CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

CSDS 600: Deep Generative Models

Yu Yin (yu.yin@case.edu)
Case Western Reserve University

https://yin-yu.github.io/

Outline

- Introductions
  - Who am I?
  - Who are you? Major, background, expectation et al
- Course Overview
- What are Generative Models?

CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

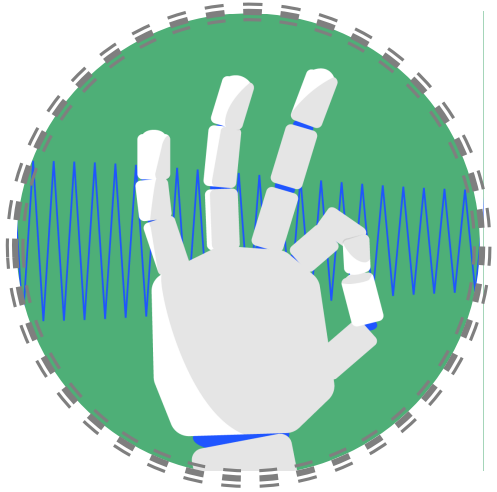
Introduction

- Who am I?
  - Yu Yin (yu.yin@case.edu), AP at Case
  - Received PhD at Northeastern University
  - Teaching instructor at Northeastern University
  - Research interests: Computer vision, Machine learning, Data mining

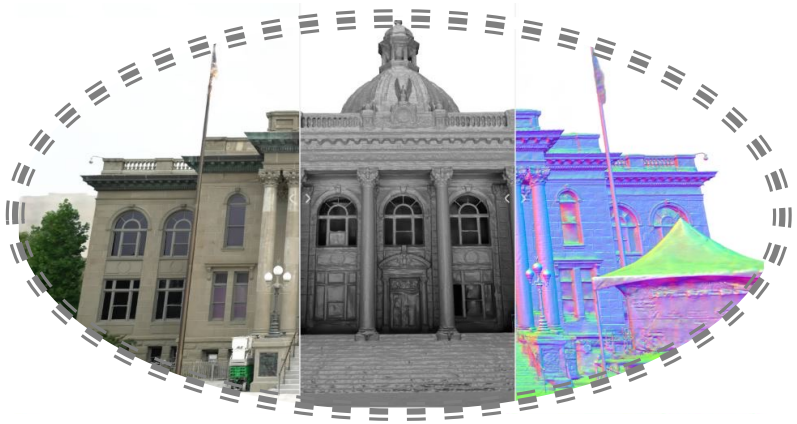
CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

More Specific Research Interests

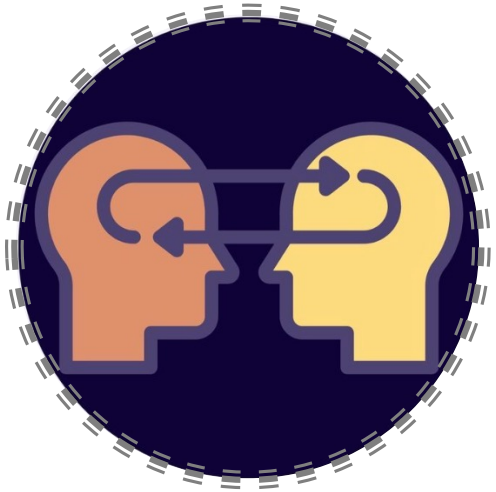
Generative AI



Multi-modal Learning



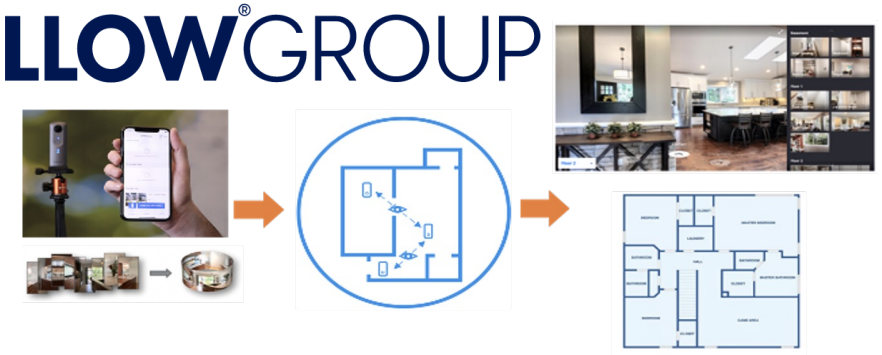
3D Vision



Transfer Learning

More research experiences

- **STR** (2019 -2023)
  - **DoD funded project for LLFR**: Proposal Main Contributor & Project Lead
  - **Leading data collection**: 200 subjects
  - **Delivery**: System prototype, 1 publication
- **Zillow** (2020)
  - Computer Vision (CV) group
  - Delivery: 1 publication, 1 patent
- **Microsoft** (2021, 2022)
  - Azure Cognitive Services: CV group
  - Delivery: 2 publications

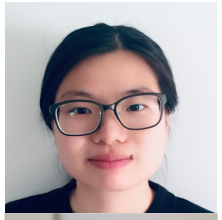


NSF I-Corps

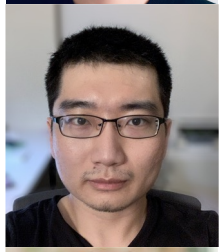
- **NSF Innovation Corps Program award**
  - **Goal:** Multi-cue facial restoration and identification for video surveillance
  - **Entrepreneurial Lead**
  - Interviewed 100 subjects to learn market demands
  - Research contribution:
    - **Patent**:

      Fu, Yun, and Yu Yin. "Frontal Face Synthesis from Low-Resolution Images." U.S. Patent Application No. 17/156,204.
    - **Publication**:

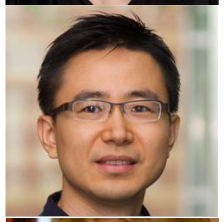
      Yu Yin, et al. "Superfront: From low-resolution to high-resolution frontal face synthesis." ACM MM. 2021.



Yu Yin, EL



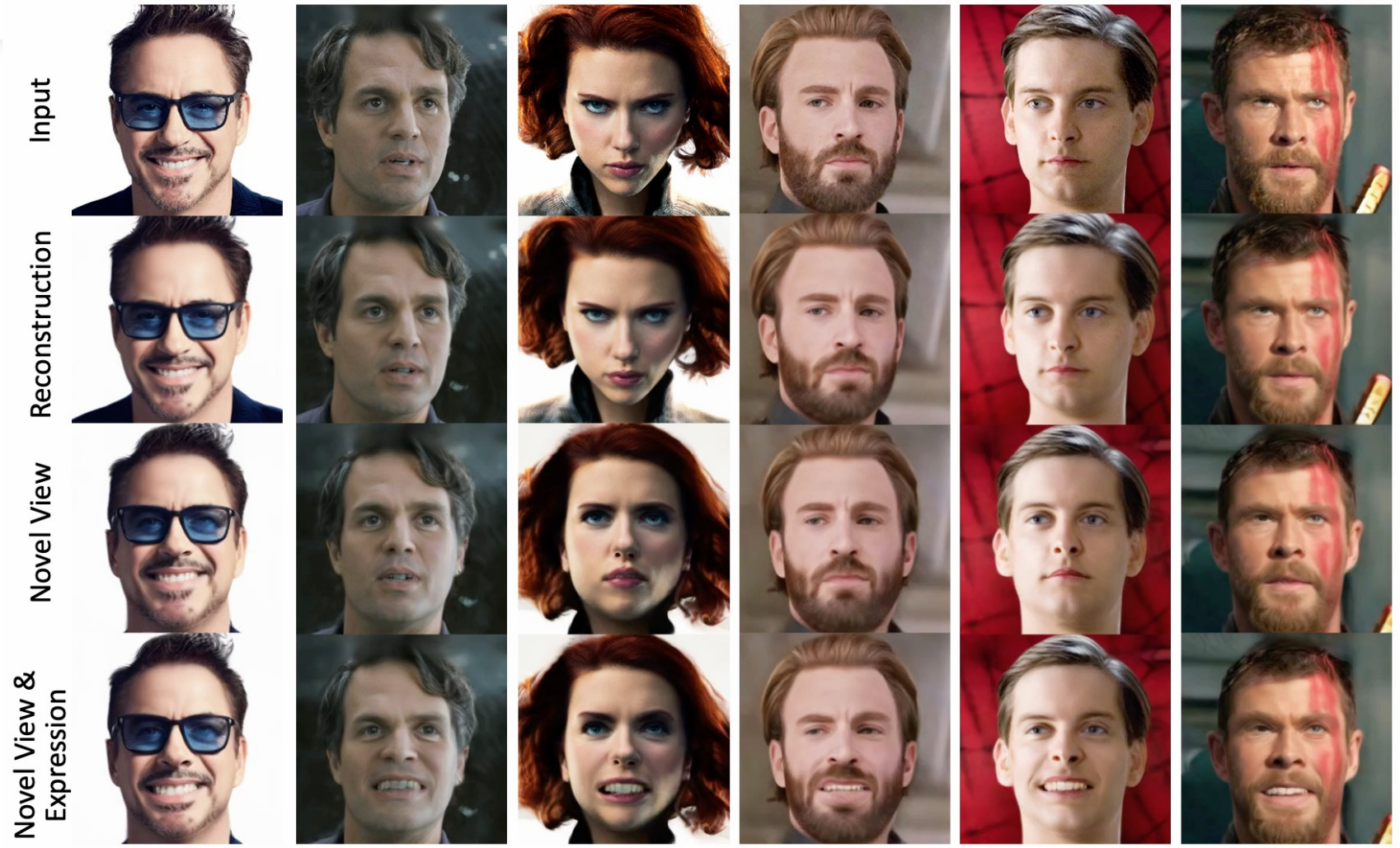
Songyao Jiang, Co-EL



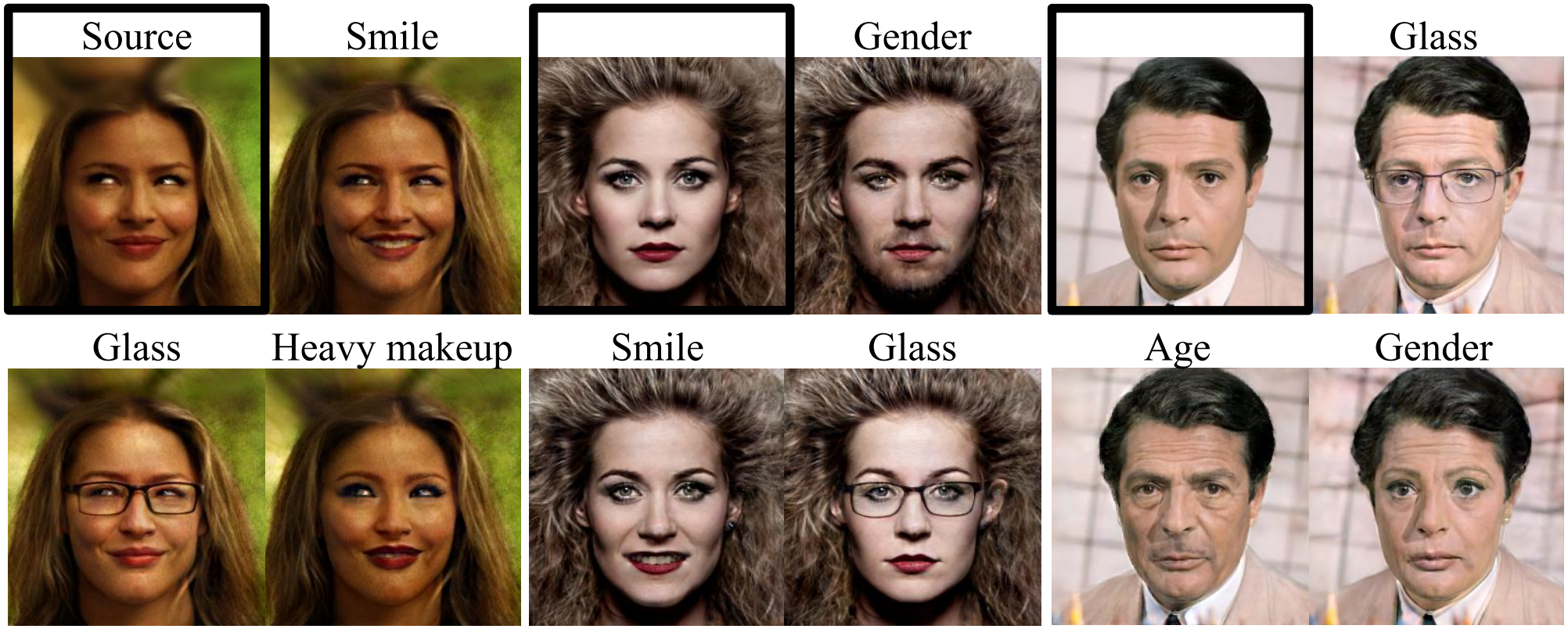
Raymond Fu, PI

Michael Stopa, IM

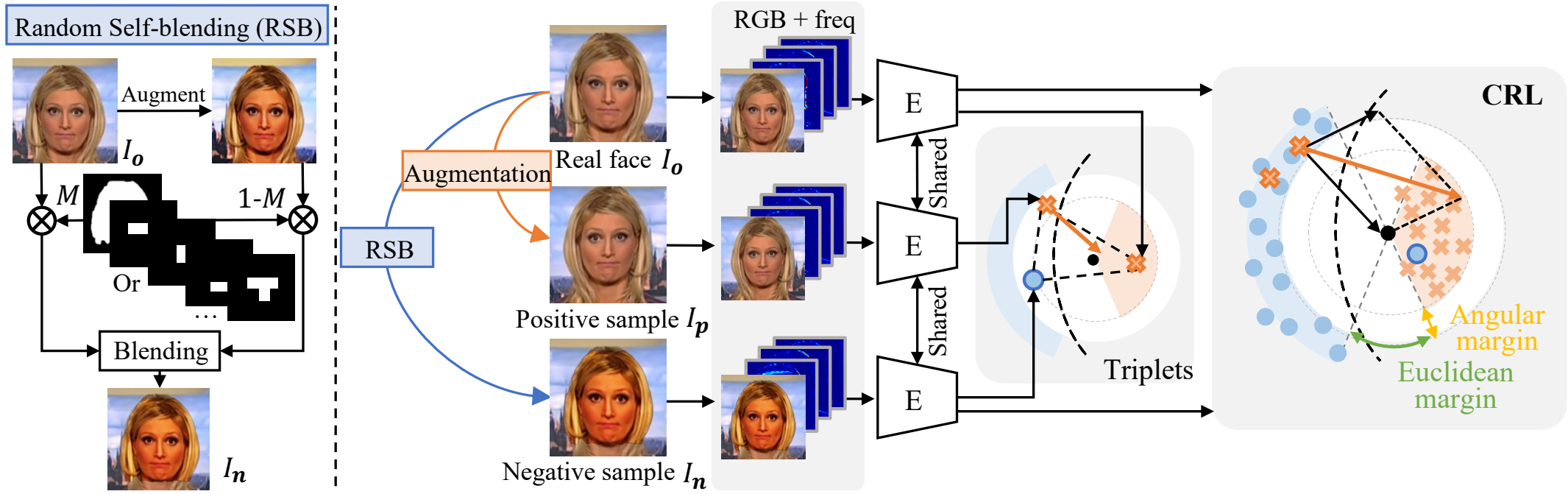
Previous work: Face Animation with single input



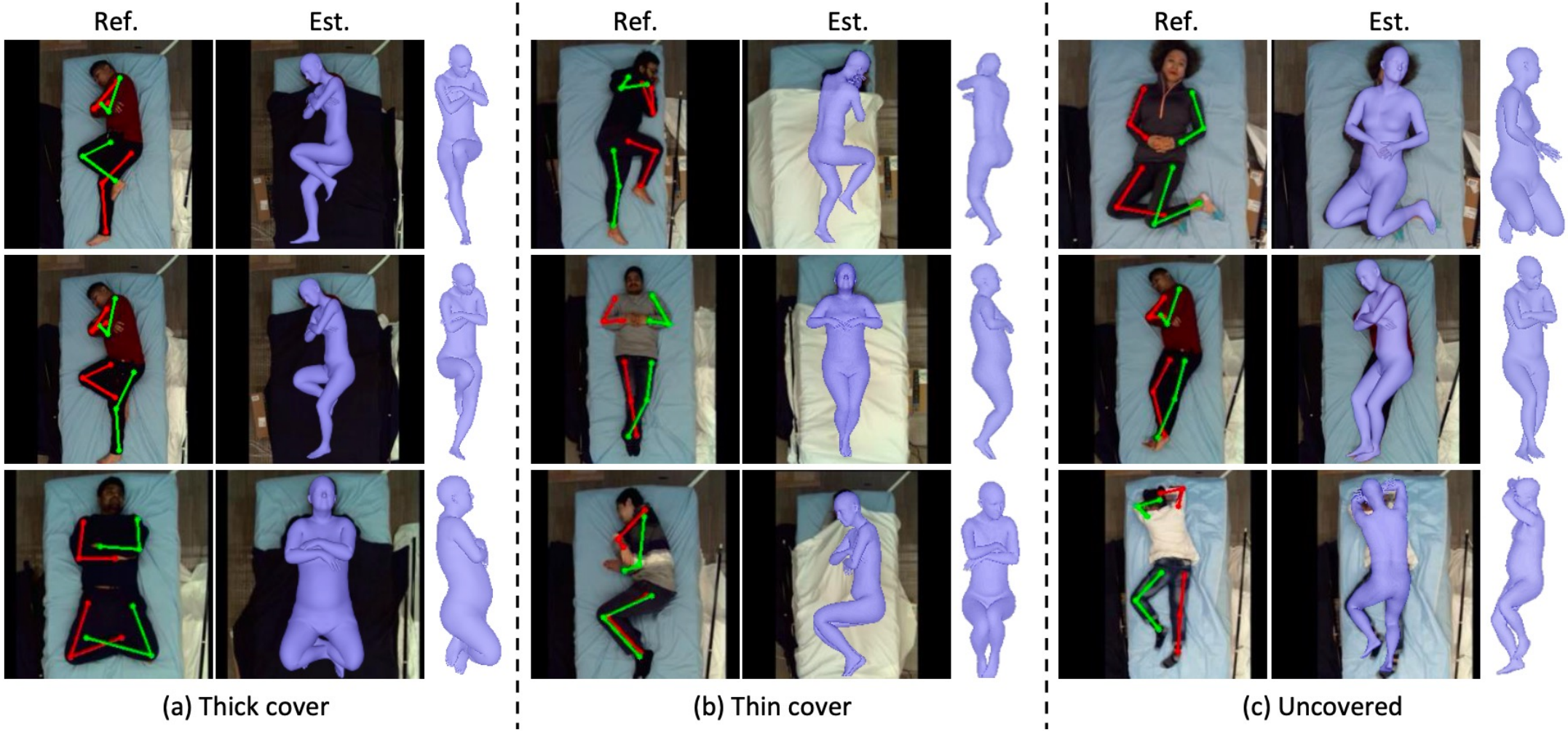
Previous work: Face Editing



Previous work: Deepfake Detection

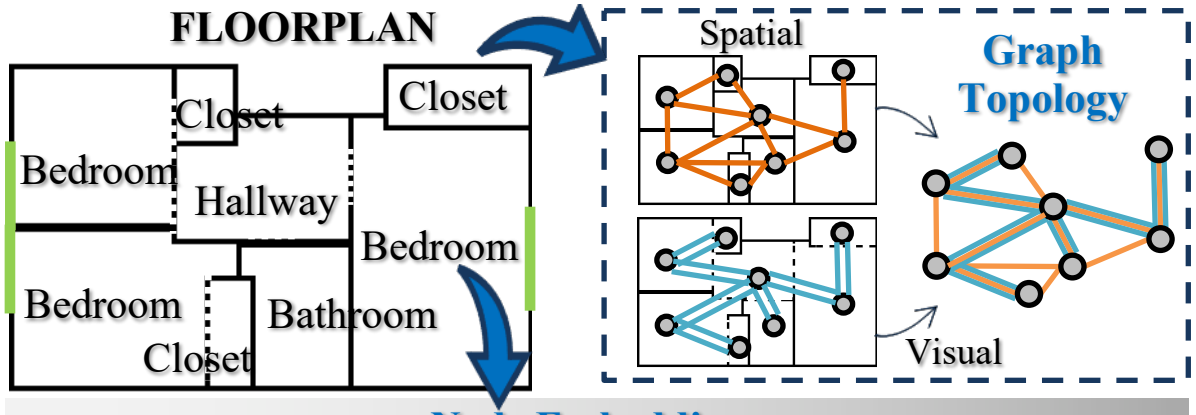


Previous work: In-bed Pose Estimation



Previous work: Generating Topological Structure of Floorplans

Green and *dash lines* in the floorplan refer to windows and doors, respectively.



Outline

- Introductions
  - Who am I?
  - Who are you? Major, background, expectation et al
- Course Overview
- What is Generative Models?

CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

Student Introductions

Outline

- Introductions
  - Who am I?
  - Who are you? Major, background, expectation et al
- Course Overview
- What are Generative Models?

Course Overview

- Lecture: Monday & Wednesday, **5:30 – 6:45 PM**, Olin 314

- Office Hours
  - Location: Olin 606
  - Times:
    - Monday: 6:45-7:45 PM
    - Wednesday: 6:45-7:45 PM
  - Feel free to schedule a meeting at a different time by sending an email.

Course Overview

- Objective: Provide students skills to design, implement, and evaluate deep generative models

- Pre-requisites
  - Basic knowledge about machine learning
  - Proficiency in a programming language (preferably python) and deep learning frameworks (e.g., Pytorch, Tensorflow)

Course Overview

Workload & Grading

- Assignments (30%)
- Mid-Term Exam (20%)
- Paper Presentation and Discussion (20%)
- Final Project and Presentation (30%)

Course Overview

- Assignments (30%)
  - 2 assignments: a mix of conceptual, mathematical, and programming problems related to the material covered in lectures
  - Test the understanding of the algorithms covered in the class

- Mid-Term Exam (20%)
  - The questions would be similar to assignments
  - Take-home exam, no class
  - Exam date and time: Oct. 16

Course Overview

- Paper Presentation (20%)
  - Discuss a recent advance in generative models or perform a literature survey on an area
  - Individual/Group presentations and reports
  - Assessment
    - Importance of the chosen topic
    - Your ability to convince in your presentation that this is an important topic
    - Justice to the topic
    - Identification of gaps or future directions in the topic chosen
    - Presentation Quality
  - Presentations during the classes

Course Overview

- Project and Presentation (30%)
  - A project using one (or more) types of the generative models discussed in the class on a topic of your choice (Consider what computing resources you might need and plan ahead)
  - Projects can be in groups (of 2) or individually.
  - Assessment
    1. Proposal presentation (10%)
    2. Final project presentation (45%) - Evaluated by both me and peer review
       - Quality of Presentation
       - Preliminary Results or Preliminary Demo
    3. Final project report (45%)
       - Professional Academic Writing
       - Reproducibility – I should be able to run your code and reproduce your results
       - Strength of Results
  - Interested students can reach out to me to explore potentials of publishing the results of the projects in a conference or a journal

Course Overview

- Grading Policy
  - Absolute Grading

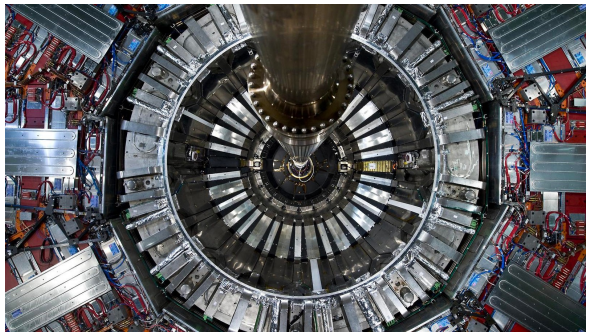
- Syllabus and all other materials will be online through Canvas

Outline

- Introductions
  - Who am I?
  - Who are you? Major, background, expectation et al
- Course Overview
- What are Generative Models?

Age of Big Unlabeled Data

Opportunity: High-throughput data and computation



Particle Accelerators

Materials Project

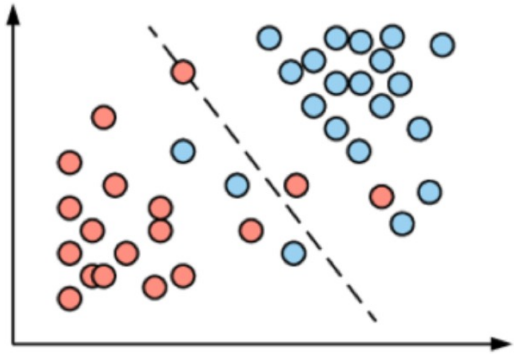
GenBank

Challenge: Supervision signals such as labels are *expensive*

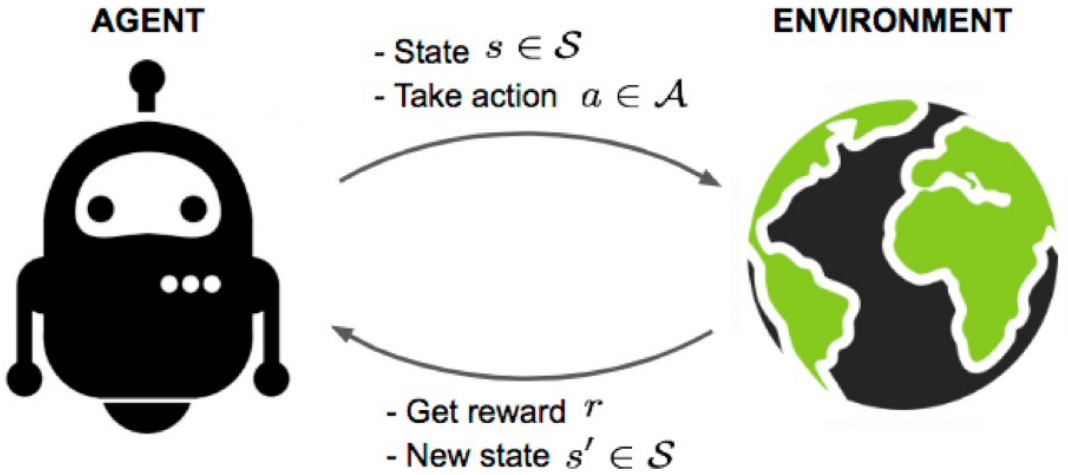
e.g., time, money, expertise, safety costs

https://dl4sci-school.lbl.gov/aditya-grover

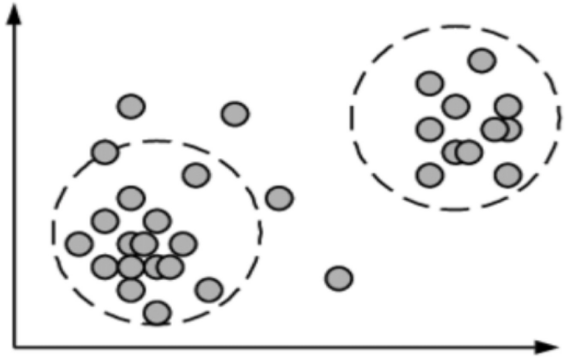
Learning With Limited Supervision



Classification
w/ active querying



Reinforcement Learning
w/ sparse rewards

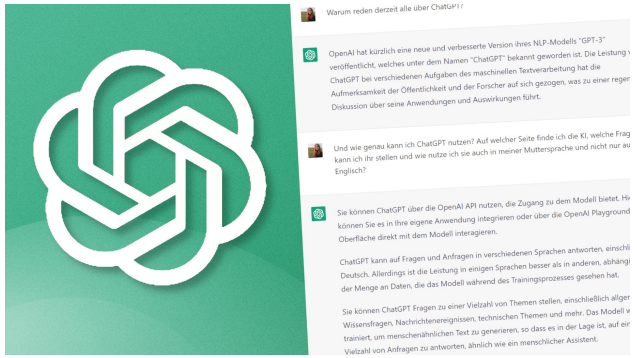


Can we learn with *no* supervision?

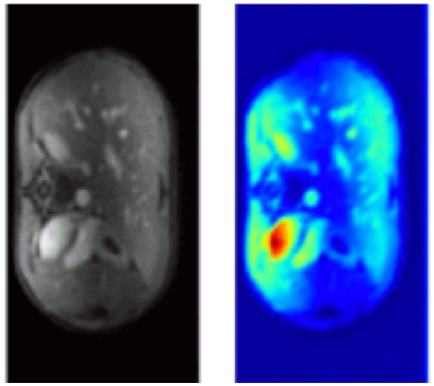
CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

Recent Advances in Unsupervised Learning

Karras et al.



OpenAI



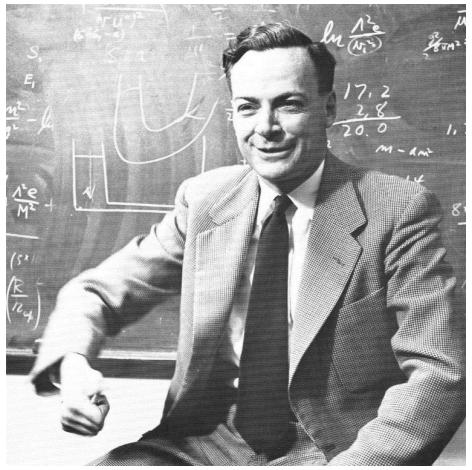
Mardani et al.

Midjourney

Generative Models

"What I cannot create, I do not understand."

---- Richard Feynman



Goal: Generate data samples similar to the ones in the training set

What are Generative Models?

- Learn and replicate the underlying distribution of a given dataset

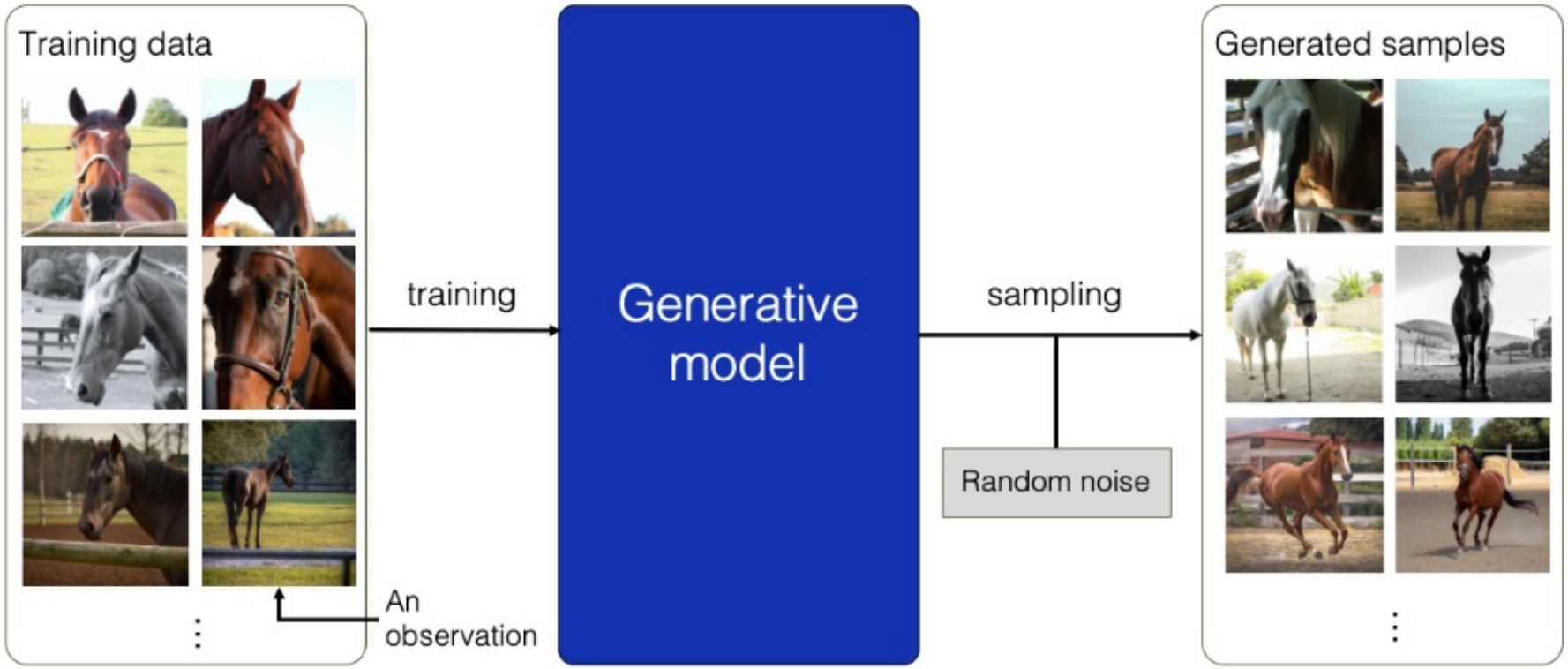


Image from https://vitalflux.com/generative-vs-discriminative-models-examples/

Discriminative vs. Generative

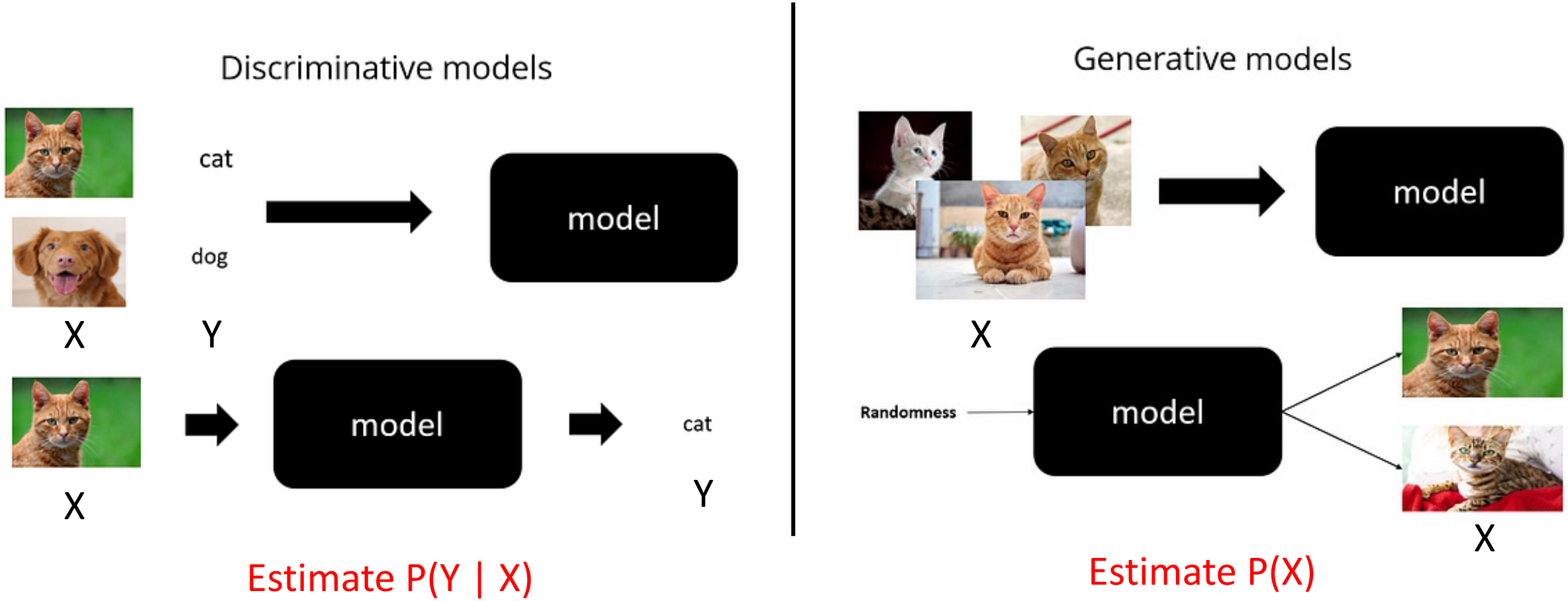


Image from https://medium.com/@amir_shakiba/introduction-to-generative-models-d17bacd3e779

Conditional Generative Models

Class **conditional generative models** are also possible:

P(X=  | Y = Bedroom)

It's often useful to condition on rich side information Y

P(X=  | Caption = "A black table with 6 chairs")

A discriminative model is a very simple conditional generative model of Y:

P(Y = Bedroom | X= )

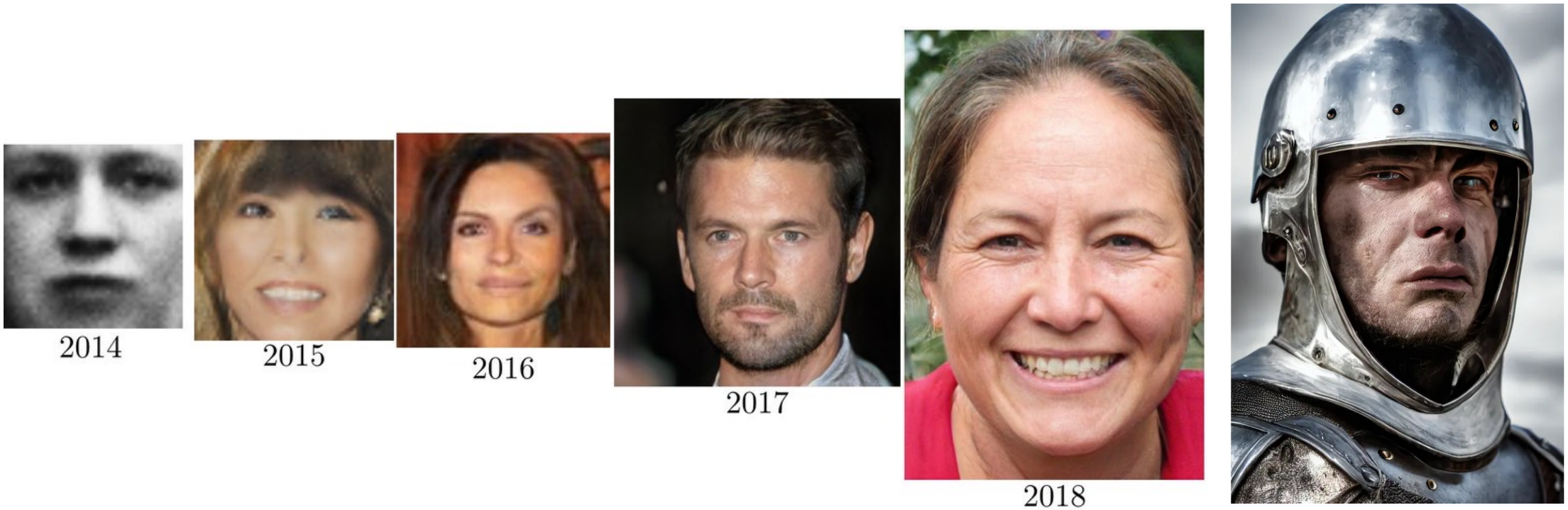
Types of Generative Models

- Autoregressive Models
- Variational Autoencoders (VAEs)
- Normalizing Flows
- Generative Adversarial Networks (GANs)
- Diffusion Models

Other Topics

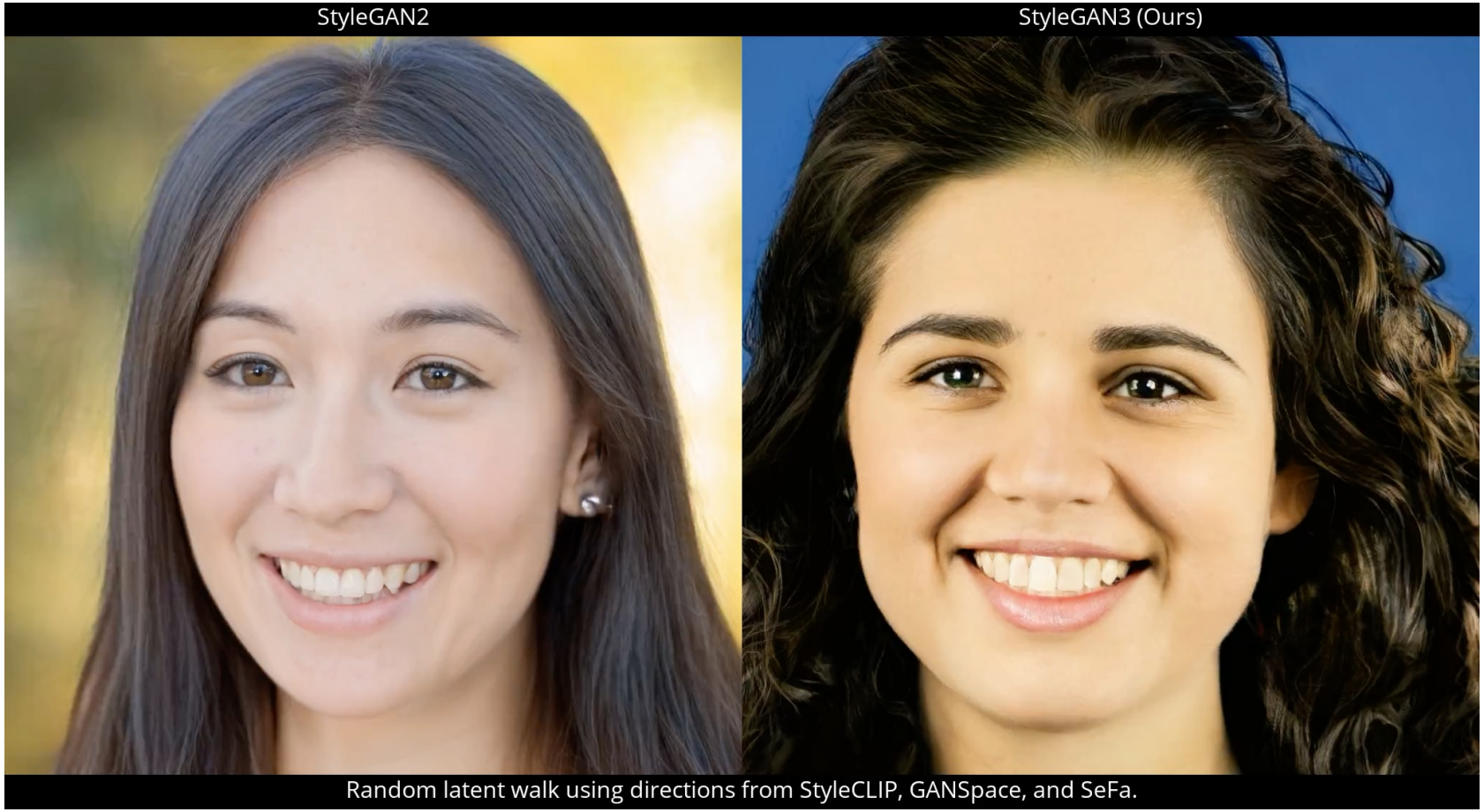
- Inversion issue in Generative Models, especially GANs
- Evaluation of Generative Models
- Generative AI with LLMs
- Image editing with generative models
- Generative models in 3D

Progress in Generative Models of Images



Ian Goodfellow, 2019 & https://stablediffusionweb.com/

Progress in Generative Models of Images



Karras et al., Alias-Free Generative Adversarial Networks, 2021

Progress in Generative Models of Images

"A photo of a Shiba Inu dog with a backpack riding a bike. It is wearing sunglasses and a beach hat." using **Imagen**

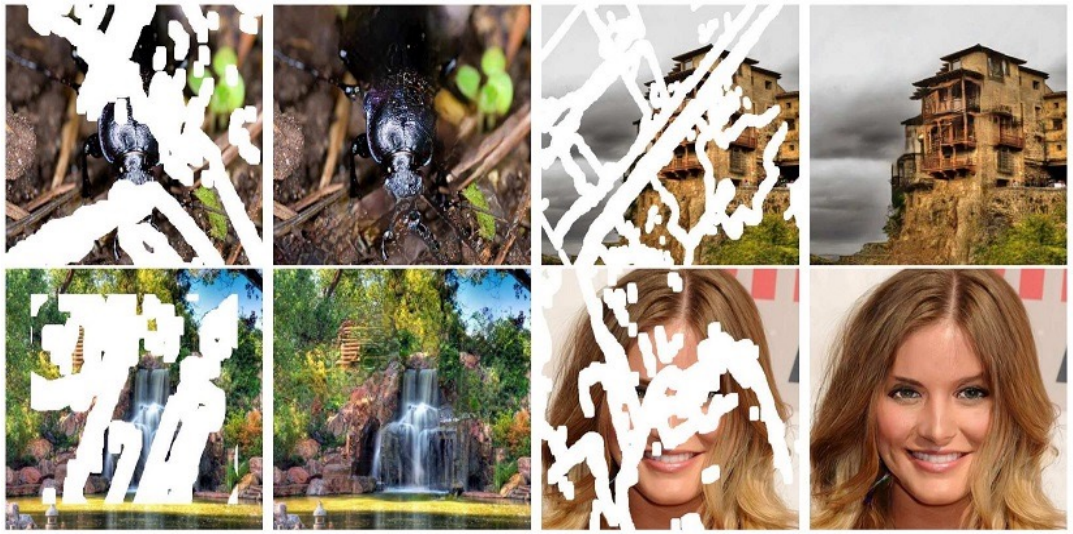
"Avocado chair" using **Midjourney** version v4 – By Jack Gallagher

Progress in Inverse Problems

P(high resolution | low resolution)

Menon et al, 2020

P(full image| mask)

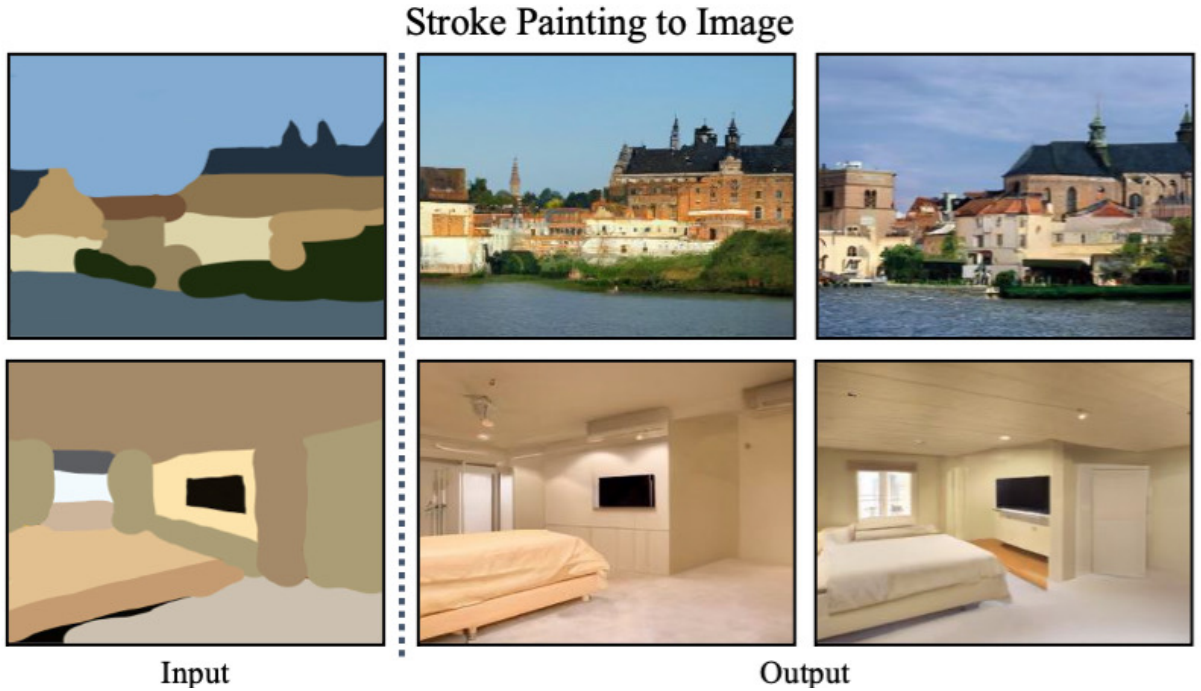


Liu al, 2018

P(color image| greyscale)

Antic, 2020

Progress in Inverse Problems



Source image
(for canny edge detection)

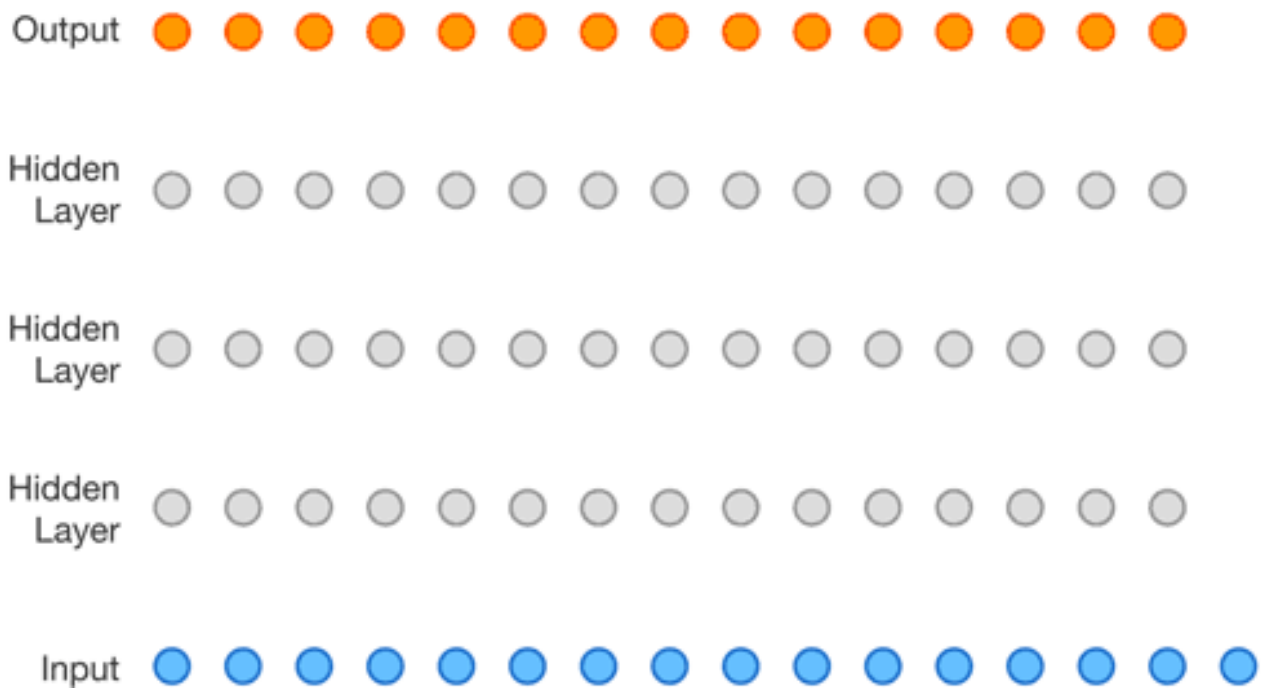
Canny edge (input)

Generated images (output)

ControlNet

WaveNet

Generative model of speech signals



Text to Speech

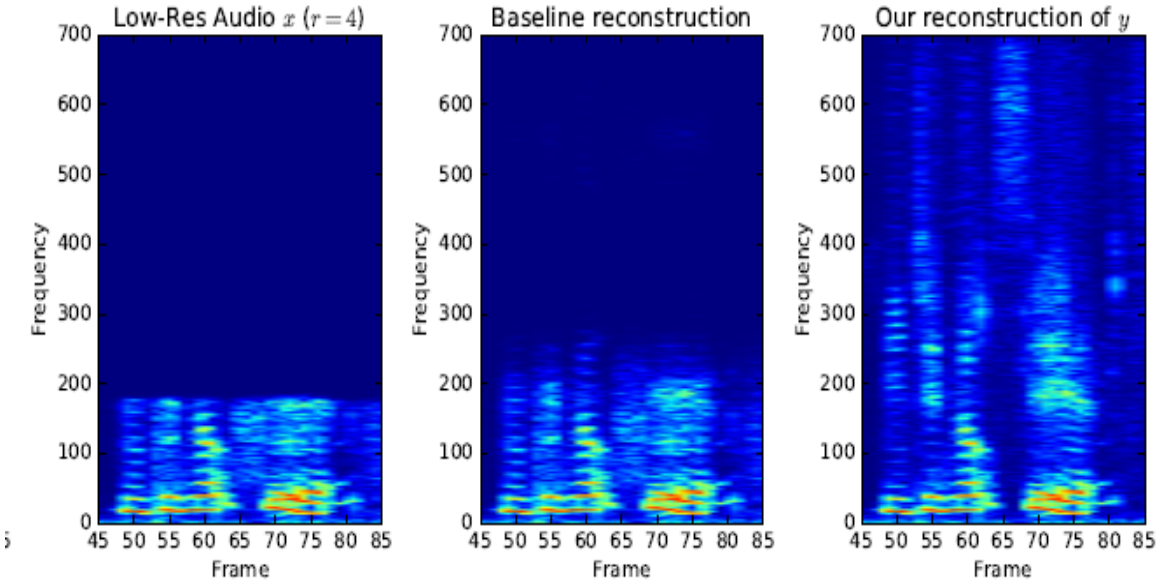
- Parametric
- Concatenative
- WaveNet

- Unconditional
- Music

van den Oord et al, 2016c

Audio Super Resolution

Conditional generative model P(high-res signal | low-res audio signal)



Low res signal

High res audio signal

Kuleshov et al., 2017

Machine Translation

Conditional generative model P(English text| Chinese text)

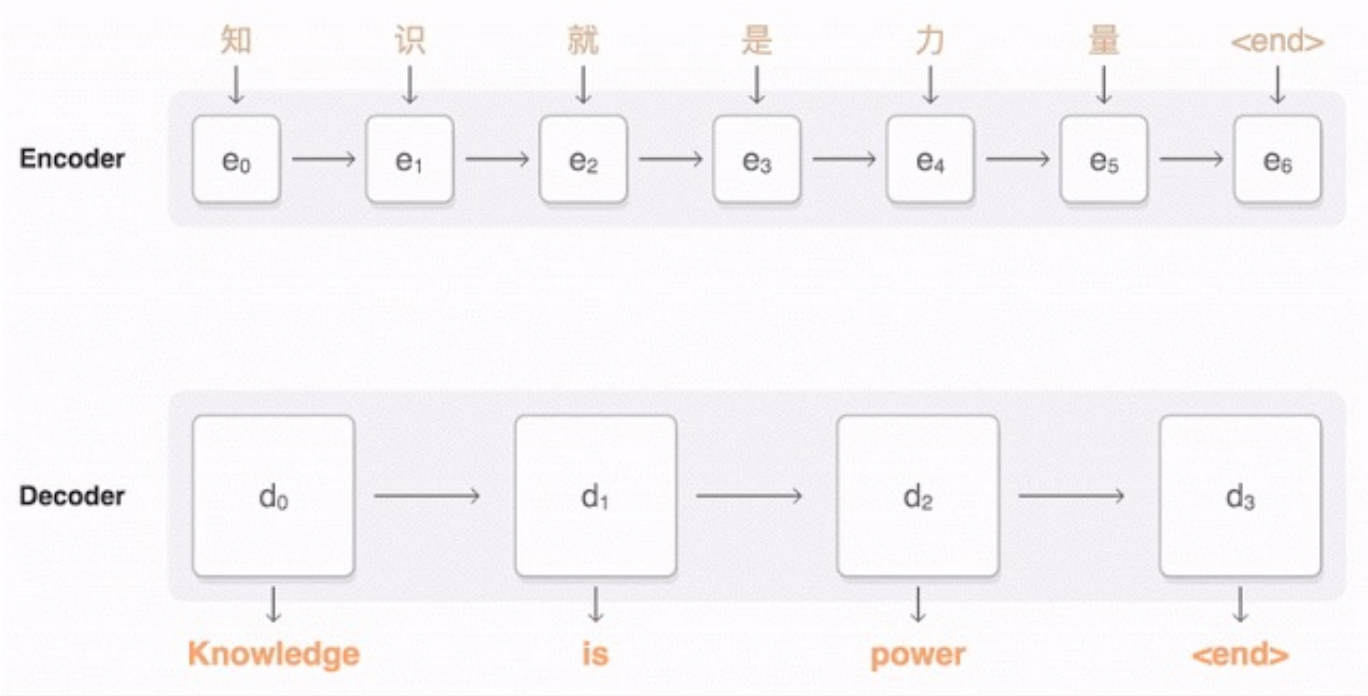


Figure from Google AI research blog.

Language Generation

Custom prompt

To get an A+ in deep generative models, students have to

Completion

To get an A+ in deep generative models, students have to be willing to work with problems that are a whole lot more interesting than, say, the ones that most students work on in class. If you're a great student, the question above can be avoided and you'll be able to do great work, but if you're not, you will need to go beyond the basics before getting good.

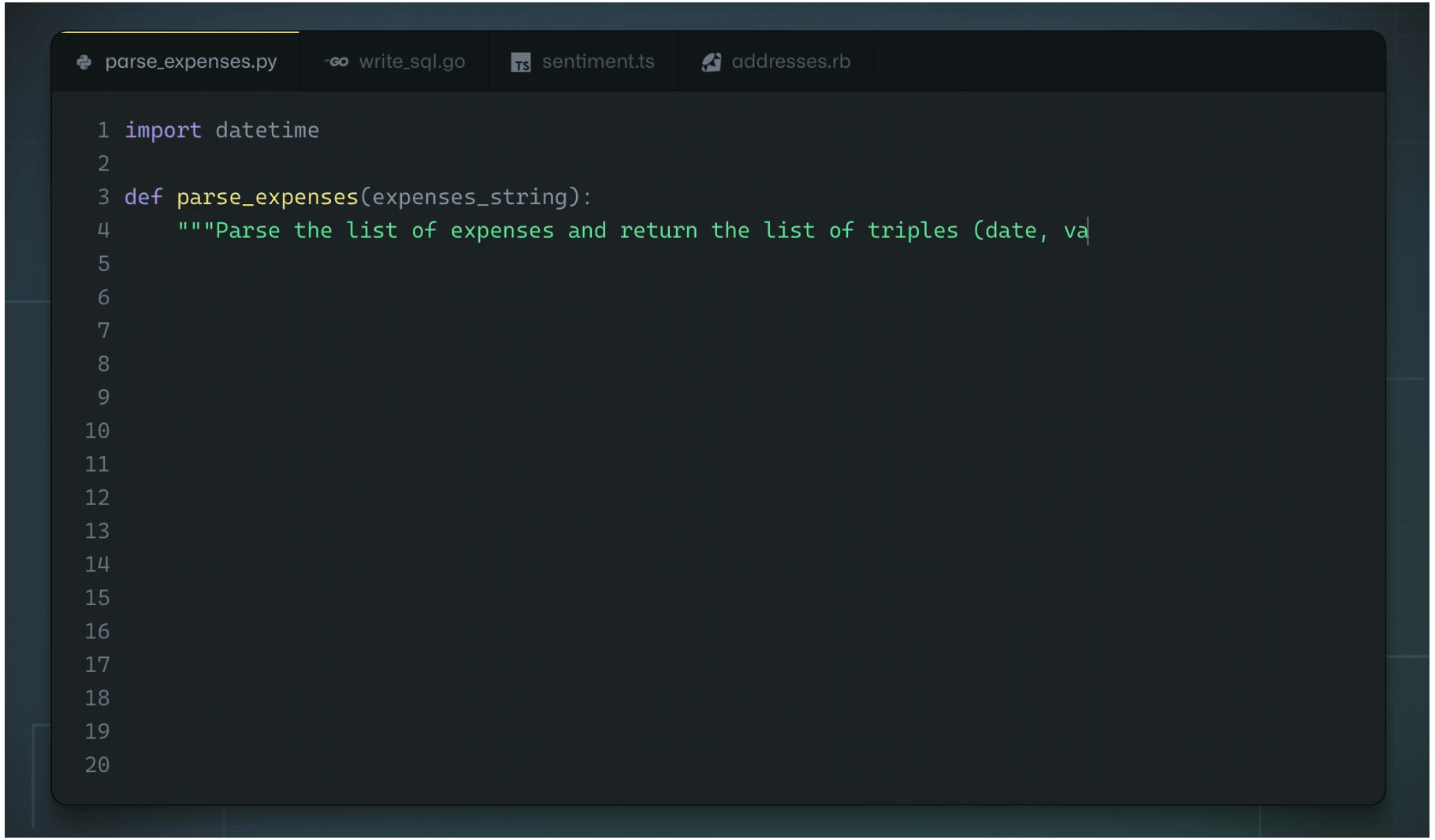
Now to be clear, this advice is not just for the deep-learning crowd; it is good advice for any student who is taking his or her first course in machine learning.

The key point is that if you have a deep, deep brain of a computer scientist, that's just as important to you.

P(next word | previous words)

Radford et al., 2019
Demo from talktotransformer.com

Code Generation



OpenAI Codex

Images and Text

TEXT PROMPT an armchair in the shape of an avocado. . . .

AI-GENERATED IMAGES

Edit prompt or view more images↓

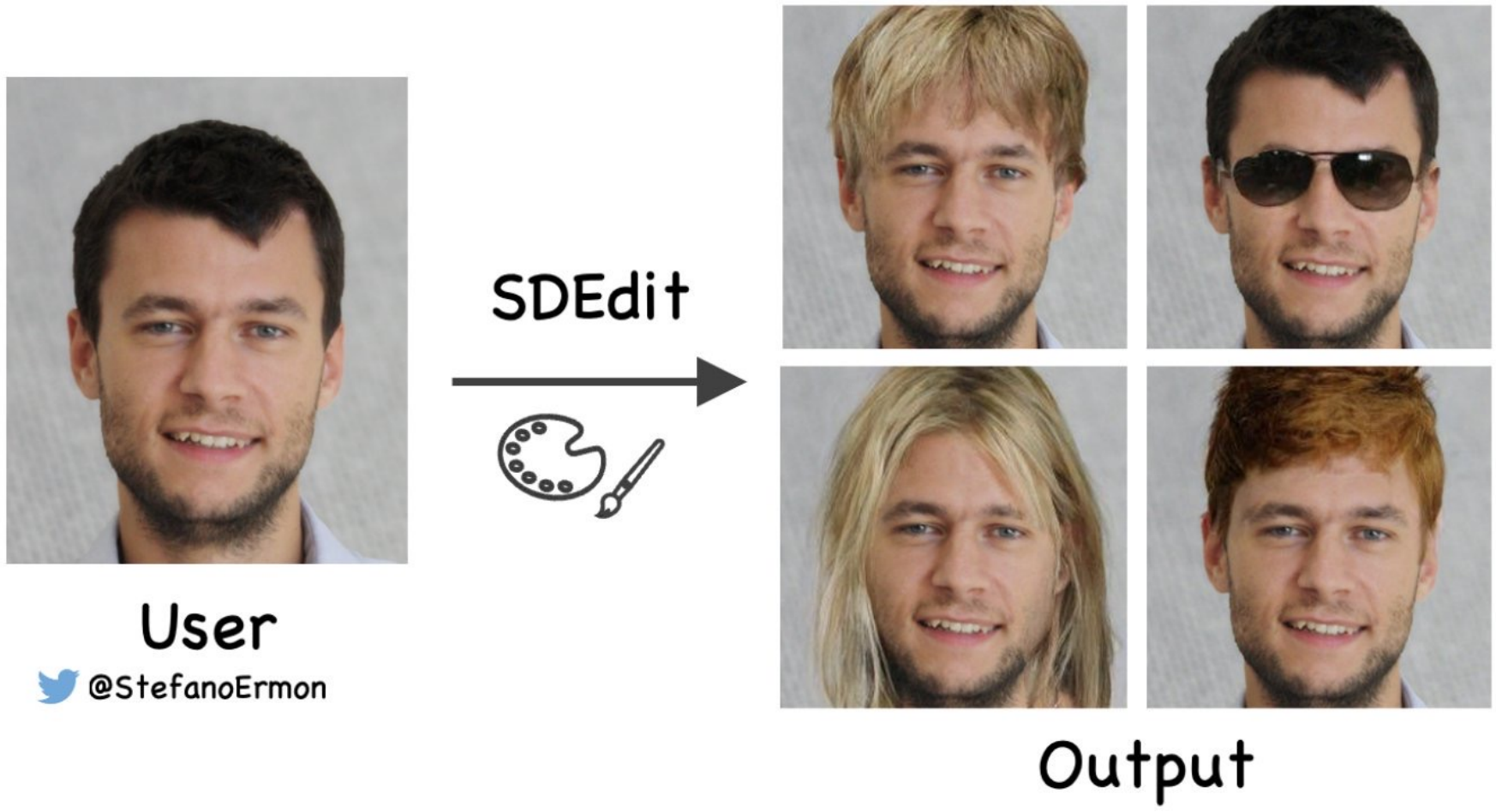
P(image | caption)

TEXT PROMPT a store front that has the word ‘openai’ written on it. . . .

AI-GENERATED IMAGES

DeepFakes

Which image is real?



DeepFakes

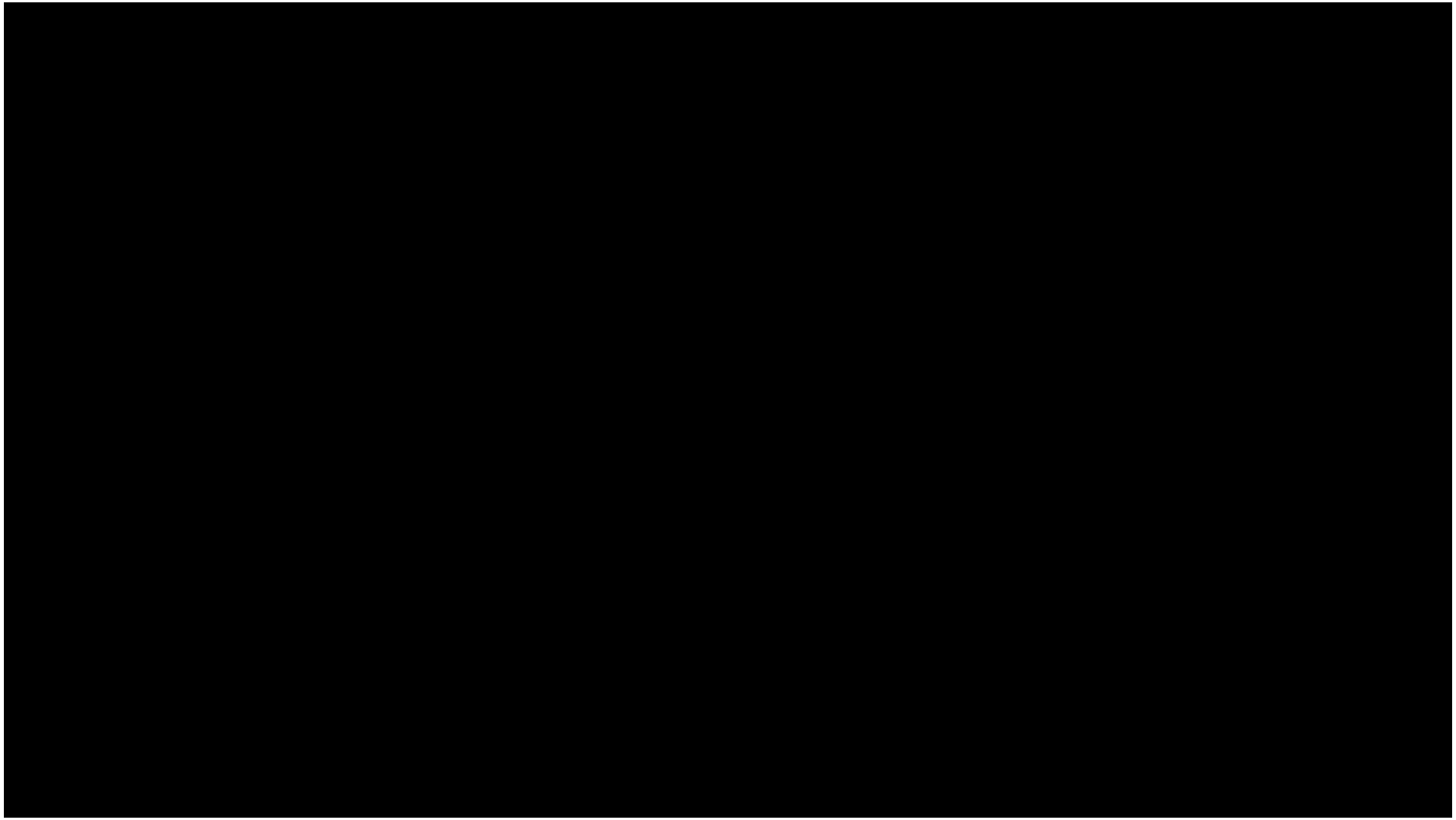


Image Translation or Style Transfer

Conditional generative model P(zebra images| horse images)

Zhu et al., 2017

Imitation Learning

Conditional generative model P(actions | past observations)

Li et al., 2017

CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

Thank You

- Questions?

- Email: yu.yin@case.edu

https://yin-yu.github.io/