

Coevolutionary GANs Training to Foster Diversity Jamal Toutouh (toutouh@mit.edu), Erik Hemberg, Una-May O'Reilly

MOTIVATION

Generative Adversarial Networks



 $\min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}(\mathcal{D}, \mathcal{G}) = \mathbb{E}_{x \sim p_{data}(x)} [\phi(\mathcal{D}(x))] + \mathbb{E}_{x \sim p_z(z)} [\phi(1 - \mathcal{D}(\mathcal{G}(z)))]$

- Simultaneous gradient updates in GAN training lead to unstable dynamics
- Similar degenerate behaviors have been studied by the coevolutionary computing for minimax optimization
 - Mode Collapse ↔ Focusing
 - **Discriminator Collapse** ↔ **Relativism**
 - Vanishing Gradients ↔ Loss of Gradients





Supplementing GAN training with established coevolutionary techniques proposed to address the unstable dynamics

RELEVANT PAPERS

- Conference (GECCO '19), July 13–17, 2019.
- Training. NeurIPS 2018 Workshop on System for ML, 2018.

MUSTANGS: DISTRIBUTED COEVOLUTIONARY GANS TRAINING

Lipizzaner is a distributed, coevolutionary framework to simultaneously train multiple GANs with gradientbased optimizers



1. J. Toutouh, E. Hemberg, and U. O'Reilly. Spatial Evolutionary Generative Adversarial Networks. In Genetic and Evolutionary Computation

2. T. Schmiedlechner, I. Yong, A. Al-Dujaili, E. Hemberg, U. O'Reilly. Lipizzaner: A System That Scales Robust Generative Adversarial Network

3. A. Al-Dujaili, T. Schmiedlechner, E. Hemberg, U. O'Reilly. Towards distributed coevolutionary GANs. AAAI 2018 Fall Symposium, 2018.





- **Linear** instead of quadratic scaling

For each iteration, each network **randomly picks a loss function** to optimize its weights according to

Near linear training times

Figure: CelebA results





