MVL-SIB: A Massively Multilingual Vision-Language Benchmark for Cross-Modal Topical Matching

Anonymous ACL submission

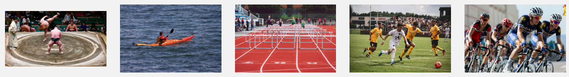
Abstract

Existing multilingual vision-language (VL) benchmarks often only cover a handful of languages. Consequently, evaluations of large vision-language models (LVLMs) predominantly target high-resource languages, underscoring the need for evaluation data for low-resource languages. To address this limitation, we introduce MVL-SIB, a massively multilingual vision-language benchmark that evaluates both cross-modal and text-only topical matching across 205 languages—over 100 more than the most multilingual existing VL benchmarks encompass. We then benchmark a range of of open-weight LVLMs together with GPT-4o(-mini) on MVL-SIB. Our results reveal that LVLMs struggle in cross-modal topic matching in lower-resource languages, performing no better than chance on languages like N'Koo. Our analysis further reveals that VL support in LVLMs declines disproportionately relative to textual support for lower-resource languages, as evidenced by comparison of cross-modal and text-only topical matching performance. We further observe that open-weight LVLMs do not benefit from representing a topic with more than one image, suggesting that these models are not yet fully effective at handling multi-image tasks. By correlating performance on MVL-SIB with other multilingual VL benchmarks, we highlight that MVL-SIB serves as a comprehensive probe of multilingual VL understanding in LVLMs.

Images-To-Sentences (I2S)

Which sentence best matches the topic of the images? The images and the sentences each belong to one of the following topics: "entertainment", "geography", "health", "politics", "science and technology", "sports", or "travel". Choose one sentence from A, B, C, or D. Output only a single letter!

Images



Sentences

A. ```Maroochydore führte am Ende die Rangfolge an, mit sechs Punkten Vorsprung vor Noos a als Zweitem.```

B. ```Es wurden keine schwere Verletzungen gemeldet, jedoch mussten mindestens fünf der zur Zeit der Explosion Anwesenden aufgrund von Schocksymptomen behandelt werden.```

C. ```Finnland ist ein großartiges Reiseziel für Bootstouren. Das „Land der tausend Seen" hat auch Tausende von Inseln – in den Seen und in den Küstenarchipelen.```

D. ```Es ist auch nicht erforderlich, dass Sie eine lokale Nummer von der Gemeinde erhalten, in der Sie leben. Sie können eine Internetverbindung über Satellit in der Wildnis v on Chicken in Alaska erhalten und eine Nummer auswählen, die vorgibt, dass Sie im sonnigen Arizona sind.```

Figure 1: Cross-modal topic matching 'Images-To-Sentence' for German with k=5 reference images.

1 Introduction

Large Vision-Language Models (LVLMs) extend Large Language Models (LLMs) to take images as inputs, leveraging their advanced language capabilities for vision-language (VL) tasks like image captioning and visual question answering (VQA). However, LVLMs are typically trained mainly on English data, leading to significant limitations despite the base LLMs' multilingual abilities. They may fail to follow instructions or struggle to interpret text within images in Non-English languages (Schneider and Sitaram, 2024; Tang et al., 2024). Although many multilingual VL benchmarks exist, they typically cover at most 10 languages (Bugliarello et al., 2022; Liu et al., 2021; Tang et al., 2024, *inter alia*). Only concurrent work has scaled VL evaluation to 100 languages using machine translation (MT) with human post-editing (Vayani et al., 2024). Nevertheless, benchmarks constructed using semi-manual MT cannot support truly low-resource languages adequately, as current MT models lack the necessary quality for these languages. Moreover, existing benchmarks primarily assess lower-level VL semantics through concrete text-image relationships, such as those found in VQA. This underscores the need for VL benchmarks that cover truly low-resource languages and evaluate more abstract VL interactions.

To address these challenges, we introduce the massively multilingual vision-language SIB (MVL-SIB) dataset, which extends the topic labels of the multi-way parallel sentences from SIB-200 (Adelani et al., 2024) by associating each topic with hand-selected images. MVL-SIB evaluates cross-modal image-text topic matching in 205 languages: LVLMs must select one of 4 candidate

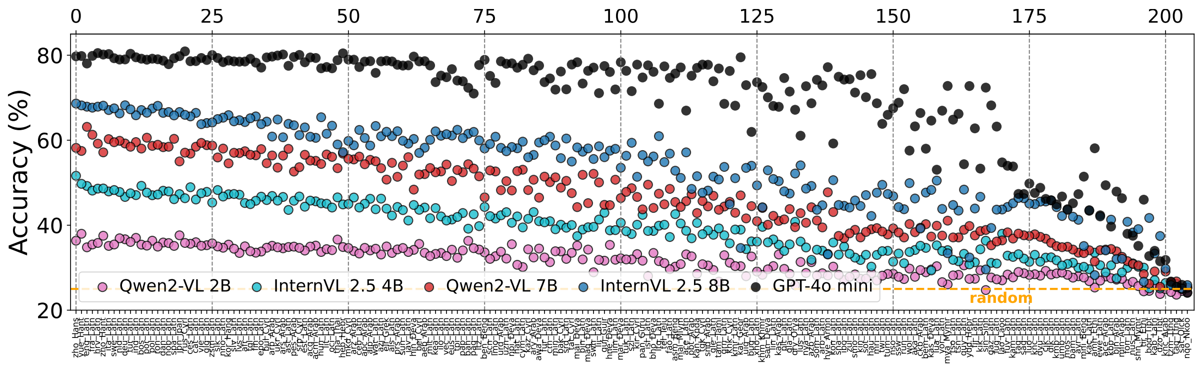


Figure 2: **Images-To-Sentences @ k=3.** The English prompt describes the cross-modal topic matching task, lists all topics, and provides both k=3 reference images and 4 sentences in the corresponding language {`eng_Latn`, . . . , `nqo_Nkoo`}. LVLMs must select the sentence of 4 options that topically fits k=3 reference images. The sentences spanning 205 languages and 7 topics are drawn from SIB-200 (Adelani et al., 2024), while images for the topics were hand-selected (cf. Appendix A.1). An example prompt is shown in Appendix A.7.3; further details are in §4. **Plot.** The x-axis orders the languages of the candidate sentences {`eng_Latn`, . . . , `nqo_Nkoo`}, respectively, by descending performance (y-axis). The top x-axis indicates the running index of each language L_i ($i \in \{1, \ldots, 205\}$).

sentences that best matches the topic of the reference images ('images-to-sentence', cf. Figure 1) or, conversely, choose one of 4 candidate images corresponding to the topic of the reference sentences ('sentences-to-image'). Figure 2 displays the 'images-to-sentence' performance for across all languages in MVL-SIB, sorted in descending order. Notably, GPT-4o-mini performs robustly on the top 125 languages. However, beyond that range its performance declines sharply, falling to chance levels in the lowest-resource languages, such as N'Koo. Bridging these gaps is crucial for developing genuinely inclusive VL technology.

Contributions. 1) MVL-SIB supports parallel VL evaluation in 205 languages on professionally translated texts, a 105 more languages than any other VL benchmark. The tasks, images-to-sentence and sentences-to-images with prefix and postfix images in context, respectively, allows for fine-grained analysis of VL interactions. We also define corresponding text-only tasks by replacing the images with the topic label to compare the VL support (by language) of LVLMs against the text-only support of their underlying LLMs. Both tasks allow to vary the number of included images to analyze the shift from single to multi-image support in LVLMs. **2)** We thoroughly evaluate LVLMs on cross-modal image-text topic matching, finding that task performance is closely associated with both model size and the size of pre-training corpora of the respective languages. We further find that only GPT-4o-mini seizes on multiple references in both cross-modal tasks. Open-weights LVLMs, moreover, favor one of the two tasks, highlighting the asymmetry in their VL support. **3)** We analyze the relationship between stand-alone text and vision-language support in LVLMs by also benchmarking LVLMs on text-only topic matching. The performance gap between matching sentences to reference images or the topic tends to be larger the better the LVLM supports the underlying language. Conversely, the spread in performance between picking the fitting image or topic for reference sentences increases the worse the vision-language support of the LVLM for the evaluated language is. **4)** We correlate MVL-SIB with established multilingual VL benchmarks on the languages shared between the respective pairs of datasets, showing that the MVL-SIB tasks align well with all VL tasks except OCR. We further show that images-to-sentence and sentence-to-image probe distinct aspects of VL interaction, as certain benchmarks correlate more strongly with one task than with the other. This analysis shows that MVL-SIB constitutes a reliable and comprehensive VL benchmark for the lower-resource languages that are not covered by other datasets.

2 Related Work

Multilingual Vision-Language Models. Researchers have extended more English-centric LVLMs like BLIP-2 or LLaVA by continuing to train on multilingual data. Google's PaLI models were the earliest closed-weight models trained on multilingual captions and VQA data; their open-weight PaliGemma followed a comparable strategy. Meanwhile, modern LLMs (e.g., Qwen 2.5, Llama 3, Gemma 2, Aya) have improved in multilingual tasks but frequently fail to respond consistently in non-English languages, particularly in low-resource settings (Schneider and Sitaram, 2024). However, foundational models tend to focus on higher-resource languages and do not fully account for broader linguistic contexts in vision-language tasks. mBLIP is the first open multilingual LVLM trained on image captions and a small set of translated instruct data in 98 languages (Geigle et al., 2024a). Pangea incorporates multicultural dimensions by blending machine-translated data, existing multilingual resources, and synthetic data, on 39 languages (Yue et al., 2024). Most recently, Geigle et al. (2025) studied the composition of training data for multilingual adaptation of LVLMs, observing that only 25-50% of the data needs to be in English. The authors apply their findings when training 'Centurio', which achieves state-of-the-art performance on 14 multilingual VL benchmarks.

Multilingual Vision-Language Benchmarks. Existing datasets span VQA, natural language inference (NLI), image captioning, outlier detection, and culturally grounded QA:

VQA. xGQA extends the questions of the GQA dataset into 8 languages (5 scripts), but the answers in English (Pfeiffer et al., 2022). MaXM offers short-form QA fully in 7 languages, pairing culturally aligned images with same-language QA pairs (Changpinyo et al., 2023). MTVQA focuses on text-heavy VQA in 9 languages (Tang et al., 2024).

Culturally-grounded VQA. CVQA collects culturally diverse images and queries in 31 languages with multiple country-specific variants (Romero et al., 2024). The concurrent ALM-Bench covers 100 languages via MT with GPT-4o followed by human post editing with both generic and cultural multiple-choice and 'true or false' questions, as well as free-form VQA (Vayani et al., 2024).

Visual Reasoning & NLI. XVNLI evaluates cross-lingual visual NLI on 5 languages (Bugliarello et al., 2022). Binary reasoning tasks include MaRVL (Liu et al., 2021) and M5B-VGR (Schneider and Sitaram, 2024), each using linguistically specific images and textual statements. M5B-VLOD presents an outlier detection challenge, where a statement holds true for all but one image (Schneider and Sitaram, 2024).

Multiple-Choice QA. Babel-ImageNet (Geigle et al., 2024c) translates ImageNet labels into nearly 300 languages for multiple-choice object classification. M3Exam and xMMMU also feature multiple-choice VQA in 9 and 7 languages, respectively.

MVL-SIB fills the gaps in existing multilingual VL benchmarks. It provides test data professionally translated to 205 languages, covering over 105 more languages than other benchmarks for which MT cannot synthesize reliable data for. Other VL datasets that eschew MT, such as culturally-grounded VQA benchmarks, typically construct more language-specific non-parallel data, that does not support comparative evaluation across languages. The cross-modal topic matching tasks can be also framed text-only (cf. §3.2), replacing the images that represent topics with the explicit topic labels. Thereby, MVL-SIB enables to ablate the vision-language support from the textual support for a language. Finally, the benchmark allows to vary the number of images provided LVLMs to analyze the support for multi-image reasoning.

3 Dataset and Tasks

3.1 Dataset

For MVL-SIB, we extend the following Flores-based datasets to create a massively multilingual, multi-way parallel VL benchmark for identifying topical associations between images and sentences.

Flores. Flores is a machine translation benchmark containing 3,001 sentences from English Wikipedia paragraphs (Team et al., 2022), professionally translated into over 200 languages.[1]

SIB-200. Adelani et al. (2024) grouped the coarse topical annotations for sentences in the DEV and DEVTEST subsets of Flores into 7 higher-level topics.[2] The resulting SIB-200 dataset is a benchmark for topical classification with 1,004 parallel examples for 205 language variants.

MVL-SIB. For each topic, we first manually collect 10 permissively licensed images that distinctly represent the topic (e.g., `sports`) with minimal overlap or ambiguity (cf. Appendix A.1). We verify that all LVLMs in our study correctly classify the topics for the images when prompted (cf. Appendix A.7.1). We next create 3 different MVL-SIB instances from each of the 1,004 SIB sentences, totaling 3,012 MVL-SIB instances. For each MVL-SIB instance, we couple the respective SIB sentence with (1) a random selection of 5 positive im-

[1] Flores splits 3,001 sentences into DEV (997), DEVTEST (1,012), and TEST (992) sets. The TEST set was not released.

[2] The topics are `entertainment`, `geography`, `health`, `politics`, `science/technology`, `sports`, and `travel`.

ages (same topic) and 4 additional sentences from the same category as the original sentence, as well as (2) 3 negative images and sentences randomly sampled from different topics compared to the starting SIB sentence. The set of sampled sentences and images by instance is maintained across languages.

3.2 Cross-modal & Text-only Topic Matching

We formulate both cross-modal and text-only topic matching tasks based on MVL-SIB. In every task, we present the model with the list of topics that images and sentences may be associated with.[2] Otherwise, it would be unclear along which dimension the model should match images and sentences. The portion of the prompt that introduces the task is provided in English, while the sentences to be topically aligned with images are presented in one of the 205 languages included in MVL-SIB. LLMs reliably perform tasks described in English, even when task-related information is conveyed in other languages (Muennighoff et al., 2022; Romanou et al., 2024). This ensures a fair comparison across all 205 languages, where MT would not accurately preserve the meaning of the prompts. We detail prompts for each task in Appendix A.2.

Cross-modal Topic Matching. Using our text-image samples (cf. §3.1), we define two cross-modal topic matching tasks: Images-To-Sentence (I2S) and Sentences-To-Image (S2I). In I2S, the model must select, from 4 candidates, the sentence that matches the topic of k reference images. Conversely, in S2I, the model chooses, from 4 options, the image that shares the topic with k reference sentences. In both tasks, we present the model with $k \in \{1, 3, 5\}$ references, respectively. These tasks evaluate the model's ability to align high-level visual and textual cues on topics.

Text-only Topic Matching. We construct two tasks by replacing the images in I2S and S2I, that represent the topics, with their corresponding labels (e.g., `sports`). The resulting unimodal tasks, Topic-To-Sentence (T2S) and Sentences-To-Topic (S2T), mirror the cross-modal tasks, I2S and S2I, respectively. For T2S, we evaluate only k=1, since repeating the topic label adds no information. These baseline tasks allow us to delineate between language support and vision-language understanding in LVLMs.

MVL-SIB offers 4 crucial advantages over prior benchmarks. 1) It supports evaluation in 205 languages, covering over 100 more languages than existing benchmarks for which MT models fail to synthesize reliable evaluation data. **2)** MVL-SIB supports ablating language understanding and multimodal reasoning of LVLMs by comparatively evaluating the mirroring text-only and cross-modal topic matching tasks (cf. §3.2). **3)** MVL-SIB enables intricate analysis of single- and multi-image VL interactions in LVLMs by allowing topics to be represented by varying numbers of images in cross-modal tasks. **4)** MVL-SIB comprises higher-level VL reasoning tasks, pairing varied images and diverse texts to test nuanced VL understanding.

4 Experimental Setup

Models. We test state-of-the-art LVLMs Qwen2-VL (Wang et al., 2024), InternVL 2.5 (Chen et al., 2024), Centurio-Qwen (Geigle et al., 2025), and GPT-4o(-mini) across available sizes.[3] Smaller LVLMs are evaluated on all languages, while larger ones (26B+) are tested on subsets (cf. §5.3). For cross-modal topic matching, we also evaluate on mSigLIP-base (Zhai et al., 2023). Trained explicitly for semantic similarity on multilingual image-caption pairs, the ViT represents a strong baseline.[4] Its prediction denotes the choice that has the highest average cosine similarity to the k references.

Image preprocessing. We downsample the images to 640×480 pixels, as the tasks rely on higher-level visual cues for topically associating images and texts rather than finer image details. This can significantly reduce the number of visual tokens input to LVLMs, enabling more efficient inference.

Hyperparameters. We decode text greedily with temperature set to 0.0 to ensure reproducibility.

Metric. We compute the share of prompts for which responses begin with the right letter. If the label is "A", a response such as "A." is also correct.

5 Results and Discussion

We categorize the languages in MVL-SIB based on their 'resourceness'. To do so, we reorganize the language groups from Joshi et al. (2020) into four tiers. We first rank the tiers w.r.t. Wikipedia size and then merge (i) the two highest-resource tiers and (ii) the two lowest-resource tiers.[5] This both better reflects current corpus availability for LLM

[3] We provide details on the LVLMs in Appendix A.3.

[4] The model is available on Huggingface at: google/siglip-base-patch16-256-multilingual.

[5] Sorting by Wikipedia size (in number of pages) swaps tiers 1 and 2; we then merge Tier 0 with the new Tier 1, as well as tiers 4 and 5.

Resourceness	High						Mid			Low					
	ENGLISH			TIER 4 (26)			TIER 3 (32)			TIER 2 (96)			TIER 1 (51)		
k References	**1**	**3**	**5**	**1**	**3**	**5**	**1**	**3**	**5**	**1**	**3**	**5**	**1**	**3**	**5**
Images-To-Sentence: *Select 1 of 4 sentences topically matching k reference images*															
mSigLIP-base	57.7	64.6	66.4	53.3	58.6	59.7	51.4	56.2	57.1	38.9	41.2	41.7	36.1	37.6	38.0
Qwen2-VL 2B	36.3	34.8	34.9	35.5	35.3	34.1	34.5	34.3	33.2	31.0	30.6	30.0	29.5	29.0	28.6
Qwen2-VL 7B	65.8	63.1	58.9	57.7	56.5	51.7	55.4	54.5	49.6	44.3	44.4	40.5	39.6	39.7	36.5
InternVL 2.5 4B	52.5	49.2	48.1	50.3	46.6	47.7	48.6	45.3	46.1	38.7	37.1	37.3	35.4	34.5	34.5
InternVL 2.5 8B	67.7	67.9	68.7	64.6	64.9	65.7	61.2	60.8	61.6	51.0	51.4	51.8	46.1	46.0	46.3
Centurio Qwen	54.8	60.0	62.4	54.2	59.2	60.6	53.4	58.1	58.9	46.6	48.9	49.2	43.0	44.2	44.7
GPT-4o-mini	**68.3**	**78.1**	**77.4**	**71.6**	**79.0**	**78.1**	**72.0**	**78.9**	**77.7**	**63.5**	**68.0**	**66.4**	**56.9**	**60.3**	**58.7**
Sentences-To-Image: *Select 1 of 4 images topically matching k reference sentences*															
mSigLIP-base	56.3	66.0	69.6	51.8	61.6	64.0	49.1	58.3	60.2	36.0	40.4	41.2	32.9	36.3	36.9
Qwen2-VL 2B	41.9	43.1	43.4	41.6	42.5	42.7	40.8	42.4	42.4	33.7	35.6	35.5	31.0	32.7	32.8
Qwen2-VL 7B	71.7	70.4	68.6	65.5	65.5	63.5	64.4	65.3	64.1	50.3	52.5	52.9	43.5	45.9	46.6
InternVL 2.5 4B	47.7	44.5	43.0	38.0	40.3	40.4	36.7	39.6	40.3	30.7	34.4	35.7	28.8	32.1	33.7
InternVL 2.5 8B	66.2	69.0	68.7	57.5	62.5	61.6	52.9	58.5	58.1	43.4	49.8	49.7	39.7	45.9	46.2
Centurio Qwen	35.3	36.1	35.6	31.1	32.9	33.3	31.0	32.8	33.1	28.7	29.7	29.8	28.1	28.7	28.7
GPT-4o-mini	**77.5**	**86.4**	**89.1**	**77.2**	**86.5**	**88.6**	**77.1**	**86.1**	**88.4**	**68.4**	**79.8**	**82.7**	**61.7**	**74.0**	**77.2**

Table 1: **Cross-modal Topic Matching:** LVLMs must select the candidate sentence (image) from 4 choices that topically align with k reference images (sentences). Prompts provided in §A.2. Languages are tiered by Wikipedia sizes (cf. §5). Number of languages in parentheses. **Metric:** share of responses starting with correct option letter. Details in §4. In each column, the best model is emphasized in **bold**, the second-best model is underlined.

Resourceness	High								Mid				Low							
	ENGLISH				TIER 4 (26)				TIER 3 (32)				TIER 2 (96)				TIER 1 (51)			
Task	**Topic-To Sentence**	**Sentences To-Topic**			**Topic-To Sentence**	**Sentences To-Topic**			**Topic-To Sentence**	**Sentences To-Topic**			**Topic-To Sentence**	**Sentences To-Topic**			**Topic-To Sentence**	**Sentences To-Topic**		
k References	**1**	**1**	**3**	**5**	**1**	**1**	**3**	**5**	**1**	**1**	**3**	**5**	**1**	**1**	**3**	**5**	**1**	**1**	**3**	**5**
Qwen2-VL 2B	56.7	86.1	96.5	98.2	49.1	78.4	92.0	95.2	45.2	73.7	89.6	93.6	36.4	53.7	69.3	76.4	33.0	46.4	60.7	68.4
Qwen2-VL 7B	85.7	89.1	95.3	97.5	81.8	84.1	93.3	96.1	80.5	84.1	93.5	96.6	63.7	67.8	81.3	86.7	55.3	59.1	73.7	79.8
InternVL 2.5 4B	81.4	90.6	98.0	**99.1**	72.7	85.2	95.8	97.9	68.2	83.7	95.5	97.7	50.2	67.8	83.4	87.9	44.6	60.7	77.4	83.1
InternVL 2.5 8B	87.0	91.7	**98.1**	99.0	83.0	86.3	96.3	98.3	79.1	82.4	94.1	96.7	65.4	66.8	81.8	86.9	58.1	58.3	74.6	80.9
Centurio Qwen	85.4	89.9	96.7	97.7	83.6	88.0	95.8	97.7	82.6	87.7	96.0	97.9	70.4	73.1	86.8	90.7	64.0	66.1	81.2	85.7
GPT-4o-mini	**88.5**	**92.4**	**98.1**	**99.1**	**89.3**	**91.6**	**98.4**	**99.3**	**89.3**	**91.6**	**98.4**	**99.3**	**80.9**	**82.1**	**93.3**	**95.7**	**73.9**	**74.3**	**88.2**	**92.1**

Table 2: **Text-only Topic Matching:** LVLMs must select the candidate sentence (topic) of 4 choices that aligns topically with the reference topics (k reference sentences). See Table 1 for further details.

pre-training (Xue et al., 2021; Kudugunta et al., 2023), and aligns with downstream performance (cf. Appendix A.3.1). We isolate English from Tier 4, since it is the pivotal language in NLP. The full per-language results by task and model are provided in Appendix A.7.

5.1 Cross-modal Topic Matching

Images-To-Sentence (I2S). The upper segment of Table 1 displays the results for I2S, in which the LVLMs must pick the candidate sentence that topically matches the k reference images.

English. The performance on I2S with English candidate sentences scales well with model size. The small Qwen2-VL 2B performs only slightly better than chance (25% vs. ca. 35%). The comparably sized Qwen2-VL 7B, InternVL 2.5 8B, and Centurio-Qwen 8B peak around 62% to 68% at various k. These models nevertheless non-negligibly trail GPT-4o-mini (78.1%). Among LVLMs, only InternVL 2.5 8B, Centurio-Qwen, and GPT-4o-mini benefit from multiple reference images. When the number of references k increases from 3 to 5, GPT-4o-mini declines slightly in performance, while InternVL and Centurio-Qwen continue to improve marginally (ca. +1%) and more notably (ca. +3-6%), respectively. All other LVLMs deteriorate materially with more reference images (ca. -3-4%). mSigLIP indeed is a strong baseline, trailing only GPT-4o-mini and InternVL 2.5 8B at k=5. The ViT yields large gains with 4 more images (+9.7%).

Tiers. The performance gap of other languages to English correlates well with their resource levels by language tier. When presented with candidate sentences in non-English high-resource languages (cf. Tier 4), GPT-4o-mini even performs better slightly better. For very low-resource languages in Tier 1, such as N'Koo or Tamazight, all models fail to perform better than chance (cf. Appendix A.7.3). Among LVLMs, only GPT-4o-mini remains overall robust for topically matching sentences of low-resource languages to images, whereas other models drop severely in performance (ca. 15-20%). While mSigLIP still performs well, it declines more notably than LVLMs on lower-resource languages.

Sentences-To-Image (S2I). The lower part of Table 1 presents the results for S2I. Here, the models select the candidate image among four options that topically fits the k reference sentences.

English. Performance again correlates well with model capacity. However, in S2I, only GPT-4o-mini significantly seizes on additional references (+13%) to excel with 89%, while models like Qwen2-VL 7B and InternVL 2.5 8B exhibit peak performance at k=1 and k=3, respectively, that taper slightly with more sentence references. Centurio-Qwen performs only slightly better than random (25% vs. ca. 35%). In S2I, mSigLIP is again very strong, second only to GPT-4o-mini at k=5. The encoder once more seizes sizable gains from additional references (+13.3%).

Tiers. In non-English evaluations, the overall trend remains similar, though absolute performance is lower. The gap between high- and low-resource language tiers is evident, as all models yield higher scores across Tiers 4 and 3. GPT-4o-mini maintains robust performance even in the most challenging Tier 1 when provided multiple references (ca.75%).

In sum, both the training protocol and the model size collectively determine whether models favor I2S or S2I. For instance, InternVL 2.5 8B outperforms Qwen2-VL 7B on I2S across the board, while trailing on S2I. Moreover, only GPT-4o-mini consistently seizes on additional references and remains largely robust to the lowest-resource languages on both tasks. This likely stems from insufficient training to enable open-weight LVLMs to perform higher-level VL reasoning with diverse texts and multiple images successfully. For instance, Centurio Qwen was mostly trained on data that prefixes images to text, resulting in low performance when images are postfixed to the context.

5.2 Text-only Topic Matching

Table 2 lists the results for text-only topic matching, in which the images are exchanged with their topic label. These tasks denominate 'upper-bounds' for their cross-modal counterparts to enable ablations of language support and VL support in LVLMs.

Topic-to-Sentence (T2S). In this task, LVLMs choose the sentence that best suits the reference topic. Barring Qwen2-VL 2B, all models perform well on the task. Notably, 5 images should capture the underlying topic well for all models (cf. Appendix A.7.1). Despite that, the gap between text-only and vision-language tasks (cf. Table 1) is sizable across all open-weights models for 'English' (ca. 20%+). In English, GPT-4o-mini and InternVL 2.5 8B achieve the highest accuracies, indicating strong topic comprehension. For non-English languages, while the overall scores are reduced, high-resource languages benefit from richer training signals compared to their low-resource counterparts – models like Qwen2-VL 2B and Centurio-Qwen 8B show a more pronounced drop in the latter, underscoring the impact of language resources.

Sentences-to-Topic (S2T). In S2T, where models choose the topic that best aligns with k reference sentences, performance scales both with model size and the number of references. The gains from additional context ($3{-}5$ k) are particularly notable for larger LVLMs. In English, GPT-4o-mini improves markedly from 92.4% at k=1 to 98.1% at k=3. In non-English languages, similar patterns emerge: high-resource languages consistently yield higher accuracies than low-resource ones, with GPT-4o-mini and InternVL 2.5 8B exhibiting the most stable improvements across varying k. This reinforces the role of model capacity and training data diversity in effective cross-lingual topic matching.

Comparing the results of cross-modal and text-only topic matching sheds further light on the VL interactions for lower-resource languages in LVLMs. The performance gap between matching sentences to reference images and matching them to textual topics tends to narrow regardless of the number of references, likely reflecting their limited textual support in LVLMs. In contrast, the discrepancy between selecting an image versus a topic for reference sentences becomes much more pronounced, especially at k=5. These findings suggest that VL support degrades more sharply than textual support for lower-resource languages in LVLMs.

5.3 Further Analyses

Task Correlation. To compare MVL-SIB with other tasks, we access the results for Qwen2-VL and InternVL 2.5 models on several established multilingual VL benchmarks from Geigle et al.

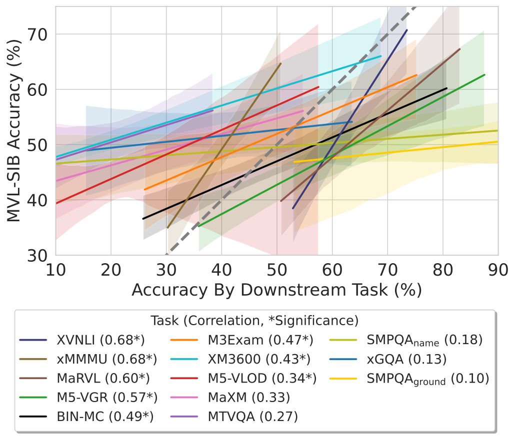


Figure 3: I2S with k=3.

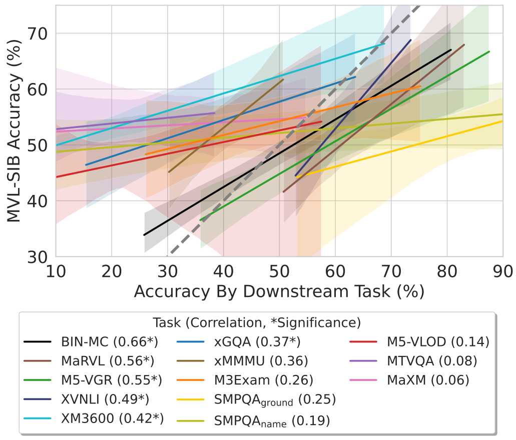


Figure 4: S2I with k=3.

Correlations Between MVL-SIB & Multilingual VL Benchmarks. Pearson correlation coefficients obtained by regressing MVL-SIB performance against performance on multilingual VL tasks on languages common to both datasets, respectively. An asterisk (*) indicates whether the coefficient is statistically significant at $p \leq 0.05$.

(2025).[6] Next, we align the results across languages for the MVL-SIB tasks and the other benchmarks. Finally, we plot the linear regressions of performance on I2S and S2I with k=3 against the performance on the VL benchmarks, respectively, pooled over models, in Figures 3 and 4.[7]

MVL-SIB positively correlates with all tasks in both the I2S and S2I evaluations. However, both the magnitude and statistical significance of these linear relationships vary across VL benchmarks. Both I2S and S2I exhibit the strongest connections with XVNLI (visual inference), BIN-MC (multiple-choice image classification), MaRVL, and M5-VGR (both visually-grounded boolean reasoning). Since these tasks restrict valid answers to a small set of fixed options (i.e., choice letters or 'yes/no/maybe'), LVLMs must engage in higher-level vision-language disambiguation rather than relying solely on lower-level visual cues to solve these tasks. In addition, xMMMU, M3Exam, and M5-VLOD are significantly related only to I2S, whereas xGQA is significantly aligned solely with S2I. The former tasks are structurally similar to I2S, typically presenting one or more images that LVLMs are given as visual context to answer multiple-choice questions. We hypothesize that only S2I is significantly correlated with xGQA, since S2I is more analogous to object-centric benchmarks. In S2I, LVLMs likely leverage targeted semantic cues (e.g., keywords or phrases) from the reference sentences to better disambiguate candidate images by topic. This behavior aligns with lower-level VL tasks such as xGQA (cross-lingual short-form QA) or BIN-MC (multiple-choice object classification), where texts and images are more deliberately connected. In contrast, MTVQA, SMPQA-name, and SMPQA-ground show only weak or statistically insignificant correlations with the MVL-SIB tasks. Since these tasks require LVLMs to comprehend text embedded in images, a low-level, fine-grained VL task, they differ substantially from the higher-level VL reasoning evaluated by MVL-SIB.

Overall, the regression analysis indicates that I2S and S2I capture distinct yet complementary aspects of VL understanding, collectively exhibiting a strong relationship with a broad set of VL tasks. This aligns with our main results (cf. Table 1), which show that different LVLMs may favor one MVL-SIB task over the other. This renders MVL-SIB as a suitable benchmark for evaluating *universal* VL understanding (across 205 languages). It also enables to ablate performance across the key axes of analysis, the task formulation (I2S vs. S2I), the language vs. vision-language support for 205 languages, and the number of images in context.[8]

Larger LVLMs. To evaluate larger LVLMs on MVL-SIB, we construct language-tier subsets that reliably estimate performance while mitigating excessive computational overhead (cf. §5). For both I2S and S2I, we identify the three languages in each tier that best replicate the average performance of

[6] We omit Centurio-Qwen, since it degrades on the S2I task. We further provide details on all the multilingual vision-language benchmarks we correlate MVL-SIB with in Appendix A.4

[7] Note that we include a constant in our regression model to bridge task-specific scales of results.

[8] In S2I, candidates could comprise more than a single image.

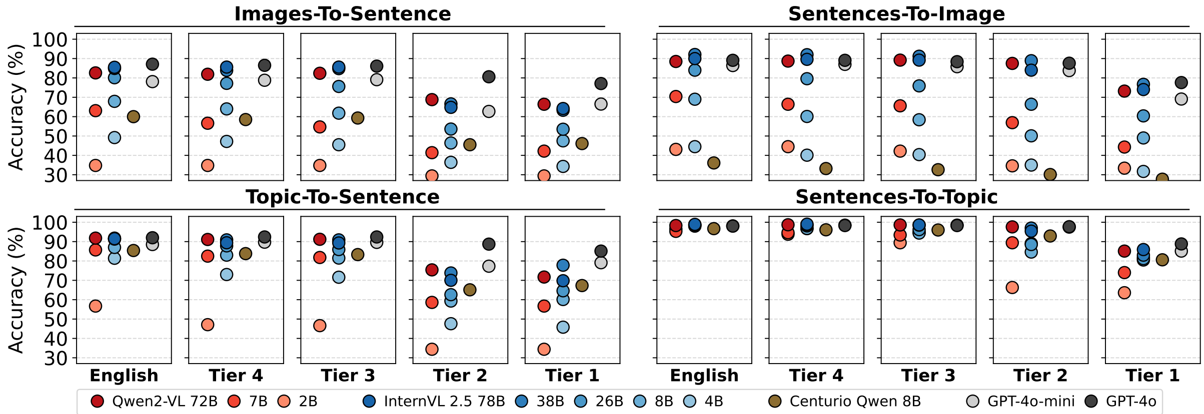


Figure 5: **Larger LVLMs on Subsampled Tiers.** We extract 3 languages per tier that mimic avg. performance full language groups (cf. §5.3) and evaluate LVLMs across all model sizes on {I2S,S2I,T2S,S2T } @ k=3 (cf. §3.2).

the tier. First, we compute the average performance per language tier for both I2S and S2I with k=3, pooling results across models (cf. §5). Then, for each tier, we select the three languages whose performance deviates least from the tier mean. The languages chosen for each tier by task are detailed in Appendix A.6. Finally, we test GPT-4o, InternVL 2.5 {26,32,72}B, and Qwen 2.5 VL {32,70}B on these subsets to assess how well larger models perform across language tiers.

Figure 5 displays the results for both I2S and S2I with k=3.[9] We observe that larger models catch up to GPT-4o on I2S and even outperform it on S2I for higher-resource languages. Since open-weight LVLMs have more limited language support for low-resource languages, GPT-4 and GPT-4o-mini outperform all other models on I2S for languages in tier 1 and 2. Increasing model capacity yields the largest gains on S2I, for which the models exceed GPT-4o up to the lowest-resource Tier 1. Moreover, unlike smaller models, all larger LVLMs (+26B) more effectively leverage multiple image references to improve performance (cf. Appendix A.6). They reap the largest benefit when the number of references increases from 1 to 3 (ca. +3%), with further improvements at $k = 5$ (ca. +1.5%). These results suggest that larger LVLMs have fundamentally better VL support irrespective of the evaluated language (cf. §5.2). The results on text-only topic matching further underscore this notion (cf. lower segment of Figure 5). Larger LVLMs near perfectly match both the reference topic to the correct sentence (ca. 90%) and references sentences to the right topic (ca. 98%). Notably, as model size increases, the performance gap between cross-modal and text-only matching narrows. Collectively, these results further indicate that VL support relative to text-only support improves with increasing model capacity in LVLMs.

6 Conclusion

We present the massively multilingual vision-language benchmark MVL-SIB for cross-modal (and text-only topic matching) in 205 languages that offers key advantages over prior multilingual VL benchmarks. Notably, it covers over 100 additional languages without relying on machine translation. MVL-SIB allows for a clear separation between textual language support and vision-language support in LVLMs by comparing performance on mirrored cross-modal and text-only tasks. Moreover, it allows us to study how LVLMs handle single-image versus multi-image formulations of cross-modal topic matching by varying the number of images provided. In our comparative evaluation of state-of-the-art LVLMs on MVL-SIB, we find that model performance is strongly correlated with both model size and the volume of available pre-training data for each language. However, all LVLMs experience a dramatic performance drop on the lowest-resource languages. Our analysis further reveals that vision-language support deteriorates disproportionately relative to language support, highlighting the need to incorporate low-resource languages into VL training. Moreover, providing multiple images does not benefit open-weight LVLMs in cross-modal topic matching, suggesting that LVLMs are not yet fully effective in multi-image tasks. Lastly, we validate that MVL-SIB correlates well with existing multilingual VL benchmarks, underscoring its reliability as a source of evaluation data for 205 languages.

[9] Qwen2-VL 72B frequently responded with the first letter of the correct topic. If that letter uniquely identifies the correct choice, the answer is considered correct.

7 Limitations

Our work faces three primary limitations. First, although a vast number of LVLMs exist, we selected a representative subset based on key criteria. Specifically, the LVLMs in our study (Qwen2-VL and InternVL, with the exception of Centurio) span a range of parameter counts typical of LLMs. Additionally, we include GPT-4o-mini in the full evaluation and GPT-4o on the subsampled language tiers. Evaluating MVL-SIB across all four tasks I2S, S2I, T2S, and S2T (cf. §3.2) at various $k \in \{1, 3, 5\}$ over 205 languages (i.e., evaluations per model and task, or 2050, in sum per model) becomes computationally intractable. This accumulates to $3 \times 205 = 615$ evaluations per model (205 for T2S as only k=1 reference topic exists) or $3 \times 205 + 205 = 2050$ evaluations in total. We therefore both provide subsets of the language tiers to evaluate on and demonstrate that evaluation only requires 1K instances to reliably estimate task performance. Second, while we strove to choose a diverse set of images to capture the full semantic range of each topic, further diversification is possible by sourcing additional images. However, due to the limited availability of openly licensed images, some topics (e.g., `politics` and `entertainment`) are represented predominantly by images that embody the topic in a more Western-centric cultural context. Hand-selecting images by topic for each language or, more broadly, cultural groups would not scale to 205 languages and would hinder the comparability of results. Our results furthermore confirm that models just as well perform on a broad range of languages spanning diverse cultural backgrounds as on English (cf. Figure 2). At the same time, LVLMs perform best on Western-centric images, mitigating any variation that would originate from using more culture-specific images. Finally, for the topic `geography`, we manually selected images that are representative within the context of SIB, as the broader definition of geography is too diffuse to capture visually.

Acknowledgements

We used AI assistance (chatGPT o3-mini) to polish the writing and the tables of the manuscript as well as to refine the code for our visualizations.

References

David Adelani, Hannah Liu, Xiaoyu Shen, Nikita Vassilyev, Jesujoba Alabi, Yanke Mao, Haonan Gao, and En-Shiun Lee. 2024. SIB-200: A simple, inclusive, and big evaluation dataset for topic classification in 200+ languages and dialects. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 226–245, St. Julian's, Malta. Association for Computational Linguistics.

Željko Agić and Natalie Schluter. 2018. Baselines and Test Data for Cross-Lingual Inference. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan.

Samuel R. Bowman, Gabor Angeli, Christopher Potts, and Christopher D. Manning. 2015. A large annotated corpus for learning natural language inference. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*, pages 632–642. The Association for Computational Linguistics.

Emanuele Bugliarello, Fangyu Liu, Jonas Pfeiffer, Siva Reddy, Desmond Elliott, Edoardo Maria Ponti, and Ivan Vulić. 2022. IGLUE: A benchmark for transfer learning across modalities, tasks, and languages. *ArXiv*, abs/2201.11732.

Soravit Changpinyo, Linting Xue, Michal Yarom, Ashish Thapliyal, Idan Szpektor, Julien Amelot, Xi Chen, and Radu Soricut. 2023. MaXM: Towards multilingual visual question answering. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 2667–2682, Singapore. Association for Computational Linguistics.

Zhe Chen, Jiannan Wu, Wenhai Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong Zhang, Xizhou Zhu, Lewei Lu, et al. 2024. Internvl: Scaling up vision foundation models and aligning for generic visual-linguistic tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24185–24198.

J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.

William Gaviria Rojas, Sudnya Diamos, Keertan Kini, David Kanter, Vijay Janapa Reddi, and Cody Coleman. 2022. The Dollar Street Dataset: Images Representing the Geographic and Socioeconomic Diversity of the World. *Advances in Neural Information Processing Systems*, 35:12979–12990.

Gregor Geigle, Abhay Jain, Radu Timofte, and Goran Glavaš. 2024a. mBLIP: Efficient bootstrapping of multilingual vision-LLMs. In *Proceedings of the 3rd Workshop on Advances in Language and Vision Research (ALVR)*, pages 7–25, Bangkok, Thailand. Association for Computational Linguistics.

Gregor Geigle, Florian Schneider, Carolin Holtermann, Chris Biemann, Radu Timofte, Anne Lauscher, and Goran Glavaš. 2025. Centurio: On drivers of multilingual ability of large vision-language model. *Preprint*, arXiv:2501.05122.

Gregor Geigle, Radu Timofte, and Goran Glavaš. 2024b. African or European swallow? benchmarking large vision-language models for fine-grained object classification. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 2653–2669, Miami, Florida, USA. Association for Computational Linguistics.

Gregor Geigle, Radu Timofte, and Goran Glavaš. 2024c. Babel-ImageNet: Massively multilingual evaluation of vision-and-language representations. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5064–5084, Bangkok, Thailand. Association for Computational Linguistics.

Drew A. Hudson and Christopher D. Manning. 2019. GQA: A New Dataset for Real-World Visual Reasoning and Compositional Question Answering. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 6700–6709. Computer Vision Foundation / IEEE.

Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online. Association for Computational Linguistics.

Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A. Shamma, Michael S. Bernstein, and Li Fei-Fei. 2017. Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. *Int. J. Comput. Vision*, 123(1):32–73. Place: USA Publisher: Kluwer Academic Publishers.

Sneha Kudugunta, Isaac Caswell, Biao Zhang, Xavier Garcia, Christopher A. Choquette-Choo, Katherine Lee, Derrick Xin, Aditya Kusupati, Romi Stella, Ankur Bapna, and Orhan Firat. 2023. Madlad-400: A multilingual and document-level large audited dataset. *Preprint*, arXiv:2309.04662.

Fangyu Liu, Emanuele Bugliarello, Edoardo Maria Ponti, Siva Reddy, Nigel Collier, and Desmond Elliott. 2021. Visually grounded reasoning across languages and cultures. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10467–10485, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Niklas Muennighoff, Thomas Wang, Lintang Sutawika, Adam Roberts, Stella Biderman, Teven Le Scao, M Saiful Bari, Sheng Shen, Zheng-Xin Yong, Hailey Schoelkopf, Xiangru Tang, Dragomir Radev, Alham Fikri Aji, Khalid Almubarak, Samuel Albanie, Zaid Alyafeai, Albert Webson, Edward Raff, and Colin Raffel. 2022. Crosslingual generalization through multitask finetuning. *Preprint*, arXiv:2211.01786.

Jonas Pfeiffer, Gregor Geigle, Aishwarya Kamath, Jan-Martin O. Steitz, Stefan Roth, Ivan Vulić, and Iryna Gurevych. 2022. xGQA: Cross-lingual visual question answering. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 2497–2511, Dublin, Ireland. Association for Computational Linguistics.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning transferable visual models from natural language supervision. *Preprint*, arXiv:2103.00020.

Angelika Romanou, Negar Foroutan, Anna Sotnikova, Zeming Chen, Sree Harsha Nelaturu, Shivalika Singh, Rishabh Maheshwary, Micol Altomare, Mohamed A. Haggag, Snegha A, Alfonso Amayuelas, Azril Hafizi Amirudin, Viraat Aryabumi, Danylo Boiko, Michael Chang, Jenny Chim, Gal Cohen, Aditya Kumar Dalmia, Abraham Diress, Sharad Duwal, Daniil Dzenhaliou, Daniel Fernando Erazo Florez, Fabian Farestam, Joseph Marvin Imperial, Shayekh Bin Islam, Perttu Isotalo, Maral Jabbarishiviari, Börje F. Karlsson, Eldar Khalilov, Christopher Klamm, Fajri Koto, Dominik Krzemiński, Gabriel Adriano de Melo, Syrielle Montariol, Yiyang Nan, Joel Niklaus, Jekaterina Novikova, Johan Samir Obando Ceron, Debjit Paul, Esther Ploeger, Jebish Purbey, Swati Rajwal, Selvan Sunitha Ravi, Sara Rydell, Roshan Santhosh, Drishti Sharma, Marjana Prifti Skenduli, Arshia Soltani Moakhar, Bardia Soltani Moakhar, Ran Tamir, Ayush Kumar Tarun, Azmine Toushik Wasi, Thenuka Ovin Weerasinghe, Serhan Yilmaz, Mike Zhang, Imanol Schlag, Marzieh Fadaee, Sara Hooker, and Antoine Bosselut. 2024. Include: Evaluating multilingual language understanding with regional knowledge. *Preprint*, arXiv:2411.19799.

David Romero, Chenyang Lyu, Haryo Akbarianto Wibowo, Teresa Lynn, Injy Hamed, Aditya Nanda Kishore, Aishik Mandal, Alina Dragonetti, Artem Abzaliev, Atnafu Lambebo Tonja, Bontu Fufa Balcha, Chenxi Whitehouse, Christian Salamea, Dan John Velasco, David Ifeoluwa Adelani, David Le Meur, Emilio Villa-Cueva, Fajri Koto, Fauzan Farooqui, Frederico Belcavello, Ganzorig Batnasan, Gisela Vallejo, Grainne Caulfield, Guido Ivetta, Haiyue Song, Henok Biadglign Ademtew, Hernán Maina, Holy Lovenia, Israel Abebe Azime, Jan Christian Blaise Cruz, Jay Gala, Jiahui Geng, Jesus-German Ortiz-Barajas, Jinheon Baek, Jocelyn Dunstan, Laura Alonso Alemany, Kumaranage Ravindu Yasas Nagasinghe, Luciana Benotti, Luis Fernando D'Haro, Marcelo Viridiano, Marcos Estecha-Garitagoitia, Maria Camila Buitrago Cabrera, Mario Rodríguez-Cantelar, Mélanie Jouitteau, Mihail Mihaylov, Mohamed Fazli Mohamed Imam, Muhammad Farid Adilazuarda, Munkhjargal Gochoo, Munkh-Erdene Otgonbold, Naome Etori, Olivier Niyomugisha, Paula Mónica Silva, Pranjal Chitale, Raj Dabre, Rendi Chevi, Ruochen Zhang, Ryandito Diandaru, Samuel Cahyawijaya, Santiago Góngora, Soyeong Jeong, Sukannya Purkayastha, Tatsuki Kuribayashi, Teresa Clifford, Thanmay Jayakumar, Tiago Timponi Torrent, Toqeer Ehsan, Vladimir Araujo, Yova Kementchedjhieva, Zara Burzo, Zheng Wei Lim, Zheng Xin Yong, Oana Ignat, Joan Nwatu, Rada Mihalcea, Thamar Solorio, and Alham Fikri Aji. 2024. Cvqa: Culturally-diverse multilingual visual question answering benchmark. *Preprint*, arXiv:2406.05967.

Florian Schneider and Sunayana Sitaram. 2024. M5 – a diverse benchmark to assess the performance of large multimodal models across multilingual and multicultural vision-language tasks. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 4309–4345, Miami, Florida, USA. Association for Computational Linguistics.

Hai-Long Sun, Da-Wei Zhou, Yang Li, Shiyin Lu, Chao Yi, Qing-Guo Chen, Zhao Xu, Weihua Luo, Kaifu Zhang, De-Chuan Zhan, and Han-Jia Ye. 2024. Parrot: Multilingual visual instruction tuning. *Preprint*, arXiv:2406.02539.

Jingqun Tang, Qi Liu, Yongjie Ye, Jinghui Lu, Shu Wei, Chunhui Lin, Wanqing Li, Mohamad Fitri Faiz Bin Mahmood, Hao Feng, Zhen Zhao, Yanjie Wang, Yuliang Liu, Hao Liu, Xiang Bai, and Can Huang. 2024. Mtvqa: Benchmarking multilingual text-centric visual question answering. *Preprint*, arXiv:2405.11985.

NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, Anna Sun, Skyler Wang, Guillaume Wenzek, Al Youngblood, Bapi Akula, Loic Barrault, Gabriel Mejia Gonzalez, Prangthip Hansanti, John Hoffman, Semarley Jarrett, Kaushik Ram Sadagopan, Dirk Rowe, Shannon Spruit, Chau Tran, Pierre Andrews, Necip Fazil Ayan, Shruti Bhosale, Sergey Edunov, Angela Fan, Cynthia Gao, Vedanuj Goswami, Francisco Guzmán, Philipp Koehn, Alexandre Mourachko, Christophe Ropers, Safiyyah Saleem, Holger Schwenk, and Jeff Wang. 2022. No language left behind: Scaling human-centered machine translation. *Preprint*, arXiv:2207.04672.

Ashish V. Thapliyal, Jordi Pont Tuset, Xi Chen, and Radu Soricut. 2022. Crossmodal-3600: A Massively Multilingual Multimodal Evaluation Dataset. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 715–729, Abu Dhabi, United Arab Emirates.

Ashmal Vayani, Dinura Dissanayake, Hasindri Watawana, Noor Ahsan, Nevasini Sasikumar, Omkar Thawakar, Henok Biadglign Ademtew, Yahya Hmaiti, Amandeep Kumar, Kartik Kuckreja, Mykola Maslych, Wafa Al Ghallabi, Mihail Mihaylov, Chao Qin, Abdelrahman M Shaker, Mike Zhang, Mahardika Krisna Ihsani, Amiel Esplana, Monil Gokani, Shachar Mirkin, Harsh Singh, Ashay Srivastava, Endre Hamerlik, Fathinah Asma Izzati, Fadillah Adamsyah Maani, Sebastian Cavada, Jenny Chim, Rohit Gupta, Sanjay Manjunath, Kamila Zhumakhanova, Feno Heriniaina Rabevohitra, Azril Amirudin, Muhammad Ridzuan, Daniya Kareem, Ketan More, Kunyang Li, Pramesh Shakya, Muhammad Saad, Amirpouya Ghasemaghaei, Amirbek Djanibekov, Dilshod Azizov, Branislava Jankovic, Naman Bhatia, Alvaro Cabrera, Johan Obando-Ceron, Olympiah Otieno, Fabian Farestam, Muztoba Rabbani, Sanoojan Baliah, Santosh Sanjeev, Abduragim Shtanchaev, Maheen Fatima, Thao Nguyen, Amrin Kareem, Toluwani Aremu, Nathan Xavier, Amit Bhatkal, Hawau Toyin, Aman Chadha, Hisham Cholakkal, Rao Muhammad Anwer, Michael Felsberg, Jorma Laaksonen, Thamar Solorio, Monojit Choudhury, Ivan Laptev, Mubarak Shah, Salman Khan, and Fahad Khan. 2024. All languages matter: Evaluating lmms on culturally diverse 100 languages. *Preprint*, arXiv:2411.16508.

Peng Wang, Shuai Bai, Sinan Tan, Shijie Wang, Zhihao Fan, Jinze Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Yang Fan, Kai Dang, Mengfei Du, Xuancheng Ren, Rui Men, Dayiheng Liu, Chang Zhou, Jingren Zhou, and Junyang Lin. 2024. Qwen2-vl: Enhancing vision-language model's perception of the world at any resolution. *Preprint*, arXiv:2409.12191.

Ning Xie, Farley Lai, Derek Doran, and Asim Kadav. 2019. Visual Entailment: A Novel Task for Fine-grained Image Understanding. *arXiv preprint arXiv:1901.06706*.

Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. 2021. mt5: A massively multilingual pre-trained text-to-text transformer. *Preprint*, arXiv:2010.11934.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, Jianxin Yang, Jin Xu, Jingren Zhou, Jinze Bai, Jinzheng He, Junyang Lin, Kai Dang, Keming Lu, Keqin Chen, Kexin Yang, Mei Li, Mingfeng Xue, Na Ni, Pei Zhang, Peng Wang, Ru Peng, Rui Men, Ruize Gao, Runji Lin, Shijie Wang, Shuai Bai, Sinan Tan, Tianhang Zhu, Tianhao Li, Tianyu Liu, Wenbin Ge, Xiaodong Deng, Xiaohuan Zhou, Xingzhang Ren, Xinyu Zhang, Xipin Wei, Xuancheng Ren, Xuejing Liu, Yang Fan, Yang Yao, Yichang Zhang, Yu Wan, Yunfei Chu, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, Zhifang Guo, and Zhihao Fan. 2024. Qwen2 technical report. *Preprint*, arXiv:2407.10671.

Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruoqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhao Huang, Huan Sun, Yu Su, and Wenhu Chen. 2023. MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI. *CoRR*, abs/2311.16502. ArXiv: 2311.16502.

Xiang Yue, Yueqi Song, Akari Asai, Seungone Kim, Jean de Dieu Nyandwi, Simran Khanuja, Anjali Kantharuban, Lintang Sutawika, Sathyanarayanan Ramamoorthy, and Graham Neubig. 2024. Pangea: A fully open multilingual multimodal llm for 39 languages. *arXiv preprint arXiv:2410.16153*.

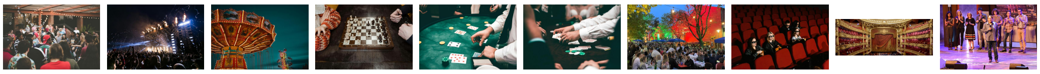
Xiaohua Zhai, Basil Mustafa, Alexander Kolesnikov, and Lucas Beyer. 2023. Sigmoid loss for language image pre-training. *Preprint*, arXiv:2303.15343.

Wenxuan Zhang, Mahani Aljunied, Chang Gao, Yew Ken Chia, and Lidong Bing. 2023. M3Exam: A Multilingual, Multimodal, Multilevel Benchmark for Examining Large Language Models. *Advances in Neural Information Processing Systems*, 36:5484–5505.

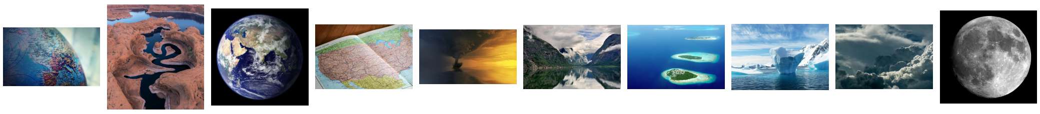
A Appendix

A.1 Images Per Topic

Entertainment



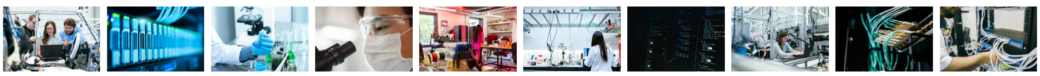
Geography



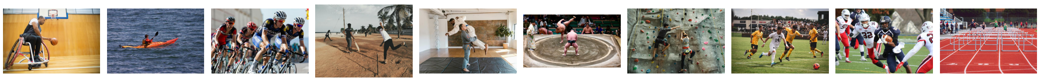
Health

Politics

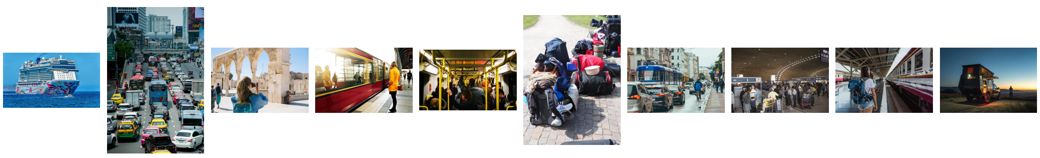
Science & Technology



Sports



Travel



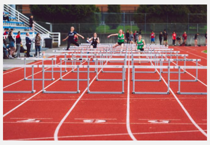
A.2 Prompts

Images-To-Sentences (I2S)

Which sentence best matches the topic of the images? The images and the sentences each belong to one of the following topics: "entertainment", "geography", "health", "politics", "science and technology", "sports", or "travel". Choose one sentence from A, B, C, or D. Output only a single letter!

Images





 # Sentences

A. ```Maroochydore führte am Ende die Rangfolge an, mit sechs Punkten Vorsprung vor Noos a als Zweitem.```

B. ```Es wurden keine schwere Verletzungen gemeldet, jedoch mussten mindestens fünf der zur Zeit der Explosion Anwesenden aufgrund von Schocksymptomen behandelt werden.```

C. ```Finnland ist ein großartiges Reiseziel für Bootstouren. Das „Land der tausend Seen“ hat auch Tausende von Inseln – in den Seen und in den Küstenarchipelen.```

D. ```Es ist auch nicht erforderlich, dass Sie eine lokale Nummer von der Gemeinde erhalten, in der Sie leben. Sie können eine Internetverbindung über Satellit in der Wildnis v on Chicken in Alaska erhalten und eine Nummer auswählen, die vorgibt, dass Sie im sonnigen Arizona sind.```

Sentences-To-Images (S2I)

Which image best matches the topic of the sentences? The sentences and the images each belong to one of the following topics: "entertainment", "geography", "health", "politics", "science and technology", "sports", or "travel". Choose one image from A, B, C, or D. Output only a single letter!

Sentences

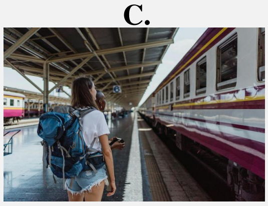
- ```Maroochydore führte am Ende die Rangfolge an, mit sechs Punkten Vorsprung vor Noosa als Zweitem.```
- ```Die Schlagmänner der mittleren Reihe, Sachin Tendulkar und Rahul Dravid, zeigten gute Leistungen und erzielten eine Partnerschaft mit 100 Runs.```
- ```Da pro Tag nur achtzehn Medaillen zur Verfügung stehen, hat es ein Anzahl an Ländern nicht auf das Podium geschafft.```
- ```Wintersportarten sind in den nördlichen Regionen am beliebtesten und Italiener nehmen an internationalen Wettkämpfen und olympischen Spielen teil.```
- ```Nach dem Rennen bleibt Keselowski mit 2.250 Punkten Spitzenreiter in der Fahrerwertung.

Images

A.

B.

C.



D.

Your answer letter:

Topic-To-Sentence (T2S)

Which sentence best matches the topic "sports"? The sentences each belong to one of the following topics: "entertainment", "geography", "health", "politics", "science and technology", "sports", or "travel". Choose one sentence from A, B, C, or D. Output only a single letter!

Sentences

A. ```Maroochydore führte am Ende die Rangfolge an, mit sechs Punkten Vorsprung vor Noos a als Zweitem.```

B. ```Es wurden keine schwere Verletzungen gemeldet, jedoch mussten mindestens fünf der zur Zeit der Explosion Anwesenden aufgrund von Schocksymptomen behandelt werden.```

C. ```Finnland ist ein großartiges Reiseziel für Bootstouren. Das „Land der tausend Seen“ hat auch Tausende von Inseln – in den Seen und in den Küstenarchipelen.```

D. ```Es ist auch nicht erforderlich, dass Sie eine lokale Nummer von der Gemeinde erhalten, in der Sie leben. Sie können eine Internetverbindung über Satellit in der Wildnis v on Chicken in Alaska erhalten und eine Nummer auswählen, die vorgibt, dass Sie im sonnigen Arizona sind.```

Your answer letter:

Sentences-To-Topics (S2T)

Which topic best matches the sentences? The sentences belong to one of the following topics: "entertainment", "geography", "health", "politics", "science and technology", "sports", or "travel". Choose one topic from A, B, C, or D. Output only a single letter!

Sentences

- ```Maroochydore führte am Ende die Rangfolge an, mit sechs Punkten Vorsprung vor Noosa als Zweitem.```
- ```Die Schlagmänner der mittleren Reihe, Sachin Tendulkar und Rahul Dravid, zeigten gute Leistungen und erzielten eine Partnerschaft mit 100 Runs.```
- ```Da pro Tag nur achtzehn Medaillen zur Verfügung stehen, hat es ein Anzahl an Ländern nicht auf das Podium geschafft.```
- ```Wintersportarten sind in den nördlichen Regionen am beliebtesten und Italiener nehmen an internationalen Wettkämpfen und olympischen Spielen teil.```
- ```Nach dem Rennen bleibt Keselowski mit 2.250 Punkten Spitzenreiter in der Fahrerwertung.```

Topics

A. sports

B. health

C. travel

D. science and technology

Your answer letter:

A.3 Further Details

Models. We test state-of-the-art LVLMs across various sizes. Smaller models are evaluated on all languages, while larger LVLMs (26B+) are tested on subsets of the MVL-SIB languages (cf. §5.3).

GPT-4o. We evaluate MVL-SIB on GPT-4o-mini-*2024-07-18* and GPT-4o-*2024-08-06*. We set the image detail in the API to 'low' because our tasks only require high-level reasoning that does not depend on finer image details.[10] Prior works show that GPT-4o is the best-performing multilingual LVLM (Schneider and Sitaram, 2024; Vayani et al., 2024).

Qwen2-VL. Qwen2-VL ties a 675M parameter vision-transformer (ViT) into Qwen2 LLMs (Wang et al., 2024). An MLP compresses adjacent 2×2 visual tokens embedded by the ViT into one token representation, which is then input to the LLM.

InternVL 2.5. Depending on the model size, InternVL uses Qwen2.5 or InternLM as its LLM backbone (Chen et al., 2024). The model embeds images either by a 6B or by a distilled 300M ViT pretrained with CLIP (Radford et al., 2021). The resulting image patch encodings are downsampled by factor 4 and fed through an MLP to the LLM.

Centurio. Centurio is the latest massively multilingual LVLM explicitly trained on 100 languages (Geigle et al., 2025), outperforming alternatives like Parrot (Sun et al., 2024) or Pangea (Yue et al., 2024). It employs Qwen2.5 as its LLM (Yang et al., 2024) and `SigLIP SO400/384` as its ViT (Zhai et al., 2023). The model mixes resolutions by stacking the encodings of the full image and those of 2×2 tiles along the features. The combined embedding is then projected via an MLP to the LLM's input space.

Besides architectures, the LVLMs most crucially differ in dataset mixtures on which they were trained. Centurio translates image-caption, VQA, OCR, and a few multi-image datasets to 100 languages with NLLB (Team et al., 2022) to mix 50:50 with the original English data. Qwen2-VL and InternVL, however, were trained on much larger, more diverse datasets that comprise sizable multi-image comparison and video understanding datasets. Moreover, assuming that the LLMs of Qwen2-VL, InternVL, and GPT-4o were pretrained on Flores, their performance would be overly optimistic.

[10] We use GPT-4o-mini because evaluating GPT-4o would be too expensive.

A.3.1 Performance by Model over Languages grouped by Language Tier

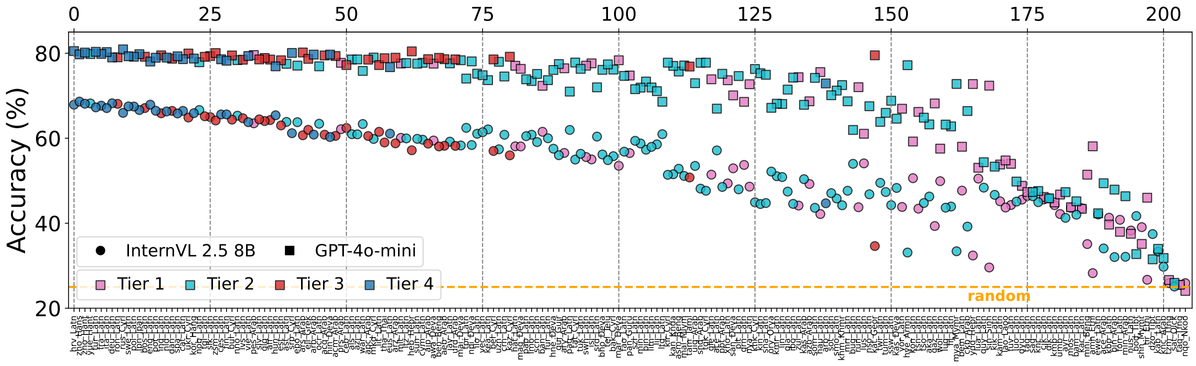


Figure 6: **Images-To-Sentences @ k=3.** The English prompt describes the cross-modal topic matching task, lists all topics, and provides both k=3 reference images and 4 sentences in the corresponding language {`eng_Latn`, . . . , `nqo_Nkoo`}. LVLMs must select the sentence of 4 options that topically fits k=3 reference images. The sentences spanning 205 languages and 7 topics are drawn from SIB-200 (Adelani et al., 2024), while images for the topics were hand-selected (cf. Appendix A.1). An example prompt is shown in Appendix A.7.3; further details are in §4. **Plot.** The x-axis orders the languages of the candidate sentences {`eng_Latn`, . . . , `nqo_Nkoo`}, respectively, by descending performance (y-axis). The top x-axis indicates the running index of each language L_i ($i \in \{1, \ldots, 205\}$). **Tiers.** The languages are grouped by tiers derived from Joshi et al. (2020) (cf. §5).

A.3.2 Calibration Analysis of Cross-Modal Topic Matching

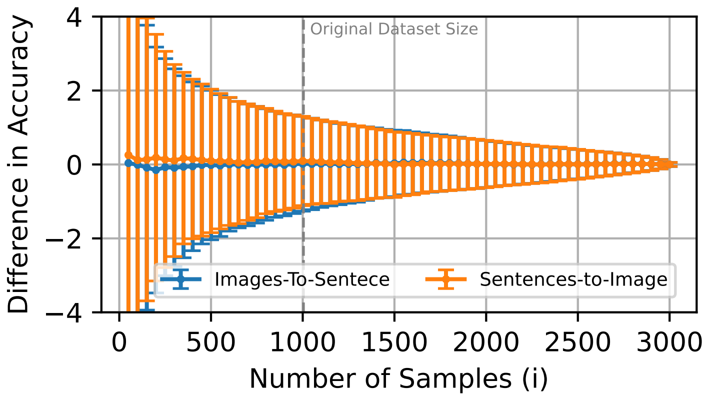


Figure 7: **Calibration Analysis of Cross-Modal Topic Matching for InternVL 2.5 8B. Analysis:** To assess the reliability of cross-modal topic matching with fewer samples than our full dataset (1004 samples per language), we randomly select 500 trajectories. We then compute performance metrics on cumulative subsets, incrementing by 50 examples at each step. The difference in performance between the full dataset and each subset is calculated to quantify the deviation at each sample size. **Plot:** The plot displays the average absolute spread in performance (averaged over all languages) along with the standard deviation for InternVL 2.5 8B. We restrict ourselves to a single open-weight LVLM, since the analysis yields identical results across all combinations of LVLMs and language tiers. **Insights:** Our analysis shows that performance stabilizes rapidly, with deviations of only about 1% observed at 1,004 instances – same size as the subset from which the dataset was created. This indicates that reliable evaluation of cross-modal topic matching can be achieved with far fewer than 3,012 samples.

A.4 Overview of Multilingual Vision-Language Benchmarks

xGQA. The xGQA dataset (Pfeiffer et al., 2022) is a cross-lingual visual question-answering resource. It extends the well-known English-only GQA dataset (Hudson and Manning, 2019) by providing manual translations of the questions in the balanced *test-dev* set. The dataset contains 9666 questions available in eight languages across five scripts, while the answers remain in English. In addition, it features 300 unique images from Visual Genome (Krishna et al., 2017).

MaXM. MaXM, introduced by Changpinyo et al. (2023), is a VQA dataset covering seven languages written in five scripts. In this dataset, both the questions and their corresponding answers are presented in the same language. The images are drawn from a subset of the XM3600 (Thapliyal et al., 2022) dataset and are selected to correspond to regions where the question-answer pair's language is spoken, ensuring both linguistic and cultural diversity.

XVNLI. The XVNLI dataset (Bugliarello et al., 2022) introduces the task of Cross-lingual Visual Natural Language Inference, where a model must determine if a textual hypothesis *entails*, *contradicts*, or is *neutral* with respect to a visual premise. This dataset spans five languages across three scripts and includes 357 unique images from Visual Genome. It is built upon a combination of the text-only SNLI (Bowman et al., 2015) dataset and its cross-lingual (Agić and Schluter, 2018) and cross-modal (Xie et al., 2019) counterparts.

MaRVL. The MaRVL dataset (Liu et al., 2021) is designed to benchmark models on Multicultural Reasoning over Vision and Language. Each sample consists of two images, a textual statement, and a binary (true/false) answer grounded in the images. Covering five languages across three scripts, MaRVL includes 4914 culturally diverse images that align with the respective languages. The images in each sample are selected to reflect the culture of the annotator who composed the textual statement in their native language.

XM3600. The XM3600 dataset (Thapliyal et al., 2022) is an extensive multilingual image captioning resource encompassing 36 languages. It contains 261375 captions across 13 scripts, with 100 unique images per language. The images are chosen to reflect the cultural background of the language, ensuring both cultural and linguistic diversity. All captions were manually produced by professional native speakers rather than being automatically generated. Due to the dataset's large size, we evaluate XM3600 using a randomly selected subset of 500 images per language.

Babel-ImageNet (multiple-choice) (BIN-MC). Babel-ImageNet (Geigle et al., 2024c) translates ImageNet's (Deng et al., 2009) labels into nearly 300 languages, allowing us to assess whether models can recognize and correctly link diverse ImageNet objects to their labels in the target language. Given the computational cost, we focus on languages that appear in only one or two other datasets, in addition to English and a select few high-resource languages, and we use 10 images per class instead of 50. We follow Geigle et al. (2024b) and frame the task as a multiple-choice problem by mining hard negative options from the complete label pool. This approach avoids the ambiguity inherent in traditional open-ended VQA formats. Negatives are selected based on the English labels, filtering out candidates not translated by Babel-ImageNet into the target language, and ultimately choosing the three most similar negative labels available.

SMPQA. Geigle et al. (2025) introduce SMPQA (Synthetic Multilingual Plot QA) as a test dataset for evaluating multilingual OCR capabilities for bar plots and pie charts, covering 11 languages and various scripts and resource levels.

M5B-VGR. The M5B-VGR dataset, presented

Task	Dataset	Visual Input	Textual Input	Target Output	Metric	#Lang.
Captioning	XM3600	Single-Image	Prompt (English)	Caption (Target Language)	CIDEr	36
Multiple-Choice Visual Question Answering	BabelImageNet-MC	Single-Image	Question (Target Language)	Letter of the correct Choice	Relaxed Accuracy	20
Text-Heavy Multiple-Choice Visual Question Answering	M3Exam MMMU xMMMU	Single or Multi-Image	Question (Target Language) Question (English) Question (Target Language)	Letter of the correct Choice	Relaxed Accuracy	7 1 7
Text-Heavy Visual Question Answering	MTVQA SMPQA - Name	Single-Image	Question (Target Language)	Word or Phrase (Target Language)	Exact Accuracy	9 11
Text-Heavy Visually Grounded Reasoning	SMPQA - Ground	Single-Image	Question (Target Language)	'yes' / 'no'	Exact Accuracy	11
Visio-Linguistic Outlier Detection	M5B-VLOD	Multi-Image	Hypothesis (Target Language)	Letter of the correct Choice	Relaxed Accuracy	12
Visual Natural Language Inference	XVNLI	Single-Image	Hypothesis (Target Language)	'yes' / 'no' / 'maybe'	Exact Accuracy	5
Visual Question Answering	MaXM xGQA	Single-Image	Question (Target Language)	Word or Phrase (Target Language) Word or Phrase (English)	Exact Accuracy	6 8
Visually Grounded Reasoning	M5B-VGR MaRVL	Multi-Image	Hypothesis (Target Language)	'yes' / 'no'	Exact Accuracy	12 6

Table 3: Summary of multilingual vision-language benchmarks we correlate MVL-SIB against. Relaxed denotes responses that start with the correct option letter (cf. 4).

by (Schneider and Sitaram, 2024), is a visually grounded reasoning benchmark akin to MaRVL. Each sample comprises two images, a textual statement, and a binary (true/false) answer based on the images. It spans 12 languages across 7 scripts and features culturally diverse photos from regions where the respective languages are spoken. The images are sampled from the Dollar Street (Gaviria Rojas et al., 2022) dataset, with 120 samples provided per language.

M5B-VLOD. The M5B-VLOD (Visio-Linguistic Outlier Detection) dataset, also introduced by (Schneider and Sitaram, 2024), consists of samples containing five images paired with a textual statement that is true for all but one image. The task is to identify the outlier image that does not match the statement. This dataset covers the same 12 languages as M5B-VGR, with images sampled in a similar manner from the same source, and provides 120 samples per language.

MTVQA. The MTVQA dataset, introduced by (Tang et al., 2024), features text-heavy visual question answering tasks. It includes expert human annotations in 9 diverse languages, comprising 6778 question-answer pairs across 2116 images. The images predominantly contain text in the corresponding language, with questions and answers closely tied to that text. These images are sourced from various publicly available datasets.

CVQA. The CVQA dataset, introduced by (Romero et al., 2024), is a multilingual and culturally nuanced VQA benchmark that includes a broad array of languages, many of which are underrepresented in NLP. It consists of 10000 questions spanning 30 countries and 31 languages, forming 39 distinct country-language pairs (for instance, Spanish appears in 7 different splits corresponding to 7 Spanish-speaking countries). The images were manually collected by human annotators to accurately depict the culture associated with each country-language pair. Each sample includes one image and a question in the respective language. Although the test set is not publicly available, the authors permit up to 5 daily leaderboard submissions for evaluation.

M3Exam. The M3Exam dataset, presented by (Zhang et al., 2023), contains real-world exam questions in 9 languages, available as either text-only or multimodal samples. For our evaluation, we only include samples that require at least one image. Due to the limited number of samples for Swhalili and Javanese, we focus on the remaining 7 languages. The selected samples consist of multiple-choice questions in the target language, accompanied by up to 8 images that may appear in both the question and the answer options, with the number of choices ranging from 4 to 8 per sample.

xMMMU. xMMMU, introduced by (Yue et al., 2024), comprises college-level multiple-choice VQA samples in seven languages. It was automatically translated using GPT4o from a randomly selected subset of 300 questions from the MMMU (Yue et al., 2023) validation split.

A.5 Prompts

We list the prompts for each dataset in our test suite used for all models in Figure 8.

SMPQA

<IMG>{QUESTION}\nAnswer the question using a single word or phrase.

CVQA

<IMG>{QUESTION}\nThere are several options:\nA. {OPTION A}\nB. {OPTION B}\nC. {OPTION C}\nD. {OPTION D}\nAnswer with the option's letter from the given choices directly.

xMMMU

{QUESTION}\nThere are several options:\nA. {OPTION A}\nB. {OPTION B}\nC. {OPTION C}\nD. {OPTION D}\nAnswer with the option's letter from the given choices directly.

MTVQA

<IMG>{QUESTION}\nAnswer the question using a single word or phrase.\nAnswer in {LANGUAGE}.

M3Exam

{QUESTION}\nOptions:\nA. {OPTION A}\nB. {OPTION B}\nC. {OPTION C}\nD. {OPTION D}\n Answer with the option's letter from the given choices directly.

BIN-MC

<IMG>Which of these choices (in English) is shown in the image?\n Choices:\nA. {CHOICE A}\nB. {CHOICE B}\nC. {CHOICE C}\nD. {CHOICE D}\n Answer with the letter from the given choices directly.

xGQA

<IMG>{QUESTION}?\nAnswer the question using a single word or phrase.\nAnswer in English.

MaXM

<IMG>{QUESTION}?\nAnswer the question using a single word or phrase.\nAnswer in {LANGUAGE}.

MaRVL

<IMG>Given the two images <IMG><IMG>, is it correct to say "{HYPOTHESIS}"? Answer yes or no.'

XVNLI

<IMG>Is it guaranteed true that "{HYPOTHESIS}"? Yes, no, or maybe? Answer in English.

M5-VGR

Given the two images <IMG><IMG>, is it correct to say "{HYPOTHESIS}"? Answer yes or no.'

M5-VLOD

Based on the 5 images <IMG><IMG><IMG><IMG><IMG> ordered from top-left to bottom-right, which image does not match the hypothesis "{HYPOTHESIS}"? Choose one from [A, B, C, D, E] and only output a single letter:

XM3600

Briefly describe the image in {LANGUAGE} in one sentence.

Figure 8: Prompts used for the different datasets of our test suite. For M3Exam and xMMMU, the questions contain images at individual positions, and also the options can consist of images. In total, a sample of M3Exam can contain up to 8 images and 8 options, and a sample of xMMMU can contain up to 4 images and 4 options.

A.6 Full Results For Subsets by Task, Model, and Language

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	**Tier**	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
acm_Arab	3	34.9	55.4	80.8	50.0	61.0	74.1	80.2	82.2	54.4	72.0	N/A
aka_Latn	2	28.7	38.0	57.2	34.6	46.3	51.3	58.3	57.0	42.4	61.4	N/A
apc_Arab	3	34.9	54.8	80.0	49.1	61.3	74.2	78.8	82.1	54.4	71.5	N/A
arb_Arab	4	35.2	57.1	81.4	51.7	62.2	75.8	80.8	82.8	55.2	73.2	N/A
azj_Latn	2	34.6	53.8	80.5	49.5	58.6	70.9	78.3	80.0	51.8	72.8	N/A
bul_Cyrl	3	35.0	56.6	79.3	50.3	65.3	77.1	79.8	81.6	53.9	71.5	N/A
ces_Latn	4	36.2	58.2	79.5	52.0	65.6	78.5	79.5	81.5	53.2	72.7	N/A
eng_Latn	4	36.3	65.8	79.8	52.5	67.7	N/A	79.7	81.7	54.8	68.3	N/A
fin_Latn	4	33.0	57.1	79.1	49.8	64.1	76.5	79.2	81.1	54.1	70.2	N/A
hat_Latn	1	32.1	50.4	77.8	43.0	54.7	66.1	74.1	74.2	54.9	69.3	N/A
hau_Latn	1	29.1	39.2	56.3	32.4	41.8	45.2	58.0	57.5	43.9	70.2	N/A
min_Arab	2	27.0	30.4	58.4	27.8	33.6	36.3	53.4	47.8	37.1	44.8	N/A
umb_Latn	1	29.1	34.8	52.5	31.1	42.5	45.2	50.8	52.0	35.4	47.4	N/A

Table 4: **Subsets of language tiers for I2S @ k=1.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	**Tier**	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
acm_Arab	3	35.2	55.1	82.4	45.6	60.6	74.0	85.5	86.7	59.7	79.3	86.0
aka_Latn	2	27.9	38.3	63.2	34.2	46.3	52.9	61.1	60.6	42.6	63.3	79.3
apc_Arab	3	34.7	54.3	82.2	45.1	60.5	73.9	84.1	85.7	59.4	78.5	85.8
arb_Arab	4	35.2	56.4	81.8	46.9	60.9	76.0	85.0	87.1	58.7	79.7	86.7
azj_Latn	2	33.1	53.4	82.3	46.2	60.9	71.4	82.7	83.5	57.9	78.6	86.5
bul_Cyrl	3	34.7	54.6	82.7	45.9	64.4	78.8	84.8	84.5	58.8	79.5	86.4
ces_Latn	4	35.6	56.8	82.2	48.9	65.6	78.4	83.6	85.3	57.3	78.6	86.3
eng_Latn	4	34.8	63.1	82.6	49.2	67.9	80.1	84.9	85.5	60.0	78.1	87.1
fin_Latn	4	33.5	56.4	81.9	45.8	65.6	77.2	83.2	84.4	59.7	78.3	86.5
hat_Latn	1	30.7	52.7	78.9	42.1	58.1	67.9	77.6	79.3	58.0	77.1	85.1
hau_Latn	1	28.7	39.0	61.6	30.2	42.2	46.8	59.8	59.8	44.5	75.6	85.1
min_Arab	2	27.4	32.6	61.0	28.9	32.1	36.4	55.9	50.2	35.9	46.3	76.1
umb_Latn	1	28.9	34.8	58.7	30.5	42.2	45.8	52.6	53.5	35.8	46.8	61.1

Table 5: **Subsets of language tiers for I2S @ k=3.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	**Tier**	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
acm_Arab	3	33.7	48.7	83.2	47.6	60.6	73.9	85.8	87.1	61.1	77.3	N/A
aka_Latn	2	27.0	34.0	64.8	34.5	47.0	52.6	60.6	60.3	43.2	60.7	N/A
apc_Arab	3	32.7	48.9	83.9	46.2	61.2	73.6	84.7	87.3	60.3	77.4	N/A
arb_Arab	4	33.3	50.6	83.2	49.3	62.3	75.6	85.9	87.7	60.2	78.2	N/A
azj_Latn	2	32.2	47.9	82.2	45.5	63.0	72.8	83.6	85.3	59.4	78.1	N/A
bul_Cyrl	3	33.2	50.1	83.1	48.5	66.0	77.6	85.2	86.5	60.2	78.3	N/A
ces_Latn	4	34.9	52.2	82.2	48.6	66.0	78.8	84.7	86.4	58.9	79.3	N/A
eng_Latn	4	34.9	58.9	83.9	48.1	68.7	N/A	84.5	87.0	62.4	77.4	N/A
fin_Latn	4	32.6	51.1	82.2	46.8	66.6	76.3	83.7	86.3	60.7	77.0	N/A
hat_Latn	1	31.1	49.5	80.0	41.3	59.4	68.7	77.0	80.4	58.5	75.6	N/A
hau_Latn	1	28.4	35.9	63.3	31.4	42.8	46.2	60.5	60.4	44.7	73.9	N/A
min_Arab	2	27.4	30.1	62.0	28.5	32.3	36.3	55.7	49.8	36.4	43.9	N/A
umb_Latn	1	28.0	31.7	59.1	31.6	42.7	46.6	52.3	53.3	35.7	45.4	N/A

Table 6: **Subsets of language tiers for I2S @ k=5.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	**Tier**	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
acm_Arab	3	46.6	81.1	91.1	71.5	80.5	84.9	91.3	89.6	83.2	89.9	92.3
aka_Latn	2	32.3	53.2	68.1	44.4	59.6	64.4	68.6	66.5	63.5	79.9	87.7
apc_Arab	3	46.0	80.9	91.3	70.0	80.4	84.6	90.9	89.1	82.7	89.4	92.5
arb_Arab	4	48.1	82.7	91.4	73.1	81.3	86.5	91.7	90.2	84.9	89.9	92.8
azj_Latn	2	42.6	80.2	90.2	65.5	76.9	82.0	89.1	87.3	82.0	89.1	91.9
bul_Cyrl	3	47.1	83.3	91.2	73.4	83.5	88.2	90.6	89.2	83.9	89.3	92.5
ces_Latn	4	50.8	82.8	90.6	75.3	85.3	89.8	90.8	90.0	84.6	90.1	92.2
eng_Latn	4	56.7	85.7	91.8	81.4	87.0	91.3	91.8	91.5	85.4	88.5	92.0
fin_Latn	4	42.3	82.0	91.4	70.5	82.3	87.1	90.2	88.1	81.8	89.2	92.3
hat_Latn	1	40.2	70.8	87.3	57.9	74.1	79.2	84.9	83.7	82.0	87.3	91.6
hau_Latn	1	32.6	52.1	66.5	40.0	53.8	57.5	70.7	66.2	66.5	87.3	91.9
min_Arab	2	28.3	42.3	68.0	32.8	41.5	41.4	63.7	56.4	49.8	62.8	86.4
umb_Latn	1	30.3	47.1	61.4	39.3	52.4	57.2	N/A	59.6	53.4	62.9	71.7

Table 7: **Subsets of language tiers for T2S @ k=1.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	**Tier**	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	43.1	62.9	80.1	36.8	53.2	68.7	79.3	77.6	29.8	76.3	N/A
arb_Arab	4	44.3	66.5	80.2	37.7	55.2	71.9	81.4	79.9	30.8	77.7	N/A
ben_Beng	3	39.8	64.4	81.2	41.2	50.9	69.0	79.1	77.2	33.0	76.4	N/A
eng_Latn	4	41.9	71.7	N/A	47.7	66.2	N/A	81.7	81.3	35.3	77.5	N/A
fao_Latn	2	33.6	52.2	75.1	30.9	43.9	56.8	73.6	69.0	29.2	75.8	N/A
kac_Latn	1	28.8	35.4	51.5	26.4	37.8	44.7	51.6	46.9	26.7	46.6	N/A
kas_Deva	2	33.3	47.1	72.4	30.2	42.4	53.0	69.0	59.3	27.7	66.5	N/A
lit_Latn	3	38.1	63.6	78.9	34.7	50.2	68.8	78.4	75.1	30.1	77.4	N/A
lua_Latn	1	31.0	43.8	57.6	29.3	41.7	48.8	57.4	55.9	27.7	54.9	N/A
mal_Mlym	2	32.1	64.3	80.1	34.1	47.0	62.3	77.0	73.2	31.0	76.2	N/A
srp_Cyrl	4	42.7	66.2	79.5	37.0	53.0	71.3	80.1	78.2	31.7	77.9	N/A
tur_Latn	4	42.7	66.6	78.5	38.7	56.4	74.4	80.5	77.6	31.0	77.4	N/A
wol_Latn	1	31.5	42.1	56.8	28.4	40.0	49.8	57.0	54.0	27.2	56.5	N/A

Table 8: **Subsets of language tiers for S2I @ k=1.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	Tier	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	44.1	65.3	89.6	38.5	58.6	76.0	92.0	89.0	31.7	85.4	88.3
arb_Arab	4	44.8	67.6	89.1	39.2	59.8	78.9	92.4	90.1	32.8	86.8	89.1
ben_Beng	3	40.4	67.0	88.9	43.4	58.1	74.5	90.8	89.2	33.9	85.7	87.9
eng_Latn	4	43.1	70.4	88.5	44.5	69.0	84.0	92.1	90.0	36.1	86.4	89.1
fao_Latn	2	35.0	55.2	87.5	34.9	50.6	68.9	89.5	85.9	30.5	86.0	89.1
kac_Latn	1	31.6	41.0	69.7	29.2	46.3	57.5	72.3	68.0	27.4	64.2	70.5
kas_Deva	2	35.1	50.1	87.0	31.9	48.9	60.6	86.5	78.0	28.9	80.4	86.1
lit_Latn	3	42.1	64.7	89.1	39.2	58.5	77.4	91.0	89.4	32.2	86.2	89.0
lua_Latn	1	33.7	46.2	74.6	33.0	50.8	61.5	78.5	77.1	27.8	70.4	79.3
mal_Mlym	2	33.7	65.3	88.0	38.1	50.8	69.7	90.7	88.2	31.0	85.1	87.8
srp_Cyrl	4	44.1	65.1	88.7	39.7	59.4	78.3	91.8	89.8	32.6	87.1	89.5
tur_Latn	4	44.6	66.7	88.3	41.3	61.0	81.7	91.8	88.8	34.3	87.0	88.5
wol_Latn	1	34.8	45.8	75.2	32.9	49.7	62.3	79.2	77.0	28.1	72.7	83.1

Table 9: **Subsets of language tiers for S2I @ k=3.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	Tier	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	44.0	65.4	91.6	39.0	57.6	77.0	93.6	91.8	32.8	88.9	N/A
arb_Arab	4	44.3	65.7	90.8	40.0	59.9	79.4	94.0	91.9	33.5	88.7	N/A
ben_Beng	3	41.1	63.2	89.3	45.9	59.4	76.4	93.0	91.5	34.8	88.4	N/A
eng_Latn	4	43.4	68.6	89.6	43.0	68.7	N/A	93.7	91.7	35.6	89.1	N/A
fao_Latn	2	35.3	56.6	90.4	36.0	50.8	71.8	92.8	89.5	30.7	87.8	N/A
kac_Latn	1	31.8	42.3	74.8	31.6	46.2	60.6	79.2	73.9	27.0	68.4	N/A
kas_Deva	2	35.3	48.8	89.0	34.1	50.0	64.7	91.3	82.9	29.4	83.2	N/A
lit_Latn	3	42.1	63.5	91.1	39.5	57.7	78.1	93.8	92.1	32.8	88.2	N/A
lua_Latn	1	34.6	48.4	80.2	34.6	49.7	63.1	85.1	82.8	28.1	75.9	N/A
mal_Mlym	2	31.6	62.7	88.9	39.0	51.8	69.3	93.2	91.7	31.5	87.1	N/A
srp_Cyrl	4	44.5	63.5	90.7	41.2	58.4	79.4	93.6	91.4	33.3	89.1	N/A
tur_Latn	4	45.6	64.8	90.3	40.9	59.9	81.4	93.3	91.3	34.8	88.5	N/A
wol_Latn	1	34.4	46.3	82.1	34.8	50.9	64.3	86.1	82.2	27.9	77.9	N/A

Table 10: **Subsets of language tiers for S2I @ k=5.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	Tier	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	77.9	85.2	91.0	87.5	83.5	83.9	91.5	91.9	88.8	91.5	N/A
arb_Arab	4	81.5	86.5	91.7	88.9	86.1	88.0	92.8	93.4	89.8	92.7	N/A
ben_Beng	3	68.8	85.2	91.2	83.7	79.5	81.6	91.2	90.8	86.4	91.1	N/A
eng_Latn	4	86.1	89.1	91.0	90.6	91.7	92.0	92.4	93.5	89.9	92.4	N/A
fao_Latn	2	52.1	72.1	86.3	71.0	72.8	76.9	86.5	83.9	80.3	90.0	N/A
kac_Latn	1	41.3	51.5	58.8	55.2	54.6	55.7	57.4	58.0	54.7	58.2	N/A
kas_Deva	2	50.1	72.4	83.5	71.7	65.1	63.4	80.4	74.2	75.0	82.9	N/A
lit_Latn	3	67.2	81.7	90.1	76.6	77.7	84.6	89.6	89.6	84.2	91.1	N/A
lua_Latn	1	49.4	58.5	65.2	61.3	61.0	59.0	64.8	68.5	63.3	65.8	N/A
mal_Mlym	2	49.0	83.0	89.7	76.9	69.1	74.3	88.7	87.2	80.6	90.5	N/A
srp_Cyrl	4	77.3	86.1	91.0	84.8	85.6	87.0	91.5	92.3	88.2	91.7	N/A
tur_Latn	4	76.4	85.0	91.3	85.1	87.5	89.4	91.7	91.6	88.2	91.5	N/A
wol_Latn	1	50.4	59.1	66.8	63.3	64.0	62.0	66.6	67.6	65.5	69.5	N/A

Table 11: **Subsets of language tiers for S2T @ k=1.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	Tier	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	94.1	93.9	98.7	97.2	95.3	96.3	98.8	99.0	96.6	98.4	98.4
arb_Arab	4	95.7	94.4	98.7	97.3	96.7	98.0	98.7	99.1	95.7	98.6	98.4
ben_Beng	3	86.9	93.9	98.7	95.3	93.0	94.7	98.6	98.4	96.5	98.5	98.3
eng_Latn	4	96.5	95.3	98.3	98.0	98.1	98.7	98.7	98.9	96.7	98.1	98.0
fao_Latn	2	69.4	86.6	97.6	88.7	87.1	92.6	97.2	97.2	94.1	98.2	98.3
kac_Latn	1	57.6	68.0	80.1	75.0	77.7	75.5	79.4	79.9	75.3	78.6	79.1
kas_Deva	2	68.6	88.4	96.5	86.9	81.2	82.3	95.4	91.4	90.5	95.9	96.7
lit_Latn	3	87.1	92.4	98.6	92.2	94.9	97.6	98.2	98.3	94.9	98.6	98.2
lua_Latn	1	66.7	76.9	86.2	82.1	83.0	82.1	86.5	88.3	81.4	87.4	92.2
mal_Mlym	2	61.1	93.3	98.7	91.2	85.5	90.7	98.3	97.7	94.0	98.3	98.1
srp_Cyrl	4	93.2	94.4	98.7	96.2	96.5	97.6	98.7	98.9	96.5	98.3	98.4
tur_Latn	4	92.5	94.9	98.8	96.6	97.2	98.6	98.6	99.0	96.0	98.6	98.1
wol_Latn	1	66.5	77.1	88.9	84.5	84.9	86.0	88.8	89.6	85.0	89.3	95.1

Table 12: **Subsets of language tiers for S2T @ k=3.**

		Qwen2-VL			InternVL 2.5					Centurio Qwen	GPT-4o	
Lang	Tier	Qwen2-VL 2B	Qwen2-VL 7B	Qwen2-VL 72B	InternVL 2.5 4B	InternVL 2.5 8B	InternVL 2.5 26B	InternVL 2.5 38B	InternVL 2.5 78B	Centurio Qwen 8B	GPT-4o-mini	GPT-4o
aeb_Arab	3	97.2	97.4	99.7	98.7	97.7	98.8	99.7	99.7	98.3	99.5	N/A
arb_Arab	4	97.6	97.4	99.6	98.6	98.5	99.3	99.7	99.7	97.6	99.3	N/A
ben_Beng	3	92.6	96.9	99.5	97.6	96.8	97.9	99.6	99.5	97.7	99.2	N/A
eng_Latn	4	98.2	97.5	99.2	99.1	99.0	99.5	99.4	99.6	97.7	99.1	N/A
fao_Latn	2	79.9	91.8	99.4	92.7	92.9	96.3	99.2	99.5	96.9	99.0	N/A
kac_Latn	1	68.0	75.0	87.4	81.5	86.0	83.9	87.1	87.3	80.9	86.3	N/A
kas_Deva	2	78.4	92.4	99.0	92.3	87.9	87.5	99.0	96.6	93.8	98.7	N/A
lit_Latn	3	93.8	96.0	99.4	95.9	98.2	99.3	99.4	99.7	97.2	99.4	N/A
lua_Latn	1	75.9	83.3	91.9	88.1	90.1	89.1	92.4	93.5	85.2	93.1	N/A
mal_Mlym	2	67.6	96.7	99.6	95.5	90.2	95.7	99.5	99.1	97.4	99.2	N/A
srp_Cyrl	4	97.2	97.4	99.4	98.0	99.0	99.4	99.6	99.6	98.2	99.4	N/A
tur_Latn	4	95.6	97.2	99.4	98.4	98.8	99.5	99.4	99.7	97.6	99.4	N/A
wol_Latn	1	77.5	84.3	94.5	90.4	92.7	92.9	95.2	95.7	91.1	94.9	N/A

Table 13: **Subsets of language tiers for S2T @ k=5.**

A.7 Full Performance by Task, Model, and Language

A.7.1 Images-To-Topics

	Topics						
Model	Entertainment	Geography	Health	Politics	Science & Tech.	Sports	Travel
QwenVL-2.5-2B	99.4	83.1	100.0	99.9	100.0	100.0	95.7
QwenVL-2.5-7B	100.0	92.9	100.0	100.0	99.8	100.0	100.0
InternVL-2.5-4B	99.5	84.5	100.0	100.0	99.7	100.0	100.0
InternVL-2.5-8B	100.0	90.0	100.0	100.0	97.5	100.0	100.0
Centurio-Qwen	99.6	90.6	100.0	100.0	99.2	100.0	100.0

Table 14: **Image-To-Topics.** For a reference image, the model must pick the correct topic out of 4 choices.

A.7.2 Images-To-Sentences

Lang.	QwenVL-2 2B			QwenVL-2 7B			InternVL-2.5 4B			InternVL-2.5 8B			Centurio-Qwen			4o-mini		
k	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5
Avg	31.8	31.4	30.7	46.6	46.4	42.4	40.9	39.0	39.3	53.1	53.3	53.7	47.8	50.5	51.1	64.2	69.2	67.7
eng_Latn	36.3	34.8	34.9	65.8	63.1	58.9	52.5	49.2	48.1	67.7	67.9	68.7	54.8	60.0	62.4	68.3	78.1	77.4
ace_Arab	27.7	27.9	27.4	33.4	34.3	31.7	29.5	28.9	29.4	34.6	34.1	34.0	38.9	37.5	38.3	48.1	49.4	46.6
ace_Latn	32.4	33.1	32.0	45.5	45.7	42.0	39.9	38.7	38.4	55.7	57.1	58.0	46.1	51.7	51.5	62.4	67.0	65.3
acm_Arab	34.9	35.2	33.7	55.4	55.1	48.7	50.0	45.6	47.6	61.0	60.6	60.6	54.4	59.7	61.1	72.0	79.3	77.3
acq_Arab	35.2	35.0	32.8	56.4	55.2	50.0	50.8	45.8	48.3	61.8	60.8	61.0	54.1	59.1	60.9	72.6	79.5	78.0
aeb_Arab	34.6	33.5	32.2	53.6	52.0	46.7	48.2	44.2	45.1	59.2	58.2	59.3	55.3	59.4	60.3	71.8	78.6	76.3
afr_Latn	35.4	35.0	34.7	56.8	57.4	52.8	48.4	45.2	46.4	61.0	64.3	64.7	55.4	60.8	61.8	71.0	78.5	77.7
ajp_Arab	35.2	34.1	33.1	55.9	55.4	49.2	49.8	46.2	48.0	60.4	58.7	60.8	53.9	59.2	60.5	72.6	78.6	76.8
aka_Latn	28.7	27.9	27.0	38.0	38.3	34.0	34.6	34.2	34.5	46.3	46.3	47.0	42.4	42.6	43.2	61.4	63.3	60.7
als_Latn	33.4	33.5	32.4	55.4	54.7	51.0	43.9	42.3	42.9	59.1	61.0	62.8	51.1	56.6	58.3	72.6	78.5	77.4
amh_Ethi	24.9	25.3	24.9	30.1	31.5	29.5	33.8	33.5	32.4	29.7	28.3	27.4	44.3	46.8	47.1	57.2	58.1	54.4
apc_Arab	34.9	34.7	32.7	54.8	54.3	48.9	49.1	45.1	46.2	61.3	60.5	61.2	54.4	59.4	60.3	71.6	78.5	77.3
arb_Arab	35.2	35.2	33.3	57.1	56.4	50.6	51.7	46.9	49.3	62.2	60.9	62.3	55.2	58.7	60.2	73.2	79.7	78.2
arb_Latn	28.7	28.7	28.9	39.2	39.5	37.5	35.4	34.0	35.2	43.5	44.7	46.3	46.3	46.2	46.9	67.8	72.9	71.6
ars_Arab	34.7	34.6	33.1	56.4	56.4	50.7	51.3	46.6	49.2	62.0	60.7	61.5	54.7	59.1	59.7	72.7	80.2	78.4
ary_Arab	33.8	32.4	31.6	52.2	50.7	45.7	45.8	43.1	43.5	56.8	56.5	57.5	54.0	57.4	58.8	70.3	76.5	74.9
arz_Arab	35.2	33.9	33.2	56.1	55.5	49.6	50.9	46.5	47.9	61.4	58.8	61.6	55.1	60.0	61.3	72.9	79.0	77.9
asm_Beng	30.7	30.2	30.0	46.1	45.5	41.0	44.6	42.6	40.4	56.4	52.8	51.6	51.5	53.1	51.8	70.8	75.7	74.1
ast_Latn	35.9	34.9	34.1	58.7	58.0	54.2	46.8	43.6	45.1	61.9	63.8	64.2	54.9	61.6	61.7	69.3	77.5	76.3
awa_Deva	34.4	34.5	33.0	48.1	46.5	41.8	43.4	40.9	39.9	63.0	59.5	57.4	53.8	58.7	58.7	70.7	77.6	76.5
ayr_Latn	29.0	28.0	28.2	33.8	33.5	31.1	30.5	30.0	29.7	40.6	41.3	41.9	35.4	34.5	35.7	44.0	47.3	46.1
azb_Arab	31.9	31.3	31.0	44.3	44.2	40.1	33.3	31.8	32.7	49.8	49.2	51.3	52.9	54.2	55.5	63.1	68.7	66.6
azj_Latn	34.6	33.1	32.2	53.8	53.4	47.9	49.5	46.2	45.5	58.6	60.9	63.0	51.8	57.9	59.4	72.8	78.6	78.1
bak_Cyrl	32.5	33.8	32.2	49.3	48.8	42.6	41.5	40.1	39.3	54.9	55.8	57.0	49.4	51.2	53.1	71.0	76.2	75.0
bam_Latn	28.0	27.7	26.9	34.7	34.3	32.4	32.3	31.2	31.5	41.9	42.0	43.6	40.5	40.2	41.4	43.4	45.2	43.9
ban_Latn	35.5	36.3	35.1	51.9	54.3	49.1	41.2	39.2	39.4	56.4	61.5	61.1	51.6	56.8	57.6	67.3	72.3	70.3
bel_Cyrl	34.1	35.6	33.7	52.8	51.9	47.2	45.9	43.6	43.8	57.7	57.0	58.5	51.9	55.7	56.2	73.0	78.6	78.2
bem_Latn	28.7	28.9	28.6	39.5	40.2	36.6	35.4	34.7	34.9	46.4	47.7	47.3	39.9	40.3	40.9	56.0	58.0	55.4
ben_Beng	32.8	33.6	32.2	48.8	46.5	41.6	47.0	44.3	43.3	62.2	58.1	55.7	51.7	55.7	54.3	72.4	78.9	77.1
bho_Deva	33.8	33.5	32.5	45.3	44.1	39.7	42.3	38.6	38.2	61.2	56.2	56.5	52.3	56.3	57.2	70.5	76.1	75.2
bjn_Arab	28.0	27.2	27.3	31.6	33.8	31.2	27.7	27.5	28.6	33.5	32.0	32.5	35.5	36.4	35.4	47.1	47.9	45.1
bjn_Latn	33.3	31.9	32.1	50.6	51.9	48.0	41.8	39.0	39.4	54.2	57.3	57.8	53.6	57.3	59.2	66.4	73.3	71.2
bod_Tibt	25.3	24.5	24.4	25.6	26.2	25.0	25.5	25.1	24.7	42.7	41.7	40.7	38.0	37.6	36.5	32.5	32.7	30.5
bos_Latn	35.9	35.4	34.1	58.6	58.0	53.8	51.7	49.3	50.3	65.9	67.1	67.6	52.9	58.0	59.8	71.4	79.2	78.2
bug_Latn	33.2	32.6	32.5	44.5	44.5	40.7	37.1	36.2	36.8	51.5	54.0	54.2	45.9	48.6	48.5	59.5	62.0	60.3
bul_Cyrl	35.0	34.7	33.2	56.6	54.6	50.1	50.3	45.9	48.5	65.3	64.4	66.0	53.9	58.8	60.2	71.5	79.5	78.3
cat_Latn	36.6	36.0	34.6	59.4	58.5	54.0	51.1	46.0	48.3	64.9	66.4	66.4	55.1	59.0	60.8	71.4	78.6	78.4
ceb_Latn	33.4	33.3	32.7	55.5	56.1	51.2	47.4	44.2	45.0	60.4	62.4	62.7	53.6	57.7	59.4	71.4	77.2	76.1
ces_Latn	36.2	35.6	34.9	58.2	56.8	52.2	52.0	48.9	48.6	65.6	65.6	66.0	53.2	57.3	58.9	72.7	78.6	79.3
cjk_Latn	29.0	29.2	29.0	36.9	36.8	34.5	32.7	32.4	33.2	43.3	45.6	46.6	38.2	38.1	38.0	43.3	46.2	45.3
ckb_Arab	29.1	29.4	29.2	41.7	39.4	36.1	35.7	35.2	35.4	39.9	39.2	39.0	40.3	40.7	41.5	64.7	66.4	64.3
crh_Latn	33.6	31.3	30.7	50.6	50.1	45.3	43.1	40.9	40.9	58.7	60.8	61.8	53.4	57.5	59.1	69.4	74.5	72.9
cym_Latn	30.5	30.1	30.2	53.3	53.1	48.6	40.9	39.5	39.9	57.7	59.7	61.7	50.9	52.8	52.5	71.6	77.8	77.0
dan_Latn	36.7	35.9	35.0	59.8	58.4	53.2	53.3	48.1	49.3	64.7	66.4	66.9	55.2	61.1	63.6	70.8	78.7	77.5
deu_Latn	36.5	36.7	35.4	61.0	59.2	53.8	51.7	46.6	47.8	67.2	68.2	67.7	54.9	60.6	63.0	70.7	79.0	78.5
dik_Latn	29.3	28.4	27.7	35.7	35.9	33.5	33.1	32.4	33.5	45.5	45.8	46.6	37.9	38.1	38.3	44.1	45.8	44.3
dyu_Latn	29.0	28.1	28.0	36.3	37.3	34.0	33.0	31.4	32.4	44.1	45.6	47.6	41.3	42.7	42.6	45.6	48.8	46.4
dzo_Tibt	23.9	23.8	24.4	24.4	25.5	24.3	24.5	25.1	24.8	39.5	37.5	36.1	38.4	37.5	36.4	30.9	31.5	29.5
ell_Grek	34.0	35.0	34.6	52.5	50.7	46.2	45.5	43.9	41.8	62.7	62.0	63.3	50.5	53.0	54.0	73.3	78.7	77.8
epo_Latn	34.1	33.8	33.2	58.0	57.9	53.3	50.3	46.7	47.4	61.7	63.8	64.8	52.6	58.7	58.1	70.0	77.1	77.0
est_Latn	34.4	34.0	32.9	55.7	56.6	51.3	48.5	44.4	43.9	61.8	63.0	64.2	53.4	57.8	57.7	70.0	78.3	77.1
eus_Latn	33.5	34.2	33.4	50.3	50.4	45.8	44.0	41.3	42.6	60.1	61.1	61.8	52.2	57.2	58.6	69.2	76.7	75.1
ewe_Latn	26.5	27.0	26.3	33.8	34.1	30.6	30.3	30.6	30.0	41.3	42.1	42.1	34.8	34.3	34.7	40.7	42.3	41.1
fao_Latn	29.9	29.6	29.7	49.3	48.5	44.7	40.3	37.2	38.8	54.4	56.8	58.8	51.8	53.7	54.4	68.9	74.7	72.4
fij_Latn	28.5	27.4	27.5	37.4	37.6	35.0	32.0	31.3	32.4	43.1	43.9	45.4	39.0	40.6	40.2	60.4	62.8	61.2
fin_Latn	33.0	33.5	32.6	57.1	56.4	51.1	49.8	45.8	46.8	64.1	65.6	66.6	54.1	59.7	60.7	70.2	78.3	77.0
fon_Latn	27.4	26.1	26.2	31.9	31.5	28.8	29.8	29.9	30.0	39.9	40.8	40.9	34.9	34.3	36.1	36.9	38.0	37.6
fra_Latn	35.9	35.5	34.1	62.7	61.3	55.3	52.2	48.1	49.4	67.0	67.7	68.3	54.2	59.5	61.2	71.7	79.8	78.7
fur_Latn	34.2	32.4	32.2	50.7	51.4	47.8	43.2	40.7	41.6	56.8	60.0	60.1	52.6	55.5	56.1	66.6	73.7	71.9
fuv_Latn	29.2	28.4	28.2	38.0	38.3	34.5	33.9	32.8	31.7	43.1	44.3	45.0	39.0	39.2	39.7	50.5	54.0	52.2
gaz_Latn	29.0	27.7	27.4	34.2	35.3	32.4	31.9	31.1	31.6	38.7	39.3	41.1	39.3	39.5	39.2	65.4	68.2	65.8
gla_Latn	28.8	28.6	28.6	42.5	43.8	40.7	36.5	35.5	35.0	46.4	47.4	49.3	47.5	49.3	49.0	67.4	71.4	69.3
gle_Latn	29.0	29.4	29.1	46.6	46.9	43.3	37.5	37.7	38.2	49.5	51.4	52.8	49.1	50.6	50.8	70.1	73.9	72.2
glg_Latn	35.0	35.3	34.0	60.2	58.8	54.2	50.7	47.8	47.3	64.0	64.0	64.9	54.8	61.3	62.6	70.1	78.9	77.8
grn_Latn	33.4	32.9	32.4	44.1	44.4	41.2	37.6	37.6	38.1	52.5	53.7	55.1	45.9	48.6	48.8	64.1	68.6	66.6
guj_Gujr	30.7	31.6	30.2	46.1	44.8	39.4	44.8	42.9	40.9	60.6	56.0	54.6	48.4	50.7	50.4	70.9	77.5	76.4
hat_Latn	32.1	30.7	31.1	50.4	52.7	49.5	43.0	42.1	41.3	54.7	58.1	59.4	54.9	58.0	58.5	69.3	77.1	75.6
hau_Latn	29.1	28.7	28.4	39.2	39.0	35.9	32.4	30.2	31.4	41.8	42.2	42.8	43.9	44.5	44.7	70.2	75.6	73.9
heb_Hebr	35.4	34.9	34.3	57.0	56.8	51.8	47.9	44.9	46.8	58.5	57.2	59.1	54.5	60.3	60.3	71.7	80.4	78.7
hin_Deva	34.2	34.2	33.0	50.7	48.4	43.6	47.0	44.8	43.7	64.9	60.3	60.7	52.2	57.8	57.6	72.3	79.7	78.1
hne_Deva	34.8	35.4	34.4	45.6	44.7	40.1	40.7	38.8	37.0	61.7	57.5	56.8	53.1	56.1	57.4	71.1	76.1	74.5
hrv_Latn	35.5	35.8	34.8	59.4	59.2	54.7	51.8	48.6	49.7	65.9	67.9	67.9	54.4	59.0	60.3	72.7	80.5	78.7
hun_Latn	35.1	33.7	32.4	56.2	54.4	49.2	48.0	45.2	46.4	64.6	65.4	66.7	56.6	61.8	62.2	71.4	76.9	76.2
hye_Armn	27.8	28.2	27.3	48.5	47.7	43.2	33.0	33.2	33.0	35.0	33.1	33.3	42.3	43.7	43.4	72.2	77.2	76.3
ibo_Latn	27.8	27.1	26.8	37.2	36.9	34.2	36.0	35.0	35.0	46.1	44.6	43.7	46.6	47.1	47.1	69.8	74.3	72.6
ilo_Latn	33.5	33.5	33.0	50.7	52.7	48.1	43.0	42.1	41.9	57.9	61.1	61.7	49.9	53.6	54.0	69.1	75.2	73.4
ind_Latn	37.8	37.0	35.8	60.2	59.6	54.4	50.6	47.1	48.3	64.1	65.9	65.8	54.6	60.7	61.3	71.9	79.5	79.5
isl_Latn	28.7	28.0	27.8	50.2	49.5	45.7	40.4	38.6	38.5	53.4	55.0	56.5	50.9	54.3	56.5	71.1	77.6	76.4
ita_Latn	35.4	35.1	33.9	62.1	60.3	55.7	51.8	47.9	49.6	67.6	67.1	67.7	53.3	59.8	61.8	71.0	80.3	79.7
jav_Latn	34.9	33.7	33.8	54.1	56.0	51.1	43.2	41.1	41.4	55.6	59.2	59.8	52.4	57.7	58.2	70.2	77.6	75.7
jpn_Jpan	37.1	37.6	36.5	56.7	55.0	50.0	51.7	47.3	47.7	65.8	66.6	67.3	53.6	60.2	60.4	72.6	79.8	79.2
kab_Latn	26.0	25.9	25.4	28.1	29.1	27.4	27.4	27.1	26.8	33.5	33.3	34.6	32.8	31.5	30.6	35.0	34.0	34.2
kac_Latn	28.6	26.8	27.2	33.7	34.1	31.1	30.1	29.6	29.8	43.5	43.5	43.7	35.8	34.5	35.2	41.1	43.4	41.6
kam_Latn	28.3	27.7	28.0	36.0	36.9	33.7	33.7	32.5	33.5	44.6	45.2	46.2	38.5	40.2	40.2	51.4	53.8	52.0
kan_Knda	28.8	28.6	29.1	44.5	42.9	38.4	42.8	40.5	38.6	56.4	51.5	48.6	45.9	49.3	46.4	71.6	77.0	76.0
kas_Arab	33.5	33.1	32.0	41.4	41.0	37.0	35.4	35.4	34.3	51.0	50.4	50.1	47.7	48.9	50.0	65.3	67.9	63.8
kas_Deva	30.4	29.6	28.9	37.5	37.3	34.4	31.9	29.3	31.3	51.1	48.3	48.7	46.3	47.0	48.3	60.8	64.6	62.2
kat_Geor	29.7	30.3	30.1	47.5	47.0	41.0	39.8	38.9	37.5	37.4	34.6	33.3	49.3	52.6	53.2	71.8	79.5	77.7
kaz_Cyrl	33.5	34.3	32.9	49.9	48.3	43.8	45.1	41.5	42.1	56.0	56.0	57.7	53.9	57.9	58.2	70.8	79.2	78.3
kbp_Latn	28.5	27.5	27.1	34.2	34.4	32.4	31.8	30.5	31.5	42.5	41.3	41.8	34.6	34.0	34.8	37.9	39.7	38.8
kea_Latn	34.5	33.2	32.6	51.6	52.5	48.4	46.5	44.0	44.5	60.3	62.1	63.1	54.0	57.6	57.4	68.3	73.7	72.1
khk_Cyrl	31.0	31.2	30.6	46.9	45.5	41.0	37.5	35.7	36.3	45.7	44.9	46.5	42.4	43.4	42.5	70.5	76.3	75.5
khm_Khmr	27.5	28.3	27.6	44.6	44.1	40.6	40.5	39.6	37.5	47.9	44.2	41.5	48.9	49.3	48.4	69.0	72.5	70.3
kik_Latn	29.8	29.4	28.8	38.3	38.6	35.7	34.8	34.3	34.3	47.4	46.7	48.2	41.9	42.8	41.7	52.4	53.3	51.6
kin_Latn	28.2	27.4	28.2	37.6	37.1	35.3	31.9	31.8	32.0	44.6	44.5	44.7	36.3	36.0	36.4	70.7	75.3	73.9
kir_Cyrl	32.9	33.2	31.7	47.6	46.6	41.1	41.4	40.1	40.2	51.9	51.4	52.1	51.4	55.2	55.4	71.4	77.8	76.8
kmb_Latn	28.1	28.7	27.9	35.1	35.1	32.0	33.0	31.7	31.5	42.2	44.3	44.7	36.7	36.1	37.4	42.5	44.9	42.5
kmr_Latn	31.0	30.4	29.2	42.8	42.9	38.1	38.9	39.0	38.0	51.1	51.1	52.4	43.0	44.1	44.5	65.3	68.1	66.9
knc_Arab	25.9	24.9	25.4	28.0	29.1	27.8	25.6	26.4	25.6	31.4	29.7	30.4	31.6	28.5	29.0	32.4	31.8	30.4
knc_Latn	29.3	28.5	28.8	37.5	37.8	34.2	34.0	33.0	33.0	44.4	45.0	46.7	37.1	38.5	37.7	46.0	47.6	46.4
kon_Latn	30.4	30.2	29.5	41.8	43.1	39.3	36.4	35.9	36.5	49.2	50.6	50.9	43.0	43.4	44.3	57.1	59.2	57.0
kor_Hang	36.6	35.4	33.6	56.6	54.5	48.6	52.5	47.3	48.0	66.9	65.9	68.2	54.3	59.5	60.5	70.8	78.8	77.7
lao_Laoo	26.6	27.0	26.0	36.8	36.4	34.0	39.6	38.0	37.5	46.5	43.7	41.1	46.8	46.2	47.0	53.4	54.8	50.6
lij_Latn	33.2	31.8	31.5	49.8	50.1	46.7	43.1	41.3	42.7	56.4	58.6	59.1	51.1	55.8	55.2	66.3	71.1	69.7
lim_Latn	33.1	31.9	31.7	50.3	50.7	47.0	42.0	38.8	39.7	56.6	58.0	60.1	51.5	55.9	57.9	66.8	71.9	69.6
lin_Latn	30.5	30.4	30.9	40.7	42.2	39.5	38.0	36.7	36.8	49.7	50.9	52.1	48.8	52.0	52.8	65.1	68.1	66.8
lit_Latn	34.9	34.4	33.5	56.9	56.6	52.3	49.9	45.1	46.6	61.2	61.5	62.8	52.2	57.5	57.9	72.0	77.2	77.6
lmo_Latn	33.9	32.7	31.3	52.2	52.7	48.1	43.2	41.6	41.4	57.0	60.8	62.0	51.7	56.9	57.3	68.2	73.5	70.9
ltg_Latn	32.5	31.5	30.4	45.7	47.4	42.6	42.3	39.9	41.0	58.7	61.0	62.8	50.3	52.9	52.6	64.8	68.6	67.0
ltz_Latn	34.5	34.3	33.3	49.7	51.5	46.8	41.6	39.8	40.7	57.2	60.0	61.5	53.1	56.9	58.8	71.7	77.7	77.0
lua_Latn	30.2	30.3	29.5	40.2	39.9	36.7	36.0	35.3	35.7	48.3	50.5	50.3	39.8	40.8	40.8	51.9	53.1	53.0
lug_Latn	28.2	27.3	27.2	36.3	36.2	33.4	32.5	31.0	32.3	43.2	43.7	44.5	37.2	37.8	36.8	59.8	63.2	60.8
luo_Latn	29.6	28.3	28.0	36.9	37.5	34.5	32.3	31.4	31.3	43.2	45.1	44.9	38.1	41.0	40.3	47.4	49.8	48.2
lus_Latn	32.7	31.3	30.8	42.5	42.8	38.5	36.3	36.2	35.5	54.0	54.1	55.6	42.5	43.7	43.5	56.9	61.1	60.2
lvs_Latn	33.5	33.5	32.9	57.5	57.0	51.9	49.6	47.2	47.4	64.6	64.7	65.0	51.8	55.3	56.1	71.8	78.5	78.0
mag_Deva	34.2	34.3	32.1	46.3	45.2	40.5	41.6	39.5	38.2	62.4	58.1	57.0	53.1	57.5	58.1	70.2	76.3	75.0
mai_Deva	34.8	35.2	34.0	45.2	44.3	39.8	39.3	37.4	37.1	63.5	58.3	57.0	54.3	57.2	57.3	71.7	78.4	77.1
mal_Mlym	30.8	30.8	29.9	46.2	44.3	38.3	42.1	40.4	37.7	55.7	51.1	48.5	45.5	47.9	48.4	70.3	77.1	74.8
mar_Deva	32.0	32.2	31.3	48.2	46.4	42.3	42.7	40.6	39.9	57.9	53.5	54.4	51.5	55.2	55.5	72.0	78.4	77.5
min_Arab	27.0	27.4	27.4	30.4	32.6	30.1	27.8	28.9	28.5	33.6	32.1	32.3	37.1	35.9	36.4	44.8	46.3	43.9
min_Latn	34.4	33.9	33.4	50.5	51.3	46.4	41.5	39.0	39.4	55.0	58.8	59.0	52.7	56.7	57.4	66.7	71.9	70.7
mkd_Cyrl	35.2	34.6	33.4	56.1	55.6	51.3	48.0	45.0	46.0	61.1	59.8	61.7	53.7	58.2	60.5	72.0	79.0	78.1
mlt_Latn	33.5	32.7	32.9	53.0	53.5	48.4	42.9	41.6	42.5	58.9	60.1	63.0	51.1	54.8	54.6	71.3	77.6	76.3
mni_Beng	29.7	27.7	27.7	32.2	33.4	31.3	30.2	30.2	29.4	37.8	35.1	34.2	37.6	36.7	38.4	50.3	51.4	49.9
mos_Latn	29.5	28.3	28.0	34.1	34.9	31.4	31.3	30.9	31.0	41.9	43.6	44.0	36.6	36.6	37.4	42.3	43.8	42.1
mri_Latn	28.2	27.0	26.6	39.2	39.3	36.4	33.9	33.0	33.5	46.8	47.6	49.2	47.7	49.2	49.4	65.6	68.7	67.9
mya_Mymr	26.9	26.1	25.5	41.6	41.1	37.3	37.1	34.1	32.5	35.2	33.4	31.1	47.0	46.8	45.6	69.9	72.8	71.0
nld_Latn	37.5	36.9	35.4	60.6	59.9	54.8	50.3	47.8	49.1	65.5	66.3	67.2	55.2	62.0	63.5	70.4	79.0	77.7
nno_Latn	36.0	34.7	33.7	58.3	58.0	52.7	50.0	46.7	46.4	62.5	65.3	65.1	55.6	60.8	62.4	70.5	78.4	77.2
nob_Latn	36.9	35.8	34.7	59.3	58.5	53.8	51.4	47.9	48.5	64.3	66.6	67.0	55.7	61.6	63.9	70.1	77.9	77.4
npi_Deva	34.6	35.5	33.8	49.3	48.2	42.8	42.9	40.6	40.9	63.2	58.4	57.8	53.7	57.2	58.3	71.0	78.1	77.5
nqo_Nkoo	24.3	24.2	24.8	25.0	25.4	24.9	25.1	25.3	24.5	26.5	25.9	25.3	26.2	24.3	25.6	24.7	24.1	24.0
nso_Latn	29.2	28.6	28.4	38.7	39.2	35.7	33.5	31.4	32.7	45.2	46.8	48.1	40.6	40.9	41.4	62.9	67.5	65.9
nus_Latn	28.0	26.7	26.4	29.8	30.9	28.8	30.3	31.0	30.3	37.2	38.3	39.2	35.1	35.5	35.9	36.3	37.5	36.1
nya_Latn	30.3	28.8	28.5	40.3	40.6	37.5	34.8	34.8	35.2	47.2	48.6	49.1	40.6	41.0	41.0	68.1	72.7	71.9
oci_Latn	34.3	34.1	32.4	56.6	56.1	51.2	47.9	44.2	46.4	60.5	63.4	64.7	52.5	57.1	57.5	72.0	76.9	77.0
ory_Orya	26.7	25.9	26.3	41.5	39.5	33.6	43.8	40.8	38.1	57.5	52.2	50.3	44.5	48.9	48.5	65.4	67.2	63.9
pag_Latn	35.2	34.6	33.4	50.5	53.3	48.9	45.1	42.6	43.1	58.6	62.0	62.5	51.8	54.9	56.1	65.3	71.0	70.0
pan_Guru	30.5	31.6	30.4	45.5	43.7	38.3	44.5	42.4	39.1	59.5	56.5	54.1	49.6	50.7	49.0	70.2	74.8	73.2
pap_Latn	34.4	34.1	33.1	52.2	52.5	47.6	46.5	42.9	44.7	56.8	60.5	61.2	51.0	55.1	55.9	67.2	73.9	72.0
pbt_Arab	32.7	32.7	30.1	48.8	47.3	42.6	37.9	36.9	37.5	49.7	48.5	49.2	48.4	51.3	50.4	69.9	75.2	73.7
pes_Arab	35.5	35.0	33.7	55.4	52.6	47.2	48.5	45.8	48.3	63.1	63.5	64.3	55.7	59.3	61.2	72.7	79.6	79.0
plt_Latn	29.4	29.7	29.8	40.2	41.4	38.1	34.2	33.9	34.7	46.0	48.0	47.9	40.5	41.5	42.2	69.0	74.6	72.0
pol_Latn	37.1	36.4	35.0	60.1	58.7	53.1	50.7	47.0	49.0	65.9	67.5	67.6	54.3	59.6	60.9	72.3	79.2	77.8
por_Latn	35.2	35.2	33.9	61.8	60.6	54.6	52.5	47.7	49.1	66.2	66.5	66.5	54.5	60.3	61.3	71.2	79.0	78.5
prs_Arab	33.8	34.3	32.5	53.9	51.4	46.1	46.8	44.3	45.2	62.5	62.1	62.3	54.4	59.1	60.1	71.6	77.8	76.5
quy_Latn	30.8	30.5	30.0	38.7	38.9	35.5	34.2	33.3	34.0	47.0	48.4	49.1	43.3	44.1	44.2	50.7	54.3	53.0
ron_Latn	35.8	35.0	33.3	59.9	58.9	54.2	50.7	47.2	49.3	67.5	68.1	69.3	54.9	59.7	61.2	73.1	79.1	78.4
run_Latn	28.1	27.3	27.4	36.4	37.1	34.1	31.1	31.0	31.8	43.5	43.8	43.6	37.6	36.6	36.0	66.7	72.0	69.8
rus_Cyrl	35.5	35.8	34.1	59.2	57.1	51.9	51.1	46.3	47.3	68.6	66.0	67.5	54.1	58.4	60.4	71.3	80.9	79.3
sag_Latn	29.4	28.7	27.6	37.7	37.6	34.9	32.8	32.1	32.9	44.4	46.3	45.8	43.1	44.9	43.5	45.8	47.3	44.9
san_Deva	31.2	29.6	29.2	39.7	40.1	37.4	36.9	36.0	34.7	55.6	52.9	51.8	51.8	54.2	54.7	68.0	70.1	69.5
sat_Olck	24.7	24.5	24.9	24.8	25.1	24.2	24.7	24.9	24.1	26.3	25.1	24.6	26.3	25.2	25.8	24.9	26.0	24.5
scn_Latn	33.4	32.1	31.5	51.5	52.9	47.5	43.2	42.1	42.9	56.2	59.1	60.0	51.8	55.5	57.0	68.9	75.1	73.4
shn_Mymr	27.2	25.7	25.1	31.1	30.5	27.8	32.3	30.2	28.3	42.3	39.1	36.9	31.7	29.5	30.3	34.3	35.1	32.2
sin_Sinh	24.3	24.7	25.0	40.0	38.9	35.1	37.6	36.4	34.9	31.7	29.6	28.7	44.0	44.8	43.7	68.3	72.4	70.4
slk_Latn	35.3	35.0	34.0	57.4	55.9	51.5	51.3	48.3	49.6	64.9	65.0	66.7	52.0	57.3	59.1	72.4	79.3	78.4
slv_Latn	34.8	34.5	32.9	57.1	57.0	52.9	51.0	47.3	48.1	63.6	64.4	65.9	53.5	57.4	58.4	71.0	78.6	77.7
smo_Latn	28.1	28.4	28.7	38.5	38.9	35.3	33.3	32.4	32.6	44.6	45.8	46.5	47.9	48.6	49.4	67.9	71.2	71.1
sna_Latn	28.3	28.3	27.8	37.5	37.5	34.5	32.5	33.1	32.5	44.6	44.8	44.7	39.2	38.5	38.6	68.5	74.9	72.8
snd_Arab	31.2	30.5	29.5	44.2	44.4	38.8	41.1	38.8	39.2	50.1	48.1	48.8	52.1	52.9	52.8	71.6	77.8	76.3
som_Latn	28.3	28.2	27.7	41.1	41.1	38.6	35.7	34.2	34.2	42.0	43.6	44.2	45.5	46.2	48.1	69.0	74.2	73.1
sot_Latn	28.5	27.4	27.2	38.8	38.4	35.9	34.5	33.0	33.0	46.7	47.1	47.4	41.6	41.6	42.4	65.5	70.1	68.4
spa_Latn	35.4	35.1	33.9	61.4	60.3	55.4	49.8	46.1	47.8	67.0	65.8	66.3	54.9	59.4	61.9	70.3	79.3	78.3
srd_Latn	33.8	32.1	30.9	50.5	50.4	46.4	41.0	39.4	39.7	58.1	60.4	62.3	52.4	55.3	56.2	66.9	72.0	70.1
srp_Cyrl	35.9	34.7	32.8	54.5	53.7	49.4	48.7	46.7	47.2	60.6	61.2	63.3	54.7	57.4	60.1	73.4	80.1	78.4
ssw_Latn	28.6	27.9	27.8	38.1	38.3	35.0	33.8	33.3	33.3	45.0	44.3	44.5	47.4	48.3	48.7	65.5	68.8	67.7
sun_Latn	35.3	34.6	34.6	54.0	54.0	50.0	45.0	43.6	44.0	57.3	59.9	60.5	53.4	58.2	58.9	69.6	77.6	75.5
swe_Latn	36.4	35.5	35.0	59.2	59.5	55.3	53.1	48.3	49.5	66.0	67.5	68.0	54.9	61.5	62.4	72.8	79.3	78.8
swh_Latn	30.1	28.9	29.0	50.8	52.1	47.9	40.0	39.6	39.5	54.0	55.6	56.2	51.1	52.8	52.1	71.5	77.1	76.5
szl_Latn	32.9	31.9	31.8	48.5	48.7	44.9	40.0	37.9	40.2	57.1	59.4	60.2	49.2	53.4	55.1	67.5	71.5	69.2
tam_Taml	27.0	26.7	26.5	44.9	43.6	38.5	41.8	39.2	36.9	54.7	50.8	49.1	49.1	51.2	49.8	70.8	76.9	74.6
taq_Latn	29.3	29.0	28.7	37.0	38.0	34.7	32.5	33.2	33.3	44.5	46.4	46.0	37.9	38.9	39.6	45.1	47.4	45.0
taq_Tfng	24.2	23.6	24.4	26.4	26.8	25.9	26.5	26.2	25.6	27.6	25.5	25.8	27.6	25.1	26.0	26.6	25.5	26.5
tat_Cyrl	32.1	33.5	32.2	48.2	47.5	42.6	42.5	40.0	41.6	53.6	55.0	55.9	49.2	52.0	53.2	70.7	77.8	77.4
tel_Telu	29.4	31.1	29.2	45.9	44.9	38.7	43.2	40.1	39.0	59.0	54.8	53.2	47.0	49.2	48.6	71.2	77.4	75.4
tgk_Cyrl	31.8	30.8	29.4	46.6	45.7	41.5	38.7	37.9	39.7	47.2	47.6	48.7	47.6	49.7	50.3	71.9	77.8	76.5
tgl_Latn	34.0	33.9	33.0	56.5	56.6	53.1	49.0	45.0	46.3	65.0	65.2	65.9	53.7	57.8	59.4	72.6	79.2	77.5
tha_Thai	35.6	36.6	35.0	55.7	53.4	48.2	48.4	46.5	48.3	60.9	59.0	60.0	55.2	59.6	61.0	74.6	78.8	78.3
tir_Ethi	24.6	24.4	25.1	27.9	28.5	27.6	30.6	30.0	29.7	27.3	26.7	26.3	42.0	43.5	43.2	46.4	46.0	42.1
tpi_Latn	33.8	32.3	31.8	52.2	53.0	48.5	43.4	41.9	43.3	60.1	62.5	63.0	53.1	56.1	56.5	67.8	74.0	70.5
tsn_Latn	28.7	28.3	27.8	37.2	37.1	34.2	32.9	31.5	32.3	44.3	43.4	45.6	42.7	43.8	42.8	63.5	66.2	63.8
tso_Latn	28.0	28.2	27.6	37.1	37.1	34.0	32.2	31.8	31.8	44.2	44.8	45.1	46.4	47.4	47.6	61.0	64.8	62.9
tuk_Latn	32.0	31.9	31.2	47.9	47.8	43.6	40.3	38.6	40.2	55.1	56.3	57.8	46.4	50.7	51.6	70.6	76.3	75.0
tum_Latn	28.4	28.5	27.7	38.0	37.8	34.7	34.5	34.0	34.5	45.5	47.3	47.7	39.8	41.4	41.9	63.2	66.0	63.8
tur_Latn	36.5	36.1	35.0	59.2	58.5	53.5	52.8	47.5	48.7	66.2	67.3	68.8	55.9	62.2	63.3	72.5	80.3	79.2
twi_Latn	29.1	28.1	27.9	38.5	38.5	34.4	34.8	33.9	34.0	48.5	49.5	48.6	43.4	43.7	43.9	60.4	63.9	61.5
tzm_Tfng	24.4	24.3	24.1	25.9	26.7	25.0	25.8	26.0	25.3	25.9	25.9	24.8	27.9	25.7	26.0	25.9	26.6	25.4
uig_Arab	28.2	28.0	26.9	42.2	41.6	37.0	34.8	34.4	33.8	55.3	53.5	53.4	46.0	45.4	46.1	68.9	73.0	70.6
ukr_Cyrl	34.2	34.8	33.5	55.6	53.6	49.0	49.0	45.8	46.7	65.3	64.9	66.4	53.1	57.9	58.5	72.5	79.9	79.1
umb_Latn	29.1	28.9	28.0	34.8	34.8	31.7	31.1	30.5	31.6	42.5	42.2	42.7	35.4	35.8	35.7	47.5	46.7	45.3
urd_Arab	33.6	33.8	33.3	50.1	48.2	44.0	44.4	42.3	42.9	59.3	58.2	58.0	52.6	57.6	56.8	72.6	78.6	77.7
uzn_Latn	32.1	32.1	31.0	50.1	50.4	45.6	46.6	43.1	44.7	56.2	57.4	59.6	51.6	56.8	56.5	71.4	78.0	77.6
vec_Latn	34.5	32.7	32.3	54.0	54.3	49.9	43.8	41.1	42.3	58.4	61.4	62.1	52.4	57.4	59.6	68.8	74.8	73.1
vie_Latn	35.2	35.2	33.6	60.0	59.4	54.1	49.8	47.5	48.7	65.9	63.8	66.4	56.6	60.9	61.5	73.5	79.3	79.1
war_Latn	34.2	34.5	33.6	54.3	55.0	50.5	46.3	43.8	45.1	61.3	63.3	64.1	51.7	55.1	55.4	70.6	75.8	74.8
wol_Latn	30.3	29.8	28.9	39.6	40.3	36.2	35.1	33.7	35.0	48.3	49.9	50.6	44.0	45.3	45.9	53.9	57.6	55.8
xho_Latn	29.7	29.1	29.2	40.9	40.9	38.6	37.2	36.1	36.3	49.5	49.4	48.9	46.9	49.7	49.8	69.7	73.6	73.0
ydd_Hebr	27.9	27.4	27.5	37.4	39.8	36.2	30.2	30.8	31.5	32.8	32.4	31.1	41.2	41.9	42.5	67.1	72.8	71.0
yor_Latn	27.6	26.6	27.0	36.8	37.5	33.4	34.8	33.9	33.8	44.9	43.8	45.2	42.9	44.7	44.9	64.4	66.9	65.7
yue_Hant	37.0	38.0	37.1	60.1	57.5	53.9	53.4	49.7	50.4	66.7	68.2	68.5	56.1	60.8	61.8	72.4	79.8	78.6
zho_Hans	37.0	36.4	36.1	61.3	58.2	54.8	55.0	51.6	51.5	66.7	68.6	68.4	55.6	61.6	64.2	72.8	79.7	79.5
zho_Hant	37.0	37.5	36.8	59.9	57.1	53.7	52.8	48.6	50.4	66.5	68.1	68.2	55.4	60.5	62.1	72.5	80.2	78.7
zsm_Latn	36.6	35.7	34.6	59.1	58.7	53.6	48.0	45.3	46.2	62.9	64.2	64.8	54.7	59.7	61.1	71.9	80.0	77.3
zul_Latn	28.1	27.9	27.4	38.0	38.1	36.5	32.8	33.3	32.4	44.9	44.2	43.8	47.0	50.0	50.9	68.6	74.4	73.4

A.7.3 Topic-To-Sentence

Lang.	QwenVL-2 2B	QwenVL-2 7B	InternVL-2.5 4B	InternVL-2.5 8B	Centurio-Qwen	4o-mini
Avg	38.6	66.5	54.5	68.0	72.4	81.5
eng_Latn	56.7	85.7	81.4	87.0	85.4	88.5
ace_Arab	29.9	46.4	34.9	42.1	54.1	66.3
ace_Latn	39.6	68.7	52.0	71.8	72.9	80.6
acm_Arab	46.6	81.1	71.5	80.5	83.2	89.9
acq_Arab	47.5	81.6	71.3	80.6	84.0	90.0
aeb_Arab	45.7	79.6	68.4	78.2	82.1	88.9
afr_Latn	49.9	82.5	70.6	82.3	83.4	89.3
ajp_Arab	46.3	81.1	70.5	80.2	83.6	89.3
aka_Latn	32.3	53.2	44.4	59.6	63.5	79.9
als_Latn	41.2	77.6	57.9	76.1	80.0	89.4
amh_Ethi	26.0	38.1	37.5	31.2	68.5	81.6
apc_Arab	46.0	80.9	70.0	80.4	82.7	89.4
arb_Arab	48.1	82.7	73.1	81.3	84.9	89.9
arb_Latn	31.2	56.1	44.9	56.6	69.8	85.1
ars_Arab	48.0	82.4	72.7	81.2	84.3	90.0
ary_Arab	43.3	77.0	64.1	75.9	81.4	88.2
arz_Arab	46.6	81.2	71.5	80.7	83.1	89.3
asm_Beng	37.0	69.6	57.8	72.0	79.2	88.2
ast_Latn	50.1	82.1	71.6	82.2	83.5	89.4
awa_Deva	41.5	74.5	59.4	78.2	81.7	88.6
ayr_Latn	32.2	46.3	37.0	52.3	54.3	62.0
azb_Arab	38.7	69.5	46.0	68.0	77.8	83.2
azj_Latn	42.6	80.2	65.5	76.9	82.0	89.1
bak_Cyrl	38.1	74.3	53.1	69.6	74.7	88.2
bam_Latn	30.9	46.6	39.4	53.3	58.8	58.5
ban_Latn	46.2	75.3	56.5	74.8	79.0	85.4
bel_Cyrl	44.2	80.4	63.7	77.7	83.9	89.4
bem_Latn	31.1	52.9	42.3	59.3	59.9	72.6
ben_Beng	38.9	77.2	64.5	76.9	81.9	88.6
bho_Deva	39.1	69.8	56.4	76.8	80.3	88.0
bjn_Arab	28.8	44.3	33.0	41.2	52.8	64.8
bjn_Latn	43.3	73.6	57.2	72.8	78.6	84.9
bod_Tibt	26.3	28.4	26.9	52.5	48.9	49.9
bos_Latn	50.8	83.6	73.5	83.6	82.9	88.8
bug_Latn	39.9	64.8	50.4	68.8	69.3	77.2
bul_Cyrl	47.1	83.3	73.4	83.5	83.9	89.3
cat_Latn	51.9	83.8	75.2	84.0	84.2	89.9
ceb_Latn	41.8	79.7	63.6	79.9	82.1	88.6
ces_Latn	50.8	82.8	75.3	85.3	84.6	90.1
cjk_Latn	31.0	50.6	40.6	55.2	57.0	60.8
ckb_Arab	30.3	61.3	43.1	49.2	58.2	82.3
crh_Latn	41.1	74.8	59.4	75.7	80.0	86.4
cym_Latn	33.6	72.8	49.1	72.8	75.9	88.5
dan_Latn	52.6	83.6	74.8	84.5	83.3	88.5
deu_Latn	52.9	85.2	77.4	85.5	84.7	89.7
dik_Latn	32.0	48.6	41.0	59.3	55.2	61.9
dyu_Latn	33.4	51.3	40.4	58.6	59.0	63.0
dzo_Tibt	24.9	24.8	24.2	46.9	46.8	45.1
ell_Grek	40.4	77.7	60.3	79.2	81.6	90.1
epo_Latn	46.2	83.1	69.3	81.1	83.2	89.2
est_Latn	42.3	79.0	66.3	77.5	80.2	89.3
eus_Latn	40.2	71.2	55.6	74.5	77.5	88.4
ewe_Latn	29.3	45.1	37.2	53.5	50.5	57.2
fao_Latn	36.0	74.0	54.3	73.5	76.0	87.4
fij_Latn	31.6	51.8	38.8	57.5	56.7	77.6
fin_Latn	42.3	82.0	70.5	82.3	81.8	89.2
fon_Latn	29.6	42.6	36.5	50.9	49.2	55.2
fra_Latn	53.3	85.2	77.7	86.4	85.9	89.9
fur_Latn	43.4	76.4	61.4	77.5	79.9	86.6
fuv_Latn	33.0	52.0	42.6	56.7	56.9	68.5
gaz_Latn	30.3	45.8	38.9	48.4	53.1	81.7
gla_Latn	32.3	59.9	44.7	59.3	71.6	85.8
gle_Latn	32.1	66.5	47.6	62.2	73.8	87.7
glg_Latn	51.2	83.7	75.8	85.0	85.2	88.8
grn_Latn	40.2	64.1	51.7	69.6	75.4	81.3
guj_Gujr	34.0	72.8	56.5	74.3	79.0	88.7
hat_Latn	40.2	70.8	57.9	74.1	82.0	87.3
hau_Latn	32.6	52.1	40.0	53.8	66.5	87.3
heb_Hebr	46.0	82.7	69.1	76.9	82.7	88.8
hin_Deva	39.9	78.9	65.2	80.3	84.1	89.1
hne_Deva	41.1	70.6	55.6	76.5	80.3	88.3
hrv_Latn	51.5	84.0	73.2	84.0	83.1	89.7
hun_Latn	44.3	80.9	66.8	81.3	81.2	89.9
hye_Armn	28.0	69.7	40.5	38.4	64.2	88.4
ibo_Latn	29.6	49.4	41.1	57.0	71.1	85.8
ilo_Latn	39.8	73.7	56.6	76.0	76.3	87.8
ind_Latn	53.4	84.2	75.0	84.7	84.4	89.3
isl_Latn	34.1	74.1	55.2	70.6	77.2	88.7
ita_Latn	53.6	84.5	75.5	86.6	86.0	89.3
jav_Latn	46.0	77.2	59.9	74.4	80.9	87.5
jpn_Jpan	52.0	81.9	76.0	86.8	84.7	90.0
kab_Latn	27.4	35.7	29.7	40.3	41.3	46.3
kac_Latn	30.8	47.0	37.7	56.3	52.6	57.3
kam_Latn	31.6	51.6	40.4	56.2	57.9	69.7
kan_Knda	30.3	69.3	52.7	68.1	77.5	87.9
kas_Arab	36.1	62.5	46.3	67.9	71.9	84.9
kas_Deva	34.6	55.2	43.5	63.3	67.4	80.8
kat_Geor	32.8	69.0	51.9	44.9	76.8	88.9
kaz_Cyrl	39.2	76.3	58.9	71.7	81.1	89.0
kbp_Latn	30.4	45.3	38.1	54.2	51.4	57.1
kea_Latn	45.9	76.3	63.8	78.2	81.1	86.2
khk_Cyrl	33.3	69.1	44.6	57.6	61.4	88.5
khm_Khmr	29.9	63.6	52.7	57.9	73.0	87.4
kik_Latn	32.4	52.5	42.6	60.0	64.9	70.8
kin_Latn	31.2	49.9	39.8	54.4	54.8	86.8
kir_Cyrl	37.9	72.1	53.6	67.7	77.9	88.0
kmb_Latn	29.3	46.6	38.6	53.3	53.0	59.2
kmr_Latn	34.7	60.1	50.5	65.1	65.9	83.4
knc_Arab	28.0	34.2	29.0	37.3	36.9	46.1
knc_Latn	33.1	52.0	42.8	59.8	60.5	64.2
kon_Latn	33.6	59.5	46.0	63.5	65.4	73.9
kor_Hang	48.6	82.6	76.0	86.0	83.3	88.9

Lang.	QwenVL-2 2B	QwenVL-2 7B	InternVL-2.5 4B	InternVL-2.5 8B	Centurio-Qwen	4o-mini
lao_Laoo	28.2	53.1	47.6	57.8	74.6	80.1
lij_Latn	42.2	75.8	60.3	76.7	79.2	86.4
lim_Latn	42.5	77.8	61.9	75.7	80.0	85.1
lin_Latn	34.6	59.4	46.5	62.5	75.6	81.2
lit_Latn	44.6	81.3	63.7	77.3	80.3	89.6
lmo_Latn	44.1	77.2	61.3	77.4	79.9	85.9
ltg_Latn	38.2	71.0	54.6	72.6	75.4	83.4
ltz_Latn	44.8	75.6	59.7	75.7	82.8	89.0
lua_Latn	34.3	56.0	45.2	61.9	62.1	69.1
lug_Latn	30.5	48.6	39.9	53.4	53.6	76.5
luo_Latn	31.7	50.0	39.9	55.3	56.7	63.1
lus_Latn	38.0	62.6	49.2	69.2	68.3	74.7
lvs_Latn	42.6	81.4	66.1	78.0	82.2	89.5
mag_Deva	40.6	71.7	56.5	77.0	80.7	87.9
mai_Deva	42.5	70.5	53.8	77.7	80.7	89.3
mal_Mlym	32.5	70.9	52.9	67.5	77.7	88.0
mar_Deva	37.1	72.8	56.4	73.2	79.9	88.6
min_Arab	28.3	42.3	32.8	41.5	49.8	62.8
min_Latn	43.4	73.1	56.5	73.4	77.5	84.4
mkd_Cyrl	45.3	82.8	69.3	79.5	82.3	89.2
mlt_Latn	41.5	76.4	57.5	74.2	82.2	89.5
mni_Beng	29.3	46.1	36.2	48.3	57.1	68.2
mos_Latn	32.1	48.0	40.5	56.7	53.1	58.9
mri_Latn	30.2	53.4	42.1	60.4	69.3	83.1
mya_Mymr	27.8	54.7	44.2	41.0	67.2	86.5
nld_Latn	54.1	84.0	76.9	85.3	85.1	89.0
nno_Latn	50.5	83.3	72.8	83.7	83.5	88.4
nob_Latn	52.3	83.8	75.2	84.0	84.2	88.7
npi_Deva	42.4	73.8	59.0	76.3	80.4	88.3
nqo_Nkoo	25.6	24.8	25.3	26.4	27.0	26.6
nso_Latn	31.9	51.8	41.0	56.9	62.6	80.3
nus_Latn	28.4	38.7	34.8	46.1	45.2	51.5
nya_Latn	31.8	56.0	43.8	61.3	62.4	85.2
oci_Latn	47.9	81.5	70.3	81.4	83.8	89.0
ory_Orya	27.5	62.4	54.2	71.5	75.9	86.4
pag_Latn	42.8	75.3	59.5	77.5	78.6	84.3
pan_Guru	34.0	68.4	56.3	73.8	79.1	88.4
pap_Latn	43.3	75.1	62.7	76.8	78.8	85.4
pbt_Arab	37.7	70.1	48.6	65.8	72.8	86.6
pes_Arab	44.5	82.0	67.5	82.5	82.7	90.0
plt_Latn	33.6	53.8	42.0	57.6	59.9	87.0
pol_Latn	51.1	83.4	73.9	85.1	84.7	89.3
por_Latn	53.9	85.5	77.5	86.4	85.4	89.1
prs_Arab	42.7	80.0	65.7	80.2	82.0	88.1
quy_Latn	34.9	57.4	42.9	61.8	67.4	68.8
ron_Latn	50.0	83.9	74.3	85.3	84.2	89.5
run_Latn	30.3	48.3	37.7	53.5	54.2	84.1
rus_Cyrl	50.0	84.4	76.3	85.4	85.3	89.2
sag_Latn	32.2	54.0	41.5	59.2	65.9	63.1
san_Deva	36.1	59.2	50.1	69.8	78.6	84.8
sat_Olck	24.9	24.6	24.6	26.1	26.3	27.6
scn_Latn	43.5	76.4	59.6	76.2	80.2	87.7
shn_Mymr	27.3	41.2	36.7	54.0	44.5	52.2
sin_Sinh	26.2	54.2	43.9	35.6	65.7	86.6
slk_Latn	48.5	83.5	73.7	83.6	83.7	89.8
slv_Latn	47.6	82.2	70.9	82.0	82.0	89.0
smo_Latn	31.8	53.1	40.7	58.9	71.5	86.1
sna_Latn	30.4	51.7	41.1	56.5	57.6	85.4
snd_Arab	35.4	64.7	50.3	65.8	79.9	88.7
som_Latn	33.6	55.2	43.3	54.9	64.7	85.0
sot_Latn	31.7	52.7	42.0	58.3	61.4	82.7
spa_Latn	52.7	84.8	76.6	85.9	85.3	89.2
srd_Latn	42.3	74.3	56.9	76.9	80.9	84.7
srp_Cyrl	46.2	82.9	71.5	80.9	83.8	89.3
ssw_Latn	30.7	48.6	39.7	53.0	69.1	82.6
sun_Latn	45.2	77.0	63.7	75.8	79.9	88.1
swe_Latn	50.8	83.9	76.1	84.8	84.0	89.6
swh_Latn	33.5	68.2	48.6	68.3	75.6	88.7
szl_Latn	42.0	76.1	57.1	76.6	78.7	87.4
tam_Taml	28.3	65.9	50.4	67.4	77.2	87.5
taq_Latn	31.9	52.7	41.8	57.7	58.2	66.0
taq_Tfng	25.8	29.4	27.3	29.1	28.6	31.7
tat_Cyrl	38.1	73.5	53.9	69.4	75.0	88.8
tel_Telu	31.5	68.7	54.6	72.0	78.8	87.7
tgk_Cyrl	33.2	68.3	50.6	61.8	74.4	88.0
tgl_Latn	43.2	80.4	69.1	83.8	82.1	89.4
tha_Thai	48.0	84.0	74.2	79.4	84.8	90.3
tir_Ethi	25.2	32.8	33.1	30.3	62.0	69.6
tpi_Latn	41.7	75.4	57.5	79.2	82.6	86.4
tsn_Latn	31.3	48.1	39.6	54.5	63.9	79.7
tso_Latn	30.4	49.4	38.3	55.2	66.5	77.4
tuk_Latn	38.0	70.3	52.3	71.1	70.9	87.6
tum_Latn	30.5	51.7	40.5	57.6	59.0	81.8
tur_Latn	49.4	82.1	74.3	84.1	84.1	89.4
twi_Latn	32.6	54.2	44.5	62.0	63.9	78.7
tzm_Tfng	25.2	27.3	25.7	27.8	29.5	29.3
uig_Arab	29.2	61.3	42.7	69.2	67.8	87.1
ukr_Cyrl	46.6	83.7	71.3	83.3	84.8	89.7
umb_Latn	30.3	47.1	39.3	52.4	53.4	62.9
urd_Arab	39.7	76.7	58.6	76.5	80.5	88.9
uzn_Latn	37.9	72.2	58.2	69.6	80.7	88.7
vec_Latn	45.7	79.9	65.3	80.0	81.6	87.7
vie_Latn	53.9	84.8	74.7	84.0	85.7	89.7
war_Latn	42.8	78.4	61.5	79.2	82.3	88.4
wol_Latn	34.8	57.8	45.4	63.0	67.5	70.2
xho_Latn	32.1	55.1	44.6	61.7	73.0	86.1
ydd_Hebr	28.7	56.6	37.2	36.2	64.9	87.5
yor_Latn	29.0	48.0	41.4	54.7	60.1	80.7
yue_Hant	53.8	84.2	77.4	85.1	85.9	89.5
zho_Hans	51.6	84.0	78.5	85.8	84.6	88.9
zho_Hant	53.8	84.2	77.6	86.2	85.7	89.8
zsm_Latn	51.5	82.5	72.7	82.8	83.4	89.3
zul_Latn	29.4	50.0	39.9	55.5	70.7	85.8

A.7.4 Sentences-To-Images

Lang.	QwenVL-2 2B			QwenVL-2 7B			InternVL-2.5 4B			InternVL-2.5 8B			Centurio-Qwen			4o-mini		
k	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5
Avg	35.1	36.8	36.8	52.8	54.5	54.4	32.2	35.4	36.6	45.8	51.8	51.7	29.2	30.3	30.5	69.2	80.2	83.0
eng_Latn	41.9	43.1	43.4	71.7	70.4	68.6	47.7	44.5	43.0	66.2	69.0	68.7	35.3	36.1	35.6	77.5	86.4	89.1
ace_Arab	29.2	30.4	30.6	36.7	38.8	41.8	26.3	28.2	30.9	29.7	31.8	30.7	26.8	26.5	26.4	53.7	70.0	73.6
ace_Latn	36.5	38.0	38.0	52.5	55.8	56.9	31.2	36.9	38.2	45.5	54.1	54.6	28.5	29.4	29.4	66.5	80.9	84.9
acm_Arab	44.1	45.3	43.9	63.8	66.6	65.2	37.9	39.0	39.9	54.1	59.6	59.7	30.5	33.2	33.5	77.5	86.6	88.8
acq_Arab	44.6	45.6	44.4	64.7	66.1	65.0	37.6	38.7	39.9	54.7	60.4	59.9	30.1	32.6	33.5	77.2	86.6	88.6
aeb_Arab	43.1	44.1	44.0	62.9	65.3	65.4	36.8	38.5	39.0	53.2	58.6	57.6	29.8	31.7	32.8	76.3	85.4	88.9
afr_Latn	40.3	41.6	42.7	64.4	64.2	63.2	36.2	40.3	40.2	54.3	60.8	60.6	31.0	32.4	32.0	76.6	86.3	88.6
ajp_Arab	43.2	44.6	44.4	63.5	65.0	63.8	36.0	38.7	39.5	53.3	60.3	59.8	30.2	33.5	34.2	77.8	85.7	88.0
aka_Latn	30.2	31.7	32.4	43.1	45.2	46.6	28.2	32.3	33.6	40.8	48.9	48.7	28.4	27.9	28.3	66.4	79.5	82.9
als_Latn	36.8	39.7	39.6	60.1	62.3	64.5	33.4	36.8	38.8	47.9	55.8	55.9	30.4	32.3	32.1	76.8	87.4	89.3
amh_Ethi	24.8	25.0	25.5	28.7	30.4	31.8	26.0	30.1	32.4	26.1	27.8	28.4	28.6	27.7	27.7	70.1	80.6	82.9
apc_Arab	43.1	44.6	45.0	64.2	64.8	63.9	36.5	37.5	38.9	53.2	59.8	59.0	30.0	33.3	33.6	76.9	86.1	89.0
arb_Arab	44.3	44.8	44.3	66.5	67.6	65.7	37.7	39.2	40.0	55.2	59.8	59.9	30.8	32.8	33.5	77.7	86.8	88.7
arb_Latn	29.6	31.5	31.0	43.3	48.0	51.6	28.5	31.5	33.2	37.5	44.4	44.4	27.5	28.2	28.1	73.4	84.8	87.4
ars_Arab	43.9	44.5	43.8	66.4	66.9	65.4	37.9	39.1	40.1	54.8	59.6	59.8	30.9	32.7	33.5	77.9	86.4	88.7
ary_Arab	42.8	43.7	43.4	60.9	63.7	63.3	35.9	37.4	38.0	49.9	56.1	55.6	29.2	32.3	33.1	76.0	85.9	88.4
arz_Arab	43.3	44.9	44.0	64.0	66.1	65.0	37.5	39.2	40.2	53.9	60.8	59.8	30.0	33.3	34.3	77.6	86.6	88.8
asm_Beng	35.2	36.9	36.5	59.6	63.9	59.9	36.6	41.9	43.8	45.9	51.4	52.4	32.3	33.2	33.9	75.8	85.8	88.1
ast_Latn	41.8	43.0	42.5	63.7	64.6	63.6	37.9	40.8	40.4	57.7	63.8	63.6	30.2	32.7	34.4	76.5	86.3	88.8
awa_Deva	40.5	42.4	43.3	61.1	64.1	61.7	35.1	38.4	41.8	54.3	62.3	62.3	31.6	33.5	32.8	75.4	86.3	88.2
ayr_Latn	28.4	30.8	31.1	35.2	38.4	39.6	26.3	29.7	31.5	36.8	43.1	43.5	26.7	27.1	26.9	46.7	63.0	67.8
azb_Arab	37.6	40.0	38.7	53.4	58.1	60.2	31.1	32.6	35.7	43.8	49.3	48.9	29.9	30.7	30.9	71.0	83.8	86.6
azj_Latn	39.7	42.3	42.6	64.8	64.4	64.8	35.3	39.1	41.3	49.6	55.7	54.8	31.0	32.5	32.5	77.5	86.4	88.8
bak_Cyrl	34.6	39.2	38.6	57.4	60.1	61.2	30.6	36.4	38.2	42.9	48.8	48.2	29.2	29.5	29.9	76.0	86.1	88.7
bam_Latn	27.7	30.3	29.7	34.2	37.3	37.6	25.8	29.4	31.2	34.1	40.8	40.3	26.1	26.8	27.0	46.3	62.4	63.3
ban_Latn	39.9	41.6	41.3	59.5	62.9	62.5	33.6	38.5	39.4	49.7	58.2	58.3	29.4	31.6	32.3	72.6	84.2	87.1
bel_Cyrl	40.6	42.8	42.2	63.2	63.3	63.4	33.1	38.1	40.1	50.2	54.4	54.4	31.0	32.2	31.6	77.1	85.8	88.8
bem_Latn	31.5	33.3	33.2	41.5	45.5	46.5	28.4	31.8	34.2	41.6	48.2	47.9	27.6	28.0	27.5	58.7	75.3	80.2
ben_Beng	39.8	40.4	41.1	64.4	67.0	63.2	41.2	43.4	45.9	50.9	58.1	59.4	33.0	33.9	34.8	76.4	85.7	88.4
bho_Deva	39.8	41.8	42.7	58.9	61.4	61.3	34.9	38.6	41.2	53.5	60.1	60.2	30.1	32.5	32.1	76.2	86.5	88.5
bjn_Arab	28.4	29.5	28.3	34.0	35.6	37.5	25.3	26.7	29.7	29.2	30.9	30.0	26.2	26.4	27.0	51.8	68.0	71.7
bjn_Latn	38.2	41.3	41.4	58.2	62.8	63.3	32.8	36.8	36.3	49.3	54.9	55.8	28.9	31.7	32.7	71.7	84.1	87.5
bod_Tibt	23.6	24.7	24.7	23.9	24.0	24.4	22.8	23.4	24.9	33.3	35.6	36.4	25.7	25.5	25.9	39.2	41.7	41.5
bos_Latn	42.8	44.7	44.5	66.5	64.9	63.3	37.2	39.4	39.2	57.5	63.4	63.3	30.7	33.2	33.7	76.9	87.1	89.2
bug_Latn	35.8	38.1	38.5	51.4	53.1	53.8	31.1	35.4	36.7	44.5	52.5	53.4	28.2	29.2	29.8	63.6	79.2	82.8
bul_Cyrl	43.2	44.3	45.8	66.0	66.7	64.9	37.8	40.2	42.2	56.2	62.4	61.2	32.3	33.8	34.1	77.1	86.8	88.4
cat_Latn	42.3	43.9	43.3	65.7	64.4	62.1	38.7	40.9	40.3	57.9	63.2	62.6	31.6	34.1	33.6	77.8	87.0	89.3
ceb_Latn	38.6	42.7	42.0	62.4	63.5	65.5	35.1	39.5	39.1	54.3	59.6	59.9	31.0	33.7	34.1	77.0	86.3	87.9
ces_Latn	42.7	43.2	42.7	65.5	64.8	63.0	37.3	39.7	41.1	60.4	65.4	64.6	31.1	32.8	34.2	77.9	86.6	88.6
cjk_Latn	30.4	33.4	32.9	38.6	40.9	42.4	27.8	30.1	31.5	36.4	43.5	43.2	27.4	27.6	27.3	48.7	64.4	69.9
ckb_Arab	30.7	31.9	31.6	45.5	47.5	48.0	28.7	31.7	34.4	32.9	36.7	36.7	27.5	26.7	26.9	70.4	83.2	85.3
crh_Latn	37.8	40.2	41.3	58.9	62.6	62.3	33.6	38.3	38.5	47.6	55.2	53.1	29.3	32.0	33.4	74.9	86.1	88.5
cym_Latn	31.9	34.6	33.8	55.4	58.1	59.3	29.2	34.0	35.2	46.3	53.2	53.7	28.5	30.5	30.5	76.3	86.7	88.8
dan_Latn	41.9	43.1	43.2	66.2	66.5	64.7	38.3	40.6	40.5	58.1	63.1	62.3	30.7	32.9	32.8	77.1	86.2	88.4
deu_Latn	42.0	42.7	43.0	68.4	67.0	64.5	41.0	41.7	41.5	60.8	64.6	63.1	32.3	32.9	33.4	77.0	86.9	88.8
dik_Latn	29.6	32.9	33.1	39.3	43.3	45.6	28.2	31.9	32.9	38.4	48.0	47.0	26.7	27.8	27.1	49.0	66.8	72.6
dyu_Latn	29.7	32.3	33.1	37.6	39.6	42.3	27.3	30.5	32.2	37.4	45.1	44.7	26.9	27.7	27.1	49.4	65.9	71.5
dzo_Tibt	23.6	23.9	24.6	21.3	20.9	21.8	21.8	23.1	23.3	30.3	31.6	32.3	26.0	26.1	25.7	37.4	38.1	36.4
ell_Grek	38.1	40.2	40.9	63.1	63.4	61.7	36.1	40.3	41.1	51.9	55.5	53.9	32.9	33.0	33.6	76.7	86.5	88.5
epo_Latn	39.6	41.2	41.2	64.7	66.0	64.2	36.4	39.8	40.3	54.0	60.9	59.5	31.2	33.4	34.3	76.8	85.9	89.3
est_Latn	36.8	38.5	38.6	62.0	64.0	63.2	33.3	39.0	39.7	49.4	59.0	58.4	31.0	31.8	32.3	77.8	85.9	87.8
eus_Latn	34.5	37.4	36.7	53.8	57.2	57.5	30.7	37.2	37.0	47.0	54.8	53.5	29.4	31.9	31.7	76.7	85.8	87.6
ewe_Latn	28.3	29.0	30.4	35.5	36.9	38.0	25.8	27.7	30.4	35.3	41.1	41.0	26.6	26.4	26.5	44.9	61.0	65.0
fao_Latn	33.6	35.0	35.3	52.2	55.2	56.6	30.9	34.9	36.0	43.9	50.6	50.8	29.2	30.5	30.7	75.9	86.0	87.8
fij_Latn	29.9	32.8	32.5	39.7	44.0	46.3	26.3	30.1	32.2	37.3	46.0	47.3	26.7	27.2	27.6	63.8	78.1	82.3
fin_Latn	37.6	39.7	40.9	64.2	64.9	64.3	34.7	39.8	39.6	55.7	61.2	60.6	30.4	32.3	33.3	78.0	86.5	88.4
fon_Latn	27.0	28.7	29.2	35.0	35.2	35.1	25.2	27.7	30.1	35.2	40.7	40.3	26.2	26.3	26.5	43.8	57.6	61.9
fra_Latn	43.2	41.7	42.6	67.8	67.6	64.5	41.1	41.9	42.8	62.6	65.7	65.7	32.8	34.0	34.4	77.1	86.6	88.8
fur_Latn	39.1	41.3	42.0	59.9	61.5	61.5	33.9	37.4	38.7	50.8	58.8	58.8	28.8	31.1	31.4	73.8	85.6	88.0
fuv_Latn	29.6	32.2	32.7	39.2	41.3	42.3	27.8	30.7	31.9	37.5	45.6	46.0	26.8	27.1	27.7	52.3	70.8	75.8
gaz_Latn	27.7	30.0	29.7	33.8	36.0	37.3	25.6	28.9	30.7	31.9	39.3	41.1	26.1	26.6	26.8	67.0	82.3	85.5
gla_Latn	29.6	31.5	30.3	44.0	46.2	47.6	28.3	31.6	32.8	37.7	42.5	43.3	28.1	29.1	29.2	72.6	83.5	86.2
gle_Latn	30.1	30.8	31.3	47.1	49.0	50.3	29.1	31.3	32.6	39.1	44.6	44.5	29.5	29.8	29.4	73.7	84.1	87.5
glg_Latn	42.3	42.3	42.9	64.1	65.7	64.4	39.8	40.8	40.8	60.1	64.8	63.9	30.8	33.9	34.0	77.6	86.4	88.8
grn_Latn	36.4	38.6	39.5	51.5	55.5	58.7	31.4	36.6	38.9	44.6	55.3	55.8	28.2	29.6	29.5	68.8	81.9	84.9
guj_Gujr	34.5	35.8	34.3	63.7	63.8	60.5	36.3	41.1	41.4	51.7	58.1	58.5	32.5	32.2	30.8	76.5	86.2	88.7
hat_Latn	37.2	38.5	39.2	54.6	57.8	58.5	32.2	37.7	38.7	46.3	56.4	57.0	30.8	32.7	32.5	75.8	85.2	87.9
hau_Latn	29.2	31.5	30.8	38.0	39.7	40.6	26.0	28.4	29.6	35.6	43.3	43.5	27.1	27.8	27.7	75.2	84.1	86.7
heb_Hebr	41.6	43.6	43.4	65.6	65.9	63.3	37.1	39.8	39.5	50.4	56.8	56.1	30.9	33.5	34.0	77.0	85.6	88.2
hin_Deva	41.9	41.9	42.5	65.4	66.0	63.0	37.8	41.4	43.8	56.6	61.6	62.1	32.3	33.8	32.9	78.1	86.2	88.5
hne_Deva	38.9	41.3	41.4	58.0	62.0	60.9	34.0	37.3	39.8	52.8	60.7	61.8	30.3	32.2	33.0	76.1	85.7	87.7
hrv_Latn	42.8	43.2	44.2	66.9	65.7	63.6	37.5	39.9	39.6	58.0	62.5	63.3	30.6	33.2	33.9	76.9	87.5	89.3
hun_Latn	38.1	40.0	41.8	64.7	64.9	64.0	35.2	38.0	38.3	51.9	57.4	56.5	30.7	32.8	33.5	77.2	86.7	88.5
hye_Armn	30.2	31.6	31.3	55.9	60.5	56.8	29.4	30.9	33.2	28.3	28.2	26.8	27.5	28.4	28.5	76.6	85.6	88.5
ibo_Latn	29.2	30.7	30.0	37.7	41.0	42.9	27.7	30.3	32.5	39.0	45.9	46.0	27.9	28.5	28.2	72.4	83.8	87.2
ilo_Latn	36.5	39.0	38.1	57.6	59.2	59.9	33.7	39.2	39.9	50.2	59.4	59.7	29.4	31.3	31.5	75.3	85.1	88.3
ind_Latn	43.8	43.2	43.5	68.4	67.6	65.9	39.0	41.2	39.8	59.7	62.7	62.2	32.3	33.2	34.1	77.9	86.4	89.0
isl_Latn	30.1	32.2	32.4	53.9	56.4	57.6	30.4	34.1	35.4	41.3	48.3	48.0	30.2	29.7	30.1	77.4	86.2	88.2
ita_Latn	43.1	43.8	43.8	67.2	66.8	63.7	39.4	41.6	41.5	60.4	64.7	62.7	31.6	32.9	34.0	77.5	86.9	88.9
jav_Latn	38.8	41.4	40.7	60.1	62.3	62.4	33.8	38.1	37.5	48.7	58.0	57.5	30.7	32.8	33.4	75.5	85.9	88.0
jpn_Jpan	43.4	44.4	43.9	68.3	68.6	66.0	40.2	42.5	41.7	59.0	65.4	63.3	30.0	31.4	32.9	78.0	86.2	88.8
kab_Latn	26.9	27.6	27.7	27.5	28.4	30.1	23.3	25.5	26.3	29.4	32.9	32.5	26.3	25.9	25.7	35.9	46.0	49.1
kac_Latn	28.8	31.6	31.8	35.4	41.0	42.3	26.4	29.2	31.6	37.8	46.3	46.2	26.7	27.4	27.0	46.6	64.2	68.4
kam_Latn	29.2	31.0	31.5	38.7	41.7	43.1	27.1	30.1	31.0	38.2	45.2	45.2	26.9	27.8	27.3	53.2	70.0	73.9
kan_Knda	28.3	29.7	28.9	58.7	61.2	58.1	33.3	38.1	39.6	44.5	48.1	49.4	30.8	31.7	29.9	76.0	84.9	87.5
kas_Arab	35.9	36.2	37.4	49.4	53.4	54.3	30.9	33.5	35.7	42.9	48.2	47.6	27.8	27.7	28.9	70.9	83.3	86.3
kas_Deva	33.3	35.1	35.3	47.1	50.1	48.8	30.2	31.9	34.1	42.4	48.9	50.0	27.7	28.9	29.4	66.5	80.4	83.2
kat_Geor	33.4	35.2	34.8	58.5	63.7	61.6	30.7	35.7	37.2	31.1	32.5	33.3	30.2	30.4	29.7	76.7	85.8	88.4
kaz_Cyrl	36.9	39.9	40.5	60.5	63.0	64.8	32.2	38.0	40.1	42.6	49.2	47.7	30.6	31.9	30.7	77.7	85.6	88.2
kbp_Latn	26.5	28.0	29.4	36.2	38.0	39.2	25.2	29.0	30.9	36.4	41.8	40.5	26.8	26.6	26.9	45.0	57.8	62.8
kea_Latn	38.2	41.1	40.5	59.2	61.4	61.0	34.6	39.5	38.9	50.5	59.2	59.2	30.2	31.1	30.5	73.3	84.5	86.5
khk_Cyrl	31.9	33.8	34.5	50.5	52.3	54.1	28.2	33.2	35.8	36.0	40.5	39.3	28.0	27.4	26.9	76.3	85.6	87.2
khm_Khmr	27.6	29.2	27.9	51.7	55.4	52.2	31.8	35.9	37.4	41.2	46.6	46.8	29.4	30.5	30.4	75.7	84.7	86.5
kik_Latn	30.4	32.8	32.5	42.1	45.8	48.1	28.5	31.6	33.6	42.0	49.9	48.7	28.3	28.1	28.1	55.9	71.6	74.8
kin_Latn	28.5	29.9	30.2	37.7	39.8	39.6	26.2	28.7	31.4	37.3	43.6	44.0	26.9	27.9	27.0	73.6	84.4	87.2
kir_Cyrl	35.8	39.3	39.5	54.0	56.9	59.5	30.4	36.5	38.5	40.9	47.5	45.9	29.5	30.0	30.3	77.0	86.1	88.2
kmb_Latn	28.4	31.0	30.9	36.2	38.8	40.6	25.6	29.7	31.6	36.1	43.2	42.7	26.6	27.0	26.5	45.0	62.8	66.4
kmr_Latn	31.5	35.2	34.6	45.3	47.9	48.1	30.2	35.4	36.4	39.4	47.8	48.3	27.2	26.7	27.1	68.8	82.6	85.4
knc_Arab	26.5	27.2	26.9	26.7	28.6	29.1	23.7	25.1	26.8	27.8	28.5	28.4	25.1	25.4	25.7	36.1	44.5	47.2
knc_Latn	29.9	31.6	32.3	41.0	43.0	43.8	27.2	30.9	32.3	39.8	47.1	47.5	27.1	27.3	27.6	50.1	67.1	72.8
kon_Latn	32.0	34.9	34.2	45.3	49.5	51.9	29.1	33.2	35.6	41.5	50.7	51.0	27.3	28.6	28.3	58.7	75.3	80.3
kor_Hang	40.5	43.3	43.3	68.1	68.4	65.4	39.9	40.7	40.7	60.5	65.1	62.9	30.5	31.9	32.9	77.3	85.9	87.9

Lang.	InternVL-2.5-4B			InternVL-2.5-4B			QwenVL-2.5-2B			QwenVL-2.5-2B			Centurio-Qwen			4o-mini		
k	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5
lao_Laoo	28.1	28.7	28.0	40.8	42.6	40.3	30.2	33.4	34.3	39.4	44.1	44.6	29.9	28.9	29.2	67.7	77.8	79.5
lij_Latn	39.5	42.7	43.3	61.6	61.9	63.1	34.8	37.6	38.5	49.6	56.9	57.3	29.7	31.3	31.2	73.9	85.0	86.9
lim_Latn	39.1	42.3	41.4	60.5	62.6	62.8	33.5	38.6	38.7	51.0	59.0	58.4	29.4	31.0	31.0	73.3	84.7	87.3
lin_Latn	31.8	34.7	34.8	46.1	49.2	51.0	28.8	34.0	34.3	41.8	51.1	50.5	29.0	30.3	30.0	67.2	81.6	85.5
lit_Latn	38.1	42.1	42.1	63.6	64.7	63.5	34.7	39.2	39.5	50.2	58.5	57.7	30.1	32.2	32.8	77.4	86.2	88.2
lmo_Latn	39.8	41.4	41.4	61.1	62.9	63.2	34.2	38.4	39.6	50.4	58.6	58.0	29.4	31.3	31.8	74.2	84.8	87.9
ltg_Latn	37.0	38.1	38.0	55.5	58.0	58.7	31.3	35.8	37.2	46.3	56.0	55.5	28.3	30.0	30.1	71.8	84.0	87.0
ltz_Latn	39.9	41.7	41.5	60.3	61.3	60.3	33.5	37.5	38.3	49.7	56.9	55.9	29.9	31.3	31.3	76.3	86.5	88.3
lua_Latn	31.0	33.7	34.6	43.8	46.2	48.4	29.3	33.0	34.6	41.7	50.8	49.7	27.7	27.8	28.1	54.9	70.4	75.9
lug_Latn	29.4	30.6	30.5	37.1	39.0	40.0	26.2	29.4	31.5	36.9	44.5	43.6	26.7	27.3	27.1	60.1	76.8	80.2
luo_Latn	31.1	32.8	33.2	38.1	40.8	42.6	26.4	30.7	32.2	37.7	46.2	47.7	26.9	27.7	26.9	49.0	66.9	73.6
lus_Latn	33.2	36.0	36.4	48.0	52.4	53.2	29.9	34.6	35.6	44.8	56.7	56.7	27.1	28.5	29.2	61.0	76.5	80.0
lvs_Latn	40.1	41.9	42.0	66.8	64.4	64.4	35.4	39.7	40.0	51.6	59.1	59.2	30.9	33.0	33.5	77.0	85.4	87.6
mag_Deva	39.9	41.7	42.7	59.0	63.7	62.2	33.5	37.6	40.9	54.9	60.5	62.5	30.4	33.0	33.3	76.5	86.2	88.3
mai_Deva	39.4	41.7	40.9	58.5	62.4	61.2	33.7	36.9	39.5	54.5	61.6	62.6	30.3	32.8	33.3	76.7	86.2	87.9
mal_Mlym	32.1	33.7	31.6	64.3	65.3	62.7	34.1	38.1	39.0	47.0	50.8	51.8	31.0	31.0	31.5	76.2	85.1	87.1
mar_Deva	36.6	37.9	39.0	61.5	62.0	59.3	35.1	38.2	40.4	49.2	55.8	57.5	31.3	32.9	32.3	76.5	86.0	88.2
min_Arab	28.1	28.6	27.7	32.2	34.2	36.4	24.3	27.0	28.9	27.8	30.1	29.5	26.2	26.0	26.0	49.5	65.3	68.8
min_Latn	38.1	40.8	40.9	58.2	60.9	61.1	32.9	37.4	37.8	48.9	55.5	55.6	29.2	31.0	32.0	71.9	83.5	86.5
mkd_Cyrl	42.4	43.0	44.1	64.9	65.1	64.2	36.0	38.7	39.7	52.6	58.5	57.2	31.4	32.9	33.5	77.0	86.4	88.2
mlt_Latn	37.8	39.0	38.2	57.4	60.8	62.4	32.8	36.2	37.8	50.0	56.9	56.7	30.6	32.1	32.9	76.1	85.5	88.0
mni_Beng	28.4	30.0	30.0	41.2	47.4	47.0	29.3	31.0	32.6	35.0	37.5	40.6	27.1	28.0	28.0	54.9	72.0	77.7
mos_Latn	28.2	31.6	32.0	37.1	39.6	42.6	26.5	29.9	31.2	35.9	42.9	43.0	27.5	27.2	27.1	47.6	64.0	69.0
mri_Latn	28.5	30.5	29.8	40.0	42.1	43.1	27.1	29.4	31.6	36.3	42.1	42.0	28.2	29.2	28.3	70.0	82.8	86.0
mya_Mymr	27.6	27.7	27.0	52.1	56.1	56.1	30.7	33.6	33.9	30.8	33.8	34.2	27.9	27.9	28.0	74.6	84.1	87.5
nld_Latn	42.1	43.1	43.7	67.0	65.8	64.4	38.8	41.8	41.2	59.4	63.4	62.3	31.9	33.1	32.5	77.6	85.5	87.8
nno_Latn	41.1	42.4	42.7	64.6	64.3	62.8	37.2	40.7	40.5	56.1	62.3	61.4	31.1	33.1	32.8	76.5	86.3	88.9
nob_Latn	41.4	43.2	43.2	66.5	66.3	65.0	38.3	40.6	40.7	57.8	63.1	62.2	30.5	33.3	33.2	76.9	86.4	89.2
npi_Deva	38.8	41.8	40.6	60.3	64.6	61.0	34.7	39.3	40.9	50.8	58.3	59.1	31.1	32.5	33.2	77.1	85.9	88.2
nqo_Nkoo	23.9	24.8	24.7	23.0	23.3	24.7	22.8	24.4	24.6	25.3	25.7	26.0	24.8	25.4	25.1	25.6	27.6	28.8
nso_Latn	29.4	31.6	31.5	40.7	41.7	42.7	27.6	29.4	31.7	38.7	47.2	46.2	26.8	27.6	27.0	64.6	79.1	83.6
nus_Latn	27.2	28.5	29.2	31.7	33.8	35.4	25.4	27.9	28.7	33.2	38.4	38.6	26.9	25.8	26.5	41.2	53.3	57.8
nya_Latn	30.1	32.7	33.3	43.0	44.9	46.6	27.5	31.1	32.5	41.2	48.5	49.1	27.3	27.9	28.1	72.4	83.2	86.5
oci_Latn	42.7	42.7	42.8	63.0	63.7	63.3	37.7	40.1	40.3	55.7	62.0	60.8	30.9	32.4	33.0	76.9	86.4	88.2
ory_Orya	25.7	26.0	26.0	59.8	60.5	61.4	34.8	39.6	40.0	47.0	52.8	55.0	31.6	29.7	29.8	75.5	85.0	86.6
pag_Latn	38.4	40.7	40.8	58.9	62.2	61.9	34.4	38.9	39.8	52.0	62.3	61.7	29.1	31.0	31.0	71.1	84.1	87.3
pan_Guru	31.7	33.7	32.6	60.6	63.0	60.2	35.6	42.2	42.7	51.3	58.9	59.5	31.0	31.4	30.5	76.1	85.7	87.7
pap_Latn	38.6	40.5	40.1	58.2	60.4	59.3	33.9	39.2	38.7	51.2	58.7	58.4	28.8	31.6	30.7	73.6	83.6	86.8
pbt_Arab	36.2	38.8	37.9	54.2	56.4	56.1	30.6	33.7	37.6	42.1	44.0	42.8	28.3	29.2	29.5	76.0	84.8	87.9
pes_Arab	42.7	43.3	42.8	66.8	66.6	64.7	37.7	39.8	41.8	55.5	60.7	59.3	31.4	33.9	34.0	77.3	86.7	88.8
plt_Latn	30.4	32.5	32.3	40.8	42.1	43.2	26.9	30.9	32.2	38.1	45.9	46.4	28.1	28.4	27.4	72.5	84.5	87.3
pol_Latn	44.0	44.4	44.3	66.8	66.4	63.2	38.0	40.1	39.6	60.6	65.1	64.6	30.7	33.6	34.2	76.1	86.5	88.6
por_Latn	42.1	42.6	42.9	68.1	68.2	65.8	40.3	42.1	42.3	61.8	66.7	65.7	32.4	33.5	34.2	77.1	86.6	88.9
prs_Arab	42.1	42.7	42.6	65.3	65.5	63.9	37.8	38.3	40.8	54.7	59.4	59.4	31.7	33.2	33.4	75.9	86.4	88.3
quy_Latn	31.8	35.2	35.1	43.3	47.2	48.0	28.3	32.4	34.4	41.2	49.3	50.2	27.4	28.3	28.2	55.9	73.1	78.5
ron_Latn	42.5	44.9	45.4	67.8	66.3	63.3	38.9	41.0	41.0	58.8	62.1	61.5	31.5	33.0	34.2	77.1	85.9	88.3
run_Latn	28.4	30.4	30.0	36.6	36.9	37.6	26.0	29.2	30.2	36.0	41.5	41.8	26.3	27.8	27.2	70.6	83.2	87.3
rus_Cyrl	43.2	43.4	43.8	68.7	69.0	64.6	40.6	42.3	42.2	61.7	66.2	64.7	31.9	33.3	33.3	76.8	86.2	88.6
sag_Latn	31.1	33.5	34.3	39.4	43.1	45.7	27.6	31.2	33.0	38.5	48.1	47.2	27.9	27.8	27.4	51.6	67.3	72.9
san_Deva	32.6	33.8	33.4	54.5	57.6	57.3	33.6	36.8	39.2	46.6	53.3	54.9	29.6	30.4	30.5	72.5	82.5	86.1
sat_Olck	24.2	23.9	24.4	22.2	22.8	23.5	22.8	23.9	24.5	24.1	24.6	24.7	24.3	25.3	25.3	25.0	25.9	25.7
scn_Latn	39.4	42.3	43.1	61.3	63.9	64.3	33.5	38.1	39.6	49.8	58.1	58.1	29.2	31.5	32.0	75.1	85.7	87.8
shn_Mymr	26.8	27.9	25.9	39.4	43.1	44.5	26.6	28.3	29.1	40.5	48.6	49.5	25.9	25.8	26.0	44.3	54.8	56.6
sin_Sinh	24.4	26.0	25.0	46.2	47.3	45.7	31.5	35.4	36.2	28.5	29.7	30.7	29.5	28.5	28.6	74.9	85.0	87.3
slk_Latn	43.3	43.0	43.1	65.6	63.6	62.1	37.5	39.8	40.2	56.2	62.0	62.1	30.6	32.5	33.3	77.3	86.8	88.4
slv_Latn	41.3	43.7	43.9	63.8	64.3	62.8	36.3	39.9	40.1	53.6	60.4	60.6	29.8	31.8	32.6	77.2	86.5	88.4
smo_Latn	30.4	32.3	32.2	41.9	45.3	48.0	26.7	30.8	32.8	38.2	46.3	47.3	28.6	29.5	29.1	72.5	84.2	87.3
sna_Latn	29.6	31.0	30.6	38.4	39.6	40.5	26.5	30.2	31.9	38.2	46.0	45.9	27.9	27.5	27.3	73.2	84.4	86.9
snd_Arab	32.4	33.5	33.0	48.2	52.9	54.8	30.9	35.2	36.8	40.0	46.1	44.5	29.8	31.6	31.4	75.8	86.1	88.2
som_Latn	28.2	30.2	30.5	37.6	38.3	38.4	27.3	29.5	31.4	35.4	40.2	40.8	27.2	26.6	27.3	71.8	83.2	86.8
sot_Latn	28.8	31.4	31.1	39.8	41.7	42.5	26.9	30.5	32.0	38.4	45.4	46.0	27.0	27.5	27.0	68.5	81.4	85.6
spa_Latn	42.6	43.0	43.7	67.7	66.6	63.6	41.4	41.9	42.1	62.0	65.8	64.4	31.7	34.8	34.2	77.0	87.1	89.2
srd_Latn	40.1	41.5	41.8	60.5	61.6	62.1	33.2	38.7	39.3	50.5	57.9	58.3	30.5	32.2	33.2	72.3	84.8	87.4
srp_Cyrl	42.7	44.1	44.5	66.2	65.1	63.5	37.0	39.7	41.2	53.0	59.4	58.4	31.7	32.6	33.3	77.9	87.1	89.1
ssw_Latn	28.5	31.2	30.6	37.0	40.2	40.0	26.3	29.8	31.2	37.1	44.4	43.8	28.2	28.2	28.0	67.5	81.0	84.7
sun_Latn	38.8	40.7	41.1	60.3	61.9	62.2	34.7	38.8	37.9	49.0	56.2	56.6	29.6	31.6	32.5	75.1	85.6	88.1
swe_Latn	41.3	43.0	42.7	66.3	66.1	63.6	38.5	41.4	41.4	59.6	63.4	62.2	31.7	33.4	33.7	77.8	86.2	88.4
swh_Latn	30.5	33.1	33.6	48.6	52.6	54.7	28.9	34.0	35.1	42.6	49.0	48.6	29.5	29.4	30.3	76.3	86.9	89.7
szl_Latn	40.4	41.2	40.3	60.5	63.0	62.1	33.5	35.6	37.8	50.9	57.7	57.3	29.0	30.9	31.3	73.9	85.0	87.4
tam_Taml	27.8	28.7	27.6	57.8	61.4	58.9	33.9	37.9	40.2	42.4	47.8	48.2	31.5	32.9	31.2	75.9	85.1	87.6
taq_Latn	30.4	32.5	32.6	41.2	44.5	44.7	27.2	31.7	33.5	38.9	47.4	47.9	27.2	27.0	26.7	51.8	68.6	72.6
taq_Tfng	24.0	24.2	24.4	24.9	25.4	27.0	23.4	24.9	25.9	27.4	28.9	29.7	24.8	25.2	25.1	28.1	32.5	33.0
tat_Cyrl	36.3	39.3	38.4	55.5	58.8	60.2	31.0	35.8	37.8	42.5	48.9	46.8	29.1	30.0	29.5	76.2	85.4	88.4
tel_Telu	31.7	32.2	31.0	59.1	60.0	56.1	35.9	39.6	39.7	49.0	56.2	56.8	32.3	31.8	31.2	75.9	85.6	87.9
tgk_Cyrl	32.8	35.2	34.6	50.8	52.4	55.6	29.4	33.6	35.8	38.0	41.1	41.7	28.4	29.0	28.6	76.0	85.8	88.0
tgl_Latn	38.3	41.9	42.6	65.3	67.5	67.4	38.5	41.6	41.3	58.7	63.1	62.2	31.2	33.7	34.2	76.3	86.1	88.3
tha_Thai	44.7	44.6	44.4	69.9	70.5	69.6	41.4	42.9	43.1	53.5	56.1	57.9	30.5	31.2	31.8	76.6	86.1	88.0
tir_Ethi	24.5	24.0	24.1	27.0	27.8	28.2	25.1	27.9	30.0	25.6	25.9	26.5	27.8	26.8	27.4	57.2	69.3	71.4
tpi_Latn	37.9	40.0	39.5	60.8	63.7	64.3	35.4	40.9	41.9	54.2	63.5	64.6	31.1	32.2	32.9	75.5	86.2	88.5
tsn_Latn	28.7	32.4	31.9	37.5	40.2	41.1	26.8	29.2	31.1	38.5	44.8	45.6	27.4	28.0	27.6	64.7	78.4	83.2
tso_Latn	28.5	30.0	30.1	38.0	39.7	40.8	26.3	29.3	30.7	38.6	44.9	45.0	27.9	29.2	28.4	61.7	76.9	81.9
tuk_Latn	34.7	37.0	38.2	53.7	57.8	61.0	30.4	35.5	37.5	42.9	51.5	50.9	28.4	29.5	30.2	75.2	85.7	88.9
tum_Latn	29.3	32.1	31.3	39.7	43.3	45.3	26.8	31.2	32.9	38.9	45.5	44.9	27.1	27.6	27.4	65.1	80.8	85.1
tur_Latn	42.7	44.6	45.6	66.6	66.7	64.8	38.7	41.3	40.9	56.4	61.0	59.9	31.0	34.3	34.8	77.3	87.1	88.7
twi_Latn	30.7	32.1	33.2	43.6	45.8	47.3	28.1	32.4	33.2	40.9	49.1	48.6	27.7	28.0	27.5	66.3	79.8	84.1
tzm_Tfng	24.2	24.4	24.1	23.6	22.7	23.3	22.7	24.3	24.3	25.0	26.4	27.1	25.1	25.1	25.4	26.9	28.7	28.9
uig_Arab	29.9	30.9	30.8	50.5	52.8	55.1	28.7	31.7	34.1	43.6	47.5	48.3	28.2	27.5	28.2	75.6	85.3	87.5
ukr_Cyrl	42.2	43.1	43.7	66.4	66.3	65.0	38.5	41.4	43.4	56.3	62.0	60.9	32.3	33.8	33.0	77.8	86.3	88.2
umb_Latn	29.4	31.9	32.4	35.7	39.4	41.0	26.1	29.3	31.3	36.2	43.8	44.8	27.5	27.2	26.9	47.8	65.6	71.4
urd_Arab	40.3	42.5	42.7	62.0	64.0	63.8	35.9	37.4	40.2	52.3	57.4	57.7	31.7	33.1	32.6	77.2	86.5	88.3
uzn_Latn	34.7	36.5	35.9	59.2	63.0	63.4	32.8	38.0	39.5	45.3	52.1	51.3	30.1	31.7	32.9	76.1	85.5	88.1
vec_Latn	41.1	42.9	43.2	63.4	63.5	62.7	35.6	39.5	39.2	54.4	60.8	61.1	29.5	32.0	32.0	75.5	86.0	88.4
vie_Latn	42.5	42.7	44.2	68.2	67.6	64.6	39.9	40.3	40.2	59.5	64.1	63.5	31.7	34.1	35.5	77.1	86.5	88.8
war_Latn	40.0	42.1	42.3	63.3	64.2	64.3	35.6	40.4	39.6	54.7	62.2	60.6	31.7	33.0	33.4	76.0	86.2	87.7
wol_Latn	31.5	34.8	34.4	42.1	45.8	46.3	28.4	32.9	34.8	40.0	49.7	50.9	27.2	28.1	27.9	56.5	72.6	77.8
xho_Latn	31.1	31.9	32.1	41.4	44.6	45.0	28.6	32.5	33.6	42.0	49.5	48.9	28.2	28.6	28.3	72.5	84.3	86.2
ydd_Hebr	27.5	28.5	28.1	39.8	43.0	45.5	25.0	29.0	31.5	28.0	29.3	29.8	26.5	28.1	28.2	72.7	84.8	86.9
yor_Latn	27.1	28.6	29.1	36.6	36.7	37.9	26.0	29.3	31.3	36.6	42.0	40.7	27.7	27.2	27.4	68.7	80.8	83.7
yue_Hant	45.0	45.0	45.2	67.3	66.6	65.0	39.1	40.5	39.8	60.1	65.6	64.1	30.0	31.4	31.3	78.7	87.1	89.0
zho_Hans	44.9	45.5	43.7	68.4	67.6	65.1	38.4	39.8	39.1	61.1	66.1	65.1	30.3	32.2	32.5	78.0	86.7	88.4
zho_Hant	45.2	45.5	44.0	67.1	67.6	65.7	38.4	39.5	39.9	60.0	65.6	64.3	29.8	32.5	31.9	77.4	86.5	88.8
zsm_Latn	41.7	43.4	43.9	67.1	66.5	64.9	36.7	39.8	38.6	56.6	61.3	60.3	32.0	33.3	33.4	77.2	85.7	89.0
zul_Latn	28.8	30.3	31.0	37.6	38.8	38.8	25.7	28.9	30.6	37.4	43.4	44.0	28.9	27.8	27.6	74.1	84.1	87.7

A.7.5 Sentences-To-Topics

Lang.	InternVL-2.5-4B			InternVL-2.5-4B			QwenVL-2.5-2B			QwenVL-2.5-2B			Centurio-Qwen			4o-mini		
k	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5
Avg	58.2	73.2	79.5	70.3	82.8	87.7	70.8	85.4	89.5	69.6	83.8	88.4	75.5	88.0	91.5	82.9	93.5	95.8
eng_Latn	86.1	96.5	98.2	89.1	95.3	97.5	90.6	98.0	99.1	91.7	98.1	99.0	89.9	96.7	97.7	92.4	98.1	99.1
ace_Arab	36.9	48.1	54.6	55.2	70.0	77.9	52.5	69.0	73.6	42.3	50.9	54.3	58.1	75.5	82.1	66.8	85.4	91.6
ace_Latn	59.8	79.6	87.2	71.8	86.6	92.6	74.6	91.4	96.2	74.4	92.6	96.9	76.3	91.5	95.7	80.5	94.9	98.2
acm_Arab	80.8	95.2	97.4	85.7	94.2	97.4	88.0	97.1	98.6	84.9	95.9	98.5	90.6	96.2	98.0	92.2	98.5	99.4
acq_Arab	81.3	95.3	97.4	86.2	94.0	96.9	88.6	97.2	98.7	85.5	96.5	98.5	89.9	95.9	97.8	92.2	98.7	99.4
aeb_Arab	77.9	94.1	97.2	85.2	93.9	97.4	87.5	97.2	98.7	83.5	95.3	97.7	88.8	96.6	98.3	91.5	98.4	99.5
afr_Latn	77.0	92.9	96.1	82.9	93.6	96.4	84.3	96.0	97.9	85.8	96.6	98.9	88.4	96.1	97.8	91.6	98.3	99.3
ajp_Arab	79.2	94.4	96.8	85.4	94.1	97.3	87.6	96.9	98.3	83.2	95.7	98.2	90.1	96.9	98.2	92.0	98.0	99.5
aka_Latn	49.1	66.8	76.6	57.5	73.9	80.9	61.5	79.9	86.8	61.4	82.2	89.3	65.6	81.3	85.5	80.2	94.5	98.2
als_Latn	61.8	84.4	91.7	79.9	91.6	96.2	74.8	91.3	95.3	77.2	94.5	98.6	83.4	95.1	97.5	91.4	98.3	99.3
amh_Ethi	24.9	29.6	33.6	37.0	48.0	54.0	49.5	64.3	71.0	28.2	33.8	35.5	68.0	83.0	88.5	85.2	96.8	98.4
apc_Arab	78.3	95.0	97.8	86.7	94.9	97.7	87.9	96.7	98.1	84.8	96.0	98.1	89.5	97.1	98.5	92.0	98.4	99.5
arb_Arab	81.5	95.7	97.6	86.5	94.4	97.4	88.9	97.3	98.6	86.1	96.7	98.5	89.8	95.7	97.6	92.7	98.6	99.3
arb_Latn	39.2	51.7	60.4	58.5	76.3	83.9	59.8	79.2	85.4	56.6	76.1	84.0	69.0	87.0	92.6	87.7	98.2	99.1
ars_Arab	81.3	95.7	97.5	86.2	94.2	97.2	89.0	97.4	98.6	85.8	96.3	98.4	90.0	96.0	97.7	92.3	98.5	99.2
ary_Arab	76.6	92.6	96.9	82.7	93.2	97.2	84.7	96.1	98.0	80.1	93.8	97.1	87.4	96.1	98.5	90.6	98.0	99.5
arz_Arab	79.2	95.3	97.4	85.5	94.4	97.2	88.2	97.4	98.9	83.7	95.7	98.2	89.6	96.8	98.2	91.9	98.6	99.4
asm_Beng	54.8	77.5	84.5	80.4	92.0	95.2	79.6	92.6	96.3	71.9	85.2	91.0	84.6	95.6	97.7	91.2	98.3	99.2
ast_Latn	79.1	95.0	97.2	83.9	94.1	96.6	86.8	96.9	98.7	87.9	97.9	99.3	88.9	97.0	98.8	90.5	98.2	99.5
awa_Deva	68.6	88.1	93.1	83.7	94.0	96.9	82.7	94.9	98.1	82.3	94.6	97.9	86.1	95.9	97.8	91.7	98.4	99.1
ayr_Latn	41.9	57.5	68.8	49.3	66.6	74.9	51.1	71.0	79.0	50.8	70.0	78.2	54.1	73.5	81.1	57.6	78.9	87.3
azb_Arab	59.5	81.3	89.4	76.4	91.1	95.4	72.3	90.2	94.1	71.2	86.5	91.6	80.1	93.2	96.6	85.8	97.2	99.0
azj_Latn	69.2	85.1	92.2	82.1	92.9	96.7	77.6	94.2	96.6	79.0	94.3	97.9	84.7	95.3	97.9	91.6	98.8	99.5
bak_Cyrl	58.5	79.2	86.9	78.9	91.5	96.5	73.4	91.1	95.3	72.2	89.7	95.9	78.2	93.2	96.4	90.6	98.6	99.5
bam_Latn	40.7	52.7	63.1	47.6	65.5	73.6	51.4	71.1	78.7	52.9	72.5	82.1	57.7	77.2	83.2	55.4	78.8	87.4
ban_Latn	68.3	87.8	93.3	78.7	92.5	96.1	79.8	94.6	97.6	79.5	95.9	98.6	82.7	95.0	98.1	86.7	98.0	99.3
bel_Cyrl	73.3	91.8	95.5	84.5	93.7	96.7	82.4	96.4	97.9	83.4	94.9	97.6	89.1	97.2	98.6	91.6	98.6	99.4
bem_Latn	46.5	63.3	73.0	54.9	74.2	80.9	59.4	79.5	85.3	60.0	81.3	89.2	63.6	79.7	83.0	71.0	90.6	96.5
ben_Beng	68.8	86.9	92.6	85.2	93.9	96.9	83.7	95.3	97.6	79.5	93.0	96.8	86.4	96.5	97.7	91.1	98.5	99.2
bho_Deva	65.6	85.9	91.8	83.5	93.6	96.9	81.9	95.4	98.1	80.5	94.6	97.9	85.5	96.4	98.3	90.7	98.5	99.2
bjn_Arab	36.6	45.8	51.2	49.8	64.6	72.4	47.9	62.5	67.3	39.6	44.5	45.6	53.1	68.3	76.6	63.4	85.5	92.1
bjn_Latn	65.1	87.3	93.7	77.4	90.1	94.5	76.9	92.8	96.5	77.0	92.0	97.1	80.4	93.9	97.6	87.7	97.8	99.6
bod_Tibt	26.9	30.1	31.8	31.7	35.3	41.0	33.4	39.5	44.1	48.6	53.1	56.9	50.0	62.5	67.6	56.8	64.4	66.4
bos_Latn	79.5	94.2	97.1	85.0	93.9	96.1	86.2	96.6	98.3	87.7	97.8	99.2	88.9	96.6	98.4	91.6	98.5	99.4
bug_Latn	58.8	77.5	85.4	69.2	85.4	91.4	71.9	89.2	94.9	71.7	89.7	95.4	71.5	89.0	93.7	77.4	93.8	97.9
bul_Cyrl	79.5	94.3	97.4	85.8	94.2	96.8	85.8	96.7	98.3	87.7	97.4	99.2	89.3	96.2	98.2	92.1	98.7	99.4
cat_Latn	82.0	95.6	97.8	85.1	94.1	96.8	87.7	96.9	98.5	88.7	97.6	99.2	89.5	96.6	98.4	92.0	98.4	99.1
ceb_Latn	65.8	85.2	91.3	81.5	92.5	96.2	80.0	93.1	96.7	83.3	96.1	98.9	86.2	95.7	98.0	90.9	98.4	99.2
ces_Latn	82.9	95.1	97.6	84.4	92.7	96.1	87.3	97.3	98.7	89.5	97.9	99.2	89.4	95.6	97.7	91.8	98.4	99.2
cjk_Latn	44.4	59.5	69.7	53.0	69.8	78.0	55.1	73.3	81.0	54.0	74.7	84.0	59.2	76.3	82.5	59.4	80.8	88.8
ckb_Arab	42.7	57.4	61.5	67.4	84.2	90.3	61.5	80.1	86.2	47.1	57.2	61.4	61.7	77.4	82.1	85.0	97.4	98.9
crh_Latn	64.2	82.2	90.2	78.8	91.6	96.3	77.4	93.7	97.0	78.0	93.4	97.4	83.2	94.9	97.7	88.9	98.1	99.4
cym_Latn	48.7	63.2	73.3	74.2	86.6	93.0	64.7	82.2	88.2	74.2	91.5	97.0	77.4	93.1	96.7	90.5	98.2	99.4
dan_Latn	80.7	94.4	97.2	84.0	93.3	95.9	85.7	96.8	98.6	88.3	97.3	98.8	89.0	96.2	97.8	92.1	98.3	99.2
deu_Latn	84.5	95.9	97.9	85.8	94.4	96.7	86.8	97.3	99.0	88.5	97.6	98.8	90.0	96.1	97.7	91.6	98.4	99.3
dik_Latn	45.2	61.8	72.1	54.2	70.8	77.2	58.4	77.7	83.3	59.9	79.8	86.7	58.0	78.9	85.9	58.5	81.3	88.4
dyu_Latn	43.4	59.0	68.6	54.0	72.5	79.8	57.3	76.8	84.6	58.2	78.1	87.3	60.3	80.4	87.5	60.0	83.1	89.9
dzo_Tibt	24.2	27.1	26.6	24.9	22.7	24.1	28.7	27.8	27.0	41.2	41.7	43.3	44.5	55.7	59.5	54.4	57.8	57.6
ell_Grek	72.9	89.5	92.7	83.6	91.9	95.6	81.2	95.0	97.8	82.8	95.6	98.7	87.1	96.8	98.3	92.3	98.6	99.4
epo_Latn	70.4	90.4	96.0	83.5	93.3	96.8	83.1	96.6	98.2	86.0	96.9	99.2	88.2	96.3	98.0	91.5	98.6	99.3
est_Latn	63.6	84.0	91.4	79.9	91.4	95.7	76.5	92.7	96.6	78.5	94.8	98.1	84.2	95.2	97.5	91.5	98.4	99.4
eus_Latn	57.9	80.5	87.5	71.8	85.8	90.6	71.7	88.1	92.6	73.8	92.7	97.3	79.4	92.6	96.0	90.3	98.3	99.5
ewe_Latn	40.2	52.9	63.9	47.7	62.7	71.3	50.9	67.6	76.6	51.5	71.8	80.9	50.3	69.8	77.3	53.7	75.4	82.6
fao_Latn	52.1	69.4	79.9	72.1	86.6	91.8	71.0	88.7	92.7	72.8	87.1	92.9	80.3	94.1	96.9	90.0	98.2	99.0
fij_Latn	46.7	59.2	68.5	53.6	71.0	77.6	57.0	77.7	83.0	56.6	73.6	81.8	61.1	80.5	86.3	78.5	92.2	96.6
fin_Latn	64.6	82.8	91.3	83.5	92.4	96.0	80.0	95.1	98.5	85.2	96.7	99.2	85.7	94.8	97.6	91.5	98.4	99.2
fon_Latn	39.0	51.2	62.3	47.9	63.1	71.0	52.3	69.5	77.4	49.3	70.2	80.2	51.5	70.8	77.7	54.1	73.8	82.2
fra_Latn	85.8	95.8	97.9	85.9	94.8	96.7	88.8	97.3	98.7	89.4	97.9	98.9	90.9	96.6	98.4	92.1	98.2	99.0
fur_Latn	70.7	90.8	95.5	79.6	91.7	96.1	81.0	95.1	97.8	83.2	97.2	98.8	84.1	95.6	98.3	87.7	97.7	99.1
fuv_Latn	44.7	60.3	71.2	53.5	70.8	77.7	57.5	78.5	84.5	57.0	78.3	86.4	58.2	79.3	84.6	65.2	87.4	93.3
gaz_Latn	36.9	48.6	58.5	45.4	61.6	70.3	50.4	70.3	78.9	45.2	63.7	72.9	55.5	74.7	81.5	82.2	96.9	99.3
gla_Latn	43.8	55.4	65.7	60.3	76.6	85.0	56.6	75.0	80.9	57.5	76.0	85.1	73.7	88.7	94.2	87.6	97.4	99.0
gle_Latn	43.9	54.5	64.6	66.4	83.3	90.0	59.7	80.2	85.0	62.0	81.0	88.7	76.9	91.7	96.0	89.9	98.1	99.0
glg_Latn	83.0	96.0	97.6	84.4	93.8	96.6	88.2	97.2	98.7	89.3	97.9	99.0	89.3	96.4	98.1	91.4	98.3	99.1
grn_Latn	61.4	80.5	88.8	70.0	87.0	92.3	74.7	91.7	95.5	73.7	91.2	96.7	77.3	92.2	95.5	82.6	95.8	98.2
guj_Gujr	52.0	68.8	74.2	82.2	93.6	96.3	80.7	94.4	96.9	77.4	92.4	96.9	83.7	95.8	98.1	90.9	98.4	99.4
hat_Latn	58.0	79.9	87.9	74.3	90.2	94.4	73.9	91.4	96.2	75.0	92.8	97.9	83.0	95.4	98.1	88.7	98.2	99.3
hau_Latn	44.0	57.2	67.9	50.4	67.3	74.9	52.0	71.9	77.9	53.4	72.1	81.8	66.3	81.2	86.7	87.8	97.5	99.0
heb_Hebr	76.4	93.1	96.1	84.5	93.7	96.4	86.0	95.5	98.3	78.8	93.3	97.2	87.7	96.2	98.2	91.5	98.4	99.2
hin_Deva	72.2	88.9	93.5	85.5	95.0	97.7	85.4	96.4	98.2	83.9	95.6	98.4	87.6	96.6	98.2	91.8	98.4	99.4
hne_Deva	64.5	85.5	91.3	83.7	94.7	97.3	82.0	95.7	98.2	80.3	94.7	98.5	84.7	95.7	98.0	90.6	98.5	99.2
hrv_Latn	79.8	94.5	97.2	85.1	93.9	96.7	86.1	96.8	98.4	87.5	97.7	99.3	88.3	96.8	98.5	92.0	98.7	99.5
hun_Latn	69.8	87.2	93.4	83.6	92.9	96.4	81.3	94.8	97.7	83.4	95.7	98.5	85.3	95.9	97.6	91.5	98.6	99.3
hye_Armn	41.7	53.5	59.4	78.3	88.9	93.6	59.1	74.4	80.5	38.7	42.0	42.6	69.3	82.9	87.4	91.7	98.4	99.5
ibo_Latn	40.6	54.7	65.9	53.7	69.6	77.5	56.7	73.1	80.0	54.8	74.4	82.5	71.7	86.2	91.4	87.3	97.0	98.6
ilo_Latn	63.0	83.1	90.7	77.1	89.3	93.9	78.5	93.8	96.7	79.2	94.8	98.6	79.4	93.5	96.3	88.9	97.5	99.1
ind_Latn	84.2	95.6	97.7	86.1	94.4	97.1	88.1	96.5	98.6	88.4	97.6	99.0	89.7	96.3	97.5	92.0	98.2	99.2
isl_Latn	46.3	59.7	69.6	72.1	85.0	90.7	67.6	87.3	91.8	69.7	83.4	89.5	81.0	94.0	96.9	91.1	98.1	99.2
ita_Latn	84.4	96.1	97.8	86.9	94.7	96.9	88.0	97.2	98.7	89.6	97.8	99.0	90.3	96.7	98.2	91.6	98.4	99.3
jav_Latn	66.8	88.0	94.3	79.1	91.8	95.7	78.8	93.6	96.8	78.3	93.4	97.8	82.9	95.3	97.7	89.9	98.1	99.4
jpn_Jpan	83.8	95.9	97.8	86.2	95.4	97.3	87.6	97.0	98.7	89.4	97.7	98.9	89.8	96.7	97.8	92.1	98.2	99.1
kab_Latn	31.8	37.2	43.9	35.6	41.3	45.0	37.5	46.9	53.4	40.8	49.3	53.8	43.5	52.8	57.7	42.8	58.2	65.3
kac_Latn	41.3	57.6	68.0	51.5	68.0	75.0	55.2	75.0	81.5	54.6	77.7	86.0	54.7	75.3	80.9	58.2	78.6	86.3
kam_Latn	44.5	57.4	66.8	52.2	68.1	75.8	53.6	71.8	79.1	55.6	75.9	83.7	59.8	77.8	82.7	65.7	86.7	92.6
kan_Knda	39.9	48.0	51.5	78.7	90.8	94.9	74.3	90.6	94.4	70.4	83.5	89.1	80.5	93.1	96.5	89.9	98.5	99.3
kas_Arab	50.3	67.9	78.7	74.7	89.9	94.5	72.9	89.9	94.1	70.4	83.2	89.9	74.7	90.2	95.1	86.7	97.1	99.4
kas_Deva	50.1	68.6	78.4	72.4	88.4	92.4	71.7	86.9	92.3	65.1	81.2	87.9	75.0	90.5	93.8	82.9	95.9	98.7
kat_Geor	54.1	72.4	80.3	80.6	91.2	95.5	69.0	88.3	93.2	45.8	52.9	56.6	81.7	94.7	97.8	91.7	98.5	99.3
kaz_Cyrl	61.1	80.4	87.8	81.8	92.7	97.3	74.8	92.6	96.2	73.3	90.7	96.7	84.7	95.6	97.5	91.8	98.5	99.5
kbp_Latn	37.4	52.2	62.5	51.7	69.2	77.5	52.2	70.3	77.5	55.7	74.7	82.7	53.8	71.2	76.7	53.7	74.6	82.4
kea_Latn	69.9	88.7	93.5	79.0	92.1	96.3	81.7	95.2	97.8	83.0	96.1	99.1	84.0	96.4	98.7	87.5	97.4	99.1
khk_Cyrl	43.9	61.6	69.5	71.4	84.4	91.1	61.6	79.5	84.8	55.0	70.7	77.5	63.1	78.0	83.1	90.5	98.4	99.5
khm_Khmr	41.9	57.7	63.1	73.4	85.1	90.6	70.0	83.0	87.3	60.1	76.9	83.1	77.1	90.4	94.9	90.3	97.9	99.5
kik_Latn	47.0	63.7	73.3	56.7	73.2	79.4	60.9	79.1	85.3	58.9	78.9	84.6	66.1	80.2	85.0	67.8	87.3	92.8
kin_Latn	42.2	56.1	67.0	51.8	68.9	77.4	55.1	74.0	81.4	53.8	73.6	83.0	57.3	74.3	79.2	88.0	98.1	99.3
kir_Cyrl	57.0	76.9	85.2	75.5	89.2	95.0	72.1	90.1	95.4	68.2	85.2	91.7	80.2	93.3	96.9	91.5	98.6	99.5
kmb_Latn	41.8	53.5	62.8	48.7	66.0	72.7	52.0	70.5	77.8	52.2	72.4	79.8	54.5	72.9	77.1	55.1	77.2	84.8
kmr_Latn	49.1	67.6	79.0	63.8	80.8	86.5	66.8	86.2	90.9	63.8	84.8	91.4	66.3	85.8	90.1	82.5	96.4	98.9
knc_Arab	30.9	37.6	42.9	36.9	49.2	57.5	36.2	46.3	51.7	33.0	40.7	41.4	35.9	48.7	57.0	41.6	59.0	68.3
knc_Latn	46.1	61.8	72.8	55.1	71.6	79.2	58.8	79.8	86.4	59.2	79.4	87.3	60.7	79.9	86.1	61.7	83.2	91.2
kon_Latn	49.7	66.2	74.9	59.8	79.3	86.3	62.8	83.6	89.8	62.7	83.2	91.9	65.6	82.3	87.8	70.1	91.0	95.8
kor_Hang	81.8	93.8	96.7	86.8	94.6	96.8	87.5	96.8	98.9	88.3	97.7	99.1	88.7	95.8	97.7	91.1	98.3	99.2

Lang.	InternVL-2.5-4B			InternVL-2.5-4B			QwenVL-2.5-2B			QwenVL-2.5-2B			Centurio-Qwen			4o-mini		
k	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5
lao_Laoo	38.6	52.4	59.0	61.2	75.9	82.9	67.8	82.4	87.9	55.3	72.5	80.7	77.5	91.4	94.8	83.8	94.4	97.3
lij_Latn	72.5	89.3	94.3	80.4	91.7	96.2	82.6	95.3	98.1	83.0	95.5	98.7	83.8	95.7	98.0	88.4	97.3	99.0
lim_Latn	70.2	89.4	94.8	78.5	92.0	95.9	82.1	95.8	97.8	80.8	95.2	98.4	82.7	96.1	98.2	87.5	97.5	99.1
lin_Latn	50.9	67.4	77.3	61.4	78.3	84.1	65.4	84.0	89.6	62.7	83.8	91.7	75.4	89.8	93.3	80.3	95.7	98.5
lit_Latn	67.2	87.1	93.8	81.7	92.4	96.0	76.6	92.2	95.9	77.7	94.9	98.2	84.2	94.9	97.2	91.1	98.6	99.4
lmo_Latn	72.1	90.6	95.2	81.2	93.0	96.4	83.7	96.0	98.6	83.3	95.8	98.7	84.3	96.0	98.0	89.1	98.0	99.2
ltg_Latn	60.4	80.0	89.7	76.4	90.6	95.2	73.6	90.8	95.9	74.0	93.5	97.9	79.5	94.3	97.4	86.3	96.8	99.1
ltz_Latn	71.0	89.3	94.6	79.3	91.2	95.4	79.5	94.8	97.5	81.6	96.2	98.8	86.3	96.7	98.5	90.6	98.6	99.3
lua_Latn	49.4	66.7	75.9	58.5	76.9	83.3	61.3	82.1	88.1	61.0	83.0	90.1	63.3	81.4	85.2	65.8	87.4	93.1
lug_Latn	41.9	54.6	63.5	49.3	67.1	74.8	52.0	70.8	78.3	52.7	73.4	82.3	54.7	73.3	79.1	72.4	91.3	96.3
luo_Latn	45.0	60.8	69.6	52.9	69.0	77.7	57.9	76.5	84.5	56.2	78.0	87.4	57.9	76.4	82.8	60.5	82.2	90.2
lus_Latn	53.5	74.1	82.5	65.8	82.8	88.4	69.7	87.4	93.1	70.6	89.5	95.5	70.0	88.0	92.8	73.9	90.9	95.8
lvs_Latn	67.8	86.4	94.2	83.8	93.0	96.4	80.5	94.0	97.1	80.7	95.2	98.1	87.2	95.3	97.8	91.3	98.4	99.3
mag_Deva	67.2	86.6	92.9	84.5	94.7	97.3	83.0	95.4	98.1	81.7	94.9	98.4	85.4	96.6	99.0	91.5	98.4	99.5
mai_Deva	66.4	86.2	92.0	83.3	94.5	97.6	82.8	95.7	98.2	82.3	95.1	98.6	85.4	96.5	98.7	91.7	98.4	99.4
mal_Mlym	49.0	61.1	67.6	83.0	93.3	96.7	76.9	91.2	95.5	69.1	85.5	90.2	80.6	94.0	97.4	90.5	98.3	99.2
mar_Deva	60.6	80.2	86.7	82.0	92.9	96.2	77.9	92.6	96.2	76.7	92.2	95.9	84.0	95.1	97.7	90.8	98.6	99.5
min_Arab	34.5	42.8	48.7	47.9	62.6	69.7	46.0	59.7	64.3	36.2	39.2	38.4	49.6	64.3	69.9	62.9	82.2	89.9
min_Latn	65.7	86.8	92.3	76.2	89.3	93.6	77.8	92.3	96.3	76.6	92.9	96.8	80.4	93.8	97.0	85.7	97.2	98.9
mkd_Cyrl	76.2	92.7	96.6	85.2	93.9	96.8	83.8	95.8	98.2	84.7	96.5	98.9	87.8	96.2	98.2	91.7	98.5	99.4
mlt_Latn	63.0	83.9	90.2	78.3	91.7	96.0	76.5	92.4	96.2	76.4	92.4	96.9	85.9	95.8	98.4	91.4	98.4	99.3
mni_Beng	36.7	51.8	58.8	62.6	79.7	85.8	57.5	72.5	79.9	46.5	57.1	63.8	60.5	79.2	87.1	68.6	88.9	95.0
mos_Latn	43.4	57.4	69.1	54.2	69.8	77.4	56.0	75.7	82.5	55.1	76.4	85.8	55.4	76.5	83.4	57.2	80.0	89.3
mri_Latn	40.6	51.7	63.1	50.7	67.2	74.9	53.9	73.4	79.4	52.4	68.3	73.5	72.2	87.6	93.0	84.1	96.6	98.7
mya_Mymr	36.4	44.1	46.7	70.9	85.2	91.0	59.1	72.7	76.9	40.6	51.7	56.8	71.2	84.7	89.7	90.0	97.9	99.2
nld_Latn	83.0	95.8	97.9	85.0	93.9	96.3	87.5	97.2	98.7	88.3	97.4	98.8	89.4	96.3	97.9	91.6	98.4	99.3
nno_Latn	78.5	93.2	96.0	83.2	92.4	95.3	84.5	96.8	98.7	86.6	96.9	98.9	88.0	97.0	98.7	91.3	98.5	99.3
nob_Latn	80.1	94.2	96.6	83.8	93.2	96.0	85.8	96.9	98.3	87.9	97.2	98.9	88.9	96.7	98.4	91.5	98.5	99.3
npi_Deva	66.1	87.0	91.5	83.5	93.6	97.3	80.8	95.0	97.9	78.4	92.7	96.6	84.3	96.3	98.1	91.2	98.3	99.2
nqo_Nkoo	25.2	28.0	28.7	30.2	33.3	37.9	27.1	30.2	32.6	27.5	30.1	31.7	24.9	30.3	34.1	27.3	33.6	38.2
nso_Latn	45.8	60.7	69.9	50.7	69.8	77.9	55.6	76.3	83.6	55.0	77.5	86.2	61.3	79.1	83.9	79.5	94.9	98.2
nus_Latn	35.8	46.5	54.0	44.0	55.1	62.9	47.9	63.3	69.6	47.0	61.9	69.2	48.7	61.4	66.7	49.4	66.2	75.5
nya_Latn	46.5	61.9	72.5	56.2	75.2	82.8	57.0	77.2	84.9	60.6	83.1	90.3	64.1	79.1	83.8	85.5	96.9	99.1
oci_Latn	79.4	94.0	97.4	83.0	93.8	96.9	86.8	97.2	98.3	87.6	97.6	99.1	89.0	96.9	98.6	91.3	98.4	99.3
ory_Orya	32.0	38.2	40.5	78.5	90.9	95.4	79.4	93.0	97.1	72.5	85.8	91.3	81.4	94.1	97.2	88.6	97.6	99.1
pag_Latn	64.2	85.0	90.8	77.8	92.3	95.8	80.5	94.7	98.1	81.5	96.4	99.0	81.4	94.6	97.5	85.1	96.9	99.1
pan_Guru	47.3	60.9	64.7	80.5	92.7	96.6	80.8	95.0	97.6	76.5	90.7	95.3	83.1	95.0	97.3	91.1	98.4	99.5
pap_Latn	70.7	89.4	94.4	78.6	92.1	95.9	81.9	94.5	97.6	82.2	96.1	98.9	81.7	94.8	97.7	87.7	97.4	99.3
pbt_Arab	57.2	78.0	85.7	76.2	89.6	94.3	70.1	87.7	92.2	66.9	82.8	88.6	77.0	91.6	95.4	90.3	98.0	99.1
pes_Arab	77.8	94.5	97.3	86.0	95.3	97.6	85.8	97.0	98.5	84.7	95.7	98.1	86.9	96.2	98.3	91.9	98.6	99.4
plt_Latn	47.0	61.7	72.3	55.0	71.8	78.9	58.4	78.1	83.8	59.1	78.3	86.6	60.6	78.4	84.8	88.7	98.2	99.2
pol_Latn	83.8	95.4	97.4	86.0	94.6	96.6	87.9	96.9	98.3	89.0	97.7	99.1	90.3	96.4	97.8	91.4	98.3	99.2
por_Latn	84.2	96.1	98.0	86.9	94.9	97.3	88.2	97.4	98.6	89.6	97.6	99.3	89.9	96.4	98.0	92.2	98.3	99.3
prs_Arab	76.3	92.4	96.1	85.5	94.8	97.3	83.6	95.6	97.6	83.6	95.3	97.7	86.8	94.9	97.1	91.6	98.1	99.1
quy_Latn	48.7	65.5	76.2	58.6	76.1	83.6	60.7	80.5	87.6	59.4	79.2	86.5	63.6	82.5	88.6	67.7	87.4	92.9
ron_Latn	79.8	95.3	97.8	86.6	94.0	96.9	86.0	97.1	98.6	88.5	98.1	99.0	88.2	96.3	98.2	91.6	98.3	99.4
run_Latn	39.9	53.0	63.5	49.3	65.9	73.6	51.6	69.3	77.9	51.6	71.0	80.5	55.2	73.8	79.4	85.7	97.2	99.4
rus_Cyrl	84.8	96.1	98.3	86.2	94.6	96.4	87.7	97.7	98.8	90.2	98.2	99.4	90.2	97.1	98.4	91.8	98.5	99.4
sag_Latn	45.1	60.4	70.5	56.1	72.4	79.4	57.5	79.4	85.5	57.2	77.3	85.5	65.3	83.2	88.4	61.1	83.6	90.6
san_Deva	53.5	72.1	79.2	75.0	87.0	91.8	75.1	90.1	94.0	74.5	88.3	92.4	81.5	94.0	96.7	86.5	97.5	99.2
sat_Olck	25.5	26.1	28.1	29.3	29.8	33.0	26.0	27.0	27.9	31.9	27.4	29.3	29.1	28.7	31.6	25.5	29.1	29.9
scn_Latn	69.0	88.9	94.0	79.3	92.6	96.8	82.1	96.1	98.3	80.9	95.1	98.5	85.6	95.8	98.5	89.2	98.0	99.5
shn_Mymr	40.5	55.1	60.8	52.6	70.5	79.2	53.7	73.2	81.8	51.9	75.8	84.0	50.6	72.1	82.7	52.7	76.8	85.3
sin_Sinh	28.3	33.5	35.0	64.1	78.4	85.1	61.8	75.5	80.5	33.8	35.8	38.3	68.2	82.7	87.3	89.2	97.8	99.3
slk_Latn	79.9	94.1	97.4	84.0	93.2	95.9	86.6	97.1	98.5	87.2	97.5	99.1	88.4	95.5	97.4	91.1	98.3	99.2
slv_Latn	75.9	93.0	97.2	83.1	93.0	96.2	83.7	96.6	98.5	85.9	97.1	99.2	86.2	95.7	98.1	91.1	98.4	99.3
smo_Latn	46.6	61.3	71.3	55.7	73.0	82.0	58.3	77.7	84.5	57.7	77.2	86.9	72.1	88.2	92.5	86.7	97.1	99.3
sna_Latn	44.5	59.0	69.6	51.0	69.0	77.1	54.7	73.2	80.7	54.1	76.3	85.1	60.6	76.5	81.4	88.2	97.5	99.2
snd_Arab	45.2	62.9	71.4	71.1	87.2	94.2	70.5	87.3	92.2	69.5	82.9	88.5	81.7	93.7	96.4	89.7	98.4	99.3
som_Latn	42.9	54.2	66.3	50.3	66.5	74.1	54.0	72.6	81.0	51.5	68.4	78.4	65.4	82.0	87.6	84.7	97.0	99.2
sot_Latn	45.5	59.3	69.5	52.6	70.1	77.9	56.0	76.5	83.7	57.2	79.2	87.3	62.6	79.5	84.2	83.4	96.2	98.9
spa_Latn	85.4	96.5	98.3	86.0	94.3	96.5	89.2	97.5	98.8	90.1	97.8	99.1	89.5	96.5	97.7	92.0	98.5	99.2
srd_Latn	70.5	89.9	95.2	79.5	91.4	95.5	81.3	94.7	97.3	81.9	96.3	98.8	84.8	95.7	98.5	86.7	97.2	99.0
srp_Cyrl	77.3	93.2	97.2	86.1	94.4	97.4	84.8	96.2	98.0	85.6	96.5	99.0	88.2	96.5	98.2	91.7	98.3	99.4
ssw_Latn	41.1	53.1	62.6	47.2	67.2	74.1	52.8	72.4	80.4	53.3	74.6	83.4	68.8	85.7	91.3	80.9	96.1	98.7
sun_Latn	68.0	87.8	93.5	79.6	91.3	95.3	80.6	94.3	97.2	80.1	94.7	98.3	82.6	95.2	98.0	90.3	98.6	99.3
swe_Latn	80.0	94.4	97.2	84.2	93.8	96.2	86.2	96.7	99.0	87.9	97.3	98.7	89.5	96.7	98.3	91.6	98.2	99.1
swh_Latn	46.1	60.0	70.5	68.3	84.4	90.4	62.6	82.5	89.3	66.8	86.2	94.0	75.7	91.5	95.6	90.8	98.5	99.2
szl_Latn	72.6	90.0	94.7	80.0	91.8	95.4	81.8	95.4	97.4	81.9	95.8	98.5	84.4	95.5	98.5	89.2	98.1	99.3
tam_Taml	35.0	45.1	50.1	77.2	88.6	93.1	72.8	88.5	92.1	67.8	81.9	86.7	80.9	92.1	95.3	89.7	98.4	99.4
taq_Latn	47.1	62.5	72.9	58.2	72.3	80.1	59.6	77.2	84.8	58.6	76.7	85.7	60.6	79.4	85.0	64.1	83.6	91.0
taq_Tfng	25.9	32.4	34.5	28.6	34.0	37.0	33.7	42.0	47.7	28.3	36.2	42.6	27.8	35.6	39.7	31.9	42.3	49.5
tat_Cyrl	58.4	76.3	84.9	79.4	90.7	95.9	73.6	90.2	94.4	71.4	88.5	93.5	78.9	93.5	96.5	90.5	98.5	99.6
tel_Telu	43.3	54.4	59.1	80.2	92.3	95.8	79.4	94.1	96.5	74.5	91.5	96.3	84.1	95.5	97.7	90.5	98.0	99.4
tgk_Cyrl	49.3	68.5	76.4	72.0	88.1	94.0	67.9	86.6	92.3	62.8	78.2	84.1	75.2	90.7	95.1	90.2	98.1	99.5
tgl_Latn	65.3	84.0	90.6	82.8	94.3	97.2	83.3	95.2	97.7	86.8	96.6	98.7	86.0	96.3	98.3	91.1	98.2	99.2
tha_Thai	81.9	94.7	96.9	87.6	95.8	98.1	88.0	97.3	98.4	83.5	95.2	97.8	90.9	97.3	98.6	92.0	98.1	99.2
tir_Ethi	24.2	27.7	30.5	34.8	43.0	50.0	44.9	59.2	66.3	26.5	31.2	31.7	64.0	81.1	86.4	73.0	90.1	94.5
tpi_Latn	67.6	83.9	90.3	78.4	90.8	94.7	79.8	95.0	98.1	81.7	95.4	99.1	86.1	96.6	99.1	88.6	98.1	99.4
tsn_Latn	43.0	58.7	69.1	49.0	67.9	75.4	53.9	73.9	80.3	55.2	76.8	85.0	63.1	80.7	85.8	78.0	94.2	97.7
tso_Latn	42.6	54.7	65.4	49.1	67.6	76.2	53.3	72.7	79.8	53.8	74.5	83.3	68.9	84.8	89.7	78.4	92.8	97.6
tuk_Latn	54.7	72.9	82.7	73.2	88.0	94.6	70.1	88.9	94.8	71.2	87.8	93.9	73.4	89.9	94.1	89.1	98.3	99.5
tum_Latn	44.1	58.9	69.1	55.1	73.8	81.0	57.0	76.6	83.3	58.6	78.8	86.9	62.7	78.8	83.9	80.7	95.0	98.5
tur_Latn	76.4	92.5	95.6	85.0	94.9	97.2	85.1	96.6	98.4	87.5	97.2	98.8	88.2	96.0	97.6	91.5	98.6	99.4
twi_Latn	48.7	65.5	75.6	57.8	73.1	79.7	61.1	79.2	84.9	62.0	81.2	88.9	64.1	79.8	84.8	80.8	94.0	97.7
tzm_Tfng	25.9	30.1	31.5	24.8	27.2	28.8	30.4	34.4	37.2	24.9	29.8	32.4	27.9	33.7	35.7	29.0	35.2	39.4
uig_Arab	40.3	51.9	55.8	70.7	88.4	93.4	61.6	81.4	88.2	72.6	88.6	94.5	73.1	87.0	91.7	89.5	98.6	99.5
ukr_Cyrl	80.3	94.5	97.3	85.9	94.2	97.1	86.4	97.4	98.4	87.7	96.7	98.6	89.6	96.0	98.2	91.5	98.2	99.4
umb_Latn	41.5	55.0	65.7	47.4	66.7	74.6	53.4	73.5	79.9	54.3	75.1	83.2	55.4	75.1	81.0	58.6	81.6	89.7
urd_Arab	65.9	82.2	89.0	83.3	93.4	97.5	81.5	94.3	97.1	79.7	92.3	96.4	85.6	95.1	96.9	91.8	98.3	99.3
uzn_Latn	55.7	72.7	81.0	75.7	90.1	94.4	73.5	89.7	94.2	70.2	89.4	95.6	82.4	94.4	97.3	91.2	98.7	99.5
vec_Latn	75.3	92.6	96.4	81.8	92.9	96.3	85.9	96.5	98.5	87.1	96.9	98.6	87.7	96.4	98.7	89.8	97.9	99.2
vie_Latn	83.0	95.9	97.9	86.3	95.3	97.6	88.2	97.1	99.0	87.8	97.2	99.1	90.0	96.5	98.3	92.6	98.5	99.3
war_Latn	68.7	87.1	93.1	80.5	92.9	96.7	81.6	94.9	97.9	83.5	96.4	98.8	86.5	96.0	98.5	89.9	98.2	99.0
wol_Latn	50.4	66.5	77.5	59.1	77.1	84.3	63.3	84.5	90.4	64.0	84.9	92.7	65.5	85.0	91.1	69.5	89.3	94.9
xho_Latn	45.8	60.6	70.3	54.0	71.2	79.1	57.2	77.4	84.0	60.5	80.6	88.3	72.7	87.6	92.3	87.5	98.0	99.2
ydd_Hebr	32.7	44.4	51.8	59.9	77.4	84.9	53.6	73.4	82.1	33.9	39.9	41.5	66.1	84.8	90.1	87.4	97.9	99.1
yor_Latn	38.7	51.4	61.3	48.2	62.6	70.1	53.9	70.1	78.6	53.6	72.2	81.1	61.4	78.2	82.4	83.0	95.6	97.9
yue_Hant	85.1	96.6	98.0	87.0	95.1	97.3	88.9	97.3	98.8	90.0	98.0	98.5	89.8	96.1	97.1	91.7	98.5	99.3
zho_Hans	86.5	97.1	98.5	87.9	95.1	97.4	88.7	97.3	98.8	90.9	97.7	98.3	90.3	96.3	98.0	91.9	98.3	99.1
zho_Hant	86.3	96.8	98.1	87.7	94.9	97.4	88.9	97.3	99.0	90.7	98.1	98.9	90.4	96.0	97.5	92.2	98.3	99.2
zsm_Latn	79.8	95.3	97.4	85.1	94.9	97.4	85.0	96.2	97.9	86.7	97.4	98.8	87.8	96.5	98.0	91.8	98.0	99.3
zul_Latn	40.3	52.8	63.4	49.1	66.8	74.9	52.6	72.1	78.5	54.0	76.2	84.2	71.1	86.0	91.2	88.6	97.6	99.1