Table 1: COCO CN and Flickr30k datasets evaluation results

Dataset	NLLB-CLIP base			NLLB-CLIP large			AltCLIP		
	R@1	R@5	R@10	R@1	R@5	R@10	R@1	R@5	R@10
COCO CN	45.21	76.77	87.51	53.05	82.67	92.64	**63.9**	**87.2**	**93.9**
Flickr30k CN	38.36	66.72	76.0	50.92	77.18	85.28	**69.8**	**89.9**	**94.7**
Flickr30k	54.6	81.2	86.9	66.9	89.4	93.5	**72.5**	**91.6**	**95.4**

Table 2: COCO Italian dataset evaluation results

Dataset	NLLB-CLIP base			NLLB-CLIP large			Italian CLIP		
	@1	@5	@10	@1	@5	@10	@1	@5	@10
COCO IT	35.5	47.1	48.7	**45.4**	**56.7**	**58.1**	37.97	50.39	52.04

A Additional experiments

We performed additional experiments to evaluate the performance of NLLB-CLIP models on single-language datasets. We used the following datasets:

1. COCO-CN [14] (test) - 1,000 image-text pairs from the MS COCO 2014 dataset manually translated into Chinese.

2. MSCOCO-IT [2] (validation) - 8,485 captions for 1,697 images from the MS COCO 2014 dataset translated into Italian.

3. Flickr30k [30] (test) - 1,000 image-text pairs from Flickr30k dataset in English.

4. Flickr30k-CN [27] - test part of the Flickr30k dataset (1,000 image-text pairs) manually translated to Chinese.

Table 1 shows the results of the experiments on COCO-CN, Flickr30k-CN, and Flickr30k-EN datasets. Since this version of AltCLIP was trained in two stages and targeted only English and Chinese, it yields noticeably higher results in all metrics. Both bilingual and multilingual results of AltCLIP show that the two-stage training is highly beneficial for aligning image and text representations. We plan to explore this approach in the future.

Table 2 shows the evaluation results on the COCO-IT dataset. To perform a fair comparison, we used the Jupyter notebooks the authors provided and changed only the parts related to loading and running the model. We used Mean Reciprocal Rank (MRR) as a comparison metric because it was used in the original paper [2]. NLLB-CLIP large outperforms the original results and sets state-of-the-art on the COCO-IT dataset.

B XTD10 results

Table 3 shows the detailed results of the experiments presented in Section 5.1.

C Crossmodal-3600 results

Table 4 shows the detailed results of the experiments presented in Section 5.2. The crossmodal-3600 dataset contains multiple captions per image. We used the first caption for the tests.

D Flores-200 datasets results

Tables 5 and 6 present the results of NLLB-CLIP base and large models on XTD200 and Flickr30k-200 datasets. Datasets description is provided in Section 2.2.

Table 3: XTD10 dataset evaluation results

Language	NLLB-CLIP base			NLLB-CLIP large			M-CLIP	MURAL	AltCLIP
	R@1	R@5	R@10	R@1	R@5	R@10	R@10	R@10	R@10
en	47.2	76.9	87.6	59.3	83.6	93.1	95.0	-	**95.4**
fr	45.6	74.8	84.8	56.7	82.3	91.1	**93.1**	-	92.9
zh	42.4	73.9	85.5	52.7	79.9	90.5	94.0	89.7	**95.1**
de	46.4	74.4	84.8	56.2	82.2	90.6	**93.1**	-	-
es	44.9	75.4	85.9	54.9	81.8	91.7	93.6	92.9	**94.1**
it	44.5	75.1	85.9	54.8	81.7	91.2	93.1	91.8	**94.2**
tr	43.5	72.8	84.2	55.8	81.6	91.2	**93.0**	89.5	-
ru	41.5	70.4	81.7	51.7	77.3	87.4	90.0	87.2	**91.8**
ko	40.1	70.5	82.8	50.2	77.2	88.7	89.0	88.1	**94.4**
pl	45.3	75.4	86.2	56.2	82.9	91.9	**94.4**	91.0	-
jp	38.1	67.3	77.3	47.6	74.3	83.4	84.2	-	**91.7**

Table 4: Crossmodal-3600 dataset evaluation results

Language	NLLB-CLIP base			NLLB-CLIP large			mSigLIP	PaLI
	R@1	R@5	R@10	R@1	R@5	R@10	R@1	R@1
ar	34.61	63.44	74.5	**42.53**	70.72	81.14	37.37	32.6
bn	32.36	60.0	70.83	**40.25**	66.36	76.78	6.25	3.31
cs	29.47	57.39	70.44	37.22	65.81	77.19	**41.64**	35.24
da	41.72	70.14	80.17	**54.06**	80.5	87.89	47.01	38.48
de	41.78	71.17	80.61	**55.11**	80.22	88.14	54.83	46.5
el	31.0	56.94	68.86	**40.03**	67.67	78.03	22.78	20.92
en	30.53	58.81	69.47	35.64	64.31	74.94	**46.21**	40.63
es	36.03	64.86	75.92	45.25	73.0	82.64	**55.04**	46.55
fa	34.28	61.14	71.5	**41.44**	69.97	79.17	40.15	35.58
fi	37.33	65.11	75.28	**48.06**	75.33	84.17	37.14	21.8
fil	25.06	50.83	62.58	**32.47**	58.33	69.31	12.93	10.04
fr	40.58	69.94	80.03	52.11	78.92	86.58	**57.08**	43.47
he	40.17	68.67	78.14	**51.94**	77.42	85.72	31.76	29.39
hi	22.61	47.03	58.78	**28.28**	53.44	64.19	8.55	7.42
hr	42.06	70.69	80.94	**54.42**	81.25	88.75	37.09	29.55
hu	38.61	65.67	75.92	**49.75**	76.69	85.5	40.2	35.49
id	41.42	68.53	79.42	**50.44**	77.22	85.06	49.42	36.75
it	38.44	66.97	77.56	49.53	75.67	84.42	**56.39**	44.76
ja	38.92	67.19	77.33	**49.44**	76.14	83.89	31.32	27.24
ko	32.83	59.25	69.47	**41.69**	67.53	76.72	34.72	25.95
mi	15.19	35.03	45.06	**19.78**	40.92	51.11	0.22	0.22
nl	35.58	62.28	73.44	43.81	72.14	81.81	**49.55**	43.79
no	34.75	62.06	72.69	43.06	71.17	80.17	**46.21**	37.35
pl	34.89	62.94	73.47	45.28	72.5	81.19	**47.36**	43.72
pt	35.03	62.72	74.25	45.06	73.47	82.36	**52.34**	42.73
quz	11.5	25.17	33.89	**14.33**	31.22	40.42	2.57	1.9
ro	41.14	70.0	80.44	**53.36**	80.56	88.42	35.6	28.82
ru	39.78	68.5	78.78	48.58	77.0	85.58	**49.89**	41.11
sv	33.44	60.97	72.39	41.58	69.28	78.92	**48.18**	40.66
sw	27.06	54.0	66.0	**34.5**	62.61	72.83	7.17	3.41
te	24.25	50.44	62.67	**29.53**	56.47	67.33	0.29	1.42
th	33.53	62.11	73.06	**42.25**	70.03	79.78	23.08	16.06
tr	36.92	65.72	76.19	**47.78**	74.19	83.14	37.38	31.47
uk	38.14	66.14	76.69	**48.11**	74.58	83.31	33.21	30.81
vi	38.86	67.31	77.89	**48.33**	76.89	85.06	41.92	21.28
zh	33.89	60.69	71.92	**41.69**	67.89	77.81	32.45	28.24
Average	33.99	61.11	71.85	**42.96**	69.65	78.87	34.87	28.46

8

Table 5: XTD200 evaluation dataset results

	NLLB-CLIP base			NLLB-CLIP large		
Language	R@1	R@5	R@10	R@1	R@5	R@10
ace_Arab	20.1	42.7	54.9	27.4	50.7	62.0
ace_Latn	40.3	70.2	82.3	50.6	76.7	86.8
acm_Arab	45.3	72.9	84.6	55.4	82.2	90.6
acq_Arab	45.5	73.4	84.0	55.5	81.6	90.3
aeb_Arab	33.8	64.1	75.2	45.9	68.8	79.2
afr_Latn	46.7	75.6	86.5	56.6	81.3	91.2
ajp_Arab	43.1	72.7	83.0	54.5	80.4	89.2
aka_Latn	33.1	60.1	74.2	42.7	69.6	81.4
als_Latn	44.3	74.8	85.2	55.1	81.0	91.0
amh_Ethi	42.3	72.1	84.0	51.8	78.1	88.7
apc_Arab	42.6	72.1	83.4	54.7	79.7	88.5
arb_Arab	45.1	74.7	85.0	57.3	83.3	90.7
ars_Arab	44.8	73.9	84.7	55.0	81.5	90.4
ary_Arab	40.8	69.1	80.0	49.7	76.8	85.6
arz_Arab	43.7	72.1	83.6	56.7	81.1	89.1
asm_Beng	42.2	71.3	82.6	51.7	79.1	89.1
ast_Latn	43.4	72.8	83.5	56.4	79.2	88.6
awa_Deva	39.0	67.1	79.2	48.5	73.6	83.2
ayr_Latn	35.4	63.4	75.9	43.4	72.2	82.4
azb_Arab	38.2	68.2	80.1	48.5	74.1	85.8
azj_Latn	43.2	75.8	85.6	55.1	80.6	89.7
bak_Cyrl	39.6	73.3	83.8	51.9	78.2	88.0
bam_Latn	32.1	60.1	71.7	39.4	66.5	77.6
ban_Latn	38.9	68.5	82.2	48.3	74.7	85.4
bel_Cyrl	44.0	74.6	85.2	55.4	82.7	91.1
bem_Latn	31.5	61.6	73.1	37.0	67.8	78.5
ben_Beng	43.4	74.6	84.9	54.8	80.9	90.5
bho_Deva	43.3	74.7	85.5	54.2	80.6	91.2
bjn_Arab	16.3	36.5	49.7	22.3	46.0	56.4
bjn_Latn	43.1	72.9	83.6	53.2	80.2	89.3
bod_Tibt	30.2	58.7	70.1	36.8	63.6	76.4
bos_Latn	44.2	76.1	86.1	56.3	81.5	91.5
bug_Latn	36.3	66.5	79.3	45.4	73.4	84.5
bul_Cyrl	44.1	75.8	85.9	56.3	81.5	91.1
cat_Latn	44.2	75.8	85.5	55.7	81.7	90.8
ceb_Latn	44.5	74.5	85.5	56.2	80.5	90.1
ces_Latn	43.9	76.5	86.4	57.2	81.7	90.3
cjk_Latn	22.4	44.7	58.0	28.4	52.0	63.5
ckb_Arab	42.0	71.4	83.1	52.2	79.3	88.4
crh_Latn	43.0	74.0	85.2	54.6	79.0	89.9
cym_Latn	38.6	68.3	80.6	49.3	76.5	87.0
dan_Latn	47.5	76.4	87.0	57.8	83.6	92.1
deu_Latn	46.4	77.2	86.4	58.3	83.1	91.9
dik_Latn	21.9	46.2	59.1	28.1	51.9	62.8
dyu_Latn	23.4	49.0	63.8	29.0	57.5	68.3
dzo_Tibt	27.7	56.7	68.8	35.7	62.4	73.9
ell_Grek	46.2	75.5	86.9	56.7	83.1	92.2
epo_Latn	45.5	74.8	86.4	55.7	81.7	90.9
est_Latn	43.2	71.5	82.3	52.9	78.8	88.5
eus_Latn	41.6	72.6	83.1	51.4	78.6	87.9
ewe_Latn	34.8	62.8	75.9	42.4	71.3	83.2
fao_Latn	42.9	71.4	83.7	53.8	79.3	89.4
fij_Latn	40.7	70.3	81.4	50.0	77.3	87.7

Table 5 – continued from the previous page

| | NLLB-CLIP base | | | NLLB-CLIP large | | |
Language	R@1	R@5	R@10	R@1	R@5	R@10
fin_Latn	43.3	71.5	82.1	52.3	78.3	88.8
fon_Latn	24.9	52.3	63.7	29.1	56.4	68.0
fra_Latn	46.0	77.5	87.3	58.0	82.2	91.7
fur_Latn	41.6	70.9	82.5	51.3	77.4	87.9
fuv_Latn	27.6	54.5	66.7	35.0	61.0	71.7
gaz_Latn	36.5	64.5	77.6	43.6	71.8	82.0
gla_Latn	40.4	71.0	81.7	50.4	78.8	87.4
gle_Latn	41.1	69.6	80.9	51.5	77.8	88.1
glg_Latn	45.3	76.4	86.6	57.9	81.9	92.2
grn_Latn	40.3	68.2	79.9	46.9	75.7	86.2
guj_Gujr	45.2	75.1	85.2	55.6	81.6	91.1
hat_Latn	43.3	74.6	85.1	54.9	80.4	89.8
hau_Latn	42.8	68.7	81.3	53.5	78.0	88.2
heb_Hebr	42.8	72.5	83.6	56.1	81.8	90.8
hin_Deva	45.1	76.5	86.2	57.1	82.9	91.6
hne_Deva	43.5	75.3	85.4	54.3	80.3	90.1
hrv_Latn	45.4	76.1	85.8	57.0	81.6	91.5
hun_Latn	44.2	76.0	85.8	54.8	81.9	90.9
hye_Armn	42.0	72.6	83.7	52.9	78.9	88.9
ibo_Latn	41.1	70.0	83.0	50.3	76.8	86.4
ilo_Latn	43.0	74.0	83.8	55.6	80.0	89.9
ind_Latn	47.3	75.4	84.8	56.0	82.0	91.1
isl_Latn	41.1	71.4	83.0	52.0	79.3	89.1
ita_Latn	45.3	76.6	87.5	56.8	82.7	92.6
jav_Latn	45.0	74.1	86.0	55.3	81.0	90.1
jpn_Jpan	44.3	73.7	84.9	53.9	81.0	90.5
kab_Latn	32.4	60.6	72.3	38.8	68.1	78.3
kac_Latn	30.4	57.3	69.3	37.5	63.1	74.9
kam_Latn	11.3	30.8	40.7	14.5	32.6	43.5
kan_Knda	44.0	74.9	85.3	57.2	81.9	91.4
kas_Arab	34.2	61.8	73.9	43.8	69.9	80.9
kas_Deva	39.3	70.4	79.6	48.5	74.7	85.8
kat_Geor	39.6	70.6	81.4	50.2	76.8	86.6
kaz_Cyrl	42.7	72.9	84.1	51.8	78.9	88.4
kbp_Latn	17.5	37.5	48.5	21.0	43.9	54.5
kea_Latn	40.4	68.8	80.7	49.7	74.5	85.7
khk_Cyrl	34.8	62.0	74.8	42.7	69.2	78.8
khm_Khmr	36.2	65.0	75.7	44.1	71.1	83.0
kik_Latn	31.1	53.4	65.7	37.1	61.5	73.6
kin_Latn	38.0	67.0	79.8	46.3	73.5	84.3
kir_Cyrl	39.4	72.8	84.6	51.7	77.6	88.0
kmb_Latn	19.9	41.6	54.0	22.4	46.2	58.9
kmr_Latn	40.4	70.1	81.5	51.6	78.4	88.8
knc_Arab	35.8	64.9	76.0	47.4	72.4	82.3
knc_Latn	24.9	49.5	64.4	29.6	54.3	68.2
kon_Latn	34.9	63.8	75.4	42.6	70.2	81.7
kor_Hang	42.5	72.6	83.5	53.6	80.3	90.0
lao_Laoo	41.9	71.6	84.4	50.5	76.9	86.6
lij_Latn	38.4	68.7	80.0	48.4	75.0	85.7
lim_Latn	45.3	75.0	85.5	54.5	81.3	91.8
lin_Latn	36.6	68.5	80.9	48.7	77.8	87.8
lit_Latn	41.5	72.1	82.8	51.7	78.6	89.1
lmo_Latn	43.8	73.3	85.2	55.1	81.3	91.5
ltg_Latn	35.1	63.9	75.3	43.1	70.0	79.5

Table 5 – continued from the previous page

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
ltz_Latn	45.7	75.7	85.5	56.5	81.2	90.9
lua_Latn	31.2	59.1	72.7	37.3	68.1	79.8
lug_Latn	34.8	63.6	76.0	41.3	70.6	82.1
luo_Latn	33.0	61.0	73.5	40.7	69.2	80.9
lus_Latn	33.6	60.0	74.0	41.4	66.9	79.3
lvs_Latn	40.2	69.5	81.4	50.2	76.1	86.8
mag_Deva	45.1	76.1	86.6	56.2	82.3	91.4
mai_Deva	44.4	75.8	86.5	53.7	81.4	90.7
mal_Mlym	42.4	73.4	85.2	53.9	81.6	90.6
mar_Deva	43.9	74.5	85.1	53.7	79.9	89.9
min_Latn	40.2	70.8	82.7	50.4	77.5	86.0
mkd_Cyrl	44.8	74.3	85.8	56.7	80.5	90.2
mlt_Latn	41.8	71.7	83.7	52.7	78.3	88.5
mni_Beng	36.4	65.8	79.6	45.1	73.2	84.8
mos_Latn	23.5	47.6	59.5	29.2	54.0	63.2
mri_Latn	39.1	69.8	80.1	48.1	76.7	86.5
mya_Mymr	41.0	71.6	83.5	49.9	78.3	88.2
nld_Latn	46.3	76.5	86.9	56.5	83.0	91.9
nno_Latn	45.0	72.4	83.5	54.6	80.7	89.2
nob_Latn	46.0	75.0	86.0	56.7	82.3	91.2
npi_Deva	43.7	74.7	85.4	54.3	80.8	90.1
nso_Latn	42.2	73.3	83.0	53.8	79.2	88.5
nus_Latn	20.8	44.2	57.2	27.9	52.1	63.2
nya_Latn	42.2	71.8	83.1	50.5	77.9	88.4
oci_Latn	42.7	74.9	85.5	54.0	80.1	90.7
ory_Orya	42.4	75.1	85.7	53.2	80.2	89.7
pag_Latn	43.1	74.0	84.7	53.3	80.1	89.6
pan_Guru	43.4	74.2	85.4	56.2	81.1	89.7
pap_Latn	44.2	72.9	84.3	51.5	79.2	89.9
pbt_Arab	40.7	72.4	82.5	51.9	77.1	87.6
pes_Arab	44.8	74.0	85.0	55.4	80.7	90.1
plt_Latn	42.6	72.3	83.9	52.0	77.5	88.5
pol_Latn	44.9	75.4	85.2	56.4	82.4	91.4
por_Latn	46.7	75.9	86.6	58.3	81.7	92.0
prs_Arab	45.3	74.6	86.5	55.3	81.2	90.8
quy_Latn	28.0	54.5	65.4	34.5	61.0	71.5
ron_Latn	45.4	75.6	86.6	57.2	82.1	91.6
run_Latn	35.7	65.7	77.0	43.2	70.1	81.9
rus_Cyrl	45.4	75.6	86.3	55.7	82.3	91.9
sag_Latn	36.1	66.0	78.1	43.5	73.6	83.1
san_Deva	32.8	60.3	74.6	42.7	69.6	79.7
scn_Latn	43.8	73.5	84.7	55.2	80.3	90.3
shn_Mymr	28.0	54.3	66.4	31.9	59.2	72.3
sin_Sinh	41.3	72.5	84.0	51.8	78.1	87.5
slk_Latn	44.6	75.9	85.9	55.5	81.3	91.5
slv_Latn	45.4	75.6	85.2	56.2	80.2	90.2
smo_Latn	41.5	72.4	84.4	53.3	79.6	88.5
sna_Latn	41.1	71.7	83.1	51.3	78.8	89.3
snd_Arab	43.9	70.1	82.8	53.0	78.2	89.3
som_Latn	41.8	71.7	84.4	50.0	77.2	88.0
sot_Latn	42.6	74.7	84.2	54.8	80.1	88.7
spa_Latn	46.3	76.8	86.5	58.3	82.5	92.7
srd_Latn	41.8	67.7	80.7	50.8	74.7	86.1
srp_Cyrl	41.7	73.1	84.7	54.1	80.6	90.9

Table 5 – continued from the previous page

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
ssw_Latn	36.5	67.7	80.4	47.6	74.1	84.9
sun_Latn	45.1	74.6	86.1	56.0	81.8	89.9
swe_Latn	47.1	76.0	87.0	57.7	82.7	91.5
swh_Latn	44.0	73.0	84.0	53.5	80.3	89.3
szl_Latn	42.0	69.7	82.4	52.2	77.9	87.7
tam_Taml	42.9	72.2	83.9	53.7	79.4	89.1
taq_Latn	20.0	41.7	53.4	24.8	50.4	62.2
taq_Tfng	18.7	40.4	54.4	24.2	47.4	59.4
tat_Cyrl	41.0	71.4	84.0	51.2	78.8	89.2
tel_Telu	43.6	75.4	86.9	55.8	81.6	90.8
tgk_Cyrl	42.5	73.2	84.3	51.2	78.5	87.7
tgl_Latn	45.0	74.6	85.9	57.7	81.4	90.6
tha_Thai	42.2	71.7	81.9	50.1	78.0	87.9
tir_Ethi	39.0	68.9	81.6	47.1	75.1	86.7
tpi_Latn	40.1	70.9	82.9	47.7	79.1	88.2
tsn_Latn	41.7	71.2	83.8	53.3	77.8	88.5
tso_Latn	41.7	69.9	81.7	51.0	77.8	87.1
tuk_Latn	37.7	70.8	81.2	47.7	75.1	85.6
tum_Latn	34.8	63.6	75.2	43.2	69.8	79.9
tur_Latn	44.2	75.2	85.2	55.9	80.8	90.7
twi_Latn	35.0	62.4	74.4	41.8	69.3	79.9
tzm_Tfng	23.1	46.6	57.8	27.4	50.4	62.2
uig_Arab	38.6	68.3	80.1	46.5	73.9	83.9
ukr_Cyrl	42.9	74.6	86.2	54.9	83.3	92.2
umb_Latn	21.4	43.1	55.4	25.2	47.4	59.1
urd_Arab	42.8	74.3	85.6	56.1	81.4	90.6
uzn_Latn	42.1	72.0	83.2	52.0	77.8	88.6
vec_Latn	45.9	75.1	86.6	57.1	81.9	91.5
vie_Latn	45.1	74.9	85.3	56.1	80.3	90.2
war_Latn	40.5	69.8	81.7	52.0	76.0	86.0
wol_Latn	27.1	54.4	67.8	36.7	62.1	73.7
xho_Latn	41.7	72.1	83.3	52.2	79.0	87.0
ydd_Hebr	42.8	72.6	84.6	52.8	79.9	90.7
yor_Latn	38.3	68.4	80.7	48.6	72.9	86.1
yue_Hant	41.4	71.1	83.1	51.1	78.0	87.8
zho_Hans	42.0	72.3	84.0	50.6	79.6	89.3
zho_Hant	37.4	66.5	79.8	47.2	73.0	84.9
zsm_Latn	44.9	75.7	84.7	55.1	82.2	90.4
zul_Latn	42.1	72.1	83.7	52.4	80.3	89.4
Average	39.4	68.6	80.0	49.1	75.2	85.3

Table 6: Flickr30k-200 evaluation dataset results

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
ace_Arab	20.1	41.3	50.3	28.0	51.3	58.8
ace_Latn	42.4	68.3	78.6	52.3	77.5	86.0
acm_Arab	48.4	74.8	82.0	61.2	83.9	88.9
acq_Arab	50.0	76.7	84.0	62.7	85.7	91.3
aeb_Arab	38.7	63.3	72.7	47.6	71.5	78.9
afr_Latn	51.0	78.2	86.8	64.2	86.8	92.8

Table 6 – continued from the previous page

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
ajp_Arab	49.2	76.1	84.1	61.4	83.9	90.0
aka_Latn	39.2	66.6	76.1	49.1	76.3	84.1
als_Latn	50.2	76.3	84.2	60.2	86.7	91.5
amh_Ethi	47.6	73.7	80.8	58.6	83.8	89.2
apc_Arab	49.4	75.6	83.5	61.6	84.9	89.6
arb_Arab	51.2	78.0	84.7	63.7	87.2	91.6
ars_Arab	49.8	76.7	84.9	62.6	85.5	91.9
ary_Arab	45.1	70.4	79.1	57.6	78.6	85.6
arz_Arab	49.7	76.7	83.9	63.2	86.1	91.6
asm_Beng	45.5	74.0	82.7	58.8	82.9	88.6
ast_Latn	50.6	75.7	83.5	63.2	85.6	90.4
awa_Deva	38.4	64.0	71.8	49.8	70.8	78.1
ayr_Latn	39.4	67.7	76.6	51.1	75.1	83.1
azb_Arab	43.3	70.0	79.2	53.4	78.9	86.3
azj_Latn	51.3	77.5	84.3	62.9	87.9	92.2
bak_Cyrl	47.6	73.8	82.3	58.4	83.5	90.3
bam_Latn	33.3	62.1	72.4	43.5	71.1	80.6
ban_Latn	40.6	69.1	78.5	51.5	77.8	85.7
bel_Cyrl	49.9	77.8	84.8	61.4	85.3	90.4
bem_Latn	37.6	64.2	73.2	48.3	72.2	81.9
ben_Beng	51.1	78.9	84.9	63.2	87.8	91.0
bho_Deva	49.4	76.9	83.4	61.6	84.5	90.0
bjn_Arab	18.7	37.7	47.0	22.2	43.0	52.4
bjn_Latn	47.6	74.9	84.4	60.0	84.5	90.0
bod_Tibt	35.6	63.2	74.0	45.8	71.1	79.6
bos_Latn	51.5	77.4	85.0	63.8	86.3	91.9
bug_Latn	40.1	66.9	76.4	49.8	74.5	83.0
bul_Cyrl	51.4	77.1	85.2	63.8	86.6	92.5
cat_Latn	50.5	78.7	84.5	64.3	86.9	92.7
ceb_Latn	51.4	77.9	85.3	63.1	85.5	91.2
ces_Latn	50.4	78.1	85.5	64.7	87.2	91.3
cjk_Latn	26.3	50.4	60.5	31.4	58.1	69.0
ckb_Arab	47.8	73.9	83.0	58.3	81.7	88.7
crh_Latn	49.0	75.4	83.6	62.1	85.6	91.3
cym_Latn	46.2	72.4	81.5	57.3	82.7	90.4
dan_Latn	53.5	79.1	85.9	63.9	87.7	92.9
deu_Latn	53.0	80.1	86.5	63.0	88.9	92.7
dik_Latn	26.7	52.4	62.1	33.4	57.9	67.4
dyu_Latn	20.2	41.6	52.8	25.8	48.5	59.1
dzo_Tibt	31.9	59.4	68.3	41.1	68.6	78.0
ell_Grek	51.4	78.8	84.3	63.6	87.9	92.2
epo_Latn	52.3	77.4	85.1	64.7	86.8	92.5
est_Latn	48.2	74.8	83.3	59.9	84.9	89.6
eus_Latn	46.1	75.2	81.6	58.1	82.1	89.8
ewe_Latn	40.0	68.1	77.0	48.4	75.6	84.4
fao_Latn	48.1	73.7	82.4	58.5	83.8	89.6
fij_Latn	41.9	69.9	78.4	55.0	80.4	86.9
fin_Latn	49.1	76.2	83.0	61.3	84.8	90.8
fon_Latn	29.7	53.3	63.0	37.7	63.3	73.3
fra_Latn	53.2	77.3	87.2	63.9	88.6	93.8
fur_Latn	46.3	72.9	81.0	59.5	83.1	90.2
fuv_Latn	31.1	58.7	68.2	38.3	63.6	74.5
gaz_Latn	37.3	65.1	74.8	50.7	77.0	84.6
gla_Latn	47.3	73.3	81.6	57.9	83.3	89.6

Table 6 – continued from the previous page

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
gle_Latn	44.3	73.4	81.3	58.7	83.9	89.9
glg_Latn	51.1	79.0	85.9	64.6	88.2	92.6
grn_Latn	44.7	69.3	79.0	54.9	79.3	85.8
guj_Gujr	51.8	76.5	84.7	64.1	86.4	91.8
hat_Latn	51.1	76.3	84.7	61.6	85.9	91.1
hau_Latn	46.8	72.1	81.2	58.0	82.3	88.7
heb_Hebr	47.6	75.6	84.9	62.3	85.5	91.9
hin_Deva	53.1	79.1	85.7	64.3	87.3	91.7
hne_Deva	51.3	78.3	83.8	63.7	85.6	90.8
hrv_Latn	50.4	77.5	85.0	64.3	86.0	91.3
hun_Latn	50.0	76.8	86.2	63.2	86.6	91.7
hye_Armn	47.8	74.9	83.3	59.0	83.1	89.2
ibo_Latn	44.0	73.9	82.3	55.3	82.9	88.7
ilo_Latn	48.7	73.7	83.5	59.1	83.8	88.5
ind_Latn	51.4	79.9	86.4	65.6	87.5	91.9
isl_Latn	46.8	74.5	83.1	58.2	83.8	90.1
ita_Latn	50.7	78.8	85.9	65.6	87.6	92.2
jav_Latn	51.1	77.6	86.2	63.0	85.6	91.2
jpn_Jpan	48.5	73.7	82.9	60.2	80.9	89.0
kab_Latn	35.4	62.2	72.5	44.6	72.2	81.7
kac_Latn	33.6	62.5	71.7	41.8	69.6	77.3
kam_Latn	10.5	24.2	32.9	12.0	28.4	37.4
kan_Knda	50.4	77.8	84.1	62.2	86.3	91.5
kas_Arab	38.4	64.4	71.6	48.4	75.3	81.6
kas_Deva	42.2	67.5	76.5	52.4	76.8	84.1
kat_Geor	43.1	69.9	78.3	54.6	78.5	85.5
kaz_Cyrl	48.3	74.4	83.1	58.4	84.0	89.0
kbp_Latn	25.5	48.1	58.2	29.8	55.1	64.6
kea_Latn	44.7	71.4	79.5	55.6	81.5	88.7
khk_Cyrl	37.6	65.1	76.6	48.5	77.1	84.8
khm_Khmr	39.9	66.7	77.2	51.6	76.3	84.8
kik_Latn	38.2	66.1	74.8	47.1	73.5	82.0
kin_Latn	44.3	72.5	80.9	56.2	81.1	88.7
kir_Cyrl	45.6	71.7	80.5	56.9	82.0	88.0
kmb_Latn	22.0	44.4	55.6	26.8	51.8	61.0
kmr_Latn	40.9	70.1	79.7	52.5	79.4	86.5
knc_Arab	37.4	64.7	74.9	49.7	74.6	82.7
knc_Latn	23.2	47.6	57.2	30.6	52.7	63.5
kon_Latn	43.4	69.3	79.4	52.9	79.0	86.3
kor_Hang	51.0	75.8	84.6	61.8	85.5	91.0
lao_Laoo	45.9	72.9	82.1	56.8	83.1	89.4
lij_Latn	45.7	71.5	80.7	57.8	81.9	87.6
lim_Latn	48.3	76.6	84.5	61.1	85.8	91.3
lin_Latn	45.6	73.4	82.1	58.2	81.7	89.7
lit_Latn	48.6	74.7	82.5	58.4	84.1	90.5
lmo_Latn	47.9	75.1	82.6	60.1	84.2	90.4
ltg_Latn	38.7	65.7	74.9	48.9	76.0	84.3
ltz_Latn	51.0	77.2	84.5	63.3	86.5	91.8
lua_Latn	37.2	66.4	76.1	46.9	76.2	84.2
lug_Latn	39.8	66.6	78.0	52.0	78.1	86.2
luo_Latn	39.5	66.8	75.4	50.7	77.4	85.0
lus_Latn	37.4	62.7	73.1	45.3	69.7	78.7
lvs_Latn	44.8	71.9	80.5	57.3	81.5	88.2
mag_Deva	52.1	78.4	85.0	63.0	85.8	91.8

14

	NLLB-CLIP base			NLLB-CLIP large		
Language	R@1	R@5	R@10	R@1	R@5	R@10
mai_Deva	51.1	76.4	85.0	62.3	86.7	91.5
mal_Mlym	49.2	75.7	83.3	59.7	85.0	90.2
mar_Deva	50.9	77.6	84.8	62.5	85.5	91.0
min_Latn	47.4	74.0	82.7	58.0	83.3	88.2
mkd_Cyrl	51.4	76.4	83.4	61.5	85.9	91.6
mlt_Latn	46.7	72.7	81.5	59.2	84.5	90.1
mni_Beng	38.6	63.8	72.3	49.0	74.3	81.6
mos_Latn	30.3	55.2	67.1	38.6	64.5	73.2
mri_Latn	42.6	72.2	81.2	55.4	82.2	88.3
mya_Mymr	45.6	70.3	79.3	55.5	80.1	84.6
nld_Latn	51.0	77.7	85.7	64.1	88.1	92.6
nno_Latn	51.4	77.1	84.8	61.2	85.2	90.8
nob_Latn	52.2	78.6	85.8	63.6	87.8	92.3
npi_Deva	49.0	74.8	82.0	60.2	83.2	88.6
nso_Latn	48.8	76.2	84.3	60.1	84.2	90.0
nus_Latn	26.0	48.3	58.6	32.9	56.8	67.8
nya_Latn	48.3	73.5	83.1	60.0	83.4	89.5
oci_Latn	50.8	76.7	84.2	64.1	86.4	91.7
ory_Orya	47.4	73.9	83.0	59.6	84.1	90.2
pag_Latn	48.0	74.8	82.9	57.7	83.3	89.2
pan_Guru	51.2	77.3	84.3	63.3	86.2	91.2
pap_Latn	49.9	75.7	83.9	61.0	85.9	90.7
pbt_Arab	49.3	73.7	81.8	61.4	83.7	89.7
pes_Arab	50.8	77.4	84.5	63.7	86.3	91.5
plt_Latn	48.3	73.8	83.7	60.6	84.6	90.4
pol_Latn	49.9	76.6	84.4	63.0	86.4	91.8
por_Latn	52.7	79.1	86.6	65.4	88.8	92.5
prs_Arab	50.5	77.0	84.4	65.0	87.0	92.1
quy_Latn	31.9	58.6	68.2	40.0	66.3	74.7
ron_Latn	51.2	77.1	84.7	64.5	87.1	92.2
run_Latn	43.0	71.6	80.3	56.3	80.6	87.5
rus_Cyrl	51.0	77.9	84.5	62.6	87.0	92.3
sag_Latn	41.0	68.7	76.7	52.1	77.2	86.1
san_Deva	38.4	63.2	73.2	48.2	73.1	80.9
scn_Latn	49.6	75.9	84.0	62.7	85.9	91.7
shn_Mymr	32.9	60.3	72.3	43.5	70.4	78.8
sin_Sinh	47.0	72.2	80.1	58.2	81.8	86.9
slk_Latn	49.2	77.2	85.6	62.4	87.1	91.7
slv_Latn	48.7	76.5	84.0	61.7	85.8	91.2
smo_Latn	48.1	73.3	82.3	58.9	83.7	90.4
sna_Latn	47.4	73.3	82.8	60.4	84.8	91.7
snd_Arab	47.0	73.9	82.2	60.4	84.0	89.5
som_Latn	46.8	75.5	83.1	60.2	84.5	88.8
sot_Latn	48.8	76.5	83.8	61.7	85.4	91.7
spa_Latn	51.8	79.4	85.7	65.4	88.7	93.4
srd_Latn	47.5	73.6	82.6	59.3	83.4	90.7
srp_Cyrl	50.4	75.6	83.8	61.7	84.9	90.7
ssw_Latn	41.9	69.1	79.8	53.6	80.1	87.2
sun_Latn	50.3	77.3	85.2	63.1	86.7	92.3
swe_Latn	52.9	78.2	85.5	62.8	86.7	91.9
swh_Latn	51.4	76.0	84.5	61.6	85.6	90.8
szl_Latn	45.4	72.3	81.6	58.2	82.9	89.9
tam_Taml	49.6	75.3	83.1	59.8	84.4	89.8
taq_Latn	20.7	40.7	49.9	24.7	46.2	55.3

Table 6 – continued from the previous page

Language	NLLB-CLIP base			NLLB-CLIP large		
	R@1	R@5	R@10	R@1	R@5	R@10
taq_Tfng	14.9	35.0	44.8	21.0	41.9	53.5
tat_Cyrl	49.2	75.3	83.2	60.0	84.4	90.6
tel_Telu	48.9	77.3	85.3	61.9	86.5	90.7
tgk_Cyrl	46.0	74.3	82.6	58.4	82.3	88.0
tgl_Latn	51.4	77.2	86.3	63.5	86.9	92.4
tha_Thai	48.0	75.3	83.8	60.0	83.7	89.9
tir_Ethi	43.2	71.6	79.2	55.6	80.6	87.2
tpi_Latn	46.6	72.3	81.0	57.2	81.5	88.1
tsn_Latn	49.9	75.5	84.0	61.5	85.3	90.7
tso_Latn	47.0	74.7	83.7	59.3	83.8	90.9
tuk_Latn	46.0	71.1	81.3	55.0	81.8	87.9
tum_Latn	39.5	67.7	76.9	47.8	75.0	83.5
tur_Latn	52.6	78.3	85.6	64.8	87.1	91.7
twi_Latn	38.8	67.2	76.5	48.9	75.4	83.8
tzm_Tfng	23.2	46.4	56.6	30.2	55.1	64.5
uig_Arab	43.0	69.7	78.7	54.2	79.9	86.1
ukr_Cyrl	52.5	77.6	84.9	62.8	86.0	91.5
umb_Latn	19.7	41.4	53.4	24.2	48.3	60.2
urd_Arab	51.5	76.1	84.9	63.5	85.9	92.1
uzn_Latn	47.4	76.1	83.2	58.5	84.6	89.9
vec_Latn	49.9	77.9	85.3	64.4	86.1	91.8
vie_Latn	50.5	77.7	85.9	62.2	86.2	92.0
war_Latn	48.7	74.3	82.1	58.9	82.4	87.6
wol_Latn	30.0	57.5	67.9	40.4	68.4	77.7
xho_Latn	48.5	74.1	83.9	61.8	84.9	90.4
ydd_Hebr	47.0	74.7	83.8	60.2	84.0	90.5
yor_Latn	42.7	70.7	79.6	54.1	80.5	87.7
yue_Hant	48.0	74.5	83.5	57.0	81.2	87.4
zho_Hans	47.9	74.5	81.4	58.0	83.0	88.3
zho_Hant	30.6	54.4	63.1	37.6	60.7	69.1
zsm_Latn	51.6	78.2	86.2	63.8	87.5	91.2
zul_Latn	47.9	75.2	83.9	61.3	85.3	91.4
Average	44.49	70.81	79.32	55.56	79.86	86.37