Learning a Visual State Representation for Generative Adversarial Imitation Learning

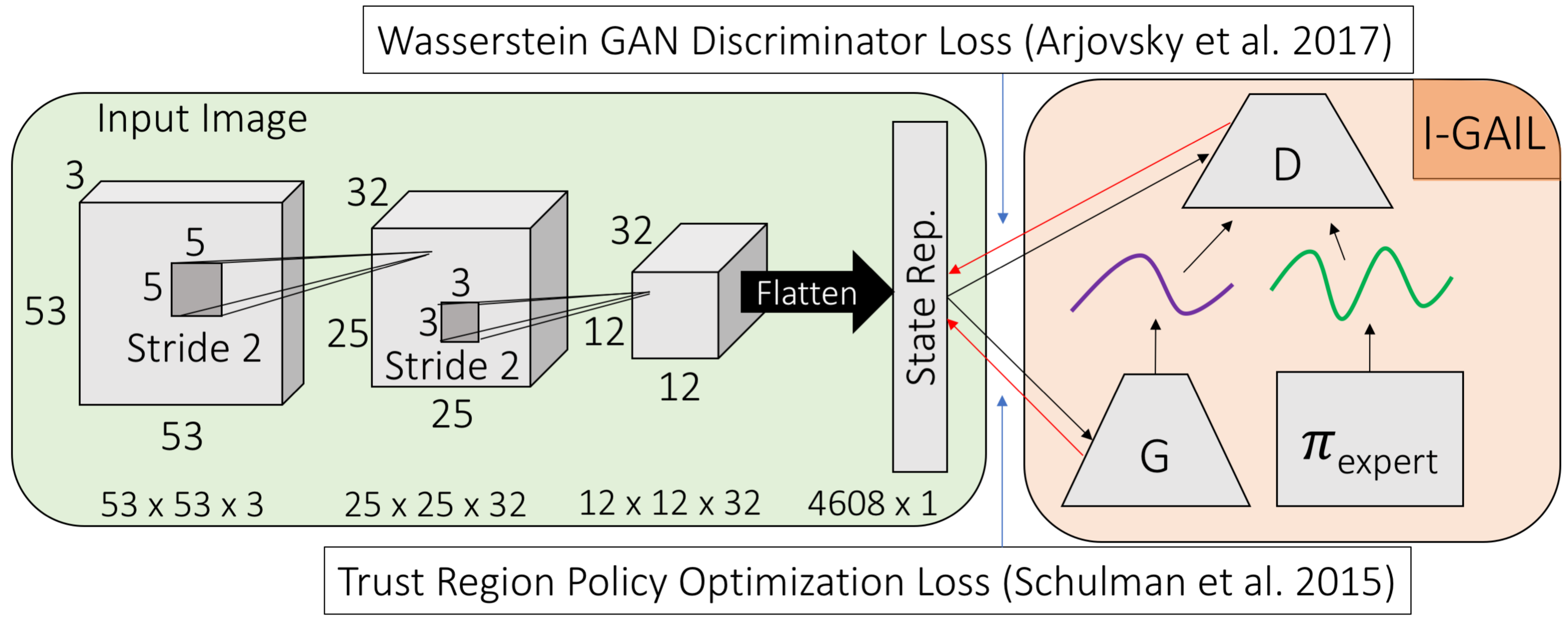
Boris Ivanovic
borisi@cs.stanford.edu

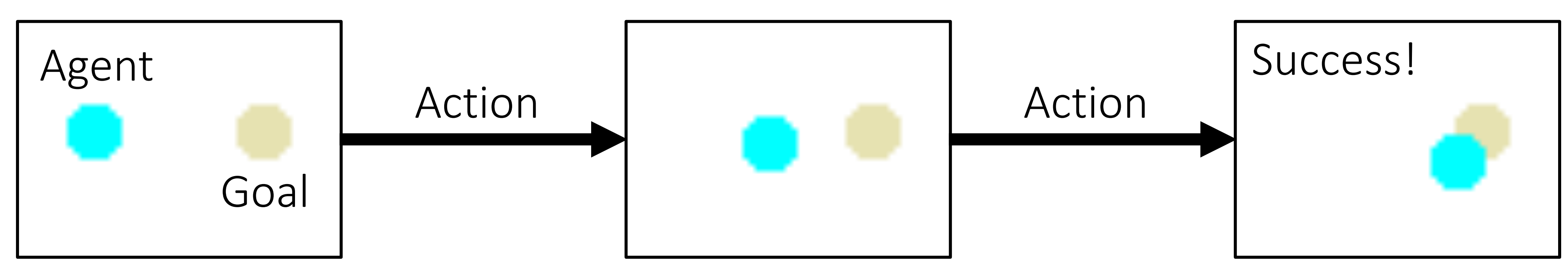
Motivation

- Generative Adversarial Imitation Learning (GAIL) [1] aims to train an agent that matches an expert's behavior, without an explicit reward signal or knowledge of the world.
- It performs well with low-level, low-dimensional state information. No convincing results on visual input yet.
- Important to enable visual input as it would allow the method to be applied to many more tasks, e.g. matching human behavior in driving.

Task Definition

- Generate a state vector representation of an input image to enable GAIL to work with images.
- Evaluate with embedding analysis (t-SNE) and learned agent performance (avg. episode length).





Approach

Environment and Data Collection

- Created and rendered 100 10x10-sized environments in Facebook's MazeBase: A sandbox for learning from games.
- Generated 100 trajectories (total) from a heading-following expert with zero-mean Gaussian noise and scaling applied to its actions, ensuring the generation of varied trajectories.

Model Definition and Training

- InfoGAIL [2] to perform policy generation (same as GAIL [1], but with the original GAN replaced by a Wasserstein GAN).
- 2-layer Convolutional Neural Network (CNN) with LeakyReLU nonlinearities for state representation.
- Trained by backpropagating the generator (G) TRPO loss and discriminator (D) Wasserstein GAN regression loss.

Performance

Agent	Avg. Episode Len.
Expert	5.39 ± 1.87
Random	155.18 ± 90.90
B. Cloning [3]	64.70 ± 117.68
InfoGAIL [2]	172.00 ± 146.42

Analysis

- CNN performs well, agent makes a beeline to the goal.
- If some noise in the agent's action causes it to miss the goal, it will never come back. => "DAgger problem"

Future Work

- CNN embedding analysis.
- Deeper dive into the agent's "DAgger problem."

References

[1] Ho, J. and Ermon, S., Generative Adversarial Imitation Learning, in *Neural Information Processing Systems (NIPS)* 2016

[2] Li, Y., Song, J., and Ermon, S., Inferring The Latent Structure of Human Decision-Making from Raw Visual Inputs, in *arXiv* 2017

[3] Pomerleau, D. A., Efficient training of artificial neural networks for autonomous navigation, in *Neural Computation* 1991