CISC6000 Deep Learning
Introduction

Dr. Yijun Zhao

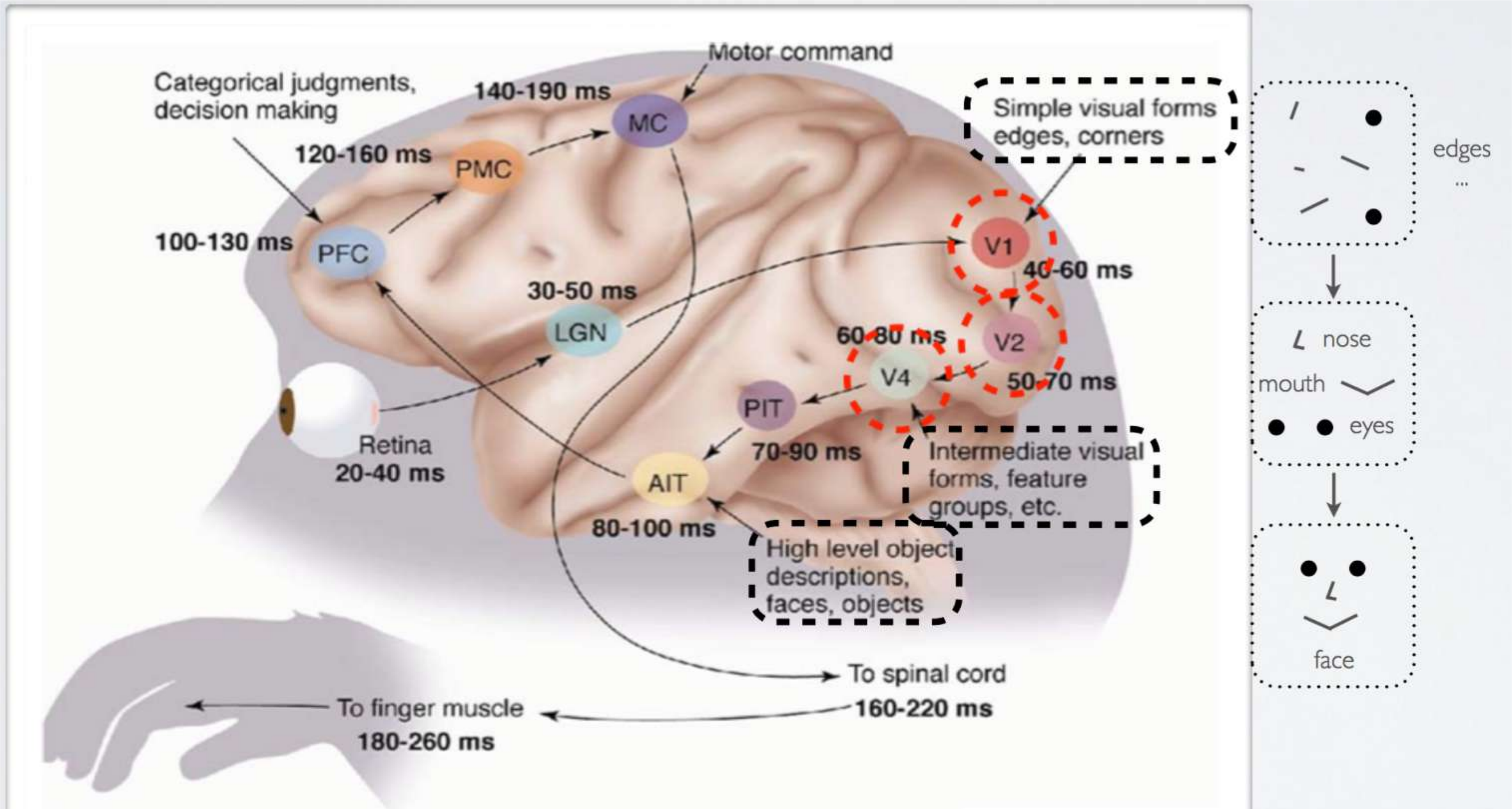
Fordham University

Puppy or bagel

Sheepdog or mop

Chihuahua or muffin

Deep Learning



[picture from Simon Thorpe]

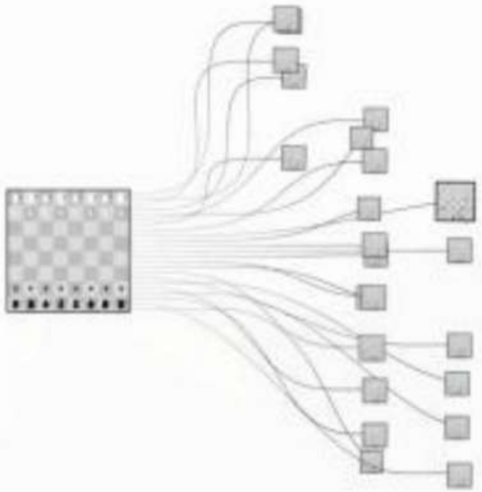
Why is computer vision hard?

The camera sees :

194	210	201	212	199	213	215	195	178	158	182	209
180	189	190	221	209	205	191	167	147	115	129	163
114	126	140	188	176	165	152	140	170	106	78	88
87	103	115	154	143	142	149	153	173	101	57	57
102	112	106	131	122	138	152	147	128	84	58	66
94	95	79	104	105	124	129	113	107	87	69	67
68	71	69	98	89	92	98	95	89	88	76	67
41	56	68	99	63	45	60	82	58	76	75	65
20	43	69	75	56	41	51	73	55	70	63	44
50	50	57	69	75	75	73	74	53	68	59	37
72	59	53	66	84	92	84	74	57	72	63	42
67	61	58	65	75	78	76	73	59	75	69	50

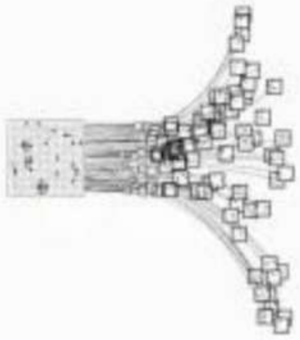
A computer first beat a chess world champion in 1997

We human lose on Go!



Chess: 10^{47}

Deep Blue, Feb 10, 1996



Go: 10^{170}

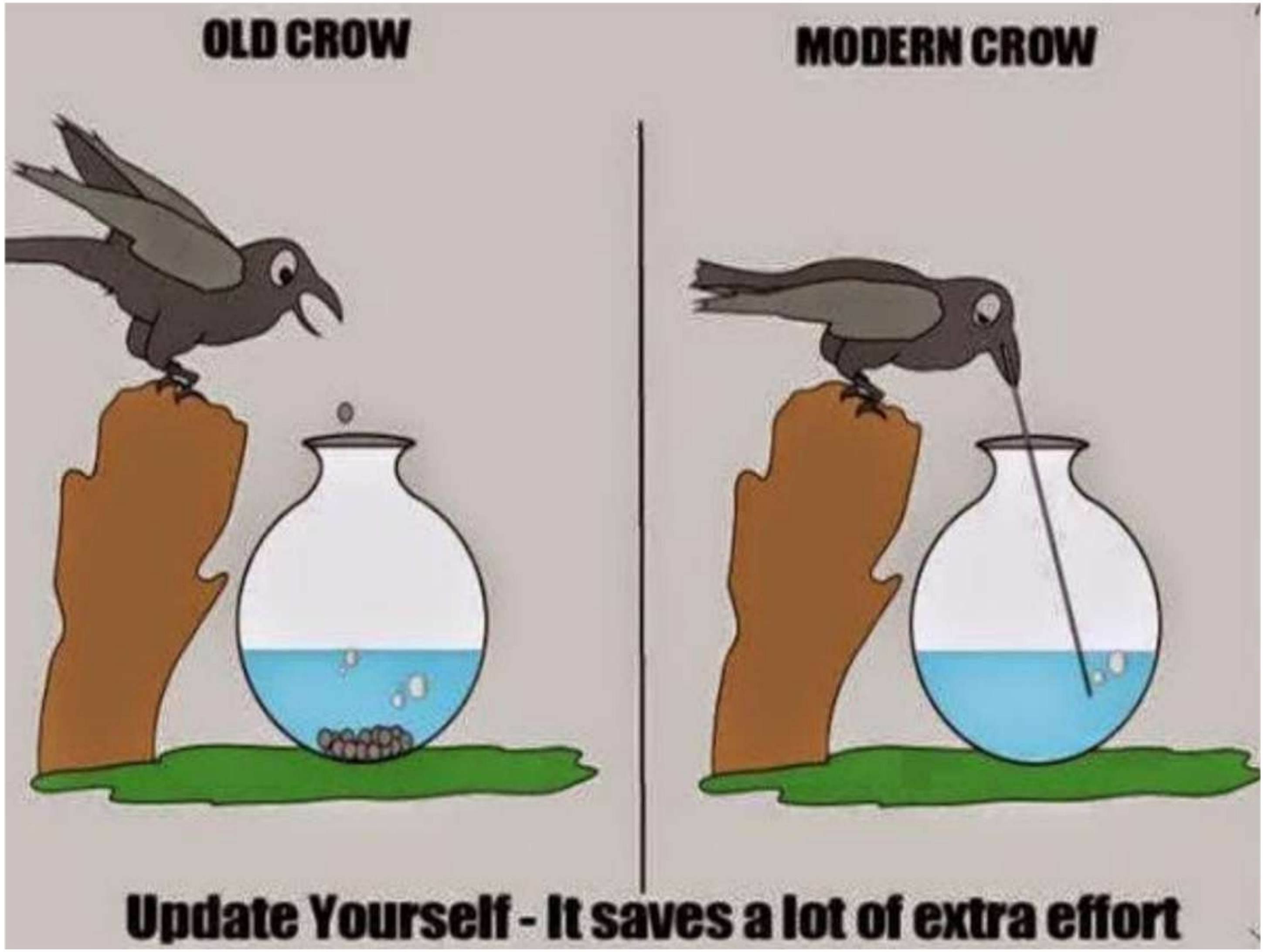
AlphaGo, March, 2016

We (will) lose on many specific tasks!

- Speech recognition
- Translation
- Self-driving
- …

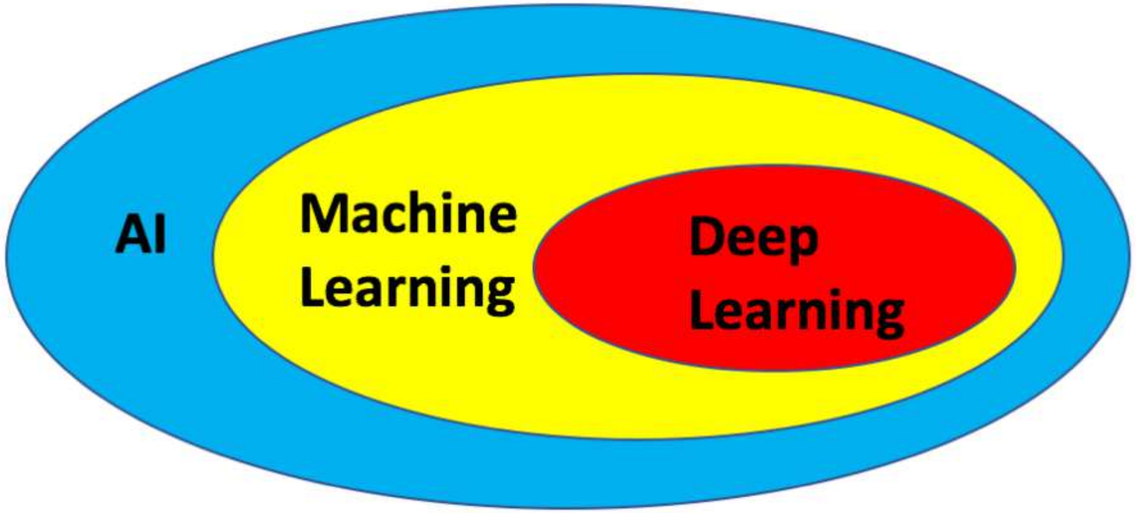
We (will) lose on many specific tasks!

- Speech recognition
- Translation
- Self-driving
- …



Deep Learning

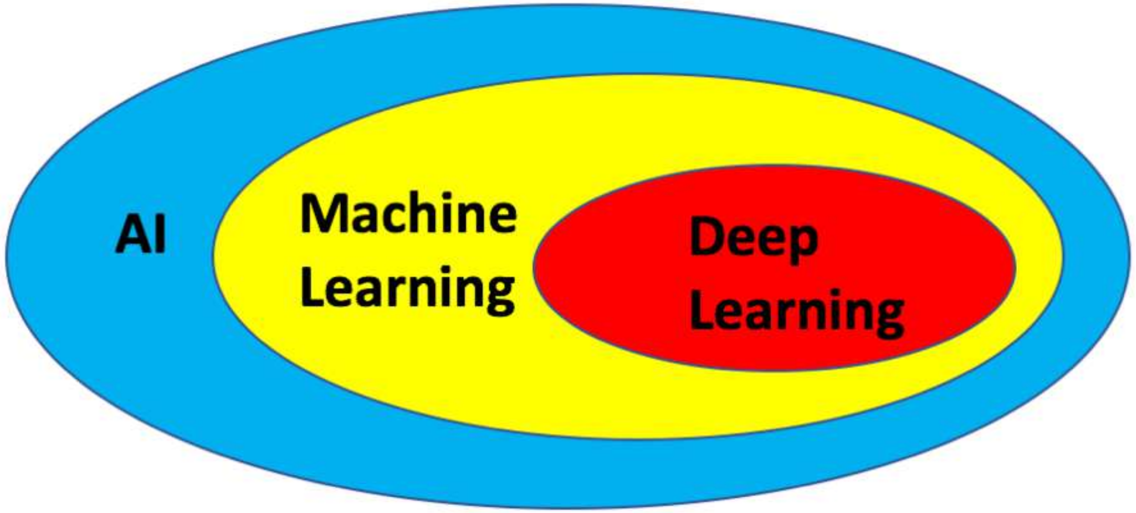
- **Artificial Intelligence**:
  - The study of how to create intelligent agents: speech recognition, robots
  - Uses models built by machine learning
- **Machine Learning**:
  - Supervised learning
  - Unsupervised learning
  - Semi-supervised learning
  - Reinforcement learning
  - Deep learning



Deep Learning

- **Artificial Intelligence**:
  - the study of how to create intelligent agents: speech recognition, robots
  - uses models built by machine learning
- **Machine Learning**:
  - Supervised learning
  - Unsupervised learning
  - Semi-supervised learning
  - Reinforcement learning

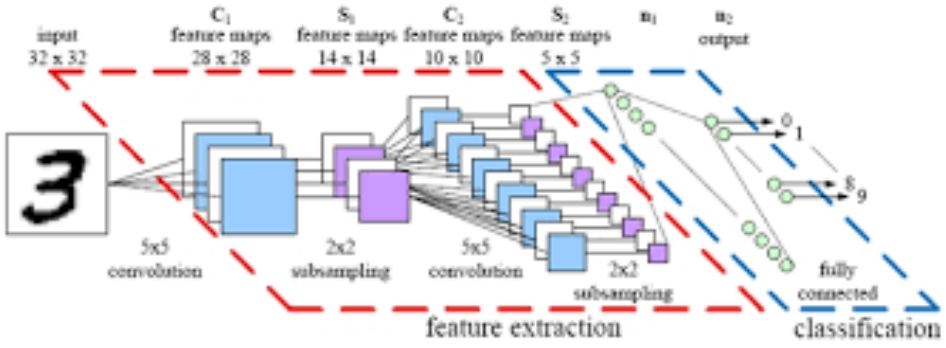
  - Deep learning
  - Collaborative learning
  - Online learning
  - Adversarial learning (GANs)



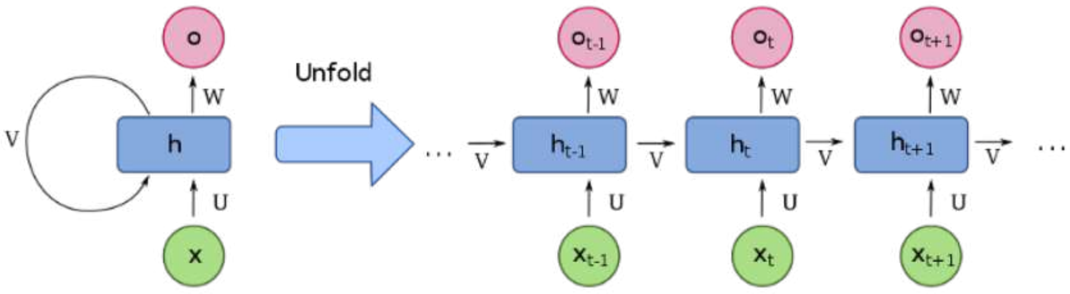
Deep Learning

Types of DL architectures covered in this course:

- **Convolutional Neural Network (CNN)**
- Autoencoder (AE)
- Recurrent Neural Network (RNN)
- Long Short-term Memory (LSTM)
- Gated Recurrent Units (GRU)
- Generative Adversarial Network (GAN)



CNN

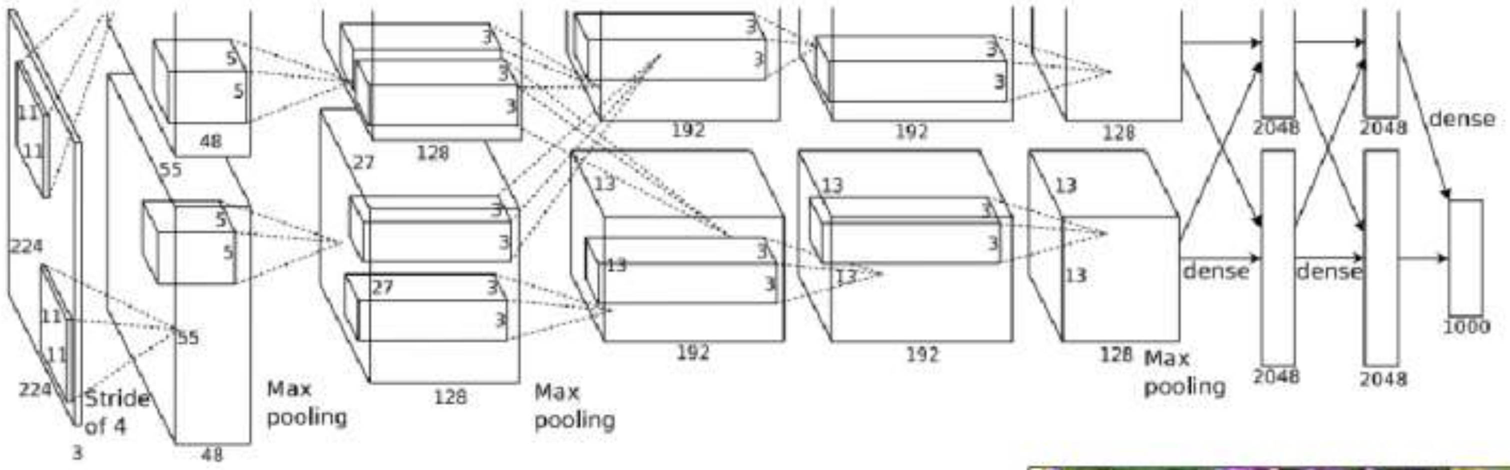


RNN

Deep Learning

Convolutional Neural Network

- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." 2012



ImageNet Competition:

- Total number of images: 14,197,122
- > 20,000 categories.

The network achieved a top-5 error of 15.4%, more than 10.8 percentage points lower than that of the runner up.

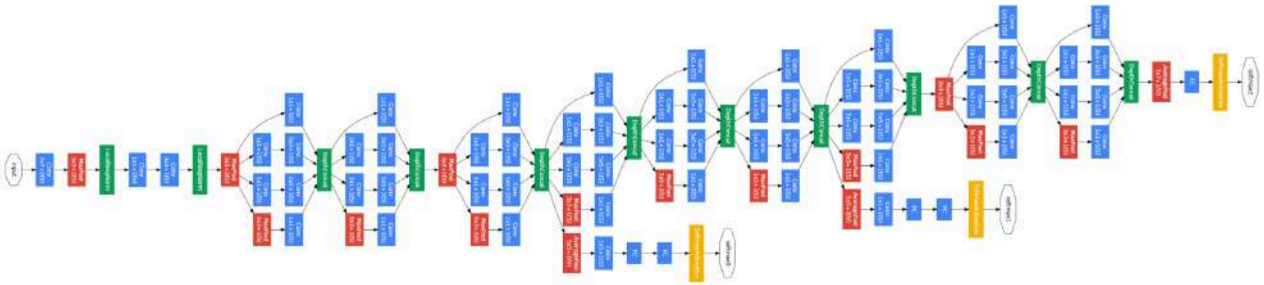
Deep Learning

Convolutional Neural Network

- VGG (Simonyan et al., 2014)

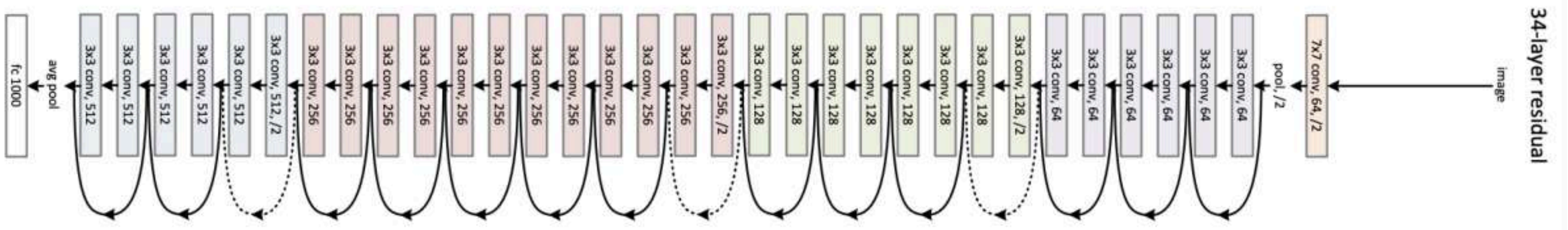
top-5 error: 7.3

- GoogleNet (Szegedy et al,, 2015)



top-5 error: 6.7

- ResNet (He et al., 2015)



top-5 error: 3.6
Better than human (5.1%)

ImageNet Competition:

- Total number of images: 14,197,122
- > 20,000 categories.

Deep Learning Applications

Convolutional Neural Network

- Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." *IEEE conference on computer vision and pattern recognition*. 2016

1 **Upload photo**

The first picture defines the scene you would like to have painted.

2 **Choose style**

Choose among predefined styles or upload your own style image.

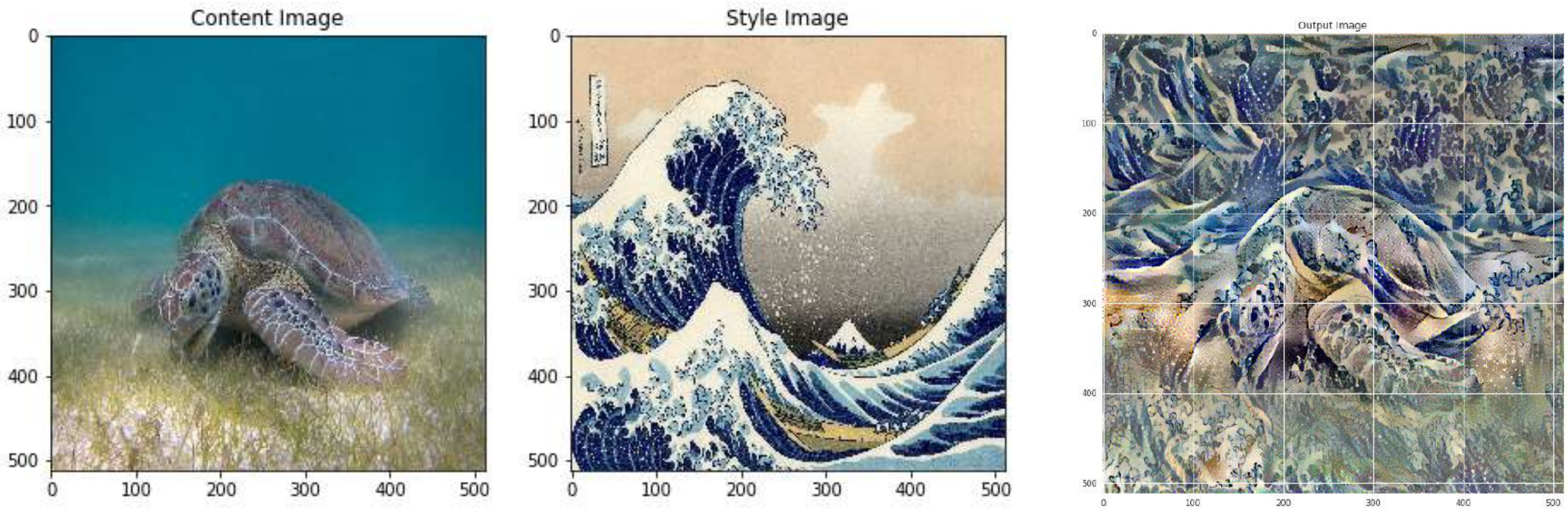
3 **Submit**

Our servers paint the image for you. You get an email when it's done.

Deep Learning Applications

Convolutional Neural Network

- Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." *IEEE conference on computer vision and pattern recognition*. 2016



Deep Learning Applications

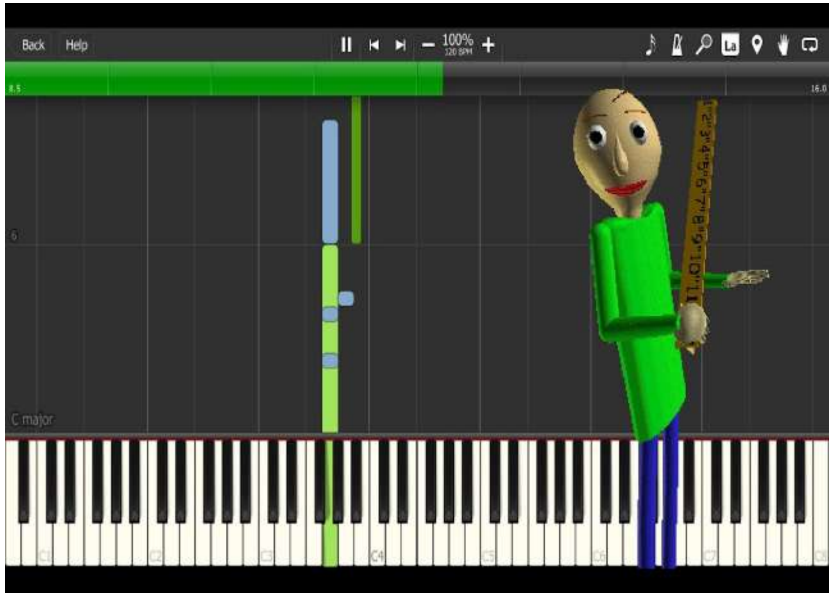
Potential Course Projects

Can you transfer music style using midi files?

- How to define a music style?
    - classic, country, rock, jazz ..etc?

- Can you classify a music with the styles?

- How to transfer the styles?



Deep Learning Applications

Convolutional Neural Network

Deep learning based super resolution: https://towardsdatascience.com/deep-learning-based-super-resolution-without-using-a-gan-11c9bb5b6cd5 2019

Deep Learning Applications

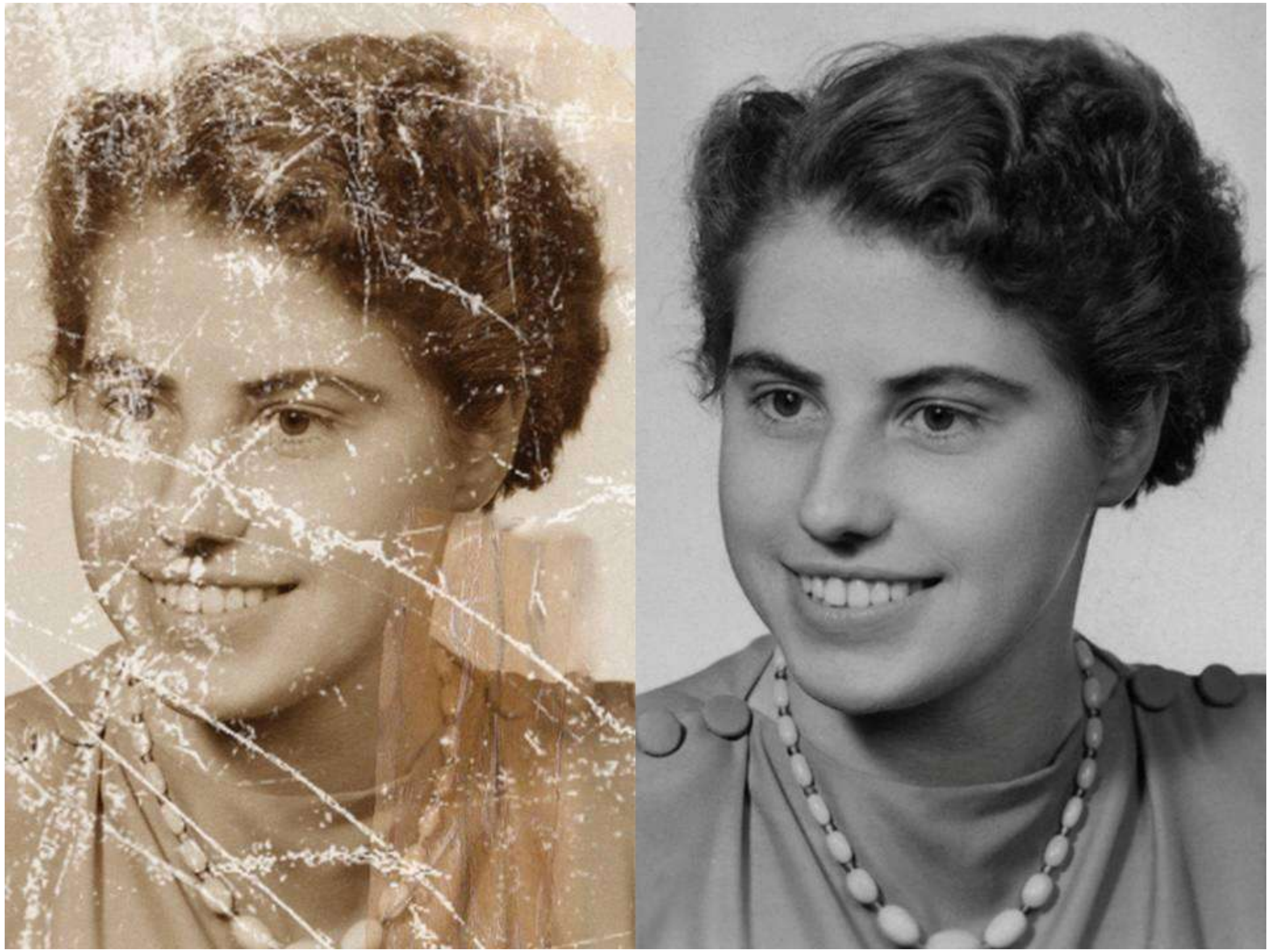
Convolutional Neural Network

Deep learning based super resolution: https://towardsdatascience.com/deep-learning-based-super-resolution-without-using-a-gan-11c9bb5b6cd5 2019

Deep Learning Applications

Convolutional Neural Network

Image Inpainting with Deep Learning 2018
https://medium.com/jamieai/image-inpainting-with-deep-learning-dd8555e56a32



Deep Learning Applications

Convolutional Neural Network

Image Inpainting with Deep Learning 2018
https://medium.com/jamieai/image-inpainting-with-deep-learning-dd8555e56a32

(

Deep Learning Applications

Convolutional Neural Network

Image Inpainting with Deep Learning

Yang, Chao, Xin Lu, Zhe Lin, Eli Shechtman, Oliver Wang, and Hao Li. "High-resolution image inpainting using multi-scale neural patch synthesis." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6721-6729. 2017.

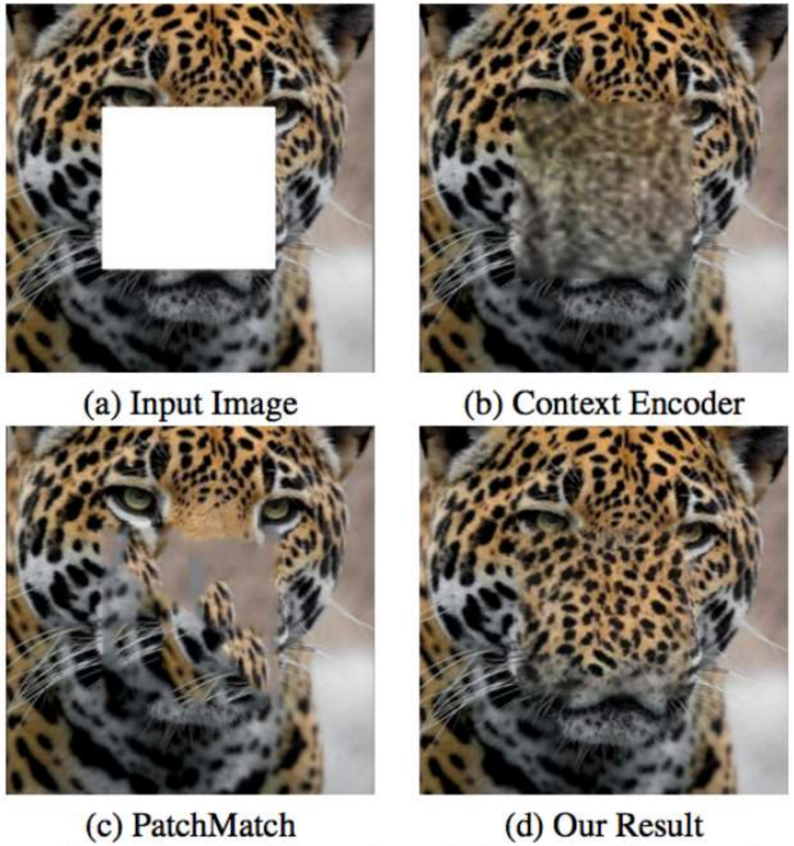
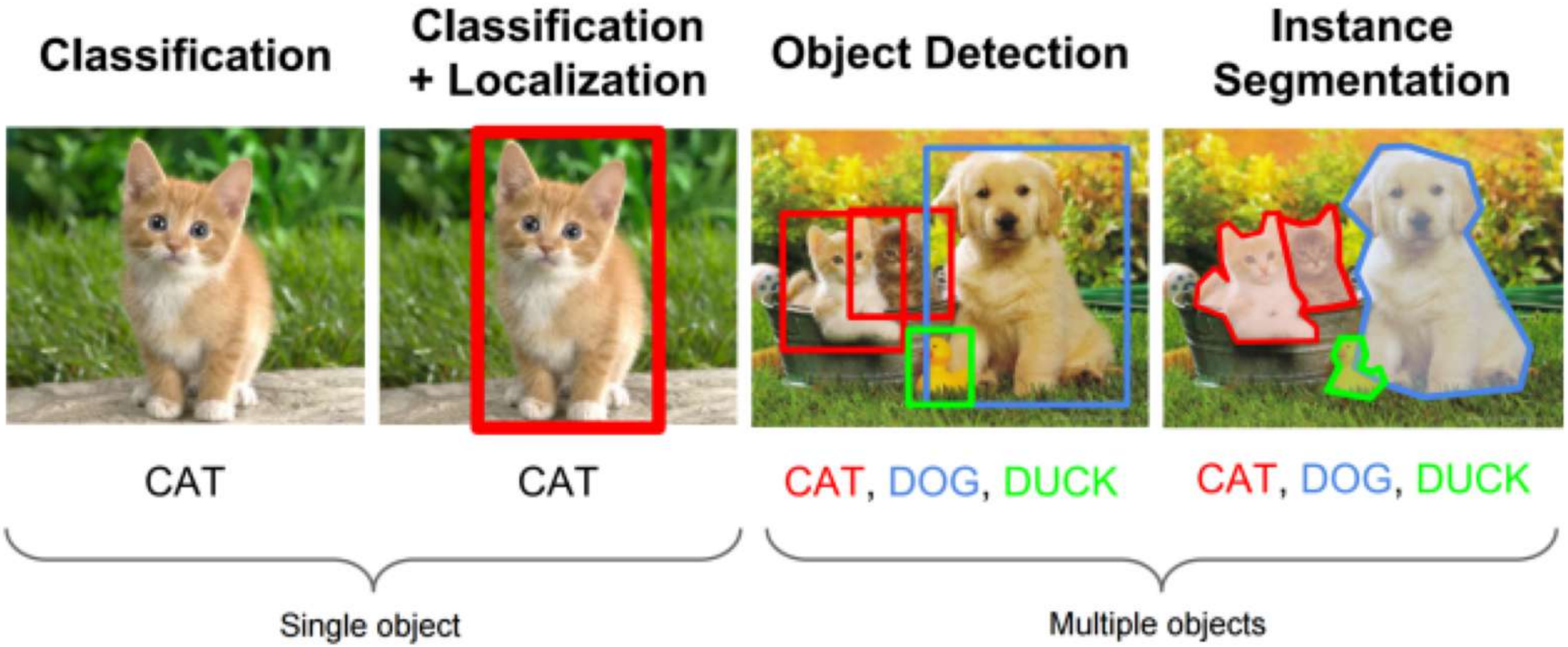


Figure 1. Qualitative illustration of the task. Given an image (512×512) with a missing hole (256×256) (a), our algorithm can synthesize sharper and more coherent hole content (d) comparing with Context Encoder [32] (b) and Content-Aware Fill using PatchMatch [1] (c).

Deep Learning Applications

Convolutional Neural Network

Object Localization: Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." *IEEE conference on computer vision and pattern recognition*. 2016 Cited by 5234



Deep Learning Applications

Convolutional Neural Network

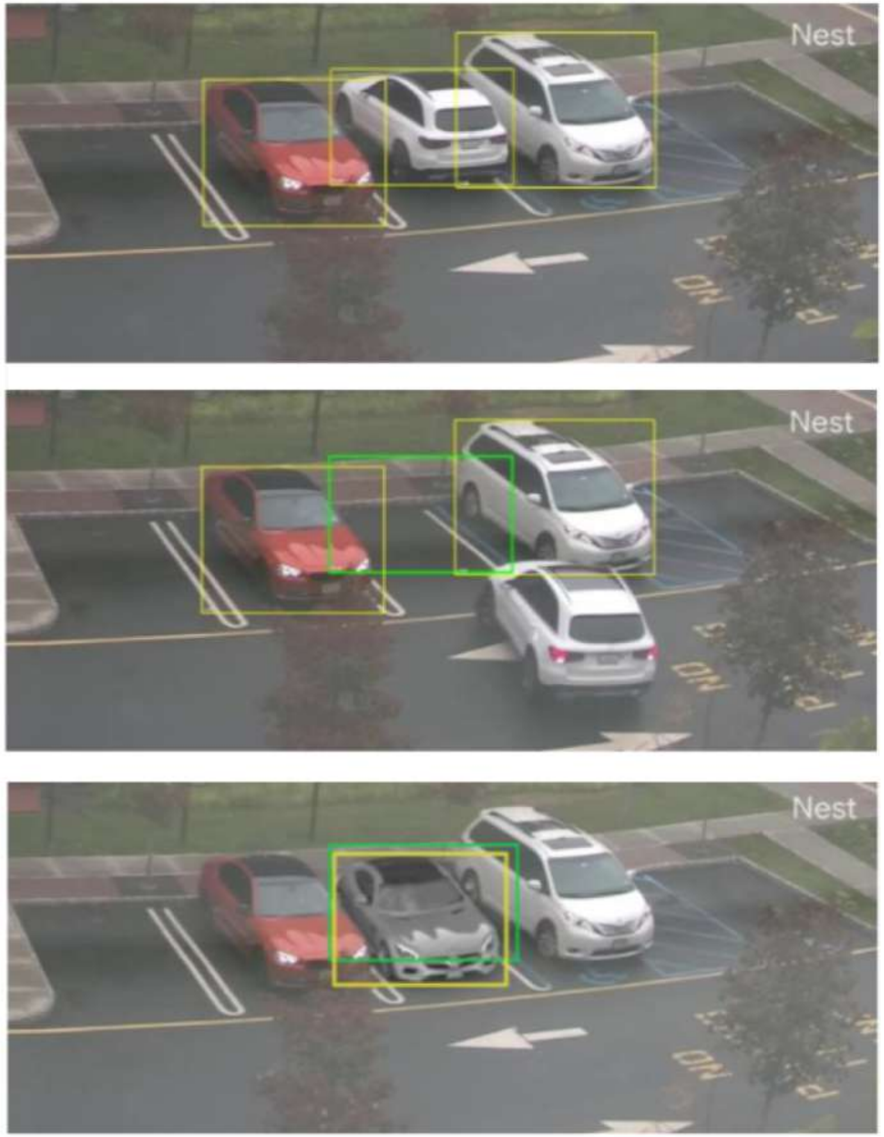
Object Localization: Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." *IEEE conference on computer vision and pattern recognition*, 2016 Cited by 5234

Deep Learning Applications

Convolutional Neural Network

Parking space detection (published course project):

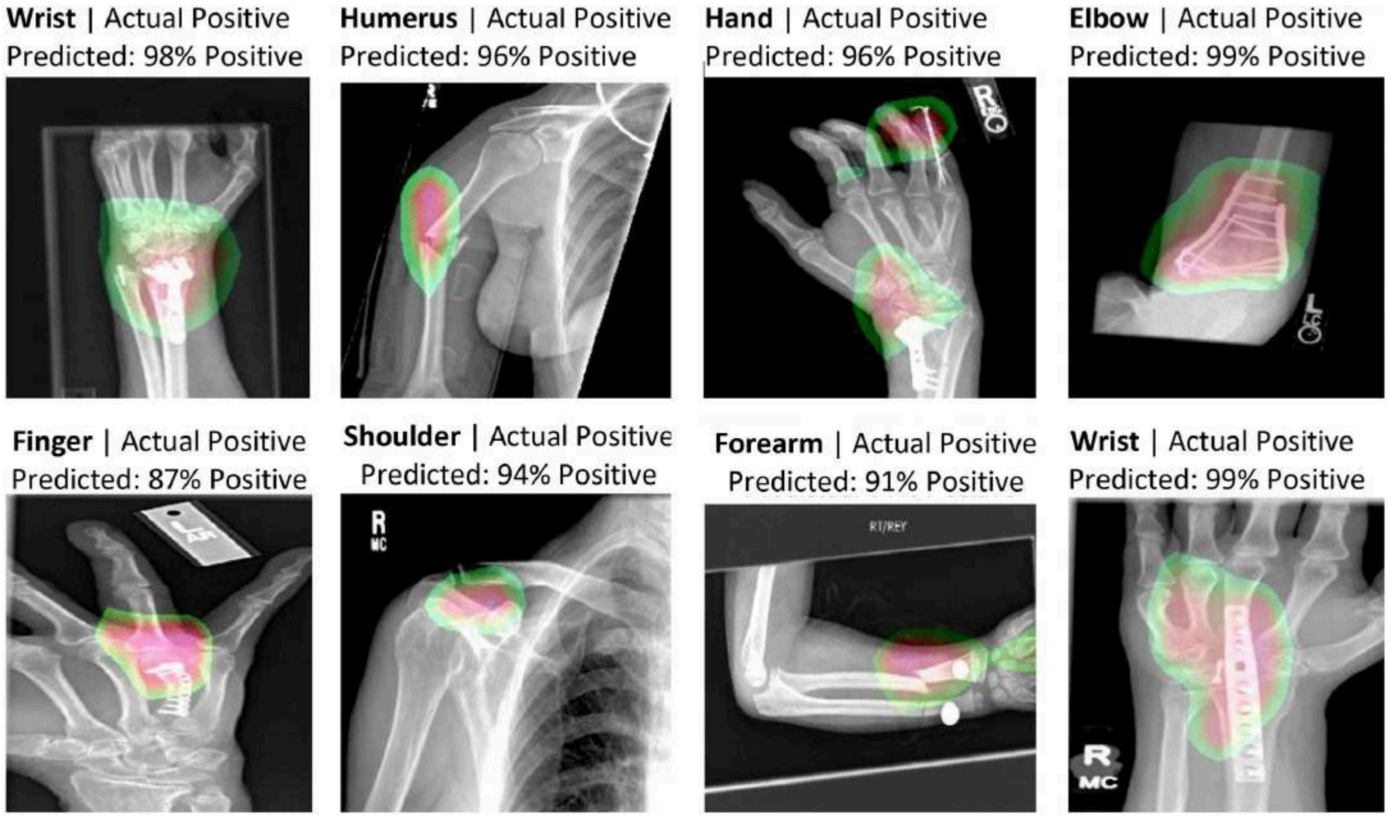
Qiangwen Xu, Menyao Sun, Bailing Fu, Yijun Zhao, “ **Deep Learning Based Parking Vacancy Detection for Smart Cities,”** Hawaii International Conference on System Sciences, 2022



Deep Learning Applications

Convolutional Neural Network

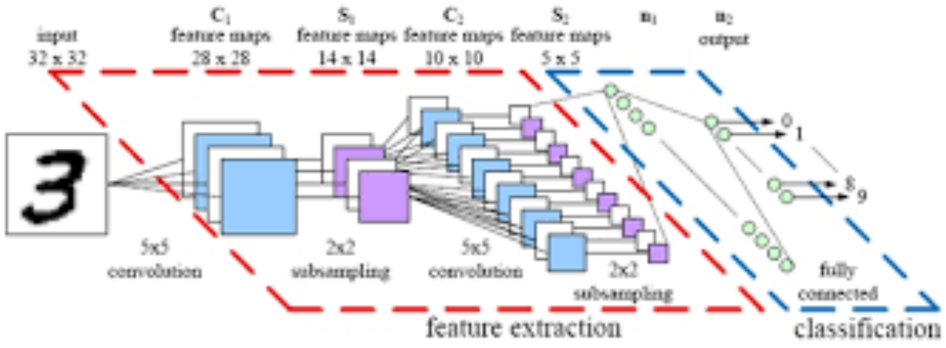
Bone X-ray Abnormality Detection (published course project):
He, Minliang, Xuming Wang, and Yijun Zhao. "A calibrated deep learning ensemble for abnormality detection in musculoskeletal radiographs." *Scientific Reports* 11, no. 1 (2021): 1-11.



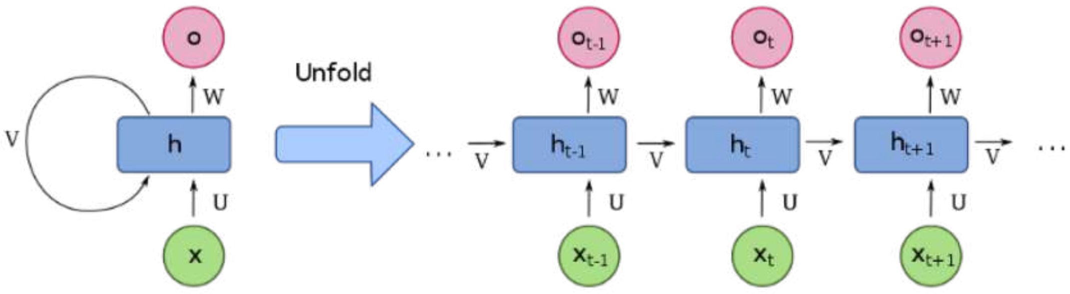
Deep Learning

Types of DL architectures covered in this course:

- Convolutional Neural Network (CNN)
- **Autoencoder (AE)**
- Recurrent Neural Network (RNN)
- Long Short-term Memory (LSTM)
- Gated Recurrent Units (GRU)
- Generative Adversarial Network (GAN)



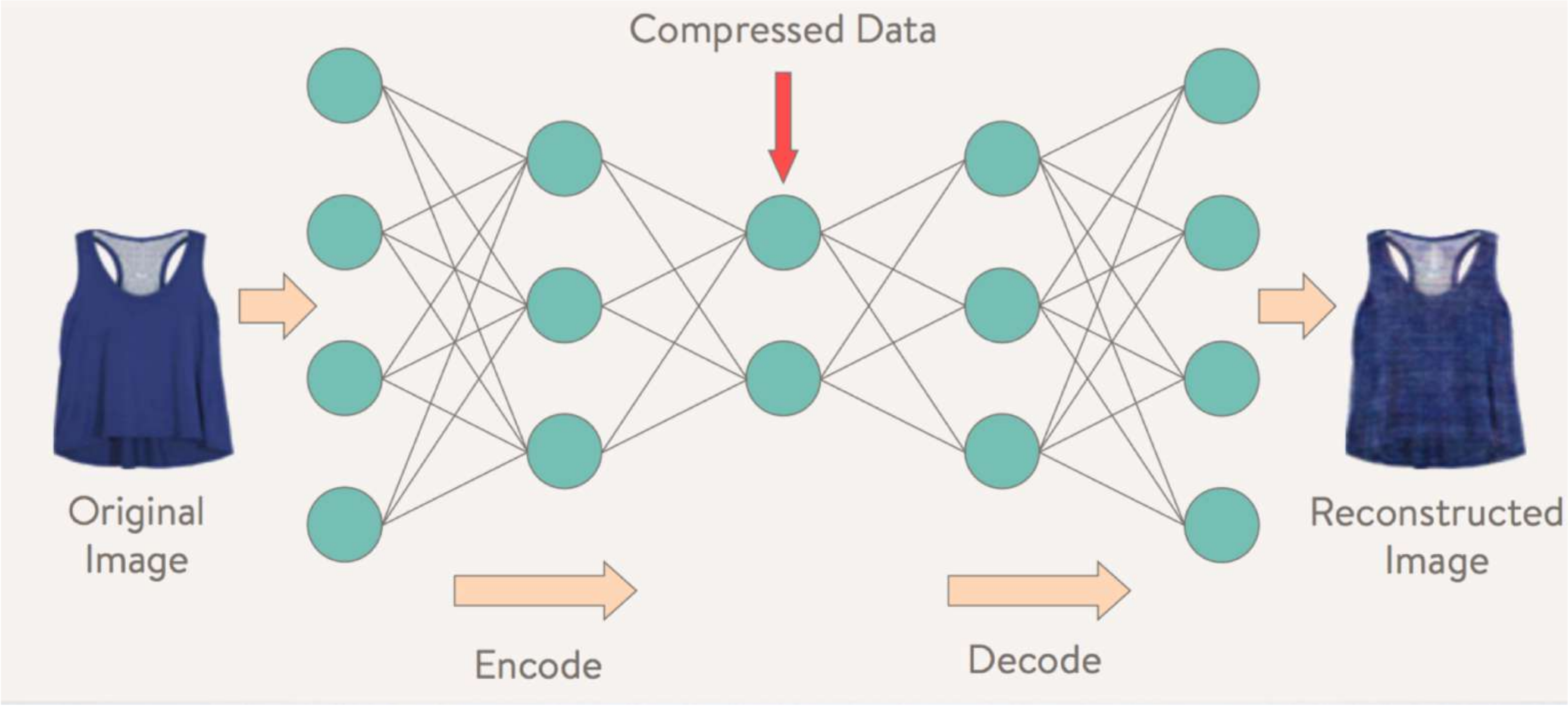
CNN



RNN

Deep Learning Applications

Autoencoder (AE)



Deep Learning Applications

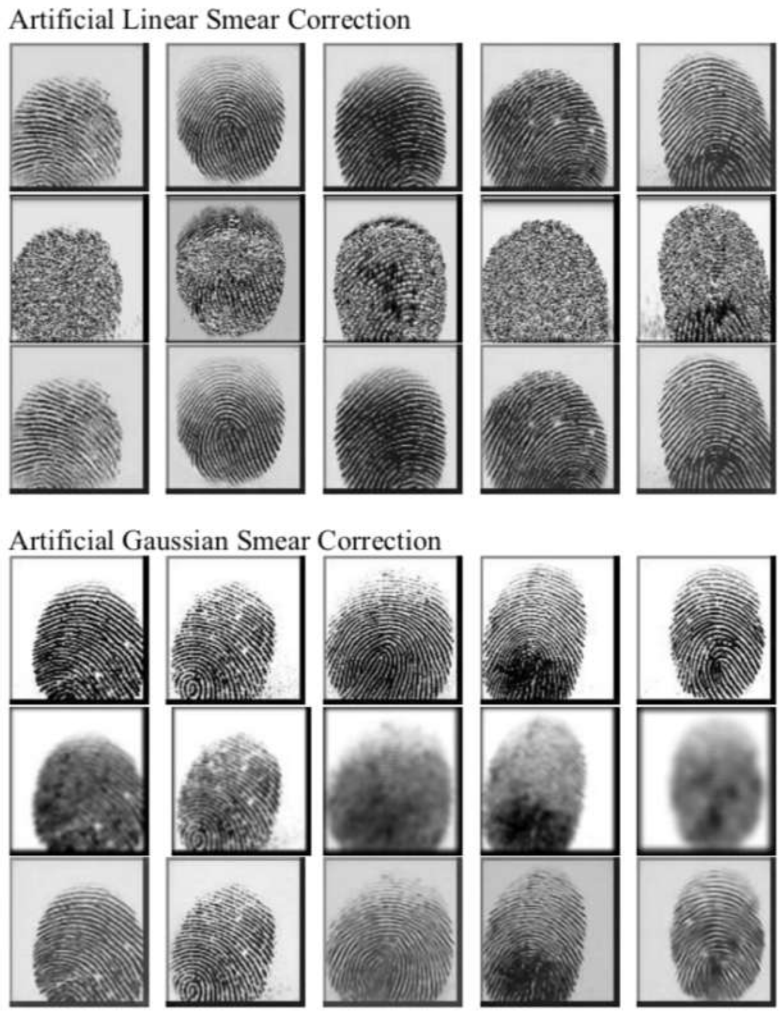
Autoencoder (AE)

•Y. Zhao, S. Lebak, "Deep Convolutional Autoencoder for Recovering Defocused License Plates and Smudged Fingerprints," *The 15th International Conference on Data Science, 2019*

Deep Learning Applications

Autoencoder (AE)

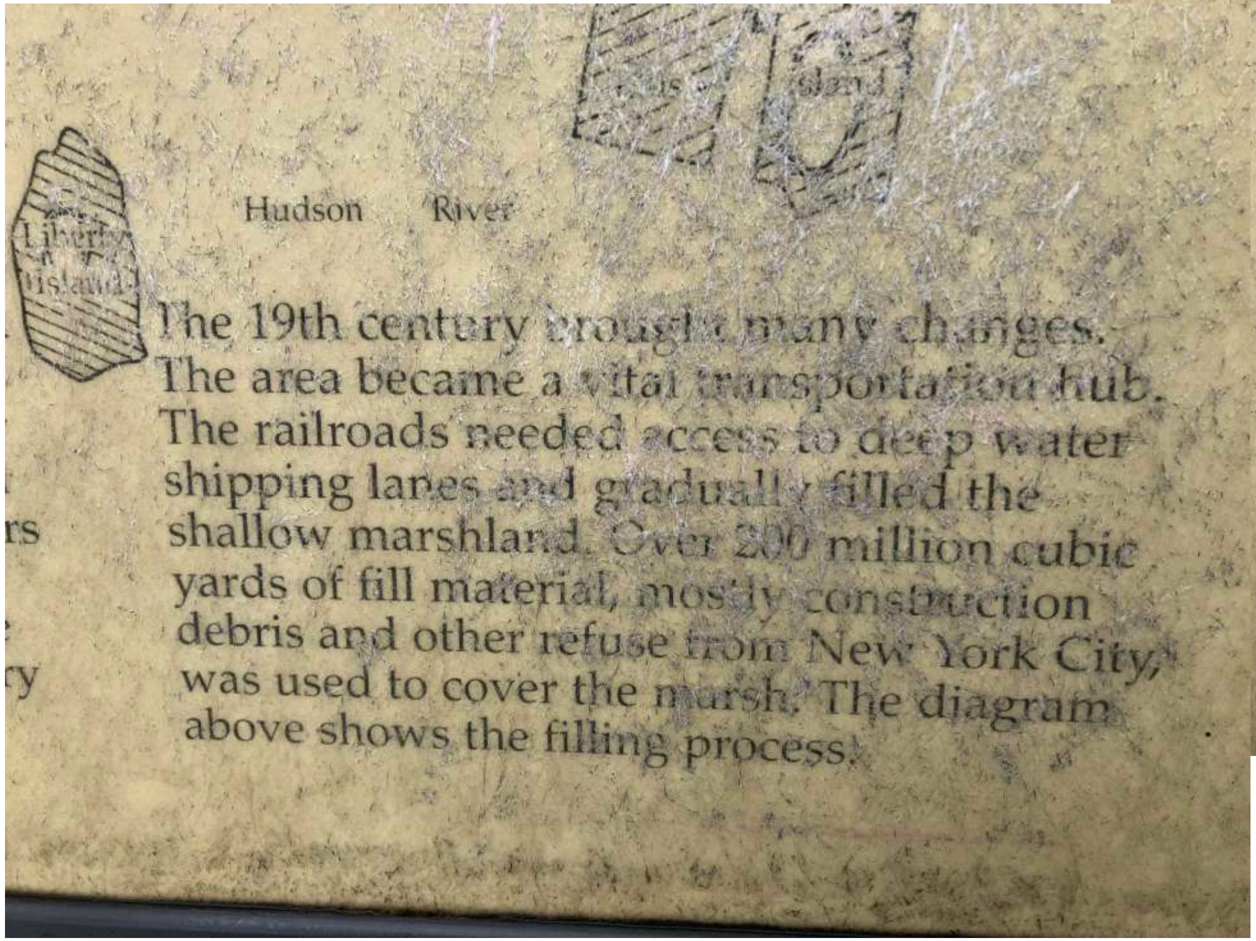
•Y. Zhao, S. Lebak, "Deep Convolutional Autoencoder for Recovering Defocused License Plates and Smudged Fingerprints," *The 15th International Conference on Data Science, 2019*

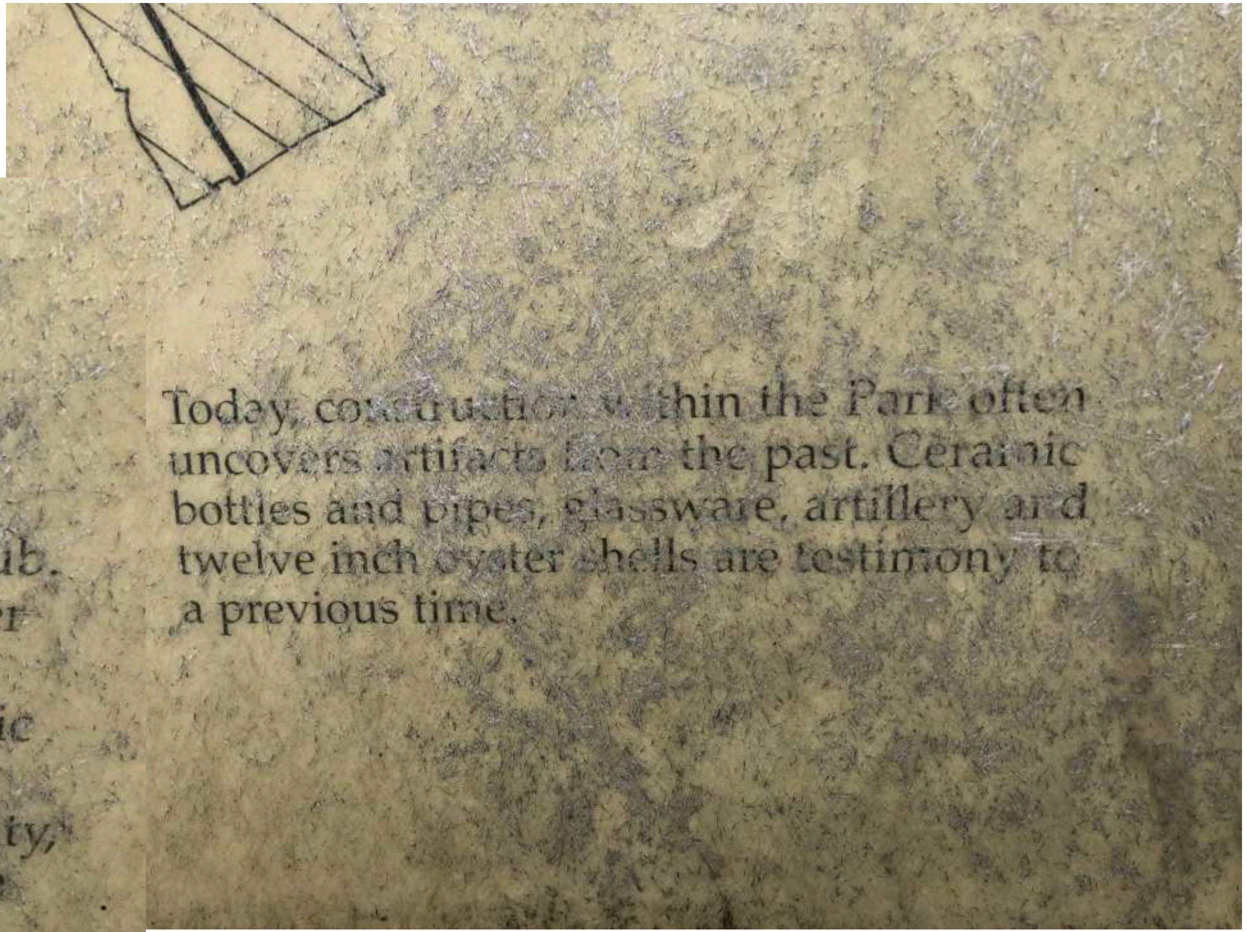


Top row: Original Images. Expected model output.
Middle row: Smudged fingerprints. Input to the model.
Bottom row: Model output.

Deep Learning Applications

Potential Course Projects





Can you recover the eroded text on the plates?

Deep Learning Applications

Potential Course Projects

Can you recover smudged footprints?

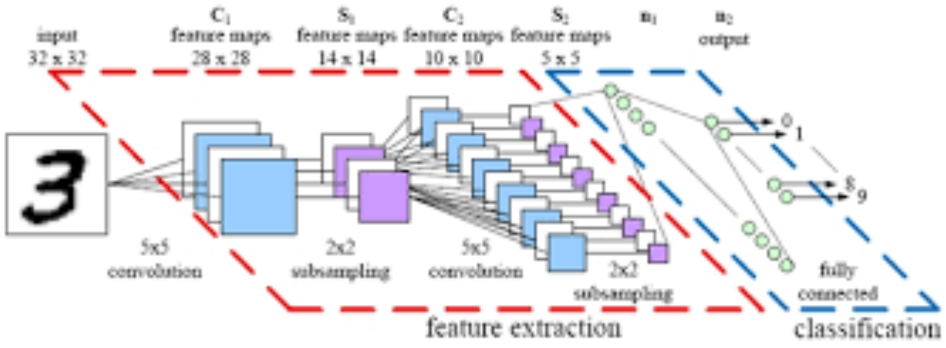
Challenge: creating the training dataset

- Search for existing public datasets
- Apply data augmentations

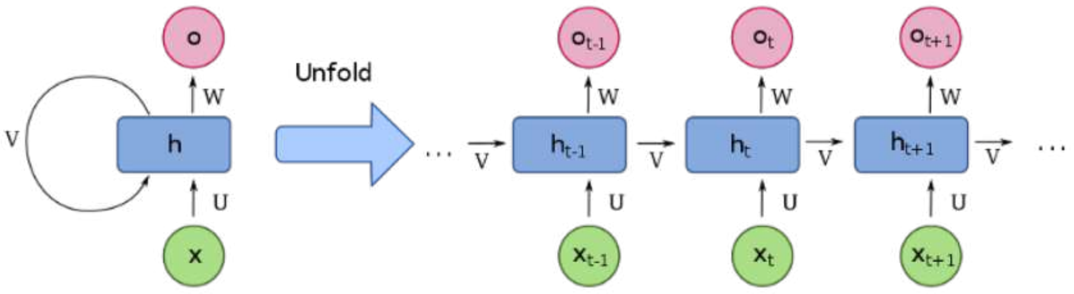
Deep Learning

Types of DL architectures covered in this course:

- Convolutional Neural Network (CNN)
- Autoencoder (AE)
- **Recurrent Neural Network (RNN)**
- **Long Short-term Memory (LSTM)**
- **Gated Recurrent Units (GRU)**
- Generative Adversarial Network (GAN)



CNN

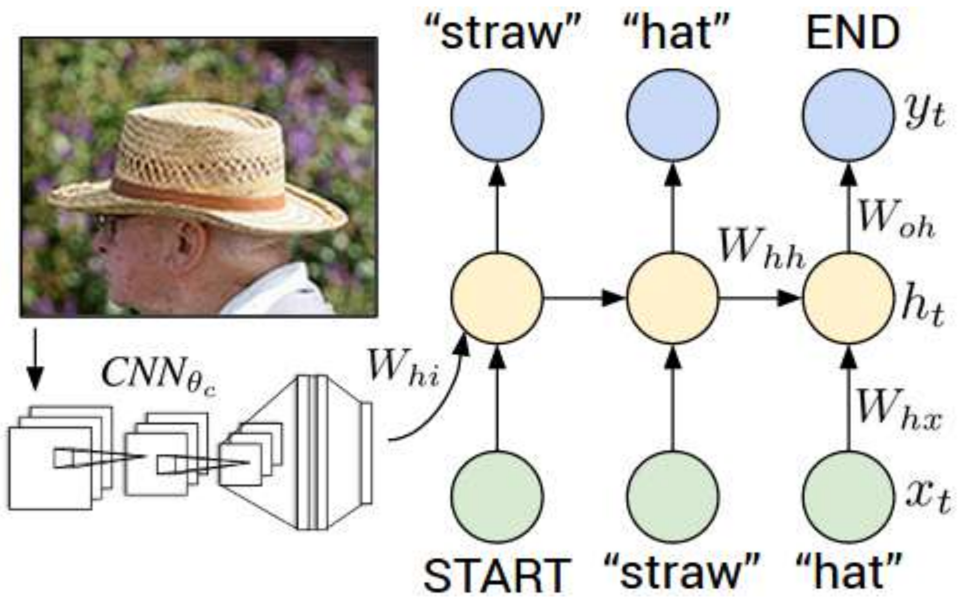


RNN

Deep Learning Applications

RNN/LSTM/GRU

Karpathy, Andrej, and Li Fei-Fei. "Deep visual-semantic alignments for generating image descriptions." *IEEE conference on computer vision and pattern recognition*, 2015 Cited by 2641



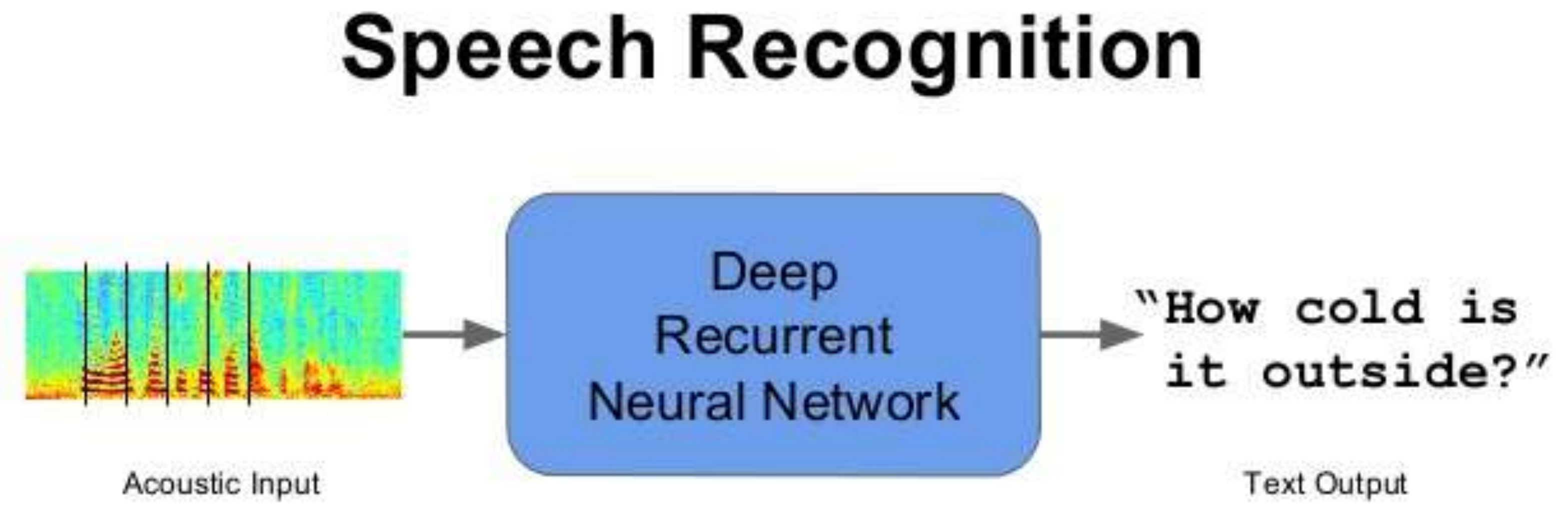
man in black shirt is playing guitar.

two young girls are playing with lego toy.

Deep Learning Applications

RNN/LSTM/GRU

Graves, Alex, Abdel-rahman Mohamed, and Geoffrey Hinton. "Speech recognition with deep recurrent neural networks." *IEEE international conference on acoustics, speech and signal processing*, 2013. Cited by 4129



Deep Learning Applications

Potential RNN/LSTM/GRU Course Projects

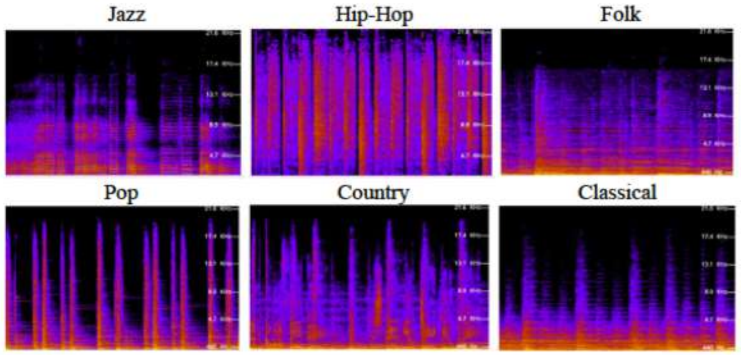
Voice identification:

- who is speaking
- is the voice authentic

Voice generation

Deep Learning Applications

Music Genre Classification (published course project): Yuan, Hui, Wenjia Zheng, Yun Song, and Yijun Zhao. "Parallel Deep Neural Networks for Musical Genre Classification: A Case Study." *IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2021.



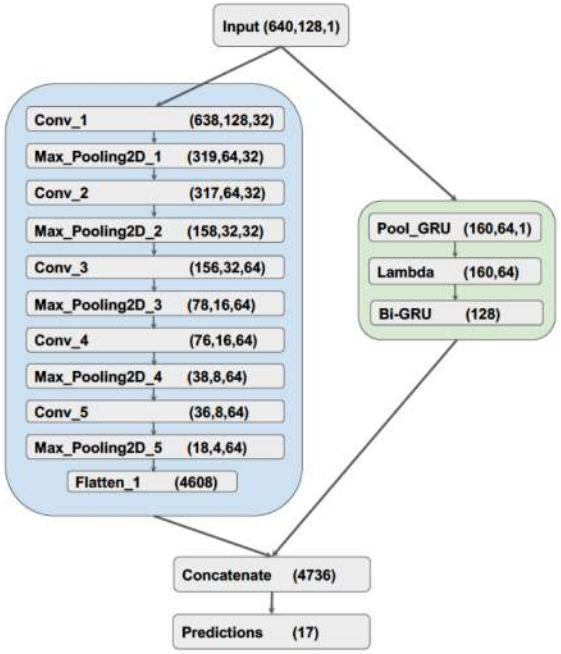
CNN

- 5 times Conv2d layers
- Filters: 32,32,64,64,64
- Kernel size:(3,1)
- Activation: ReLU
- Max Pooling: 2x2
- Flatten: (None,4608)

Bi-GRU

- Deploy a Max Pooling layer first to downsize the images
- Bidirectional GRU (64 units)
- Flatten to one-dimension vector as input (None,128)

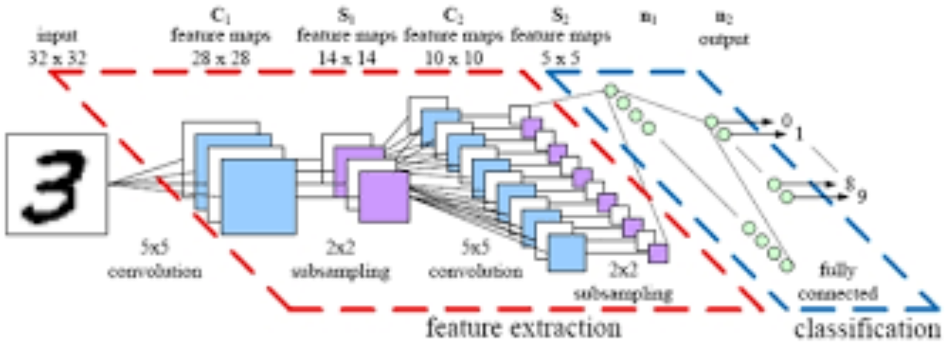
Merge layer
- Concatenate outputs from CNN and Bi-GRU
- Dense layer: 17 final classes



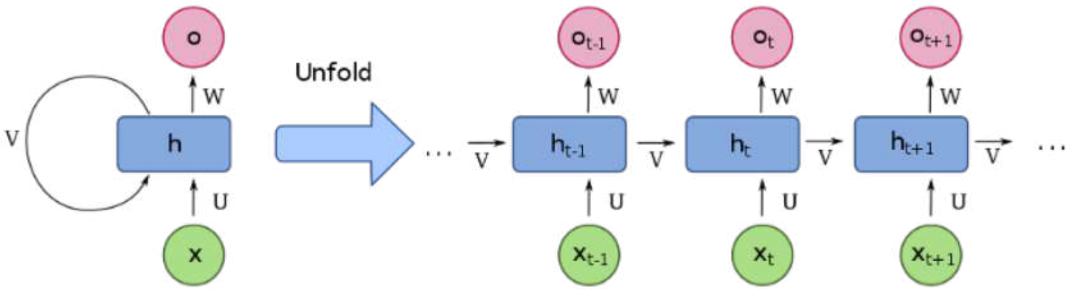
Deep Learning

Types of DL architectures covered in this course:

- Convolutional Neural Network (CNN)
- Autoencoder (AE)
- Recurrent Neural Network (RNN)
- Long Short-term Memory (LSTM)
- Gated Recurrent Units (GRU)
- **Generative Adversarial Network (GAN)**



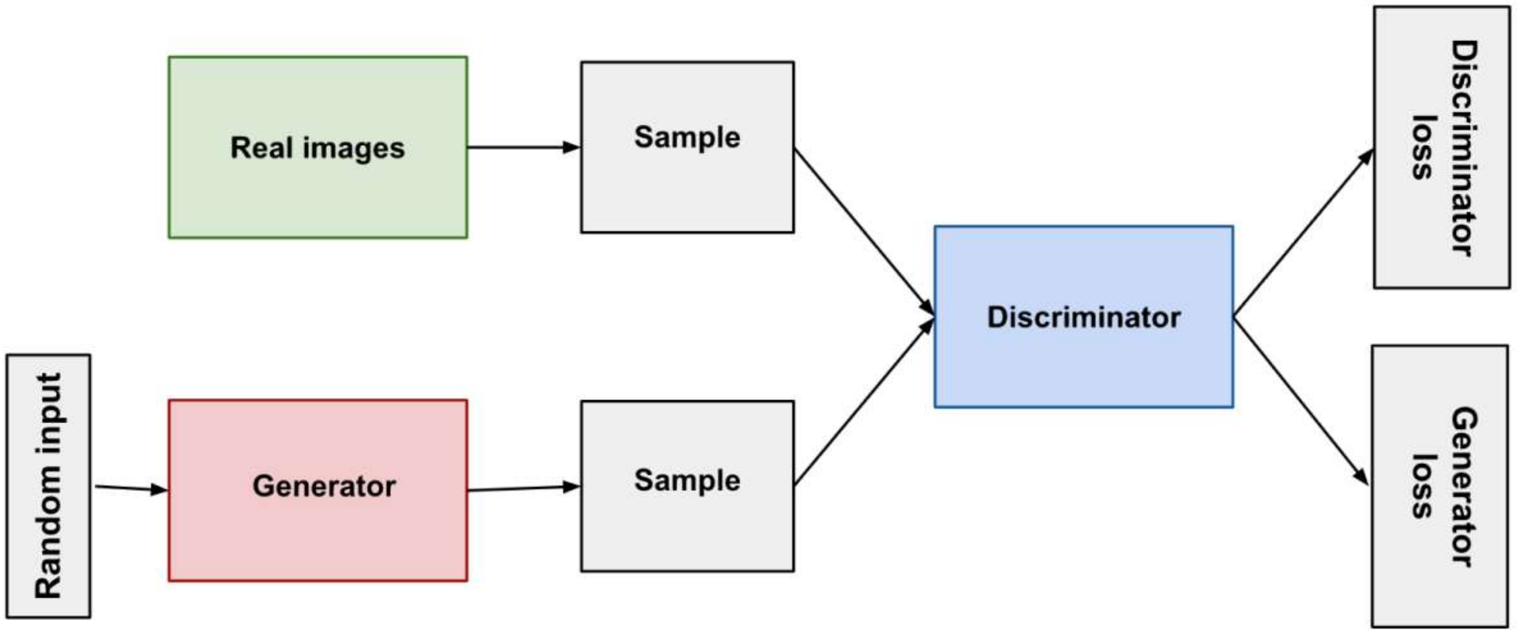
CNN



RNN

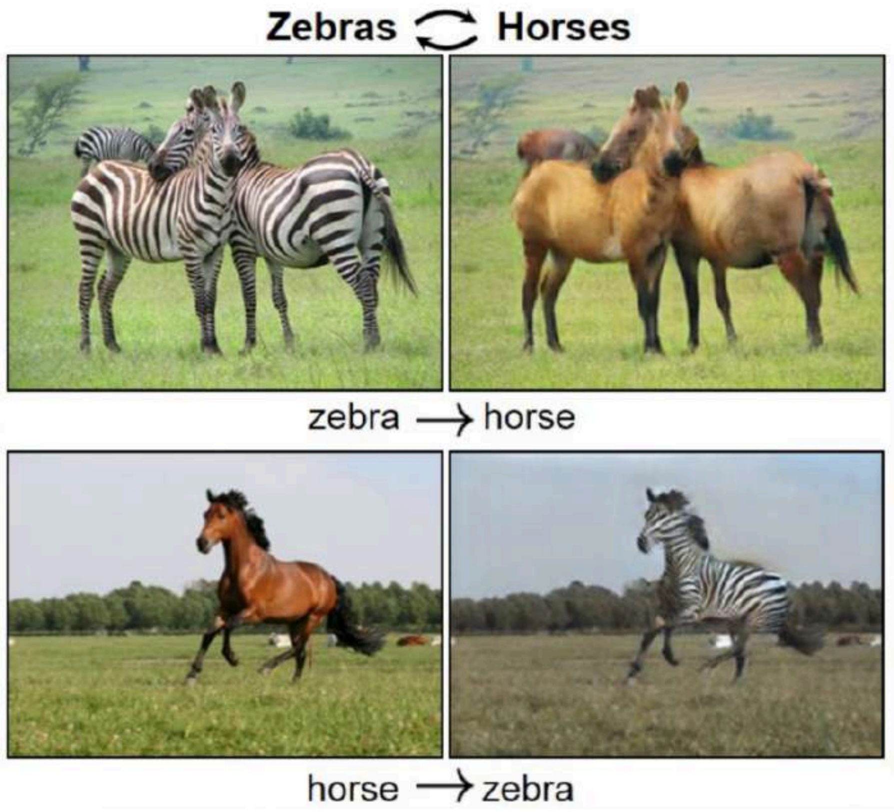
Generative Adversarial Networks (GANs)

- Introduced by Goodfellows et al., 2016



Deep Learning Applications

Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *IEEE international conference on computer vision*, 2017.



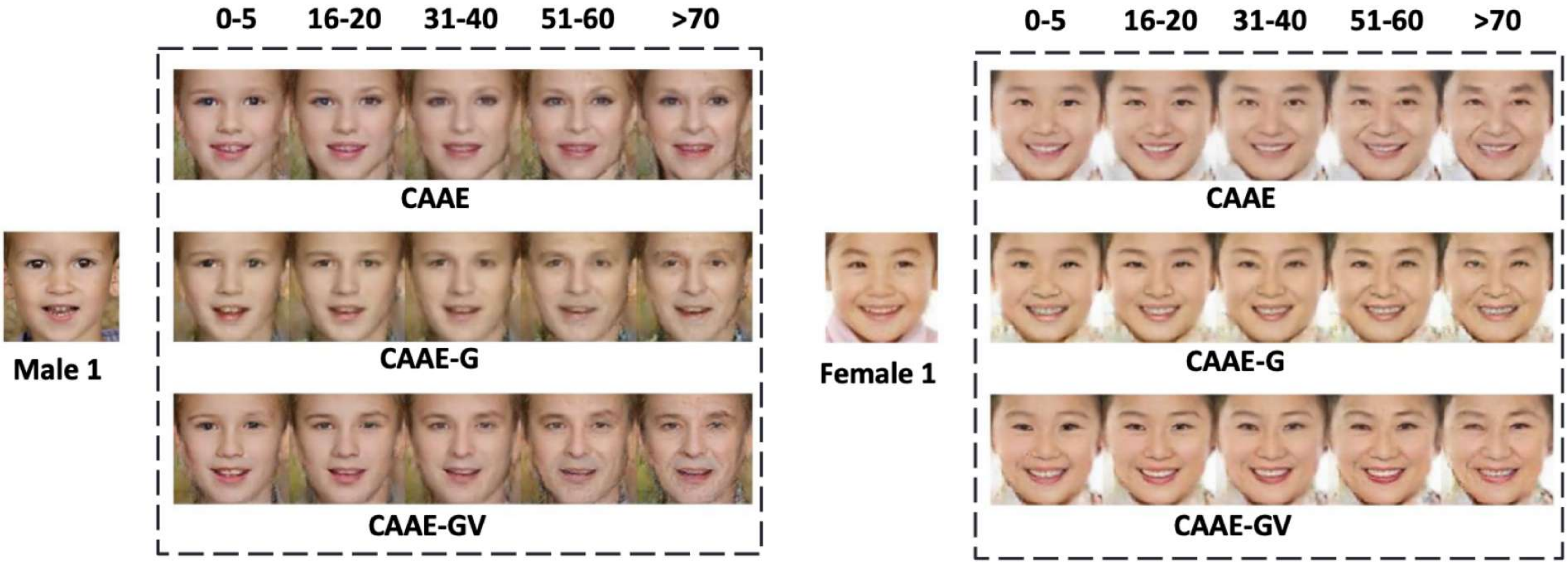
Karras, Tero, Timo Aila, Samuli Laine, and Jaakko Lehtinen. "Progressive growing of gans for improved quality, stability, and variation." *arXiv preprint arXiv:1710.10196* (2017).

Realistic yet fictional portraits of celebrities generated from originals using GANs

Deep Learning Applications

Face Age Progression (published course project):

Xiao, Yao, and Yijun Zhao. "Preserving Gender and Identity in Face Age Progression of Infants and Toddlers." In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pp. 1-8. IEEE, 2021.



Sample published course projects from previous CISC 6000 class:

- **CNN, ResNet, and DenseNet**

A calibrated deep learning ensemble for abnormality detection in musculoskeletal radiographs
Minliang He, Xuming Wang , Yijun Zhao
Scientific Reports **volume 11**, 2021

- **GANS**

Preserving Gender and Identity in Face Age Progression of Infants and Toddlers
Yao Xiao, Yijun Zhao
IEEE International Joint Conference on Biometrics (IJCB), 2021

- **CNN and LSTM**

Parallel Deep Neural Networks for Musical Genre Classification: A Case Study
Hui Yuan, Wenjia Zheng, Yun Song, Yijun Zhao
IEEE COMPSAC, 2021

- **MRCNN**

Deep Learning Based Parking Vacancy Detection for Smart Cities
Qianwen Xu, Mengyao Sun, Bailing Fu, Yijun Zhao
Hawaii International Conference on System Sciences, 2022

What’s the key to get your course project published?
Novelty, novelty, novelty.

1. **A new idea to solve a practical problem**

Deep Learning Based Parking Vacancy Detection for Smart Cities

2. **Improve on an existing architecture**

Preserving Gender and Identity in Face Age Progression of Infants and Toddlers

3. **Combine existing solutions to get better results.**

A calibrated deep learning ensemble for abnormality detection in musculoskeletal radiographs

4. **Customize existing approach as a case study.**

Parallel Deep Neural Networks for Musical Genre Classification: A Case Study

Course Homepage:

https://storm.cis.fordham.edu/~yzhao/fall2022/CS6000_syllabus.htm

OR

Syllabus on Blackboard