

A Reinforcement Generative Adversarial Network for Multilingual Neural Machine Translation

Yiou Huang^a, Hui Liu^{b,*}

^aDepartment of Mathematical Sciences, University of Liverpool, Liverpool, United Kingdom

^bSchool of Foreign Languages, Dalian University of Technology, Dalian, China

ARTICLE INFO

Keywords:

Chinese and English machine translation
adaptive context-aware
policy gradient reinforcement

ABSTRACT

Multilingual neural machine translation (MNMT) has achieved remarkable progress in terms of performance across a wide range of benchmarks. However, traditional MNMT methods are largely constrained by a unidirectional decoding strategy, which is either left-to-right or right-to-left. This limitation restricts their capacity to fully comprehend contextual semantics and often results in translation imbalances. To address these challenges, this paper proposes a reinforcement generative adversarial networks for multilingual neural machine translation (RGAN). RGAN integrates a novel generative network designed to optimize the decoding sequence of the conventional transformer architecture. By dynamically adjusting the decoding order, RGAN effectively mitigates output imbalances while enhancing its capability to capture richer contextual information. Embedded within a Generative Adversarial Network (GAN) framework, the generator undergoes continuous refinement through adversarial training, thereby progressively generating higher-quality translations. Moreover, RGAN employs a policy gradient reinforcement learning technique to tackle the inherent limitations associated with processing discrete text data in natural language processing tasks. The proposed framework significantly enhances translation fluency and accuracy in both Chinese and English languages.

1. Introduction

The rapid advancement of information technology and the internet has created unprecedented global opportunities [1, 2, 3, 4]. However, language barriers continue to impede the full realization of globalization. While human translation can help bridge communication gaps, its inefficiency in terms of processing time fails to meet the growing demands for seamless global interaction. In contrast, machine translation offers a more convenient solution, enabling quick language switching and supporting diverse translation scenarios, thereby promoting efficient communication. In recent years, deep learning has breathed new life into natural language processing (NLP) [5, 6, 7]. By integrating deep learning techniques with traditional machine translation methods, researchers have developed neural machine translation (NMT), which has now emerged as the leading approach in the field. NMT has outperformed traditional statistical machine translation models across numerous language pairs, establishing itself as a dominant force in the translation landscape [8, 9].

Current NMT models for Chinese-English translation can be categorized into two main approaches: attention-based NMT [10, 11] and transformer-based NMT [9, 12]. (1) Attention-based NMT: These models leverage attention mechanisms that allow the system to dynamically focus on

different parts of the input sentence during translation. By continuously adjusting the attention based on the relevance of each input word to the word being translated, these models achieve more precise translations. (2) Transformer-based NMT: Relying primarily on self-attention mechanisms, transformer models can capture global dependencies between input and output sequences. Unlike traditional models, transformers do not require recursion and can be processed in parallel, resulting in higher efficiency during both training and inference. Due to their exceptional performance, transformer models have become one of the leading methods in NMT and have set new benchmarks for machine translation success [13, 14, 15, 16]. For example, Ko et al. proposed a method for translating similar low-resource languages without the help of the parallel corpora by fine-tuning pre-trained high-resource translation models [17]. They utilized a weighted combination of losses from various tasks and iteratively trained translation models from English to multiple low-resource languages and vice versa. After a few iterations of training, their models outperformed the results obtained from the combination of mBART and Mass pre-trained models on most languages [18]. Aharoni et al. explored large-scale multilingual translation models [19].

Although NMT methods have achieved remarkable progress, they typically adopt a left-to-right sequential decoding approach to generate target translations. This characteristic of sequential unidirectional decoding confines each output word to the previously generated outputs, thus restricting the decoder's ability to fully consider semantic information on the target side during the translation process. Consequently, this limitation leads to inaccuracies in translation and lack of fluency in the translated text.

DOI: <https://doi.org/10.70891/JAIR.2026.030001>

ISSN of JAIR: 3078-5529

License: CC-BY 4.0, see <https://creativecommons.org/licenses/by/4.0/>

*Corresponding author

sgyhua40@liverpool.ac.uk (Y. Huang); liuhui1126@dlut.edu.cn (H.

Liu)

To address the unidirectional nature of the decoder order of the NMT methods and the output imbalance problem, a generative adversarial network with an adaptive optimisation is proposed for Chinese and English translation (RGAN). Specifically, a new generative network is designed to improve the decoder order of original transformers, which improves the output imbalance problem of the original model and enhances the ability to capture contextual semantic information at the same time. In the generative adversarial network framework, the generative capability of the generative network is continuously strengthened by adversarial training of the generator and discriminator, prompting the generator to generate higher quality translations. Rooted in the reinforcement learning, a policy gradient approach is designed to improve the disadvantages of discrete text data in natural language processing. Finally, experiments show that RGAN has a good effect on the ability to mine semantic knowledge and the fluency of translations.

The main contributions are summarized as follows:

- Traditional multilingual neural machine translation models are primarily constrained by a unidirectional sequential decoding approach, such as left-to-right or right-to-left. To address this, RGAN introduces a novel adaptive decoding order that initiates the prediction from the middle of the sentence. By first generating the right-hand sequence followed by the left-hand side, the model effectively mitigates output imbalances and enhances its capacity to capture richer contextual semantic information.
- RGAN integrates an enhanced Transformer architecture into a generative adversarial network framework to drive continuous performance improvement. The generative network is optimized to produce translations that are virtually identical to human-authored samples, while the convolutional neural network-based discriminator is trained to discern between machine-generated and ground-truth sequences. This adversarial synergy encourages the generator to progressively yield more robust, fluent, and high-quality translations.
- To overcome the inherent limitations of processing discrete text data in natural language processing, RGAN employs a policy gradient reinforcement learning technique. By treating the generative model as an agent, the framework utilizes the discriminator’s output as a reward signal to guide the optimization process. Furthermore, Monte Carlo Tree Search is implemented to evaluate the action values of intermediate states, ensuring the model is optimized for global translation accuracy and fluency.

The subsequent sections of RGAN are conducted as follows: Section II elaborates an in-depth exposition of RGAN. Section III offers a comprehensive depiction of experiment evaluations. Lastly, Section IV concludes RGAN.

2. Methodology

As shown in Fig. 1, a novel adaptive generative adversarial network is proposed for Chinese and English machine translation (RGAN). Specifically, RGAN improves the decoding order of neural machine translation models, transforming the enhanced translation model into a generative network to enhance its ability to capture contextual semantic information. Furthermore, within the framework of generative adversarial networks, the adversarial training between the generator and discriminator continuously strengthens the generation capabilities of the generative network, prompting the generator to produce higher-quality translations. Additionally, leveraging policy gradient methods from reinforcement learning helps mitigate the shortcomings of discrete text data in natural language processing.

2.1. Translation modelling of RGAN

The goal of RGAN is to automatically translate source sentence ($X = \{x_1, x_2, \dots, x_n\}$) into the target sentence ($Y = \{y_1, y_2, \dots, y_m\}$). RGAN predicts the intermediate translation $Z_k = (y_{m-k+1}, \dots, y_m, [m], y_1, \dots, y_{m-k})$ starting from the middle of the sentence, and then utilize Z_k to construct the target translation Y , where $[m]$ is the m -th word of Z_k . Specifically, RGAN firstly predicts a word y_{m-k+1} , and then predicts the right-hand side of the sentence (y_{m-k+1}, \dots, y_m) until the symbol $[m]$ is encountered, at which point it switches to predicting the left-hand side of the sentence (y_1, \dots, y_{m-k}). Mathematically, having:

$$\begin{aligned}
 P_{\theta}(Z_k|X) &= P_{\theta}(y_{m-k+1}|X) \\
 &\times \prod_{m-k+1 < i < n}^T P_{\theta}(y_i|X; y_k, \dots, y_{i-1}) \\
 &\times P_{\theta}([M]|X; y_{m-k+1}, \dots, y_m) \\
 &\times \prod_{1 < j < m-k}^T P_{\theta}(y_j|X; y_1, \dots, y_{j-1}, y_{m-k+1}, \dots, y_m)
 \end{aligned} \tag{1}$$

where i and j stand for the i -th and the j -th words of Y , respectively.

2.2. Generative Adversarial Network

RGAN mainly contains the generative and discriminative networks, both engaged in adversarial training. The generative network, based on real sample sets, aims to enhance the similarity between generated samples and real samples, while the discriminative network’s purpose is to distinguish whether the translation results come from generated samples or real samples, and provide feedback to the generative network to guide its learning and optimization.

Generative Network: In RGAN, the generative Network G is utilized to obtain target sentences, which are virtually identical to human translations given the source sentence. To boost translation efficiency, RGAN adopts the Transformer architecture. This model not only improves translation speed but also enhances the performance of machine translation.

Due to the non-sequential input network structure being employed, word positions need to be annotated before word

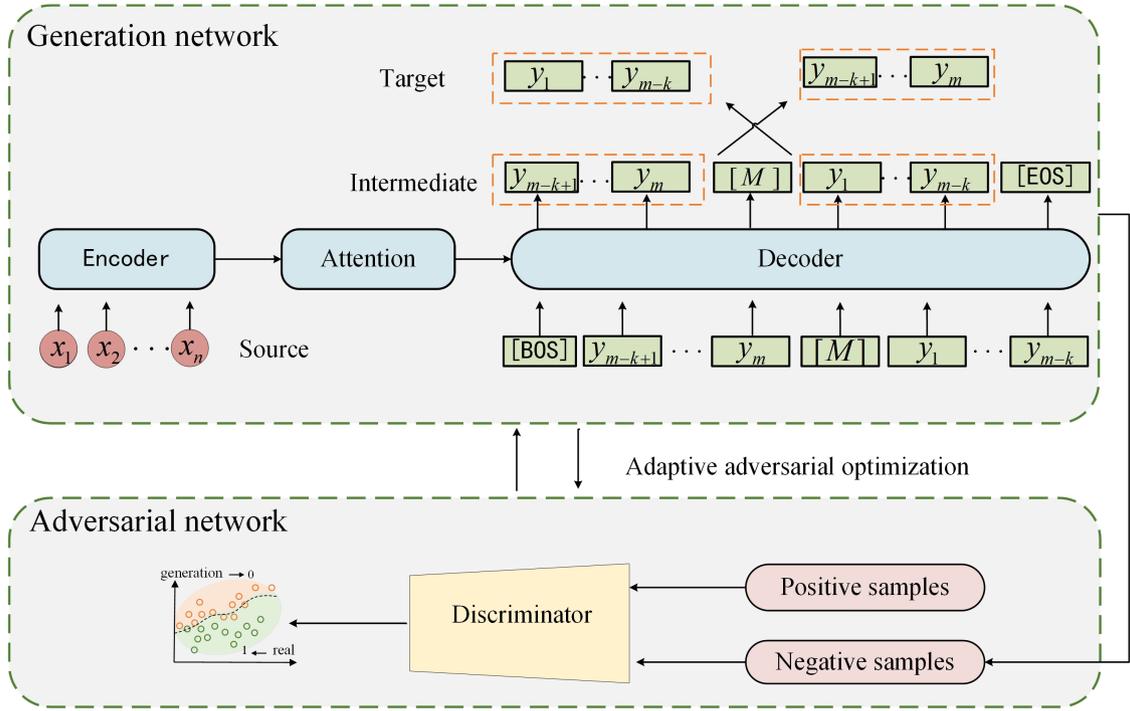


Figure 1: The illustration of RGAN. RGAN contains the generative network, the adversarial network and policy gradient reinforcement learning.

vectors enter the model:

$$\begin{aligned} PE_{pos,2i} &= \sin(pos/10000^{2i/d_{\text{mod}}}) \\ PE_{pos,2i+1} &= \cos(pos/10000^{2i/d_{\text{mod}}}) \end{aligned} \quad (2)$$

where d_{mod} denotes the length of word vectors. After obtaining word embeddings with positional information. Then, these word embeddings with positional information are fed into self-attention mechanism to obtain self-attention scores for each word:

$$Attention(q_i, v_i, k_i) = \text{softmax}\left(\frac{q_i k_i^T}{\sqrt{d_k}}\right) v_i \quad (3)$$

where q_i , v_i , and k_i denote the query vector, the key vector, and the value vector of the i -th word x_i . d_k stands for the dimension of k_i

$$q_i = z_i w^Q, \quad k_i = z_i w^K, \quad v_i = z_i w^V \quad (4)$$

where w^Q , w^K , and w^V denote self-attention parameters. z_i denotes the embedding feature of x_i obtained via the encoder of the transformer.

The Transformer completely discards the architectures based on recurrent and convolutional networks. The self-attention mechanism better captures global information and significantly enhances training speed through parallel processing. However, the Transformer’s output, which employs weighted averaging, leads to attention dispersion, neglecting the relationships between adjacent words. To this end, improvements are made to the decoding sequence of the Transformer to improve the model’s performance to capture

dependencies in short sentences. The enhanced Transformer model is then utilized as a new generative model, trained adversarially alongside a discriminator network. This process aims to boost the robustness of neural machine translation models.

Adversarial Network: The discriminator architecture is based on Convolutional Neural Networks. Specifically, the discriminator D strives to discern generated and human-translated sentences, contingent upon the source language sentence. Given that D optimizes concurrently with the G , it is commonly viewed as a fluid target. Considering the varying lengths of sequences obtained via the generative network, the length of output sequences are initially elongated as T . Afterward, matrices X and Y are formulated rooted in the source language and the target language, correspondingly:

$$\begin{cases} X_{1:T} = x_1; x_2, \dots; x_T \\ Y_{1:T} = y_1; y_2, \dots; y_T \end{cases} \quad (5)$$

Then, we obtain features C_{ji} of $X_{1:T}$ via the convolutional kernel with a length of L .

$$C_{ji} = f(W_j \otimes X_{i:i+L-1} + b) \quad (6)$$

in which W_j stands for parameters in the convolutional kernel. f stands for the activation function. To fully capture information in the sequence, we utilize kernels with various window sizes to extract information where the maximum of each representation is concatenated to conduct the representation C_x of the source languages. Maxpooling is applied to

the feature maps of each convolutional kernel to produce the final representations.

$$C_x = \max\{c_{j1}, c_{j2}, \dots, c_{jT-l+1}\} \quad (7)$$

Similarly, the representation C_y for the target sequence in the same manner. Based on C_x and C_y , we compute the probability of the target sentence belonging to a real sample:

$$P = \Phi(V[C_x; C_y]) \quad (8)$$

where V is the transformation function used to concatenate C_x and C_y .

2.3. Adaptive Optimization of RGAN

GAN is an unsupervised learning approach where parameter updates are based on a continuous space. However, natural language processing data is discrete. Hence, the concept of policy gradients from reinforcement learning is introduced to combine reinforcement learning with GANs. Specifically, the generative model is regarded as the agent in reinforcement learning, while the discriminator evaluates sequences and provides feedback to guide the learning of the generative model. Additionally, parameter updates in the model are facilitated using gradient-based algorithms. In the gradient policy of the model, Monte Carlo tree search is employed to enable real-time evaluation of incomplete sequences.

$$J(\theta) = \sum_{y_{1:T}} G_\theta(y_{1:T}|x) \cdot R_D^{G_\theta}(y_{1:T-1}, x, y_T) \quad (9)$$

where $y_{1:T}$ and x stand for the generation and the source sentences, respectively. $R_D^{G_\theta}$ denotes the action-value function in reinforcement learning. We use the discriminator's output, which represents the probability of being human-labeled, to denote the value of $R_D^{G_\theta}$:

$$R_D^{G_\theta}(y_{1:T-1}, x, y_T) = D(x, y_{1:T}) - b(x, y_{1:T}) \quad (10)$$

in which $b(x, y_{1:T})$ is a deviation term.

However, the discriminator can only provide reward values for complete sequences, and there is no clear interpretation for the results obtained from $D(x, y_{1:T})$, making it impossible to evaluate the action values of intermediate states. Therefore, it is necessary to rely on policy gradient algorithms G_θ . When the generated sequence is not complete, Monte Carlo search is used to generate subsequent complete word sequences. The search ends when the sentence is sampled to its end or when the generation of the complete sequence from the initial state (the sequence generated at the previous $T - 1$ time step) reaches the maximum length set. To reduce the variance of the reward signal, N Monte Carlo searches are performed:

$$\{y_{1:T_1}^1, \dots, y_{1:T_N}^N\} = MC^{G_\theta}((y_{1:t}, x), N) \quad (11)$$

in which $\{y_{1:T_1}^1, \dots, y_{1:T_N}^N\}$ denote N possible word sequences generated according to the policy G_θ . The final feedback function for the entire sentence sequence is:

$$R_D^{G_\theta} = \begin{cases} D(x, y_{1:T}) - b(x, y_{1:T}) & t = T \\ \frac{1}{N} \sum_{n=1}^N D(x, y_{1:T_n}^n) - b(x, y_{1:T_n}^n) & t < T \end{cases} \quad (12)$$

During adversarial training, the discriminator is utilized as a reward function to incentivize the generator's learning, thereby further enhancing the generator's performance. Once the generator receives updates, the model retrains the discriminator, aiming to minimize its loss function.

$$E_{x,y \in P_{data}}[\log D(x, y)] - E_{x,y \in G_\theta}[\log(1 - D(x, y))] \quad (13)$$

The parameter optimization of the generative network is:

$$\nabla J(\theta) = \sum_{t=1}^T R_D^{G_\theta}(y_{1:t-1}, x, y_t) \cdot \nabla_\theta \log(G_\theta(y_{1:t-1}|x)) \quad (14)$$

In the adversarial training process, the generative and discriminative networks are alternately updated. G is updated by leveraging policy gradient methods from reinforcement learning and using the rewards obtained from Monte Carlo search based on the discriminative network D , aiming to confuse the discriminator by generating sentences close to human translations. The discriminative network D is trained jointly on real human-translated data and data generated by the generative network, enhancing its discriminative ability and thus achieving the adversarial effect.

In summary, this paper proposes RGAN, a comprehensive reinforcement generative adversarial framework tailored for multilingual neural machine translation. The architecture redefines the decoding paradigm by transitioning from a traditional unidirectional sequence to an adaptive middle-out order, thereby enabling more balanced and contextually aware translation generation. By integrating this enhanced transformer-based generator within a GAN framework, the model benefits from the global semantic evaluation provided by the discriminator. The implementation of policy gradient reinforcement learning further bridges the gap between discrete token prediction and continuous sequence optimization. Collectively, these components form a robust translation system capable of producing high-fidelity outputs across diverse language pairs. The primary advantages of the proposed RGAN are summarized as follows: (1) The adaptive middle-out decoding strategy breaks the constraints of sequential order, allowing the model to effectively leverage bi-directional semantic dependencies (2) Through the adversarial game between the generator and the discriminator, the model shifts from local token-level matching to global sequence fluency, resulting in more human-like translations. (3) The application of policy gradients and Monte Carlo Tree Search successfully addresses the non-differentiability of text data, ensuring more stable training and superior convergence in complex translation tasks.

Table 1

The model parameters across different datasets.

Hyper-parameters	NIST02	NIST03	NIST06
Encoder Layers	6	6	6
Decoder Layers	6	6	6
Embedding Size	512	512	512
Learning Rate	0.001	0.005	0.001
Dropout	0.2	0.3	0.4

3. Results and discussion

3.1. Setup

Dataset: A general Chinese-English translation dataset [20, 21] from the LDC corpus is used to verify the performance of RGAN, where 10,000 sentence pairs are considered as the training data, NIST02, NIST03, and NIST06 datasets are considered as the test data, and the NIST 2005 dataset is used as the validation data chosen. For monolingual data, 50,000 Chinese monolingual data and 50,000 English monolingual data are extracted from the Gigaword corpus.

Metric: Consistent with prior researches [22], the performance is evaluated via Bilingual Evaluation Understudy (BLEU) that calculates overlapping n-gram number to evaluate its fluency and accuracy of the translation. BLEU score belongs to $[0, 1]$, where a value closer to 1 indicates greater.

Implementation Details: In RGAN, the hardware used includes an Intel Core i7-9700 CPU with 16GB of RAM. The software operates on Windows 10 Professional edition, programmed in Python 3.5 using the PyTorch framework. For the generative model, this chapter adopts the Transformer model architecture, with the main parameters shown in Table 1. The learning rate is set to 0.005. The Adam is leveraged to optimize RGAN, setting the word embedding dimension to 512. To prevent overfitting during training, a dropout mechanism is introduced and set to 0.3. The parameters trained with Maximum Likelihood Estimation (MLE) are used to initialize the adversarial network model. BLEU score is employed as the evaluation metric for machine translation results. All experimental models are implemented in PyTorch.

3.2. Comparison with baselines

Comparison baselines: Seven baselines are compared with RGAN on three datasets, including RNNSearch [8], RNNSEARCH, kNN-TL [9], Transformer, DynamicConv, SKD-NANMT[23], BNMT [12].

Comparison results: Table 2 presents a comparison to prior works on Chinese-English translation dataset in terms of BLEU. Specifically, RGAN outperforms the other methods in all the NIST dataset (NIST02, NIST03, NIST06) as well as in the average score. RGAN achieves the highest BLEU scores, with 0.7329, 0.7023, and 0.6969 for NIST02, NIST03, and NIST06, respectively, and an overall average score of 0.7107. This significant improvement over previous methods suggests that the RGAN model is more effective

Table 2

Comparison of RGAN to prior works on Chinese-English translation dataset in terms of BLEU.

Methods	NIST02	NIST03	NIST06	Average
RNNSearch	0.3576	0.3393	0.3324	0.3431
RNNGAN	0.3901	0.3766	0.3672	0.3779
Transformer	0.4063	0.3895	0.3862	0.3940
DynamicConv	0.5260	0.4907	0.4965	0.5044
kNN-TL	0.6582	0.6015	0.5901	0.6166
SKD-NANMT	0.6912	0.6788	0.6907	0.6869
BNMT	0.6123	0.5825	0.6000	0.5983
RGAN	0.7329	0.7023	0.6969	0.7107

in capturing the nuances of language and generating translations that are closer to human quality. The other methods listed in the table, such as RNNSearch, RNNSearch+GAN, Transformer, DynamicConv, kNN-TL, SKD-NANMT, and BNMT, all show relatively lower BLEU scores compared to RGAN. Notably, the baseline RNNSearch performs the least effectively with scores of 0.3576, 0.3393, and 0.3324 for NIST02, NIST03, and NIST06, respectively. The addition of GAN to RNNSearch (RNNSearch+GAN) improves the scores, but they still lag behind the other methods. The Transformer model, which has been a significant advancement in NMT, shows competitive scores, but it still falls short compared to the RGAN. DynamicConv and kNN-TL also demonstrate good performance, but they do not surpass the RGAN in any of the years or the average score. The SKD-NANMT and BNMT models, which are other neural machine translation approaches, show strong results, particularly SKD-NANMT with a score of 0.6912 in NIST02, but again, they do not match the performance of RGAN. The superior performance of RGAN can be attributed to its advanced architecture and training methodology. The combination of adversarial training with generative adversarial networks (GANs) and the continuous evaluation metric training (CEMT) likely contributes to the model’s ability to generate more fluent and accurate translations. The adversarial training contributes to the model learn to obtain translations that are more diverse and robust, while the continuous evaluation metric training ensures that the model is optimized specifically for the BLEU score, which is the metric used in this comparison.

In conclusion, the RGAN model demonstrates a clear advantage over prior works in terms of BLEU scores on the Chinese-English translation dataset. This indicates that the RGAN is a highly effective method for machine translation tasks and has the potential to become a new SOTA in NMT. Further research and development of models like RGAN could lead to even greater improvements in machine translation quality and more effective bridging of language barriers.

3.3. Ablation Study

This section conducts two ablation experiments about RGAN on all the NIST dataset (NIST02, NIST03, NIST06).

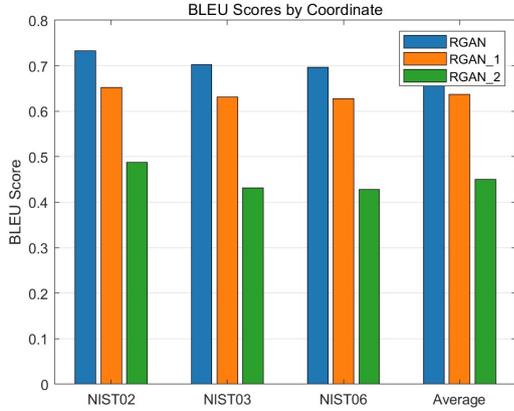


Figure 2: Ablation experiments of each component in RGAN.

Specifically, RGAN_1 uses the tradition transformer as the generation network RGAN to achieve chinese-english translation. RGAN_2 only uses improvement transformer framework to achieve chinese-english translation without the generative adversarial mechanism. As illustrated in Figure 2, the following three observations can be drawn:

Observation 1: Synergy between Architecture and Training Paradigm. The full RGAN model consistently outperforms both RGAN_1 and RGAN_2 across all test sets. This confirms that the integration of the adaptive middle-out decoding strategy and the generative adversarial framework creates a synergistic effect, where the structural advantages of the decoder are fully exploited through adversarial optimization.

Observation 2: Essential Role of the Generative Adversarial Mechanism. A comparison between RGAN and RGAN_2 reveals that the absence of the adversarial training process leads to a noticeable decline in translation quality. This suggests that the discriminator provides critical global semantic feedback, enabling the generator to produce more fluent and human-like translations that exceed the capabilities of traditional maximum likelihood estimation.

Observation 3: Effectiveness of the Adaptive Decoding Strategy. RGAN_2 demonstrates superior performance compared to RGAN_1, indicating that the middle-out decoding order is more effective than the traditional unidirectional approach. By initiating prediction from the center of the sequence, the model effectively captures richer multi-directional contextual information and mitigates the inherent output imbalance found in standard sequential models.

3.4. Dimensionality Analysis of embeddings

The dimensionality of word embeddings significantly affects the translation quality, as reflected by the BLEU scores in Table 3. As the embedding dimension increases from 64 to 512, the BLEU scores improve steadily across all NIST datasets, indicating that higher-dimensional embeddings capture richer semantic information, leading to better translation performance. However, beyond 512 dimensions,

Table 3

Dimensionality analysis of embeddings across different datasets.

Dimension	NIST02	NIST03	NIST06
64	0.7115	0.6845	0.6711
128	0.7125	0.6952	0.6774
256	0.7299	0.7000	0.6899
512	0.7329	0.7023	0.6969
1024	0.7312	0.6985	0.6912
2048	0.6985	0.6774	0.6745

Table 4

Dimensionality analysis of embeddings across different layers.

Layers	NIST02	NIST03	NIST06
2	0.6175	0.6284	0.6078
3	0.6754	0.6528	0.6447
4	0.6929	0.6733	0.6678
5	0.7100	0.6985	0.6912
6	0.7329	0.7023	0.6969
7	0.7297	0.6993	0.6965

the performance either stabilizes or slightly declines, suggesting diminishing returns and potential overfitting with excessively high dimensions, such as 1024 and 2048. This indicates an optimal embedding dimension range around 512 for balancing translation quality and computational efficiency.

The number of layers in the encoder and decoder has a substantial impact on translation performance, as evidenced by the BLEU scores in Table 4. Initially, increasing the number of layers from 2 to 6 results in a consistent improvement across all NIST datasets, indicating that deeper models can capture more complex linguistic patterns and dependencies, thereby enhancing translation quality. However, after reaching 6 layers, the performance starts to plateau or slightly decrease, particularly for NIST02 and NIST03, suggesting that additional layers beyond this point may introduce redundancy or overfitting without significant gains in translation accuracy. This highlights that 5 to 6 layers might be the optimal configuration for balancing performance and model complexity in this setting.

3.5. Discussion

The experimental results and ablation studies underscore the efficacy of RGAN in multilingual neural machine translation. The primary advantage of the proposed framework lies in the synergy between the adaptive middle-out decoding strategy and the adversarial training paradigm. Unlike traditional unidirectional models, our approach enables the decoder to capture bi-directional contextual dependencies more effectively, thereby mitigating the information bottleneck and output imbalance inherent in standard sequential generation. Furthermore, the integration of policy gradient reinforcement learning allows the model to optimize discrete text sequences directly, ensuring that the generated

translations are not only grammatically correct but also semantically fluent and closer to human-level expression.

Despite its superior performance, several limitations of the current RGAN framework warrant further consideration. First, the introduction of the GAN-based architecture and Monte Carlo Tree Search (MCTS) significantly increases the computational complexity during the training phase. Compared to the vanilla Transformer, RGAN requires a larger memory footprint and more extensive training epochs to achieve convergence, which may limit its deployment in resource-constrained environments. Second, the model exhibits a degree of sensitivity to hyper-parameter configurations, such as embedding dimensions and the number of layers. As observed in the parameter analysis, an improperly scaled architecture can lead to either underfitting or a performance plateau, necessitating careful empirical tuning for different language pairs.

Moving forward, several promising directions for future research emerge from this study. To address the computational overhead, we aim to explore lightweight GAN variants and more efficient reward estimation techniques to accelerate the training process without compromising translation quality. Additionally, extending the RGAN framework to ultra-low-resource language pairs could further validate its robustness and generalization capabilities in more challenging translation scenarios. We also plan to investigate the integration of pre-trained language models within the RGAN generator to leverage large-scale unlabeled data, potentially pushing the boundaries of multilingual neural machine translation even further.

4. Conclusion

A generative adversarial network with adaptive optimization approach for neural machine translation is proposed in this paper. The improvement lies primarily in enhancing the decoding sequence and integrating adversarial network optimization into the neural machine translation model. Firstly, regarding the enhancement of the decoding sequence, the original neural machine translation model employs a unidirectional decoding process from left to right. However, this decoding order restricts the parallelism of the model algorithm and limits its capability to capture contextual information, resulting in output imbalance issues. In the improved decoder, predictions start from the middle word, followed by the generation of words on the right, and finally, words are generated on the left. This improvement greatly enhances the robustness to capture contextual knowledge. Another optimization lies in using the improved Transformer model architecture as the new generative network and training it together with the discriminator network through adversarial training to enhance machine translation performance. Additionally, during adversarial training, the policy gradient algorithm from reinforcement framework is introduced to avoid the drawbacks of discrete textual data in natural language processing.

References

- [1] S. Gui, C. Shao, Z. Ma, X. Zhang, Y. Chen, Y. Feng, Non-autoregressive machine translation with probabilistic context-free grammar, Proceedings of the Annual Conference on Neural Information Processing Systems, New Orleans, LA, USA, 2023.
- [2] S. Conia, D. Lee, M. Li, U. F. Minhas, Y. Li, Enhancing machine translation experiences with multilingual knowledge graphs, Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press, Vancouver, Canada, 2024, pp. 23781–23783.
- [3] P. Li, J. Gao, J. Zhang, S. Jin, Z. Chen, Deep reinforcement clustering, IEEE Transactions on Multimedia, 2023, vol. 25, pp. 8183–8193.
- [4] J. Gao, M. Liu, P. Li, J. Zhang, Z. Chen, Deep multiview adaptive clustering with semantic invariance, IEEE Transactions on Neural Networks and Learning Systems, 2024, vol. 35, no. 9, pp. 12965–12978.
- [5] J. Gao, M. Liu, P. Li, A. A. Laghari, A. R. Javed, N. Victor, T. R. Gadekallu, Deep incomplete multiview clustering via information bottleneck for pattern mining of data in extreme-environment IoT, IEEE Internet Things Journal, 2024, vol. 11, no. 16, pp. 26700–26712.
- [6] P. Li, A. A. Laghari, M. Rashid, J. Gao, T. R. Gadekallu, A. R. Javed, S. Yin, A deep multimodal adversarial cycle-consistent network for smart enterprise system, IEEE Transactions on Industrial Informatics, 2023, vol. 19, no. 1, pp. 693–702.
- [7] J. Gao, P. Li, A. A. Laghari, G. Srivastava, T. R. Gadekallu, S. Abbas, J. Zhang, Incomplete multiview clustering via semidiscrete optimal transport for multimedia data mining in IoT, ACM Transactions on Multimedia Computing, Communications and Applications, 2024, vol. 20, no. 6, pp. 158:1–158:20.
- [8] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 2015.
- [9] S. Liu, X. Liu, D. F. Wong, Z. Li, W. Jiao, L. S. Chao, M. Zhang, kNN-TL: k-nearest-neighbor transfer learning for low-resource neural machine translation, Proceedings of the Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Toronto, Canada, 2023, pp. 1878–1891.
- [10] B. Zhang, D. Xiong, J. Xie, J. Su, Neural machine translation with GRU-gated attention model, IEEE Transactions on Neural Networks and Learning Systems, 2020, vol. 31, no. 11, pp. 4688–4698.
- [11] W. He, Y. Wu, X. Li, Attention mechanism for neural machine translation: A survey, Proceedings of the IEEE Information Technology, Networking, Electronic and Automation Control Conference, IEEE, Xi’an, China, 2021, pp. 1485–1489.
- [12] Y. Zhang, A. Garg, Y. Cao, L. Lew, B. Ghorbani, Z. Zhang, O. Firat, Binarized neural machine translation, Proceedings of the Annual Conference on Neural Information Processing Systems, New Orleans, LA, USA, 2023.
- [13] D. Dong, H. Wu, W. He, D. Yu, H. Wang, Multi-task learning for multiple language translation, Proceedings of the Annual Meeting of the Association for Computational Linguistics, The Association for Computer Linguistics, Beijing, China, 2015, pp. 1723–1732.
- [14] S. Wu, J. Yu, J. Chen, X. Deng, W. Zhou, Improving open-domain dialogue response generation with multi-source multilingual commonsense knowledge, Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press, Vancouver, Canada, 2024, pp. 19252–19260.
- [15] B. Zoph, K. Knight, Multi-source neural translation, Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, The Association for Computational Linguistics, San Diego, CA, USA, 2016, pp. 30–34.
- [16] M. Johnson, M. Schuster, Q. V. Le, M. Krikun, Y. Wu, Z. Chen, N. Thorat, F. B. Viégas, M. Wattenberg, G. Corrado, M. Hughes, J. Dean, Google’s multilingual neural machine translation system: Enabling zero-shot translation, Transactions of the Association for Computational Linguistics, 2017, vol. 5, pp. 339–351.
- [17] W. Ko, A. El-Kishky, A. Renduchintala, V. Chaudhary, N. Goyal, F. Guzmán, P. Fung, P. Koehn, M. T. Diab, Adapting high-resource

- NMT models to translate low-resource related languages without parallel data, Proceedings of the Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Virtual Event, 2021, pp. 802–812.
- [18] K. Song, X. Tan, T. Qin, J. Lu, T. Liu, MASS: Masked sequence to sequence pre-training for language generation, Proceedings of the International Conference on Machine Learning, PMLR, Long Beach, CA, USA, 2019, pp. 5926–5936.
- [19] R. Aharoni, M. Johnson, O. Firat, Massively multilingual neural machine translation, Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Minneapolis, MN, USA, 2019, pp. 3874–3884.
- [20] B. Zhang, P. Williams, I. Titov, R. Sennrich, Improving massively multilingual neural machine translation and zero-shot translation, Proceedings of the Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Virtual Event, 2020, pp. 1628–1639.
- [21] C. Zhu, H. Yu, S. Cheng, W. Luo, Language-aware interlingua for multilingual neural machine translation, Proceedings of the Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Virtual Event, 2020, pp. 1650–1655.
- [22] L. Chen, S. Ma, D. Zhang, F. Wei, B. Chang, On the pareto front of multilingual neural machine translation, Proceedings of the Annual Conference on Neural Information Processing Systems, New Orleans, LA, USA, 2023.
- [23] M. Liu, Y. Bao, C. Zhao, S. Huang, Selective knowledge distillation for non-autoregressive neural machine translation, Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press, Washington, DC, USA, 2023, pp. 13246–13254.