IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 April 2023, accepted 18 May 2023, date of publication 29 May 2023, date of current version 6 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3280549

RESEARCH ARTICLE

An Automatic Error Detection Method for Machine Translation Results via Deep Learning

WEIHONG ZHANG
Foreign Languages School, Zhengzhou Business University, Gongyi, Henan 450012, China
e-mail: zhanghelen3@zbu.edu.cn

ABSTRACT Nowadays, the rapid development of natural language processing has brought great progress for the area of machine translation. Various deep neural network-based machine translation approaches have been more and more general. However, there still lacks effective automatic error detection approaches for machine translation results. To bridge such gap, this paper proposes an automatic error detection method for machine translation results via deep learning. The training data is synthesized using the deep generative model proposed in this paper, which is used for the training of the foreign trade English grammatical error correction model. Then, the grammatical error correction model is used to correct the source sentences in the learner's corpus, and the corrected target sentences and the manually annotated standard sentences are formed into ''error-correct'' sentence pairs, which are fed back to the error generation model for alternate training. By establishing a link between the grammatical error detection model and the grammatical error correction model, the error detection and correction capability of the model is improved. Experiments on datasets such as GTRSB show that the proposed error detection method significantly improves the stealthiness of the trigger while ensuring the effectiveness of the backdoor attack, and at the same time enables the trigger to resist certain data augmentation operations.

INDEX TERMS Automatic error detection, machine translation, natural language processing, deep learning.

I. INTRODUCTION

In recent years, deep learning techniques have made rapid progress and have become one of the most popular technologies [1]. Previously, the field of machine learning relied heavily on expert models, leading to specific extensions to each application scene [2], [3]. And domain experts had to tailor a set of feature extraction methods for the current scene to achieve the feature extraction function for specific scene data [4], [5]. However, the artificially set features are generally simple and low-dimensional and do not have high abstract description ability, leading to unsatisfactory results in subsequent training applications [6], [7]. Since the emergence of deep learning, the effect of various feature extraction in it has greatly surpassed the traditional expert setting mode [8]. Deep learning can extract data features at different levels of abstraction in the training dataset by repeatedly training and learning at multiple levels on large-scale data, and finally presenting and applying them in the form of models [9]. However, the problem of a lengthy and inefficient deep learning training process has been one of the dilemmas facing deep learning [10]. It often takes a lot of time and computational resources to train a usable deep model for a specific application [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Liudong Xing.

Moreover, the trend of scaling up deep learning models nowadays, if the training process of deep learning models cannot be accelerated at the system level, it will bring constraints to researchers of deep learning model algorithms and hinder the design work of upper-level models. At the same time, deep learning systems have also shown vulnerability to attacks and lack of robustness, and there have even been reported incidents of threats to personal safety due to defects in deep learning systems. Since its introduction, deep learning methods have received a lot of attention from the industry. Grammatical Error Correction (GEC) is a traditional part of computer-aided language research and an important task in the field of natural language processing, which aims to identify and correct grammatical errors by analyzing the grammatical dependencies and logic among the components of the input utterance. In the process of learning English as a Second Language (ESL), ESL learners need not only to quickly identify and correct spelling errors but also to ensure grammatical correctness.

Only within the constraints of grammar can semantic expressions be more accurate, clear, and effective. In our current education process, learners still learn collocations mainly from textbooks and dictionaries, and students need to rely on memory to determine which collocations are reasonable and the correct meaning of collocations, which is undoubtedly a great burden for learners [12]. In the actual writing and communication process, learners are confused by unskilled collocations and make mistakes, such as constructing collocations by translating directly from their native language to English without considering their grammatical correctness, which makes the learners' English look unnatural. Teachers also teach collocations based on some dictionaries, but in English, the laws of collocations are so varied that dictionaries can hardly cover all collocations, and teachers are powerless to cover some uncommon collocations, on the other hand, the compilation of dictionaries requires not only great human and material resources, but also sufficient English expertise, so there are few existing collocation dictionaries.

Increased people in the domestic economic and trade sector need to understand and grasp the international economic situation and market trends timelier and accurately to expand foreign trade and facilitate economic exchanges [13]. At the same time, more people from foreign countries also want to understand China's economic policies, laws and regulations, market demand, and investment environment [14]. As a branch of Business English, Foreign Trade English belongs to the category of English for Special Purposes (ESP) [15]. Foreign trade English and general English have the general commonality of language, but there are differences in the scope of use, stylistic features, lexical meaning, etc [16]. Foreign trade English belongs to special purpose English, and its language has its characteristics [17]. So the general translation theory is not fully suitable for foreign trade English translation [18]. As an important branch of business English, foreign trade English has an important significance in promoting China's foreign trade economic activities [19]. Based on the two characteristics that the translation of foreign trade English involves a wide range and is very time-sensitive, we cannot spend weeks or even years to refine the translation like other types of translation [20]. But we must race against time to deliver the information as soon as possible [21]. Therefore, how to quickly complete the translation of foreign trade English is a problem with both theoretical and practical significance [22].

Based on this, this paper proposes the automatic detection method of foreign trade English translation based on deep learning model, which can quickly complete the translation of foreign trade English [23]. And the method can improve foreign trade English and its translation from theoretical research to a new level [24]. Main contributions of this paper can be summed up as three aspects:

- The significance and importance of automatic error detection for translation are discussed in this paper.
- This paper proposes an automatic error detection method for machine translation results via deep learning.
- The proposed method is evaluated via simulation experiments on scenarios constructed by real-world data.

This paper is divided into six sections, and the main contents of each section are specified as follows [25]. The first section is an introduction, which mainly introduces the research background related to foreign trade English translation and its significance. The second section is related work, which mainly introduces the status of domestic and foreign research on foreign trade English translation and the research status of related processing technology. The third section introduces the theoretical knowledge related to deep learning model, and on this basis, an automatic error detection model based on deep learning is designed. The fourth section constructs the automatic error detection model for English translation errors and sets the parameters of the data enhancement strategy. The fifth section is the result analysis, and this section mainly focuses on simulation and simulation experiments to verify the effectiveness of the automatic error detection model of foreign trade English translation based on deep learning model. The sixth section is the conclusion and outlook, and this section briefly discusses the contributions and limitations of this paper as well as the future development directions.

II. RELATED WORK

With the continuous development of natural language processing technology, automatic correction methods for English grammatical errors are also being innovated. There are different advantages and disadvantages among different methods [26]. Early grammatical error detection relied on handwriting rules and customization of error grammar templates, while later work focused on supervised learning, construction of error-annotated corpora, and the use of feature engineering methods with maximum entropy-based classification for error detection. The rule-based approach can target the correction of grammatical errors and significantly improve the accuracy of correction, but it requires the construction of many rule bases, and conflicts between rules may occur. The classification-based GEC approach can effectively correct errors of coronals, prepositions, and other types, but the types of errors corrected are relatively few; using the neural machine-translation-based approach can better deal with the long-distance dependency problem and improve the performance of GEC models, but it may face the problem of high model training complexity.

Hinton G et al. were the first to propose the use of neural networks to solve the error detection problem by training a sequence labeling model based on bidirectional LSTM for word embedding modeling, using the whole sentence as the context, and outputting the probability distribution of all labels in turn [27]. By adding a secondary language modeling target to the original neural sequence labeling structure to learn the surrounding words of each word in the prediction dataset, this multiple training target architecture is found to be particularly useful for sequence labeling tasks, especially error detection tasks, by adding a secondary language modeling target to the original neural sequence labeling structure, which greatly improves the accuracy of sequence labeling tasks by learning more general features about language and context. Xu Y et al. use three data-driven system design approaches based on a large-scale native language and learner corpus and discuss the evaluation criteria of the language error detection system. Grammatical error correction is a mature topic in the field of natural language processing, and with the development of deep neural networks, scholars at home and abroad have become increasingly concerned with this topic [28]. At present, experts and scholars at home and abroad have achieved good results in the field of grammar error correction, and there are also many error correction websites for public use. The common error correction methods mainly include classification-based grammar error correction methods, statistical-based grammar error correction methods, and translation architecture-based grammar error correction methods.

A deep learning model backdoor attack is when an attacker injects a backdoor into the model using, for example, data poisoning, so that the model outputs predefined false results for samples containing specific patterns while behaving normally for clean samples. These specific patterns are also called backdoor triggers. Based on the differences between backdoor triggers and backdoor injection methods, existing backdoor attack methods for deep learning models can be classified as visible backdoor attacks, invisible backdoor attacks, clean labeled backdoor attacks, black box backdoor attacks, physical backdoor attacks, and other backdoor attacks, etc. Gu et al. first defined backdoor attacks for deep learning models and proposed a backdoor attack method for outsourced training and migration learning [29]. The basic idea of this method is that the attacker adds some specific pixel points to part of the training data as backdoor triggers and assigns error labels to them to obtain toxic data; then, these toxic data are mixed with other training data for model training to complete backdoor injection. To further improve the stealthies of the backdoor attack, Zhao Z et al. generate backdoor triggers based on the idea of generic adversarial perturbation, while using 12 parametric constraints on the size of the triggers [30].

The invisible backdoor attack method improves the stealthiness of the attack, but the labels of the poisoned sample and the original sample are inconsistent. Therefore, artificial detection of backdoor attacks can be achieved by judging the correspondence between samples and labels. To mitigate or eliminate the backdoor in the model, researchers have investigated the defense methods of backdoor attacks. Existing backdoor attack defense methods for deep learning models can be divided into input preprocessing-based defense methods, model diagnosis-based defense methods, model reconstruction-based defense methods, and other defense methods. After the rise of deep learning, numerous z-automatic detection models based on deep learning have been introduced, making the correct rate of automatic detection and recognition greatly increased. Deep learning-based c-error detection recognition technology has also gradually become the mainstream technology in the field of detection recognition, and has also rapidly become a research hotspot in the field. The globalization of the world economy has deepened, and international business communication has taken on a trend that occurs from time to time. English has become an essential language communication tool in the world economy. Statistics report that almost 90% of people use English every day as their first or second language for business communication [31]. However, foreign trade English translators are facing a dilemma when it comes to supply and demand. On the one hand, due to the gradual expansion of international exchange and cooperation, the workload of foreign trade business increases, and translators always have a lot of heavy work to do on their hands; on the other hand, the application of some modern media means like e-commerce, Internet, etc. in international business relations requires translators to provide high-quality business information quickly and accurately, and under such conditions, translators are required to improve the quality and speed of translation.

According to Pliushch D, there is a close connection and influence between economic and trade English translation and the prototype theory and metaphor of cognitive linguistics [32]. In the translation process, attention should be paid to the selection of different vocabulary, tones, and formats according to the original language text. Bahriddinovna T G points out through his research that it is not easy to translate economic and trade English materials into standard Chinese and at the same time be able to express the meaning of the original text appropriately [33]. Translators not only need to master the characteristics and principles of business English translation, but also need to be familiar with the context of business English. These research results have discussed foreign trade English translation from different perspectives, and all of them are reasonable from their perspectives. It is entirely necessary and possible for foreign trade English to establish a set of scientific and reasonable translation standards to guide the practice of foreign trade English translation. In all these works, numerous examples are presented to express the authors' views on analyzing the linguistic features and specific translation techniques of foreign trade English. The establishment of a set of scientific and reasonable translation standards to guide the practice of foreign trade English translation. So it is very necessary to raise foreign trade English and its translation from some new perspectives to a new level of theoretical research.

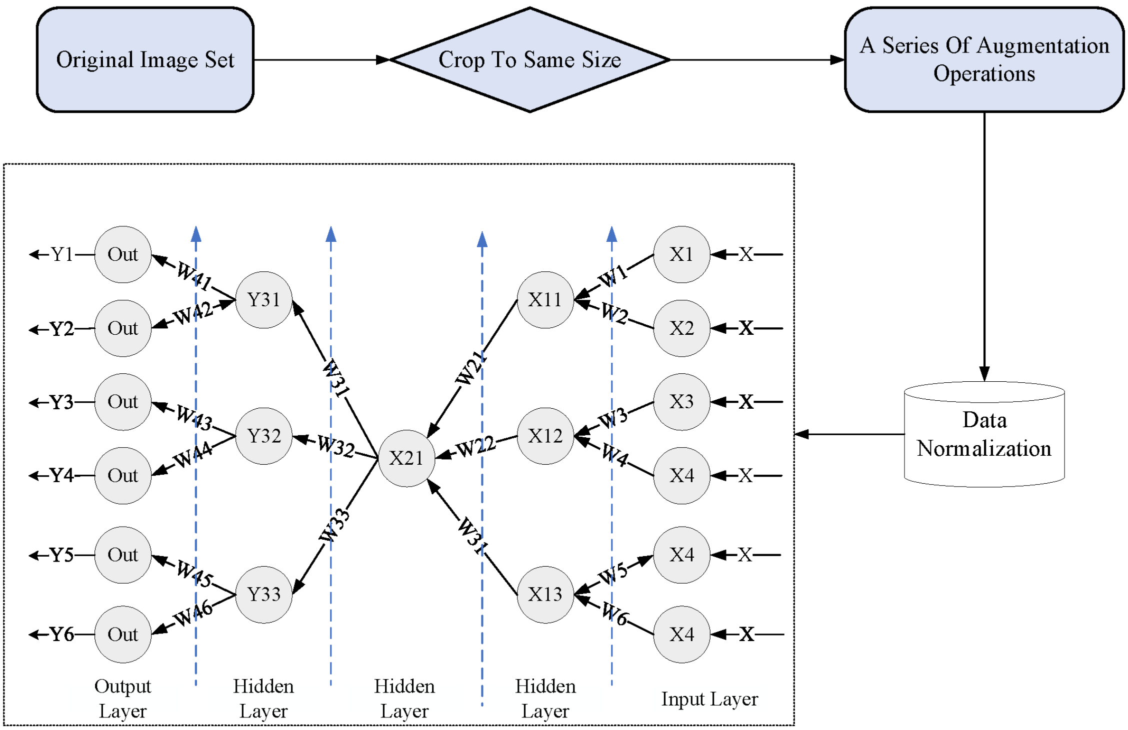

FIGURE 1. Random enhancement process of samples.

III. PRELIMINARY

With the advent of the era of artificial intelligence, deep learning has developed hotly, showing good performance in computer vision, medical care, natural language processing, autonomous driving, speech recognition, and other fields. Deep learning methods, as a branch of machine learning methods, are different from traditional methods in that they can automatically learn the association between tasks and features through deep neural networks and many data samples so that the extraction of complex features can be achieved based on simple features [34]. Deep learning is a simulation of the learning mechanism of the brain, and the newer development of hardware such as GPU also guarantees its implementation, thus achieving better results and even becoming the most respected research direction among scholars. The process of object discrimination in the human brain is roughly like this: the eye serves as the entrance of the signal, and the signal is transmitted to the brain with the help of retinal nerves, which form directional edge features and abstract basic shape features in the lower and lower layers of the brain, respectively, and discriminate the abstract signal through the higher layers to identify the object. Similarly, deep learning is a neural network that extracts low-level features such as edge information of data with the lower layer of the network, then extracts abstract high-level features of data with the higher layer of the network, and finally discriminates the data by the last layer of the network, which reflects the "depth" of deep learning in the process of extracting data features.

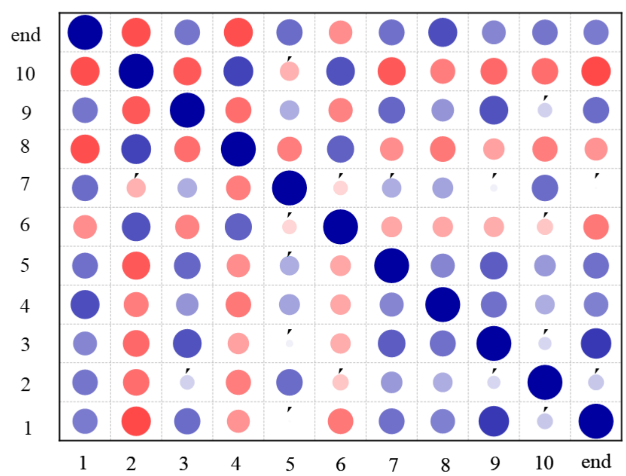

FIGURE 2. Alignment between sequences.

The current model variation scheme for deep learning models has shortcomings such as incomplete variation rule capability and weak applicability of the framework. One of the key advantages of fuzzy testing and other vulnerability mining techniques is that they are fast and can be tested for multiple iterations in a short period, and test case generation is a key step in this process, but the efficiency of its generation is particularly important. Model generation efficiency mainly includes the speed of model generation and the effectiveness of the model generation, because the error model that cannot run only affects the efficiency of testing [35]. Deep learning models are different from testing traditional software with inputs such as text. The generation of models includes the stages of writing, training, and testing of learning programs, and only by choosing reasonable stage variants can the efficiency of model generation be guaranteed. The diversity of models largely determines the effect of testing. Different model inputs may call different codes in the deep learning framework, so the diversity of models determines the degree of code coverage of the framework to be tested. This metric mainly considers the completeness of the model generation set, i.e., the ability to form legal or illegal but valid models by combining different model parameters and structures, etc., and is not limited to the initial seed library provided, but can use the limited initial seeds to eventually generate a collection of models that converge to an approximation.

$$Tan(x) = \frac{1}{e^{x} + e^{-x}} - 2 \tag{1}$$

$$L_i = [l_{i1}, l_{i2}, \ldots, l_{ik}] \tag{2}$$

where L is the length of the neural unit and i is the current time step. The activation function *Tanh* maps real numbers to the interval a between [−1,1] and is a hyperbolic tangent function centered at zero. The effect of this activation function is proportional to the degree of difference between the features. The smaller the difference between the features, the less effective the activation of the *Tanh* is and the weaker the learning ability of the model.

The main task of the preparation phase is to generate toxic data containing backdoor triggers. The method takes a clean sample D_C and a predefined backdoor trigger trig as input and adds the backdoor trigger trig to the frequency domain of the sample D_C. Also, a predefined error label y_T is assigned to the new sample and toxic data D_T is obtained, i.e., $D_T = ((\mathrm{D}_C\ , trig\ ), \mathrm{y}_T)$. This method sets the size of the backdoor trigger smaller than the size of the original sample. Therefore, the backdoor trigger can be injected multiple times to improve the robustness of the backdoor attack while ensuring the backdoor trigger concealment. In the training phase, the attacker merges the generated toxic data D_T with the clean data D_C to obtain the toxic training dataset D_P, i.e., $D_P = D_C \cup D_T$.

$$\min_{\theta} = \sum_{(x,y)}^{n} L\left(f(\mathrm{x})\right) \tag{3}$$

The attacker uses this dataset to train the target model $f(x)$. After training, $f(x)$ outputs predefined error results y_T for samples that contain backdoor triggers, but normal results for clean samples. The random augmentation of samples is mainly designed to address the shortage of training data in certain scenarios. A series of augmentation operations can be performed on the original data, such as rotation, cropping, scaling, and zooming, to expand the scarce original data set for the model training task. Random augmentation can be applied to the training of neural networks in a similar way. Before using an image for model training, a random data enhancement, including cropping, random flipping, resizing, etc., is used to process the original image. This is useful to guide the model to grasp the true class features of the data and improve generalization ability. After the random enhancement is completed, the images need to be resized to fit the data input criteria of the training model. Different models have different input criteria, and finally, the preprocessing operation is completed by normalization. The sample images can be input to the model for training only after the preprocessing operation is completed, as shown in Figure 1.

Shuffle in machine learning and deep learning stands for the operation of disorganizing the data set of a training model. The original data, with a balanced sample, is arranged in a certain order, e.g., the first half is of one category and the second half is of another category. However, the arrangement of the data after scrambling has certain randomness, so that the next time the samples are read in order, the likelihood of getting any type of data is the same [36]. Shuffling is a training technique because machine learning assumes that the data used for the training itself satisfies an independent identical distribution, after which the probability of occurrence of any type of sample is the same. Mixing prevents overfitting of the data by the model and allows the model to learn the most realistic features that the data possesses. If the data used for training is not scrambled, the model will repeatedly learn the features of the data in sequential order and will soon reach a state of overfitting, and will most likely learn only the sequential features of the data, resulting in a model that lacks generalization ability. 120 of the 200 data items are class A and the remaining 80 are class B. Trained under unmixing conditions, the model can learn 100 as the cut-off point, i.e., the first half as A and the second half as B, in a very short learning process, but does not learn the real useful class features of the data.

IV. AUTOMATIC DETECTION MODEL DESIGN FOR ENGLISH TRANSLATION ERRORS

Foreign trade is relative to domestic trade, and it refers to cross-border trading activities between goods produced or services provided in one country and goods produced or services provided in another country, and foreign trade English is the English used in this field. Foreign trade English, also called economic and trade English, specifically includes various English documents involved in import and export trade affairs, including commercial letters, contracts, documents, product specifications, various documents, etc. As a kind of special English, foreign trade English belongs to the category of English for Special Purposes (ESP), which is an English form with the continuous development of the world economy and international trade [37]. Usually, the vocabulary used in foreign trade English is very formal, with more noun phrases and sentences that are complex in structure. Foreign trade English is different from ordinary English. Its language is strict and concise, often using precise vocabulary and fixed structures to convey accurate information. To better

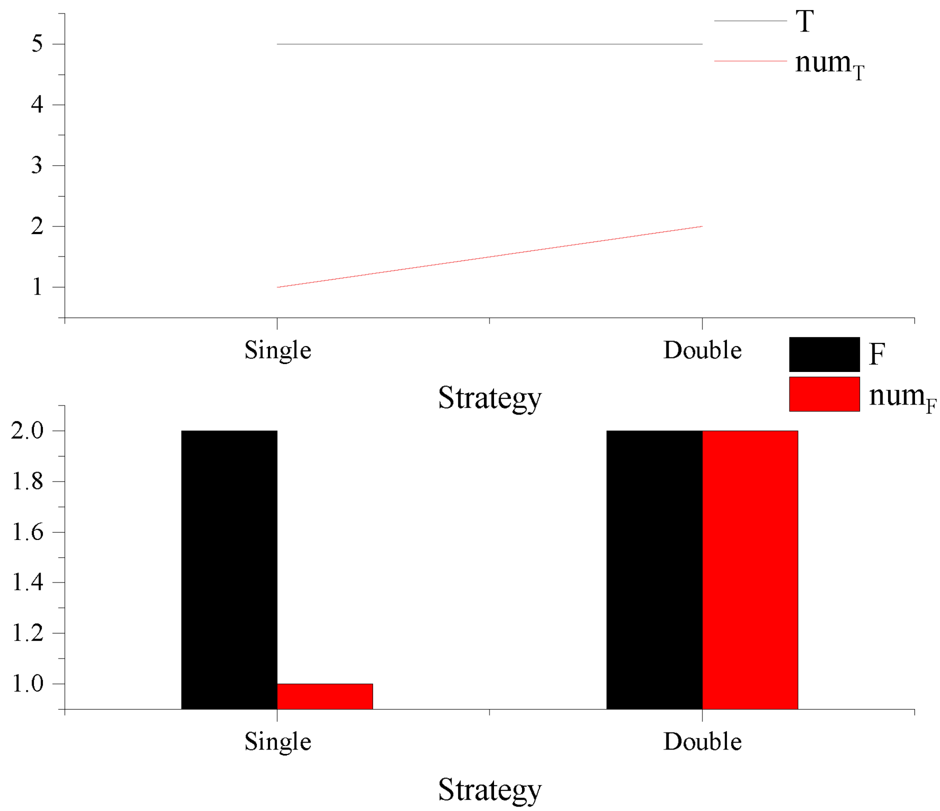

FIGURE 3. **Data enhancement policy parameter settings.**

understand foreign trade English and further discuss its translation standards, we must first understand the characteristics of foreign trade English. The meaning of some words in foreign trade translation is different from that in literary translation, and the words with this feature are called semi-professional words, which are words converted from ordinary words. The use of semi-professional words in foreign trade English is relatively flexible, and attention should be paid to translating their meanings according to specific contexts.

In investigating the mispronunciation detection model, this paper also refers to machine translation, a task in text processing. The length correspondence of the input and output sequences of machine translation is uncertain, and the input and output sequences do not conform to a one-to-one correspondence in temporal order, so a temporal classification algorithm cannot be used. The core of this method is to calculate the weight of the current decoding vector and all encoding vectors, which describes the correlation between the decoding vector and each encoding vector and determines the ‘‘attention'' of the current decoding step to the information encoded at different positions in the original sequence, as shown in Figure 2. Next, the weights calculated in the previous step are used to weight and sum up all the encoding vectors to obtain the context vector corresponding to the current decoding step. Finally, additional alignment information is provided for decoding by stitching the context vectors together. The attention mechanism not only provides additional information for the decoding step, but the mechanism also establishes a direct connection between any encoding node and decoding node, which well solves the long-distance context dependency problem and enables the model to get better learning, and greatly improves the translation quality, which is especially effective on long sentences, making neural network-based machine translation surpass statistical-based machine translation methods for the first time, and becoming a major research hotspot in the field of natural language processing.

In the parallel text collection stage, the author collected many computers security-related bilingual texts. Firstly, all these bilingual texts were converted to word format, and then the original text was separated from the translated text and imported into Abbyy Aligner for bilingual alignment, firstly selecting automatic alignment and then manually adjusting the incorrect alignment of the software, and then selecting the output as a TMX format file after completion. After that, use the online free terminology extraction website, Wordfan Terminology Treasure, to extract terms from the bilingual aligned text in TMX format to generate a bilingual aligned glossary, and then you can import the bilingual aligned glossary and the previously generated translation memory into

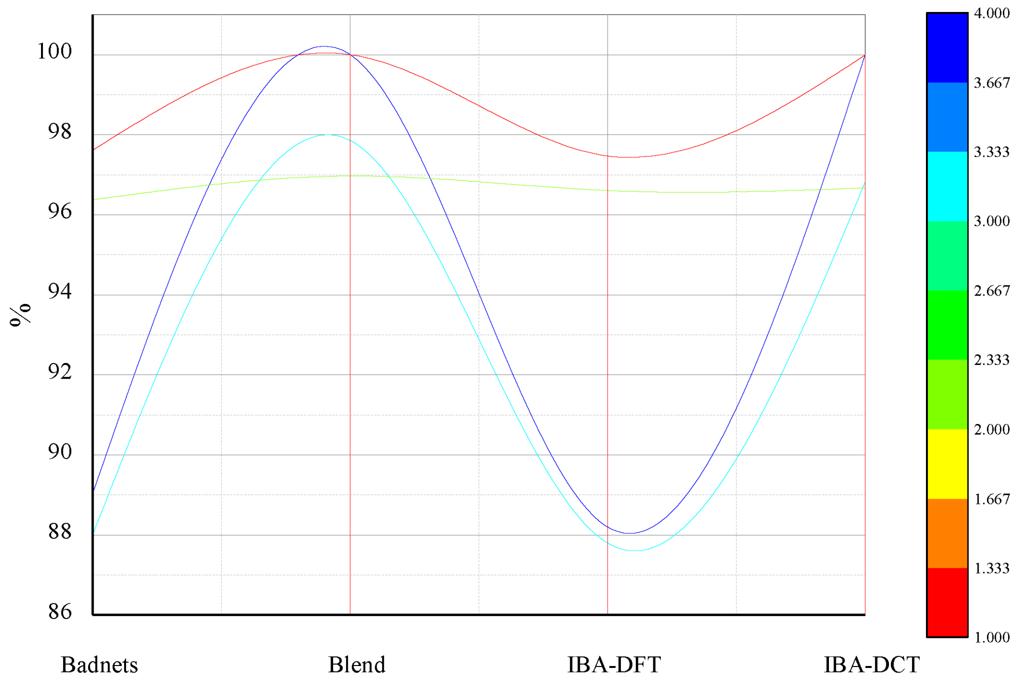

FIGURE 4. Backdoor attack effect comparison results.

the translation software. The bilingual-aligned glossary and memory are equivalent to the translator having a bilingual parallel corpus in the same specialized field as the text to be translated, which can directly improve the efficiency and quality of post-translation editing.

The feature distribution of each audio frame is fitted by Gaussian Mixed Model (GMM), and then the initial probability of phonemes, the observation probability of audio features to phonemes, and the transfer probability between phonemes are modeled by Hidden Markov Model (HMM). First, the phonemes on each audio frame are equally assigned as the initialization of the model parameters, then the GMM-HMM parameters are updated by iteration, and then the updated parameters are used to obtain the optimal decoding result corresponding to the current observation sequence as the next round of labeling by the Viterbi algorithm, and the process is iterated until the parameters of the HMM model converge to complete the alignment. When deep learning models are applied to speech, the phenomenon of overfitting often occurs due to the limited amount of data. Researchers in various fields have proposed a range of data augmentation approaches for creating additional training data. In the speech domain, it is common practice to do variable speed processing on top of the original speech. In this paper, the training data is less than 10 hours, so there is a strong need to use data augmentation to improve the robustness of the training model. Multiple time-domain and frequency-domain masks can be used on the myelogram. In the three myelograms, the first row is the original audio, the audio in the second row uses one time-domain mask and one frequency-domain mask, and the audio in the third row uses two time-domain masks and two frequency-domain masks. The corresponding parameter settings are shown in Figure 3.

In this article, first, create a new word document, then copy the original text into the original text box on the left side of the Google Translate interface, and the translation will be automatically completed on the right side, and then copy the translation into the word document to get the machine translation. Google Translator original text box supports inputting 5000 English letters at one time, and the above steps will be repeated many times to complete the pre-translation. At present, compared with the machine translation based on data statistics, the neural network-based machine translation will automatically adjust the translation according to the grammatical rules of the input language, so that the translation will be smooth and conform to the expression habits of the input language, and the machine translation engine will make appropriate breaks and sequencing adjustment [38]. The original translated materials are mostly scientific and technical English texts, news texts, and processing guides, which have fixed formats and standardized language, and the completion of machine translation is relatively high. However, there are still some problems that need attention, such as there are many punctuation errors in machine-translated

TABLE 1. Results of memory comparison before and after model optimization.

Word vector dimension	Before optimization (GB)	After optimization (GB)	Optimization ratio (%)
128	4.01	0.35	91.27%
256	10.61	0.89	91.61%
512	34.03	1.53	95.50%
1024	24.93	1.02	96.21%

translations. For some specialized terms in some fields, machine translation cannot achieve consistency in terminology throughout the translation; for the phenomenon of multiple meanings of words in English, the machine translation engine does not choose the accurate expression according to the context. In the processing of long and difficult sentences, sometimes the machine translation engine cannot sort out the logical order of the original text. The translation will be ambiguous with the original text. These problems are common problems in machine translation, which need to be corrected manually. The results of memory comparison before and after model optimization are shown in Table 1.

V. RESULTS AND ANALYSIS

To address the problems of stealthiness and poor robustness of backdoor attacks of deep learning models, two backdoor trigger injection algorithms based on frequency domain transform are proposed, namely, backdoor trigger injection algorithm based on discrete Fourier transform and backdoor trigger injection algorithm based on discrete cosine transform [39]. Digital watermarking technology is an information hiding technology that can embed copyright, logo, image, and other information in an invisible way into the carrier data such as video, audio, image, and text. Since the embedded watermark information is usually below the minimum granularity discovery range that the human sensory system can withstand, it can prove the source of data and protect the integrity of data while ensuring the concealment of watermark information. Inspired by digital watermarking techniques, this section proposes a backdoor attack method based on frequency domain transformation by fully considering the effectiveness, stealthies, and robustness of backdoor attacks with deep learning models [40]. The method uses data poisoning to add imperceptible perturbations to a small amount of training data without manipulating the training process and parameter information of the model. To verify the effectiveness of the backdoor attack, this experiment uses toxic samples generated by Badnets, Blend, IBA-DFT, and IBA-DCT to embed backdoors into the target model respectively, and then evaluates the effect of the attack in terms of both the prediction accuracy (CSA) of clean samples and the success rate (ASR) of the attack, and the evaluation results are shown in Figure 4.

Figure 4 shows that the attack success rates of IBA-DFT and IBA-DCT are 96.97% and 96.67% on the GTSRB dataset, and the prediction accuracies of 96.38% and 96.61% on the clean samples. However, on the CIFAR10 dataset, the attack success rate, and the prediction accuracy of clean samples of IBA-DFT and IBA-DCT are better than those of Yadav and Vishwakarma [41]. the attack success rate reaches 100%. Overall, the proposed methods in this paper have good attack results and can meet the requirements of most application scenarios. During the model learning process, to dynamically track the modeling learning effect, the learning effect of the model will be tested after a certain level of training. The test method is that after completing the training, the BLEU value will be evaluated based on the translation results. We integrate the BLEU evaluation session directly into the NMT. This is done as follows: First, a copy script is used to measure the BLEU value on the server where the NMT system is located. Then, the operating system package (imports) is imported into the source code of the training model and the following method is executed in the requested translated source code: os.systemO. This command corresponds to the command executed in SMT, but the location of the Perl script is different. Therefore, we store the evaluation results as a file in the evaluation of the translation results. Only one line stores the content: BLEU = 34.35. Calculate the model size before and after the model memory optimization. If we have a maximum training round of 100,000 and a model saving frequency of 1000, the statistics before and after optimization are shown in Figure 5. The space occupied by model saving increases very quickly as the dimensionality of the word vector increases, while the space occupied after optimization is very small.

If the learning rate is too small, the network loss will drop very slowly or fall into a local minimum. If the learning rate is too large, then the parameter updates are very large, producing an unstable learning curve, or the loss starts to increase directly. After the data is processed, the following work is required to transform the natural language processing problem into a machine learning problem, which starts with digitization. Instead of directly inputting text sentences from the English-Indian or Chinese-Indian parallel corpus into the neural network as text, we first represent each word in the corpus in the form of a vector, and during training, the actual input is a vector set representing the text. The original and most direct representation of word vectors is the one-hot encoding method, which sets the vector length as the size of the vocabulary, and the vector has 0 and 1 values so that when a word is represented, the index position of the vector corresponding to the word is 1 and the other positions are 0. This method of representing word vectors is very simple and easy to use.

The translation of foreign trade English is consistent with the original in terms of semantics and style (language format,

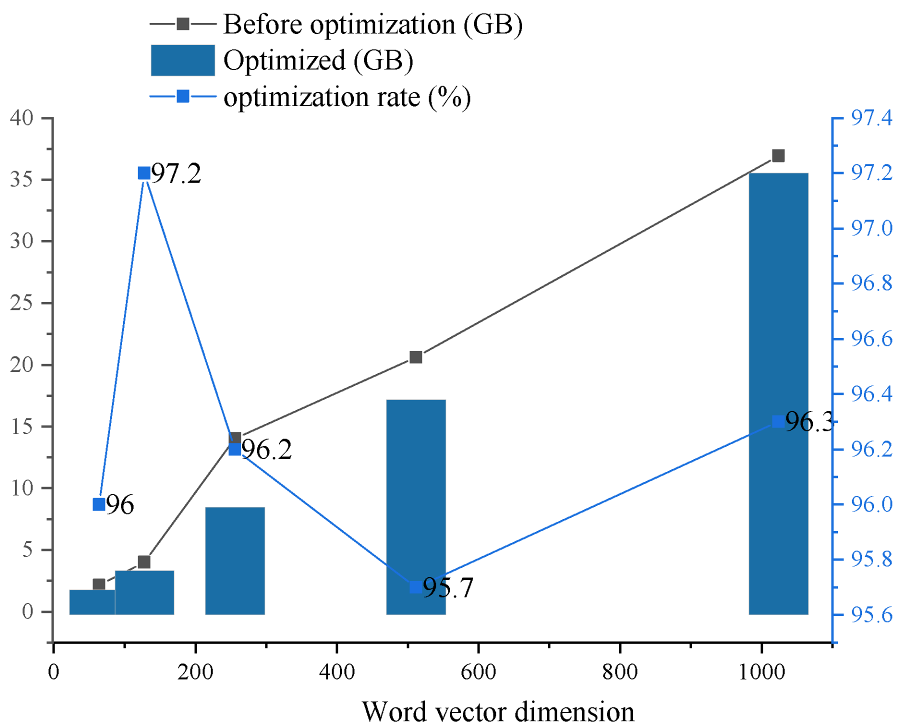

FIGURE 5. Comparison results before and after model storage optimization.

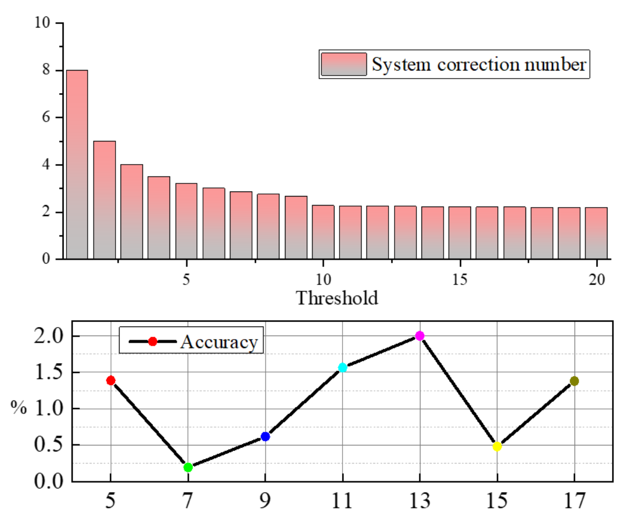

FIGURE 6. Matching test experiment results.

language style, word choice, etc.), i.e., both content and form are consistent. However, the application of this criterion differs in the practice of foreign trade English translation and literary translation due to different genres or text types [42]. The focus of foreign trade translation is to achieve information equivalence, and it is of course best if the language forms can be consistent, but if the expressions of the two languages are too different to maintain the consistency of the language format, the form should be ignored and the essential content of the original text should be expressed. In the process of converting the original language content to the translated language content, foreign trade English translators choose accurate words to express the exact concepts, correct objects, names, and accurate numbers and units. Foreign trade often occurs between trading partners of different countries and cultures, including contracts, insurance, cargo transportation, payment methods, etc. The contents are specific and serious, some of which have legal effects and cannot be arbitrary. Any carelessness or inaccurate expressions can lead to serious consequences. The translation is in line with foreign trade style norms in terms of wording, syntax, format, tone, etc. The wording is appropriate and the translation can maintain the characteristics of the original style and language. The tone should be fair, objective, and strict.

For the user, the user needs to input the sentence and the number of the sentence and choose whether only detection or error correction is required, and the output file of the detection result should be given, and the system writes the result into the file after detection or error correction. The detection of collocation is relatively simple in this paper, when the collocation is classified, it is looked up in the relevant collocation database, and if the number of occurrences of the collocation is greater than a certain threshold, it is considered a correct collocation, otherwise, it is considered as unreasonable collocation. In this class, for the interface of sentence checking, the user needs to input the sentence, the number corresponding to the sentence, choose whether to perform error correction and the system output file stream, after checking, the checking

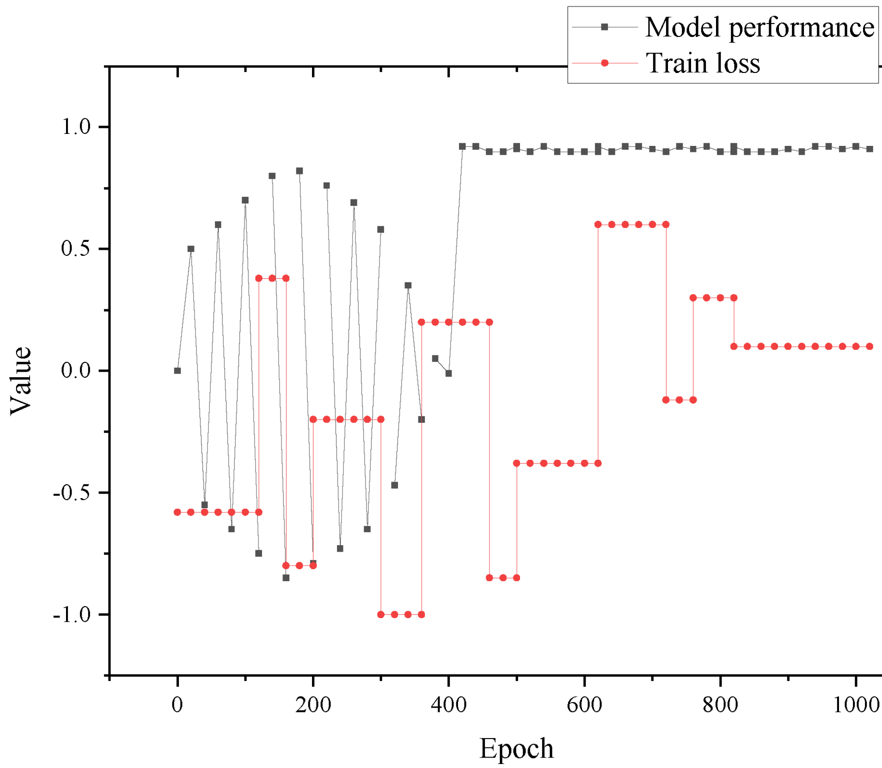

FIGURE 7. Display of model training results.

result will be written to the output file stream. Spell check the sentence and return the correct sentence, and write the spell check result to the output file stream, for the test interface, you can implement the test of various collocation detection in this method. In this subsection, the collocation detection results are experimentally verified. Collocation detection is relatively simple in the system, if the number of collocation occurrences is greater than a certain threshold, the collocation is considered reasonable, if it is less than the threshold, it is considered unreasonable. The test set used in this paper is 1458 sentences from the test set provided by NUCLE, which contains 3764 grammatical errors in the test set, including 512 collocation errors. The results of the collocation detection experiments are shown in Figure 6.

After a long period of development and trade practice, many customary expressions and fixed words and phrases have been formed in foreign trade English, which have been recognized by the industry. As a translator in this field, it is necessary to abide by these customary expressions, so that the translation can accurately convey the gist of the original text and ensure that the reader can correctly understand it. Even if a few translations are incorrect or not perfect, they can still be accepted by businesspeople in the field of foreign trade. Therefore, do not and should not try to change them, because it is useless, futile, and not in line with the industry practice, and more seriously, it is easy to cause unnecessary misunderstandings and differences, leading to trade disputes.

The weights of the features are trained by the MERT algorithm, which first compares and counts the number of errors present in the source statement and the output sentence of the GEC system. The purpose of the algorithm is to find the target sentence with the lowest total number of errors from a series of candidate corrections, where the parameters of the algorithm are optimal. The BERT pre-training model can better solve the long-distance dependency in the context and improve the performance of the model. Second, it can accelerate the training speed of the pre-trained model, reduce the risk of overfitting in small data sets, and improve the language comprehension of the model. In this paper, the BERT-based grammatical error detection model is trained together with the CoLA data set after merging the tagging data, and the final model performance reaches 90% accuracy on the test set, as shown in Figure 7.

This paper focuses on the correction of common error types. The model designed in the paper can successfully detect and correct some common English grammar errors, but there are still some errors that cannot be detected well. The experimental results show that the model has a good correction effect on the noun, preposition, coronin errors, and verb form errors. Among them, the accuracy of correcting verb form errors can reach more than 81.25%. In terms of error type detection, the model has high accuracy in identifying error types for noun singular-plural forms and verb forms.

In this paper, by creating confusion sets of different lexical forms, the synthesized data include errors in verbs, nouns, and other related lexical forms Therefore, the model has a strong advantage in its ability to correct and detect such errors.

VI. CONCLUSION

In this paper, an automatic detection model for foreign trade English translation errors is constructed based on deep learning. The neural network structure of foreign trade English translation is determined and improved in the paper. In terms of data acquisition, single language data with high quantity and quality are obtained as much as possible. The constructed grammatical error detection-correction model improves the error detection and correction ability of the model by 20.12%. By cleaning the raw data, we ensure that the data used for training tests are clean but not over-cleaned. The foreign trade English translation model was trained and optimized for testing on the GTRSB and CIFAR10 datasets, and the results illustrate the usability of the model. The API is designed into a widely used framework, and the training tasks run on the framework achieve the goal of obtaining I/O performance gains without modifying the module code. This paper is a huge research area and limited by time and effort, the work done in this paper can be described as the beginning. During the implementation of the collocation detection and error correction system, many difficulties were encountered, and we are deeply aware of the shortcomings of the work in this paper in some aspects. The collocation database should be improved with more corpus in future research work.

REFERENCES

[1] A. Namdari, M. A. Samani, and T. S. Durrani, "Lithium-ion battery prognostics through reinforcement learning based on entropy measures," *Algorithms*, vol. 15, no. 11, p. 393, Oct. 2022, doi: 10.3390/a15110393.

[2] J. Lee and M. Mitici, "Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics," *Rel. Eng. Syst. Saf.*, vol. 230, Feb. 2023, Art. no. 108908, doi: 10.1016/j.ress.2022.108908.

[3] V. Zouhar, M. Popel, O. Bojar, and A. Tamchyna, "Neural machine translation quality and post-editing performance," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Punta Cana, Dominican Republic, M. Moens, X. Huang, L. Specia, and S. W. Yih, Eds. Stroudsburg, PA, USA: Association for Computational Linguistics, Nov. 2021, pp. 10204–10214, doi: 10.18653/v1/2021.emnlp-main.801.

[4] Z. Guo, K. Yu, A. Jolfaei, G. Li, F. Ding, and A. Beheshti, "Mixed graph neural network-based fake news detection for sustainable vehicular social networks," *IEEE Trans. Intell. Transp. Syst.*, early access, Jul. 7, 2022, doi: 10.1109/TITS.2022.3185013.

[5] D. Lee, J. Ahn, H. Park, and J. Jo, "IntelliCAT: Intelligent machine translation post-editing with quality estimation and translation suggestion," in *Proc. Joint Conf. 59th Annu. Meeting Assoc. Comput. Linguistics, 11th Int. Joint Conf. Natural Lang. Process. (ACL)*, H. Ji, J. C. Park, and R. Xia, Eds. Stroudsburg, PA, USA: Association for Computational Linguistics, Aug. 2021, pp. 11–19, doi: 10.18653/v1/2021.acl-demo.2.

[6] Y. Li, H. Ma, L. Wang, S. Mao, and G. Wang, "Optimized content caching and user association for edge computing in densely deployed heterogeneous networks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 6, pp. 2130–2142, Jun. 2022.

[7] T. Hong-Viet, N. Van-Vinh, and N. Hoang-Quan, "Improving machine translation quality with denoising autoencoder and pre-ordering," *J. Comput. Inf. Technol.*, vol. 29, no. 1, pp. 39–56, 2021. [Online]. Available: http://cit.fer.hr/index.php/CIT/article/view/5316

[8] P. Zhang, B. Yang, H. Wei, D. Liu, K. Fan, L. Si, and J. Xie, "Competency-aware neural machine translation: Can machine translation know its own translation quality?" in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, Abu Dhabi, United Arab Emirates, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds. Stroudsburg, PA, USA: Association for Computational Linguistics, Dec. 2022, pp. 4959–4970. [Online]. Available: https://aclanthology.org/2022.emnlp-main.330

[9] S. K. Mahata, A. Garain, D. Das, and S. Bandyopadhyay, "Simplification of English and Bengali sentences for improving quality of machine translation," *Neural Process. Lett.*, vol. 54, no. 4, pp. 3115–3139, Aug. 2022, doi: 10.1007/s11063-022-10755-3.

[10] I. Rivera-Trigueros, "Machine translation systems and quality assessment: A systematic review," *Lang. Resour. Eval.*, vol. 56, no. 2, pp. 593–619, Jun. 2022, doi: 10.1007/s10579-021-09537-5.

[11] S. Eo, C. Park, H. Moon, J. Seo, and H. Lim, "Word-level quality estimation for Korean-English neural machine translation," *IEEE Access*, vol. 10, pp. 44964–44973, 2022, doi: 10.1109/ACCESS.2022.3169155.

[12] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electron. Markets*, vol. 31, no. 3, pp. 685–695, Apr. 2021.

[13] G. Ranganathan et al., "A study to find facts behind preprocessing on deep learning algorithms," *J. Innov. Image Process.*, vol. 3, no. 1, pp. 66–74, Apr. 2021.

[14] Z. Zhou, Y. Su, J. Li, K. Yu, Q. M. Jonathan Wu, Z. Fu, and Y. Shi, "Secret-to-image reversible transformation for generative steganography," *IEEE Trans. Dependable Secure Comput.*, early access, Oct. 27, 2022, doi: 10.1109/TDSC.2022.3217661.

[15] L. Zhao, Z. Yin, K. Yu, X. Tang, L. Xu, Z. Guo, and P. Nehra, "A fuzzy logic-based intelligent multiattribute routing scheme for two-layered SDVNs," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 4, pp. 4189–4200, Dec. 2022.

[16] S. Xia, Z. Yao, Y. Li, and S. Mao, "Online distributed offloading and computing resource management with energy harvesting for heterogeneous MEC-enabled IoT," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6743–6757, Oct. 2021.

[17] L. Zhao, Z. Bi, A. Hawbani, K. Yu, Y. Zhang, and M. Guizani, "ELITE: An intelligent digital twin-based hierarchical routing scheme for softwarized vehicular networks," *IEEE Trans. Mobile Comput.*, early access, May 31, 2022, doi: 10.1109/TMC.2022.3179254.

[18] L. Yang, Y. Li, S. X. Yang, Y. Lu, T. Guo, and K. Yu, "Generative adversarial learning for intelligent trust management in 6G wireless networks," *IEEE Netw.*, vol. 36, no. 4, pp. 134–140, Jul. 2022, doi: 10.1109/MNET.003.2100672.

[19] Q. Zhang, Z. Guo, Y. Zhu, P. Vijayakumar, A. Castiglione, and B. B. Gupta, "A deep learning-based fast fake news detection model for cyber-physical social services," *Pattern Recognit. Lett.*, vol. 168, pp. 31–38, Apr. 2023.

[20] Z. Guo, K. Yu, A. Jolfaei, F. Ding, and N. Zhang, "Fuz-spam: Label smoothing-based fuzzy detection of spammers in Internet of Things," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 11, pp. 4543–4554, Nov. 2022.

[21] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 766–775, Apr. 2020.

[22] C. Chen, Z. Liao, Y. Ju, C. He, K. Yu, and S. Wan, "Hierarchical domain-based multicontroller deployment strategy in SDN-enabled space–air–ground integrated network," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 6, pp. 4864–4879, Dec. 2022.

[23] Z. Cai, X. Zheng, and J. Yu, "A differential-private framework for urban traffic flows estimation via taxi companies," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6492–6499, Dec. 2019.

[24] Z. Guo, Y. Shen, S. Wan, W. Shang, and K. Yu, "Hybrid intelligence-driven medical image recognition for remote patient diagnosis in Internet of Medical Things," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 12, pp. 5817–5828, Dec. 2022.

[25] Z. Cai and Q. Chen, "Latency-and-coverage aware data aggregation scheduling for multihop battery-free wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1770–1784, Mar. 2021.

[26] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," *IEEE Trans. Image Process.*, vol. 13, no. 1, pp. 87–99, Jan. 2004.

[27] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[28] Y. Xu, J. Du, L. Dai, and C. Lee, "A regression approach to speech enhancement based on deep neural networks," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 1, pp. 7–19, Jan. 2015.
[29] Z. Zhao and H. Wang, "MaskGEC: Improving neural grammatical error correction via dynamic masking," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 1, pp. 1226–1233.
[30] N. Rusk, "Deep learning," *Nature Methods*, vol. 13, no. 1, p. 35, 2016.
[31] H. Hong-Hua, "Research on teaching reform of 'foreign trade correspondence' for business English majors—Based on regional features," *Sino-US English Teach.*, vol. 17, no. 9, pp. 275–279, 2020.
[32] D. Pliushch, "Gravity model of foreign trade between Ukraine and EU countries," *Eur. J. Econ. Manag.*, vol. 6, no. 1, pp. 16–23, 2020.
[33] T. G. Bahriddinovna and J. N. Akhmatovna, "The coverage of the history of domestic and foreign trade relations of the emirate of Bukhara in the memoirs of ambassador PI Demezon and IV Vitkevich," *IJAR*, vol. 6, no. 6, pp. 49–51, 2020.
[34] V. Martyniuk and Y. Muravska, "Forming a foreign trade partnership strategy in the context of strengthening national economic security: A case study of Ukraine," *Forum Scientiae Oeconomia*, vol. 8, no. 2, pp. 5–24, 2020.
[35] M. Kaneko and M. Komachi, "Multi-head multi-layer attention to deep language representations for grammatical error detection," *Computación y Sistemas*, vol. 23, no. 3, pp. 883–891, Oct. 2019.
[36] J. Fu, Y. Chiba, T. Nose, and A. Ito, "Language modeling in speech recognition for grammatical error detection based on neural machine translation," *Acoust. Sci. Technol.*, vol. 41, no. 5, pp. 788–791, 2020.
[37] Y. Wang, Y. Wang, K. Dang, J. Liu, and Z. Liu, "A comprehensive survey of grammatical error correction," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 5, pp. 1–51, Oct. 2021.
[38] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: model-based, AI-based, or both?" *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.
[39] N. Sharma, R. Sharma, and N. Jindal, "Machine learning and deep learning applications—A vision," *Global Transitions*, vol. 2, no. 1, pp. 24–28, 2021.
[40] G. Eraslan, Ž. Avsec, J. Gagneur, and F. J. Theis, "Deep learning: New computational modelling techniques for genomics," *Nature Rev. Genet.*, vol. 20, no. 7, pp. 389–403, Jul. 2019.
[41] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: A review," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4335–4385, Aug. 2020.
[42] A. P. Pandian, "Performance evaluation and comparison using deep learning techniques in sentiment analysis," *J. Soft Comput. Paradigm*, vol. 3, no. 2, pp. 123–134, 2021.

WEIHONG ZHANG was born in Henan, China, in 1974. She received the bachelor's and master's degrees from Zhengzhou University, in 1999 and 2007, respectively. She has published a total of 13 articles. Her research interests include machine translation and translation teaching.