

BVSLP at WMT 2026: An Empirical Study of Multilingual Transfer for Low-Resource Arabic-Centric Translation

Palak Arora^{1,2}, Syed Afroz Ahmed³, Mansi Jangid¹, Bharti Nathani¹, Nisheeth Joshi¹

¹Speech and Language Processing Lab, Centre for Artificial Intelligence, Banasthali Vidyapith, Rajasthan, India

²School of Computing, DIT University, Dehradun, Uttarakhand, India

³College of Technology, City University, Ajman, UAE

palak.arora.pa55@gmail.com, a_afroz@hotmail.com, mansijangid@banasthali.in, nbharti@banasthali.in, nisheeth.joshi@rediffmail.com

Abstract

This paper presents HEMANT framework, the BSVLP submission to the WMT 2026 Low-Resource Arabic–Asian Machine Translation Shared Task. The system addresses ten bidirectional translation directions involving Arabic, English, Hindi, Bangla, Indonesian, and Urdu. HEMANT integrates language-specific Unicode normalization, spelling correction, named-entity recognition, knowledge-base-assisted entity translation, transliteration, multilingual transfer learning, and parameter-efficient adaptation of the NLLB-200 model using LoRA. Two directional multilingual models were trained for many-to-Arabic and Arabic-to-many translations. The official results showed considerable variation across language pairs. Arabic–Hindi produced the strongest relative performance, while English–Arabic remained the most challenging pair. In several directions, COMET scores were more competitive than BLEU and TER, suggesting that semantic adequacy was often better preserved than lexical overlap. The findings highlight both the potential and limitations of entity-aware multilingual adaptation for low-resource Arabic–Asian translation.

1 Introduction

Machine translation has advanced considerably with pretrained models capable of transferring knowledge across languages and translation directions. However, translation quality remains uneven for low-resource language pairs, particularly when languages differ in morphology, syntax, writing systems, and availability of parallel corpora. These challenges are evident in Arabic–

Asian translation, where Arabic must be translated to and from languages such as English, Hindi, Bangla, Indonesian, and Urdu, which belong to different linguistic families and use diverse scripts.

The WMT 2026 Low-Resource Arabic–Asian Machine Translation Shared Task provided a common evaluation framework for ten bidirectional translation directions involving Arabic and these five languages. Arabic itself presents additional difficulties because of its rich morphology, clitic attachment, gender and number agreement, definiteness marking, and orthographic variation. Named entities further complicate translation because person names, locations, organizations, and institutions may require conventional translation, transliteration, or preservation depending on the target language.

This paper presents HEMANT (Highly Efficient Machine-Assisted Natural Translation), the BSVLP submission to the shared task. The system is built on the NLLB-200 distilled 600M multilingual translation model and integrates language-specific Unicode and spelling normalization, corpus cleaning, POS-assisted named-entity recognition, knowledge-base-supported entity translation, rule-based transliteration, and entity-aware corpus augmentation.

The translation framework was organized into two Arabic-centric multilingual systems: a many-to-Arabic model and an Arabic-to-many model. Parameter-efficient adaptation was performed using LoRA, while temperature-based sampling was used to reduce corpus imbalance during multilingual training.

The study examines both the effectiveness and limitations of this approach across all ten directions. Particular attention is given to language-specific performance differences,

named-entity handling, script variation, and the contrast between lexical and semantic evaluation metrics. The results indicate that multilingual transfer provides a useful foundation for Arabic–Asian translation, although further language-pair-specific adaptation is necessary for uniformly strong performance.

The remainder of the paper is organized as follows. Section 2 reviews related work on multilingual neural machine translation, low-resource Arabic and Asian translation, named-entity translation, transliteration, and parameter-efficient model adaptation. Section 3 describes the BSVLP-HEMANT system. Section 5 concludes the paper and outlines directions for future research.

2 Literature Review

Neural machine translation has progressed from recurrent models to Transformer architectures, with Vaswani et al. (2017) establishing self-attention as the dominant paradigm. Vocabulary sparsity and limited parallel data were addressed through subword modelling (Sennrich et al., 2016a) and back-translation (Sennrich et al., 2016b), later extended through large-scale synthetic-data studies by Edunov et al. (2018). Transfer learning further improved low-resource NMT, with Zoph et al. (2016), Nguyen and Chiang (2017), and Kim et al. (2019) demonstrating effective knowledge transfer across related and unrelated languages.

Multilingual NMT was advanced by Firat et al. (2016), Johnson et al. (2017), and Aharoni et al. (2019), although Zhang et al. (2020) identified problems such as negative transfer and off-target generation. Multilingual pretraining through mBART (Liu et al., 2020) and NLLB (NLLB Team, 2022) further strengthened low-resource translation. Parameter-efficient adaptation was explored through adapters (Bapna & Firat, 2019), LoRA (Hu et al., 2021), and language-family adapters (Chronopoulou et al., 2023).

For Indic MT, Samanantar (Ramesh et al., 2022), IndicTrans2 (Gala et al., 2023), and Basit et al. (2024) demonstrated the value of language-specific resources. Arabic MT studies highlighted morphological and dialectal challenges (Almansor & Al-Ani, 2017; Shapiro & Duh, 2018; Sajjad et al., 2020; Nagoudi et al., 2022). Named-entity translation has similarly benefited from explicit entity modelling (Gahbiche-Braham et al., 2013;

Ugawa et al., 2018; Modrzejewski et al., 2020), neural transliteration (Grundkiewicz & Heafield, 2018), and knowledge-aware approaches (Hu et al., 2022; Zeng et al., 2023; Liang et al., 2024).

3 Methodology

The BSVLP system, named HEMANT (Highly Efficient Machine-Assisted Natural Translation), was developed for the WMT 2026 Low-Resource Arabic–Asian Machine Translation Shared Task. The task involved bidirectional translation between Arabic and five languages: English, Hindi, Bangla, Indonesian, and Urdu. Thus, the complete experimental setting consisted of ten translation directions, including the five language-to-Arabic directions and their corresponding Arabic-to-language directions. The proposed framework combined multilingual neural machine translation with language-specific text normalization, spelling correction, named-entity recognition, knowledge-base-assisted entity translation, transliteration, multilingual transfer learning, and parameter-efficient adaptation of the NLLB-200 model.

The methodology was based on the observation that direct fine-tuning of a multilingual pretrained model on noisy and unevenly distributed low-resource parallel data may not fully exploit the model’s multilingual capacity. Orthographic inconsistencies, rare forms, named-entity variation, script differences, and corpus imbalance can substantially increase data sparsity. Consequently, HEMANT incorporated preprocessing and entity-aware augmentation before model adaptation.

The corpus for fine-tuning the models was provided by the organizers (Laskar et al., 2026). This had 20,000 parallel sentences each direction.

3.1 Corpus Processing

The first stage involved corpus organization and profiling. The official training, development, and development-test files were stored separately for each language pair and translation direction. Each corpus was analyzed for sentence counts, duplicates, average and maximum sentence lengths, script distribution, mixed-language content, numerical expressions, exact source–target copies, and alignment inconsistencies. Source and target files were also checked to ensure equal line counts. This preliminary analysis helped identify potentially noisy sentence pairs and provided information about corpus imbalance across language directions.

All text was converted to UTF-8 and normalized using Unicode Normalization Form C. General preprocessing included removal of leading and trailing whitespace, repeated spaces, invalid control characters, malformed byte sequences, visible HTML/XML tags, and unnecessary punctuation duplication. Meaningful numbers, dates, percentages, currencies, acronyms, and named entities were preserved. Lower-casing and script unification were avoided because these operations could remove information useful for named-entity recognition and target-language generation.

Language-specific normalization was then applied. Arabic preprocessing included tatweel removal, elimination of inconsistent optional diacritics, punctuation normalization, correction of malformed character sequences, and standardization of spacing around Arabic and Western digits. Aggressive character normalization was deliberately avoided so that distinct lexical forms were preserved. Urdu was processed separately because it contains language-specific characters such as ھ, ں, ڑ, ٲ, ٴ, and ٷ. The Urdu pipeline preserved these characters while normalizing tatweel, punctuation, Farsi Yeh, Bari Yeh, spacing, and invalid zero-width characters. Hindi preprocessing handled Devanagari Unicode variants, nukta forms, malformed combining marks, punctuation, and Romanized text. Bangla preprocessing addressed vowel-sign inconsistencies, conjunct consonants, malformed combining sequences, excessive Latin-script content, and possible Assamese contamination. English and Indonesian were normalized more conservatively, with case, abbreviations, proper names, reduplicated Indonesian forms, and valid borrowed expressions retained.

A spelling-correction module was applied before subword tokenization. The module combined manually defined language-specific rules, corpus-frequency information, and Levenshtein distance. For a potentially misspelled token, candidate corrections were generated from a language-specific lexicon. Candidates within an edit distance of two were considered, and the final correction was selected based on edit distance, frequency in the training corpus, and contextual compatibility. The lexicons were compiled from the official corpora and supplementary resources such as Wikipedia text, news sources, government terminology lists, gazetteers, and dictionaries

where permitted. The correction process was deliberately conservative to avoid modifying legitimate rare words or proper names. Named entities were therefore protected from automatic spelling correction whenever possible.

3.2 Model Training

Following normalization, the corpora were filtered for alignment and quality. Sentence pairs were removed when either side was empty, contained only punctuation or numbers, exceeded 256 NLLB subword tokens, displayed an extreme source–target length ratio, belonged to the wrong language, consisted predominantly of symbols or markup, or represented exact duplicates. A broad length-ratio interval between one-third and three was used because the six languages differ significantly in morphology, tokenisation, and average sentence length. The development and development-test datasets were handled conservatively, with only clear encoding or structural problems corrected.

Named-entity handling formed an important component of HEMANT. A hybrid NER approach combined BiLSTM-based part-of-speech information with gazetteers, contextual rules, regular expressions, capitalisation cues, and script-sensitive patterns. Proper-noun POS categories were used as candidate signals for named entities, particularly in Hindi, Bangla, and Urdu. English and Indonesian additionally benefited from capitalisation patterns, whereas Arabic and Urdu relied more heavily on lexical and contextual markers. Recognised entities were classified into categories such as Person, Organisation, Location, Date, Time, Money, Percentage, and miscellaneous named expressions.

The detected entities were then passed to a multilingual Named Entity Translation module. Each entity was searched in a knowledge base containing conventional translations and transliterations for frequently occurring names, locations, institutions, government bodies, airports, universities, currencies, and common abbreviations. When an approved target-language equivalent existed, that form was used. When no translation was available, a standard or rule-based transliteration was generated. This distinction was important because organisation and location names often have established translations, whereas personal names generally require transliteration. The processed named entities were used to

generate entity-aware sentence variants, which were added to the original corpus as supplementary training examples.

For subword segmentation, the system retained the original SentencePiece tokenizer distributed with NLLB-200 with 600 million parameters. No new BPE vocabulary was trained because this would have broken compatibility with the pretrained model’s embedding and output layers.

The translation framework used two Arabic-centric multilingual models. The first model handled English, Hindi, Bangla, Indonesian, and Urdu to Arabic, allowing the decoder to specialise in Arabic generation. The second handled Arabic to the five target languages, enabling the encoder to develop a shared Arabic representation. Temperature-based sampling with an exponent of 0.7 was applied to reduce the dominance of larger corpora during multilingual training.

The base NLLB-200 distilled 600M model was adapted using Low-Rank Adaptation. LoRA modules with rank 16, alpha 32, and dropout 0.05 were inserted into the `q_proj`, `k_proj`, `v_proj`, and `out_proj` attention layers. Training used AdamW, weight decay of 0.01, ten epochs, batch size 1, gradient accumulation of 16, label smoothing of 0.1, FP16 precision, gradient checkpointing, and maximum source and target lengths of 256 subword tokens. The effective batch size was 16.

4 Evaluation

The BSVLP system participated in all ten primary translation directions of the WMT 2026 Low-Resource Arabic–Asian Language Translation Shared Task. Table 1 summarizes the official scores reported in the attached result sheets.

The results indicate that BSVLP’s performance was strongly language-dependent rather than uniformly direction-dependent. The system did not consistently perform better either into Arabic or from Arabic. Instead, the strongest results were obtained for Arabic-Hindi language pair, while the weakest results occurred in both English-related directions.

The BSVLP system showed considerable variation across the ten translation directions. For English-Arabic, BSVLP ranked 9th of nine systems, achieving BLEU 12.4680, chrF++ 44.8541, TER 81.5223, COMET 80.7693, and mBERT 82.0453. The relatively smaller COMET gap compared with BLEU suggests partial

preservation of meaning despite weaker lexical and surface correspondence.

For Hindi-Arabic, BSVLP ranked fifth of eight with BLEU 13.2825. Its chrF++ of 48.1994 and COMET of 81.8869 were comparatively stronger, indicating better semantic adequacy than the BLEU ranking alone suggests. A similar pattern occurred for Bangla-Arabic, where BSVLP ranked fourth of six but achieved the third-best chrF++ (47.1008) and COMET (81.2549). Indonesian-Arabic ranked fourth of four, although its COMET of 81.1868 exceeded IndicMT, again indicating stronger semantic than surface-level performance. For Urdu-Arabic, BSVLP ranked fifth of six, with COMET 81.4107 and chrF++ 47.0654 showing moderate competitiveness.

Among the reverse directions, Arabic-English was the weakest, ranking 10th of ten with BLEU 20.1319 and COMET 78.8771. In contrast, Arabic-Hindi was BSVLP’s strongest result, ranking fourth of ten with BLEU 24.6720, TER 66.1223, and mBERT 84.3182. For Arabic-Bangla, the system ranked fifth of seven but achieved the third-highest COMET of 77.6456. Arabic-Indonesian ranked fourth of four, although its COMET of 81.8981 was almost identical to IndicMT. Arabic-Urdu ranked fifth of six with BLEU 20.9917 and COMET 73.8080.

5 Conclusion

This paper presented HEMANT framework, the BSVLP submission to the WMT 2026 Low-Resource Arabic–Asian Machine Translation Shared Task. The system addressed ten bidirectional translation directions involving Arabic, English, Hindi, Bangla, Indonesian, and Urdu. It combined language-specific normalization, spelling correction, named-entity recognition, knowledge-base-assisted entity translation, transliteration, multilingual transfer learning, and LoRA-based adaptation of the NLLB-200 distilled 600M model.

The official results showed substantial variation across language pairs. Arabic–Hindi emerged as the strongest pair, with Arabic-to-Hindi obtaining the best relative position among the BSVLP submissions. Bangla-related directions also showed encouraging semantic performance, particularly on COMET. In contrast, English–Arabic was the weakest pair in both directions. Several into-Arabic directions showed comparatively lower BLEU and TER performance

- but more competitive COMET scores, indicating that the system often preserved semantic content even when its lexical or surface realization differed from the reference.
- BSVLP results can be characterized as moderate but informative. They demonstrate the feasibility of combining multilingual transfer with entity-aware processing for low-resource Arabic–Asian translation while highlighting the need for stronger language-pair-specific adaptation. Future work will investigate improved multilingual data balancing, neural transliteration, transformer-based NER, and targeted optimization of weaker directions, particularly English–Arabic. Ablation studies will also be useful for quantifying the individual contribution of preprocessing, named-entity handling, and LoRA adaptation.
- ## References
- Aharoni, R., Johnson, M., & Firat, O. (2019). Massively multilingual neural machine translation. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (pp. 3874–3884). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1388>
- Almansor, E. H., & Al-Ani, A. (2017). Translating dialectal Arabic as low resource language using word embedding. In Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017 (pp. 52–57). INCOMA Ltd. https://doi.org/10.26615/978-954-452-049-6_008
- Bapna, A., & Firat, O. (2019). Simple, scalable adaptation for neural machine translation. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (pp. 1538–1548). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1165>
- Basit, A., Azeemi, A. H., & Raza, A. A. (2024). Challenges in Urdu machine translation. In Proceedings of the Seventh Workshop on Technologies for Machine Translation of Low-Resource Languages (pp. 44–49). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.loresmt-1.4>
- Chronopoulou, A., Stojanovski, D., & Fraser, A. (2023). Language-family adapters for low-resource multilingual neural machine translation. In Proceedings of the Sixth Workshop on Technologies for Machine Translation of Low-Resource Languages (pp. 59–72). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.loresmt-1.5>
- Edunov, S., Ott, M., Auli, M., & Grangier, D. (2018). Understanding back-translation at scale. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (pp. 489–500). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D18-1045>
- Firat, O., Cho, K., & Bengio, Y. (2016). Multi-way, multilingual neural machine translation with a shared attention mechanism. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 866–875). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N16-1101>
- Gabbiche-Braham, S., Bonneau-Maynard, H., & Yvon, F. (2013). Traitement automatique des entités nommées en arabe: Détection et traduction [Automatic processing of Arabic named entities: Detection and translation]. Traitement Automatique des Langues, 54(2), 101–132. <https://aclanthology.org/2013.tal-2.5/>
- Gala, J. P., Chitale, P. A., AK, R., Gumma, V., Doddapaneni, S., Aswanth, K. M., Nawale, J. A., Sujatha, A., Puduppully, R., Raghavan, V., Kumar, P., Khapra, M. M., Dabre, R., & Kunchukuttan, A. (2023). IndicTrans2: Towards high-quality and accessible machine translation models for all 22 scheduled Indian languages. Transactions on Machine Learning Research, 2023. <https://openreview.net/forum?id=vfT4YuzAYA>
- Grundkiewicz, R., & Heafield, K. (2018). Neural machine translation techniques for named entity transliteration. In Proceedings of the Seventh Named Entities Workshop (pp. 89–94). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-2413>
- Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. In International Conference on Learning Representations. <https://openreview.net/forum?id=nZeVKeeFYf9>
- Hu, J., Hayashi, H., Cho, K., & Neubig, G. (2022). DEEP: DEnoising entity pre-training for neural machine translation. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Volume 1, Long Papers (pp. 1753–1766). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.123>
- Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M.,

- Corrado, G., Hughes, M., & Dean, J. (2017). Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics*, 5, 339–351. https://doi.org/10.1162/tacl_a_00065
- Kim, Y., Gao, Y., & Ney, H. (2019). Effective cross-lingual transfer of neural machine translation models without shared vocabularies. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1246–1257). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1120>
- Laskar, S. R., Alam, F., Paul, B., Ahmad, I., Lydia, M. S., & Dadure, P. K. (2026). Findings of the WMT 2026 shared task on low-resource Arabic-Asian language translation. In P. Koehn, B. Haddow, T. Kocmi, & C. Monz (Eds.), *Proceedings of the Eleventh Conference on Machine Translation* (in press). Association for Computational Linguistics.
- Liang, T., Wang, X., Yang, M., Yang, Y., Shi, S., & Tu, Z. (2024). Addressing entity translation problem via translation difficulty and context diversity. In *Findings of the Association for Computational Linguistics: ACL 2024* (pp. 11628–11638). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.findings-acl.691>
- Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). Multilingual denoising pre-training for neural machine translation. *Transactions of the Association for Computational Linguistics*, 8, 726–742. https://doi.org/10.1162/tacl_a_00343
- Modrzejewski, M., Exel, M., Buschbeck, B., Ha, T.-L., & Waibel, A. (2020). Incorporating external annotation to improve named entity translation in NMT. In *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 45–51). European Association for Machine Translation. <https://aclanthology.org/2020.eamt-1.6/>
- Nagoudi, E. M. B., Elmadany, A., & Abdul-Mageed, M. (2022). TURJUMAN: A public toolkit for neural Arabic machine translation. In *Proceedings of the Fifth Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Qur'an QA* (pp. 1–11). European Language Resources Association. <https://aclanthology.org/2022.osact-1.1/>
- Nguyen, T. Q., & Chiang, D. (2017). Transfer learning across low-resource, related languages for neural machine translation. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing, Volume 2: Short Papers* (pp. 296–301). Asian Federation of Natural Language Processing. <https://aclanthology.org/I17-2050/>
- NLLB Team, Costa-jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., Kalbassi, E., Lam, J., Licht, D., Maillard, J., Sun, A., Wang, S., Wenzek, G., Youngblood, A., Akula, B., Barrault, L., Gonzalez, G. M., Hansanti, P., ... Wang, J. (2022). No language left behind: Scaling human-centered machine translation [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2207.04672>
- Ramesh, G., Doddapaneni, S., Bheemaraj, A., Jobanputra, M., AK, R., Sharma, A., Sahoo, S., Diddee, H., J, M., Kakwani, D., Kumar, N., Pradeep, A., Nagaraj, S., Deepak, K., Raghavan, V., Kunchukuttan, A., Kumar, P., & Khapra, M. S. (2022). Samanantar: The largest publicly available parallel corpora collection for 11 Indic languages. *Transactions of the Association for Computational Linguistics*, 10, 145–162. https://doi.org/10.1162/tacl_a_00452
- Sajjad, H., Abdelali, A., Durrani, N., & Dalvi, F. (2020). AraBench: Benchmarking dialectal Arabic–English machine translation. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 5094–5107). International Committee on Computational Linguistics. <https://doi.org/10.18653/v1/2020.coling-main.447>
- Sennrich, R., Haddow, B., & Birch, A. (2016a). Improving neural machine translation models with monolingual data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers* (pp. 86–96). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P16-1009>
- Sennrich, R., Haddow, B., & Birch, A. (2016b). Neural machine translation of rare words with subword units. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers* (pp. 1715–1725). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P16-1162>
- Shapiro, P., & Duh, K. (2018). Morphological word embeddings for Arabic neural machine translation in low-resource settings. In *Proceedings of the Second Workshop on Subword/Character Level Models* (pp. 1–11). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-1201>
- Ugawa, A., Tamura, A., Ninomiya, T., Takamura, H., & Okumura, M. (2018). Neural machine translation incorporating named entity. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 3240–3250). Association for Computational Linguistics. <https://aclanthology.org/C18-1274/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (Vol. 30, pp.

5998–6008). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>

Zeng, Z., Wang, R., Leng, Y., Guo, J., Xie, S., Tan, X., Qin, T., & Liu, T.-Y. (2023). Extract and attend: Improving entity translation in neural machine translation. In Findings of the Association for Computational Linguistics: ACL 2023 (pp. 1697–1710). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.findings-acl.107>

Zhang, B., Williams, P., Titov, I., & Sennrich, R. (2020). Improving massively multilingual neural machine translation and zero-shot translation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 1628–1639). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.148>

Zoph, B., Yuret, D., May, J., & Knight, K. (2016). Transfer learning for low-resource neural machine translation. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (pp. 1568–1575). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D16-1163>

Table 1: Evaluation Results of Low-Resource Arabic-Asian Language Translation Shared Task 2026 for BVSLP

Sub-task	Translation Direction	BLEU	chrF++	TER ↓	COMET	BERT/mBERT
1A	English-Arabic	12.468	44.8541	81.5223	80.7693	82.0453
1B	Hindi-Arabic	13.2825	48.1994	82.6278	81.8869	82.5518
1C	Bangla-Arabic	12.9225	47.1008	82.3228	81.2549	82.3782
1D	Indonesian-Arabic	12.7467	46.3116	81.8052	81.1868	82.2082
1E	Urdu-Arabic	12.9352	47.0654	82.1182	81.4107	82.3541
2A	Arabic-English	20.1319	47.3402	69.4364	78.8771	77.9019
2B	Arabic-Hindi	24.672	50.0713	66.1223	68.8786	84.3182
2C	Arabic-Bangla	13.7988	48.0035	77.2467	77.6456	82.222
2D	Arabic-Indonesian	22.8092	53.0951	65.3573	81.8981	84.2562
2E	Arabic-Urdu	20.9917	47.5932	70.3681	73.808	82.4422