Attention Block 2.0 1.5 1.0 0.5 Δ_{cofa} 0.0 -0.5-1.0-1.5-2.00 1 2 3 4 5 6 8 9 10 11 layer