

New input

Generated questions

$I$ :



$Q$ :

What is the animal doing?

$A$ :

standing

VQG  
Model

$Q_1$ : What are the zebras doing?  
 $Q_2$ : What activities is the animal doing?  
 $Q_3$ : How the animal doing for sleep?  
...

VQA  
Model



Input

Infer

Compute  
Loss

Loss

$A_p$

Loss

Loss

$A_1; A_2; A_3; \dots$