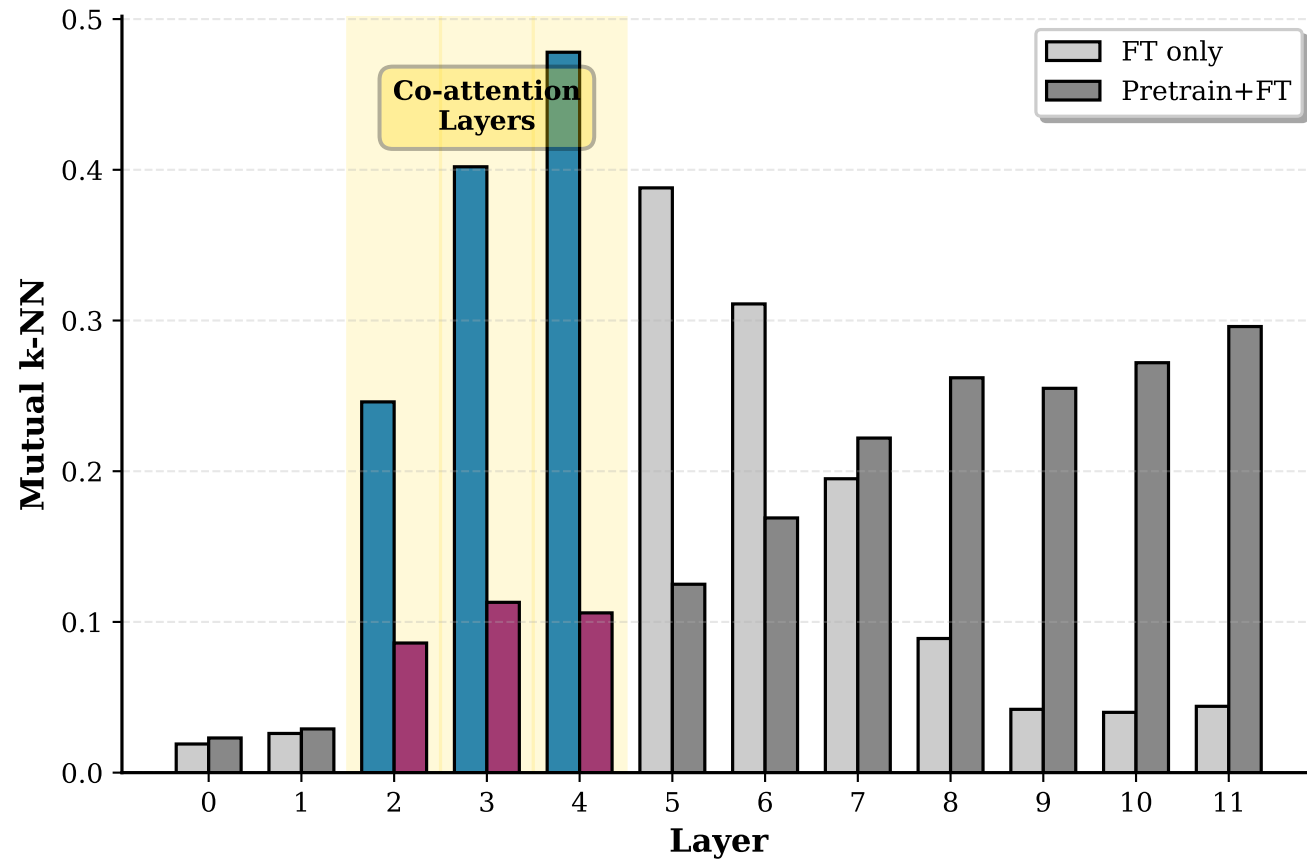


Early Fusion: Co-attention Layer Alignment



Mid Fusion: Co-attention Layer Alignment

