Analysis of Chinese Speech Adaptive Translation Model Combining Deep Learning Technology and CAD Technology

Hua Xia[1] and Bin Li[2]

[1]School of Literature and Education, Anhui Open University, Hefei, Anhui 230022, China, xiahua@ahou.edu.cn
[2]School of Information and Architecture, Anhui Open University, Hefei, Anhui 230022, China, hfbinlee@126.com

Corresponding author: Hua Xia, xiahua@ahou.edu.cn

Abstract. The environment for training and recognition in Chinese speech recognition under computer-aided design may vary due to the difference of channel and background noise. When the trained model cannot well represent the test data, the recognition rate of the system will drop sharply. The computer-aided design method focuses on using a small amount of Chinese voice data to improve the performance of the system in the test environment. In this paper, we choose the BiLSTM CRF word separation model under deep learning as the improved benchmark model, and combine the Bert language pre-training module to enhance the performance of Chinese word separation. Combining the deep learning sample transfer learning theory and the improved sampling strategy, an adaptive translation model for intelligent Chinese domain is constructed. The experimental results show that Bert Chinese word segmentation model is superior to other word segmentation models in different data sets and has the best word segmentation performance, which can provide reliable support for the application experiment of this model. The test results show that this method can achieve high speech recognition accuracy and good application results.

Keywords: translation methods; intelligence; Chinese language; domain adaptation; deep learning; computer aided design
DOI: https://doi.org/10.14733/cadaps.2023.S10.1-12

1 INTRODUCTION

After nearly 50 years of research, computer-aided Chinese speech recognition teaching has made great progress. Since the 1980s, the current computer-assisted Chinese speech recognition technology has replaced the traditional template matching technology, making a breakthrough in speech recognition. Liang et al. [1] proposed a new objective function analysis model with a design level. An optimization framework for auxiliary polymers was defined, and the required property polymers were subjected to linear analysis, resulting in optimization results. However, in practical applications, the training and recognition environment often varies due to the difference of channel and background noise. Or the trained model cannot well represent the test data, resulting in a sharp decline in the recognition rate of the system. The most direct solution is to re-collect data to train all models in a new environment or for a specific speaker, but this requires collecting and processing a large amount of data, which is time-consuming and impractical. Wolf et al. [2] conducted ergonomic product design under model framework constraints. In the user's product interaction analysis, a digital human body model was tested. However, the Chinese language possesses strong domain relevance, which can reduce the performance of neural network word separatio. Although this problem can be alleviated by taking a large range of data labeling for different domains, this will undoubtedly increase the resource consumption significantly, which is not proportional to the rewar.

Barreto et al. [3] proposed an editor-based management output model. A simple and centralized analysis of the computer-assisted software architecture was conducted. The proposed method is based on virtual reality for re layout and design. The feature extraction of speech signal does not use conventional methods, but directly uses the spectrum diagram as the model input. Referring to the best network configuration model in image recognition, the depth convolution neural network V GG, can see the long-time correlation information. Compared with RNN (Recurrent Neural Network), it has better robustness. At the output end, it can be combined with the connected time series classification (CTC) scheme to realize the training of the whole model from the input end to the output end. The Industrial Revolution (IR) is mainly related to how changes in enterprise operations affect the manufacturing systems of the enterprises. [4]. The sound waveform signal is translated into the pinyin sequence of standard Chinese. In the language model, the pinyin sequence is converted into Chinese text through hidden Markov model. Set up protocol servers such as HTTP, provide voice recognition application program interface API, and STM32MCU embedded client calls API to realize voice recognition function through Internet network. Digitalization has invaded all fields of human activities, including innovation. A more substantive innovation than previous generations of tools. Marion and Fixson [5] studied the digitalization of new product development (NPD), a subset of innovation.

Therefore, many scholars have conducted research and achieved certain results for adaptive methods in Chinese language domain, but the current research system in this direction is still in the development stage, which is not mature and complete enough, and many methods have limited usefulness. In this paper, we will improve on the basis of previous research results and the mainstream Chinese language adaptive translation framework by incorporating a Bert language pre-training module in the BiLSTM-CRF word separation model to improve the performance of Deep learning for Chinese language word separation. On this basis, this paper also combines sample migration theory and sampling strategy to construct an intelligent adaptive translation method for Chinese language, and validates it by corresponding experiments.

2 RELATED WORK

The traditional teacher-student teaching method is crucial and effective in Chinese phonetic translation, but it is not enough. Computer assisted instruction is a useful supplement to this teaching method, which is not limited by time and place compared with the former. Different learning strategies can be formulated according to different people. However, due to the limitations of their own level, it is difficult for self-learners to find errors and correct incorrect pronunciation completely by themselves. At the same time, the rapid development of computers has brought a new technology, called computer-assisted language learning, to help people learn languages. The initial computer-aided learning was mainly applied to the training of writing ability and understanding ability. With the development of computer voice technology, more and more researchers begin to pay attention to pronunciation learning. At present, robust speech recognition and speaker adaptive technology are very active topics in the field of speech recognition. Large vocabulary keyword recognition algorithm, speech recognition confidence evaluation algorithm. Class-based language model and adaptive language model, as well as deep learning of natural speech understanding. The research direction is also more and more focused on the oral dialogue system. At present, the research of speaker adaptive technology has made considerable progress, and some mature technologies have emerged. Such as deep learning technology, maximum likelihood linear regression algorithm, Bayes adaptive estimation algorithm. The current research focuses on how to realize online unsupervised learning and multi-method comprehensive adaptive learning algorithm. Language model is also an important aspect of current research. How to combine statistical language model with knowledge-based language model is one of the focuses of its research. In short, the research of deep learning speech recognition is developing in depth and breadth.

At the early stage of translation subdivision research, some scholars made standard matching of word sequences in utterances according to linguistic rules or customary rules that could be matched for subdivision operations, and the cut-up method exhibited faster speed and efficiency, easy modification and flexible operation, but with high dependence on dictionaries and rule systems. Other scholars point out that the rule-based syncopation method requires manual construction of the lexicon system, but the update speed cannot keep up with the speed of new word generation, and the rule framework cannot adapt to the translation of multiple meanings. Azman et al. [6] and reverse the element development logic of product form design. The manufacturing of engineering fibers under different composite materials is described. Xue et al. [7] proposed fuzzy rule core task image processing. The employment function fuzzy theory rules for model design are determined. Computer aided task flow is performed by associating product elements with image models. Jones et al. [8] proposed a conceptual matching topology with mutual participation of resource interests. Use the conceptualized type matching of learning resources for advanced technology available resource structure matching. Papakostas et al. [9] proposed a conceptual framework for data processing through functional development of product systems developed using additive technology. Dimitrijevi ć Et al. [10] conducted an objective system evaluation of clothing enterprises under the computer optimal model. The development and research content of the CAD system was determined, and important experimental processes were determined through effective parameter analysis in the new model ecosystem.

3 CONSTRUCTION OF INTELLIGENT CHINESE LANGUAGE ADAPTIVE TRANSLATION MODEL

3.1 Word Separation Method for Chinese Language Combined with Deep learning

Aiming at the speed of BISTM model, the application of extended convolution neural network in Chinese word segmentation is discussed. Specific research contents include: exploratory experiments around the sequence annotation model BILSTM+CRF. In Chinese word segmentation, if we only consider the single characteristics of characters, we cannot understand the meaning of characters at a deeper level. In this paper, the root information is used in Chinese word segmentation to extract the semantic information of Chinese root. In this paper, we use a method that combines root feature and character feature. The set of Chinese words is written as $L=\{B,M,E,S\}$, and the words in the set are classified and grouped into sentence head, sentence middle, sentence end and individual words. Let the statement of sentence length a be $N=\{n_1,n_2,...,n_a\}$, and its optimal annotation sequence be $M^*=\{m_1^*,m_2^*,...,m_a^*\}$, which can be solved according to Equation (3.1).

$$M^* = ard\max_{M\in L^a} p(M|N) \tag{3.1}$$

where M and N are the probability expressions of the corresponding label sequences as $p(M|N)$. The labeled sequences of the word separation model obtained from the neural network training are given a score i.e. $s(x,y)$, and the corresponding probabilities are obtained after normalizing the scores as shown in Equation (3.2).

$$p(M|N)=\exp(s(x,y^{'}))/\sum_{y^{'}}\exp(s(x,y^{'})) \tag{3.2}$$

The key link of the method for Chinese word separation is the correspondence between the input sequence and its tag sequence. The most applied word separation model currently has some shortcomings, and this paper will improve on the basis of the BiLSTM-CRF word separation model, as Figure 1 shows the flow chart of the neural network Chinese word separation framework.

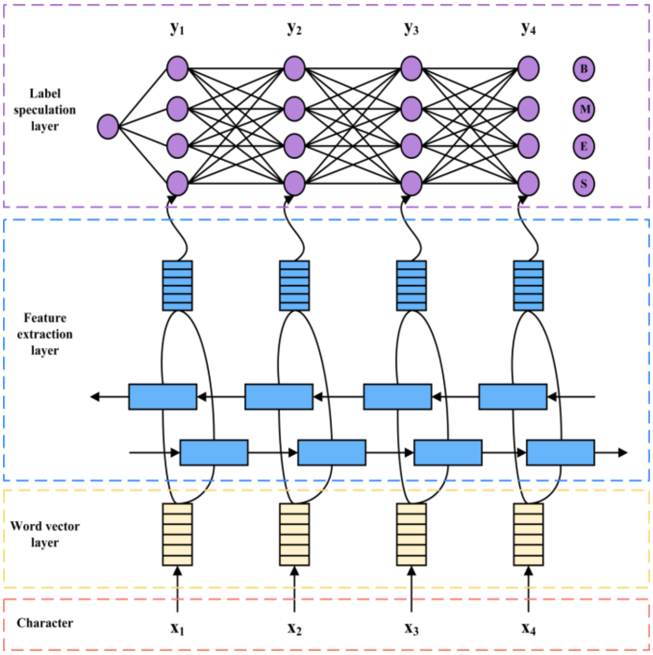

Figure 1: Flowchart of the neural network Chinese word separation framework.

The figure shows that Chinese language word separation by neural network model will be performed first by transforming Chinese characters into distributed vector input before feature extraction. The neural network node connections are developed between layers, but the nodes in the same layer do not generate connections, which is not conducive to handling the relationship problem between Chinese characters. In this paper, the LSTM model structure is used to make connections between Chinese characters in the model. the threshold recursion mechanism of the LSTM model enables the selection of information retention and forgetting, and realizes that dependent connections can be maintained over a longer period of time. The structure adds modules for forgetting gates and memory retention in addition to output and input ports, and the first three can be realized by equation (3.3).

$$\sigma(n)=1/(1+e^{-n}) \tag{3.3}$$

The output value of $\sigma(n)\in R[0,1]$, the minimum value, means that no message can pass, and the maximum value means that any message can pass.

The forgetting gate can determine the amount of information from the previous moment of the memory module that is present in the current time information in the memory module, and its formula is shown in (3.4).

$$f_t=\sigma(W_f\cdot[h_{t-1},n_t]+b_f) \tag{3.4}$$

Where n_t expresses the input information and the previous moment's information in the hidden layer is noted as h_{t-1} . The new information added to the current memory module requires the input gate to issue the decision of information update and then obtain the alternative memory module by virtue of the calculation. The input gate formula is shown in (3.5).

$$a_t = \sigma(W_a \cdot [h_{t-1}, n_t] + b_a) \tag{3.5}$$

The present alternative memory module calculation formula is shown in (3.6).

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, n_t] + b_C) \tag{3.6}$$

One of the alternative memory modules is noted as $\tilde{C}_t$.

The formula for calculating the memory module at this moment is shown in (3.7).

$$C_t = f_t * C_{t-1} + a_t * \tilde{C}_t \tag{3.7}$$

The output gate calculation equation is shown in (3.8).

$$O_t = \sigma(W_O \cdot [h_{t-1}, n_t] + b_O) \tag{3.8}$$

The final result obtained by the model is the output gate and cell state determined together according to Equation (3.9).

$$h_t = O_t * \tanh(C_t) \tag{3.9}$$

The Chinese language translation process requires context-based translation, so the LSTM model adopts a bi-directional structure, i.e., Bi-LSTM, which can perform input processing from the beginning to the end of a sentence and from the end to the beginning of a sentence simultaneously.

After extracting the features need to speculate on the label sequence, the original framework of this paper based on the introduction of the CRF layer as a constraint to avoid some errors, the probability calculation in this layer is shown in Equation (3.10).

$$p(M|N) = e^{s(N,M)} / \sum_{\tilde{M} \in M_N} e^{s(N,M)} \tag{3.10}$$

$$s(N, M) = \sum_{a=0}^{i} Z_{M_a, M_{a+1}} + \sum_{a=1}^{i} P_{a, M_f} \tag{3.11}$$

where the model output label score matrix is denoted as P and the transfer matrix is denoted as Z .

3.2 Adaptive Translation Methods for Chinese Language Domains

Domain adaption is a special case that arises based on migration learning. There is a certain correlation between knowledge and knowledge, and new knowledge that can be understood and learned by the already learned knowledge, and migration learning is to use this to transfer the already existing knowledge models, information to the target knowledge. Compared with previous machine learning, the data for training and testing of migration learning do not follow the same distribution, but are different. The learning in the new domain is also applied from the old domain based on the similarity between data, training, and models, as shown in Figure 2 for a comparison of previous machine learning and migration learning processes. In this paper, the TrAdaboost algorithm in migration learning is selected according to the requirements, which obtains migration learning performance with the update of weights by Boosting technique.

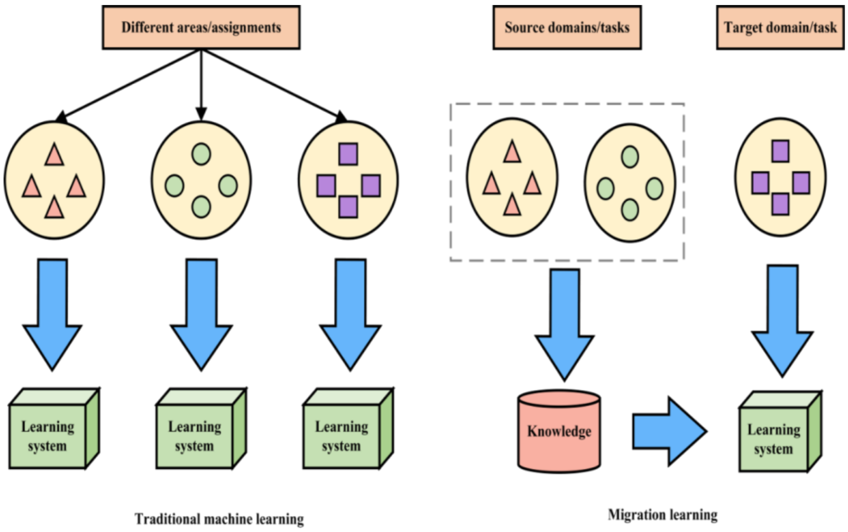

Figure 2: Comparison of previous machine learning and migration learning processes.

Among the domain adaptive methods, the TrAdaboost algorithm and the two methods that achieve the purpose by differential selection have been relatively well studied, and both have been shown to be effective in improving the domain adaptiveness of translation models through experiments. However, there is a large gap between the effect of the former in Deep learning and its effect in statistical translation models. The latter, on the other hand, will achieve the purpose of increasing the word separation effect in a short period of time at the initial application, but the great distribution difference exhibited between the selected target and source regions should have a destructive effect on the data distribution to a greater extent, causing the Chinese word separation effect to deteriorate. Therefore, this paper adopts the approach of Chinese sentence similarity calculation based on X meta-vectors inspired by both approaches to enhance the validity of data selection and strengthen the coverage of unregistered word samples.

The meanings of the same word in Chinese language in different domains may have extended meanings suitable for that domain, or some words have specific adaptive environments that rarely occur in other domains, both of which largely affect the effect of machine translation on the Chinese language's word separation. The approach adopted in this paper requires first obtaining some data with value in the target domain and labeling them, and incorporating them into the training dataset to enhance the domain adaptation of the translation model. The important aspect of this process is to avoid the degradation of machine translation performance. The method in this paper obtains the required data according to the similarity between data, and then selects the sample sampling strategy with similar source area data characteristics and the maximum randomness of annotation by combining the randomness composition of data annotation.

4 EXPERIMENTAL RESULTS AND ANALYSIS ON THE APPLICATION OF INTELLIGENT CHINESE LANGUAGE ADAPTIVE TRANSLATION MODEL

In the experimental application of the intelligent Chinese language adaptive translation model, the performance of the model can be evaluated by the following three indicators, namely, accuracy, recall and F-value, the first two of which can be mutually constrained and the higher the value the better, while F-value is a comprehensive measure of the quality of word annotation.

$$P = c / (c + e) \tag{4.1}$$

$$R = c / N \tag{4.2}$$

$$F = (\beta^2 + 1) \times P \times R / (\beta^2 \times P + R) \tag{4.3}$$

The N number of correct words is c and the number of incorrect words is e . Accuracy and recall are expressed as P and R , respectively, and β expresses the focus of F-values on the other two metrics.

4.1 Experimental Results of Word Separation Performance of Chinese Language Combined with Deep Learning

The BiLSTM-CRF word separation model is used as the benchmark for experimental comparison, and the Bert Chinese word separation model will be introduced in this paper to compare the Chinese word separation performance with the BiLSTM-CRF word separation model that incorporates the attention mechanism and direct vector splicing. The experimental data results of different Chinese word separation models in the CTB6 dataset are shown in Figure 3. It can be seen that several other word separation models, whether binary or ternary vectors, show higher word separation results than the benchmark model, which indicates that linking historical data and adding semantic content in the benchmark model is an effective way to improve the word separation effect of the model in the Chinese language system. The experimental values of the model with direct splicing are lower than those of the model with the attention mechanism, regardless of how many elements the vectors belong to, compared with the model with the attention mechanism. However, it should be noted that the ternary vector results of the attention mechanism model improve less than the binary vector with the benchmark model. The experimental results of the Bert Chinese word separation model taken in this paper perform the best among all models and show a strong Chinese word separation performance.

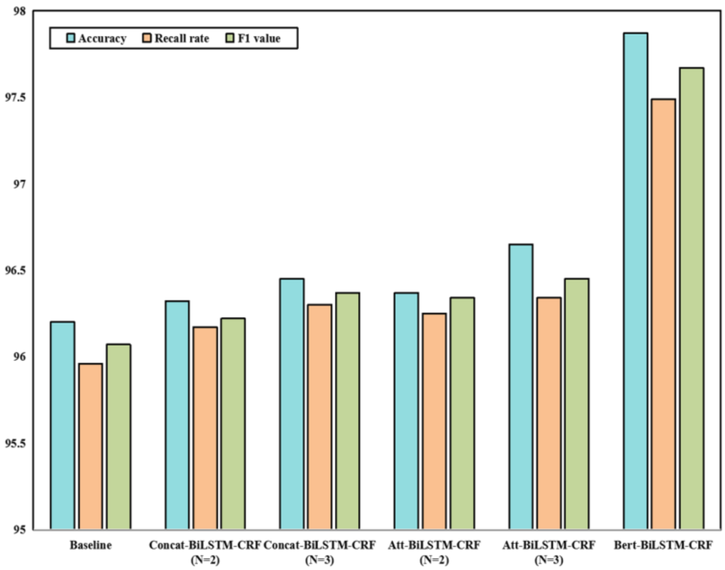

Figure 3: Experimental data results of different Chinese word separation models in the CTB6 dataset.

In order to verify the effectiveness level of the word separation performance of the models adopted in this paper, the word separation results of different Chinese word separation models in the MSRA dataset are shown in Figure 4. The results show that the word separation results of the other three models are higher than those of the benchmark model, the performance of the direct splicing model in the ternary vector is higher than that of the attention mechanism model, and the word separation effects between the word separation models in this paper and the other models show a certain disparity. This indicates that the improved Chinese language word separation model in this paper can improve the accuracy of Chinese language analysis in a targeted way while maintaining the advantages of the benchmark model, and can provide a reliable data base and basis for subsequent experiments.

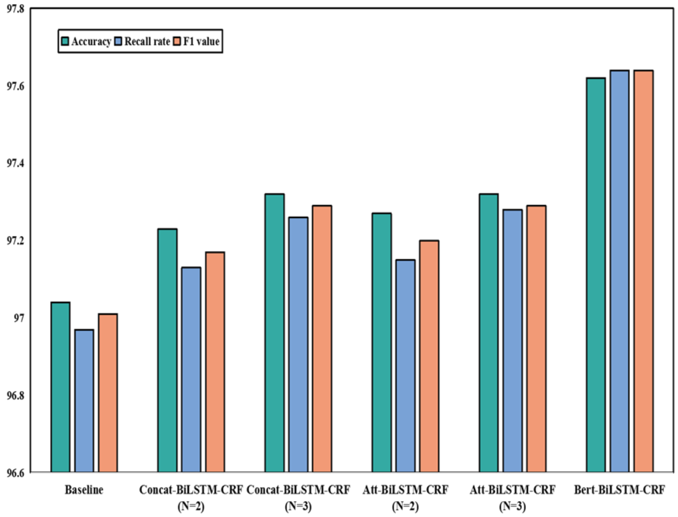

Figure 4: Results of different Chinese word separation models in MSRA dataset.

4.2 Experimental Results of Adaptive Translation Method in Chinese Language Domain

The rise of social network platforms, novel culture, animation culture and other domains provide training datasets for intelligent Chinese translation, and also show the challenges that intelligent Chinese translation models need to face, and the Chinese sub-domains in the news domain will not be the only basis for sub-domains. This part of the experiment will include two parts of domain adaptive translation model for Chinese language domain, namely, news domain to social domain and news domain to novel domain. The data sources in the news domain are mainly the People's Daily datasets, the social platform datasets are from Weibo and Zhihu datasets, and the novel domain datasets are selected from novel texts that have been paid more attention in recent years. Before conducting the experiments, the unlabeled data in the social domain and the novel domain were obtained according to the sampling strategy to obtain the top four hundred and fifty Chinese sentences, and then labeled and corrected into the training dataset for iterative training. The experimental results from the news domain to the social domain are shown in Figure 5.

The experimental results show that the comprehensive index F value of Chinese word separation is 90.46% when sample migration has not yet occurred, and the increase of each iteration after the model improves the value of F. This indicates that the measures taken by the translation model in this paper to improve the domain adaptation of Chinese word separation are effective. The increase of samples after several iterations did not affect the improvement of the system's word separation performance, but the improvement of word separation effect kept slowing down. This reflects that the samples obtained according to the sampling strategy are of certain value and do not cause negative sample migration, maintaining the distribution of the initial training data. The number of iterations in the experiment is limited, and if we continue to increase the number of samples and iterations, the adaptive performance of the translation model in the Chinese language domain will remain somewhat enhanced. In addition, the word recall rate within the dataset from news domain to social domain also increased after the iterations, especially the recall rate of unlogged words increased more. It reflects the high accuracy of the model's Chinese language cut-off effect in the in-set words of the target domain and no negative impact on the original distribution of the training data. The increase in the recognition rate of unlanded words indicates its enhanced performance in generalization ability and improved domain adaptive ability.

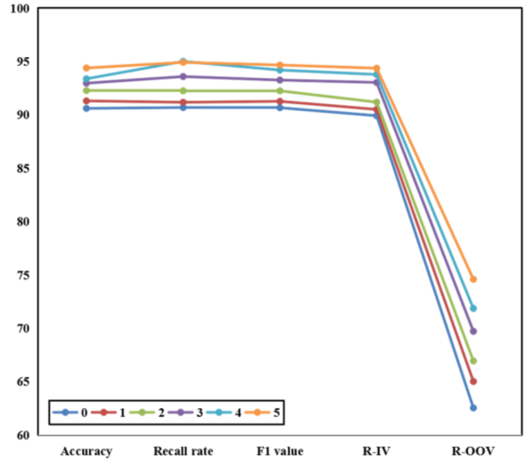

Figure 5: Experimental results of news domain to social domain.

The results of the experiments from the news domain to the novel domain are shown in Figure 6. The novel domain has a smaller dataset than that of the social domain due to the limitation of text content length, but the same four hundred and fifty target domain data are selected for the annotation process. The results show that the adaptivity of news domain to fiction domain also increases with the number of iterations and the number of samples, and its recall rate of unlanded words is 1.49% higher than the increase of news domain to social domain, which shows better performance.

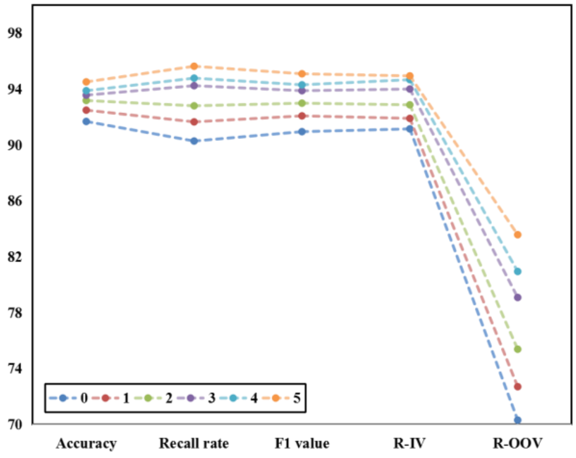

Figure 6: Experimental results of news domain to fiction domain.

In this paper, we also add comparison experiments to verify the effectiveness of the intelligent Chinese language adaptive translation approach by diversifying the objects of the comparison experiments, which are random sample selection, the translation model with the presence of supervised approach and other models. As shown in Figure 7, the results of the comparison experiment between Bert Chinese word separation model and random samples are shown. The comparison experiment is to verify the effectiveness of the model sampling strategy, and the results show that the sampling strategy of Bert Chinese word separation model shows an upward trend in the effect of its Chinese word separation performance under the increasing sample size, i.e., the way maintains a continuous selection of samples with higher value. There is a decline in the curve of random sample selection, i.e., it causes a negative migration problem, and the overall performance is inferior to that of the Bert Chinese word separation model.

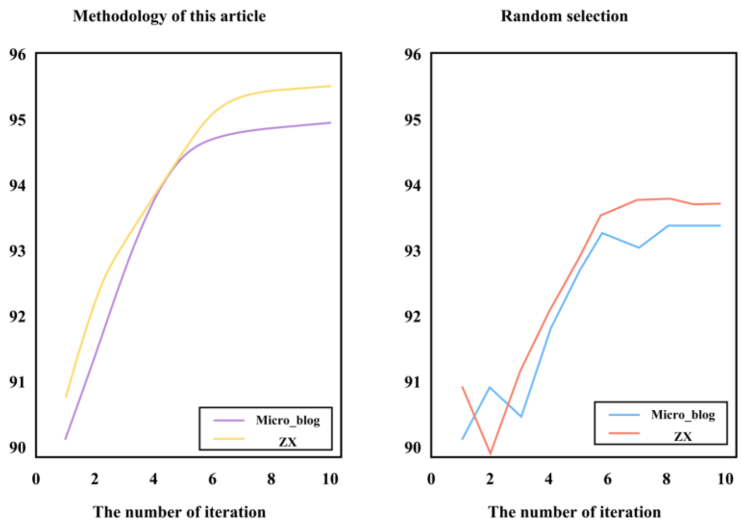

Figure 7: Experimental results of Bert's Chinese word separation model and random sample comparison.

5 CONCLUSION

Chinese word segmentation is a basic task in natural language processing technology. Traditional machine learning methods have achieved good results in Chinese word segmentation, and deep learning technology has developed rapidly in recent years. Many natural language processing tasks have made breakthroughs after adopting deep learning technology. Deep learning avoids manual extraction of data features, and deep network structure can extract more abstract features from a large amount of data. In view of the existing problems of Chinese word segmentation, this paper focuses on the research of Chinese word segmentation using the deep learning method. In this paper, Bert pre-training language module is introduced into the basic framework of the BiLSTM CRF word separation model to build a two-layer neural network to achieve the acquisition and connection of Chinese context information. Combining the sample transfer learning theory with the improved sampling strategy, an intelligent Chinese adaptive translation method is constructed. The performance of Chinese word segmentation is verified before the experimental test of model translation application. The validation results show that the Bert Chinese word segmentation model has higher segmentation effect than the benchmark model, and has improved segmentation effect in different data sets. Compared with other word segmentation methods in the experiment, it has certain advantages, and can provide reliable data support for subsequent experiments. In the comparative experiment, the intelligent Chinese adaptive model shows a good ability to move forward, and its comprehensive index F value and recall rate continue to increase with the expansion of the number of iterations and sample size. Whether it is from the news field to the social field or from the news field to the news field, it has enhanced the generalization ability and domain adaptive ability of the model. In other comparative experiments, the experimental results show that the model in this paper can maintain the distribution characteristics of the source area data, save data label consumption resources, and show better adaptive effect. However, the types of experimental data samples in this paper are limited, and experiments of applying models in different data sets should be added to further refine and improve the performance of translation models.

6 ACKNOWLEDGEMENTS

This work was supported by the Key Program of the Teaching Reform Research Anhui Province under Grant 2021jxjy020; The Key Program of the Foundation for Young Talents in the Colleges and Universities of Anhui Provincial under Grant GXYQ2021042; The Credit Bank Research Team of Anhui Open University under Grant 2022KYTD02; The Key Program of the Scientific research project of colleges and universities in Anhui Province under Grant 2022AH052676.

Hua Xia, https://orcid.org/0000-0001-7878-7585
Bin Li, https://orcid.org/0000-0002-2085-0024

REFERENCES

[1] Liang, X.; Zhang, X.; Zhang, L.; Liu, L.; Du, J.; Zhu, X.; Ng, K.-M.: Computer-aided polymer design: Integrating group contribution and molecular dynamics, Industrial & Engineering Chemistry Research, 58(34), 2019, 15542-15552. https://doi.org/10.1021/acs.iecr.9b02769

[2] Wolf, A.; Wagner, Y.; Oßwald, M.; Miehling, J.; Wartzack, S.: Simplifying computer aided ergonomics: A user-product interaction-modeling framework in CAD based on a taxonomy of elementary affordances, IISE transactions on occupational ergonomics and human factors, 9(3-4), 2021, 186-198. https://doi.org/10.1080/24725838.2021.1941433

[3] Barreto, J.-C.-L.; Cardoso, A.; Lamounier, J.-E.-A.; Silva, P.-C.; Silva, A.-C.: Designing virtual reality environments through an authoring system based on CAD floor plans: A methodology and case study applied to electric power substations for supervision, Energies, 14(21), 2021, 7435. https://doi.org/10.3390/en14217435

[4] Pereira, P.-M.-V.; Jauregui, B.-J.-M.: Smart design engineering: a literature review of the impact of the 4th industrial revolution on product design and development, Research in Engineering Design, 31(2), 2020, 175-195. https://doi.org/10.1007/s00163-020-00330-z

[5] Marion, T.-J.; Fixson, S.-K.: The transformation of the innovation process: How digital tools are changing work, collaboration, and organizations in new product development, Journal of Product Innovation Management, 38(1), 2021, 192-215. https://doi.org/10.1111/jpim.12547

[6] Azman, M.-A.; Asyraf, M.-R.-M.; Khalina, A.; Petrů, M.; Ruzaidi, C.-M.; Sapuan, S.-M.; Suriani, M.-J.: Natural fiber reinforced composite material for product design: A short review, Polymers, 13(12), 2021, 1917. https://doi.org/10.3390/polym13121917

[7] Xue, L.; Yi, X.; Lin, Y.-C.; Drukker, J.-W.: An approach of the product form design based on gra-fuzzy logic model: A case study of train seats, International Journal of Innovative Computing, Information and Control, 15(1), 2019, 261-274. https://doi.org/10.24507/ijicic.15.01.261

[8] Jones, B.; Hildreth, D.; Chen, D.; Baran, I.; Kim, V.-G.; Schulz, A.: Automate: A dataset and learning approach for automatic mating of cad assemblies, ACM Transactions on Graphics (TOG), 40(6), 2021, 1-18. https://doi.org/10.1145/3478513.3480562

[9] Papakostas, N.; Newell, A.; Hargaden, V.: A novel paradigm for managing the product development process utilising blockchain technology principles, CIRP Annals, 68(1), 2019, 137-140. https://doi.org/10.1016/j.cirp.2019.04.039

[10] Dimitrijević, D.; Spaić, O.; Đurić, Ž.; Urošević, S.; Nikolić, M.: CAD/CAM system implementation criteria in the process generating of optimal and efficient models for clothing industry, Industria Textila, 71(5), 2020, 467-472. https://doi.org/10.35530/T071.05.1741