Hindawi
Mobile Information Systems
Volume 2022, Article ID 8478760, 11 pages
https://doi.org/10.1155/2022/8478760

*Research Article*

# Research on Transfer Learning-Based English-Chinese Machine Translation

**Li Jiaxin and Wang Peiting**

*Hebei University of Architecture, Zhangjiakou 075132, China*

Correspondence should be addressed to Li Jiaxin; ljx1027@hebiace.edu.cn

Received 6 January 2022; Revised 6 February 2022; Accepted 11 February 2022; Published 11 March 2022

Academic Editor: Hasan Ali Khattak



In order to solve the problems of time-consuming and frequent mistranslations in traditional translation methods, this study designed an English-Chinese machine translation method based on transfer learning. On the basis of analyzing the basic principles and specific strategies of English-Chinese machine translation, the traditional neural machine translation methods are analyzed, and then the translation process is optimized by transfer learning. On the basis of preprocessing English-Chinese translation text data, features of English-Chinese translation text are extracted, features of English-Chinese translation text are rapidly classified by feature transfer learning, and machine models of English-Chinese translation are constructed based on the classification results. The objective of feature transfer learning is to reuse the past knowledge obtained in the form of dataset to be utilized for another target data. The findings of the experiment show the effectiveness of the proposed method in achieving the design expectation. The benefit of the proposed method includes a short translation time and fewer mistranslations.

## 1. Introduction

As one of the fields with the longest history of computer technology application, machine translation has attracted the attention of linguists, philosophers, psychologists, scientists, engineers, and other scholars from different fields since its birth because it involves multiple disciplines [1]. In its 70 years of development, machine translation has experienced a process from rise to peak, from slump to opening up new research, confirming its strong industrial application and academic research value [2].

Since 1990, various significant interests have been witnessed in the emergence of many bilingual and multilingual corpora [3]. Research papers on the new dynamics of machine translation were published by IBM researchers in the late 1980s and early 1990s. These papers gave a detailed explanation of the traditional machine translation method based on dictionaries and conversion rules and the instance machine translation method based on a parallel corpus. The role of these research papers is remarkable in promoting the consolidation and expansion of machine translation theory and the practice and formation of new methods and rules.

The rise and development of machine translation after World War II mainly benefited from the invention and application of the first batch of computers [4]. At the same time, the development of cryptography during World War II and the interest in language research led people to realize that machine translation can be seen as a process of coding in one language and decoding in another. Relevant scholars believe that any language problems faced by translation can be attributed to lexical meaning problems and grammatical structure problems [5]. Semantic problems mainly include polysemy, homomorphism, ambiguous meaning, and ambiguous meaning. Structural problems mainly include morphological structure, syntactic structure, and textual structure [6]. In addition, extra linguistic issues (language emotion, tone, background knowledge, etc.) are also important factors affecting translation quality.

Linguistic-based ontologies can be glossaries, dictionaries, controlled vocabularies, taxonomies, folksonomies, thesauri, or lexical databases [7]. Linguistic-

based ontologies established semantics understanding for natural language. Many works have been done on ontology learning approaches. The ontology is used as a key element for domain knowledge. The ontologies are used as a language model based on a neural network [8].

Translation has become a comprehensive, complex, and difficult task for human beings. Machine translation has developed from the original rule-based translation method to the current neural translation method that simulates the human brain. Although there is still much room for improvement in performance, its high development potential and broad application space cannot be ignored [9].

Therefore, a machine translation method based on semantic selection is designed in reference [10], which studies the translation technical route by listing English units integrating semantic information and combining the characteristics of semantic information. Then, GIZA++ is used as the translation cornerstone, words are aligned with Berkeley aligner, reverse transcription grammar is introduced to describe the structural correlation features of the Chinese language form and the English translation language form, and the machine translation method is completed through sentence static and static collocation and ambiguity solving. In reference [11], designing a machine translation method for English long sentences based on interdependent syntactic analysis and sequence labeling this method is based on the analysis of the interdependence syntactic rules that were used, respectively, to match with the sequence labeling method based on conditional random field, and then on coarse-grained and fine-grained segmentation, segmentation of long sentences works together. In reference [12], an English-Chinese translation method based on the improved SEQ2SEQ model is designed. According to the method, at present, machine translation mainly optimizes and evaluates the Indo-European language family and seldom optimizes Chinese. Likewise, the SEQ2SEQ model, which is the most useful and efficient attention-based neural machine translation model in the field of machine translation, does not consider the grammatical transformation between different languages. And therefore, this method suggests an optimized English-Chinese translation model, applying different text preprocessing and embedding layer parameter initialization methods and enhancing/improving SEQ2SEQ model structure by adding a translation layer for syntactic changes between encoder and decoder. In this way, the parameters and training time of the model can be reduced by 20%, and the translation performance can be increased by 0.4 BLEU. With the use of the conversion layer, the SEQ2SEQ improves the translation performance by 0.7~1.0 BLEU.

However, in practical application, it is found that the above traditional methods have different degrees of the time-consuming translation process and many mistranslations. To solve this problem, this study designs a transfer learning-based English-Chinese machine translation method.

## 2. Theoretical Analysis of English-Chinese Machine Translation

*2.1. Basic Principles of English-Chinese Machine Translation.* Machine translation is a science that studies how to translate one language into another language automatically through an electronic computer [13]. The machine has an automatic translation system. During the translation procedure, the original text is read into the machine, and the automatic translation system stored in the machine starts translating. After the translation, the automatic printer of the machine typed the translation. This newly developed automatic English-Chinese translation system by us consists of dictionary, word meaning analysis system, lexical analysis system, and word order adjustment system. The schematic diagram of the English-Chinese machine translation system is shown in Figure 1.

Step 1: Read English. The English sentences are to be translated, punched through the card punch, read into the computer by the automatic card reader, and allocated to the specified storage unit for storage.

Step 2: Look up a dictionary. The dictionary in question here is a machine dictionary, generally including a comprehensive dictionary, subdictionary and phrase dictionary, and other dictionaries. The comprehensive dictionary gives various grammatical features, semantic features, and further processing information of each English word. Grammatical features include nouns, including singular nouns, plural nouns, proper nouns, common nouns, living nouns, and inanimate nouns; verbs, including transitive verbs, intransitive verbs, infinitives, past tense, and participles; and the so-called semantic features, including simple words and polysemy. Further processing information includes the address of various subdictionaries and processing methods.

Subdictionary refers to noun dictionary, verb dictionary, adjective dictionary, and other subdictionaries. The subdictionary is actually the continuation of the integrated dictionary to solve the problems that cannot be solved in the zinc dictionary. Phrase dictionaries deal with the lexical problems of English. A phrase dictionary includes features such as the part of speech and the meaning of a phrase [14].

Step 3: Lexical analysis system mainly solves the problem of one word with multiple categories. English, like other languages, has many kinds of words and needs a special system. The method of distinguishing one word from another is mainly by morphological marks and collocation of parts of speech in context. The morphological signs in English are not as rich as those in Russian, French, and other languages, so it can only solve part of the phenomenon of multiple types of words and mainly needs to rely on the collocation of context. For example, "have" can be a verb or an auxiliary verb. It cannot be solved by morphological signs but only by context, that is, have + is a past

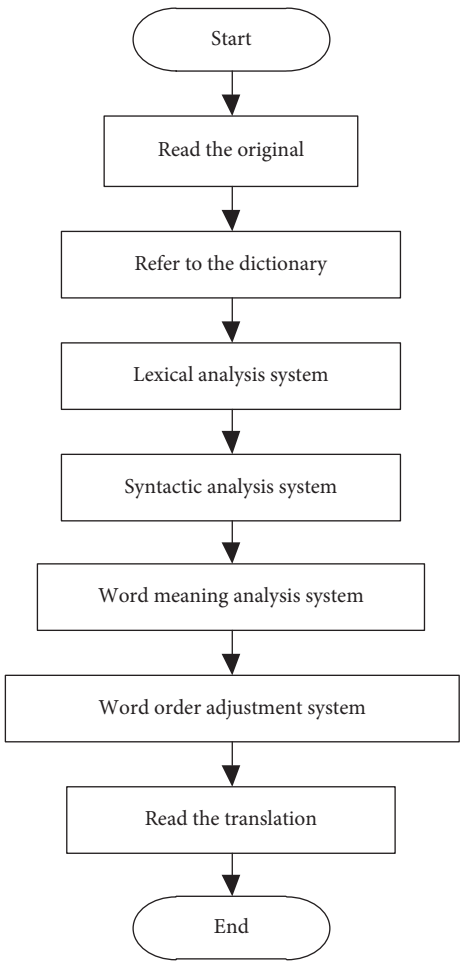


Figure 1: Schematic diagram of English-Chinese machine translation system.

participle and have + Ved), is a verb. To can be a preposition or an infinitive. There is no sign of morphological change, and it can only be solved by the collocation of context; that is, to 10 N is a preposition, to + V is an infinitive symbol.

Step 4: Syntactic analysis system and word order adjustment system: syntactic analysis is based on lexical analysis. Only after the parts of speech in an English sentence are determined can syntactic analysis be performed correctly. The task of the syntactic analysis system is to determine the syntactic function of English words in sentences. According to the traditional grammar, it is to determine the subject, predicate, attribute, object, and adverbial of each word in the sentence to lay the foundation for adjusting the word order. Syntactic analysis and word order adjustment are carried out simultaneously in English-Chinese machine translation. In other words, once the English word order is analyzed by the syntactic analysis system, it is immediately adjusted to the Chinese word order.

Step 5: Word meaning analysis system mainly solves the problem of polysemy [15]. In English, as in other languages, a word often has different meanings in different contexts. This requires a word meaning analysis system to analyze English words carefully. It should be pointed out that semantic analysis is still one of the main difficulties in machine translation research around the world. This is because the semantic analysis is more complex than syntactic analysis. A word often has several or even dozens of meanings and is often distinguishable only in its context. Whether word meaning analysis is scientific or not directly affects the quality of translation. At present, there is still no ideal analysis method.

Step 6: Read out the English sentence after the automatic translation system translation, through the Chinese output program to read the Chinese translation processing and then through the electronic computer automatic printing equipment automatically typed.

### 2.2. Specific Strategies of English-Chinese Machine Translation

*2.2.1. Strategies for Literal Translation.* Take the original sentence, which is a literal translation, but it cannot be simply identified with a letter or phrase. As far as cognitive linguistics is concerned, languages can be translatable to each other. Native speakers of different languages share some similarities in their language expression habits, and an expression can be understood and recognized by people of two cultures at the same time. In English-Chinese translation strategies, sentence meaning, cultural connotation, and grammatical relations should be consistent, reflecting both the meaning of the original sentence and the form so as to achieve readers' understanding. Literal translation retains the meaning of the original language to a certain extent [16].

*2.2.2. Set Translation Strategies.* The existence of set translation is the theoretical basis of translation, which can activate the subconscious culture and make it easier for readers to understand the meaning of the sentence, and this subconscious expression is understood by more people. The essence of set translation is the conversion of interlingual meaning, that is, the code conversion between linguistic symbols, especially the method of parallel translation. The translation should correspond to idioms, and only on the premise of accurate semantic expression can it be accepted by readers and maximize the translation effect. In order to "bring the original text into the heart," the translator must analyze the original text from the aspects of meaning, color, culture, and history and pay attention to the problems in translation. When translating cognitive thinking into the target language, the translator should look for phrases consistent with the original word in long-term memory and use the original word to express the concept understood. Therefore, in English-Chinese translation, sentence patterns should be fully adjusted so that the two languages can be flexibly converted as much as possible [17].

To maintain the parallelism of the structure, the semantic meaning of the source language should not be destroyed or distorted, and similarly, the metaphorical color of the

original text should not be altered very much. Semantic understanding refers to the meaning of words arranged in a syntactically acceptable way. Such a statement is in more consonance with the psychology of the readers and can raise their resonance. Original metaphors are replaced with the use of common target words in this way. This needs the translator to strictly compare the meaning of the source metaphor with the translation metaphor. It is not suitable for use in case of any differences, despite the form being the same.

*2.2.3. Free Translation Strategies.* The lack of concept is a difficulty in the source language culture, which is characterized by lack of concept and ambiguity. To ensure the acceptance of different cultures, translators often choose the same language to express the corresponding meaning with the target language of free translation on the basis of maintaining the semantics of the original text [18]. Replace untranslatable metaphors with the most appropriate literal interpretation, discarding the rhetoric of the original text and using everyday language. If a suitable metaphor is not found in the translation to express the original meaning, then the most common translation is a free translation, and no matter what the form of the translation is, the translation method is very different from the sentence pattern and semantics of the original text. The main purpose of translation is to help readers understand the meaning of the original text, while the form of free translation is to deal with the text flexibly according to the translation rules. The accurate meaning of a translation should be reflected in a certain context. Therefore, appropriate translation methods should be selected according to specific translation scenarios.

## 3. Analysis of Traditional Neural Machine Translation Methods

Neural machine translation is a translator used to make translation possible of words from one language to any other language. Neural machine translation became a buzzword in the advanced era of technology because many multinational organizations have adopted neural machine translation engines for internal and external interaction. The most common example of neural machine translation is Google and Baidu translate. The neural machine translation comprehends different types of machine translation where artificial neural networks predict a series of numbers [19]. The structure of the neural machine translation model is shown in Figure 2.

The introduction of deep learning networks into the field of machine translation constitutes a neural machine translation process. If neural machine translation models are classified according to network performance, two modules can be obtained, namely, decoder and encoder. The former involves the tree search algorithm, and the connection part of the two modules is the attention force mechanism. While the latter often applies important technologies such as word vectors and long- and short-term memory [4] networks.

The encoder transforms the input source language text into vector representation in vector space through a neural network.

The source input text expression for neural machine translation is as follows:

$$X = \{x_1, x_2, x_3, \ldots, x_i, \ldots, x_n\}, \tag{1}$$

where $n$ represents the number of words in a sentence.

At the initial stage of neural network coding, the embedding layer of the encoder encodes all the terms of the source input text and converts them into the word vector expression shown as follows:

$$W^i = x_i \times R^{|V|}, \tag{2}$$

where $V$ represents the size of the source language vocabulary. Through continuous neural network coding, all word vectors $W^i$ will generate a hidden state $h_i$. The hidden state contains the current lexical information. Taking the current hidden state as the input of the neural network at the next moment, the hidden state information at that time will be saved with the effectiveness of the gate valve mechanism. Therefore, when encoding the final word $x_i$, the implicit state will contain all the implicit state information of the sentence. Since most of the hidden states are extracted semantic information, text vectors need to be input into the recurrent neural network to define the encoder's hidden state output. The structure of the recurrent neural network is shown in Figure 3.

In Figure 3, the weight from the input layer to the forward coding layer is $W_1$, the weight from the reverse coding layer is $W_2$, the self-cyclic weight of the forward coding layer is $W_3$, the self-cyclic weight of the reverse coding layer is $W_4$, the weight from the forward coding layer to the output layer is $W_5$, and the weight from the reverse coding layer to the output layer is $W_6$. The six weights mentioned above are used repeatedly in the whole process, and the noncyclic characteristics of the structure are guaranteed by the removal of the direction of arrows of the forward and reverse coding layers.

The decoder performs the output operation after changing the intermediate vector output by the encoder to the target vector. If the termination identifier of the source input text is received by the encoder, the encoding stops, and the decoding phase starts. The vector representation of the source input text is set as the initialization decoder unit [20].

In essence, decoding is the process of calculating the conditional probability of the output of all target words in known parts. According to the implicit state $s_j$ of the current decoder and the first $j-1$ successful output target words, the conditional probability of the $j$-th target word can be calculated, and its calculation formula is as follows:

$$P\left(y_j | y_{j-1}, x_i\right) = g\left(s_j\right), \tag{3}$$

where the nonlinear function is expressed as $g(\cdot)$ and $s_j$ is calculated by the following formula:

$$s_j = f\left(y_{j-1} \times s_{j-1}\right). \tag{4}$$

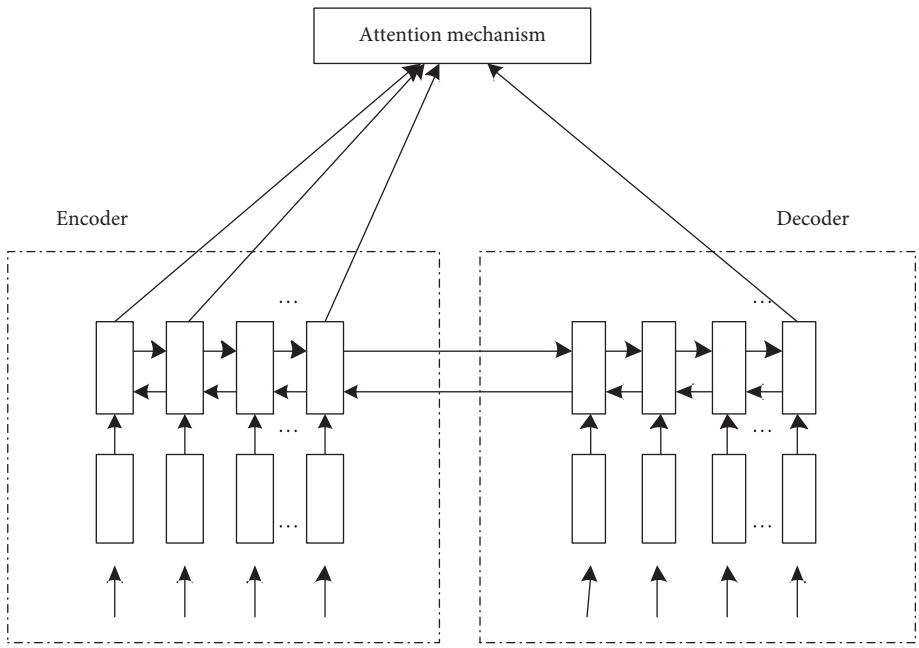


Figure 2: Schematic diagram of neural machine translation model.

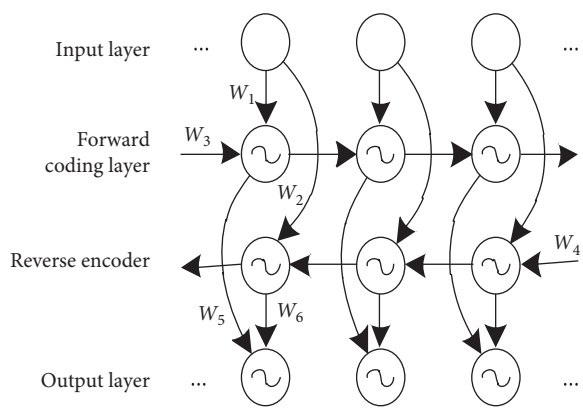


Figure 3: Schematic diagram of recurrent neural network structure.

By learning the conditional distribution, the model can output the target words with the maximum likelihood with the tree search algorithm on the premise of knowing a source input text, namely, the translation output.

## 4. Optimization of English-Chinese Machine Translation Design Based on Transfer Learning

### 4.1. Preprocessing and Feature Extraction of English-Chinese Text Data

*4.1.1. Preprocessing of English-Chinese Translation Text Data.* In order to make the extracted text features more obvious, it is necessary to perform numerical and normalized preprocessing on the collected English-Chinese text data before extracting English-Chinese text features.

Step 1: Numerical processing. Because English-Chinese text data are nonnumerical attributes, distance calculation cannot be implemented in the K-means clustering algorithm, so attributes of each dimension of text data need to be converted into numerical values. The frequency presented by attributes of different dimensions is used to replace the initial attributes to realize numerical transformation so as to prevent the clustering error caused by unequal distances between values of the same attributes in the transformation process [21].

Step 2: Normalization. There is a large gap between the values of different dimensions in the text data. In order to make more effective use of the data of different dimensions, it is necessary to implement normalized processing for the data of different dimensions in this paper, and the processing process is as follows:

$$X = \frac{x - N_{\min}}{N_{\max} - N_{\min}} \times 100, \tag{5}$$

where $N_{\max}$ and $N_{\min}$ represent the highest and lowest data in a dimension, respectively, and $x$ represents the data to be normalized.

*4.1.2. Extract Text Features from English-Chinese Translation.* In the process of English-Chinese translation, the sentence to be translated is first converted into a string of characters and punctuation. Sentences can be divided into phrases made up of phrases, which in turn can be broken down into characters. Before the application of the English-Chinese translation strategy, the VSM document classification method is relied on to represent text information and transform original sentences into patterns that can be

understood by computers. The sentence to be translated usually includes various characters and punctuation marks, and the characteristic item of the sentence is the minimum structure in the string. In addition, the number of feature items contained in a sentence is given weight values according to the importance of the sentence [22].

According to the above feature terms and vector values, the context features of sentences are extracted. The feature extraction algorithm is used to map the optimal context of sentences to the English-Chinese translation process. In the process of English-Chinese sentence translation, the total number of translation contexts and the total number of semantic translation types should be set so as to achieve the translation process of standard translation contexts. The number of sentences that can reach the standard translation context needs to be based on the number of translation contexts and the probability of semantic translation to get the final result.

Therefore, this study selects the TF-IDF algorithm to extract text features from preprocessed English-Chinese translation text data to form an English-Chinese translation text feature sample set. Text features are extracted by calculating the weight of text proximity and the word frequency in English-Chinese translation, where the weight of the words in the English-Chinese translation text can be obtained by the product of IDF and TF. On this basis, the best text of the words can be extracted, and the formula is

$$C(X, n) = \mathrm{IDF} \times \mathrm{TF}, \tag{6}$$

where $n$ indicates the document number. The task of IDF is to improve the key degree of words that appear less frequently and the difference in texts, and its calculation formula is as follows:

$$\mathrm{IDF} = \log \frac{M}{m}, \tag{7}$$

where $m$ represents the number of words contained in all documents and $M$ represents the number of documents contained in English-Chinese translated texts. TF stands for feature frequency, and its expression is

$$\mathrm{TF} = \frac{\log(d+1)}{\log N}, \tag{8}$$

where after the document is processed, $N$ represents the total number of words in the document and the number of words is represented by $d$.

According to the above results, the appropriate and inappropriate matrices for translation context are obtained. By analyzing the translation context matrix, the optimal sentence context and the evaluation criteria of sentence semantics and context are obtained. Because the result of optimal context is the same as a result of correlation degree mapping, through the above research, the results of context feature extraction are obtained, which provides a basis for the design of the English-Chinese translation process.

*4.2. Fast Classification of English and Chinese Text Features Based on Transfer Learning.* Transfer learning is a machine learning method in which a pretrained model is reused for another task. Transfer learning is related to problems such as multitasking learning and concept drift, and it is not a specific field of machine learning. However, transfer learning is very popular in some deep learning problems, such as when there are a large number of resources needed to train the deep model or when there are a large number of datasets to pretrain the model. Transfer learning only works when the depth model feature in the first task is a generalization feature [23]. This kind of transfer in deep learning is called inductive transfer. The idea is to narrow down the search for possible models in a favorable way by using a model that applies to different but related tasks.

With the emergence of more and more machine learning application scenarios, the existing supervised learning with better performance requires a large amount of annotated data, which is a boring and costly task, so transfer learning has attracted more and more attention.

There are three ways to realize transfer learning:

(a) *Transfer Learning*. Freeze all convolutional layers of the pretraining model and train only the customized full connection layer.

(b) *Extract Feature Vector*. First, calculate the feature vectors of the convolution layer of the pretraining model for all the training and test data, then abandon the pretraining model, and only train the customized simplified version of the fully connected network.

(c) *Fine-Tune*. Freeze part of the convolutional layers of the pretraining model (usually most of the convolutional layers are close to the input because these layers retain a large amount of underlying information) or even do not freeze any network layer and train the remaining convolutional layers (usually part of the convolutional layers close to the output) and the full connection layer.

We can leverage knowledge in the source model to enhance learning in the target task. In addition to providing the ability to reuse established models, transfer learning can also assist in achieving learning objectives in the following ways:

(a) Improved baseline performance: When we augment the knowledge of isolated learners (also known as uninformed learners) with knowledge from the source model, baseline performance may improve as a result of this knowledge transfer.

(b) Model development time: Leveraging knowledge from the source model helps to fully learn the target task, compared to learning from the target model from scratch. This in turn will lead to improvements in the total time required to develop or learn the model.

(c) Improved final performance: Higher final performance can be achieved by using transfer learning.

Therefore, this study uses the transfer learning process to quickly classify English and Chinese text features. The transfer learning process obtains the classification target by updating and supplementing the pretraining model. AlexNet is a widely used transfer learning model.

(a) Convolution layer: after the above-mentioned English-Chinese text features are input to the convolution layer, the convolution filter is applied to calculate the local features of the layer by adding bias, which can be obtained by the activation feature function. The process is as follows:

$$X_0^l = \sum_{i \in f} X_i^{l-1} \lambda_{io}^l + B^l, \tag{9}$$

where $l$ represents the number of convolution layers, $f$ represents the convolution filter, $X_i^{l-1}$ represents the feature input to the convolution layer, $X_0^l$ represents the output result of the convolution layer, $B^l$ represents the bias term in the $l$ layer, and $\lambda$ represents the convolution kernel.

(b) Pooling layer: this layer is used to reduce the data dimension of the feature of the convolution layer to avoid the overfitting phenomenon. For the data in the filter, the mean pooling method is used to obtain the internal mean, and the quadratic features are collected by formula (10):

$$X_0^l = \frac{\sum_{i \in f} X_i^{l-1} + B^l}{k}. \tag{10}$$

(c) Full connection layer: the calculation method of this layer is the weight calculation method of the multilayer neural network, which is described by formula (11):

$$X_0^l = f\left(\sum_{i \in l-1} \omega_{ji} X_i^{l-1} + B^l\right). \tag{11}$$

In the formula, $\omega_{ji}$ represents the weight from the feature quantity $j$ to the feature quantity $i$ of English and Chinese text. The feature quantity $j$ exists in the $l$ layer, and the feature quantity $i$ exists in the $l+1$ layer.

The model has five convolution modules and three fully connected modules, and third fully connected layer outputs the number of features identified in English and Chinese texts. ReLU function is then used as the neuron activation function of the two types of modules to optimize the gradient dispersion phenomenon generated in the training process so as to complete the classification of Chinese and English text features according to the similarity of features through the above process.

The higher the similarity of English and Chinese text features is, the more similar their categories will be. Therefore, Chinese and English text features are transferred to the similarity calculation process, assuming that $Q^*$ is the transfer matrix, which can be defined as

$$Q^* = f(d) \cdot X_0^l, \tag{12}$$

where $f(d)$ represents the function. Then, the annotated word vector $T_{ij}$ is multiplied by the similarity measure transfer value $w_{ij}'$, and the amount of English and Chinese text transfer is obtained as follows:

$$M_{ij} = \frac{T_{ij} \times w_{ij}^*}{Q^*}. \tag{13}$$

In order to unify the result of the final migration vector, the vector is normalized, the threshold value is set, and the exceeded threshold value is retained to obtain the unknown annotation set.

*4.3. Translation of English and Chinese Texts.* Based on the feature transfer quantity obtained above, this study constructs a machine model of English-Chinese translation through the application of the fuzzy matching algorithm.

In the process of translation strategy design, the sentences to be translated and their corresponding translation results are described, respectively. The sentence to be translated is divided into several phrases by parsing the string into which the sentence is converted, and the form of the phrase is a continuous sequence. For each phrase contained in the phrase set, the phrase with high similarity is searched in the phrase table. The results obtained by similarity are compared with the preset similarity threshold value. If the phrase table contains phrases that meet the similarity threshold value, the sentence translation can be directly completed by using the existing knowledge. However, if there is no phrase that meets the similarity requirement in the phrase table [11], it indicates that there is no corresponding translation content in the phrase table, and this type of phrase can be described as an unknown phrase. In the design of the English-Chinese translation model, it is necessary to extend unknown phrases in the phrase table and output the final English-Chinese translation results by using the found content.

For unknown phrases that are not in the phrase list, it is difficult to complete sentence translation directly. Therefore, by using fuzzy matching technology, we can find the same content in the phrase table as the extended phrase. After that, the original phrase is replaced with the found translation content to generate English-Chinese translation results. Since fuzzy matching can find more similar phrases and output more translation results, it is necessary to use a combinatorial classifier to identify the sentence with the best translation effect from all English-Chinese translation results. Through the above, the translation model shown in Figure 4 is generated.

In the application of the English-Chinese translation model, one of the key links is to obtain the similarity of sentences to be translated. To obtain the sentence similarity, it is often necessary to count the keywords contained in the sentence through some operations. After obtaining the same

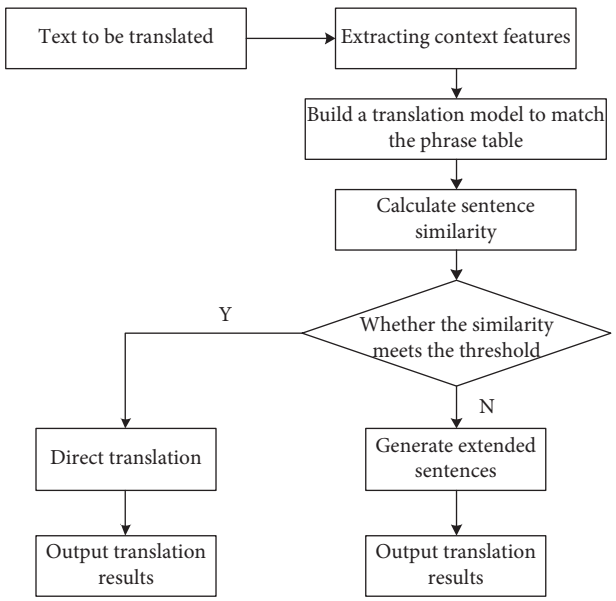


Figure 4: Translation model.

keywords, the sentence structure similarity result can be obtained. In addition to the keywords of sentences, it is also necessary to take word order, word form, and other sentence structure features into further consideration to obtain accurate similarity values when obtaining sentence similarity. To obtain sentence morphological similarity is to analyze the similarity degree of word morphology, which depends on the number of words shared between sentences and the length of sentences.

If the probability of a feature term appearing in two sentences is greater than once, it indicates that the number of feature terms is the minimum number of occurrences. The total length of two sentences is obtained, and the result is obtained by the total number of feature items. Once two sentences contain more identical words, it indicates that the sentences have a high degree of similarity.

The standard sequence of the sentence is consistent with the lexical order of the sentence, while the lexical sequence number of the sentence is in reverse order. In this case, when there is no common keyword between the two sentences, the result of obtaining sentence similarity is 0. When the set contains 1 cooccurrence word, the similarity acquisition result is expressed as 1/2. In the process of obtaining sentence similarity in translation, the acquisition results of word shape similarity and word order similarity are the key influencing factors, thus reflecting sentence information as a whole. Introducing the feature of sentence length to analyze, in the process of analysis, the length difference between two sentences is large, which means that the two sentences obtained have a relatively low similarity.

The final result is obtained by obtaining the ratio of the absolute values of the two sentences to the normal difference by taking the absolute value function of the difference between the length of one sentence and the length of the other and subtracting the scale value by 1. In this paper, the analytic hierarchy process (AHP) is used to obtain the different weights of sentence structure similarity, and the factors affecting sentence translation are decomposed into objectives and criteria so as to obtain the analysis results by qualitative and quantitative methods.

In addition to the above analysis, the semantic similarity of sentences also needs to be analyzed because the actual meaning contained in the sentence to be translated reflects the concrete semantic information. By analyzing the concept nodes contained in semantics, we can obtain the content words corresponding to the concept nodes. The concept similarity of sentences is obtained by using concept nodes, and the same semantic similarity is obtained. Count the number of the same nodes contained in the two conceptual nodes in the sentence, and get the semantic coincidence degree of the sentence.

After the processing of function words, subjects and objects are selected from the remaining keywords to form a set. Then, the semantic distance between concepts is used to obtain the semantic similarity of sentences. After the weight assignment of the concept relation is completed, the shortest path algorithm is used to obtain the shortest path value between the concept node and the root node, and the semantic similarity of sentences is obtained. The final sentence similarity is obtained by considering synthetically sentence structure similarity and sentence semantic similarity.

In order to ensure the final translation quality, the similarity threshold should be set according to the analysis results. Once the similarity reaches the threshold, it is proved that the current matching content is the phrase with the highest accuracy, thus terminating the matching with other phrases and improving the efficiency of English-Chinese translation.

Based on sentence similarity mining, this paper generates analogical similarity in English-Chinese translation. In the process of designing English-Chinese translation strategies, the method of word alignment between English and Chinese translation is designed on the basis of multilevel templates. Example translation model is used to analyze the differences between the original text and the translated text and clarify the word alignment relationship. After analyzing the application effect of the currently designed translation strategies, we can ensure that the translation quality meets the requirements after the application of English-Chinese translation strategies. After applying the multilevel template, words and phrases are divided into different levels, and the corresponding relationship between English and Chinese sentences is clearly reflected by combining the results of lexical annotation. This method only needs part of the corpus and a bilingual dictionary to complete English-Chinese translation, which reduces the complexity of English-Chinese translation and improves the quality of translation.

## 5. Experiment and Result Analysis

In order to verify the practical application performance of the above designed English-Chinese machine translation method based on transfer learning, the following experimental process is designed. Taking the corpus of an English-

Table 1: Statistical table of experimental corpus information.

| Name | Number |
|---|---|
| Number of sentences | 2000 |
| Extract the number of sentences for the longest noun phrase | 535 |
| Number of sentence structures | 562 |
| Maximum number of noun phrases | 2701 |
| Average sentence length | 13 |
| Average sentence structure length | 15 |
| Average length of the longest noun phrase | 8 |

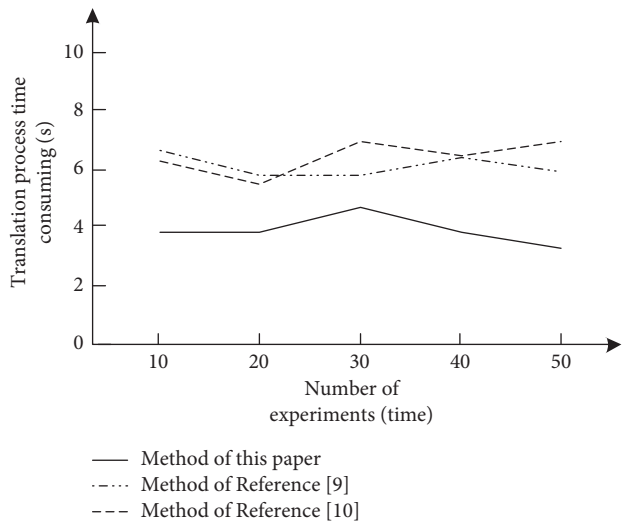


Figure 5: Comparison of translation time between different methods.

Table 2: Comparison of mistranslation times between different methods.

| Number of experiments/time | Method of this paper | Method of reference [11] | Method of reference [12] |
|---|---|---|---|
| 10 | 1 | 7 | 9 |
| 20 | 3 | 8 | 7 |
| 30 | 0 | 9 | 7 |
| 40 | 2 | 8 | 10 |
| 50 | 4 | 13 | 7 |

Chinese translation platform as an example, part of the corpus is randomly selected as experimental objects. The relevant information of the corpus is shown in Table 1.

In order to improve the explicability of experimental results, the method of English long sentence machine translation based on dependency parsing and sequence annotation (method of reference [11]) and the method of English long sentence translation based on improved SEq2SEQ model (method of reference [12]) were taken as the comparison group. The method of this paper was taken as the experimental group for the comparative test.

First of all, the translation process time of different methods is used as an indicator for comparison and verification, and the results are shown in Figure 5.

According to the results shown in Figure 5, in the 10th experiment, the translation process of the method of this paper takes 3.8 s, and that of the method of reference [11] takes 6.7 s. The translation process of the method of reference [12] takes 6.1 s. In the 30th experiment, the translation process of the method of this paper, the method of reference [11], and the method of reference [12] took 4.6 s, 6.0 s, and 7.1 s, respectively. In the 50th experiment, the translation process of the method of this paper, the method of reference [11], and the method of reference [12] took 3.9 s, 6.1 s, and 7.3 s, respectively. And in the whole experiment, the translation of the method of this paper takes up to 4.6 s, indicating that the method of this paper has high timeliness.

Then, the mistranslation times of different methods are used as indicators for comparative verification, and the results are shown in Table 2.

The findings of the 10th experiment in Table 2 show that the number of mistranslations of the method of this paper was 1, and the number of mistranslations of the method in reference [11] and the number of mistranslations of the method in reference [12] were 7 and 9, respectively. In the experiment no. 30, the mistranslations of the method of this paper were 0, 9, and 7, respectively, for the method in reference [12] and the method in reference [5]. In the 50th

experiment, the mistranslations of the method of this paper were 4 times, 13 times of the method in reference [11], and 7 times of the method in reference [12]. On the grounds of the above results, it can be analyzed that the number of mistranslations of the method of this paper is lower, which indicates its greater effectiveness and usefulness.

## 6. Conclusion

In the current study, an English-Chinese machine translation method based on transfer learning is designed to fix the problem of time-consuming and frequent mistranslations. The neural machine translation methods are analyzed, and the translation process is optimized by transfer learning. For the English-Chinese translation process, the features of English-Chinese translation text are extracted and classified by feature transfer learning. The machine models of English-Chinese translation are constructed based on the classification results. Experimental results show that the proposed method has a short translation time and fewer mistranslations. The results indicated that the machine translation, from the original "mechanical brain" to the deep learning neural network machine translation model, has witnessed the unremitting efforts of human beings to explore knowledge. Since the machine translation has many challenges as it deals with the core issues of the highest and deepest level of human intelligence, exploring and trying to interpret the human brain can be a good aspect of further research.

At present, machine translation is actively exploring model research and development, pretranslation source language processing, posttranslation editing, and accumulated useful experience. Since the core of machine translation technology is statistics-based, building a large database is crucial for machine translation. For text types lacking sufficient corpus, such as literary texts, philosophical texts, and texts of low-resource languages, machine translation can cooperate with human translation to improve translation efficiency and guarantee translation quality. Although the research of machine translation is still in its infancy in many aspects, the wide application and interdisciplinary research conducted from its birth to the present show that, as an emerging industry, machine translation enjoys vigorous and broad development prospects in both application and scientific research.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

[1] L. I. Qiang, F. Wong, C. H. A. O. Sam, Y. Han, T. Xiao, and J. Zhu, "Template-driven neural machine translation," *Chinese Journal of Computers*, vol. 42, no. 03, pp. 566–581, 2019.

[2] J. Mizera Pietraszko and P. Pichappan, "Advances in Intelligent Systems and Computing] Lecture Notes in Real-Time Intelligent Systems," *Study of the establishment of a reliable English-Chinese machine translation system based on artificial intelligence*, vol. 10, no. 2, pp. 13–23, 2018.

[3] K. W. Church, "Emerging trends: APIs for speech and machine translation and more," *Natural Language Engineering*, vol. 24, no. 6, pp. 951–960, 2018.

[4] A. Khan and A. Sarfaraz, "RNN-LSTM-GRU based language transformation," *Soft Computing*, vol. 23, no. 24, pp. 13007–13024, 2019.

[5] A. Vaswani, S. Bengio, E. Brevdo, F. Chollet, and J. Uszkoreit, *Tensor2Tensor for Neural Machine Translation*, vol. 16, no. 5, pp. 328–333, 2018.

[6] M. X. Chen, O. Firat, A. Bapna et al., "The Best of Both Worlds: Combining Recent Advances in Neural Machine Translation," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, July, 2018.

[7] C. Roussey, F. Pinet, M. A. Kang, and O. Corcho, "An introduction to ontologies and ontology engineering," in *Advanced Information and Knowledge Processing*, pp. 9–38, Springer, NY, USA, 2011.

[8] P. Lomov, M. Malozemova, and M. Shishaev, "Training and application of neural-network language model for ontology population," in *Proceedings of the Computational Methods in Systems and Software*, pp. 919–926, Springer, Cham, 2020.

[9] H. Choi, K. Cho, and Y. Bengio, "Fine-grained attention mechanism for neural machine translation," *Neurocomputing*, vol. 284, no. 5, pp. 171–176, 2018.

[10] H. Zhang, "Research on machine translation based on semantic selection," *Automation & Instrumentation*, vol. 5, no. 8, pp. 29–32, 2018.

[11] J. An, "Machine translation of long English sentence based on dependency parsing and sequence labeling," *Journal of Lanzhou University of Technology*, vol. 44, no. 1, pp. 100–103, 2018.

[12] X. Xiao, S. Li, W. Yu, J. Liu, and B. Liu, "English-Chinese translation based on an improved seq2seq model," *Computer Engineering and Science*, vol. 41, no. 7, pp. 1257–1265, 2019.

[13] P. Tran, D. Dinh, L. E. Tan, and L. H. B. Nguyen, "Linguistic-Relationships-Based Approach for Improving Word Alignment," *ACM transactions on Asian language information processing*, vol. 17, no. 1, pp. 1–16, 2018.

[14] F. Aqlan, X. Fan, and A. Alqwbani, "Arabic–Chinese neural machine translation: romanized Arabic as subword unit for Arabic-sourced translation," *IEEE Access*, vol. 17, no. 4, pp. 38–45, 2019.

[15] S. Barzegar, B. Davis, and S. Handschuh, "Multilingual Semantic Relatedness Using Lightweight Machine Translation," in *Proceedings of the IEEE International Conference on Semantic Computing*, IEEE, Laguna Hills, CA, USA, January, 2018.

[16] L. Liu, "Way to better English-Chinese machine translation based on contrastive study of machine and human translation of the financial times," *Journal of Yichun University*, vol. 42, no. 1, pp. 74–78, 2020.

[17] F. Zamora Martinez and M. J. Castro Bleda, "Efficient embedded decoding of neural network language models in a machine translation system," *International Journal of Neural Systems*, vol. 28, no. 9, pp. 1–15, 2018.

[18] H. Long, V. H. Pham, and D. Dinh, "Improving neural machine translation with AMR semantic graphs," *ICIC Express Letters*, vol. 17, no. 10, pp. 301–311, 2021.

[19] J. E. Ortega, R. Castro Mamani, and K. Cho, "Neural machine translation with a polysynthetic low resource language," *Machine Translation*, vol. 34, no. 4, pp. 325–346, 2020.

[20] L. Xing, "Design of translation automatic generation system for English-Chinese machine translation," *Modern Electronics Technique*, vol. 41, no. 24, pp. 86–89, 2018.
[21] Q. Ren, Y. Su, and N. Wu, "Research on Mongolian-Chinese machine translation based on the end-to-end neural network," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 18, no. 01, pp. 46–59, 2020.
[22] J. Qiu, Y. Liu, Y. Chai et al., "Dependency-Based Local Attention Approach To Neural Machine Translation," *Computers, Materials and Continua*, vol. 59, no. 2, pp. 547–562, 2019.
[23] L. I. Sha and C. Qi, "Computer aided translation model simulation based on bilingual E-chunk," *Computer Simulation*, vol. 36, no. 12, pp. 345–348+352, 2019.