a) OLMo-7B + CLIP • Layer 4

Top NNs

LN-Lens

1. "bus license plate says fck **884** " from LLM L4
2. "a bus license plate reading pwl **413** " from LLM L4

...

Embedding Matrix		LogitLens	
1. ="	(0.09)	1. imator	(5.80)
2. _C	(0.08)	2. france	(5.68)
3. ");	(0.08)	3. Replies	(5.47)
...		...	

b) Qwen2-7B + SigLIP • Layer 16

Top NNs

LN-Lens

1. "batter holding bat in blue **jerz** ee" from LLM L16
2. "girl wearing a **ts** hirt" from LLM L16
3. "part of a white **swimming** board" from LLM L16

...

Embedding Matrix		LogitLens	
1. _HELPER	(0.11)	1. *)(	(14.91)
2. □	(0.10)	2. □□	(13.02)
3. izzato	(0.10)	3. _unregi...	(12.91)
...		...	