

$t_2$ :  
man

$t_5$ :  
hat

生成器

隐层向量

$h_1^l$

$h_2^l$

$h_3^l$

$h_4^l$

$h_5^l$

共享编码器

(Transformer编码器)

词向量与  
视觉特征



A



with

a

