

a) LLaMA3-8B + DINOv2 • Layer 4



Top NNs

| LN-Lens | |
|---|---------------|
| 1. "large elephant facing forward " | from LLM L8 |
| 2. "slow motion shot of a married couple who runs in the park , t | from LLM L8 d |
| 3. "the woman is leaping into t..." | from LLM L8 |
| ... | |

| Embedding Matrix | | LogitLens | |
|------------------|--------|-------------|--------|
| 1. Luu | (0.07) | 1. UnitTest | (9.60) |
| 2. [] | (0.07) | 2. hausen | (9.40) |
| 3. Binding... | (0.07) | 3. unken | (9.23) |
| ... | | ... | |

b) OLMo-7B + DINOv2 • Layer 24



Top NNs

| LN-Lens | |
|--|--------------|
| 1. "black and white clock on adve..." | from LLM L8 |
| 2. "large stone tower with gold clocks " | from LLM L16 |
| 3. "a tower with small clocks " | from LLM L8 |
| ... | |

| Embedding Matrix | | LogitLens | |
|------------------|--------|-----------|---------|
| 1. clocks | (0.11) | 1. clock | (10.78) |
| 2. clock | (0.11) | 2. dial | (8.80) |
| 3. Clock | (0.10) | 3. -clock | (8.76) |
| ... | | ... | |