




ALF-T5: Adaptive Neural Machine Translation Framework for Constructed and Natural Languages


Matheus J. G. Silva ^{1¶}

¹ Anhanguera University, Brazil ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: 

Submitted: 09 October 2025

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Summary

Neural machine translation (NMT) typically demands abundant parallel corpora, restricting its reach to widely spoken languages and limiting exploration of low-resource and constructed languages (conlangs). **ALF-T5** is a Python framework that enables **few-shot**, **bidirectional** translation between English and target languages, natural or constructed, by leveraging transfer learning, parameter-efficient fine-tuning (LoRA via PEFT), and data augmentation. It supports interactive usage, batch translation, automatic evaluation (BLEU, METEOR), and architectural experimentation, making it useful for research, teaching, and prototype development.

Statement of need

Most modern NMT toolkits assume large parallel corpora and substantial compute. This restricts experimentation with under-resourced languages and constructed languages, and limits exploration of translation strategies in extremely low-data regimes.

ALF-T5 fills this gap by enabling translation model training from as few as **10–50 example pairs**. The project emphasizes usability (CLI + Python API), reproducibility (built-in metrics and tests), and modularity (pluggable augmentation and experiment components). Because it integrates with the widely used Transformers and PEFT ecosystems, ALF-T5 is usable on modest hardware and useful for both educators and researchers.

Technical approach

Model and loss formulation

Let the source sequence be $X = (x_1, x_2, \dots, x_n)$ and the target sequence be $Y = (y_1, y_2, \dots, y_m)$. The model factorizes the conditional probability as

$$P(Y | X) = \prod_{t=1}^m P(y_t | y_{1:t-1}, X).$$

We train using the standard cross-entropy loss:

$$\mathcal{L}_{\text{CE}}(X, Y) = - \sum_{t=1}^m \log P(y_t | y_{1:t-1}, X).$$

To support bidirectional translation, we train both directions and combine losses:

$$\mathcal{L}_{\text{bi}} = \mathcal{L}_{\text{CE}}(X \rightarrow Y) + \mathcal{L}_{\text{CE}}(Y \rightarrow X).$$

(Using $y_{1:t-1}$ instead of $y_{<t}$ avoids characters that can be converted or escaped by the toolchain.)

Parameter-efficient tuning (LoRA / PEFT)

ALF-T5 uses LoRA (low-rank adaptation) implemented through the PEFT framework: the pretrained base model weights remain frozen and we learn small low-rank adapter matrices. This drastically reduces the number of trainable parameters and lowers memory / compute requirements compared to full fine-tuning (Hu et al., 2021). The adapters are inserted at attention / projection layers and trained while the backbone stays unchanged.

Data augmentation

To amplify scarce parallel data we use lightweight augmentations:

- Case variation (lowercase / uppercase / title)
- Word-order permutations within short phrases
- Vocabulary recombination (mixing subword segments or analogous tokens)

These synthetic variants increase the effective dataset size and help generalization in low-data settings.

Usage & features

- **Interactive mode:** CLI REPL for single-sentence translation with optional confidence scores.
- **Batch mode:** File and list translation for large-scale processing.
- **Automatic evaluation:** BLEU and METEOR computed during or after training.
- **Experiment framework:** Grid search over hyperparameters (LoRA rank, learning rate) with result logging.
- **Model inference:** Load a saved translator and translate new text via a simple API.

Citations

This project builds on a number of foundational works and libraries, including T5 (Raffel et al., 2020), the Transformer architecture (Vaswani et al., 2017), LoRA (Hu et al., 2021), BLEU (Papineni et al., 2002), METEOR (Banerjee & Lavie, 2005), and the Hugging Face Transformers library (Wolf et al., 2020).

62 Figures

```

Showcase

[30]: translation = translator.translate("the man did walk to his house with the dog running", direction="e2c")
      translation

[30]: 'erno voram pa duomo qui barmo rumam'

[31]: translation = translator.translate("erno voram pa duomo qui barmo rumam", direction="c2e")
      translation

[31]: 'man is walking to house with dog running'

```

Figure 1: Few-shot translation example in a constructed language (45 training pairs).

```

Tests

[12]: english = translator.translate("Quando é que foi possível extrair o objetivo principal desse artigo?", direction="c2e")
      portuguese = translator.translate("Why this action can establish causal links with the project?", direction="e2c")

      print(f"Portuguese to English: {english}")
      print(f"English to Portuguese: {portuguese}")

      Portuguese to English: When was it possible to achieve the main objective of this article?
      English to Portuguese: Por que a atividade pode estabelecer relações causais com o projeto?

[18]: translator.plot_training_history()

[21]: translator.history["bleu_scores"]

[21]: [{'epoch': 5,
      'corpus_bleu': 0.3727643147106651,
      'mean_bleu': 0.33965593822255324},
      {'epoch': 10,
      'corpus_bleu': 0.4196489353123995,
      'mean_bleu': 0.38185075305590166},
      {'epoch': 15,
      'corpus_bleu': 0.42741013803730493,
      'mean_bleu': 0.3849941277280052},
      {'epoch': 20,
      'corpus_bleu': 0.4392373169506965,
      'mean_bleu': 0.39789867099558707}]

[22]: english = translator.translate("Quando o artigo foi publicado?", direction="c2e")
      portuguese = translator.translate("Why was this important?", direction="e2c")

      print(f"Portuguese to English: {english}")
      print(f"English to Portuguese: {portuguese}")

      Portuguese to English: When was the article published?
      English to Portuguese: Por que é importante?

```

Figure 2: Portuguese-English translation performance (BLEU / METEOR).

63 Acknowledgements

64 Thanks to the authors of T5 and LoRA, and to the Hugging Face and PEFT teams. Special
65 thanks to Paulo Pirozelli for providing the Pirá dataset used in evaluation, and to early testers
66 who provided feedback.

67 Banerjee, S., & Lavie, A. (2005). METEOR: An automatic metric for MT evaluation with
68 improved correlation with human judgments. *Proceedings of the ACL Workshop on Intrinsic*
69 *and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, 65–72.
70 <https://aclanthology.org/W05-0909.pdf>

71 Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021).
72 LoRA: Low-rank adaptation of large language models. *arXiv Preprint arXiv:2106.09685*.
73 <https://arxiv.org/abs/2106.09685>

74 Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic
75 evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association*
76 *for Computational Linguistics*, 311–318. <https://doi.org/10.3115/1073083.1073135>

77 Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu,

- 78 P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer.
79 *Journal of Machine Learning Research*, 21(140), 1–67. [http://jmlr.org/papers/v21/20-074.](http://jmlr.org/papers/v21/20-074.html)
80 [html](http://jmlr.org/papers/v21/20-074.html)
- 81 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Łukasz, &
82 Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing*
83 *Systems*, 30. <https://arxiv.org/abs/1706.03762>
- 84 Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T.,
85 Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Platen, P. von, Ma, C., Jernite, Y., Plu,
86 J., Xu, C., Scao, T. L., Gugger, S., ... Rush, A. M. (2020). *Transformers: State-of-the-art*
87 *natural language processing*. <https://arxiv.org/abs/1910.03771>

DRAFT