



CNN

$I \times 1$ Tensor

Pretrained image features from ResNet

Fully Connected Layer



Activation (ReLU)



Multiplication

Transform Gate

Carry gate

What

$W \times 1$

Glove word vectors

holiday

$W \times 1$

is

$W \times 1$

the

$W \times 1$

room

$W \times 1$

decorated

$W \times 1$

for?

$W \times 1$

Fully Connected Layer

Transform Gate

Carry gate

Bias omitted for clarity

N way softmax