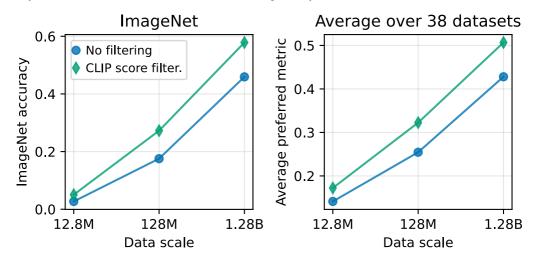a) Filtering uncurated datasets at various scales can improve the zero-shot accuracy of pre-trained models

ImageNet

Average over 38 datasets

No filtering
CLIP score filter.

b) Incorporating curated data from external sources can further enhance model performance

LAION-A +2.9

5 ext. datasets +4.2

ImageNet accuracy gain (percentage points) from CLIP score filtering (1.28B scale)