GANS FOR GENERATIVE ART

Laura Miller

OVERVIEW

Discriminative modeling has driven most advancements in deep learning to date, but generative modeling is thought to be the next frontier

- Discriminative modeling
  - Estimates p(y|x): the probability of a label y given observations x (classification, regression, etc.)
  - Learns how to value data
  - Example: classifying images of paintings made by Van Gogh versus other impressionist artists
- Generative modeling
  - Estimates p(x) the probability of observing observation x
  - Learns how the data was created and can generate synthetic data
  - Example: generating new impressionist paintings that did not previously exist

OVERVIEW

- The Generative Adversarial Network (GAN) has been the most promising generative model
- GAN has a **latent space**, a compressed representation of what it has learned
  - GAN learns mappings between input vectors (latent codes) and specific outputs
  - Create interpolation videos by traversing the space from one latent code to another
- **Project output:** two Generative Adversarial Networks (StyleGAN2-ADA) with latent spaces that can be used to create novel videos for promotional or artistic purposes
  - **malliGAN:** generates new instances of food images
  - **grappleGAN**: generates new instances of people grappling (jiu jitsu)

OUTLINE

“What I cannot create, I do not understand.”

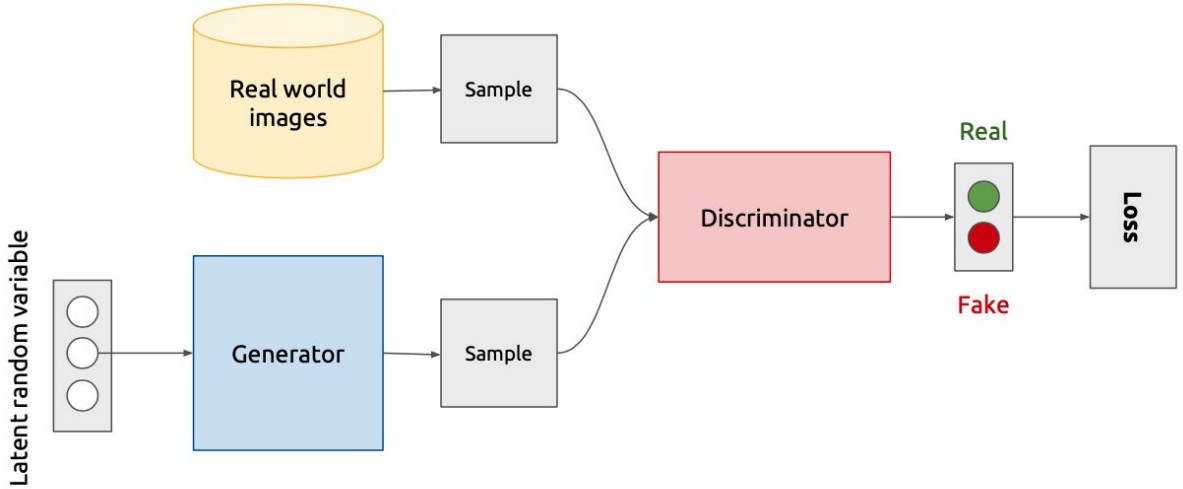
—RICHARD FEYNMAN

BUSINESS APPLICATION

- The learned latent space of styleGAN2-ADA can be used to generate digital art from limited amounts of data
  - malliGAN
    - Promotional videos for Malli, a Los Angeles based pop-up restaurant
  - grappleGAN
    - Promotional videos for a jiu jitsu tournament, or themed artworks to sell as NFTs

BACKGROUND

Generative adversarial networks (conceptual)



BACKGROUND

- GANs were first proposed by Ian Goodfellow et. al. in 2014
- DCGAN (2015)
  - Uses CNN architecture for G and D
- StyleGAN (2018)
  - Generates high resolution images with style information at each layer
- StyleGAN2 (2019)
  - Removes droplet artifacts and other irregularities
- StyleGAN2-ADA (2020)
  - Dynamically changes the amount of augmentations as needed
  - Ideal for smaller datasets

DATA

malliGAN

392 images of food dishes prepared by Malli

- Provided directly by business owner

grappleGAN

2,147 images of two people actively grappling

- Scraped from Instagram and Flickr

METHODS

01

PREPARE DATA

Gather images and convert to square, 1024x1024, sRGB. Convert to TFRecords

02

MODEL

Train StyleGAN2-ADA architecture for multiple days

03

EVALUATE

Quantitative and qualitative assessments of GAN performance (examine FID and Generator output)

04

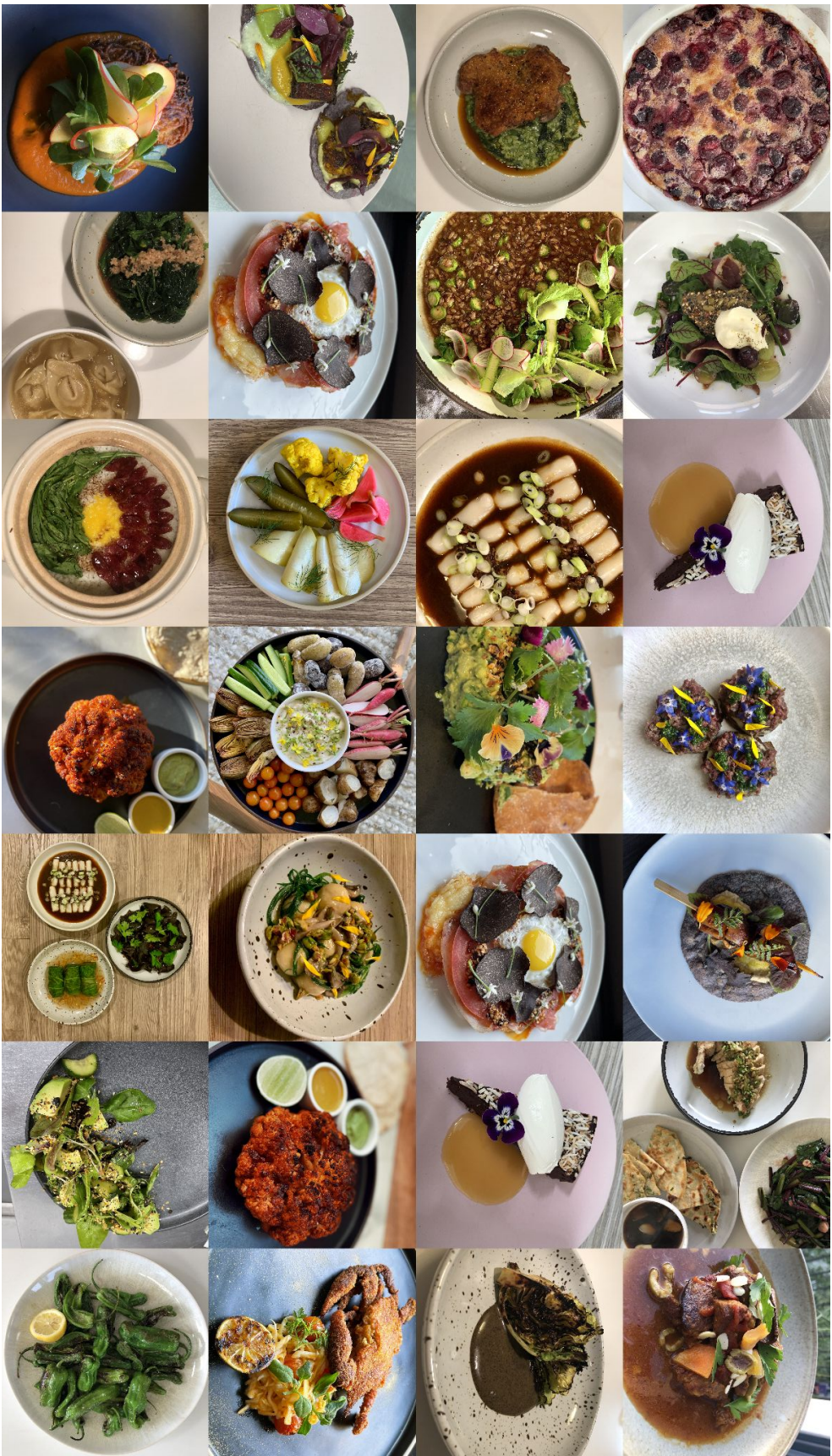
INTERPOLATE

Create cool interpolation videos for NFTs and other marketing purposes

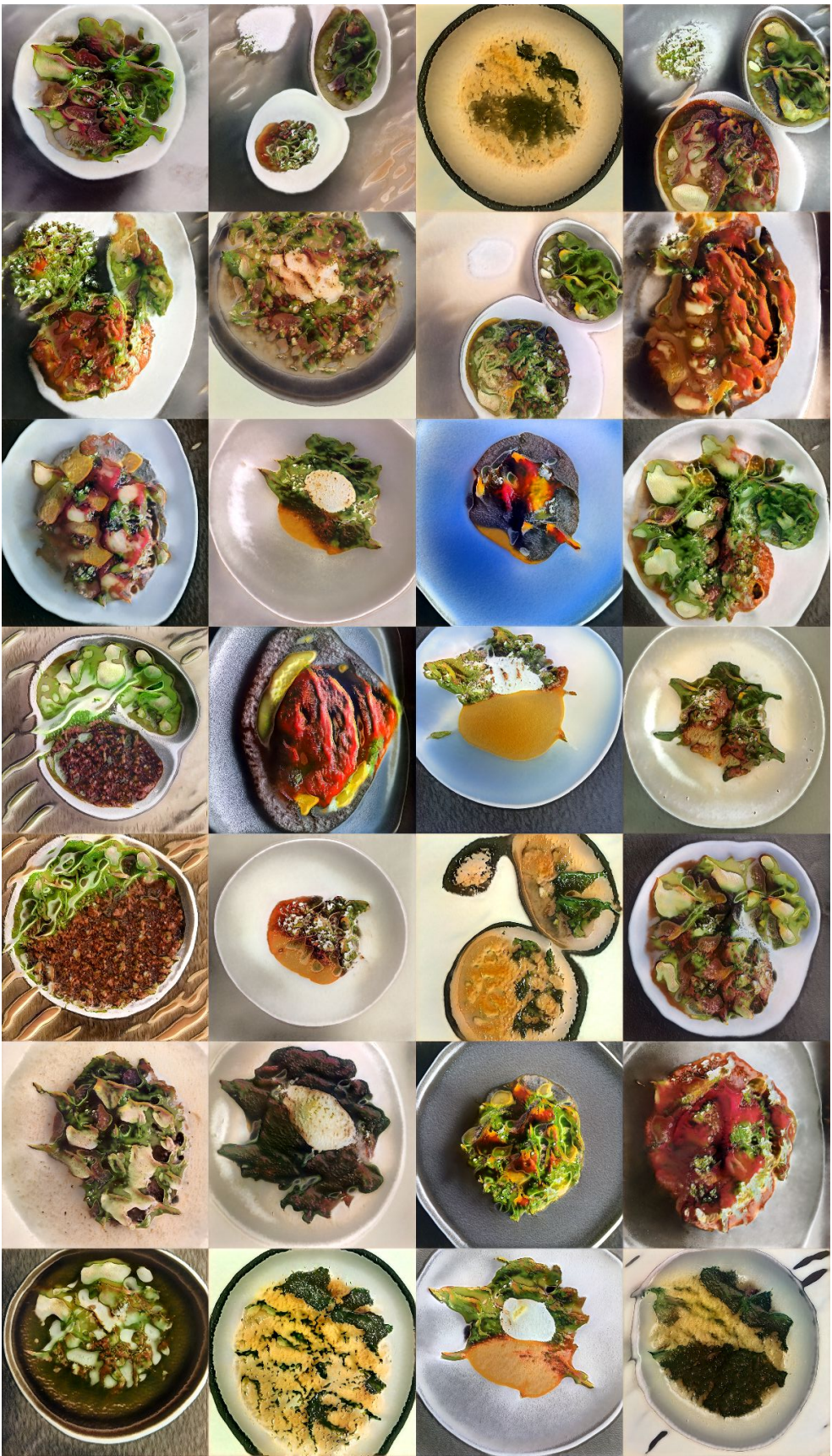
RESULTS

QUANTITATIVE AND QUALITATIVE ASSESSMENT

malliGAN Reals

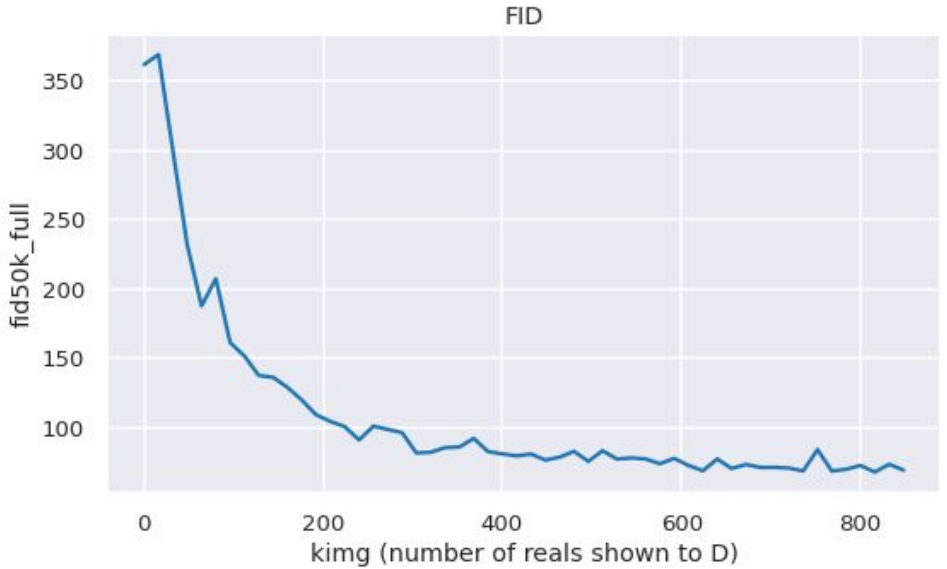


malliGAN Fakes

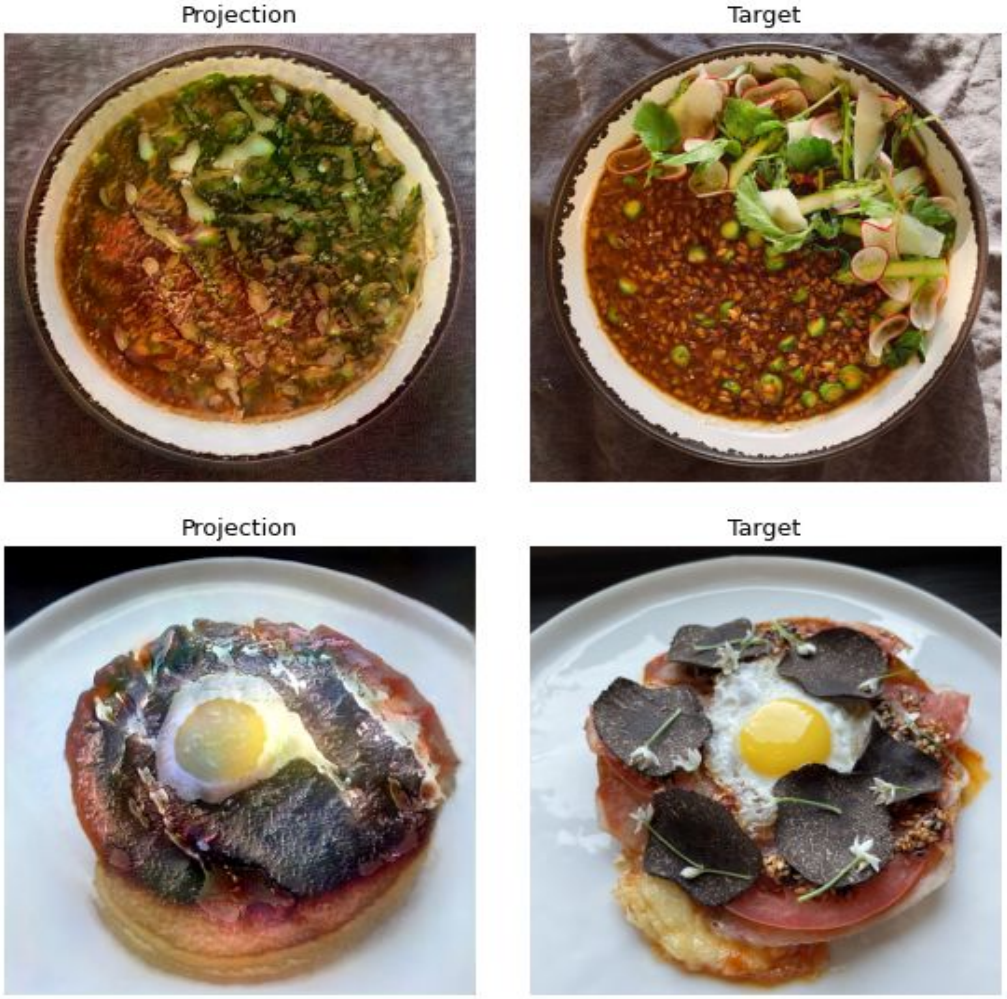


malliGAN FID

- Frechet Inception Distance: the distance between the real and generated distributions (p-data and p-model)
- Begins to plateau around 600 Kimg



malliGAN Projections

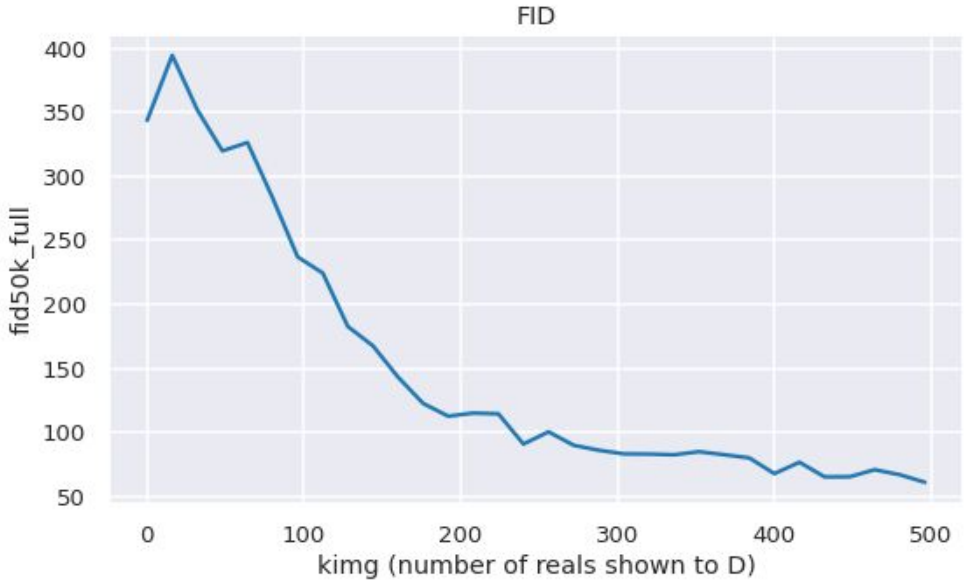


grappleGAN Reals

grappleGAN Fakes

grappleGAN FID

- Is still decreasing slightly by time training was terminated
- Would expect only slight improvements with more training time



grappleGAN Projections

CONCLUSIONS

This project produced two POC GANs to use as part of a digital art practice

- malliGAN: A means to generate promotional interpolation videos, which would become more realistic with more training images
- grappleGAN: A study of how a computer “sees” grappling. Can be used to generate NFTs or other digital artworks

The training datasets should be adequate in size (2K or more samples) and low in diversity for best results

NEXT STEPS

- Set the hyperparameter p to the most recently used value when resuming training (on Google Colab) in order to get stronger augmentations throughout training
- Revisit malliGAN with more data
- More modular data for grappleGAN
- styleGAN3!

THANK YOU!

Does anyone have any questions?

laura@warrahdesign.com
