Review on Machine Translation Model under Low Resource Condition

Yihang Feng

Quanzhou University of Information Engineering, Quanzhou, 362000, China
yxf1004@sru.edu

Abstract. This paper provides a comprehensive review of the development, challenges, and future prospects of machine translation. It covers the evolution from rule-based systems to neural machine translation (NMT) models, using recurrent neural networks (RNNS) and convolutional neural networks (CNNS) as examples to illustrate ways to solve problems such as accuracy and fluency. Multilingual translation, domain adaptation, and decoding acceleration are also discussed as potential development areas. Despite the progress, challenges remain, such as dealing with rare words and long sentences. The paper emphasizes the importance of conducting research in a variety of language disciplines to overcome these limitations. Overall, machine translation will continue to evolve, with the goal of achieving greater accuracy, efficiency, and intelligence in the future.

Keywords: Machine Translation, Natural Language Processing, Low Resource Condition

1 Introduction

Machine translation (MT) is the process of automatically translating text from one language to another using computer algorithms. The development of the MT model dates back to the 1950s, when researchers began experimenting with rule-based systems that relied on dictionaries and grammar rules. However, these early models had limitations in producing accurate translations, and it was not until the advent of statistical machine translation (SMT) in the 1990s that significant progress was made.

Despite advances in machine translation technology, there are still two major issues that need to be addressed: accuracy and fluency. Accuracy refers to the ability of the machine translation model to produce a translation that is faithful to the original text, while fluency refers to the ability of the model to produce a translation that is naturally smooth and easy to read. These problems have been addressed by the development of NMT models, which use deep learning algorithms to improve translation quality. The NMT model uses an encoder-decoder structure, in which the encoder encodes the source language sentence into a vector and the decoder converts the vector into the target language sentence. The attention mechanism is an important

B. H. Ahmad (ed.), *Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)*, Advances in Intelligent Systems Research 180,
https://doi.org/10.2991/978-94-6463-370-2_30

part of the NMT model, which helps the decoder focus on the relevant part of the source language sentence at each decoding step. The application value and importance of MT model is huge [1]. They can be used to translate large amounts of text quickly and accurately, so they are widely used in various fields such as business, government, and academia. For example, companies can use the MT model to translate product specifications, contracts, and other business documents to better communicate with international customers. Governments can use the MT model to translate legal documents, policy documents and other important documents to better communicate with other countries. The academic community can use the MT model to translate research papers, academic journals, and other academic literature to better communicate with international peers.

In addition, MT models can help break down language barriers and facilitate cross-cultural communication. They can help people better understand other countries and cultures, and promote cooperation and exchange between different countries. MT, by processing natural languages and translating them into other natural languages, plays an important role in linguists, computer scientists and more.For example, in the tourism industry, MT models can help tourists better understand the local culture and history and communicate with the locals. Governments can use machine translation models to translate legal documents, policy documents, and other important documents to better communicate with other countries. Academia can use machine translation models to translate research papers, academic journals, and other academic literature to better communicate with international peers.

In this study, we aimed to investigate the effectiveness of NMT models in addressing accuracy and fluency issues in MT. We will evaluate the performance of multiple state-of-the-art NMT models on various translation tasks and compare their results with those obtained using SMT models [2]. Our findings will provide insights into the strengths and weaknesses of different MT approaches and help guide future research efforts in this area.

2 Solution

2.1 Frame

The problems of accuracy and fluency in machine translation have been solved by the development of NMT models. The NMT model uses an encoder-decoder structure, in which the encoder encodes the source language sentence into a vector and the decoder converts the vector into the target language sentence. The attention mechanism is an important part of the NMT model, which helps the decoder focus on the relevant part of the source language sentence at each decoding step. The NMT model uses deep learning algorithms to improve the quality of translation, thus solving the problem of accuracy and fluency in machine translation. In addition, NMT models can also improve translation quality by using a large number of parallel corpora. These parallel corpora contain corresponding texts between source and target languages and can be used to train NMT models. In short, NMT model is one of the most advanced techniques in the field of machine translation, and it has achieved great success in practical applications. The development of the NMT model is an important milestone in the field of machine translation. It not only solves the problems of accuracy and fluency in machine translation, but also improves the speed and efficiency of translation. NMT models have achieved great success in various machine translation tasks, including text translation, speech translation and image translation. In addition, NMT models can also improve translation quality by using a large number of parallel corpora. These parallel corpora contain corresponding texts between source and target languages and can be used to train NMT model However, NMT models still have some challenges and limitations. For example, NMT models require large amounts of training data to achieve optimal performance. In addition, NMT models can have difficulty processing rare words and long sentences. Therefore, future research directions include how to better use limited training data to train NMT models and how to improve NMT models to deal with problems such as rare words and long sentences. In short, NMT model is one of the most advanced techniques in the field of machine translation, and it has achieved great success in practical applications. As the technology continues to evolve, we believe that NMT models will play an even more important role in the future.

2.2 Basic solution

NMT model uses an encoder-decoder structure, where the encoder and decoder are two neural networks. The encoder encodes the source language sentence into a vector that contains the semantic information of the source language sentence. The decoder converts this vector into a target language sentence. In this process, the decoder uses an attention mechanism to focus attention on the relevant parts of the source language sentence. The training process for NMT models usually includes the following steps:

Data prepossessing: The text data of source language and target language are cleaned, word segmentation, ionization, etc., so as to facilitate subsequent training and evaluation [3].

Build thesaurus: Transform the per-processed text data into a digital representation to build thesaurus in the source and target languages.

Training model: NMT model is trained using per-processed text data and constructed vocabularies. The training process usually requires the use of hardware acceleration devices such as G pus in order to speed up training.

Evaluation model: The trained NMT model is evaluated using a test set to calculate the model's performance indicators in terms of translation accuracy and fluency.

Adjust the model: The NMT model is adjusted and optimized based on the evaluation results to improve the translation quality [4].

Deploy the model: Deploy the trained NMT model into a production environment to facilitate automated translation services.

In short, NMT model is one of the most advanced techniques in the field of machine translation, and it has achieved great success in practical applications. As the technology continues to evolve, we believe that NMT models will play an even more important role in the future.

2.3 Detailed solution

To further improve the accuracy and fluency of machine translation, attention mechanisms are introduced into encoder-decoder structures. The attention mechanism can make the model pay more attention to the parts of the source language sentences that are relevant to the target language sentences, thus improving the translation quality. In addition, several other techniques have been applied to machine translation. Bidirectional recurrent neural network (Bidirectional RNN) is a common neural network structure that can take into account both past and future information about the input sequence, thereby improving the expressiveness of the model. In machine translation, bidirectional RNNS can be used in encoders and decoders to improve the accuracy and fluency of translation. Specifically, bidirectional RNNS allow the encoder and decoder to model input sequences in both directions of the source and target languages, respectively, to better capture contextual information in the sequence. This allows the model to understand the input sequence more accurately and produce a more fluid and natural translation. In addition, two-way RNNS can further improve translation quality by using attention mechanisms in encoders and decoders. In summary, bidirectional RNNS are a very effective technique to help machine translation models better understand input sequences and produce more accurate, fluid translations [5].

Another solution focuses on the lack of easy context information in machine translation. This method first inputs the current sentence of the source language and the previous sentence of the source language independently into the source language sentence encoder and context encoder, and then uses the attention mechanism to map the output of the two encoders into the final context. At the same time, the subject representation encoder based on Bi-GRU and CNN maps the source language sentence input after word embedding into the topic representation. Finally, the fused sentence representation and the topic representation are decoded by two tandem attention mechanisms, respectively. The results show that the proposed method can improve the performance of text level neural machine translation compared with the benchmark system [6]. Table 1 summarizes the findings of researchers working on machine translation under the above low-resource conditions.

3 Machine translation under low resource conditions

3.1 Development process and history

Since 1947 Warren Weaver has proposed the possibility of machine translation [7]. Although the concept of translation has existed for thousands of years in human history, machine translation has only been developed for more than 70 years. Throughout the development of machine translation, the course of twists and turns and intriguing, it can be said that reviewing the history of machine translation will have a good inspiration for in-depth understanding of relevant technical methods, and even to understand the development of the entire field of natural language processing [8]. Machine translation under low resource conditions refers to the use of a small number of parallel corpora or even no parallel corpora to train machine translation models in the absence of a large number of parallel corpora. In such cases, the performance of machine translation models tends to be greatly limited because there is not enough data to train the models. The goal of these methods is to train machine translation models with a small number of parallel corpora or even no parallel corpora, and to achieve better performance without a large number of parallel corpora. But then in modern times, Professor Zhang Huaao, for the situation of low resources, how to efficiently use the existing bilingual data, fully tap its performance, and so on Research is carried out on the problem of raising the sexual capacity of low resources by machine translation [9]. She proposed a method of capturing knowledge from source end to target end and from target end to source end at the same time, which is called double direction knowledge steaming. knowledge-distillation (KD)is an algorithm that transfers the "knowledge" of the teacher model to the student model [10]. She also used the scarce parallel language to pass from source to target and target to source. The data are enhanced by the end translation side mixing and forward and backward knowledge distillation methods. Experiments were performed on 3 speech pairs (6 directions) of low source data sets, and the relative strength baseline system obtained an increase in the mean 2 B Low-Enriched Uranium(LEU) fraction [11]. Then, Liu Wuying and Wang Lin of Guangdong University of Foreign Studies studied the scarcity of bilingual sentence pairs, which led to the inability of some deep learning-based machine translation algorithms to achieve better performance in low-resource machine translation. Therefore, aiming at the problem of language resource construction in low-resource machine translation, the idea of corpus cycle promotion is proposed [12]. They used the trained machine translation model to incrementally construct bilingual pseudo-corpus by translating monolingual sentences twice. Then the initial bilingual resources are enhanced with bilingual pseudo-corpus. Finally, based on the enhancement of bilingual resources, the above process is repeated until the trained machine translation model meets the performance requirements. As a result, the experimental results in multiple languages prove that the expanded language resources can effectively improve the performance of low-resource machine translation. Different from the above two, Yu Zhiqiang et al., from Kunming University of Science and Technology, used an improved loss function to reduce the impact of low-quality prototype sequences on the model while using a gating mechanism to control information flow at the decoding end. Experimental results on multiple datasets show that compared with the baseline model, the proposed method can effectively improve the performance of machine translation in low-resource scenarios [13].

Table1. People who have influenced the improvement of machine translation above.

Who	Targeted problem	Improvement	Achieved effect
Warren Weaver	The difficulty of human translation.	The concept of machine translation is proposed.	It provides theoretical support for statistical

			machine translation.
Zhang Huaao	How to use the existing bilingual data efficiently and fully tap its performance.	The method of knowledge-distillation is proposed to enhance the data.	The average LEU score of the relative strength baseline system is much higher.
Liu Wuying, Wang Lin	The scarcity of bilingual sentence pairs causes some machine translation algorithms to fail to achieve better performance in low-resource machine translation.	Using the trained machine translation model, the bilingual pseudo-corpus is gradually constructed by translating monolingual sentences twice, and the above process is repeated.	The expanded language resources can effectively improve the performance of low-resource machine translation.
Yu Zhiqiang's team	Due to the shortage of parallel corpus resources, the prototype sequence cannot be matched or the quality of the sequence is poor.	An improved loss function is used to reduce the impact of low quality prototype sequences on the model, and a gating mechanism is used at the decoding side to control the information flow.	Compared with the baseline model, their proposed method can effectively improve the performance of machine translation in low-resource scenarios.

3.2 In detail

Take the method of corpus cycling for example, Machine translation is an algorithmic process of retelling source natural language semantics using target natural language form. Rule-based machine translation algorithm is time-consuming and difficult to ensure the self-consistency among many rules, so it is difficult to popularize it into practical machine translation applications. One way of thinking is to continue to promote deep learning. This paper explores the language resource construction algorithm that makes full use of monolingual corpus, and realizes artificial fabrication of false corpus. In this way, the bilingual sentence pair resources can be expanded incrementally by using the appropriate scale of bilingual seed resources and ultra-large scale monolingual resources, and the training of machine translation model can be promoted by increasing the corpus step by step [14]. The incremental pseudo corpus in the cyclic promotion of corpus comes from the trained machine translation model, and the machine translation model is trained from the newly created pseudo corpus enhanced training set, which is a self-advancing closed-loop idea based on the homogeneous machine translation model.

4 Multilingual translation problems

4.1 Development process and history

Early machine translation systems were primarily rules-based, using human-written rules to convert source language text into target language text. However, these systems struggle to handle complex natural language structures and require a lot of manual work.At present, there is a neural machine translation model for multilingual pairs. However, due to limited resources, it cannot be applied to all languages in the world. Practice one translation model for each pair of words [15]. In order to improve the performance of network multilingual translation system in terms of translation speed and matching rate, Qi Wei et al., from Guangdong Technical Normal University, proposed a design of network multilingual timely translation system based on big data analysis. Feature extraction algorithm was used to extract semantic features of network multilingual, and combined with the design of network multilingual timely translation algorithm, software design of the system was completed. The results of simulation test show that the network multi-language timely translation system based on big data analysis has better performance in terms of translation speed and matching rate [16]. In addition, Xie's team is aware of the problem that existing methods tend to directly mix all language corpus as training corpus, failing to take advantage of the correlation and similar information between multiple languages. To solve this problem, they proposed a course learning method based on language correlation degree to improve the overall performance and convergence speed of multi-language neural machine translation, so that the model training from the whole training to a series of training courses, reducing the difficulty of training. The results show that the proposed method is superior to the multilingual baseline translation system and can shorten the training time by up to 64% [17]. Multilingual neural machine translation uses a single encoder-decoder model to model simultaneous translation between multiple languages. Multilingual neural machine translation can not only promote the knowledge transfer between related languages, improve the translation quality of low-resource languages, but also realize the translation between unseen language pairs. In view of this situation, Liu Junpeng pointed out that multi-language neural machine translation still has the problems of insufficient language diversity modeling ability and poor translation quality of unknown languages. They propose a varied-dimension bilingual adapter model based on the existing adapter model, and add a bilingual adapter between each sub-layer of the Transformer model to extract the unique features of each language pair, and adjust the specific language expression space at both ends of the encoder and decoder by changing the adaptor hidden layer dimension. The experimental results show that the variable dimension bilingual adapter model can significantly improve the performance of multilingual translation [18]. Table 2 summarizes the findings of the above multilingual translation researchers.

Table2. People who are good at studying multilingual problems above

Who	Targeted problem	Improvement	Achieved effect
Qi Wei's team	Due to the influence of string length on the network multilingual timely translation system, the translation speed of the system is slow and the matching rate is low	The hardware design of the system is completed through the design of network multilingual timely translation server and network multilingual lexical analyzer. Feature extraction algorithm is used to extract semantic features of network multi-language	The design of network multi-language timely translation system is realized, and the performance of the network multi-language timely translation system based on big data analysis has been improved in terms of translation speed and matching rate
Yu Dong's team	Most of the existing methods directly mix all language corpus as training corpus, and fail to take advantage of the correlation and similar information between multiple languages	Two measures of language correlation are proposed: singular vector canonical correlation analysis is used to rank different languages and cosine similarity is used to rank different sentences in a particular language	The proposed method is superior to the multilingual baseline translation system and can shorten the training time by up to 64%
Liu Junpeng's team	Multilingual neural machine translation can not only promote knowledge transfer between related languages but also realize translation between unknown language pairs.	A bilingual adapter is added between each sub-layer of the model to extract the unique features of each language pair.	Monolingual adapter model can improve the translation quality of unseen language pairs without affecting the performance of multilingual translation.

4.2 In detail

The problem of slow multi-language translation and low matching rate. This is a common problem, and there are many factors that can affect the speed and accuracy of multilingual translation. The difficulty of translation varies from language to language, and translation between some languages may be more difficult than between others. The reason why the network multilingual translation system can be recognized by most users is because many developers are in the design process [19]. Based on the above research background, some people use the network multilingual timely translation system design of big data analysis.The design of the network multi-language timely translation server is divided into two steps: first training and then decoding. Training is to collect the network multi-language data in the huge network multi-language database for solving the maximum probability, and decoding is to find the solution with the maximum probability by using the training results.Network multi-language matching rate test, network multi-language matching rate can reflect the translation accuracy of network multi-language translation system, respectively using the document network multi-language timely translation system, document network multi-language timely translation system and network multi-language timely translation system based on big data analysis, to test the network multi-language matching rate.

5 Discussion

5.1 Feelings

Machine translation model is a very useful technology that can help people better understand other countries and cultures and promote cooperation and communication between different countries. With the continuous development of technology, machine translation models have made great progress and are playing an increasingly important role in practical applications [20]. However, machine translation models still have some challenges and limitations, such as weak processing power for rare words and long sentences. Future research directions include how to better use limited training data to train machine translation models, and how to improve NMT models to handle these problems. In modern times, many researchers have realized the limitations of solutions to multilingual problems and made efforts to achieve the results of the present hundred flowers.

5.2 Advantages and disadvantages

The machine translation model is a very useful technology that can help people better understand other countries and cultures, and promote cooperation and communication between different countries. With the continuous development of technology, machine translation models have made great progress and are playing an increasingly important role in practical applications. However, machine translation models still have some challenges and limitations, such as weak processing power for rare words and long sentences. Future research directions include how to make better use of limited training data to train machine translation models and how to improve NMT models to handle these problems [21]. At present, the self-attention mechanize-based Transformer model has become the mainstream model for machine translation tasks, and has achieved optimal translation performance on several public test sets. At the same time, great progress has been made in the fields of unsupervised machine translation, multilingual translation, and speech translation. However, machine translation still has some challenges and limitations, such as weak processing power for rare words and long sentences. Future research directions include how to make better use of limited training data to train machine translation models and how to improve NMT models to handle these problems. Machine translation has always been a focus of research in the field of natural language processing, from the earliest rule-based machine translation to the mapping between languages that rely on neural networks to learn end-to-end. This website provides a tutorial, the purpose is to provide a systematic introduction to the basic knowledge and modeling methods of machine translation, and on this basis, some of the leading techniques of machine translation are discussed. The bottleneck problem that hinders the development of machine translation is still structural ambiguity and semantic ambiguity, which can be attributed to three major points: the complexity of translation, the complexity of natural language itself and the limitations of machine translation. In the future, more research results on various language disciplines such as syntax, semantics and even pragmatics are needed to solve these problems.

5.3 The Future

Self-attention-based Transformer model: It has become a mainstream model for machine translation tasks, achieving optimal translation performance on multiple public test sets1. Future research directions include how to make better use of limited training data to train machine translation models, and how to improve NMT models to handle some challenges and limitations, such as weaker processing of rare words and long sentences.

Unsupervised machine translation: Learning mappings between languages can be challenging with little or no resources. At present, unsupervised machine translation usually adopts iterative back-translation1. Future research directions include how to make better use of limited training data to train machine translation models and how to improve NMT models to handle these problems.

Multi-language translation: Multi-language model to realize the translation between multiple languages through a model can effectively reduce the deployment cost of multi-language translation. Simultaneous translation of one source language into many different target languages is one of the most common scenarios for multilingual translation.

Domain adaptation: In neural machine translation, domain migration through fine-tuning is a common approach. However, unconstrained fine-tuning requires very careful hyper-parameter tuning, otherwise it is easy to over fit on the target domain, resulting in performance degradation on the generic domain 1

Decoding acceleration: Lightweight models/non-auto-regressive decoding: Over-parameterized (very large scale) models can effectively improve the performance of neural machine translation, but the large storage overhead and high computational complexity make such models impossible to deploy directly on edge devices [22]. Non-auto-regressive neural machine translation (NAT) systems greatly increase the speed of inference by breaking auto-regression and generating all target words in parallel.

Other directions: For example, how to solve the multi-peak problem, how to improve the computational complexity of the decoding side, etc., are all problems that need to be solved in future machine translation.

6 Conclusion

A machine translation model is a technique that uses computer algorithms to automatically translate text from one language to another. It can help people better understand other countries and cultures, and promote cooperation and exchange between different countries. With the continuous development of technology, machine translation models have made great progress and are playing an increasingly important role in practical applications. However, machine translation models still have some challenges and limitations, such as weak processing power for rare words and long sentences. Currently, the self-attention-based Transformer model has become the mainstream model for machine translation tasks, achieving optimal translation performance on multiple public test sets. At the same time, great progress has been made in the fields of unsupervised machine translation, multilingual translation, and speech translation. However, machine translation still has some challenges and limitations, such as weak processing power for rare words and long sentences. Future research directions include how to make better use of limited training data to train machine translation models and how to improve NMT models to handle these problems. The bottleneck problem that hinders the development of machine translation is still structural ambiguity and semantic ambiguity, which can be attributed to three major points: the complexity of translation, the complexity of natural language itself and the limitations of machine translation. In the future, more research results on various language disciplines such as syntax, semantics and even pragmatics are needed to solve these problems. In short, as the technology continues to evolve, we believe that machine translation models will be more widely used in the future, and will continue to improve to become more accurate, efficient and intelligent.

References

1. Fan Yunge,Lv Yuhui,Han Hongqi:Deep learning prediction model of taxi demand based on context-aware attention mechanism,41-49(2023)

2. Li Han, Zhang Dongsheng, Han Haotian,Liu Chunyan: A method and system for creating a neural machine translation NMT model, 2019
3. Feng Xiongbo,Huang Yuxin,Lai Hua,Gao Yumeng: Research on low resource cross-language summarization method based on multi-strategy reinforcement learning,101-102(2023)
4. Chen Tingting: Research on machine-oriented NMT English translation system, 69-72(2022)
5. Donh Xishuang,Qian Lijun: Semi-supervised bidirectional RNN for misinformation detection,2022
6. Chen Xiwen,Yu Zhengtao,Gao Shengxiang, Wang Zhenhan: Text level neural machine translation with thematic information,11-12(2023)
7. Muasci: Neural Machine Translation for Low-Resource Languages: A Survey, 2021
8. Ankush Garg,Mayank Agarwal: Machine Translation: A Literature Review,1-2(2018),17 pages, 3 figures
9. Zhang Huaao: Optimization and implementation of low resource machine translation system,1-2(2022)
10. Hinton Geoffrey, Vinyals Oriol, Dean Jeff. Distilling the knowledge in a Neural Net-work,2015
11. Zhang Huaao:Optimization and implementation of low resource machine translation system,(45-46,)2022
12. Liu Wuying,Wang Lin: Corpus cycling promotes low-resource neural machine translation,94-95(2023)
13. Yu Zhiqiang,Yu Zhengtao, Huang Yuxin, Guo Junjun,Xian Yantuan: Low resource neural machine translation based on multi-strategy prototype generation,1-2(2023)
14. Liu Wuying, Wang Lin: Corpus cycling promotes low-resource neural machine translation,2-3(2023)
15. Xiao Tong, Zhu Jingbo:Machine Translation: Fundamentals and Models,26-27(2021),NEUNLPLab
16. Qi Wei,Niu Huan,Xiao Lei: Design of network multilingual timely translation system based on big data analysis,154-158(2021)
17. Yu Dong,Xie Wanying,Gu Shuhao,Feng Yang: Multilingual neural machine translation based on language association degree curriculum learning,24-30(2022)
18. Liu Junpeng,Su Jinsong,Huang Degen: Multilingual neural machine translation with specific language adaptation modules,17-23(2022)
19. Qi Wei,Niu Huan,Xiao Lei: Design of network multilingual timely translation system based on big data analysis,154-158(2021)
20. Globalism:Thoughts on promoting multilingualism in cyberspace in the post-pandemic era, 237-242(2021)
21. Zheng Yangmin,Lu Jie: The development of machine translation and its impact on the translation industry,277-278(2022)
22. Yin Fenglin,Zhao Yixin: Challenges and responses: The development of machine translation and the future of translation rights exercise,45-54(2021)