3043, Florence, Italy, July 2019.
- [33] H. Choi, K. Cho, and Y. Bengio, "Fine-grained attention mechanism for neural machine translation," *Neurocomputing*, vol. 284, pp. 171–176, 2018.

- [34] S. Abujar, A. K. M. Masum, A. Bhattacharya, S. Dutta, and S. A. Hossain, "English to Bengali neural machine translation using global attention mechanism," vol. 3, pp. 359–369, in *Proceedings of the Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020*, vol. 3, pp. 359–369, Springer Singapore, Kolkata, India, May 2021.

RETRACTED