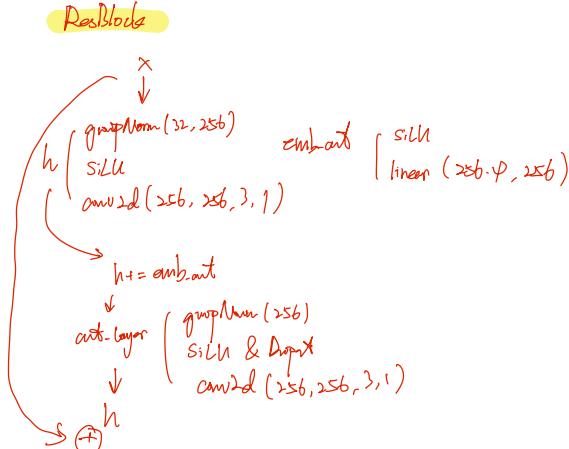


input-blocks  $\Rightarrow$  [ canid (4, 256, 3, 1), ]

$dh = 1 \cdot 256$   
 $t\_emb\_dim = 256 \cdot 4$   
 $attn = 256 \cdot 32 \approx 256$



$ch = 1 \cdot 256$   
 $head\_ch = 32 \Rightarrow num\_heads = 256/32 = 8$   
 $dim\_head = 256/8 = 32$   
 $transfoer\_depth = 1, context\_dim = 512$

