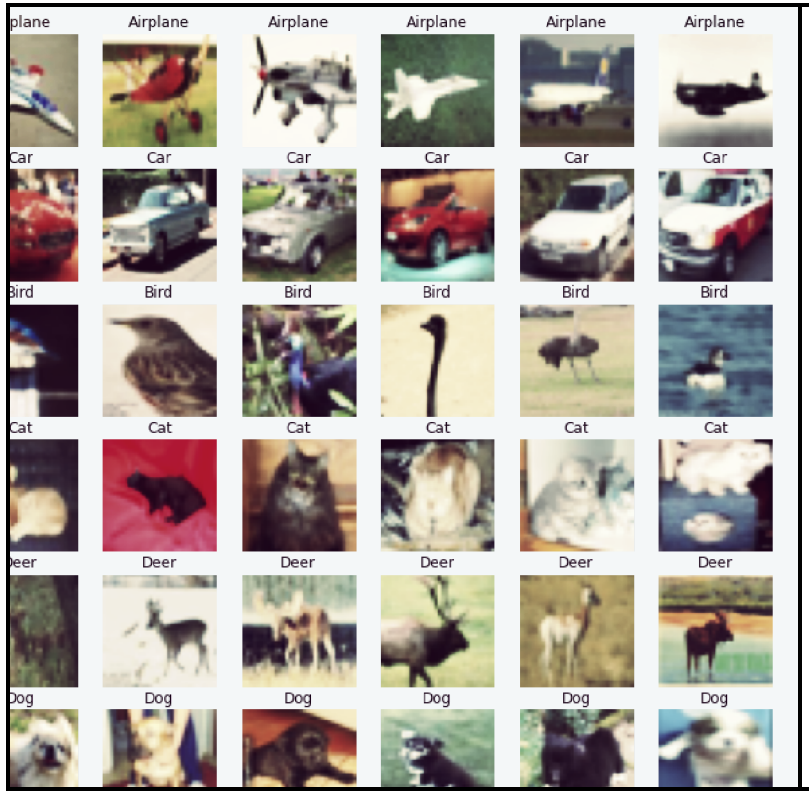
Conditional Generation



Can AI Generate Images Based On What I Tell It?

Applications?

- Video Game Art
- Historical Reenactments
- Product Design

Experimental Details

Dataset

- CIFAR-10
- 10 Classes
- 50K Training Set Images
- 10K Testing Set Images
- Normalize to [-1, 1]

Training Process

- Adam Optimizer with Beta1 = 0.0, Beta2=0.999
- Learning Rate initially set to 0.0002, with equal number of discriminator updates and generator updates
- Batch Size: 256

Metrics

- Metrics must measure both the quality and diversity of the output
- Main: Kernel Inception Distance
- Secondary: FID, IS
- Generate 10K images and compare statistics with CIFAR10 validation set (10K)

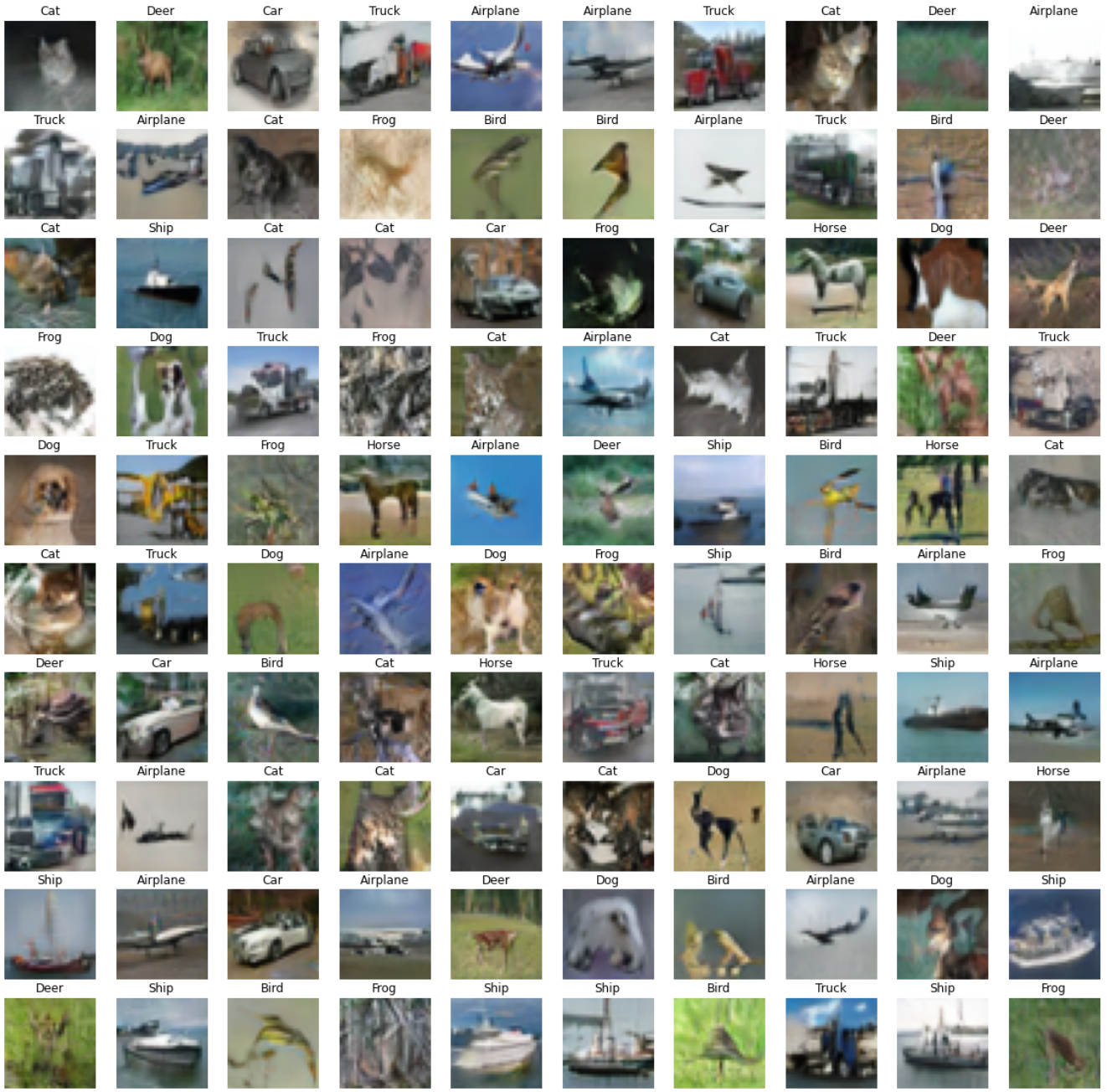
Summary of Results

Conditional DCGAN
KID: 0.0223
FID: 36.349
IS: 6.877

Label Smoothing and R1 Regularization
KID: 0.01937
FID: 32.52
IS: 7.064

Spectral Normalization
KID: 0.01856
FID: 30.165
IS: 7.148

SNGAN with Projection
KID: 0.01144
FID: 19.988
IS: 7.988



Truncation Trick

Truncation trick was applied to improve the quality of the final GAN output, at some cost to image diversity

Summary of Results

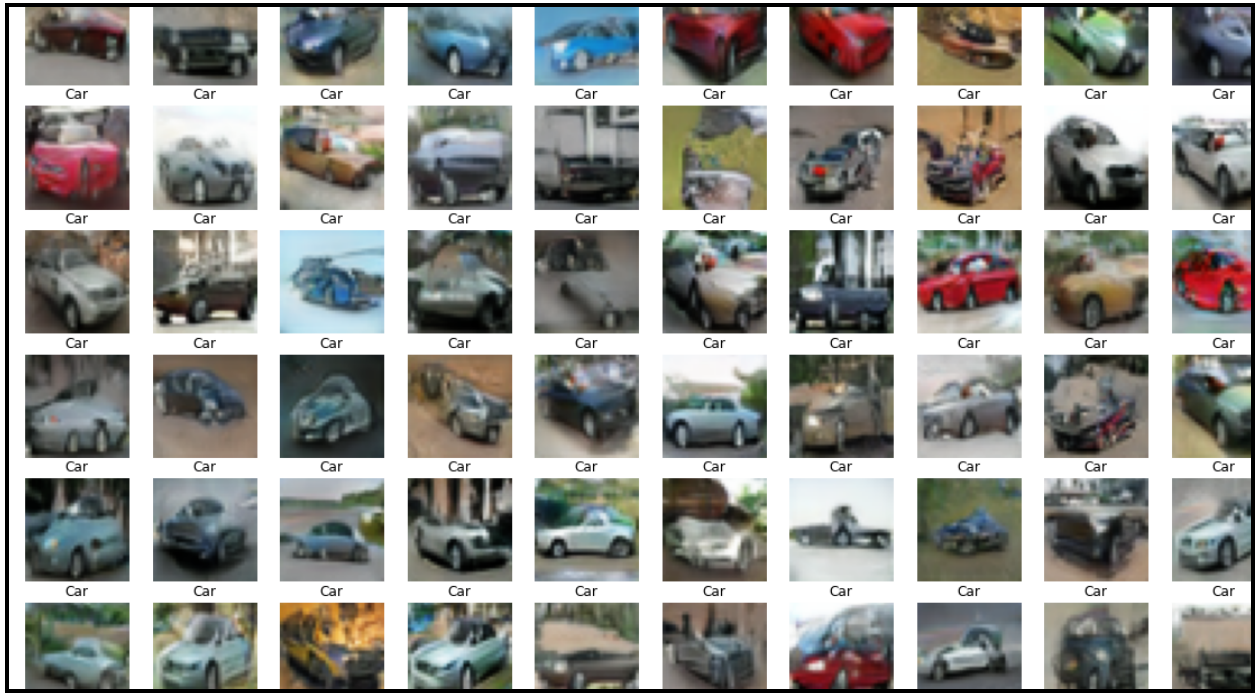
Animals are Harder to Generate

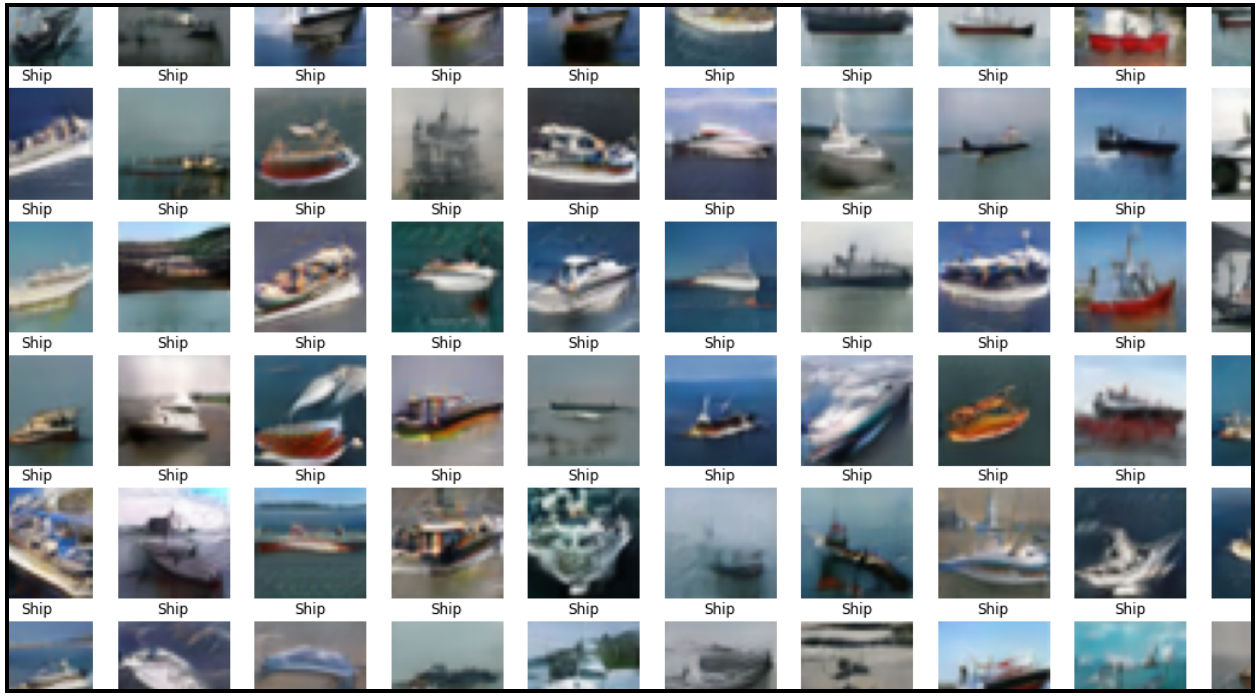
GAN appears to get the “texture” right, but has trouble getting the correct proportions and location of certain animal features

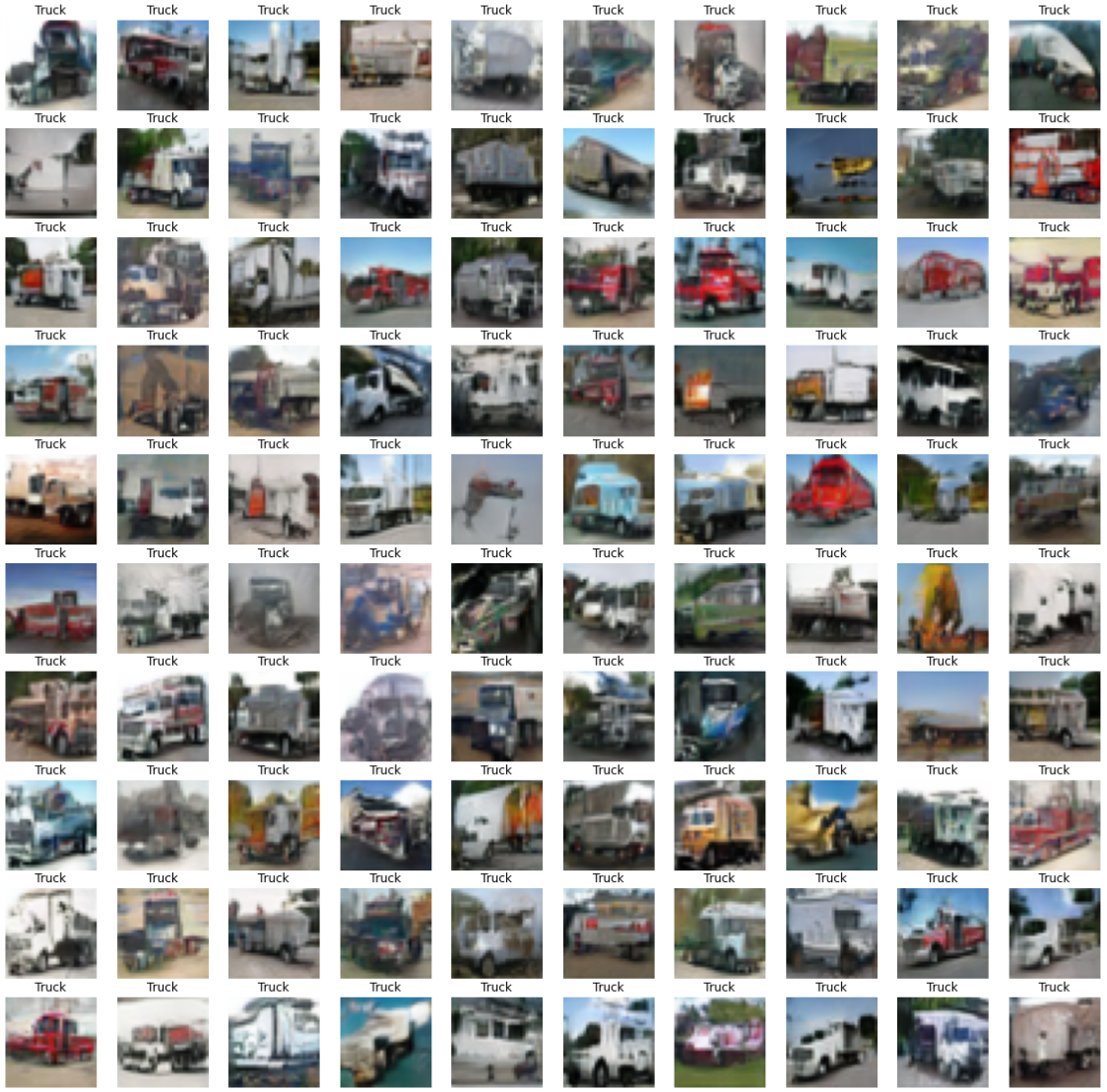
Does the GAN Suffer From Mode Collapse?

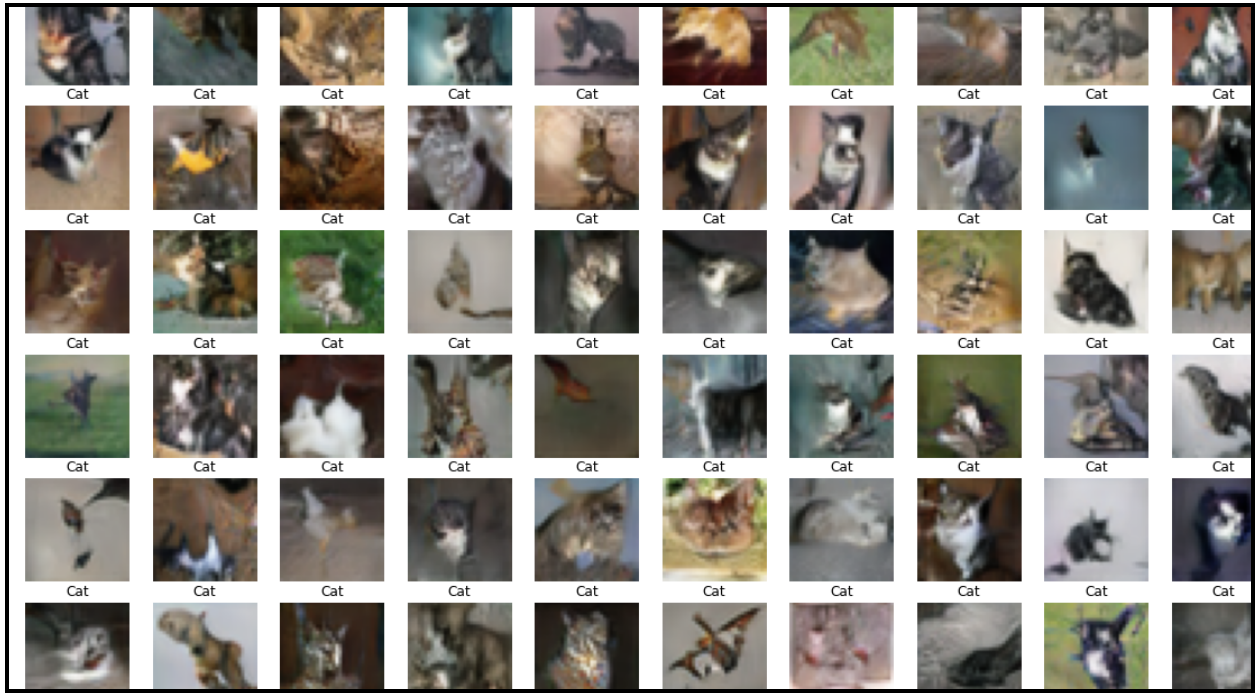
Most images are diverse even with truncation, but GAN can generate similar looking results at times

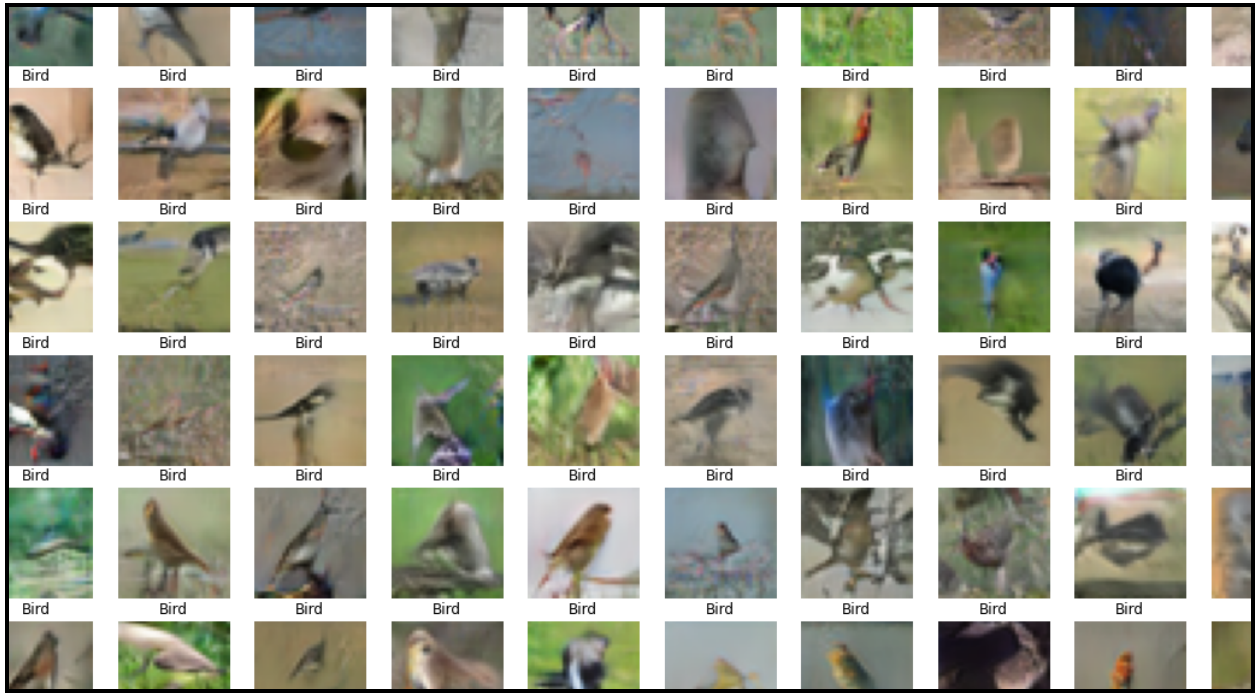

Airplane

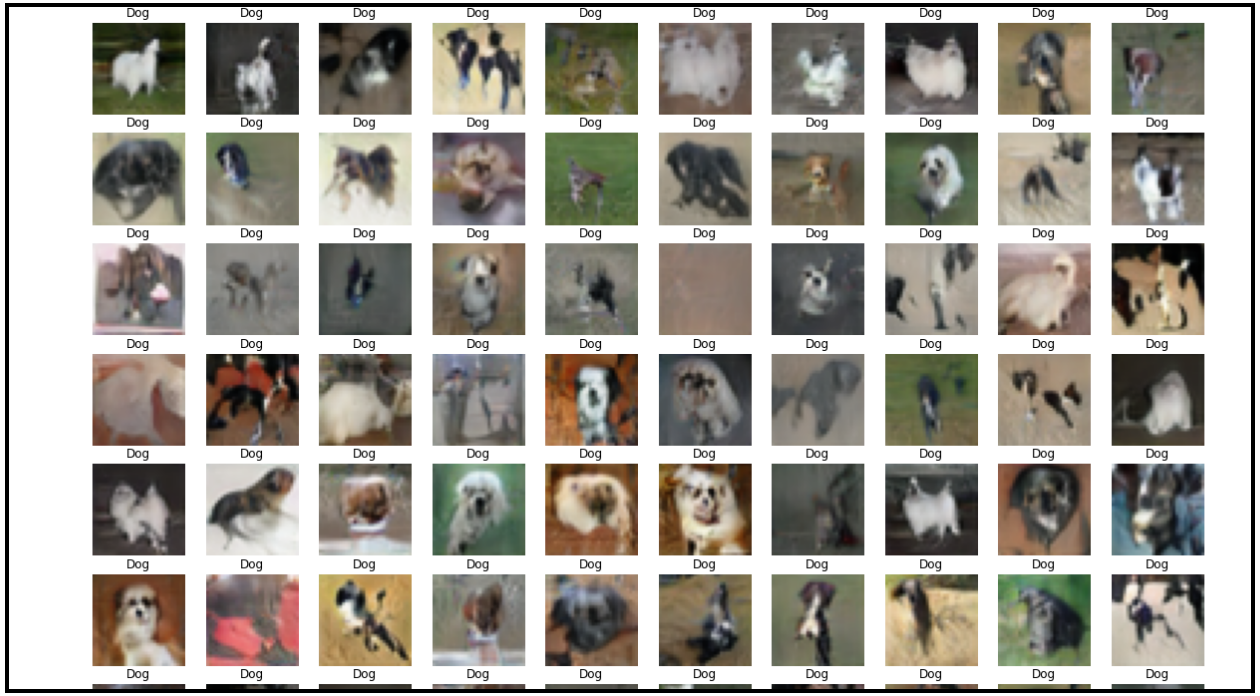

Car

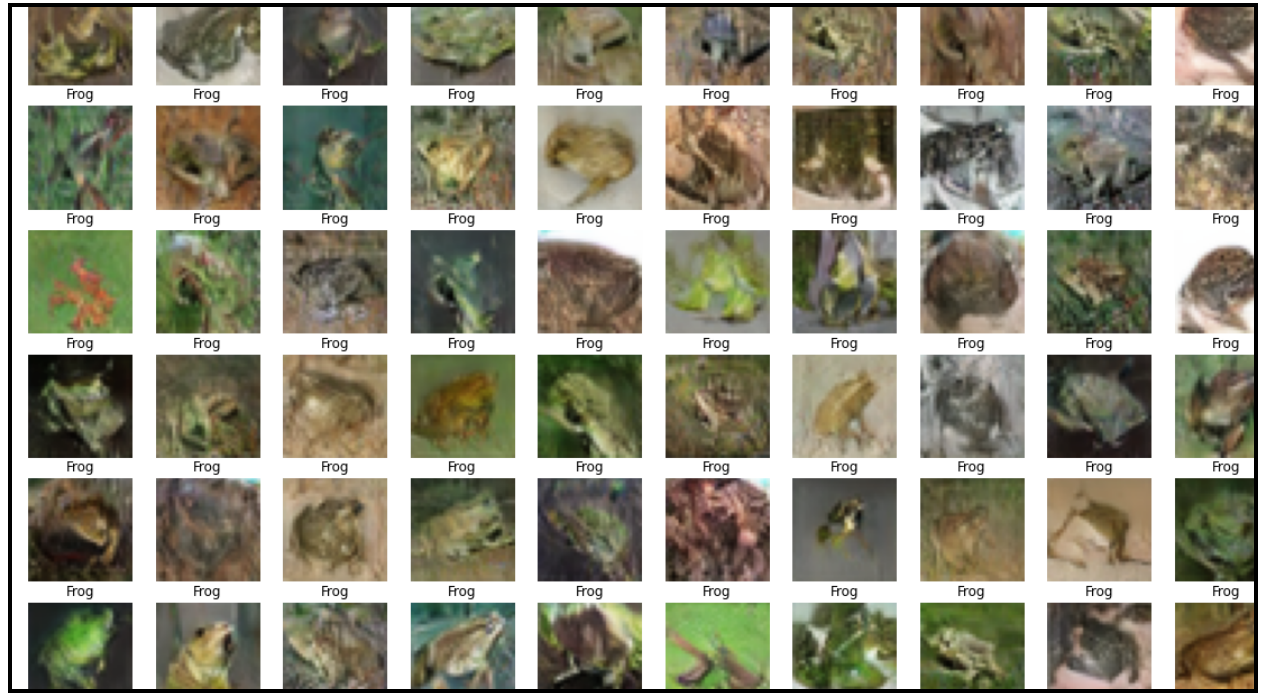

Ship

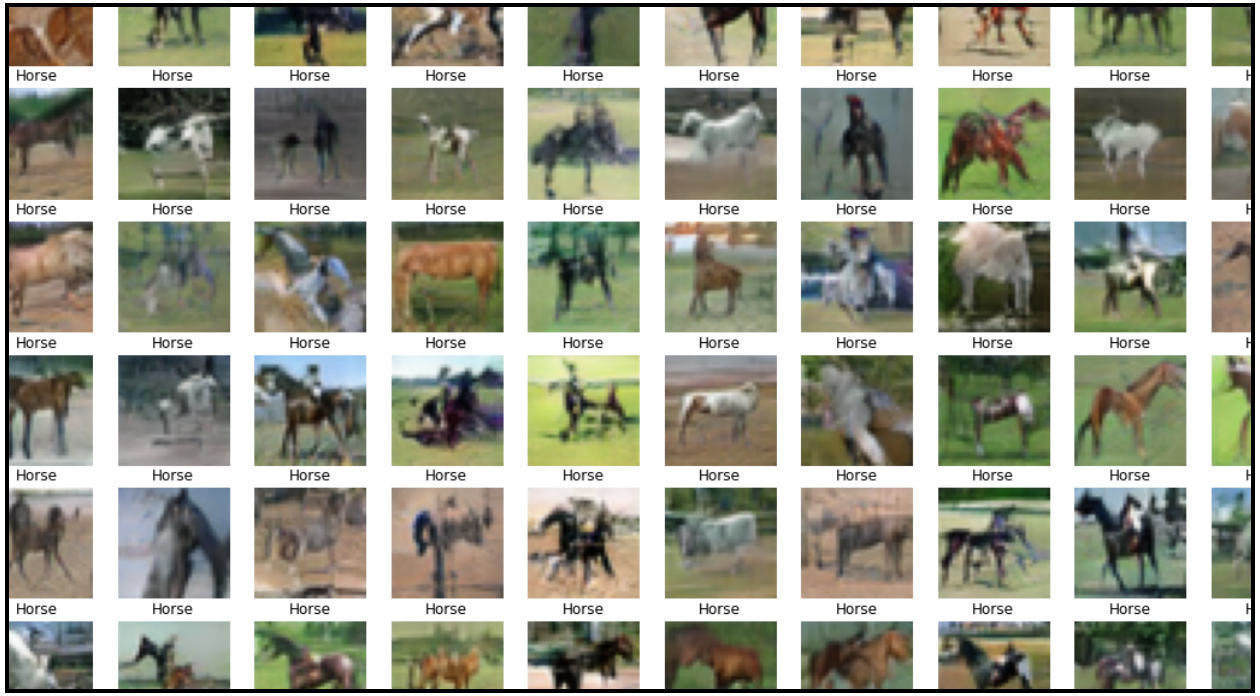

Truck


Cat


Bird


Dog


Frog


Horse


Deer