Deep Learning-Based Automatic Cultural Translation Method for English Tourism

Jianguo Liu[1]*, Ruohan Liu[2]

Foreign Languages College, Henan University of Science and Technology, Luoyang 471000, Henan, China[1]

Marxist Academy, Henan University of Science and Technology, Luoyang 471000, Henan, China[1]

School of Electronic Engineering, Xi'an Post and Communications University, Xi'an 710100, Shaanxi, China[2]

***Abstract*—The general LSTM-based encoder-decoder model has the problems of not being able to mine the sentence semantics and translate long text sequences. This study presents a neural machine translation model utilizing LSTM with improved attention, incorporating multi-head attention and multi-skipping attention mechanisms into the LSTM baseline model. By adding multi-head attention computation, the syntactic information in different subspaces can be mined, and then attention can be paid to the semantic information in the sentence sequences, and then multiple attentions are computed on each head separately, which can effectively deal with the long-distance dependency problem and perform better in the translation of long sentences. The proposed model is analysed and compared using the WMT17 Chinese and English datasets, newsdev2017 and newstest2017, and the results show that the proposed model improves the BLEU score of the automatic translation of Tourism English Culture and solves the problem of low scores in long sentence translation.**

Keywords—LSTM-based encoder-decoder model; tourism English culture; automatic translation; enhanced attention mechanism

I. INTRODUCTION

The process of economic globalisation has further brought about the globalisation of language, and English has become the only protagonist of linguistic globalisation, and English language learning has gradually become a matter of concern [1]. English, as a part of the cultural composition of tourism, has become a necessary skill for people travelling abroad across borders. In order to understand tourism English more conveniently, natural language processing technology based on artificial intelligence algorithms has entered people's life [2]. Natural Language Processing (NLP) is to transform the language humans usually communicate and the text they see into what machines can understand [3]. Natural language processing technology has a wide range of applications, including machine translation [4], sentiment classification [5], robot dialogue [6], text classification [7], etc. With the popularity of deep learning, deep neural networks began to be introduced into NLP tasks and made great progress, while machine translation based on deep neural networks received great attention, and researchers embedded deep neural networks into machine translation tasks, which led to the improvement of the quality of automatic English translation [8]. Deep neural network-based machine translation can effectively promote the future economic and social development, thus enhancing people's satisfaction and sense of access. Therefore, the study of machine translation of tourism English based on deep neural networks is a meaningful research direction, which is of great significance for globalised economic exchange and cultural output [9].

The primary objective of the application of deep learning technology in the automatic translation of tourism English culture is to acquire a comprehensive understanding of the structure and laws of language through neural network models, thereby enhancing the quality and efficiency of translation [10]. Currently, the automatic translation methods of tourism English culture based on deep learning include Seq2Seq model [11], Attention mechanism [12], Transformer model [13], Pre-training language model [14], Multi-modal data translation [15], Zero Resource Translation [16], Online learning and incremental learning of Neural Machine Translation [17] [18] and so on. Although Neural Machine Translation has made greater progress and is better than Statistical Machine Translation on some public datasets, Neural Machine Translation is still far from the effect of human translation, and there are still the following challenges and problems [10]: 1) data sparsity problem; 2) model optimisation problem; 3) large-scale vocabularies and rare words problem.

This text proposes a method for autonomous cultural translation of Tourist English, addressing the issues of attention computation and model optimization in encoder-decoder architectures utilizing recurrent neural networks, specifically through the implementation of an LSTM-enhanced attention mechanism. This paper's primary contributions are: 1) the introduction of an automatic language translation model and a neural machine translation framework; 2) the investigation and design of a neural machine translation model utilizing an LSTM-enhanced attention mechanism; and 3) a comparative analysis of the proposed model employing the Tourism English dataset.

The structure of this paper is organized as follows: Section I discusses foundational techniques, covering automatic translation systems and neural machine translation frameworks. Section II outlines the key challenges in translating tourism English, such as data sparsity and issues with long-sequence translations. Section III introduces the enhanced attention mechanism based on LSTM, emphasizing multi-head and multi-hop attention for improved performance. Section IV describes the experimental framework, including datasets, evaluation methods, and model comparisons using BLEU scores. Section V presents the findings and analysis, demonstrating the model's advantages. Last Section VI concludes with insights and suggestions for future work.

*Corresponding Author

II. Relevant Theoretical Techniques

A. Automatic Language Translation Models

The Automatic Language Translation Model [19] is a tool that uses artificial intelligence techniques to automatically convert text from one language to another by means of a computer programme, as shown in Fig. 1.

Automatic language translation models are usually based on deep learning techniques, especially neural networks, such as Recurrent Neural Networks (RNN), Long Short-Term Memory Networks (LSTM) and Transformer models, as shown in Fig. 2. They are capable of handling complex linguistic structures and expressions and have achieved good translation quality in several public reviews.

To mitigate the issues associated with one-hot encoding, including context independence and excessive dimensionality, the Neural Probabilistic Language Model (NPLM) [20] derives word vectors by learning word distributions, as illustrated in Fig. 3.

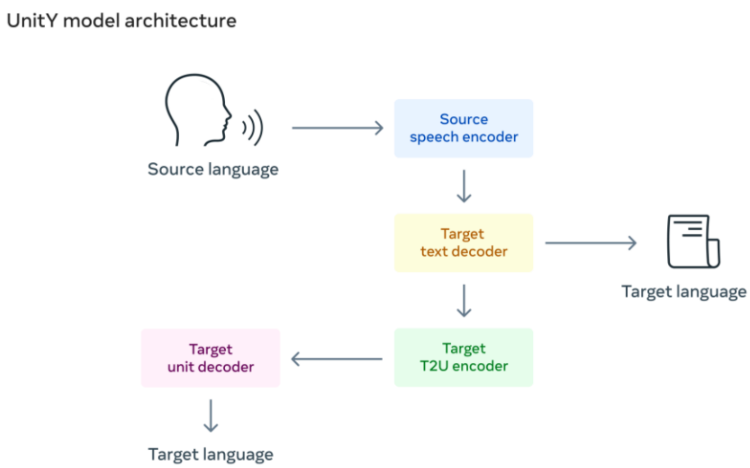

Fig. 1. Model of automatic language translation.

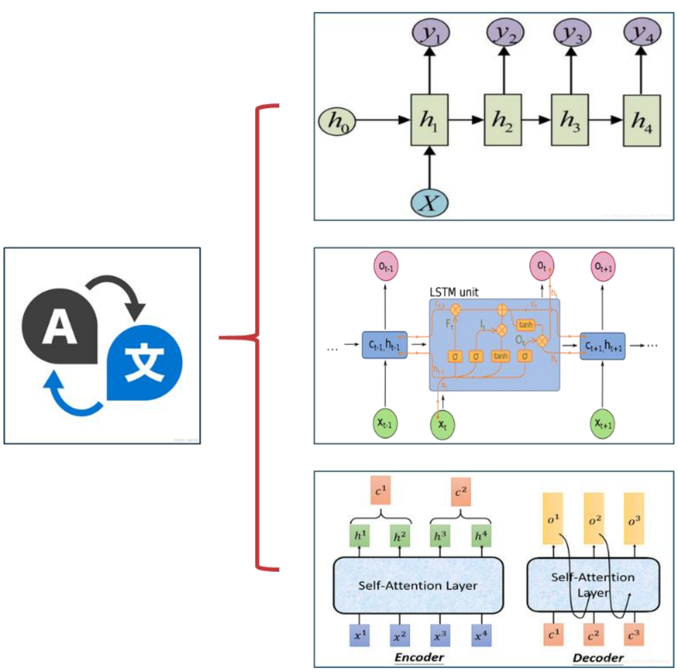

Fig. 2. Classification of automatic language translation models.

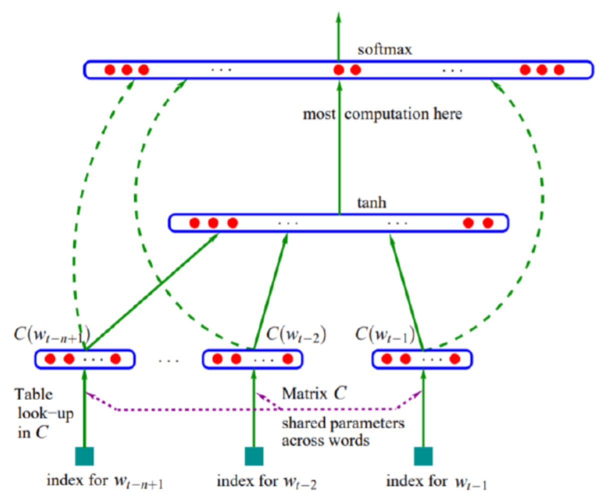

Fig. 3. NPLM structure.

In the NPLM structure, given a text sequence $(w_1, w_2, \cdots, w_t, \cdots, w_T)$, where w_t is the word in the word list V. The objective function is to build the optimal model f, calculated as follows:

$$f\left(w_t,\cdots,w_{t-n+1}\right)=\hat{P}\left(w_t \mid w_1,\cdots,w_{t-1}\right) \tag{1}$$

$$\sum_{i=1}^{|V|} f\left(i;w_{t-1},\cdots,w_{t-n+1}\right)=1 \tag{2}$$

The structure is divided into two parts, the first part is to build a mapping from any word w_i to a vector in the word list $V \;\; C(i)$, the second part has a feed forward neural network g to fit $f\left(i;w_{t-1},\cdots,w_{t-n+1}\right)$ where $f\left(i;w_{t-1},\cdots,w_{t-n+1}\right)=g\left(i,C\left(w_{t-1}\right),\cdots,C\left(w_{t-n+1}\right)\right)$, the training objective is to maximise the following equation:

$$L=\frac{1}{T}\sum_t \log f\left(w_i,w_{t-1},w_{t-N+1}\right)+R(\theta) \tag{3}$$

where R is the regular term and θ is the parameter of the feedforward neural network g .

B. *Neural Machine Translation Framework and Classification*

1) Text feature representation: Based on the representation of word vectors, the textual feature representation is subsequently obtained, i.e. the Embedding operation [21], the specific structure of which is shown in Fig. 4. The Embedding layer is often used in the first layer of the neural machine translation model, and its role is to map the input sequences into dense vectors of lower dimensions, which are able to characterise the word information effectively.

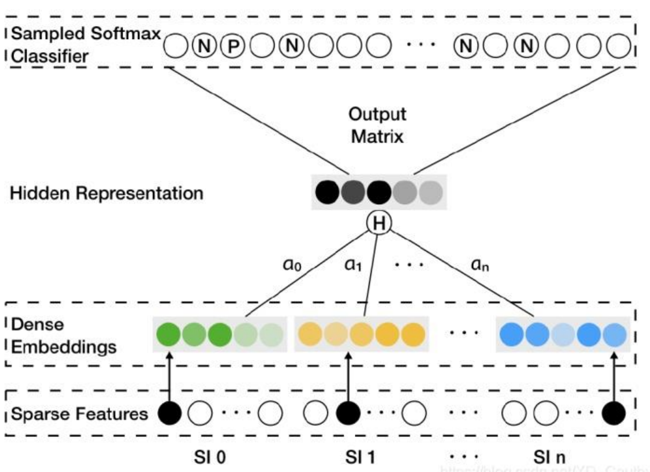

Fig. 4. Embedding operation.

2) Encoder-decoder structure: Many neural machine translation models are constructed using the Encoder-Decoder framework [22], which is also referred to as the sequence-to-sequence architecture. The fundamental concept of neural machine translation is exemplified by this framework, which involves the conversion of a source text sequence into a mathematical problem. The mathematical problem is then solved to produce a target text sequence. Refer to Fig. 5 for the specific structure.

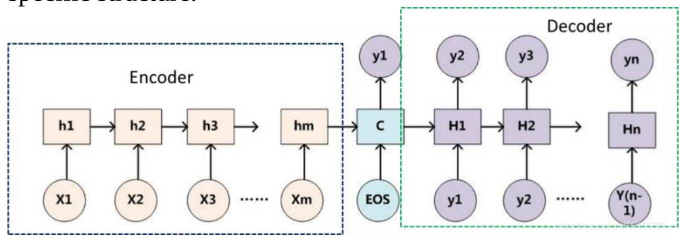

Fig. 5. Encoder-Decoder structure.

Fig. 5 illustrates that the Encoder-Decoder architecture comprises two components: the Encoder, which processes a word from the source sentence at each time step, extracting the informational properties of the source sequence. Subsequent to several time steps, all words will be condensed into the encoder's hidden states, resulting in a context vector C. The decoder receives inputs comprising the context vector C, the prior hidden states, and the previously anticipated output. The outcome of each phase serves as input for the subsequent step.

In this procedure, the decoder functions as a language model; however, this model is conditional, constructed based on the context vector C, therefore referred to as a "Conditional Language Model." The expression is stated as follows:

$$p(y)=\prod_{t=1}^{n} p\left(y_t \mid y_1,y_2,\cdots,y_{n-1},C\right) \tag{4}$$

where the output target sequence is $y=\left(y_1,y_2,\cdots,y_n\right)$.

III. LSTM Enhanced Attention Mechanism and Applications

A. *LSTM Enhanced Attention Mechanism*

This study proposes an upgraded neural machine translation model utilizing an LSTM-based attention mechanism, which is an improvement of the RNN encoder-decoder architecture. The model comprises three components:

1) Encoder: This component employs a bidirectional LSTM neural network (Bi-LSTM) [23], which combines the sentence information of the source sequence with the future textual information. The word embedding vectors are input into the Bi-LSTM to encode the complete source sentence, which is subsequently processed by the augmented attention module;

2) Enhanced attention module: this component receives all the encoder's concealed states after the source sentence sequence is encoded at the encoder side and calculates them in conjunction with the decoder's current state to generate the dynamic context vector for the current moment. This section encompasses both Multi-Head Attention [24] and Multi-Hop Attention [25] methods.

- The multi-attention mechanism is mainly designed to fully mine the sentence information in different subspaces in the model, and the specific structure is shown in Fig. 6.

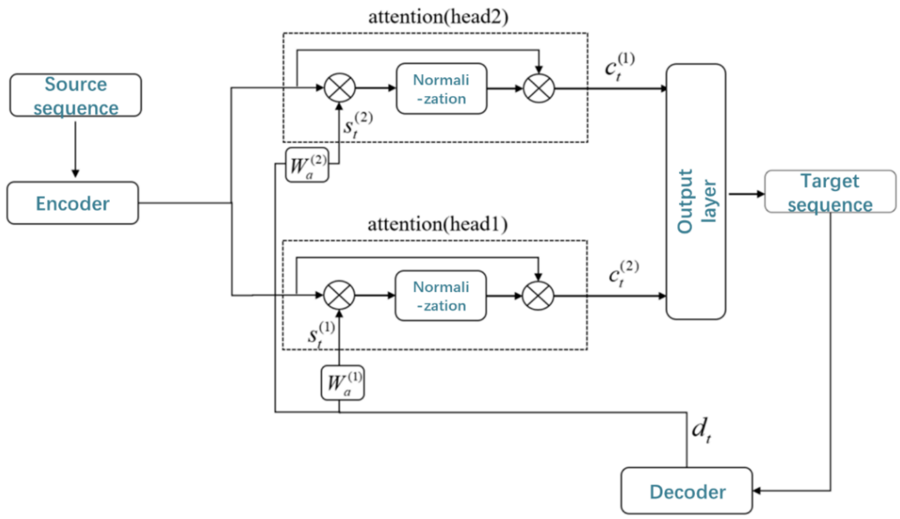

Fig. 6. Multi-attention mechanism structure.

- The multi-hop attentional mechanism mainly performs multiple attentional computations on each HEAD to extend the model's representational capability. The specific structure is shown in Fig. 7.

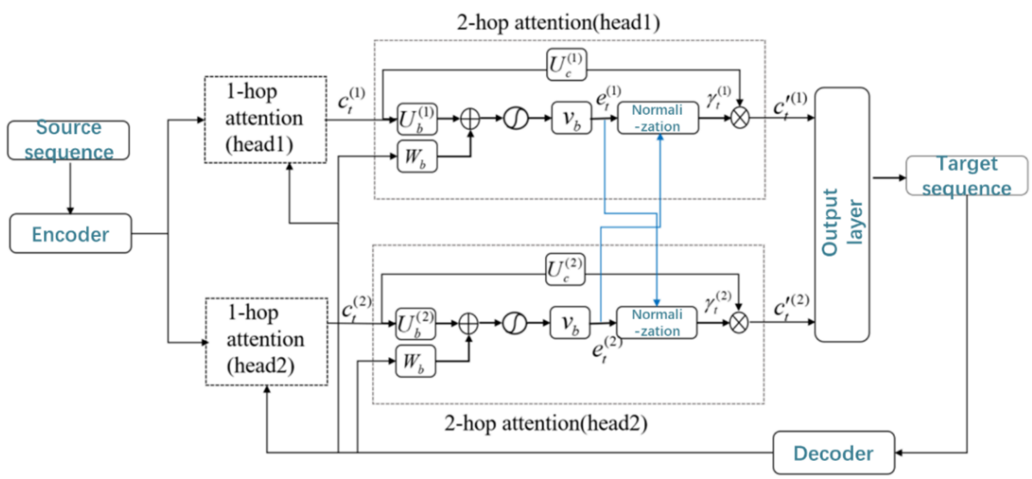

Fig. 7. Structure of multi-hop attention mechanism.

3) Decoder: This part is using LSTM network and receives the hidden state information from the encoder.

The overall structure of the model is shown in Fig. 8.

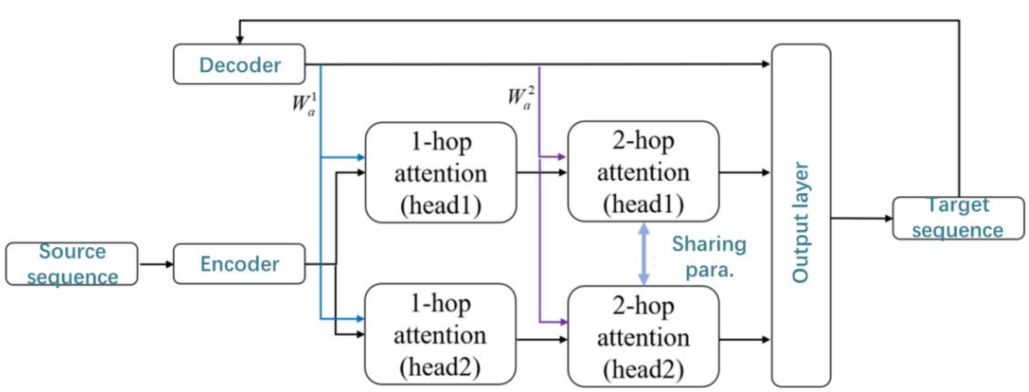

Fig. 8. General structure of the model.

B. Modelling Steps

The operational procedures of the LSTM-based neural machine translation model utilizing an increased attention mechanism are as follows:

1) Input the source sentence sequence $X=\left(x_1,x_2,\cdots,x_T\right)$ into the encoder with Bi-LSTM, the forward LSTM $\vec{f}_{LSTM}$ reads the sentence sequence sequentially from x_1 to x_T and computes the forward hidden state $\left(\vec{h}_1,\vec{h}_2,\cdots,\vec{h}_T\right)$; the backward LSTM $\overleftarrow{f}_{LSTM}$ reads the sentence sequence sequentially from x_T to x_1 and computes the backward hidden state $\left(\overleftarrow{h}_1,\overleftarrow{h}_2,\cdots,\overleftarrow{h}_T\right)$, and splices the forward hidden state and the backward hidden state to get the final hidden state h_t , which is computed by the following formula:

$$h_t=\left[\vec{h}_t;\overleftarrow{h}_t\right]=\left[LSTM_{encoder}\left(x,\vec{h}_{t-1}\right);LSTM_{encoder}\left(x,\overleftarrow{h}_{t+1}\right)\right] \quad (5)$$

2) Deliver the hidden state h_t to the Enhanced Attention Mechanism module;

3) Calculate the 1-hop attention score. Based on the hidden state of the target sequence with respect to the previous time node, the output of the LSTM at the current moment is obtained d_t :

$$d_t=LSTM_{encoder}\left(\hat{y}_{t-1},d_{t-1}\right) \quad (6)$$

4) Based on the trained matrix $W_a^{(k)}$, get the hidden state $s_t^{(k)}$ at the current moment:

$$s_t^{(k)}=W_a^{(k)}d_t \quad (7)$$

where k represents the k th head.

5) Calculate the context vector $c_t^{(k)}$ for the k th head:

$$c_t^{(k)}=soft\max\left(s_t^{(k)}H_{encoder}^T\right)H_{encoder} \quad (8)$$

6) Calculate the attention fraction of 2-hop:

$$e_t^{(k)}=v_b^T\tanh\left(U_b^{(k)}c_t^{(k)}+W_b s_t^{(k)}\right) \quad (9)$$

7) Normalise the attention score for each HEAD to $\gamma_t^{(k)}$:

$$\gamma_t^{(k)}=\frac{\exp\left(e_t^{(k)}\right)}{\sum_{n=1}^{N}\exp\left(e_t^{(n)}\right)} \quad (10)$$

where N represents the total number of heads.

8) The trained parameters $U_t^{(k)}$, $\gamma_t^{(k)}$ and $c_t^{(k)}$ are used to compute the context vector $c_t^{n(k)}$ at the current moment with the following formula:

$$c_t^{n(k)}=\gamma_t^{(k)}U_c^{(k)}c_t^{(k)} \quad (11)$$

9) The context vector $c_t'^{(k)}$ is spliced with the output of LSTM d_t , and the text feature vector is obtained by training parameters with Tanh activation function layer:

$$o_t=\tanh\left(W_o\left[d_t;c_t'^{(1)};c_t'^{(2)};\cdots;c_t'^{(k)}\right]\right) \quad (12)$$

10) The decoder inputs the final text vector o_t to the output layer to get the model prediction result, which is calculated as follows:

$$p\left(y_t \mid y_1,y_2,\cdots,y_{t-1},X\right)=soft\max\left(o_t\right) \quad (13)$$

Fig. 9 illustrates the computational process of the neural machine translation model, which is based on the LSTM enhanced attention mechanism.

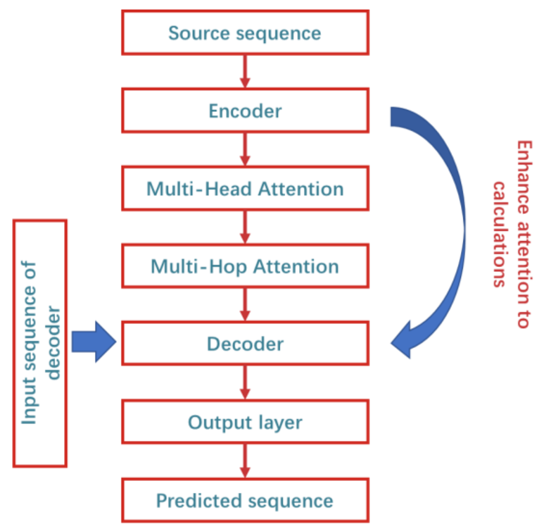

Fig. 9. Computational process of neural machine automatic translation model.

IV. Model Experiments and Analyses

A. Data Sets

The training dataset used for the experiments in this section is the WMT17 Chinese-English (WMT17zh-en) dataset [26], which is used to train the neural machine translation model based on the LSTM enhanced attention mechanism proposed in this chapter, and newsdev2017 and newstest2017 are used as the validation and test sets [27], respectively, and the introduction about them is shown in Table I.

TABLE I. Data Information

Data type	Name (of a Thing)	Magnitude
Training set	WMT17zh-en	227k
Validation set	newsdev2017	4k
Test set	newstest2017	2k

The model is generally set to a fixed text length, so the pad operation is used to supplement sentences of shorter length, as shown in Fig. 10.

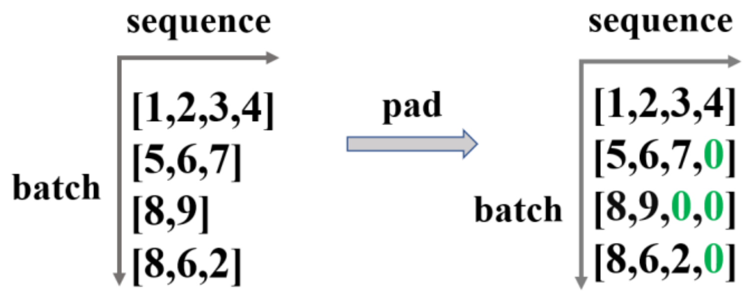

Fig. 10. Pad operation.

B. Indicators for Assessing Translation Effectiveness

In this paper, we adopt an automatic machine translation evaluation method, i.e. BLEU (Bilingual evaluation understudy) [28]. BLEU is the calculation of a similarity score between a given translation generated by a machine translation system, and a reference translation, which is used to measure the performance of this machine translation system, where the range of this score is [0,1]. The specific calculation formula is as follows:

$$BLEU = BP \cdot \exp\left(\sum_{n=1}^{N} w_n \log P_n\right) \tag{14}$$

$$BP = \begin{cases} 1 & c > r \\ e^{(1-r/c)} & c \le r \end{cases} \tag{15}$$

where w_n is the weight for different n-grams, P_n is the weight of the corresponding n-element word in the sequence of reference answers, c is the length of the candidate sentence, and r is the number of words in common between the model-translated sentence and the reference answer sentence.

C. Environmental Settings

This experiment is a deep neural network based translation model, the required experimental environment is shown in Table II, the deep learning framework used for the model experiments in this paper is Pytorch 1.8.1, which contains a large number of libraries internally for the convenience of the researcher. Bi-LSTM [29] and Conv S2S model [30] are used to compare with the proposed model, and the specific model parameter settings are shown in Table III.

V. Analysis of Results

Table IV presents the BLEU scores for the Bi-LSTM, Conv S2S model, and the neural machine translation model utilizing an LSTM-enhanced attention mechanism across three datasets: WMT17zh-en, newsdev2017, and newstest2017. Table IV illustrates that the BLEU scores of the neural machine translation model utilizing an LSTM-enhanced attention mechanism are the highest across the three datasets: WMT17zh-en, newsdev2017, and newstest2017, with scores of 22.86, 23.64, and 23.14, respectively.

TABLE II. Configuration of the Experimental Environment

No.	Causality	Attribute value
1	CPU	Intel Intel(R) Xeon(R)
2	GPUs	Ge Force RTX 2080Ti
3	memory	24G
4	programming language	Python 3.8.3
5	Deep learning frameworks	Pytorch 1.8.1
6	operating system	Linux

TABLE III. Parameter Settings for Experimental Comparison Models

No.	Modelling	Parameterisation
1	Bi-LSTM	Stacked 4-layer LSTM with 1000 hidden layers each, the dimension of word embedding is 1000; the number of neurons in the attention mechanism layer is 1000; the optimiser is SGD, the initial learning rate is 0.005, the batch size is 128 and the number of iterations is 50
2	Conv S2S	The dimensions of the hidden units of the encoder and decoder are 512; the optimiser is SGD; the dropout probability is set to 0.2, the initial learning rate is 0.005, the batch size is 128 and the number of iterations is 50
3	Proposed Method	2-layer Bi-LSTM with word embeddings of dimension 1024, optimiser SGD; initial learning rate 0.005, batch size 128, number of iterations 50

In order to observe the differences between the three models more intuitively, the data in Table IV were visualised as bar charts, as shown in Fig. 11. As can be seen from Fig. 11, on the WMT17zh-en data, the neural machine translation model based on the LSTM enhanced attention mechanism has a higher value of 1.44 BLEU and 0.56 BLEU than the Bi-LSTM and Conv S2S models, respectively; on the newsdev2017 data, the model proposed in this paper has a higher value of 1.44 BLEU and 0.56 BLEU than the Bi-LSTM and Conv S2S models 0.75 BLEU and 0.4 BLEU values, respectively; on newstest2017 data, the model proposed in this paper outperforms Bi-LSTM and Conv S2S models by 1.28 BLEU and 0.36 BLEU values, respectively.

TABLE IV. EXPERIMENTAL COMPARISON MODEL OF BLEU SCORES

No.	Modelling	WMT17zh-en	newsdev2017	newstest2017
1	Bi-LSTM	21.42	23.89	21.86
2	Conv S2S	22.28	23.24	22.78
3	Proposed Method	22.86	23.64	23.14

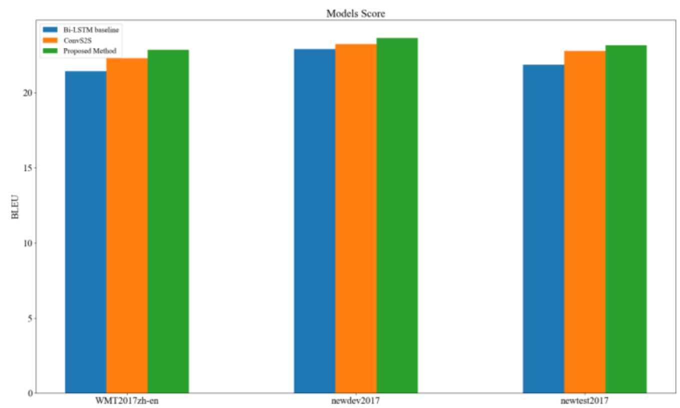

Fig. 11. BLEU scores for three models.

In order to analyse the BLEU scores of the three models under different tourism English sentence lengths, this paper investigates 11 length ranges of tourism English long sentences, and the specific results are shown in Fig. 12. Fig. 12 illustrates that despite the neural machine translation model utilizing the LSTM-enhanced attention mechanism exhibiting a lower BLEU score compared to the Conv S2S model within the sentence length range of [20,25], it surpasses the BLEU score of the Conv S2S model for sentence lengths around 80 and exceeding 90.

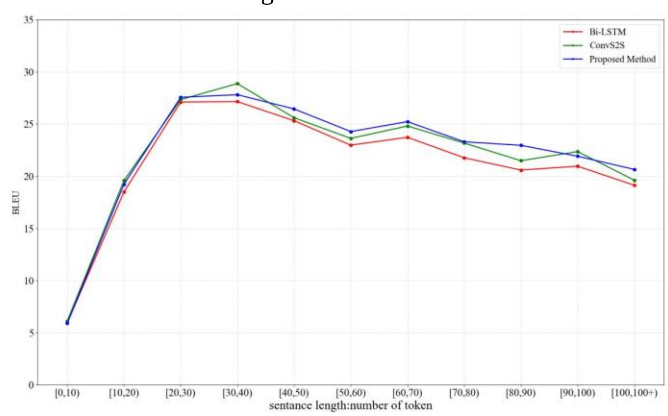

Fig. 12. BLEU scores of the three models for different sentence lengths.

VI. CONCLUSION AND OUTLOOK

This work addresses the issue of automatic translation within the context of English culture, highlighting the shortcomings of the LSTM-based encoder-decoder model, and presents a neural machine translation model that incorporates an increased attention mechanism based on LSTM. By examining pertinent theoretical methodologies and delineating the challenges of English automatic translation, a neural machine translation model utilizing an LSTM-enhanced attention mechanism is developed through the implementation of multi-hop attention computation and multi-head attention procedures. The experimental part uses WMT17 Chinese and English datasets, newsdev2017 and newstest2017, and introduces the criteria of machine translation evaluation to measure the effect of translation quality through BLEU. The experimental comparative analysis demonstrates the effectiveness of the proposed enhanced attention mechanism, especially for translating long sentences.

Despite its effectiveness, the proposed LSTM-based enhanced attention model for tourism English translation has several limitations. First, the reliance on the WMT17 dataset may limit the model's applicability to other domains or language pairs, as the dataset might not cover diverse linguistic features or cultural nuances. Second, while the multi-head and multi-hop attention mechanisms improve long-sequence translation, the model's complexity increases significantly, leading to higher computational costs and longer training times. Third, the translation quality heavily depends on the availability of high-quality, domain-specific training data, which remains a challenge in many low-resource contexts. Lastly, the model's performance was evaluated solely with BLEU scores, which might not fully capture the subtleties of cultural translation, such as idiomatic expressions or contextual accuracy. These limitations suggest the need for further improvements in model generalizability, efficiency, and evaluation methods.

To overcome the limitations identified, future research could focus on the following areas: Future research should focus on expanding datasets to include diverse linguistic and cultural contexts, beyond just tourism English. This can involve collecting data from multiple language pairs and incorporating low-resource languages to improve the model's adaptability. A richer dataset will ensure that the translation system captures various idiomatic expressions and cultural nuances, making the model more universally applicable and effective in handling diverse real-world scenarios.

To address the computational burden, future work could explore optimizing the model's architecture to be more lightweight without sacrificing performance. Techniques such as sparse attention mechanisms or pruning can reduce resource usage and training times. This would make the model more scalable and suitable for deployment in environments with limited computing resources, such as mobile devices or embedded systems.

Incorporating more comprehensive evaluation metrics is essential to fully capture translation quality. Beyond BLEU scores, qualitative metrics should be used to assess how well the model handles idiomatic expressions and cultural subtleties. Adding human evaluation to the testing process can provide valuable insights into the model's contextual accuracy and overall fluency, ensuring more reliable and culturally sensitive translations.

References

[1] An G , Tan D A L .Enhancing cross-cultural communication for Chinese tourists in non-native English-speaking destinations: a study of sociolinguistic competence and politeness challenges[J].Global Chinese, 2024, 10(2):215-240.DOI:10.1515/glochi-2023-0041.

[2] Laskar S R , Manna R , Pakray P B S .A Domain Specific Parallel Corpus and Enhanced English-Assamese Neural Machine Translation[J].computacion y sistemas, 2022, 26(4):1669-1687.

[3] Goldberg Y .A Primer on Neural Network Models for Natural Language Processing[J].Computer Science, 2016.DOI:10.1613/jair.4992.

[4] Takahashi K , Sudoh K , Nakamura S .Automatic Machine Translation Evaluation using a Source and Reference Sentence with a Cross-lingual Language Model[J].Journal of Natural Language Processing, 2022, 29(1):3-22.DOI:10.5715/jnlp.29.3.

[5] Gupta S , Bouadjenek M R , Robles-Kelly A .PERCY: A post-hoc explanation-based score for logic rule dissemination consistency assessment in sentiment classification[J].Knowledge-Based Systems, 2023.DOI:10.1016/j.knosys.2023.110685.

[6] Ishiguro H .A Preliminary Study on Realising Human-Robot Mental Comforting Dialogue via Sharing Experience Emotionally[J].Sensors , 2022, 22.DOI:10.3390/s22030991.

[7] Zhong T .A generic multi-level framework for building term-weighting schemes in text classification[J].The Computer Journal, 2024.DOI:10.1093/ comjnl/bxae068.

[8] Kramov A , Pogorilyy S .Usage of the Speech Disfluency Detection Method for the Machine Translation of the Transcriptions of Spoken Language[J]. NaUKMA Research Papers. computer science, 2023.DOI:10.18523/2617-3808.2022.5.54-61.

[9] Bilianos D , Mikros G .Sentiment analysis in cross-linguistic context: how can machine translation influence sentiment classification?[J]. Literary & linguistic computing: Journal of the Alliance of Digital Humanities Organizations, 2023.

[10] Zhou X , Jia W , Shi C .Automatic Translation of English Terms for Computer Network Security Based on Deep Learning[J]. 2024, 20(3s):598-609.

[11] YANG Donghua, ZOU Development, WANG Hongzhi, WANG Jinbao. SparQL query prediction based on Seq2Seq model[J]. Journal of Software, 2021.DOI:10.13328/j.cnki.jos.006171.

[12] Gemechu E , Kanagachidambaresan G R .English-Afaan Oromo Machine Translation Using Deep Attention Neural Network[J].Optical memory & neural networks, 2023(3):32.DOI:10.3103/S1060992X23030049.

[13] Liu W , He Y , Lan T W Z .Research on system combination of machine translation based on Transformer[J].high technology letters, 2023, 29(3):310-317.

[14] Thin D V , Hao D N , Nguyen L T .Vietnamese Sentiment Analysis: An Overview and Comparative Study of Fine-tuning Pretrained Language Models[J].ACM transactions on Asian and low-resource language information processing, 2023(6):22.

[15] Ma F .Construction and Evaluation of College English Translation Teaching Model Based on Multimodal Integration[J].Applied Mathematics and Nonlinear Sciences, 2024, 9(1).DOI:10.2478/amns-2024-1774.

[16] Huang P , Zhao J , Sun S L Y .Knowledge enhanced zero-resource machine translation using image-pivoting[J].Applied Intelligence: the International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2023, 53(7):7484-7496.

[17] Uzma F , Shafry M R M , Adnan A .A multi-stack RNN-based neural machine translation model for English to Pakistan sign language translation[J].Neural computing & applications, 2023.DOI:10.1007/s00521-023-08424-0.

[18] Madi S , Baba-Ali A R .A new hybrid incremental learning system for an enhanced KNN algorithm (hoKNN)[J].Evolving Systems, 2024, 15(3):1001-1019. DOI:10.1007/s12530-023-09531-y.

[19] Edmundson H P , Oettinger A G .Automatic Language Translation[J].Mathematics of Computation, 2011, 15(74).DOI:10.2307/2004259.

[20] Maharjan J , Garikipati A , Singh N P , Cyrus, L. , Sharma M, Ciobanu M. OpenMedLM: prompt engineering can out-form fine-tuning in medical question- answering with open-source large language models[J].Scientific Reports, 2024, 14(1).DOI:10.1038/s41598-024-64827-6.

[21] Li N , Gao C , Jin D , Liao Q. Disentangled Modeling of Social Homophily and Influence for Social Recommendation[J].IEEE Transactions on Knowledge and Data Engineering, 2023(6):35. doi:10.1109/TKDE.2022.3185388.

[22] Gao X , Li X , Qi G Y .GELU-LSTM-encoder-decoder fault prediction for batch processes based on the global -local percentile method[J].The Canadian Journal of Chemical Engineering, 2024, 102(6):2208-2227.DOI:10.1002/cjce.25170.

[23] Xu Y , Liu T , Du P .Volatility forecasting of crude oil futures based on Bi-LSTM-Attention model: the dynamic role of the COVID-19 pandemic and the Russian-Ukrainian conflict[J].Resources Policy, 2024, 88.DOI:10.1016/j.resourpol.2023.104319.

[24] Garg M , Ghosh D , Pradhan P M .Multiscaled Multi-Head Attention-Based Video Transformer Network for Hand Gesture Recognition[J].IEEE signal processing letters, 2023.DOI:10.1109/LSP.2023.3241857.

[25] Xinyu H , Tongxuan Z , Guiyun Z .MultiHop attention for knowledge diagnosis of mathematics examination[J].Applied Intelligence: the

International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2023, 53(9):10636-10646.

[26] Zheng X, Chen H L, Ma Y Q, Wang Q. A neural machine translation model incorporating dependent syntax and LSTM[J]. Journal of Harbin Institute of Technology, 2023, 28(3):20-27.DOI:10.15938/j.jhust.2023.03.003.

[27] Sharaff A , Chowdhury T R , Bhandarkar S .LSTM based Sentiment Analysis of Financial News[J].SN Computer Science, 2023, 4:1-8.DOI:10.1007/s42979- 023-02018-2.

[28] Wołk, Krzysztof, Marasek K .Enhanced Bilingual Evaluation Understudy[J].Computer ence, 2014.DOI:10.12720/lnit.

[29] Ramadhan T I , Ramadhan N G , Supriatman A .Implementation of Neural Machine Translation for English-Sundanese Language using Long Short Term Memory (LSTM)[J].Building of Informatics, Technology and Science (BITS), 2022.DOI:10.47065/bits.v4i3.2614.

[30] Tiwari G , Sharma A , Sahotra A , Kapoor R. English-Hindi Neural Machine Translation-LSTM Seq2Seq and ConvS2S[C]//2020 International Conference on Communication and Signal Processing (ICCSP).2020.DOI:10.1109/ICCSP48568.2020.9182117.