CSC 2541:
Generative AI for Images

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

Generative AI

- Recent dramatic advances in AI techniques for generating text, speech, music, images, video, artistic styles, textures, 3D structures …
- This course will focus on images
- Long history of AI techniques for generating images:
  - Variational Autoencoders (VAEs)
  - Normalizing flows
  - Autoregressive methods
  - Generative Adversarial Networks (GANs)
- This course will cover some of this foundational material, but will focus on recent advances

Recent Advances in Image Generation

- Diffusion Models
- Latent Diffusion Models
- Conditional Diffusion Models
- Score Matching
- Accelerated sampling
- Image editing
- Neural Differential Equations
- Generative Transformer models

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

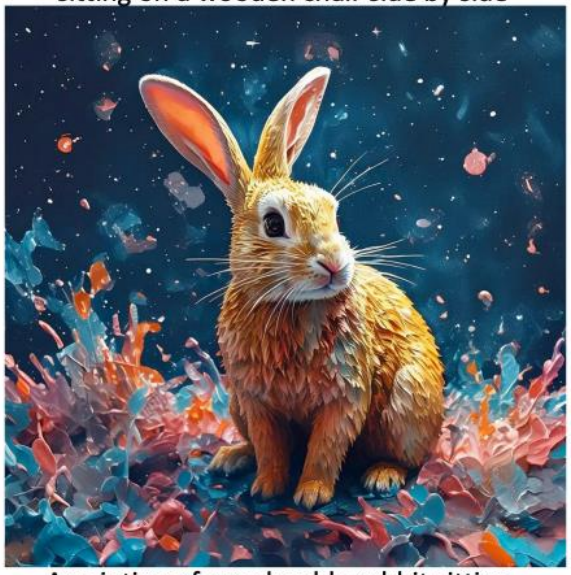
A bread, an apple, and a knife on a table

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

Paper Presentations

- Each week will focus on one or two topics, as listed on the course web page (soon).
- You can vote for your choice of topic/week (soon).
- I will assign you to a week (soon).
- Papers on each topic 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.

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).
- 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 Presentatations

- 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).

Marking Scheme

- Paper presentation: 20%
- Course project Proposal: 20%
- Project presentation: 20%
- Project report and code: 40%

Prerequisites

- Solid introduction to deep learning (eg, csc 413/2516)
- Solid familiarity with neural nets and CNNs
- Solid background in linear algebra
- Multivariate calculus and probability
- Differential equations would be helpful
- Programming skills (eg, Tensorflow or Pytorch if you plan an implementation project)
- Mathematical maturity will be assumed.

More information

- See the course website.
- Accessible through my home page:
  www.cs.toronto.edu/~bonner
- Announcements will be made through Quercus and Piazza.

Volunteers Needed for Sept 20

- We need 7 or 8 students to volunteer to present papers on Sept 20, the first week of presentations.
- The papers will cover foundational methods.
- Advantages of being first:
  - More support.
  - No overlap with course project or project proposal deadlines.
- If you are interested, please send me 2 or 3 paper choices ASAP.

Suggested Papers for Presentation

- **Pixel Recurrent Neural Networks** Autoregressive image generation
- **Attention is all you need** The original paper on Transformers
- **An Image is Worth 16x16 Words** The original paper on Vision Transformers
- **NICE: Non-Linear Independent Components Estimation** Flow-based image generation
- **Variational Inference with Normalizing Flows** The paper that coined the term "Normalizing Flow"
- **Deep Conditional Generative Models** Using Variational Autoencoders to generate images subject to conditions
- **Generative Adversarial Networks** The original paper on GANs
- **Conditional Adversarial Networks** Using GANs to generate images subject to conditions
- **Deep Residual Learning for Image Recognition** The original paper on ResNets.

Other papers may be added soon. Feel free to suggest more.