International Journal for Housing Science and Its Applications
Publish September 8, 2025. Volume 47, Issue 1 Pages 1-11

https://doi.org/10.70517/ijhsa47101

A Strategy for Improving the Accuracy of Context-Aware Translation Based on Deep Reinforcement Learning in English Translation

Ping Yin[1] and Juanyin Liu[1,*]
[1] Teaching Department of Basic Courses, Hebei Vocational University of Industry and Technology, Shijiazhuang, Hebei, 050091, China
Corresponding authors: (e-mail: eiessay777@126.com).

Abstract The development of globalization has led to increasingly stringent requirements for translation accuracy. This paper designs a cross-language English translation model based on deep reinforcement learning and translation quality assessment, and selects the Transformer architecture with multi-layer encoder-decoders. Through the reward and punishment mechanism of the intelligent NMT system, real-time probability calculations of contextual information are performed to select the most appropriate words at the semantic level as components of the target sentence. Combined with a supervised quality assessment module, the translated text is scored, and the next word selection is guided. Experiments show that after 934 iterations, the BLEU score stabilizes around 97.62%. The F1 score reaches 99.87% after 316 iterations, and the accuracy achieves a stable value of 90.74% after 815 iterations. In two-class cross-language English translation tasks, the model's average BLEU scores were 90.67% and 93.08%.

Index Terms deep reinforcement learning, contextual information, agent, quality assessment, English translation

I. Introduction

With the continuous development of artificial intelligence technology, machine translation is gradually becoming an indispensable tool in people's daily lives. Machine translation is a key task in natural language processing, with the aim of enabling computers to automatically translate one natural language into another [1]. The concept of machine translation can be traced back to the 17th century, but it wasn't until the 1950s that concrete proposals for its implementation emerged. Among these, the most notable was Weaver's memorandum, which marked the beginning of machine translation research [2]-[4]. In this memorandum, Warren Weaver proposed using computers for translation, particularly suggesting the integration of statistical knowledge, cryptography, information theory, logic, and linguistics to address issues of linguistic ambiguity [5], [6]. Since then, machine translation has undergone a series of developments; however, to this day, it remains an extremely challenging problem that has yet to be fully resolved.

Machine translation has gone through three key stages: rule-based, statistical learning-based, and deep learning-based. However, due to the complexity and ambiguity of language, machine translation systems may produce errors or inaccuracies during the translation process, necessitating further optimization and improvement [7]-[9]. Additionally, in traditional machine translation, sentences are often translated in isolation without considering their contextual relationships within the broader text [10], [11]. However, language serves as a medium for information transmission, where each sentence is interconnected. Therefore, translating sentences in isolation may lead to semantic shifts and confusion, failing to accurately convey the original meaning [12]. Against this backdrop, context-aware research has emerged as a hot topic in the field of machine translation.

Context-aware translation utilizes the contextual information surrounding a sentence to more accurately understand and translate it, including the context before and after the sentence, the logical relationships between sentences, and the internal logical structure of the entire text [13], [14]. By incorporating context-aware technology, machine translation systems can better understand the semantic and logical structure of the entire text, thereby improving translation accuracy and fluency [15], [16].

Although context-aware translation theoretically enhances the quality of machine translation, it still faces several challenges in practical applications. First, effectively capturing the contextual relationships between sentences remains a significant challenge. Second, the issue of long-range dependencies between sentences is another critical challenge. Additionally, there is a lack of large-scale corpora. Throughout history, scholars have offered their insights. Reference [17] explored the mechanism of the BERT model in context-aware neural machine translation, which involves concatenating context sequences into a longer sequence, encoding the sequence using the BERT model, and achieving good BLEU scores in several English translation tasks. Reference [18] applied an adjusted BERT model to analyze long sentence structures and extract semantic relationships in machine translation systems, thereby optimizing the long-distance dependency issue in neural network machine translation systems. By combining an attention-guided mechanism, it mitigated ambiguity and omission quality issues in translation. Literature [19] combines context-aware mechanisms, sequence-to-sequence models, and self-attention mechanisms to enhance machine translation's understanding of textual contextual relationships. Through pre-training and large-scale corpus extraction of lexical and syntactic structural relationships, it utilizes contextual information and introduces an adaptive adjustment mechanism to perform real-time adjustments to the system, thereby improving the translation accuracy of machine translation. Literature [20] starts from the source language side and constructs a pre-training method based on global context (including self-supervised pre-training tasks and pre-training models). The model consists of a global encoder and a sentence encoder-decoder. This method significantly improves the BLEU score of the context-based machine translation model. Literature [21] optimizes context-aware neural machine translation by preserving the decoder state during the translation of the current sentence, calculating attention vectors based on this state, and incorporating the target-side context as decoder embeddings into the neural machine translation model. Literature [22] constructs an AI translation model that uses deep learning to establish semantic relationships between the source-side language and target-side language, combining contextual information to improve translation accuracy and fluency. Reference [23] utilizes a transformer-type neural network with attention mechanisms to eliminate ambiguity in machine translation, supported by deep learning. It combines context embedding and syntactic-semantic analysis to construct a context-aware machine translation system, thereby improving translation accuracy and fluency. Reference [24] developed a hierarchical context encoder enabling the hierarchical full attention network structure to utilize multiple context sentences, achieving good BLEU scores in three English translation corpora experiments, thereby enhancing the performance of context-aware neural machine translation. Reference [25] designed a graph-based encoder to enhance context awareness in neural machine translation, primarily by encoding co-reference relationships in text, achieving a BLEU score of 0.9. Meanwhile, [26] embedded an input coreference model into machine translation using context features, leveraging coreference features in the input for translation decisions, resulting in a BLEU score 1 point higher. This demonstrates that the machine translation results under this processing method exhibit extremely high alignment with reference translations. [27] developed a Japanese-English dialogue corpus, which also includes cross-sentence context, and applied it to improve context-aware machine translation.

Deep reinforcement learning combines the advantages of deep learning and reinforcement learning, enabling it to address more complex problems to a certain extent. It employs deep neural networks to handle machine translation tasks [28]. In reinforcement learning, deep neural networks can serve as the policy network for an agent, selecting target language sentences based on input source language sentences, while reinforcement learning algorithms are used to train the neural network's parameters to generate superior translation results [29], [30].

This paper uses a Transformer with multi-level encoders and decoders as the basic structure of a cross-language English translation model. By combining cross-language contrastive learning loss calculations, it improves the performance of context complex information perception and semantic association capture mapping. Through document topic distribution analysis, as well as minimizing change loss and semantic distance calculations, the model associates phrase vector representations between the source and target ends, integrating contextual information. The problem of evaluating semantic deep reinforcement learning strategies is quantified as a step-accumulated reward optimization problem, and an evaluation mechanism is introduced to estimate the state-action value function values, thereby identifying the optimal model training path and enhancing cross-language English translation accuracy.

II. Building a cross-language English translation model architecture under deep reinforcement learning

II. A. Building a cross-language English translation model

II. A. 1) Model Framework

Figure 1 shows the overall framework of the proposed cross-language English translation model based on deep learning and context awareness. The model takes a pair of parallel sentences as input and uses the encoder-decoder of the cross-language Transformer to calculate the normal cross-entropy loss. The proposed model compares parallel corpora and monolingual corpora to learn from them, thereby reducing the expressive gap between different languages and improving translation quality. Additionally, this paper introduces an effective alignment augmentation technique that minimizes training discrepancies by calculating the contrastive loss between aligned pairs (positive samples) and randomly selected unaligned pairs (negative samples), thereby improving training quality. The proposed model fully leverages knowledge from all supervision directions, minimizing the representational distance between similar sentences through cross-lingual contrastive learning, ensuring that the model adequately learns cross-lingual representations of similar sentences in a shared space.

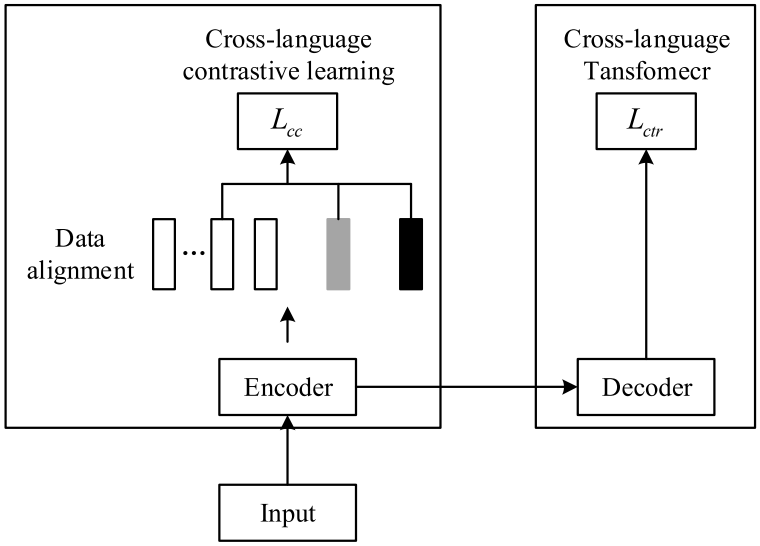

Figure 1: shows the overall framework of the model

II. A. 2) Cross-language Transformer

Cross-language English translation models aim to learn a many-to-many mapping function f for translating from one language to another. To distinguish between different languages, this paper adds an additional language identification tag before each sentence to differentiate between the source and target sides. The Transformer architecture is widely used in natural language processing (NLP) due to its ability to effectively capture contextual information and semantic relationships in text predictions. Therefore, the cross-language English translation model proposed in this paper, which is based on deep learning and context awareness, is built on the Transformer architecture.

Cross-lingual Transformers can implicitly learn shared representations across different languages and perceive complex relationships between sentences based on context. The cross-lingual Transformer proposed in this paper features 10 encoder layers and 10 decoder layers to increase model capacity, thereby better learning semantic relationships between different languages. To simplify the training process of deep models, normalization is applied to the encoder and decoder for word embedding and pre-normality residual connections.

Let the cross-language corpus collection be $L=\{L_1,L_2,\cdots,L_M\}$, where M is the number of cross-languages. Let D_{ij} denote the parallel corpus dataset formed by the corpora L_i and $L_j(i,j\in M)$, and D denote the set of all parallel corpus datasets. The cross-lingual Transformer training loss is defined as the cross-entropy loss function, specifically defined as:

$$L_{ce}=\sum_{x^i,x^j\in D}-\log P_\theta(x^i\mid x^j) \tag{1}$$

In the equation: L_{ce} is the cross-language Transformer training loss; x^i is a sentence in the corpus L_i; x^j is a sentence in the corpus L_j; θ is the parameter of the cross-language Transformer model; P_θ is the predicted probability.

II. A. 3) Cross-language comparative learning

The proposed model introduces cross-language contrastive learning loss, which maps different languages to a shared semantic space, thereby improving the learning of correlations between different languages. The core idea of contrastive learning is to minimize the representation gap between similar sentences and maximize the representation difference between unrelated sentences.

Given a bilingual translation pair $(x^i,x^j)\in D$, where (x^i,x^j) is a positive sample. Next, randomly select a sentence y^j from the corpus L_j to form a negative sample (x^i,y^j). Note that $L_j=L_i$ may exist. The goal of cross-lingual contrastive learning is to minimize the following loss function:

$$L_{ctr}=-\sum_{x^i,x^j\in D}\log\frac{e^{\sin^+(R(x^i),R(x^j))/\tau}}{\sum_{y^j}e^{\sin^-(R(x^j),R(y^j))/\tau}} \tag{2}$$

In the equation: L_{ctr} is the cross-language contrastive learning loss; $sin(\cdot)$ is the similarity function for calculating the similarity between different sentences; + and - represent positive and negative signs, respectively; $R(\cdot)$ is the average aggregated encoding output function for any sentence; τ is a control parameter used to distinguish the difficulty of positive and negative samples. Additionally, the similarity function between two sentences, $sim(\cdot)$ is calculated using the cosine similarity of the average aggregated encoding output. To simplify the process, negative samples are sampled from the same training batch. Then, by maximizing the softmax term $\mathrm{sim}^{+}(R(x^{i}), R(x^{j}))$ and comparing the loss, the semantic projections of two similar sentences are forced to be close to each other. Meanwhile, the softmax function also minimizes the mismatched pair $\mathrm{sim}^{-}(R(x^{i}), R(x^{j}))$, ensuring that the semantic projections of two dissimilar sentences are separated further apart.

In the training process of the proposed deep learning-based and context-aware cross-language English translation model, the model can be optimized by jointly minimizing the contrastive training loss and the Transformer training loss:

$$L = L_{ce} + \lambda |s| L_{ctr} \tag{3}$$

In the formula: λ is the balancing coefficient for balancing the two training losses; $|s|$ is the average sequence length. Since L_{ctr} is calculated at the sentence level, while L_{ce} is calculated at the corpus level, L_{ctr} should be multiplied by the average sequence length $|s|$.

II. B. Adding a bilingual recurrent autoencoder to the thematic context

II. B. 1) Contextual representation modeling based on thematic information

The meaning of a phrase is influenced by its context. When placed in different contexts, this paper can determine its specific meaning with a high degree of probability. To further strengthen the bilingual representation of phrases, this paper incorporates contextual information into the representation learning of phrases.

Following the trend of topic-based machine translation, this paper uses the topic distribution of a document to represent the context of phrases in the document. However, in neural networks, the main issue is how to represent data. Here, for computational convenience, each topic is treated as a "word," and each topic is also represented by an n-dimensional vector. Similar to the vector representation of words, the vector representations of all topics form a topic embedding representation matrix $L_z \in R^{n\times|Z|}$.

Since the topic representation of a document is a probability distribution over all topics, this paper performs a weighted sum of the vector representations of topics based on their probabilities to obtain the topic representation of the document:

$$dc = \sum_{z} p(z \mid d) \cdot \vec{z} \tag{4}$$

In this context, z and $\vec{z}$ correspond to the topic and its embedded representation, respectively; dc denotes the topic of the document.

During training, this paper can use the method mentioned above to obtain the context representation of the phrase pair (f, e), where dc_f and dc_e represent the topic context of phrases f and e, respectively. Since the context is obtained from monolingual documents, this paper can use additional corpora to better train the topic model. It is worth noting that during testing, this paper can only obtain the source document's topic context information. Machine translation is the process of translating a text from one natural language to another. After translation, although the form of semantic expression changes, the essence of the semantics and topics remains unchanged. At the same time, since the topic modeling of the source text and the target text are independent, the topic distributions of the two models may not be in the same vector space. Based on the above conditions, to obtain the target-side topic distribution, this paper can learn a transformation relationship between the source-side and target-side topic contexts. Specifically, for parallel phrase pairs (f, e) and their topic contexts (dc_e, dc_f), this paper learns the transformation relationship between topic contexts by minimizing the change loss:

$$E_{tcm}(f \mid e;\theta) = \frac{1}{2} \| dc_e - f(W_{f2e}^{(3)} \cdot dc_f + b_{f2e}^{(3)}) \|^2 \tag{5}$$

Among these, $W_{f2e}^{(3)} \in R^{n\times n}$ and $b_{f2e}^{(3)} \in R^{n\times 1}$ are the parameters to be learned. In this way, when the model is tested, this paper can use the learned transformation relationship to transform the source-end topic context to the target-end, thereby obtaining the target-end topic context.

II. B. 2) Bilingual semantic constraints

The model in this paper is based on the premise that parallel phrase pairs have identical semantics. The semantic equivalence of phrases is conditional, meaning that under the same thematic context, the source phrase and the target phrase have identical semantics. Therefore, under the same thematic context, the representations of the two phrases can be mutually supervised during learning, with the source phrase representation serving as the true representation of the target phrase, and vice versa.

First, this paper uses a fully connected network to obtain the phrase representation p_{dc} that is known in the thematic context:

$$p_{dc} = g(W^{(4)}[p;dc] + b^{(4)}) \tag{6}$$

Here, $W^{(4)} \in R^{n\times 2n}$ and $b^{(4)} \in R^{n\times 1}$ are network parameters with learning. Where p is the phrase representation learned using RAE, dc is the topic context corresponding to the phrase obtained from formula (6), and p_{dc} is the phrase representation with topic context information.

The learning process for obtaining the final short phrase representations at the source and target ends is independent, so the corresponding phrase representations with topic information may be in different vector spaces. Even if the phrases have the same semantics, the semantic distance between them cannot be directly calculated in this paper. Similarly, the calculation of semantic distance involves two directions: the semantic distance between the source-end phrase representation mapped to the target-end and the target-end phrase representation, and the semantic distance between the target-end phrase representation transformed to the source-end and the source-end phrase representation. Taking the source-end mapped to the target-end as an example, the semantic loss is defined as:

$$E_{sem}(f \mid e;\theta) = \frac{1}{2} \| e_{dc} - f(W_{f2e}^{(5)} f_{dc} + b_{f2e}^{(5)}) \|^2 \tag{7}$$

Here, $W_{f2e}^{(5)} \in R^{n\times n}$ and $b_{f}^{(5)} \in R^{n\times 1}$ are the parameters to be learned.

Then, this paper also calculates the semantic loss by maximizing the semantic distance. For positive samples (f,e) and negative samples $(f,e^{'})$, I can obtain:

$$E_{sem}^{*}(f \mid e;\theta) = max\{0, E_{sem}(f \mid e;\theta) - E_{sem}(f \mid e';\theta) + 1\} \tag{8}$$

Here, $e^{'}$ is a negative sample, which is another translation of f or obtained by randomly replacing words in e. Inter-sample, this paper can obtain $E_{sern}^{*}(e \mid f;\theta)$.

II. C. Translation model architecture based on deep reinforcement learning

II. C. 1) Reinforcement Learning

The basic idea behind reinforcement learning (RL) is that an agent (AG) selects an action (A) based on the current interaction environment (E), after which the environment transitions with a certain probability and provides the agent with a reward (R). The agent repeats this process with the goal of maximizing the reward. Unlike traditional machine learning models, reinforcement learning does not use training data with explicit labels, and the training process is more exploratory in nature. Therefore, reinforcement learning is more suitable for complex decision-making tasks.

Reinforcement learning tasks are typically abstracted as Markov decision processes (MDPs): an agent operates in an environment E, with a state space X, where each state $x \in X$ represents the agent's perception of the environment; The set of all actions the agent can take forms the action space A. When an action $a \in A$ is applied to the current state x, the transition probability function P causes the environment to transition from the current state to another state with a certain probability. Simultaneously, the environment provides feedback to the agent in the form of rewards or penalties based on the reward function. Reinforcement learning tasks can be formalized as a Markov decision process quadruple $E = < X, A, P, R >$, where $P: X \times A \times X \mapsto \mathbb{R}$ specifies the state transition probabilities, $R: X \times A \times X \mapsto \mathbb{R}$ specifies the rewards and punishments, and when the reward function is only related to state transitions, $R: X \times X \mapsto \mathbb{R}$.

When all four elements of the Markov decision process are known, it is called model-known, and learning in a model-known environment is called model-based learning. When the model is known, the expected cumulative reward for any policy can be estimated. Based on the cumulative function, there are state value functions and state-action value functions:

$$V_T^{\pi}(x) = \mathrm{E}_{\pi}\left[\frac{1}{T}\sum_{t=1}^{T} r_t \mid x_0 = x\right] \tag{9}$$

$$Q_T^{\pi}(x,a) = \mathrm{E}_{\pi}\left[\frac{1}{T}\sum_{t=1}^{T} r_t \mid x_0 = x, a_0 = a\right] \tag{10}$$

Since MDP has Markov properties, i.e., the next moment state of the system is determined only by the current state, the value function takes the form of Bellman's equation:

$$\begin{aligned} V_T^\pi(x) &= \mathrm{E}_\pi\left[\frac{1}{T}\sum_{t=1}^{T} r_t \mid x_0 = x\right] \\ &= \mathrm{E}_\pi\left[\frac{1}{T}r_1 + \frac{T-1}{T}\frac{1}{T-1}\sum_{t=2}^{T} r_t \mid x_0 = x\right] \end{aligned} \tag{11}$$

After simple conversion, the state-action value function can be obtained:

$$Q_T^\pi(x,a) = \sum_{x'\in X} P_{x\to x'}^a\left(\frac{1}{T}R_{x\to x'}^a + \frac{T-1}{T}V_{T-1}^\pi(x')\right) \tag{12}$$

From the above, it can be seen that when the model is known, the policy evaluation problem is transformed into a dynamic programming problem. That is, first solve for the single-step cumulative reward at each time step, then solve for the two-step cumulative reward at each state, and iterate until the T -step cumulative reward at each state is obtained.

II. C. 2) Translation Model Framework Structure

To address exposure bias and translation diversity issues, this paper proposes a cross-language English translation model (NMT-Transformer) based on reinforcement learning and machine translation quality assessment. NMT-Transformer introduces an evaluation mechanism at the sentence level to guide the model's predictions, ensuring they do not fully converge on the reference translation. Figure 2 shows the specific framework structure of the model. It primarily consists of two modules: machine translation and machine translation quality assessment. The model's translation module adopts an encoder-decoder architecture consistent with the Transformer, while the assessment module employs a sentence-level machine translation quality assessment model.

In the machine translation process, the NMT system acts as an intelligent agent in reinforcement learning, continuously interacting with the environment to obtain information about the current state of the environment. Specifically, the probability $P(y_t \mid x, \hat{y}_{<t})$ of the source sentence x and the generated target sentence context at time step t, where $\hat{y}_{<t}$ denotes the target sentence predicted by the model prior to time step t. The agent makes decisions on the next word to select based on the current state of the environment, simultaneously obtaining the reward value after executing the word selection operation in the current state and transitioning to the next state. Through reinforcement learning, the agent ultimately identifies the optimal translation strategy. Based on the model structure, the machine translation task is described as follows: Given parallel corpora, train a machine translation model M_θ with parameters θ; The machine translation model M_θ translates the given source sentence sequence $x = (x_1, x_2, ..., x_n)$ into the target sentence sequence $y = (y_1, y_2, ..., y_n)$ where n and m are the lengths of the source and target sentence sequences, respectively; at time step t, the state y_t defines the target sentence $y_t = (y_1, y_2, ..., y_t)$ generated by the translation model M_θ at the current time step t; the action a_t is defined as selecting the next word y_{t+1} in the current environment; given the translation data and its HTER score, the machine translation quality assessment model Q_φ with training parameters φ, the quality assessment model Q_φ after supervised training serves as the generator of the reward function to provide a quality score $Score_{QE}(\hat{y}_t)$ for unseen translations, The machine translation model M_θ interacts with the environment under the guidance of $Score_{QE}(\hat{y}_t)$ to generate the next word y_{t+1}.

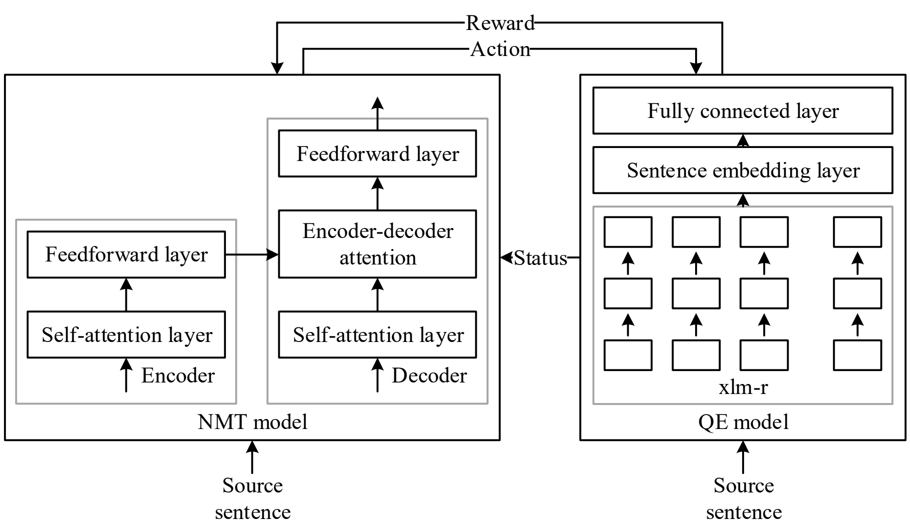

Figure 2: NMT-Transformer translation model

III. Performance verification and application analysis of cross-language English translation models

III. A. Analysis of the model training process

III. A. 1) Comparison of model performance on different test datasets

Different corpus resources have varying effects on model training performance. In this section, we select different corpus resources as test datasets and set up model comparison experiments. To provide a clearer comparison of the experiments, the models and methods based on different types of corpus resources are compared in the experiments. These three types of corpus resources are categorized into three groups: using only bilingual parallel corpora, using only monolingual corpora, and a mixture of bilingual parallel corpora and monolingual corpora. For convenience, these three types are labeled as "A," "B," and "C," respectively. Specifically, the baseline methods based on "Type A" corpus include ConvS2S and Transformer; the baseline methods based on "Type B" corpus include "RNN+back translation" and "unsupervised PBSMT"; and the methods based on "Type C" corpus include the baseline method "duality machine translation" and the model proposed in this paper, "Transformer+NMT." Table 1 presents the model performance results across different test datasets in the experiments. The proposed model achieves the highest BLEU translation accuracy among the six models by combining bilingual parallel corpora with monolingual corpus resources for training. In English-to-Chinese translation, the BLEU scores of the proposed model are 29.41%, 29.11%, and 29.54%; in Chinese-to-English translation, the BLEU scores are 38.89%, 37.97%, and 38.62%. Through a comparison of cross-language English translation methods based on three different types of corpora, it is clearly demonstrated that the proposed model achieves superior translation performance.

Table 1: Models performance results on different test datasets in the experiment

Corpus type	Model	BLEU values of the model on different datasets (%)					
		$NMT_{EN\to Zh}$			$NMT_{Zh\to EN}$		
		NIST04	NIST08	NIST12	NIST04	NIST08	NIST12
A	ConvS2S	23.37	23.14	24.02	28.95	28.12	28.76
	Transformer	26.15	26.22	26.95	32.62	33.54	33.63
B	RNN+ reverse translation	25.02	25.16	26.08	31.71	30.36	33.54
	Unsupervised PBSMT	26.64	26.58	27.47	35.56	33.05	35.68
C	Dual machine translation	26.96	27.17	28.13	36.79	35.81	37.35
	Transformer+NMT	29.41	29.11	29.54	38.89	37.97	38.62

III. A. 2) Analysis of Model Performance Changes in Translation Tasks

Figure 3 illustrates how the performance of the proposed model varies with the number of iterations in the Chinese-to-English and English-to-Chinese translation tasks on the NIST2020 validation dataset. Figure 4 shows how the perplexity values of the generated corpora vary with the number of iterations in the Chinese-to-English and English-to-Chinese translation tasks on the NIST2022 test dataset. Whether it is BLEU or the perplexity value of the generated corpus, both reach a relatively stable state after the second iteration of the model. After 8 iterations, the English-to-Chinese translation accuracy of the model reaches 38.91%, and the perplexity value of the generated Chinese corpus decreases to 75.82. The Chinese-to-English translation accuracy reached 29.79%, and the generated English corpus perplexity value decreased to 80.13%. Achieving high accuracy and a significant reduction in generated corpus perplexity values after only 8 iterations suggests that the model architecture is effective and can be further tested for performance and specific translation tasks.

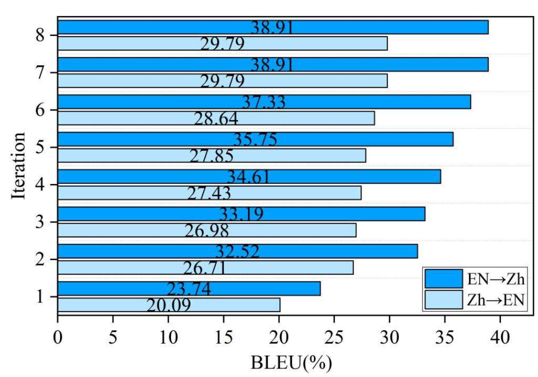

Figure 3: The variation of model performance with the number of iterations

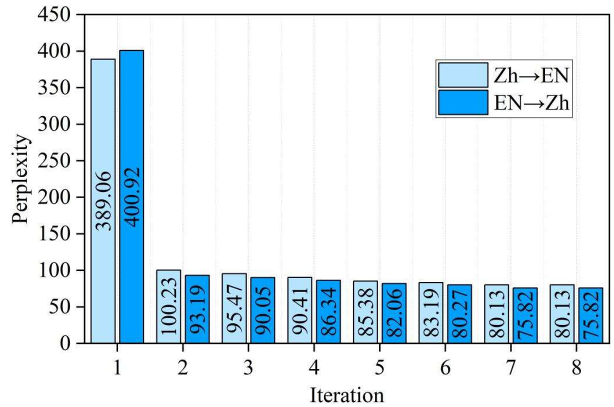

Figure 4: Generate expected confusion degree varying with number of iterations

III. B. Model Performance Evaluation

III. B. 1) Comparison of BLEU scores for different models

After initially assessing the effectiveness of the Transformer+NMT cross-language English translation model architecture described in this paper, we selected similar cross-language translation models—XLM, mBERT, and SeamlessM4T—as comparison models and trained them on large datasets. Figure 5 shows the specific changes in BLEU scores for each model as the number of epochs increases. After 934 iterations, the BLEU score of the model in this paper reached a stable 97.62%. After 965 iterations, the XLM model reached a stable 67.25%. The mBERT model's BLEU score stabilized at 72.31% after 960 iterations. The SeamlessM4T model's BLEU score stabilized around 77.57% after 912 iterations. The proposed model's final BLEU score exceeded 95%, demonstrating very high translation accuracy and significantly outperforming the comparison models. This validates the advantages of the deep reinforcement system and the context-aware thematic perception module.

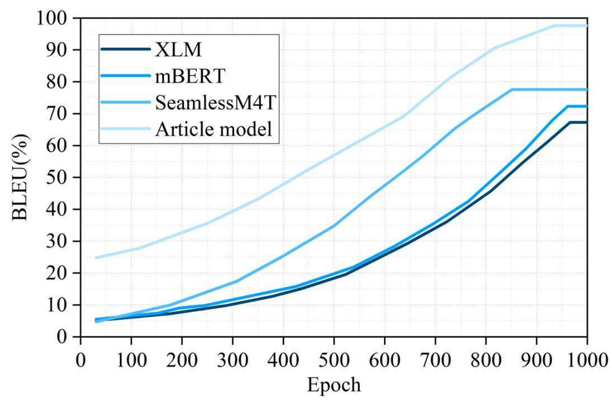

Figure 5: BLEU value of the model varies with epochs

III. B. 2) Module Performance Testing and Results Analysis

Validate the effectiveness of the model architecture in this paper on a module-by-module basis. The cross-language English translation model of the reinforcement learning-based NMT system plays a crucial role in predicting the target sentence. Accuracy is the percentage of correctly predicted target sentences out of the total number of samples. While accuracy can assess overall correctness, it is not an ideal metric for evaluating results in cases of sample imbalance. Definition and understanding of the F1 score: The F1 score is defined based on each category and encompasses two key concepts: precision (P) and recall (R). Accuracy refers to the probability that a sample predicted as positive is actually positive. Recall refers to the probability that a sample that is actually positive is predicted as positive. Figure 6 shows the changes in F1 score and accuracy of the model containing only the NMT system as the training epoch increases. The F1 score reaches its peak of 97.87% at 953 iterations, while accuracy peaks at 80.92% at 976 iterations. Both the F1 score and accuracy of the NMT system model using reinforcement learning exceed 80%, indicating that the prediction of target sentences is effective, thereby validating the application effectiveness of the NMT system.

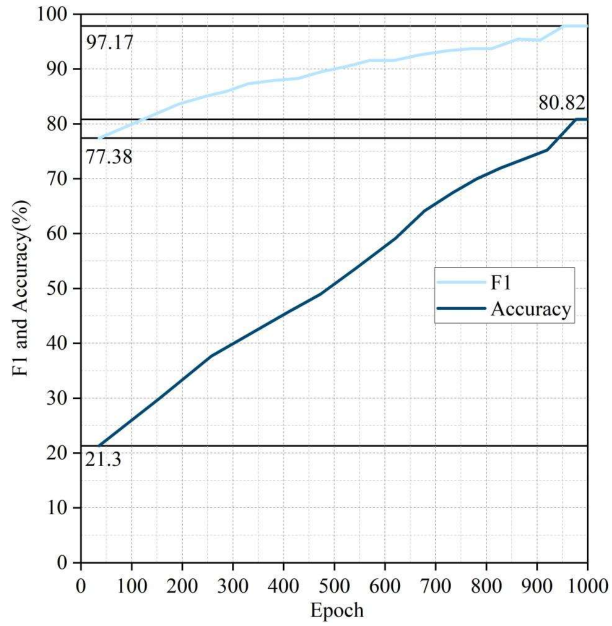

Figure 6: F1 value and Accuracy vary with the training epoch

Figure 7 shows the specific changes in the F1 score and accuracy of the proposed model as the number of iterations increases, after incorporating context quality assessment. After introducing the context quality assessment module, the F1 score of the proposed model ranged from a minimum of 90.43% to a maximum of 99.87%, and stabilized after 316 iterations. The accuracy stabilized at 90.74% after 815 iterations. Overall, the context-aware architecture based on deep reinforcement learning effectively handles sample training and cross-language English translation.

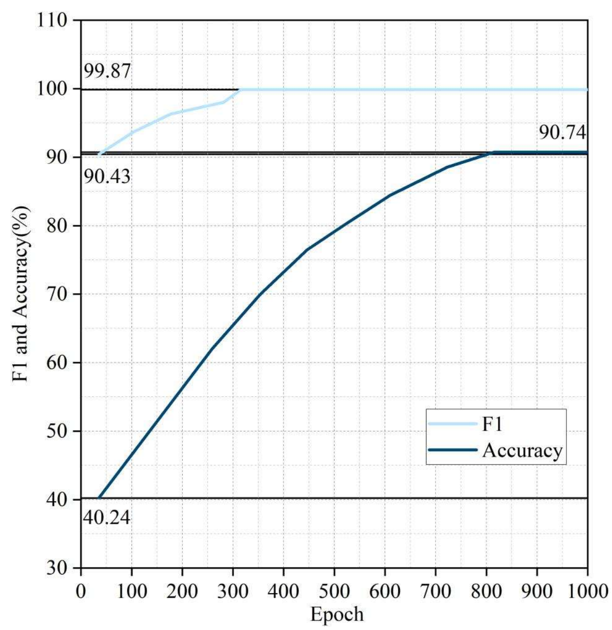

Figure 7: F1 value and Accuracy of article model vary with number of iterations

III. C. Model Translation Task Evaluation Results

III. C. 1) Evaluation results for different translation tasks

The model's actual cross-language English translation performance was evaluated through Chinese-to-English and German-to-English translation tasks. After completing model training, the top 30,000 high-frequency words from the bilingual corpus were selected as the vocabulary list, with a maximum sentence length of 60 words for the training corpus. Multiple models were selected as comparison models to compare translation results. Table 2 shows the BLEU evaluation results for the

Chinese-to-English translation task. Table 3 shows the BLEU evaluation results for the German-to-English translation task. In both the Chinese-to-English and German-to-English translation tasks, the BLEU scores of the proposed model exceeded 90%, with a maximum of 94.18%. The average scores were 90.67% and 93.08%, respectively. This demonstrates that the proposed model achieves high translation accuracy in actual cross-language English translation tasks.

Table 2: BLEU evaluation results for Zh-EN translation tasks

Zh-EN	BLEU (%)					
	NIST04	NIST06	NIST08	NIST010	NIST012	AVG
Pbsmt	76.37	85.28	84.24	79.81	77.81	80.45
GlobalAtt	80.92	88.95	88.51	82.67	80.76	84.36
Sennrich-deponly	80.19	89.51	87.28	83.15	81.24	84.35
LocalAtt	81.25	89.35	88.84	82.93	81.13	84.84
FlexibleAtt	80.98	89.37	88.67	83.02	80.99	84.47
SDRNMT	81.56	88.23	89.09	83.95	81.92	85.01
Transformer+NMT	90.35	90.16	91.65	90.24	90.94	90.67

Table 3: BLEU evaluation results for DE-EN translation tasks

DE-EN	BLEU (%)					
	NIST04	NIST06	NIST08	NIST010	NIST012	AVG
Pbsmt	79.82	89.13	86.18	81.07	79.74	83.04
GlobalAtt	84.37	92.97	90.45	83.93	82.69	87.06
Sennrich-deponly	83.64	93.53	89.22	84.42	83.17	86.91
LocalAtt	84.17	93.36	90.78	84.19	83.06	87.01
FlexibleAtt	84.43	93.39	90.61	84.28	82.92	87.02
SDRNMT	85.01	92.25	91.03	85.21	83.85	87.39
Transformer+NMT	93.28	94.18	93.59	91.25	92.89	93.08

III. C. 2) Visualization of sentence translation based on implicit topic representation

Figure 8 illustrates the process of the Chinese-to-English translation task. The original Chinese text on the vertical axis is "Today is Monday, and we are testing the performance of a cross-language English translation model." During the Chinese-to-English translation of this sentence, it can be observed that the model calculates the reward and penalty function values at the word level to determine the final target English sentence. For example, in the word "today," the reward function value for 'Taday' is the highest at 0.4, exceeding the reward values of other translated words. Therefore, this word is translated as "Taday," and the translation process continues sequentially for the next word until the target sentence is output.

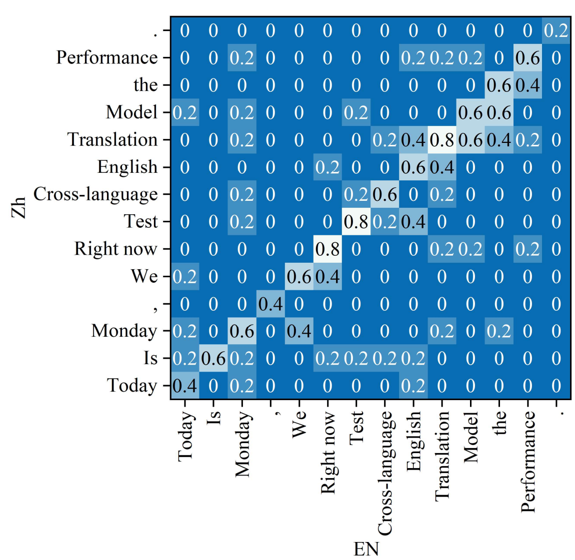

Figure 8: The process of Chinese-English translation tasks

IV. Conclusion

This paper utilizes the Transformer cross-language basic structure and a reinforcement learning-based NMT system to construct a cross-language English translation model and evaluate its translation performance. The model's BLEU score reached a stable value of 97.62% after 934 iterations, outperforming the other three comparison models. The F1 score ranged from [90.43, 99.87]%, with the highest accuracy reaching 90.74%. The BLEU score exceeds 90% in various English translation tasks, indicating good translation performance. Considering that this paper only tested two actual English translation tasks, to fully validate the translation performance of the designed model, future studies could set up more English translation tasks in different languages to verify the reliability of the results presented in this paper.

References

[1] Al-khresheh, M. H., & Almaaytah, S. A. (2018). English proverbs into Arabic through machine translation. International Journal of Applied Linguistics and English Literature, 7(5), 158-166.

[2] Mondal, S. K., Zhang, H., Kabir, H. D., Ni, K., & Dai, H. N. (2023). Machine translation and its evaluation: a study. Artificial Intelligence Review, 56(9), 10137-10226.

[3] Gordin, M. D. (2020). The forgetting and rediscovery of Soviet Machine Translation. Critical Inquiry, 46(4), 835-866.

[4] DeCamp, J. A. (2022, September). Thoughts on the History of Machine Translation in the United States. In Proceedings of the 15th Biennial Conference of the Association for Machine Translation in the Americas (Volume 2: Users and Providers Track and Government Track) (pp. 427-437).

[5] Melby, A. K. (2019). Future of machine translation: Musings on Weaver's memo. The Routledge handbook of translation and technology, 419-436.

[6] SALMANOVA, A. (2019). THE FINE LINE BETWEEN CRYPTOLOGY AND MACHINE TRANSLATION. development, 8, 1.

[7] Cao, J., Li, M., Li, Y., Wen, M., Cheung, S. C., & Chen, H. (2022). SemMT: a semantic-based testing approach for machine translation systems. ACM Transactions on Software Engineering and Methodology (TOSEM), 31(2), 1-36.

[8] Turganbayeva, A., Rakhimova, D., Karyukin, V., Karibayeva, A., & Turarbek, A. (2022). Semantic connections in the complex sentences for post-editing machine translation in the Kazakh language. Information, 13(9), 411.

[9] Rivera-Trigueros, I. (2022). Machine translation systems and quality assessment: a systematic review. Language Resources and Evaluation, 56(2), 593-619.

[10] Naveen, P., & Trojovský, P. (2024). Overview and challenges of machine translation for contextually appropriate translations. Iscience, 27(10).

[11] Choi, H., Cho, K., & Bengio, Y. (2017). Context-dependent word representation for neural machine translation. Computer Speech & Language, 45, 149-160.

[12] Seligman, M. (2019). The evolving treatment of semantics in machine translation. Adv. Empir. Transl. Stud. Dev. Transl. Resour. Technol., 53.

[13] Sugiyama, A., & Yoshinaga, N. (2019, November). Data augmentation using back-translation for context-aware neural machine translation. In Proceedings of the fourth workshop on discourse in machine translation (DiscoMT 2019) (pp. 35-44).

[14] Sang, Z. (2018). How does the context make a translation happen? An activity theory perspective. Social Semiotics, 28(1), 125-141.

[15] Zhang, B., Xiong, D., Su, J., & Duan, H. (2017). A context-aware recurrent encoder for neural machine translation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 25(12), 2424-2432.

[16] Ding, C., Fu, Z., Yang, Z., Yu, Q., Li, D., & Huang, Y. (2023). Context-aware linguistic steganography model based on neural machine translation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 32, 868-878.

[17] Wu, X., Xia, Y., Zhu, J., Wu, L., Xie, S., & Qin, T. (2022). A study of BERT for context-aware neural machine translation. Machine Learning, 111(3), 917-935.

[18] Chen, S. (2025). Semantic relationship extraction of English long sentences and quality optimization of machine translation based on BERT model. Journal of Computational Methods in Sciences and Engineering, 14727978251322656.

[19] Jiang, C., & He, Y. (2025). Construction and Evaluation of Context Aware Machine Translation System. Procedia Computer Science, 261, 529-537.

[20] Chen, L., Li, J., Gong, Z., Duan, X., Chen, B., Luo, W., ... & Zhou, G. (2021). Improving Context-Aware Neural Machine Translation with Source-side Monolingual Documents. In IJCAI (pp. 3794-3800).

[21] Yamagishi, H., & Komachi, M. (2019, October). Improving context-aware neural machine translation with target-side context. In International Conference of the Pacific Association for Computational Linguistics (pp. 112-122). Singapore: Springer Singapore.

[22] Wu, Y., & Xu, N. (2024, October). Contextual Aware AI Translation Models in Cross Language Information Retrieval. In 2024 3rd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI) (pp. 164-169). IEEE.

[23] Guo, D. (2025). Deep learning-driven context-aware English translation for ambiguous sentences. International Journal of Information and Communication Technology, 26(15), 41-56.

[24] Yun, H., Hwang, Y., & Jung, K. (2020, April). Improving context-aware neural machine translation using self-attentive sentence embedding. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 05, pp. 9498-9506).

[25] Ohtani, T., Kamigaito, H., Nagata, M., & Okumura, M. (2019, November). Context-aware neural machine translation with coreference information. In Proceedings of the Fourth Workshop on Discourse in Machine Translation (DiscoMT 2019) (pp. 45-50).

[26] Vu, H. H., Kamigaito, H., & Watanabe, T. (2024). Context-Aware Machine Translation with Source Coreference Explanation. Transactions of the Association for Computational Linguistics, 12, 856-874.

[27] Rikters, M., Ri, R., Li, T., & Nakazawa, T. (2021). Japanese–English conversation parallel corpus for promoting context-aware machine translation research. Journal of Natural Language Processing, 28(2), 380-403.

[28] Guo, Q. (2025). Deep Reinforcement Learning Method for Optimizing English Translation Algorithms. International Journal of High Speed Electronics and Systems, 2540568.

[29] Xiao, N. (2025). Innovation and Evaluation of Machine Translation Models Combining Reinforcement Learning Algorithms and RNN. Procedia Computer Science, 261, 821-828.

[30] Kumar, A., Pratap, A., Singh, A. K., & Saha, S. (2022). Addressing domain shift in neural machine translation via reinforcement learning. Expert Systems with Applications, 201, 117039.