GAN and RL: A Perspective

Deepayan, Rajdeep, Sudarshan

University of Maryland, College Park

December 4, 2018

Motivation

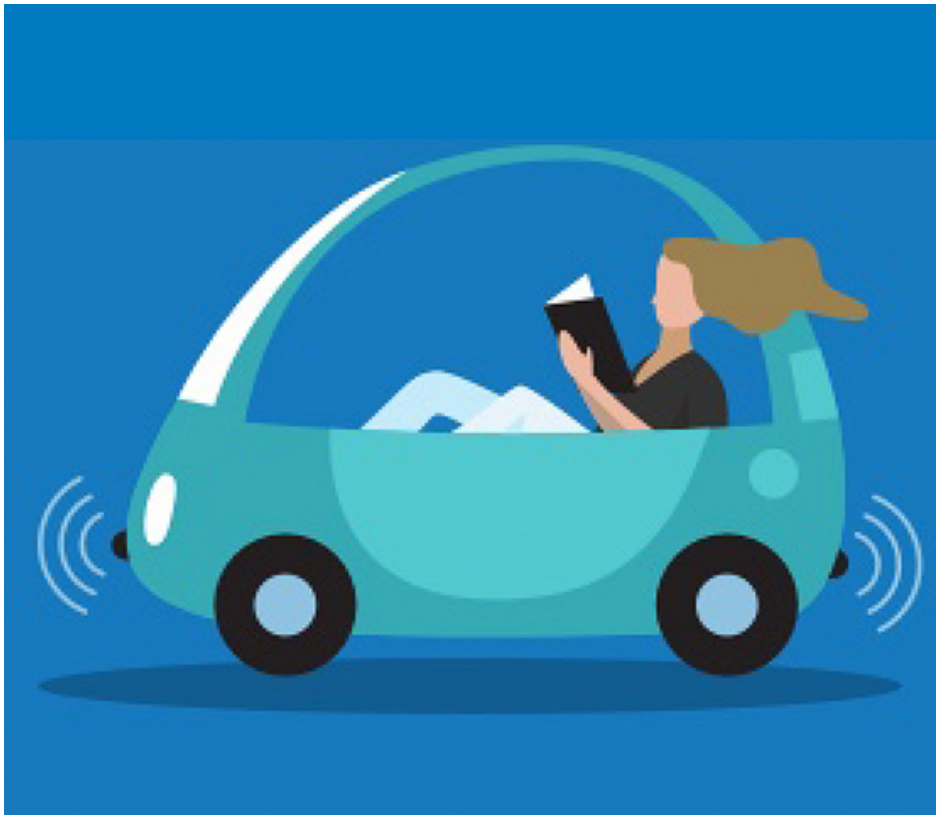


Figure: Self-Driving Cars [Erie]

- What's the deal about GAN?
- Was it RL that Alpha-Go used to defeat Lee Sedol?
- Can we possibly devise an idea linking these two?

Generative Adversarial Networks (GAN)

Figure: Fake Faces [NVidia]

- The high-level idea
- Can something still go wrong with GANs?
- (Yes, instability, mode collapses, ...)
- Enter, Wasserstein GAN.
- Earth-Mover Distance

Objective: $min_G max_D E_{x \sim P_r}[log(D(x))] + E_{\tilde{x} \sim P_g}[log(1 - D(\tilde{x})]$

Architecture

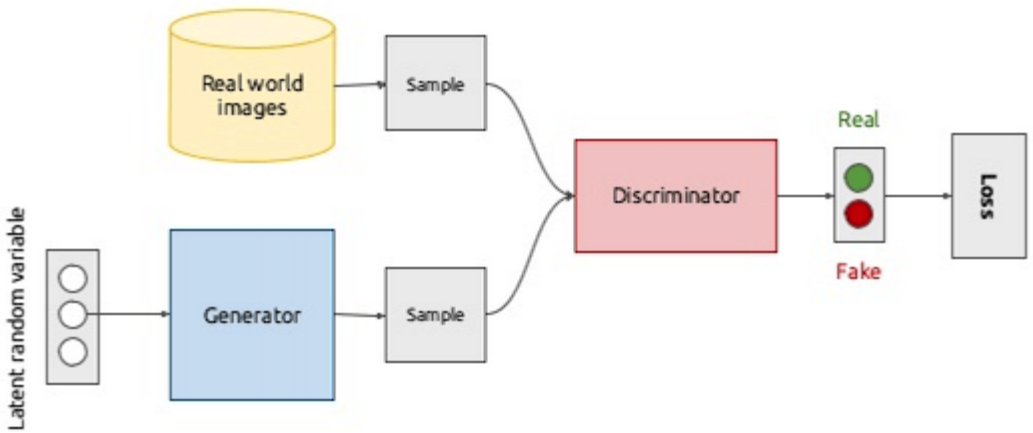


Figure: Skeleton of a GAN model [UPC 2016]

- Even WGAN by itself isn't good enough!
- Lipschitz constraint $\Leftrightarrow$ Weight clipping
- Think model capacity, unbounded gradient
- Thus, Gradient Penalty (WGAN-GP)

Objectives of the current study

- Classification Accuracy $\uparrow$ with GANs
- Augmenting available datasets for RL models

(Don't ask what RL is. The professor just taught it last week. You should know)

Reinforcement Learning (RL)

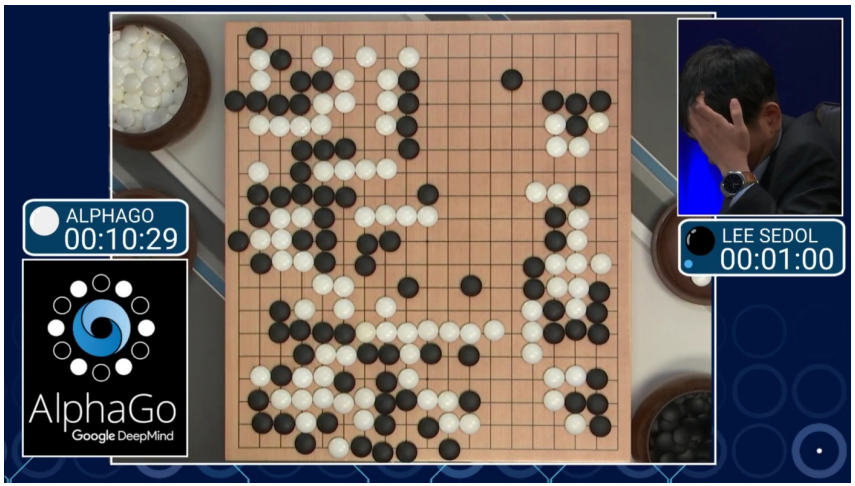


Figure: Go v/s Lee [Shelly Palmer]

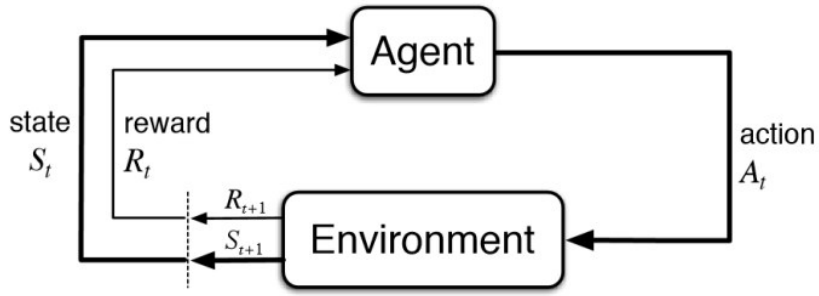


Figure: RL Architecture [Barto/Sutton]

A Pre-Trained RL Model

- Model Courtesy - Stanford's CS238 course (Surag et al.)
- MCTS (Monte-Carlo Tree Search) for Policy Improvement
- Policy Iteration through Self-Play

High Level Challenges

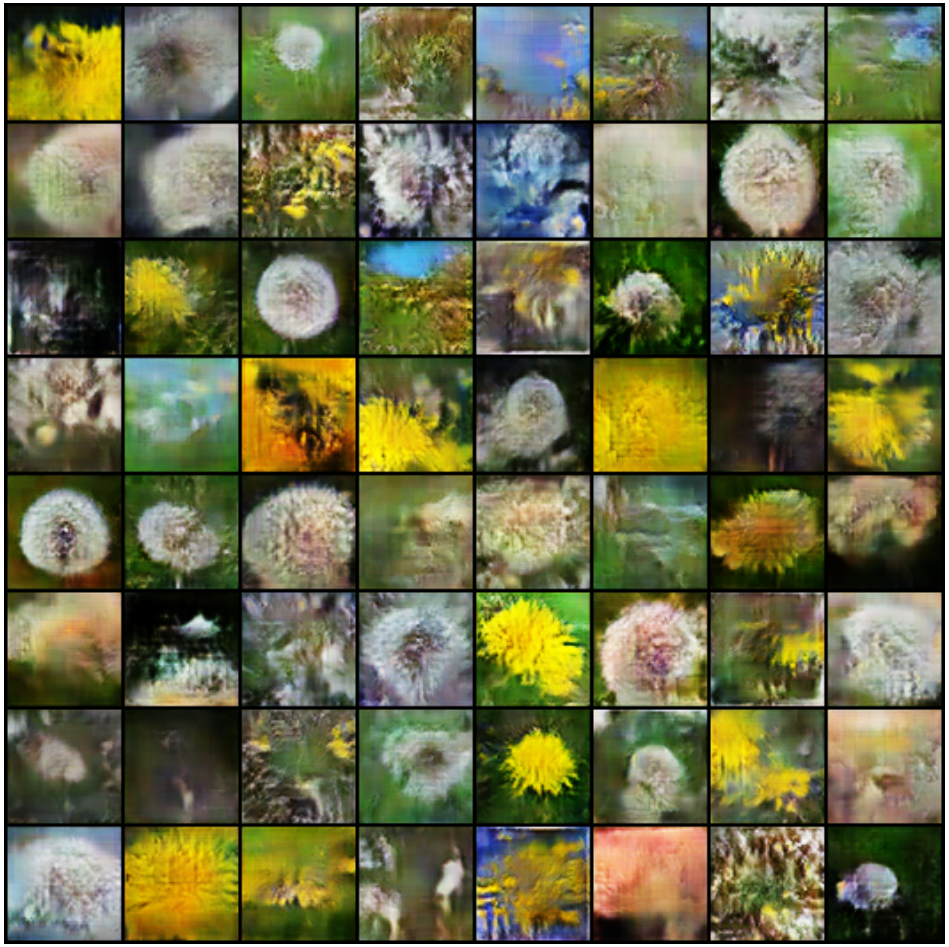
- Connecting the dot between GAN images and RL inputs
- Classification Accuracy?
- Training an RL model from scratch.
- Application to Self-Driving Cars

Experiments

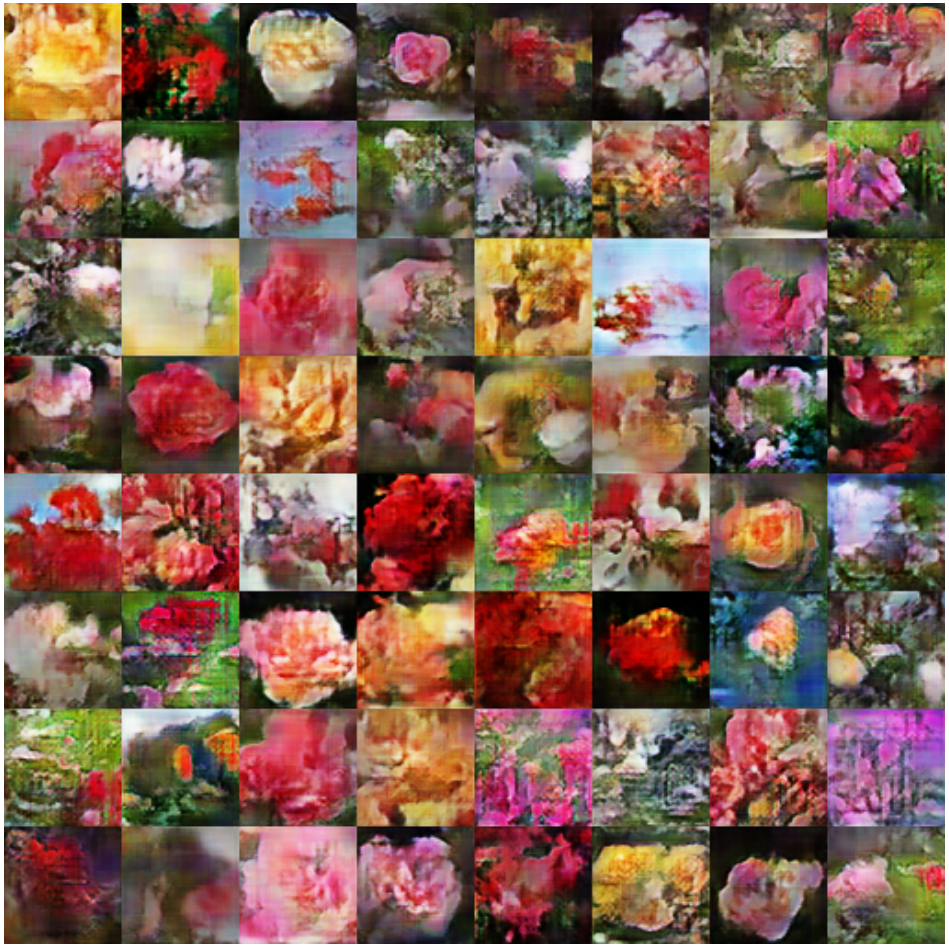
- Applied WGAN to Flower Dataset
- Co-GAN did not produce good results with small batch size
- Applied WGAN separately to each class

Generated Daisies

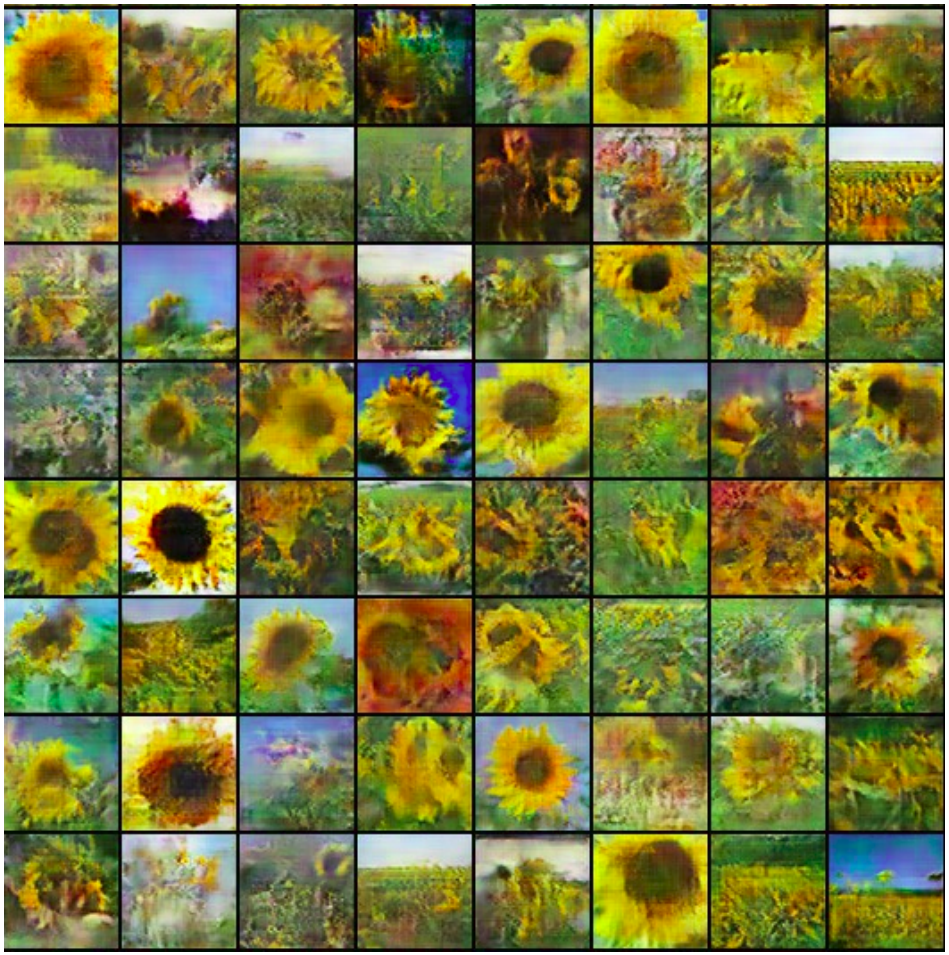
Generated Dandelions



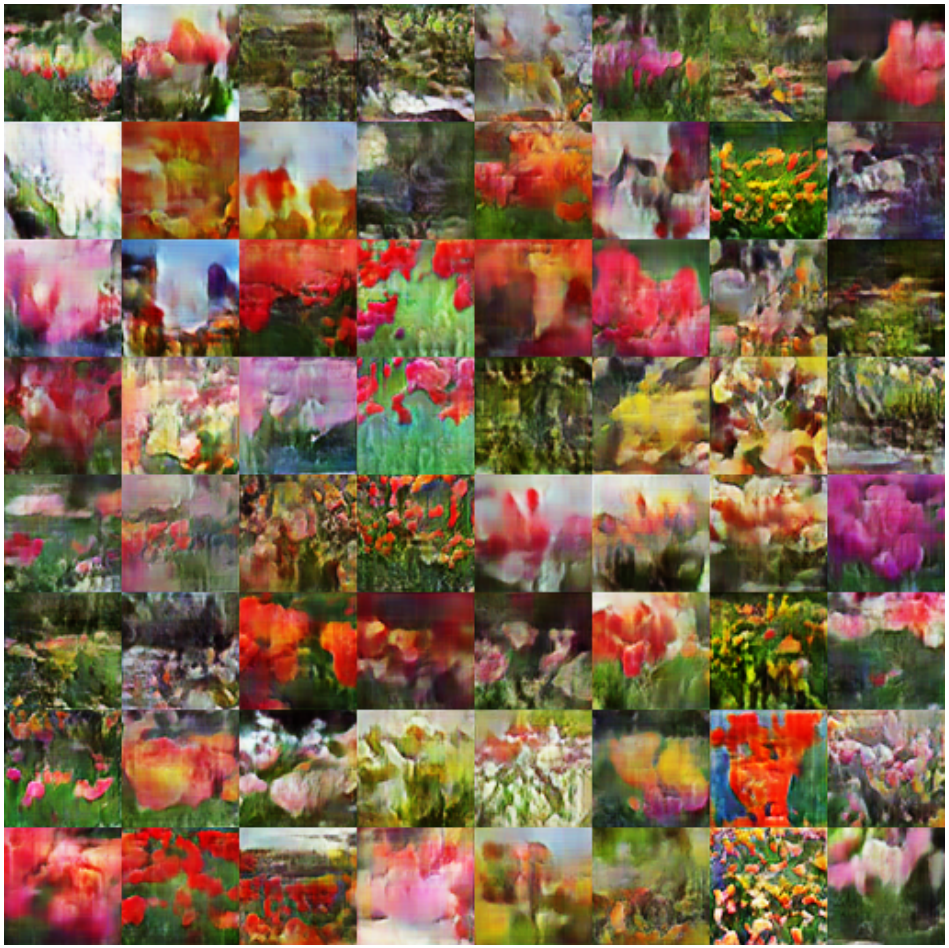
Generated Roses



Generated Sunflower

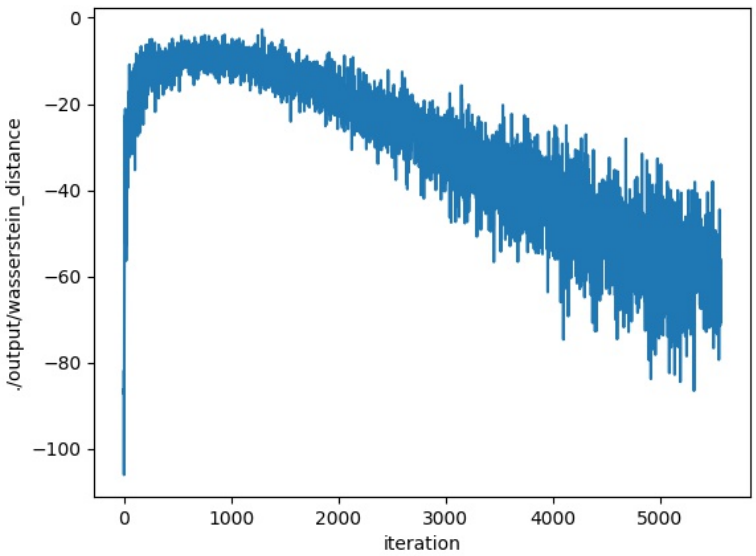


Generated Tulips

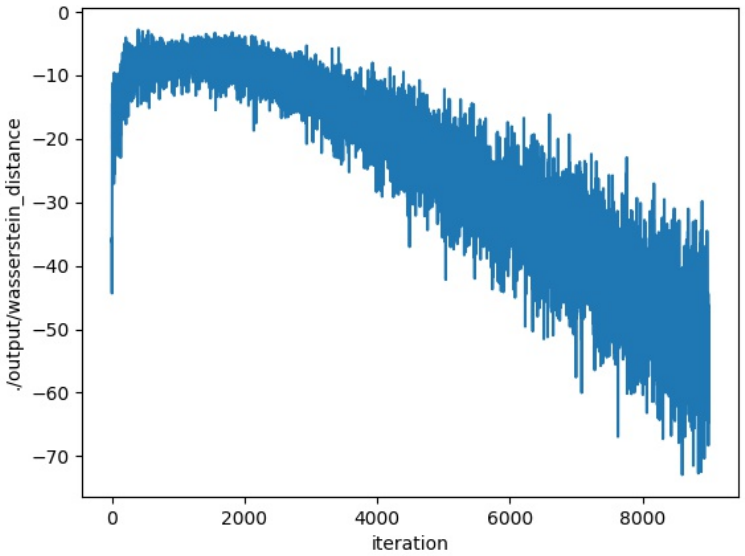


Wasserstein Distance Convergence

(1).jpg



(2).jpg



Results

- Achieved accuracy of 85.2 using GAN
- Used a DNN classifier to calculate the accuracy of the vavlidation set till 100 epochs.

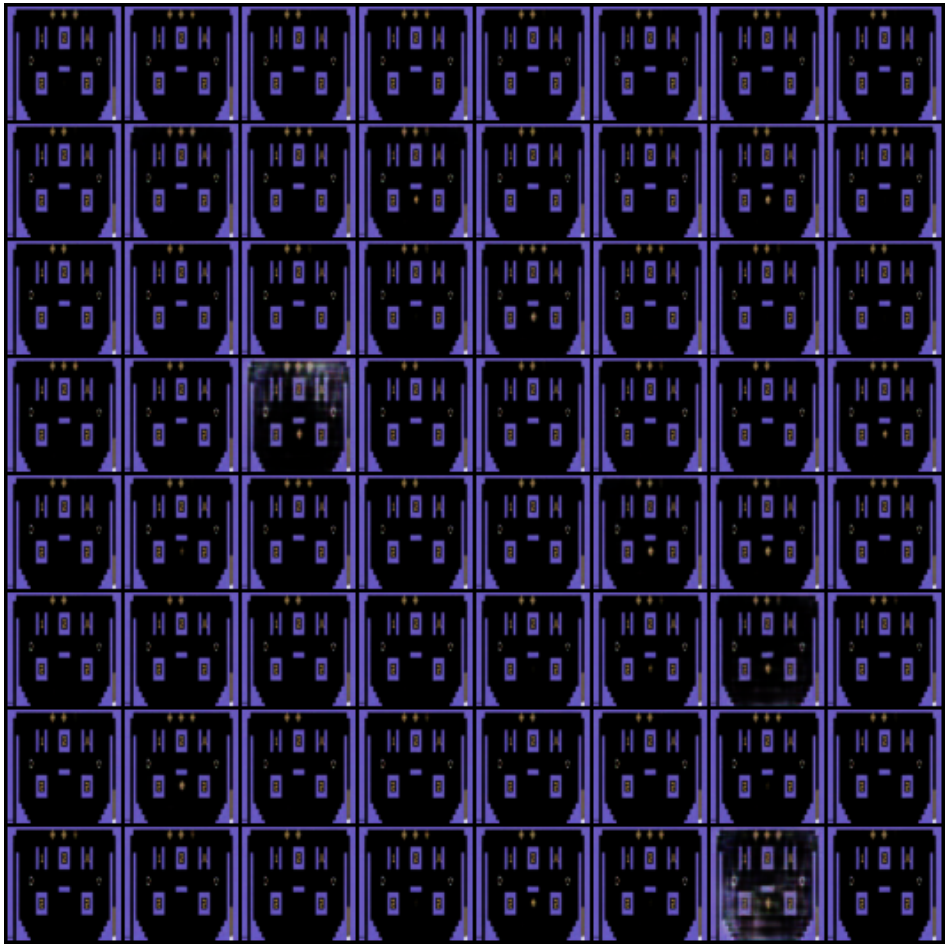
Experimental Challanges

- Initial Learning rate was poor due to weak Critic at the beginning
  - Used more critic iterations initilaly to correct this
- Wasserstein distance moves too slowly
  - Changing learning rates help
- Long training time(spent 500 worth of credits)

Next Steps

- Use conditional GANs on the cloud
  - It might improve tulips, roses, etc. by improving basic features detection
- Select only images which the critic approves (have to see how it would work)
- Figure out the gaps in our RL approach

Generated Pinball images



Questions?

The ball's in your court. Thanks for listening.