



Introduction: How to Create an Image?

Jun-Yan Zhu

16-726, Spring 2021

Teaching Staff

Jun-Yan Zhu



Carnegie
Mellon
University



- Computer Vision, Computer Graphics, Machine Learning, Computational Photography
- Love pets (cat & dog), gaming (mostly FIFA these days)

Cat Paper Collection

As reported by Cisco, 90% of net traffic will be visual, and indeed, most of the visual data are cat photos and videos. Thus, understanding, modeling, and synthesizing our feline friends becomes a more and more critical research problem these days, especially for our cat lovers.

Cat Paper Collection is an academic paper collection that includes computer graphics, computer vision, and machine learning papers that produce experimental results related to **cats**. If you would like to add/remove an article, please send an email to **Jun-Yan Zhu** (junyanz at cs dot cmu dot edu). We thank all the authors for their contribution and support.

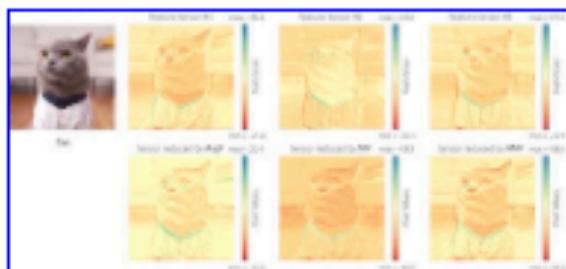
See also [GitHub](#) | [CSV file](#)



Flow-edge Guided Video Completion

Chen Gao, Ayush Saraf, Jia-Bin Huang, Johannes Kopf
In ECCV 2020

[\[Paper\]](#) [\[Project\]](#)



Matching Guided Distillation

Kaiyu Yue, Jiangfan Deng, Feng Zhou
In ECCV 2020

[\[Paper\]](#) [\[Project\]](#)



Strong 3D Printing by TPMS Injection

Xin Yan, Cong Rao, Lin Lu, Andrei Sharf, Haisen Zhao, Baoquan Chen
In IEEE TVCG 2019

[\[Paper\]](#)

Viraj Mehta

- Interested in control, generative modeling, nuclear fusion
- From Austin, TX
- Undergrad in math at Stanford
- Fly airplanes
- Was on jeopardy once



Yufei (Judy) Ye



Born and raised in Beijing



13-17 Tsinghua CS

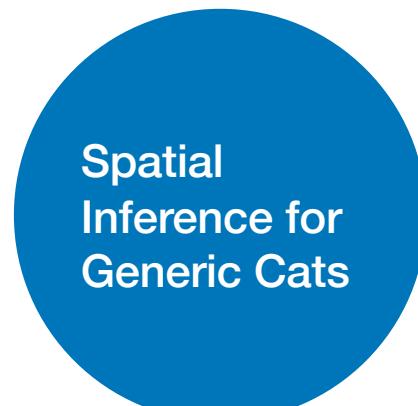


17-19 CMU MSR

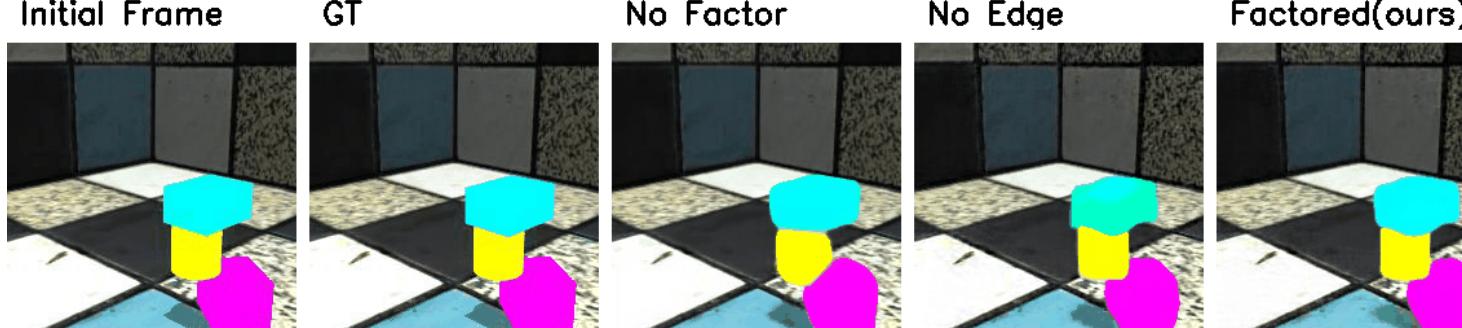


19-??? CMU PhD

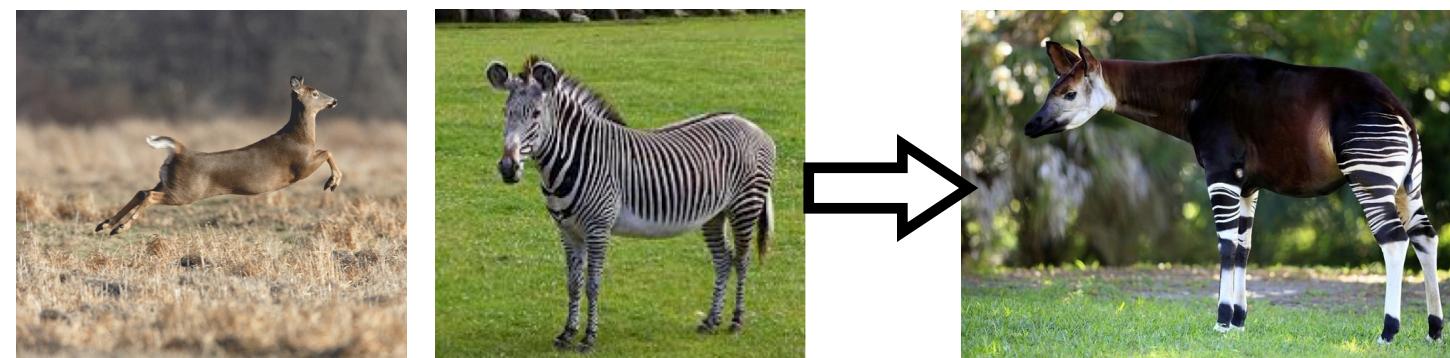
RI PhD student with Abhinav Gupta, Shubham Tulsiani
Sports lover: badminton, soccer, running, etc.



Weakly supervise 3D reconstruction using GAN.



Generate videos via scene graph using VAE.



Generalize to novel categories using GNN.



Introduction: How to Create an Image?

Jun-Yan Zhu

16-726, Spring 2021

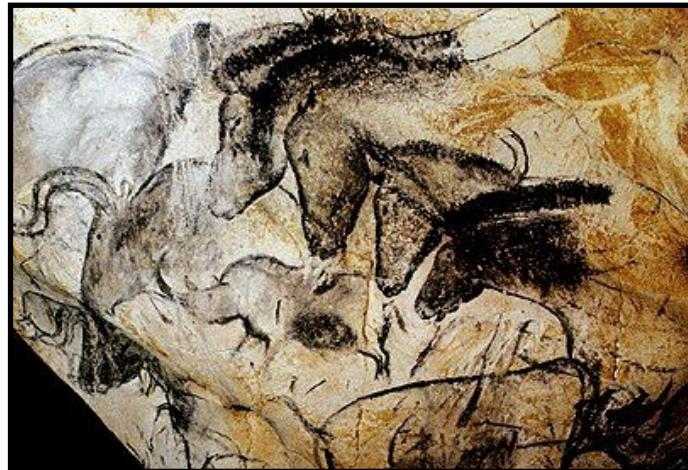
Visual Content Creation

Cave art

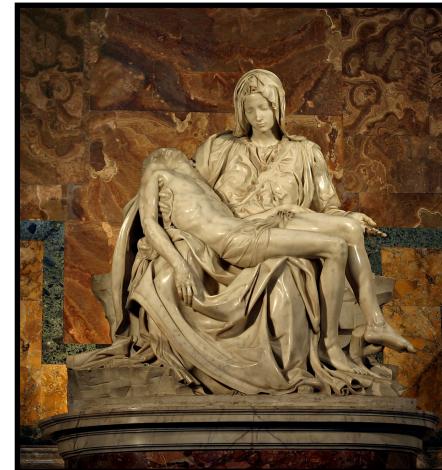


Visual Content Creation

Cave art



Sculpture



Painting



Time

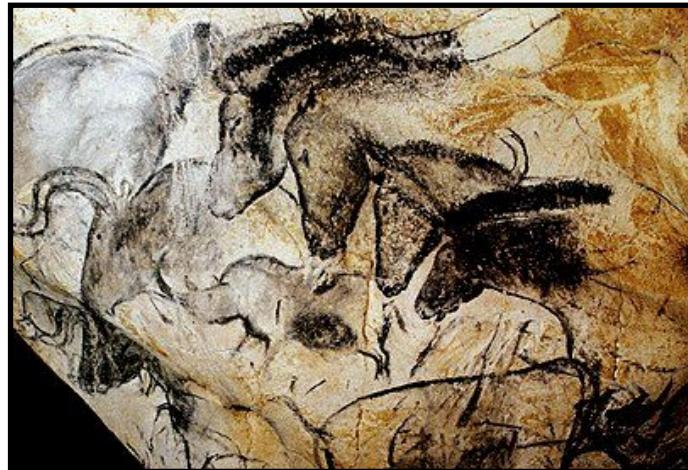
32,000 BC

1498

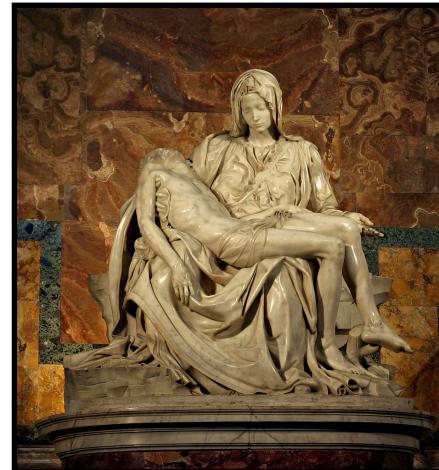
1872

Visual Content Creation

Cave art



Sculpture



Painting



Computer Graphics



Time

32,000 BC

1498

1872

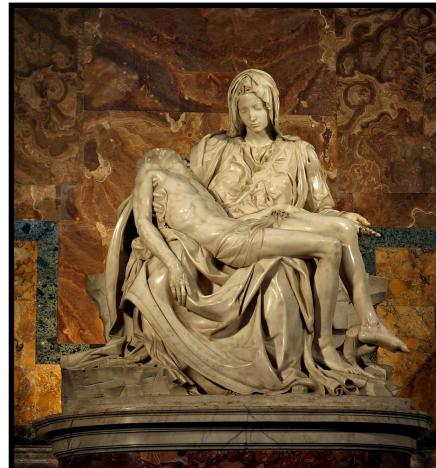
2012

Who is creating visual content?

Cave art



Sculpture



Painting



Computer Graphics



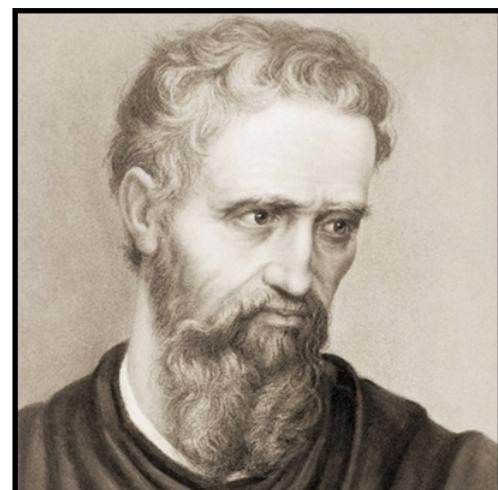
Time

32,000 BC

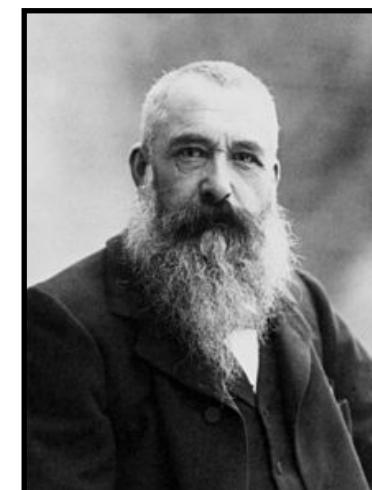
1498

1872

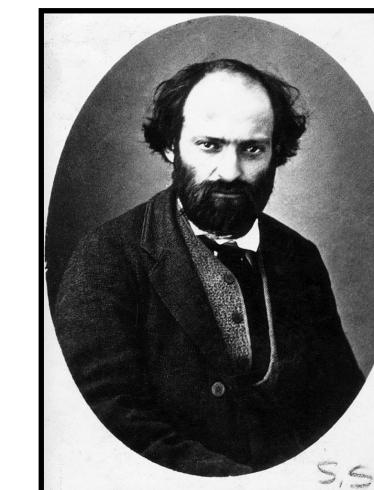
2012



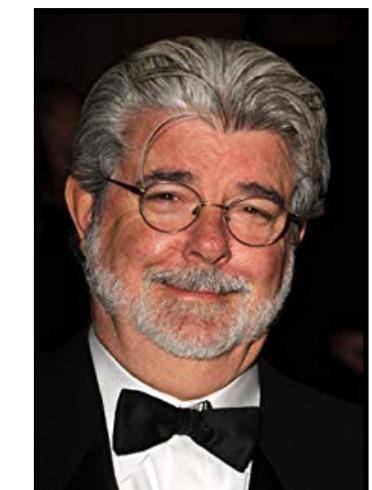
Michelangelo



Claude Monet



Paul Cezanne

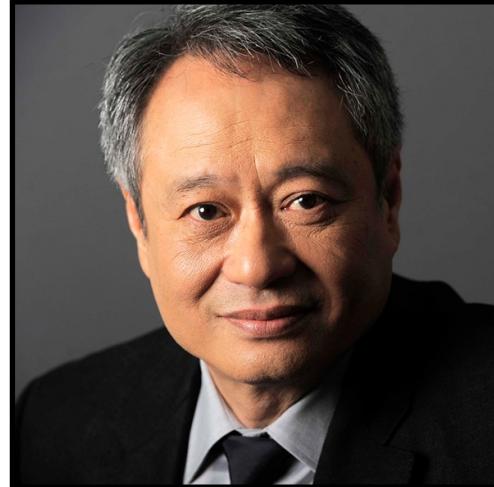


George Lucas

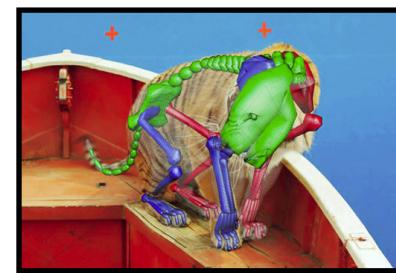


Ang Lee

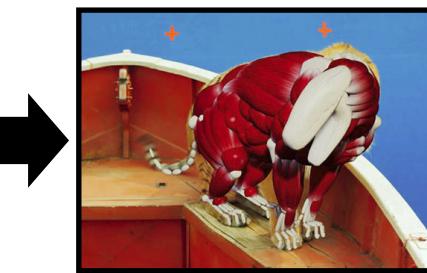
Who is creating visual content?



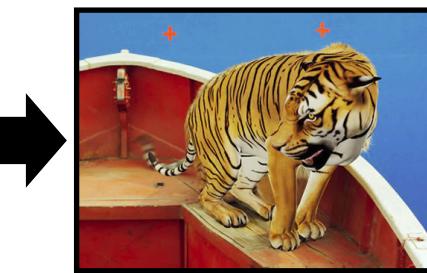
Ang Lee



Skeleton



Geometry



Texture



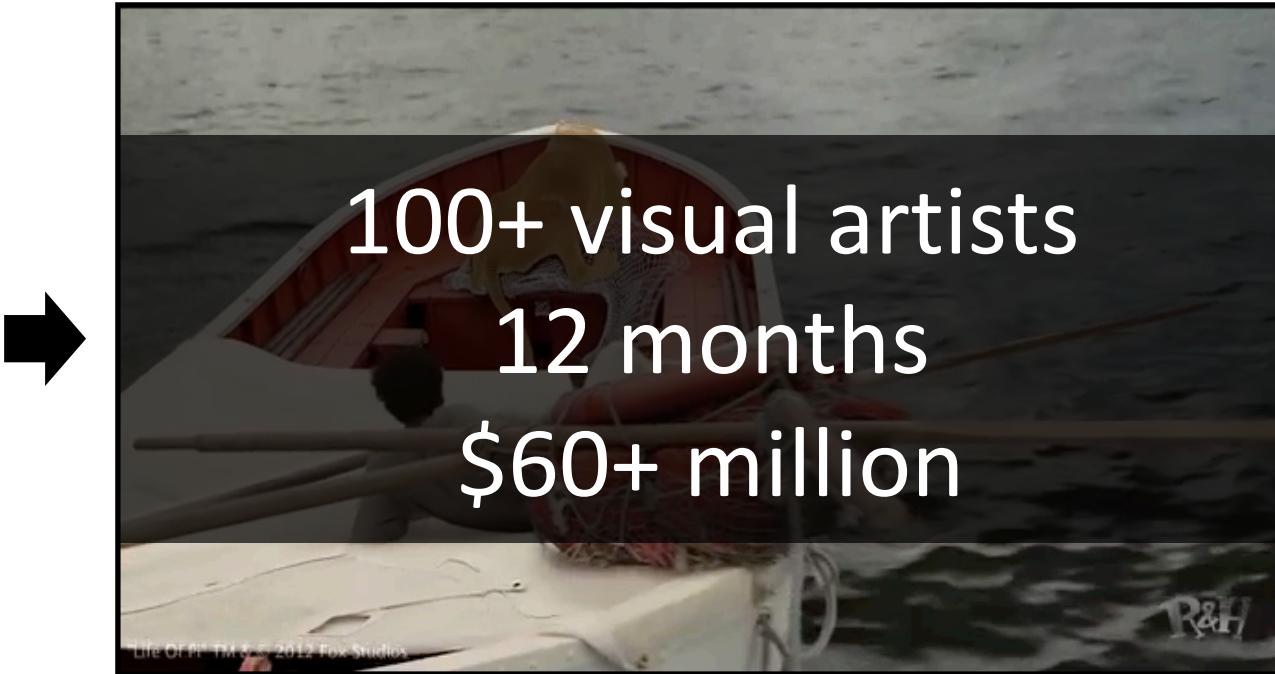
Details



Image



Idea



"Life OF PI" TM & © 2012 Fox Studios



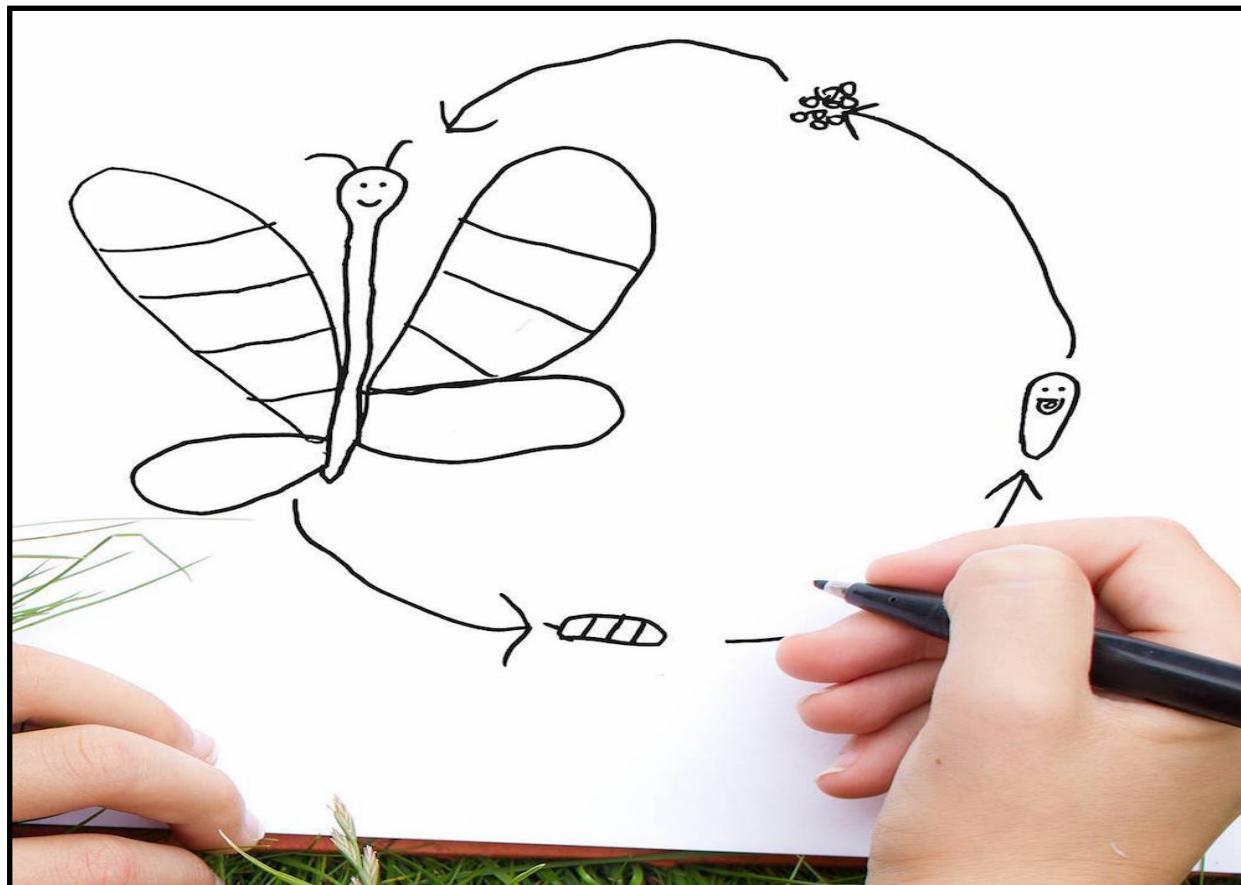
Visual Content

Who is creating visual content?

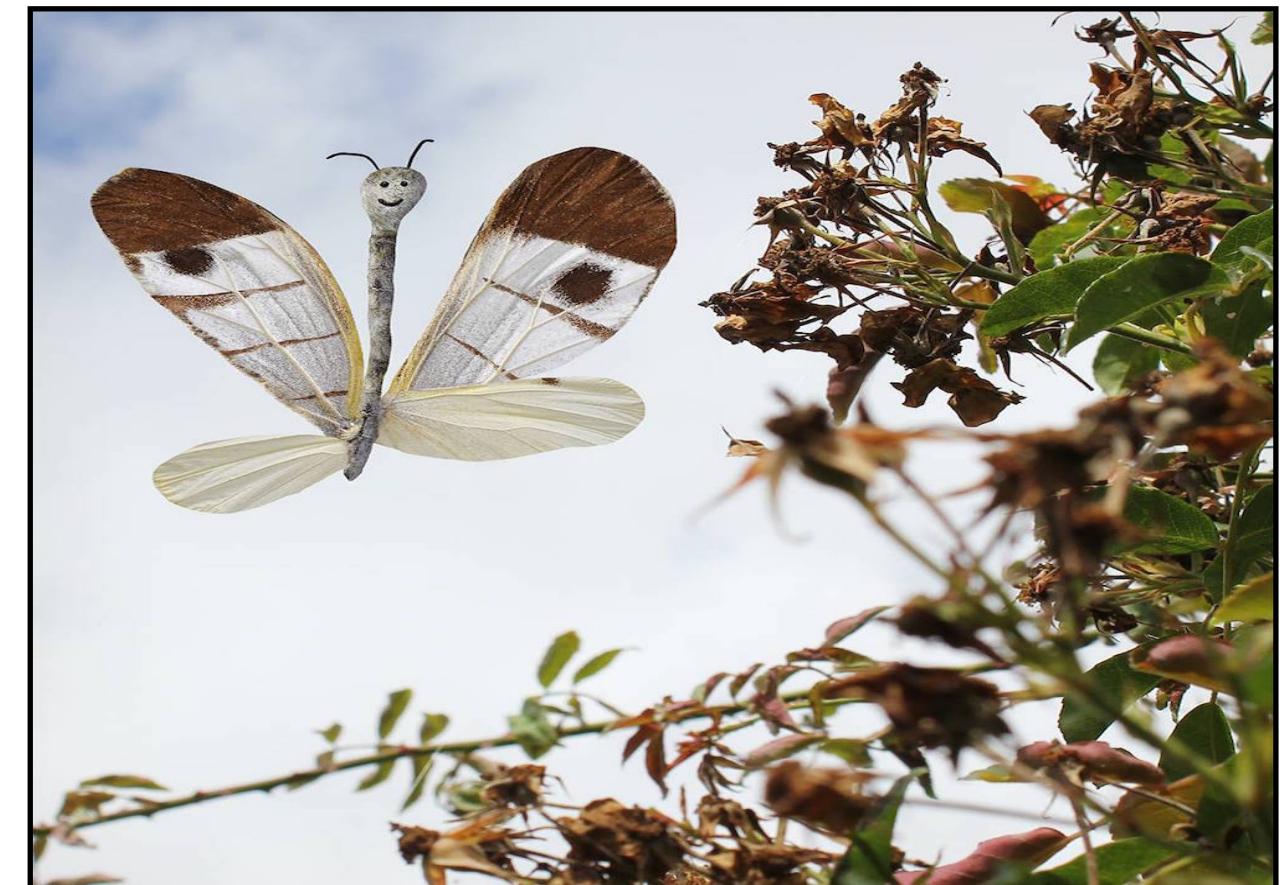
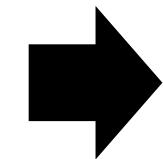
Homework

09/27/2003

Who is creating visual content?



Kid's drawing



Photoshop result by his father

Creating Visual Realism Manually



CG office



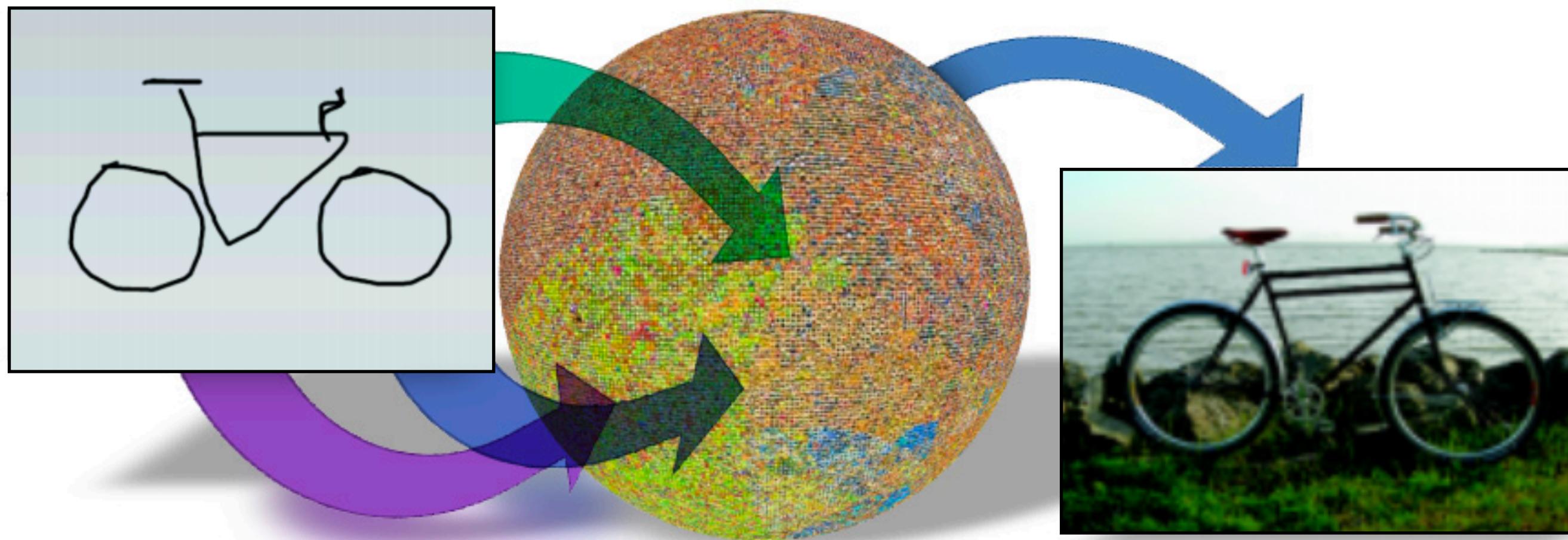
CG office (more details)



My advisor's office

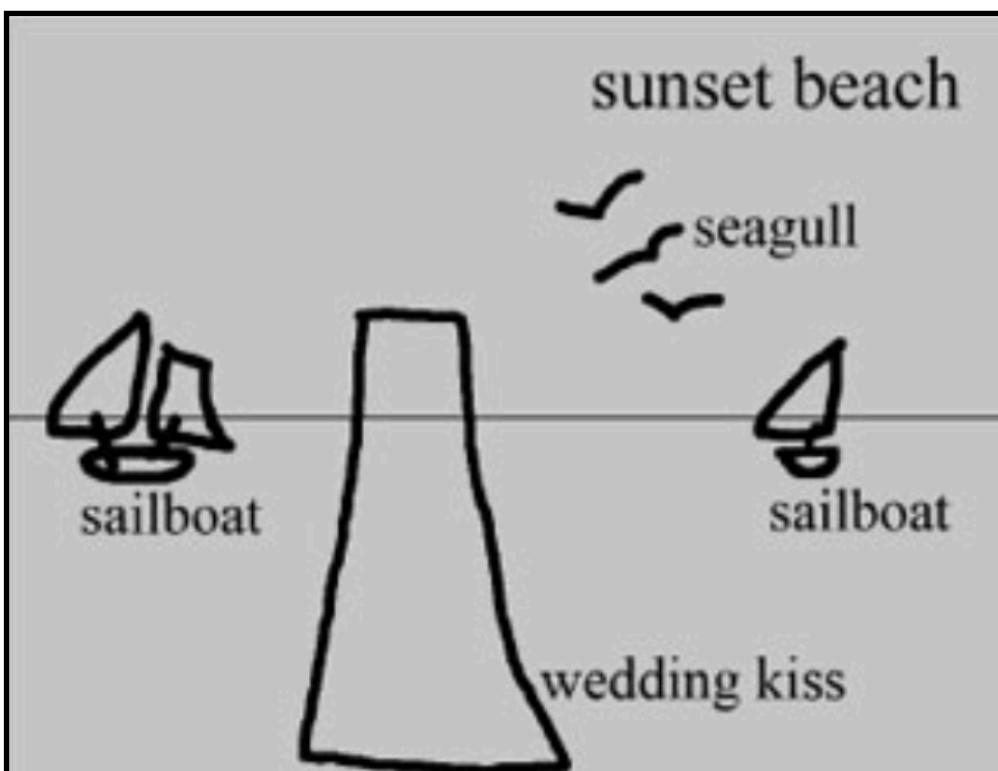
Data-Driven Graphics (2000s)

Graphics → Image Retrieval



Data-Driven Graphics (2000s)

Compositing multiple parts



User Input



Database images



Output

Data-Driven Graphics (2000s)

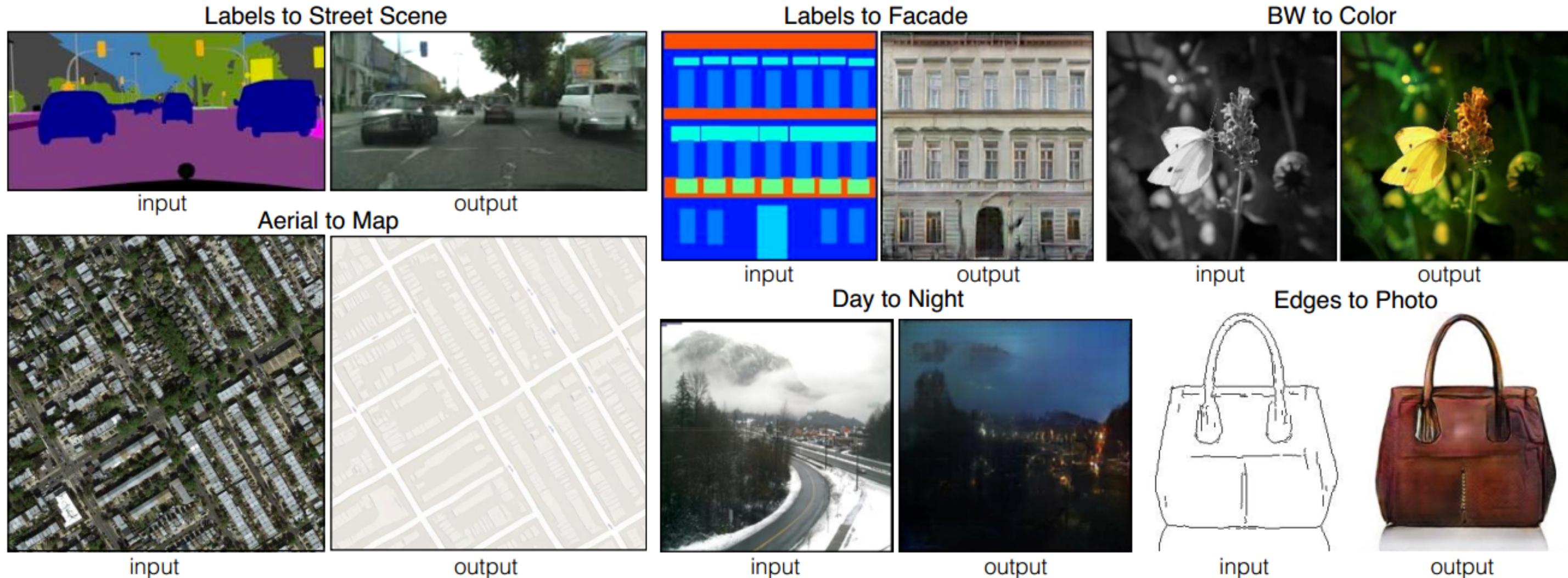


- Hard to combine pieces
- No understanding of visual realism

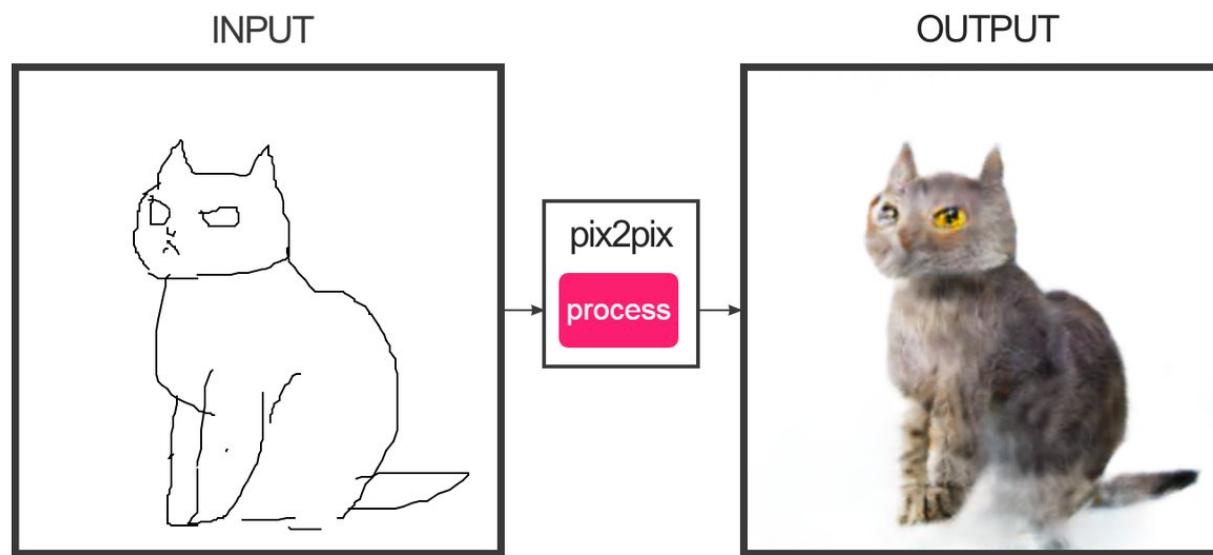
Help everyone
easily create visual content

Teach machines
how to create realistic content

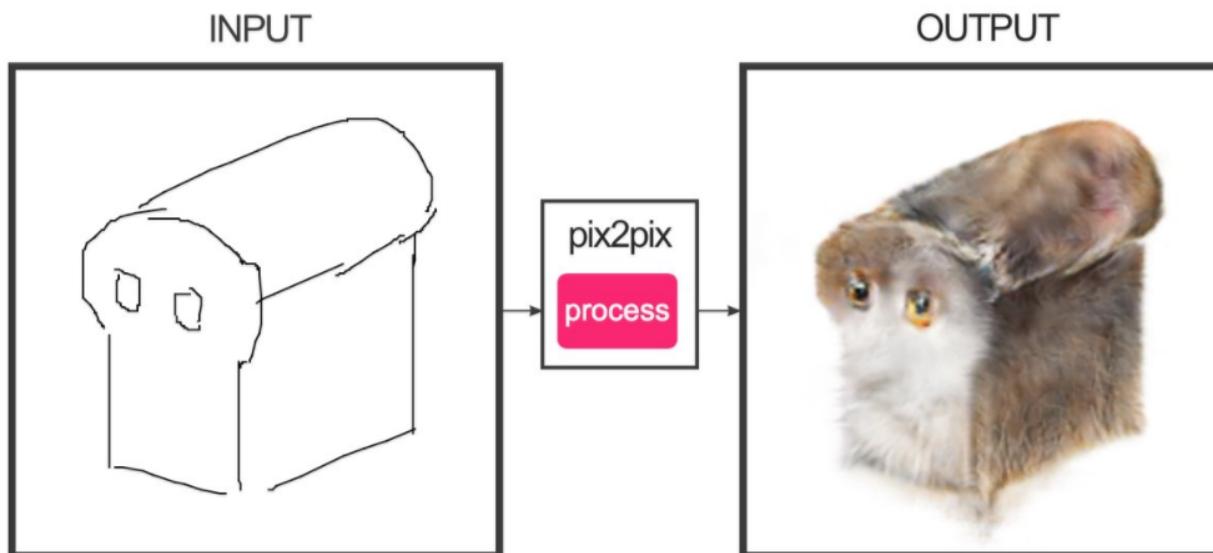
Image-to-Image Translation with pix2pix



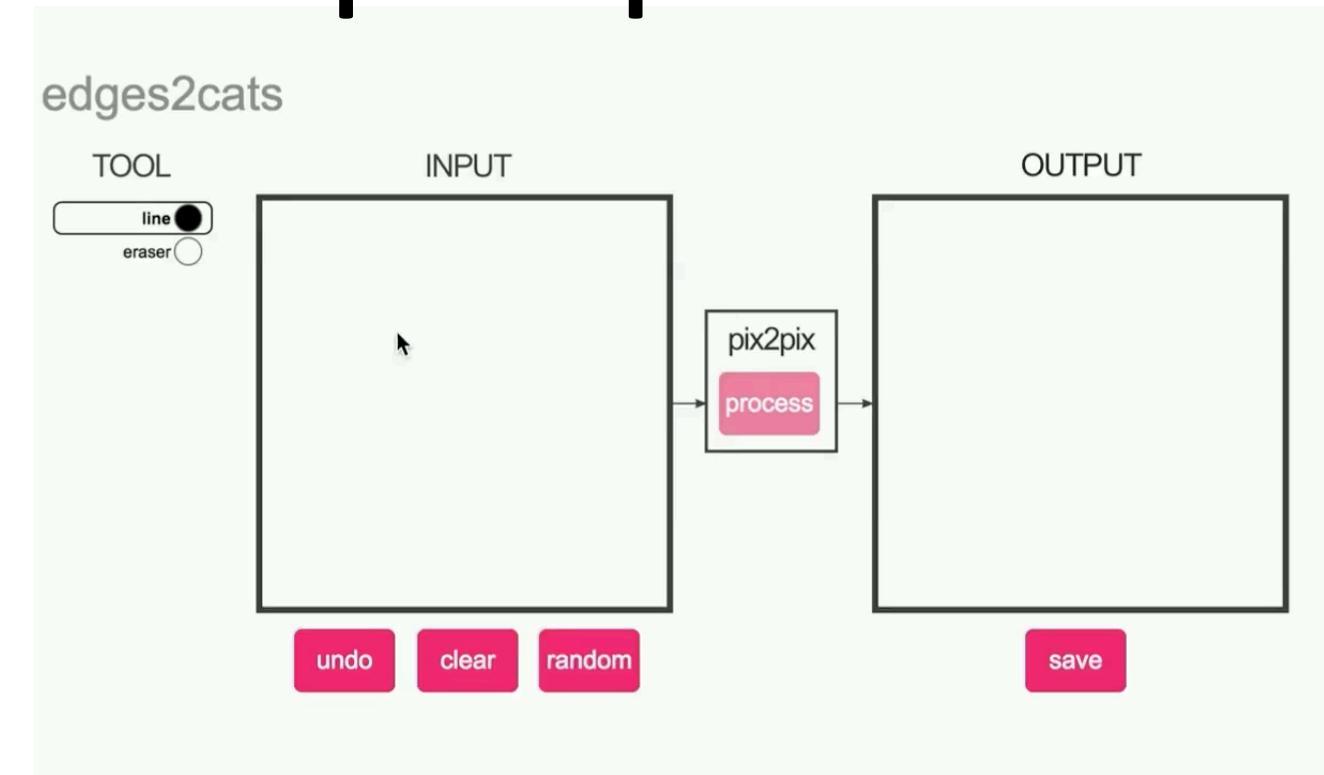
#edges2cats with pix2pix



@gods_tail



Ivy Tasi @ivymyt



@matthematician

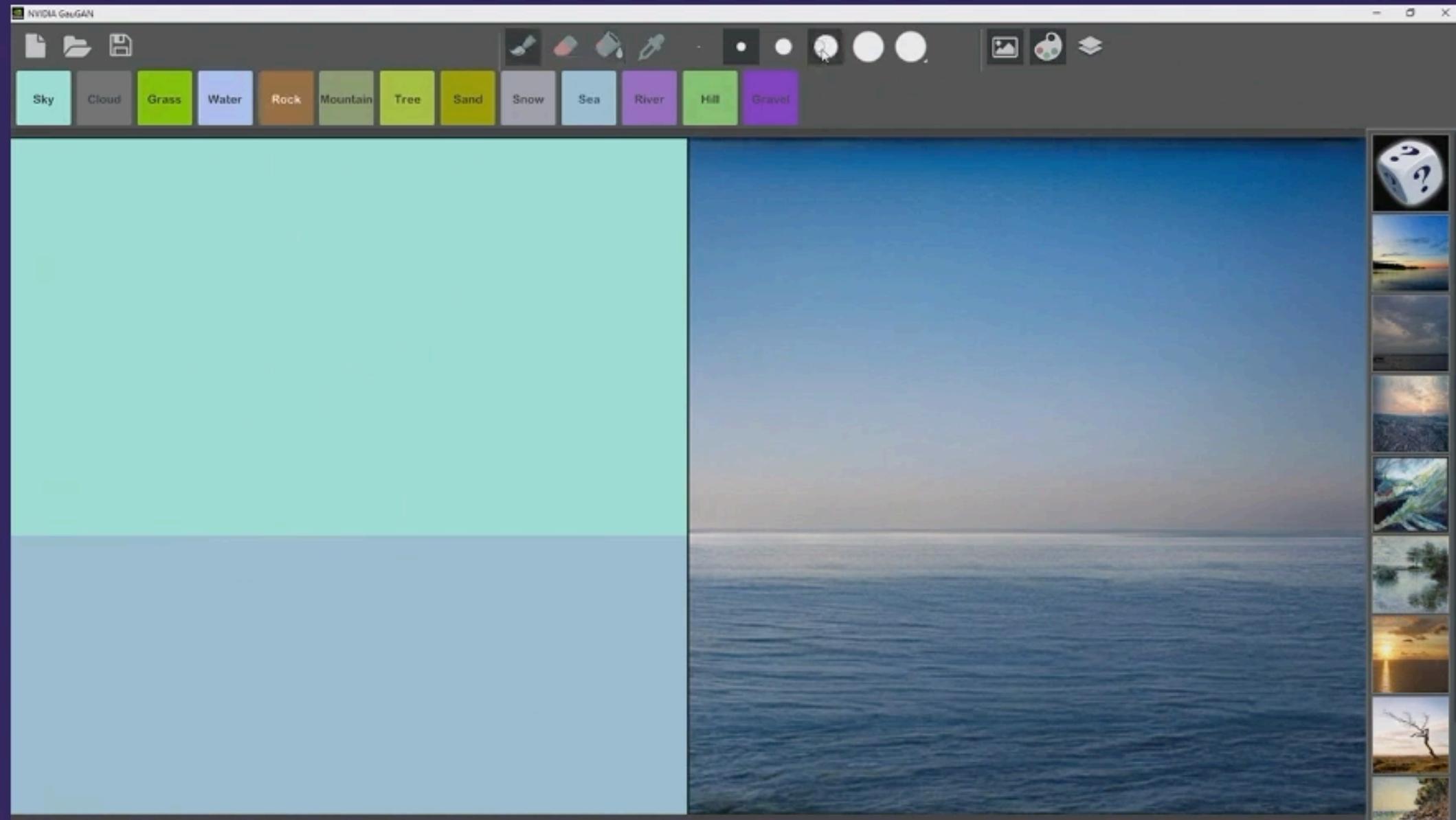


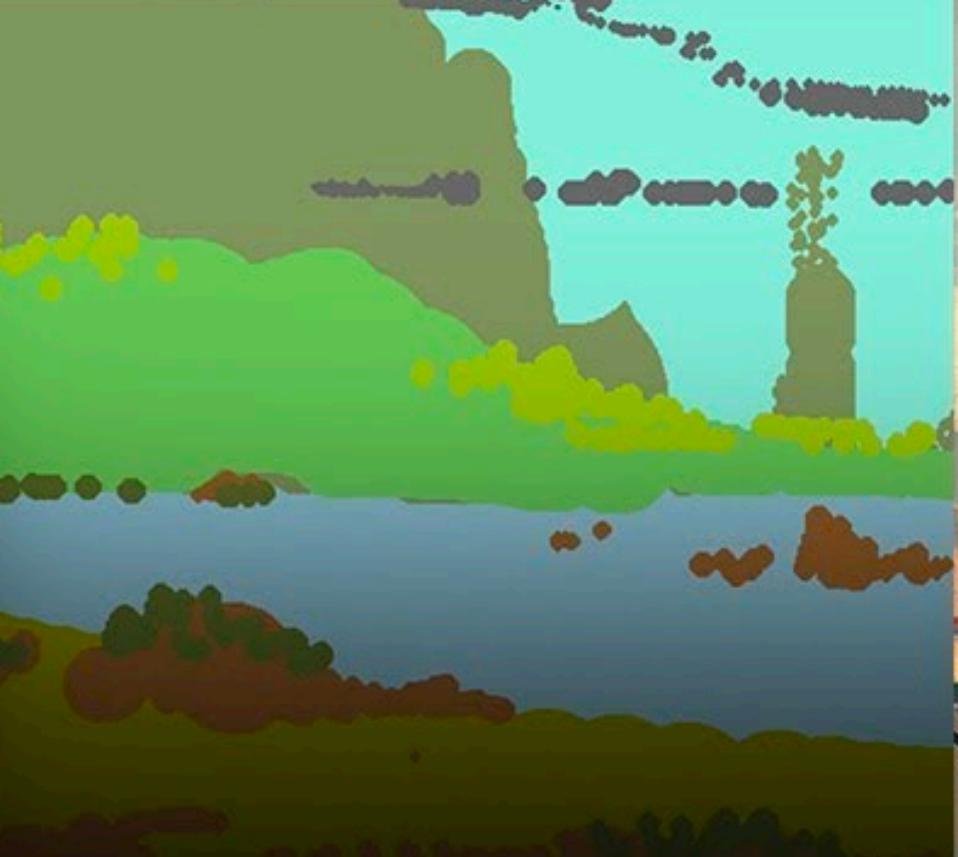
Vitaly Vidmirov @vvid

By Christopher Hesse

<https://affinelayer.com/pixsrv/>

GauGAN [Park, Liu, Wang, Zhu. 2019]





By Darek Zabrocki, Concept Designer and Illustrator

Collection Style Transfer



Photograph ©Alexei Efros



Monet



Van Gogh



Cezanne

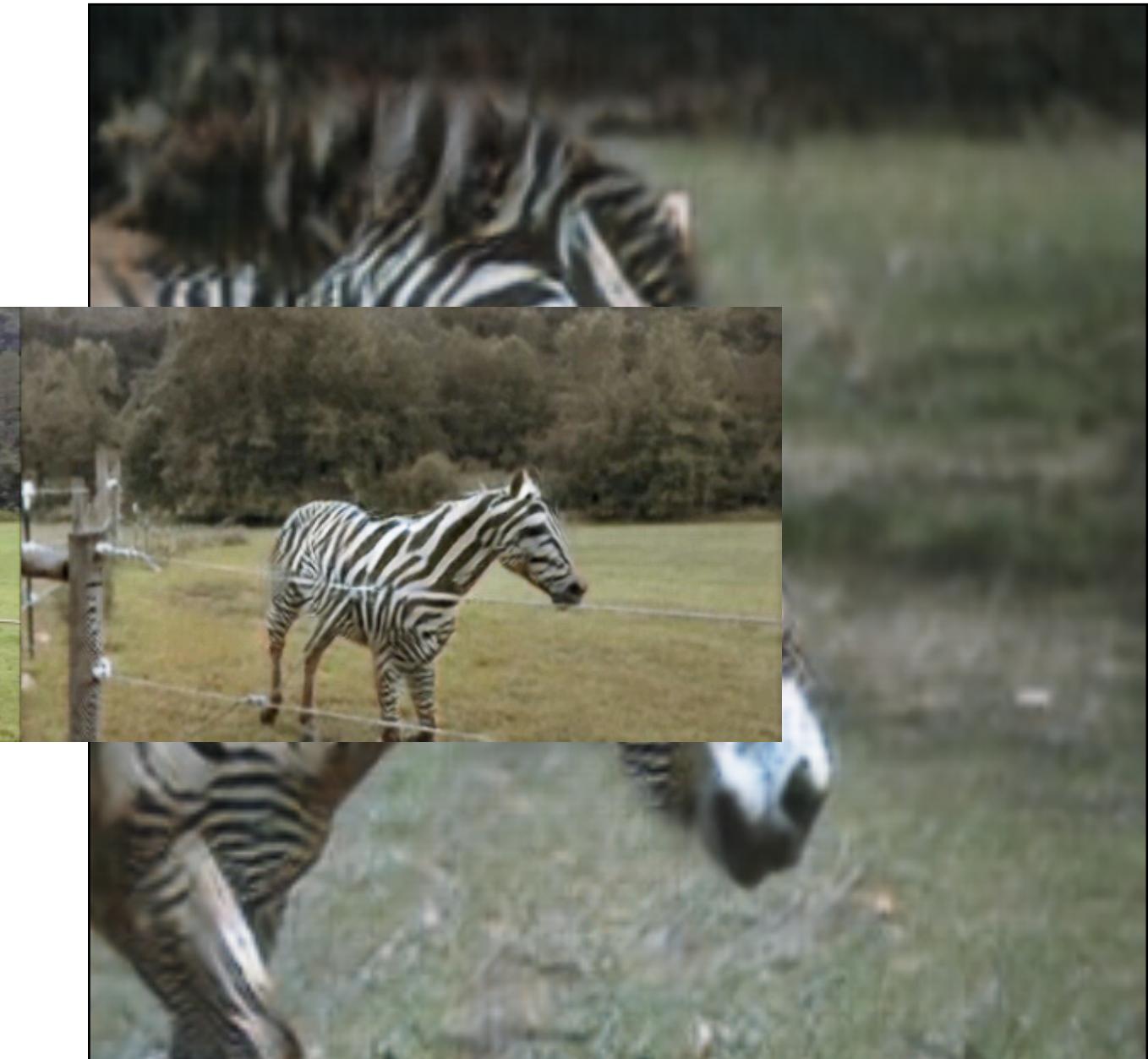


Ukiyo-e

Monet's paintings → photographic style



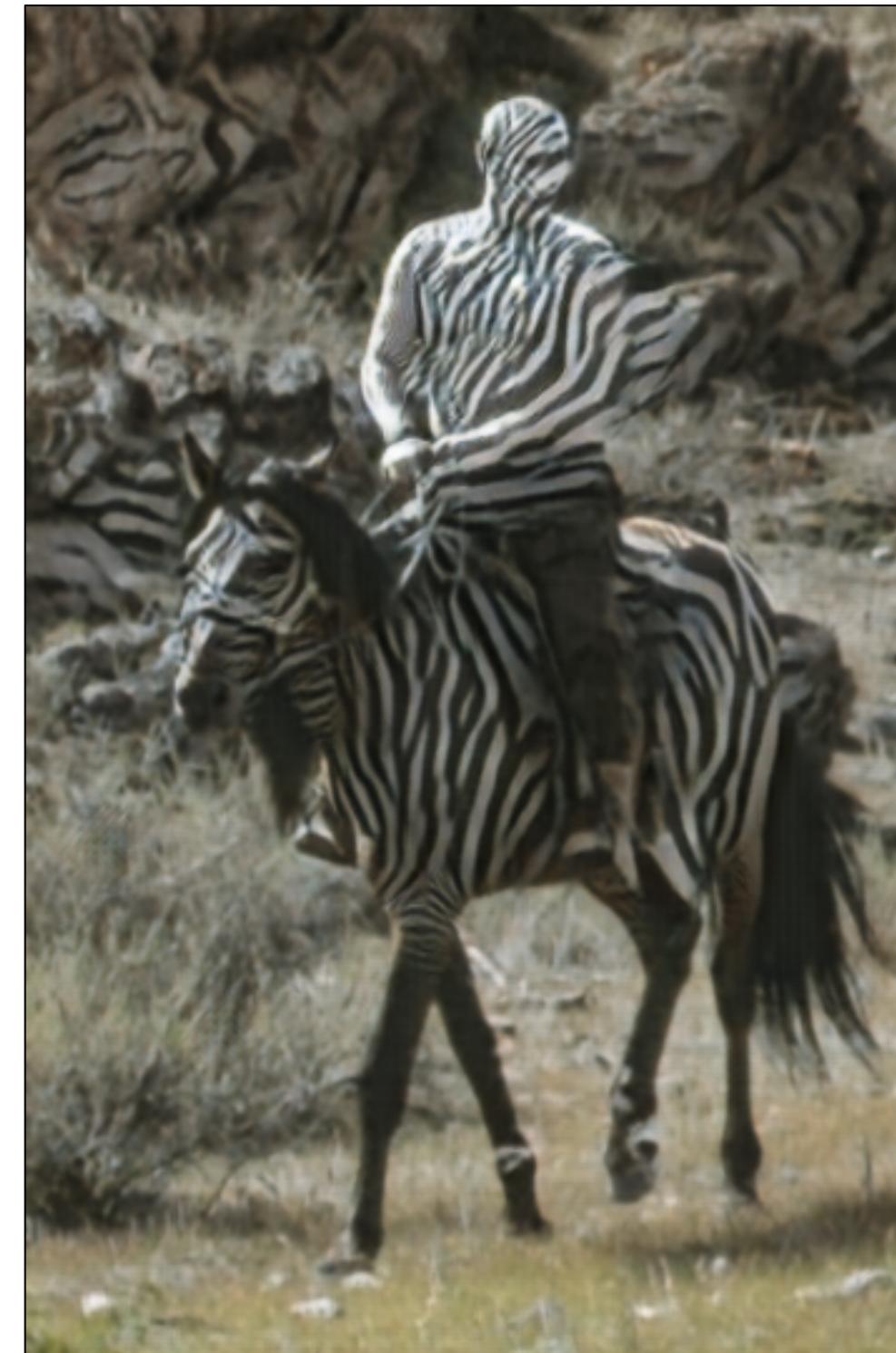
Horse → Zebra



Failure case



Failure case



Swapping Autoencoder For Deep Image Manipulation

Taesung Park¹, Jun-Yan Zhu², Oliver Wang², Jingwan Lu², Eli Shechtman², Alexei Efros¹, Richard Zhang²

¹UC Berkeley, ²Adobe Research



<not_an_ads>

Photoshop 2021 “Neural Filters”



</not_an_ads>

Research Highlight

Synthesizing High-res Portraits



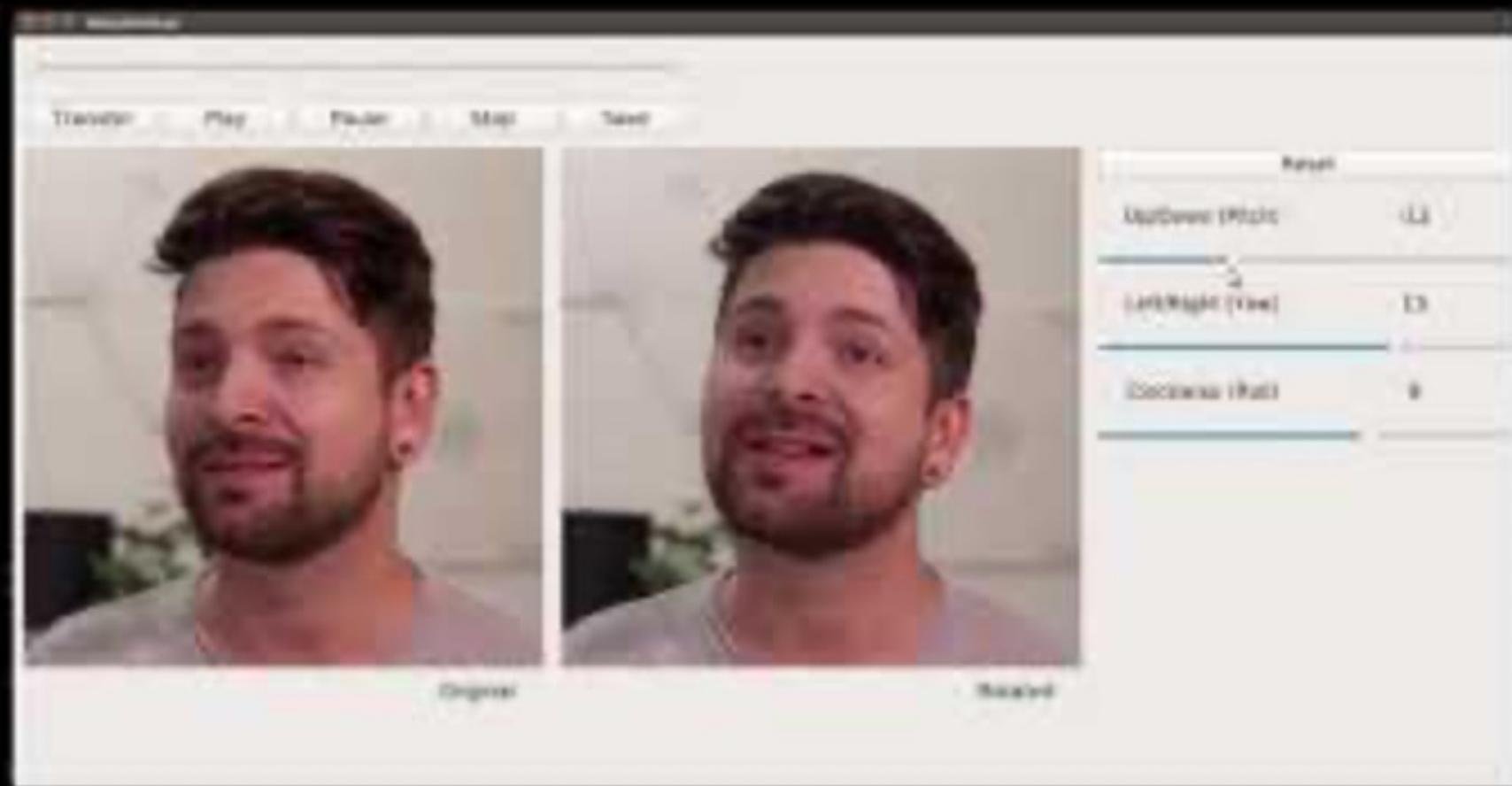
[Kerras et al. CVPR 2020] @ NVIDIA

Everybody Dances Now



Neural Talking-Head Synthesis

Interactive Demo Interface

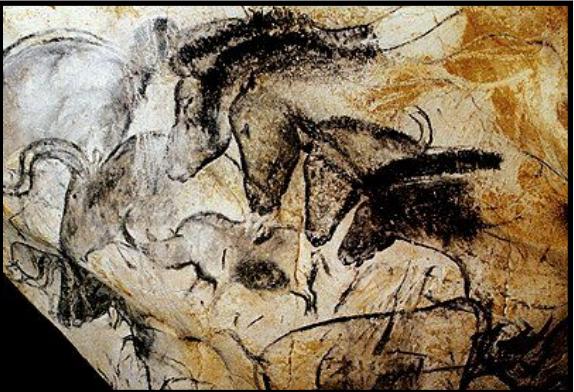


[Wang et al. arXiv 2020] @ NVIDIA

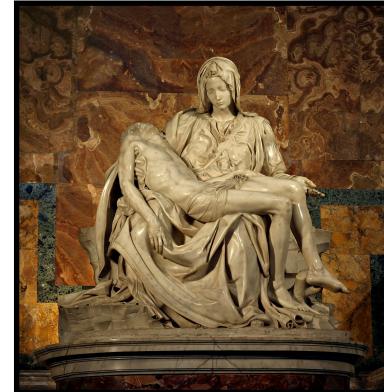
NeRF in the Wild



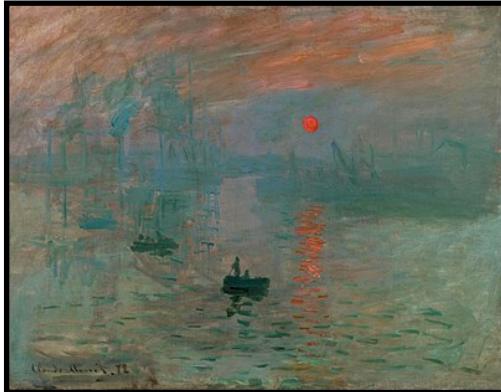
Cave art



Sculpture



Painting



Computer Graphics

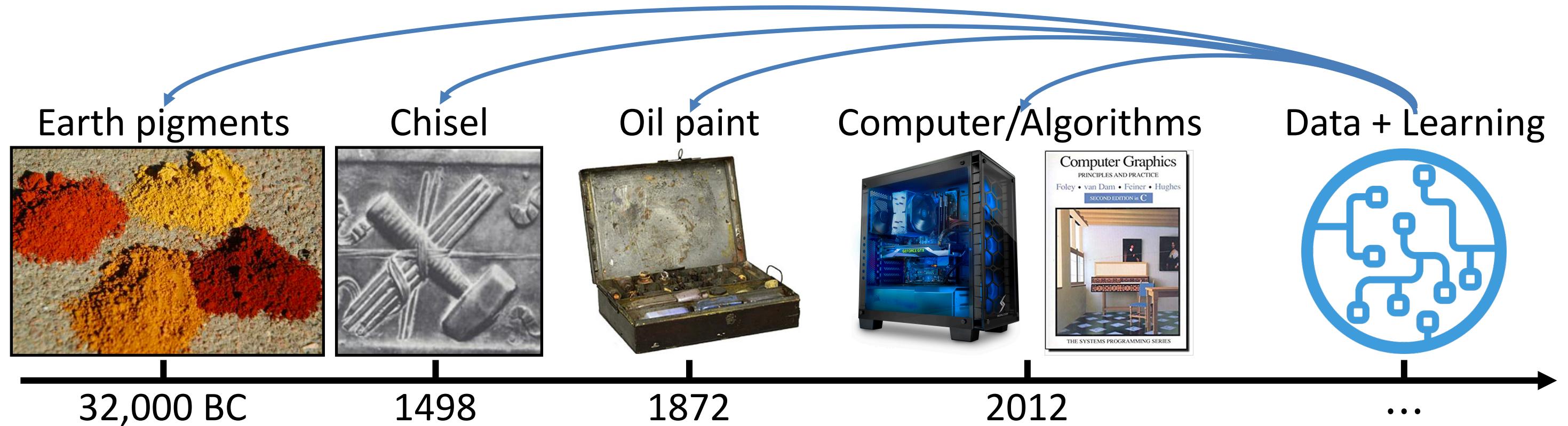


32,000 BC

1498

1872

2012



Course preview

- A modern machine learning perspective
- Widely-used learning algorithms
- Interactive content creation tools

Logistics

Course objectives

1. You will get a foundation in image editing and synthesis.
 - Texture synthesis and style transfer.
 - Face modeling and synthesis.
 - Image colorization and inpainting.
 - Video generation and editing.
 - Image-to-image translation.
 - Image and video editing. (warping, morphing, compositing)
 - Image and video forensics.

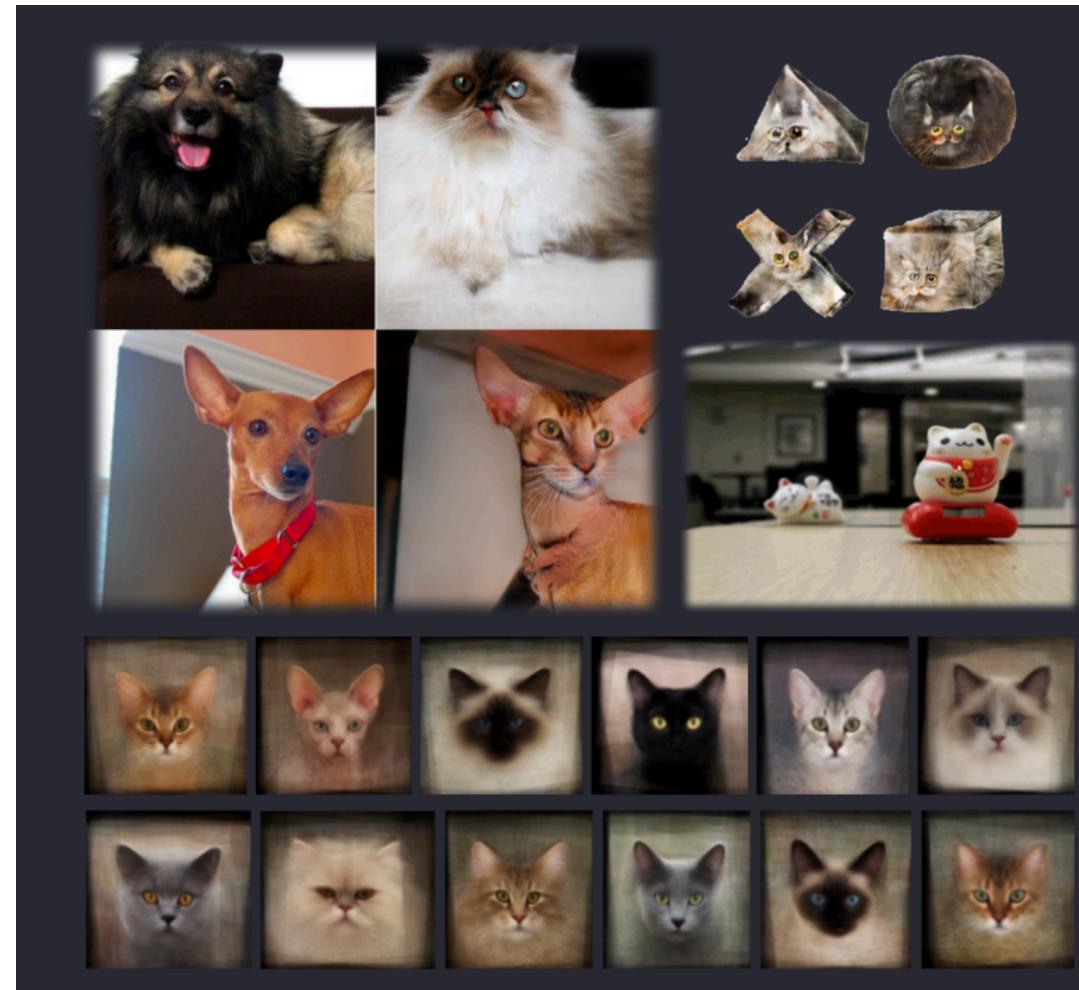
Course objectives

2. You will get a foundation of machine learning concepts

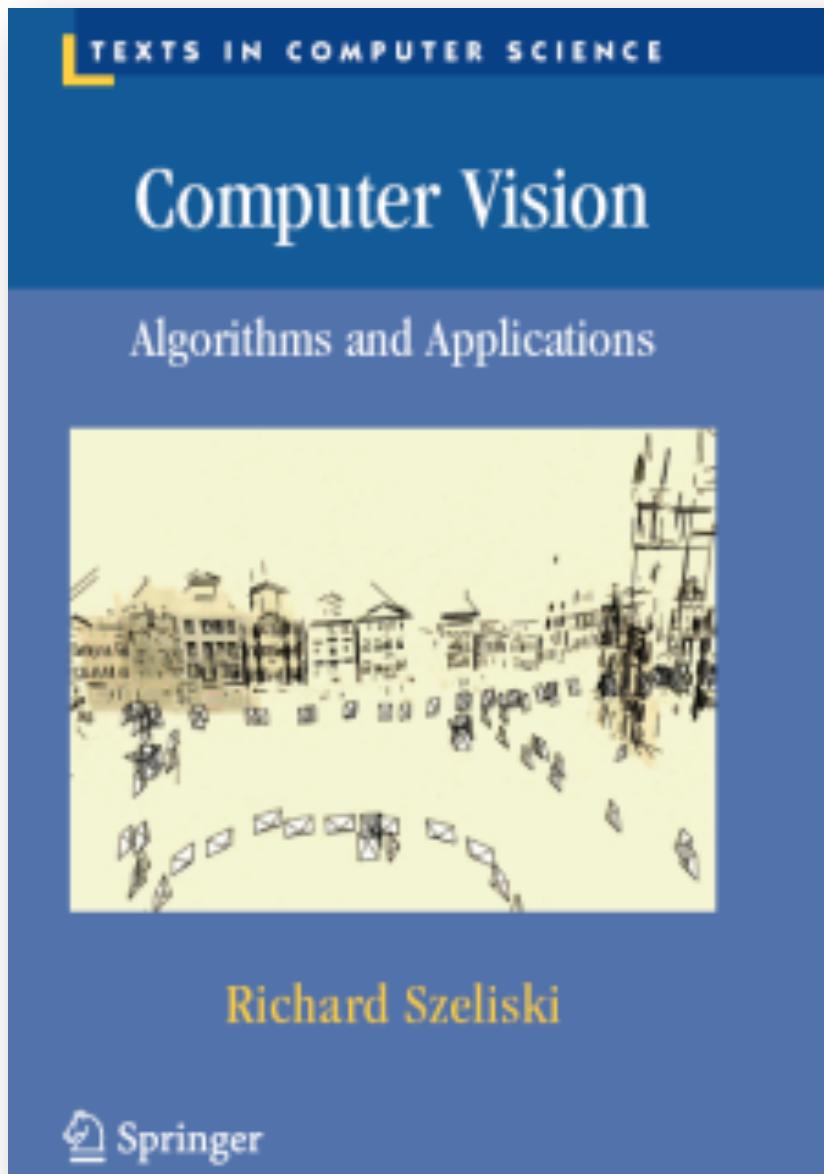
- (fast) Nearest neighbor search.
- Principal component analysis, Gaussian Mixture model. Markov Random Field (MRF)
- Convolutional neural networks.
- Deep generative models: Auto-encoder, Generative Adversarial Networks, Flow-based models, Variational Auto-encoder, PixelCNN, Energy-based models.
- Conditional generative models.

Course objectives

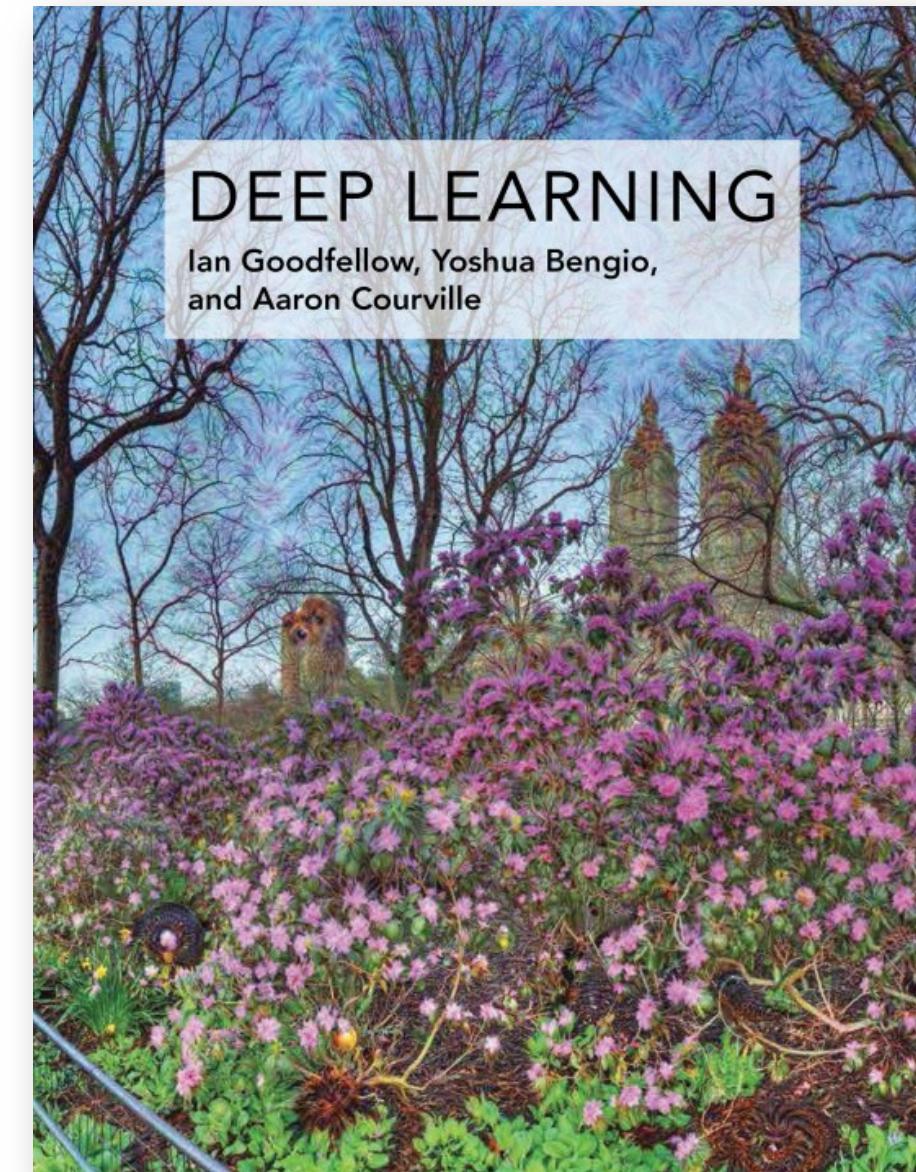
3. You will have some cool results using your own photos



Textbook



<https://szeliski.org/Book/>
(2021 edition")



<https://www.deeplearningbook.org/>
(2016 edition)

Grading

- Emphasis on programming projects (**65%**).
 - Classic: 1. image alignment. 2. image blending
 - Deep learning: 3. neural style transfer. 4. GANs and conditional GANs.
 - 5. reconstructing and editing an image with GANs.
- Late Policy for programming assignments.
 - Five (5) emergency late days for semester, to be spent wisely
 - 10% of penalty per day afterwards
- One paper presentation (**10%**):
 - 10-20 min, 1-2 people in a group.
 - Need to answer questions about this paper from now on.
- Final Project (**25%**)
 - A webpage-based report + a presentation.
 - No late day.
 - 2-3 people per group.

For each project

- Derive the math, implement stuff from scratch (+ starter code), and apply it to your own photos
- Every person does their own project (except final)
- Reporting via web page (+ submit code to Canvas)
- Afterwards, vote for class favorite(s)! Gift!
- Programming Language:
 - Python/PyTorch
 - you can use other languages, but you are on your own

Academic Integrity

- Can discuss projects, but don't share code
- Don't look up code or copy from a friend
- If you're not sure if it's allowed, ask
- Acknowledge any inspirations
- If you get stuck, come talk to us

Getting help outside of class

- Course Web Page
 - <https://learning-image-synthesis.github.io/>
- Discussion board:
 - Piazza.com
- Office hours (EST)
 - Viraj: 5:30-6:30 pm Monday [[link](#)]
 - Yufei: 9-10 am Tuesday [[link](#)]
 - Jun-Yan: 8-9 pm Wednesday [[link](#)]

The screenshot shows the course website for 16-726 Learning-Based Image Synthesis at Carnegie Mellon University for the Spring 2021 semester. The header features the university's logo and the course title "16-726 Learning-Based Image Synthesis" with the subtitle "Spring 2021". Below the header, there is a navigation bar with links for HOME, SCHEDULE, LECTURES, ASSIGNMENTS, and MATERIALS. The main content area displays a photograph of a river flowing through a canyon, with the caption "source" underneath it. The course description is provided below the image, detailing the objectives of synthesis research, ranging from classical learning methods to deep learning models, and the practical applications students will learn to build.

Carnegie Mellon University
16-726 Learning-Based Image Synthesis
Spring 2021

HOME SCHEDULE LECTURES ASSIGNMENTS MATERIALS

16-726 Learning-Based Image Synthesis / Spring 2021

Time: Mondays, Wednesdays 4:00 pm - 5:20 pm ET

Location: remote-only (see Canvas for Zoom links)

source

Course Description

This course introduces machine learning methods for image and video synthesis. The objectives of synthesis research vary from modeling statistical distributions of visual data, through realistic picture-perfect recreations of the world in graphics, and all the way to providing interactive tools for artistic expression. Key machine learning algorithms will be presented, ranging from classical learning methods (e.g., nearest neighbor, PCA, Markov Random Fields) to deep learning models (e.g., ConvNets, deep generative models, such as GANs and VAEs). We will also introduce image and video forensics methods for detecting synthetic content. In this class, students will learn to build practical applications and create new visual effects using their own photos and videos.

Why you should NOT take this class

- Project-based class
 - No canned problem sets.
 - Not theory-heavy.
 - will read many research papers.
 - Open-ended by design.
- Need time to think, not just hack
 - **Creativity** is a class requirement.
- Not worth it if you don't enjoy it.

Now... reasons TO take this class

- Not too many similar courses at other places.
- You get to create pictures and unleash your creative potential.
- Interested in grad school and research? ☺
- Interested in industry jobs? ☺

Thank You!



16-726, Spring 2021

<https://learning-image-synthesis.github.io/>