

Heatmap visualization of attention weights for the 'Layer' (0-11) vs 'Head' (0-11) matrix. The color scale ranges from 0.000 (dark purple) to 0.030 (light yellow). High attention is visible at (Layer 5, Head 8) and (Layer 5, Head 9).

