Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

A Study on Improving Semantic Consistency of Translation Systems by Combining Dynamic Computing Methods in English Corpus

Li Zhang[1,]✉

[1] *School of Humanities and Design, Henan Open University, Zhengzhou, Henan, 450046, China*

ABSTRACT

Aiming at the dilemma of corpus-based intelligent English translation, the article proposes an English neural machine translation method based on depth-separable convolution, which combines with the dynamic computation method to improve the semantic consistency of the translation system for semantic alignment and fusion. In order to verify the training effect of the proposed convolutional neural network model combined with the dynamic computation method, comparison experiments with one-way and two-way network models and baseline model with different cut-off granularity are conducted respectively. In order to better examine its performance in practical translation applications, online translation, machine translation and systematic methods are utilized for comparison. The BLUE values of this paper's model for Chinese-English data translation in four different granularities of words, syllables, subwords and characters are 21.41%, 21.91, 29.25% and 20.40%, respectively. In 100,000, 200,000 and 500,000 training English-Chinese bilingual parallel corpus, the training time consumed by the model in this paper is 9.58 h, 15.94 h and 32.69 h. In practical application, the decibel range of the noise reduction of the translation system method designed by the research is distributed in [1.62 ~ 1.89], the average value of coherence is 91.1%, and the average compression rate and the average stability of the BLEU scores are 93.84% and 98.38%, respectively, and the results are better than the comparison methods.

Keywords: intelligent English translation, dynamic computing, semantic coherence, convolutional neural network modeling

1. Introduction

Language is an important tool for human communication, and translation is a bridge for communication between different languages. With the development of globalization and the increase

✉ Corresponding author.
E-mail address: zhangli2@haou.edu.cn (L. Zhang).
Received 05 November 2024; Revised 20 December 2024; Accepted 15 March 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-196

of communication among countries, the translation industry becomes more and more important. In the process of translation, a professional corpus and translation system can help translators to improve the quality and efficiency of translation [1-4]. English Translation Corpus is a large-scale English translation corpus for translation and linguistic research,which contains a large number of English texts,including various types of texts,such as news,novels,scientific and technical literature,business documents and so on,which can provide more accurate and comprehensive information about vocabulary,phrases,grammar and sentence patterns for translators[5-8]. By consulting the corpus, translators can find suitable translation examples for better understanding and application. Besides corpus, translation system is also one of the commonly used tools in translation industry [9-11].

Due to the differences between different languages, semantic expressions in translation often face some challenges, and it is crucial to express semantics accurately in the translation process. Translation systems are able to store previously translated sentences and passages, and subsequently match and apply them in new translation tasks so as to realize the accurate expression of semantics [12-15]. Since semantic expression in English translation is a complex and important issue, during the translation process, translators need to face the challenges of vocabulary, phrases, sentence structure, grammar and culture [16-18]. In order to accurately express the semantics, translators need to have good linguistic competence, cross-cultural communication skills, and adopt appropriate translation strategies, while the translation system can realize the accurate communication between the source language and the target language to achieve the ultimate goal of translation [19-22].

The article takes the transformer neural machine translation model based on the attention mechanism as a prototype, maps the source language sentences to the target language sentences through the intermediate vectors in the semantic space, and combines the word vectors and positional encoding to represent the positional information of the words in the text sequence more clearly. On this basis, a low-resource neural machine translation method based on translation memory information enhancement is proposed to generate high-quality pseudo-prototype sequences in a multi-strategy way by combining the traditional retrieval method and the proposed pseudo-prototype sequence generation method. In addition to the improvement of the traditional encoder-decoder framework to accommodate the prototype input, the proposed method also uses an improved loss function to weaken the effect of low-quality prototype sequences on the model. Finally, it is combined with a dynamic computational approach to enable the model to fully utilize the feature information of each layer of the neural network and train to obtain a better translation model.

2. Design of Transformer Neural Translation Model Based on Combined Dynamic Computing

2.1. The Dilemma of Corpus-Based Intelligent English Translation

2.1.1. Intelligent English Translation Research Technical Attributes to be Improved. At present, cross-cultural communication is getting deeper and deeper, and the influence of culture on English translation is becoming more and more prominent. This highlights the reality that China's foreign language ability needs to be improved. In particular, there is a lack of Chinese-English aligned corpora, which directly affects the construction of China's artificial intelligence corpus for international communication. Although many domestic universities and research institutes have built Chinese learners' English corpus, some English translators do not have enough knowledge about corpus construction and in-depth application of corpus, and their knowledge structure is relatively biased. They still focus on English translation knowledge and ignore the technical attributes of intelligent English translation under corpus integration. This directly leads to the lack of theoretical research on the integration of corpus and intelligent English translation, and the intelligent English translators do not have any knowledge of the technology of intelligent English translation in the pre-translation, mid-

translation and translation process. Intelligent English translators do not use corpora before, during and after translation. Intelligent English translators do not use corpora in a targeted way before, during and after translation.

2.1.2. The level of professional education intelligence needs to be improved. Intelligent English translation research based on corpus needs the help of professional education, including the use of corpus for education and teaching by teachers of English translation majors, or scholars in the industry to study the logic of corpus and English translation education, as well as the fit between intelligent English translation. However, at this stage, some teachers of English translation majors do not have a good grasp of the underlying logic and effective practice paths of the integration of intelligence and English translation education and corpus, resulting in the corpus not being widely popularized in the teaching of English translation majors, and electronic dictionaries and paper dictionaries are still the main tools for translation majors in their English translation practice, and the overall level of intelligence in English translation education is yet to be upgraded. Although many translation learners have a large English vocabulary reserve, they are not able to choose words accurately in English translation, which is due to the fact that the corpus has not been able to fulfill the function of providing cultural context and background in professional education.

2.2. Neural Machine Translation

2.2.1. Neural Machine Translation Fundamentals. Machine translation is an important research task in NLP, which investigates how computers can be used to translate source language sentences into target language sentences. With the development of deep learning technology, the emergence of neural networks has driven the technological development of various downstream tasks in the field of NLP. Neural machine translation based on deep learning gradually emerged and gradually surpassed the then mainstream statistical machine translation. Neural machine translation has advantages such as simple structure and the ability to capture long dependencies in sentences, so it gradually dominates the field of machine translation and becomes the next-generation machine translation paradigm.

The most original and classical structure in the field of neural machine translation is the encoder-decoder framework. The encoder-decoder architecture is a sequence-to-sequence model. An encoder reads a source language sentence and compresses the text sequence into a vector of fixed dimensions at each hidden state, a process known as encoding. The decoder reads this vector and generates sequences of target language words sequentially, and the output process of the decoder can be seen as the reverse work of the encoder, the encoder-decoder framework is shown in Fig. 1.The principle of the encoder-decoder architecture of the NMT can be seen as the mapping of the source language sentences to the target language sentences through intermediate vectors in the semantic space.

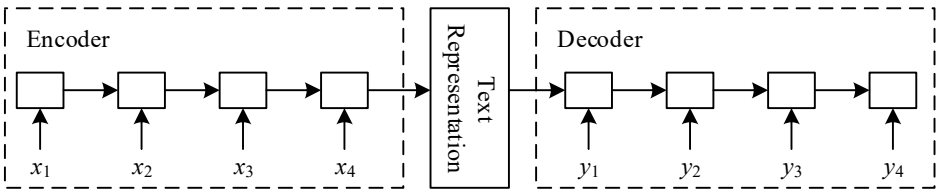

Fig. 1. The Encoder-Decoder architecture

Given a parallel corpus C consisting of a set of parallel sentence pairs (x, y), NMT usually uses maximum likelihood estimation as the objective function for model training with the following functional expression:

$$\mathrm{L}_0 = \sum_{(x,y)\in C} -\log p(y \mid x;\theta) \tag{1}$$

where x is a source language sentence, y is a target language sentence, and θ is a set of parameters to be learned. Given a source language sentence, the probability of occurrence of the target language sentence is:

$$p(y \mid x;\theta) = \prod_{j=1}^{m} p\left(y_j \mid y_{<j}, x;\theta\right) \tag{2}$$

where m is the number of words in y, y_j is the word generated by the NMT model at the current time step, and $y_{<j}$ is the word generated by the NMT model at the previous time step.

2.2.2. Transformer model. The most important part of the structure of the Transformer model is the multi-head self-attention module, which is also the key to improving the quality of the model's translation for long sentences. The advantage of the multi-head self-attention mechanism is its ability to capture multiple dependencies. In addition, because Transformer is able to parallelize computation on the GPU, model training is much faster and the quality of translation is greatly improved. It is due to the multifaceted advantages of the Transformer model that it is unstoppable to become the mainstream model in the field of NMT, and at the same time it is blossoming in other downstream tasks in NLP.The Transformer model is shown in Fig. 2.The Transformer model adopts an encoder-decoder structure, where the input of the encoder is the source-language sentence and the input of the decoder is the target language sentence.

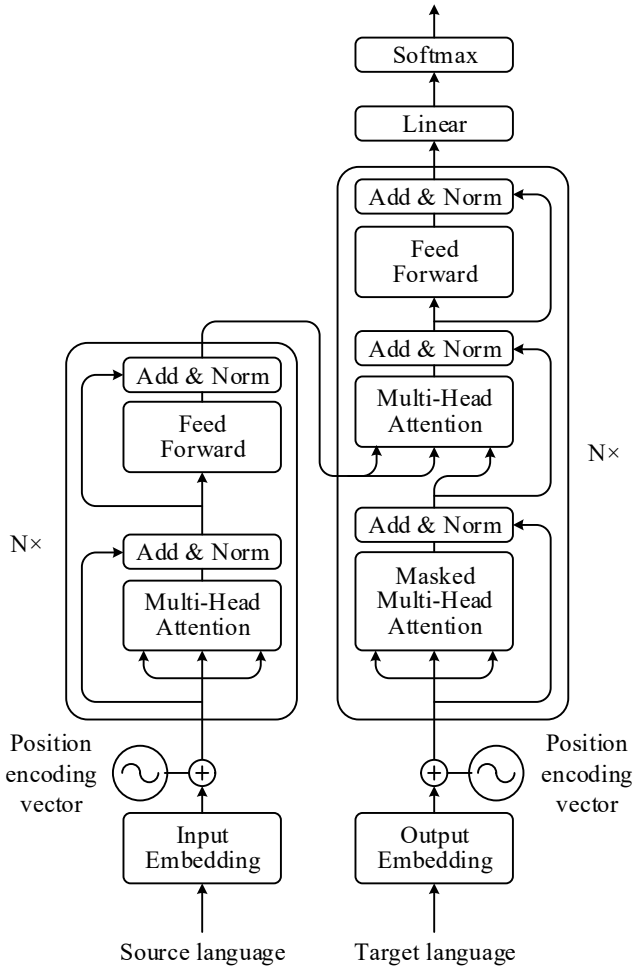

Fig. 2. The Transformer model

The Transformer model combines word vectors and positional coding to represent the positional information of words in a text sequence more clearly.The RNN model indirectly represents the positional information between words through the time series, and the Transformer model avoids the model's invariance to the order of the text sequence through positional coding. Therefore, the complete input to the encoder of the Transformer model is obtained by summing the position encoded vectors and the word vectors. In Transformer, the final choice is to add positional encoding using the $\sin$ -function and the $\cos$ -function, and each position can be encoded using Equation (3):

$$\begin{cases} PE(pos,2i)=\sin\left(pos/1000^{2i/d_{\mathrm{mod}\,el}}\right) \\ PE(pos,2i+1)=\cos\left(pos/1000^{2i/d_{\mathrm{mod}\,el}}\right) \end{cases} \tag{3}$$

where pos represents the position and i represents the dimension, and each dimension of the position encoding corresponds to a sinusoidal wave function.

The self-attention mechanism of the Transformer model uses dot product attention to compute the correlation coefficients. In the self-attention mechanism, the query matrix Q, the key matrix K and the value matrix V are all the same for which the attention is computed as shown below:

$$Attention(Q,K,V)=Soft\max\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4}$$

where d_k is the dimension of the key matrix K. Since the computation of attention is done by matrix multiplication, and the characteristic of GPU is to be able to perform such mathematical computations as matrix multiplication quickly.

The multi-head attention mechanism of the Transformer model is to divide the original query matrix Q, key matrix K, and value matrix V into h parts by linear transformation, and then perform attention computation for each of the Q, K, and V parts individually, and finally splice $head_i$. The multi-head attention mechanism can be represented as:

$$MultiHead(Q,K,V)=Concat(head_1,\cdots,head_h)W^O \tag{5}$$

where the i st attention head is computed $head_i=Attention\left(QW_i^Q,KW_i^K,VW_i^V\right)$. The multi-head attention mechanism can be viewed as multiple networks running in parallel using different views on the set of key-values and mapping them to different subspaces of the output representation, enriching the representation of the semantic information.

In addition to this, the Transformer model uses a feed-forward neural network with the aim of mapping the representation after the multi-head attention mechanism into a new space. The specific computation of the feed-forward neural network is as follows:

$$FFN(x)=\max\left(0,xW_1+b_1\right)W_2+b_2 \tag{6}$$

where W_1, W_2, b_1 and b_2 are the parameters of the Transformer model.

Residual network and layer normalization are also important parts of the Transformer model, residual network is introduced to solve the problem of gradient explosion or gradient vanishing and layer normalization is introduced to avoid training instability. The Multihead Self-Attention Model with Mask is used in the decoder of the Transformer model to mask future information that has not yet been predicted by the NMT model, this is to avoid the model from observing future information during the training process. The Multihead Self-Attention Module with Mask ensures that the prediction result at the current position can only depend on the known outputs prior to that position.

2.3. Neural machine translation model based on translation memory information enhancement

Given the low-resource scenario, the insufficient number of translation memories leads to the ineffectiveness of transformer translation memory-based methods. In this chapter, we propose a low-resource neural machine translation method based on translation memory information enhancement by combining the traditional transformer retrieval method with the proposed pseudo-prototype sequence generation method in a multi-strategy approach to generate high-quality pseudo-prototype sequences in combination with the dynamic computation method. To ensure the usability of the prototype sequence, we utilize the prototype generation method based on the combination of multiple strategies for translation memory information enhancement.

2.3.1. Hybrid prototype matching combined with distributed representation matching. A translation memory is a collection $\{(s_l, t_l) : l = 1, \ldots, L\}$ of L pair of parallel sentences, where s_l is the source sentence and t_l is the target sentence. For a given input sentence x, we first use keyword matching to retrieve it from the translation memory. In view of the low number of long sentences in the low-resource scenario, N-gram matching is not used as the keyword matching method, but fuzzy matching is used for keyword matching, and fuzzy matching is defined as:

$$FM(x, s_i) = 1 - \frac{ED(xs_i)}{\max(|x|, |s_i|)} \tag{7}$$

where $ED(x, s_i)$ is the edit distance between x, s_i and $|x|$ is the sentence length of x.

Unlike keyword-based matching methods, distributed representation matching retrieves sentences based on the distances between their vector representations, which is somehow a means of similarity retrieval using semantic information, and thus provides a different retrieval perspective than keyword matching. Distributed representation matching based on cosine similarity is defined as:

$$EM(x, s_i) = \frac{h_x \cdot h_{s_i}}{\|h_x\| x \| s_{s_i}\|} \tag{8}$$

where h_x and h_{s_i} are the vector representations of x and s_i, respectively, and $\|h_x\|$ is the metric of vector h_x. In order to achieve fast computation, in this chapter, the multilingual pre-training model mBERT is first used to obtain the vectorial representations of sentences x and s_i, and then based on the representations, the similarity matching is performed using the faiss tool.

To combine the advantages of utilizing the two methods, when fuzzy matching is able to obtain the optimal matching source sentence s_{best}, a set of top-k matching results $s' = \{s_1, s_2, \ldots, s_k\}$ is obtained by using distributed representation matching, e.g., $s_{best} \in s'$, and the target-end sentence t_{best} corresponding to s_{best} is selected as the prototype sequence. When fuzzy matching fails to retrieve a matching source sentence or $s_{best} \notin s'$, the optimal matching source sentence s_{best} is retrieved by distributed representation matching.

2.3.2. Translation model based on prototype sequence incorporation. By designing a hierarchical combinatorial approach, prototype matching and prototype generation methods are used to enhance the quantity and quality of retrieved prototype sequences. However, in terms of prototype utilization, the traditional transformer prototyping approach is not tailored to low-resource environments in the neural network architecture. To address this problem, we improve the codec architecture and present it in this section. First, we review the standard neural machine translation model based on the traditional attention mechanism: given a source language input sentence $x = \{x_1, x_2, \ldots, x_m\}$, the traditional model is defined as:

$$P(y \mid x;\theta) = \prod_{i=1}^{N} P\left(y_i \mid y_{<i}; x; \theta\right) \tag{9}$$

where θ is the model parameters and $y_{<i}$ is the already obtained translation. After incorporating the prototype sequence $t = \{t_1, t_2, \ldots, t_n\}$, the new translation model is expressed as:

$$P(y \mid x;t;\theta) = \prod_{i=1}^{N} P\left(y_i \mid y_{<i}; x; t; \theta\right) \tag{10}$$

In order to better incorporate the prototype sequence, the overall structure of the model proposed in this chapter is designed as follows: the encoding side adopts a dual encoder structure, which can receive both sentence input and prototype sequence input, and then encode the input into the corresponding hidden state representation; the decoding side incorporates a gating mechanism, which utilizes the self-learning ability of neural network to achieve the optimization of the ratio between sentence information and prototype information, and to control the flow of information in the decoding process. In order to reduce the impact of too low quality prototype sequences on the decoding process, the loss function is optimized. The prototype sequence $t = \{t_1, t_2, \ldots, t_n\}$ is obtained based on a combination of strategies in the prototype generation method.

3. Experimental analysis of improved translation models combined with dynamic computing

3.1. Comparative study of translation models

3.1.1. Comparison of model performance at different slicing granularities. In order to verify the performance of each translation model when different translation models adopt different English slicing granularity, the experiments will be conducted by using the traditional RNN, LSTM, GRU, CNN models as well as the pre-trained language models from the replacement models of bidirectional networks BIRNN, BILSTM, BIGRU, BICNN as the model coding layer, and and as well as the self-attention network based Transformer machine translation model and the improved translation model combining dynamic computation in this paper, respectively, to translate a Chinese-English-English-Chinese parallel corpus of an English sentence with four different granularities of words, syllables, subwords, and characters, and the resulting BLEU scores of the 10 models are shown in Table 1.

Table 1. BLEU scores of translation models in different segmentation granularity

Model	Translation mode	BLEU /%			
		English granularity			
		Word	Syllable	Subword	Character
RNN	Chinese-English	17.31	17.84	18.38	15.42
	English-Chinese	16.86	17.17	17.56	14.73
LSTM	Chinese-English	17.65	18.18	18.90	15.90
	English-Chinese	17.26	17.95	18.12	15.29
GRU	Chinese-English	18.58	20.39	21.60	15.29
	English-Chinese	18.17	19.70	20.95	14.51
CNN	Chinese-English	17.37	18.26	19.85	16.25
	English-Chinese	17.06	18.07	18.57	15.96
BIRNN	Chinese-English	19.86	19.49	20.89	17.48
	English-Chinese	18.97	18.01	19.60	16.68
BILSTM	Chinese-English	20.02	20.62	21.39	17.60
	English-Chinese	19.79	20.35	19.93	16.67
BIGRU	Chinese-English	19.53	19.95	20.35	17.55
	English-Chinese	18.40	19.39	19.85	17.17
BICNN	Chinese-English	19.17	19.36	20.63	17.28
	English-Chinese	18.25	17.93	19.15	16.07
Transformer	Chinese-English	19.09	20.76	21.65	18.35
	English-Chinese	18.56	19.83	20.35	17.16
Textual model	Chinese-English	21.41	21.91	23.25	20.40
	English-Chinese	20.66	20.84	21.88	19.43

As can be seen from the results in Table 1, the BLEU values of the prototypical sequence incorporation model combined with distributed representation matching proposed in this paper are higher than those of the traditional RNN, LSTM, GRU, and CNN models, as well as the bi-directional network BIRNN, BILSTM, BIGRU, and BICNN models, and the Transformer machine translation model for the tests of different slicing granularities in English. The BLEU values of this model are 23.25% and 21.88% in the subword granularity of English cutoff, respectively. Compared with the most basic CNN model, the BLEU scores of this model are improved by 3.90% and 3.29%, respectively; comparing with the Transformer machine translation model, the BLEU scores of this model are improved by 1.60% and 1.53% in the direction of English-Chinese and Chinese-English translation, respectively. In character granularity, this model improves the BLEU scores in English-Chinese and Chinese-English by 3.60% and 3.61%, respectively, compared with the basic CNN model; compared with the Transformer machine translation model, this model improves the BLEU scores in English-Chinese and Chinese-English by 2.25% and 2.27%, respectively. Comprehensive analysis shows that this model can achieve high translation results under four different granularities, no matter in the bidirectional translation process of English-Chinese or Chinese-English translation.

3.1.2. Comparison of training time. In order to more intuitively see the training effect of the proposed convolutional neural network model combined with the dynamic computing method. Experiments are conducted to compare and analyze the training time of this model with the above nine translation models in 100,000, 200,000 and 500,000 training English-Chinese bilingual parallel corpus, and the comparison results are obtained as shown in Figure 3.

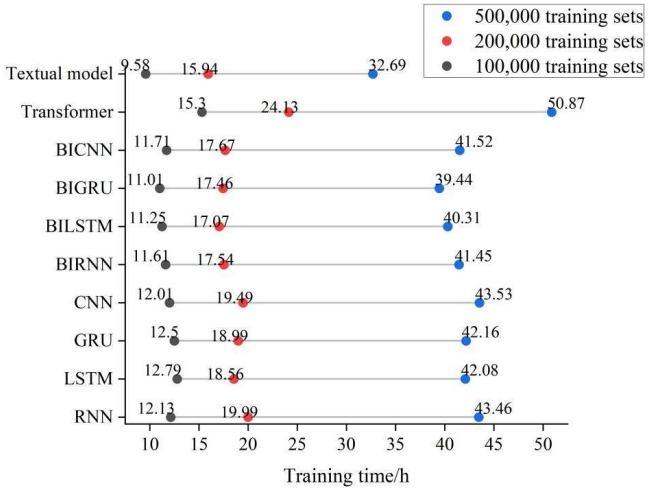

Fig. 3. Comparison of training time of different models

As can be seen from the comparison results in Fig. 3, in the training dataset of 100,000, the training time of the present model is 9.58h, the training time of the traditional convolutional neural network RNN model is 12.13h, the training time of the optimized bi-directional network BIRNN model is shortened to 11.61, and the training time of the Transformer machine translation model is 15.3h; in the 200,000 training datasets, the training time of this model is 15.94h, that of the RNN model is 19.99h, that of the BIRNN model is 17.54h, and that of the Transformer model is 24.13h; in the training dataset of 500,000 datasets, the training time of this model is 32.69h, which is shorter than that of the original Transformer machine translation model by It can be seen that the improved model combined with dynamic computing proposed in this paper can effectively improve the training speed and computational efficiency, and significantly shorten the training time.

3.1.3. Model performance comparison. In order to test the performance of the prototype sequence incorporation model incorporating distributed representation matching, this section conducts experiments on the effect of different encoding styles on the model performance.

In order to study the effect of different encoding methods on the model performance and verify the effectiveness of the translation model proposed in the paper, we use different networks as the encoding layer in this section to compare the effect of the model's parallel sentence pair translation under different encoding methods. Different comparison models are trained separately using Chinese-English data containing 200,000 positive and negative samples while keeping other settings unchanged, and the performance of different models is compared by precision, recall and F1 value scores. The detailed experimental results are shown in Fig. 4, where the comparison models all use the parameters that obtain the best performance.

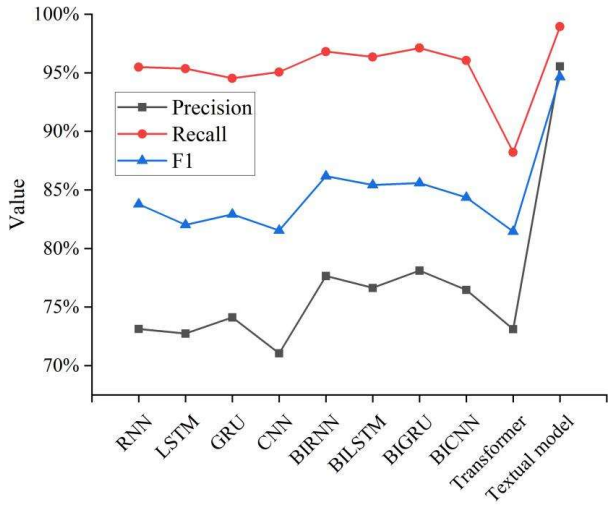

Fig. 4. Experimental results of different models in small-scale corpus

The experimental results in Fig. 4 show that the bidirectional network structure outperforms the model with the unidirectional network structure due to the fact that the bidirectional network can better characterize the contextual information of the input text; The F1 value of the Transformer-based model decreased by 2.34% compared to the RNN-based model because its complex network structure resulted in too many model parameters, and limited by the size of the training data, the model was not sufficiently trained and had not yet reached the fit; the improved neural machine translation model based on the translation memory information augmentation substantially outperforms the performance of the other models. Attributed to the fact that data size plays a crucial role in the performance of neural networks, the model augmented with translation memory information uses a huge amount of data to learn better contextualized vector representations for word representations, thus generating better semantic representations for the input text.

3.1.4. Comparison of changes in losses. For the English-Chinese neurotranslation task, the method proposed in this paper and the unidirectional network structure model, the bidirectional network structure model, and the baseline model are used in the corpus for training, respectively. The change of loss during the training process is shown in Fig. 5.

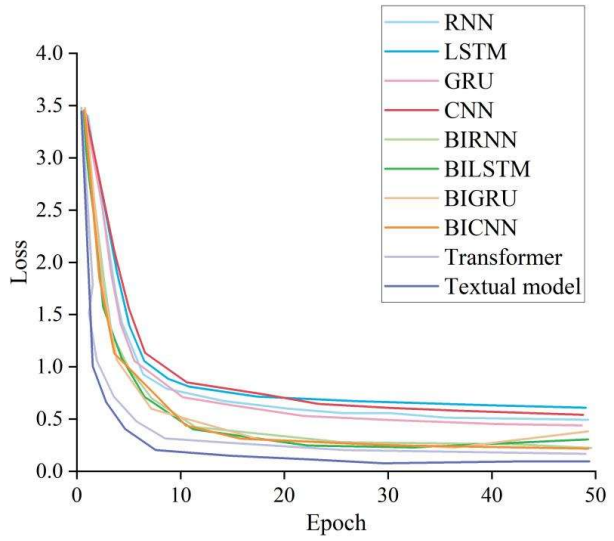

Fig. 5. Changes in English Chinese Translation Task Losses

Through the loss change graph, it can be seen that the loss function curve of the method proposed in this paper declines faster than the four unidirectional network structure models RNN, LSTM, GRU, CNN and the four bidirectional network structure models BIRNN, BILSTM, BIGRU as well as the baseline model Transformer in the early stage of training because the translation model can learn more semantic information from the large-scale corpus in the early stage of training, so the model can converge faster in the early stage of training. Because the translation model can learn more semantic information from the large-scale corpus in the early stage of training, the model can converge faster in the early stage of training, and when the training reaches about 10 epochs, the loss function curve starts to decline slowly and then stabilizes at about 20 epochs. Because most of the semantic information has been learned in the early stage of training, the speed of acquiring information in the later stage is gradually weakened, so the curve declines more slowly. The above changes can be seen, the fluctuation of the curve is relatively within the controllable range and shows a monotonous downward trend, indicating that the data enhancement method proposed in this paper is relatively stable in the training process, compared with other models in both the speed of convergence and the final convergence results have been improved to a certain extent.

3.2. Translation machine model decoding effect

3.2.1. Experimental configuration. Transformer, which incorporates dynamic computation, is used as the English-Chinese neural machine translation model. The number of encoder and decoder layers of the model are both 6, the number of attention heads is 8, the dropout ratio is set to 0.1, and the initial learning rate is 0.001.

In addition, the ReLU activation function is chosen to enhance the nonlinear characteristics of the model during the training process, and the cross-entropy loss function is used to optimize the model in order to improve the accuracy of translation.

3.2.2. Analysis of experimental results. In order to see more clearly the attention weights of the source and target languages in each part of the translation process, a heat map is used here to illustrate the final decoding effect. The hotspot effect of Chinese-English machine translation is shown in Fig. 6, the darker the color of the squares in this figure means the higher the probability of the candidate words it provides, and the highlighted area in the diagonal line indicates that the model correctly aligns the words or phrases of the input and output sequences during the translation process, which indicates that the model pays attention to the Chinese words in the corresponding position in the majority of the cases when it translates each English word, which also proves that the model's translation accuracy.

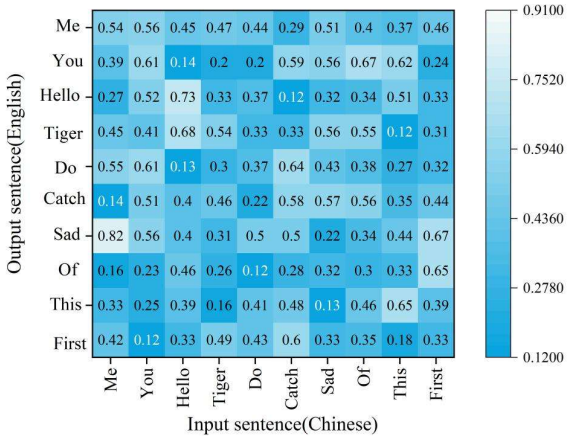

Fig. 6. Heat map in Chines-English machine translation

As can be seen from Figure 6, when translating the word "You", the Chinese word "you" and its vicinity pay more attention, and "Sad" and "sad" have a high alignment accuracy. According to this diagram, it can be seen that there should be no significant deviation between the subsequent translations and the standard translations.

4. Performance Analysis of Corpus English Machine Translation System

4.1. Translation system performance comparison

4.1.1. Comparison of noise reduction performance and coherence of translation systems. In order to verify the performance of the corpus English translation system designed by the study, the study conducts experimental verification in the CoNLL2020 dataset. Online translation and machine translation are utilized to compare with the system approach, and the noise reduction performance and context coherence after noise reduction are taken as the comparison indexes, as shown in Fig. 7 and Fig. 8 which show the comparison results of noise reduction performance and coherence of the three approaches, respectively.

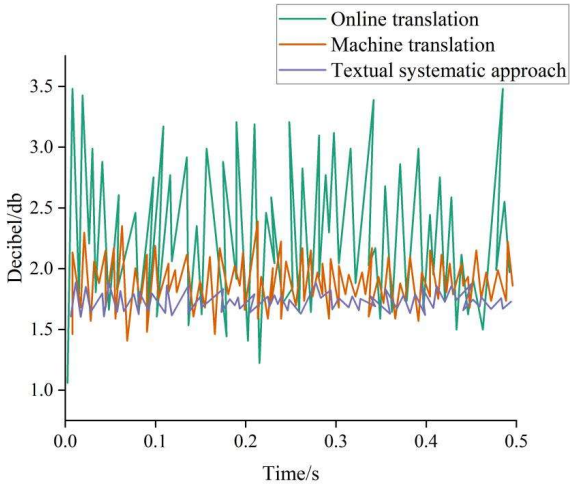

Fig. 7. Comparison of noise reduction performance of three methods

As can be seen from Fig. 7, the noise reduction effect of online translation is the worst, and the decibel range after noise reduction is distributed in [1.04~3.46]; the decibel range after noise reduction of machine translation is distributed in [1.57~2.33]; and the decibel range after noise reduction of systematic method is distributed in [1.62 ~1.89].

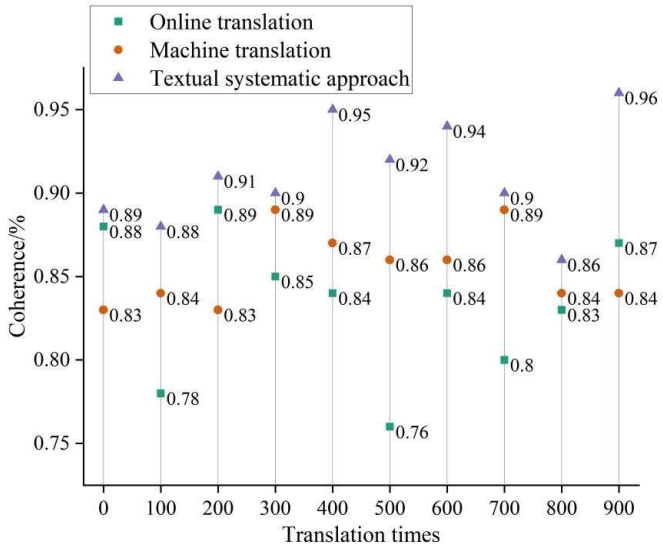

Fig. 8. Comparison of the coherence of three methods

As can be seen from Fig. 8, the average values of coherence for the system approach, machine translation and online translation are 91.1%, 85.5% and 83.4%, respectively. This shows that the systematic approach has some advantages in the noise reduction process.

4.1.2. Comparison of system performance in statement compression. In order to verify the performance of the designed system in utterance compression, the study utilizes three methods for comparison, as shown in Fig. 9 and Fig. 10 are the results of compression rate and compression stability comparison of the three methods respectively.

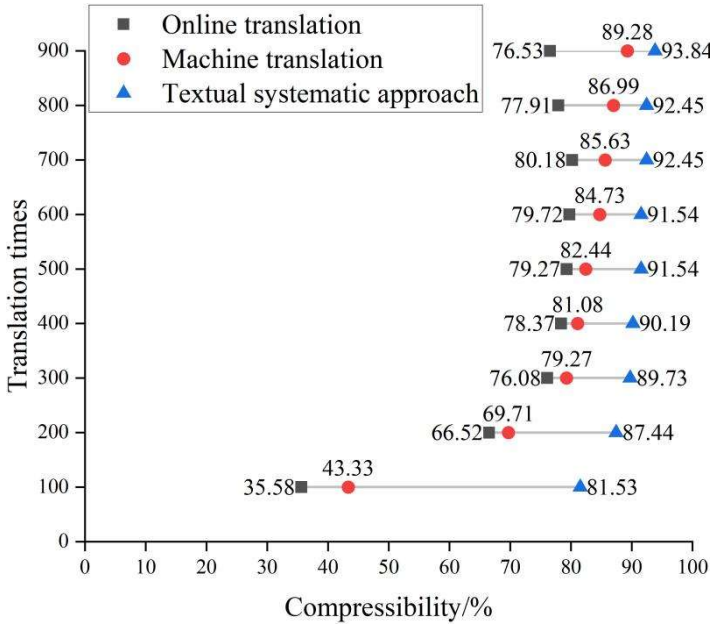

Fig. 9. Comparative results of compressibility of three methods

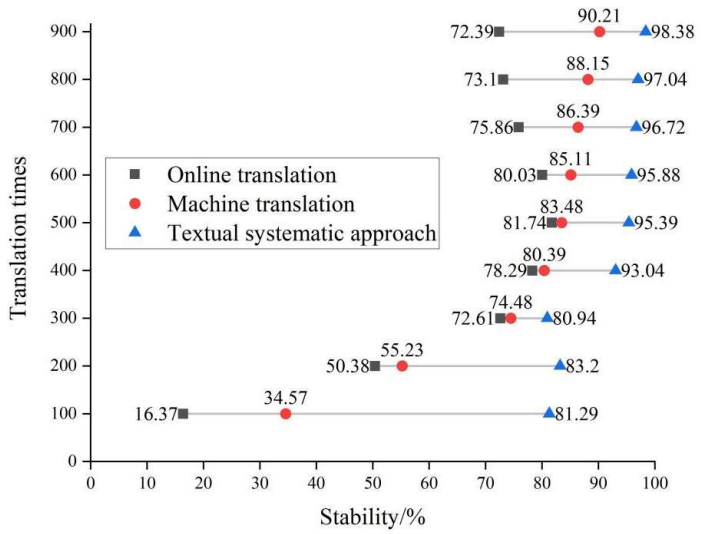

Fig. 10. Comparative results of stability of three methods

As can be seen from Fig. 9, the average value of compression rate for online translation is 72.24%, and the average value of compression rate for machine translation and systematic approach is 78.05% and 90.08%, respectively. As can be seen from Fig. 10, the average values of compression stability for online translation, machine translation and systematic approach are 66.75%, 75.33% and 91.32%, respectively. This indicates that the systematic approach proposed in this paper also has high performance in the process of English utterance compression.

4.2. Experiment on the actual application performance of the translation system

In order to further verify the practical application performance of the systematic method, the study compares its translated utterances with the results of manual translation, and also takes the fluency of the translated warnings as an indicator, and the results of the comparison between the translation results and the fluency are shown in Fig. 11.

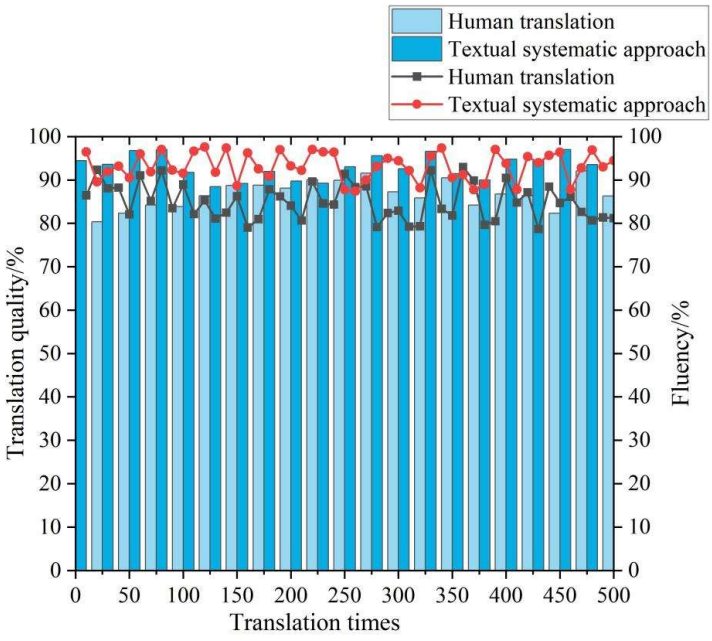

Fig. 11. Quality comparison between manual translation and systematic methods

As can be seen from Fig. 11, the average quality of manual translation is 86.83%, and the translation quality of the systematic method is evaluated to be 93.09%, which is 6.26% higher than the average quality of manual translation; the average fluency of manual translation is 85.01%, and that of the systematic method is 93.17%, which is 8.18% higher than the quality of manual translation. This shows that the English-Chinese translation method combined with dynamic computing proposed in this paper has strong applicability.

5. Conclusion

In order to improve the performance of the corpus-based English machine translation system, the study integrates the dynamic computing algorithm with the self-attention mechanism to improve the semantic consistency of the translation system, which is used to design a better English machine translation system. Meanwhile, a series of comparison experiments are conducted on the translation system, and the experimental conclusions are as follows.

Under four different granularities of words, syllables, subwords and characters, the BLUE values of the present model in Chinese-English material translation are 21.41%, 21.91%, 29.25% and 20.40%, respectively, which are 2.32%, 1.15%, 1.60% and 2.05% higher than that of the prototypical Transformer model, and even more significantly higher than that of the traditional RNN, LSTM, GRU, CNN models and the bidirectional network BIRNN, BILSTM, BIGRU, BICNN models. It shows that the model can achieve higher translation results in both English-Chinese and Chinese-English bidirectional translation processes, or under which granularity.

The improved model combined with dynamic computation proposed in this paper can effectively shorten the training time, which is 9.58h, 15.94h and 32.69h in 100,000, 200,000 and 500,000 training English-Chinese bilingual parallel corpus, respectively, which is shortened by 6.02h, 9.99h and 18.18h compared with the Transformer model, indicating that the combination of dynamic computation can significantly improve the training speed and computational efficiency.

With the parameters used to obtain the best performance for each model, the model in this paper still achieves the highest performance with accuracy, recall and F1 to of 0.974, 0.950 and 0.948, respectively. Its loss variation is also the fastest converging among the models, with the final convergence result stabilizing at 0.078.

In the practical application of translation system, the decibel range of this paper's systematic method after noise reduction is distributed in [1.62 ~ 1.89], and the decibel range of online translation and machine translation is distributed in [1.04 ~ 3.46] and [1.57 ~ 2.33]; the average values of coherence of the three methods are 91.1%, 85.5%, and 83.4%, respectively; the average value of compression rate is 72.24%, 78.05% and 90.08%; the average values of compression stability are 66.75%, 75.33% and 91.32%, respectively. The translation quality of the systematic method in this paper is evaluated as 93.09%, which is 6.26% higher than the average quality of human translation; the translation fluency is 93.17%, which is 8.18% higher than the quality of human translation.

Funding

This work was supported by 2024 Science and Technology Key Project of Henan Province: Research on the Translation and Dissemination Path of National Intangible Cultural Heritage in the Henan Section of the Yellow River Basin (242102320298) and by Henan Province Philosophy and Social Science Planning Project: Research on the Dissemination Path of National Intangible Cultural Heritage in the Henan Section of the Yellow River Basin Based on Parallel Corpora (2024CYY030).

References

[1] Brisset, A. (2017). Globalization, translation, and cultural diversity. Translation and Interpreting Studies, 12(2), 253-277.

[2] Xie, S. (2018). Translation and globalization. In The Routledge handbook of translation and politics (pp. 79-94). Routledge.

[3] Akpaca, S. M., Minaflinou, E., & Afolabi, S. (2020). Translation in the Era of Globalisation. International Journal of Linguistics and Communication, 8(2), 13-21.

[4] Alves, F., & Vale, D. C. (2017). On drafting and revision in translation: A corpus linguistics oriented analysis of translation process data. Annotation, exploitation and evaluation of parallel corpora, 89-110.

[5] Di Gangi, M. A., Cattoni, R., Bentivogli, L., Negri, M., & Turchi, M. (2019). Must-c: a multilingual speech translation corpus. 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. 2012-2017). Association for Computational Linguistics.

[6] Zhang, F. (2021). Application of data storage and information search in english translation corpus. Wireless Networks, 1-11.

[7] del Mar Sánchez Ramos, M. (2020). Teaching English for Medical Translation: A Corpus-Based Approach. Iranian Journal of Language Teaching Research, 8(2), 25-40.

[8] Akkoyunlu, A. N., & Kilimci, A. (2017). Application of corpus to translation teaching: practice and perceptions. International Online Journal of Education and Teaching, 4(4), 369-396.

[9] Li, Q., Wu, R., & Ng, Y. (2022). Developing culturally effective strategies for Chinese to English geotourism translation by corpus-based interdisciplinary translation analysis. Geoheritage, 14(1), 6.

[10] Laviosa, S. (2021). Corpus-based translation studies: Theory, findings, applications (Vol. 17). Brill.

[11] Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., ... & Dean, J. (2017). Google's multilingual neural machine translation system: Enabling zero-shot translation. Transactions of the Association for Computational Linguistics, 5, 339-351.

[12] Hoang, J. (2019). Translation technique term and semantics. Applied Translation, 13(1), 16-25.

[13] Yang, H. (2024). Optimized English Translation System Using Multi-Level Semantic Extraction and Text Matching. IEEE Access.

[14] Zhang, B., & Liu, Y. (2022). Construction of English translation model based on neural network fuzzy semantic optimal control. Computational Intelligence and Neuroscience, 2022(1), 9308236.

[15] Sun, C. A., Mu, J., Xiao, M., Liu, H., & He, P. (2025). Semantic Structure Invariance-Based Metamorphic Testing for Machine Translation Systems. IEEE Transactions on Reliability.

[16] Fu, L. (2017, April). On semantic equivalence in English-Chinese translation. In 2017 International Conference on Society Science (ICoSS 2017) (pp. 164-167). Atlantis Press.

[17] Hanczakowski, A. (2018). Translating Emotion from Italian to English: A Lexical-Semantic Analysis. The AALITRA Review, (13), 21-34.

[18] Beizaee, M., & Suzani, S. M. (2019). A semantic study of english euphemistic expressions and their Persian translations in Jane Austen's novel "Emma". International Academic Journal of Humanities, 6(1), 81-93.

[19] Liu, Y., Chen, S., & Yang, Y. (2025). Semantic alignment: A measure to quantify the degree of semantic equivalence for English–Chinese translation equivalents based on distributional semantics. Behavior Research Methods, 57(1), 51.

[20] Pu, X., Mascarell, L., & Popescu-Belis, A. (2017, April). Consistent translation of repeated nouns using syntactic and semantic cues. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers (pp. 948-957).

[21] Cao, L., & Fu, J. (2023). Improving efficiency and accuracy in English translation learning: Investigating a semantic analysis correction algorithm. Applied Artificial Intelligence, 37(1), 2219945.

[22] Liang, J. (2024, February). English Translation Model and System Design Based on Semantic Similarity. In 2024 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) (pp. 1083-1087). IEEE.