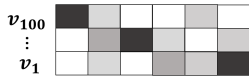
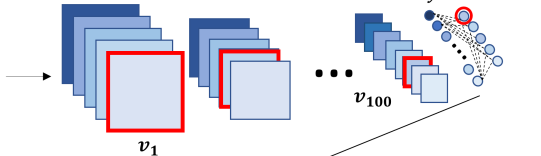


Input image



Decision-relevant features



$D_r = \{\text{red, crown}\}$

Relevance Vector Calculator

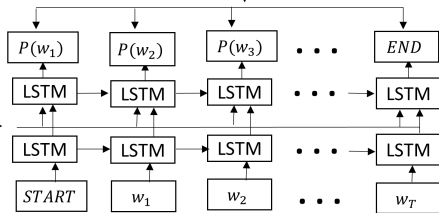
D
Dictionary

\mathbf{z} - Relevance Vector

Image features
from CNN

Multiply by
 $\text{ReLU}(\alpha^c)$

Weighted
image features



Ground truth word sequence