Application Of Improved Neural Network Technology In English Translation: Refining And Reducing Redundancy Through Self Attention Mechanism

Qingyue Su*

School of English, Hainan College of Foreign Studies; Wenchang Hainan 571300 China

** Corresponding author. E-mail: qingyue_su56@outlook.com; suqingyue1001@163.com*

Received: May. 24, 2026; Accepted: Aug. 03, 2026

In recent years, neural network-based approaches have revolutionized the field of machine translation (MT), with Transformer models and their self-attention mechanisms becoming the dominant paradigm. This research presents an improved neural network approach for English translation, focusing on refining output quality and reducing redundancy through an enhanced self-attention mechanism. The Monarch Butterfly Optimized Bidirectional Encoder Representations from Transformers with self-attention (MBO-BERT-SAtt) is a model that dynamically weighs contextual information to produce fluent and lexically diverse translations while minimizing repetitive elements. The research utilizes a 2000 -pair English-Japanese parallel translation dataset, a large-scale bilingual corpus combining Japanese-English parallel texts. Data preprocessing includes normalization and tokenization of long sentence pairs to improve model training efficiency. The model incorporates a dynamic attention head weighting mechanism, which selectively prunes redundant heads during training to optimize translation performance and computational resources. For evaluation, key metrics such as BLEU, and METEOR are used to quantify translation accuracy, fluency, and redundancy reduction, respectively. The Python implementation of the proposed approach achieves 97.4% recall, 98.5% accuracy, 96.6% F1-score, 70.5 BLEU, 67.6 METEOR, 1.0 s translation time, and 5.4% error rate. Human evaluation further confirms enhanced readability and contextual coherence. This approach demonstrates that refining neural translation models through targeted SAtt mechanisms and redundancy-aware training pipelines leads to more precise, natural, and efficient English translations. The results highlight the potential of combining neural network advancements with linguistic insights for next-generation MT systems capable of handling complex language phenomena such as repetition.

Keywords: Self-Attention Mechanism, Translation Quality Enhancement, Redundancy Reduction, Translation Refinement, English Translation, Bidirectional Encoder Representations from Transformers with self-attention (MBO-BERT-SAtt)

http://dx.doi.org/10.6180/jase.202611_34.063

1. Introduction

Machine translation (MT) is the driving force behind many aspects of how people communicate globally. Historically, conventional algorithms tend to struggle with issues of context, accuracy and repetition [1]. Neural networks combined with attention mechanisms enhance language understanding, enabling more fluent, logical, and accurate translations through improved contextual representation and processing [2]. Whereas early rule-based and statistical models required the use of large amounts of training data to develop, this requirement has limited the scalability and adaptability of these models [3]. AI-enabled MT has en-

tered into the realm of education, significantly enhancing the abilities of translators to be productive and efficient [4]. The need for multilingual communication is ever-growing as the world continues to become more connected, while at the same time, the use of AI raises ethical and cultural questions [5]. Panoramic attention mechanisms and data augmentation from both bilingual and monolingual language corpora lead to further improvement of translation quality, especially when processing longer sentences [6].

NMT has advanced global communication by replacing traditional methods with deep learning models that provide adaptable, context-aware translation while overcoming grammar and data scarcity challenges [7]. Encoder-decoder architectures improved semantic representation and contextual understanding for translating complex and lengthy texts [8]. Prompting strategies improve machine translation by optimizing examples, templates, and transfer learning using large language models [9]. Develops an MBO-BERT-SAtt framework to improve translation accuracy, reduce redundancy, enhance contextual coherence, and generate efficient translations for complex bilingual sentence structures.

- Proposes an MBO-BERT-SAtt model improving Japanese-English translation accuracy, contextual understanding, and redundancy reduction.

- Enhanced preprocessing improves input quality, feature stability, translation efficiency, and consistent output through normalization and tokenization.

- Enhanced self-attention improves contextual relevance, minimizes semantic redundancy, and dynamically optimizes attention-head weights.

- Compares translation performance using BLEU, METEOR, and human evaluation of fluency, readability, and contextual accuracy.

System organization: The system is organized into five sections. Section 1 presents the introduction and research motivation. Section 2 reviews related works and existing translation techniques. Section 2 details the proposed MBO-BERT-SAtt methodology. Section 3 provides experimental results and discussion. Section 4 concludes the study and highlights future research directions.

2. Materials and methods

Research analysis highlights translation advancements in accuracy, semantics, and reliability, while identifying limitations in datasets, computational costs, and cross-language generalization challenges. Recent Transformer-based translation studies improved contextual representation, yet redundancy reduction, adaptive attention weighting, and semantic consistency remain challenging, motivating the proposed MBO-BERT-SAtt framework. Existing translation methods improve representation but face limitations in adaptive attention, redundancy reduction, and semantic alignment, motivating the proposed MBO-BERT-SAtt framework. Table 1 summarizes key MT studies, highlighting objectives, methods, results, and limitations across translation approaches.

2.1. Research gap

The NMT has made advances with new models, but problems exist. For instance, a translation system that uses ChatGPT has problems because it does not have access to many languages and APIs [10]. A Multimodal MMNMT system has difficulty aligning image and text, and small datasets make it difficult to use [11]. CNN-NMT is not able to translate idiomatic phrases or domain-specific text [12]. Low-resource DNNs face limitations due to insufficient or generated translation corpora availability [13]. MBO-BERT-SAtt enhances translation robustness, semantics, and adaptability through optimized attention-based learning. The proposed MBO-BERT-SAtt framework introduces adaptive MBO-driven attention-head optimization for redundancy-aware contextual learning, improving semantic consistency and translation quality beyond conventional hybrid models.

MBO-BERT-SAtt dynamically optimizes self-attention weights using MBO to enhance translation quality, semantic consistency, and contextual representation by suppressing redundant heads and strengthening informative features beyond conventional Transformer models.

Japanese-English corpus preparation includes preprocessing with Z-score normalization, tokenization, and noise removal for efficiency. Noise removal improves data quality by eliminating errors, while Z-score normalization enhances feature consistency, stability, transparency, and reproducibility during training. BERT captures context, SAtt refines attention, and MBO optimizes parameters for improved translation, as shown in Fig. 1.

2.2. Data collection

The English–Japanese Parallel Translation dataset (Kaggle dataset) contains 2,000 English–Japanese sentence pairs with diverse features, enabling NMT evaluation of accuracy, fluency, and redundancy reduction using balanced training, validation, and test splits. Adaptive preprocessing, including normalization, noise removal, and subword

Table 1. Highlights of key MT studies, including objectives, methodologies, datasets, performance results, and constraints

Ref.	Objective	Method	Results	Limitations
[10]	Evaluate ChatGPT translation performance	ChatGPT	Standard NMT consistently outperforms ChatGPT.	Language restrictions, API constraints, and reliance on benchmarks.
[11]	Enhance multimodal translation	Multimodal NMT	Outperforms baseline methods.	Dependence on image–text alignment, small dataset size, and generalization issues.
[12]	Handle long sentences and improve robustness	Convolutional Neural Network-based Machine Translation (CNNMT)	Performance improved from 47.9% to 89.7%.	Dataset dependency and reduced effectiveness for idiomatic text.
[13]	Evaluate English translation in low-resource languages	DNN with linguistic feature integration	BLEU improvement of +3.97 (Ch–En) and +2.64 (Tib–Ch).	Reliance on generated data and domain-specific corpora.
[14]]	Improve English translation using an ML–DL hybrid	Hybrid Transformer-NMT framework	Outperforms previous systems.	High computational cost, large dataset requirements, and domain specificity.
[15]	Enhance translation coherence	Gradient-based enhanced Double Deep Q-Network	96.1% recall, 97.2% accuracy, and 95.4% F1-score.	High computational cost and low efficiency on unusual texts.

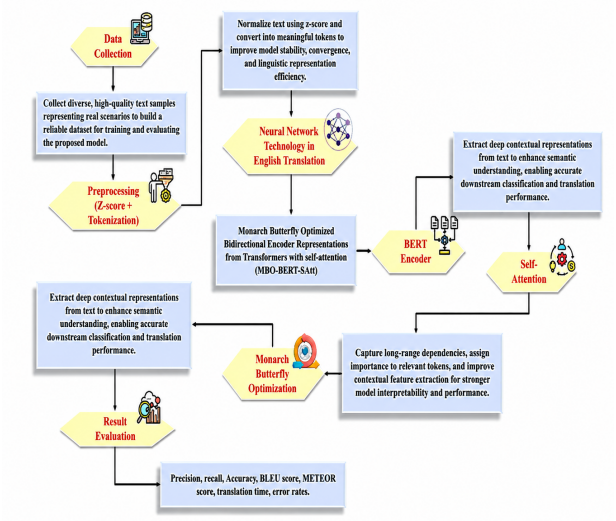

Fig. 1. The methodology flow diagram for English translation

tokenization, improves bilingual data quality, consistency, and model stability. Evaluation on English–Japanese sentence pairs and WMT17 confirms the robustness, scalability, and generalization capability of the proposed translation framework.

2.3. Data preprocessing

Data preparation normalizes, tokenizes, and cleans inputs, enabling faster MBO-BERT-SAtt training with accurate translations.

2.3.1. *Z-score normalization using stabilized features for consistent learning*

Z-score normalization standardizes feature distributions, preventing gradient skew and improving SAtt weight stability during training, as done using the formula given by Eq. (1).

$$v' = \frac{v - \mu}{\sigma} \tag{1}$$

Where v is the original feature value, μ the mean, σ the standard deviation, and v' is the standardized value, improving stability, optimization, and translation quality

Tokenize using structured text for efficient processing Tokenization structures text units, reducing noise and redundancy while enhancing MBO-BERT-SAtt contextual representation efficiency. MBO optimizes attention weights through balanced exploration and exploitation, reducing convergence issues and enhancing contextual representation learning.

2.4. MBO-BERT-SAtt to integrating strengths for superior translation

The MBO-BERT-SAtt model integrates MBO, enhanced BERT, and refined self-attention to improve translation quality, capture dependencies, enhance coherence, and reduce redundancy. The computational complexity analysis quantifies training cost, inference efficiency, memory consumption, and scalability, demonstrating that MBO integration introduces moderate optimization overhead while maintaining efficient translation performance.

2.4.1. *Using BERT to Improve Contextual Understandings during Translation*

BERT provides contextual semantic representations through pre-training and fine-tuning, while MBO optimizes attention weights to improve alignment, reduce redundancy, and generate accurate, coherent, and semantically consistent translations across diverse language scenarios. Self-attention (SAtt) for capturing dependencies for improved translation

SAtt captures long-term dependencies, dynamically focusing linguistic features to improve translation accuracy and quality. Redundant attention heads are identified based on low contribution scores, and adaptive pruning dynamically reduces their weights during training to improve efficiency, reduce redundancy, and enhance contextual representation quality. SAtt uses Query, Key, and Value representations to evaluate token relationships and generate contextual representations. The attention function is expressed in Eq. (2).

$$\text{Attention}(P, L, U) = \text{soft max}\left(\frac{PL^T}{\sqrt{c_l}}\right)u \tag{2}$$

P - Query matrix for contextual relevance. L - Key matrix provides positional characteristics. U - Value matrix containing feature information. PL^T : Projected language traits that have been transposed. u : Value matrix containing contextual feature representations. Vector similarity is measured using the dot-product $(P, L, U).c_l$ - Normalizer prevents gradient magnitude explosions. Candidate solution populations guide neural parameter updates through search-driven optimization, enabling adaptive weight adjustment, improved convergence stability, and enhanced translation performance during iterative training optimization cycles. Softmax stabilizes attention weights, while adaptive SAtt refines token relationships, reduces redundancy, and enhances contextual representations. Dynamic attention weighting optimizes attention heads, enhancing semantic representation, reducing redundancy, and improving translation accuracy while efficiently capturing long-range dependencies with scalability. Fig. 2 illustrates enhanced SAtt reducing redundancy through dynamic head weighting and attention-weighted output generation.

Eight SAtt heads capture syntactic and semantic links for accurate translations. Standardized multihead attention parameterization employs empirically selected attention heads and embedding dimensions to ensure reproducibility, stable convergence, and consistent translation performance across diverse linguistic domains. Residual connections, normalization, and masking ensure stable autoregressive translation with improved contextual quality. MBO

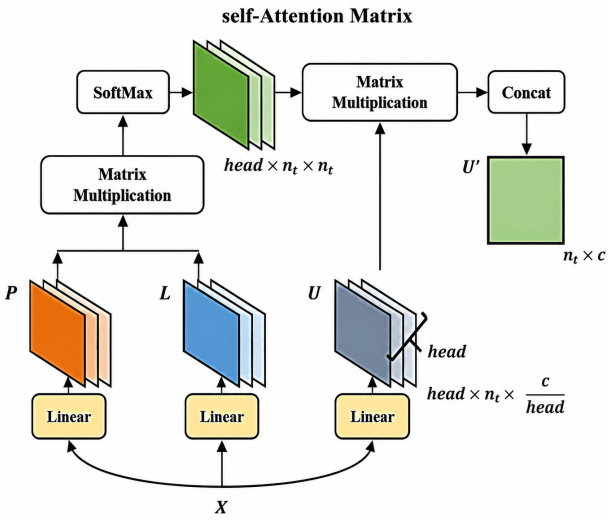

Fig. 2. The architectural flow of the increased SAtt system with dynamic attention head weightings

maintains efficient translation with moderate overhead, while adaptive decoding and attention pruning improve domain terminology handling, efficiency, and contextual representation quality.

2.4.2. *MBO to optimize attention weights for efficiency*

MBO optimizes self-attention by dynamically adjusting attention-head weights, reducing contextual redundancy, and dividing candidate solutions into two adaptive groups for balanced exploration and optimization based on the weights of attention into Land I and Land II in Eq. (3) and Eq. (4).

$$LI = np^{(1)} \times C(r \times np) \tag{3}$$

$$LII = np - np^{(1)} = np^{(2)} \tag{4}$$

np - is the ceiling function, C - attention-based weighting coefficient r - random scaling control factor. LI - candidate set assigned to Land I. LII - candidate set designated for Land II. The bipopulation strategy introduces moderate computational overhead while maintaining an effective exploration-exploitation balance, improving optimization stability and scalability for large-scale translation tasks. The new population is updated by migration and recombination mechanisms in Eq. (5).

$$W_{J,I}^{T+1} = W_{R1,I}^{T}, r = c \times d \tag{5}$$

$W_{J,I}^{T+1}$ Represents the candidate's updated weight. $W_{R1,I}^{T}$ Is the neighbor's reference weight, r Is the migration ratio. Updates candidate weights using neighbor influence and adaptive randomization for optimization. Con-

trolled Lévy-flight updates regulate stochastic perturbations through adaptive step-size adjustment, improving optimization stability, reducing variability, and ensuring reproducible convergence across repeated optimization runs. Fig. 3 illustrates MBO initialization, cost evaluation, migration, and Lévy-flight updates, optimizing attention weights to produce cohesive, fluent, and semantically accurate English translations.

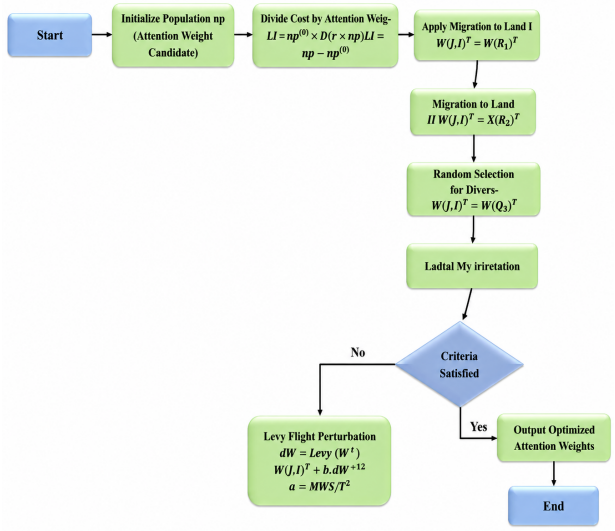

Fig. 3. The flowchart for the MBO method, which optimizes attention weights

The MBO-BERT-SAtt framework combines BERT, self-attention, and MBO to improve translation quality, while lightweight pruning reduces computational cost and supports efficient real-time translation. Generates fluent, accurate translations with contextual coherence, preserved meaning, improved readability, and minimal repetitive phrases.

MBO-BERT-SAtt integrates preprocessing, BERT, self-attention, and MBO optimization to enhance translation quality and reduce redundancy.

3. Results and discussion

This section discusses the results of proposed vs existing methods for neural network technology in refining English translation and reducing redundancy through the SAtt mechanism.

3.1. System configuration

The MBO-BERT-SAtt model was implemented in Python using PyTorch, Transformers, and GPU support, with preprocessing and MBO optimization ensuring stable and efficient training. Hyperparameters for BERT fine-tuning and

Algorithm 1. MBO-BERT-SAtt Translation Framework

Require: Dataset $\mathcal{D}$ and number of iterations T
Ensure: Final translated output
Step 1: Initialize MBO-BERT-SAtt
Initialize $\text{BERT}_{\text{model}}$, attention heads, and MBO population with random weights
Step 2: BERT Encoding
Model $\leftarrow \text{BERT}_{\text{model}}$
Step 3: Self-Attention
function SELFATTENTION(P, L, U, c_l)
  scores $\leftarrow PL^{\mathsf{T}} / \sqrt{c_l}$
  weights $\leftarrow$ softmax(scores)
  return weights$U$
Step 4: MBO Optimization
function MBOOPTIMIZE(population, T)
  for $t = 1$ to T **do**
    $(L_I, L_{II}) \leftarrow$ SplitPopulation(population)
    population $\leftarrow$ Migrate(L_I, L_{II})
    population $\leftarrow$ Recombine(population)
    population $\leftarrow$ LevyFlightUpdate(population, t)
    population $\leftarrow$ Elitism(population)
  return Best(population)
Step 5: MBO-BERT-SAtt Training
function TRAIN($\mathcal{D}$)
  data $\leftarrow$ Preprocess($\mathcal{D}$)
  contextual_emb $\leftarrow \text{BERT}_{\text{model}}$.Encode(data)
  for each attention head h **do**
    $\text{SA}_h \leftarrow$ SELFATTENTION($P_h, L_h, U_h, c_{l,h}$)
  optimized_weights $\leftarrow$ MBOOPTIMIZE(MBO_population, 50)
  Apply optimized weights to attention heads
  final_output $\leftarrow$ TranslationDecoder(SA)
  return final_output

MBO were selected through empirical tuning and validation experiments to achieve stable convergence, improved translation performance, and experimental reproducibility.

3.2. Comprehensive Model Evaluation

MBO-BERT-SAtt performance is analyzed using Fig. 4, Fig. 5 and Fig. 6 demonstrating consistent quality, low redundancy, and domain adaptability.

Visualization results demonstrate reduced attention redundancy and improved contextual focus, where adaptive head weighting enhances meaningful token interactions across varying sentence structures, improving semantic consistency and translation quality in diverse linguistic scenarios.

3.3. Performance evaluation

This section compares proposed and existing methods, including SMT, NMT, Transformer-based NMT, Hybrid RL [14], and GB-EDDQN approaches [15], using multiple evaluation metrics. Three bilingual English-Japanese experts independently evaluated translations for fluency, adequacy, semantic accuracy, and readability using a five-point Likert scale.

Human evaluation involves expert bilingual evaluators

Table 2. Human Evaluation Methodology and Assessment Criteria

Parameter	Description
Number of Evaluators	3 bilingual English-Japanese translation experts
Evaluation Criteria	Fluency, Adequacy, Semantic Accuracy, Overall Readability
Scoring Protocol	Independent evaluation using a 5-point Likert scale (1 = Poor, 5 = Excellent)
Final Score	Average of all evaluator ratings
Inter-rater Reliability	Fleiss' Kappa

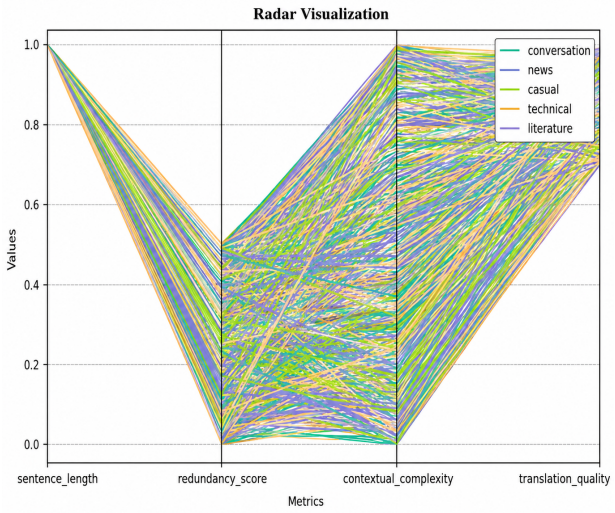

Fig. 4. Parallel coordinates visualization was built to correlate translation metrics with domain-specific linguistic behavior

using structured scoring criteria for fluency, adequacy, and meaning preservation, supported by inter-rater agreement measures to ensure reliability and consistency of assessments, as shown in Table 2.

Table 3 compares the proposed MBO-BERT-SAtt framework with Transformer-based models on the WMT17 benchmark. The proposed model achieves the highest BLEU, METEOR, and Accuracy, demonstrating superior translation performance and generalization.

Table 4 presents five independent experimental runs of the proposed MBO-BERT-SAtt framework. The consistent mean and low standard deviation confirm its robustness, stability, and reproducibility across experiments.

Table 5 presents statistical metrics with low deviations and narrow confidence intervals, confirming stable, consistent, robust, and reliable performance of the proposed framework.

Table 6 presents the statistical significance of the ablation study. Significant p-values and large effect sizes confirm the contribution of each component to the overall performance.

Table 7 presents the ablation study of the proposed framework. Incrementally adding BERT, Self-Attention, and MBO consistently improves translation accuracy and redundancy reduction, demonstrating each component's contribution.

Table 8 presents the statistical significance of performance improvements over baseline models. Low p -values and large effect sizes confirm that the observed gains are statistically significant.

4. Metrics explanation

Precision and recall evaluate translation correctness and coverage, while Fig. 5, Fig. 6 and Fig. 7 and Table 9 show that the MBO-BERT-SAtt model achieves 70.5 BLEU, 67.6 METEOR, 1.0 s translation time, and 5.4% error rate, demonstrating accurate and efficient translation. Accuracy, Recall, and F1-score complement BLEU and METEOR by evaluating translation correctness, semantic preservation, and overall prediction consistency within the proposed machine translation framework

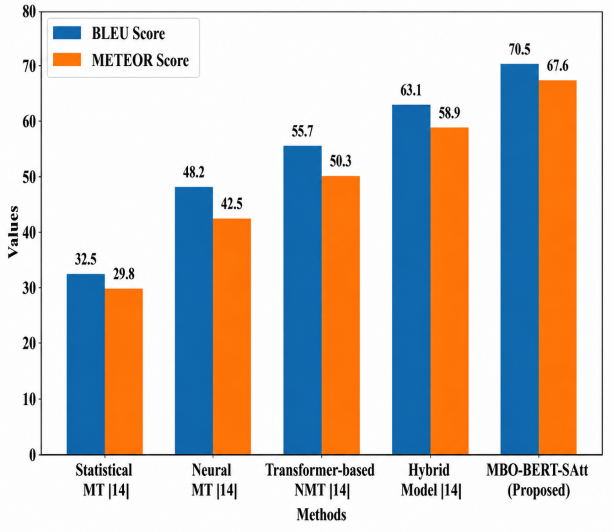

Fig. 5. The BLEU and METEOR values for the proposed methods

Table 10 demonstrates that MBO-BERT-SAtt outperforms GB-EDDQN with higher recall, accuracy, and F1-score, improving contextual dependency learning and translation performance. Multi-metric evaluation com-

Table 3. Performance Comparison on the WMT17 Japanese-English Benchmark

Model	BLEU	METEOR	Accuracy (%)
Transformer-base	38.6 ± 1.2	35.2 ± 1.1	87.4 ± 0.8
Transformer-big	42.8 ± 1.0	39.1 ± 0.9	89.2 ± 0.7
Hybrid Model [14]	49.2 ± 0.9	45.6 ± 0.8	92.8 ± 0.5
MBO-BERT-SAtt (Proposed)	53.7 ± 0.7	49.8 ± 0.6	95.1 ± 0.4

Table 4. Results from Five Independent Experimental Runs

Run	BLEU	METEOR	Accuracy (%)	Recall (%)
1	69.8	66.9	98.2	97.1
2	70.3	67.4	98.4	97.3
3	71.1	68	98.7	97.6
4	70.4	67.5	98.5	97.5
5	70.9	68.2	98.6	97.5
Mean ± SD	70.5 ± 0.5	67.6 ± 0.5	**98.5** ± 0.2	97.4 ± 0.2

Table 5. Statistical Robustness Analysis of the Proposed MBO-BERT-SAtt Framework

Metric	Mean	SD	95% Confidence Interval
BLEU	70.5	0.5	69.88-71.12
METEOR	67.6	0.5	66.98-68.22
Accuracy (%)	98.5	0.19	98.26-98.74
Recall (%)	97.4	0.21	97.14-97.66

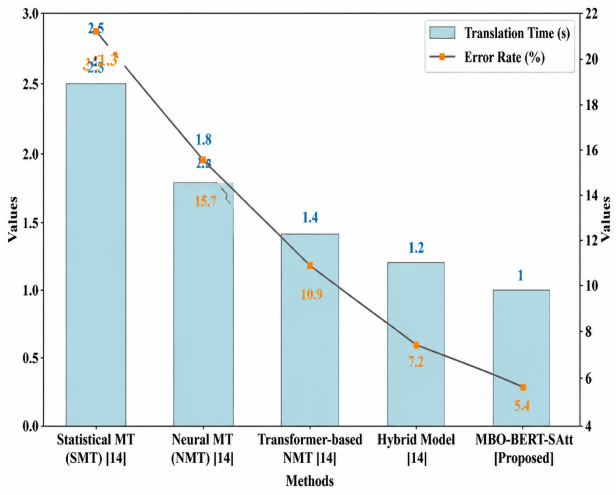

Fig. 6. The Translation Time and Error Rate values for the proposed methods

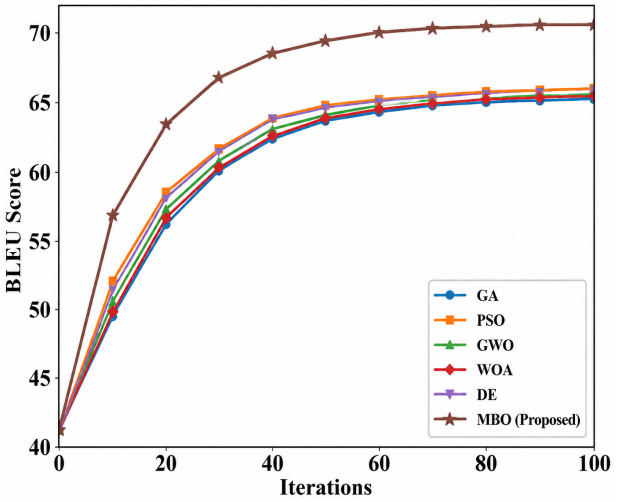

Fig. 7. Convergence Analysis of MBO and Metaheuristics

bined with human-in-the-loop assessment improves reliability by capturing both quantitative performance and qualitative semantic accuracy, enabling robust and business-relevant translation validation.

Fig. 7 compares MBO with GA, PSO, GWO, WOA, and DE, showing faster convergence and higher BLEU scores, demonstrating superior optimization efficiency for translation tasks.

4.1. Discussion

The hybrid ML-DL framework for translations requires a large number of computational resources, and it needs large datasets too. Since the framework is built in such a way that the models may perform poorly on very domain- and context-specific, rarely seen, or confusing texts [14]. MBO-BERT-SAtt enhances rare vocabulary translation through optimized attention despite the computational demands of neural models. BLEU and METEOR assess translation quality, while evaluations on 2,000 English-

Table 6. Statistical Significance Analysis of the Ablation Study

Comparison	BLEU Gain	Accuracy Gain (%)	p-value	Cohen's d	Significant
BERT Only vs. Full Model	12.3	4.2	< 0.001	2.74	Yes
BERT + Self-Attention vs. Full Model	5.7	2.4	0.003	1.59	Yes
BERT + MBO vs. Full Model	8.0	3.1	0.001	1.97	Yes

Table 7. Ablation Study of the Proposed MBO-BERT-SAtt Framework

Model Configuration	BERT	Self-Attention	MBO	Translation Accuracy (%)	Redundancy Reduction (%)
Baseline	×	×	×	89.41	82.17
Baseline + BERT	✓	×	×	92.08	86.34
Baseline + BERT + Self-Attention	✓	✓	×	94.63	90.27
MBO-BERT-SAtt (Proposed)	✓	✓	✓	**97.84**	**95.62**

Table 8. Statistical Significance Analysis of Performance Improvements

Comparison	BLEU Difference	METEOR Difference	p-value	Cohen's d	Significant
Hybrid vs Proposed	7.4	8.7	<0.001	2.31	Yes
GB-EDDQN vs Proposed	3.7	5.3	0.002	1.68	Yes
Transformer vs Proposed	29.3	29.1	<0.001	4.15	Yes

Table 9. MBO-BERT-SAtt considerably improves translation accuracy and efficiency compared to benchmarks

Model	BLEU Score	METEOR Score	Translation Time (s)	Error Rate (%)
Statistical MT (SMT) [14]	32.5	29.8	2.5	21.3
Neural MT (NMT) [14]	48.2	42.5	1.8	15.7
Transformer-based NMT [14]	55.7	50.3	1.4	10.9
Hybrid Model [14]	63.1	58.9	1.2	7.2
MBO-BERT-SAtt [Proposed]	70.5	67.6	1.0	5.4

Table 10. The proposed model outperforms GB-EDDQN in all performance criteria

Methods	Recall (%)	Accuracy (%)	F1-score(%)
GB-EDDQN [15]	96.1	97.2	95.4
MBO-BERT-SAtt [Proposed]	97.4	98.5	96.6

Japanese pairs and WMT17 demonstrate generalization. Future multilingual testing with pretrained models, tokenization, and fine-tuning will enhance robustness, scalability, and adaptability and reduce overfitting risks.

5. Conclusion

To improve the English translation of the model, researchers developed an MAO-BERT-SAtt hybrid ML/DL framework that aims to address the issues faced by existing MT systems when translating text from a specific domain or with ambiguous meanings that do not have many resources available in binary representation. Most current ML-based MT systems, e.g., statistical MT, Neural MT, Transformer-based NMT, Hybrid Model [14], and GB-EDDQN [15], require much greater computational resources and a considerable quantity of large datasets, and frequently yield lower performance rates for words that are rarely seen in the context of an extremely ambiguous or context-sensitive sentence. To provide a solution to these problems, the MBO-BERT-SAtt uses a combination of the BERT attention mechanism together with MBO to provide a stronger level of comprehension of context with much less reliance on immense datasets. Results from an experimental evaluation of the performance of the MBO-BERT-SAtt were excellent. With recall, accuracy, and F1-score values of 97.4%, 98.5%, and 96.6%, respectively, it exceeded that of the baseline models and provided for improved computational efficiencies. MBO-BERT-SAtt requires further adaptation for low-resource and specialized domains. Future improvements focus on real-time translation optimization, support for morphologically rich languages, and adaptive learning integration to handle evolving vocabulary, con-

textual meanings, and complex translation requirements across diverse linguistic environments. Collectively, the MBO-BERT-SAtt has the potential to be an effective, efficient, and cost-effective MT tool that could help fill the voids left by previous MT resources.

6. Acknowledgment

Not Applicable

7. Declarations

8. Funding:

Authors did not receive any funding.

9. Conflicts of interests:

Authors do not have any conflicts.

10. Data availability statement:

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

11. Code availability:

Not Applicable.

12. Authors' contributions:

Qingyue Su, is responsible for designing the framework, analyzing the performance, validating the results, and writing the article.

References

[1] L. He, J. Guo, and J. Lin, (2022) *"Design of Shared Internet of Things System for English Translation Teaching Using Deep Learning Text Classification"* **Wireless Communications and Mobile Computing *2022***: 3576419. DOI: 10.1155/2022/3576419.

[2] L. Su, (2024) *"Study on an Intelligent English Translation Method Using an Improved Convolutional Neural Network Model"* **International Journal of e-Collaboration *20***(1): 1–14. DOI: 10.4018/IJeC.357556.

[3] Y. Yuxiu, (2024) *"Application of Translation Technology Based on AI in Translation Teaching"* **System and Soft Computing *6***: 200072. DOI: 10.1016/j.sasc.2024.200072.

[4] Y. A. Mohamed, A. Khanan, M. Bashir, A. H. H. Mohamed, M. A. Adiel, and M. A. Elsadig, (2024) *"The Impact of Artificial Intelligence on Language Translation: A Review"* **IEEE Access *12***: 25553–25579. DOI: 10.1109/ACCESS.2024.3366802.

[5] P. L. Liu and C. J. Chen, (2023) *"Using an AI-Based Object Detection Translation Application for English Vocabulary Learning"* **Educational Technology & Society *26***(3): 5–20. DOI: 10.30191/ETS.202307_26(3).0002.

[6] X. Liu, J. He, M. Liu, Z. Yin, L. Yin, and W. Zheng, (2023) *"A Scenario-Generic Neural Machine Translation Data Augmentation Method"* **Electronics *12***(10): 2320. DOI: 10.3390/electronics12102320.

[7] A. Ekuerhare and O. P. Udoka, (2024) *"The Effect of Globalization on Language Services and Translation Strategies"* **International Journal of Advanced Multidisciplinary Research Studies *4***(6): 507–516. DOI: 10.62225/2583049X.2024.4.6.3467.

[8] R. Mundlamuri, G. R. Gunnam, N. K. Mysari, and J. Pujuri, (2025) *"The Evolution of AI: From Classical Machine Learning to Modern Large Language Models"* **IEEE Access *13***: DOI: 10.1109/ACCESS.2025.3621344.

[9] B. Zhang, B. Haddow, and A. Birch. "Prompting Large Language Model for Machine Translation: A Case Study". In: *International Conference on Machine Learning*. PMLR, 2023, 41092–41110.

[10] J. Son and B. Kim, (2023) *"Translation Performance from the User's Perspective of Large Language Models and Neural Machine Translation Systems"* **Information *14***(10): 574. DOI: 10.3390/info14100574.

[11] J. Su, J. Chen, H. Jiang, C. Zhou, H. Lin, Y. Ge, Q. Wu, and Y. Lai, (2021) *"Multi-Modal Neural Machine Translation with Deep Semantic Interactions"* **Information Sciences *554***: 47–60. DOI: 10.1016/j.ins.2020.11.024.

[12] F. Ma, (2022) *"Application of Convolutional Neural Network Based on Deep Learning in College English Translation Teaching Management"* **Mobile Information Systems *2022***: 9673133. DOI: 10.1155/2022/9673133.

[13] X. Jiang, (2022) *"Research on the Analysis of Correlation Factors of English Translation Ability Improvement Based on Deep Neural Network"* **Computational Intelligence and Neuroscience *2022***: 9345354. DOI: 10.1155/2022/9345354.

[14] N. Q. Zhao, (2025) *"AI-Driven English Translation: Leveraging Machine Learning and Deep Learning for Enhanced Accuracy"* **International Journal of Information and Communication Technology *26***(18): 1–17. DOI: 10.1504/IJICT.2025.10071429.

[15] Q. Guo, (2025) *"Deep Reinforcement Learning Method for Optimizing English Translation Algorithms"* **International Journal of High Speed Electronics and Systems**: DOI: 10.1142/S0129156425405686.