Fuel or Fool: Decoding Your Food Choices

Pratham Mehta (pratham.mehta@gatech.edu), **Xiangsheng Gu** (xgu304@gatech.edu),
Yu-Ling (Renee) Lu (ylu603@gatech.edu), **Zeezoo Ryu** (zeezooryu@gatech.edu)

Georgia Tech College of Computing

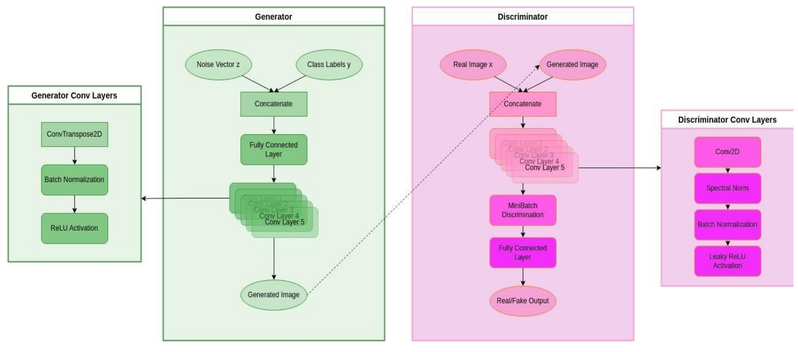
Introduction

- Unhealthy diets high in sugar, sodium, fat, and cholesterol, coupled with lack of exercise, are major causes of chronic diseases like obesity, diabetes, heart disease, and mental health issues.
- We propose using a conditional Generative Adversarial Network (cGAN) to generate personalized healthy and unhealthy food images.
- The goal is to promote healthier eating habits by providing appealing healthy options and discouraging unhealthy choices through visual cues.

Methodology

- Dataset: Food-11 on Kaggle - 16643 food images grouped in 11 major food categories. Manually label them as healthy/unhealthy.
- Conditional GAN (cGAN): both generator and the discriminator are conditioned on the label indicating whether the food is healthy or unhealthy.
- Fine-tuning: use Frechet Inception Distance (FID) as a measure to test out different approaches, including minibatch discrimination, DCGAN, Wasserstein loss, and different image resolutions.

Figure 1. cGAN model architecture



Results

- Showed improvements in model performance using the designed fine-tuning approach.
- Best food images: DCGAN + minibatch discrimination with an image resolution of 128x128 pixels.
- Despite achieving promising results, our model still exhibited some limitations, including mode collapse and white noises.
- Persistent oscillations in the graph (Figure 2) indicate the inherent instability of GAN, needing improved architectures or regularization.

Table 1. Results on different model configurations

Model Config.	Resolution	Epochs	Batch Size	FID
Baseline cGAN + Adversarial Loss	64×64	1000	32	311.53
cGAN + W-Loss	64×64	3000	64	204.18
DCGAN + W-Loss	64×64	10000	32	203.82
DCGAN + MD + W-Loss	64×64	10000	32	202.57
DCGAN + MD (Best Image) + W-Loss	128×128	3000	32	418.12

Figure 2. Test Losses Over Epochs

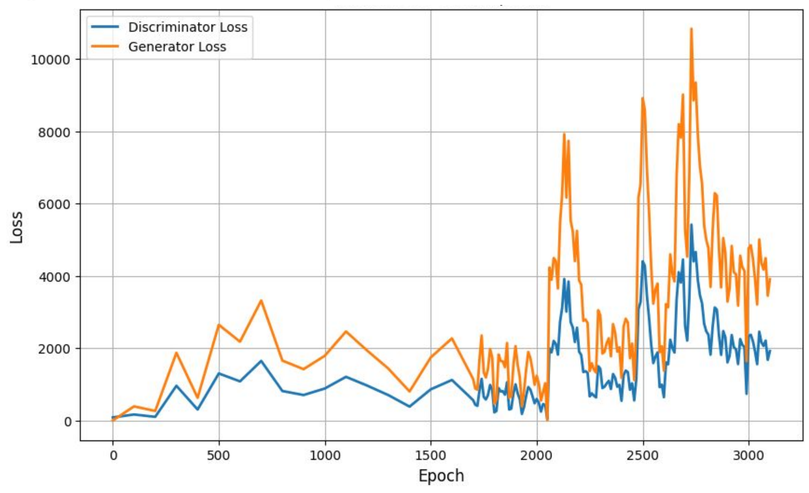


Figure 3. Generated images by the different attempts of the cGAN models

(a) Sample Images from Training Set
(b) 64x64 without Minibatch Discrimination
(c) 64x64 with Minibatch Discrimination
(d) 128x128 Best Generated Image
(e) 128x128 Right before Gradient Issues
(f) 128x128 Image after Gradient Issues

Conclusion & Future Work

- Further fine-tuning on a larger, higher resolution dataset to generate more realistic images.
- Enhancing the performance by exploring alternative architectures and techniques, such as PacGAN or bigGAN frameworks.
- Utilizing a high-performance computing system to train higher resolution images and complex model.
- Experimenting with a diffusion model to reduce mode collapse and unstable training.
- To conclude, while some challenges remain in generating realistic images using the GAN architectures, this work will contributes to future research in deep learning-based food image synthesis that can promote healthier dietary choices.