Research Article

Application of Convolutional Neural Network Based on Deep Learning in College English Translation Teaching Management

Fei Ma ⓘ

Sanquan College of Xinxiang Medical University, Xinxiang, Henan, China

Correspondence should be addressed to Fei Ma; 14582153@sqmc.edu.cn

With the increased demand for English translators in recent years, more scholars and researchers have begun to focus on college English translation education. The demand for computers to perform automatic natural language translation is at an all-time high, presenting another potential for machine translation (MT) research. Although neural machine translation (NMT) technology has advanced significantly, it still has several flaws. The reply NN-based MT technology is not optimal in the field of long sentence translation, and there are situations such as missing translation and overtranslation. So this paper proposes college English translation based on a convolutional neural network (CNN). This paper proposes that CNNs have stronger feature extraction and information processing capabilities than other deep NNs. It successfully solves the situation of missing translation and over-translation, lowering the quality of college English translation instruction. The experimental results of this research reveal that when the length of an English sentence grows longer, the translation accuracy based on the reply NN drops to 47.9%, while the translation accuracy based on the CNN is 89.7%. It can be seen that college English translation based on CNN can be integrated into English translation teaching as an innovation of traditional teaching methods, so as to create a learning environment for students and enhance learning effects. MT based on CNN has greatly improved students' performance, which is conducive to improving the quality of college English translation teaching.

1. Introduction

English classroom teaching is one of the tasks that must be broken through. In today's international financial and commercial exchanges, the requirements for comprehensive English application ability are constantly increasing. Among them, college English translation ability plays a key role. In recent years, the demand for English translation talents has been continuously increasing, and the problem of translation education has gradually attracted the attention of the academic circles. In college English education, translation education has always been concerned, but its educational effect is not ideal.

As a result, English teachers and scholars are concerned about how to improve students' translation abilities. This study suggests that English majors be taught to translate using a college English translation approach based on CNNs. However, NMT brings benefits to the field of MT. However,

NMT only converts between natural languages using a single NN, which has a number of flaws. This research uses an MT model based on CNNs to increase MT performance.

The translation is determined as a compulsory course for English majors and is considered to be a very practical course. The cultivation of translation ability is one of the comprehensive language abilities. With the constant development of computer information technology, according to Yao, people's lives have been significantly altered. Computer-aided translation technology (CAT) is currently being used as an important supplementary tool, including attempts to adapt it to the teaching process [1]. Wang studied the creation of English translation training using computer-aided translation software in depth. With the help of corpus statistics software, he performed a complete data analysis of the translation and the original text [2]. Najjar discussed the translation of the Qur'an in Chinese and English and studies the morphological transformation of

hyperbolic patterns. He sought to determine the impact of translation strategies on the quality of the translation of research data [3]. In cultural translation, Zhang discovered that people see translation as a cross-cultural communicative activity. In diverse cultural situations, translation thought will likewise shift dramatically. The translation industry places a high value on cultural differences, as evidenced by recent translation studies [4]. Fitriani's research tried to classify and describe grammatical faults in English-translated sentences in terms of syntax and morphology. The surface strategy taxonomy is used to guide the analytical approach [5]. Susini found that implicit meaning is one of the linguistic phenomena that need to be overcome in translation. His research aims to investigate the structure in which implicit meaning is realized in English [6]. Although scholars agree that English translation is very important, there is a lack of research on how to improve the effect of college English translation.

MT is the process of converting sentences from a source language to a target language using a computer. MT helps people communicate in a variety of languages. Chen proposed the Residual Encoder-Decoder CNN, which was inspired by deep learning (RED-CNN). His proposed RED-CNN model produced excellent results [7]. Kruthiventi found the discovery that understanding and predicting human visual attention mechanisms has received neuroscience attention. Its large receptive field can understand semantics at multiple scales while taking into account the global context [8]. Lu created a deep learning technique that uses graph cut refinement to partition the liver in CT scans automatically. For precise refining of the initial segmentation, it leverages graph cuts and previously learned probability maps [9]. Deep learning is widely employed in a variety of fields and has produced numerous results.

In recent years, scholars have begun to consider from a new perspective of deep learning regarding many of the main modules retained in MT tasks. Many studies have shown that deep learning can effectively solve various bottlenecks of previous MT methods. As a result, NMT received attention immediately after it was proposed, greatly improving its performance in just a few years. The innovation of this paper is to analyze how the CNN plays a role in MT technology, thereby improving the teaching quality of college English translation. Experimental analyses of deep learning-based revertive networks and CNNs are carried out.

2. CNN MT Technology Based on Deep Learning

The ultimate goal of language learning is communication, and translation is one of the most important forms of cross-cultural communication for foreign language students. As a basic skill for comprehensive language use, translation ability is very necessary for English majors. Because the future employment of college students will require them to work in various fields, learn the latest achievements of other countries in various fields, or show China to the world the latest achievements in many fields, so college English translation is crucial for them.

MT is one of the best ideas of scientists and all human beings in the early days of computers. It has generated new vitality under the urgent needs of world economic development and cultural exchange [10]. MT has gone through a tortuous development path and has made great progress. The current NN MT technology includes revertive NN, recurrent NN, feedforward NN, and CNN. However, from the current situation, there are still many problems in MT that need to be solved as soon as possible [11]. The structure of NN MT is shown in Figure 1.

Deep learning is a relatively recent research area in machine learning. It was added to machine learning to help it achieve its original goal of artificial intelligence. Figure 1 shows how deep learning has advanced rapidly in the disciplines of image processing and speech recognition since 2013. Deep learning has recently been discovered to be more effective at reducing the problem of linear inseparability, lack of correct semantic representation, and functional design. It can help statistical MT overcome the challenges of fully utilizing nonlocal context, data sparseness, and error propagation. MT is a popular topic in academia right now [12].

2.1. Algorithm Based on Revertive NN.
In the traditional feedforward NN model, the adjacent layers are fully connected, while the nodes in each layer are disconnected. A revertive NN is a network that contains both feedforward and feedback pathways. Its feedforward pathway is similar to the traditional feedforward NN model. The feedback pathway can reflect the output of some neurons to itself at a later time as the input of a new time [13]. The schematic diagram of the revertive NN is shown in Figure 2.

As demonstrated in Figure 2, feed forward NNs are the simplest type of artificial NN. According to the different layers of feed forward NN, it can be divided into single-layer feed forward NN and multilayer feed forward NN. The revertive NN, unlike the standard feed forward NN, pays attention to the timing input. It adds a cyclic circle to the hidden layer, and this cyclic circle signifies that the hidden layer's output from the previous moment is used as the hidden layer's current input. The calculation formula of the revertive NN h_t is as follows:

$$h_t = \sigma_h (W_h a_t + U_h h_{t-1} + b_h). \tag{1}$$

This way of connecting the revertive NN has obvious advantages for the sequence input U_h. The network can better capture timing information and obtain context information. Therefore, the reply NN is especially suitable for the fields of speech, text, and video processing that focus on timing. The field of speech includes speech recognition and language conversion. The text domain includes text classification and MT. The video domain includes video recognition and video classification [14].

It is precise because of the special connection method of the revertive NN that the revertive NN has a serious problem; that is, it cannot solve the long-term dependence. This is due to the continuous multiplication of gradients in the revertive NN when the error is backpropagated.
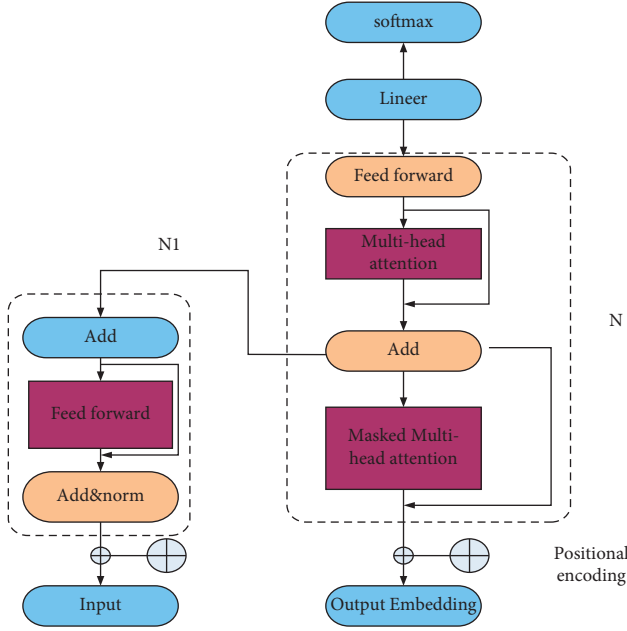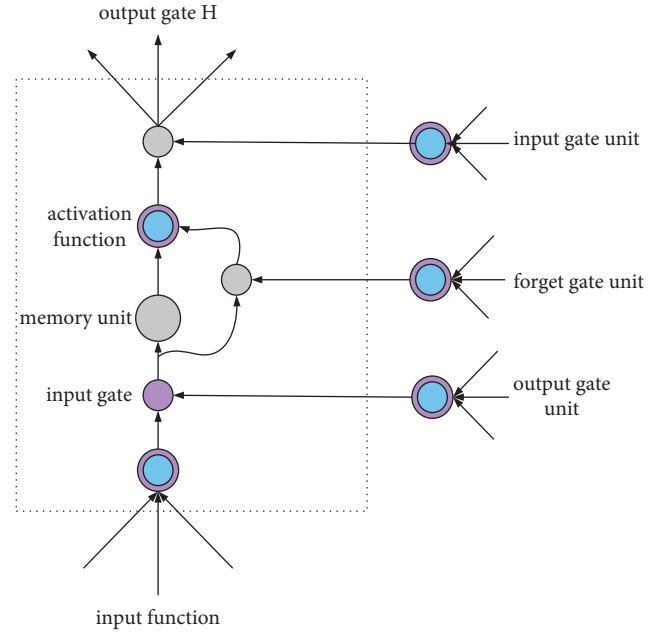
FIGURE 1: The architecture of NN MT.
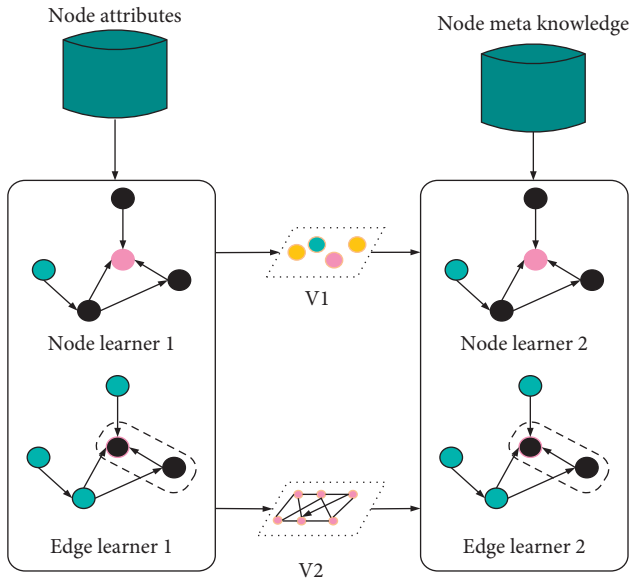

FIGURE 3: LSTM network.

FIGURE 2: Revertive NN.

Therefore, when the gradient is less than 1, the gradient is prone to disappear, and when the gradient is greater than 1, it is prone to gradient explosion [15]. The Long-Short Term Memory (LSTM) was created to overcome the revertive NN's long-term reliance problem, and the network was enhanced over the classic revertive NN. In science and technology, LSTMs have a wide range of applications. And, more are all tasks that LSTM-based systems may learn. Figure 3 depicts the LSTM network.

The LSTM network differs from a standard revertive NN in that the original revertive NN unit is changed into a CEC memory unit, as shown in Figure 3.

To tackle the problem of gradient dispersion, the CEC memory unit's summation technique allows the gradient to

be kept while the mistake is conveyed. In the LSTM network, three gates are introduced, whereas the forget gate determines whether or not the input should be forgotten [16, 17]. The forget gate's job is to pick which parts of a long-term memory (the output from the previous unit module) to keep and which to discard.

The input gate i_t is as follows:

$$i_t = \sigma_g \left(W_i a_i + u_i h_{t-1} + b_i \right). \tag{2}$$

The forget gate f_t is as follows:

$$f_t = \sigma_g \left(W_f a_t + U_f h_{t-1} + b_f \right). \tag{3}$$

The output gate o_t is as follows:

$$o_t = \sigma_g \left(W_o a_t + U_o h_{t-1} + b_o \right). \tag{4}$$

Here, i is the input gate, f is the forgetting gate, and o is the output gate. After the calculation of the three gates is completed, the memory unit is updated. After the widespread use of LSTM networks, many variants of traditional revertive NNs have followed, including GRU networks [18].

GRU is a well-performing LSTM network variation. It has a simpler structure than an LSTM network and produces excellent results. The GRU network is different from the LSTM network which contains three gates. There are two gates in it: an update gate and a reset gate. The update gate determines the percentage of input and memory in the hidden layer's output. There are two gates in it: an update gate and a reset gate. The update gate determines the percentage of input and memory in the hidden layer's output [19].

It consists of two gates: an update gate and a reset gate. The percentage of input and memory in the hidden layer's output is determined by the update gate.

The update gate and reset gate are calculated by

$$z_t = \sigma(W_z a_t + U_z h_{t-1}), \tag{5}$$

$$r_t = \sigma(W_r a_t + U_r h_{t-1}). \tag{6}$$

There are many defects such as missing translation, mistranslation, overtranslation, and so on, which lead to the deterioration of translation results [20].

2.2. MT Based on CNNs.

CNNs are a type of feedforward neural network (NN) with deep structures and convolutional computations. It is one of the most representative deep learning algorithms. Then, a convolutional neural-based MT model is proposed for the long sentence translation defect as shown in Figure 4.

As shown in Figure 4, CNNs (CNN) are similar to regular NNs in that they are made up of neurons that have trainable and learnt weights and bias constants. CNN models are usually built on feedforward NN models. The difference is: the "hidden layers" of traditional NNs are replaced by "convolutional layers," "pooling layers," and "completely linked layers." This unique structure enables CNNs to perform exceptionally well as shown in Figure 5.

As shown in Figure 5, the activation layer is followed by the convolution layer, which is similar to the role of the activation function in the NN. It generally uses the ReLu activation function and can also be other activation functions such as Sigmoid.

The calculation formula of the neuron is as follows:

$$F(a) = f\left(\sum_{i,j}^{n} \theta_{(n-i)(n-j)} a_{ij} + b\right), \tag{7}$$

Among them, θ is the parameter of the convolution kernel.

Some researchers proposed a new model of NMT in 2016 and successfully applied CNNs (CNN) to MT. The Encoder-Decoder model based on the reply NN is similar in concept.

A CNN of the model is responsible for the Encoder encoding to convert the source sentence into an intermediate vector. Another CNN is responsible for Decoder decoding, decoding the intermediate vector to generate the target sentence. The high-level CNN abstractly extracts long-distance information, realizes NMT, and has achieved outstanding achievements, which has brought significant development to the field of MT by CNNs.

2.2.1. Encoder Coding Based on CNN.

The translation is divided into two parts: understanding and expression. Encoder encoding is responsible for understanding, and Decoder decoding is responsible for expression. The input layer represents the input to the model. The input of the Encoder model is the sequence number of the word corresponding to the vocabulary after the segmentation and generalization of the source sentence as shown in Figure 6.

As shown in Figure 6, the vocabulary is constructed based on the training data set. The embedding layer refers to the process of representing words as vectors. After the embedding layer, an input sequence number is
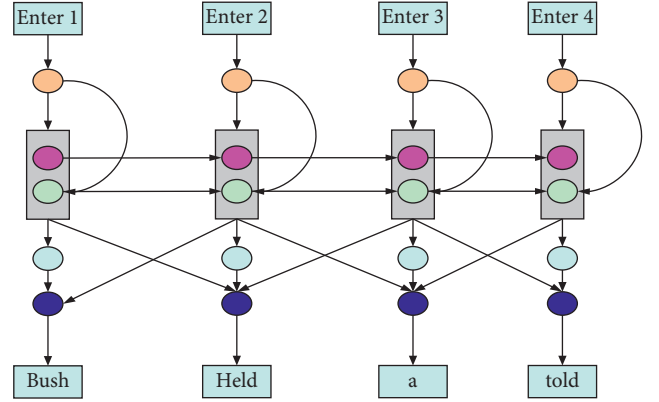

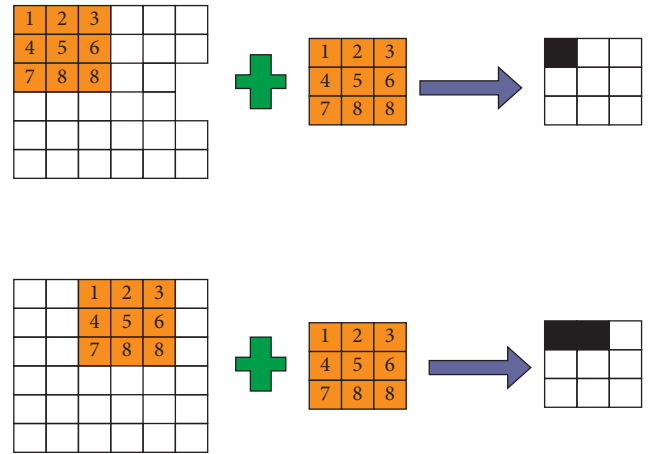FIGURE 4: Convolutional neural-based MT model.
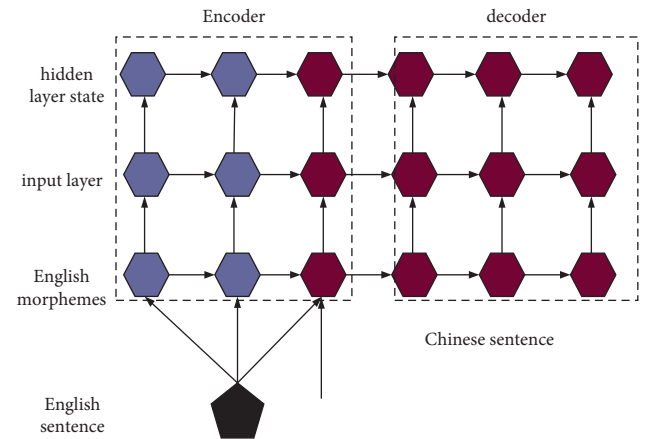

FIGURE 5: Convolution calculation process.

FIGURE 6: Encoder-decoder block diagram.

correspondingly converted into a vector of a specified dimension. Embedding layer $e_{e,t}$ is calculated as follows:

$$e_{e,t} = e_e(a_t), \tag{8}$$

The embedding layer uses random initialization to assign the initial value a_t to the word vector. After training, the required word vector e_e is obtained to realize the representation of words with vectors.

The proposal of the hidden layer is proposed along with the concept of a multilevel network, which mainly solves a linear inseparable problem. The bidirectional hidden layer actually contains two layers of network, one is a forward LSTM network and the other is a reverse LSTM network. The bidirectional LSTM network aims to capture the input sentence information more comprehensively from two different perspectives of positive order and reverse order, so as to achieve a better understanding of the input sentence.

To aid in generating the output of the current moment, the current moment calculation is added to the output of the LSTM network at the previous instant. The network captures the above-given information $h_{f,t}$ in positive sequence, such as the following formula:

$$h_{f,t} = h_f\left(e_{e,t}, h_{f,t-1}\right). \tag{9}$$

The network captures the above information in reverse order, as in the following formula:

$$h_{b,t} = h_b\left(e_{e,t}, h_{b,t+1}\right). \tag{10}$$

The output $e_{e,t}$ of the forward LSTM network is fused with the output $h_{b,t+1}$ of the reverse LSTM network to obtain the output of the bidirectional hidden layer. It captures the context information and obtains the final vector representation, which is calculated as follows:

$$h_{bt} = f\left(h_{f,t}, h_{b,t}\right). \tag{11}$$

This $h_{f,t}$ vector will serve as the input to the next hidden layer. f takes the concatenation function.

In the discipline of natural language processing, deep networks attempt to extract abstract high-level features of language in order to gain a deeper comprehension of language. The network is calculated using the following formula:

$$h_{s,t} = h\left(h_t, h_{s,t-1}\right). \tag{12}$$

In practical applications, it can be considered to continue to increase the number of hidden layers $h_{s,t-1}$ to obtain higher-level features according to requirements.

2.2.2. Decoder Decoding Based on CNN.

The Encoder is pretrained through a NN to determine the initial value of W. Its goal is to make the input value equal to the output value. After Encoder encoding, the sequence of the output of the hidden layer corresponding to the input of the source sentence at each moment is obtained. It uses the attention mechanism network to compute contextual information. After Encoder encoding, the kernel state of the hidden layer is obtained. The Decoder part of the model is shown in Figure 7.

Figure 7 shows that the input of the decoder model is different in the training learning phase and the testing application phase. In the process of model testing, because the correct target sentence is unknown, the input of the model at the first moment is still a special word, which represents the beginning of the decoding process. The input
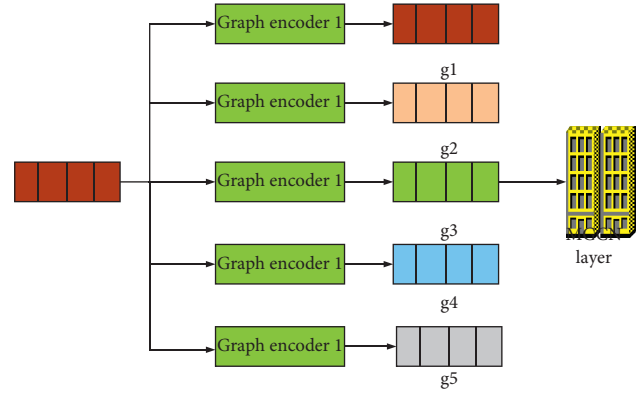

FIGURE 7: Decoder network structure.

at each other moment is the final output of the network at the previous moment.

To a certain extent, it is used for dimensionality reduction. The principle of dimensionality reduction is matrix multiplication. In convolutional networks, it can be understood as a special fully connected layer operation. This is the role of the embedding layer. The embedding layer refers to the process of representing words with vectors, which is the same as the embedding layer of the Encoder network and corresponds to different vocabulary. After the embedding layer, an input sequence number is correspondingly converted into a vector of a specified dimension. The calculation of the embedding layer $e_{d,t}$ is shown in the following formula:

$$e_{d,t} = e_d\left(b_t\right). \tag{13}$$

In the hidden layer, the forward LSTM network is employed, and the number of layers can be modified depending on the situation. Two layers of LSTM networks are hidden in the hidden layer of the MT model's block principle-based Decoder network. The LSTM network's kernel is launched by the Encoder network's second hidden layer's kernel.

Attention models are commonly employed in natural language processing, image identification, and audio recognition, among other deep learning applications. It is an important technology that ought to be studied in depth. The attention mechanism network is found in the Decoder network's hidden layer. The output of the hidden layer at the previous moment of the Decoder network serves as the basis for the attention mechanism. It constantly mixes the output of the Encoder network to calculate and provide context information appropriate for the present state of the Decoder network.

For the calculation of the current network context information,

$$e_{ti} = a_f\left(s_{f,(t-1)}, h_{s,i}\right), \tag{14}$$

$$\alpha_{ti} = \frac{\exp\left(e_{ti}\right)}{\sum_{j=1}^{T} \exp\left(e_{tj}\right)}, \tag{15}$$

$$c_t = \sum_{j=1}^{T} \alpha_{ti} h_{s,i}. \tag{16}$$

And, all e_{tj} are normalized to get the contribution of the output of the Encoder network at each moment to the calculation of the hidden layer of the Decoder network at the current moment. According to the contribution calculation, the context information c_t corresponding to the current moment is obtained.

The output of the embedded layer at the current moment, the network output is calculated as follows:

$$s_{f,t} = s_f \left(e_{d,t}, s_{f,t-1}, c_t \right). \tag{17}$$

Each layer of the multilayer LSTM network is roughly the same, the network structure is the same, and the difference lies in the specific input s_f and output $s_{f,t-1}$.

The output of the final hidden layer is as follows:

$$s_{s,t} = s_s \left(s_{f,t}, s_{s,t-1}, c_t \right). \tag{18}$$

The normalized exponential function is commonly known as the softmax function. In multiclassification, it is a generalization of the binary classification function sigmoid. The goal is to present the findings of multiclassification as probabilities. The output layer is a completely linked layer with an output dimension equal to the target language vocabulary's dimension vocabulary.

The final output is activated by the Softmax function, as shown in the following formula:

$$o_t = o \left(s_{s,t} \right). \tag{19}$$

After Encoder encoding and Decoder decoding, back-propagation updates the network parameters, fits the training data, and captures MT features.

2.2.3. Cross Entropy Loss.

In classification issues, the cross-entropy loss function is frequently utilized. Cross-entropy is frequently utilized as a loss function in NN classification issues.

In classification issues, the cross-entropy loss function is frequently utilized. Cross-entropy is frequently utilized as a loss function in NN classification issues. In machine learning and deep learning, entropy is a measure of the uncertainty of random variables. Assume an is a discrete random variable with a finite number of potential values, and the following is its probability distribution:

$$p \left(A = a_i \right) = p_i, \ i = 1, 2, \ldots, n. \tag{20}$$

Then, the entropy of random variable a is calculated as follows:

$$H(A) = - \sum_{i=1}^{n} p_i \log p_i. \tag{21}$$

Here, the base of log is 2 or the natural logarithm, and if p_i is 0, the value of log0 is defined as 0. The distance between two random distributions is measured by cross-entropy.

Assuming that random variable A has a value set of U and that its two distributions are p and q, the cross entropy of the two random distributions is determined as follows:

$$CEH(p, q) = - \sum_{a \in U} p(a) \log q(a). \tag{22}$$

In deep learning, cross-entropy is frequently used to calculate the similarity between the network output and the ground truth. It calculates the cross-entropy loss to determine the current error and adjusts the parameters accordingly.

2.2.4. BLEU Score.

A variety of automatic evaluation standards for translation technologies have been proposed. At present, the widely used and recognized evaluation standard is to use the BLEU algorithm for scoring and discrimination.

The BLEU algorithm is calculated as follows:

$$\text{BLEU} = BP \times \exp \left(\sum_{n}^{N} w_n \log p_n \right), \tag{23}$$

where N represents the number of N-gram models, and w_n represents the weight of the corresponding N-gram model, usually $1/N$. p_n in the formula represents the matching accuracy of its corresponding model. The BP is shown in the following formula:

$$BP = \begin{cases} 1, \text{ if } c \succ r, \\ \exp \left(1 - \dfrac{r}{c} \right), \text{ if } c \leq r. \end{cases} \tag{24}$$

The length of the to-be-evaluated translation is C, and the length of the reference translation is r. The length penalty factor BP is connected to the relationship between c and r and is a piecewise function.

Since any n-gram model does not match, the BLEU value in this case is 0, which is meaningless. Therefore, the BLEU algorithm is not suitable for measuring the translation of a single sentence but is suitable for evaluating the translation of many sentences.

3. College English Translation Teaching Experiment Based on CNN

3.1. Comparative Experiment of CNN and Other English Translation Models.

The experiment employs Ubuntu14.04LTS as the experimental environment and LuaJIT for NN initialization and NMT model development. The NMT model, which is based on a CNN, is then tweaked and built accordingly. This section mostly compares and examines the outcomes of similar investigations.

The application data come from the machine text translation public dataset of the Global AI Challenge. The original dataset contains a total of 10 million Chinese-English parallel sentence pairs. It limits the source and target language sentences in the training data to no more than 60 words.
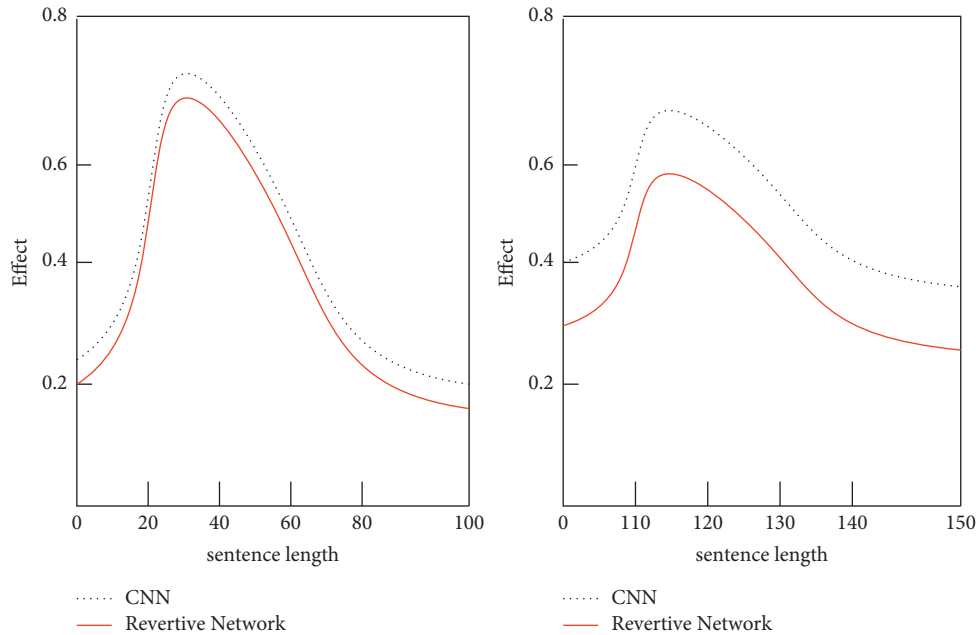
FIGURE 8: Translation performance of different models on different sentence length distributions.

The two MT models were trained under different sentence lengths, and the efficiency of the two translation models on the data set was compared to ensure that the parameters during the training process remained unchanged. The comparison experiment of the two models is shown in Figure 8.

As illustrated in Figure 8, MT technology based on the reply NN performs well in short sentence translation. However, the translation quality results in extended sentences that are unsatisfactory. MT technology based on revertive NNs has a severe difficulty with long sentence translation capacity. It has a significant impact on MT quality and is an area that has to be improved.

According to the findings of the experiments, the CNN described in this study can partially solve the problem that NMT is sensitive to sentence length. Simultaneously, the multisequence coding method incorporates lexical and syntactic information into the NN to further guide the generation of translations by using the CNN to encode related sequences other than the source language sentences in parallel. Thus, the translation performance is improved to a certain extent.

In the experiment, the revertive NN and the CNN-based translation model are compared, and the experimental data are used to identify the model's relevant benefits and characteristics. The following is shown in Table 1.

Table 1 shows the following: the calculation of the output at the present moment is dependent on the output at the previous moment due to the revertive NN's timing mechanism. Context information stacking in long-term sequences is prone to information loss and information disturbance, resulting in low quality translation results. The statements are dynamically segmented according to the specific conditions of each source statement. Compared with the source sentence, the sentence block successfully removes redundant information interference, and the sentence length is

TABLE 1: Comparison of accuracy rates between revertive NNs and CNN-based translation models.

Sentence length	Revertive NN (%)	CNN (%)	Error (%)
5	76.8	94.7	17.9
10	64.9	90.3	25.4
15	58.0	89.9	31.9
20	47.9	89.7	41.8

shortened. The network can better understand the meaning that the source sentence wants to express and improve the effect of MT.

The identical English texts are utilized as the test set when the experimental model is created. Corresponding model tests were carried out using different experimental models. It analyzes the quality and performance of the model by comparing the different translations generated by the model. In this paper, the quality of sentences of different length levels is analyzed, and the semantic extraction and semantic expression as well as the fluency of sentences are compared, as shown in Figure 9.

As shown in Figure 9, from the translation effect of the two translation models, the English translation model based on CNN shows better translation quality than other models in terms of semantic extraction and semantic expression of sentences. Moreover, the contextual relationship of related words in the sentence and the choice of translation also show certain advantages. It greatly guarantees the fluency of sentences while delivering the semantics correctly, illustrating the ability of LSTM for deep semantic encoding and long-term memory.

3.2. Comparison Experiment of MT Technology. This study reveals the translation results of MT models using BLEU scores.
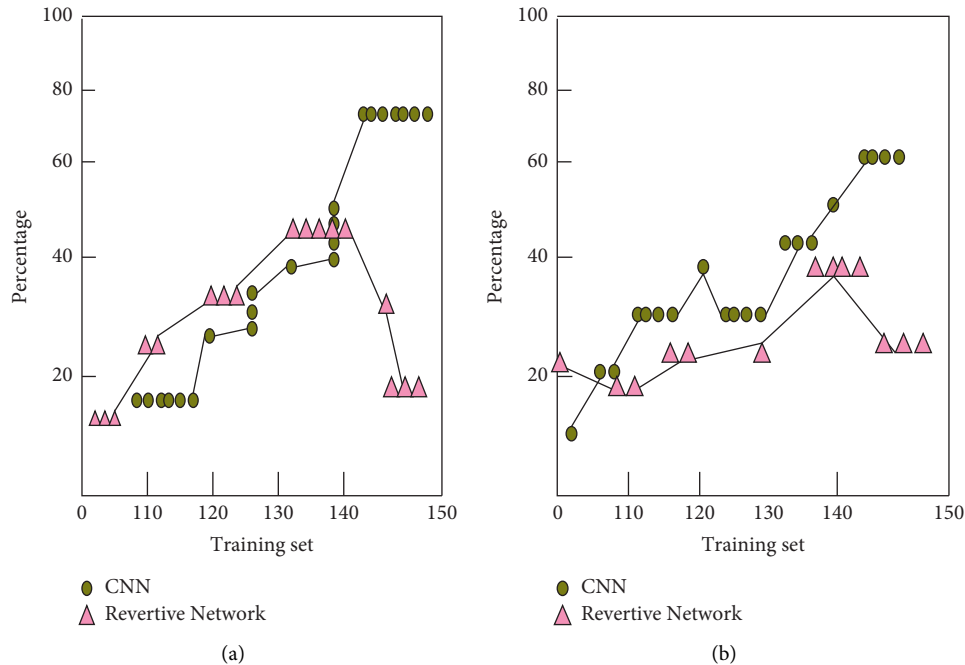
FIGURE 9: Comparison of sentence translation quality at different length levels. (a) Comparison of semantic extraction and semantic expression capabilities of the two models (b) Comparison of translation fluency between the two models.

Overfitting can be avoided by using dropout (neuron dropout probability). It is the most appropriate dropout value based on the existing data and application context. The model cross-entropy loss and the BLEU score of the training dataset in the first 5 iterations were lower than when the dropout was set to 0.2, 0.5, and 0.7, as shown in Tables 2 and 3.

Dropout can help avoid overfitting to a certain extent, as shown in Tables 2 and 3. Since the number of hidden layers is set to two, the number of model layers is reduced. At the same time, it sets dropout to 0.2 based on comparing experimental results, model training time, and overfitting risk. It limits overfitting to a certain level.

Table 4 shows the experimental findings of the MT model comparison in this paper:

Table 4 shows the cross-entropy loss of each translation model on the training dataset, although the fit is poor. The BLEU scores on the training dataset were similar when each model stopped training. The MT model based on the CNN provides a better translation effect than the regression NN model. The experimental findings objectively indicate the efficiency of the MT model based on CNNs.

3.3. Effect of CNN in College English Translation Teaching Management.

This paper conducts experiments on two English major classes *A* and *B*. There are 40 students in each class, the time period is half a semester, and the experimental comparison results are midterm exam results. Class *A* is taught by traditional manual translation, and class *B* is taught by MT teaching management mode based on CNN.

This paper investigates the students' preference for the teaching management of MT and traditional manual

TABLE 2: Comparison of cross-entropy loss under different dropouts (Chinese-English).

Learning rate	0.2	0.5	0.7
1	5.152	4.719	4.546
2	4.624	4.176	4.052
3	4.406	3.930	3.885
4	4.251	3.765	3.800
5	4.156	3.649	3.742

TABLE 3: Comparison of BLEU scores under different dropouts.

Learning rate	0.2	0.5	0.7
1	0.211	0.314	0.362
2	0.344	0.424	0.450
3	0.398	0.458	0.470
4	0.418	0.475	0.479
5	0.427	0.487	0.482

TABLE 4: Comparing the experimental results of the two models.

Contrast	Reply	Convolution
Training set LOSS	3.175	3.234
Training set BLEU	0.567	0.586
Test set BLEU	0.350	0.347

translation teaching management by two classes, as shown in Tables 5 and 6.

As shown in Tables 5 and 6, for the teaching management mode of MT, there are 35 students who like it very much, accounting for 43.75%. Only 3 students expressed

TABLE 5: Teaching management of MT.

Degree of liking	Number of people	Percentage	Effective percentage
Like very much	35	43.75	43.75
Like	23	28.75	28.75
Neither like nor hate	10	12.50	12.50
Hate	9	18.0	18.0
Very annoying	3	6.0	6.0

TABLE 6: Teaching management of traditional teacher translation.

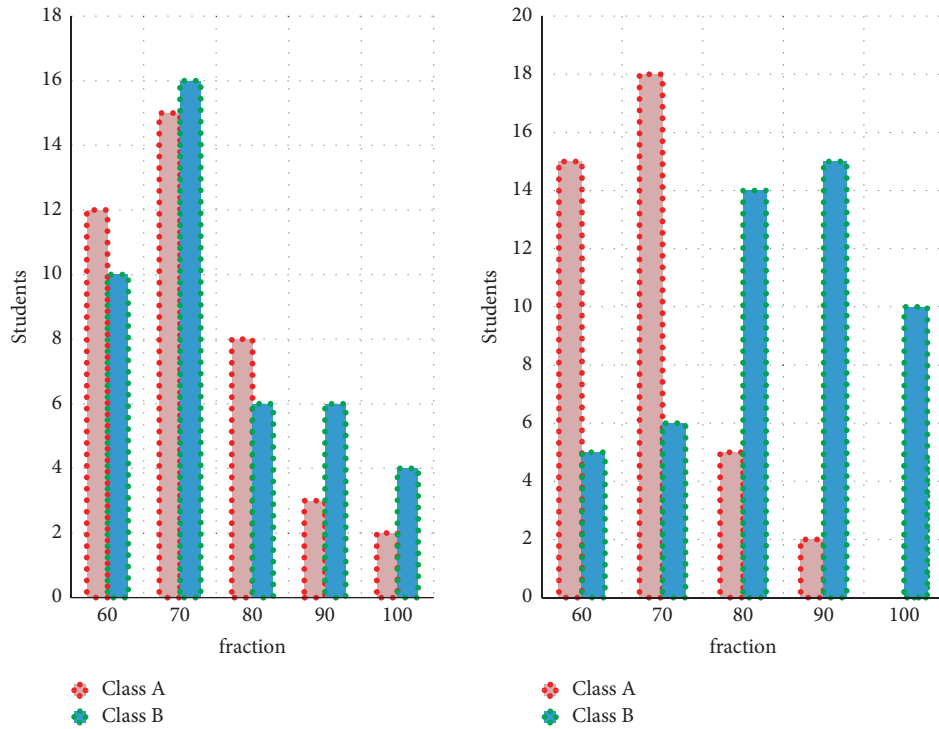| degree of liking	Number of people	Percentage	Effective percentage
Like very much	10	12.50	12.50
Like	9	11.25	11.25
Neither like nor hate	21	26.25	26.25
Hate	26	32.50	32.50
Very annoying	14	17.50	17.50

FIGURE 10: Comparison of the scores of the two classes before and after the test.

dislike, accounting for 6.0%. But for the teaching management mode of traditional teacher translation, only 10 students expressed that they like it very much, accounting for 12.50%. 14 students expressed that they disliked it very much, accounting for 17.50%. It can be seen that the teaching management of traditional teacher translation is not popular.

This article compares the scores of the two classes after one semester, as shown in Figure 10.

As shown in Figure 10, at the beginning of the semester, the average grades of Class A and Class B were compared, and it was found that the students in both classes were at the same level of translation. However, after a semester of study, their translation skills have improved a lot, especially in English. The application of different teaching methods is the reason for this difference. Therefore, it can be concluded that through the application of this method in teaching, students' translation level has been fully improved.

4. Conclusions

With the increasing demand for professional knowledge of English translation in society, college English translation education is becoming more important at major schools and universities. The usage of neural networks in English translation systems has been discovered to improve the quality and accuracy of the English translation. Traditional statistical machine translation, on the other hand, can no

longer match the needs of modern civilization. Through these analyses, CNNs have been found to help improve the accuracy of MT. The experimental comparison between CNN and the revertive MT model is carried out in the experimental section. The translation ability of CNN is found to be stronger when the sentence lengths are different. Finally, the translation model based on CNNs is carried out for the English translation class, and the teaching comparison is carried out. Finally, it is found that the translation model based on CNNs has good teaching quality. Due to the author's lack of ability, there are still some flaws in many aspects, and the author strives to do better in the next work.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

References

[1] S. Yao, "Application of computer-aided translation in English teaching," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 12, no. 08, pp. 105–117, 2017.

[2] X. Wang, "Building a parallel corpus for English translation teaching based on computer-aided translation software," *Computer-Aided Design and Applications*, vol. 18, no. S4, pp. 175–185, 2021.

[3] I. Najjar, K. A. Kadhim, and S. Al-Heeh, "Morphological shift of hyperbolic patterns in the Quran, with reference to English translation," *FORUM/Revue internationale d'interprétation et de traduction/International Journal of Interpretation and Translation*, vol. 19, no. 2, pp. 148–163, 2021.

[4] X. . Zhang, "A study of cultural context in Chinese-English translation," *Region - Educational Research and Reviews*, vol. 3, no. 2, pp. 11–14, 2021.

[5] E. Fitriani, "Errors found on the English translation sentences of the third-year students of English at IKIP budi utomo," *Journey (Journal of English Language and Pedagogy)*, vol. 3, no. 1, pp. 36–45, 2020.

[6] M. Susini, "Overcoming implicit meaning in Indonesian-English translation," *RETORIKA: Jurnal Ilmu Bahasa*, vol. 6, no. 1, pp. 63–68, 2020.

[7] H. Chen, Y. Zhang, M. K. Kalra et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[8] S. S. S. Kruthiventi, K. Ayush, and R. V. Babu, "DeepFix: a fully convolutional neural network for predicting human eye fixations," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4446–4456, 2017.

[9] F. Lu, F. Wu, P. Hu, Z. Peng, and D. Kong, "Automatic 3D liver location and segmentation via convolutional neural network and graph cut," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 171–182, 2017.

[10] Y. Fang, Z. Chi, W. Yang, J. Liu, and Z. Guo, "Blind visual quality assessment for image super-resolution by CNN," *Multimedia Tools and Applications*, vol. 77, no. 10, pp. 1–18, 2018.

[11] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, and D. Kong, "Automatic abdominal multi-organ segmentation using deep CNN and time-implicit level sets," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 3, pp. 1–13, 2017.

[12] C. Yuan, X. Li, Q. Wu, J. Li, and X. Sun, "Fingerprint liveness detection from different fingerprint materials using CNN and principal component analysis," *Computers, Materials & Continua*, vol. 53, no. 4, pp. 357–371, 2017.

[13] H. A. Haenssle, C. Fink, R. Schneiderbauer et al., "Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, 2018.

[14] I. Evo and A. Avramovi, "CNN based automatic object detection on aerial images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 740–744, 2017.

[15] O. Köksal, "Investigation of cultural competence development processes of English translation and interpretation department students," *Dil ve Dilbilimi Çalışmaları Dergisi*, vol. 16, no. 2, pp. 1070–1083, 2020.

[16] H. Wen, "Intelligent English translation mobile platform and recognition system based on support vector machine," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 7095–7106, 2020.

[17] J. Zhihao, "Simulation of ocean surface temperature based on audio signal collection and accuracy of trade English translation," *Arabian Journal of Geosciences*, vol. 14, no. 16, pp. 1614-1615, 2021.

[18] P. Huang, "On chinglish in Chinese-English translation and its countermeasures: taking translation of modern Chinese prose as an example," *OALib*, vol. 08, no. 06, pp. 1–8, 2021.

[19] Y. Tan, X. Qiao, and Y. Ma, "A study on Chinese-English translation of international publicity of the northern shaanxi tourism in China from the perspective of reception aesthetics," *Open Journal of Modern Linguistics*, vol. 11, no. 03, pp. 415–423, 2021.

[20] D. D. Sokoloff, M. S. Nuraliev, A. A. Oskolski, and M. V. Remizowa, "Gynoecium evolution in angiosperms: monomery, pseudomonomery, and mixomery," *Moscow University Biological Sciences Bulletin*, vol. 72, no. 3, pp. 97–108, 2017.