

- [1] S. Ranathunga, E.-S. A. Lee, M. P. Skenduli, R. Shekhar, M. Alam, and R. Kaur, “Neural machine translation for low-resource languages: A survey,” 2021.
- [2] S. Thillainathan, S. Ranathunga, and S. Jayasena, “Fine-tuning self-supervised multilingual sequence-to-sequence models for extremely low-resource nmt,” in *2021 Moratuwa Engineering Research Conference (MERCCon)*. IEEE, 2021, pp. 432–437.
- [3] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, “Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension,” *arXiv preprint arXiv:1910.13461*, 2019.
- [4] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens *et al.*, “Moses: Open source toolkit for statistical machine translation,” in *Proceedings of the 45th annual meeting of the association for computational linguistics companion volume proceedings of the demo and poster sessions*, 2007, pp. 177–180.
- [5] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” *arXiv preprint arXiv:1409.3215*, 2014.
- [6] Y. Liu, J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, and L. Zettlemoyer, “Multilingual denoising pre-training for neural machine translation,” *Transactions of the Association for Computational Linguistics*, vol. 8, pp. 726–742, 2020.
- [7] Y. Tang, C. Tran, X. Li, P.-J. Chen, N. Goyal, V. Chaudhary, J. Gu, and A. Fan, “Multilingual translation with extensible multilingual pretraining and finetuning,” *arXiv preprint arXiv:2008.00401*, 2020.
- [8] L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, and C. Raffel, “mt5: A massively multilingual pre-trained text-to-text transformer,” *arXiv preprint arXiv:2010.11934*, 2020.

- [9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *arXiv preprint arXiv:1810.04805*, 2018.
- [10] R. Dabre, C. Chu, and A. Kunchukuttan, “A survey of multilingual neural machine translation,” *ACM Computing Surveys (CSUR)*, vol. 53, no. 5, pp. 1–38, 2020.
- [11] P. Tennage, P. Sandaruwan, M. Thilakarathne, A. Herath, S. Ranathunga, S. Jayasena, and G. Dias, “Neural machine translation for sinhala and tamil languages,” in *2017 International Conference on Asian Language Processing (IALP)*. IEEE, 2017, pp. 189–192.
- [12] P. Tennage, P. Sandaruwan, M. Thilakarathne, A. Herath, and S. Ranathunga, “Handling rare word problem using synthetic training data for sinhala and tamil neural machine translation,” in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.
- [13] P. Tennage, A. Herath, M. Thilakarathne, P. Sandaruwan, and S. Ranathunga, “Transliteration and byte pair encoding to improve tamil to sinhala neural machine translation,” in *2018 Moratuwa Engineering Research Conference (MER-Con)*. IEEE, 2018, pp. 390–395.
- [14] A. Pramodya, R. Pushpananda, and R. Weerasinghe, “A comparison of transformer, recurrent neural networks and smt in tamil to sinhala mt,” in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2020, pp. 155–160.
- [15] T. Fonseka, R. Naranpanawa, R. Perera, and U. Thayasivam, “English to sinhala neural machine translation,” in *2020 International Conference on Asian Language Processing (IALP)*. IEEE, 2020, pp. 305–309.
- [16] R. Naranpanawa, R. Perera, T. Fonseka, and U. Thayasivam, “Analyzing subword techniques to improve english to sinhala neural machine translation,” *International Journal of Asian Language Processing*, vol. 30, no. 04, p. 2050017, 2020.

- [17] B. Janarthanasarma and T. Uthayasanker, “A survey on neural machine translation for english-tamil language pair.”
- [18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *arXiv preprint arXiv:1706.03762*, 2017.
- [19] N. Arivazhagan, A. Bapna, O. Firat, D. Lepikhin, M. Johnson, M. Krikun, M. X. Chen, Y. Cao, G. Foster, C. Cherry *et al.*, “Massively multilingual neural machine translation in the wild: Findings and challenges,” *arXiv preprint arXiv:1907.05019*, 2019.
- [20] A. Arukgoda, A. Weerasinghe, and R. Pushpananda, “Improving sinhala-tamil translation through deep learning techniques.” in *NL4AI@ AI\* IA*, 2019.
- [21] L. Nissanka, B. Pushpananda, and A. Weerasinghe, “Exploring neural machine translation for sinhala-tamil languages pair,” in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2020, pp. 202–207.
- [22] N. Kalchbrenner and P. Blunsom, “Recurrent continuous translation models,” in *Proceedings of the 2013 conference on empirical methods in natural language processing*, 2013, pp. 1700–1709.
- [23] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” *arXiv preprint arXiv:1409.0473*, 2014.
- [24] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [25] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
- [26] K. Epaliyana, S. Ranathunga, and S. Jayasena, “Improving back-translation with iterative filtering and data selection for sinhala-english nmt,” in *2021 Moratuwa Engineering Research Conference (MERCon)*. IEEE, 2021, pp. 438–443.

- [27] H. Choudhary, A. K. Pathak, R. R. Saha, and P. Kumaraguru, “Neural machine translation for english-tamil,” in *Proceedings of the third conference on machine translation: shared task papers*, 2018, pp. 770–775.
- [28] T. Banerjee, A. Kunchukuttan, and P. Bhattacharyya, “Multilingual indian language translation system at wat 2018: Many-to-one phrase-based smt,” in *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation: 5th Workshop on Asian Translation: 5th Workshop on Asian Translation*, 2018.
- [29] R. Aharoni, M. Johnson, and O. Firat, “Massively multilingual neural machine translation,” *arXiv preprint arXiv:1903.00089*, 2019.
- [30] D. Dong, H. Wu, W. He, D. Yu, and H. Wang, “Multi-task learning for multiple language translation,” in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2015, pp. 1723–1732.
- [31] M.-T. Luong, Q. V. Le, I. Sutskever, O. Vinyals, and L. Kaiser, “Multi-task sequence to sequence learning,” *arXiv preprint arXiv:1511.06114*, 2015.
- [32] M. Johnson, M. Schuster, Q. V. Le, M. Krikun, Y. Wu, Z. Chen, N. Thorat, F. Viégas, M. Wattenberg, G. Corrado *et al.*, “Google’s multilingual neural machine translation system: Enabling zero-shot translation,” *Transactions of the Association for Computational Linguistics*, vol. 5, pp. 339–351, 2017.
- [33] B. Zoph and K. Knight, “Multi-source neural translation,” *arXiv preprint arXiv:1601.00710*, 2016.
- [34] O. Firat, K. Cho, and Y. Bengio, “Multi-way, multilingual neural machine translation with a shared attention mechanism,” *arXiv preprint arXiv:1601.01073*, 2016.
- [35] T.-L. Ha, J. Niehues, and A. Waibel, “Toward multilingual neural machine translation with universal encoder and decoder,” *arXiv preprint arXiv:1611.04798*, 2016.

- [36] O. Firat, B. Sankaran, Y. Al-Onaizan, F. T. Y. Vural, and K. Cho, “Zero-resource translation with multi-lingual neural machine translation,” *arXiv preprint arXiv:1606.04164*, 2016.
- [37] Y. Lu, P. Keung, F. Ladhak, V. Bhardwaj, S. Zhang, and J. Sun, “A neural interlingua for multilingual machine translation,” *arXiv preprint arXiv:1804.08198*, 2018.
- [38] S. M. Lakew, M. Federico, M. Negri, and M. Turchi, “Multilingual neural machine translation for zero-resource languages,” *arXiv preprint arXiv:1909.07342*, 2019.
- [39] G. Blackwood, M. Ballesteros, and T. Ward, “Multilingual neural machine translation with task-specific attention,” *arXiv preprint arXiv:1806.03280*, 2018.
- [40] Y. Wang, J. Zhang, F. Zhai, J. Xu, and C. Zong, “Three strategies to improve one-to-many multilingual translation,” in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 2955–2960.
- [41] V. Goyal, S. Kumar, and D. M. Sharma, “Efficient neural machine translation for low-resource languages via exploiting related languages,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, 2020, pp. 162–168.
- [42] S. M. Lakew, A. Erofeeva, M. Negri, M. Federico, and M. Turchi, “Transfer learning in multilingual neural machine translation with dynamic vocabulary,” *arXiv preprint arXiv:1811.01137*, 2018.
- [43] S. M. Lakew, M. Cettolo, and M. Federico, “A comparison of transformer and recurrent neural networks on multilingual neural machine translation,” *arXiv preprint arXiv:1806.06957*, 2018.
- [44] B. Zoph, D. Yuret, J. May, and K. Knight, “Transfer learning for low-resource neural machine translation,” *arXiv preprint arXiv:1604.02201*, 2016.
- [45] R. Dabre, T. Nakagawa, and H. Kazawa, “An empirical study of language relatedness for transfer learning in neural machine translation,” in *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation*, 2017, pp. 282–286.

- [46] T. Q. Nguyen and D. Chiang, “Transfer learning across low-resource, related languages for neural machine translation,” *arXiv preprint arXiv:1708.09803*, 2017.
- [47] G. Neubig and J. Hu, “Rapid adaptation of neural machine translation to new languages,” *arXiv preprint arXiv:1808.04189*, 2018.
- [48] A. F. Aji, N. Bogoychev, K. Heafield, and R. Sennrich, “In neural machine translation, what does transfer learning transfer?” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 7701–7710.
- [49] B. Ji, Z. Zhang, X. Duan, M. Zhang, B. Chen, and W. Luo, “Cross-lingual pre-training based transfer for zero-shot neural machine translation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 115–122.
- [50] T. Kocmi and O. Bojar, “Efficiently reusing old models across languages via transfer learning,” *arXiv preprint arXiv:1909.10955*, 2019.
- [51] M. Maimaiti, Y. Liu, H. Luan, and M. Sun, “Multi-round transfer learning for low-resource nmt using multiple high-resource languages,” *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 18, no. 4, pp. 1–26, 2019.
- [52] Y. Kim, Y. Gao, and H. Ney, “Effective cross-lingual transfer of neural machine translation models without shared vocabularies,” *arXiv preprint arXiv:1905.05475*, 2019.
- [53] M. Maimaiti, Y. Liu, H. Luan, and M. Sun, “Enriching the transfer learning with pre-trained lexicon embedding for low-resource neural machine translation,” *Tsinghua Science and Technology*, p. 1, 2020.
- [54] A. Imankulova, R. Dabre, A. Fujita, and K. Imamura, “Exploiting out-of-domain parallel data through multilingual transfer learning for low-resource neural machine translation,” *arXiv preprint arXiv:1907.03060*, 2019.
- [55] C. Chu, R. Dabre, and S. Kurohashi, “An empirical comparison of domain adaptation methods for neural machine translation,” in *Proceedings of the 55th Annual Meeting*

of the Association for Computational Linguistics (Volume 2: Short Papers), 2017, pp. 385–391.

- [56] G. Luo, Y. Yang, Y. Yuan, Z. Chen, and A. Ainiwaer, “Hierarchical transfer learning architecture for low-resource neural machine translation,” *IEEE Access*, vol. 7, pp. 154 157–154 166, 2019.
- [57] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving language understanding by generative pre-training,” 2018.
- [58] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “Roberta: A robustly optimized bert pretraining approach,” *arXiv preprint arXiv:1907.11692*, 2019.
- [59] S. Clinchant, K. W. Jung, and V. Nikoulina, “On the use of BERT for neural machine translation,” in *Proceedings of the 3rd Workshop on Neural Generation and Translation*, 2019, pp. 108–117.
- [60] J. Yang, M. Wang, H. Zhou, C. Zhao, W. Zhang, Y. Yu, and L. Li, “Towards making the most of bert in neural machine translation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 9378–9385.
- [61] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, “Pre-trained models for natural language processing: A survey,” *Science China Technological Sciences*, pp. 1–26, 2020.
- [62] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, “Exploring the limits of transfer learning with a unified text-to-text transformer,” *arXiv preprint arXiv:1910.10683*, 2019.
- [63] Z. Chi, L. Dong, S. Ma, S. H. X.-L. Mao, H. Huang, and F. Wei, “mt6: Multilingual pretrained text-to-text transformer with translation pairs,” *arXiv preprint arXiv:2104.08692*, 2021.
- [64] E.-S. A. Lee, S. Thillainathan, S. Nayak, S. Ranathunga, D. I. Adelani, R. Su, and

- A. D. McCarthy, “Pre-trained multilingual sequence-to-sequence models: A hope for low-resource language translation?” *arXiv preprint arXiv:2203.08850*, 2022.
- [65] F. Guzmán, P.-J. Chen, M. Ott, J. Pino, G. Lample, P. Koehn, V. Chaudhary, and M. Ranzato, “The flores evaluation datasets for low-resource machine translation: Nepali-english and sinhala-english,” *arXiv preprint arXiv:1902.01382*, 2019.
- [66] L. Madaan, S. Sharma, and P. Singla, “Transfer learning for related languages: Submissions to the wmt20 similar language translation task,” in *Proceedings of the Fifth Conference on Machine Translation*, 2020, pp. 402–408.
- [67] S. Cahyawijaya, G. I. Winata, B. Wilie, K. Vincentio, X. Li, A. Kuncoro, S. Ruder, Z. Y. Lim, S. Bahar, M. L. Khodra *et al.*, “Indonlg: Benchmark and resources for evaluating indonesian natural language generation,” *arXiv preprint arXiv:2104.08200*, 2021.
- [68] A. Bapna, N. Arivazhagan, and O. Firat, “Simple, scalable adaptation for neural machine translation,” *arXiv preprint arXiv:1909.08478*, 2019.
- [69] Z. Liu, G. I. Winata, and P. Fung, “Continual mixed-language pre-training for extremely low-resource neural machine translation,” *arXiv preprint arXiv:2105.03953*, 2021.
- [70] P.-J. Chen, A. Lee, C. Wang, N. Goyal, A. Fan, M. Williamson, and J. Gu, “Facebook ai’s wmt20 news translation task submission,” *arXiv preprint arXiv:2011.08298*, 2020.
- [71] R. H. Susanto, D. Wang, S. Yadav, M. Jain, and O. Htun, “Rakuten’s participation in wat 2021: Examining the effectiveness of pre-trained models for multilingual and multimodal machine translation,” in *Proceedings of the 8th Workshop on Asian Translation (WAT2021)*, 2021, pp. 96–105.
- [72] I. Beltagy, K. Lo, and A. Cohan, “Scibert: A pretrained language model for scientific text,” *arXiv preprint arXiv:1903.10676*, 2019.



- [73] E. Alsentzer, J. R. Murphy, W. Boag, W.-H. Weng, D. Jin, T. Naumann, and M. McDermott, “Publicly available clinical bert embeddings,” *arXiv preprint arXiv:1904.03323*, 2019.
- [74] S. Gururangan, A. Marasović, S. Swayamdipta, K. Lo, I. Beltagy, D. Downey, and N. A. Smith, “Don’t stop pretraining: adapt language models to domains and tasks,” *arXiv preprint arXiv:2004.10964*, 2020.
- [75] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, “Biobert: a pre-trained biomedical language representation model for biomedical text mining,” *Bioinformatics*, vol. 36, no. 4, pp. 1234–1240, 2020.
- [76] R. Zhang, R. G. Reddy, M. A. Sultan, V. Castelli, A. Ferritto, R. Florian, E. S. Kayi, S. Roukos, A. Sil, and T. Ward, “Multi-stage pre-training for low-resource domain adaptation,” *arXiv preprint arXiv:2010.05904*, 2020.
- [77] L. K. Hansen and P. Salamon, “Neural network ensembles,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 12, no. 10, pp. 993–1001, 1990.
- [78] H. Chen, S. Lundberg, and S.-I. Lee, “Checkpoint ensembles: Ensemble methods from a single training process,” *arXiv preprint arXiv:1710.03282*, 2017.
- [79] R. Sennrich, B. Haddow, and A. Birch, “Edinburgh neural machine translation systems for wmt 16,” *arXiv preprint arXiv:1606.02891*, 2016.
- [80] R. Sennrich, A. Birch, A. Currey, U. Germann, B. Haddow, K. Heafield, A. V. M. Barone, and P. Williams, “The university of edinburgh’s neural mt systems for wmt17,” *arXiv preprint arXiv:1708.00726*, 2017.
- [81] K. Imamura and E. Sumita, “Ensemble and reranking: Using multiple models in the nict-2 neural machine translation system at wat2017,” in *Proceedings of the 4th Workshop on Asian Translation (WAT2017)*, 2017, pp. 127–134.
- [82] A. Fernando, S. Ranathunga, and G. Dias, “Data augmentation and terminology integration for domain-specific sinhala-english-tamil statistical machine translation,” *arXiv preprint arXiv:2011.02821*, 2020.

- [83] M. Rajitha, L. Piyarathna, M. Nayanajith, and S. Surangika, "Sinhala and english document alignment using statistical machine translation," in *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*. IEEE, 2020, pp. 29–34.
- [84] F. Farhath, S. Ranathunga, S. Jayasena, and G. Dias, "Integration of bilingual lists for domain-specific statistical machine translation for sinhala-tamil," in *2018 Moratuwa Engineering Research Conference (MERCOn)*. IEEE, 2018, pp. 538–543.
- [85] R. Sennrich and B. Zhang, "Revisiting low-resource neural machine translation: A case study," *arXiv preprint arXiv:1905.11901*, 2019.
- [86] A. Fan, S. Bhosale, H. Schwenk, Z. Ma, A. El-Kishky, S. Goyal, M. Baines, O. Celebi, G. Wenzek, V. Chaudhary *et al.*, "Beyond english-centric multilingual machine translation," *Journal of Machine Learning Research*, vol. 22, no. 107, pp. 1–48, 2021.
- [87] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 2002, pp. 311–318.
- [88] R. Futrell, K. Mahowald, and E. Gibson, "Quantifying word order freedom in dependency corpora," in *Proceedings of the third international conference on dependency linguistics (Depling 2015)*, 2015, pp. 91–100.
- [89] M. Anand Kumar, V. Dhanalakshmi, K. Soman, and S. Rajendran, "A sequence labeling approach to morphological analyzer for tamil language," *IJCSE) International Journal on Computer Science and Engineering*, vol. 2, no. 06, pp. 1944–195, 2010.