



视觉目标检测



CNN & FNN

$v_2$

$v_5$

短语检测

A man

a hat

A man with a hat

重构模型

重构目标: A man with a hat .

或: Ein Mann mit einem Hut .

重构解码器

共享编码器

$e_1$

$v_2$

$e_3$

$e_4$

$v_5$

$e_6$

词嵌入层

$Z_w$ : A <E2> with a <E5> .