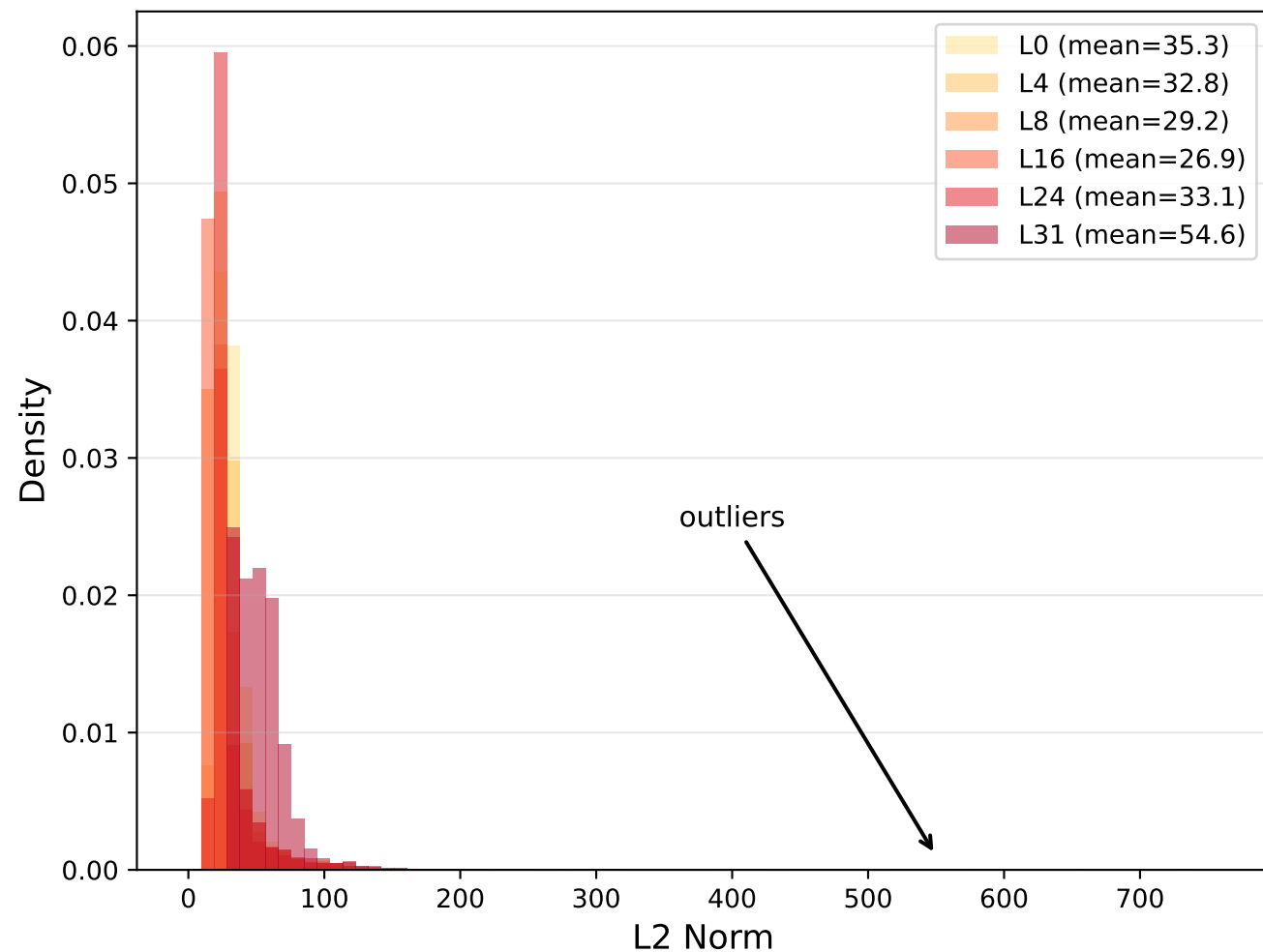


OLMo-7B + CLIP: L2 Norm Distribution Across Layers

Vision Tokens



Text Tokens

