Journal of Electronic Research and Application, 2025, Volume 9, Issue 3
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Research on the Development Status and Difficulties of Machine Translation in the Era of Artificial Intelligence

Ruijing Xu, Qi Bai*

Anhui Institute of Information Technology, Wuhu 241100, Anhui, China

**Author to whom correspondence should be addressed.*

Abstract: Machine translation builds a bridge for cross-language communication by realizing text conversion between different languages. However, there are still many challenges in achieving context-accurate translations. These mainly include how to accurately capture subtle information in context, effectively resolve the ambiguity of polysemous words, properly translate idiomatic expressions, accurately reflect cultural differences, and correctly use terms in specific fields. This article reviews the existing platforms and the latest research results in the field of machine translation, deeply explores the above-mentioned key difficulties, and explores the introduction of artificial intelligence technology. The aim is to improve the overall performance of machine-translation systems, facilitate smoother communication and understanding among people from different cultural backgrounds, further eliminate language barriers, and promote the in-depth integration and development of global multiculturalism.

Keywords: Machine translation; Artificial intelligence; Context analysis

Online publication: 6 June, 2025

1. Introduction

Machine translation (MT) is an automated process of converting source language into target language through computational models and algorithms [1], aiming to eliminate language barriers and promote international communication and cooperation. Its application in fields such as business and academia has significantly improved the efficiency of information transfer. Current research focuses on improving translation accuracy, efficiency, and context-understanding ability. Li improved the accuracy of the neural machine translation (NMT) model by optimizing the long-short-term memory network [2]. Lai *et al*. proposed the CFDT method [3], which uses a discourse-context information-screening mechanism to further optimize the translation effect.

Machine-translation methods mainly include rule-based machine translation (RBMT), statistical machine translation (SMT) [4–9], and neural machine translation (NMT). Although significant progress has been made in

MT technology, it still faces many challenges, including context understanding, idiom translation, differences in grammatical structures, processing of complex sentences, translation of technical terms, emotion conveyance, and real-time response. Solving these challenges is of great significance for improving translation performance, promoting cross-language communication, and integrating diverse cultures. This article will deeply explore the above-mentioned difficulties, hoping to provide theoretical support for the further development of machine translation technology.

2. Context and contextual understanding

Context understanding in machine translation is a core challenge, mainly reflected in aspects such as pronoun resolution, cultural differences, and semantic ambiguity. Pronoun resolution requires the system to accurately identify referential relationships. For example, in the sentence "Mary saw Lucy, and she waved at her," it is difficult to determine the referents of "she" and "her" without context, which may lead to translation errors. Cultural differences are reflected in language norms. For example, the Chinese term "Xiansheng" can refer to a male scholar or be used as an honorific for a woman (such as "Yangjiang Xiansheng"), while there is no exact equivalent in English like "Mr" or "Sir," and it needs to be judged based on the context. The problem of semantic ambiguity is particularly prominent in idiom translation. For example, "face the music" should be understood as "bear the consequences" rather than its literal meaning.

To address these challenges, it is necessary to improve natural-language understanding ability, integrate cultural knowledge, and establish an effective context-modeling mechanism. Current research focuses on optimizing neural-network architectures, improving pre-training methods, and integrating external knowledge sources to enhance the system's context-awareness. The breakthrough of these technologies is of great significance for achieving accurate and natural machine translation.

3. Idiom translation

Idiom translation is a core challenge in machine translation, mainly manifested in four aspects: cultural differences, metaphorical meanings, language structures, and context dependence. Cultural differences lead to semantic gaps. For example, it is difficult to accurately convey the cultural connotation of "Huashetianzu" through literal translation [10]. Identifying metaphorical meanings requires the system to break through literal limitations and establish a deep-semantic mapping model. The particularity of language structures, such as the four-character structure of Chinese idioms and the alliteration in English idioms, increases the complexity of translation. Especially when dealing with humorous expressions, it is necessary to understand the cultural references of word games. Context dependence requires the system to have multi-level context-analysis capabilities to accurately identify the semantic orientation of idioms. For example, "break the ice" has multiple meanings in different contexts.

Future research should focus on constructing cultural-cognitive models, innovating semantic-analysis techniques, and improving context-understanding ability. By integrating multi-modal information and deep-learning technology, it is expected to achieve the unity of accuracy and naturalness in idiom translation and promote the application of machine-translation systems in cultural dissemination and cross-language communication [11].

4. Differences in grammatical structures

Differences in grammatical structures are a core challenge faced by machine translation, mainly reflected in word order, verb morphology, grammatical cases, agreement, and prepositional systems. Word-order differences require the system to have the ability to restructure syntax. For example, when converting the English "SVO" structure to the Chinese "STVO" structure. The complexity of verb-morphology changes varies from language to language [12]. The verb-conjugation system in inflectional languages such as Spanish is more complex than that in English, and it is necessary to accurately map tenses and persons. The grammatical case system is crucial in languages like German. It is necessary to identify the syntactic function of nouns and select the correct case form. Agreement requirements involve the matching of the gender and number of adjectives and nouns. For example, the morphological changes in "niño feliz" and "niña feliz" in Spanish. The differences in prepositional systems are reflected in the expression of spatial and temporal relationships [13]. For example, the correspondence between "in" in English and "dans" in French.

To overcome these challenges, machine-translation systems need to establish deep-level grammatical-analysis models, integrate syntactic rules and semantic information, and achieve cross-language syntactic-structure conversion. This requires the system to not only have explicit knowledge of grammatical rules but also obtain the implicit ability to recognize syntactic patterns through deep learning, to generate fluent translations that conform to the grammatical norms of the target language. Future research should focus on developing more accurate syntactic-analysis algorithms and more flexible syntactic-restructuring mechanisms to improve the accuracy and naturalness of machine translation.

5. Translation of long and complex sentences

When dealing with sentences containing multiple clauses, long sentences, or complex sentence structures, machine-translation systems often have difficulty correctly understanding the sentence structure, resulting in grammatical errors or logical confusion. For example, for the English sentence "The man who was walking in the park, which was very large and beautiful, suddenly stopped when he saw the bird." If the long sentence is not properly split and hierarchically analyzed during translation into Chinese, the translation may be ambiguous, unnatural, or lack coherence. The above phenomenon highlights the challenge of complex sentence structures for machine translation. In complex sentence structures, subordination and coordination are core factors [14]. Many languages use subordinate clauses, relative clauses, and conjunctions to increase the semantic depth and hierarchical complexity of sentences. Translating such structures requires machine-translation systems to accurately capture hierarchical relationships to ensure that the target-language text can express logical relationships, maintain naturalness and fluency, and still faithfully convey the meaning and grammatical features of the original text.

To effectively deal with the above-mentioned complex grammar and syntactic phenomena, machine-translation systems should adopt more complex syntactic-analysis algorithms, deeply understand language rules and structures, and train the model using high-quality training data that covers a variety of complex sentence patterns. At the same time, with the continuous development of neural machine translation and natural-language-processing technology, the integration strategy of using context information, language features, and syntactic analysis is constantly optimizing the performance of machine-translation systems. These research and developments are expected to significantly improve the ability of machine translation to handle complex sentence structures and grammatical conversions, to achieve more accurate, coherent, and natural translation outputs.

6. Translation of technical terms in specific fields

Machine translation faces three core challenges in the conversion of technical terms: semantic specificity of the field, dynamic evolution of terms, and complexity of language structures. Firstly, technical terms have strict field-reference and conceptual uniqueness, but their cross-language mapping is often restricted by cultural-load differences and lexical gaps. For example, the medical term “stroke” needs to be distinguished as “Zhongfeng” (stroke in medicine) or “Jida” (hit in general semantics) according to the context. The legal term “due process” needs to go beyond literal translation through the analysis of legal connotations. This requires the system to construct a three-dimensional disambiguation mechanism that integrates field-knowledge graphs, multi-modal term libraries, and context-association models to achieve accurate conceptual correspondence. Secondly, technical terms in science and technology are proliferating exponentially. For example, in the field of information technology, “cloud computing” has spawned sub-concepts such as “virtualization.” The update cycle of traditional translation methods is difficult to match the speed of concept formation. It is necessary to establish a dynamic update framework based on field-standard literature and real-time corpus mining and form an open-evolution system through term-life-cycle management [15].

Effective solutions need to integrate three modules: dynamic term maintenance, deep-semantic analysis, and resource-optimization allocation, and construct an intelligent translation system with continuous learning ability. The core lies in achieving the organic coordination of field-knowledge embedding, context-awareness, and term calculation.

7. Summary and outlook

This article introduces the current situation and research results of machine translation and summarizes many difficulties in machine translation, such as context understanding, idiom translation, differences in grammatical structures, processing of complex sentences, and translation of technical terms. Further in-depth research is needed to address these issues. It is hoped that with the support of artificial-intelligence technology, machine translation can become more accurate and efficient.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Naveen P, Trojovsky P, 2024, Overview and Challenges of Machine Translation for Contextually Appropriate Translations. iScience, 27(10): 110878.

[2] Li Z, 2024, Research on the Optimization of Neural Machine Translation Models based on the Continuous Bag-of-words Model. Automation & Instrumentation, (11): 48–52.

[3] Lai H, Zhang Y, Guo J, et al., 2024, Context-Filter for Document-Level Neural Machine Translation. Journal of Kunming University of Science and Technology (Natural Science Edition), (06): 75–85.

[4] Stoyanova K, Stoyanov D, Khorev V, et al., 2024, Identifying Neural-network Structures Explained by Personality Traits: Combining Unsupervised and Supervised Machine Learning Techniques in Translational Validity Assessment. The European Physical Journal Special Topics, 2024: 1–21.

[5] Wang W, Jiao W, Huang T, et al., 2025, On the Shortcut Learning in Multilingual Neural Machine Translation. Neurocomputing, 615: 128833.

[6] Tao Y, Tao D, 2021, Research on Machine Translation Algorithms based on DNN and Rule Learning. Computer Measurement & Control, 29(01): 150–153 + 158.

[7] Cai X, Wen B, 2021, Statistical Analysis of Machine Translation Error Types in Chinese-to-English Translation Taking Chinese-to-English Translation of Publicity Texts as an Example. Journal of Zhejiang Sci-Tech University (Social Sciences Edition), 46(02): 162–169.

[8] Qin W, 2020, Research on Neural Machine Translation with Incorporated Rule Knowledge, dissertation, Soochow University.

[9] Liu X, 2020, Optimization Methods for Statistical Machine Translation Results from the Perspective of the “Skopos Theory,” dissertation, Beijing Foreign Studies University.

[10] Ding Q, Cao H, Cao Z, et al., 2025, Dual-perspective Fusion for Word Translation Enhancement. Information Fusion, 117: 102815.

[11] Sun H, Kong B, 2024, Sustainable Improvement and Application of Multilingual English Translation Quality Using T5 and MAML. Discover Artificial Intelligence, 4(1): 98.

[12] Li J, Hu X, 2024, A Better or Worse Communicator? Comparing Human and Machine Translation in Source-language Shining Through Across Registers. Lingua, 312: 103834.

[13] Sethiya N, Maurya KC, 2025, End-to-End Speech-to-Text Translation: A Survey. Computer Speech & Language, 90(C): 101751.

[14] Han Q, 2024, Research on the Accuracy of Machine Translation in Cross-cultural Communication based on Embedded Neural Networks. International Journal of High-Speed Electronics and Systems, 2024: 2540125.

[15] Lim WZ, Vylomova E, Kemp C, et al., 2024, Predicting Human Translation Difficulty with Neural Machine Translation. Transactions of the Association for Computational Linguistics, 12: 1479–1496.

Publisher’s note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.