

Image of
environment (s_t)

Repeating Convolution,
Batch Normalization,
ReLU blocks

Shared fully
connected layers

Fully connected layer
with 5 outputs

Softmax

$\pi(a_1|s_t)$

...

$\pi(a_5|s_t)$