

## Research Article

# Research on Business English Translation Architecture Based on Artificial Intelligence Speech Recognition and Edge Computing

Yunwei Xu 

Nanchang Jiaotong Institute, Nanchang 330100, China

Correspondence should be addressed to Yunwei Xu; 07054@ecjtu.edu.cn

Received 2 February 2021; Revised 15 April 2021; Accepted 5 July 2021; Published 18 August 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yunwei Xu. 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.

In today's society, the continuous deepening of international cultural integration has become the background of the times. China has become more and more closely connected with the world, and many physical or online news media have become a platform for China to receive world information and spread Chinese culture. Business English translation is therefore valued by translation researchers and translators. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing. First of all, considering the relevance and complementarity between speech and text modalities, this paper uses the deep neural network feature fusion method to effectively fuse the extracted monomodal features and perform speech recognition. Secondly, adopt the edge computing method to establish the business English translation system architecture. Finally, the simulation test analysis verifies the efficiency of the business English translation framework established in this paper. Compared with the existing methods, our proposal improved the accuracy than others at least 10% and the time of model building also decreased obviously. The purpose of this research is to discuss how to deal with the many differences between the source language and the target language, and how to enhance the readability of the translation and meet the reader's cultural cognition and needs.

## 1. Introduction

Recently, international business activities in China have been increasing for many years. A deep understanding of the business translation theory has important practical significance and principles [1, 2]. From the actual situation of business translation, there are many difficulties in understanding and expression. International business translation has received increasing attention as communication tools. In the foreseeable future, businessmen should be able to communicate directly with their foreign partners in English [3, 4]. Government officials and decision makers, who are not businessmen, should be able to review business documents in English. In short, business translation in the era of globalization is very necessary and important [5, 6].

Although people are still too focused on specific translation skills and detailed functional applications, there is still a lack of systematic theoretical translation practices that can guide business [7, 8]. Business translation can provide a fertile ground for translation theory and provide a lot of help

to enrich and develop it [9, 10]. In this sense, it is necessary to promote commercial translation to improve the translation theory and its improvement. Business translation plays an important role in various business situations. This requires translators to fully understand the relevant business knowledge and have the correct understanding of words, sentences, and paragraphs in the translation process, as well as to find out the exact meaning of words or sentences in different business situations [11, 12]. Since 1980, research on translation versatility has been explosive. Many linguists have studied generality from different perspectives, including syntactic, pragmatic, cognitive, and functional perspectives [13, 14]. Jordan divides ESP into two types: one is the purpose of academic English, which refers to the English and communication used to complete study and academic research. The other is professional English, which refers to English suitable for a certain industry, which is more practical and professional. Nick Briefer put forward a famous theory about the category of business English [15, 16]. He claimed that business English includes language knowledge, communication

skills, professional content, management skills, and cultural awareness, in other words, marketing and investment, foreign insurance, international finance, publicity, etc. The universality of translation is an inherent feature, expressed in terms of simplified explicitness and standardization, and has nothing to do with language and cultural differences and the translator's personal preferences [17, 18]. Compared with English texts, based on English translation texts, we can find that almost all English translation texts have several characteristics. These universal theorists have given simplification, explicitness, and normalization.

With the development of descriptive translation studies (DTS), translation is no longer considered a static text transmission process language from the source language to the target language for a long time. Descriptive translation studies (DTS) broadens the horizons of translation studies by shifting the focus from goal orientation to target culture rather than prescriptive translation studies [19, 20]. It provides a theoretical framework for corpus-based translation. China has become more and more closely connected with the world, and many physical or online news media have become a platform for China to receive world information and spread Chinese culture. Business English translation is therefore valued by translation researchers and translators. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing.

In our work, we design and develop a business English translation architecture based on artificial intelligence speech recognition and edge computing to deal with the many differences between the source language and the target language, and to enhance the readability of the translation and meet the reader's cultural cognition and needs. The main contributions of our proposal are as follows:

- (1) It considers the relevance and complementary between speech and text modalities
- (2) It uses the deep neural network feature fusion method to effectively fuse the extracted monomodal features
- (3) It adopts the edge computing method to establish the business English translation system architecture

The rest of this paper is organized as follows. Section 2 discusses the business English translation framework, and edge computing technology is induced in Section 3. Business English translation based on speech recognition methods are discussed in Section 4. Section 5 shows simulation result analysis and discussion, and Section 6 concludes the paper with summary and future research directions.

## 2. Business English Translation Framework

With the help of relevant background and language knowledge, business translators should not only have reliable language knowledge but also should have economic, legal, management, trade, and other aspects [21–30]. Then, all this

knowledge will be applied to provide the background of the content involved in the translation and help the translator to complete the task better and faster. Business theory and practice are two important parts of business English, which means that the language involved should be professional. Without the following foundation, you cannot fully understand business terms [24]. At the same time, due to the huge differences in Western culture between China and China, people have different values and ideologies; translators may encounter expressions that they have never seen before or are difficult to understand. In this case, we can apply the schema theory to cultural background, language environment activation, content structure, and other structure activation to help translators merge what they have learned and have a better understanding of the content [25]. The architecture of business English translation is shown in Figure 1.

With the globalization of the world economy, more and more companies are committed to participating in the international market competition. Therefore, more and more scholars are devoted to the research of business translation [26, 27]. A lot of content related to business translation also appeared. However, there is no uniform definition that is agreed by experts from all walks of life. In this part, the author tries to find out the characteristics and common points of specific previous research and gives the definition of business translation of this article. Business English is proposed under the theoretical framework of ESP, which is a kind of English business activities specially used for international professional subjects [28]. Jordan divides ESP into two types: one is the purpose of academic English, which refers to the English and communication used to complete study and academic research. The other is professional English, which refers to English suitable for a certain industry, which is more practical and professional. Nick Briefer put forward a famous theory about the category of business English [29, 30]. He claimed that business English includes language knowledge, communication skills, professional content, management skills, and cultural awareness. In other words, marketing and investment, foreign insurance, international finance, publicity, etc. are all covered in the scope of business English [31–37]. Therefore, it can be seen that business English is a kind of practical language, a medium of business activities, and business-related English needs and contains various business activities. In essence, business English is a combination of business activities and English. Therefore, business English can be defined as the English used in various business activities.

This society needs business English translators to communicate with each other, to have cross-cultural communication awareness, to grasp the characteristics of business translation, and to avoid contradictions and conflicts caused by cultural differences [38–41]. This paper explains the characteristics of cross-cultural factors in business English translation and study the method for cross-cultural phenomena in business English translation. This will contribute to strengthen awareness of cross-cultural in English translation and makes the ordinary people to have a better realization of various works around the word.

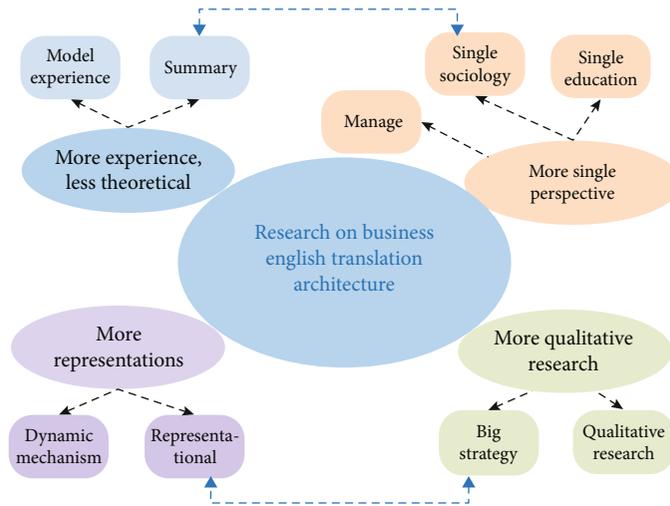


FIGURE 1: Schematic diagram of the architectural concept of business English translation.

### 3. Edge Computing Technology

The popularization of mobile terminals such as smartphones, tablet computers, and smart vehicles has had a huge impact on mobile and wireless networks, triggering challenges for global mobile networks. Mobile devices have certain computing and storage capabilities, but they still suffer from low storage capacity, low bandwidth, and high latency [31, 32]. Mobile cloud computing (MCC) integrates cloud computing and mobile computing, provides powerful functions for mobile devices, and provides them with storage, computing, and energy resources provided by centralized clouds. However, facing countless mobile devices, MCC faces huge challenges. For example, problems such as high latency, low coverage, and data transmission lag. Especially in the case of next-generation mobile networks (such as 5G), it is particularly difficult to deal with these challenges. In addition, MCC is not suitable for scenarios where real-time applications guarantee high-quality services [33, 34].

Edge computing is the combination of mobile edge network and Internet services in the future 5G network. Its biggest feature is to sink core services to the edge of the mobile network, thereby improving the quality of various types of services, reducing waiting time, and providing a better user experience. MEC provides cloud computing functions in the RAN, allowing direct mobile communication between the core network and end users, while connecting users directly to the nearest edge network enabled with cloud services [35]. Deploying on base stations can enhance computing power and effectively avoid bottlenecks and system failures.

### 4. Research on Business English Translation Based on Speech Recognition

Machine translation is a deep learning technology combined with NLP using one of the most active, the most hopeful one direction. The original fully rely on people compiling rule-based machine translation method, then based on statistical

method of SMT now nerve NMT machine translation, machine translation technology over the past 60 years has been updated, especially after the 2012 deep learning technology into the field of vision; the accuracy of machine translation is constantly refreshed; today, it is mainly taking stock of all kinds of deep learning the present situation of the application of machine translation on the inside, giving some representative papers to learn.

The biggest advantages of neural translation (NMT) based on deep learning technology are as follows: (1) It adopts an end-to-end structure and does not need to extract features artificially. (2) The network structure design is simple, and there is no need to carry out complex design work such as word segmentation, word alligation, and syntactic tree design. Meanwhile, the disadvantages of this method are also obvious: (1) Deep learning training often has hundreds of millions of samples. (2) Training a model requires a special GPU cluster, and it takes a few days or even a week to get a result. The iterative update of the model is very slow. Nal Kalchbrenner and Edward Grefenstette put forward in 2013, based on coding decoding structure framework, a new kind of machine translation in the source language sentence, using a convolution neural network to map it into a continuous dense implicit vector; using a recursive neural network as a decoder, the implicit vector decode sentences into the target language. Doing a little bit is that you can use the RNN to deal with different lengths of the input sentence, which try to capture all of its historical information, but a gradient disappeared and gradient explosion occurred because of RNN. There is no way to capture long-term dependencies.

Speech recognition is an interdisciplinary subject, involving many fields, such as artificial intelligence, pattern recognition, signal processing, and information theory. In information theory, the statistical method of speech recognition is modeled as a noise channel model [36, 37]. Each component is associated with the component of the general noise channel model at the bottom of the figure. The speech recognition process includes the speaker and the speech

recognizer. The speaker will consciously generate a sequence of words. A schematic diagram of the architecture of the convolutional neural network is shown in Figure 2.

Then, the word sequence is sent to the acoustic (corresponding to the noise channel) channel composed of a voice generator and an acoustic processing unit. The voice generator generates a voice signal of the word sequence in the acoustic environment, and then, the acoustic processing unit signals the voice signal processing to generate acoustic features; the speech decoder receives a sequence of feature vectors [42] and infers the word sequence that is closest to the original word sequence; the acoustic processing unit and the language decoder together form the speech recognizer.

An artificial neural network (ANN) is an information processing model inspired by the complex neural network of the brain. Simply put, the model consists of multiple layers of neurons. In order to obtain the most ideal model output results, it is necessary to constantly adjust the relevant parameters between neurons [38, 43]. A schematic diagram of business English translation based on speech recognition is shown in Figure 3.

Therefore, the artificial neural network model needs to be combined with the corresponding algorithm to solve the problem of how to adjust the weight parameters. The BP (back propagation) algorithm solves this problem well [39]. The basic principle of the BP algorithm includes the forward propagation of sample features at all levels. At the same time, the learning rules of the steepest descent method are used to adjust the weights of related neurons in the model through the feedback value of back propagation, so as to minimize the output error of the model. The BP algorithm is based on neurons, and its composition is divided into three parts: the input layer, hidden layer with different numbers, and output layer. The specific structure and algorithm are shown in Section 2. In recent years, neural networks have been rapidly developed and used, and they have performed well in the fields of pattern recognition, artificial intelligence, and automatic control [40].

In the first few layers of the convolutional neural network, the data nodes are transformed into a three-dimensional matrix. Unlike the fully connected layer, only some nodes are connected to adjacent layers, which greatly reduce the number of parameters and reduces subsequent models' training time. Usually a convolutional neural network consists of the following structure:

$$\begin{aligned} w_i &= \sqrt[n]{M_i}, \\ \lambda_{\max} &= \sum_{i=1}^n \frac{(BW)_i}{nw_i}. \end{aligned} \quad (1)$$

The input layer is the input of the neural network. In the image-oriented convolutional neural network, it represents the pixel matrix converted after a picture is read by a computer.

$$Xt = \sum_{i=1}^k \alpha_i X_{t-i} + \sum_{i=1}^k \beta_i Y_{t-i} + u_{it}. \quad (2)$$

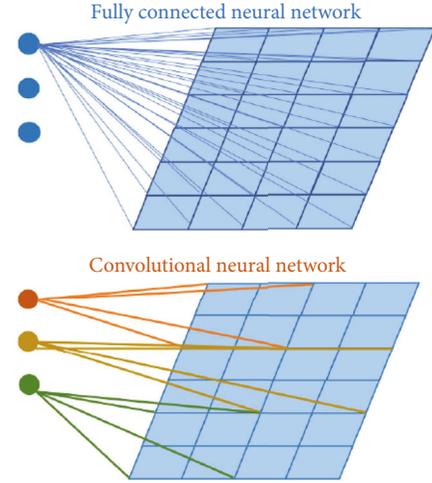


FIGURE 2: Schematic diagram of the architecture of the convolutional neural network.

The pixel matrix is a three-dimensional matrix. The length and width represent the size of the image, and the depth represents the number of color channels of the image. When the input picture is a black and white picture, the depth is 1; when the input picture is a color picture, the depth is 3.

$$\begin{aligned} F &= \frac{(SSE_r - SSE_R)/k}{SSE_R/(T - 2k)} \sim F_{(k, T-2k)}, \\ \Delta f_r &= \left[ \frac{\partial f_r}{\partial x_1}, \frac{\partial f_r}{\partial x_2}, \dots, \frac{\partial f_r}{\partial x_n} \right]. \end{aligned} \quad (3)$$

Starting from the input layer, the three-dimensional matrix is transformed into another three-dimensional matrix through a different network structure until the final fully connected layer. The convolutional layer is the most important part of the convolutional neural network and the key to extracting features.

$$\begin{aligned} f &= f(x), \\ x'_j &\leq x_j \leq x''_j \quad (j = 1, 2, \dots, n), \\ g_j &\leq g''_j \quad (j = 1, 2, \dots, n), \\ h'_j &\leq h_j(x) \quad (j = 1, 2, \dots, n), \\ w'_j &\leq w_j(x) \leq w''_j \quad (j = 1, 2, \dots, n). \end{aligned} \quad (4)$$

Generally, two-dimensional convolution is used in image recognition, that is, a discrete two-dimensional filter (also called a convolution kernel) and a two-dimensional image are convolved. The specific process is that the two-dimensional filter passes through all positions on the two-dimensional image, and do inner product with the pixel and its domain pixels at each position, and

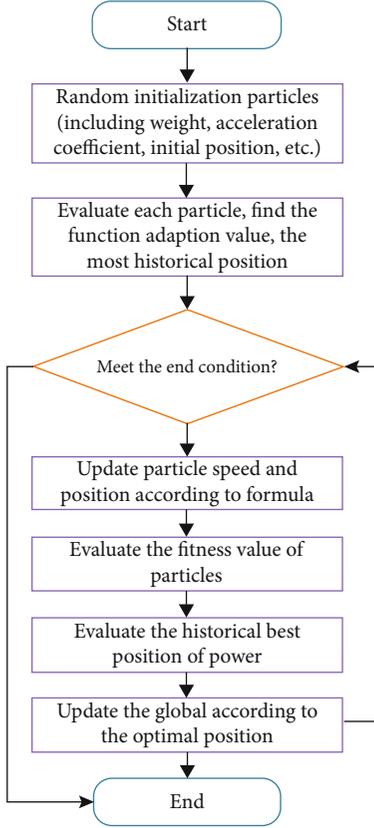


FIGURE 3: Schematic diagram of business English translation based on speech recognition.

different convolution kernels correspondingly extract different features.

$$x^{(j+1)} = x^{(j)} + s_j d^{(j)}. \quad (5)$$

Through the product of the above conditional probabilities, the joint probability of the word sequence is obtained. The above function can be further decomposed into two parts: feature mapping and calculating conditional probability distribution.

$$\begin{aligned} |f^{(j)} - f^{(j-1)}| &\leq \tau, \\ |f^{(j)} - f^{(b)}| &\leq \tau. \end{aligned} \quad (6)$$

In feature mapping, map each input word to a feature vector through the mapping matrix.  $C$  represents the feature vector corresponding to the  $Z$ th word in the dictionary, and  $Y72$  represents the dimension of the feature vector. The process will be the vector obtained by feature mapping.

$$Q(x, q) = \frac{f}{f_0} + \sum_j p_x(x_j) + q \left[ \sum_j p_g(g_j) + \sum_j p_h(h_j) + \sum_j p_w(w_j) \right]. \quad (7)$$

A previous work used RNN for temporal classification by combining RNN in a hybrid system. However, in addition to inheriting the shortcomings of HMM, this method does not make full use of the full potential of RNN for long-term sequence modeling. Therefore, it seems more feasible to directly train RNN for temporal classification tasks. CTC achieves this by allowing the network to perform label prediction at any point in the input sequence, as long as the overall order of the labels is correct. Since CTC does not care about the alignment of the target label with the input sequence and directly outputs the probability of the complete label sequence, this allows CTC to use the network as a timing classifier without data segmentation and external postprocessing [41].

A new method is proposed in this paper. The GAN translate network model, namely, CSGAN NMT model, is composed of two against models, one is the generation model  $G$ , which adopts the traditional attention-based neural network model, and where the source language sentence is translated into the target language; secondly, the discriminant model  $D$ , where the judgement sentence is translated by machine translation or people. Its design has two kinds of methods, namely, adopt the CNN or RNN design, and finally find that CNN's effect is better than LSTM, because a negative signal interference exists for LSTM training. In this paper, MLE is used to pretrain generator  $G$ , and then samples generated by  $G$  and real samples are used to pretrain  $D$ . When  $D$  reaches a certain accuracy rate, confrontational training will be entered. The part of GAN is basically the same as that of SeqGAN, which adopts the policy gradient method +MC search. Compared with these existing methods, our methods can do well in different reference standard environments though it may be overfitting. Compared with the existing methods, our proposal has three main advantages: firstly, it has a lower time consumed than others, which indicates that our proposal can translate as quickly as possible and in time. Secondly, our proposal has a higher accuracy than others, which shows that our proposal can perform well. Thirdly, our proposal can be adapt to any situations with English and business. However, the main limitation of our proposal is that it needs a huge computation space, which indicates that our proposal has a strong requirement for computation and it may not be easy to realize.

## 5. Case Study of Business English Translation

**5.1. Data Case Source.** The experiment in this section uses the deep learning model introduced in the paper to establish a business English speech translation model. In the training process of this experiment, the input of the deep learning model is input by frame, and the dimension of each frame of data is different according to the different feature parameters, and the corresponding output label of each frame of data is obtained after forward propagation. The dimension of the output layer is the same as the total number of labels, and the probability of each node output represents the probability of the corresponding label of the node in the frame of data. When the deep learning model is initially constructed, all weights will be initialized. After the input data enters the

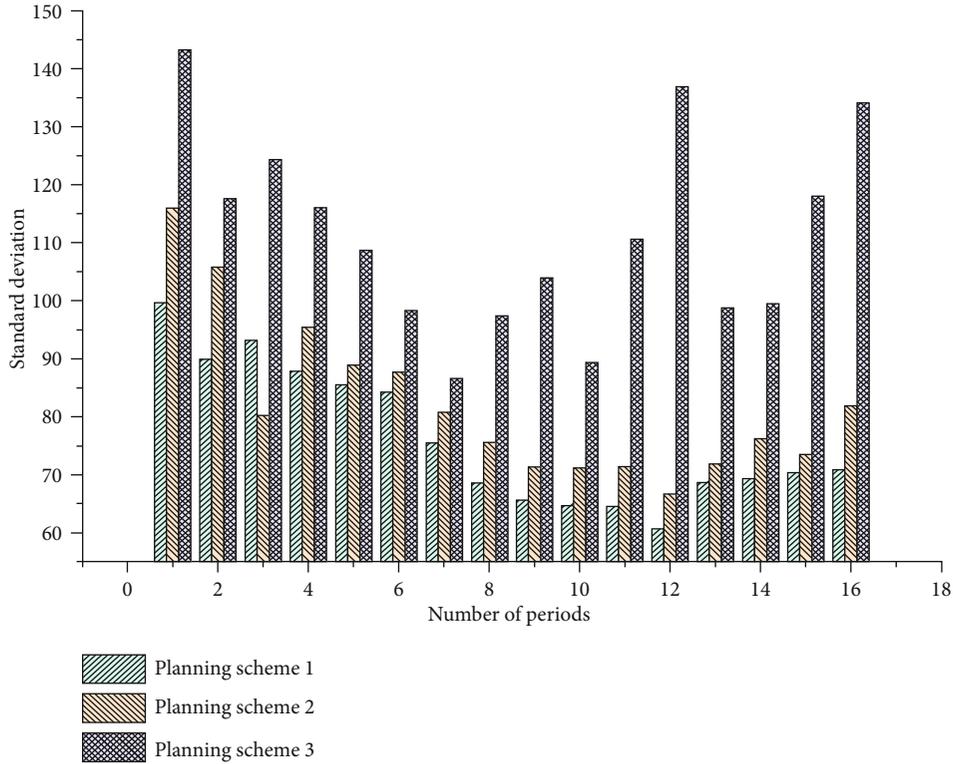


FIGURE 4: Schematic diagram of business English translation alignment label data distribution.

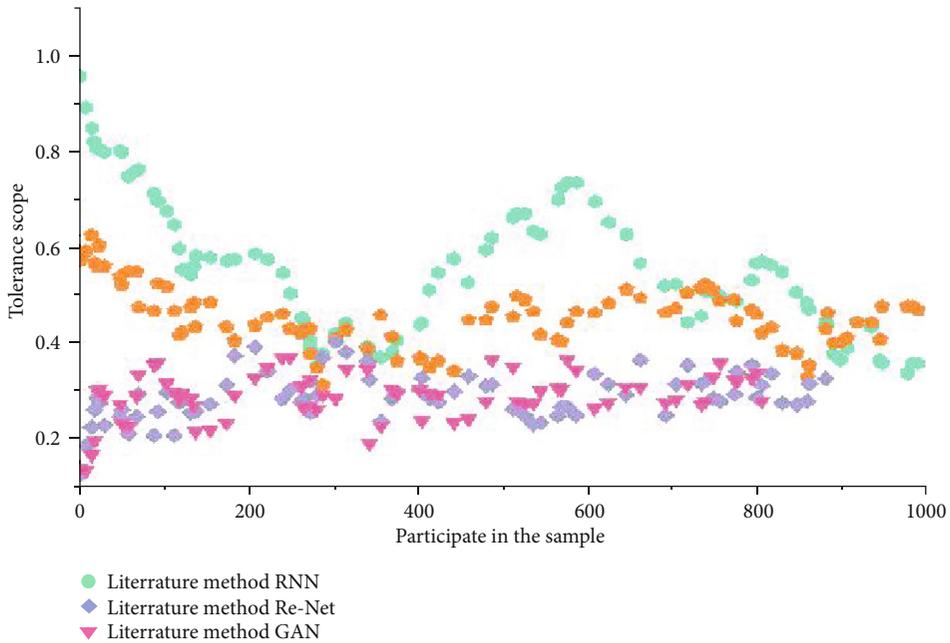


FIGURE 5: Comparison results of business English translation accuracy under different models.

network, it will be forward propagated according to the initial weight, and the output obtained is the predicted value.

The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that

each weight and the parameters of each gate are updated. The weights are changed to complete a training process. The static optimization architecture implemented in this article and the static thermal path identification based on this optimization method have enabled the system to obtain a greater performance improvement. This chapter will specifically

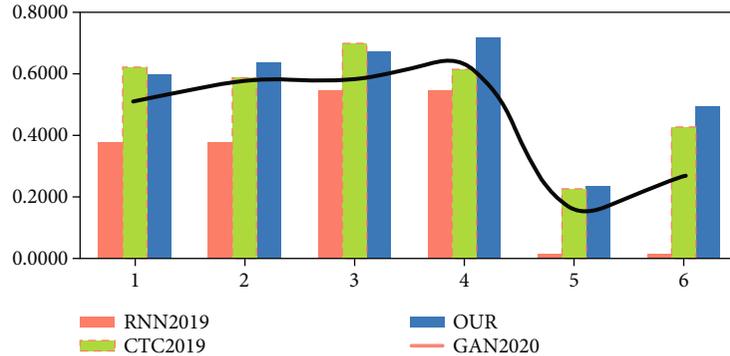


FIGURE 6: Comparison of accuracy of business English translation under different datasets.

introduce the resulting performance speedup and analyze the root cause of the performance improvement based on experimental data.

## 5.2. Application Effect Evaluation

**5.2.1. Accuracy.** Different feature parameters are extracted according to different experimental requirements. The feature parameters used in this experiment include MFCC, PLP, and FBANK features. For far-field voice data, we need to use microphone array technology to improve the quality of far-field voice. By beamforming the MDM data, the enhanced voice is obtained before subsequent processing. In order to obtain more accurate alignment label data, multiple GMM-HMM acoustic models are trained as described in the experiment, and label alignment is repeated, and finally, the alignment label data with higher accuracy is obtained. Figure 4 shows the distribution diagram of alignment label data for business English translation.

In planning scheme 1, the first alignment method to obtain the alignment label data is to averagely segment the training samples according to the number of sentence states and use the initial alignment label data to train a simple monophony model.

In planning scheme 2, after the acoustic model is obtained, the optimal path in the state diagram is selected according to the speech frame and the existing acoustic model, and each frame is matched to the state diagram to obtain the corresponding state of each frame.

In planning scheme 3, more accurate alignment label data was obtained.

According to this idea, we repeatedly train new models and repeatedly align them and finally get more accurate alignment label data. In order to prove the effectiveness of the deep neural network on the convergence speed of the model, we compared our model with existing methods. In order to ensure the reliability of the comparison, the same parameter settings are used for all models and run in the same environment. The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that each weight and the parameters of each gate are updated. The weights are changed to complete a

TABLE 1: Comparison table of results of performance conversions.

Real prediction	Positive	Negative
Positive	TP	FN
Negative	FP	TN

training process. It can be seen from Figure 5 that the accuracy of different models in business English translation is still different. The accuracy of the validation set changes significantly. As the training progresses, the accuracy gradually increases and the range is more obvious. When the epoch is greater than 6 and continues to increase, the changes in the validation set and training tend to be flat and reach a stable state. Figure 5 shows the accuracy comparison of several models.

**5.2.2. Time Consuming.** If the epoch continues to increase, the accuracy changes on the training set and validation set are no longer obvious, which shows that there is still a certain overfitting. In order to ensure the reliability of the comparison, the model uses the same parameter settings and runs in different environments. Figure 6 shows the accuracy comparison under several environments and standards.

We use the confuse matrix and F1-score to evaluate the translate model performance; the matrix can be defined as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (8)$$

For experimental results, the results of performance conversion are shown in Table 1.

$$\begin{aligned} \text{f1-score} &= \frac{2 * \text{precise} * \text{recall}}{\text{precise} + \text{recall}}, \\ \text{precise} &= \frac{TP}{TP + FP}, \\ \text{recall} &= \frac{TP}{TP + FN}. \end{aligned} \quad (9)$$

It can be seen from Figure 7 that the accuracy of the same model in business English translation under different environments and standards is still different. The reference

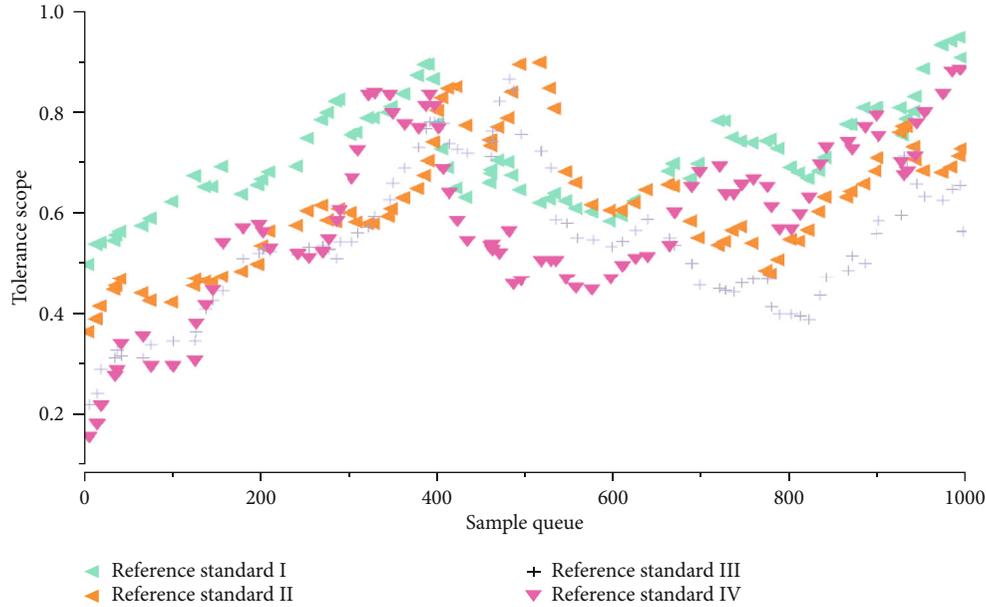


FIGURE 7: Comparison of accuracy of business English translation under different environments and standards.

standard is as follows: (1) conform to the expression habits of the target language, (2) when there are no words in the target language, keep the original words or create new words according to the meaning, (3) complete and accurate expression of the original information, without core semantic errors, and (4) the term is consistent with the common standards or practices of the industry and profession in the target language. Figure 6 shows the four methods results on different datasets. It indicated that our proposal obtained the best accuracy on 4 datasets of 6. This shows that there is still some overfitting. After the acoustic model is obtained, the optimal path in the state diagram is selected according to the speech frame and the existing acoustic model, and each frame is matched to the state diagram to obtain the corresponding state of each frame, thereby obtaining more accurate alignment label data. According to this idea, we repeatedly train new models and align them repeatedly and finally get more accurate alignment label data. During the initial construction of the deep learning model, all weights are initialized. After the input data enters the network, it will be forward propagated according to the initial weight, and the output obtained is the predicted value. The error between the predicted value and the true value is calculated through the loss function, and then, the time-based backpropagation algorithm is used to update the gradient in the direction where the loss function is falling, so that each weight and the parameters of each gate are updated. The weights are changed to complete a training process.

And we also compared three methods proposed in the last three years such as GNNE, RNNE, and CNNE with our proposal to investigate the effectiveness of our methods. The results can be shown in Figure 8:

As shown in Figure 8, the green line represents our proposal, and on all datasets, we can see that our proposal is better than others except D2, which all methods obtain the same

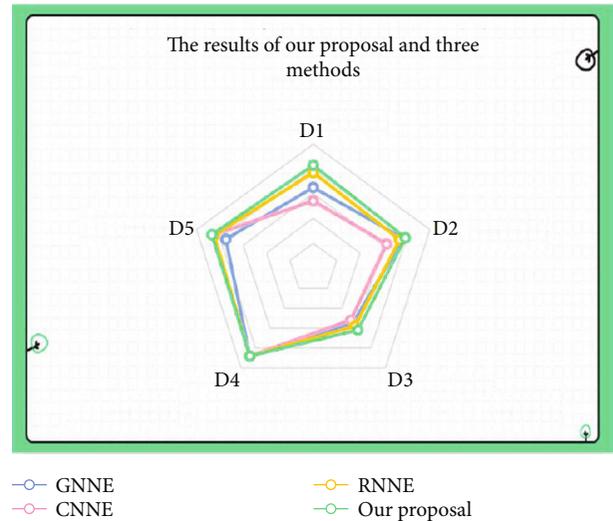


FIGURE 8: The results of our proposal and three methods on F1-score.

results. It indicated that our proposal can perform well than other three methods.

## 6. Conclusion

International business translation has received increasing attention as communication tools. In the foreseeable future, businessmen should be able to communicate directly with their foreign partners in English. Government officials and decision makers, who are not businessmen, should be able to review business documents in English. In short, business translation in the era of globalization is very necessary and important. In other words, marketing and investment, foreign insurance, international finance, publicity, etc. are all

covered in the scope of business English. Therefore, it can be seen that business English is a kind of practical language, a medium of business activities, and business-related English needs and contains various business activities. In essence, business English is a combination of business activities and English.

Therefore, business English can be defined as the English used in various business activities. Aiming at the shortcomings of current business English translation research, this paper designs and develops a business English translation architecture based on artificial intelligence speech recognition and edge computing. In this case, we can apply schema theory to the cultural background, the activation of the language environment, the content structure, and the activation of other structures to help translators merge the content they have learned and have a better understanding. We will also continue to devote ourselves to relevant research, hoping to provide a meager power for business English translation. Although our method has achieved good translation accuracy at present, it is still unable to achieve considerable accuracy in the face of complex speech environment, and the training time of the model is long, which cannot meet the purpose of real-time translation. In the future, we will further optimize the speech translation model to improve the training speed of the model while ensuring the accuracy.

## 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 they have no conflicts of interest.

## Acknowledgments

The research was supported by the Youth Project of University Humanities and Social Science Research of Jiangxi Province (No. JC19209) and the Project of 13th Five-Year Plan for Education and Science of Jiangxi Province (No. 19YB387).

## References

- [1] J. Zhang, L. Zhou, F. Zhou et al., "Computation-efficient offloading and trajectory scheduling for multi-UAV assisted mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2114–2125, 2020.
- [2] K. Y. Zhan, J. H. Lewis, K. J. Vasil et al., "Cognitive functions in adults receiving cochlear implants: predictors of speech recognition and changes after implantation," *Otology & Neurotology*, vol. 41, no. 3, pp. e322–e329, 2020.
- [3] X. Xiong, K. Zheng, L. Lei, and L. Hou, "Resource allocation based on deep reinforcement learning in IoT edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1133–1146, 2020.
- [4] T. Wang, Y. Liang, Y. Yang et al., "An intelligent edge-computing-based method to counter coupling problems in cyber-physical systems," *IEEE Network*, vol. 34, no. 3, pp. 16–22, 2020.
- [5] P. Wagner, "Carnapian and Tarskian semantics," *Synthese*, vol. 194, no. 1, pp. 97–119, 2017.
- [6] B. Trukeschitz, A. Hajji, J. Litschauer et al., "Translation, cultural adaptation and construct validity of the German version of the Adult Social Care Outcomes Toolkit for informal Carers (German ASCOT-Carer)," *Quality of Life Research*, vol. 30, no. 3, pp. 905–920, 2021.
- [7] N. Tratnik, "Generalized cut method for computing the edge-Wiener index," *Discrete Applied Mathematics*, vol. 282, pp. 222–233, 2020.
- [8] P. X. W. Zou, G. Zhang, and J. Wang, "Understanding the key risks in construction projects in China," *International Journal of Project Management*, vol. 25, no. 6, pp. 601–614, 2007.
- [9] I. Speck, M. C. Ketterer, S. Arndt, A. Aschendorff, T. F. Jakob, and F. Hassepass, "Comparison of speech recognition and localization ability in single-sided deaf patients implanted with different cochlear implant electrode array designs," *Otology & Neurotology*, vol. 42, no. 1, pp. e22–e32, 2021.
- [10] V. F. Uricchio, R. Giordano, and N. Lopez, "A fuzzy knowledge-based decision support system for groundwater pollution risk evaluation," *Journal of Environmental Management*, vol. 73, no. 3, pp. 189–197, 2004.
- [11] Z. Song, "English speech recognition based on deep learning with multiple features," *Computing*, vol. 102, no. 3, pp. 663–682, 2020.
- [12] S. P. Singh, A. Nayyar, R. Kumar, and A. Sharma, "Fog computing: from architecture to edge computing and big data processing," *Journal of Supercomputing*, vol. 75, no. 4, pp. 2070–2105, 2019.
- [13] J. Shehorn, O. Strelcyk, and P. Zahorik, "Associations between speech recognition at high levels, the middle ear muscle reflex and noise exposure in individuals with normal audiograms," *Hearing Research*, vol. 392, article 107982, 2020.
- [14] R. de Vries, F. Haas, and working group for revision of the Dutch Blood Transfusion Guideline 2011, "English translation of the Dutch Blood Transfusion Guideline 2011," *Clinical Chemistry*, vol. 58, no. 8, pp. 1266–1267, 2012.
- [15] D. Raynaud, S. Gessner, and B. Mota, "Andalò di Negro's De compositione astrolabii: a critical edition with English translation and notes," *Archive for History of Exact Sciences*, vol. 73, no. 6, pp. 551–617, 2019.
- [16] Y. Qu and H. Liu, "The basic problem of contemporary water resources environment research: translation and analysis from the English practice report on water resources," *Journal of Coastal Research*, vol. 104, 2020.
- [17] Y. R. Oh, K. Park, H. B. Jeon, and J. G. Park, "Automatic proficiency assessment of Korean speech read aloud by non-natives using bidirectional LSTM-based speech recognition," *Etri Journal*, vol. 42, no. 5, pp. 761–772, 2020.
- [18] A. Mehrabi, M. Siekkinen, and A. Yla-Jaaski, "Edge computing assisted adaptive mobile video streaming," *IEEE Transactions on Mobile Computing*, vol. 18, no. 4, pp. 787–800, 2019.
- [19] Z. Liu, "A study on English translation of tourism publicity in coastal cities from the perspective of cross-cultural communication," *Journal of Coastal Research*, vol. 115, no. sp1, 2020.
- [20] H. Liao, Z. Zhou, W. Kong et al., "Learning-based intent-aware task offloading for air-ground integrated vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–13, 2020.
- [21] H. Li, H. Xu, C. Zhou, X. Lu, and Z. Han, "Joint optimization strategy of computation offloading and resource allocation in

- multi-access edge computing environment,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10214–10226, 2020.
- [22] M. Lee, J. Lee, and J. H. Chang, “Ensemble of jointly trained deep neural network-based acoustic models for reverberant speech recognition,” *Digital Signal Processing*, vol. 85, pp. 1–9, 2019.
- [23] C. Lee, S. Hong, S. Hong, and T. Kim, “Performance analysis of local exit for distributed deep neural networks over cloud and edge computing,” *Etri Journal*, vol. 42, no. 5, pp. 658–668, 2020.
- [24] A. M. Kubiak, J. Rennies, S. D. Ewert, and B. Kollmeier, “Prediction of individual speech recognition performance in complex listening conditions,” *Journal of the Acoustical Society of America*, vol. 147, no. 3, pp. 1379–1391, 2020.
- [25] D. Kim, H. Lee, H. Song, N. Choi, and Y. Yi, “Economics of fog computing: interplay among infrastructure and service providers, users, and edge resource owners,” *IEEE Transactions on Mobile Computing*, vol. 19, no. 11, pp. 2609–2622, 2020.
- [26] Y. H. Jung, S. K. Hong, H. S. Wang et al., “Speech recognition: flexible piezoelectric acoustic sensors and machine learning for speech processing (Adv. Mater. 35/2020),” *Advanced Materials*, vol. 32, no. 35, article 2070259, 2020.
- [27] J. M. Hung, X. Li, J. Wu, and M. F. Chang, “Challenges and trends in developing nonvolatile memory-enabled computing chips for intelligent edge devices,” *IEEE Transactions on Electron Devices*, vol. 67, no. 4, pp. 1444–1453, 2020.
- [28] D. Hülsmeier, A. Warzybok, B. Kollmeier, and M. R. Schädler, “Simulations with FADE of the effect of impaired hearing on speech recognition performance cast doubt on the role of spectral resolution,” *Hearing Research*, vol. 395, article 107995, 2020.
- [29] Y. Huang, X. Song, F. Ye, Y. Yang, and X. Li, “Fair and efficient caching algorithms and strategies for peer data sharing in pervasive edge computing environments,” *IEEE Transactions on Mobile Computing*, vol. 19, no. 4, pp. 852–864, 2020.
- [30] X. Huang, R. Yu, S. Xie, and Y. Zhang, “Task-container matching game for computation offloading in vehicular edge computing and networks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–14, 2020.
- [31] B. Hu and J. Li, “An edge computing framework for power-train control system optimization of intelligent and connected vehicles based on curiosity-driven deep reinforcement learning,” *IEEE Transactions on Industrial Electronics*, vol. 68, no. 8, pp. 7652–7661, 2021.
- [32] M. P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, “Multimodal data analysis of epileptic EEG and rs-fMRI via deep learning and edge computing,” *Artificial Intelligence in Medicine*, vol. 104, article 101813, 2020.
- [33] H. Xing, L. Liu, J. Xu, and A. Nallanathan, “Joint task assignment and resource allocation for D2D-enabled mobile-edge computing,” *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4193–4207, 2019.
- [34] J. T. Holder, N. C. Dwyer, and R. H. Gifford, “Duration of processor use per day is significantly correlated with speech recognition abilities in adults with cochlear implants,” *Otology & Neurotology*, vol. 41, no. 2, pp. e227–e231, 2020.
- [35] R. Du, Y. Liu, L. Liu, and W. Du, “A lightweight heterogeneous network clustering algorithm based on edge computing for 5G,” *Wireless Networks*, vol. 26, no. 3, pp. 1631–1641, 2020.
- [36] J. Dornhoffer, Y. F. Liu, E. E. Zhao, E. L. Camposeo, T. A. Meyer, and T. R. Mccrackan, “Factors that influence second-side cochlear implant speech recognition outcomes,” *Otology & Neurotology*, vol. 42, 2021.
- [37] S. Colby and B. McMurray, “The role of listening effort during degraded speech recognition: a comparison of the dual-task and pupillometry paradigms,” *Journal of the Acoustical Society of America*, vol. 146, no. 4, pp. 3050–3051, 2019.
- [38] D. Liu, J. Xu, P. Zhang, and Y. Yan, “A unified system for multilingual speech recognition and language identification,” *Speech Communication*, vol. 127, pp. 17–28, 2021.
- [39] S. Yang, Z. Gong, K. Ye, Y. Wei, Z. Huang, and Z. Huang, “EdgeRNN: a compact speech recognition network with spatio-temporal features for edge computing,” *IEEE Access*, vol. 8, pp. 81468–81478, 2020.
- [40] J. Zhang and D. Tao, “Empowering things with intelligence: a survey of the progress, challenges, and opportunities in artificial intelligence of things,” *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7789–7817, 2021.
- [41] M. S. Hossain and G. Muhammad, “An audio-visual emotion recognition system using deep learning fusion for a cognitive wireless framework,” *IEEE Wireless Communications*, vol. 26, no. 3, pp. 62–68, 2019.
- [42] A. Boukerche and V. Soto, “Computation offloading and retrieval for vehicular edge computing,” *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–35, 2020.
- [43] K. Borowiak, C. Maguinness, and K. V. Kriegstein, “Dorsal-movement and ventral-form regions are functionally connected during visual-speech recognition,” *Human Brain Mapping*, vol. 41, no. 4, pp. 952–972, 2020.