International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 4572-4581

https://doi.org/10.70517/ijhsa464301

Strategies for comparing Chinese and English abstracts of medical papers under the intelligent CARS model

Peijun Liu[1,*]
[1] Department of General Education, Foreign Language Teaching and Research Section, West Anhui Health Vocational College, Lu'an, Anhui, 237000, China
Corresponding authors: (e-mail: princecom2012@163.com).

Abstract As an important carrier of scientific communication, the abstract part of medical literature carries the key function of conveying core information. However, the differences in academic writing traditions between China and the West have led to significant differences in the semantic expression, rhetorical structure and information organization of Chinese and English medical paper abstracts. This study combines semantic role annotation technology and CARS model to explore the semantic expression differences between Chinese and English medical abstracts, and proposes a corresponding translation optimization method. By constructing a dataset containing 13,791 medical papers with a total of 249,762 corpora, the improved LSTM-CRF model is used for semantic role annotation, and the modified CARS model is applied to analyze the rhetorical structure of abstracts. The experimental results show that the LSTM-CRF model proposed in this paper performs well in the semantic role annotation task, with a precision rate of 86.6%, a recall rate of 87.1%, and an F1 value of 86.8%, which is an improvement of more than 9% compared with the comparison model. The speech step analysis shows that Chinese dissertation abstracts are used 8124 times in speech step 2, which is more than twice as many as 4016 times in English dissertation abstracts. In the translation performance evaluation, the BLEU value of the translation model incorporating semantic role features is improved by 5.47% to 7.50% compared with the comparison model, and the TER metric is reduced by 0.257 to 0.452. In the semantic component recognition experiments, the recognition accuracies of the eight major types of medical semantic components are over 90%. The results demonstrate that the combination of semantic role annotation and CARS model can effectively identify the expression differences between Chinese and English medical abstracts and significantly improve the quality of machine translation.

Index Terms Semantic role annotation, CARS model, medical abstract, LSTM-CRF model, machine translation, semantic expression differences

I. Introduction

A title epitomizes a paper and reflects the main argument of a paper, while an abstract summarizes the main content of the literature without comment [1]. In order to promote the exchange of academic information and results at home and abroad, the International Organization for Standardization's International Standard ISO 214 states that "every paper, short article and thesis in an academic journal should be accompanied by an abstract" [2], [3]. With the increasing frequency of international exchanges, the databases of major international searching organizations have become increasingly dependent on the English abstracts of papers [4]. As a result, most academic journals now require both Chinese and English titles and abstracts for papers formed by scientific research and related disciplines [5].

With the advancement of internationalization, as well as the establishment and popularization of computer databases, in order to facilitate literature tracking and retrieval, the abstract, which is the essence of medical papers, has inevitably become an element of the format specification of medical papers, and is the touchstone of scientific research results and world communication [6]-[9]. The medical profession itself has many disciplines and diverse categories, while the medical vocabulary is highly specialized. Among them, the English translation of Chinese medicine vocabulary is particularly difficult, and the translation of such vocabulary should not only reflect their own medical connotations, but also must take into account their cultural deposits [10]-[13]. Therefore, the translation of abstracts of medical articles is more difficult, and as far as the current translation of abstracts of medical papers is concerned, the main problem is still the weak language foundation of the translators, the lack of accurate choice of vocabulary and grammar, and the tendency to express themselves in Chinese English, which leads to improper translation [14]-[17]. Whether an excellent paper can be accepted by foreign journals or included in databases, the English translation of the abstract plays a crucial role [18]. Therefore, in the process of translation, in addition to ensuring the transmission of basic information, attention should also be paid to overcoming the structural differences between the original language and the target language to ensure that the translated text is equivalent to the original meaning [19]-[21].

This study proposes a semantic difference mining and translation optimization method for medical paper abstracts that integrates semantic role annotation and CARS model. First, a large-scale corpus of Chinese and English medical paper abstracts is constructed to provide sufficient data base for the study. Second, the traditional CARS model is adapted to the characteristics of the medical field, and the rhetorical steps specific to medical papers, such as defining terms and elaborating theoretical foundations, are added to make it more suitable for the structural analysis of medical abstracts. Meanwhile, the improved LSTM-CRF semantic role labeling model is designed to enhance the recognition of medical terminology and complex semantic relations by introducing attention mechanism and feature fusion technology. On this basis, the difference characteristics of Chinese and English medical abstracts in terms of speech step distribution, semantic role configuration and information organization mode are systematically analyzed to provide theoretical basis for translation optimization. Ultimately, the semantic analysis results are integrated into the machine translation system to achieve more accurate and natural cross-linguistic conversion through multi-level semantic feature guidance.

II. CARS model

In his book Genre Analysis, Swales defines a discourse genre as “a series of communicative events with a specific communicative purpose that can be jointly identified by a discourse community”. Thus, discourse is a social activity with a communicative purpose that is inextricably linked to the way it is realized in the text, influencing and defining the choices of discourse in terms of content and linguistic style. Swales developed the CARS model of discourse structure analysis, which takes discourse steps and strides as a point of departure, for describing the rhetorical structure of the abstract section of an academic paper.

The CARS model consists of three discourse steps, i.e., establishing the field of study, establishing the research position, and occupying the research position, each of which consists of a number of steps [22]. Although the CARS model is universally applicable, the model has some limitations due to certain differences in the structure of abstracts of scientific papers from different disciplines. Therefore, some researchers have made appropriate modifications to the CARS model to make it more applicable to discipline-specific corpus analysis. The author also found some missing steps of the CARS model when studying the corpus of medical papers. For example, the rapid development of the medical field has given rise to a large number of emerging scientific and technological terms, and the conceptual system is constantly developing and enriching, and many abstracts adopt the steps of defining terms and elaborating theoretical foundations. As the research competition in the medical field is extremely fierce, some authors try to better attract readers and peers' attention through the step of elaborating the value and significance of the research. According to the above characteristics, the author makes appropriate amendments to the CARS model and applies it to this study. The CARS model for analyzing the genre of abstracts in the medical field is shown in Table 1.

Table 1: The CARS model of the medical field paper abstract

Language step	Language step content	Step	Step content
M1	Establish research	S1	Center issues
		S2	Summarize the topic content
		S3	Definition term
		S4	Elaboration theory
		S5	Retrospective study
M2	Establish research status	S1a	Counterargument
		S1b	Point out
		S1c	Ask questions
		S1d	Previous study
		S2	Make positive reasons
M3	Research position	S1a	Overview
		S1b	Report research status
		S2	Make assumptions
		S3	Report main discovery
		S4	Research significance

III. LSTM-CRF based semantic role labeling

(1) Modeling framework

LSTM network layer, CRF and post-processing are used to form the LSTM-CRF role labeling model during the experiment. The processed experimental corpus is input to the LSTM model for training to obtain semantic role labeling. In the process of model training, the first step is to vectorize the corpus to generate word vectors as inputs to the LSTM network layer, and to learn the feature representation of the input sequences through this network layer. Then, the output values computed by LSTM are used as input data for the CRF layer, and the correct semantic role labels are obtained after CRF and post-processing steps.The LSTM model framework is shown in Fig. 1.

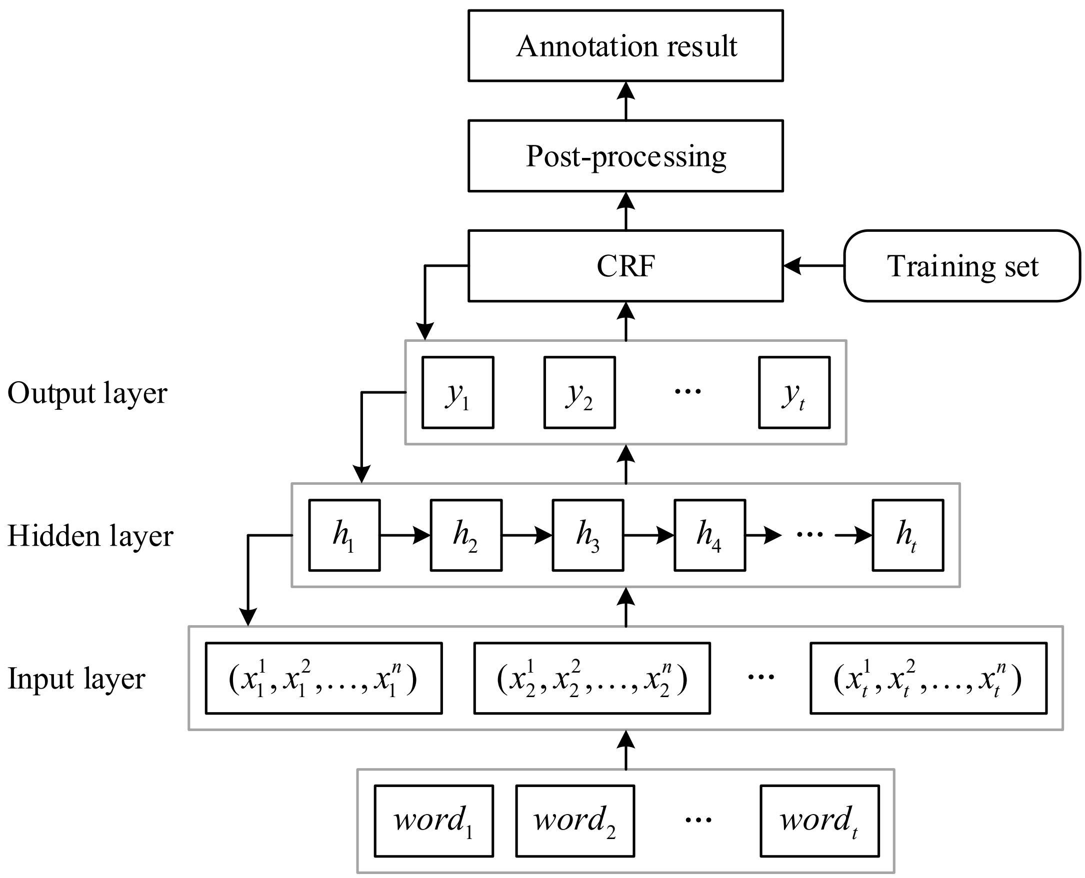

Figure 1: Framework diagram of the LSTM model

The word in the figure, represents the input data at the current moment t:

$$\{x^1, x_1^2, \ldots\ldots, x_1^n\} \tag{1}$$

represents the vector representation of the input data at the current moment t obtained after preprocessing, n denotes the dimension of this vector, t the output value of the network at the moment is represented by y_t, and H denotes the number of nodes in the hidden layer, and the three parts of the input layer, the hidden layer, and the output layer constitute the LSTM network.

(2) LSTM network layer

The structure of LSTM model is shown below in Fig. 2. The LSTM model [23] is improved based on RNN, and a memory unit is added on the basis of RNN model, which is able to fully utilize the inter-relationships between words in the whole text and the captured inter-relationship information is used in the feature expression processing for each word. The LSTM layer consists of two parts: the network construction and the network training.

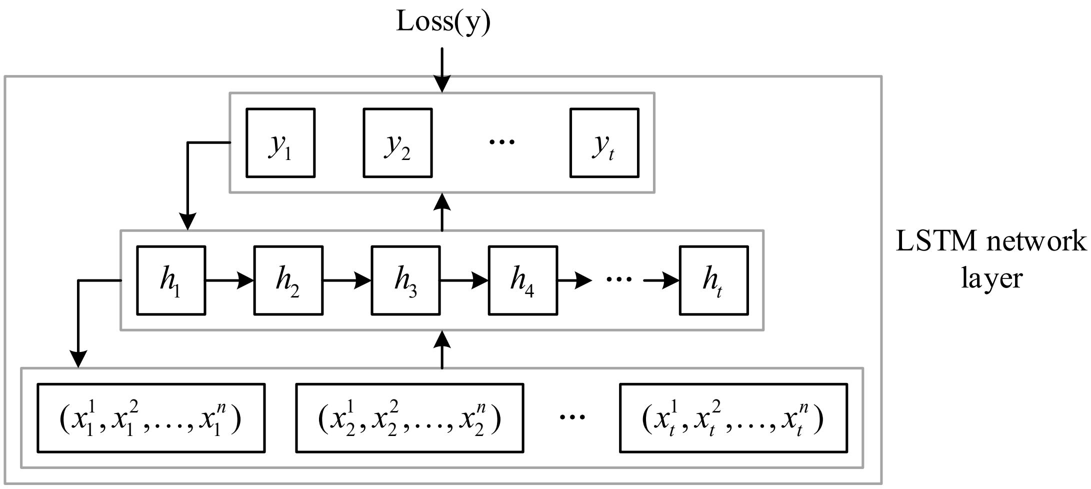

Figure 2: LSTM network Layer structure Diagram

The first step is to build the network, in this stage, the connection weights of each gate are initialized, and then the value 1 is assigned to the initial value of the memory state, this is to facilitate the subsequent network calculation. The formula for calculating the connection weights of each gate is as follows:

$$\tanh : W \sim Uniform(-\frac{\sqrt{6}}{\sqrt{n_{in}+n_{ou}}}, \frac{\sqrt{6}}{\sqrt{n_{in}+n_{ou}}}) \tag{2}$$

$$sigmoid : W \sim Uniform(-4\times\frac{\sqrt{6}}{\sqrt{n_{in}+n_{ou}}}, 4\times\frac{\sqrt{6}}{\sqrt{n_{in}+n_{ou}}}) \tag{3}$$

$$ReLU : W \sim N(0,1) \tag{4}$$

The formula n_{in} represents the number of input nodes in the network layer and n_{ou} represents the number of output nodes.

The words that have been vectorized are input in the order in which they appear in the sentence, and then the values of $b_i^t, b_f^t, b_c^t, b_o^t, b_h^t$ are computed using the following formula:

$$a_l^{'} = \sum_{i=1}^{l} w_{i_l} x_i^{'} + \sum_{h=1}^{H} w_{h_l} b_h^{t-1} + \sum_{c=1}^{C} w_{c_l} s_c^{t-1} \tag{5}$$

$$b_l^t = f(a_l^t) \tag{6}$$

$$a_\phi^{'} = \sum_{i=1}^{l} w_{i\phi} x_i^{'} + \sum_{h=1}^{H} w_{h\phi} b_h^{t-1} + \sum_{c=1}^{C} w_{c\phi} s_c^{t-1} \tag{7}$$

$$b_\phi^{'} = f(a_\phi^t) \tag{8}$$

$$a_c^t = \sum_{i=1}^{l} w_{ic} x_i^t + \sum_{h=1}^{H} w_{hc} b_h^{t-1} \tag{9}$$

$$s_c^t = b_\phi^t s_c^{t-1} + b_l^t g(a_c^t) \tag{10}$$

$$a_\omega^t = \sum_{i=1}^{l} w_{i\omega} x_i^t + \sum_{h=1}^{H} w_{h\omega} b_h^{t-1} + \sum_{c=1}^{C} w_{c\omega} s_c^{t-1} \tag{11}$$

$$b_\omega^t = f(a_\omega^t) \tag{12}$$

$$b_c^t = b_o^t h(s_c^t) \tag{13}$$

The last thing is to obtain the feature output of the LSTM layer, which is the value s_k^t of the output layer at each moment t obtained after into a while set a_p . Repeat the previous steps until all samples are fed into the input.

Once the network is constructed the network training process begins, that is, after all the data in the first sequence has been processed the training process of the network begins. The loss function of the model is first calculated by comparing the output of the network layer with the original labels of the words, and then the connection weights between the gates are updated, and the updating process is done in the direction of the decreasing loss function for the gradient learning of the model. The network training process first calculates a total loss function for the input data sequence, and then calculates the error before the output, hidden, and input layers of the model, and finally uses the combination of the error value of each layer and the learning rate to calculate the gradient of each unit for assigning a new connection weight to each layer. The error value of each layer is not only related to the output of the current moment, but also related to the output of the next moment, but the last input data of the model does not have the output value of the next moment, in order to be able to accurately calculate the last moment of the input data of the next moment is directly set to one.

The quantized text data is input to the LSTM network layer, and after processing, the feature vectors of the semantic roles of each word are obtained, and the input data of each moment will have an impact on the input data of the next moment, and the ap feature vectors obtained after processing of the LSTM network layer are then input to the CRF for sequence annotation, and then the annotated data will be post-processed into the correct semantic role labeling.

(3) CRF layer

Conditional Random Field is a probabilistic model for sequence annotation, which assumes that the output sequence is a Markov Random Field, i.e., the current state is only related to the previous state and not to the other states. Given an input sequence, the model is able to compute the output of the model based on this particular input sequence, and this output is the conditional probability distribution of the outgoing target sequence.The input of the CRF layer is the vector of LSTM outputs, which is labeled with the labeled sequences as the target vectors, and the final labeled sequences are obtained using the Viterbi algorithm, and the structure of the CRF is shown in Figure 3.

In the actual annotation process there may be more than one word is the same semantic role labeling, in this case if only use Arg0 these tags are not a way to accurately label semantic roles, the use of IOBES labeling strategy can provide an accurate labeling method. As can be seen from the above structure diagram, the labels output by CRF are containing prefixes B-, I-, etc. The beginning, middle and ending words of semantic roles are labeled using B-ArgX, I-ArgX, E-ArgX, respectively, and semantic roles consisting of a single word are labeled as S-ArgX, while non-semantic role labels are denoted with O.

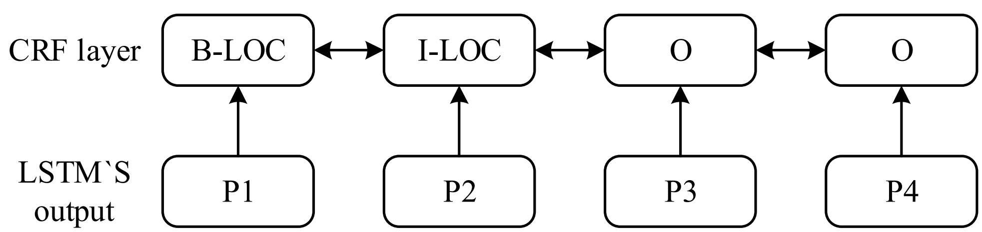

Figure 3: CRF Structure Diagram

IV. Experimental parameters and comparative modeling of translation performance

(1) Data set

The dataset used in this study comes from academic journal websites, and the abstract parts in Chinese and English medical papers were obtained by crawling software. The crawled thesis data contains 13,791 dissertations and journal papers, with a total corpus of 249,762 entries, which are used to mine the semantic expression differences between Chinese and English abstracts and provide a data base for translation optimization.

(2) Experimental Environment

All experiments in this study use hardware are trained on Tesla V100s 32GB GPUs compute card. PyTorch version 1.12.0 was used, Python version number 3.7, and PyCharm was used for program writing.

(3) Comparison Experiments

In order to verify the translation effectiveness of this paper's semantic role annotation combined with CARS model in medical papers, the translation performance of this paper's method is compared with other methods. The comparison methods include Prompt-Tuning, Prefix-Tuning, Infix-Tuning, Full-Tuning.

Prompt-Tuning: is a continuous prompting method that adds prompt vectors only in front of the input word embeddings. Freezes the parameters of the model and relies on the knowledge of the multilingual GPT model to adapt to the translation task.

Prefix-Tuning: is a continuous cueing method that adds a continuous cue before each layer of the model, optimizing a small continuous vector of the form [$Prefix$, x, y] of the form [$Prefix$, x, y] . Freezing the parameters of the model relies on the knowledge of the multilingual GPT model adapted to the translation task.

Infix-Tuning: similar to Prefix-Tuning, but placing consecutive cues in the middle of two sentences of the form [x, $Infix$, y]. Freezing the parameters of the model relies on the knowledge of the multilingual GPT model to adapt to the translation task.

Full-Tuning: use the Chinese-English legal parallel and pseudo-parallel corpus for fine-tuning, and input the corpus to the mGPT model for training by splicing the corpus into the form of [x <s> y].

(4) Evaluation Metrics

(1) BLEU (Bilingual Evaluation Understudy) evaluation metrics are a common method used in natural language processing to evaluate the quality of machine translation. Its core idea is to quantify the quality of machine translation by comparing the reference translation and the translation generated by the machine translation model, the formula is as follows:

$$BLEU = BP \times \exp(\sum_{n=1}^{N} w_n \log p_n) \tag{14}$$

where n denotes n meta-grammar, p_n denotes the proportion of phrases in n meta-grammar, w_n denotes the weighting factor, and N denotes the maximum grammatical order.

2) TER (Translation Error Rate), this metric measures the number of actions required to edit a translated passage that matches the reference translation. It is easy to use, language-independent, and corresponds to the amount of post-editing effort, as defined below:

$$TER = \frac{Edit\ \ frequency}{The\ average\ editing\ number\ of\ \ the\ translation} \tag{15}$$

The available editing operations include four: insertion, deletion, word substitution, and phrase panning, each at a cost of 1. Punctuation marks are also treated as words, and case inconsistencies are counted as errors.

V. Mining the difference between Chinese and English language steps in medical paper abstracts

In this study, based on the CARS model, the corpus of abstracts from the collected Chinese and English medical papers is fully utilized to realize a multi-layer comparative study of the abstracts of academic papers in Chinese and English journals by taking full advantage of the corpus's large text, multi-functionality, rigorous corpus selection, and the combination of quantitative and qualitative research. The collected corpus of medical papers was categorized into Chinese Journal Abstracts (CJ), English Journal Abstracts (EJ), Chinese Dissertation Abstracts (CD), and English Dissertation Abstracts (ED). The distribution characteristics of the speech steps in each corpus were counted, aiming at mining the differences between Chinese and English semantic expressions through the CARS model. The results of normalized frequency distribution of speech steps of medical dissertation abstracts are shown in Table 2.

The data of speech steps in the table are normalized. The results show that the performance of medical Chinese-English dissertation abstracts and journal paper abstracts is not consistent. First of all, as far as the dissertation is concerned, the frequency of both speech steps I and II of Chinese dissertation abstracts is higher than that of the corresponding English dissertation abstracts. Especially for step 2, the frequency of the former is 8124 times, while that of the latter is only 4016 times, which is more than twice as much as that of the latter. Second, as far as journal papers are concerned, the frequency of all three language steps in Chinese journal paper abstracts is less than that in English journals. In particular, although the difference between Chinese journal abstracts and English journal abstracts in the first two steps is nearly double, the difference is close to the proportion of the difference between the total number of steps in the two sub-corpora, which is due to the amount of text in the text itself. However, the frequency of using speech step three in Chinese journal paper abstracts is considerably less than that in English journal abstracts, with the latter exceeding the former by nearly four times.

The data show that in the semantic expression of medical abstracts, authors of Chinese dissertations or journals spend a great deal of time in the abstracts on “establishing the field of study” and “establishing the status of the study”. On the other hand, English dissertations, on the contrary, emphasize more on the specific research they have done.

Table 2: The distribution result of the normalized frequency of the language step

	Degree paper		Journal paper	
	CD	ED	CJ	EJ
Language step 1	19380	12874	5091	8445
Language step 2	8124	4016	1498	2339
Language step 3	7829	8413	1376	5219
Opening speech	29	44	16	50
Other	21	3	13	4
Total	35383	25508	7994	16057

VI. Abstract Semantic Role Annotation Performance and Semantic Component Recognition

In order to verify the effectiveness of the semantic role annotation model in this paper, the following annotation models are selected for comparison: the CNN-CRF model, the IDCNN-CRF model, the BERT-CRF model, and the RNN_CRF model. The experimental results of semantic role annotation performance testing of different models are shown in Table 3.

The experimental results show that the precision rate, recall rate and F1 value of semantic role annotation of this paper's model in the above medical paper dataset are better than those of the comparison models. The annotation precision rate, recall rate, and F1 value of this paper's model are 0.866, 0.871, and 0.868, respectively, and the precision rate is improved by 8% to 15.1%, the recall rate is improved by 10.6% to 15.7%, and the F1 value is improved by more than 9% at most compared with the comparison model. It indicates that the model in this paper is better able to grasp the global information and can effectively extract features compared to CNN_CRF, IDCNN_CRF, BERT_CRF, and BERT_CRF models. In addition, in terms of training time, the model in this paper is also significantly shortened compared to the comparison model, and the training can be completed in only 13 epochs. Combining the accuracy and efficiency of the models, it is concluded that the LSTM-CRF model proposed in this paper has the best performance in the medical paper dataset.

Table 3: Different model semantic role annotation performance test results

Model	Precision	Recall	F1	Time (epoch)
CNN-CRF	0.782	0.765	0.773	27
IDCNN-CRF	0.786	0.763	0.774	24
BERT-CRF	0.736	0.72	0.728	19
RNN-CRF	0.715	0.714	0.715	31
Ours	0.866	0.871	0.868	13

The richness of the features set in the experiment resulted in a large feature space and a low frequency of occurrence of some semantic components in medical paper abstracts. In this section of the experiment, semantic components with a frequency of more than 50 occurrences in the abstracts of Chinese and English medical papers were counted, including eight items: anatomical structure, pathological manifestations, symptoms and signs, examination indexes, pharmacological interventions, etiological factors, diagnostic conclusions, and prognostic descriptions. Using the LSTM-CRF model constructed in this paper, these semantic components are labeled and evaluated, and the results of the semantic component labeling and evaluation of medical paper abstracts are shown in Table 4.

As can be seen from the table, the recognition precision rate, recall rate and F1 value of anatomical structure, pathological manifestations, symptoms and signs, examination indexes, pharmacological interventions, etiological factors, diagnostic conclusions, and prognostic descriptions are all above 90%. The results show that for Chinese and English medical paper abstracts, the model role labeling in this paper has a high accuracy rate and a good recognition effect.

Table 4: The semantic component notes the evaluation results

Semantic component	Precision (%)	Recall (%)	F1 (%)
Anatomical structure	92.94	95.13	94.02
Pathological manifestation	95.85	90.36	93.02
Symptoms and signs	93.82	93.31	93.56
Inspection index	92.20	90.47	91.33
Drug intervention	94.45	90.48	92.42
Etiology	90.14	94.97	92.49
Diagnostic conclusion	95.26	94.47	94.86
Prognostic description	92.04	91.45	91.74

VII. Analysis of translation effects based on semantic role labeling

In order to test the translation effect of the English machine translation system incorporating semantic role annotation in medical paper abstracts, this section sets up a comparison experiment in the Chinese-English direction. The system is developed with a dependency tree-to-string based machine translation system as the base prototype, and the English syntactic analyzer is used to obtain the English syntactic analysis tree by the English syntactic analyzer developed by the lab itself.

In the experiment, we used 40,000 pairs of Chinese-English bilingual sentences collected above as the training set for this experiment, and randomly selected 10,000 pairs of Chinese-English bilingual sentences as the test set. For the translated sentences, scores were calculated using BLEU and TER on the translation results obtained from the test set. In order to test the effectiveness of the English-Chinese machine translation system incorporating semantic role features, the experimental results were compared and analyzed, and the comparison models included Prompt-Tuning, Prefix-Tuning, Infix-Tuning, and Full-Tuning.

Figure 4 demonstrates the experimental results of the comparison of the translation effect of medical dissertation abstracts. As can be seen from the figure, on the test set, the model incorporating the semantic role features of this paper has a higher BLEU than the comparison model by 5.47% to 7.50%, and the TET metrics are reduced by 0.257 to 0.452 compared with the comparison method, and the translation model incorporating the semantic role labeling features has a better overall translation performance. This is because the semantic role feature model fused with English semantic role features improves the ordering ability of syntactic structure tree. The experimental results show that the accuracy of English-Chinese machine translation is improved using the fused semantic role feature model.

Figure 4: The results of the translation were compared

VIII. Conclusion

In this study, through the combination of semantic role annotation and CARS model, we successfully mined the significant differences in semantic expression between Chinese and English medical paper abstracts, and realized the effective improvement of translation quality. The constructed LSTM-CRF semantic role labeling model shows excellent performance, with an F1 value of 86.8% on the medical paper dataset, which is improved by 9.5

percentage points compared with the traditional CNN-CRF model, while the training efficiency is significantly improved, and the model training can be completed in only 13 epochs. The results of the speech step analysis reveal the fundamental difference between Chinese and English abstracts in terms of rhetorical structure, with the Chinese dissertation abstracts focusing more on the elaboration of the research background, while the English abstracts highlight the description of the specific research content more prominently. The semantic component recognition experiments verified the practicality of the model, with the recognition accuracy of over 92% for eight categories of core medical concepts such as anatomical structure, pathological manifestations, symptoms and signs, including the diagnostic conclusions of which the recognition accuracy is as high as 95.26%. In the translation performance evaluation, the translation system incorporating semantic role features performs well, and the BLEU score is improved by 7.50% compared with the baseline model, effectively reducing the translation error rate. These experimental results fully demonstrate the important value of semantic role annotation technology in the optimization of medical literature translation, and provide a feasible technical path to improve the quality of international dissemination of medical academic achievements. The method can not only accurately identify semantic differences, but also effectively guide the translation process, which has good application prospects and popularization value.

Acknowledgements

The research was founded within the 2013004Srp project entitled Comparative Analysis of Chinese and English Abstracts of Medical Papers Based on CARS Model, which is a Key Research Project in Humanities and Social Sciences of Higher Education Institutions in Anhui.

The research was founded within the 2013004Srp project, titled “Comparative Analysis of Chinese and English Abstracts of Medical Papers Based on CARS Model.” which is a key research project in Humanities and Social Sciences of Higher Education Institutions in Anhui, supported by the Anhui Education Department.

References

[1] Saini, R. K., Issac, M., Kumar, K. J. D., Chaudhury, S., Sharma, R., Dangi, A., & Chail, A. (2023). Selecting a thesis topic: a postgraduate’s dilemma. Industrial Psychiatry Journal, 32(1), 19-23.
[2] Alspach, J. G. (2017). Writing for publication 101: Why the abstract is so important. Critical care nurse, 37(4), 12-15.
[3] Khatri, B. B. (2022). Writing an effective abstract for a scientific paper. Nepalese Journal of Development and Rural Studies, 19(01), 1-7.
[4] Letchford, A., Preis, T., & Moat, H. S. (2016). The advantage of simple paper abstracts. Journal of Informetrics, 10(1), 1-8.
[5] Andrade, C. (2011). How to write a good abstract for a scientific paper or conference presentation. Indian journal of psychiatry, 53(2), 172-175.
[6] Millar, B. C., & Lim, M. (2022). The role of visual abstracts in the dissemination of medical research. The Ulster medical journal, 91(2), 67.
[7] Fischer, I., & Steiger, H. J. (2020). Toward automatic evaluation of medical abstracts: The current value of sentiment analysis and machine learning for classification of the importance of PubMed abstracts of randomized trials for stroke. Journal of Stroke and Cerebrovascular Diseases, 29(9), 105042.
[8] Hwang, T., Aggarwal, N., Khan, P. Z., Roberts, T., Mahmood, A., Griffiths, M. M., ... & Khan, S. (2024). Can ChatGPT assist authors with abstract writing in medical journals? Evaluating the quality of scientific abstracts generated by ChatGPT and original abstracts. Plos one, 19(2), e0297701.
[9] Cohen, J. F., Korevaar, D. A., Gatsonis, C. A., Glasziou, P. P., Hooft, L., Moher, D., ... & Bossuyt, P. M. (2017). STARD for Abstracts: essential items for reporting diagnostic accuracy studies in journal or conference abstracts. bmj, 358.
[10] Hays, M., Andrews, M., Wilson, R., Callender, D., O'Malley, P. G., & Douglas, K. (2016). Reporting quality of randomised controlled trial abstracts among high-impact general medical journals: a review and analysis. BMJ open, 6(7), e011082.
[11] Parlak, B., & Uysal, A. K. (2020). On classification of abstracts obtained from medical journals. Journal of Information Science, 46(5), 648-663.
[12] Abdollahpour, Z., & Gholami, J. (2018). Rhetorical structure of the abstracts of medical sciences research articles. La Prensa Medica Argentina, 105(2), 1-5.
[13] Daniele, F. (2019). Performance of an automatic translator in translating medical abstracts. Heliyon, 5(10).
[14] Martikainen, H. (2017). A functional approach to translation quality assessment: Categorizing sources of translational distortion in medical abstracts. Linguistica Antverpiensia, New Series–Themes in Translation Studies, 16.
[15] Kyrychuk, L. (2019). STRATEGIC APPROACHESTO ACADEMIC PAPER ABSTRACT TRANSLATION. Research Trends in Modern Linguistics and Literature, 2, 65-76.
[16] Ling, X. I. A., Xiaoli, X. I. A. N. G., Yiman, L. I., Hongli, W. U., Yibo, L. I., & Shuang, L. I. U. (2024). Current application status of machine translation in medical paper abstracts and collaborative pathway between authors and editors. Chinese Journal of Scientific and Technical Periodicals, 35(12), 1757.
[17] Martikainen, H. (2018). Modal Markers as Potential Sources of Distortion in Translated Medical Abstracts. Educational Sciences: Theory and Practice, 18(4), 917-934.
[18] Kuzmina, O. D., Fominykh, A. D., & Abrosimova, N. A. (2015). Problems of the English abbreviations in medical translation. Procedia-Social and Behavioral Sciences, 199, 548-554.
[19] Zang, D., & Liu, C. (2023). Exploring the clinical translation intensity of papers published by the world’s top scientists in basic medicine. Scientometrics, 128(4), 2371-2416.
[20] Weikang, L. (2019). Review of previous studies on the translation of ABSTRACT of traditional Chinese medicine papers. International Journal of New Developments in Engineering and Society, 3(1).

[21] Liu, K., Yin, H., & Cheung, A. K. (2025). Interactional metadiscourse in translated and non-translated medical research article abstracts: a corpus-assisted study. Perspectives, 33(3), 542-562.

[22] Hijjatul Qamariah & Sri Wahyuni. (2017). HOW A RESEARCH ARTICLE INTRUDUCTION STRUCTURED?THE ANALYSIS OF SWALES MODEL (CARS) ON ENGLISHRESEARCH ARTICLE INTRODUCTIONS. GEEJ (Getsempena English Education Journal),4(2),

[23] Imen Jarraya,Safa Ben Atitallah,Fatimah Alahmed,Mohamed Abdelkader,Maha Driss,Fatma Abdelhadi & Anis Koubaa. (2025). Corrigendum “SOH-KLSTM: A hybrid Kolmogorov-Arnold network and LSTM model for enhanced lithium-ion battery health monitoring” [journal of energy storage, volume 122 (2025), 116,541]. Journal of Energy Storage,124,116868-116868.