Sequence models

Feedforward models

- (Fixed) order of computation

  - Lower to upper layers

- Once we have the result

  - Discard all activations
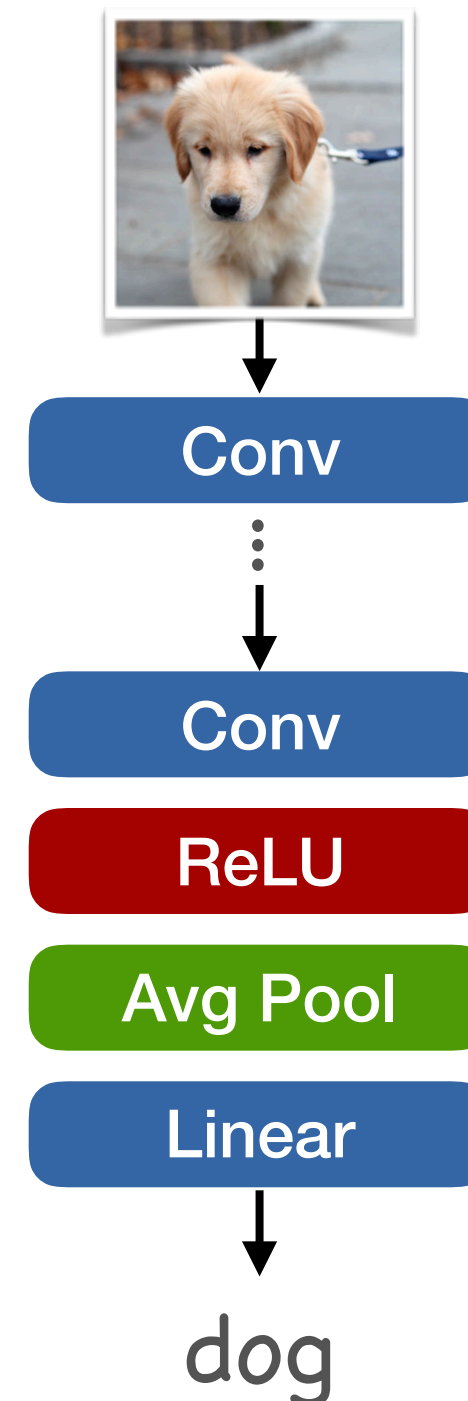
Would you use this to drive a car?

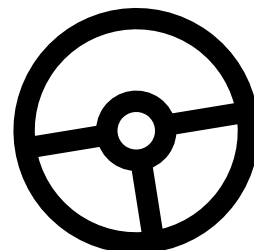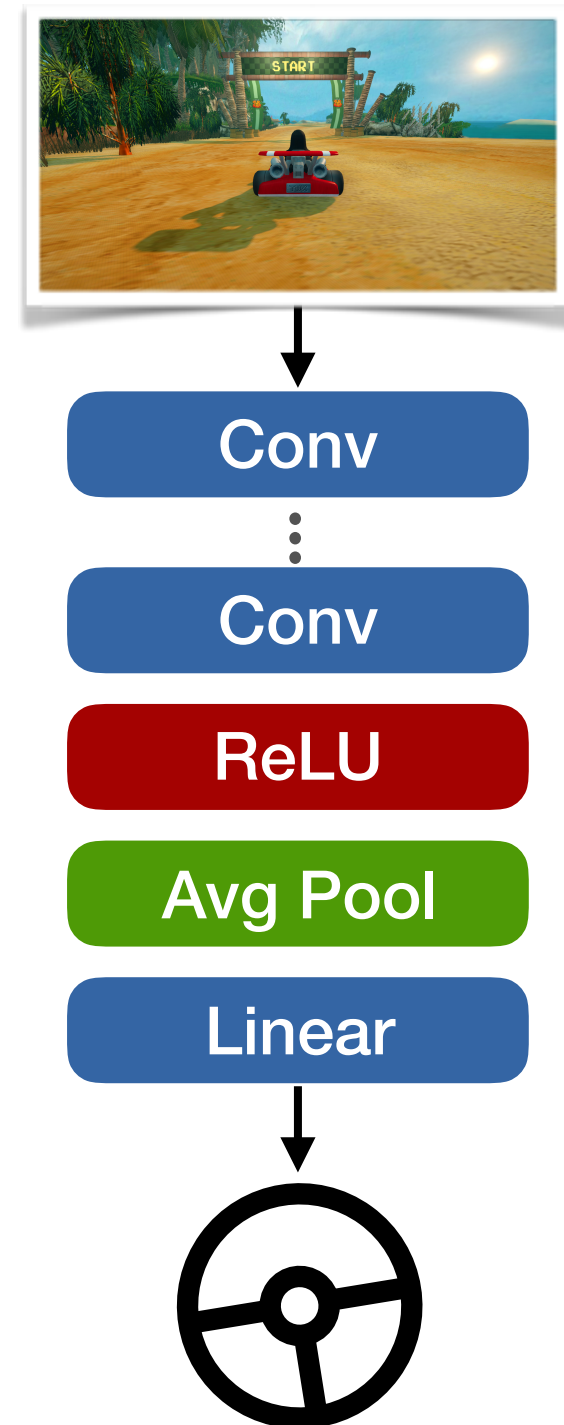# Would you use this to drive a car?

Conv

⋮

Conv

ReLU

Avg Pool

Linear

Would you use this to drive a car?

- Hopefully not

  - Independent decision for each frame

  - No state or memory

How should we keep state around?