CSC 2541:
Generative AI for Images

Anthony Bonner
www.cs.toronto.edu/~bonner

https://stability.ai/news/stable-diffusion-sdxl-1-announcement

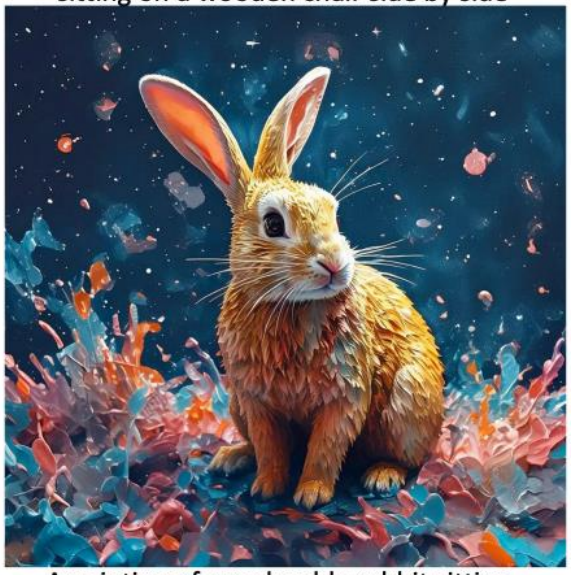
A bread, an apple, and a knife on a table

a robot cooking dinner in the kitchen

A teddy bear and a stuffed raccoon sitting on a wooden chair side by side

A heart made of wood

an old man with green eyes and a long grey beard



A painting of an adorable rabbit sitting on a colorful splash

[Dai et al, "Emu: Enhancing Image Generation Models Using Photogenic Needles in a Haystack", arXiv 2023]

Course Structure

- Seminar course with a major project.
- Study papers from the literature.
- First 2 classes:
  lectures on background material.
- Next 8 classes:
  student presentations of papers.
- Last 2 classes:
  project presentations

Course Projects

- Anything you like about generative AI for images.
- Must have a major technical component.
- See course website for project ideas.
- May be done individually or in teams of up to 3.
- More will be expected from team projects.
- Make sure you have the computational resources to do the project.

Course Projects

- Project proposals (2-4 pages) due October 15
- Project presentations (4 minutes per team member) in class November 22 and 29.
- Project reports (4-8 pages) tentatively due on December 13.
- All team members receive the same grade.
- See course website for details.

Paper Presentations

- Each week will focus on one or two topics, as listed on the course web page.
- You can vote for your choice of topic/week (soon).
- I will assign you to a week (soon).
- Papers on each topic are/will be listed on the course web page.
- If you have a particular paper you would like to add to the list, please let me know.

Tentative Weekly Topics

1. Foundations and Background (Sept 20)
2. Autoregressive Models (Sept 27)
3. Generative Transformer Models (Oct 4)
4. Diffusion Models and Score Matching
5. Conditional Diffusion Models and Image Editing
6. Latent Diffusion models and Accelerated Sampling
7. Transformer Architectures for Diffusion Models
8. Neural Differential Equations

Paper Presentations

- Goal: high quality, accessible tutorials.
- 8 weeks and 60 students = 7 or 8 students per week and 13 minutes per student (including questions and transition).
- 2-week planning cycle:
  - 2 weeks before your presentation, meet me after class to discuss and assign papers.
  - The following week, meet me or the TA online for a practice presentation (required).
  - Present in class under strict time constraints.

Team Presentations

- Papers may be presented in teams of two or more with longer presentations (13 minutes per team member).
- Unless a paper is particularly difficult or long, a team will be expected to cover a group of related papers (one paper per team member).
- A team may cover one of the listed papers and one or more of its references (but see me first).

Volunteers Needed for Sept 27

- We need 7 or 8 students to volunteer to present papers on Sept 27, the second week of presentations.
- The papers will cover autoregressive methods.
- They are relatively non-mathematical.
- This is a good time to present, before the semester gets too busy.
- If you are interested, please send me 2 or 3 paper choices ASAP.

Suggested Papers for Presentation

- Neural Discrete Representation Learning Describes VQ-VAE, a variational autoencoder with discrete latent variables.
- Conditional Image Generation with PixelCNN Decoders Image generation with input prompts.
- PixelCNN++: Improving the PixelCNN with Discretized Logistic Mixture Likelihood and Other Modifications Describes an improved implementation of PixelCNN
- WaveNet: A Generative Model for Raw Audio Develops techniques for audio that are frequently applied to images.
- Hierarchical image generation:
  - Generating Diverse High-Fidelity Images with VQ-VAE-2
  - PixelCNN Models with Auxiliary Variables for Natural Image Modeling
  - Generating High Fidelity Images with Subscale Pixel Networks and Multidimensional Upscaling
  - Hierarchical Autoregressive Image Models with Auxiliary Decoders