



OPEN Japanese waka translation supported by internet of things and artificial intelligence technology

Rizhong Shen

With the advancement of internet of things (IoT) and artificial intelligence (AI) technology, access to large-scale bilingual parallel data has become more efficient, thereby accelerating the development and application of machine translation. Given the increasing cultural exchanges between China and Japan, many scholars have begun to study the Chinese translation of Japanese waka poetry. Based on this, the study first explores the structure of waka and the current state of its Chinese translations, analyzing existing translation disputes and introducing a data collection method for waka using IoT. Then, an optimized neural machine translation model is proposed, which integrates a Bidirectional Long Short-Term Memory (Bi-LSTM) network, vertical Tree-LSTM, and an attention mechanism into the Transformer framework. Experimental results demonstrate that the three optimized models—Transformer + Bi-LSTM, Transformer + Tree-LSTM, and Transformer + Tree-LSTM + Attention—outperform the baseline Transformer model on both public and waka datasets. The BLEU scores of the models on the public dataset were 23.71, 23.95, and 24.12, respectively. Notably, on the waka dataset, the Transformer + Tree-LSTM + Attention model achieved the highest BLEU score of 20.65, demonstrating a significant advantage in capturing waka's unique features and contextual information. This study offers new methods to enhance the quality of Chinese-Japanese translation, promoting cultural exchange and understanding.

Keywords Japanese waka, Internet of things, Neural machine translation, Transformer, Long short-term memory network

As the economies of China and Japan become increasingly interconnected, cultural exchanges between the two countries have flourished. As a literary form uniquely characteristic of Japan, the translation of Japanese waka has attracted considerable scholarly attention¹. In the context of the information explosion, traditional human translation methods are insufficient to meet the overwhelming demand for large-scale data processing. As a result, leveraging machine learning techniques to enhance translation efficiency and process large volumes of data more effectively has become an inevitable societal trend. In recent years, under the “Internet+” strategy, China’s Internet of Things (IoT) technology system has significantly advanced, with related industries experiencing rapid development². The progress of artificial intelligence (AI) technology has further facilitated access to large corpora, providing valuable resources for building statistical language models and improving the performance of machine translation (MT) systems³.

Studying language models has long been a critical component of neural machine translation (NMT) research. As NMT continues to evolve, various innovations have emerged, and NMT has been widely applied across industries. Liu et al. addressed the issue of data sparsity in NMT by proposing a method that integrated several data augmentation techniques, including low-frequency word replacement, back-translation, and a grammar correction module. Their experimental results demonstrated that this approach enhanced translation performance in both resource-rich and resource-poor environments, with particularly notable improvements in the latter due to the augmentation of training corpora⁴. Klimova et al. reviewed the application of NMT in foreign language learning and provided pedagogical recommendations. The study found that NMT effectively enhanced learners’ productive and receptive language skills, particularly for advanced L2 learners⁵. Tonja et al. explored how leveraging source-language monolingual datasets from low-resource languages could enhance NMT system performance. Through self-training and fine-tuning, the researchers achieved significant improvements in BLEU score in the Wolaytta-English translation task⁶. While these studies enhance NMT performance through various methods, they do not specifically address the translation of Japanese waka.

School of Foreign Languages, Quanzhou Normal University, Quanzhou 362000, Fujian, China. email: shenrizhong@qztc.edu.cn

Building on this background, this study examines the structure of Japanese waka and the current state of its Chinese translation, analyzes existing translation controversies, and introduces an IoT-based data collection method. In response to the limitations of traditional AI-based NMT models, an optimized translation model is proposed that incorporates Bidirectional Long Short-Term Memory (Bi-LSTM) networks, vertical Tree-LSTM, and an attention mechanism into the Transformer framework. Experimental validation of these optimized models is performed using both public datasets and a specialized waka dataset, yielding conclusions that demonstrate improvements in translation quality. This study offers innovative methods and theoretical support for the translation of Japanese waka. This study is well-structured, beginning with an introduction that outlines the research background and objectives. “Challenges and complexities in translating Japanese Waka Poetry” section explores the challenges and complexities of translating Japanese waka poetry, discussing academic debates on waka translation and its unique intricacies within the context of machine translation. “Construction of NMT models” section details the development of the NMT model, emphasizing data collection methods that leverage IoT and AI technologies and the model’s optimization through positional encoding techniques. “Results and discussion” section outlines the experimental design and data sources, offering a detailed analysis of how adjusting the number of Transformer encoder layers impacts model performance. Finally, “Conclusion” section concludes the study by summarizing key findings and providing thoughtful recommendations for future research.

Challenges and complexities in translating Japanese Waka Poetry

The dispute of Japanese waka translation

Japanese waka possesses distinctive linguistic features that make its translation into Chinese particularly challenging⁷. In classical Japanese literature, “poetry” primarily refers to ancient Chinese court poetry—a literary form written in classical Chinese and composed according to traditional standards, with a strong emphasis on rhyme. “Songs,” by contrast, generally refer to longer forms of poetry and oral literature native to Japan, including genres such as long poems (chōka), short poems (tanka), and linked-verse poetry (renga)⁸. Waka follows a strict phonetic structure, with the typical tanka form consisting of five lines arranged in a syllabic pattern of 5, 7, 5, 7, and 7. This fixed syllabic pattern governs the rhythm of the poem and is closely tied to the tonal and phonetic characteristics of the Japanese language, enhancing its rhythmic aesthetic. The primary challenge of translating waka stems from the significant differences in phonology and syntax between Chinese and Japanese. Japanese, as an agglutinative language, relies heavily on functional words such as particles and auxiliary verbs to structure sentences, whereas Chinese, as an isolating language, employs a more concise syntactic structure. This typological difference makes it challenging to preserve both the form and content of waka in translation, often resulting in the loss or distortion of semantic nuances and emotional depth. Additionally, waka often employs rich metaphors, wordplay, and concise language to convey intricate emotions and vivid imagery. For instance, seasonal words commonly found in traditional waka, such as “snow” or “cherry blossoms,” are not merely descriptive of nature but also carry specific emotional connotations rooted in cultural context. Such imagery poses a significant challenge for translators, particularly when these words lack equivalent cultural meanings in Chinese or direct counterparts. Thus, translators must balance fidelity to the original content with the accurate conveyance of emotions, often requiring creative compromises.

In China, debates over the translation of waka into Chinese ultimately center on the choice between domestication and foreignization⁹. Proponents of domestication argue that waka often contain varying degrees of content and complex syntax, suggesting that translators should prioritize capturing the original meaning and syntax rather than rigidly adhering to the original form¹⁰. In contrast, advocates of foreignization stress the importance of preserving the cultural essence embedded in waka, supporting the retention of its original prosodic structure and poetic format¹¹. For example, a well-known waka from the classic collection Hyakunin Isshu reads “秋の田の、かりほの庵の、苫をあらみ、わが衣手は、露にぬれつつ,” which can be translated literally as “In the autumn fields, the thatched hut I stay in has a sparse roof, and my sleeves are often wet with dew.” This waka depicts a scene of solitude, using the seasonal word “autumn” to evoke a sense of the passage of time and the fleeting nature of life. In translation, the phrase “苫をあらみ” (a sparse thatch roof) symbolizes hardship and impermanence, but this metaphor is challenging to fully convey in Chinese. Moreover, the original waka’s rhythmic structure is difficult to preserve while maintaining its meaning. Similarly, in a waka from the Man’yōshū, “しらつゆに、風の吹きしく、秋の野は、つらぬきとめぬ、玉ぞ散りける” the word “しらつゆ” (white dew) serves as a metaphor for the transience and fragility of life. A domesticated translation might render this as “The autumn wind blows across the field, white dew covering the ground, like pearls scattered beyond grasp,” which flows more naturally in Chinese and helps readers understand the emotional resonance. However, this approach sacrifices the original structure, particularly the distinctive 5–7 syllabic rhythm. The fundamental differences in vocabulary, syntax, and cultural imagery between Japanese and Chinese make it challenging to balance form and content in translation. The linguistic features of waka are not only governed by intricate phonetic patterns but also exhibit a high degree of metaphorical richness and cultural specificity in word choice and emotional expression^{12–14}.

Challenges in translating Japanese Waka Poetry with machine translation

The unique characteristics of Japanese waka poetry pose significant challenges for conventional machine translation methods. First, waka follows a fixed syllabic pattern of 5-7-5-7-7, which imparts a distinct rhythm and phonetic beauty to the poetry. However, traditional machine translation models process data at the word or sentence level and lack the capability to model syllabic structures. As a result, these models struggle to preserve the rhythmic qualities and formal aesthetics inherent in waka during translation. Additionally, waka incorporates seasonal words and natural imagery that carry specific cultural connotations and emotional expressions. Failure to accurately capture these cultural nuances in translation can lead to semantic distortion or loss. Furthermore,

the highly condensed language style of waka, combined with its reliance on contextual meaning, places a higher demand on semantic interpretation. The agglutinative nature of the Japanese language means that particles and auxiliary verbs play a crucial role in conveying meaning and emotion, which traditional translation methods often fail to address adequately. This can lead to semantic misrepresentation or fragmented translation outputs.

To address these challenges, this study proposes an optimized neural machine translation (NMT) model that enhances processing capabilities for the specific features of waka by integrating lateral Bi-LSTM, vertical Tree-LSTM, and attention mechanisms. Specifically, to model and preserve the phonetic structure, the proposed method introduces a Tree-LSTM module to the Transformer model, capturing hierarchical relationships between waka sentences through syntactic tree structures. Additionally, during the training process, a syllabic length constraint is incorporated to guide the model in generating translations that adhere to the waka form. This approach not only improves the semantic accuracy of the translation output but also helps preserve, to some extent, the rhythmic beauty of waka. To effectively understand and transfer cultural imagery, key symbolic words in waka (such as seasonal terms and natural imagery) are semantically annotated during the data preprocessing phase to ensure that the model recognizes their cultural significance. The optimized attention mechanism further refines the model's focus on these annotated terms, significantly enhancing its ability to capture cultural imagery. To address the challenge of semantic analysis in the complex contextual environment of waka, the combination of Bi-LSTM and Tree-LSTM improves the model's understanding of context both horizontally and vertically. The Bi-LSTM captures global contextual information through bidirectional sequence processing, while the Tree-LSTM leverages syntactic structure to capture hierarchical relationships between words. Together, these two components effectively address the limitations of traditional methods in contextual modeling.

Construction of NMT models

Waka data collection using IoT and artificial intelligence technologies

The collection of waka data in Japan presents a range of complex challenges. Waka, a significant component of Japanese classical literature, differs substantially from modern Japanese and other text types due to its historical origins and unique linguistic characteristics. The classical Japanese language used in waka is distinguished by intricate grammatical structures and an extensive vocabulary, which considerably complicate its digitization and processing. The frequent incorporation of cultural metaphors, seasonal vocabulary, and historical contexts in waka further intensifies the challenges for machine processing, often resulting in misinterpretations or the loss of critical information. Thus, data collection must not only focus on recognizing the nuances of classical language but also prioritize the accurate preservation and transmission of cultural significance. Additionally, the rarity of waka, particularly in the form of classical manuscripts, introduces further difficulties. Issues such as copyright restrictions and the diverse formats of source materials create additional barriers to obtaining high-quality data. Conducting waka data collection and validation in a multicultural and multilingual context demands rigorous filtering of data sources and cross-database comparisons to ensure both accuracy and completeness.

The development of IoT technology has significantly enhanced the collection and processing of language data. The core concept of IoT is to enable interconnected communication between humans and objects, as well as between objects themselves, through sensors and network devices. This allows for the automatic collection, storage, and transmission of data on a large scale^{15–17}. By leveraging IoT technology, text data from diverse sources can be gathered quickly and accurately. This is particularly effective in cross-linguistic and cultural contexts, providing an efficient method for the automated collection of large-scale waka data.

In this study, the data collection focuses on waka found in three major works of classical Japanese literature:

- ① The Tale of Genji: A long novel from the Heian period, regarded as a pinnacle of ancient Japanese literature. It is not only a cultural symbol of Japan but also a rich repository of waka, celebrated for their literary value, linguistic artistry, and cultural significance.
- ② The Man'yōshū: Japan's earliest known anthology of waka poetry, encompassing a vast array of poems on diverse topics, from aristocrats to commoners. It represents the origins of Japanese literature and provides invaluable insights into early Japanese culture.
- ③ The Ise Monogatari: Another influential work like The Tale of Genji, presented in the form of uta monogatari (poem-tales). In this narrative, waka play a central role, enriching the text with profound emotional expression and an elegant linguistic style.

By leveraging IoT technology, including intelligent sensors, Optical Character Recognition (OCR) tools, and cloud computing platforms, the automation and accuracy of data collection can be significantly improved. This facilitates the efficient digitization of complex classical texts. The integration of OCR technology with multilingual support and semantic annotation methods is particularly critical in addressing the unique linguistic and cultural characteristics of waka, ensuring more effective data extraction and processing. The workflow for collecting waka data using IoT technology is illustrated in Fig. 1.

Step 1: Data source confirmation.

Selecting high-quality data sources is critical for building a reliable corpus. This study identifies Japanese classical literary works, including The Tale of Genji, Man'yōshū, and The Tales of Ise, as primary data sources. These texts are celebrated for their literary significance and their representation of the diversity and complexity inherent in classical Japanese waka. During the data confirmation phase, authoritative resources such as the National Literature Database of Japan and other credible digital archives are utilized to ensure reliability and comprehensiveness. Priority is given to academically verified and revised editions to minimize the risk of errors and inconsistencies.

Step 2: Data digitization and extraction.

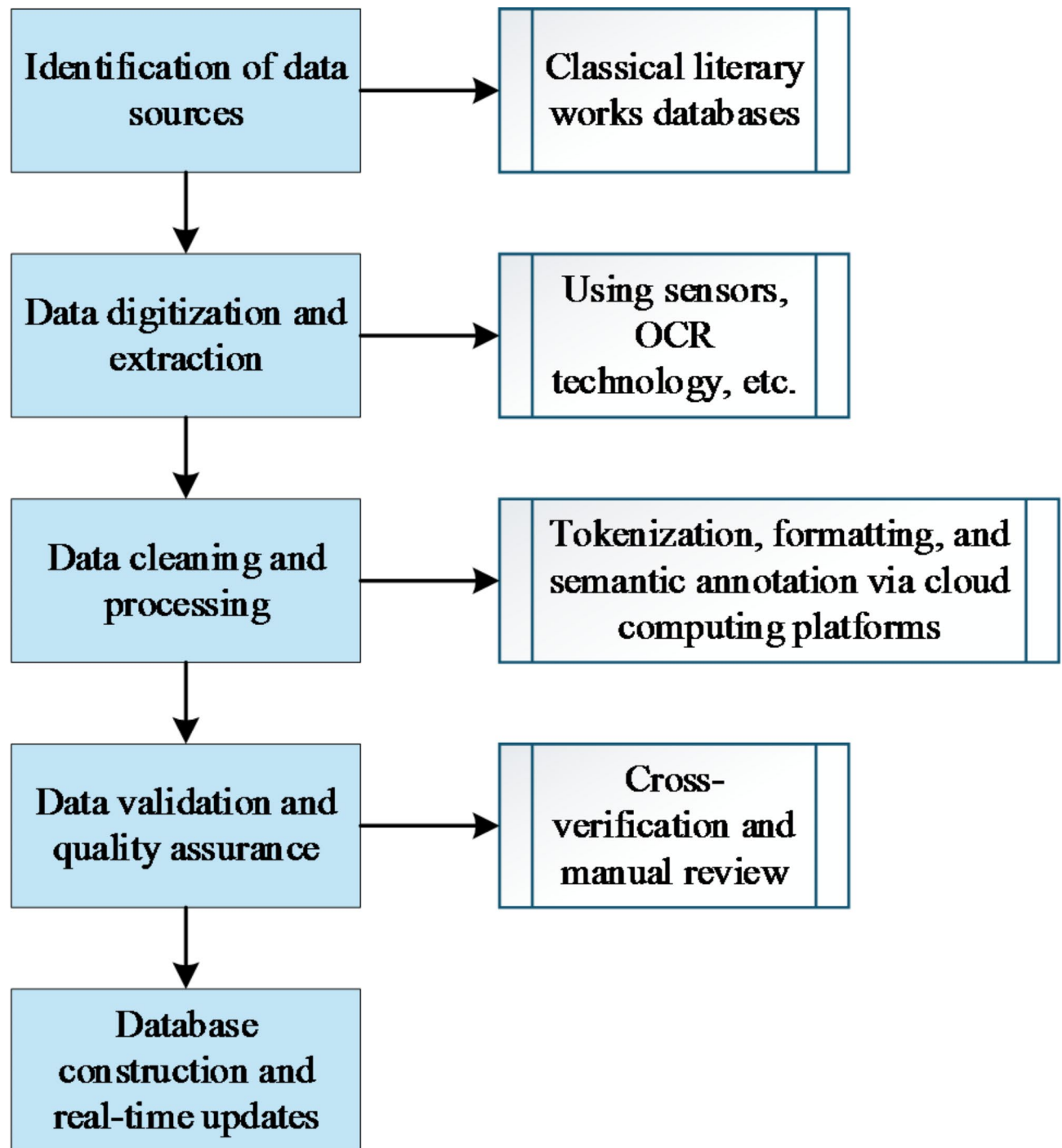


Fig. 1. Workflow for waka data collection using IoT technology.

Automation of data digitization and extraction is achieved using IoT technologies. Intelligent sensors, OCR tools, and advanced scanning devices are employed to convert waka from physical or image-based formats into structured text data. To improve OCR accuracy, multilingual OCR tools are integrated with optimized language models specifically trained to handle classical Japanese expressions found in waka. The digitization process is supported by automated monitoring systems connected via IoT devices, enabling real-time checks for completeness and format consistency. Post-digitization, a combination of manual reviews and automated comparisons ensured an OCR accuracy rate exceeding 98%. Additional validation targeted classical vocabulary and the distinctive textual formats characteristic of waka.

Step 3: Data cleaning and processing.

Once digitized, the waka texts underwent cleaning and processing on cloud computing platforms. Tokenization is conducted using advanced morphological analysis tools designed for classical Japanese, allowing

for precise identification of word boundaries. Text formatting is standardized to ensure consistent punctuation and structure, meeting the input requirements of machine translation models. Rule-based semantic annotation methods are applied to account for waka's unique linguistic features, highlighting elements such as emotional expressions, cultural symbols, and temporal references. Semantic annotations are verified using Inter-Annotator Agreement, achieving a Kappa coefficient exceeding 0.85. This rigorous process ensured reliable annotations, providing robust contextual support for training machine translation models.

Step 4: Data validation and quality assurance.

The validation phase combined cross-validation with manual reviews to enhance reliability. Cross-referencing with other digitized waka databases helped identify duplicates, inconsistencies, and omissions. Linguistic experts conducted sample evaluations of the corpus, focusing on completeness, alignment accuracy, and preservation of cultural context in the waka. These evaluations established an average corpus accuracy of 99% or higher. Errors, such as mistranslations or semantic gaps, are promptly identified and rectified during this phase.

Step 5: Database construction and real-time updates.

After validation, a scalable waka corpus database is constructed to support multimodal input, including text, audio, and image-based waka data. The database organizes information in a structured format, ensuring accessibility and usability. Real-time update capabilities are integrated to detect and synchronize newly added waka documents automatically via IoT technologies. For example, updates to a literary database triggered automated workflows for data extraction and processing, maintaining the corpus's timeliness and scalability. Regular evaluations of query efficiency, data integrity, and system responsiveness ensured the database's stability and effectiveness for large-scale applications.

This study implements an innovative AI-driven approach to waka data collection by integrating IoT technologies. The system automates the collection and processing of waka data from multiple authoritative literary databases. Through real-time monitoring and automated validation, it effectively resolves format inconsistencies and errors in waka texts, significantly enhancing data accuracy and reliability. Moreover, the multi-layered optimization enables by AI technologies streamlined the validation and processing stages, supporting the construction and maintenance of large-scale corpora. This approach ensures data freshness and relevance, establishing a robust foundation for advanced machine translation research on classical Japanese texts. It also provides a comprehensive corpus to train and experiment with machine translation models, enabling these models to adapt and perform effectively when addressing waka from diverse literary styles and contexts.

NMT model optimized by positional encoding

Neural network models are a crucial tool in AI, simulating the connections and information transmission processes between neurons in the human brain. These models are widely used across various AI tasks, including image recognition, speech recognition, natural language processing, and machine translation. In many applications, neural networks are employed to handle natural language tasks. The sinusoidal positional encoding in traditional Transformer-based NMT models computes positional information using predefined functions but does not incorporate contextual information^{18,19}. To improve the performance of the Transformer model, this study proposes two approaches to directly capture the positional encoding information of the source language sequence: horizontal Bi-LSTM and vertical Tree-LSTM.

NMT with integrated Bi-LSTM

In the Long Short-Term Memory (LSTM) network, time steps can retain memory of previous information for extended periods, making it well-suited for handling long sequences of text. The structure of its unit is shown in Fig. 2²⁰.

Figure 2 illustrates the structure of an LSTM unit. An LSTM unit comprises a memory cell and three gate units: the input, forget, and output gates. At time step t , the information updating process of the LSTM can be described by Eqs. (1–6)^{21–23}:

$$a_t = \phi[Q_a(m_{t-1}, X_t) + w_a] \quad (1)$$

$$b_t = \phi[Q_b(m_{t-1}, X_t) + w_b] \quad (2)$$

$$c_t = \phi[Q_c(m_{t-1}, X_t) + w_c] \quad (3)$$

$$U_t = \tanh[Q_i(m_{t-1}, X_t) + w_i] \quad (4)$$

$$U_t = b_t \odot u_{t-1} + a_t \odot U_t \quad (5)$$

$$s_t = n_t \odot \tanh(u_t) \quad (6)$$

Here, ϕ denotes the sigmoid activation function, \odot represents element-wise multiplication, and X_t , m_{t-1} , and u_{t-1} refer to the input, memory, and output values at time step t , respectively. Q_a , Q_b , Q_c , and Q_i represent the weight matrices, and w_a , w_b , w_c , and w_i are the bias terms. The variable a_t represents the input gate, which determines whether the information at time t is stored in the current cell state or filtered out. b_t is the forget gate, responsible for discarding irrelevant information from the previous time step, while c_t is the output gate, controlling the output of the current hidden layer. The memory cell is denoted by u_t , and the hidden state by s_t . By utilizing the memory and forget gates, the LSTM network updates the state of the memory cell at each time step, effectively addressing long-term dependencies in text data. The Bi-LSTM extends the traditional LSTM by stacking a forward and a backward LSTM, enabling the model to capture both past and future contextual information. Bi-LSTM is particularly effective in extracting latent semantic information embedded in text sequences²⁴.

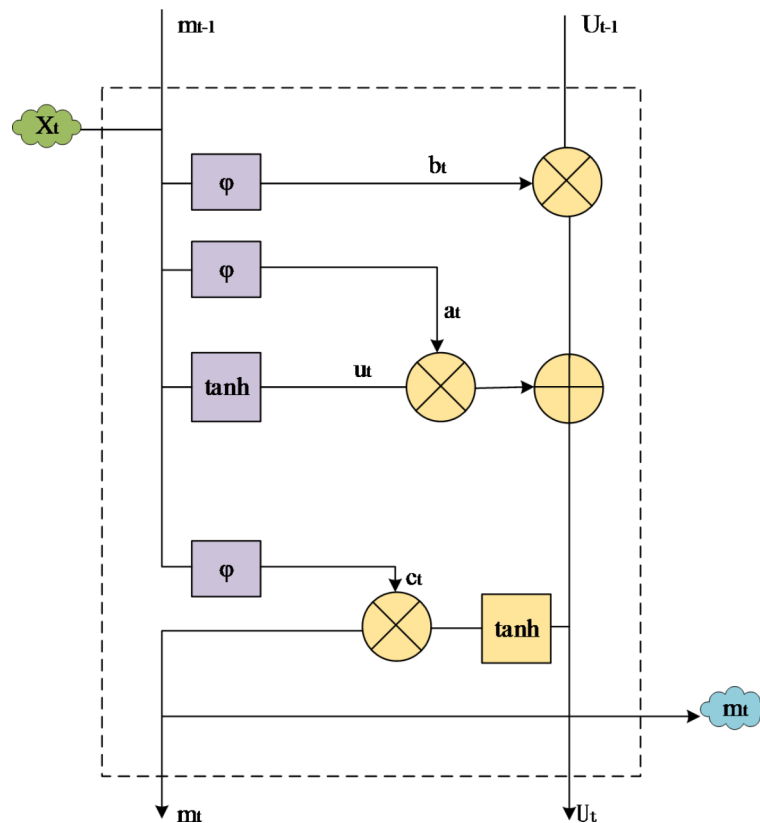


Fig. 2. LSTM Unit Structure.

In the Transformer model, integrating Bi-LSTM helps capture positional information for each word in a sentence from a horizontal perspective. The word embeddings are fed into the Bi-LSTM model for training, leveraging its memory properties. One key advantage of Bi-LSTM is its ability to capture long-range dependencies, addressing the issue of long-distance dependencies in sentences. A single-layer Bi-LSTM consists of two LSTM layers processing information in opposite directions. After training, the output vector h_t integrates the positional information of each word. This output is then used as the input sequence for the translation model, as shown in Eqs. (7)–(9):

$$h_{Lt}, c_{Lt} = LSTM_L(x_t, h_{L(t-1)}, c_{L(t-1)}) \quad (7)$$

$$h_{Rt}, c_{Rt} = LSTM_R(x_t, h_{R(t-1)}, c_{R(t-1)}) \quad (8)$$

$$h_t = [h_{Lt}, h_{Rt}] \quad (9)$$

Here, h_t represents the hidden state of the model at time step t , c_t represents the cell state, and x_t represents the input at time t . The subscripts L and R indicate the left-to-right and right-to-left directions, respectively.

NMT with integrated Tree-LSTM

Similar to the LSTM unit, the Tree-LSTM unit consists of an input gate, an output gate, a memory cell, and a hidden state. However, unlike the LSTM unit, which has a single forget gate, each child node in a Tree-LSTM unit is associated with its own forget gate. This allows the parent node to selectively integrate hidden information from each child node^{25,26}. In an LSTM unit, the hidden state is computed based on the current input and the hidden state from the previous time step. In contrast, the hidden state in a Tree-LSTM unit is calculated using the current input and the hidden states from multiple child Tree-LSTM units^{27–29}. The structures of the LSTM and Tree-LSTM are illustrated in Fig. 3a,b, respectively.

Due to the limitation that the positional encoding in Transformers and the positional encoding obtained through Bi-LSTM do not inherently capture the syntactic structure of sentences, Tree-LSTM is introduced to model and analyze the syntactic structure of source language sentences. To improve the accuracy of positional encoding in representing positional information, this study incorporates Tree-LSTM. Additionally, a self-attention mechanism is employed to calculate the dependencies between words in a sentence. In this mechanism, the input consists of the hidden vectors of all the child nodes of the current parent node, the word vectors of the child nodes, and the word vector of the parent node itself. The word vectors are used to compute the attention weights between the parent node and the child nodes, while the hidden vectors are used to calculate the weighted sum vector h , as shown in Fig. 4.

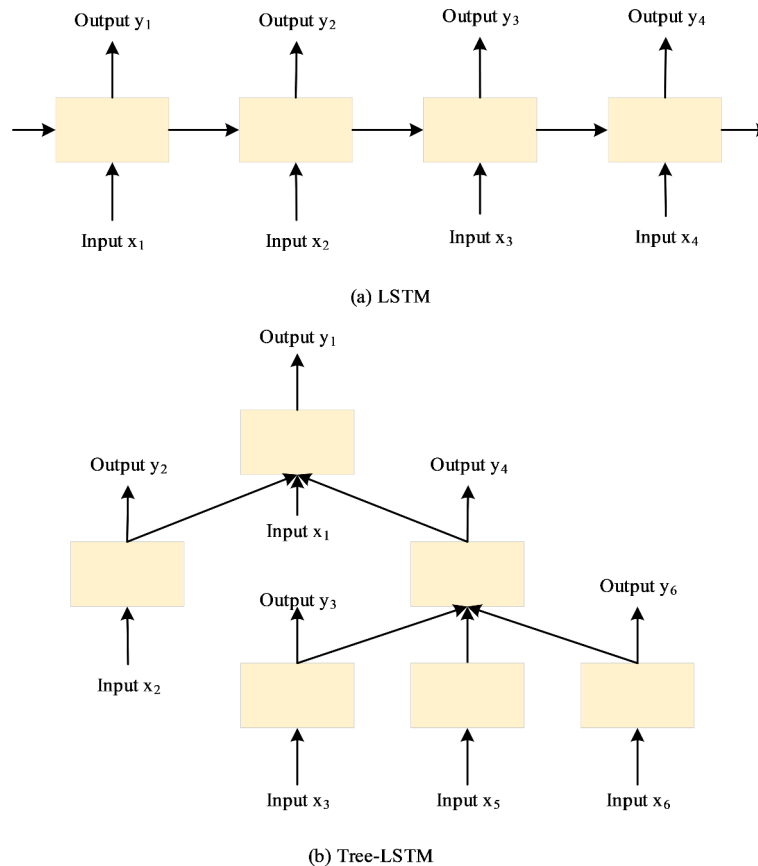


Fig. 3. A Comparison Between LSTM and Tree-LSTM Structures.

Given an input sequence $X = \{x_1, x_2, \dots, x_n\}$, the dependency γ_i between the head node and its child node i is computed via dot product. The calculated γ_i is then multiplied by the corresponding hidden vector h_i of the child node, and the results are summed to obtain the syntactic structure representation of the head node for this subtree. The weighted sum of the hidden vectors of all child nodes in the subtree is then passed through an activation function to yield h , as described in Eqs. (10)–(12).

$$\gamma_i = \frac{\exp(z_i^T z)}{\sum_{k=1}^n \exp(z_k^T z)} \quad (10)$$

$$g = \sum_{i=1}^n \gamma_i h_i \quad (11)$$

$$\tilde{h} = \tanh(W^{(t)}g + b^{(t)}) \quad (12)$$

In this context, γ_i represents the dependency between the head node and its i -th child node; z represents the word vector of the head node; z_i represents the word vector of the i -th child node; h_i denotes the hidden vector of the i -th child node; n represents the total number of child nodes; g signifies the syntactic structure representation of the head node. By assigning different attention weights to the child nodes, the weighted sum h is obtained, providing the hidden vector of the head node. This vector is used as input to the translation model. The source language sequence is generated by traversing each dependency tree from bottom to top, with all node output vectors serving as input sequences for the translation model.

Results and discussion

Experimental design and data sources

The experiment is divided into two parts. In Experiment 1, the performance of the proposed NMT model, which optimizes positional encoding, is validated using public datasets. The datasets used for training include the Workshop on Machine Translation 17 (WMT17) and the China Workshop on Machine Translation (CWMT). The validation set consists of newsdev2017, and the test sets include newstest2017, cwmt2018, and newstest2018. Prior to conducting the experiment, several data preprocessing steps are performed, including data cleaning,

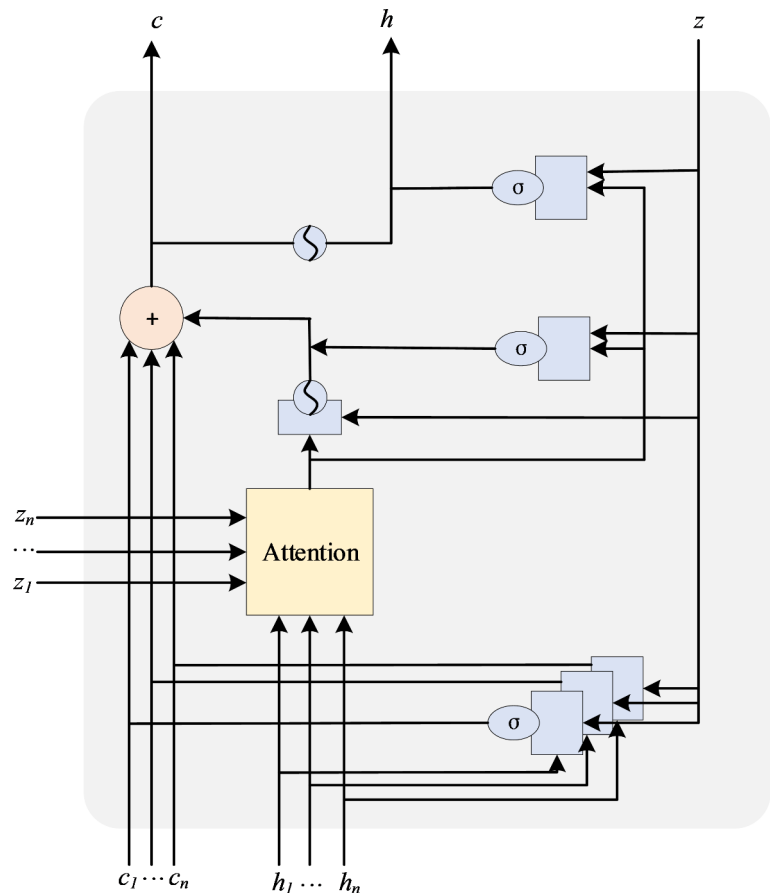


Fig. 4. Schematic Diagram of the Tree-LSTM Structure Based on Attention Mechanism.

Experimental environment	Configuration	Parameter name	Value
Operating System	Windows10	Word Vector Dimension	512
CPU	Intel Core i7	Hidden Layer Dimension	512
GPU	NVIDIA GeForce RTX 2060	Number of Attention Heads	8
RAM	32GB	Feedforward Network Dimension	2048
Disk Space	1T	Learning Rate	0.5
Programming Language	Python3.6	Maximum Sequence Length	50
Deep Learning Framework	Tensorflow	Batch Size	32

Table 1. Experimental environment and parameter settings.

tokenization, and symbol handling. The experimental environment and parameter settings are presented in Table 1.

In Experiment 2, the performance of the proposed model in translating Japanese Waka poetry was evaluated using datasets collected via the IoT. To assess the model’s adaptability and limitations in processing waka translations, a total of 5000 waka samples were collected from three classical Japanese literary works: The Tale of Genji, Man’yōshū, and The Tales of Ise. These samples encompass various types and styles of waka. They were divided into training, testing, and validation sets in an 8:1:1 ratio. The Japanese waka test data used in Experiment 2 are presented in Table 2.

The following two examples highlight the model’s strengths and weaknesses in handling waka translations. The waka “秋の田の、かりほの庵の、苦をあらみ、わが衣手は、露にぬれつつ” (Hyakunin Isshu) demonstrates the model’s proficiency in translating natural imagery and direct emotional expression. This waka, which describes a temporary straw hut in the autumn fields and the hardships of life, is effectively translated as “秋天的田野上, 临时搭起的茅草屋, 屋顶茅草稀疏, 我的衣袖, 常常被露水打湿,” preserving the emotional atmosphere and basic meaning. Although some nuances, such as rhyme structure, are lost, the model maintains a high level of fidelity to the content and imagery. In contrast, the waka “しらつゆに、風の吹きしく、秋の野は、つらぬきとめぬ、玉ぞ散りける” (Man’yōshū) reveals the model’s difficulty with deep metaphors and cultural context. In

No.	Original Waka	Translation	Notes
1	秋の田の、かりほの庵の、苫をあらみ、わが衣手は、露にぬれつつ	In the autumn fields, a temporary thatched hut, with sparse thatch on the roof, my sleeves are often wet with dew	Direct emotional expression and natural imagery, successfully translated by the model
2	しらつゆに、風の吹きさしく、秋の野は、つらぬきとめぬ、玉ぞ散りける	The autumn wind blows fiercely, white dew covers the fields, and the jewel-like dew drops scatter without being caught	Contains deep metaphors and cultural context, the model struggles to fully convey the philosophical meaning and emotions of impermanence
3	わが恋は、春の夜空の、星の数ほど、絶え間なく光りて、永遠に輝かむ	My love, like the countless stars in the spring night sky, shines endlessly and will shine forever	The model handles direct metaphors and emotional expression well, with an accurate translation
4	さくら花、ひとひらの花の、散りゆくまで、忍ぶ心を、知りてこそ思え	Cherry blossoms, falling one petal at a time, until the last one falls, only then do I understand the heart of endurance	The subtle emotions and cultural context present challenges, but the model conveys the basic meaning
5	露の世は、夢の世にも似て、はかないことの、さだめにぞあるなりける	The world of dew is like a dream, transient and destined to be fleeting	The translation retains the poem's atmosphere, though the depth of the metaphor is somewhat diminished
6	我が身一つ、人に知られず、山の端にて、泣く泪はほととぎすの声	Alone, unnoticed by others, at the mountain's edge, my tears flow with the sound of the cuckoo's call	The model successfully retains delicate expressions and natural sounds in the translation
7	よしの山、秋の山をば、見わたせば、いにしえよりの、紅葉すぐしも	In Yoshino Mountain, gazing over the autumn landscape, the red leaves have been this vivid since ancient times	Direct imagery and temporal expressions are effectively translated, though the emotional nuance is limited
8	ながらへば、忍ぶることの、弱りては、見し人のほほえむ顔	If I can endure for long, the faces of those I have seen will appear even more precious	The model conveys the emotional content and meaning, but lacks depth in cultural context
9	君がため、春の野に出でて、野草摘む、われを思うと、花に声をかけよ	For you, I venture into the spring fields to pick wildflowers, and when I think of you, I call to the flowers	The model captures the poetic essence, but some meanings lose their cultural connection.
10	ただひとつ、心の中に、住むべき、花ぞ散りゆく、つまさえ見えず	Only one heart resides within me, though the flowers fall, we can no longer see each other	Symbolic and philosophical expressions are difficult for the model to fully convey

Table 2. Sample waka test data.

this poem, “しらつゆ” (white dew) metaphorically refers to the transience and fragility of life. While the model translates it as “The autumn wind blows fiercely, white dew covers the fields, and the jewel-like dew drops scatter without being caught,” the profound philosophical and existential themes conveyed in the original text are less apparent in the translation. Due to the specific imagery and cultural significance of waka, the model struggles to fully capture the original poem's reflection on impermanence and the fleeting nature of life.

Both experiments employ the BLEU score as the evaluation metric to assess the quality of machine translation outcomes. The BLEU score measures translation accuracy by comparing machine-generated translations to human reference translations, focusing on the overlap between the two. A higher BLEU score indicates that the machine translation results more closely align with the reference translations, thus reflecting superior translation quality³⁰.

Adjusting the number of transformer encoder layers

In the Transformer model, an increase in the number of encoder layers enables the model to capture more semantic information. To examine the impact of the number of encoder layers on translation performance, experiments were conducted using the WMT17 and CWMT datasets. The results are presented in Fig. 5.

In Fig. 5, when the number of encoder layers is fewer than 6, the model's translation performance improves as the number of layers increases. However, once the number of encoder layers reaches 6 or more, the BLEU score begins to stabilize. Based on these experimental results, the optimal number of encoder layers for the model is set to 6.

Model performance validation in experiment 1

To validate the performance of the proposed model, the experiment compared three models—Transformer + Bi-LSTM, Transformer + Tree-LSTM, and Transformer + Tree-LSTM + Attention—with traditional models such as Transformer, Convolutional Sequence to Sequence (ConvS2S), and Google Neural Machine Translation (GNMT) using public datasets. The results are illustrated in Fig. 6.

As shown in Fig. 6, the proposed translation models outperform other models in translation tasks on public datasets. The BLEU score of the Transformer + Bi-LSTM model, which integrates Bi-LSTM into the Transformer, averages 23.71. The Transformer + Tree-LSTM model, which integrates Tree-LSTM, achieves a BLEU score of 23.95. By further enhancing the Transformer + Tree-LSTM model with a self-attention mechanism to calculate the contribution of each child node to the parent node, the BLEU score increases to 24.12, marking an improvement of 0.71% over the Transformer + Tree-LSTM model. To further demonstrate the effectiveness of the proposed model, the BLEU score changes of the three proposed models are compared with those of the baseline models, as shown in Fig. 7.

In Fig. 7, the three proposed models exhibit strong BLEU scores across different datasets. The Transformer + Tree-LSTM + Attention model stands out, achieving an average BLEU score increase of 1.13, which is significantly higher than the other two models. This suggests that the model, which combines Tree-LSTM with a self-attention mechanism, is more effective at capturing sentence structure and context, leading to improved translation quality. The improvement in translation performance can be attributed to the integration of Bi-LSTM and Tree-LSTM structures, which enable the model to better capture long-range dependencies and manage complex sentence structures. This results in more fluent and natural translations. Specifically, in the Transformer + Tree-LSTM + Attention model, the inclusion of a self-attention mechanism allows the model to calculate the contribution of various sub-nodes to the parent node. This further enhances contextual coherence and ensures the accurate transmission of information. Overall, the combination of Bi-LSTM, Tree-LSTM, and

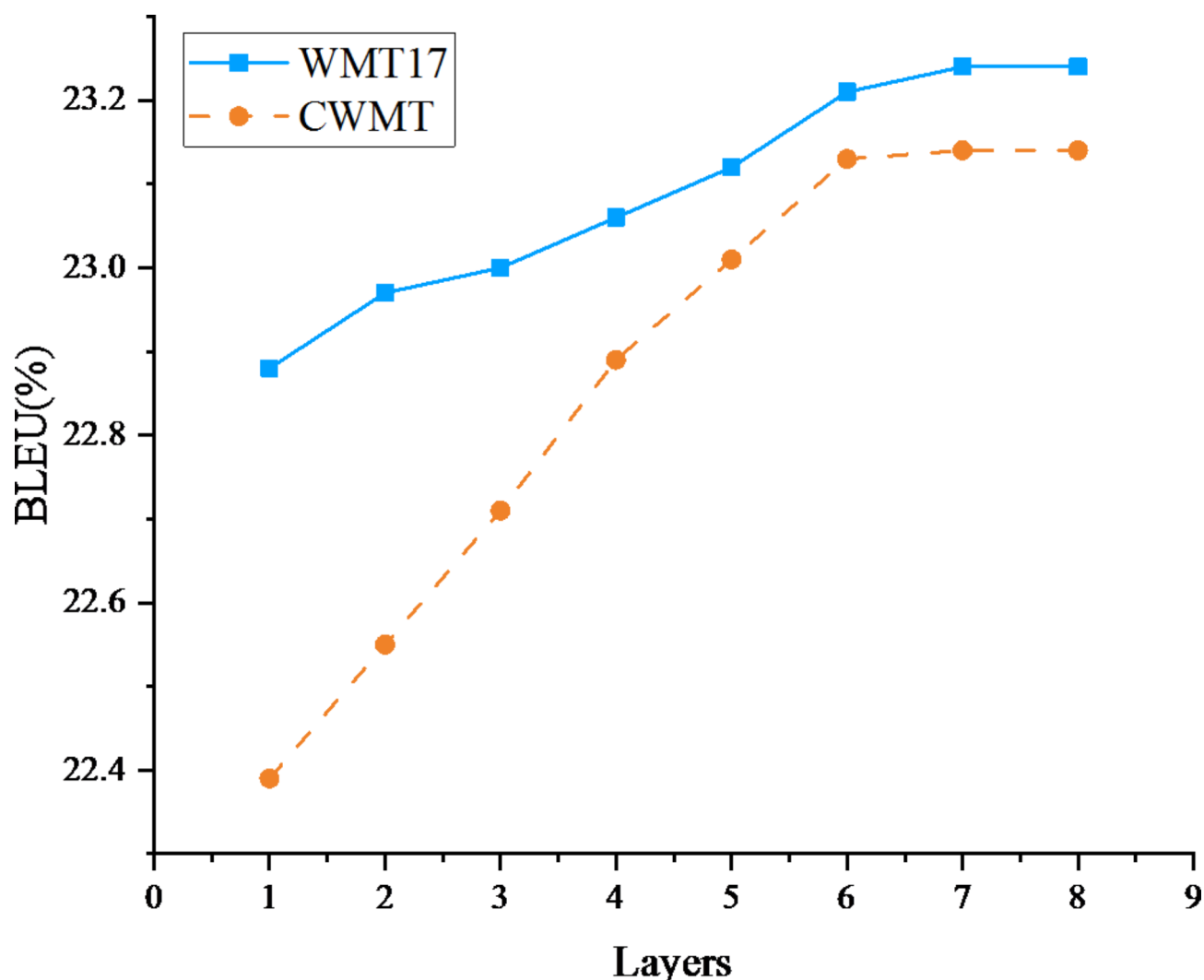


Fig. 5. Validation results of encoder layer numbers.

the self-attention mechanism has significantly enhanced NMT performance, confirming the proposed model's effectiveness.

Model performance validation in experiment 2

Finally, the translation performance of the models was validated using the collected Japanese Waka dataset. The BLEU scores and hyperparameter tuning times of the three proposed models, compared to the baseline model, are shown in Fig. 8.

As shown in Fig. 8, all three improved models outperform the baseline Transformer model in terms of BLEU scores on the Japanese Waka dataset. Specifically, the Transformer + Tree-LSTM + Attention model achieves the highest BLEU score of 20.65, demonstrating its significant advantage in capturing the characteristics and contextual information of Waka. Although the hyperparameter tuning time increased slightly, the gains in BLEU scores make this time investment reasonable, showcasing a high cost-performance ratio. In conclusion, the Transformer + Tree-LSTM + Attention model more effectively captures the linguistic characteristics and poetic structure of Classical Japanese when translating Waka, leading to translations with improved grammatical accuracy and contextual consistency. Compared to the standard Transformer model, the enhanced model more accurately conveys the cultural nuances and emotional undertones of Waka, successfully mitigating common issues such as vocabulary omissions and semantic ambiguity. While hyperparameter tuning requires additional time, the improvements in BLEU scores and translation quality demonstrate the cost-effectiveness of this approach, underscoring its practical advantages in Waka translation.

Conclusion

To improve the translation of Japanese Waka, this study first explores the form of Waka and the current state of its Chinese translation, analyzing existing translation controversies and introducing a method for Waka data collection based on IoT. Building on this foundation, the NMT model was optimized by integrating Bi-LSTM,

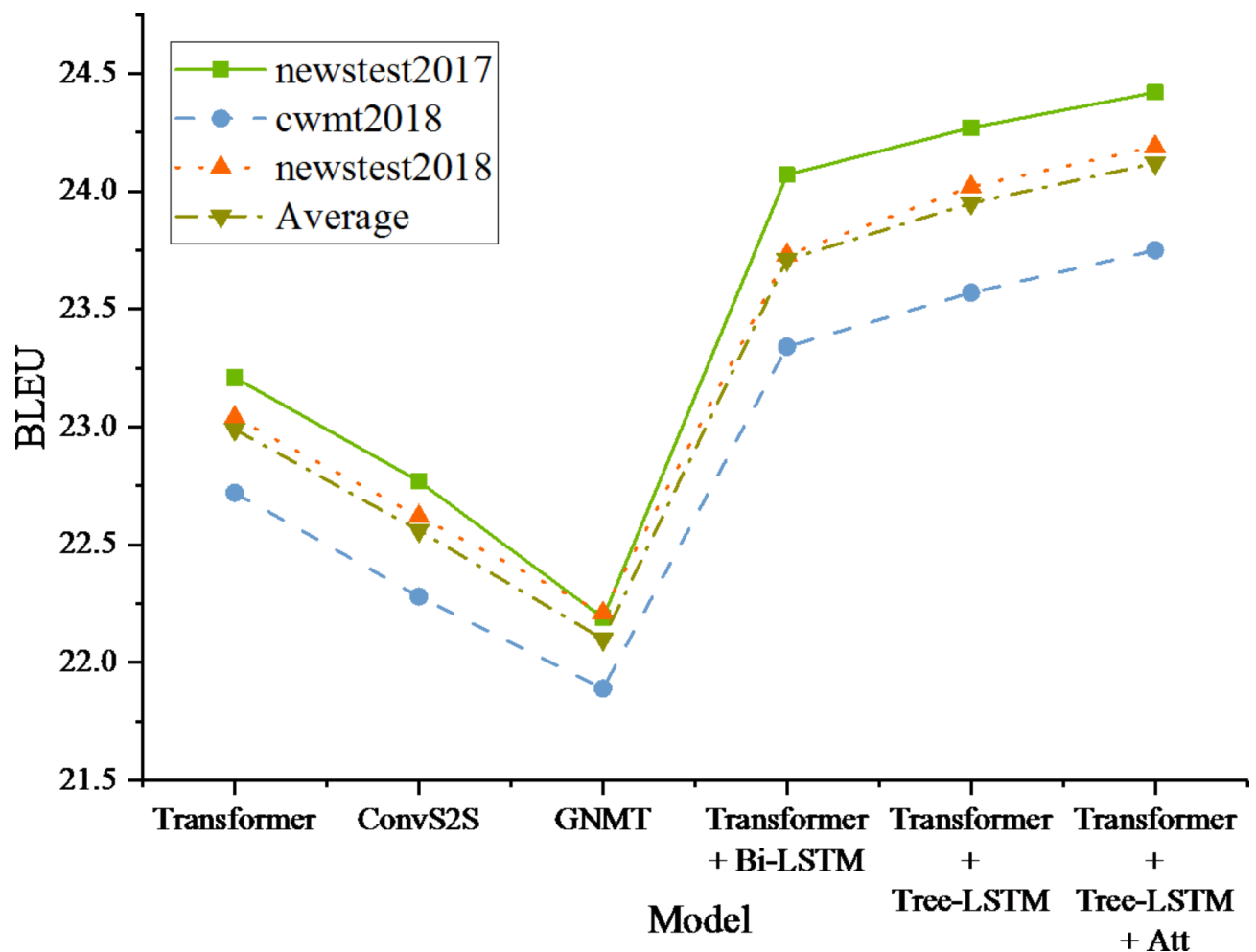


Fig. 6. Performance comparison of different models.

Tree-LSTM, and attention mechanisms, resulting in a new translation model. The performance of these three optimized models was validated through experiments, leading to the following conclusions:

1. On public datasets, the proposed Transformer + Bi-LSTM, Transformer + Tree-LSTM, and Transformer + Tree-LSTM + Attention models outperform other models, with average BLEU scores of 23.71, 23.95, and 24.12, respectively, all surpassing the baseline Transformer model.
2. Compared to the baseline Transformer, the Transformer + Tree-LSTM + Attention model achieves an average BLEU score increase of 1.13, significantly higher than the other two models. This demonstrates that the Transformer + Tree-LSTM + Attention model, by incorporating Tree-LSTM and attention mechanisms, is better able to capture sentence structure and contextual information, thereby significantly improving translation quality.
3. All three improved models outperform the baseline Transformer in terms of BLEU scores on the Japanese Waka dataset. Specifically, the Transformer + Tree-LSTM + Attention model achieves the highest BLEU score of 20.65, indicating its notable advantage in capturing the features and contextual information of Waka, despite the increase in hyperparameter tuning time.

However, this study has several limitations. Firstly, due to the complex cultural connotations and linguistic characteristics of Japanese Waka, while the model demonstrates good performance in terms of BLEU scores, it may still fall short in fully capturing the subtle emotional and cultural nuances of Waka, which could lead to translations that lack naturalness and aesthetic quality. Secondly, the dataset is relatively small. While some Waka data were collected via IoT, the dataset lacks the richness found in other language pairs, which may affect the model's generalization ability. Moreover, the process of hyperparameter tuning is relatively time-consuming, posing challenges for real-time translation systems. Future research can address these issues by further expanding the Waka dataset and incorporating more bilingual parallel corpora to improve the model's training effectiveness and robustness. Additionally, more complex model architectures, such as using pre-trained models combined with transfer learning, can be explored to better capture the semantic and emotional information in Waka. Lastly, the diversity of Waka translations should be studied, with efforts to integrate cultural factors into

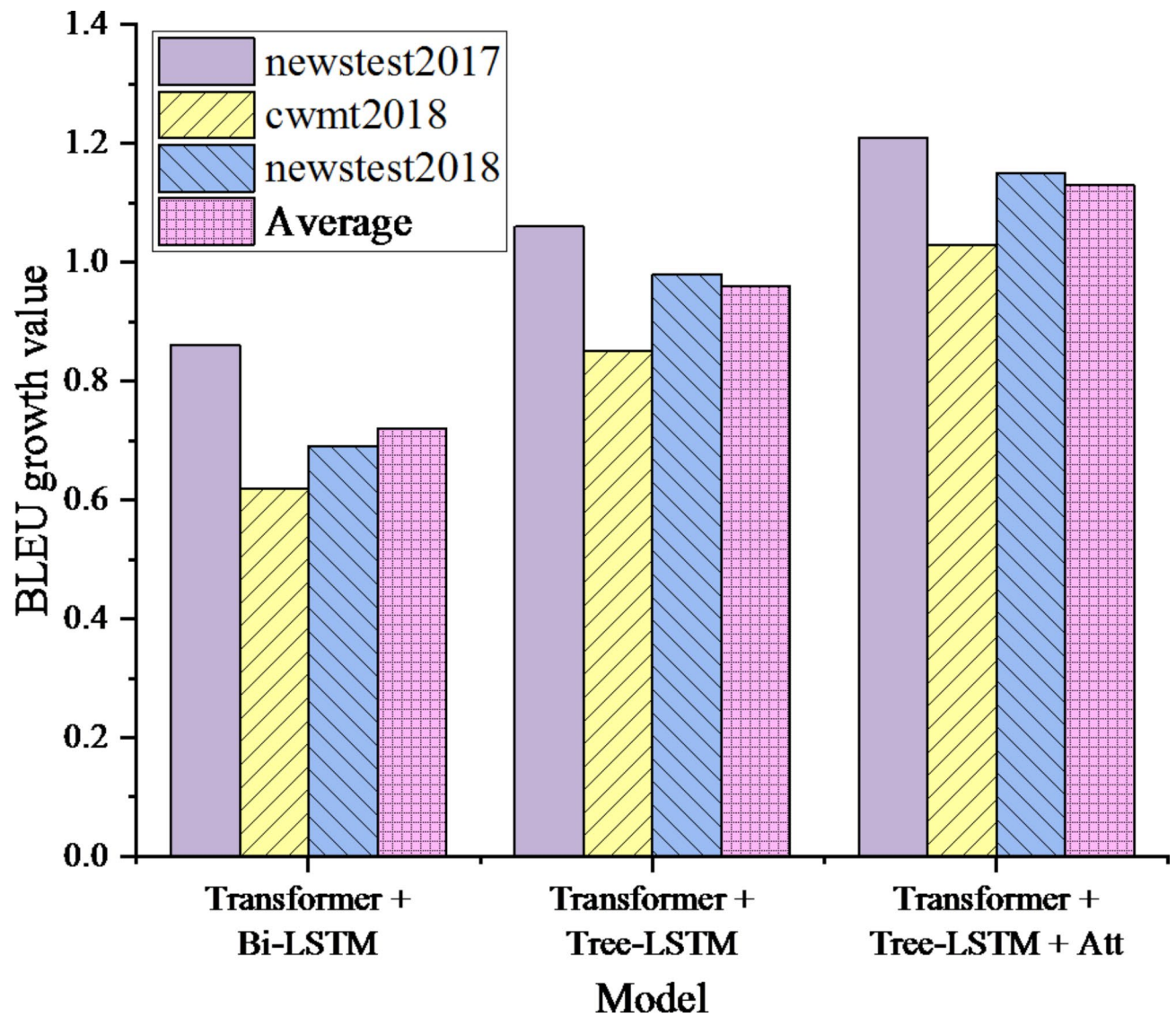


Fig. 7. BLEU score comparison across different datasets.

the design of the translation model to enhance the cultural adaptability and aesthetic quality of the translations, thereby providing higher-quality translation services for Sino-Japanese cultural exchanges.

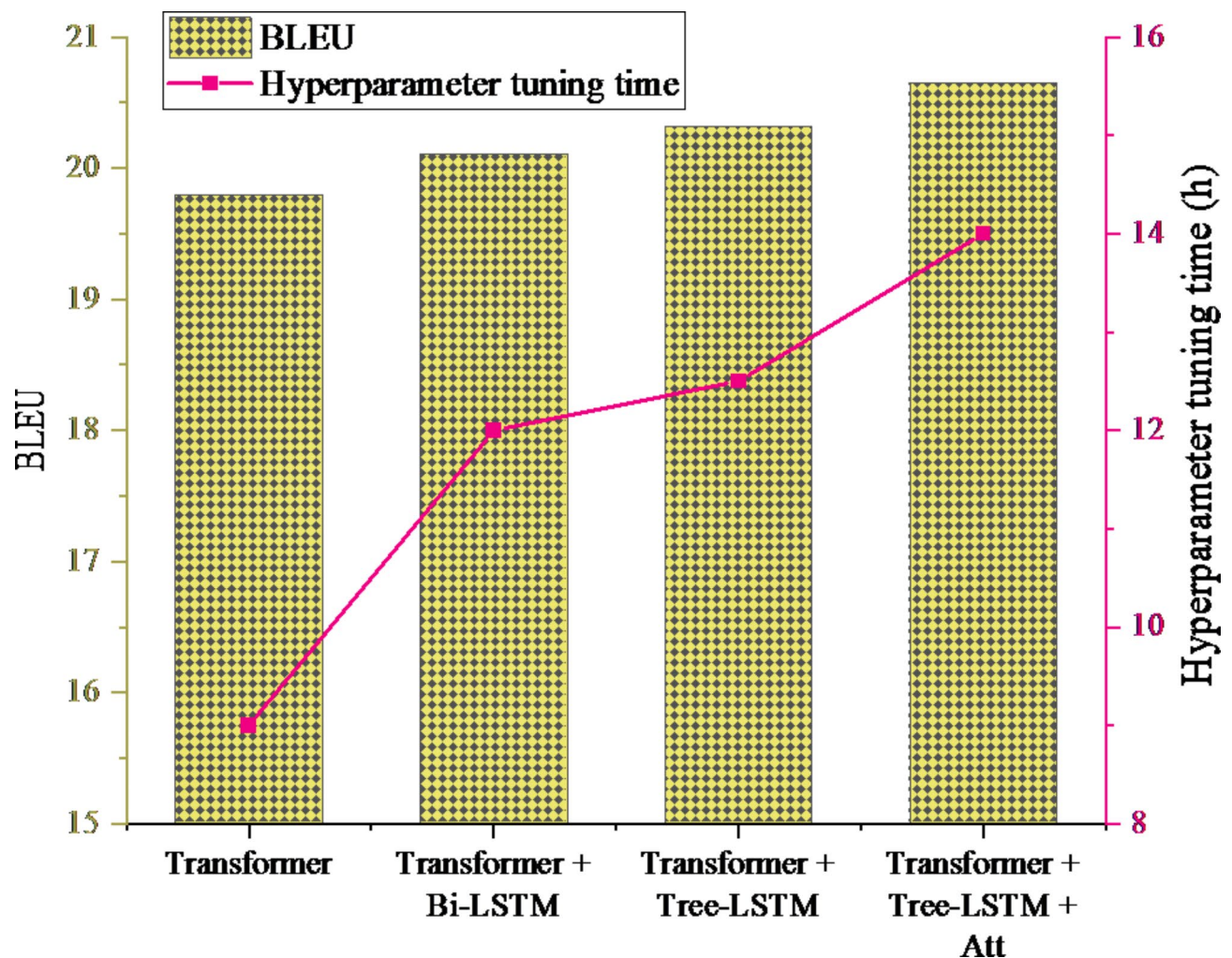


Fig. 8. Translation performance validation on the Japanese Waka Dataset.

Data availability

The human participants/human dataset were not directly involved in the manuscript. The datasets used and/or analyzed during the current study are available from the corresponding author Rizhong Shen on reasonable request via e-mail shenrizhong@qztc.edu.cn.

Received: 22 June 2024; Accepted: 1 January 2025

Published online: 06 January 2025

References

- Fujito, S. et al. Japanese bunching onion line with a high resistance to the stone leek leafminer, *Liriomyza chinensis* from the 'Beicong' population: Evaluating the inheritance of resistance. *Euphytica* **217** (2), 1–8 (2021).
- Xing, L. Reliability in internet of things: Current status and future perspectives. *IEEE Internet Things J.* **7**(8), 6704–6721 (2020).
- Kaul, V., Enslin, S. & Gross, S. A. History of artificial intelligence in medicine. *Gastrointest. Endosc.* **92**(4), 807–812 (2020).
- Liu, X. et al. A scenario-generic neural machine translation data augmentation method. *Electronics* **12**(10), 2320 (2023).
- Klimova, B. et al. Neural machine translation in foreign language teaching and learning: A systematic review. *Educ. Inform. Technol.* **28**(1), 663–682 (2023).
- Tonja, A. L. et al. Low-resource neural machine translation improvement using source-side monolingual data. *Appl. Sci.* **13**(2), 1201 (2023).
- Wang, F. & K Washbourne, J. Technical and scientific terms in poetry translation: The tensions of an 'anti-poetical' textual feature. *J. Spec. Transl.* **38**, 55–74 (2022).
- Bose, H. Influence of Alfred Thayer Mahan on Japanese maritime strategy. *J. Def. Stud.* **14**(12), 49–68 (2020).
- Pan, S. Y. et al. Tea and tea drinking: China's outstanding contributions to the mankind. *Chin. Med.* **17**(1), 1–40 (2022).
- Song, W. China's normative Foreign policy and its multilateral engagement in asia. *Pac. Focus.* **35**(2), 229–249 (2020).
- Mikhailovna, D. E. Commentaries and commentary modes in Japanese Literary tradition based on the examples of the classical poetry anthologies. *Russian Japanol. Rev.* **4**(2), 70–93 (2021).
- Haqzazarova, S. Translation and Literary Influence on the work of Abdulla Sher. *Tex. J. Multidiscip. Stud.* **5**, 92–93 (2022).
- Fani, A. The allure of untranslatability: Shafi'i-Kadkani and (not) translating persian poetry. *Iran. Stud.* **54**(1–2), 95–125 (2021).
- Al-Awawdeh, N. Translation between creativity and reproducing an equivalent original text. *Psychol. Educ. J.* **58**(1), 2559–2564 (2021).

15. Mouha, R. A. R. A. Internet of things (IoT). *J. Data Anal. Inform. Process.* **9**(02), 77 (2021).
16. Zhou, I. et al. Internet of things 2.0: Concepts, applications, and future directions. *IEEE Access.* **9**, 70961–71012 (2021).
17. Abosaif, A. N. & Hamza, H. S. Quality of service-aware service selection algorithms for the internet of things environment: A review paper. *Array* **8**, 100041 (2020).
18. Mohamed, S. A. et al. Neural machine translation: Past, present, and future. *Neural Comput. Appl.* **33**, 15919–15931 (2021).
19. Xu, W. & Carpuat, M. EDITOR: an edit-based transformer with repositioning for neural machine translation with soft lexical constraints. *Trans. Assoc. Comput. Linguist.* **9**, 311–328 (2021).
20. Lindemann, B. et al. A survey on long short-term memory networks for time series prediction. *Proc. Cirp* **99**, 650–655 (2021).
21. Wang, C. et al. A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells. *Int. J. Hydrog. Energy* **47**(18), 10395–10408 (2022).
22. Gauch, M. et al. Rainfall-runoff prediction at multiple timescales with a single long short-term memory network. *Hydrol. Earth Syst. Sci.* **25**(4), 2045–2062 (2021).
23. Farrag, T. A. & Elattar, E. E. Optimized deep stacked long short-term memory network for long-term load forecasting. *IEEE Access.* **9**, 68511–68522 (2021).
24. Wu, P. et al. Sentiment classification using attention mechanism and bidirectional long short-term memory network. *Appl. Soft Comput.* **112**, 107792 (2021).
25. Hastuti, R. P., Suyanto, Y. & Sari, A. K. Q-learning for shift-reduce parsing in Indonesian Tree-LSTM-Based text generation. *Trans. Asian Low-Resource Lang. Inform. Process.* **21**(4), 1–15 (2022).
26. Wang, L., Cao, H. & Yuan, L. Gated tree-structured RecurNN for detecting biomedical event trigger. *Appl. Soft Comput.* **126**, 109251 (2022).
27. Wang, L., Cao, H. & Yuan, L. Child-Sum (N2E2N) Tree-LSTMs: an interactive child-Sum Tree-LSTMs to extract biomedical event. *Syst. Soft Comput.* **6**, 200075 (2024).
28. Wu, E. et al. Treeago: Tree-structure aggregation and optimization for graph neural network. *Neurocomputing* **489**, 429–440 (2022).
29. Chen, G. Timed failure propagation graph construction with supremal language guided Tree-LSTM and its application to interpretable fault diagnosis. *Appl. Intell.* **52**(11), 12990–13005 (2022).
30. Lateef, H. M. et al. Evaluation of domain sulfur industry for DIA translator using bilingual evaluation understudy method. *Bull. Electr. Eng. Inf.* **13**(1), 370–376 (2024).

Author contributions

Rizhong Shen: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, project administration, funding acquisition.

Declarations

Competing interests

The authors declare no competing interests.

Ethics statement

The studies involving human participants were reviewed and approved by School of Foreign Languages, Quanzhou Normal University Ethics Committee (Approval Number: 2022.0506967). The participants provided their written informed consent to participate in this study. All methods were performed in accordance with relevant guidelines and regulations.

Additional information

Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025