國立高雄科技大學
National Kaohsiung University of Science and Technology

Introduction to Artificial Intelligence and CV Applications

Speaker: 蘇佳益
Advisor: 陳聰毅
國立高雄科技大學建工校區電子工程系

https://github.com/chiayisu/Artificial_Intelligence_Course

Agenda

- Course Introduction
- Artificial Intelligence
- Machine Learning
- PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models
- SynSin: End-to-end View Synthesis from a Single Image
- PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization
- Wish You Were Here: Context-Aware Human Generation
- Detectron2: A PyTorch-based modular object detection library
- References

Course Introduction

Course Information

- This course will cover
  - Some AI applications
  - Machine Learning algorithms
  - Deep Learning algorithms
  - NLP models
  - NLP applications

Tutorials for Implementing Algorithm

- Python Tutorials
  - http://cs231n.github.io/python-numpy-tutorial/
  - http://web.stanford.edu/class/cs224n/readings/python-review.pdf
- Pytorch Tutorial
  - https://pytorch.org/tutorials/
- Tensorflow Tutorial
  - https://www.tensorflow.org/tutorials

Online Lectures

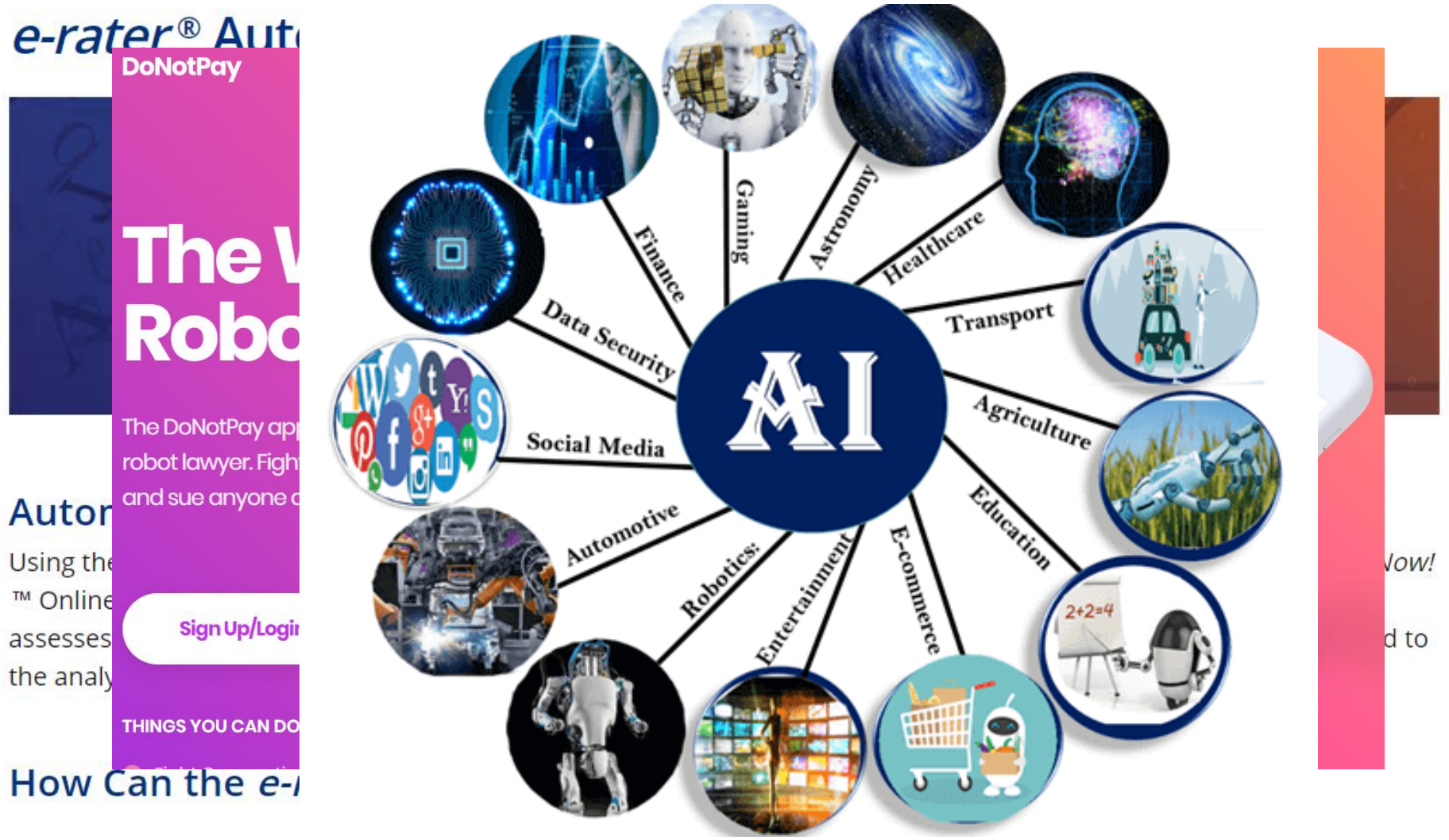
- CS230 Deep Learning – Stanford University
  - https://cs230.stanford.edu/
- EECS 498-007 / 598-005 Deep Learning for Computer Vision – University of Michigan
  - https://web.eecs.umich.edu/~justincj/teaching/eecs498/FA2020/
- CS221: Artificial Intelligence: Principles and Techniques – Stanford University
  - https://stanford-cs221.github.io/autumn2019/#schedule
- CS224n: Natural Language Processing with Deep Learning – Stanford University
  - http://web.stanford.edu/class/cs224n/
- CS231n: Convolutional Neural Networks for Visual Recognition – Stanford University
  - http://cs231n.stanford.edu/
- 11-411: Natural Language Processing – CMU
  - http://demo.clab.cs.cmu.edu/NLP/#overview
- Neural Network for NLP – CMU
  - http://www.phontron.com/class/nn4nlp2019/description.html

Free Books

- Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning.
  - https://www.deeplearningbook.org/
- Dan Jurafsky and James H. Martin, Speech and Language Processing.
  - https://web.stanford.edu/~jurafsky/slp3/

Artificial Intelligence

AI is every where.



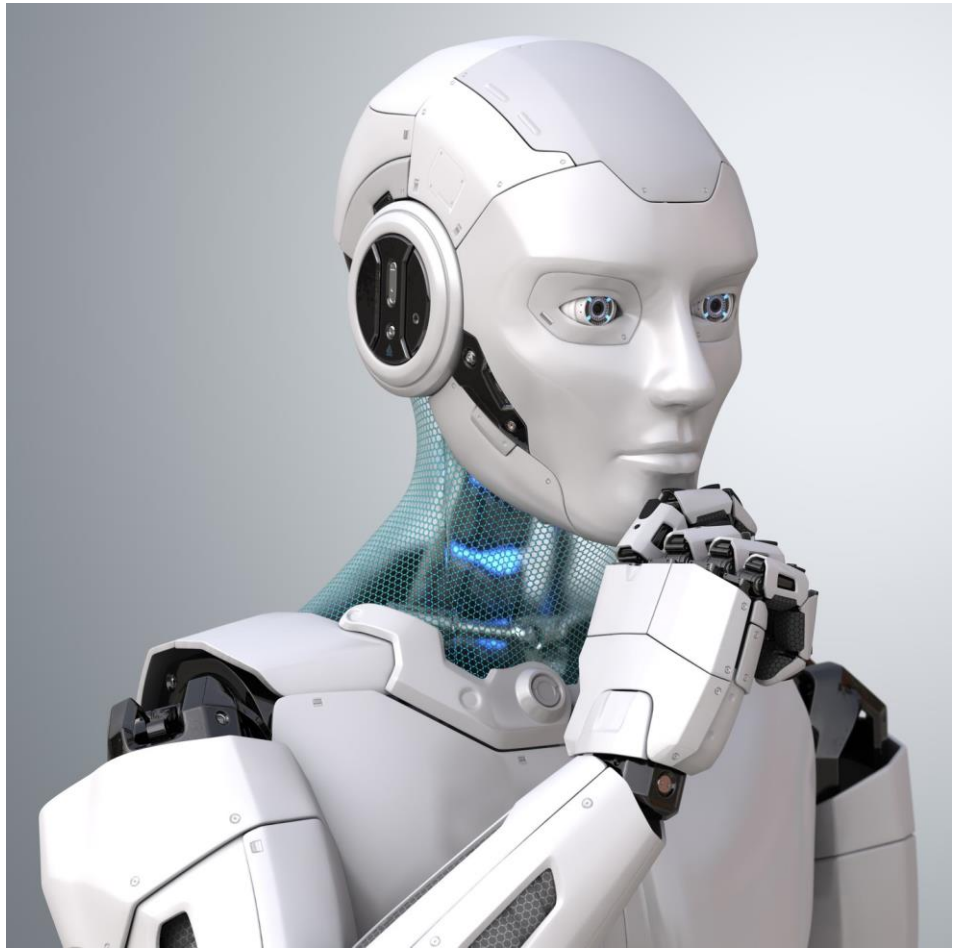
Turing Test

If Interrogators cannot differentiate the difference between human and computer, then test passes.

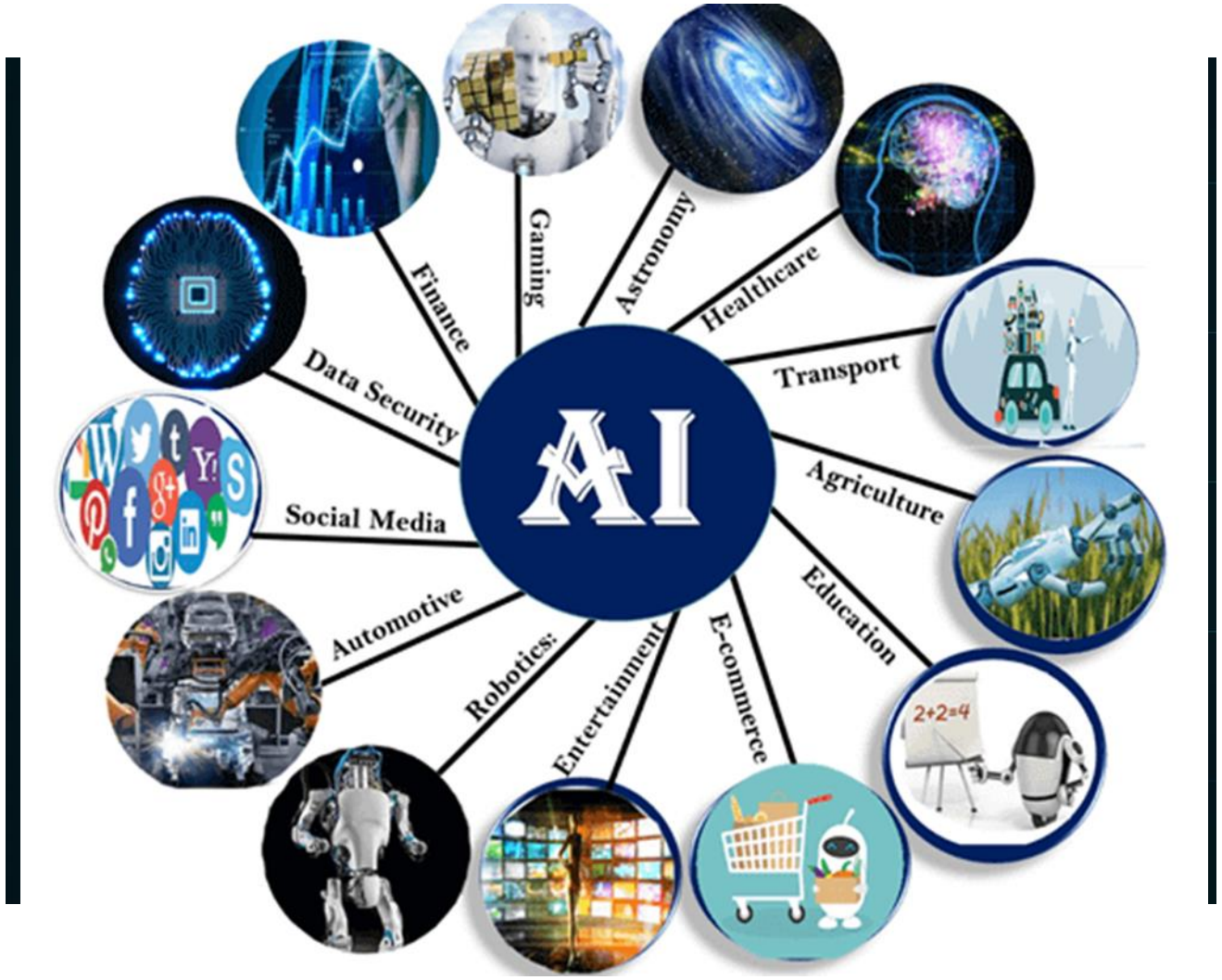
Two types of AI

- Strong AI
- Weak AI

Difference between Strong AI and Weak AI



Two Views of AI

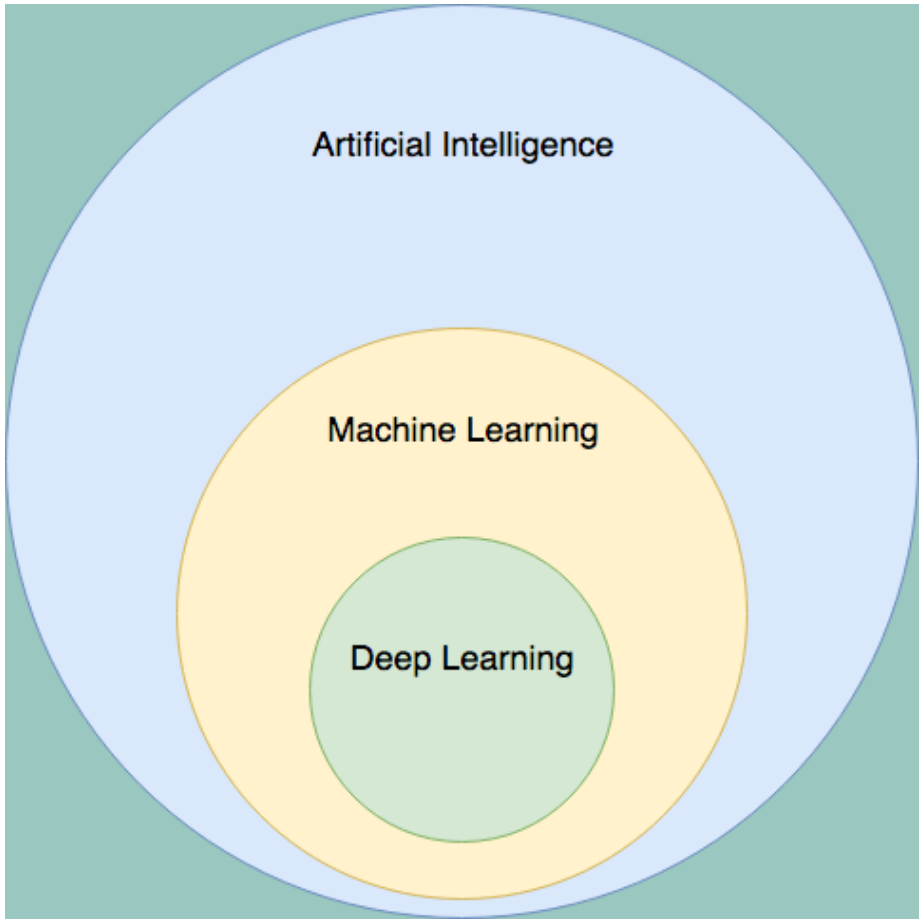


What is the intelligence?

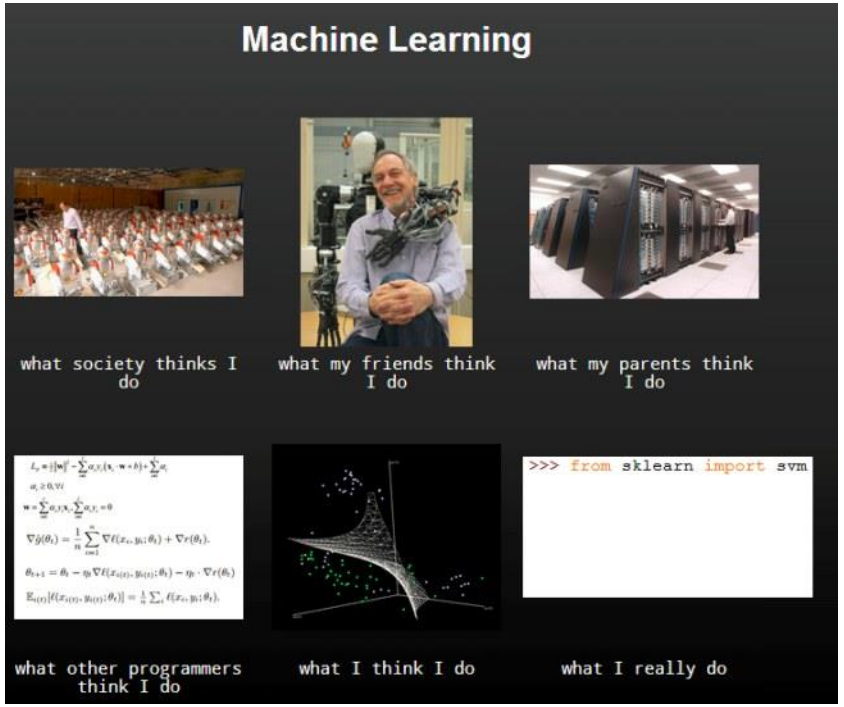
Machine Learning

What is Machine Learning?

Venn Diagram of Machine Learning



Definition of Machine Learning



- Tom Mitchell(1998) : A computer is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, with experience E.

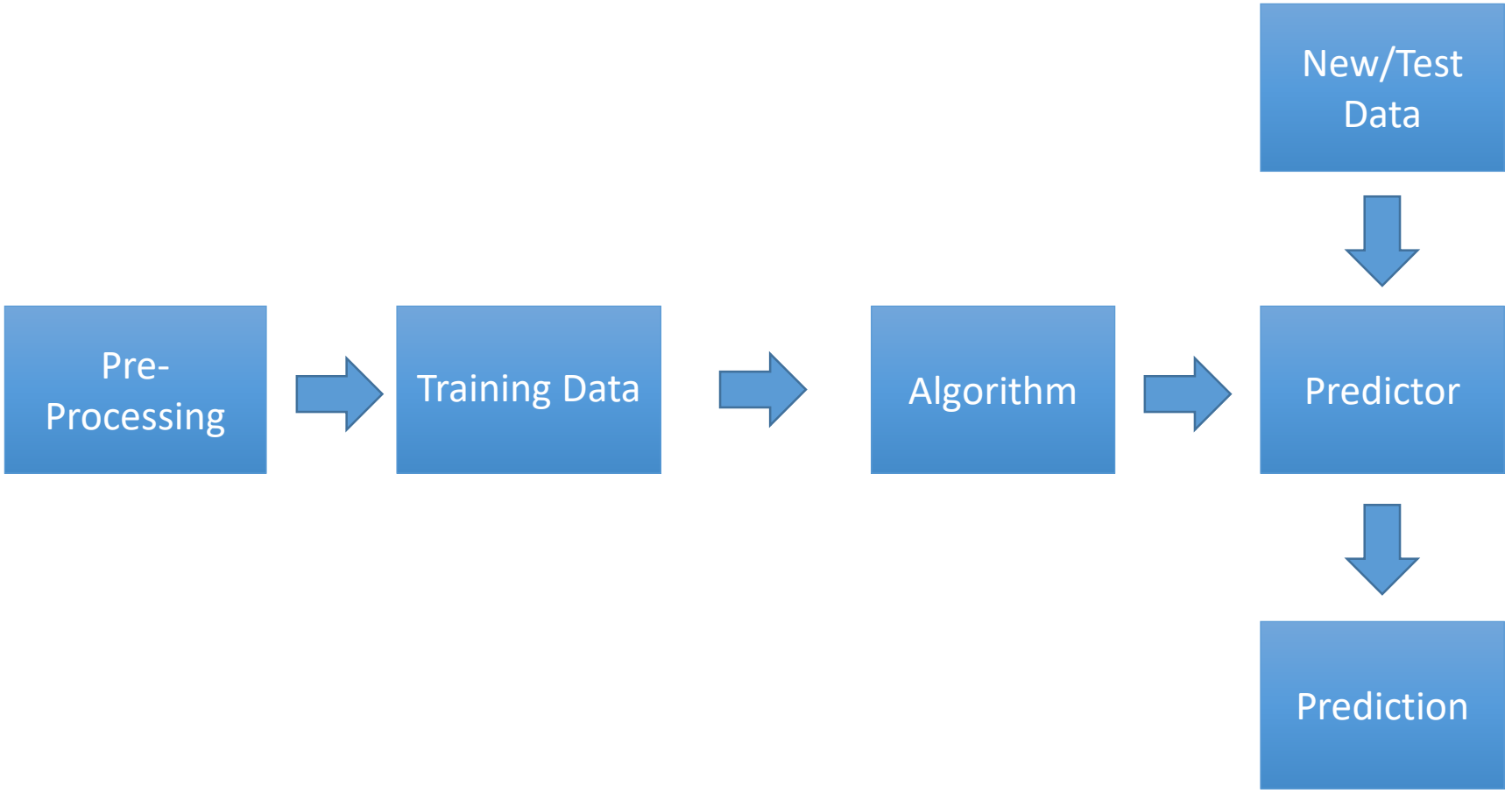
Why Machine Learning?

- 您購買的東西已經送達超商，請您來領取。 (Spam? Not Spam?)
- 我們這有非常好吃的蔬菜，請您來購買。(Spam? Not Spam?)

Types of Machine Learning

- Supervised Learning
  - Regression
  - Classification
- Unsupervised Learning
  - Clustering
- Reinforcement Learning
- Self-Supervised Learning (More Recently)

Flow of Machine Learning



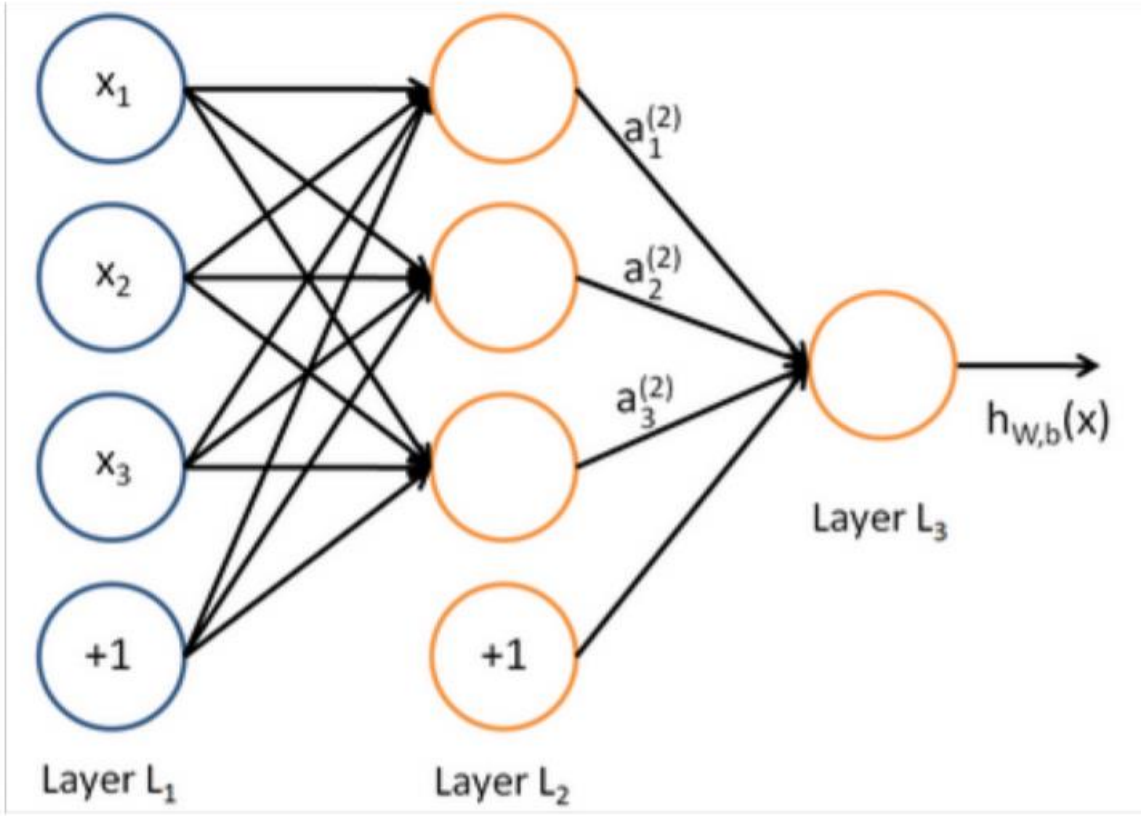
Classification

- Supervised Learning
- Trained with Labeled Data
- Algorithm
  - Naïve Bayes
  - Logistic Regression
  - KNN
  - SVM
  - Neural Network
- Application
  - Text Classification
  - Image Classification

Clustering

- Unsupervised Learning
- Trained with Unlabeled Data
- Algorithm of Clustering
  - K-means
  - Affinity Propagation
  - Neural Network
- Application
  - Recommendation System
  - Word Sense Induction

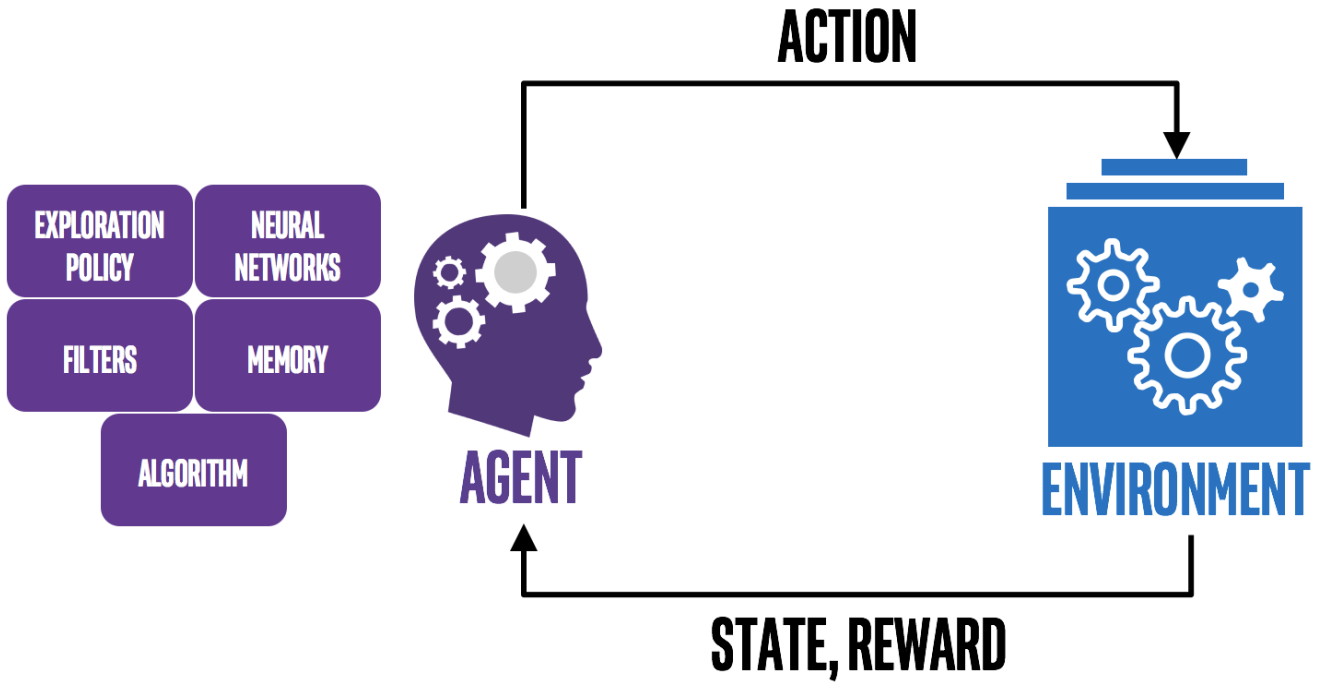
Deep Learning



$$h_{w,b}(x) = f(w^{\mathsf{T}}x + b)$$

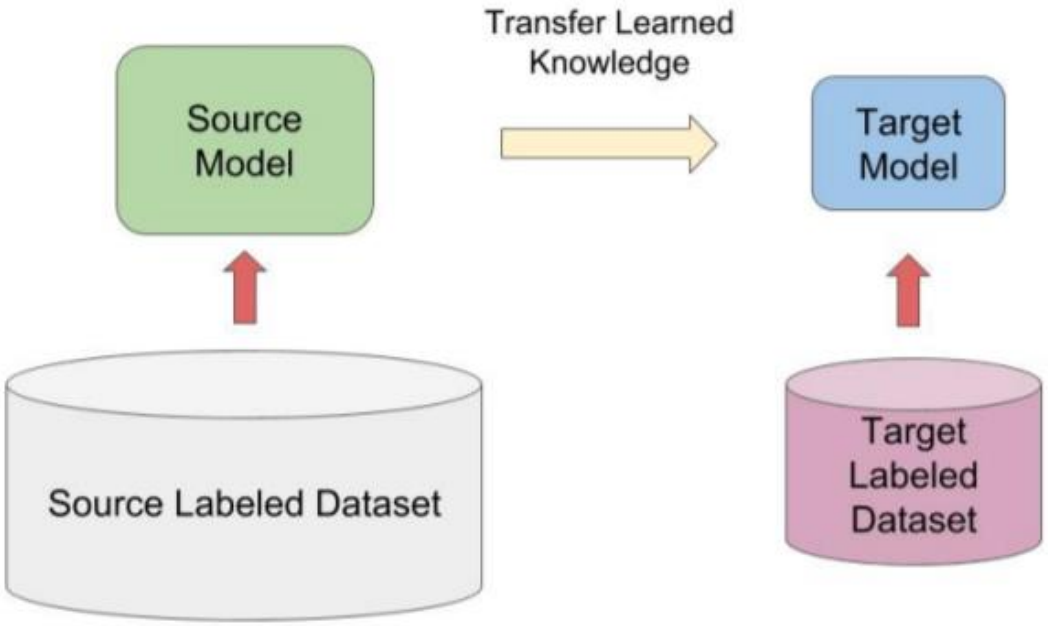
$$f(z) = \frac{1}{1+e^{-z}}$$

Reinforcement Learning



- Application
  - AlphaGo

Self-Supervised Learning

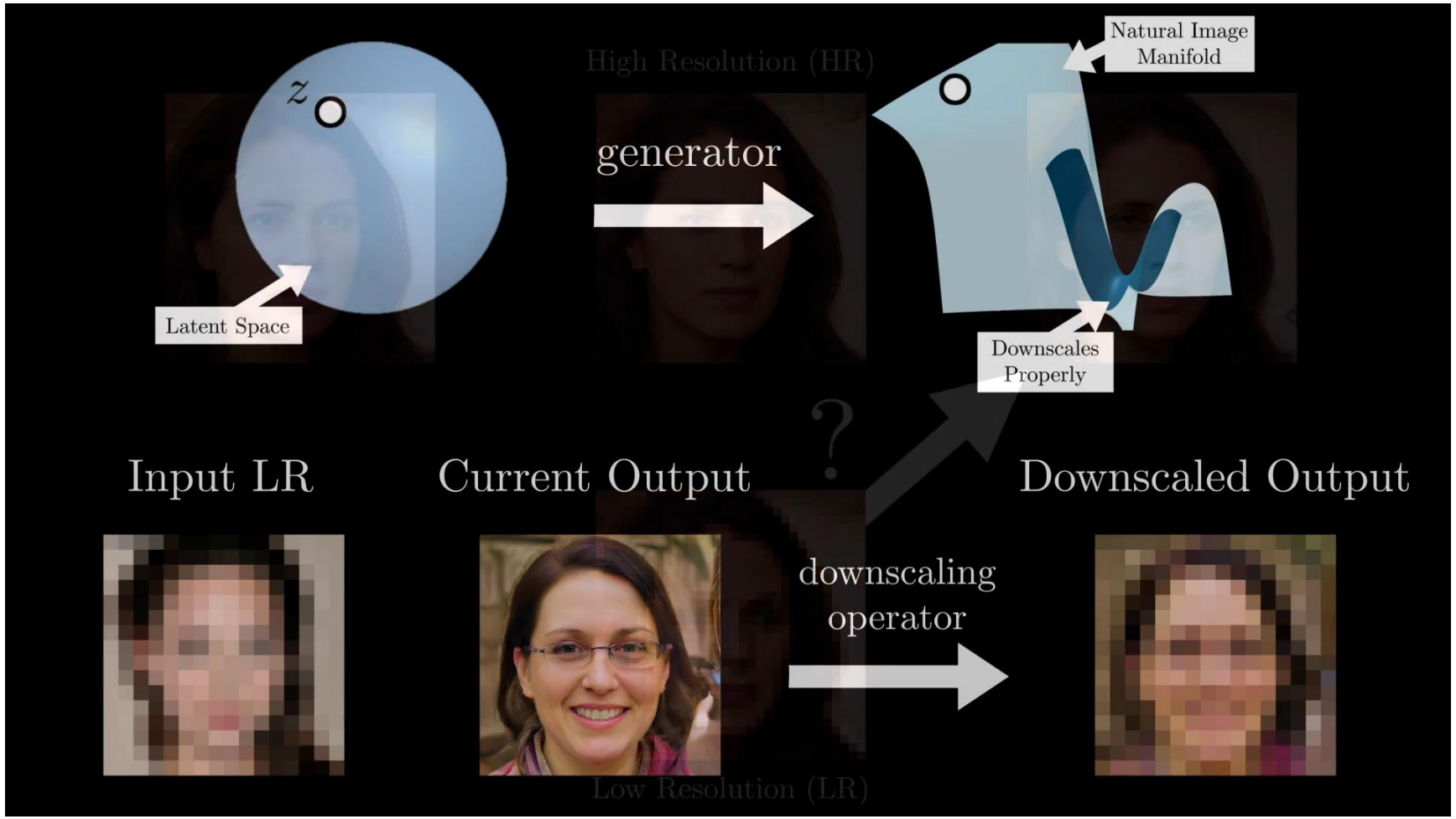


- Source: http://web.stanford.edu/class/cs224n/slides/cs224n-2020-lecture14-contextual-representations.pdf

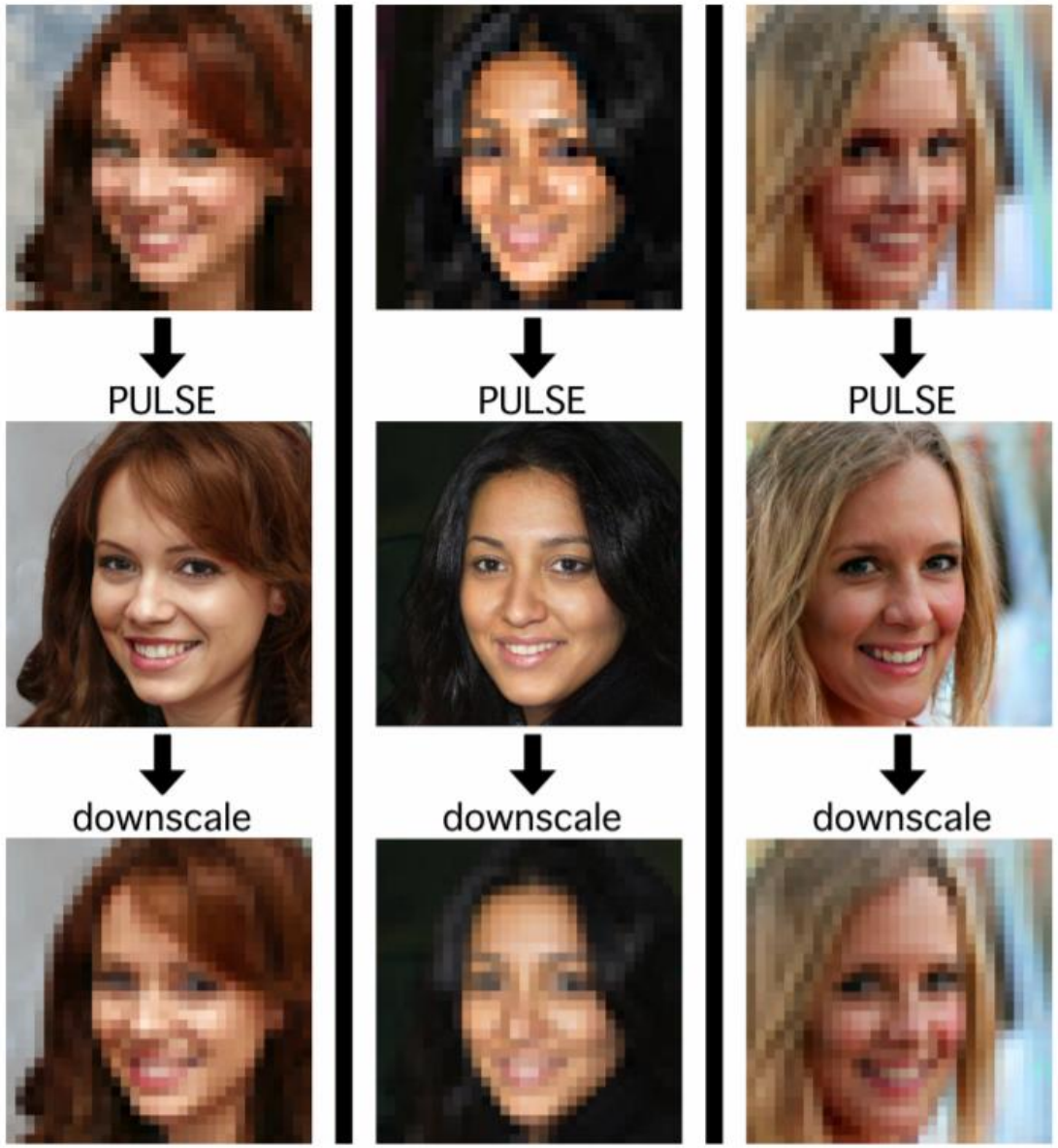
PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models

Menon et al., CVPR, 2020.

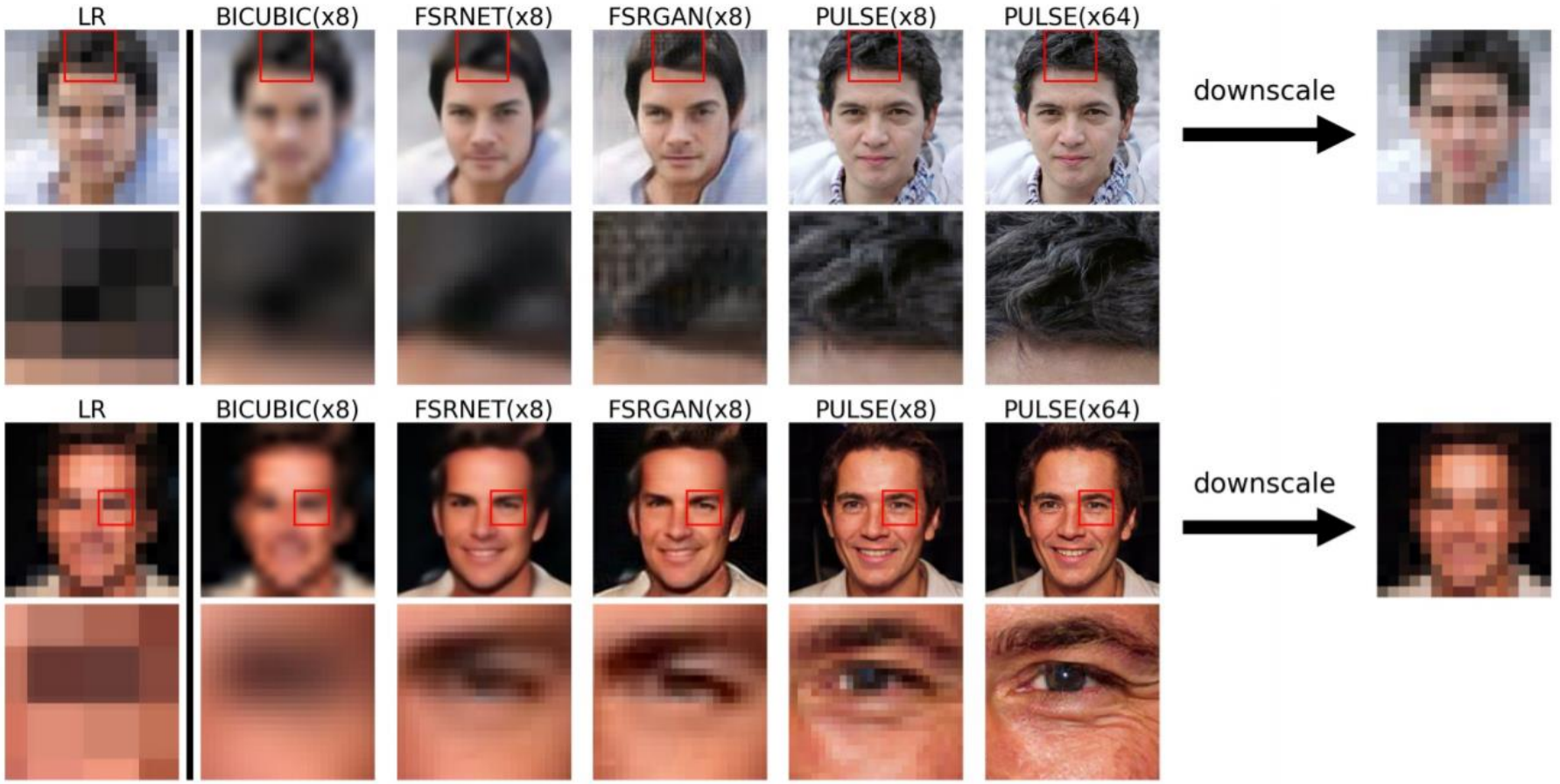
Method


High Resolution (HR)
z
Latent Space
generator
Natural Image Manifold
Downscales Properly
?
Input LR
Current Output
Downscaled Output
downscaling operator
Low Resolution (LR)

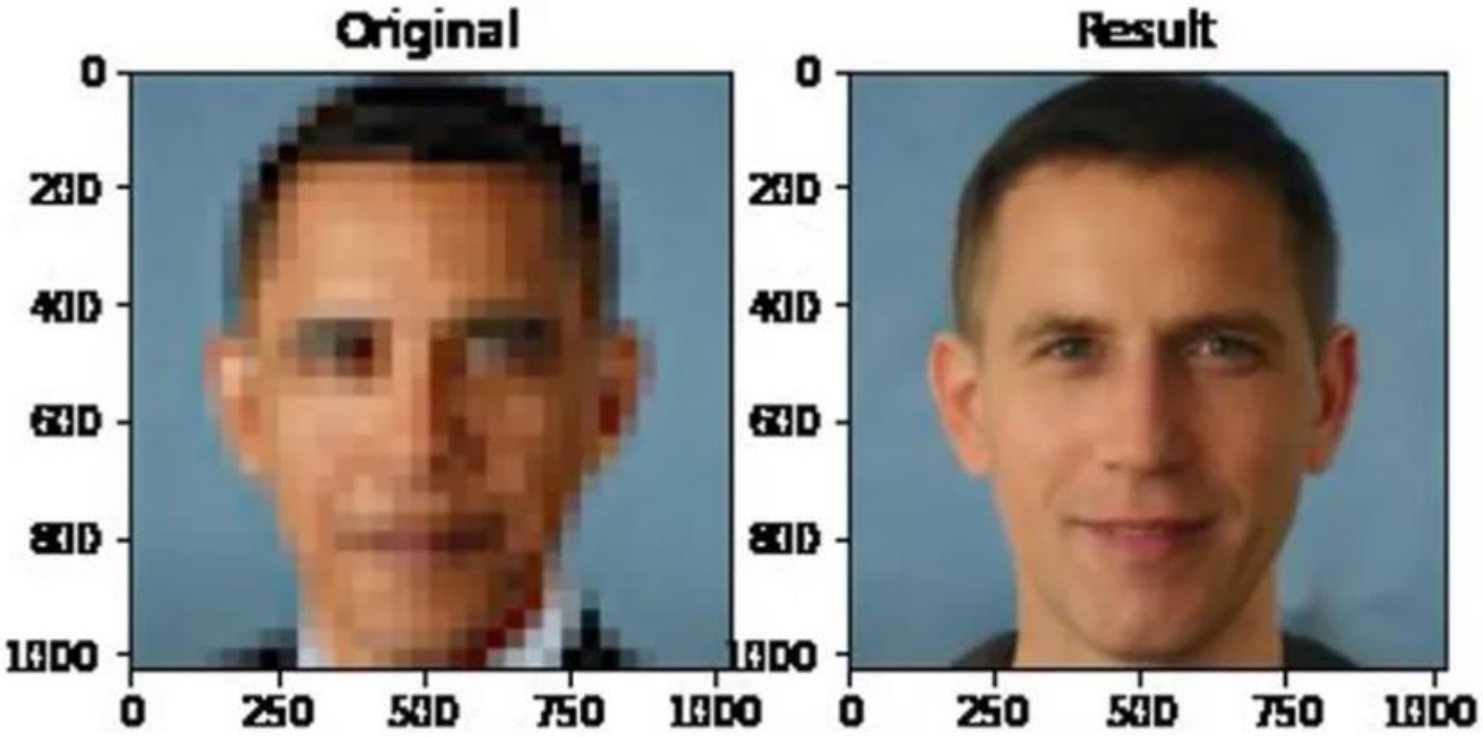
Results



Comparison with Other Methods



Bias in this Research (Obama)



- Source: https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias

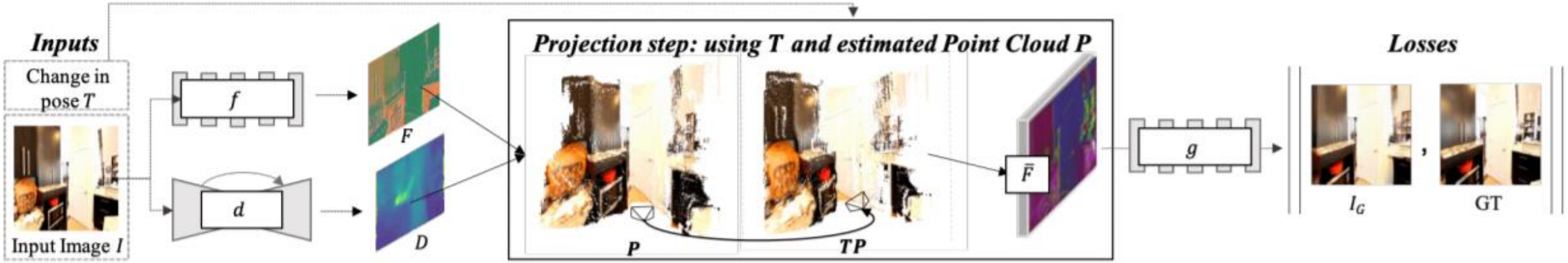
Reasons to the Bias

- Researchers
- Algorithm
- Data

SynSin: End-to-end View Synthesis from a Single Image

Wiles et al., CVPR, 2020.

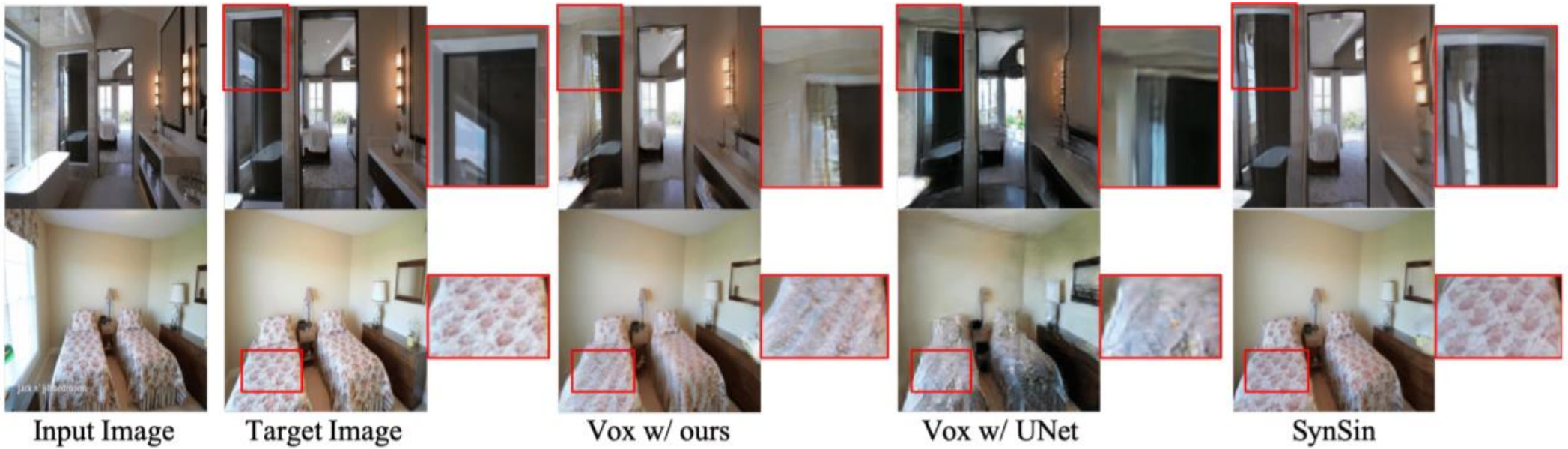
Method


Inputs
Change in pose T
Input Image I
f
d
F
D
Projection step: using T and estimated Point Cloud P
P
TP
$\bar{F}$
g
Losses
I_G
GT

Results


Input Image
Learned 3D point cloud with trajectory overlaid
Generated views along the trajectory
Input Image
Learned 3D point cloud with trajectory overlaid
Generated views along the trajectory

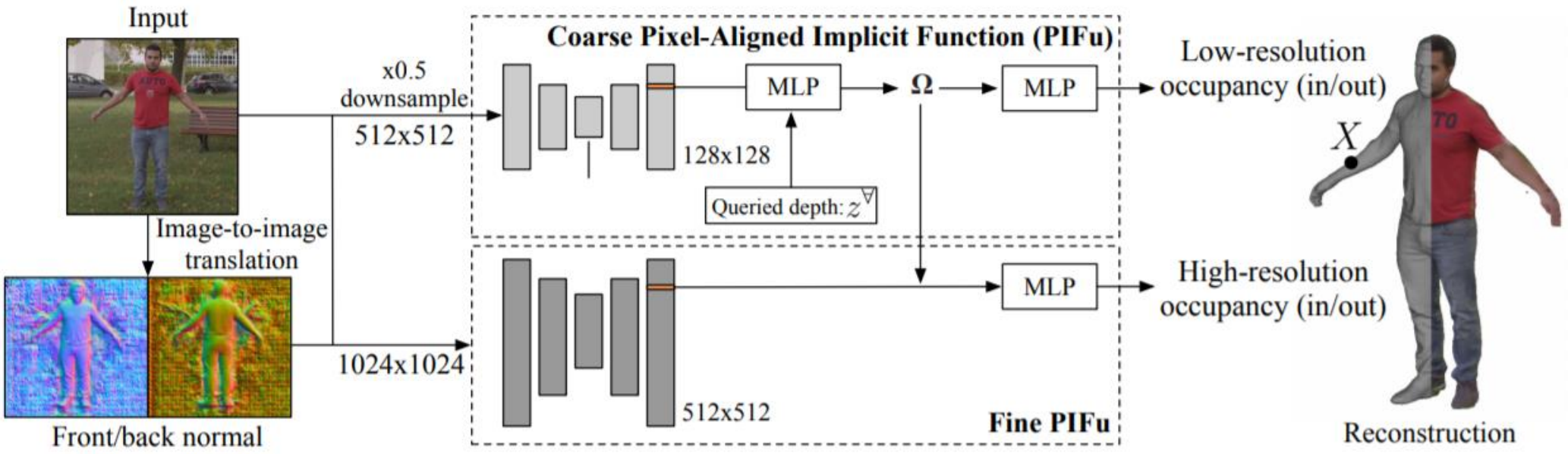
Comparison with other Methods



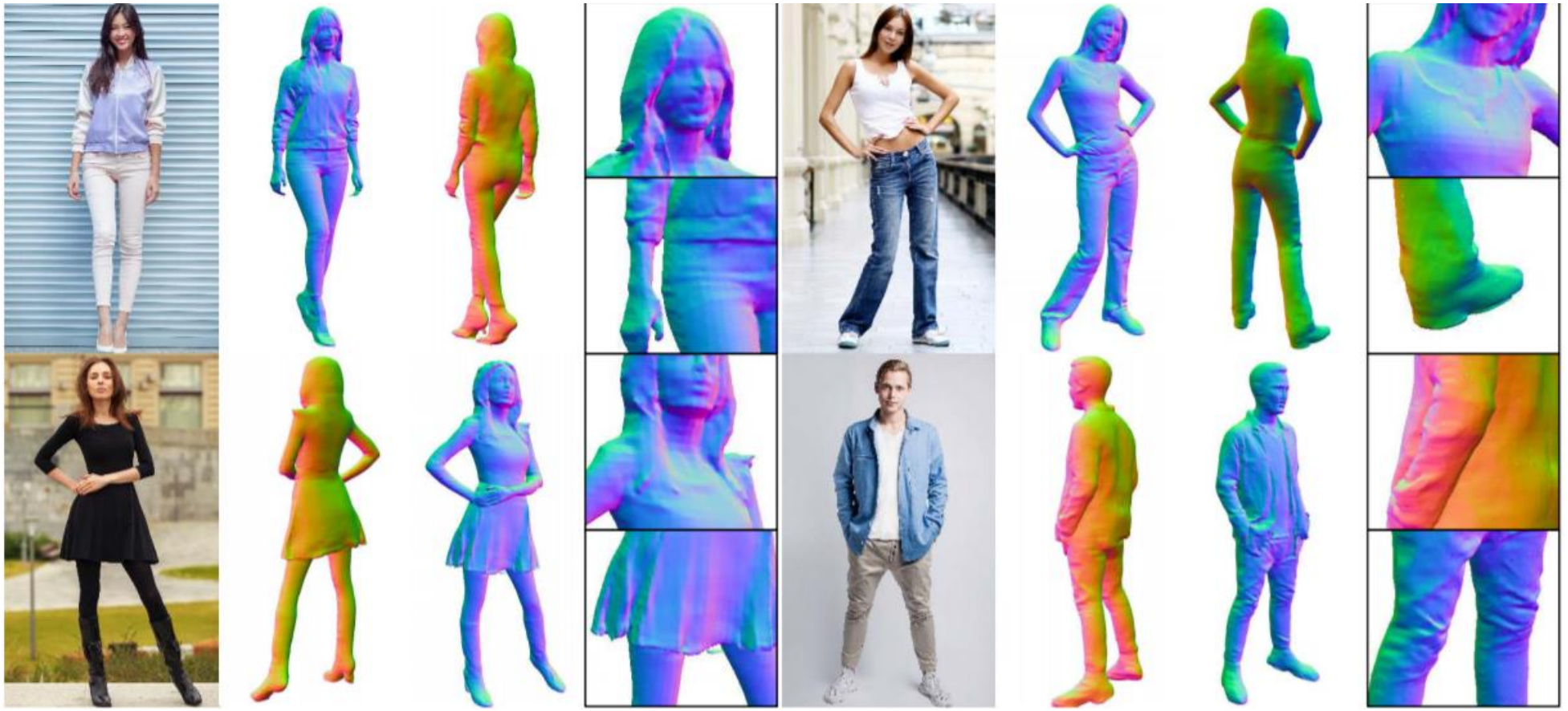
PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization

Saito et al., CVPR, 2020.

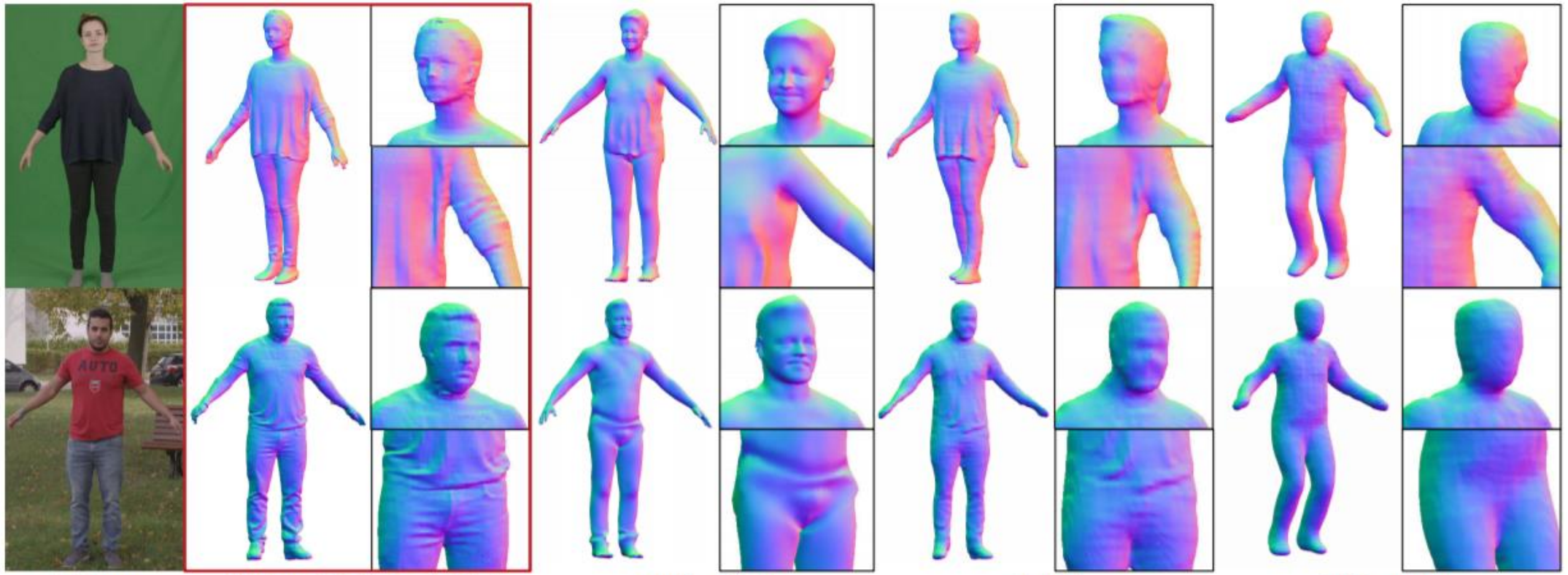
Methods


Input
x0.5
downsample
512x512
Image-to-image
translation
1024x1024
Front/back normal
Coarse Pixel-Aligned Implicit Function (PIFu)
MLP
Ω
MLP
128x128
Queried depth: z
MLP
512x512
Fine PIFu
Low-resolution
occupancy (in/out)
High-resolution
occupancy (in/out)
X
Reconstruction

Results



Comparison with other Methods



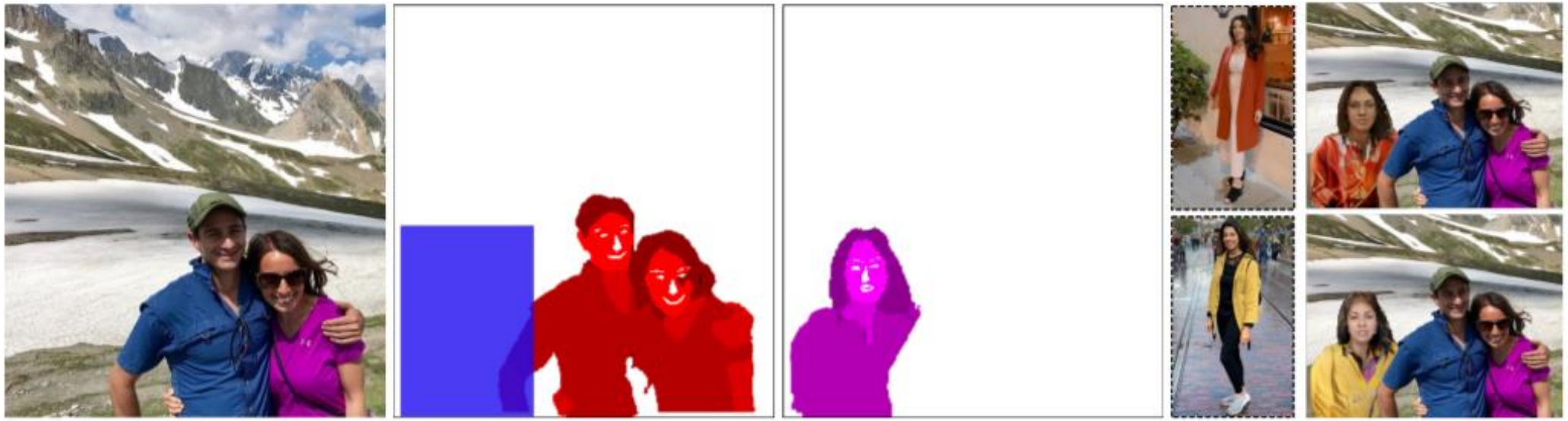
Wish You Were Here: Context-Aware Human Generation

Gafni et al., CVPR, 2020.

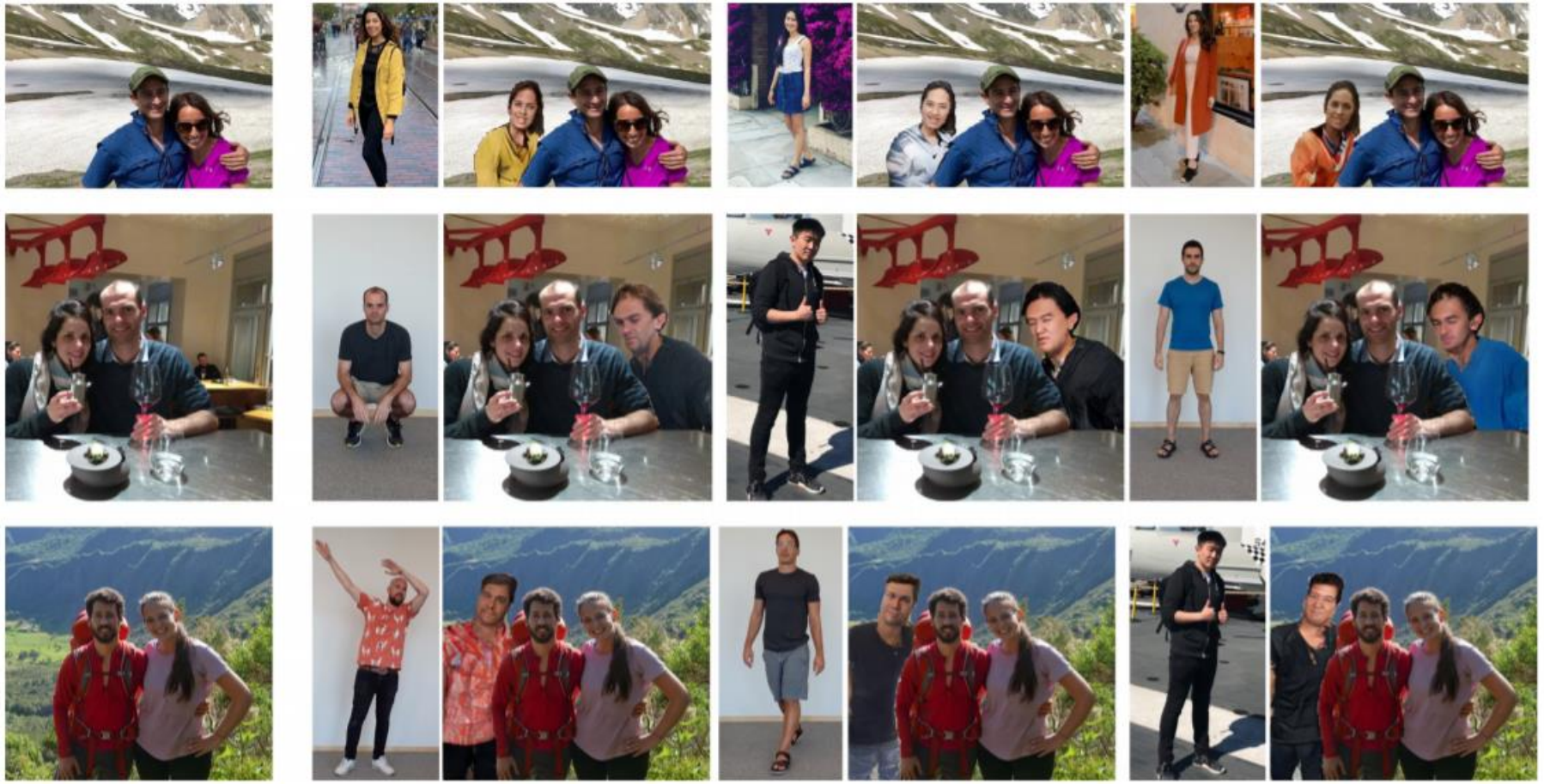
Methods


Essence Generation Network
Multi-Conditioning Rendering Network
Face Refinement Network

Methods



Result (Add person to source image)



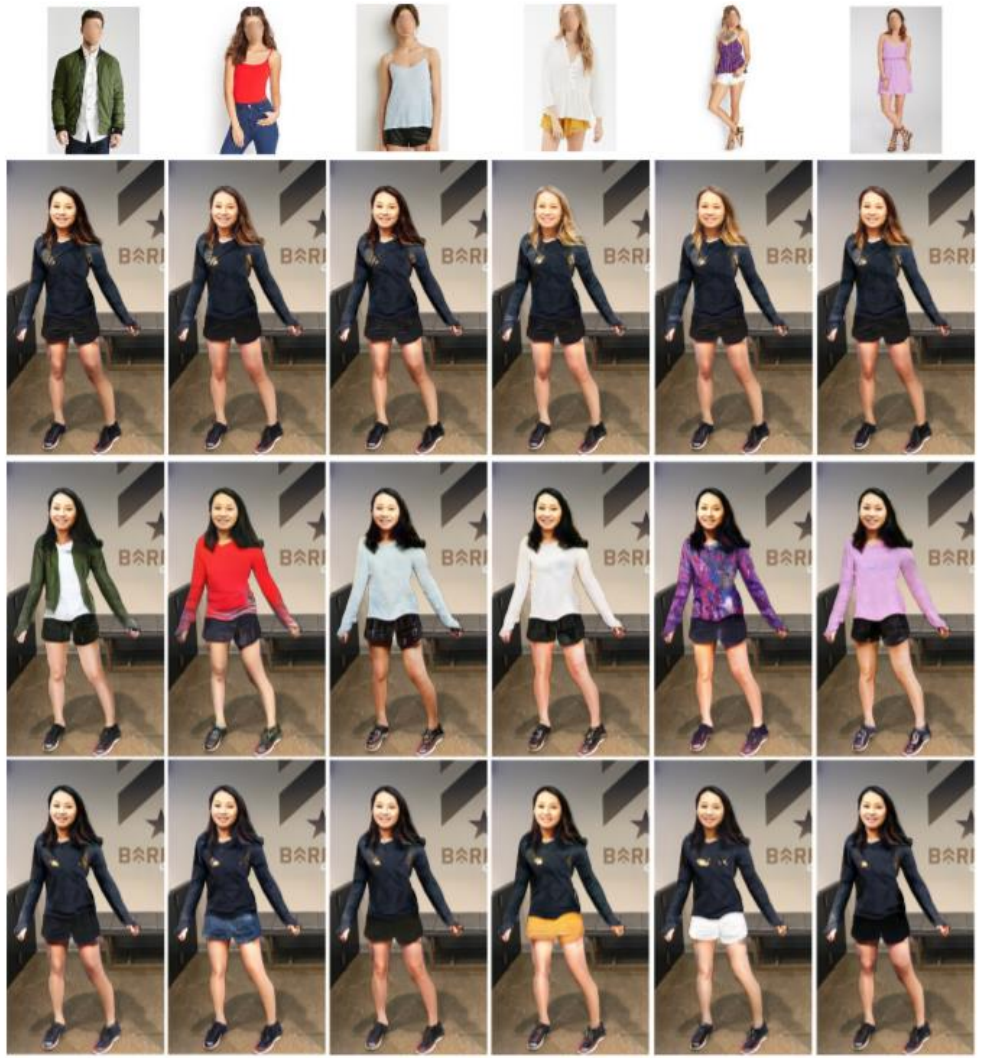
Result (Replace with hair, shirt and pants)

Target

Hair

Shirt

Pants



Detectron2: A PyTorch-based modular object detection library

Facebook

DEMO Videos

- https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library-/

Introduction

- Framework
  - Pytorch
- Modular, Extensible Design
- Models
  - Faster R-CNN, Mask R-CNN, RetinaNet, DensePose, Cascade R-CNN, Panoptic FPN, and TensorMask etc.
- Tasks: Object Detection with Box, Instance Segmantation Masks, Human Pose Prediction, Sematic Segmentation, Panoptic Segmentation

Semantic, Instance, Panoptic Segmentation

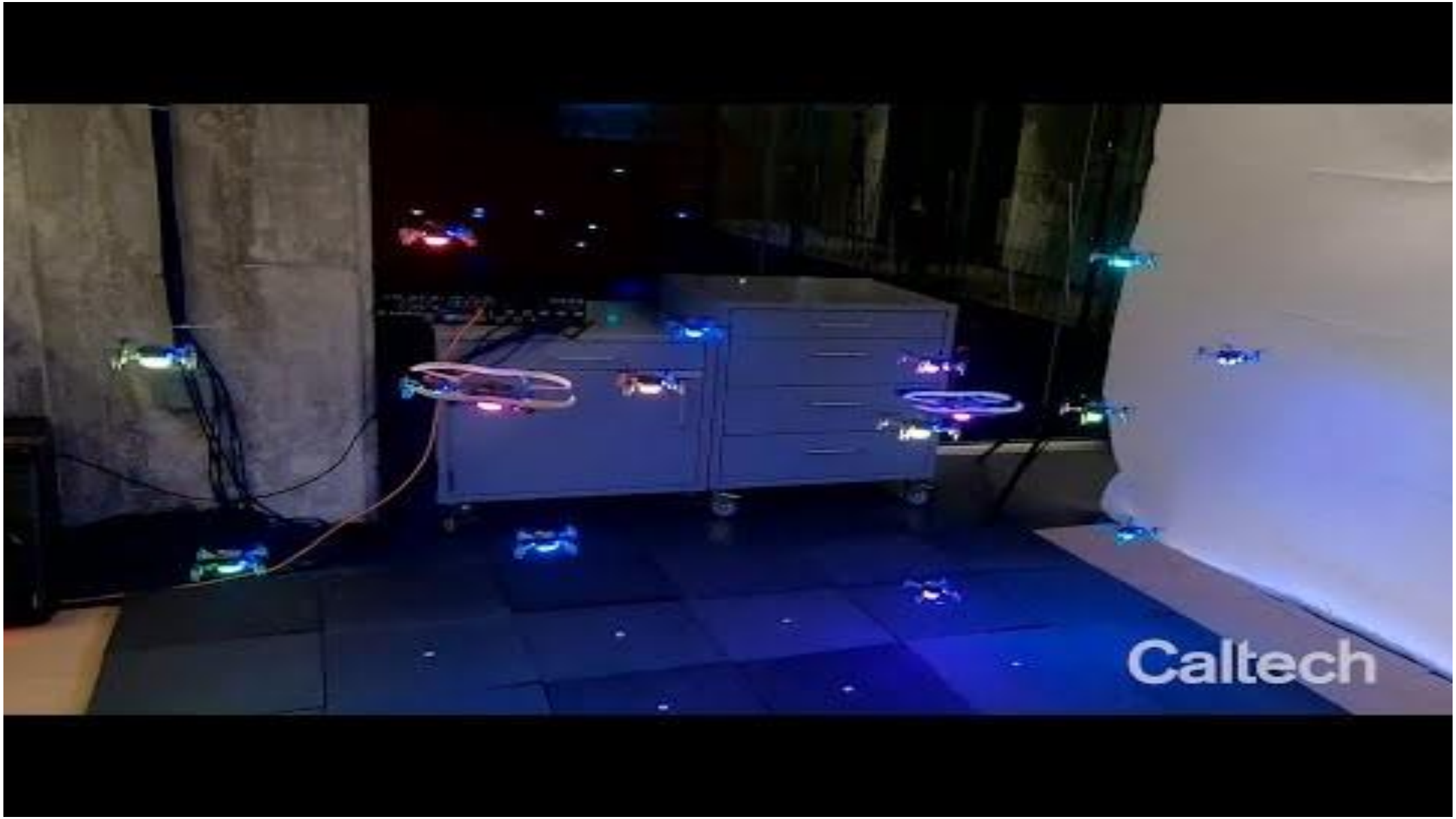
- Semantic Segmentation
  - Classify the objects into right Category
- Instance Segmentation
  - Segment each object separately
- Panoptic Segmentation
  - The combination of Semantic Segmentation and Instance Segmentation

Github Link

- https://github.com/facebookresearch/detectron2

Close-Proximity Flight of Sixteen Quadrotor Drones

CalTech


Caltech

References

- Rethinking Weak Vs. Strong AI, Forbes.
- Detectron2: A PyTorch-based modular object detection library, Facebook Blog.
- Percy Liang and Dorsa Sadigh, CS221: Artificial Intelligence: Principles and Techniques, Stanford University.
- Introduction to Panoptic Segmentation: A Tutorial, Technical Fridays.