ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Research on the Module Design of AI+Project-based French Translation Course

Cen Peng[1], *Ying Chen*[1,*] *and Dinglin Cui*[2]

[1] School of Foreign Languages, ZhiXing College of Hubei University, Wuhan, Hubei, 430011, China
[2] Graduation Guidance, Shanghai, 201306, China

SUMMARY: *The current study initially presents a teaching methodology incorporating artificial intelligence in French translation education. Then, we build an unsupervised NMT framework based on pre-training by adopting the Transformer network along with natural language processing tools (particularly BERT). Furthermore, BERT and word embeddings are integrated to generate a two-level representation for the Chinese-to-French NMT framework, thereby facilitating smart French translation. Based on the results, the proposed model successfully utilizes the knowledge learned from pre-trained language models to extract domain-specific characteristics and improve translation performance in various domains, while its utilization proves effective in handling multi-domain translation issues. Besides, after applying the pedagogic model and the proposed method, there is substantial enhancement in the participants' translation skills, as the mean difference in translation scores between the experimental group and the control group in the pre- and post-tests is significantly larger in the former ($P = 0.000$, < 0.05).*

KEYWORDS: *Neural Machine Translation Model; BERT; Teaching Model; Transformer Architecture; French Translation*

1 Introduction

With the continuous development of economic globalization, in addition to learning the mother tongue well, it is especially important to master a foreign language with a wide range of applications. In the current social context, China and France are increasingly frequent exchanges and cooperation in the fields of economy, culture and education, and it is crucial to do a good job in teaching applied French translation to eliminate language barriers and realize cross-cultural exchanges as the basis of business cooperation and educational cooperation [1-4]. At present, China is actively promoting the teaching of French translation, and compared with the teaching in the past has made a lot of gratifying achievements. However, from the current situation, China's French translation teaching work still exists more problems, such as the teaching objectives are not positioned, the teaching content lacks practicality, the teaching method is old-fashioned, the practical teaching link needs to be improved, and the teaching evaluation mechanism is single [5-7].

Artificial Intelligence (AI), as one of the hotspots in today's technology field, is merging with the education industry to produce a brand new technology field. Through the addition of AI, the automation, intelligence, and immediacy of language translation and teaching

*CHENYing6457@163.com
https://doi.org/10.65102/is2026178 Published: 30 April, 2026

evaluation are realized, effectively assisting the teaching of translation. Application of AI in French Teaching. Literature [8] developed a Python-based AI motion-sensing teaching system integrating gesture detection and French language learning, which instantly evaluates students' movements and gives them guidance, especially beneficial for kinesthetic learners. Students using the system consolidated their memory through body movements and showed better learning retention in a delayed test after 14 days. Literature [9] This constructs a Chinese-French translation automatic assessment system based on a large language model, over fine-tuning, chain of thought and other methods, combined with two-way converter model quantitative assessment and Qwen2.5 to generate scoring rationale, to achieve accurate and interpretable automated judgments of translation quality. Literature [10] designed an intelligent French translation training and evaluation system based on deep learning, which can provide personalized translation exercises according to students' abilities, and effectively evaluate students' translation level, improve their translation ability to give real-time feedback, but also significantly enhance the learning interest and enthusiasm, injecting a new impetus for French teaching. Literature [11] discusses the application of ChatGPT-4 in the English-French translation classroom, and students generally regard it as a useful lexical and grammatical reference tool, which promotes students' critical thinking and editing ability, but also points out that its output often needs manual proofreading to correct the stiffness or inaccuracy. Teaching experiments in the literature [12] have shown that university students using generative AI have significant gains in French language proficiency and motivation, but also face challenges such as lack of cultural depth and content accuracy, which require optimization of the design and enhancement of teacher guidance and data protection.

Based on most existing studies, it can be observed that project-based teaching provides clear benefits for language education. Literature [13] comparing the impact of this teaching approach on the performance of learners in French courses has found out that those taught using this teaching approach scored much better than those in regular French classes with an interaction effect between their level of motivation and type of learning. This learning technique was developed for the purpose of medical education in the 1950s but has later been applied to other subjects; yet, it has been rarely used in language education, especially in French [14, 15]. This approach is appealing to many teachers because of its focus on effective transferable learning and autonomous development of students. It has proven itself to be useful for the development of translation competencies, too. Literature [16] has incorporated this teaching approach into a Chinese to English translation course at university where most students agreed that this approach effectively improved their metacognitive awareness and translation competency skills. Literature [17] has adopted this teaching approach in teaching business translation in two Yemeni universities and reported that translation competency of the experiment group has improved significantly compared to the control group. Literature [18] has adopted this teaching approach in teaching business translation in two Yemeni universities and reported that translation competency of the experiment group has improved significantly compared to the control group. Literature [19], conducted by means of classroom action research, has found that implementation of the proposed teaching approach has significantly improved translation competency among the students. Moreover, it can be used to boost learners' participation in class, integrate theories and practices, and help develop students' professional skills [20].

The article first designed a teaching model for French translation courses and proposed a hybrid five-step teaching method on-line and off-line. After that, the Transformer self-attention mechanism is utilized to achieve end-to-end neural machine translation effect; then the BERT training method based on natural language processing and unsupervised pre is introduced to break the records of many natural language processing (NLP) tasks through simple fine-tuning.

The current research seeks to solve the problem of Chinese－French neural machine translation due to the lack of parallel sentences between two languages and insufficient comprehension of the source language within the encoder. This paper presents a technique for Chinese-French neural machine translation that uses dual representations generated through the employment of BERT and word embedding. Specifically, the attention mechanism is applied to accomplish adaptive dynamic fusion of dual representations and improve source language representation learning.

2 AI+Project-based French Translation Course Module Design

2.1 Teaching Model Setting for French Translation Courses

(1) Training Objectives

This course is a professional elective course for French majors. It starts in the 6th semester and lasts for 16 weeks, with a total of 32 hours. It aims to cultivate French-speaking talents with knowledge and ability of computer-aided translation in line with the needs of the industry.

(2) Teachers and Teaching Materials

There are four teachers in this course, all of whom are majored in French, have rich practical experience in French translation, and continue to pay attention to, study and research on translation technology and the teaching of translation technology, so as to fully prepare for the opening of the introductory course of computer-assisted translation.

(3) Setting of Sections and Chapters

This course is divided into three major sections, namely, CAT tool foundation, information search, and translation platform, CAT tools include peripheral tools and core tools. CAT tools include peripheral tools and core tools. Tools refer to pre-translation and post-translation tools such as text recognition, text editing, scanning, printing, etc. Core tools refer to translation memory, terminology database, corpus, etc. Information search is the process of searching for information and removing the false from the true. The ability to acquire new knowledge through tools. Translation platform integrates translation core tools and is divided into local desktop translation platform and cloud platform.

(4) Course Features

According to the chapter setting of the self-compiled lecture notes, three main features are highlighted. First, real project-driven. Driven by the school's projects such as “News from the Academy”, translation by famous translators and German certification, students can make progress in the real world. Secondly, students participate actively. In the implementation of translation projects, different students act as project managers for each project; for large-scale projects, a department manager leads several project managers to start the translation work.

(5) Teaching Mode

The online-offline hybrid five-step teaching method is shown in Figure 1. It is the third prize winning project of the School Teaching Achievement Award, and its strong vitality and operability have been proved in previous course practices. This course follows this teaching method. “Dual Platform of Resources”: online learning through the platform course construction, recording theory and practical videos; offline learning is based on project-based practical exercises, with teachers answering questions and assisting students in computer-aided translation practice. The “Five Steps of Teaching” includes independent study before class, classroom quiz for students, classroom presentation for students, classroom face-to-face teaching for teachers, and extended evaluation after class. Taking the three teaching dimensions of “mixed teaching resources, mixed teaching activities and mixed teaching evaluation” as the entry point, this teaching mode covers the whole process of introductory computer-assisted translation lectures.

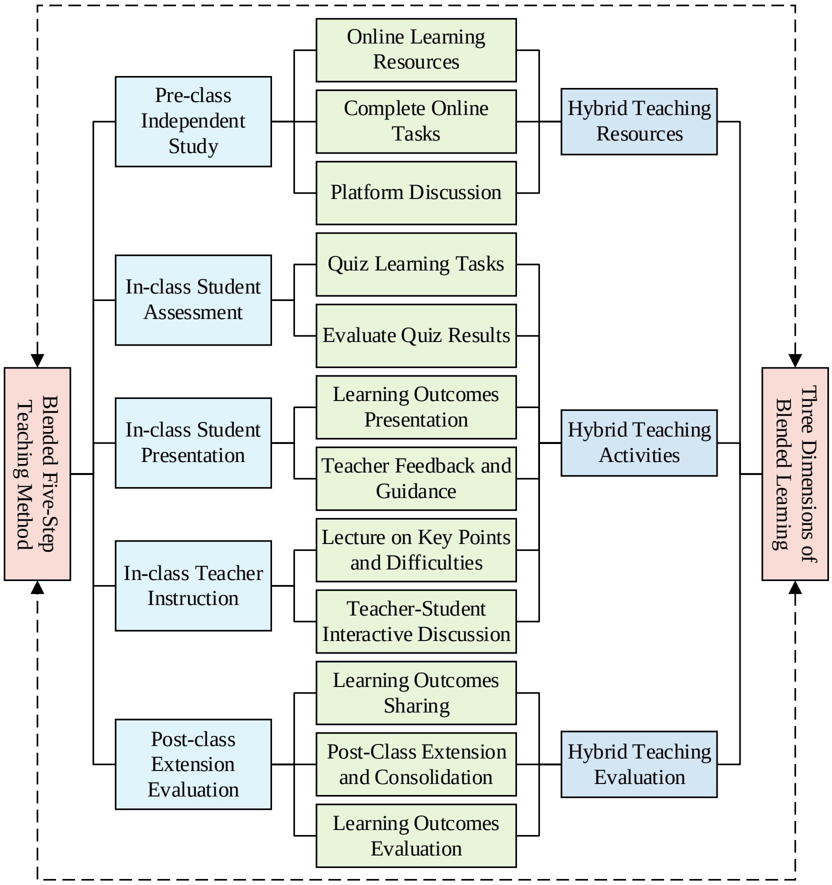

Figure 1: Online and offline hybrid five-step teaching method

(6) Course Evaluation and Assessment

The evaluation and assessment of this course is based on diagnostic and formative evaluation, supplemented by summative evaluation. Specifically, the total grade = paper grade × 60% + usual grade × 40%. The final grade, i.e., the paper grade, examines students' knowledge of the CAT course and is based on subjective questions; the usual grade includes pre-course online learning, in-class attendance, classroom practice, weekly assignments, group projects, which focuses on detecting whether students have completed their studies, whether they are proficient in the use of CAT tools, whether the quality of their translation works meets the requirements, whether they can accumulate their own corpus, terminology library, memory library and carry out reasonable maintenance, and whether they can carry out different translation tools. Whether they can maintain their own corpus, terminology database and memory and carry out reasonable maintenance, and whether they can transform between different translation tools and translation platforms.

2.2 Neural Machine Translation Model and Architecture

2.2.1 Transformer Architecture

Figure 2 illustrates the model architecture and initialization process of the parameters for XLM (Cross-Language Model), which has proven to be highly effective in pre-training and is broadly used in RN or CNN areas. Self-Attention is applied in order to develop an end-to-end NMT (Neural Machine Translation). By calculating attention scores between each token and all other tokens in a sentence, intra-sentential dependencies are learned, which means that internal structure of the input will be captured. The Transformer Encoder-Decoder is made up of many stacked layers. In particular, the Transformer encoder is composed of N identical blocks each of which comprises two components: first, Self-Attention; second, Position-Wise Feed-Forward Networks. The Transformer decoder is also made up of N identical blocks, each of which consists of three subcomponents: first, Masked Multi-Head Self-Attention; second, cross-attention which links back to the previous output of the decoder; third, Position-Wise Feed-Forward Networks. Residual connections are applied to the Transformer encoder and decoder blocks. Residual connection is represented by the following formula:

$$h^{l} = h^{l-1} + f_{sl}\left(h^{l-1}\right) \tag{1}$$

where h^{l} denotes the output of the l th sublayer and $f_{sl}(\cdot)$ denotes the function function of the layer.

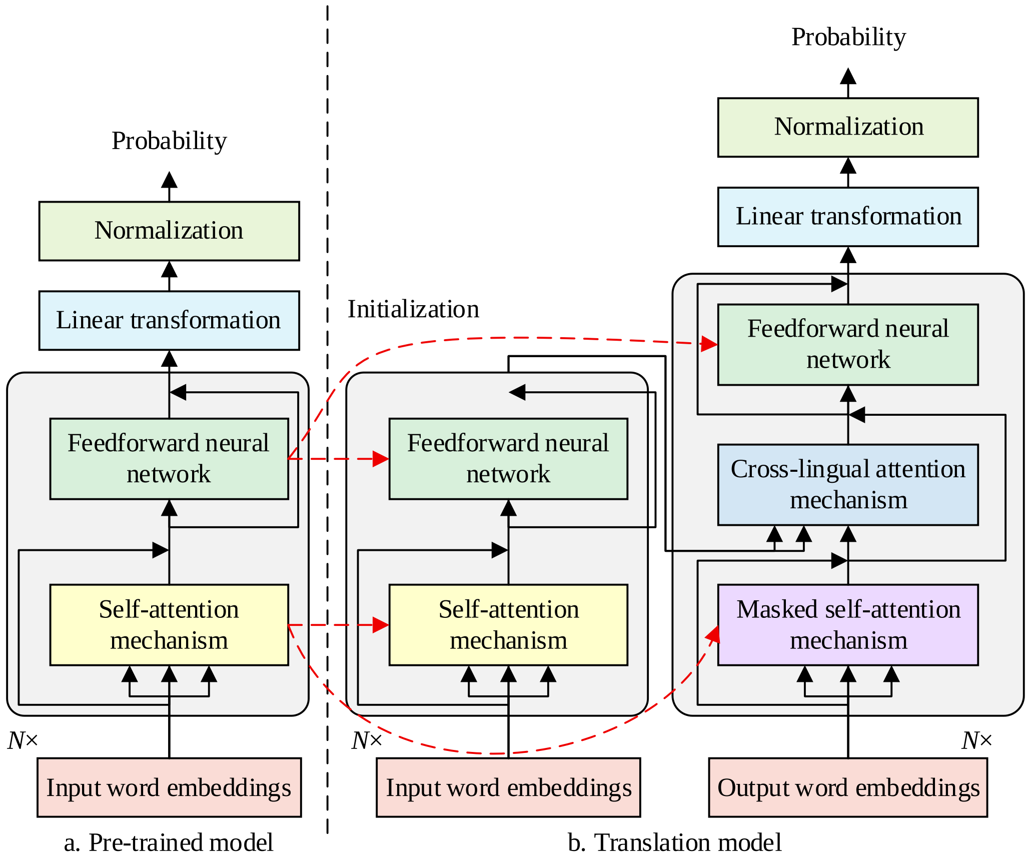

Figure 2: Transformer model structure and the initialization process of XLM

2.2.2 BERT and XLM

BERT is a powerful natural language processing, unsupervised pre-training method that has broken numerous records for natural language processing NLP tasks with simple fine-tuning. The core idea of BERT is to pre-train a deep bi-directional Transformer encoder, which means that it processes a word taking into account the information of the words that precede it and those that follow it in order to obtain the contextual semantics. The pre-training of BERT has 2 training tasks:

(1) Masked Language Model MLM training, which randomly masks out a certain percentage of input sentence tokens and predicts these masked out words based on the context.

(2) Prediction of the following sentences, a task that predicts the content of the following sentences based on the existing sentences, e.g., the Q&A QA task in NLP and the natural language inference NLI task. Cross-language modeling (XLM) is trained based on 2 main points:

One is to use byte pairs from 2 languages to encode BPE shared vocabularies, and the same method is used to obtain shared vocabularies, which achieves very good experimental results. The difference is that in the cross-language pre-training experiments, the sampling probabilities $\{q_i\}_{i=1,\cdots,M}$ of the sentences in different languages obey a polynomial distribution before BPE processing. The purpose of sampling in this way is to smooth the large and small language corpora and prevent the words of the small language corpora from being segmented from the character level during the BPE processing, in order to balance the number of words of different languages in the shared word list:

$$q_i = \frac{p_i^\alpha}{\sum_{j=1}^{M} p_j^\alpha},\ p_i = \frac{n_i}{\sum_{k=1}^{M} n_k} \tag{2}$$

where α is taken as 0.5, p_i denotes the ratio of the number of sentences in the i th language to the number of all sentences in the corpus, M denotes the number of language types, and n_i denotes the number of sentences in the i th language.

Secondly, the Translation Language Model (TLM) is proposed, which in essence extends the MLM task to bilingualism by linking parallel sentence pairs of two languages together to form a continuous text stream, which is used as the input of MLM. This allows the model to utilize the cross-linguistic information provided by the parallel utterances to predict masked words, while encouraging the model to learn alignment information from the bilingual context as a way of pre-training the word vectors and parameters of the language model.

2.3 A French Neural Machine Translation Approach Incorporating BERT and Word Embedding

A neural machine translation method for French language fusing BERT and word embedding dual representations, based on Transformer, is an encoder-decoder architecture. The model performs BERT pre-training language model and word embedding dual representation on the source language sequences respectively, and splices the two representations and then performs linear transformation to establish a simple dynamic fusion between the two representations, and obtains a new representation vector containing the two parts of the information as the input to the encoder, and then, after the module of the self-attention mechanism within the encoder, the two sources of information are fused in a deep dynamic fusion, and finally, using the Finally, the neural machine translation model is trained using the dual representation of BERT and word embedding. The French neural machine translation model fusing BERT and word embedding is shown in Fig. 3.

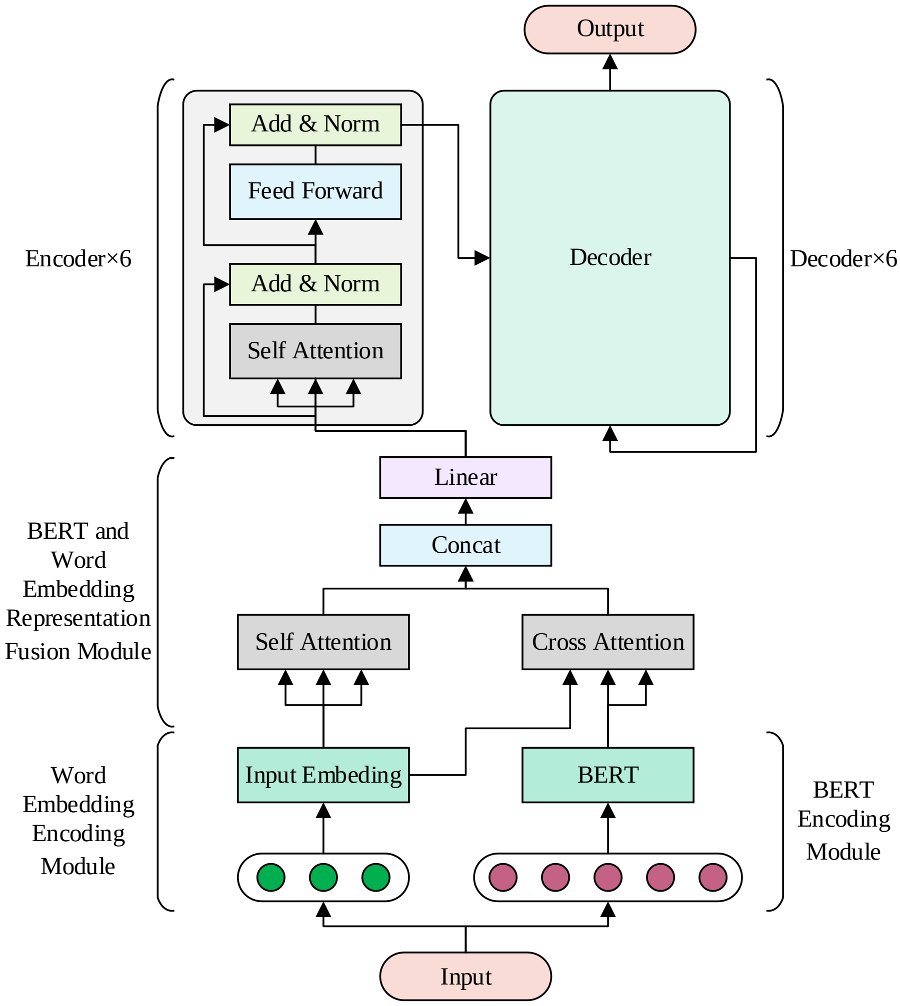

Figure 3: A French Translation Model Based on BERT and Word Embeddings

(1) Word Embedding Module

Word embedding part relies on the word-embedding method based on transformer without any additional modification, where the input text will be divided into words using the word embedding dictionary and then feed into the word-embedding unit to generate the word embedding vector $E_{embedding}$ of the input text.

(2) BERT Encoding Module

BERT is a language model trained by self-supervised learning method through a large amount of corpus, through a large amount of corpus learning can give a better feature representation of the word, the use of the feature representation vector to participate in the training can be realized in a large number of corpus to learn the linguistic information to migrate to the specified task. The network architecture used in this language model is a multi-layer transformer structure, which can represent all the contextual information based on the left and right sides during the encoding process compared to RNN and CNN networks.

Since the training of BERT pre-trained language model requires a large amount of monolingual corpus as well as a large amount of computational resources, the method in this chapter uses Google's publicly available Chinese BERT pre-trained language model, which has achieved excellent performance in a number of Chinese NLP tasks, proving that the model has a strong encoding ability for Chinese sequences. The input sequence is obtained by splitting the input text according to the BERT lexicon $x=\left(x_1,\cdots,x_n\right)$, and after inputting the input sequence to the BERT pre-training model, an implicit state vector is outputted at each layer of the model, and the method of this chapter uses the implicit state vector h_6 outputted at the last layer to be the output of the part of the part. As the output $E_{bert-out}$ of that part.

(3) BERT and Word Embedding Representation Fusion Module

Using $E_{bert-out}$ and word embedding representations $E_{embedding}$ for the calculation of cross-attention mechanism, the word embedding part of the output $E_{embedding}$ as $Query$, $E_{bert-out}$ as Key to compute the attention weight, and $E_{bert-out}$ as $Value$ and attention weight, so that the BERT pre-trained model representations are linked by the word embedding representations, and after applying the cross-attention mechanism, the BERT pre-trained model representations are linked by the word embedding representations to make the BERT pre-trained model representations. Value and the attention weights are multiplied so that the BERT pre-trained model representations are linked by the word embedding representations, and after applying the cross-attention mechanism so that $E_{bert-out}$ is constrained by $E_{embedding}$, a new representation of $E'_{bert-out}$ is obtained. There are:

$$Query = E_{embedding} \tag{3}$$

$$Value = Key = E_{bert-out} \tag{4}$$

$$Attention\left(Query, Key, Value\right) = soft\max\left(\frac{QueryKey^T}{\sqrt{d_k}}\right)V \tag{5}$$

$$E'_{bert-out} = Attention\left(Query, Key, Value\right) \tag{6}$$

A self-attention mechanism calculation is performed for characterization enhancement, which is calculated as follows:

$$Query = Value = Key = E_{embedding} \tag{7}$$

$$E'_{embedding} = Attention\left(Query, Key, Value\right) \tag{8}$$

The new text sequence hidden state vector $E_{bert-embedding}$ is obtained by splicing $E'_{bert-out}$ and $E'_{embedding}$ and then linearly transforming the dimensions:

$$E_{contact} = contact\left(E'_{bert-out}, E'_{embedding}\right) \tag{9}$$

$$E_{bert-embedding} = Linear\left(E_{contact}\right) \tag{10}$$

(4) Encoder module

The BERT and word embedding representation fusion module obtains a representation vector $E_{bert-embedding}$ containing the information of $E'_{bert-out}$ and $E'_{embedding}$, the two parts of which are not connected. When $E_{bert-embedding}$ enters the first layer of the encoder a single self-attention mechanism computation is performed, causing the two otherwise independent parts to establish a connection, yielding $E'_{bert-embedding}$. i.e:

$$Query = Value = Key = E_{bert-embedding} \tag{11}$$

$$E'_{bert-embedding} = Attention\left(Query, Key, Value\right) \tag{12}$$

The $E'_{bert-embedding}$ obtained after the computation of self-attention mechanism realizes the dynamic fusion of $E_{bert-out}$ and $E_{embedding}$. The $E'_{bert-embedding}$ goes through a feed-forward neural network to get the output H_1 of the first layer of the encoder, and then goes through multiple coding layers to finally get the final output of the encoder, which is computed as:

$$H_1 = FNN\left(E'_{bert-embedding}\right) \tag{13}$$

$$h_t = Attention\left(H_{t-1}, H_{t-1}, H_{t-1}\right), t > 1 \tag{14}$$

$$H_t = FNN\left(h_t\right), t > 1 \tag{15}$$

(5) Decoder Module

The decoder accepts the hidden state vector H output by the encoder as input. In this paper, the structural design of dynamic fusion of $E_{bert-out}$ and H at the decoder side using the representation fusion method proposed in this chapter ended up with degradation in translation performance, so the method in this chapter does not do additional design at the model decoder.

3 Performance Analysis of AI+Project-based French Translation Course

3.1 Experimental setup

(1) Experimental Data

This paper validates the effectiveness of this paper's method on a French-Chinese (ELRA) multi-domain translation dataset, from which four specialized domains, IT (Information Technology), Koran (Qur'an), Law (Legal) and Medical (Medicine) are selected for the experiments, with the training set: test set: validation set = 7:2:1. In this paper's experiments, French and Francophone share a single lexicon with the size of the lexicon is 10 million bilingual segments covering close to 400 language pairs. In the experiments covered in this paper, the source language is French, and the pre-trained model using French data in this paper is a French neural machine translation model integrating BERT and word embedding.

(2) Comparison Model

In this paper, a total of eight comparative models, namely, TMixed, LMixed, WMT19, WMT19-FT, Dual-Encoder-in, Dual-Encoder-out, WDFS, and BERT-NMT, are selected for comparison and analysis with the model in this paper.

3.2 Analysis of the training effect of the model

The learning rate is the tuning parameter in machine learning and statistical methods that controls the process of all updates taking place in one epoch to ensure that the loss function can reach its minimum point.

The learning rate formula for this experiment is:

$$r_{rate} = d_{\mathrm{mod}el}^{-0.5} \cdot \min\left(step\_num^{-0.5}, step\_num \cdot warmup\_steps^{-1.5} \right) \tag{16}$$

The described algorithm describes a learning-rate strategy that starts with linear growth during warm-up and continues with decaying in relation to the inverse square root of the iteration number. Since the initial values of parameters in the models are set randomly, a high value of the learning rate may create instability in this process. Thus, the duration of warm-up reaches 4000 iterations, and for this reason, the learning rate has a lower value at the first stages. After warm-up, a reduced constant value of the learning rate is used, which results in faster convergence. The relationship between the learning rate and iteration number is shown in Fig. 4, when the Steps value is from 0 to 1965 times, the learning rate grows the most, and the learning rate is the maximum value of 0.00066 when the Steps value is 1965. the learning rate starts to decline gradually after more than 2,000 iterations, and the learning rate has been reduced to 0.00032 when the Steps value is 10,000.

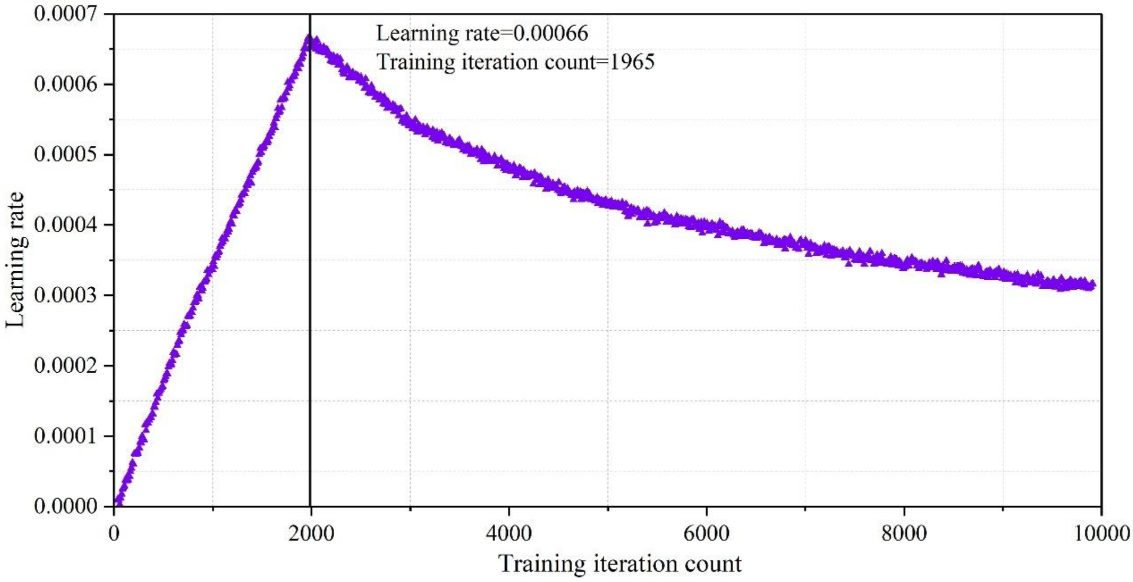

Figure 4: Learning rate as a function of training iteration count

The variation of Loss value with the number of iterations is shown in Fig. 5, where the Loss value decreases rapidly from 0 to 669, and the decline decreases after 1542 iterations, approaching 0 infinitely at 9900 iterations.

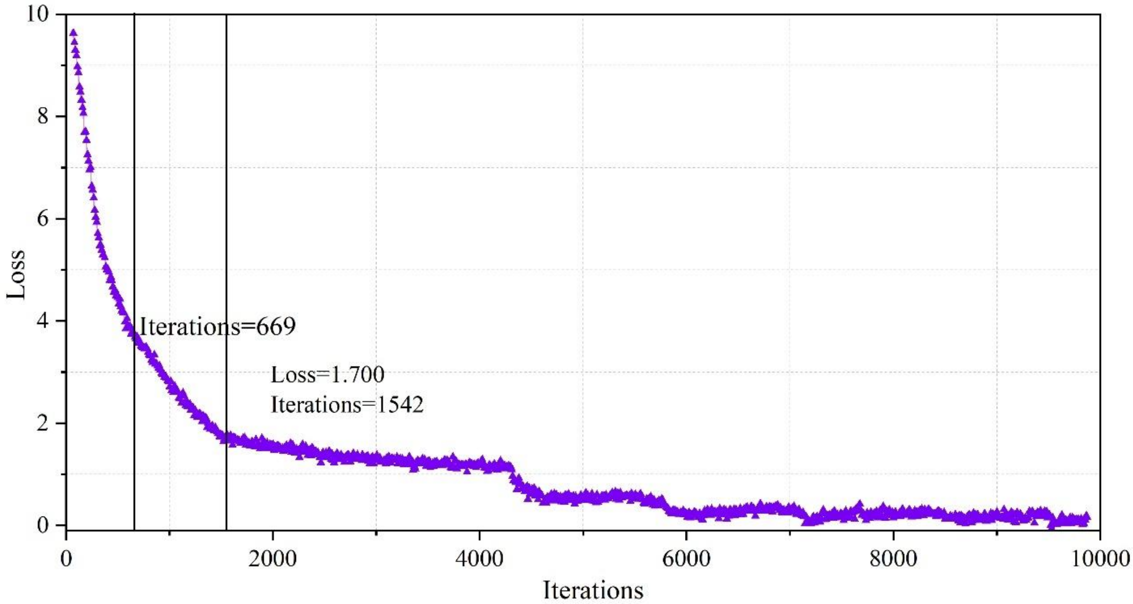

Figure 5: The variation of the Loss value with the number of iterations

When solving problems within an experiment, some indicators are usually defined to measure the effectiveness of the final model according to the actual needs, and the Acc value is one of the indicators. The change of Acc value with the number of iterations is shown in Fig. 6. The Acc value rises rapidly from 0 to 591 iterations, and the increase decreases after 1550 iterations and reaches a maximum value of 0.9827 at 9900 iterations. The accuracy rate starts to slow down after 1550 iterations, and reaches a maximum value of 0.9827 at 9800 iterations. The accuracy rate increases from 1.550 to 591 iterations. After that the rise begins to slow down and has converged to the maximum value at 9800 times.

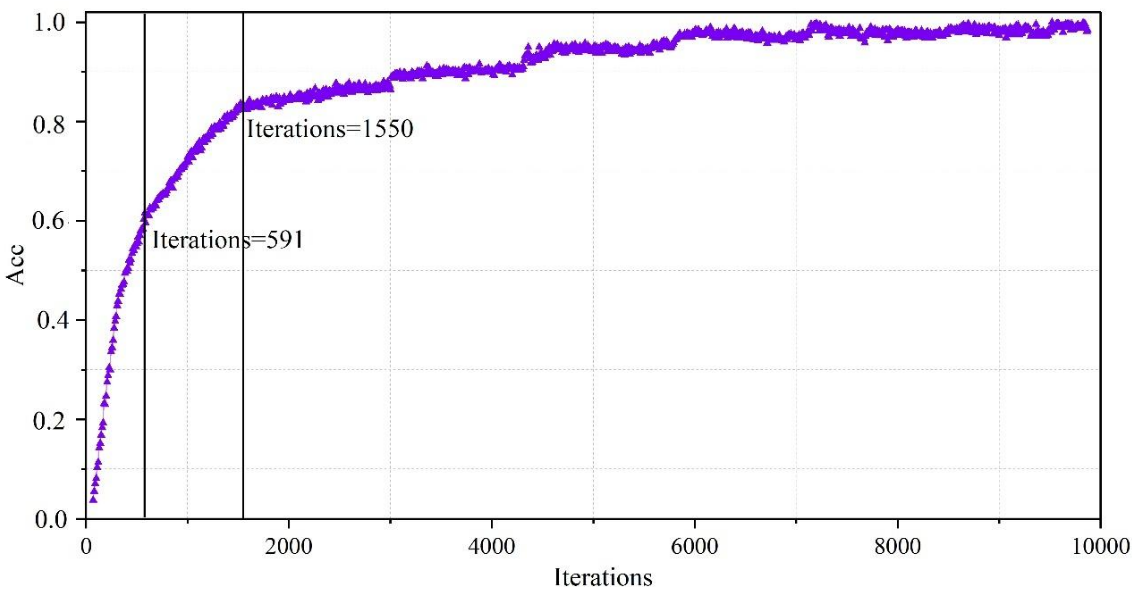

Figure 6: The change of Acc value with iteration times

The experimental results on ELRA French-Chinese translation directions are shown in Table 1. The results show that this paper's model performs well on the IT, Koran, Law and Medical domains, improving 10.52, 7.796, 12.157 and 12.884 BLEUs compared to the TMixed model, respectively; and 11.786, 9.138, 13.99 and 13.546 BLEUs compared to the LMixed model, respectively.

In terms of individual domains, the IT, Law and Medical domains have larger amount of data, so the models in this paper show good translation performance in these domains and have significant advantages over TMixed and LMixed. The Koran domain, on the other hand, has very little training data and is a really low-resource domain, which is not enough to support the training of this paper's model, so the BLEU score is lower. The LMixed and TMixed models, although they use all the training data, are “contaminated” with the Koran domain, which leads to a decrease in the translation performance. Meanwhile, TMixed has better performance compared to LMixed, which indicates that the Transformer framework has stronger translation capability compared to the LSTM encoding-decoding framework. In addition, this paper finds that the performance of the BERT-NMT model is closest to the model in this paper.

Table 1: Results of the ELRA French-Chinese Translation Experiment

Model	IT	Koran	Law	Medical	Mean
LMixed	32.587	14.370	50.022	45.005	35.496
TMixed	33.853	15.712	51.855	45.667	36.772
WMT19	37.927	16.369	45.646	40.011	34.988
WMT19-FT	41.277	19.708	56.189	49.131	41.576
Dual-Encoder-in	35.623	18.164	54.127	48.114	39.007
Dual-Encoder-out	39.101	19.743	56.413	49.649	41.227
WDFS	41.364	18.757	59.506	54.224	43.463
BERT-NMT	41.680	21.506	59.551	54.346	44.271
Ours	44.373	23.508	64.012	58.551	47.611

In order to verify that the method proposed in this paper can improve translation performance, this paper performs a visualization analysis of BERT representations. In this paper, 200 French sentences are randomly selected from each domain in the test set, and the representations of different sentences encoded by BERT are downscaled and projected into the plane space for observation using the algorithm of this paper.

The visualization results of BERT after encoding the sentences in the BERT-NMT model are shown in Fig. 7; the visualization results of this paper's model after encoding the sentences are shown in Fig. 8. It can be found that the representation between all the sentences in the BERT-NMT model is relatively scattered, and there are some sentences overlapping between the domains. Such results indicate that for some sentences, it may be difficult to encode domain-specific features under the fine-tuning of the BERT-NMT model, suggesting that BERT provides a generalized representation, thus limiting its enhancement in multi-domain translation to some extent. In contrast, in the visualization results of the model proposed in this paper, sentence representations of different domains are effectively distinguished in the low-dimensional space, which indicates that the model in this paper is able to learn domain-specific features and recognize them.

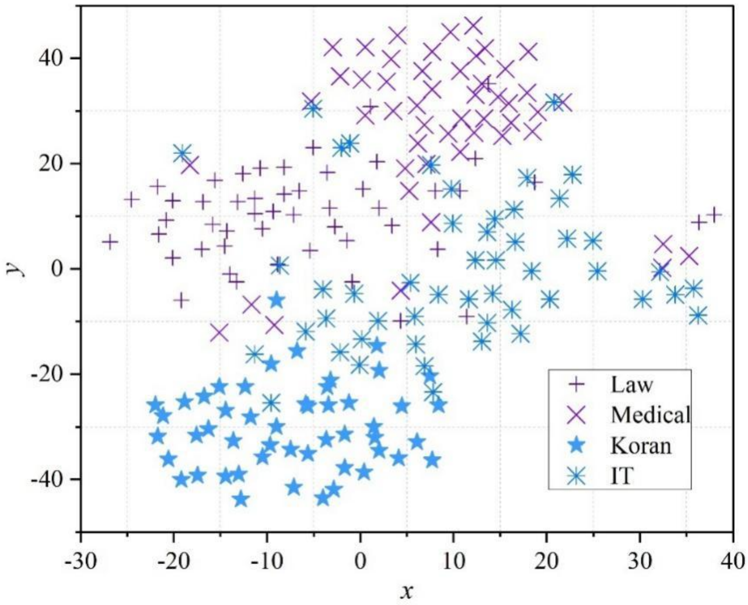

Figure 7: Visualizations of BERT's sentence encoding in BERT-NMT model

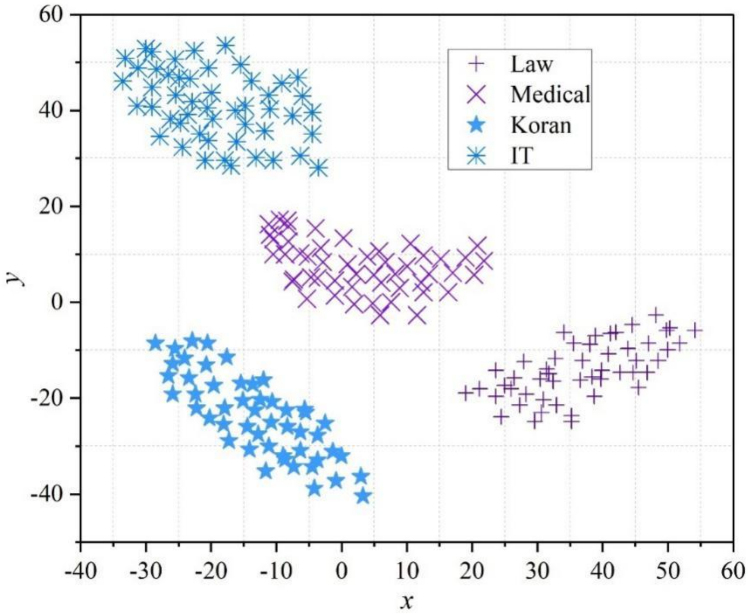

Figure 8: The visualized results of sentence encoding in this model

3.3 Analysis of factors affecting model performance

3.3.1 Impact of BERT Sentence Vector Representation on Model Performance

The input of supervised contrastive learning requires sentence vector representations of source language sentences encoded by BERT and the corresponding domain labels, and different sentence vector representations will have some effects in calculating the contrastive learning loss in backpropagation when updating the model parameters. Therefore, in this experiment, this paper aims to investigate the effect of different sentence vector representations of BERT on the model translation performance. Specifically, this paper will compare the performance of the BERT model under the following four output sentence vector patterns: FT-CLS, FT-ALL, FT-Concat and FT-ALLConcat.

The experimental results for different sentence vector representations are shown in Table 2. The Concat representation has the highest BLEU scores in IT, Koran, Law and Medical domains. This indicates that splicing the [CLS] representation of BERT with the average sentence representation (FT-Concat) when using the sentence vector representation of BERT tends to be better than other methods. In addition, this paper calculates the average BLEU scores of each model in the four domains, and it can be observed that the differences in the BLEU scores in different sentence vector representations are extremely small, which suggests that BERT's sentence vector representations have a certain degree of generality. In order to maintain the best performance of the model, FT-Concat is finally chosen as the BERT sentence vector representation of the model in this paper.

Table 2: Results of different sentence vector representation experiments

Model	IT	Koran	Law	Medical	Mean
FT-CLS	43.143	22.407	55.117	60.278	45.236
FT-ALL	42.869	22.420	55.048	60.230	45.142
FT-Concat	43.370	22.474	55.554	60.998	45.599
FT-ALLConcat	42.479	22.310	55.708	60.509	45.251

3.3.2 Contrasting the effect of learning temperature coefficients on model performance

The temperature coefficient, as an important parameter of contrast learning, determines how much the contrast learning loss focuses on difficult negative samples. Therefore, this paper needs to experimentally explore the effect of temperature coefficient of contrast learning on the model translation performance.

The effect of learning temperature coefficient on model performance is shown in Figure 9. The results show that when the models in this paper are trained together without supervised contrast learning (0°C), the average BLEU of the models is at its lowest value. When the temperature coefficient is set to 0.2, there is a large increase in the model's translation performance, and as the temperature coefficient reaches 0.3, the model's translation performance is at its highest, and a subsequent increase in the temperature coefficient results in a decline in translation performance followed by a period of recovery and then a continuous decline. This trend may reflect the sensitivity of the model to a certain temperature coefficient, as well as the fact that too high or too low a temperature coefficient may lead to different degrees of decline in the model's translation performance. Therefore, supervised contrast learning proves to be an effective mechanism to enhance the domain adaptive capability of the model in this paper's model. By studying the correlation between the temperature coefficient and the translation performance, a suitable range of values for the temperature coefficient is derived as [0.1,0.5]. With a temperature coefficient within this range, the model can better distinguish sentences in different domains, reduce intra-class differences, and achieve better translation performance. In the experiments of the model in this paper for the ELRA dataset, the translation performance of the model reaches its highest point when the temperature coefficient is 0.3.

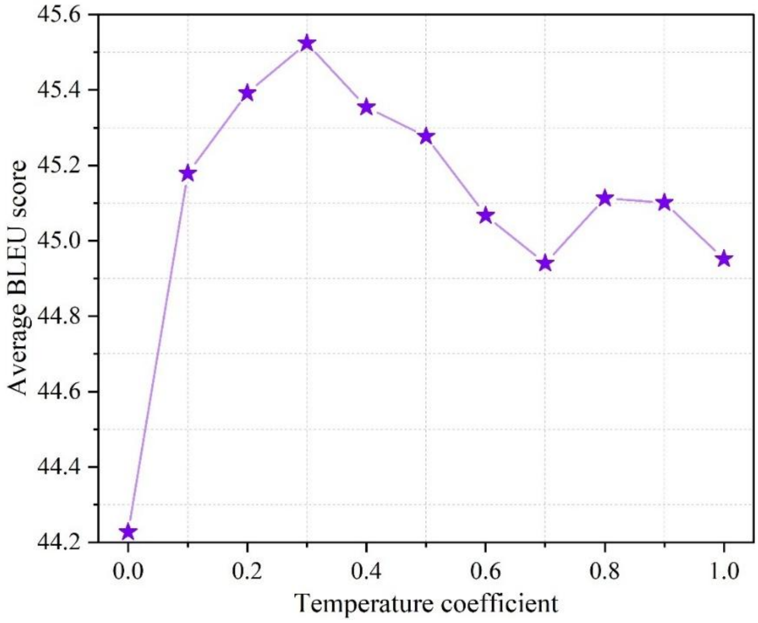

Figure 9: Study on the Effect of Temperature Coefficient on Model Performance

4 Analysis of the results of a real case test of the French translation model

(1) Research object

In the current study, two complete classes containing 72 third-year undergraduates specializing in French language were used, where both classes underwent translation courses conducted by the same instructor with two lecture hours weekly, identical course content, similar schedules for conducting lectures, and similar gender distribution along with educational background. In addition, there was no statistical significance ($p = 0.6737$) in translation proficiency between the two groups, suggesting that they had similar levels of translation proficiency before treatment. As regards the design of the study, the experimental group consisted of 35 students, whereas the control group contained 37 students, wherein the experimental group used an online/offline blended learning approach compared to the traditional approach of translation instruction among students in the control group.

(2) Research Questions

1) Does the online-offline blended instructional approach influence students' translation proficiency? If so, in what ways? If not, what are the reasons?

2) What factors explain the effect of the online-offline blended approach on students' translation proficiency?

4.1 Comparison of the pre-test for translation between the experimental and control classes

To compare the baseline translation proficiency of the two classes before the intervention, an independent-samples t-test was conducted on their pre-test scores. The corresponding results for both groups are presented in Table 3. The findings show that the mean pre-test scores of the experimental and control classes were 12.2893 and 12.1908, respectively, with a difference of only 0.0985. This suggests that the two classes had similar translation proficiency prior to the experiment.

Table 3: Pre-test results of translation for two classes

	Class	Number of cases	Average value	Standard error	Standard error of mean
Pretest	Experimental class	35	12.2893	2.068	0.3739
	Control class	37	12.1908	2.2927	0.3922

To obtain a more reliable conclusion, an independent-samples t-test was conducted. The results for the translation pre-test of the two classes are reported in Table 4. The findings indicate that the Sig. (two-tailed) value is 0.6737, which is greater than 0.05, suggesting that no statistically significant difference existed in the pre-test scores of the two classes. In other words, their translation proficiency was comparable prior to the experiment.

Table 4: Independent Sample t-test for Translation Performance

		Levene's test		T-test for equality of means						
		F	Sig	t	Freedom	Sig (two-tailed)	Mean difference	SD	95% confidence interval: Lower limit	95% confidence interval: Superior limit
Pretest	Hypothesis of equal variance	0.1834	0.6798	0.3156	72	0.6737	0.0985	0.5517	-0.9331	1.2514
	Nonassumption of equal variance							0.5536	-0.9206	1.2631

4.2 Comparison of Translation Posttests between Experimental and Control Classes

The findings about the translation performances of the two groups after the post-test are illustrated in Table 5. According to the findings, the post-test scores of the experimental group and the control group are 16.9834 and 12.1354, respectively, resulting in a gap of 4.8480. Comparing with the gap before the experiment of 0.0985, it can be seen that there is a great difference between the two groups after the experiment. To test the result, an independent-samples t-test was done again. As illustrated in Table 6, the Sig. value of the post-test in two tails is 0, lower than 0.05, implying that there is a significantly different score of translation ability between the two groups in the post-test.

Table 5: Post-test results of translation performance in two classes

	Class	Number of cases	Average value	Standard error	Standard error of mean
Posttest	Experimental class	35	16.9834	1.8265	0.3419
	Control class	37	12.1354	2.2182	0.3941

Table 6: T-test for independent samples of translation scores in two classes

		Levene's test		T-test for equality of means						
		F	Sig	t	Freedom	Sig (two-tailed)	Mean difference	SD	95% confidence interval: Lower limit	95% confidence interval: Superior limit
Posttest	Hypothesis of equal variance	0.0324	0.8809	8.9624	72	0.0000	4.8480	0.5548	-0.9313	1.2723
	Nonassumption of equal variance							0.5389	-0.9468	1.2865

4.3 Comparison of Translation Pretest and Posttest Scores

The difference between the performance of translation before and after the experimental intervention is determined by the paired sample t-test. Tables depicting the results of the pre-tests and post-tests are provided in Table 7 below. P-values for the pre-post test difference are 4.6941 and 0.0554 for the experimental and control group, respectively. According to the obtained results, it is clear that learners in the experimental group have demonstrated a remarkable improvement while translating, but learners in the control group have not significantly improved their translation skills.

Table 7: Statistical results of pre-test and post-test paired samples

		Number of cases	Average value	Standard error	Standard error of mean
Experimental class	Pre-test	35	12.2893	2.068	0.3739
	Post-test	35	16.9834	1.8265	0.3419
Control class	Pre-test	37	12.1908	2.2927	0.3922
	Post-test	37	12.1354	2.2182	0.3941

To examine changes in the control class between the translation pre-test and post-test, a paired-samples t-test was performed on the corresponding scores. The statistical results are presented in Table 8. The significance values for the experimental and control classes were 0.0000 and 0.4385, respectively. These findings indicate that the experimental class showed a statistically significant improvement after the intervention (P=0.000<0.05), whereas no significant change was found in the control class (P=0.4385>0.05). This suggests that, following one semester of instruction, students in the experimental class made marked progress in translation competence, further supporting the effectiveness of the French translation course module integrating BERT and XLM.

Table 8: Results of paired sample t-test before and after translation

Pre-test and post-test scores	Paired difference					t	Freedom	Sig (two-tailed)
	Mean	SD	Standard error of mean	95% confidence interval				
				Lower limit	Superior limit			
Experimental class	-4.6941	1.7292	0.289	-5.0509	-3.8445	-14.205	35	0.0000
Control class	-0.0554	1.4221	0.2471	-0.7342	0.3162	-0.757	37	0.4385

5 Conclusion

The article constructs an online-offline hybrid five-step teaching model based on French, and uses Transformer and BERT natural language processing methods to construct an initial cross-language model (XLM); then fuses BERT and word embedding with a French neural machine translation model to realize the design of an AI+project-based French translation course. The results show:

(1) The learning rate of this paper's model increases with the number of training iteration rounds Steps to reach the maximum value when 1965 times, the decline of Loss value slows down after 1,542 iterations, the ACC value reaches more than 0.8 after 1,550 iterations, and the model's translation performance reaches the highest point when the temperature coefficient is 0.3.

(2) The model in this paper allows BERT to learn domain-specific feature representations

more efficiently, which in turn leads to better performance improvement of the multi-domain translation model. In addition, in the low-dimensional space, this paper's model can effectively differentiate sentences from different domains, learn domain-specific features and recognize them, and its recognition performance is better than the existing comparison models.

(3) The translation model in this paper can significantly improve students' French translation level after it is applied to teaching, which proves that the online-offline hybrid five-step teaching model based on AI+project-based proposed in this paper is effective in the application of the French translation course, and it can be popularized and used in real teaching.

About the Author

Cen Peng was born in 1987 in Wuhan, Hubei Province, China. She obtained her bachelor's degree in French from Guangdong University of Foreign Studies in China. She obtained her master's degree in French from Wuhan University in China. She is currently a lecturer in the School of Foreign Languages, Zhixing College of Hubei University, and has been teaching French for a long time.

Ying Chen was born in 1986 in Xiangyang, Hubei Province, China. She obtained her bachelor's degree in French from Wuhan University in China. She obtained her master's degree in Language Teaching, French as a Foreign Language and French as a Second Language from the University of Perpignan, France. She is currently a lecturer in the School of Foreign Languages, Zhixing College of Hubei University, and has been teaching French for a long time.

Dinglin Cui was born in 1991 in Wuhan, Hubei Province, China. She obtained her bachelor's degree from Central Queensland University in Australia. Her main research direction is Application Development.

References

[1] Wu, Y. Q., Li, S. S., & Zeng, J. (2024). Language learners’ stereotypical representations of France: Insights for teaching French as L3 with implications for minority foreign language crisis in China. Language Teaching Research, 13621688241304965.

[2] Gallego-Hernández, D. (2025). Business translation research from the West and China: A comparative bibliometric analysis. Babel.

[3] Shen, L. (2025). Constrained communication of motion events: Exploring shared cognitive constraints in Chinese-French translation and description tasks. Review of Cognitive Linguistics.

[4] Shen, L., & Xie, W. (2026). Socio-Cognitive Correlation Between Translation Attitudes and Translation of Motion Event Expressions: Examining its Presence in Chinese-French Translation and the Role of Measurement Approaches and Proficiency Levels. Vigo International Journal of Applied Linguistics, (23).

[5] Zhong, W., Zhao, T., & Xu, M. (2020). Professional interpreting translation education in the Chinese mainland: History, achievements, challenges and future prospects. Babel, 66(6), 883-901.

[6] Shan, Z., & Xu, H. (2023). Teacher beliefs about teaching French as a foreign language in a Chinese university: a multilingual perspective. Porta Linguarum Revista

Interuniversitaria de Didáctica de las Lenguas Extranjeras, (VIII), 169-182.

[7] Hang, S. U. (2024). A Study of French Language Teaching in Primary Schools in China. International Journal of Linguistics, Literature and Translation, 7(3), 59-69.

[8] Huang, T. H., & Wang, L. Z. (2023). Artificial intelligence learning approach through total physical response embodiment teaching on French vocabulary learning retention. Computer Assisted Language Learning, 36(8), 1608-1632.

[9] Liu, J., & Wu, H. (2025, June). Quality Assessment of Chinese-French Translation Based on Language Model. In 2025 IEEE 5th International Conference on Software Engineering and Artificial Intelligence (SEAI) (pp. 1-5). IEEE.

[10] Pu, W. (2025, May). The Application of Deep Learning-Based Intelligent Translation Technique Training and Evaluation System in French Language Teaching. In International Conference on Intelligent Technology for Educational Applications (pp. 129-138). Singapore: Springer Nature Singapore.

[11] El Jemli, S., & LAMIMI, I. J. (2025). Computer-Assisted Translation and ChatGPT: A Case Study in Teaching English-French Translation at Euromed University of Fès. International Journal of Language and Literary Studies, 7(4), 416-425.

[12] Zhao, X., & Fang, L. (2025). EXPLORING THE APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE IN UNIVERSITY FRENCH TEACHING. World Journal of Educational Studies, 64.

[13] Saragih, A. H., & Simbolon, N. (2019, December). The Use of Project Based Learning in Teaching French. In 4th Annual International Seminar on Transformative Education and Educational Leadership (AISTEEL 2019) (pp. 292-295). Atlantis Press.

[14] Nakada, A., Kobayashi, M., Okada, Y., Namiki, A., & Hiroi, N. (2018). Project-based learning. Journal of the Medical Society of Toho University, 65(4), 157-163.

[15] Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. International journal of educational research, 102, 101586.

[16] Hastürkoğlu, G., & Özer, Ö. B. (2020). A Project-Based Approach in Translation Classroom: From the Students’ Perspectives. International Journal of Languages' Education and Teaching, 8(2), 335-344.

[17] Al-Sowaidi, B. (2021). Use of project-based training in teaching business translation. European Journal of Education and Pedagogy, 2(4), 11-16.

[18] Li, N. (2023). Project-based learning in translation teaching: A case of audiovisual translation project. The Educational Review, USA, 7(6).

[19] Retnaningsih, W. (2023). Project based learning in teaching translation: Effect on students’ performance. Journal of Educational Management and Instruction (JEMIN), 3(1), 1-8.

[20] Tambunan, E. E., Ramadhani, Y. R., & Sibuea, B. (2024). The impact of project-based

learning on collaborative and critical thinking skills of students in translation course. ETANIC Journal of English language teaching and applied linguistics, 2(1), 65-79.