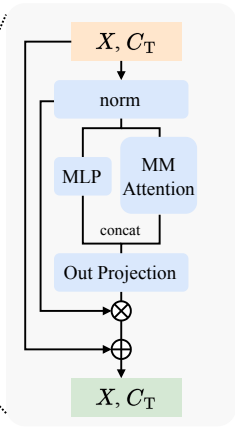
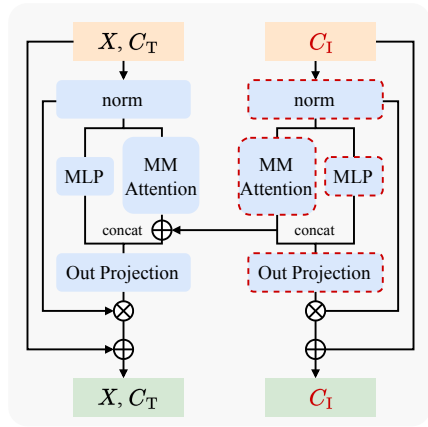


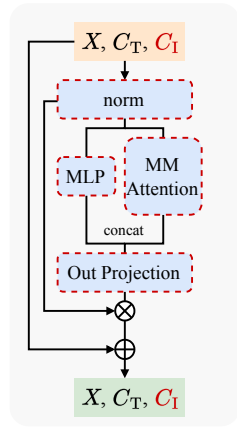
(a) Diffusion Transformer



(b) Original  
No image condition



(c) Integrate image condition  
by Adding



(d) Integrate image condition  
by MM-Attention