

Xiaoyu-MT: Kunming University of Science and Technology at WMT 2026 Chinese-Southeast Asian Multilingual MT Task

Zhenhan Wang¹ Fengzhao Fan¹ Kaiwen Tan¹
Zhengtao Yu¹ Shengxiang Gao¹ Cunli Mao¹
Siqi Zhang¹ Shuting Jiang¹ Yuxin Huang^{1*}

¹Kunming University of Science and Technology

20250014@kust.edu.cn, 2806494388@qq.com, kwtan@kust.edu.cn

ztyu@hotmail.com, gaoshengxiang.yn@foxmail.com, maocunli@163.com

zhangsiqi.mo@foxmail.com, shutting_jiang22@163.com, huangyuxin2004@163.com

Correspondence: huangyuxin2004@163.com

Abstract

This paper presents Xiaoyu-MT, our submission system for the WMT 2026 Chinese-Southeast Asian Multilingual MT Task, focusing on bidirectional translation between Chinese and seven Southeast Asian languages. To address limited parallel resources and imbalanced multilingual data distributions, we select Gemma-4-12B-it as the base model and construct multilingual training data by combining official WMT resources with additional monolingual and bilingual corpora. Xiaoyu-MT adopts a multi-stage training pipeline consisting of Continual Pre-Training (CPT), Supervised Fine-Tuning (SFT), and Direct Preference Optimization (DPO). During SFT, we introduce a language-aware Prompt encoding mechanism to explicitly model source language, target language, and their directed translation relationships through structured Prompt representations and Prompt compression. DPO further optimizes translation preferences using preference data constructed from reference and model-generated translations. Official evaluation results on WMT 2026 show that Xiaoyu-MT achieves an overall BLEU score of 34.73, a COMET score of 81.87, and a Final Result score of 61.80, demonstrating its effectiveness for multilingual and low-resource machine translation scenarios.

1 Introduction

This paper presents Xiaoyu-MT, our machine translation system submitted to the WMT 2026 Chinese-Southeast Asian Multilingual Machine Translation Task. The task focuses on bidirectional translation between Chinese and seven Southeast Asian languages, namely Thai, Vietnamese, Lao, Burmese, Khmer, Indonesian, and Malay. Due to the limited availability of parallel resources, substantial linguistic differences, and imbalanced data distributions across languages and translation directions,

the task poses considerable challenges to the language modeling and cross-directional adaptation capabilities of multilingual machine translation systems.

In recent years, the rapid development of Large Language Models (LLMs) has created new opportunities for multilingual machine translation. Benefiting from the language understanding and cross-lingual generalization capabilities acquired through large-scale pre-training, LLMs have demonstrated strong performance on multilingual machine translation tasks (Cui et al., 2025; Zhu et al., 2024; Bu et al., 2024; Hudi et al., 2024). Meanwhile, research on LLMs for Southeast Asian languages has also continued to advance. For example, SeaLLMs 3 (Zhang et al., 2025) is specifically optimized for Southeast Asian languages and has achieved competitive performance on tasks including translation, knowledge, reasoning, and instruction following. Existing studies have shown that, in low-resource translation scenarios, how to effectively utilize limited linguistic resources and improve the multilingual generalization capability of models remains an important challenge in multilingual machine translation (Wang et al., 2021; Qin et al., 2025).

To address the limited availability of resources for the target languages and the insufficient adaptation of models to multilingual environments, CPT has been shown to be an effective approach for enhancing the language modeling capabilities of pre-trained language models. Zheng et al. (Zheng et al., 2024) explored cross-lingual continual pre-training to extend pre-trained LLMs to new languages, while LLaMAX (Lu et al., 2024) further enhanced the multilingual translation capabilities of LLaMA-family models through large-scale multilingual continual pre-training. These studies demonstrate that CPT can effectively strengthen a model’s ability to model target languages and their linguistic characteristics, thereby providing a stronger linguistic foundation for multilingual task adaptation.

*Corresponding author.

During the SFT stage, existing studies have explored various approaches to adapting multilingual LLMs from perspectives including data construction, representation learning, and parameter-efficient fine-tuning. Multilingual instruction-tuning datasets, such as the Aya Dataset (Singh et al., 2024), enhance multilingual capabilities by providing high-quality supervision across a wide range of languages. Representation-based approaches explicitly establish cross-lingual connections in the latent space to improve multilingual modeling (Ye et al., 2025). In addition, parameter-efficient methods such as LoRA (Hu et al., 2022), Adapter (Houlsby et al., 2019), and Prompt Tuning (Lester et al., 2021) provide effective approaches for adapting language models to downstream tasks. These studies offer different technical approaches for multilingual LLM adaptation. However, in scenarios involving multiple languages and bidirectional translation directions, how to provide models with explicit language and directional conditions remains an open issue worth further investigation.

Preference optimization has also recently demonstrated the potential of DPO for machine translation. Vajda et al. (Vajda et al., 2025) investigated the application of DPO to Slovenian machine translation and demonstrated its effectiveness in improving the translation quality of the GaMS-9B-Instruct model. Compared with conventional reinforcement learning methods that rely on a separate reward model, DPO directly optimizes the model’s generation distribution using preference data, providing a simple and effective approach for further improving translation quality.

Overall, to address the challenges of low-resource conditions, substantial linguistic differences, and varying adaptation requirements across translation directions in bidirectional translation between Chinese and seven Southeast Asian languages, we develop Xiaoyu-MT, a multi-stage multilingual machine translation system based on Gemma-4-12B-it. Xiaoyu-MT combines data construction and augmentation, Continual Pre-Training, Supervised Fine-Tuning, and Direct Preference Optimization to progressively improve the model’s language modeling capability for the task-related languages and its adaptability to multilingual translation. Furthermore, during the SFT stage, Xiaoyu-MT introduces a language-aware Prompt encoding mechanism that explicitly models the input language, target language, and their

directed translation relationship, providing more explicit conditional information for different translation directions. By combining multi-stage training with language-aware conditional modeling, Xiaoyu-MT aims to fully exploit the multilingual capabilities of LLMs while further enhancing task adaptability and translation generation quality in low-resource Southeast Asian languages and multilingual bidirectional translation scenarios.

2 Preliminary Experiment

To select the base model for Xiaoyu-MT, we first conduct a preliminary evaluation of several open-source large language models on bidirectional translation between Chinese and Southeast Asian languages. Considering their zero-shot translation performance, multilingual modeling ability, instruction-following capability, and inference efficiency, we ultimately select Gemma-4-12B-it as the foundation model for Xiaoyu-MT.

2.1 Translation Capabilities of Different Models

We evaluate the zero-shot translation capabilities of different open-source large language models on the devtest set of FLORES+¹. Four representative open-source models are selected for evaluation: Llama-3.1-8B-Instruct², Gemma-3-12B-it³, Gemma-4-12B-it⁴, and Qwen3.5-9B⁵. We evaluate their bidirectional translation performance between Chinese and seven Southeast Asian languages, including Thai, Vietnamese, Lao, Burmese, Khmer, Indonesian, and Malay. For each translation direction, we randomly sample 100 instances for evaluation. We use sacreBLEU and COMET-22⁶ as the primary evaluation metrics, which measure lexical similarity between the generated and reference translations and their semantic consistency, respectively. The zero-shot evaluation results are presented in Table 1.

It should be noted that this zero-shot experiment is intended to provide a preliminary comparison of the zero-shot multilingual translation capabilities of different base models and to inform the subsequent model selection. Since the FLORES+ devtest

¹https://huggingface.co/datasets/openlanguagedata/flores_p/us

²<https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct>

³<https://huggingface.co/google/gemma-3-12b-it>

⁴<https://huggingface.co/google/gemma-4-12B-it>

⁵<https://huggingface.co/Qwen/Qwen3.5-9B>

⁶<https://huggingface.co/Unbabel/wmt22-comet-da>

Models	sacreBLEU/COMET-22							
	id-zh	km-zh	lo-zh	ms-zh	my-zh	th-zh	vi-zh	Avg.
Gemma-4-12B-it	43.27/87.70	32.22/85.36	34.20/85.61	40.43/87.63	30.88/84.19	39.91/88.15	41.56/87.99	37.64/86.66
Gemma-3-12B-it	42.69/86.99	31.98/84.19	34.37/84.72	38.68/87.25	28.21/ 84.80	40.67/87.57	40.05/87.70	36.66/86.10
Llama-3.1-8B-Instruct	34.93/84.72	20.02/78.77	9.08/68.26	32.18/84.18	17.93/77.94	26.43/81.91	33.40/84.48	24.85/80.84
Qwen3.5-9B	41.36/86.88	28.34/84.74	32.00/84.50	40.05/86.27	13.16/72.28	31.96/86.43	40.74/87.55	32.52/84.09

Models	sacreBLEU/COMET-22							
	zh-id	zh-km	zh-lo	zh-ms	zh-my	zh-th	zh-vi	Avg.
Gemma-4-12B-it	26.63/ 88.47	19.73/80.68	24.67/81.34	19.96/86.82	18.87/85.83	38.64/87.57	33.39/88.05	25.98/85.54
Gemma-3-12B-it	27.91/88.26	13.26/74.90	18.59/77.36	17.51/86.64	6.32/75.74	37.31/87.12	31.39/87.82	21.76/82.55
Llama-3.1-8B-Instruct	19.07/84.12	4.10/52.93	1.24/33.27	15.68/83.20	4.17/57.43	20.40/76.53	26.34/85.12	13.00/67.51
Qwen3.5-9B	24.62/88.17	14.85/77.03	12.90/70.95	17.26/86.32	13.25/83.95	36.05/86.53	32.59/87.67	21.65/82.95

Table 1: Zero-shot Translation Results of Different Base Models on FLORES+

set and the WMT 2026 test set differ in terms of data scale and distribution, the results reported in the table are primarily used to reflect the relative translation capabilities of different models under a unified zero-shot setting and are not intended to directly estimate their final performance on the WMT 2026 test set.

The results show that the models exhibit varying performance across different translation directions. Overall, Gemma-4-12B-it demonstrates strong overall performance on bidirectional translation between Chinese and Southeast Asian languages, particularly on several low-resource translation directions. Considering its zero-shot translation performance, multilingual modeling capability, instruction-following ability, and inference efficiency, we ultimately select Gemma-4-12B-it as the base model for Xiaoyu-MT. We use the following unified translation prompt for all models:

Translate the following sentence from <source language> into <target language>. Output only the translation.

2.2 Data Construction

After selecting the base model, we construct multilingual training data for the subsequent training stages. Given the relatively limited availability of high-quality parallel corpora between Chinese and Southeast Asian languages, in addition to the training data provided by the WMT 2026 Chinese-Southeast Asian Multilingual MT Task, we further incorporate data from our existing Southeast Asian multilingual machine translation database. From this database, we select 50,000 monolingual sentences and 100,000 bilingual parallel sentence pairs. By integrating these data resources, we construct

bidirectional translation training data covering Chinese and seven Southeast Asian languages, namely Thai, Vietnamese, Lao, Burmese, Khmer, Indonesian, and Malay, providing the data foundation for subsequent continual pre-training and supervised fine-tuning.

Considering that multilingual parallel corpora for low-resource Southeast Asian languages may contain language identification errors, misaligned sentence pairs, duplicate samples, and anomalous text, we first perform unified data cleaning and quality control on the data collected from multiple sources. Specifically, we use the `lid.176.bin` language identification model to independently detect the source and target languages and remove sentence pairs with incorrect language labels, mixed languages, or inconsistencies with the expected target languages. We also perform duplicate detection to remove highly repetitive samples and reduce their impact on the training data distribution. In addition, we filter anomalous samples based on sentence length and abnormal symbol patterns, and split or remove excessively long or otherwise unsuitable samples when necessary.

Since the amount of training data varies substantially across translation directions, directly mixing all data during training may cause high-resource directions to account for a larger proportion of the training process, thereby weakening the model’s learning of low-resource directions. We therefore further adjust the data distribution across translation directions. While maintaining data quality, we increase the proportion of effective samples from low-resource directions during training to alleviate data imbalance in multilingual joint training.

After data collection, cleaning, deduplication, and distribution adjustment, we obtain multilin-

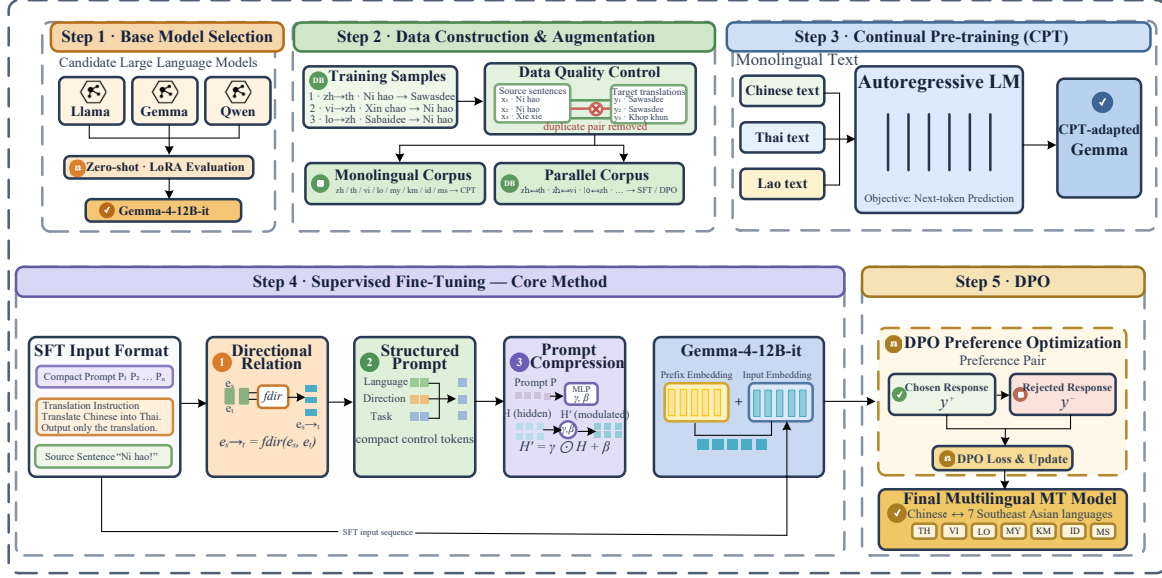


Figure 1: Overall Framework of the Multi-Stage Multilingual Machine Translation System.

gual training data covering bidirectional translation between Chinese and the seven Southeast Asian languages. These data provide a unified foundation for the subsequent CPT and SFT stages.

3 Methodology

3.1 System Overview

For the bidirectional machine translation task between Chinese and seven Southeast Asian languages, we develop Xiaoyu-MT, a multi-stage multilingual machine translation system based on Gemma-4-12B-it. Xiaoyu-MT adopts a stage-wise training strategy consisting of CPT, language-aware SFT, and DPO, where CPT, SFT, and DPO are responsible for enhancing language modeling capabilities, adapting the model to the translation task, and optimizing the generation distribution, respectively. During SFT, we further introduce a language-aware Prompt encoding mechanism to explicitly model the source language, target language, and their directed translation relationship. An overview of Xiaoyu-MT is shown in Figure 1.

During SFT, the language-aware Prompt encoding mechanism serves as an intermediate representation connecting language conditions with the translation generation model. It first represents the source and target languages separately and explicitly models the directed translation relationship between them. The language, translation direction, and task conditions are then organized into a structured Prompt and mapped to a fixed-length

conditional representation through Prompt compression. The compressed Prompt is provided to the large language model together with the source text, allowing the model to retain shared multilingual knowledge while obtaining explicit conditions for the current translation direction. Given a source language s , a target language t , and source text x , the language-aware Prompt encoding process can be formulated as:

$$P_{s \rightarrow t} = G_{\text{compress}} \left(G_{\text{prompt}} \left(F_{\text{dir}}(E(s), E(t)), e_{\text{task}} \right) \right). \quad (1)$$

Here, $E(\cdot)$ denotes the language representation function, $F_{\text{dir}}(\cdot)$ denotes the directed translation relationship modeling function, and e_{task} denotes a task-level representation derived from the content, semantic characteristics, and domain information of the current input text. $G_{\text{prompt}}(\cdot)$ denotes the structured Prompt encoding function, while $G_{\text{compress}}(\cdot)$ denotes the Prompt compression function. $P_{s \rightarrow t}$ denotes the compressed Prompt corresponding to the current translation direction and input conditions.

The compressed Prompt is then provided to the translation model together with the source text to generate the target-language translation:

$$y \sim \pi_{\theta}(y \mid x, P_{s \rightarrow t}). \quad (2)$$

Here, π_θ denotes the translation model being optimized.

3.2 Continual Pre-Training

Although Gemma-4-12B-it has acquired multilingual knowledge during pre-training, the coverage of different languages in the pre-training data is not uniform. Its ability to model the vocabulary, syntax, and linguistic patterns of some low-resource Southeast Asian languages can therefore be further improved. Before supervised fine-tuning, we first perform continual pre-training using the curated monolingual corpora in Chinese and the seven Southeast Asian languages.

Unlike directly adapting the model to the translation task using parallel corpora, the primary objective of the CPT stage is to further enhance the model’s fundamental language modeling capabilities for the target languages rather than directly learning specific translation directions. By continuing training on monolingual corpora in the target languages, the model can further learn their vocabulary distributions, syntactic structures, and linguistic expression patterns, thereby providing richer linguistic knowledge for subsequent cross-lingual translation.

Given a monolingual training sequence:

$$u = (u_1, u_2, \dots, u_T), \quad (3)$$

where u_j denotes the j -th token in the monolingual training sequence and T denotes the sequence length, we adopt the standard autoregressive language modeling objective, which predicts the current token based on its preceding context:

$$\mathcal{L}_{\text{CPT}} = -\frac{1}{T} \sum_{j=1}^T \log p_\theta(u_j \mid u_{<j}). \quad (4)$$

Here, $u_{<j}$ denotes the context preceding position j . By minimizing $\mathcal{L}_{\text{CPT}}$, the model can further adapt to the linguistic characteristics of Chinese and Southeast Asian languages while retaining its existing knowledge, thereby providing a stronger linguistic foundation for subsequent translation task adaptation.

3.3 Language-Aware Prompt Encoding

3.3.1 Directed Translation Relationship Modeling

An important distinction between multilingual machine translation and general multilingual text gen-

eration is that translation relationships are inherently directional. For the same language pair, Chinese $\rightarrow$ Thai and Thai $\rightarrow$ Chinese correspond to different conditional information and generation targets. If only independent source- and target-language representations are used, the model must implicitly infer the current translation direction from the combination of the two language representations, which may lead to confusion between different translation tasks.

We therefore first represent the source and target languages separately. For a given translation direction $s \rightarrow t$, we obtain:

$$e_s = E(s), \quad e_t = E(t), \quad (5)$$

where s and t denote the source and target languages, respectively, and $E(\cdot)$ denotes the language representation function.

We then explicitly represent the relationship from the source language to the target language through a directed relationship modeling function:

$$e_{s \rightarrow t} = F_{\text{dir}}(e_s, e_t). \quad (6)$$

Here, $F_{\text{dir}}(\cdot)$ denotes the directed translation relationship modeling function, and $e_{s \rightarrow t}$ denotes the directional representation of the current translation direction.

Since translation is directional, the two directions of the same language pair have different conditional representations:

$$e_{s \rightarrow t} \neq e_{t \rightarrow s}. \quad (7)$$

For example, for Chinese and Vietnamese, the model establishes different directional representations for Chinese $\rightarrow$ Vietnamese and Vietnamese $\rightarrow$ Chinese. By explicitly introducing directed relationship representations, the model does not need to rely solely on independent language identifiers to implicitly distinguish translation directions, thereby providing clearer task conditions for subsequent Prompt encoding.

3.3.2 Structured Language-Aware Prompt Encoding

After obtaining the directional representation, we organize conditional information into a structured Prompt. Rather than simply concatenating multiple language labels and task descriptions, we divide the Prompt conditions into four components according to their functional roles:

$$P_{\text{lang}}, P_{\text{dir}}, P_{\text{task}}, P_{\text{int}}, \quad (8)$$

Here, P_{lang} denotes the source- and target-language identity information involved in the current translation task, which specifies the language of the input text and the expected target language. P_{dir} denotes the directed translation relationship determined by the source and target languages, which distinguishes different translation directions, such as Chinese $\rightarrow$ Vietnamese and Vietnamese $\rightarrow$ Chinese. P_{task} denotes task-level conditional information derived from the content, semantic characteristics, and domain information of the input text, which characterizes the content and domain characteristics of the current input instance. P_{int} denotes the interaction information between the language-direction conditions and the input text, which further characterizes the task conditions of the input text under a specific translation direction.

To organize these conditions into a unified Prompt representation, we first feed the directed translation relationship representation and task-level representation into the structured Prompt encoding function:

$$P_{\text{raw}}^{s \rightarrow t} = G_{\text{prompt}}(e_{s \rightarrow t}, e_{\text{task}}). \quad (9)$$

Here, $P_{\text{raw}}^{s \rightarrow t}$ denotes the raw Prompt obtained through structured encoding, $G_{\text{prompt}}(\cdot)$ denotes the structured Prompt encoding function, and e_{task} denotes the task-level representation derived from the content, semantic characteristics, and domain information of the current input text.

From the perspective of functional composition, the structured Prompt can be represented as:

$$P_{\text{raw}}^{s \rightarrow t} = [P_{\text{lang}}; P_{\text{dir}}; P_{\text{task}}; P_{\text{int}}]. \quad (10)$$

Here, structured language-aware Prompt encoding is not simply a matter of adding language labels to the Prompt. Instead, it assigns explicit functional roles to different types of conditional information within the Prompt representation. In this way, the model can simultaneously access language identity, translation direction, and input-text information, thereby reducing potential confusion among different conditional signals.

3.3.3 Prompt Compression

Although additional conditional information can provide richer linguistic and task-related cues, directly concatenating a large number of conditional tokens before the source text increases the input

sequence length and introduces additional computational overhead during inference. We therefore introduce a Prompt compression mechanism to map the structured conditional information into a fixed-length Prompt representation. Let the raw conditional representation obtained through structured encoding be:

$$P_{\text{raw}}^{s \rightarrow t} \in R^{L \times d}, \quad (11)$$

where L denotes the length of the original conditional sequence and d denotes the model hidden dimension. We aggregate the raw conditional representation using the conditional Prompt compression function:

$$P_{s \rightarrow t} = G_{\text{compress}}(P_{\text{raw}}^{s \rightarrow t}). \quad (12)$$

This produces a fixed-length Prompt:

$$P_{s \rightarrow t} \in R^{K \times d}, \quad K \ll L. \quad (13)$$

Here, $G_{\text{compress}}(\cdot)$ denotes the Prompt compression module, and K denotes the pre-defined Prompt length.

After compression, $P_{s \rightarrow t}$ no longer preserves the original conditional sequence on a token-by-token basis. Instead, a learnable encoding process aggregates language, translation-direction, and task information into a limited number of Prompt representations. Therefore, the model can still obtain the conditional information required for the current translation direction without significantly increasing the input sequence length.

Finally, the compressed Prompt is provided to the translation model together with the source text:

$$y \sim \pi_{\theta}(y \mid x, P_{s \rightarrow t}). \quad (14)$$

Here, x denotes the source text, and $P_{s \rightarrow t}$ denotes the conditional Prompt corresponding to the current translation direction. This design balances the expressiveness of conditional information with the additional computational overhead.

3.4 Direct Preference Optimization

After completing SFT, we further employ Direct Preference Optimization (DPO) to optimize the translation generation distribution. Unlike SFT, which directly fits the reference translations, DPO constructs preferred and rejected translation pairs, enabling the model to further learn quality differences among alternative translation outputs.

3.4.1 Preference Data Construction

To avoid using exactly the same supervision signal for DPO and SFT, we construct preference data from data that are not used during SFT. For each source sentence x and its corresponding translation direction $s \rightarrow t$, we first use the SFT model to generate a translation:

$$\hat{y} \sim \pi_{\text{SFT}}(y \mid x, P_{s \rightarrow t}). \quad (15)$$

We then pair the model-generated translation $\hat{y}$ with the corresponding reference translation y^* to construct the preference dataset:

$$D_{\text{DPO}} = \{(x_i, P_{s_i \rightarrow t_i}, y_i^*, \hat{y}_i)\}_{i=1}^M. \quad (16)$$

The reference translation y_i^* is used as the chosen response, while the translation $\hat{y}_i$ generated by the SFT model is used as the rejected response:

$$y_w = y^*, \quad y_l = \hat{y}. \quad (17)$$

This construction uses the reference translation as a high-quality translation and the model-generated translation as a relatively weaker candidate, thereby providing the model with a direct preference signal for translation quality.

3.4.2 DPO Optimization

Given a preference tuple (x, P, y_w, y_l) , DPO directly optimizes the relative probabilities of the current policy model π_θ and the reference model π_{ref} . The optimization objective is:

$$\begin{aligned} \mathcal{L}_{\text{DPO}} = & -E_{(x, P, y_w, y_l) \sim D_{\text{DPO}}} \left[\log \sigma \left(\beta \left(\right. \right. \right. \\ & \log \frac{\pi_\theta(y_w \mid x, P)}{\pi_{\text{ref}}(y_w \mid x, P)} \\ & \left. \left. \left. - \log \frac{\pi_\theta(y_l \mid x, P)}{\pi_{\text{ref}}(y_l \mid x, P)} \right) \right) \right]. \end{aligned} \quad (18)$$

Here, π_θ denotes the translation model being optimized, π_{ref} denotes the reference model, and y_w and y_l denote the chosen and rejected translations, respectively. $\sigma(\cdot)$ denotes the sigmoid function, and β is a hyperparameter controlling the strength of preference optimization.

During training, the model is optimized to increase the relative probability of the high-quality reference translation y_w while decreasing the relative probability of the model-generated translation y_l . Unlike conventional RLHF pipelines that

require an additional reward model, DPO directly optimizes the model parameters using paired preference data, thereby further adjusting the generation distribution of the translation model.

For this task, this process further strengthens the model’s preference for high-quality translations after SFT has established the basic translation mappings. It is therefore intended to reduce translation errors, language confusion, and generation biases, while improving the final translation quality across different directions between Chinese and the seven Southeast Asian languages.

4 Experiments

4.1 Experimental Setup

We use Gemma-4-12B-it as the base model and sequentially perform CPT, SFT, and DPO following the training pipeline described above, progressively enhancing multilingual language capabilities, adapting the model to the translation task, and optimizing its generation distribution. During the CPT stage, we perform continual pre-training on cleaned and curated monolingual corpora in Chinese and the seven Southeast Asian languages, allowing the base model to further adapt to the multilingual data distribution involved in this task. After CPT, the resulting model parameters are used to initialize the SFT stage. We then perform supervised fine-tuning on the constructed multilingual parallel corpora and incorporate the language-aware Prompt encoding mechanism to provide the model with conditional information corresponding to the current input and translation direction. The learning rate is set to 3×10^{-4} , the batch size is set to 4, and the model is trained for 3 epochs during SFT.

After completing SFT, we further perform DPO training. The DPO stage uses the SFT model as the initial policy model and employs the corresponding reference model to compute the relative preference between the policy and reference models. The preference data are constructed from data that are not used for SFT, where reference translations are used as chosen responses and relatively lower-quality translations generated and filtered by the SFT model are used as rejected responses. In this way, the model’s translation generation distribution can be further optimized without introducing an additional independent reward model, encouraging the model to generate translations that are more consistent with the quality of the reference translations.

During inference, a corresponding language-aware Prompt is constructed for each test instance according to the specified source and target languages and provided to the final model together with the source text. We use beam search for decoding with a beam size of 4, balancing translation quality and inference efficiency.

4.2 Evaluation Metrics

To comprehensively evaluate the model’s performance on bidirectional machine translation between Chinese and Southeast Asian languages, we use sacreBLEU and COMET-22 as the primary evaluation metrics. sacreBLEU measures the degree of n -gram overlap between the generated translation and the reference translation, while COMET-22 employs a pre-trained model to evaluate the semantic consistency among the source sentence, candidate translation, and reference translation. These metrics provide complementary assessments of translation quality from the perspectives of lexical matching and semantic similarity.

4.3 Results

To evaluate the adaptation capability of Xiaoyu-MT for bidirectional translation between Chinese and Southeast Asian languages, we first conduct validation experiments on an internal test set constructed during the model development stage. Since this internal test set is primarily used for model analysis and rapid iteration during training rather than replacing the official comprehensive evaluation, we select a subset of translation directions covering different language types and translation scenarios for experimental validation. Among the 14 bidirectional translation directions, we select 4 directions and compare Xiaoyu-MT with the zero-shot translation results of the base model. As shown in Table 2, the experimental results demonstrate that Xiaoyu-MT achieves performance improvements across all evaluated directions after multi-stage training with CPT, language-aware Prompt SFT, and DPO. These results indicate that the training strategies adopted in Xiaoyu-MT effectively enhance the model’s adaptation capability to target languages and different translation directions. Specifically, the language-aware Prompt encoding mechanism explicitly models the source language, target language, and their directed translation relationship, providing more explicit directional conditions for the model and improving generation quality in multilingual translation scenarios.

Model	Metric	zh-vi	zh-my	zh-lo	zh-km
Xiaoyu-MT	sacreBLEU	38.51	22.48	31.70	25.29
	COMET-22	88.30	86.12	82.00	81.93

Table 2: Translation Results of Xiaoyu-MT on Representative Translation Directions

Metric	Xiaoyu-MT
BLEU	34.73
COMET	81.87
Throughput Rank	5
Throughput Score	3.5
Quality Score	58.30
Final Result	61.80

Table 3: Official Evaluation Results of Xiaoyu-MT on the WMT 2026 Chinese-Southeast Asian Multilingual Machine Translation Task

Furthermore, for the official evaluation of the WMT 2026 Chinese-Southeast Asian Multilingual Machine Translation Task, we perform inference with Xiaoyu-MT on the released test set and submit all generated translations to the task organizers. The official evaluation results are presented in Table 3. According to the official evaluation, Xiaoyu-MT achieves an overall BLEU score of 34.73 and COMET score of 81.87 across the bidirectional translation directions. In addition, the system obtains a Throughput Rank of 5 and a Throughput Score of 3.5, reflecting its competitive inference efficiency. Considering both translation quality and efficiency, Xiaoyu-MT achieves a Quality Score of 58.30 and a Final Result score of 61.80. These results demonstrate that Xiaoyu-MT can effectively handle bidirectional translation between Chinese and seven Southeast Asian languages, achieving a favorable balance between translation quality and inference efficiency in multilingual and low-resource translation scenarios.

5 Conclusion

This paper presents Xiaoyu-MT, the machine translation system submitted by the Kunming University of Science and Technology team to the WMT 2026 Chinese-Southeast Asian Multilingual MT Task. For bidirectional translation between Chinese and seven Southeast Asian languages, Xiaoyu-MT is built on Gemma-4-12B-it and adopts a multi-stage training framework consisting of data construction and augmentation, CPT, SFT, and DPO. During SFT, we introduce a language-aware Prompt encoding mechanism to explicitly model the source

language, target language, and their directed translation relationship, providing more effective conditional information for different translation directions. Preliminary evaluations on representative internal test directions show that Xiaoyu-MT consistently improves over the zero-shot baseline, demonstrating the effectiveness of the proposed training strategy. Furthermore, we perform inference on the official WMT 2026 test set and submit the generated translations for official evaluation. The official results show that Xiaoyu-MT achieves an overall BLEU score of 34.73 and a COMET score of 81.87, with a Throughput Rank of 5, a Throughput Score of 3.5, a Quality Score of 58.30, and a Final Result score of 61.80. These results demonstrate the effectiveness of combining multi-stage training with language-aware conditional modeling for adapting large language models to multilingual and low-resource machine translation scenarios.

Acknowledgments

We thank all reviewers for their helpful comments. This work financially supported by the National Natural Science Foundation of China (62666040, 62366027, 62306129, 62266027, U23A2038, U24A20334), Yunnan Fundamental Research Projects (202601CI070148, 202401CF070121, 202401BC070021), Major Science and Technology Special Projects of Yunnan Province (202602AD080015, 202502AD080014, 202502AD080016, 202402AG050007, 202402AD080002), Kunming University of Science and Technology Construction Joint Project (2023 01BE070001-027), Major Science and Technology Achievement Transformation Project of Yunnan Provincial Department of Education (FWCY-ZD2025006), Intergovernmental International Science, Technology and Innovation Cooperation (2025YFE0121700)

References

- Mengyu Bu, Shuhao Gu, and Yang Feng. 2024. [Improving multilingual neural machine translation by utilizing semantic and linguistic features](#). In *Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024*, volume ACL 2024 of *Findings of ACL*, pages 10410–10423. Association for Computational Linguistics.
- Menglong Cui, Pengzhi Gao, Wei Liu, Jian Luan, and Bin Wang. 2025. [Multilingual machine translation with open large language models at practical scale: An empirical study](#). In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2025 - Volume 1: Long Papers, Albuquerque, New Mexico, USA, April 29 - May 4, 2025*, pages 5420–5443. Association for Computational Linguistics.
- Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. 2019. [Parameter-efficient transfer learning for NLP](#). In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 2790–2799. PMLR.
- Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. [Lora: Low-rank adaptation of large language models](#). In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.
- Frederikus Hudi, Zhi Qu, Hidetaka Kamigaito, and Taro Watanabe. 2024. [Disentangling pretrained representation to leverage low-resource languages in multilingual machine translation](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy*, pages 4978–4989. ELRA and ICCL.
- Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. [The power of scale for parameter-efficient prompt tuning](#). In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3045–3059, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Yinqian Lu, Wenhao Zhu, Lei Li, Yu Qiao, and Fei Yuan. 2024. [Llamax: Scaling linguistic horizons of LLM by enhancing translation capabilities beyond 100 languages](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024, Miami, Florida, USA, November 12-16, 2024*, volume EMNLP 2024 of *Findings of ACL*, pages 10748–10772. Association for Computational Linguistics.
- Libo Qin, Qiguang Chen, Yuhang Zhou, Zhi Chen, Yinghui Li, Lizi Liao, Min Li, Wanxiang Che, and Philip S. Yu. 2025. [A survey of multilingual large language models](#). *Patterns*, 6(1):101118.
- Shivalika Singh, Freddie Vargus, Daniel D’souza, Börje Karlsson, Abinaya Mahendiran, Wei-Yin Ko, Herumb Shandilya, Jay Patel, Deividas Mataciunas, Laura O’Mahony, Mike Zhang, Ramith Hettiarachchi, Joseph Wilson, Marina Machado, Luisa Souza Moura, Dominik Krzeminski, Hakimeh Fadaei, Irem Ergün, Ifeoma Okoh, and 14 others. 2024. [Aya dataset: An open-access collection for multilingual instruction tuning](#). In *Proceedings of*

the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024, pages 11521–11567. Association for Computational Linguistics.

Dario Vajda, Domen Vres, and Marko Robnik-Sikonja. 2025. [Improving llms for machine translation using synthetic preference data](#). In *Proceedings of the 2nd Language Understanding in the Human-Machine Era Workshop, LUHME 2025, Bologna, Italy, October 26, 2025*, pages 67–73. Universidade do Porto/University of Eastern Finland/Universidad de Alcalá.

Rui Wang, Xu Tan, Renqian Luo, Tao Qin, and Tie-Yan Liu. 2021. [A survey on low-resource neural machine translation](#). In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021*, pages 4636–4643. ijcai.org.

Yangfan Ye, Xiaocheng Feng, Zekun Yuan, Xiachong Feng, Libo Qin, Lei Huang, Weitao Ma, Yichong Huang, Zhirui Zhang, Yunfei Lu, Xiaohui Yan, Duyu Tang, Dandan Tu, and Bing Qin. 2025. [Cc-tuning: A cross-lingual connection mechanism for improving joint multilingual supervised fine-tuning](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2025, Vienna, Austria, July 27 - August 1, 2025*, pages 19036–19051. Association for Computational Linguistics.

Wenxuan Zhang, Hou Pong Chan, Yiran Zhao, Mahani Aljunied, Jianyu Wang, Chaoqun Liu, Yue Deng, Zhiqiang Hu, Weiwen Xu, Yew Ken Chia, Xin Li, and Lidong Bing. 2025. [Seallms 3: Open foundation and chat multilingual large language models for southeast asian languages](#). In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2025 - System Demonstrations, Albuquerque, New Mexico, USA, April 29 - May 4, 2025*, pages 96–105. Association for Computational Linguistics.

Wenzhen Zheng, Wenbo Pan, Xu Xu, Libo Qin, Li Yue, and Ming Zhou. 2024. [Breaking language barriers: Cross-lingual continual pre-training at scale](#). In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024*, pages 7725–7738. Association for Computational Linguistics.

Wenhao Zhu, Hongyi Liu, Qingxiu Dong, Jingjing Xu, Shujian Huang, Lingpeng Kong, Jiajun Chen, and Lei Li. 2024. [Multilingual machine translation with large language models: Empirical results and analysis](#). In *Findings of the Association for Computational Linguistics: NAACL 2024, Mexico City, Mexico, June 16-21, 2024*, volume NAACL 2024 of *Findings of ACL*, pages 2765–2781. Association for Computational Linguistics.