

Retraction

Retracted: LSTM-Based Attentional Embedding for English Machine Translation

Scientific Programming

Received 29 August 2023; Accepted 29 August 2023; Published 30 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] L. Jian, H. Xiang, and G. Le, "LSTM-Based Attentional Embedding for English Machine Translation," *Scientific Programming*, vol. 2022, Article ID 3909726, 8 pages, 2022.

Research Article

LSTM-Based Attentional Embedding for English Machine Translation

Lihua Jian ¹, Huiqun Xiang^{2,3} and Guobin Le^{2,3}

¹School of International Education, Hunan University of Medicine, Huaihua 418000, Hunan, China

²Changsha Vocational and Technical College, Changsha 410200, Hunan, China

³School of Foreign Languages, Huaihua University, Huaihua 418000, China

Correspondence should be addressed to Lihua Jian; 202110080102@hunnu.edu.cn

Received 23 December 2021; Revised 9 January 2022; Accepted 17 January 2022; Published 16 March 2022

Academic Editor: Baiyuan Ding

Copyright © 2022 Lihua Jian et al. 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.

In order to reduce the workload of manual grading and improve the efficiency of grading, a computerized intelligent grading system for English translation based on natural language processing is designed. An attention-embedded LSTM English machine translation model is proposed. Firstly, according to the characteristics of the standard LSTM network model that uses fixed dimensional vectors to represent words in the encoding stage, an English machine translation model based on LSTM attention embedding is established; the structure level of the English translation scoring system is constructed. A linguistic model of the English translation scoring system is established, and the probability distribution of a particular sentence sequence or word sequence of the translated text is statistically calculated using the model. The results show that the English machine translation model based on LSTM attention embedding proposed in this study can enhance the representation of the source language contextual information and improve the performance of the English machine translation model and the quality of the translation compared with the English machine translation models constructed by existing neural network structures, such as standard LSTM models, RNN models, and GRU-Attention translation models.

1. Introduction

With the development of computer technology and the maturity of artificial intelligence technology, machine translation is gradually replacing human translation and occupying a larger proportion in the translation field. At present, there are four main types of machine translation [1–3]. Among them, neural network-based machine translation models can alleviate the problem of feature design of high-dimensional data and improve the expressiveness of the model by building neural network classifiers when dealing with high-dimensional complex data, which has become the most popular and effective language translation model nowadays [4, 5].

In the literature [6], an English translation scoring system based on hidden Markov model is used, combining Markov model and Viterbi comparison system to input similar words between the translation and the reference

translation, match the similar words to calculate the proximity between them, and then compare the similarity between the translated utterances, and according to the comparison results, achieve the translation scoring [7]. The accuracy of the scoring results of this system is high, but the computation is large and time-consuming. Corpus-based English translation scoring system designed in [8] obtains word alignment ratios by analyzing word collocations in the structure of corpus materials, compares the word collocations and structure of the input translations, and scores the translations. The scoring results of this system have large errors and the process of word collocation analysis is complicated. Translation models such as those based on LSTM, RNN, and GRU-Attention neural networks have been widely used in the field of English machine translation [8–11] using neural networks with different structures to study the translation effect of English machine translation in the field of component products and other areas and to

achieve intelligent English machine translation. The results of English machine translation in areas such as component products were studied using different structures of neural networks, and intelligent English machine translation was achieved. However, the abovementioned English machine translation models based on neural network structures all suffer from the problem of unsatisfactory translation results due to the loss of long-distance information in the process of transmission due to long-distance dependence and therefore need to be improved [12].

To address the problems in existing marking systems, an intelligent computerized marking system for English translation based on natural language processing is designed. Through simulation experiments, the system is compared with the current scoring system and manual scoring method, and it is verified that the designed scoring system has high operational stability and accuracy, and the overall performance is better than the current scoring system.

2. Design of a Computerized Intelligent Scoring System for English Translation

2.1. Hierarchical Construction of English Translation Scoring System. The hierarchical relationship of each module is shown in Figure 1.

At the initial stage of the system, students' English translations are entered through the translation data collection module and processed by the collection module to produce a standardised format of the database file [13].

2.2. English Translations Scoring System. The overall framework of the natural language processing-based English translation system is shown in Figure 2. The user uploads a translation through the user side and, after the computer's natural language intelligence processing and information interaction, inputs it into the system's English translation scoring model.

2.3. Models in This Paper. LSTM is a special recurrent neural network model that solves the long sequence dependence problem in recurrent neural networks by adding memory units, input gates, output gates, and forgetting gates and improves the ability of recurrent neural networks to process long sequence data [14].

The transformer model also consists of an encoder group and a decoder group. An encoder or decoder group consists of multiple encoder modules or decoder modules stacked on top of each other. Each module consists of a multi-head attention and a fully connected feed-forward layer. Since the RNN is abandoned, another method is needed to remember the location information of the input sequence. A positional embedding is used in the transformer model to add a relative position to each element of the input sequence, and this position information is then used as a representation of each word [15, 16].

According to the above analysis of the LSTM network model, the output vector in the coding stage of the LSTM network model has a fixed dimension, so it uses the same

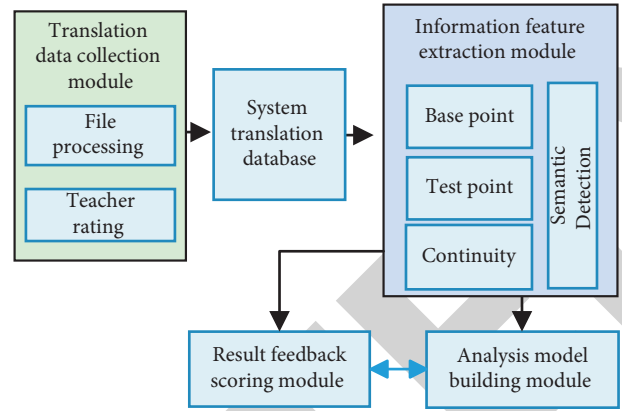


FIGURE 1: Rating system hierarchy.

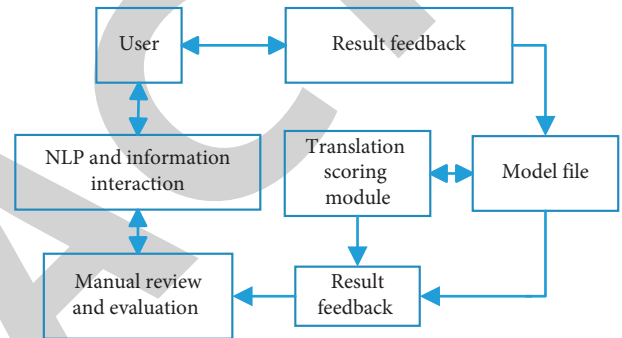


FIGURE 2: Overall architecture of the translation scoring system.

dimensional vector for any length of the source language sequence to encode. In actual English machine translation, the input English sequences are of variable length, which makes it easy to use the standard LSTM model for English machine translation, and the model does not fit the English input sequences perfectly, thus making the translation effect unsatisfactory. Moreover, due to the different focus of translation, the use of a fixed dimensional representation of the input model sequence, i.e., the same level of attention to the sequence, is obviously not conducive to improving the quality of the translation. Therefore, in order to solve the above problems, an attention mechanism is embedded in the LSTM network [17], and an English machine translation model based on LSTM attention embedding is proposed.

First, a set of multiple vectors is used instead of a fixed dimension for representing the source language sequence. Then, by dynamically selecting the background vectors during the target sequence generation process, the translation model is improved to pay more attention to the parts with high relevance to the source language during the translation process, which in turn improves the translation performance of the model [18]. The LSTM English machine translation model embedded with attention mechanism consists of three parts: encoder, decoder, and attention mechanism, as shown in Figure 3.

The next hidden state at the target side of the model is calculated in the same way as the LSTM decoder part, as in the following equation:

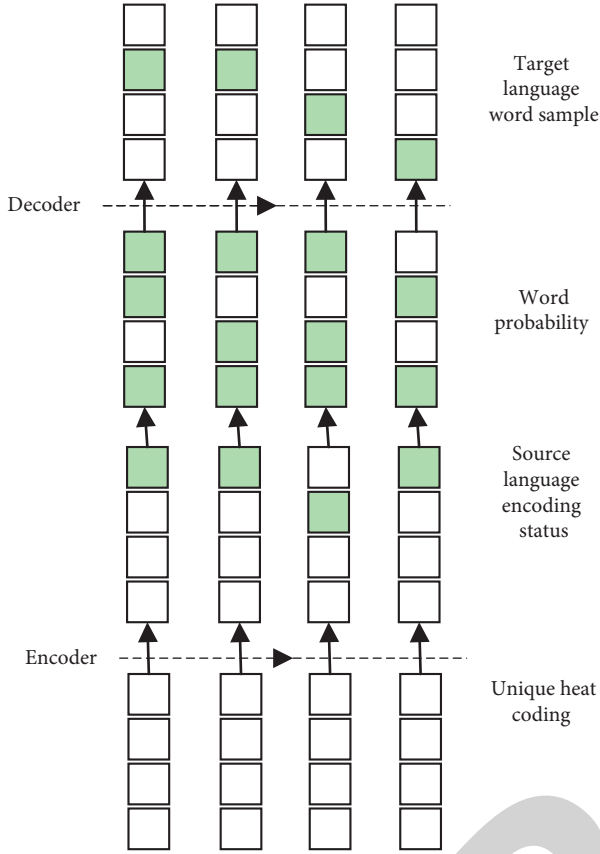


FIGURE 3: LSTM translation model with embedded attention mechanism.

$$Z_i + 1 = \sigma(c_i, u_i, z_i), \quad (1)$$

where u_i denotes the i -th word in the target language sequence and c_i denotes the background vector of word i . Since the background vectors of the LSTM model with the embedded attention mechanism are a set of multiple vectors, rather than being uniformly fixed [19], each word in the target language sequence can find a unique background vector corresponding to it.

Let the state of the implicit layer at encoder j be h_j , then its corresponding background vector can be calculated by

$$C_i = \sum_{j=1}^r a_{ij} h_j, \quad (2)$$

where a_{ij} represents the weight, i.e., the attention value of the i -th word in the target language sequence to the j -th word in the source language sequence, which can be calculated by the following equations:

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^T \exp(e_{ik})}, \quad (3)$$

$$e_{ij} = a(z_i, h_j), \quad (4)$$

where a is a function that measures the match between the current hidden state z_i of the target language sequence and

the hidden state h_j of the source language sequence and can be calculated by

$$e_{ij} = v^T \tanh(W_{zzi} + W_h h_j), \quad (5)$$

where v , W_z , and W_h denote the model parameters to be learned.

By embedding an attention mechanism in the LSTM network, the model can be weighted with different weights on the source language side, which solves the long-range dependency problem of standard LSTM models and thus improves the model performance.

3. Implementation of an English Translation Scoring System

3.1. Language Models for English Translation Scoring Systems. Statistical language models can give the probability distribution of a particular sentence sequence or word sequence in a translation [20–22]. To simplify the computation and reduce the complexity, a ternary model is introduced. Let the preferred set embedded in the ternary language model be V and the ternary combination be (u, v, w) , corresponding to a parameter $q(w | u, v)$ with full $w \in v \cup \{\text{STOP}\}$ and $u, v \in v \cup \{*\}$. $q(w | u, v)$ represents the probability that a single word w follows a word u and v when the binary combination is known. The probability distribution of the ternary language model for a given translated sentence $x_1 x_2 \dots x_n$ is given by

$$p(x_1 x_2 \dots x_n) = \prod_{i=1}^n q(x_i | x_{i-2} x_{i-1}). \quad (6)$$

The restrictions that need to be met are

$$q(w | u, v) \geq 0 \quad \sum_{w \in v \cup \{\text{STOP}\}} q(w | u, v) = 1. \quad (7)$$

The maximum likelihood estimation algorithm is used to solve for $q(w | u, v)$, which corresponds to the following equation:

$$q(w | u, v) = \frac{c(u, v, w)}{c(u, v)}, \quad (8)$$

where $c(u, v, w)$ represents the frequency of occurrence of (u, v, w) in the translation training set and $c(u, v)$ is the frequency of occurrence of (u, v) in the translation training set [23].

To address the problem that not all ternary combinations that do not appear in the translation training set have a probability of zero, a smoothing algorithm is introduced to obtain the descriptive formula for the language model as

$$q(w | u, v) = \lambda_1 * q(w | u, v) + \lambda_2 * q(w | v) + \lambda_3 * q(w), \quad (9)$$

where λ_1, λ_2 , and λ_3 represent the smoothing factors and satisfy $\lambda_1, \lambda_2, \lambda_3 \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1$; $q(w | v)$ represents the probability of word w occurring after the word v when word v is known; and $q(w)$ represents the total probability of word w occurring.

3.2. Similarity Calculation and Scoring of English Translations.

In order to calculate the similarity between the user's translation result and the standard answer, the similarity of keywords is introduced and the word similarity is calculated by the following formula [24]:

$$\text{sim Word}(A, B) = \frac{\text{Same}(A, B)}{\text{Num}(A) + \text{Num}(B)}, \quad (10)$$

where $\text{sim Word}(A, B)$ is the word similarity between sentences A and B , $\text{Same}(A, B)$ represents the number of identical words in sentences A and B , and $\text{Num}(A)$ and $\text{Num}(B)$ represent the number of words in sentences A and B , respectively.

The characteristic keyword similarity is calculated, the particle swarm optimized BP network is used to fit the calculation, and the calculation result is compared with the set scoring standard.

4. Experimental Results and Analysis

4.1. Experimental Environment and Parameter Settings. In order to verify the effectiveness of the proposed LSTM attentional embedding-based English machine translation model, the study built an LSTM English machine translation system on the TensorFlow framework [25]. The parameters of the LSTM neural network are set as follows: the vocabulary size is 30 000, the word vector dimension and the number of nodes in the implicit layer are 512, the number of LSTM network layers is 2, the column search width is 3, the learning rate is 0.1, the dropout is 0.5, and the batch size is 128. The decoding stage is based on the column search algorithm.

4.2. Dataset Sources and Preprocessing. The study chose the International Spoken Language and its Translation Review Contest (IWSLT) 2019 data, which has a small data size, as the dataset for this experiment, including 220,000 Chinese-English parallel utterance pairs, pairs of test set data, and pairs of development data [26]. Since the LSTM attentional embedding-based English machine translation model could not be trained and learned directly on the IWSLT 2019 dataset, word vector transformation of the dataset was also required [27]. The study performed a word separation process on the data and then used CBOW to factorize the separated data.

4.2.1. Split Word Processing. As (IWSLT) 2019 dataset contains Chinese and English parallel utterance pairs, the Chinese and English word separation methods are different; therefore, the study carried out word separation for the Chinese and English of the experimental dataset separately [28]. For Chinese word separation, a statistical-based word separation method was used. Firstly, a word is regarded as a combination of several fixed words according to the composition form of Chinese words; then, the probability of word generation is judged according to the frequency of co-occurrence between words in the context of an utterance, i.e., the credibility of the word; finally, a threshold is set

according to the credibility of the word to form the word composition condition and determine the word separation. In the case of English, since the basic unit of English is the word, it is only necessary to split the word directly according to the space. However, since English sentences contain stop words, they also need to be deactivated during the word separation process [29]. The English deactivation process consists of three main steps: firstly, capitalisation of the English language, then space splitting of the words and symbols at the end of the sentence, and finally, generalisation of the word sentence using the special noun special bond method.

4.2.2. Word Vector Representation. Victorian representation of words means digitising linguistic symbols so that language numbers can be fed into a model for training and learning [30]. The study uses CBOW for the factorized representation of words. Suppose the size of the dictionary is v , and an index set 111 of one-to-one correspondence between the word and the integers in the dictionary is established. If there exists a test sequence with length T , time window size m , and word J at time t , the probability of CBOW maximizing the background work to generate a central word is given by

$$J = \prod_{t=1}^T P(w^{t-m}, \dots, w^{t-1}, w^{t+1}, \dots, w^{t+m}). \quad (11)$$

Taking the negative logarithm of the above equation gives the loss function, i.e., the maximum likelihood estimate of equation (11) can be calculated by minimizing the following equation:

$$J = \sum_{t=1}^r \log P(w^t | w^{t-m}, \dots, w^{t+1}, \dots, w^{t+m}). \quad (12)$$

Assuming that the background word vector is denoted as v and the central word vector is denoted as u , then by CBOW training, for each word indexed as i in the lexicon, the vector of that word as a background word is obtained (v_i) and the vector as a central word can be denoted u_i .

4.3. Evaluation Indicators. BLEU value is selected as the index to evaluate the translation quality of the translation model. The larger its value is, the higher the translation quality is. The calculation method of the BLEU value is shown as follows:

$$\text{BLEU} = \text{BP} \times \exp\left(\sum_{n=1}^N w_n \log_{10} P_n\right), \quad (13)$$

where BP is the penalty factor; N is the longest tuple length, usually 4; n is the number of tuples; w_n is the tuple n weight; and P_n is the tuple n ratio [12].

4.4. Model Validation. In order to verify the performance of the proposed LSTM attention-embedded translation model, the experimental dataset was trained with the standard LSTM model and the attention-embedded LSTM model,

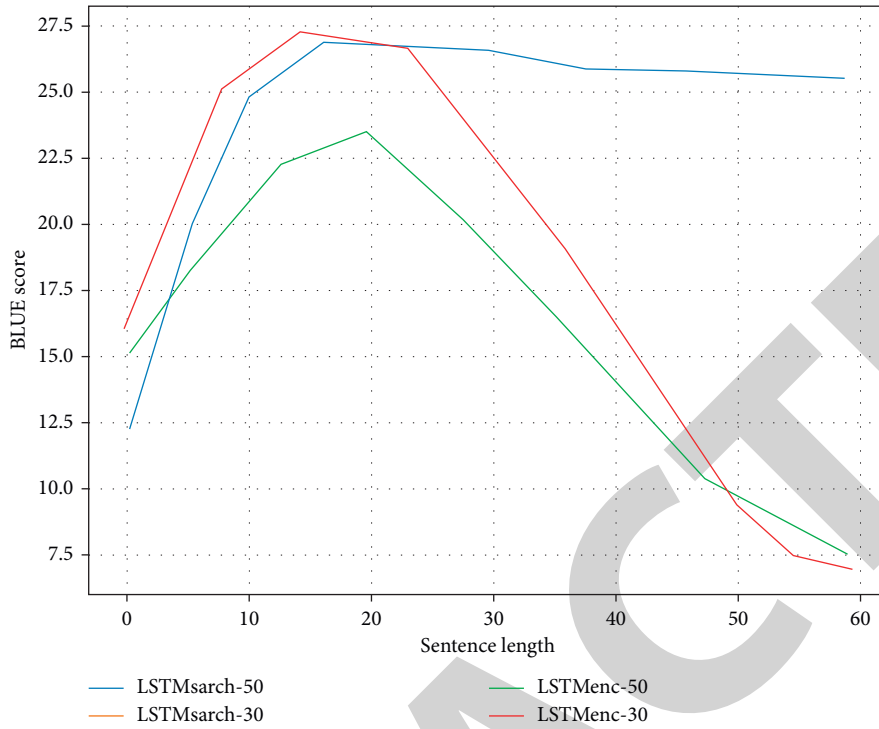


FIGURE 4: Performance of the attention mechanism (RNN to LSTM).

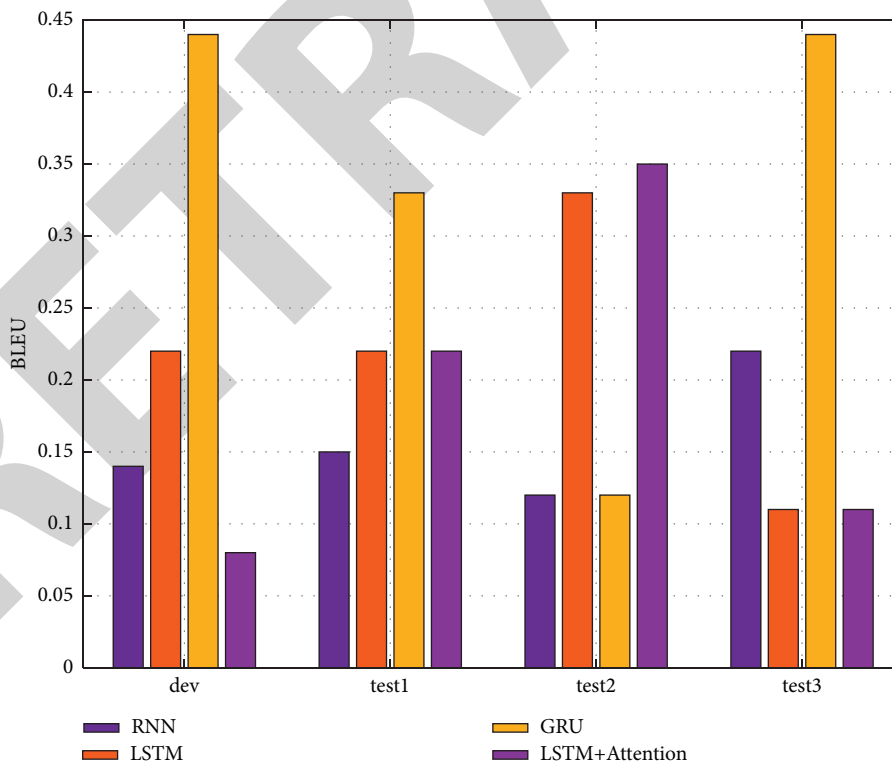


FIGURE 5: BLEU values of different translation models on the experimental dataset.

and the results are shown in Figure 4. The numbers in Figure 4 represent the different network layers in the network model we designed. Compared with the standard

LSTM model, the attention-embedded LSTM model has a higher BLEU value, indicating that the attention-embedded

TABLE 1: Comparison of scoring values by scoring method.

DE	RM	SC/C
1	RA	87.4
	RB	84.7
	RC	87.6
2	RA	74.6
	RB	76.9
	RC	74.7
3	RA	72.6
	RB	69.8
	RC	72.9

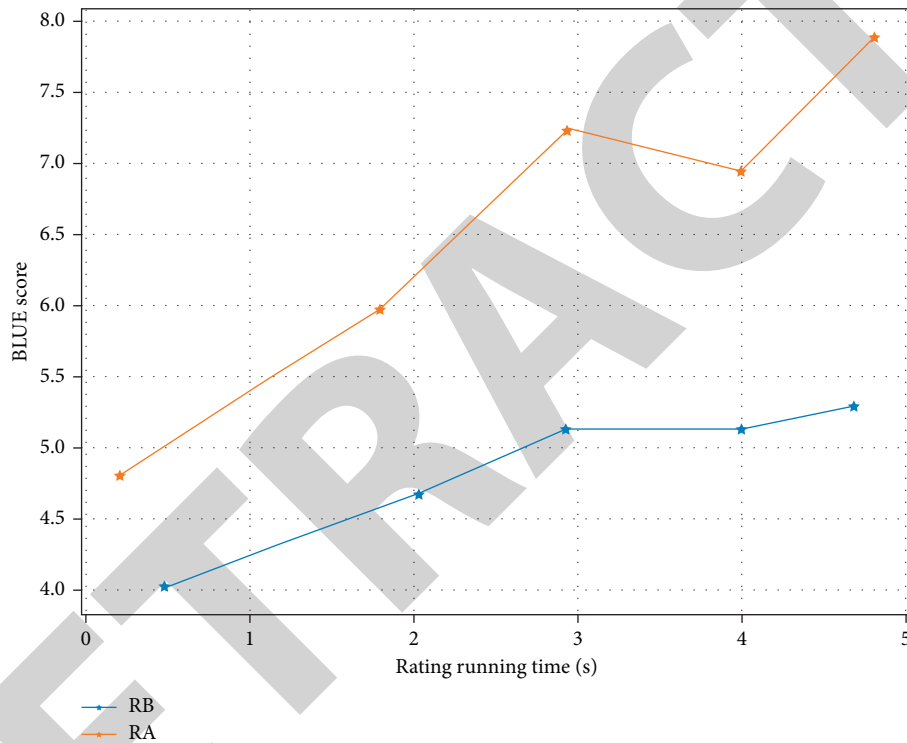


FIGURE 6: Comparison of scoring system run times.

LSTM model is more effective in translating long sentences and the model performance is effectively improved.

4.5. Model Comparison. In order to verify the translation effectiveness of the proposed English machine translation model of LSTM attention, the experimental analysis results are shown in Figure 5. As can be seen from the figure, the BLEU values of the proposed English machine translation model based on LSTM attention embedding on both the development machine and the test set are higher than those of the traditional LSTM, RNN, and GRU-Attention English machine translation models, indicating that, compared to the comparison translation models, the proposed translation model improves translation by our network. The proposed translation model improves the performance and translation quality by our network.

4.6. Scoring Effect. In Table 1, DE denotes the English translation document to be scored; RM denotes the scoring method; RA, RB, and RC denote the designed system, the existing scoring system, and the manual scoring method, respectively; SC denotes the score in points and is denoted by the letter C.

According to the data in Table 1, the scoring results of the designed system are closer to the manual scoring results, with a minimum difference of 0.1 C and a maximum difference of 0.3 C. This indicates that the scoring error of the designed English translation scoring system is smaller and the scoring performance is better. Experiments were conducted using the designed system and the existing scoring system to compare the running time of the scoring process, and the experimental results are shown in Figure 6. In Figure 6, RA and RB denote the runtime of the designed system and the existing scoring system, respectively.

According to Figure 6, the range of fluctuation of the scoring runtime curve of the designed system is smaller than that of the runtime curve of the existing scoring system, which indicates that the designed system is more stable in operation. For the translation sample, the scoring time of the designed system was 4.7 s, while that of the existing scoring system was 6.1 s. For the translation sample, the scoring times of the designed system and the existing scoring system were 4.9 s and 5.9 s, respectively, which indicates that the scoring time of the designed system was significantly lower than that of the existing scoring system for the same translation sample, indicating that the scoring efficiency of the designed system was higher.

5. Conclusions

The proposed English machine translation model based on LSTM attention embedding is innovative in that it enhances the representation of source language contextual information by introducing an attention mechanism into the standard LSTM English translation model, thereby improving the performance of the English machine translation model and the quality of the translated text. The result is better than the standard LSTM model and the traditional RNN and GRU-Attention English machine translation models and can be used in real English machine translation. The experimental results show that the overall performance of the designed system is better than that of the traditional system, indicating its strong practicality.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] C. Su, H. Huang, S. Shi, P. Jian, and X. Shi, "Neural machine translation with Gumbel tree-LSTM based encoder," *Journal of Visual Communication and Image Representation*, vol. 71, Article ID 102811, 2020.
- [2] Y. Ma, J. Yu, B. Ji, J. Chen, S. Zhao, and J. Chen, "Three-way decisions based RNN models for sentiment classification," in *Proceedings of the International Joint Conference on Rough Sets*, pp. 247–258, Springer, Bratislava, Slovak Republic, September 2021.
- [3] N. A. Mohamed, M. A. Zulkifley, A. A. Ibrahim, and M. Aouache, "Optimal training configurations of a CNN-LSTM-based tracker for a fall frame detection system," *Sensors*, vol. 21, no. 19, p. 6485, 2021.
- [4] J. Wang, J. Zhu, and Z. Yu, "Person re-identification based on attention clustering and long short-term memory network," *Journal of Electronic Imaging*, vol. 30, no. 3, Article ID 033014, 2021.
- [5] C. H. Cao, Y. N. Tang, D. Y. Huang, G. WeiMin, and Z. Chunjong, "IIBE: an improved identity-based encryption algorithm for wsn security," *Security and Communication Networks*, vol. 2021, Article ID 8527068, 8 pages, 2021.
- [6] A. K. Mandal, R. Sen, S. Goswami, and B. Chakraborty, "Comparative study of univariate and multivariate long short-term memory for very short-term forecasting of global horizontal irradiance," *Symmetry*, vol. 13, no. 8, p. 1544, 2021.
- [7] C. Shi, S. Liu, S. Ren et al., "Knowledge-based semantic embedding for machine translation," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 2245–2254, Berlin, Germany, August 2016.
- [8] M. A. Hasan, F. Alam, S. A. Chowdhury, and N. Khan, "Neural machine translation for the Bangla-English language pair," in *Proceedings of the 2019 22nd International Conference on Computer and Information Technology (ICCIT)*, pp. 1–6, IEEE, Dhaka, Bangladesh, December 2019.
- [9] V. Goyal and D. M. Sharma, "LTRC-MT simple & effective Hindi-English neural machine translation systems at WAT 2019," in *Proceedings of the 6th Workshop on Asian Translation*, pp. 137–140, Hong Kong, China, November 2019.
- [10] Y. Xia, T. He, X. Tan, F. Tian, D. He, and T. Qin, "Tied transformers: neural machine translation with shared encoder and decoder," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 5466–5473, 2019.
- [11] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 160–167, IEEE, Wanzhou, China, November 2018.
- [12] M. Sato, J. Suzuki, and S. Kiyono, "Effective adversarial regularization for neural machine translation," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 204–210, Florence, Italy, July 2019.
- [13] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2016.
- [14] P. An, Z. Wang, and C. Zhang, "Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection," *Information Processing & Management*, vol. 59, no. 2, Article ID 102844, 2022.
- [15] K. Chen, R. Wang, M. Utiyama, E. Sumita, and T. Zhao, "Syntax-directed attention for neural machine translation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, New Orleans, LA, USA, April 2018.
- [16] C. Ma, A. Tamura, M. Utiyama, T. Zhao, and E. Sumita, "Forest-based neural machine translation," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 1253–1263, Melbourne, Australia, July 2018.
- [17] Q. Xia, Z. Li, M. Zhang et al., "Syntax-Aware neural semantic role 1," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 7305–7313, 2019.
- [18] J. Yang, R. Yang, H. Lu, C. Wang, and J. Xie, "Multi-entity aspect-based sentiment analysis with context, entity, aspect memory and dependency information," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 18, no. 4, pp. 1–22, 2019.
- [19] M. Nguyen, G. H. Ngo, and N. F. Chen, "Hierarchical character embeddings: learning phonological and semantic representations in languages of logographic origin using recursive neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 461–473, 2019.
- [20] 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.
- [21] X. Wei, Y. Hu, L. Xing, Y. Wang, and L. Gao, “Translating with bilingual topic knowledge for neural machine translation,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 7257–7264, 2019.
- [22] M. Zhang, Z. Li, G. Fu, and M. Zhang, “Dependency-based syntax-aware word representations,” *Artificial Intelligence*, vol. 292, Article ID 103427, 2021.
- [23] W. Jiang, Z. Li, and M. Zhang, “Syntax-enhanced ucca semantic parsing,” *Beijing Da Xue Xue Bao*, vol. 56, no. 1, pp. 89–96, 2020.
- [24] H. Li, D. Zeng, L. Chen, Q. Chen, M. Wang, and C. Zhang, “Immune multipath reliable transmission with fault tolerance in wireless sensor networks,” *Bio-inspired Computing - Theories and Applications*, Springer, in *Proceedings of the International Conference on Bio-Inspired Computing: Theories and Applications*, pp. 513–517, October 2016.
- [25] H. Zhang, A. Ng, and R. Sproat, “Fast and accurate reordering with ITG transition RNN,” in *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 1454–1463, Santa Fe, NM, USA, August 2018.
- [26] Y. Li, J. Li, and M. Zhang, “Improving neural machine translation with latent features feedback,” *Neurocomputing*, vol. 463, pp. 368–378, 2021.
- [27] C. Ma, A. Tamura, M. Utiyama, E. Sumita, and T. Zhao, “Syntax-based transformer for neural machine translation,” *Journal of Natural Language Processing*, vol. 27, no. 2, pp. 445–466, 2020.
- [28] H. Fadaei and H. Faili, “Using syntax for improving phrase-based SMT in low-resource languages,” *Digital Scholarship in the Humanities*, vol. 35, no. 3, pp. 507–528, 2020.
- [29] L. Shi, W. Rong, S. Zhou, N. Jiang, and Z. Xiong, “A dual channel class hierarchy based recurrent language modeling,” *Neurocomputing*, vol. 418, pp. 291–299, 2020.
- [30] 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.