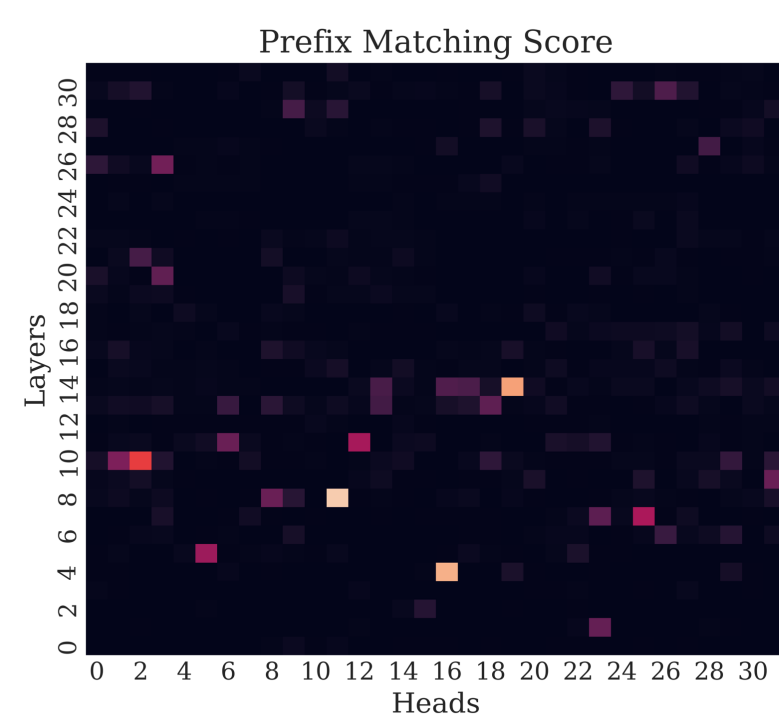
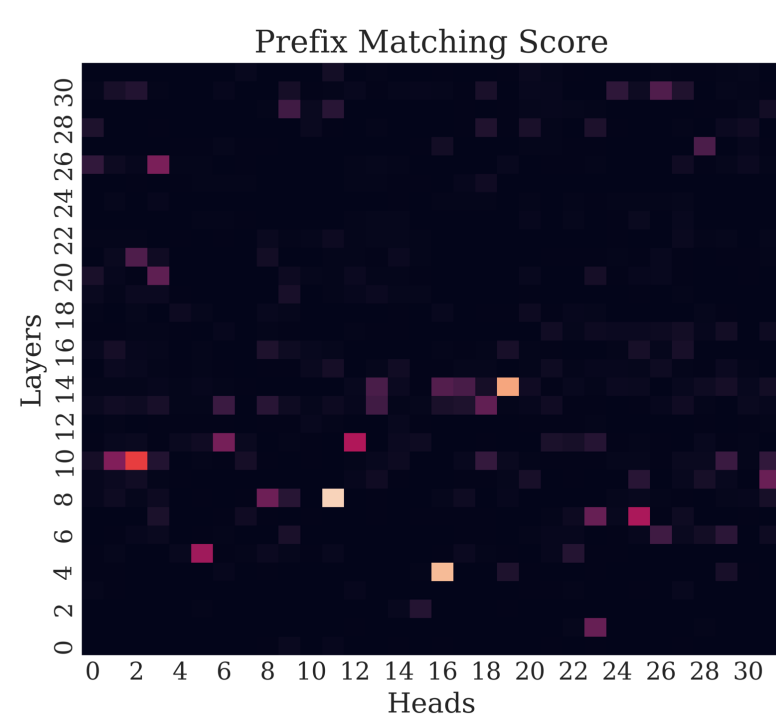


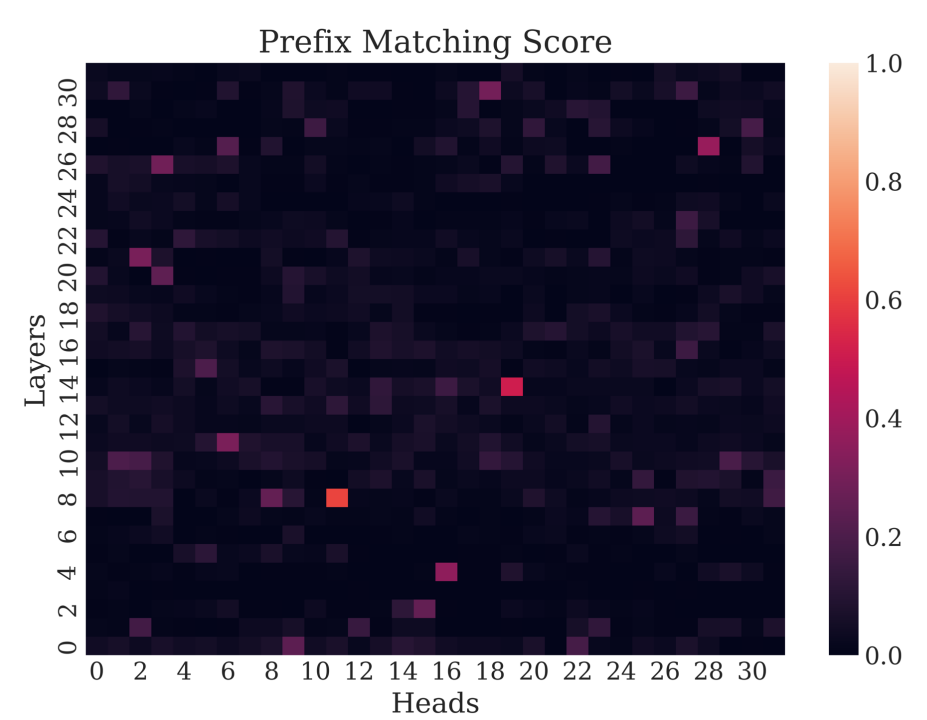
(a) Llama-3.1-8B
0.981



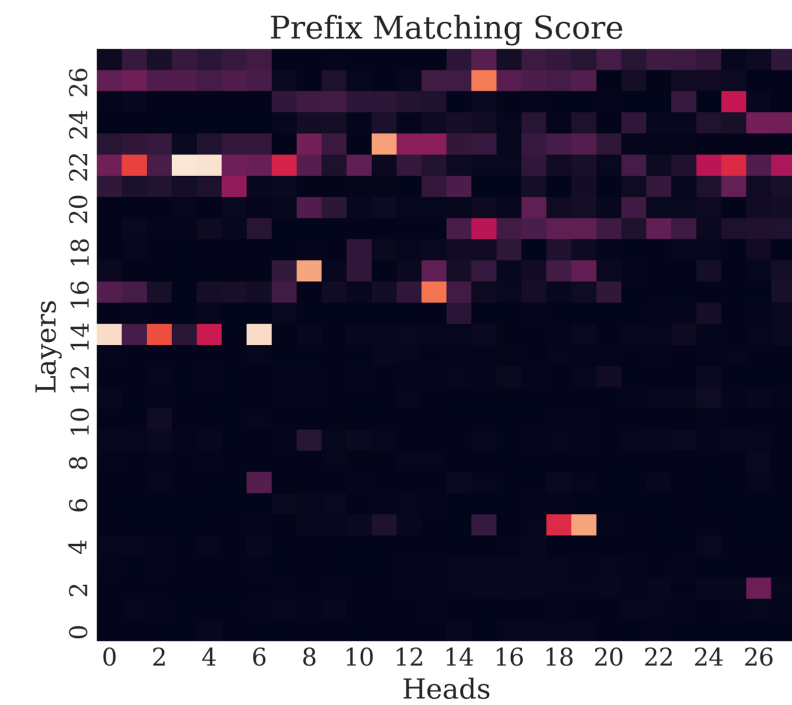
(b) Llama-3.2-11B-Vision-Instruct
(all tokens)
0.916



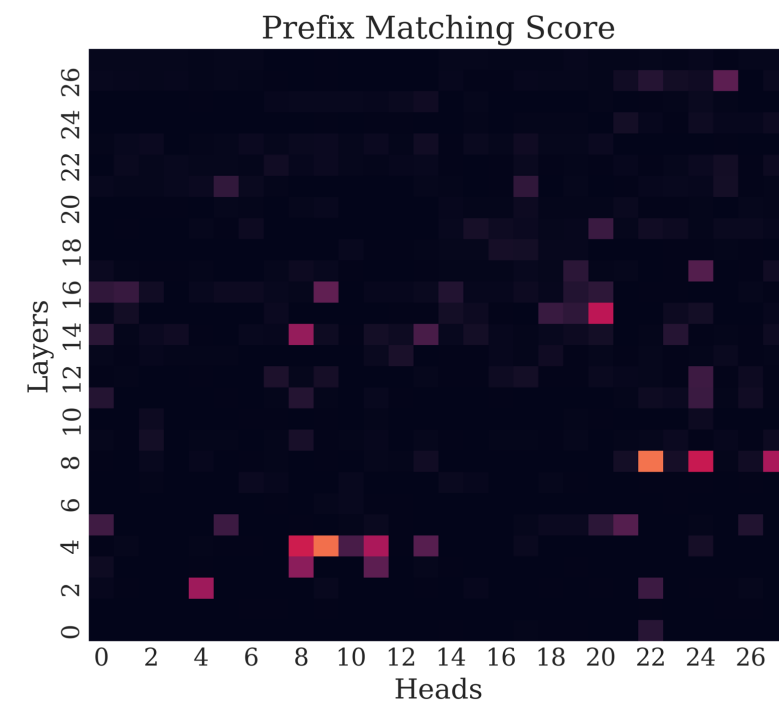
(c) Llama-3.2-11B-Vision-Instruct
(text tokens)
0.937



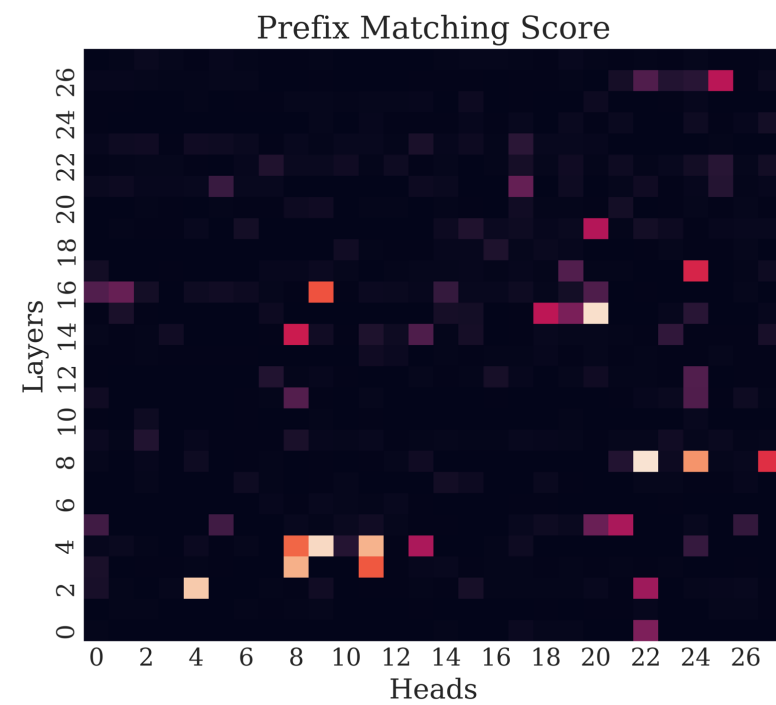
(d) Llama-3.2-11B-Vision-Instruct
(img tokens)
0.612



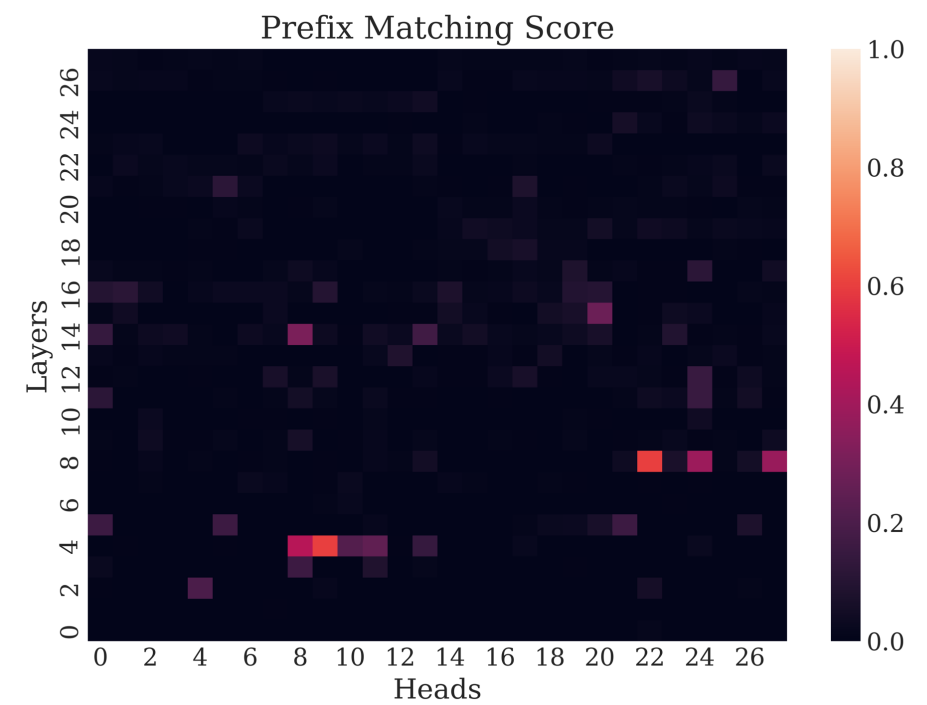
(e) Qwen2.5-7B-Instruct
0.987



(f) Qwen2.5-VL-7B-Instruct
(all tokens)
0.705



(g) Qwen2.5-VL-7B-Instruct
(text tokens)
0.979



(h) Qwen2.5-VL-7B-Instruct
(img tokens)
0.600