

## Retraction

# Retracted: Automatic Calibration Algorithm for English Text Translation Based on Semantic Features

### Security and Communication Networks

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

Copyright © 2024 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

### References

- [1] F. Liu, "Automatic Calibration Algorithm for English Text Translation Based on Semantic Features," *Security and Communication Networks*, vol. 2021, Article ID 3870656, 8 pages, 2021.

## Research Article

# Automatic Calibration Algorithm for English Text Translation Based on Semantic Features

Fengzhen Liu 

HuangHuai University, Zhumadian 463000, China

Correspondence should be addressed to Fengzhen Liu; 20070479@huanghuai.edu.cn

Received 11 August 2021; Revised 14 September 2021; Accepted 20 September 2021; Published 30 September 2021

Academic Editor: Yinghui Ye

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

At present, the existing methods of English article flip calibration neglect to extract English semantic features, which leads to errors in English flip results and has a great impact on the accuracy and time consumption of translation sentence calibration. Therefore, a semantic feature-based automatic text flipping calibration algorithm is proposed. According to the features of semantic information in machine translation, a semantic grammar tree is constructed to complete the machine turning of English articles. The CART decision tree attribute is obtained, and the random forest method is introduced to extract the input matrix and output matrix of the corpus feature as samples to determine the spatial attribute feature of the mistranslated sentences. Choose 10000 English sentences about human body parts as the experimental object and design the simulation experiment. The experimental results show that the minimum and maximum accuracy rates are 95.4% and 100.0%, respectively. The proposed algorithm is time-consuming, and the KSMR value is lower than that of the traditional method. It is proved that the error rate of English article flipping is significantly reduced.

## 1. Introduction

With the development of Internet technology, a lot of English translation software has been born. It is more convenient for people to obtain the English translation results [1]. According to its own algorithm, the English translation software provides the English translation results by searching the semantic vocabulary of the whole web. However, these English translation results are unreliable and cannot be used directly, and a lot of manual proofreading is needed later. There are two drawbacks to manual proofreading.

The template matching model for automatic calibration of Chinese-English translation is constructed by using the differentiated semantic modification method, the semantic tree for automatic calibration of Chinese-English translation is established by combining the syntax analysis method, the semantic similarity feature quantity of Chinese-English translation is extracted, and the semantic allocation and sentence subject word analysis in Chinese-English translation are carried out according to different combinations of

semantic similarity. The tree lexical-semantic database of Chinese-English translation is established by using the machine learning method, and the structure is automatically adjusted according to the semantic modification objectives in the tree lexical-semantic database, so as to realize the automatic calibration of Chinese-English translation and subject word registration, calculate the optimal semantic correlation feature quantity of each clause, and use the machine learning algorithm for automatic optimization to realize the automatic calibration of Chinese-English translation. However, the above traditional methods ignore the extraction of English semantic features, resulting in large errors in English machine translation results, which greatly impacts the accuracy and time consumption of translation sentence calibration. Sun et al. [2] first conducted an empirical study on the relationship between UNMT and UBWE/CMLM. The empirical results show that the performance of UBWE and CMLM has significant impact on the performance of UNMT. Based on this, we propose a new UNMT architecture with a cross-language presentation protocol to capture the interaction between UBWE/CMLM

and UNMT during UNMT training. Experimental results on multiple language pairs show that the proposed UNMT model significantly improves over the corresponding state-of-the-art UNMT baseline. Heo et al. [3] presents a multimodal neural machine translation system, which uses text and related images to translate Korean images into English subtitles. The data in the experiment was a set of unlabeled images containing only bilingual subtitles. To train the system with a supervised learning method, we propose a weak labeling method, which selects keywords from image titles using the feature selection method. Keywords are used to determine the image label roughly.

## 2. Related Work

Li and Qi [4] proposed and designed a machine-aided translation model based on Bilingual E-Chunk. Using semantic smoothing technology to solve the problem of data sparsity, the computer-aided translation model based on the E-Chunk module is optimized by the maximum likelihood estimation algorithm. The probability matching cost of the preprocessed corpus is calculated to find an optimal semantic recognition path; combined with smoothing technology and matching probability calculation, the machine-aided translation model of bilingual E-Chunk is designed. According to the assisted translation model, the bilingual translation process is obtained to realize the accurate translation processing of translation materials. Cui [5] developed a new English translation computer intelligent proofreading system, including search module, English translation module, English translation calibration module, and behavior log. Use the search module to complete the search of the basic meaning of the words to be proofread and the subject content, master the user's behavior data through the behavior log, and optimize the system. The software part of the system adopts the computer intelligent proofreading method based on the improved phrase translation model to find the correct words to replace the words to be proofread and to realize the intelligent proofreading of the English translation. Luo [6] proposed an automatic calibration method for Chinese-English translation based on machine learning and semantic feature extraction.

An automatic calibration algorithm based on semantic features is proposed to solve the defects of the previous methods.

## 3. A Machine Translation Approach to English Text Based on Semantic Cognitive Features

According to the machine translation features of semantic information [7], a semantic syntax tree [8] is constructed to realize machine translation of English articles. The specific steps are as follows:

- Step 1: list semantic units according to semantic information and get semantic syntax patterns based on semantic syntax tree;
- Step 2: align words

Step 3: calculate the translation probability of the word after the translation method processing, use lexicalized weighting formula to calculate the translation probability to obtain phrase similarity and as a new feature

Step 4: the new features will be added to the decoder and output translation results; Figure 1 shows the roadmap of translation technology from different semantic cognitive perspectives.

As shown in Figure 1, listing semantic units, acquiring semantic grammatical patterns, and word alignment play an essential role in translation methods. In addition, from different semantic cognitive perspectives, technical translation routes can be referred to realize the problem of machine translation of English articles.

**3.1. Enumeration Semantic Unit.** There is no obvious correspondence between words in different languages. In different languages, sentences express the same meaning, meaning different languages can translate each other. The semantic units of each language compose the sentences of the corresponding language. Words, words, phrases, and variables constitute the semantic units.

First, listing semantic units requires consideration of text and images based on learning independent context vectors from text encoders [9] and image encoders [10], respectively; secondly, in the process of learning spaces for text and images, sentences are constrained to be closer to the meaning of the sentence in the spaces, that is, to be semantically identical. Finally, the concepts of semantic unit and semantic unit representation are set up, specifically as follows:

Set the language collection  $R = \{L | L = 1, 2, \dots, N\}$  consisting of  $N$  languages. In this article, language  $L$  includes Chinese and English.

The sentence is the same sentence meaning in every different language, that is, the sentence meaning in every different language.

Semantic unit  $CV$  is the unit representing the same meaning in each different language, and  $(P_l CV)$  is the semantic unit representing the same meaning in each different language.

Sentence meaning  $(CC)$  is a semantic unit  $(CV)$ .  $(P_l CV)$  is a semantic unit representation in different languages and a sentence in different languages.

**3.2. Semantic Syntax Tree.** The semantic grammar tree represents the semantic grammar hierarchy between concepts [11]. Syntactic category is the top of the tree, syntactic category name is the branch node, and semantic category name is the node below the third level. There will be a lot of single semantic grammar content, such as with text description of the picture. The image description generation method describes and annotates the natural language of the input image generation, thus associating the language with the image.

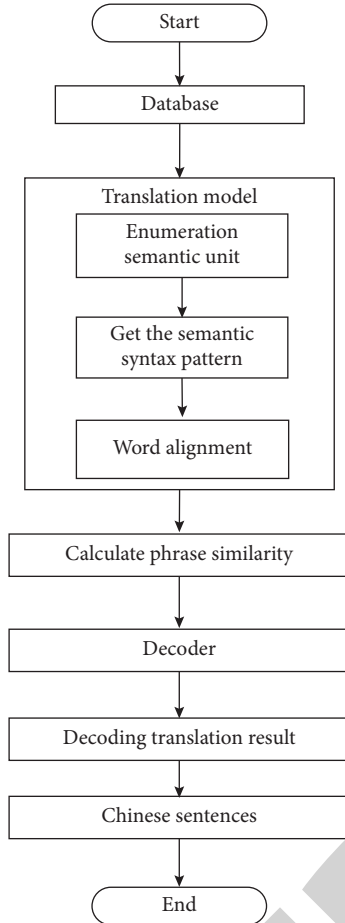


FIGURE 1: Translation flowchart from different semantic cognitive perspectives.

According to the synonym word forest, the semantic category names are set by comprehensively considering other semantic concept hierarchies [12]. Figures 2 and 3 are part of the English and Chinese semantic grammar trees.

In Figure 2, there may be other nodes in the path, represented by dotted lines.

#### 4. English Text Machine Automatically Calibrates Mistranslations

Based on the results obtained in the previous section, combined with the stochastic forest method, the automatic calibration of English article mistranslation is realized. Stochastic forest is a kind of machine learning method based on bagging. It is an ensemble learning method with the decision tree algorithm, has good application in many fields, and can achieve higher prediction accuracy.

**4.1. CART Decision Tree Attribute Integration.** Decision tree (DT) is an integrated algorithm [13], which is a decision analysis method based on the known probability of occurrence of various situations, and by forming a decision tree, it obtains the probability that the expected value of the

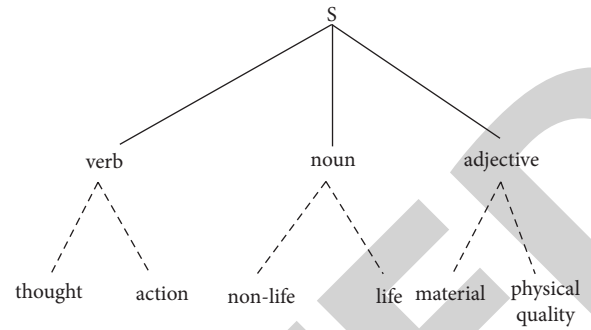


FIGURE 2: English semantic syntax tree.

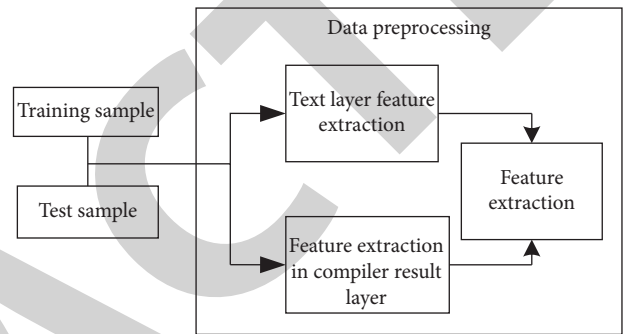


FIGURE 3: Detection method framework.

net present value is greater than or equal to zero, evaluates the risks of the project, and determines the feasibility thereof. This is a graphical approach that intuitively uses probability analysis. Because this decision branch is drawn as a graph, it is very similar to the branch of the tree and is called a decision tree. In machine learning [14], a decision tree is a predictive method representing the mapping between an object's attributes and values. This process involves not only dividing the playing field but also constructing the decision tree. The generation of CART (classification and regression tree) is a process of recursively constructing a binary decision tree, choosing the optimal function according to the criterion of Gini index minimization, and determining the optimal binary partition point of the function.

Given that there is class  $N$  and the probability of sample points belonging to class  $n$  is  $P'$ , then the Gini index is defined as follows:

$$J(p) = \sum_{N=1}^n P' (1 - P'_k). \quad (1)$$

For the set  $R$  of machine translation of English articles, the Gini index formula is

$$R = 1 - \sum_{N=1}^n (|C_N|)^2, \quad (2)$$

In equation (2),  $C_N$  is a subset of English text translation belonging to class  $N$ .

The Gini coefficient of set  $D$  under the condition of feature  $A$  is expressed as follows:

$$J(D, A) = |D| \cdot C_N \cdot R. \quad (3)$$

In the generation of the CART decision tree, the Gini index is calculated and the minimum Gini index is taken as the criterion of feature selection. The decision tree is generated recursively from the root node, and the training set is divided into subsets with basically correct classification.

**4.2. Random Forest Algorithm Detection Framework.** The data set contains  $r$  samples, Randomly select a sample and place the sample in the sample set. Put the sample back into the original record to ensure that the next sample is selected. A random sample  $r$  can be obtained with  $r$  sample set of  $r$  samples. Then, a  $R$  sample set containing  $A$  training samples is sampled, and a basic learner is trained according to each sample set and then combined.

Based on random forest [15], arbitrary attributes are introduced into decision tree training, as shown in Figure 3.

**4.3. Feature Extraction.** In this paper, the proportion of nonalphanumeric characters and the matching result based on the feature code are added according to the detection feature of NeoPI. After character operation conversion, the proportion of nonalphanumeric characters is generally larger than that of ordinary pages. Extract the longest continuous string record length, usually less than 300.

The expression of the text coincidence index  $IC$  is

$$IC = \sum_{i=0}^{300} \frac{F_i}{N} (N - 1), \quad (4)$$

where  $i$  represents the value of ascii corresponding to the character and  $N$  represents the number of all characters and is the frequency of the occurrence of character  $i$ .

The calculation formula of information entropy is as follows:

$$S = -IC \cdot \lg(p(x))^{-1}, \quad (5)$$

where  $p(x)$  represents the probability function of character  $x$ .

The compression ratio of the file after code confusion is greater than that of the normal file. After character conversion and encryption, the ratio of nonalphanumeric characters to the total number of special characters and the total number of text characters is greater than that of ordinary files [16, 17]. The rule in the sample matching signature text is marked as 1; otherwise, it is marked as 0.

Although there are many types of Web shells and ways to escape from them, they all depend on the basic structure of the shell, namely, data transfer and execution transfer. At present, there is not a perfect method to locate the steps of data transmission. The method of 1-gram is used to segment the sample at the level of code text, and the continuous string and the constant string ending with opening bracket are used as a phrase, and the frequency of each phrase in the sample is counted.

The opcode is the bytecode of the result layer of the PHP compilation and is the intermediate language after the script compilation. After the PHP script reads the script string, it converts it to a language fragment tag through the lexical parser and then to a language fragment tag. Finally, the Zend virtual machine executes each opcode to get the run results. Script corresponding to the opcode text can be directly obtained through the PHP VLD extension.

The binary grammar method (2-gram) is used to segment the opcode text. Two adjacent opcodes are split into a phrase. The phrase frequency matrix in each sample is calculated, and the word frequency matrix based on the sequence of operation codes is obtained as the eigenvector. The feature set is obtained using the feature extraction method of the previous part, combining the code text layer and the compile result layer. However, the dimension of the feature set may be very high, which will bring great computational pressure to the automatic calibration of mistranslated sentences. In the data preprocessing stage, feature selection can effectively reduce the dimension of the feature and the training complexity of later methods.

Feature selection [18, 19] is one of the hotspots in machine learning and pattern calibration. The Fisher linear discriminant method is one of the filtering feature selection methods based on distance measurement, which has the advantages of small calculation, high accuracy, good operability, and short calculation time.

By the Fisher linear discriminant algorithm for feature selection and feature dimension reduction, the central idea is to find a suitable projection axis, and make the samples on the axis of projection. The projection points of the same class should be as close as possible and should be as far away as possible from different types of projection points, that is, the smaller the deviation in the class, the farther the class is alienated, and its functions are as follows:

$$M = K \cdot \delta \cdot W. \quad (6)$$

In formula (6),  $w$  represents the projection direction,  $K$  represents the interclass dispersion matrix,  $W$  represents the intraclass dispersion matrix, and  $\delta$  represents the discriminant coefficient. The higher  $M$  is, the better  $\delta$  is. Fisher feature selected the feature as the projection axis, calculated the discriminant value of feature direction, and ordered each feature from large to small according to its discriminant value. The larger the discriminant value is, the more effective the feature classification is. In this paper, the extracted features are sorted from large to small according to the Fisher discriminant value, and the important features are selected in an appropriate proportion to form a new feature set for future method training.

The input matrix and output matrix after feature extraction are used as samples for training. Random forest is an integrated learning device with the ability to analyze complex interactive characteristics and quickly calibrate machine mistranslations of English articles.

**4.4. Automatic Calibration of Mistranslated Statements Based on Thresholds.** After the content of machine translation mistranslation sentences in English articles obtained above

is calibrated with the score  $f(\text{score})$ , the spatial attribute characteristic quantity of mistranslation sentences is determined. The two-dimensional information entropy of each difference mark is optimized to determine the calibration threshold of the difference function. The steps are described below.

Assume that  $x$  represents the weighted average vector of the distribution of the mistranslation statement node  $j$ ,  $\bar{x}$  represents the average vector of the differential feature distribution near  $j$ , and  $\rho$  represents the joint probability density of the distribution region of the intrusion data. Then, the score  $f(\text{score})$  is calibrated according to the content of the differential features. The calculation formula of the spatial feature quantity of the mistranslation statement distribution is

$$e' = u \mp \frac{j_{dr} \cdot m'_{hk}}{x - \bar{x}} \oplus f(\text{score}). \quad (7)$$

In formula (7),  $u$  describes the original information of mistranslated statements contained in different feature distribution regions.

Assuming that  $X_{\text{Characteristic}}$  is the global comprehensive feature of the distribution of mistranslation statements and  $o_{\text{thresholdvalue}}$  is the threshold value of the two-dimensional entropy of mistranslation statements, and equation (8) is used to optimize the two-dimensional information entropy of different features:

$$b' = \frac{X_{\text{Characteristic}} \oplus o_{\text{thresholdvalue}}}{e'}. \quad (8)$$

Let  $\phi$  be the different probability density vectors of each node in the distribution region of each difference feature and its neighborhood and  $\varepsilon$  be the weight vector of the two-dimensional information entropy of each difference feature; then, equation (9) is used to determine the calibration threshold of the mistranslation statement:

$$o''_{\text{thresholdvalue}} = \frac{\phi \oplus \varepsilon}{e'} + g \cdot b'_{\text{max}}, \quad (9)$$

where  $g$  is the constraint condition for the calibration of mistranslation statements and  $b'_{\text{max}}$  is the maximum two-dimensional entropy of differential features.

Assuming that  $z$  describes the distribution ratio of machine-translated sentences in English articles in each different feature distribution region [20],  $z_{\text{similarity}}$  describes the similarity between different types of differential feature sets, and  $n''$  describes the number of subsets contained in the distribution region  $p$  of mistranslation statements, then the automatic calibration is carried out by the following equation:

$$G_{\text{proofreading}} = \frac{z \mp z_{\text{similarity}}}{n''} \otimes o''. \quad (10)$$

According to the above steps, the automatic calibration of mistranslation sentences in English articles is completed.

## 5. Experimental Analysis

10000 sentences of English materials about human body parts were selected as the experimental subjects to find out the basic information about human body words in 1000 sentences of English materials, as shown in Table 1.

The English translation accuracy and recall rate of the method in this paper were tested according to four different situations, which are as follows:

- ① The translation result is successfully output, and the translation is accurate, consistent with the original Chinese translation
- ② The successful output of the translation result is different from the original Chinese translation, but the meaning is the same, which is an accurate translation
- ③ The translation result is successfully output, but the translation is wrong
- ④ Translation failure means that no translation result is output

The accuracy calculation formula is as follows:

$$P = \frac{\text{Number of times to output an accurate translation}}{\text{Total number of output translations}} \\ = \frac{a + b}{a + b + c} \times 100\%. \quad (11)$$

The recall rate is calculated as follows:

$$R = \frac{\text{Number of accurate translations}}{\text{The number of occurrences of human body words}} \times 100\%. \quad (12)$$

The test results of accuracy and recall rate of the method in this paper are shown in Table 2.

According to Table 2, the lowest accuracy of the method in this paper is 95.4% and the highest accuracy is 100.0%. The lowest recall rate was 96.5%, and the highest recall rate was 100.0%. The above experimental data show that the proposed method can effectively translate target words with high accuracy.

In order to further verify the effectiveness of the proposed method, a computerized intelligent proofreading system for English translation and a method for automatic calibration of Chinese-English translation based on machine learning and semantic feature extraction proposed in [6] are used to translate 1000 English sentences in the selected news field. The number of correctly translated sentences in sentences with words less than or equal to 30 words and more than 30 words is counted, respectively. Table 3 shows the test results of the three methods.

It is found from Table 3 that the number of correct sentences in the proposed method is higher than that in the other two methods. In the sentences whose aligner is less than or equal to 30 words, there are 9920 correct sentences in this method, and 9900 correct sentences in the sentences whose aligner is more than 30 words. Experimental results

TABLE 1: Basic information about human body words.

English human verbs	The number of sentences containing the word	The number of occurrences of this word
Heel	1278	80
Lip	965	75
Finger	1452	155
Mouth	1055	72
Ear	1000	136
Brain	1000	89
Leg	1000	185
Shoulder	982	39
Blood	1269	27
Tongue	1389	20

TABLE 2: Test results of accuracy and recall rate of the method in this paper.

English human verbs	Case ① occurrences/ times	Case ② occurrences/ times	Case ③ occurrences/ times	Case ④ occurrences/ times	Accuracy (%)	Recall rate (%)
Heel	165	105	10	6	98.4	97.4
Lip	243	109	10	4	97.2	99.2
Finger	253	75	15	1	100.0	97.3
Mouth	28	18	2	0	96.8	98.8
Ear	247	138	9	1	97.2	97.5
Brain	5	1	0	0	100.0	100.0
Leg	7	0	0	0	100.0	100.0
Shoulder	19	10	1	1	98.7	96.5
Blood	30	5	0	1	100.0	100.0
Tongue	97	49	7	2	95.4	97.2

TABLE 3: Statistical results of the correct number of translations of the three methods.

Methods	Corpus	Word alignment	Number of sentences translated correctly
Reference [5] methods	Closed set of 10000 sentences	Aligner less than or equal to 30 words	4230
	Closed set of 10000 sentences	Aligner is greater than 30 words	4850
Reference [6] methods	Closed set of 10000 sentences	Aligner less than or equal to 30 words	4200
	Closed set of 10000 sentences	Aligner is greater than 30 words	3670
In this paper, methods	Closed set of 10000 sentences	Aligner less than or equal to 30 words	9920
	Closed set of 10000 sentences	Aligner is greater than 30 words	9900

show that the proposed method has higher accuracy in sentence translation and provides significant reference value for English text flipping.

Using English articles of different themes as the basic corpus, the translation time of different methods is compared, and the experimental results are shown in Figure 4:

From the comparison of the sentence translation in Figure 4, we can see that the proposed method has some advantages in time consumption, especially for sports, military, and science materials. Taking science and technology as an example, in the same 196 sample data, the translation time of Cui [5] translation method is about 120 s, the translation method of Luo [6] is over 110 s, and the total translation time of the design method is about 69 s. It is important to the automatic translation of proper nouns and technical nouns, especially to the simultaneous interpretation, which can effectively avoid misunderstanding and disputes between the two parties.

KSMR value is one of the important indexes to measure the comprehensive performance of a translation method. The lower the KSMR value, the lower the error rate and the stronger the performance. Under the conditions of the four translation methods, the changes of KSMR values are statistically analyzed according to the six corpus samples mentioned above. The specific statistical analysis results are shown in Table 4:

The KSMR value is inversely proportional to the ease of translation; the more proper nouns a corpus contains, the more complex the translation is, the lower the KSMR value is. Taking the corpus of life class as an example, the difference of KSMR values of 4 methods is not obvious. Taking the most difficult scientific corpus as an example, the translation method of Cui [5] has a KSMR value of 1.254, and the translation method of Luo [6] has a KSMR value of 0.982, which is better than the above two methods. This is mainly because the smoothing technique and matching probability find the best translation path with the lowest cost.

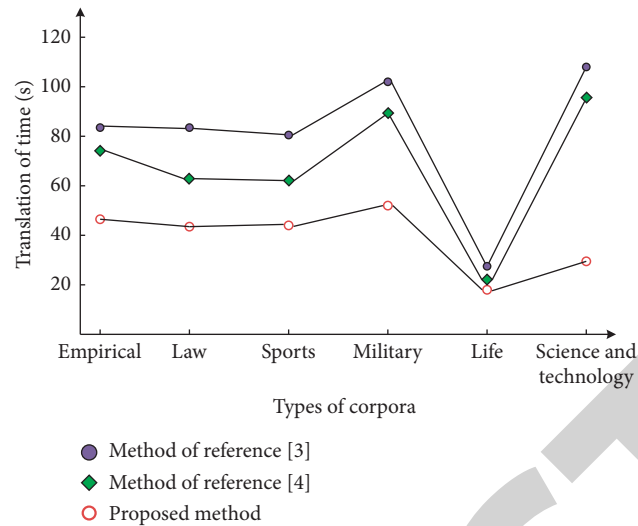


FIGURE 4: Comparison of translation time of different methods.

TABLE 4: Statistical analysis results of KSMR values.

Project	Sample size	Methods of reference [5]	Methods of reference [6]	The proposed method
The empirical	156	0.260	0.177	0.101
Sports	171	0.577	0.421	0.260
The law	168	0.347	0.265	0.115
Military	269	0.912	0.825	0.561
Life	265	0.257	0.136	0.091
Science and technology	196	1.254	0.982	0.774

## 6. Conclusion

In order to optimize the accuracy of English article flipping, an automatic calibration algorithm based on semantic features is proposed. To improve the accuracy of MT in English articles, a semantic grammar tree is constructed in this study, which combines the semantic information with the MT features. In order to correct translation automatically, the stochastic forest method is introduced to extract the input matrix and output matrix of corpus features, and the two-dimensional information entropy of each difference mark is optimized. The high-precision calculation of calibration threshold can effectively improve the accuracy of calibration. To verify the effectiveness of the proposed algorithm, simulation experiments are designed. Experimental results show that the proposed algorithm can improve the accuracy of English article translation significantly, and the algorithm takes less time, and KSMR is lower than the traditional method. The existing deep learning methods have achieved good results, and there is room for improvement in the accuracy of this paper. Therefore, in future work, we will incorporate deep learning algorithms to improve the model's performance.

## Data Availability

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

## Conflicts of Interest

The author declares that there are no conflicts of interest.

## References

- [1] S. Bi, "Intelligent system for English translation using automated knowledge base," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 5, pp. 1–10, 2020.
- [2] H. Sun, R. Wang, K. Chen, M. Utiyama, E. Sumita, and T. Zhao, "Unsupervised neural machine translation with cross-lingual language representation agreement," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1170–1182, 2020.
- [3] Y. Heo, S. Kang, and D. Yoo, "Multimodal neural machine translation with weakly labeled images," *IEEE Access*, vol. 7, pp. 54042–54053, 2019.
- [4] S. Li and C. Qi, "Computer aided translation model simulation based on bilingual E-chunk," *Computer Simulation*, vol. 36, no. 12, pp. 345–348+352, 2019.
- [5] D. Cui, "Design of intelligent computer proofreading system for English translation," *Modern Electronics Technique*, vol. 42, no. 4, pp. 179–182, 2019.
- [6] F. Luo, "Research on automatic calibration of Chinese-English translation based on machine learning," *Automation & Instrumentation*, vol. 250, no. 8, pp. 152–155, 2020.
- [7] N. Bharti, N. Joshi, I. Mathur, and P. Katyayan, "Quality-based ranking of translation outputs," *IT Professional*, vol. 22, no. 4, pp. 21–27, 2020.
- [8] P. Liina, "The neural basis of combinatory syntax and semantics," *Journal of Science*, vol. 366, no. 6461, pp. 62–66, 2019.



- [9] I. Lizarralde, C. Mateos, A. Zunino, T. A. Majchrzak, and T.-M. Grønli, "Discovering web services in social web service repositories using deep variational autoencoders," *Information Processing & Management*, vol. 57, no. 4, Article ID 102231, 2020.
- [10] B. Vennemann and T. Rsgen, "A dynamic masking technique for particle image velocimetry using convolutional autoencoders," *Experiments in Fluids*, vol. 61, no. 7, pp. 1–11, 2020.
- [11] J. Grudzinska and M. Zawadowski, "Continuation semantics for multi-quantifier sentences: operation-based approaches," *Fundamenta Informaticae*, vol. 164, no. 4, pp. 327–344, 2019.
- [12] R. Denaux and J. M. Gomez-Perez, "Vecsigrafo: corpus-based word-concept embeddings Bridging the statistic-symbolic representational gap in natural language processing," *Semantic Web*, vol. 14, no. 2, pp. 1–28, 2019.
- [13] P. Nancy, S. Muthurajkumar, S. Ganapathy, S. V. N. Santhosh Kumar, M. Selvi, and K. Arputharaj, "Intrusion detection using dynamic feature selection and fuzzy temporal decision tree classification for wireless sensor networks," *IET Communications*, vol. 14, no. 5, pp. 888–895, 2020.
- [14] X. Wu, X. Yuan, C. Duan, and J. Wu, "A novel collaborative filtering algorithm of machine learning by integrating restricted Boltzmann machine and trust information," *Neural Computing & Applications*, vol. 31, no. 5, pp. 1–8, 2019.
- [15] S. Ramalingam and K. Baskaran, "An efficient data prediction model using hybrid Harris Hawk Optimization with random forest algorithm in wireless sensor network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 20, pp. 1–25, 2020.
- [16] T. Vantuch, M. Prilepok, J. Fulneček, R. Hrbáč, and S. Mišák, "Towards the text compression based feature extraction in high impedance fault detection," *Energies*, vol. 12, no. 11, p. 2148, 2019.
- [17] Y. Tao, X. Yang, X. Meng, Y. Wang, Y. Yin, and G. Dong, "Hybrid cryptosystem based on plaintext related computational ghost imaging encryption and elliptic curve algorithm," *Journal of Optics*, vol. 22, no. 6, Article ID 065701, 2020.
- [18] K. S. Kyaw and S. Limsiroratana, "Optimization of text feature selection process based on advanced searching for news classification," *International Journal of Swarm Intelligence Research*, vol. 11, no. 4, pp. 1–23, 2020.
- [19] M. Labani, P. Moradi, and M. Jalili, "A multi-objective genetic algorithm for text feature selection using the relative discriminative criterion," *Expert Systems with Applications*, vol. 149, Article ID 113276, 113285 pages, 2020.
- [20] F. H. Lin, Y. F. Liu, H. J. Lee, C. H. C. Chang, I. P. Jaaskelainen, J.-N. Yeh et al., "Differential brain mechanisms during reading human vs. machine translated fiction and news texts," *Scientific Reports*, vol. 9, no. 1, pp. 1–10, 2019.