

Part 3/6: Training Algorithm

## Practical Tips

### 1 Data & Activation

- ✓ Normalize inputs to `[-1, 1]`
- ✓ Use `tanh` activation for generator output
- ✓ `LeakyReLU` in D, `ReLU` in G

### 2 Normalization & Regularization

- ✓ Batch normalization (except last G, first D layer)
- ✓ Label smoothing: `real=0.9, fake=0.1`
- ✓ Add noise to discriminator inputs

### 3 Optimization

- ✓ Use Adam optimizer with  $\beta_1 = 0.5$

### Architecture Guide

#### Generator

- Dense/Conv Layers
- ReLU
- Batch Norm
- ...
- No BN (Last Layer)
- tanh Output

#### Discriminator

- No BN (First Layer)
- LeakyReLU
- Batch Norm
- ...
- LeakyReLU
- Sigmoid Output

### ⚠ Monitoring

Watch D loss carefully - if it drops to 0, G won't learn properly

✓ Monitor D loss - shouldn't go to 0

✓ Recommended Optimizer

Adam with  $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$