

Seeing is Believing: Exploring Convolutional Neural Networks

Yinan Guo

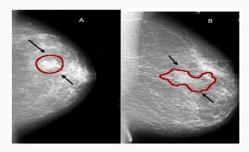
Statistics & Data Science





What is CNN? How DL transformed CV?

01 Object Detection



[1]

[3]

03 Image Reorganization



02 Image Segmentation



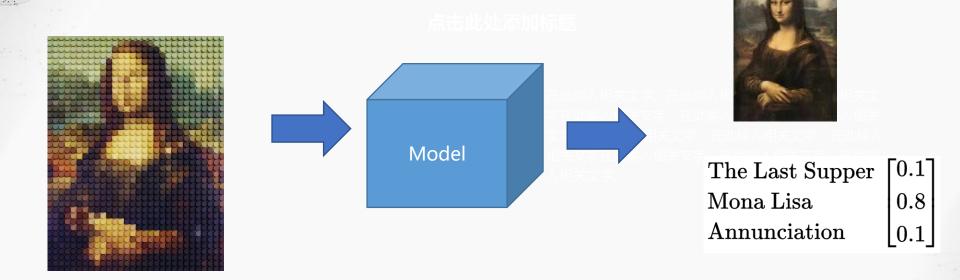
[2]

Generative models



[4]

The basics



We usually accept input images as pixel-wise matrix. i.e. a 1080 x 1080 x 3 RGB image.



Two Concepts

Feature Extraction and Convolution



Features: Eyes, Hand, Position of body, Color, etc.

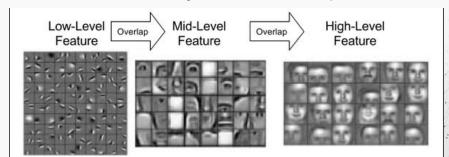
Past: Manual Feature Extraction

Domain Knowledge Define Features

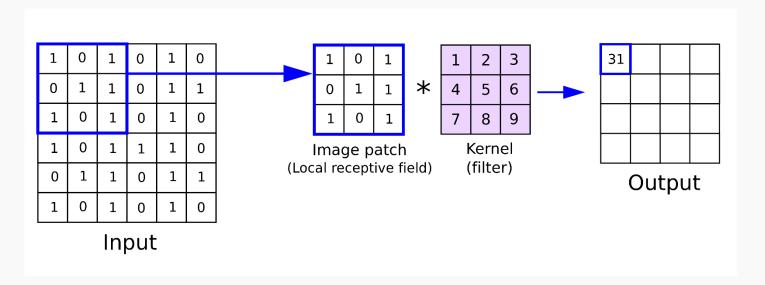
Label Features and Classify



Feature Extraction by Convolution operation

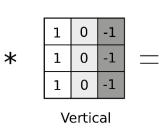


The Convolution Operation

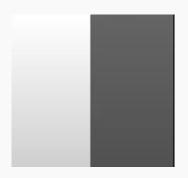


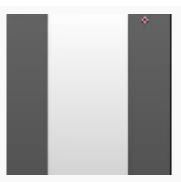
The Convolution Operation

| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
|----|----|----|----|---|---|---|---|
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |
| 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 |

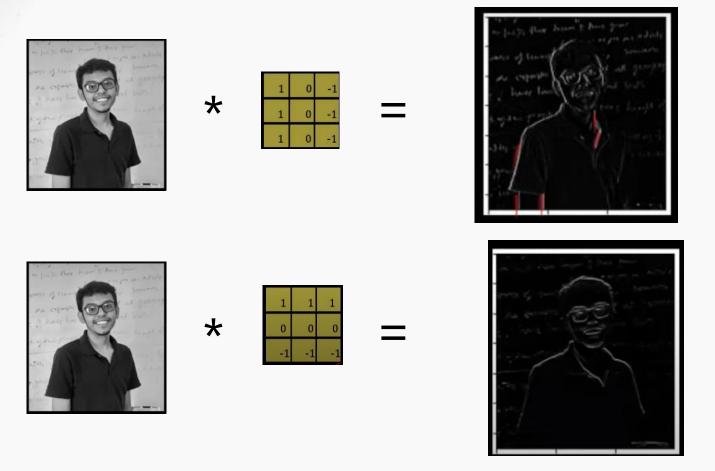


| 0 | 0 | 30 | 30 | 0 | 0 |
|---|---|----|----|---|---|
| 0 | 0 | 30 | 30 | 0 | 0 |
| 0 | 0 | 30 | 30 | 0 | 0 |
| 0 | 0 | 30 | 30 | 0 | 0 |
| 0 | 0 | 30 | 30 | 0 | 0 |
| 0 | 0 | 30 | 30 | 0 | 0 |

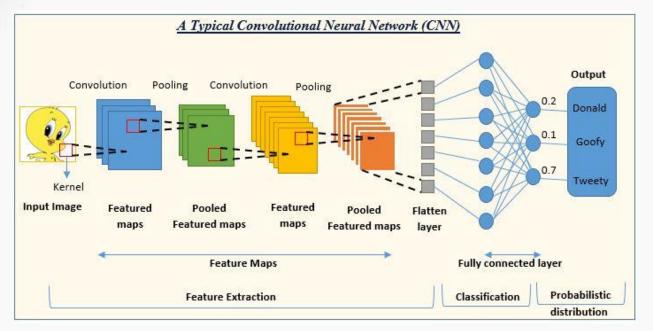


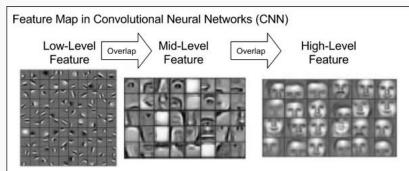


[6]



Example Credit: https://www.youtube.com/watch?v=gLwX3zHkims&t=333s



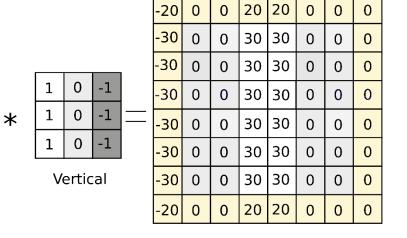


[7]

Padding

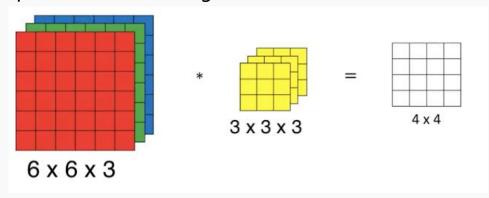
Edge of the original input is not used as often as the central ones.

| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|---|----|----|----|----|---|---|---|---|---|
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 10 | 10 | 10 | 10 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |



Convolution Layer (Padding, Stride)

Example of 1 filter, Padding =0, Stride =1



| 0 0 |
|--------------|
| 0 0 |
| 0 0 1 0 -1 |
| 0 0 * 1 0 -1 |
| 0 0 1 0 -1 |
| 0 0 |
| 0 0 Vertical |
| 0 0 |
| |

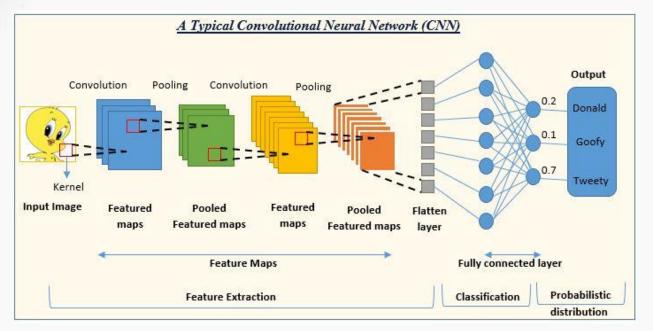
[6]

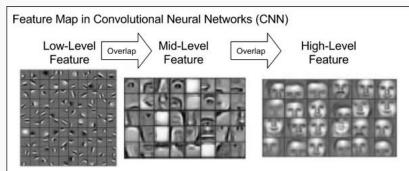
Pooling Layer (Filter size, Stride)

| 1 | 3 | 1 | 2 | | | |
|---|---|---|---|----------------|---|---|
| 2 | 9 | 1 | 1 | Max Pooling | 9 | 2 |
| 1 | 5 | 2 | 1 | f = 2 s = 2 | 6 | 3 |
| 3 | 6 | 3 | 2 | | | |

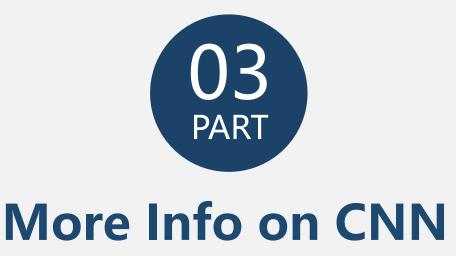
Max pooling, Average pooling, etc.

[6]





[7]

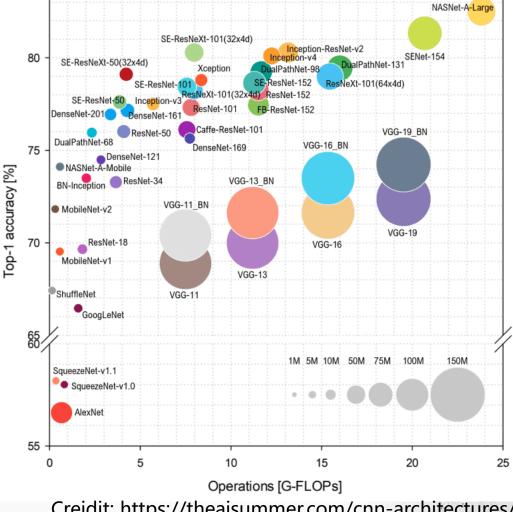


State of Art Models

Research Direction: Classification Accuracy

$$(2015+)$$
 Layers >= 20

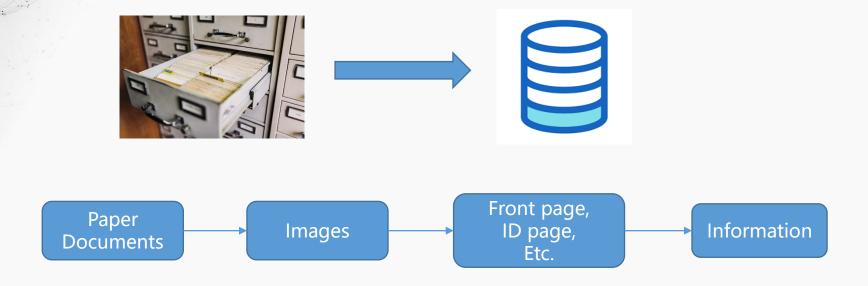
Dynamical Isometry and a Mean Field Theory of CNNs: How to Train 10,000-**Layer Vanilla Convolutional Neural Networks**



Creidit: https://theaisummer.com/cnn-architectures/



Application 1: Document Organization



Problem: Time, Money, Accuracy

Application 2: Al cloud scale

Problem: Detect unlabeled vegetables

Cloud based Al Chip based Al





Large Scale Network Applications

Research Direction: Generative models

Nvidia DLSS

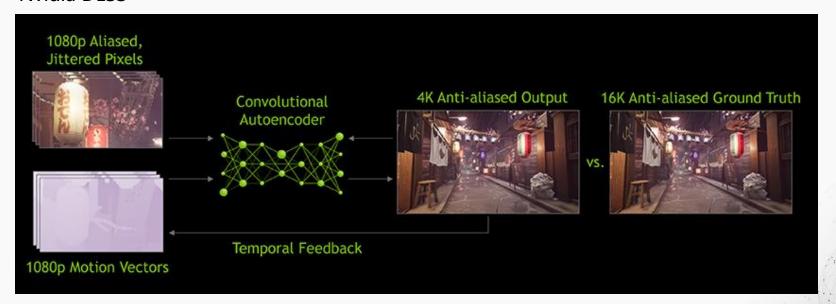




Image References:

- [1] Ragab DA, Sharkas M, Marshall S, Ren J. 2019. Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* 7:e6201 https://doi.org/10.7717/peerj.6201
- [2] Vyas, K. (2020, April 26). *Object segmentation*. Medium. Retrieved April 25, 2023, from https://medium.com/visionwizard/object-segmentation-4fc67077a678
- [3] Gayde, W. (2017, April 19). *Passengers leaving the US will have to pass facial recognition scanners at all international airports in the future*. TechSpot. Retrieved April 25, 2023, from https://www.techspot.com/news/69005-passengers-leaving-us-have-pass-facial-recognition-scanners.html
- [4] Steve, P. author B. (2016, March 24). *Face2Face: Real-time face capture Steve Diggins*. Steve Diggins. Retrieved April 25, 2023, from http://stevediggins.com/2016/03/23/face2face-real-time-face-capture/

[5] Mohammad B., Robert B., *A Critical Study on the Recent Deep Learning Based Semi-Supervised Video Anomaly Detection Methods*, https://arxiv.org/pdf/2111.01604.pdf

[6] Anh H. Reynolds. (2017, October 15). *Convolutional Neural Networks (cnns)*. Anh H. Reynolds. Retrieved April 25, 2023, from https://anhreynolds.com/blogs/cnn.html

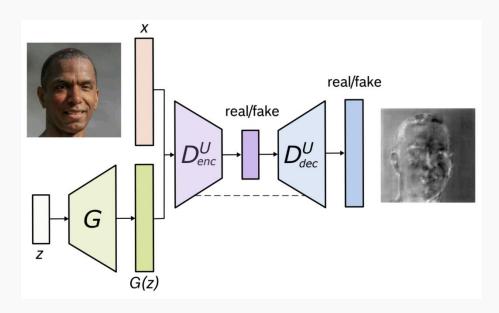
[7] Shah, S. (2022, March 15). *Convolutional Neural Network: An overview*. Analytics Vidhya. Retrieved April 25, 2023, from https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-overview/



Large Scale Network Applications

Research Direction: Generative models





My Past Projects: Al cloud scale

Research Direction: Mobile CNN

