Findings of WMT 2024 Shared Task on Low-Resource Indic Languages Translation

Partha Pakray
NIT Silchar

Santanu Pal
Wipro AI Lab45

Advaitha Vetagiri
NIT Silchar

Reddi Mohana Krishna
NIT Silchar

Arnab Kumar Maji
North-Eastern Hill University

Sandeep Kumar Dash
NIT Mizoram

Lenin Laitonjam
NIT Mizoram

Lyngdoh Sarah
North-Eastern Hill University

Riyanka Manna
Amrita Vishwa Vidyapeetham Amaravati

Abstract

This paper presents the results of the low-resource Indic language translation task, organized in conjunction with the Ninth Conference on Machine Translation (WMT) 2024. In this edition, participants were challenged to develop machine translation models for four distinct language pairs: English–Assamese, English-Mizo, English-Khasi, and English-Manipuri. The task utilized the enriched IndicNE-Corp1.0 dataset, which includes an extensive collection of parallel and monolingual corpora for northeastern Indic languages. The evaluation was conducted through a comprehensive suite of automatic metrics—BLEU, TER, RIBES, METEOR, and ChrF—supplemented by meticulous human assessment to measure the translation systems' performance and accuracy. This initiative aims to drive advancements in low-resource machine translation and make a substantial contribution to the growing body of knowledge in this dynamic field.

1 Introduction

The low-resource Indic language translation field has witnessed significant advancements, particularly marked by the success of last year's Indic MT Shared Task. This initiative, organized alongside the Eighth Conference on Machine Translation (WMT) 2023[1] (Pal et al., 2023), demonstrated the potential and necessity of focusing on low-resourced languages. Building on the momentum and achievements of last year's task, we are pleased to continue our efforts with the Indic MT Shared Task for the Ninth Conference on Machine Translation (WMT) 2024[2].

Low-resource Indic languages represent a vast and diverse array of languages spoken across India. Despite their deep cultural and linguistic heritage, these languages face significant challenges due to limited resources and institutional support. The obstacles are multifaceted, including smaller speaker populations, minimal governmental backing, insufficient documentation, and restricted access to modern technological tools.

India is celebrated for its linguistic diversity, with many languages spoken throughout the subcontinent. The Eighth Schedule of the Indian Constitution officially recognizes 22 languages, granting them substantial governmental support and resources. However, numerous other languages, particularly those spoken by indigenous and minority communities, often remain marginalized and under-supported. These low-resource languages encounter additional barriers, such as the absence of standardized scripts, limited lexical resources, and a dearth of linguistic research. These factors, combined with the lack of formal educational resources and declining inter-generational transmission, threaten their preservation and vitality. As a result, many of these languages risk becoming endangered, underscoring the urgent need for targeted efforts to document, revitalize, and sustain them in the face of ongoing challenges.

Given these challenges, our initiative is dedicated to documenting, revitalizing, and supporting low-resource Indic languages through innovative technological solutions. The previous year's Indic MT Shared Task concentrated on four language pairs: English–Assamese, English–Mizo, English–Khasi, and English–Manipuri — utilizing the enriched IndicNE-Corp1.0 dataset (Pal et al., 2023). The success of this task highlighted the critical need for sustained efforts in this domain. Our ongoing objective is to foster advancements in machine translation and natural language processing tailored to these languages.

The evaluation of this task employs a comprehensive set of metrics, incorporating both automatic measures—such as BLEU (Papineni et al., 2002),

[1] https://www2.statmt.org/wmt23/indic-mt-task.html

[2] https://www2.statmt.org/wmt24/indic-mt-task.html

Proceedings of the Ninth Conference on Machine Translation, pages 654–668
November 15-16, 2024

TER (Snover et al., 2006), RIBES (Isozaki et al., 2010), METEOR (Banerjee and Lavie, 2005), and ChrF (Popović, 2015)—and rigorous human assessments. This dual approach ensures a thorough evaluation of the translation systems' performance, accuracy, and cultural fidelity.

Through this ongoing initiative, we aim to make a significant contribution to the preservation of linguistic diversity and cultural heritage, thereby supporting the rights and identities of minority language communities in India. By leveraging cutting-edge technologies, we strive to create a lasting impact and propel the field of low-resource language translation forward, ensuring these languages not only survive but thrive in the digital age.

2 Languages

2.1 Khasi Language and Its Dialects

Khasi belongs to the Austro-Asiatic family of languages spoken in the central and eastern regions of Meghalaya. Before 1813, the Khasi lacked its own script. During the period from 1813 to 1814, the Bengali script was employed to translate the Bible into Khasi, owing to the widespread literacy in Bengali at that time. By 1816, some translated versions of the Gospel of Matthew had been printed and distributed among Khasi speakers who were literate in Bengali. However, it was not until 1841, with the arrival of a Welsh missionary, that the Roman script was introduced, and translations were subsequently made into the standard dialect, specifically the Sohra variety.

Khasi exhibits significant dialectal diversity. Grierson (1904) identified four dialects of Khasi: Standard Khasi, Pnar or Synteng, Lyngngam, and War. Acharya (1971) reaffirmed Grierson's classification and noted the existence of additional sub-dialects, such as Bhoi, spoken in the northern open lands of Meghalaya. Bareh (1977) offers a more comprehensive list of Khasi dialects, primarily based on their geographical distribution:

- Amwi in the southern Jaiñtia hills,
- Shella in the southern Khasi hills,
- Warding in the south of the Khasi hills,
- Myriaw, Nongkhlaw, Nongspung, Maram, and Mawiang in the mid-western area of the Khasi hills,
- Cherra in the mid-southern hills,
- Mylliem, Laitlyngkot, Nongkrem, and Lyniong-Khasi in the central parts,
- Jowai in the central Jaintia hills,
- Bhoi in the north-east Khasi hills,
- Manar, Nongwah, and Jirang in the north Khasi hills,
- Khatarblang (Mawpran) in the mid-southern region, and
- Nongstoin and Langrin in the west Khasi region.

Bareh further adds that several sub-dialects exhibit variations within each group, particularly in phonology. Among these, Amwi is considered the most typical dialect. Compared to other dialects, Amwi appears to be the most rudimentary and is generally not intelligible to speakers of neighbouring dialects such as Jowai or Khad ar Blang. Amwi is said to be more agglutinative in form, potentially preserving its Mon-Khmer heritage. While its grammar resembles Jowai's, notable differences exist in morphology and phonology. Despite these distinctions, the Amwi speakers are familiar with their neighbouring dialects and can adopt them for communication.

Bareh (1977) categorizes the aforementioned dialects into three major branches:

1. **Eastern dialects**:
    - Jowai (Central Highlands),
    - Amwi and the War dialects (in the south), and
    - Bhoi Synteng in the north.
2. **Central dialects**:
    - Nongphlang or Nonglum, Cherra, and related dialects such as Nongkrem, Mylliem, Nongkhlaw, Nongspung, Rambrai, Mawsynram, Maram, Laitlyngkot, Mawphlang, etc.,
    - Bhoi East (in the north), consisting of Mawrong, Bhoi Lymbong, etc., and
    - Bhoi West (in the north), consisting of Manar, etc.,
    - War Shala (in the south), and
    - Warding (in the south).
3. **Western dialects**:

- Nongstoin
- Lyngam
- Langrin

Addressing the abovementioned dialects, Bareh notes that numerous sub-dialects exhibit phonological variations within each group. Daladier (2007:341), cited in Sidwell (2009), comments on the Mon-Khmer language group, which includes Khasi, noting that it comprises three main branches. Although now standardized and formalized through written use, Khasi retains conservative unwritten dialects, particularly in the War region. Other notable dialects include Pnar and War, with War further subdivided into four sub-dialect groups: Nongtalang, Amvi, Tremblang, and Shella. The subclassification of Pnar dialects remains largely unexplored. Additionally, Pnar-War and War-Khasi dialects are spoken in several Jaintia villages. The War dialects of Khasi are divided into two groups: War-Khasi and War-Jaintia, spoken in the southeast corners of the Khasi and Jaintia Hills districts, respectively. Grierson (1904) also discusses the War dialects.

For the shared task, we have utilized the Sohra (Cherra) dialect of Khasi as the standard form for translation purposes. This dialect, recognized for its historical significance and broad usage in educational and religious contexts, has been established as the standardized variant of Khasi following its formalization through the introduction of the Roman script in 1841. By employing the Sohra dialect, we ensure consistency and accessibility for participants, reflecting the widely accepted linguistic norm within the Khasi-speaking community.

2.2 Introduction: About the Manipuri Language

Manipuri, also known as Meiteilon, is a Sino-Tibetan language predominantly spoken in the northeastern Indian state of Manipur. It is recognized as one of the 22 scheduled languages of India and serves as the lingua franca among various ethnic communities in the region, fostering communication and cultural exchange.

The language boasts a rich literary heritage, with a history of written texts dating back to ancient times. Manipuri uses the Meitei script, also known as Meitei Mayek, alongside the Bengali script for writing purposes. Despite its cultural significance, Manipuri faces linguistic preservation and modernization challenges, particularly in the digital era. There is a pressing need for computational resources and tools to support the language, which is vital for its continued use and growth.

In recent years, there has been growing interest in developing natural language processing (NLP)(Allen, 2003) tools and resources for under-resourced languages like Manipuri. However, several challenges persist in this area for the Manipuri language (Gyanendro Singh et al., 2016). One of the primary issues is the limited availability of annotated corpora and linguistic resources, which are essential for training robust machine learning models. This scarcity hinders the development of accurate NLP applications such as machine translation (Pal et al., 2023), sentiment analysis (Singh and Singh, 2017), and speech recognition (Gyanendro Singh et al., 2016).

Another significant challenge is the complexity of the Manipuri script and its morphological structure. The language exhibits rich inflectional morphology, making it difficult to apply standard NLP techniques that are typically designed for resource-rich languages like English. Moreover, the lack of standardization in digital representation further complicates computational processing, as existing tools often struggle with script conversion and text normalization.

Current research efforts are focused on addressing these challenges by creating linguistic resources, developing language-specific algorithms, and adapting existing NLP frameworks to better accommodate the unique characteristics of Manipuri. However, much work remains to be done to bridge the gap between Manipuri and other well-resourced languages in the digital domain.

2.3 Introduction: About the Assamese Language

Assamese, an Indo-Aryan language, is predominantly spoken in the northeastern Indian state of Assam. It serves as the official language of Assam and plays a crucial role as a lingua franca among various ethnic groups in the region, facilitating communication and cultural exchange. Assamese is also one of the 22 scheduled languages of India, underscoring its significance in the country's linguistic landscape.

The Assamese language has a rich literary tradition, with its roots extending back to the early medieval period. The script used for Assamese is derived from the ancient Brahmi script, and over time, it evolved into its current form. Despite its historical and cultural importance, Assamese faces challenges in the modern era, particularly in the field of language technology. The development of computational tools and resources for Assamese is critical for its preservation and growth, especially in an increasingly digital world.

2.4 Introduction: About the Mizo Language

Mizo, a member of the Tibeto-Burman language family, is predominantly spoken in the northeastern Indian state of Mizoram. It serves as the primary language of communication among the Mizo people and is also spoken by various ethnic groups in neighbouring states and regions, including Manipur, Tripura, Assam, and even parts of Myanmar and Bangladesh. Mizo is recognized for its tonal nature and distinct phonological features, which make it a unique language in the Tibeto-Burman group.

The language has a rich oral tradition, encompassing folktales, songs, and cultural narratives that reflect the heritage of the Mizo people. The development of the written form of Mizo began in the late 19th century with the introduction of the Roman script by Christian missionaries, which facilitated the transcription of the language and the creation of written literature. Today, Mizo has a well-established literary tradition, with a substantial body of work ranging from poetry to modern prose. Despite its cultural significance, Mizo faces linguistic preservation and development challenges, particularly in the context of modern technological advancements and digital communication.

3 Low-Resource Indic Language Translation 2024 Shared Task

3.1 Overview and Task Description

Building upon the resounding success of the "Shared Task: Low-Resource Indic Language Translation" at WMT 2023, which witnessed enthusiastic participation from around the globe, we are excited to announce the continuation of this initiative at the Ninth Conference on Machine Translation (WMT 2024). The advances in machine translation (MT) have significantly enhanced the performance of translation systems, especially with the adoption of techniques such as multilingual translation and transfer learning. Despite these advancements, extending coverage to diverse low-resource languages remains a formidable challenge due to the scarcity of parallel data needed to train robust MT systems.

The WMT 2024 Indic Machine Translation Shared Task addresses this challenge by focusing on low-resource Indic languages from diverse language families. This year, the task emphasizes the following language pairs: English-Assamese, English-Mizo, English-Khasi, and English-Manipuri. Additionally, there was an intended focus on English-Nyishi; however, this category was cancelled due to issues with training data. Similarly, other planned language pairs under the category with very limited training data, such as English-Bodo, English-Mising, and English-Kokborok, were also cancelled for this year.

3.2 Categories

This year's task features two main categories based on the availability of training data:

3.2.1 Category 1: Moderate Training Data Available

- English ⇔ Assamese (en-as)
- English ⇔ Mizo (en-lus)
- English ⇔ Khasi (en-kha)
- English ⇔ Manipuri (en-mni)

3.3 Goal

The central objective of this shared task is to develop machine translation systems that produce high-quality translations despite the constraints posed by limited data availability. Participants are encouraged to explore several innovative strategies, including:

- **Monolingual Data Utilization**: Effectively leveraging monolingual data to enhance translation quality.
- **Multilingual Approaches**: Investigating the benefits of cross-lingual transfer for low-resource language pairs.
- **Transfer Learning**: Adapting models trained on resource-rich language pairs to target low-resource languages.
- **Innovative Techniques**: Experimenting with novel methods specifically tailored to low-resource settings.

3.4 Data

3.4.1 Training

The datasets used for this task include parallel and monolingual corpora for Assamese, Khasi, Mizo, and Manipuri, drawn from the IndicNE-Corp1.0 dataset. While the dataset for Nyishi was planned, it remains unavailable this year due to data quality issues.

3.4.2 Testing

For the testing section, we have created 1000 language pair sentences for each of the following language pairs:

- English ⇔ Assamese (en-as)
- English ⇔ Mizo (en-lus)
- English ⇔ Khasi (en-kha)
- English ⇔ Manipuri (en-mni)

The first 500 sentences are provided in English to be translated into the specific target language, and the last 500 sentences are provided in the target language to be translated into English.

3.5 Evaluation

The evaluation will be conducted using both automatic and human evaluation methods to ensure a comprehensive assessment of the translation systems. Automatic evaluation metrics include BLEU, TER, RIBES, METEOR, and ChrF. In addition, native speakers will perform human evaluations to assess the quality of the translation more rigorously.

4 Dataset

4.1 Training

The dataset for the WMT 2024 Shared Task on Low-Resource Indic Language Translation is primarily based on the IndicNE-Corp1.0 dataset [3]. This corpus was built by aggregating datasets from previous research, including significant contributions from (Laskar et al., 2020) (Laskar et al., 2022), (Khenglawt et al., 2022), and (Laitonjam and Ranbir Singh, 2021). The compiled datasets encompass both parallel and monolingual corpora across four languages: Assamese, Mizo, Khasi, and Manipuri.

[3] https://data.statmt.org/wmt23/indic-mt/

In earlier studies, we focused on developing parallel and monolingual corpora for English ⇔ Assamese (en-asm) (Laskar et al., 2020, 2022), English ⇔ Mizo (en-lus) (Khenglawt et al., 2022), English ⇔ Khasi (en-kha) (Laskar et al., 2021), and English ⇔ Manipuri (en-mni) (Laitonjam and Ranbir Singh, 2021). The data was sourced from a variety of online platforms including the Bible, multilingual dictionaries (such as Xobdo and Glosbe), multilingual question papers, PMIndia (Haddow and Kirefu, 2020), web pages, blogs, and online newspapers.

Table 1 shows the detailed statistics of the parallel datasets used for training and validation for each language pair.

Type	Sentences	Tokens (eng)	Tokens (target)
Assamese	50,000	969,623	825,063
Mizo	50,000	981,468	1,062,414
Khasi	24,000	729,930	875,545
Manipuri	21,687	390,730	330,319

Table 1: Parallel data statistics for train and validation.

In addition to the parallel corpora, we also made monolingual data available for each language, which is presented in Table 2.

Language	Size (MB)	Sentences	Tokens
Assamese	805	2,624,715	49,232,154
Mizo	145	1,909,823	27,936,225
Khasi	104	182,737	22,140,361
Manipuri	716	2,144,897	36,514,693

Table 2: Monolingual data statistics for Assamese, Mizo, Khasi, and Manipuri languages.

4.2 Testing

The testing dataset for the 2024 shared task was meticulously curated to present a substantial challenge beyond previous years' datasets. It comprised 1000 samples for each language pair, spanning four distinct and diverse domains: News, Travel, Sports, Entertainment, and Business. This domain-specific distribution aimed to comprehensively evaluate models' performance across varied and complex linguistic contexts, reflecting real-world translation demands. A collaborative approach was employed to create these testing samples, involving four specialized teams, each dedicated to one domain. These teams were provided 1000 English sentences, which they translated into their assigned target languages. The translation teams were instructed to maintain high fidelity to the source mate-

Language Pair	Domain	Source Sentences	Target Sentences	Task
en-as	Sports and Travel	500	500	English to Assamese
en-lus	Sports and Travel	500	500	English to Mizo
en-kha	Sports and Travel	500	500	English to Khasi
en-mni	Sports and Travel	500	500	English to Manipuri
en-as	Entertainment and Business	500	500	Assamese to English
en-lus	Entertainment and Business	500	500	Mizo to English
en-kha	Entertainment and Business	500	500	Khasi to English
en-mni	Entertainment and Business	500	500	Manipuri to English

Table 3: Domain-specific distribution of the test dataset for each language pair.

rial while ensuring the translations were idiomatic and contextually appropriate for each domain.

The test set release process was intentionally staged to introduce additional complexity and rigour. In the first phase, 500 English sentences were released, requiring participants to translate these into the target languages. This forward translation task required participants to demonstrate their models' proficiency in capturing nuances and domain-specific terminology in the target languages. In the second phase, 500 sentences in the target languages were provided, requiring translation back into English. This reverse translation task assessed the models' ability to accurately render the meaning, tone, and subtleties of the original sentences in English, thus testing bidirectional translation capability. The combined forward and reverse tasks aimed to evaluate the accuracy, fluency, and idiomatic correctness of the translations. The careful selection of diverse domains and the structured release of the test set was intended to challenge the generalization capabilities of the participating models. The goal was to ensure that only the most robust models, capable of handling a wide range of real-world scenarios, would excel.

This approach ensures a rigorous and multi-faceted evaluation, capturing the subtleties of each language pair's translation performance across different domains.

5 Participants and System Descriptions

In this shared task, total of 12 teams registered and contributed, as indicated in table 8, the released dataset have been distributed among participants. In table 7, we have compiled the system outputs submitted by participants, encompassing both primary and contrastive submission types.

DLUT-NLP (Ju et al., 2024)**:** The participant for low-resource translation tasks involving English-Assamese, English-Mizo, English-Khasi, and English-Manipuri language pairs. It utilized a transformer-based model, with monolingual data for pre-training and parallel data for fine-tuning. Enhancements included back-translation, oversampling, and model averaging, along with knn-mt technology during inference, supported by a datastore created from parallel data.

Language Pair	Submissions
English - Assamese	11 (primary), 6 (contrastive)
English-Mizo	10 (primary), 5 (contrastive)
English-Khasi	10 (primary), 6 (contrastive)
English-Manipuri	10 (primary), 6 (contrastive)

Table 4: Number of participants in the low-resource Indic language translations

A3-108 (Yadav et al., 2024)**:** The team tackled low-resource machine translation by implementing control mechanisms in transformer-based NMT models. They encoded the target sentence length as a control token in the source sentence for eight language pairs: English-Assamese, Manipuri, Khasi, and Mizo. Four variations of this encoding were tested against baseline models. Two systems were submitted for each language pair: a primary system using control tokens based on the target-to-source token length ratio, and a contrastive baseline system without control tokens. All models were trained on the provided dataset.

SRIB-NMT (Patil et al., 2024)**:** The team participated in the WMT-24 challenge for translating English to four low-resource Indic languages. They used transformer models for both their primary and contrastive systems. The primary system involved pre-training language models on large amounts of text data before fine-tuning them for translation. The contrastive system improved upon this by further fine-tuning a pre-trained translation model using a technique called LoRA, resulting in better translation quality.

YES-MT (Bhaskar and Krishnamurthy, 2024)**:**

The team participated, focusing on four language pairs: English to Assamese, Khasi, Manipuri, and Mizo. Their primary systems used Transformer models trained from scratch. In contrast, contrastive systems applied transfer learning with fine-tuning techniques like LoRA and Supervised Fine-Tuning (SFT) on pre-trained models such as IndicTrans2 and LLaMA 3. Their experiments explored the effectiveness of these approaches, including quantization, in enhancing translation quality for low-resource languages.

HW-TSC (Wei et al., 2024)**:** The team participated in the WMT-24 challenge for translating English to four low-resource Indic languages. They used transformer models for both their primary and contrastive systems. The primary system involved pre-training language models on large amounts of text data before fine-tuning them for translation. The contrastive system improved upon this by further fine-tuning a pre-trained translation model using a technique called LoRA, resulting in better translation quality.

CycleL (Sören Dréano, 2024)**:** The team developed a novel self-supervised Neural Machine Translation (NMT) model called CycleGN. Unlike traditional NMT models, CycleGN doesn't require parallel data. It utilizes Cycle Consistency Loss (CCL) and Masked Language Modeling (MLM) for training. The model was tested on low-resource language pairs Spanish-Aragonese and Spanish-Asturian using PILAR datasets as part of the WMT24 Shared Task. Despite computational challenges and early training termination, the results demonstrated the potential of self-supervised learning for low-resource translation scenarios.

NLIP_Lab-IIITH (Sahoo et al., 2024)**:** The participated team aiming to improve Manipuri and Khasi translations. They utilized mBART and IndicTrans2 models as baselines, incorporating data augmentation techniques like backtranslation and data filtering with fine-tuned LaBSE. Despite limited data, iterative fine-tuning on enhanced datasets led to significant improvements in translation quality, as measured by BLEU, chrF, and TER metrics.

MTNLP-IIITH (P M et al., 2024)**:** The team tackled the WMT24 Low-Resource Indic NMT challenge for Manipuri and Khasi, employing mBART and IndicTrans2 models. To overcome data scarcity, they implemented backtranslation and LaBSE-based data filtering. Despite computational constraints, iterative fine-tuning on the processed data yielded substantial enhancements in translation quality as assessed by BLEU, chrF, and TER metrics.

SPRING-IITM (Sayed et al., 2024)**:** The team developed a robust translation model for four low-resource Indic languages: Khasi, Mizo, Manipuri, and Assamese. They expanded their training corpus using back translation on monolingual datasets and fine-tuned the pre-trained NLLB 3.3B model for Assamese, Mizo, and Manipuri, achieving superior performance over the baseline. For Khasi, they introduced special tokens and trained the model on a custom Khasi corpus, demonstrating significant improvements in translation quality for all four languages.

JUNLP: The participant focused on developing a translation system for four low-resource Indic languages: Assamese, Manipuri, Mizo, and Khasi, which are widely spoken in India's North Eastern zone. They combined all language data into a single system using Transformer architecture, enabling translation from English to any of these languages within the same framework. Their approach addresses the challenges posed by the scarcity of data for these languages.

SRPH-LIT (Roquea et al., 2024)**:** The team from Samsung R&D Institute Philippines joined the WMT 2024 Low-Resource Indic Language Translation task, focusing on the translation of the following pairs: English ⇔ Assamese, English ⇔ Mizo, English ⇔ Khasi, and English ⇔ Manipuri. In both directions, they adopt the standard sequence-to-sequence Transformer model for translation. The following techniques are data augmentation by back-translation, noisy channel reranking, and checking a multilingual model, which is trained on all the combined language pairs.

ADAPT-MT (Gajakos et al., 2024)**:** The ADAPT-MT team participated in the WMT 2024 Low-Resource Indic Language Translation task, focusing on Assamese-to-English and English-to-Assamese. They leveraged Large Language Models (LLMs) as their base systems, employing strategies like fine-tuning with WMT data, few-shot prompting, and efficient data extraction techniques to enhance translation quality. Their approaches were evaluated using BLEU, ChrF, WER, and COMET metrics, showing effective improvements in translating low-resource languages.

Team Name	Name of University/Lab/Industry/Group
AI Lab-IITI	Indian Institute of Technology Indore
NLIPLab-IITH	**Natural Language and Information Processing Lab at IIT Hyderabad, India**
GUIT-NLP	Gauhati University
ATULYA-NITS	National Institute of Technology, Silchar
Lokkhi	Central Institute of Technology
CFILT-IITB	Indian Institute of Technology Bombay
CNLP-NITMZ	NIT MIZORAM
NITS-CNLP	National Institute of Technology, Silchar
DCU-ADAPT	Dublin City University
onemt	IIIT-H
CL-IIITM	Indian Institute of Information Technology
A3-108	**International Institute of Information Technology - Hyderabad**
SRIB-NMT	**Samsung Research Institute**
JUNLP	**Jadavpur University**
GNLP	GKV
BVSLP	Banasthali Vidyapith
LangMavericks	IIT Madras
BITS-P	Birla Institute of Technology & Science, Pilani
DLUT-NLP	**Dalian University of Technology**
JC-beginners	NJIT
GUIT-NLP	Gauhati University
SHARK	Independent Researcher
bjfu	Beijing Forestry University
Yes-MT	**IIIT Hyderabad**
MTNLP-IIITH	**LTRC, IIIT Hyderabad, India**
SRPH-LIT	**Samsung Research Philippines**
MUNI-NLP	Masaryk University
CycleL	**Dublin City University**
JUMT	Jadavpur University
mbzuai-uhh	MBZUAI, Universität Hamburg
Nexus	Z-AGI Labs
SPRING-IITM	**Indian Institute of Technology, Madras**
BV-SLP	Banasthali Vidyapith
HW-TSC	**Huawei Technologies Co., Ltd.**
SAILors	University of New Haven
NLIPLab_IITH	Natural Language and Information Processing Lab
ADAPT-MT	**ADAPT Centre, Dublin City University**

Table 5: The following table provides an overview of the teams registered for the low-resource Indic language translation task at WMT24 and the datasets provided to them. Participation varied across different language pairs, and only 12 teams in bold completed submissions of both system outputs and system descriptions.

6 Results and Discussion

Results for both directions of the four language pairs in WMT 2024 are detailed as follows: English-Assamese in Table 6, English-Mizo in Table 10, English-Khasi in Table 12, and English-Manipuri in Table 8. This section provides the evaluation scores for teams that submitted system outputs and corresponding papers.

Quantitative results are evaluated using established metrics: BLEU, TER, RIBES, ChrF, and METER. BLEU measures the precision of n-grams in candidate translations relative to reference translations. TER quantifies the number of edits required to align the candidate translation with the reference. RIBES evaluates the correlation between the rank orders of words in candidate and

Team	Test Set	BLEU	TER	RIBES	METEOR	ChrF
DLUT-NLP	en_to_as_primary	0.0723	85.17	0.183	0.2205	0.3786
	as_to_en_primary	0.05	81.7	0.1361	0.2907	0.3398
A3-108	en_to_as_contrastive	0	100.46	0.0347	0.0587	0.1817
	as_to_en_contrastive	0	96.44	0.0378	0.0677	0.1803
	en_to_as_primary	0	99.79	0.0243	0.05134	0.1773
	as_to_en_primary	0	96.19	0.0322	0.0671	0.1883
SRIB-NMT	en_to_as_primary	0.0132	101.83	0.071	0.0744	0.2215
	as_en_en_contrastive	0.2959	34.92	0.3505	0.7409	0.6488
YES-MT	en_to_as_contrastive	0.2568	54.63	0.306	0.5029	0.6518
	en_to_as_primary	0	101.78	0.0105	0.0292	0.1123
HW-TSC	en_to_as_primary	0.2516	55.43	0.2963	0.5124	0.6569
	as_to_en_primary	**0.3228**	32.71	0.3625	0.7606	0.6593
CycleL	en_to_as_primary	0	123.02	0.0029	0.0061	0.0886
	as_to_en_primary	0	101.81	0.0075	0.0249	0.0994
NLIP_Lab-IIITH	en_to_as_primary	0.2058	62.65	0.2674	0.4539	0.6021
	as_to_en_primary	0.1685	55.11	0.242	0.5746	0.5286
	en_to_as_contrastive	0.185	65.79	0.2583	0.433	0.5891
	as_to_en_contrastive	0.1547	58.12	0.2312	0.5326	0.5053
SPRING-IITM	en_to_as_contrastive	**0.2726**	52.79	0.3032	0.513	0.652
	as_to_en_contrastive	0.2669	39.08	0.3308	0.7066	0.6048
JUNLP	en_to_as_primary	0	134.69	0	0.0059	0.0563
SRPH-LIT	en_to_as_primary	0	1195.25	0	0.0001	0.1852
	as_to_en_primary	0	104.67	0.0175	0.0513	0.166
ADAPT-MT	en_to_as_primary	0.1612	65.96	0.2641	0.3927	0.5673
	as_to_en_primary	0.318	33.56	0.3778	0.7537	0.6551
	as_to_en_contrastive	0.3227	33.63	0.372	0.7563	0.6573

Table 6: Performance of teams in the WMT24 low-resource Indic language translation task for the English-Assamese language pair, measured across multiple metrics.

Team	Test Set	Adequacy	Fluency	Overall Rating
DLUT-NLP	en_to_as_primary	2.5	3	2.75
	as_to_en_primary	1.8	2.4	2.1
A3-108	en_to_as_contrastive	0.6	1	0.8
	as_to_en_contrastive	0.1	0.2	0.15
	en_to_as_primary	0.1	0.2	0.15
	as_to_en_primary	0	0	0
SRIB-NMT	en_to_as_primary	0.4	0.6	0.5
	as_en_en_contrastive	3.6	4.1	3.85
YES-MT	en_to_as_contrastive	4.3	4.5	4.4
	en_to_as_primary	0	0	0
HW-TSC	en_to_as_primary	4.1	4	4.05
	as_to_en_primary	4.6	4.7	4.65
CycleL	en_to_as_primary	0	0	0
	as_to_en_primary	0	0	0
NLIP_Lab-IIITH	en_to_as_primary	4.2	4.1	4.15
	as_to_en_primary	4.1	4.1	4.1
	en_to_as_contrastive	3.4	4.1	3.75
	as_to_en_contrastive	3.4	3.5	3.45
SPRING-IITM	en_to_as_contrastive	4.6	4.6	**4.6**
	as_to_en_contrastive	4.3	4.3	4.3
JUNLP	en_to_as_primary	0	0	0
SRPH-LIT	en_to_as_primary	0	0	0
	as_to_en_primary	0	0	0
ADAPT-MT	en_to_as_primary	4.2	4.4	4.3
	as_to_en_primary	4.7	4.7	4.7
	as_to_en_contrastive	4.8	4.8	**4.8**

Table 7: Human evaluation of teams in the WMT24 low-resource Indic language translation task for the English-Assamese language pair, assessed based on Adequacy, Fluency, and Overall Rating.

Team	Test Set	BLEU	TER	RIBES	METEOR	ChrF
DLUT-NLP	en_to_mni_primary	0.0077	96.554	0.0697	0.0711	0.2863
	mni_to_en_primary	0.0315	87.21	0.1297	0.2131	0.3166
A3-108	en_to_mni_contrastive	0	101.73	0.0084	0.0179	0.1401
	mni_to_en_contrastive	0.002	96.45	0.029	0.0615	0.1865
	en_to_mni_primary	0	101.55	0.0072	0.0166	0.1415
	mni_to_en_primary	0	96.5	0.0271	0.0635	0.1889
SRIB-NMT	en_to_mni_primary	0	104.1	0.0191	0.0307	0.1889
	mni_to_en_contrastive	0.1889	53.05	0.2917	0.5943	0.571
Yes-MT	en_to_mni_primary	0	104.03	0.007	0.0214	0.1102
	en_to_mni_contrastive	0.0259	84.47	0.1312	0.1605	0.4438
HW-TSC	en_to_mni_primary	0.0211	87.93	0.1077	0.1406	0.4218
	mni_to_en_primary	**0.2877**	42.16	0.3532	0.6646	0.6106
CycleL	en_to_mni_primary	0	ERROR	0.0054	ERROR	ERROR
	mni_to_en_primary	0	ERROR	0.00542	ERROR	ERROR
NLIP_Lab-IIITH	en_to_mni_primary	0.0258	88.53	0.1176	0.1391	0.4062
	mni_to_en_primary	0.1106	67.02	0.2303	0.4557	0.4935
	en_to_mni_contrastive	**0.0279**	87.22	0.1235	0.1235	0.414
	mni_to_en_contrastive	0.1159	67.49	0.2319	0.4416	0.4748
MTNLP-IIITH	en_to_mni_primary	0	94.77	0.0737	0.0822	0.3325
	mni_to_en_primary	0.0362	94.79	0.1136	0.1873	0.2777
	en_to_mni_contrastive	0.0064	96.46	0.0628	0.0724	0.3191
	mni_to_en_contrastive	0.0484	101.76	0.1087	0.194	0.2662
SPRING-IITM	en_to_mni_contrastive	0.027	84.6	0.1185	0.1567	0.4428
	mni_to_en_contrastive	0.2088	48.77	0.3031	0.61	0.5364
JUNLP	en_to_mni_primary	0	101.25	0.0044	0.0239	0.1471
SRPH-LIT	en_to_mni_primary	0	940.98	0	0.0001	0.1568
	mni_to_en_primary	0	103.28	0.0046	0.0396	0.1729

Table 8: Performance of teams in the WMT24 low-resource Indic language translation task for the English-Manipuri language pair, measured across multiple metrics.

reference translations. ChrF assesses the character n-gram F-score, and METER offers a learned metric for translation quality evaluation.

Furthermore, linguistic experts proficient in the target language pairs were engaged for manual evaluations. Twenty sample sentences from the primary submission type were randomly selected for each language pair. Human evaluators assessed the candidate translations based on three criteria: adequacy, fluency, and overall rating. Adequacy gauges how well the candidate translation captures the meaning of the reference. Fluency assesses whether the candidate translation constitutes a well-formed sentence in the target language, independent of its correspondence to the reference. Overall rating integrates both adequacy and fluency to comprehensively evaluate translation quality.

For example, if the reference translation is “The cat sat on the mat,” a candidate translation such as “The feline rested on the carpet” is deemed adequate as it preserves the meaning of the reference. In contrast, a candidate translation like “The cat ran across the street,” although fluent, is considered inadequate due to the introduction of new information not present in the reference.

The human evaluation parameters are rated on a scale of 0–5, with higher scores reflecting superior quality. The final adequacy, fluency, and overall rating scores are the average ratings assigned to individual test sentences.

Discussion

For the English-Assamese language pair team, SPRING-IITM achieved a high BLEU score, low TER and an overall rating of 4.6 in the human evaluation. They expanded their training corpus using back translation on monolingual datasets and fine-tuned the pre-trained NLLB 3.3B model. For the Assamese-English language pair team, HW-TSC reaches a higher BLEU score, which is even more than the en-as pair, lower TER and team ADAPT-MT gains a higher overall rating of 4.8 in human evaluation.

For the English-Manipuri language pair team, NLIP_Lab_IIITH achieved a higher BLUE score in automatic evaluation and overall rating in human evaluation compared to the other teams. They utilized mBART and In-dicTrans2 models as baselines, incorporating data augmentation techniques like back translation and data filtering with fine-

Team	Test Set	Adequacy	Fluency	Overall Rating
DLUT-NLP	en_to_mni_primary	1.95	3.6	2.75
	mni_to_en_primary	1.9	2.2	2.05
A3-108	en_to_mni_contrastive	1.5	2.2	1.85
	mni_to_en_contrastive	1.0	1.15	1.075
	en_to_mni_primary	1.15	3.45	2.3
	mni_to_en_primary	1.0	1.05	1.025
SRIB-NMT	en_to_mni_primary	1.75	2.4	2.075
	mni_to_en_contrastive	3.95	3.75	3.85
YES-MT	en_to_mni_primary	1.1	2.15	3.25
	en_to_mni_contrastive	4.4	4.2	4.3
HW-TSC	en_to_mni_primary	4.1	4.6	4.35
	mni_to_en_primary	4.8	4.4	**4.6**
CycleL	en_to_mni_primary	1.25	3.65	2.45
	mni_to_en_primary	1.0	1.0	1.0
NLIP_Lab-IIITH	en_to_mni_primary	2.35	3.95	3.15
	mni_to_en_primary	3.1	3.65	3.375
	en_to_mni_contrastive	3.3	4.2	**3.75**
	mni_to_en_contrastive	3.2	3.35	3.275
MTNLP-IIITH	en_to_mni_primary	3.1	3.7	3.4
	mni_to_en_primary	1.0	1.0	1.0
	en_to_mni_contrastive	1.6	2.3	1.95
	mni_to_en_contrastive	1.0	1.0	1.0
SPRING-IITM	en_to_mni_contrastive	3.25	3.75	3.5
	mni_to_en_contrastive	3.8	4.06	3.93
JUNLP	en_to_mni_primary	2.7	2.4	2.55
SRPH-LIT	en_to_mni_primary	1.65	2.3	1.975
	mni_to_en_primary	1.0	1.0	2.0

Table 9: Human evaluation results for teams in the WMT24 low-resource Indic language translation task for the English-Manipuri language pair. The results are presented for Adequacy, Fluency, and Overall Rating on a scale from 0 to 5.

Team	Test Set	BLEU	TER	RIBES	METEOR	ChrF
DLUT-NLP	en_to_lus_primary	0.0075	98.17	0.0725	0.1395	0.2426
	lus_to_en_primary	0.0233	86.79	0.0895	0.2622	0.3162
A3-108	en_to_lus_contrastive	0	92.32	0.0406	0.0978	0.18
	lus_to_en_contrastive	0	97.75	0.0195	0.0544	0.1633
	en_to_lus_primary	0	92.84	0.0328	0.0906	0.173
	lus_to_en_primary	0	96.18	0.0181	0.0587	0.1826
SRIB-NMT	en_to_lus_primary	0	102.98	0.0361	0.062	0.1646
	lus_to_en_contrastive	0.1127	64.94	0.2026	0.4784	0.4482
YES-MT	en_to_lus_primary	0	97.19	0.0445	0.0802	0.1282
	en_to_lus_contrastive	0.0468	73.07	0.176	0.4087	0.4151
HW-TSC	en_to_lus_primary	0.0189	86.38	0.1074	0.1962	0.2873
	lus_to_en_primary	0.0492	76.27	0.1492	0.3646	0.3769
CycleL	en_to_lus_primary	0	101.76	0.008	0.0477	0.1645
	lus_to_en_primary	0	100.48	0.0064	0.0311	0.1487
NLIP_Lab-IIITH	en_to_lus_primary	0.0303	81.89	0.1479	0.2575	0.3396
	lus_to_en_primary	0.0603	76.34	0.1739	0.3716	0.3893
	en_to_lus_contrastive	0	98.53	0.0277	0.0807	0.1792
	lus_to_en_contrastive	0.0849	70.28	0.1819	0.4374	0.4188
SPRING-IITM	en_to_lus_contrastive	**0.066**	66.06	0.1746	0.495	0.4979
	lus_to_en_contrastive	**0.1849**	53.19	0.2684	0.588	0.5044
JUNLP	en_to_lus_primary	0	98.71	0.0589	0.0837	0.1502
SRPH-LIT	en_to_lus_primary	0.0025	94.46	0.0255	0.0834	0.1891
	lus_to_en_primary	0	108.95	0.014	0.0421	0.1431

Table 10: Performance of teams in the WMT24 low-resource Indic language translation task for the English-Mizo language pair, measured across multiple metrics.

Team	Test Set	Adequacy	Fluency	Overall Quality
DLUT-NLP	en_to_lus_primary	0.6	0.5667	0.5833
	lus_to_en_primary	2.7	2.7	2.7
A3-108	en_to_lus_contrastive	0	0	0
	lus_to_en_contrastive	0	0	0
	en_to_lus_primary	0	0	0
	lus_to_en_primary	0	0	0
SRIB-NMT	en_to_lus_primary	0	0	0
	lus_to_en_contrastive	4.3667	4.6667	4.5167
YES-MT	en_to_lus_primary	0	0	0
	en_to_lus_contrastive	2.65	2.8	2.725
HW-TSC	en_to_lus_primary	0.2333	0.1333	0.1833
	lus_to_en_primary	4.2333	4.3	4.2667
CycleL	en_to_lus_primary	0	0	0
	lus_to_en_primary	0	0	0
NLIP_Lab-IIITH	en_to_lus_primary	3.0333	3.2667	3.15
	lus_to_en_primary	3.3333	3.4333	3.3833
	en_to_lus_contrastive	0	0	0
	lus_to_en_contrastive	4.6	4.7	4.65
SPRING-IITM	en_to_lus_contrastive	4.5333	4.5667	**4.55**
	lus_to_en_contrastive	4.7667	4.8333	**4.8**
JUNLP	en_to_lus_primary	0	0	0
SRPH-LIT	en_to_lus_primary	0	0	0
	lus_to_en_primary	0	0	0

Table 11: Updated human evaluation results for teams in the WMT24 low-resource Indic language translation task for the English-Mizo language pair, based on Adequacy, Fluency, and Overall Quality scores.

Team	Test Set	BLEU	TER	RIBES	METEOR	ChrF
DLUT-NLP	en_to_kha_primary	0.0665	78.17	0.1583	0.2939	0.3512
	kha_to_en_primary	0.0253	81.7	0.1223	0.2834	0.2953
A3-108	en_to_kha_contrastive	0.0108	92.92	0.087	0.1209	0.1905
	kha_to_en_contrastive	0	105.76	0.0094	0.0403	0.1358
	en_to_kha_primary	0.011	87.69	0.0873	0.1589	0.2296
	kha_to_en_primary	0	107.7	0.0071	0.0359	0.1348
SRIB-NMT	en_to_kha_primary	0.0054	103.72	0.0821	0.0969	0.1778
	kha_to_en_contrastive	0.042	80.29	0.1205	0.3283	0.318
Yes-MT	en_to_kha_primary	0.0029	159.36	0.0489	0.0511	0.1139
	en_to_kha_contrastive	0.0696	80.74	0.2167	0.2797	0.3541
HW-TSC	en_to_kha_primary	0.0454	87.75	0.1509	0.2134	0.2747
	kh_en_primary	0.0315	79.83	0.1275	0.3044	0.3137
CycleL	en_to_kha_primary	0.0038	91.86	0.0696	0.1399	0.2245
	kha_to_en_primary	0	132.21	0.0062	0.0264	0.0973
NLIP_Lab-IIITH	en_to_kha_primary	0.0475	87.16	0.1406	0.2205	0.2894
	kha_to_en_primary	0.0108	92.83	0.0742	0.1612	0.2488
	en_to_kha_contrastive	0.0521	88.9	0.1515	0.2173	0.288
	kha_to_en_contrastive	0.0312	81.23	0.1263	0.3007	0.312
MTNLP-IIITH	en_to_kha_primary	0.0492	84.79	0.1595	0.2589	0.3316
	kha_to_en_primary	0.0049	ERROR	0.25108	ERROR	ERROR
	en_to_kha_contrastive	0.0359	103.49	0.1106	0.1649	0.2333
	kha_to_en_contrastive	0.006	106.6	0.0487	0.102	0.1731
SPRING-IITM	en_to_kha_contrastive	**0.1212**	63.31	0.1864	0.4453	0.4455
	kha_to_en_contrastive	**0.1047**	61.43	0.2172	0.5042	0.4271
JUNLP	en_to_kha_primary	0	138.36	0.0079	0.0094	0.0344
SRPH-LIT	en_to_kha_primary	0.0044	126.94	0.0533	0.0879	0.1425
	kha_to_en_primary	0	109.81	0.0106	0.0407	0.1336

Table 12: Performance of teams in the WMT24 low-resource Indic language translation task for the English-Khasi language pair, measured across multiple metrics.

Team	Test Set	Adequacy	Fluency	Overall Quality
DLUT-NLP	en_to_kha_primary	2.43	3.2	2.815
	kha_to_en_primary	2.83	3.53	3.18
A3-108	en_to_kha_contrastive	0.33	0.6	0.465
	kha_to_en_contrastive	0.33	0.36	0.345
	en_to_kha_primary	0.33	0.7	0.515
	kha_to_en_primary	1	1	1
SRIB-NMT	en_to_kha_primary	0.33	0.46	0.395
	kha_to_en_contrastive	3.36	3.6	3.48
Yes-MT	en_to_kha_primary	0.33	0.33	
	en_to_kha_contrastive	2.3	2.5	2.4
HW-TSC	en_to_kha_primary	0.43	0.5	0.465
	kha_to_en_primary	4	4.23	4.115
CycleL	en_to_kha_primary	0.33	0.4	0.365
	kha_to_en_primary	0.33	0.4	0.365
NLIP_Lab-IIITH	en_to_kha_primary	1.66	1.76	1.71
	kha_to_en_primary	2.66	2.93	2.795
	en_to_kha_contrastive	2.26	2.3	2.28
	kha_to_en_contrastive	3.47	3.23	3.35
MTNLP-IIITH	en_to_kha_primary	1.93	1.93	1.931
	kha_to_en_primary	2.83	2.83	2.83
	en_to_kha_contrastive	0.76	0.76	0.76
	kha_to_en_contrastive	1.76	1.83	1.795
SPRING-IITM	en_to_kha_contrastive	4.56	4.93	**4.745**
	kha_to_en_contrastive	4.93	4.96	**4.945**
JUNLP	en_to_kha_primary	1	1	1
SRPH-LIT	en_to_kha_primary	0	0	0
	kha_to_en_primary	0	0	0

Table 13: Human evaluation results for teams in the WMT24 low-resource Indic language translation task for the English-Khasi language pair, based on Adequacy, Fluency, and Overall Quality scores.

tuned LaBSE.Team HW-TSC achieved higher BLEU score as well as overall rating in human evaluation for the Manipuri-Englishwhich is significantly higher when compared to the en-mni language pair. They employed a contrastive system which improved upon this by further fine-tuning a pre-trained translation model using a technique called LoRA, resulting in better translation quality.

For the English-Mizo language pair team, SPRING-IITM outperforms all the teams in both directions of the language pairs with higher BLEU and overall ratings in human evaluation. The team developed a robust translation model for four low-resource Indic languages: Khasi, Mizo, Manipuri, and Assamese. They expanded their training corpus using back translation on monolingual datasets and fine-tuned the pre-trained NLLB 3.3B model.

Team SPRING-IITM surpassed all the teams in for the both directions of the language pairs for English-Khasi in automatic and human evaluation. They expanded their training corpus using back translation on monolingual datasets and fine-tuned the pre-trained NLLB 3.3B model. For Khasi, they introduced special tokens and trained the model on a custom Khasi corpus.

Conclusion

The outcomes of the participating teams in the WMT 2024 translation task for four language pairs have been meticulously evaluated using both automated and human metrics. This year's shared task on low-resource Indic language translation utilised the IndicNE-Corp1.0 dataset from WMT 2023, while a newly developed test set was introduced, characterized by a higher difficulty level than the previous year. This enhanced test set aims to better assess the translation capabilities of the models across the participating languages.

The dataset features four under-resourced languages—Assamese, Mizo, Khasi, and Manipuri—from the northeastern region of India. Future initiatives will focus on expanding the dataset by adding more northeastern Indic languages and increasing the corpus size.

We will be incorporating additional languages in the next iteration of the shared task, including English ⇔ Nyishi (en-nshi), English ⇔ Bodo (en-bodo), English ⇔ Mising (en-mrp), and English ⇔ Kokborok (en-trp). This expansion aims to enhance the scope of linguistic diversity, allowing participants to engage with a broader range of low-resource languages.

Acknowledgements

Gratitude is extended to all participants for their active and enthusiastic engagement in this shared task. Appreciation is also expressed to the organisers and reviewers of the Ninth Conference on Machine Translation (WMT) 2024. We also appreciate the Center for Natural Language Processing (CNLP) & Artificial Intelligence (AI) laboratories resources and the research atmosphere at the Dept of CSE, NIT Silchar.

References

James F. Allen. 2003. *Natural language processing*, page 1218–1222. John Wiley and Sons Ltd., GBR.

Satanjeev Banerjee and Alon Lavie. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 65–72, Ann Arbor, Michigan. Association for Computational Linguistics.

Yash Bhaskar and Parameswari Krishnamurthy. 2024. Yes-mt's submission to the low-resource indic language translation shared task in wmt 2024. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Neha Gajakos, Prashanth Nayak, Rejwanul Haque, and Andy Way. 2024. Adapt-mt submissions to wmt 2024 low-resource indic language translation task. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Loitongbam Gyanendro Singh, Lenin Laitonjam, and Sanasam Ranbir Singh. 2016. Automatic syllabification for Manipuri language. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 349–357, Osaka, Japan. The COLING 2016 Organizing Committee.

Barry Haddow and Faheem Kirefu. 2020. Pmindia – a collection of parallel corpora of languages of india.

Hideki Isozaki, Tsutomu Hirao, Kevin Duh, Katsuhito Sudoh, and Hajime Tsukada. 2010. Automatic evaluation of translation quality for distant language pairs. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 944–952, Cambridge, MA. Association for Computational Linguistics.

Chenfei Ju, Junpeng Liu, Kaiyu Huang, and Degen Huang. 2024. Dlut-nlp machine translation systems for wmt24 low-resource indic language translation. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Vanlalmuansangi Khenglawt, Sahinur Rahman Laskar, Santanu Pal, Partha Pakray, and Ajoy Kumar Khan. 2022. Language resource building and English-to-mizo neural machine translation encountering tonal words. In *Proceedings of the WILDRE-6 Workshop within the 13th Language Resources and Evaluation Conference*, pages 48–54, Marseille, France. European Language Resources Association.

Lenin Laitonjam and Sanasam Ranbir Singh. 2021. Manipuri-English machine translation using comparable corpus. In *Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages (LoResMT2021)*, pages 78–88, Virtual. Association for Machine Translation in the Americas.

Sahinur Rahman Laskar, Abdullah Faiz Ur Rahman Khilji Darsh Kaushik, Partha Pakray, and Sivaji Bandyopadhyay. 2021. EnKhCorp1.0: An English–Khasi corpus. In *Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages (LoResMT2021)*, pages 89–95, Virtual. Association for Machine Translation in the Americas.

Sahinur Rahman Laskar, Abdullah Faiz Ur Rahman Khilji, Partha Pakray, and Sivaji Bandyopadhyay. 2020. EnAsCorp1.0: English-Assamese corpus. In *Proceedings of the 3rd Workshop on Technologies for MT of Low Resource Languages*, pages 62–68, Suzhou, China. Association for Computational Linguistics.

Sahinur Rahman Laskar, Riyanka Manna, Partha Pakray, and Sivaji Bandyopadhyay. 2022. A domain specific parallel corpus and enhanced english-assamese neural machine translation. *Computación y Sistemas*, 26(4):1669–1687.

Abhinav P M, Ketaki Shetye, and Parameswari Krishnamurthy. 2024. Mtnlp-iiith: Machine translation for low-resource indic languages. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Santanu Pal, Partha Pakray, Sahinur Rahman Laskar, Lenin Laitonjam, Vanlalmuansangi Khenglawt, Sunita Warjri, Pankaj Kundan Dadure, and Sandeep Kumar Dash. 2023. Findings of the WMT 2023 shared task on low-resource Indic language translation. In *Proceedings of the Eighth Conference on Machine Translation*, pages 682–694, Singapore. Association for Computational Linguistics.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the*

40th Annual Meeting of the Association for Computational Linguistics, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Pranamya Patil, Raghavendra HR, Aditya Raghuwanshi, and Kushal Verma. 2024. Srib-nmt's submission to the indic mt shared task in wmt 2024. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Maja Popović. 2015. chrF: character n-gram F-score for automatic MT evaluation. In *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 392–395, Lisbon, Portugal. Association for Computational Linguistics.

Matthew Theodore Roquea, Carlos Rafael Catalana, Dan John Velascoa, Manuel Antonio Rufinoa, and Jan Christian Blaise Cruza. 2024. Samsung r&d institute philippines @ wmt 2024 indic mt task. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Pramit Sahoo, Maharaj Brahma, and Maunendra Sankar Desarkar. 2024. Nlip-lab-iith low-resource mt system for wmt24 indic mt shared task. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Hamees Sayed, Advait Joglekar, and Srinivasan Umesh. 2024. Spring lab iitm's submission to low resource indic language translation shared task. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Loitongbam Gyanendro Singh and Sanasam Ranbir Singh. 2017. Word polarity detection using syllable features for manipuri language. In *2017 International Conference on Asian Language Processing (IALP)*, pages 206–209.

Matthew Snover, Bonnie Dorr, Rich Schwartz, Linnea Micciulla, and John Makhoul. 2006. A study of translation edit rate with targeted human annotation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, pages 223–231, Cambridge, Massachusetts, USA. Association for Machine Translation in the Americas.

Noel Murphy Sören Dréano, Derek Molloy. 2024. Exploration of the cyclegn framework for low-resource languages. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Bin Wei, Jiawei Zhen, Zongyao Li, Zhanglin Wu, Jiaxin Guo, Daimeng Wei, Zhiqiang Rao, Shaojun Li, Yuanchang Luo, Hengchao Shang, Jinlong Yang, Yuhao Xie, and Hao Yang. 2024. Machine translation advancements of low-resource indian languages by transfer learning. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.

Saumitra Yadav, Ananya Mukherjee, and Manish Shrivastava. 2024. A3-108 controlling token generation in low resource machine translation systems. In *Proceedings of the Low-Resource Indic Language Translation Shared Task, 9th Conference on Machine Translation (WMT24), EMNLP*, November 15-16, 2024, Miami, Florida, USA.