EVAN RACAH

EXPLORING "WORLD MODELS"

(HA ET AL., 2018)

WHY WORLD MODELS?

- Focus on the representation

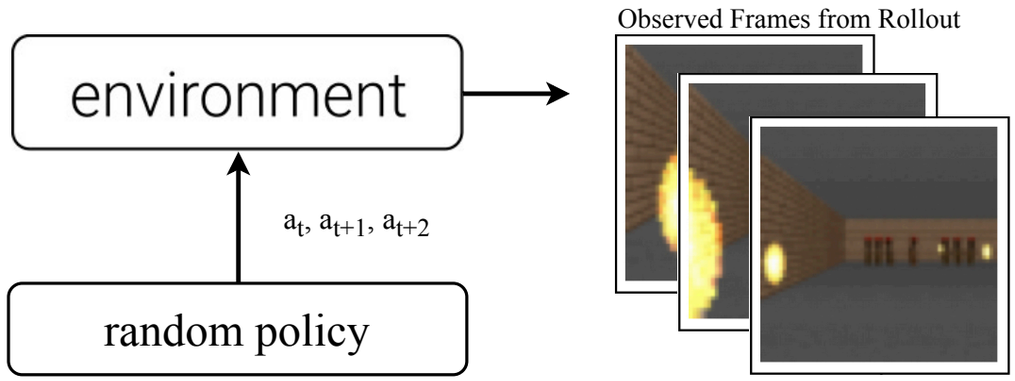
- Decoupling of the model components

- Action-conditioned prediction in latent space

- Many avenues of exploration and extension

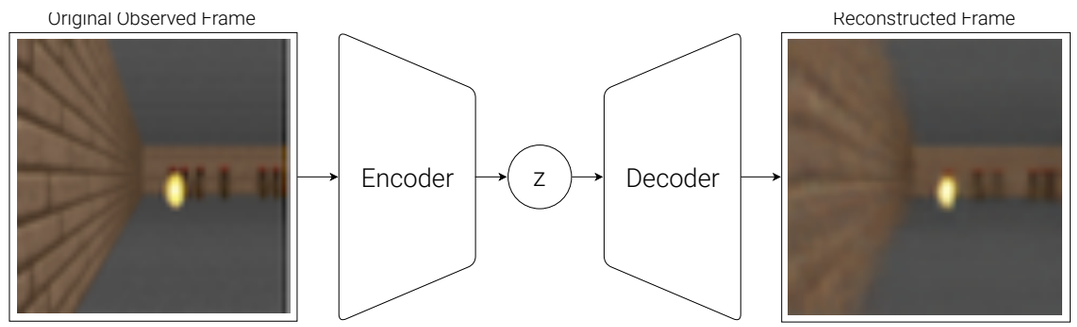
THEIR APPROACH

1



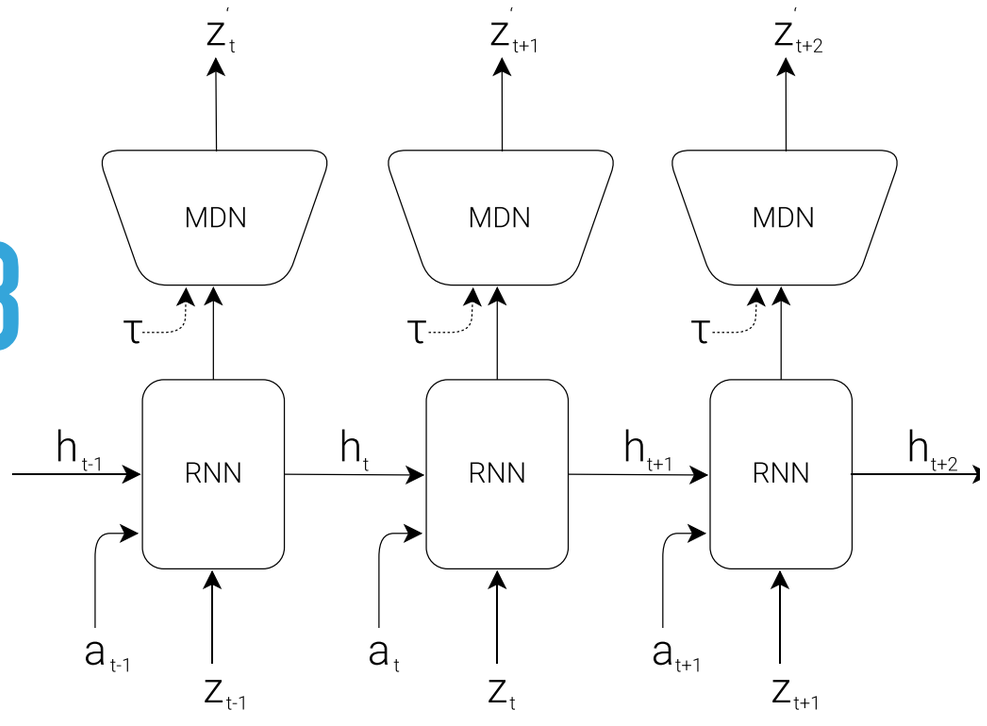
Generate actions, frames from random policy (10,000 rollouts)

2



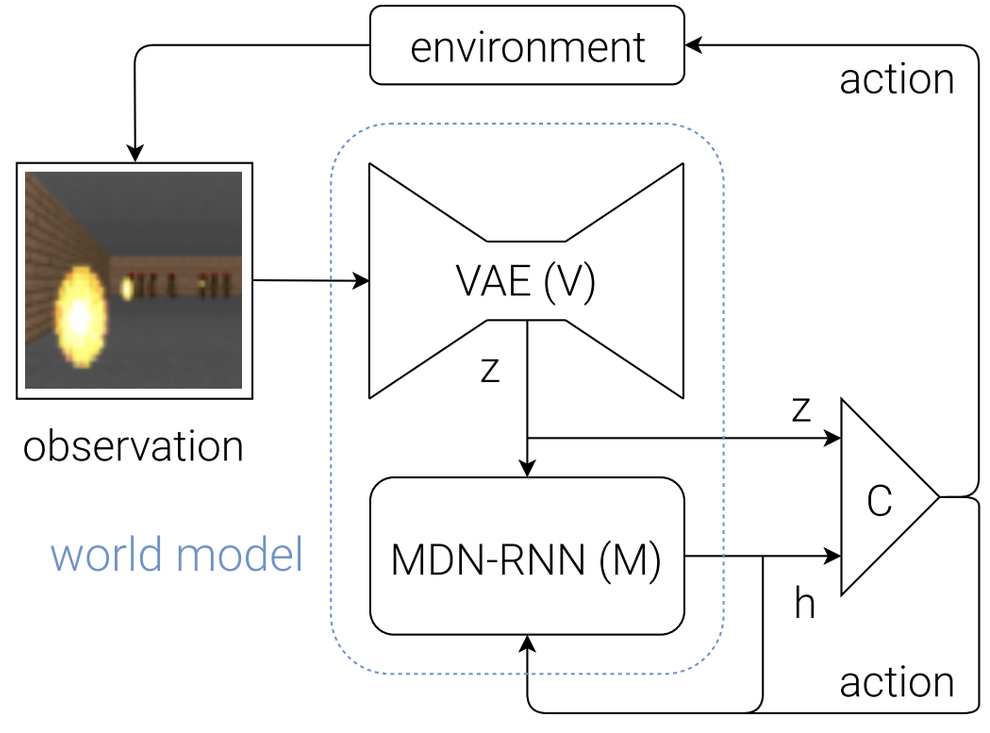
Train VAE to encode these frames into z

3



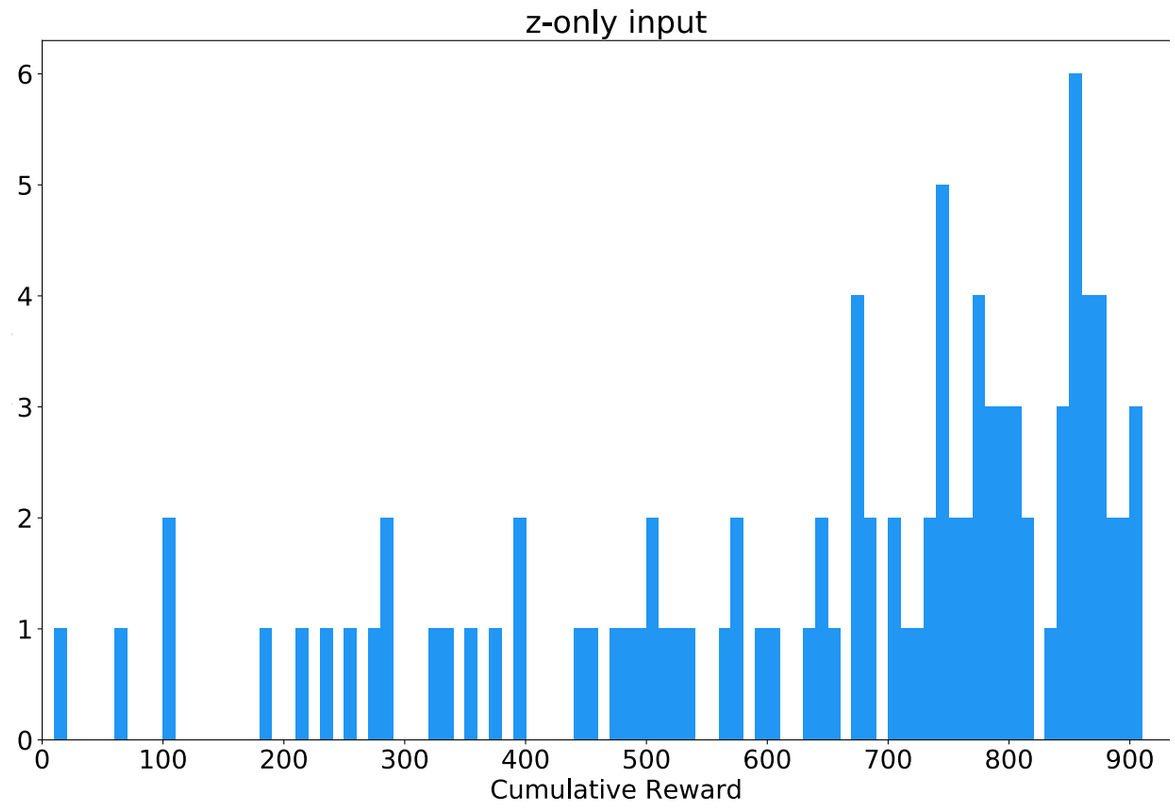
Learn to predict future z, given previous z and the action that caused it from random policy

4

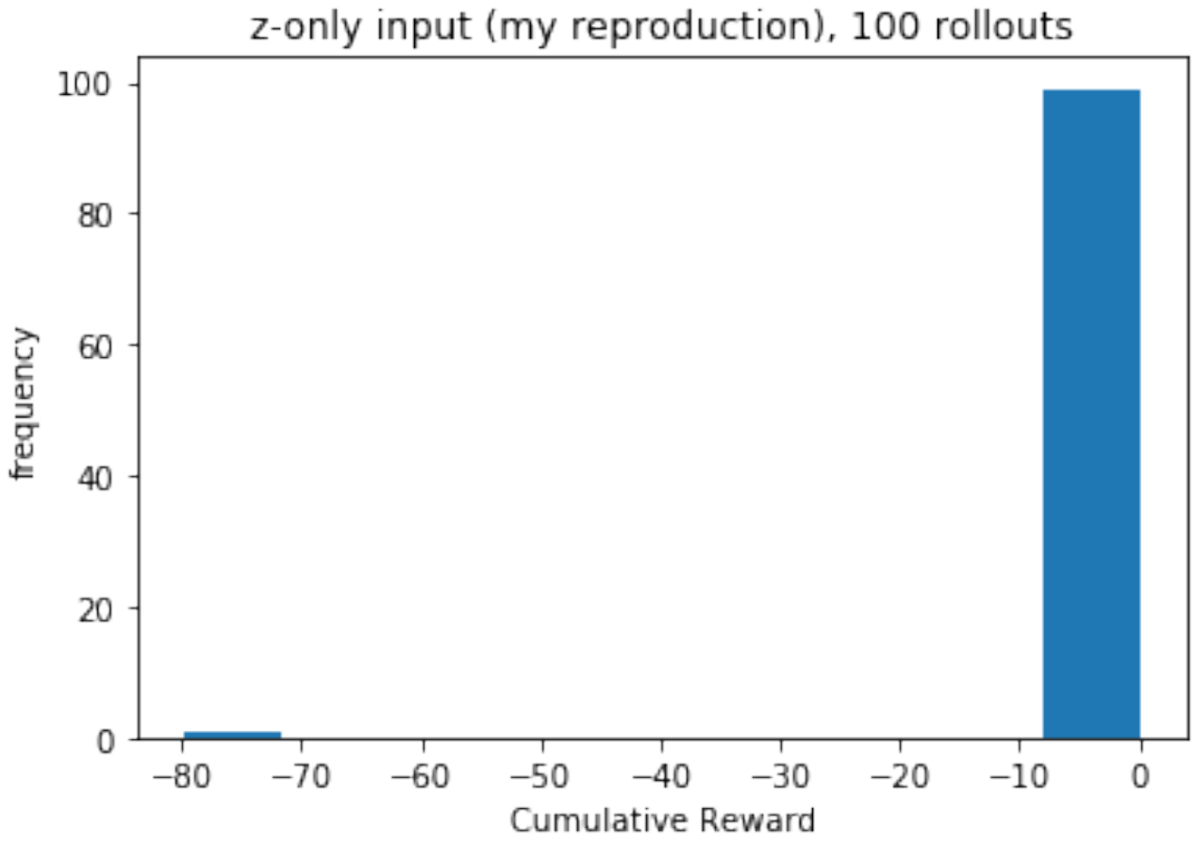


Learn best action using z from VAE and h from RNN as input using evolutionary algorithms

THEIR RESULTS VS. MINE



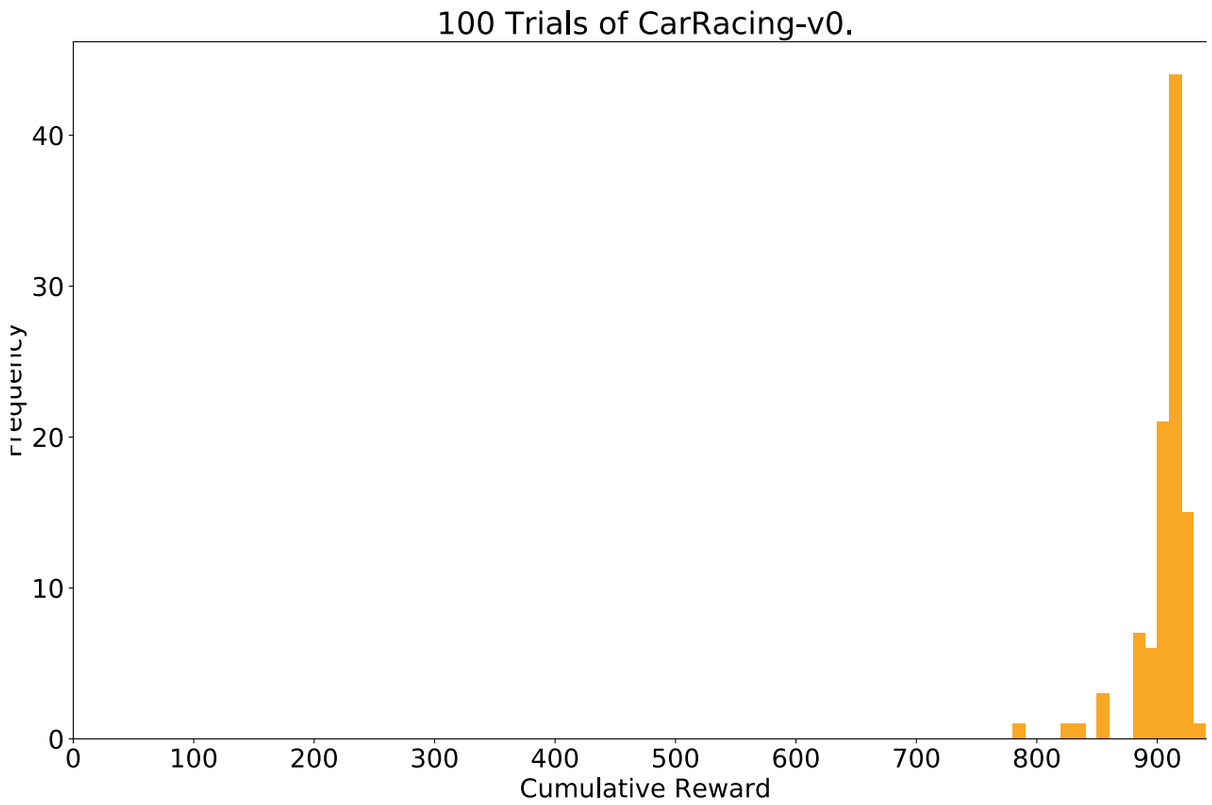
Their reward distribution for 100 rollouts with only input from VAE to controller



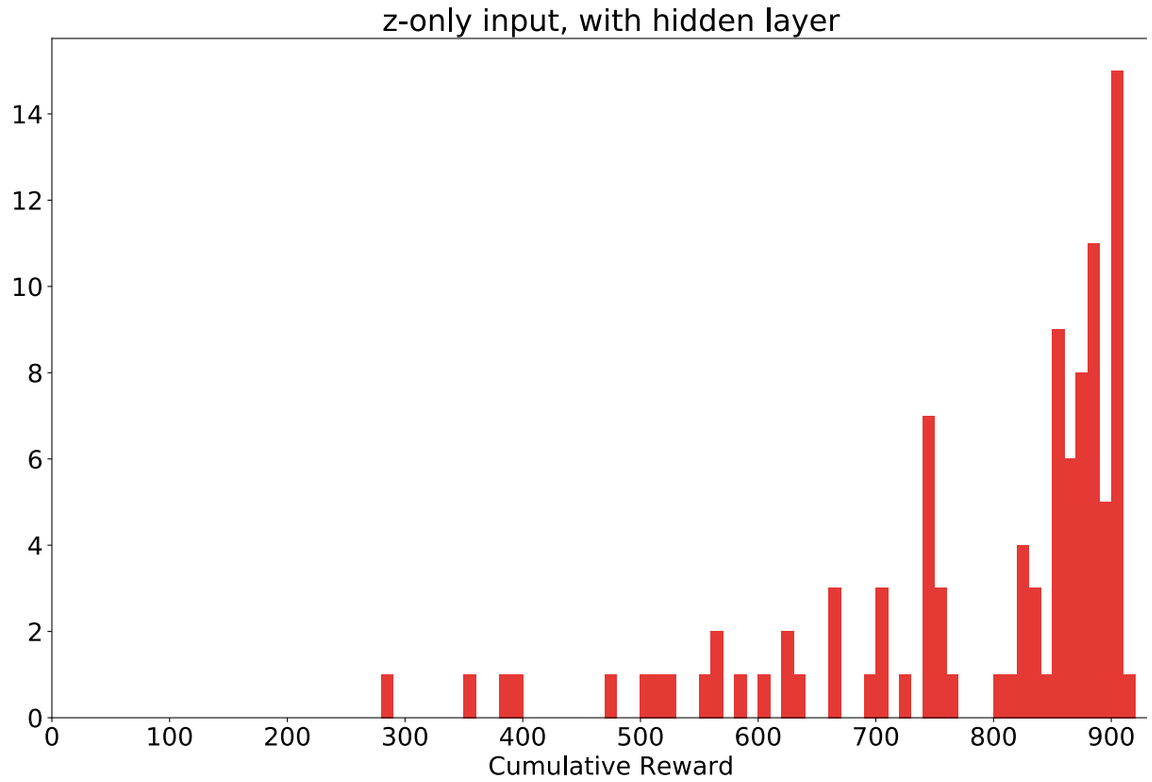
My reproduction of this with the time I had

RESULTS/DISCUSSION

MORE OF THEIR RESULTS



100 rollouts with z from VAE and h from LSTM



Just z from VAE, but added complexity to controller

HOW DID IT GO?

- What more needs to be done to reproduce
  - Add in features from LSTM
  - Run evolutionary search longer to match their performance
- Future Questions to Address
  - Policy gradient (just for controller and end-to-end)
  - Why can’t VAE see actions too?
  - Add some disentangling inductive bias to VAE?
  - Does it work for Atari?