PROJECT DESCRIPTION | MACHINE TRANSLATION

30x more efficient

Diffusion Language Model for Translation

English > es | English > fr | English > ru | English > zh

A **27B-parameter diffusion language model** for translation with a **canvas size of 1024 tokens**. It denoises the whole canvas in **30-40 steps on average**, which works out to about **30 tokens per forward pass**, roughly 30x fewer sequential passes than decoding one token at a time. It translates from English into Spanish, French, Russian and Chinese. On the Cocodoc test set it reaches **69.4 chrF++** overall, ahead of both the Microsoft Translator API (66.7) and GPT-5-mini (68.5).

27B
parameters, diffusion LM

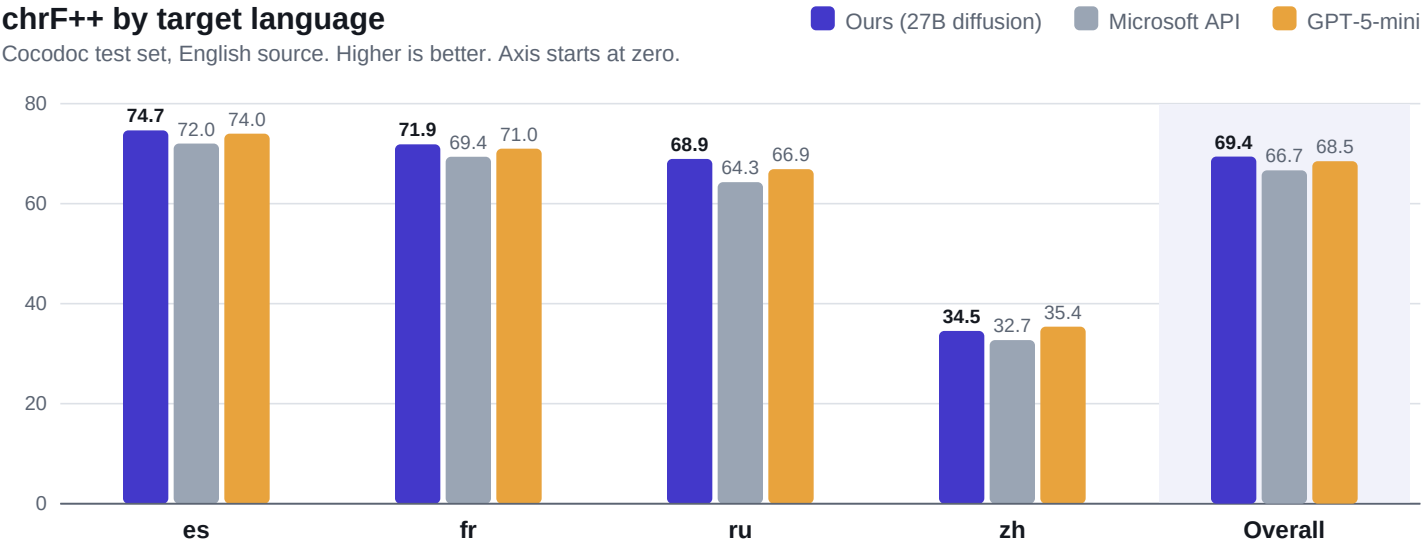
1024
tokens, canvas size

30-40
avg. denoising steps

~30
tokens per forward pass

chrF++ by target language

Cocodoc test set, English source. Higher is better. Axis starts at zero.



Scores and deltas

Delta = ours minus baseline, in chrF++ points.

	Ours	Microsoft API	GPT-5-mini	vs Microsoft	vs GPT-5-mini
en > es	**74.65**	72.02	74.00	+2.63	+0.65
en > fr	**71.87**	69.38	71.00	+2.48	+0.87
en > ru	**68.94**	64.30	66.90	+4.64	+2.04
en > zh	**34.55**	32.72	35.40	+1.83	-0.85
Overall	**69.42**	**66.66**	**68.49**	**+2.76**	**+0.93**

Generation length / token efficiency

Table to be added: gen_len and tokens per forward pass, by target language.

Metric: chrF++. Test set: Cocodoc. Overall = score over all four target languages. GPT-5-mini scores as reported (rounded).