

Per-layer attention mass  $\rightarrow$  image tokens

