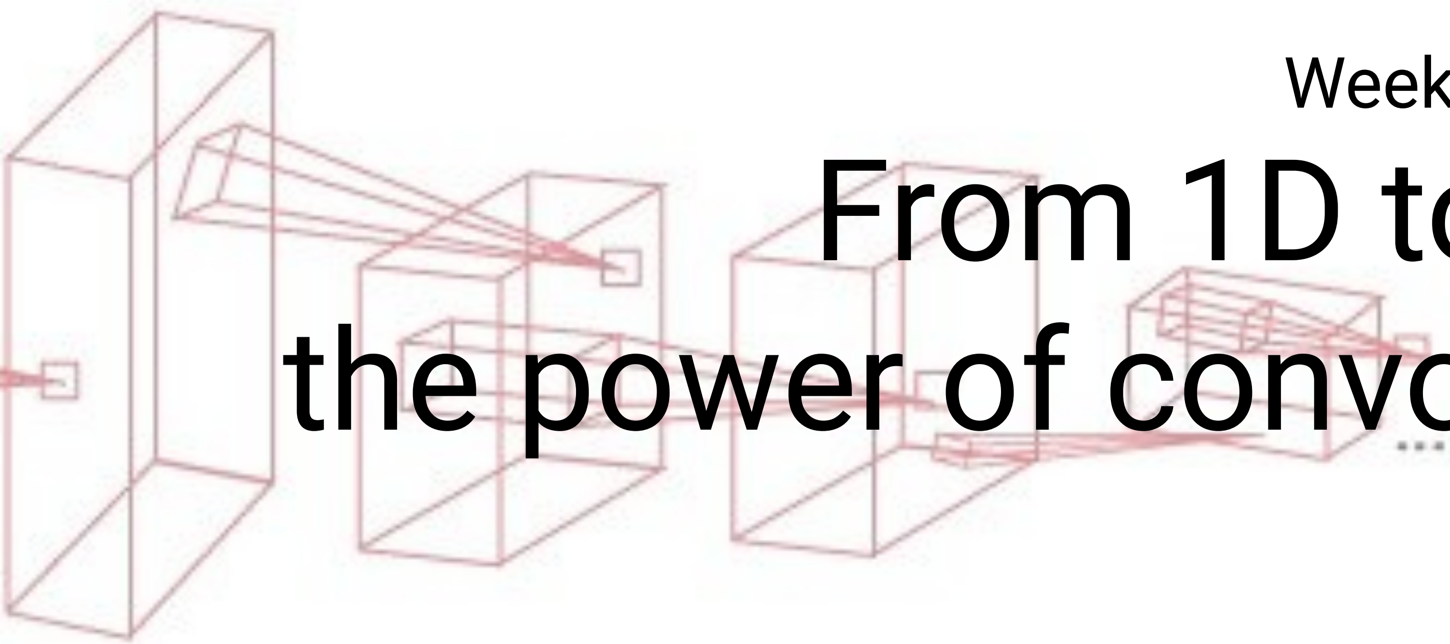
Class core values

1. Be **respect**ful to yourself and others
2. Be **confident** and believe in yourself
3. Always do your **best**
4. Be **cooperative**
5. Be **creative**
6. Have **fun**
7. Be **patient** with yourself while you learn
8. Don’t be shy to **ask “stupid” questions**
9. Be **inclusive** and **accepting**

Week 4, Lecture 1

From 1D to 2D & the power of convolution

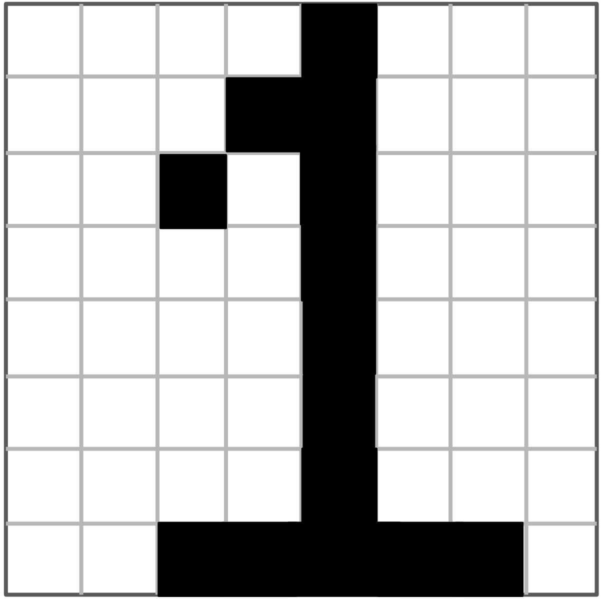


Learning Objectives

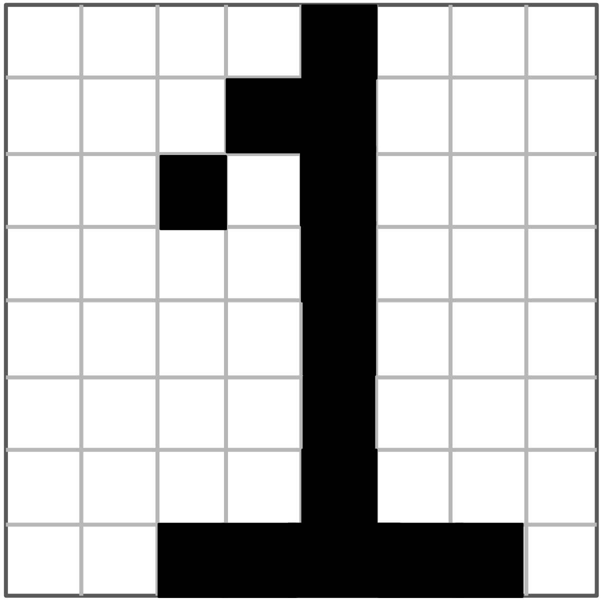
1. Describe the specific requirements of learning on image data
2. Explain the basic concept of convolution
3. Describe the specific terminology related to convolutional neural networks
4. Understand the applications of CNNs in image analysis

P1.

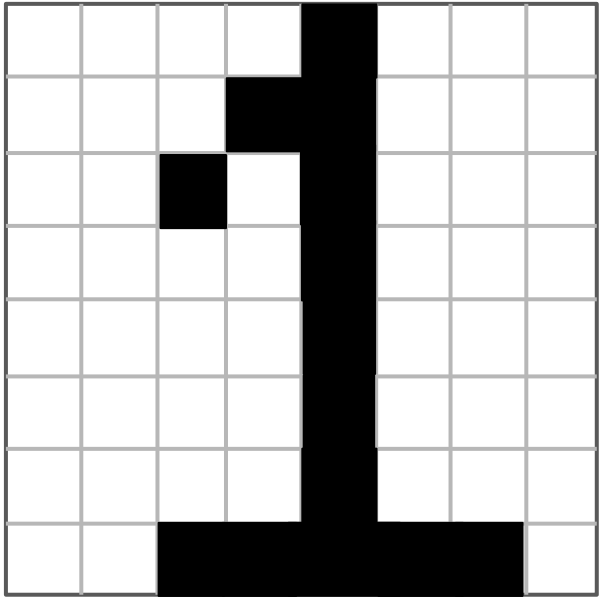
Can we use a dense neural net for image data?



Images as input data



Images as input data are matrices of numbers

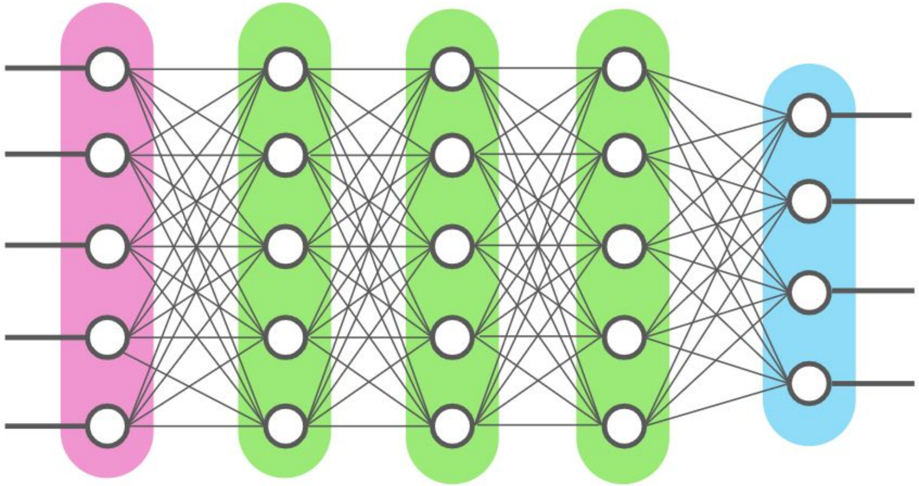


=

0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	1	1	1	1	1	0

Inputting image data to neural nets ?

0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	1	1	1	1	1	0



Inputting image data to neural nets ?

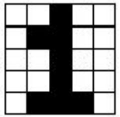
0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	1	1	1	1	1	0

$$\begin{matrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{matrix}$$

Inputting image data to neural nets ?

0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	1	0	0	0
0	0	1	1	1	1	1	0

0
0
0
1
0
0
0
0
0
0
1
1
0
0
0



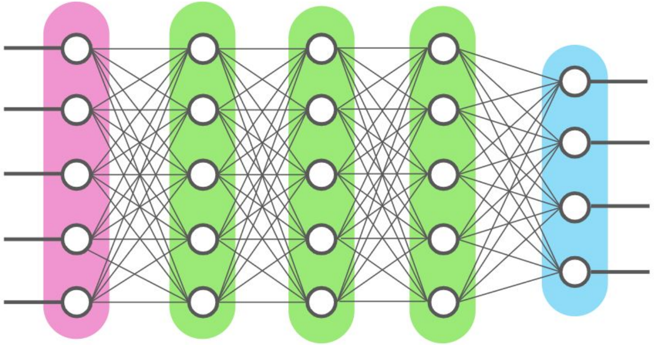
In-class activity

How many connections?

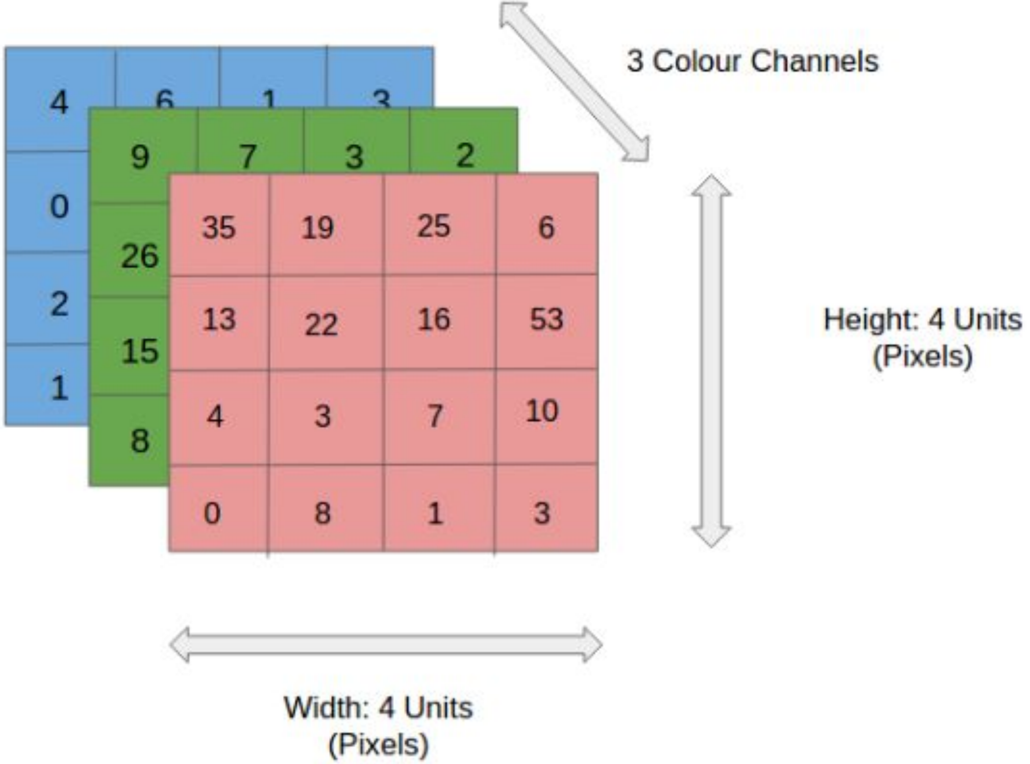
100

100

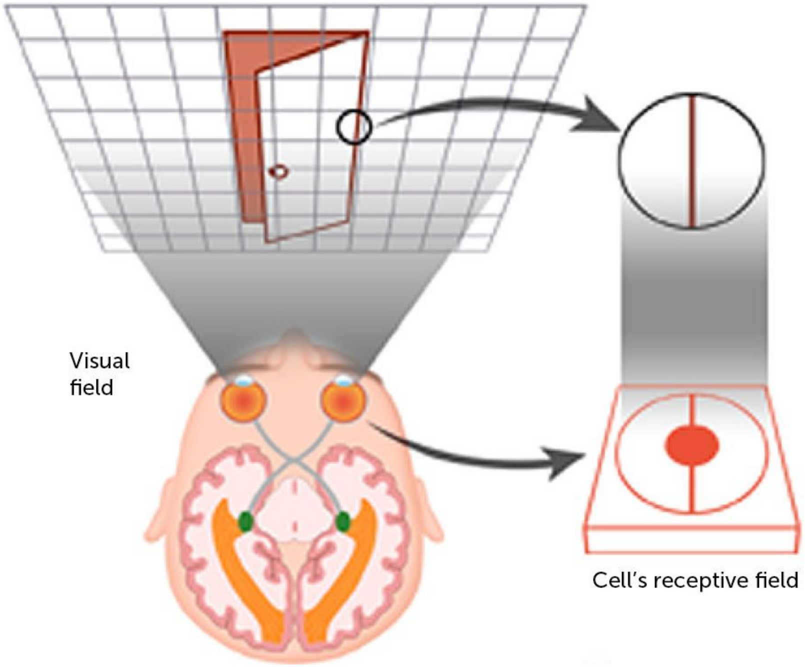
0
1
0
0.5
.
.
.
1



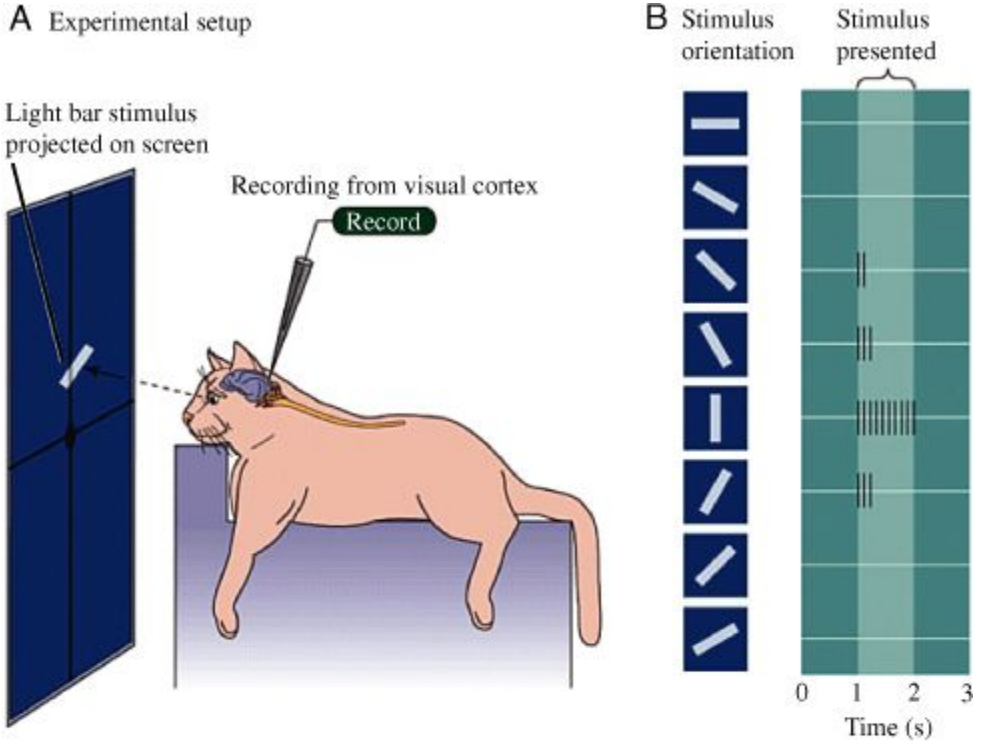
Real images often have 3 channels with many pixels



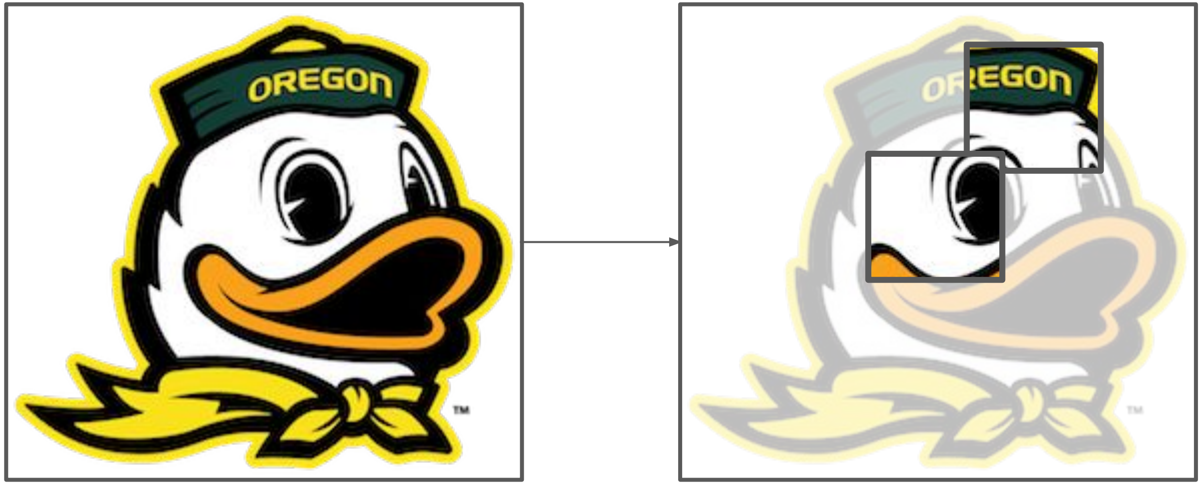
P2.
Learning from our visual cortex



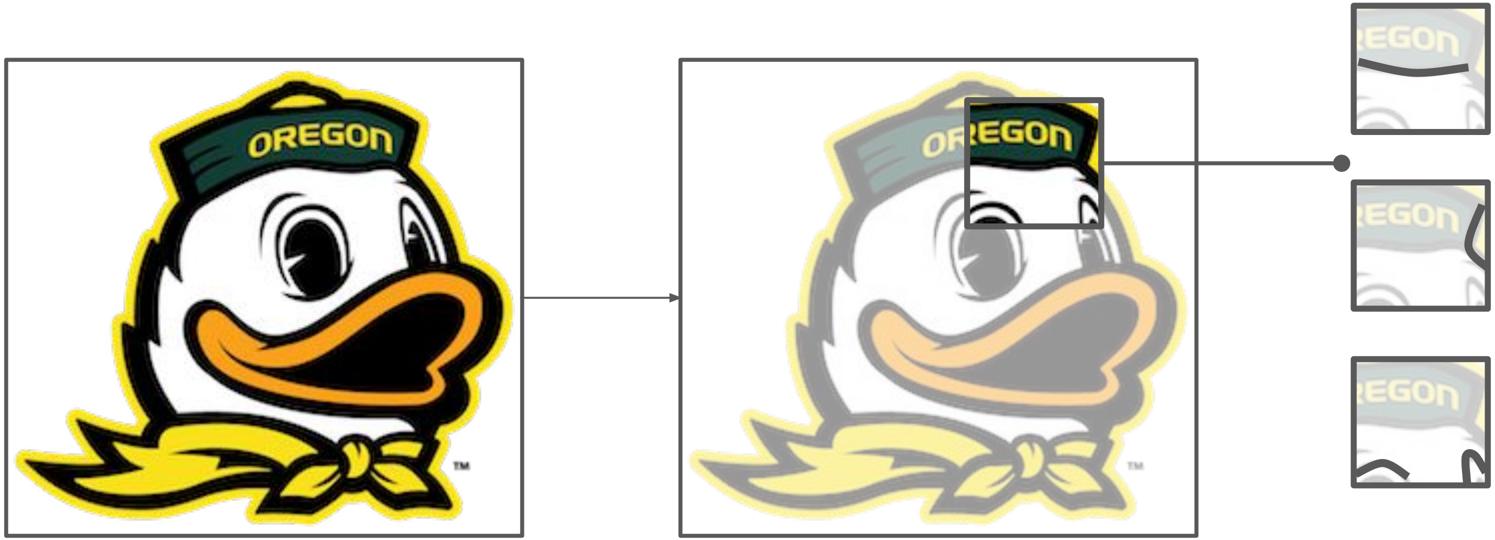
Hubel and Wiesel’s experiment on vision and the idea of *receptive field*



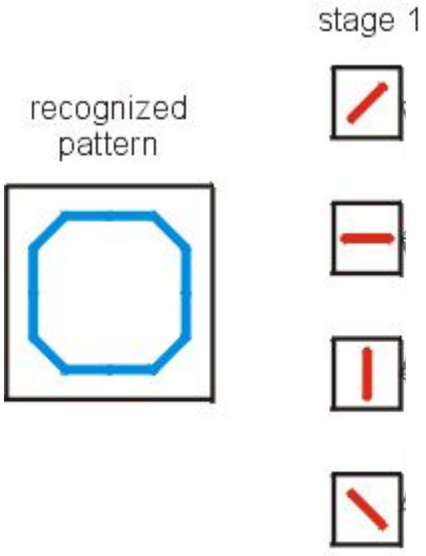
Hubel and Wiesel's experiment on vision and the idea of *receptive field*



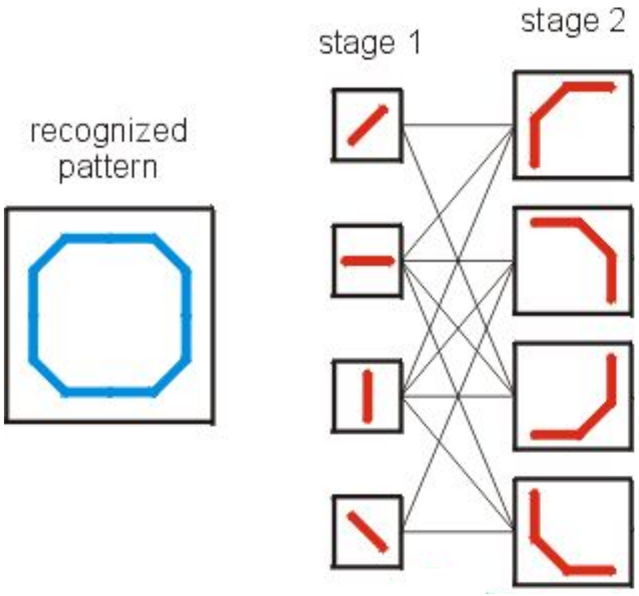
Hubel and Wiesel's experiment on vision and the idea of *receptive field*



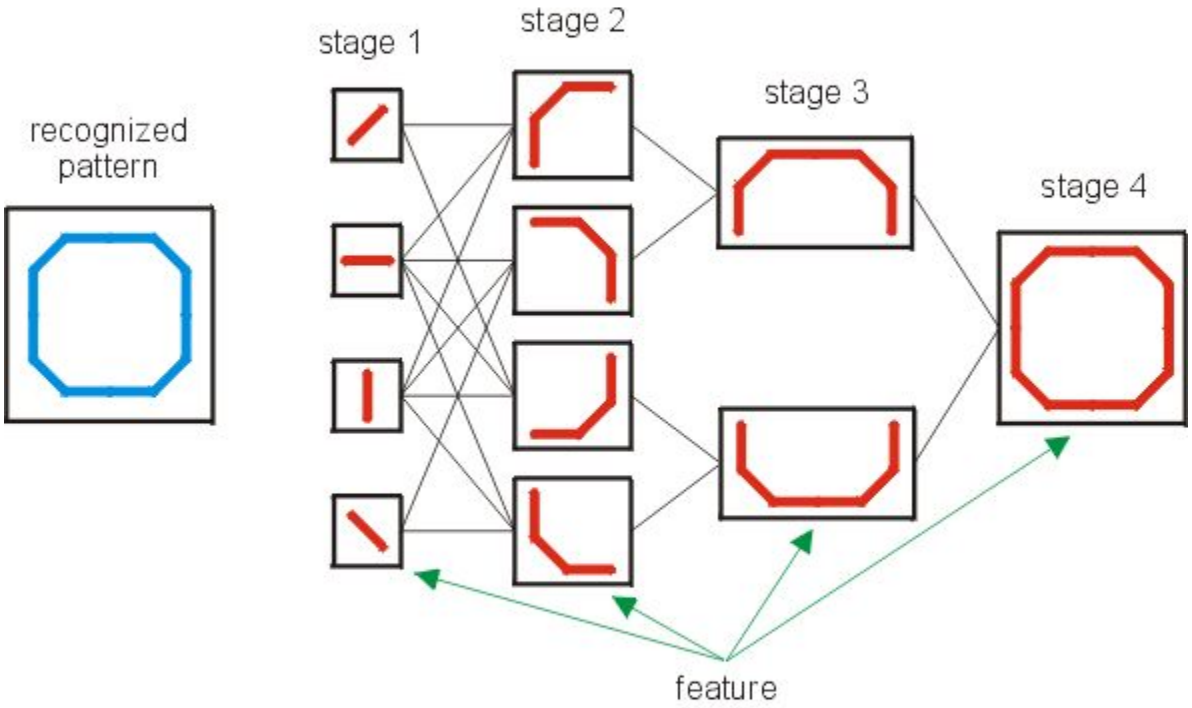
Introduction of *neocognitron*: Features and feature extraction



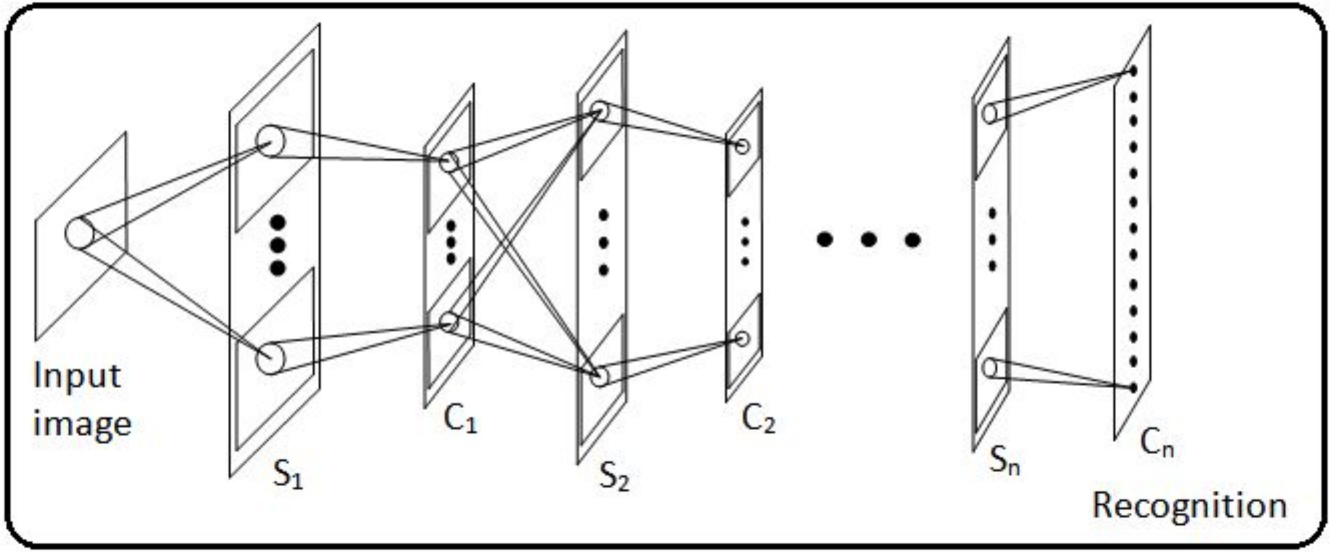
Introduction of *neocognitron*: Features and feature extraction



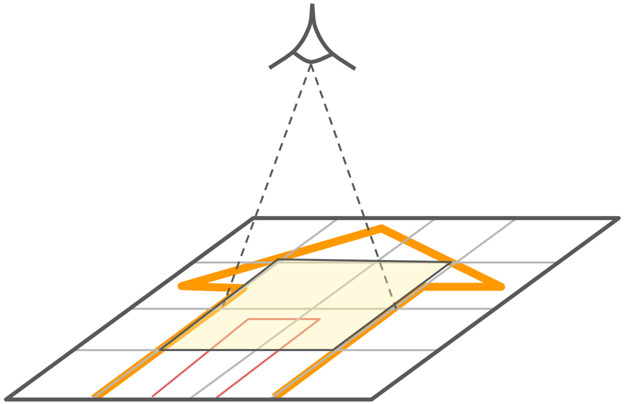
Introduction of *neocognitron*: Features and feature extraction



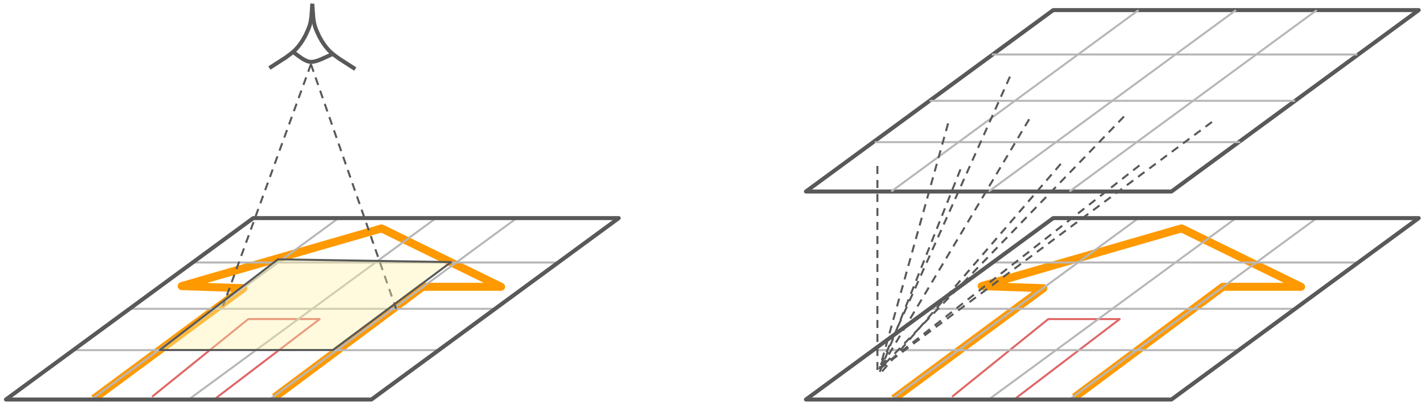
Introduction of *neocognitron*



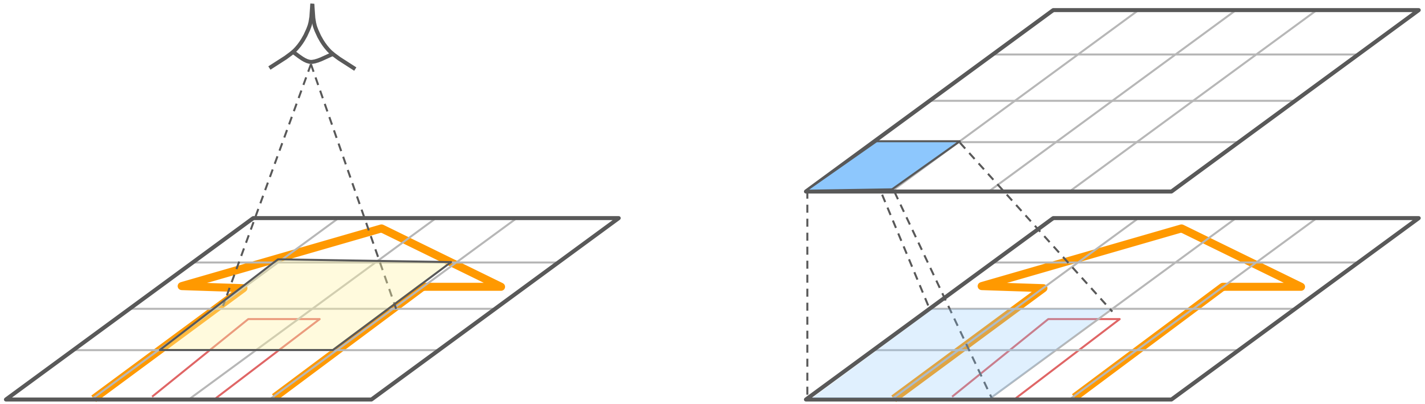
Convolution, a breakthrough in image learning



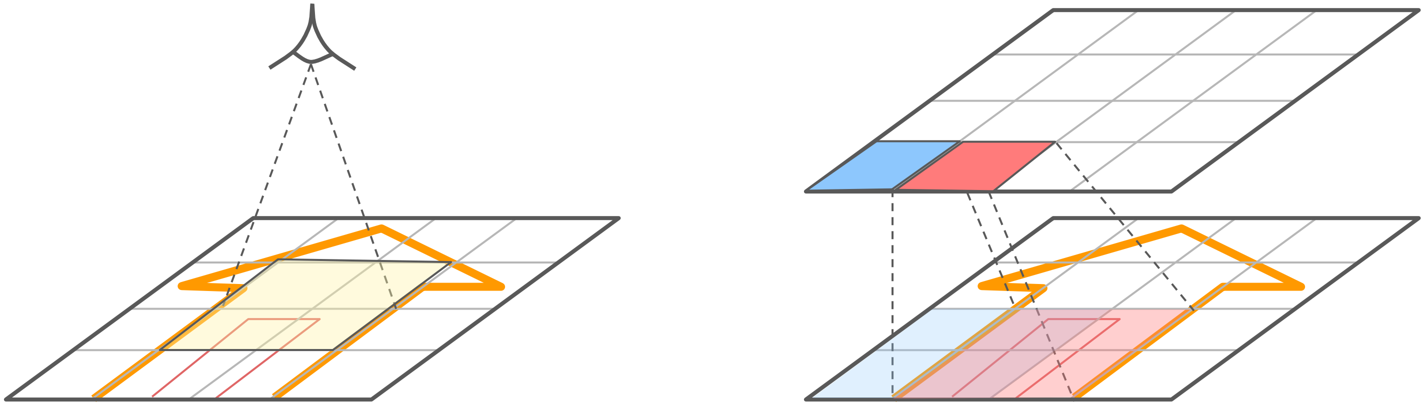
Convolution, a breakthrough in image learning



Convolution, a breakthrough in image learning



Convolution, a breakthrough in image learning



Convolution is a mathematical calculation over the matrix that is similar to feature extraction

0	2	3	1	1	0
3	7	8	4	5	1
5	8	7	5	2	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

Convolution is a mathematical calculation over the matrix that is similar to feature extraction

0	2	3	1	1	0
3	7	8	4	5	1
5	8	7	5	2	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

0	1	0
0	1	0
0	1	0

Convolution is a mathematical calculation over the matrix that is similar to feature extraction

0x0	2x1	3x0	1	1	0
3x0	7x1	8x0	4	5	1
5x0	8x1	7x0	5	2	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

17

0	1	0
0	1	0
0	1	0

The amount of movement of the convolutional filter over the image is called *stride*

0	2x0	3x1	1x0	1	0
3	7x0	8x1	4x0	5	1
5	8x0	7x1	5x0	2	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

17	18

Stride = 1

The amount of movement of the convolutional filter over the image is called *stride*

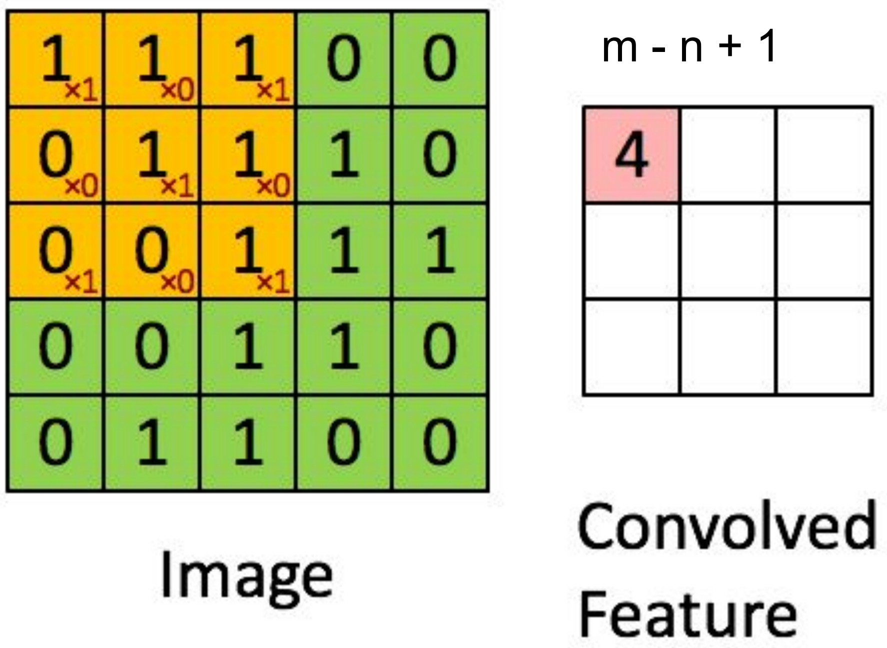
0	2x0	3x1	1x0	1	0
3	7x0	8x1	4x0	5	1
5	8x0	7x1	5x0	2	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

Stride = 1

Stride = 2

0	2	3x0	1x1	1x0	0
3	7	8x0	4x1	5x0	1
5	8	7x0	5x1	2x0	1
5	5	8	9	8	3
3	5	9	8	5	1
2	2	5	4	3	0

Convolution, a breakthrough in image learning



Convolution in reality is a 3D operation

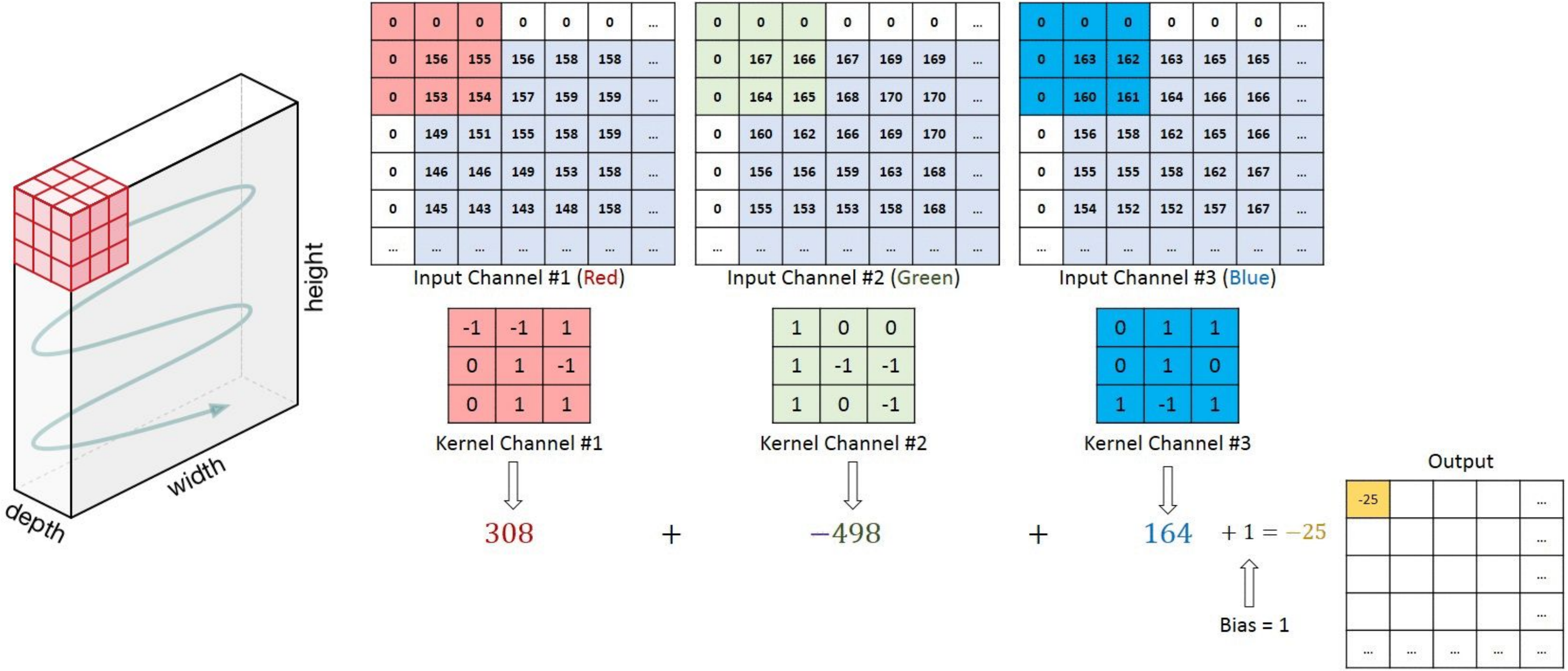
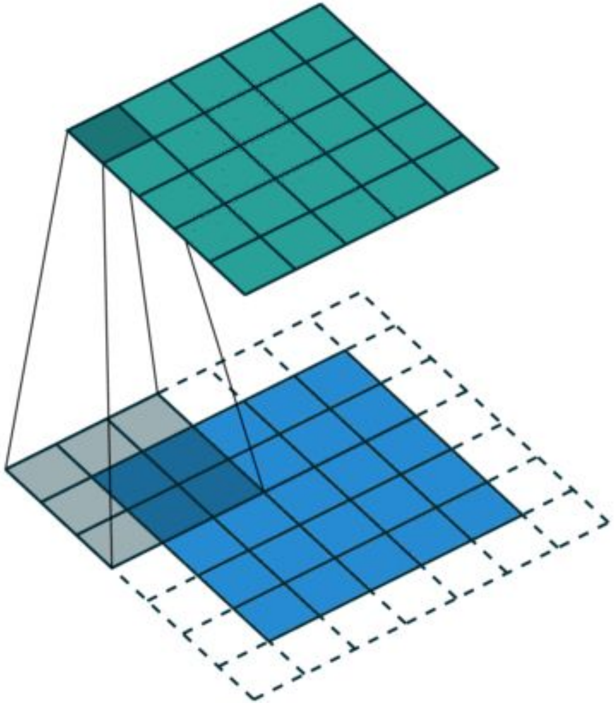


Image from TDS - comprehensive guide to CNN

In-class activity

Perform a simple convolution



Padding gives more room to the filter and retains the size after convolution

0	0	0	0	0	0	0	0
0	0	2	3	1	1	0	0
0	3	7	8	4	5	1	0
0	5	8	7	5	2	1	0
0	5	5	8	9	8	3	0
0	3	5	9	8	5	1	0
0	2	2	5	4	3	0	0
0	0	0	0	0	0	0	0

"Same" padding

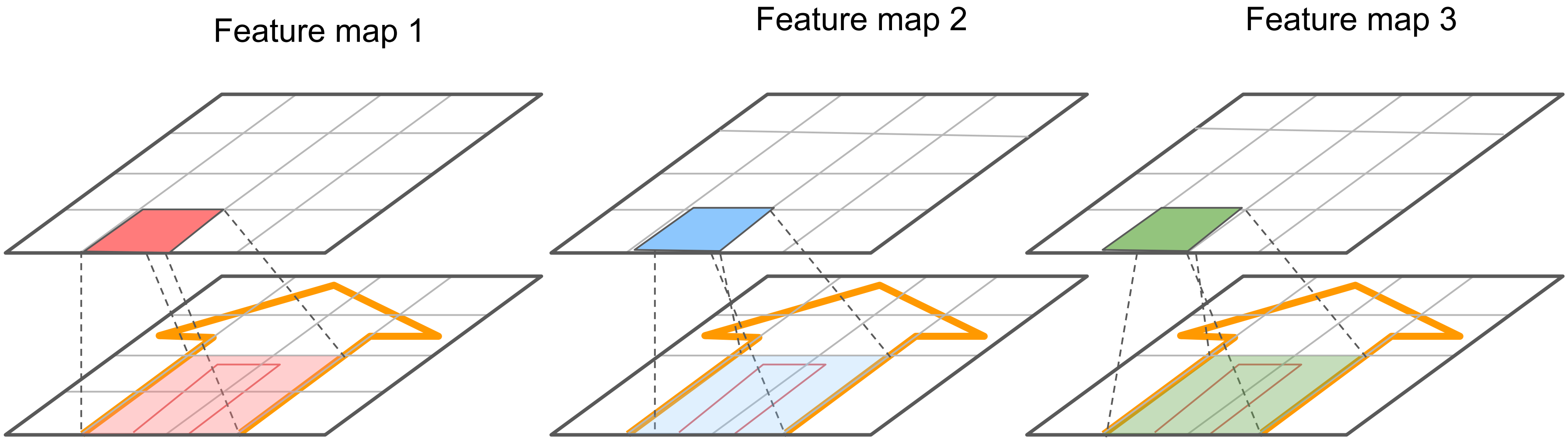
vs

no padding
(aka "valid" padding)

Padding gives more room to the filter and retains the size after convolution

0	0	0	0	0	0	0	0
0	0	2	3	1	1	0	0
0	3	7	8	4	5	1	0
0	5	8	7	5	2	1	0
0	5	5	8	9	8	3	0
0	3	5	9	8	5	1	0
0	2	2	5	4	3	0	0
0	0	0	0	0	0	0	0

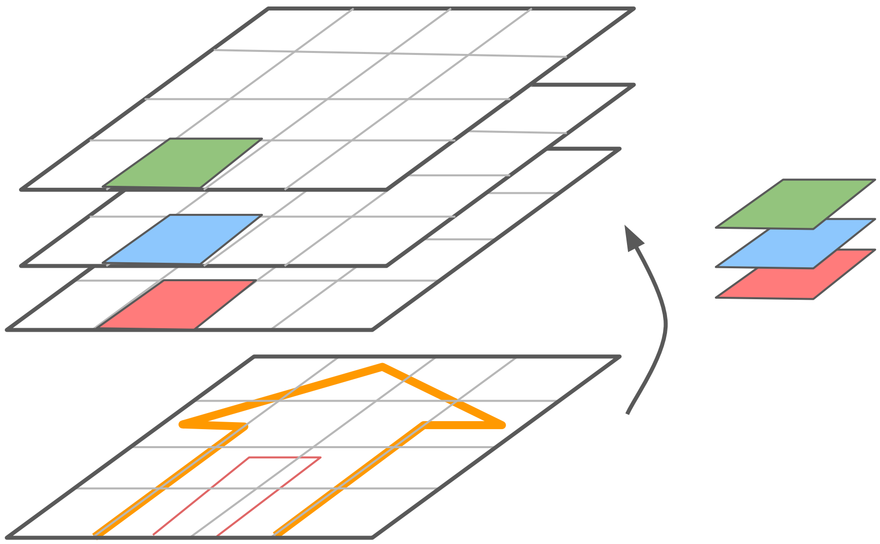
Each convolution is a filter/kernel that generates one feature map



In-class activity

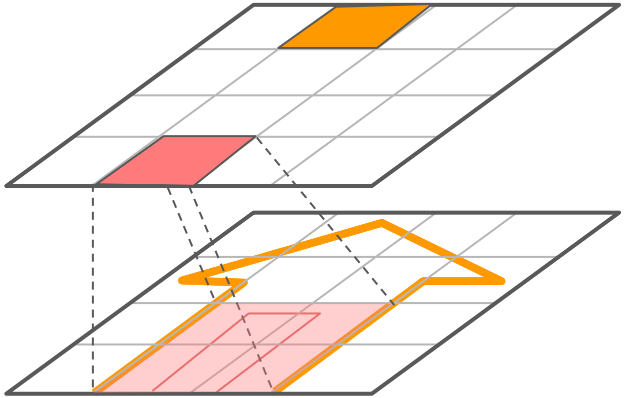
Visualizing the effect of kernels

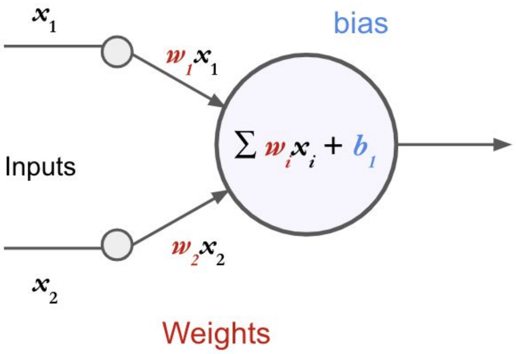
Each convolution is a filter/kernel that generates one feature map



Convolution filters are weights in a convolutional layer

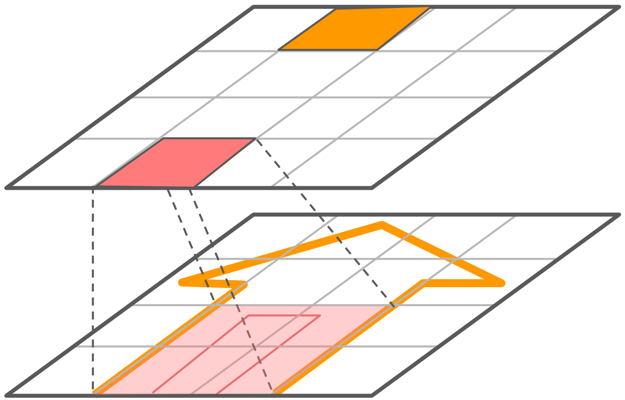
Feature map 1





All neurons in one feature map share same parameters, aka are generated with same kernel

Feature map 1

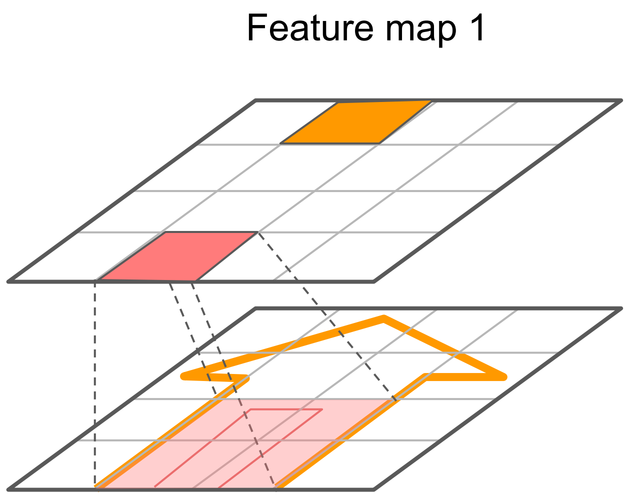


Same weights:

- Weights = convolution filter

Same bias

Weight sharing enhances efficiency and learning

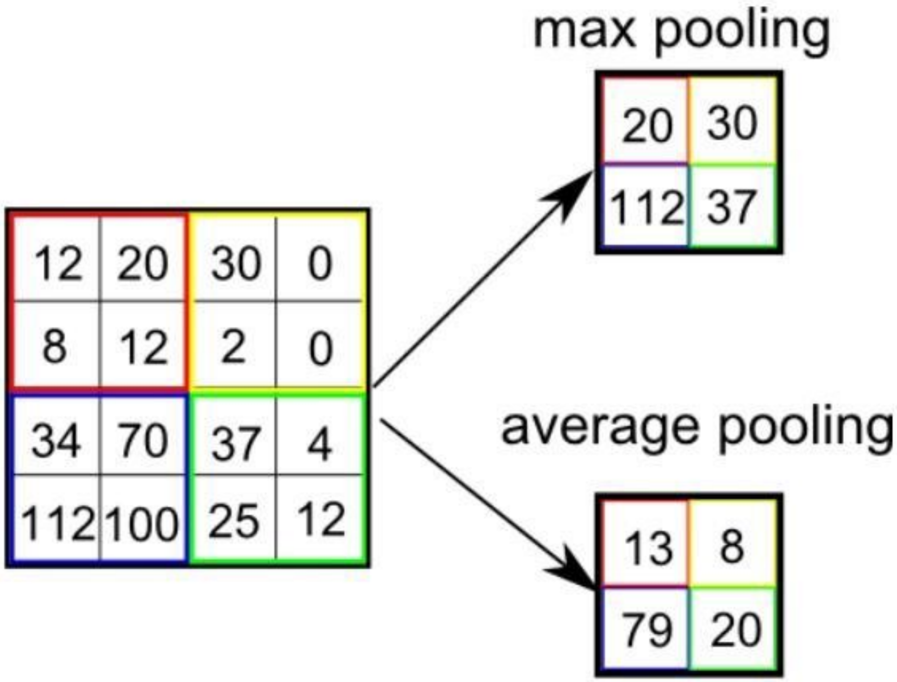


Same weights:

- Weights = convolution filter

Same bias

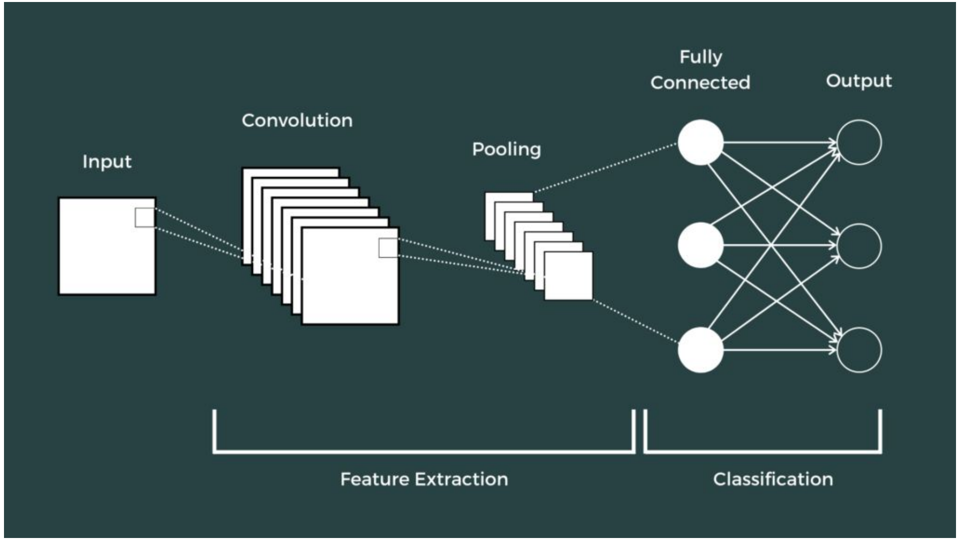
Pooling layers allow for reduction of the convolved features



2x2 pool
Stride = 2

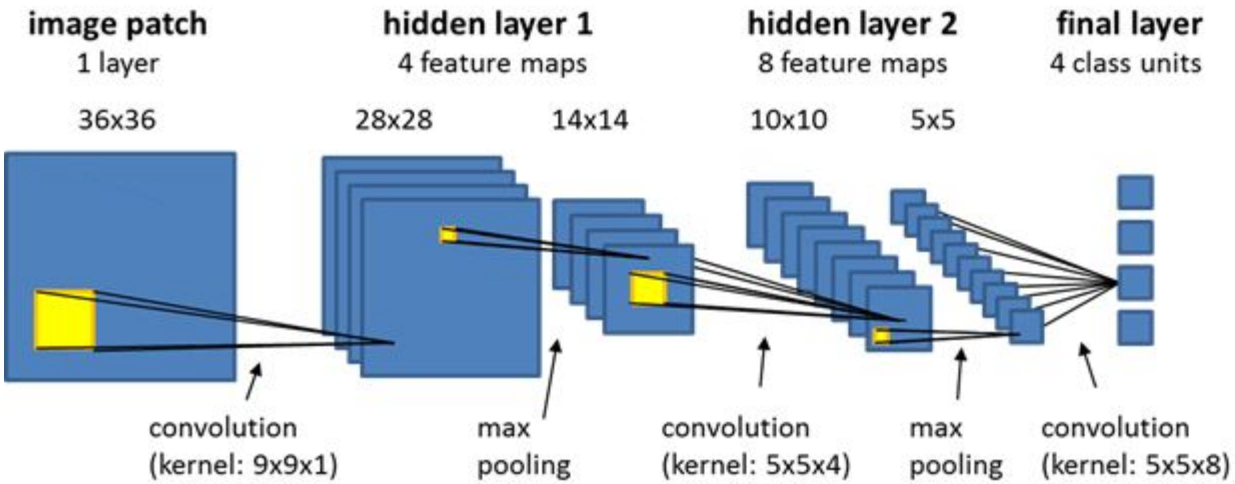
In-class activity

Shape of layers and architecture of model



In-class activity

Shape of layers and architecture of model



Translation/Rotation and the concept of *invariance*

CNNs are not truly invariant to rotation or translation (equivariant to translation)

Data augmentation is necessary

P4.

Some common CNN architectures

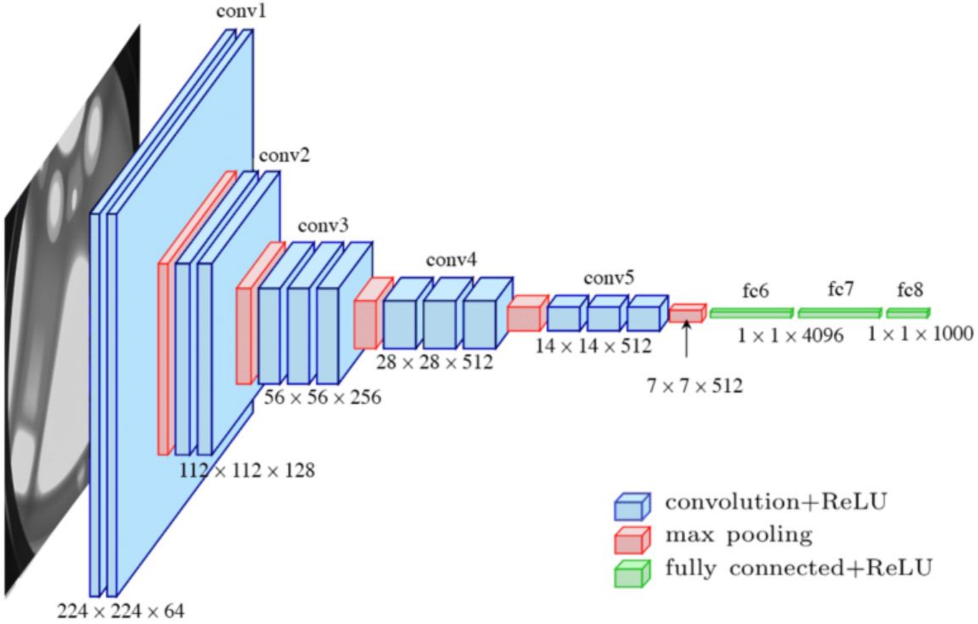


Image from Data Analytics: Different types of CNNs

The power of a curated dataset and competitions

IMAGENET

14,197,122 images, 21841 synsets indexed

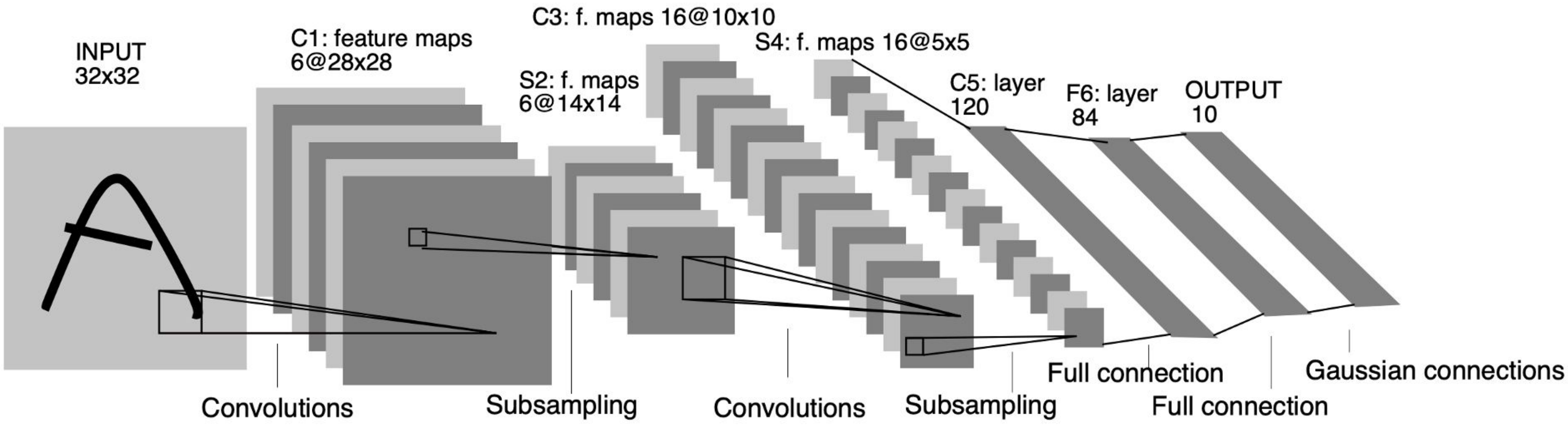
Home Download Challenges About

Not logged in. Login I Signup

ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. The project has been instrumental in advancing computer vision and deep learning research. The data is available for free to researchers for non-commercial use.

The power of a curated dataset and competitions

LeNet-5


INPUT
32x32
C1: feature maps
6@28x28
S2: f. maps
6@14x14
C3: f. maps 16@10x10
S4: f. maps 16@5x5
C5: layer
120
F6: layer
84
OUTPUT
10
Convolutions
Subsampling
Convolutions
Subsampling
Full connection
Full connection
Gaussian connections

AlexNet

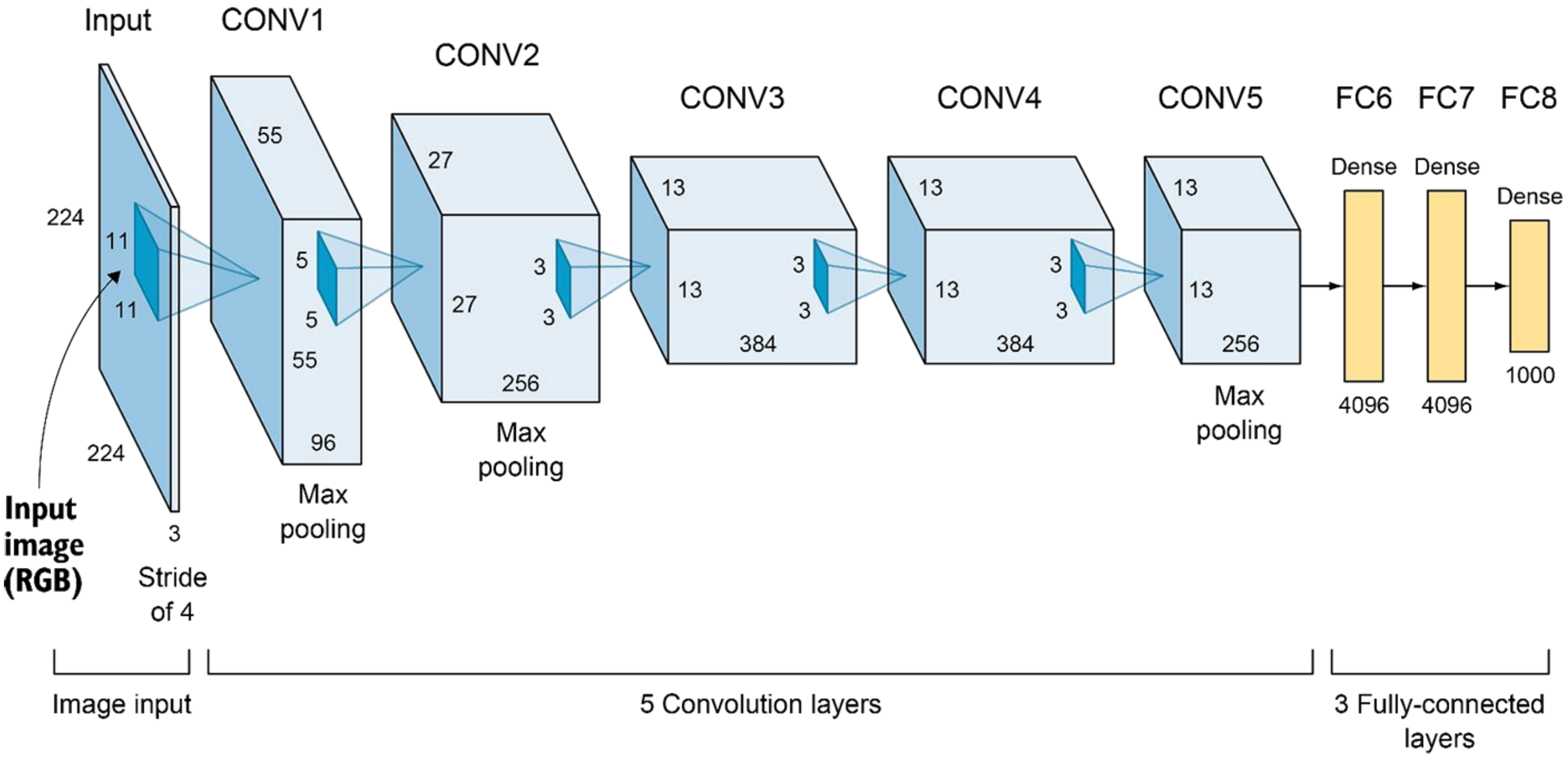
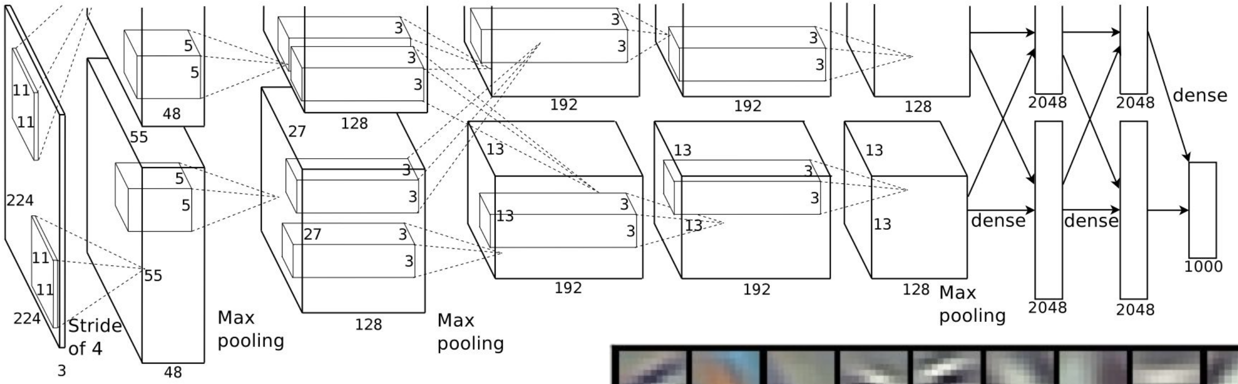
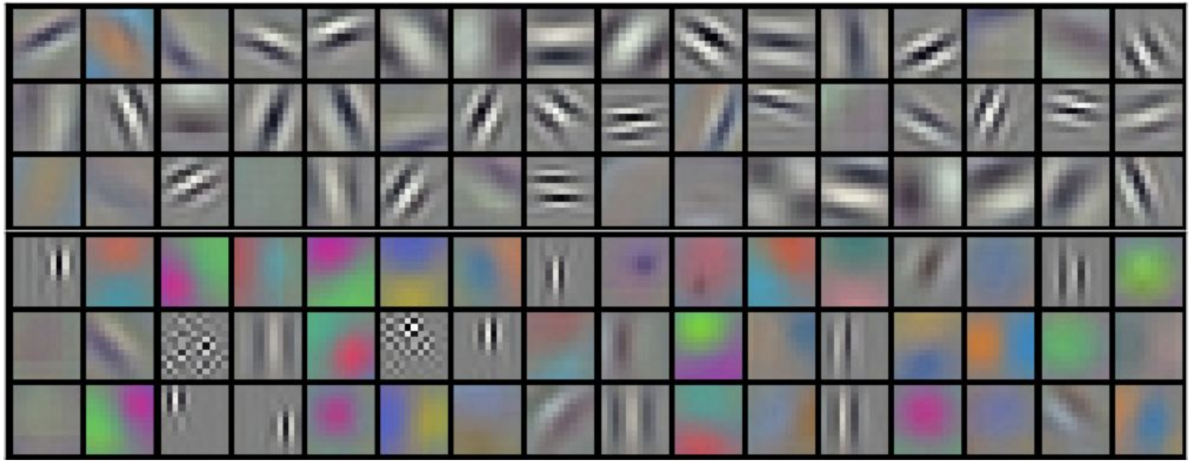


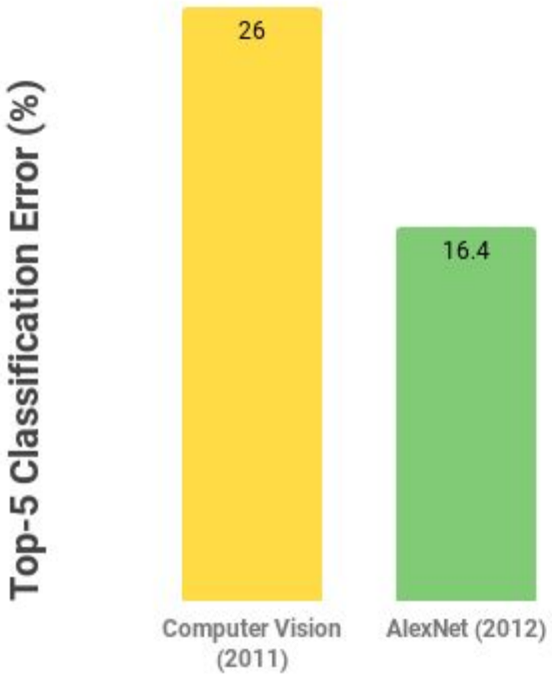
Image from Deep Learning for vision systems

AlexNet

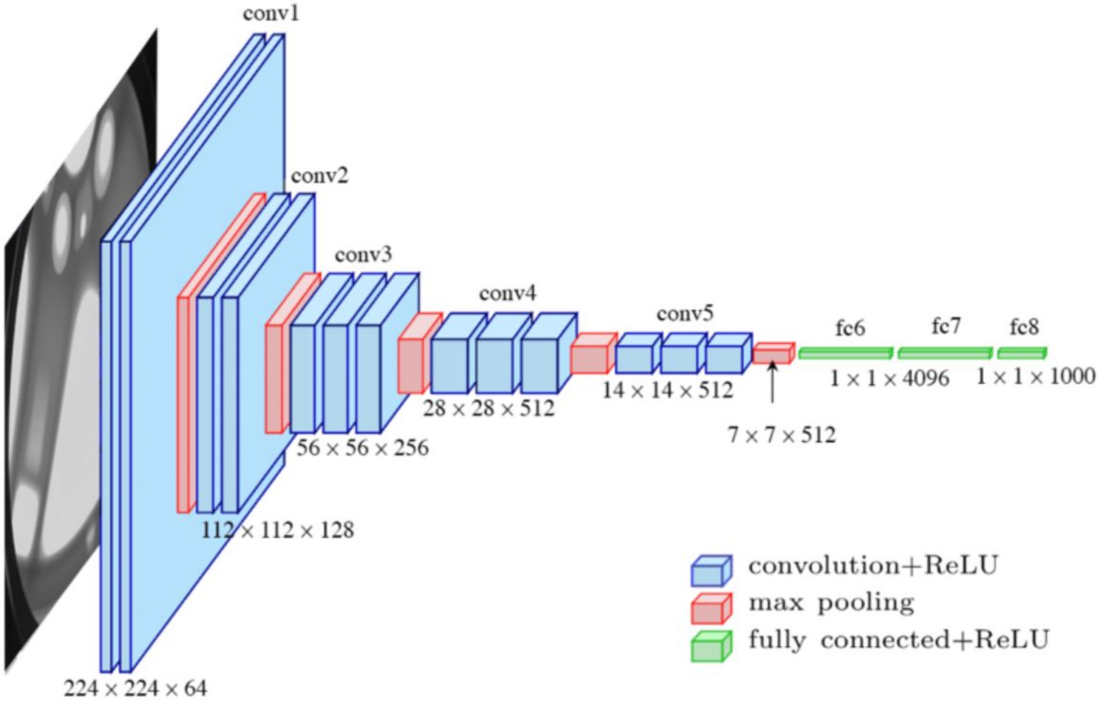

224
224
3
11
11
11
11
Stride
of 4
55
55
5
5
5
5
48
48
Max
pooling
27
27
3
3
3
3
3
3
128
128
128
Max
pooling
13
13
3
3
3
3
192
192
13
13
3
3
3
3
192
192
13
13
128
128
Max
pooling
dense
dense
dense
2048
2048
2048
2048
1000



The power of a curated dataset and competitions



VGGNet


conv1
conv2
conv3
conv4
conv5
fc6
fc7
fc8
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 × 512
1 × 1 × 4096
1 × 1 × 1000
convolution+ReLU
max pooling
fully connected+ReLU

VGGNet

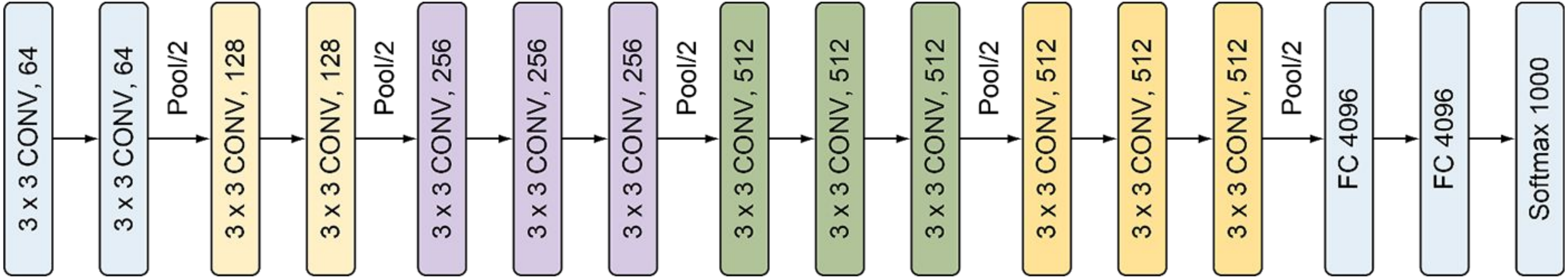


Image from Deep Learning for vision systems

VGGNet

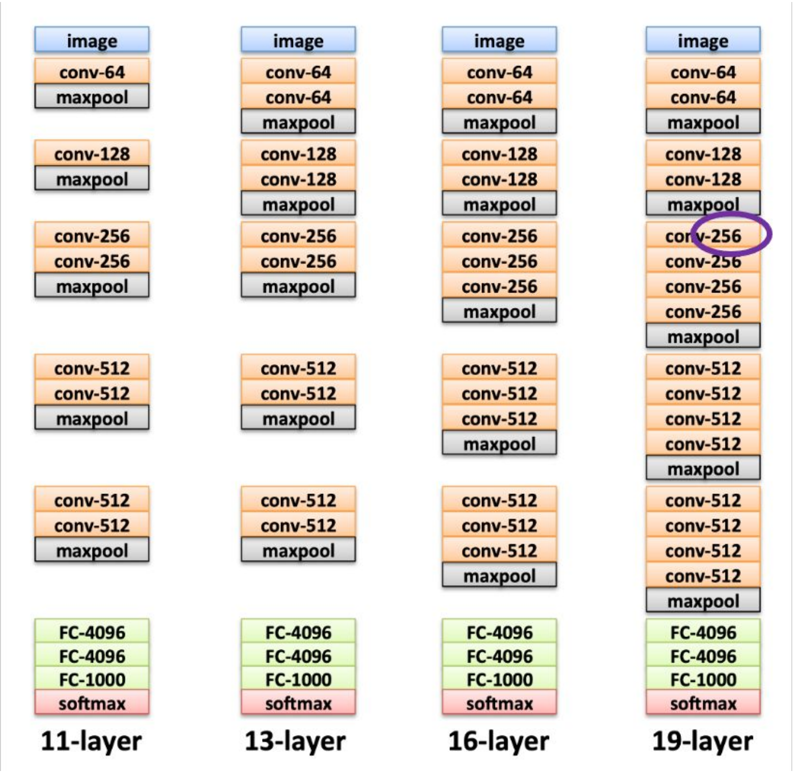
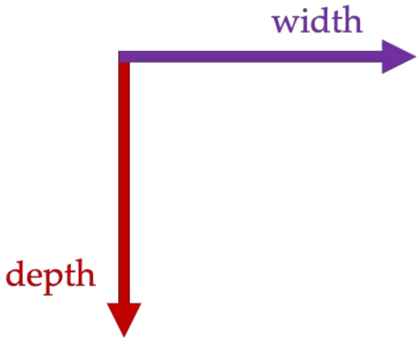
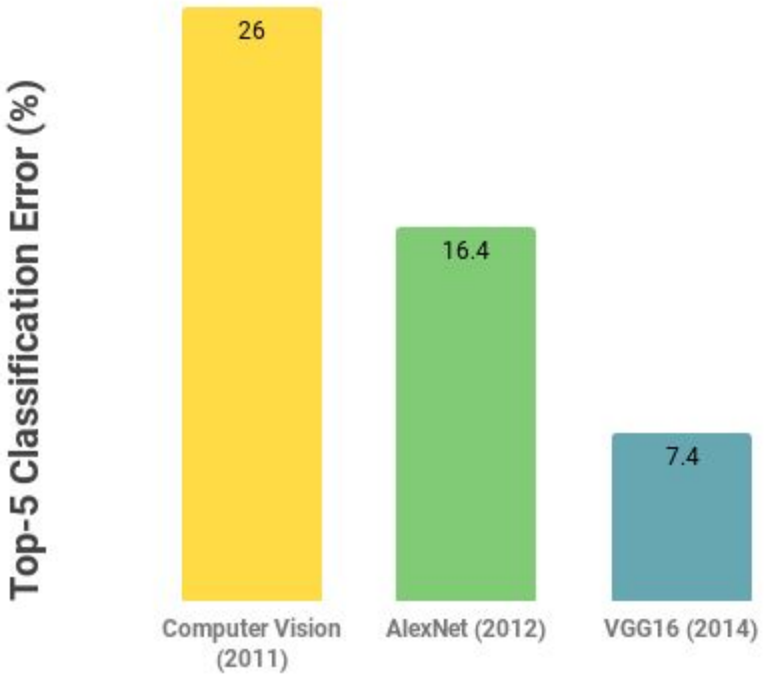


image
conv-64
maxpool
conv-128
maxpool
conv-256
conv-256
maxpool
conv-512
conv-512
maxpool
conv-512
conv-512
maxpool
FC-4096
FC-4096
FC-1000
softmax
11-layer
image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
maxpool
conv-512
conv-512
maxpool
conv-512
conv-512
maxpool
FC-4096
FC-4096
FC-1000
softmax
13-layer
image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
conv-256
maxpool
conv-512
conv-512
conv-512
maxpool
conv-512
conv-512
conv-512
maxpool
FC-4096
FC-4096
FC-1000
softmax
16-layer
image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
conv-256
conv-256
maxpool
conv-512
conv-512
conv-512
conv-512
maxpool
conv-512
conv-512
conv-512
conv-512
maxpool
FC-4096
FC-4096
FC-1000
softmax
19-layer

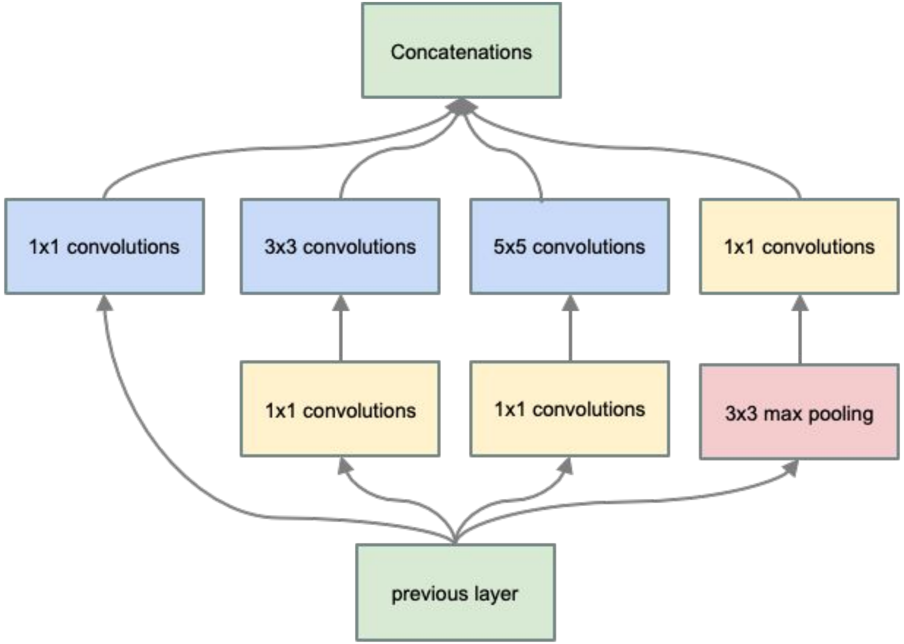


width
depth

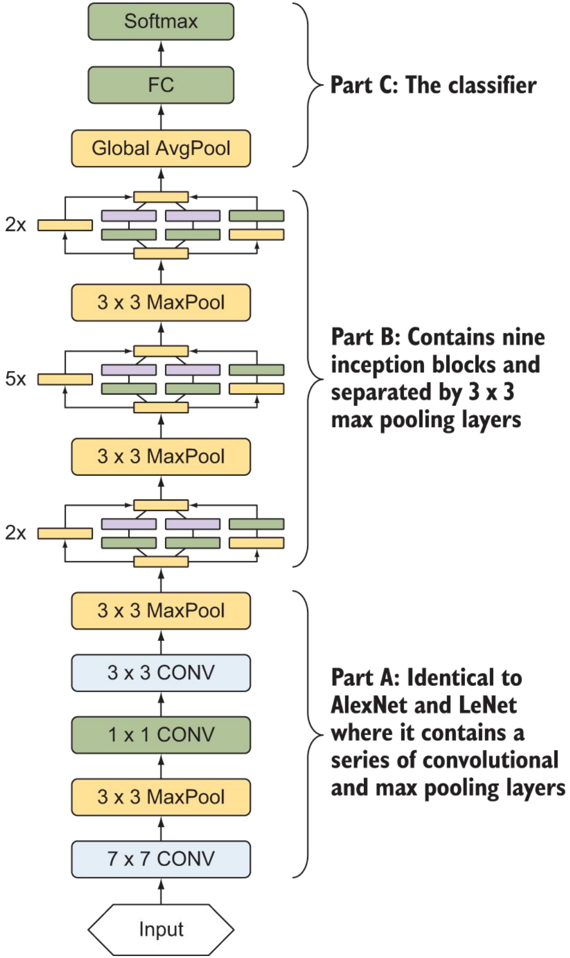
The power of a curated dataset and competitions



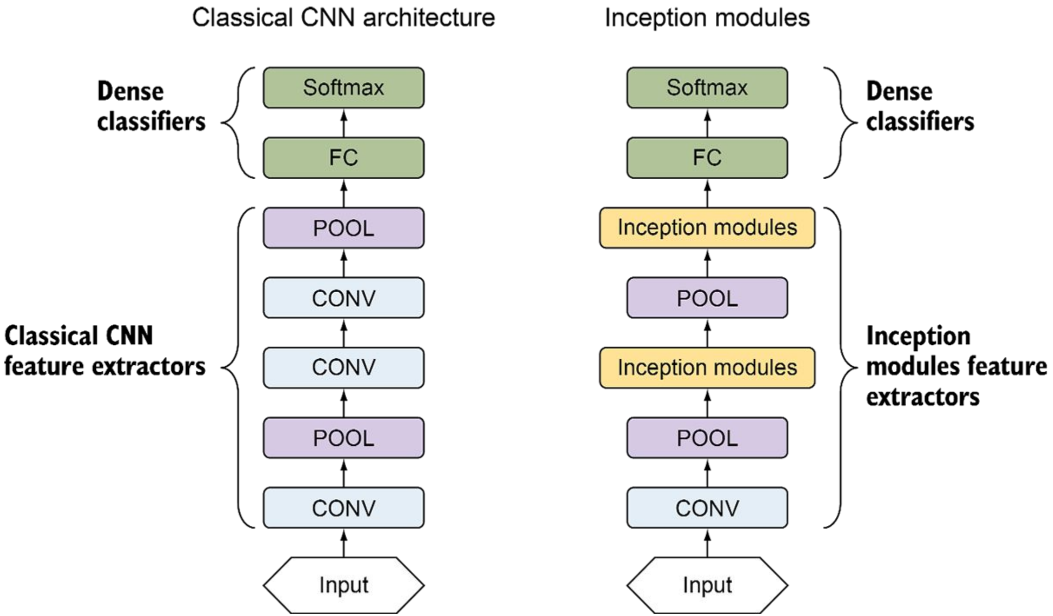
GoogLeNet – Inception



Inception module



GoogLeNet – Inception



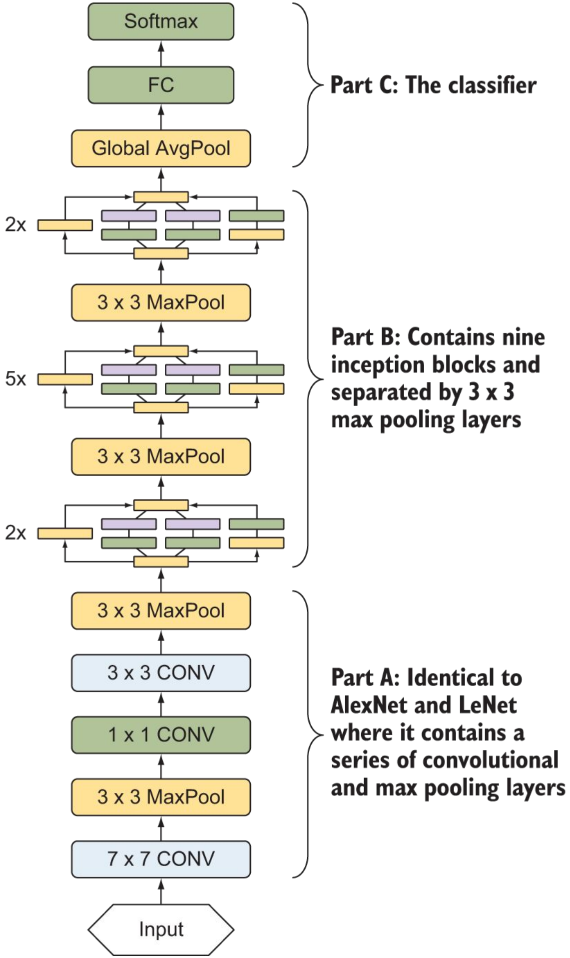
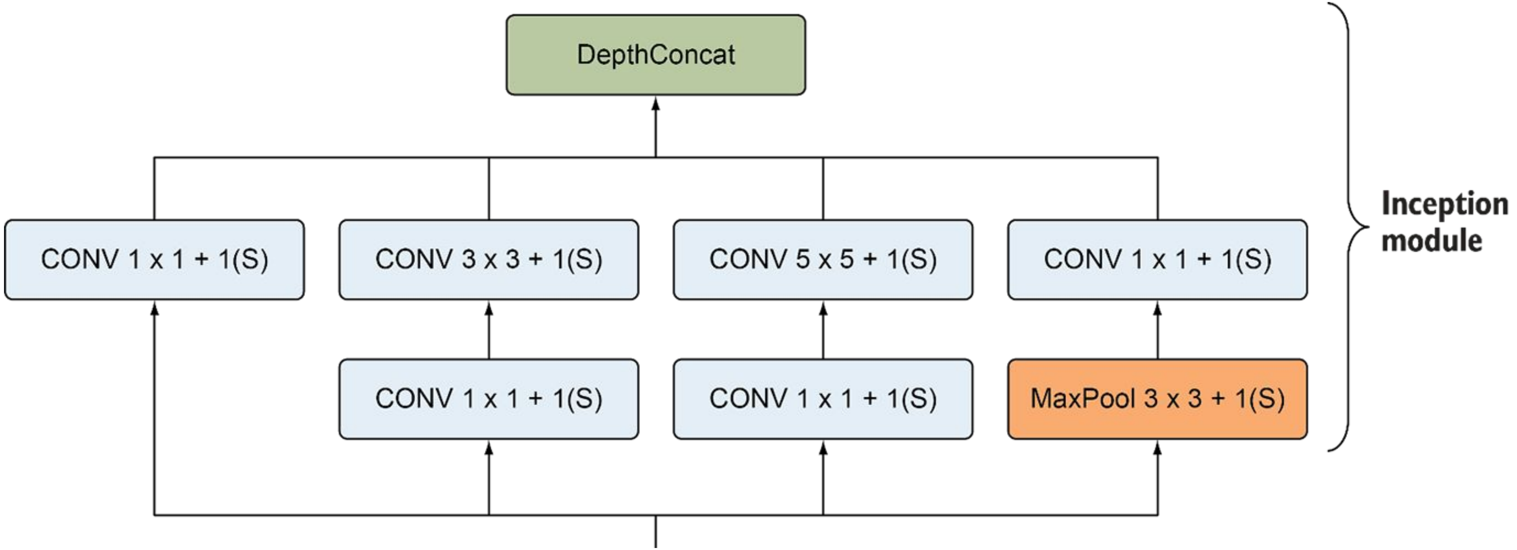


Image from Deep Learning for vision systems

GoogleLeNet – Inception



The power of a curated dataset and competitions

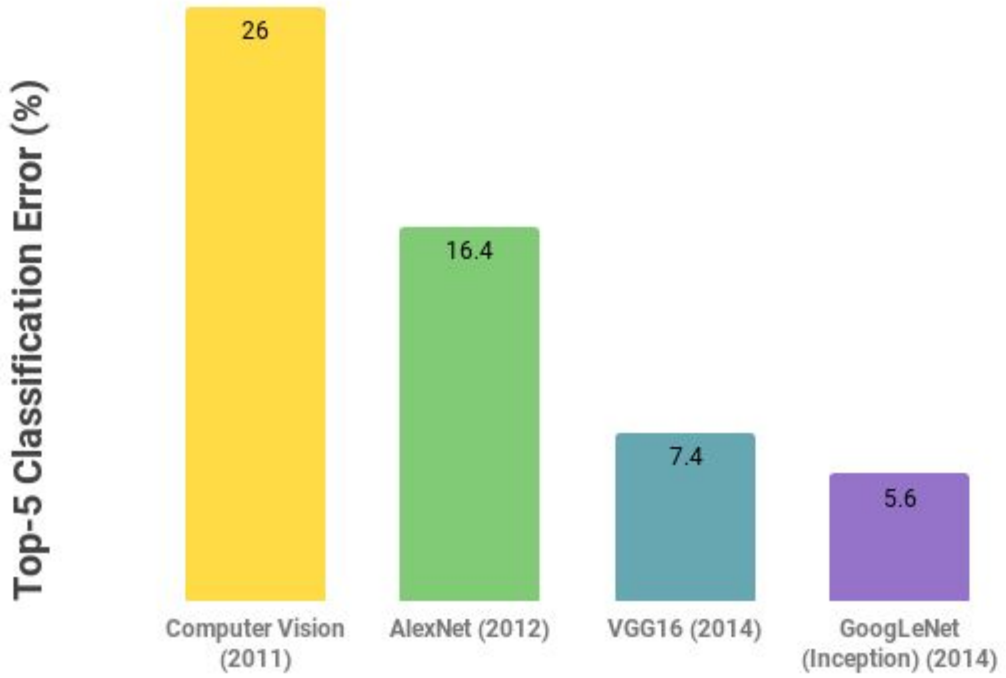
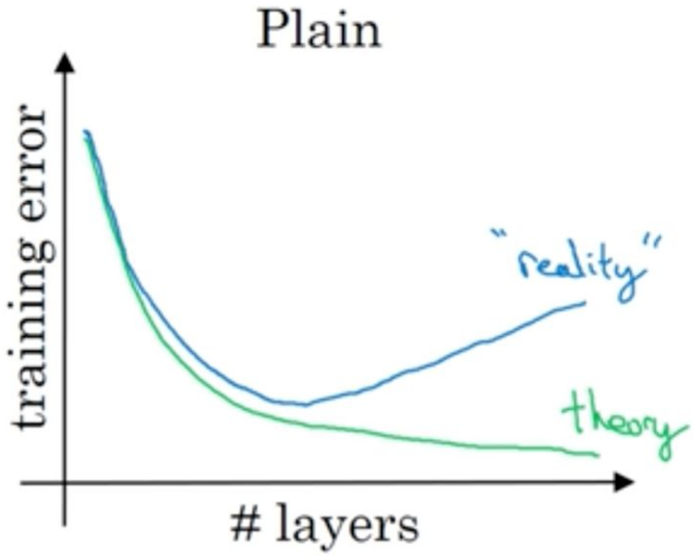
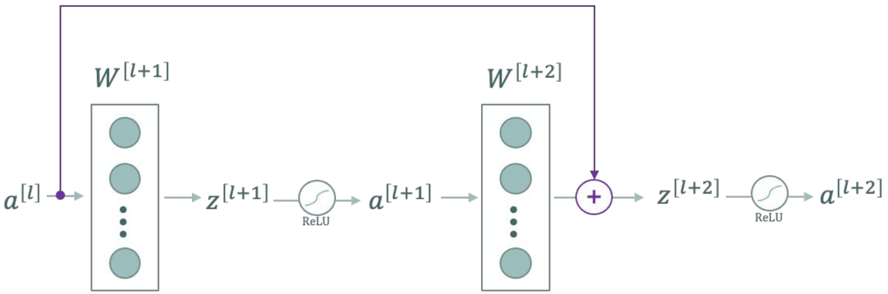


Image from medium - timeline of transfer learning models

ResNet: the problem with deep layers



ResNet: the problem with deep layers



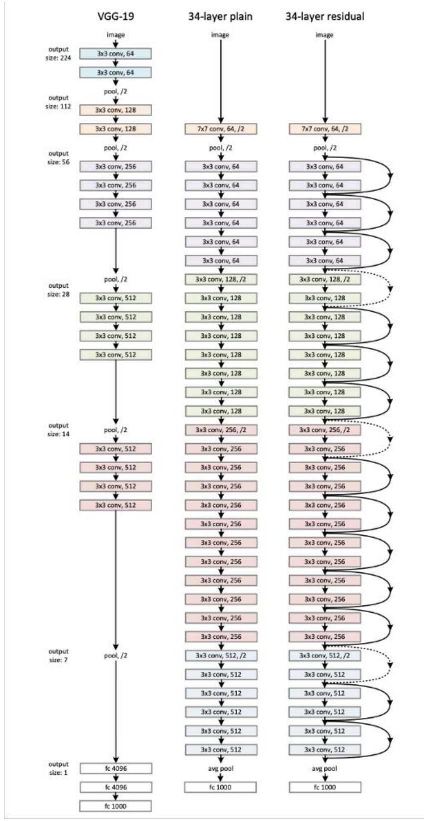
$$z^{[l+1]} = W^{[l+1]} a^{[l]} + b^{[l+1]}$$

$$a^{[l+1]} = \mathrm{ReLU}(z^{[l+1]})$$

$$z^{[l+2]} = W^{[l+2]} a^{[l+1]} + b^{[l+2]}$$

$$\cancel{a^{[l+2]} = \mathrm{ReLU}(z^{[l+2]})}$$

$$a^{[l+2]} = \mathrm{ReLU}(z^{[l+2]} + a^{[l]})$$



ResNet: the problem with deep layers

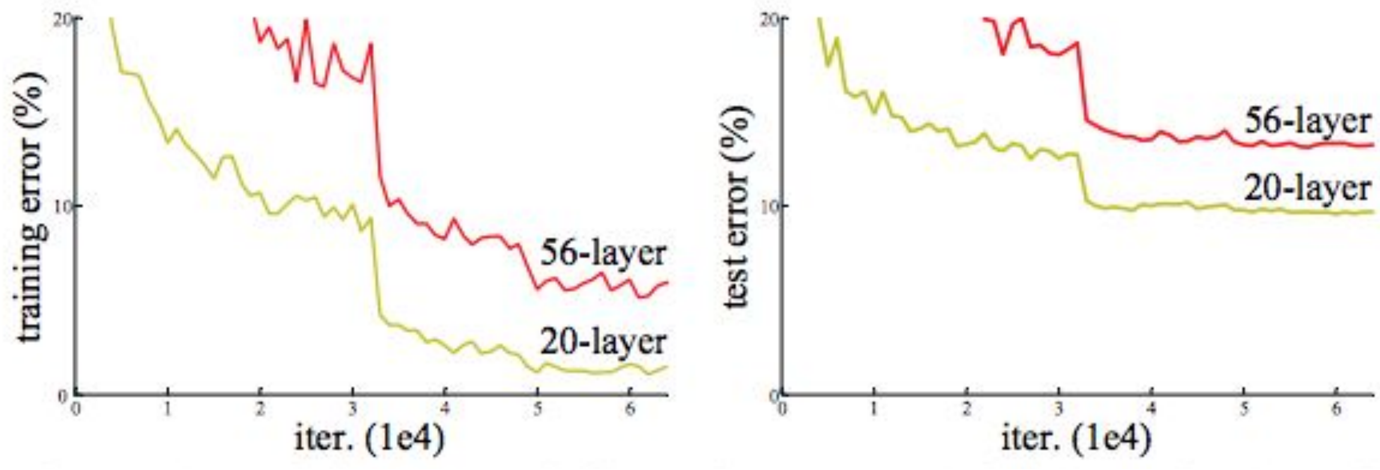
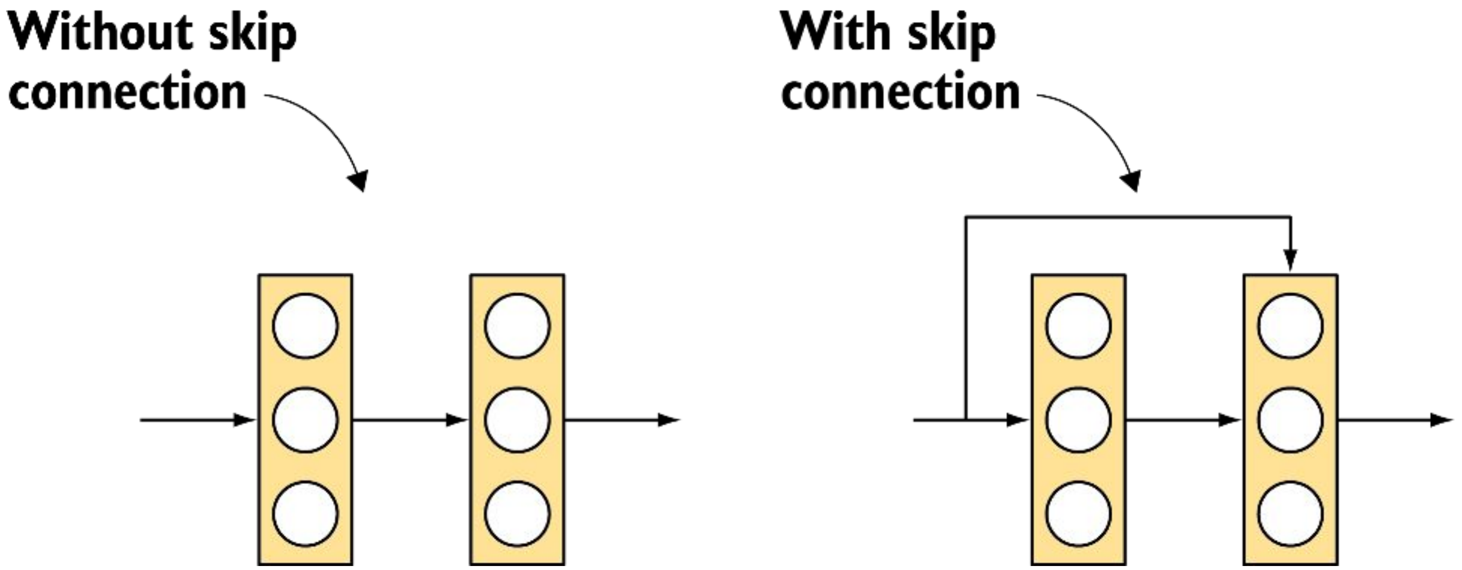
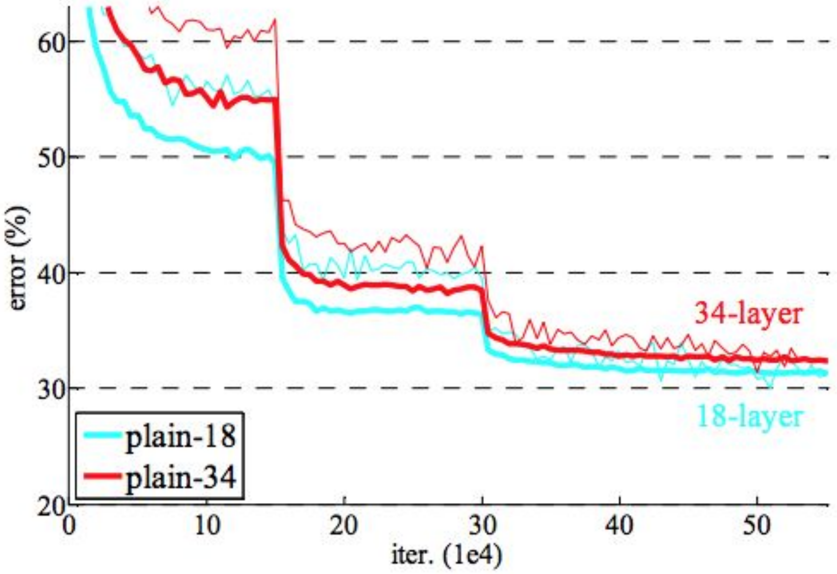


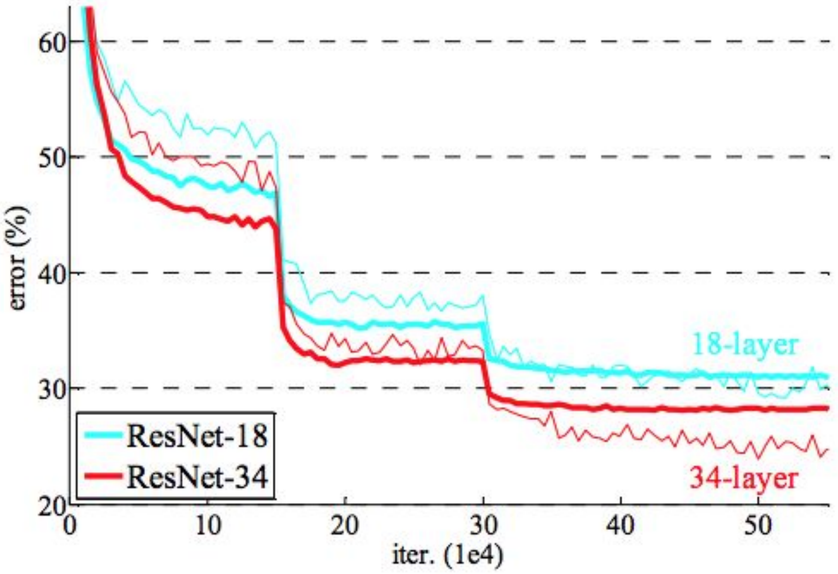
Image from TDS - introduction to ResNet

ResNet: Residual networks and the idea of skip



ResNet: Residual networks and the idea of skip





ResNet: Residual networks and the idea of skip

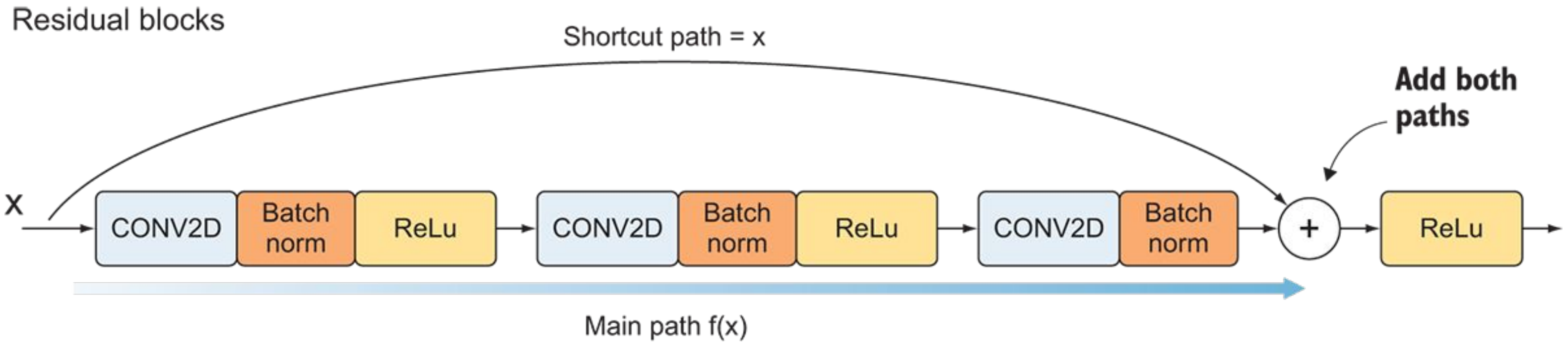
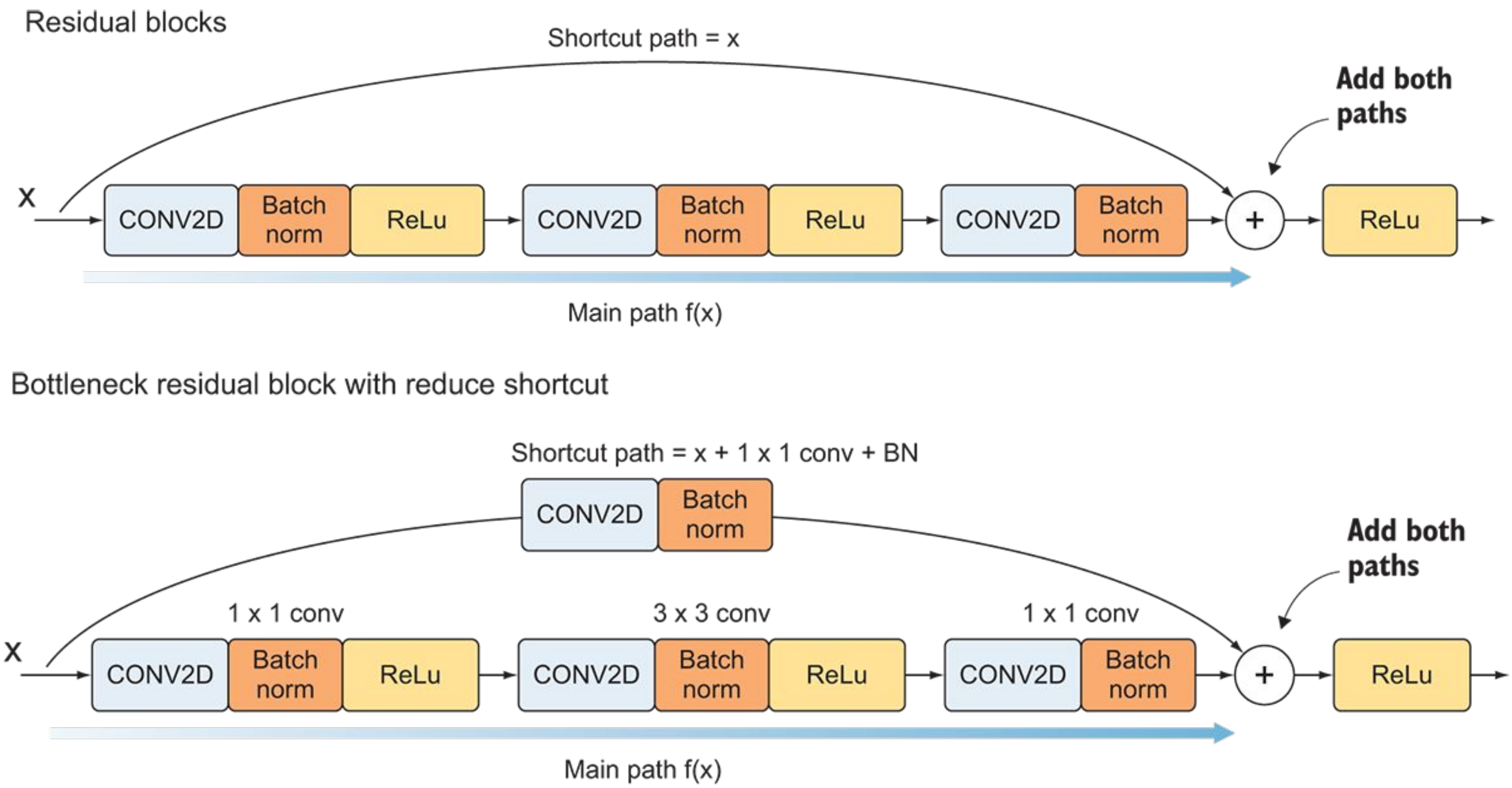
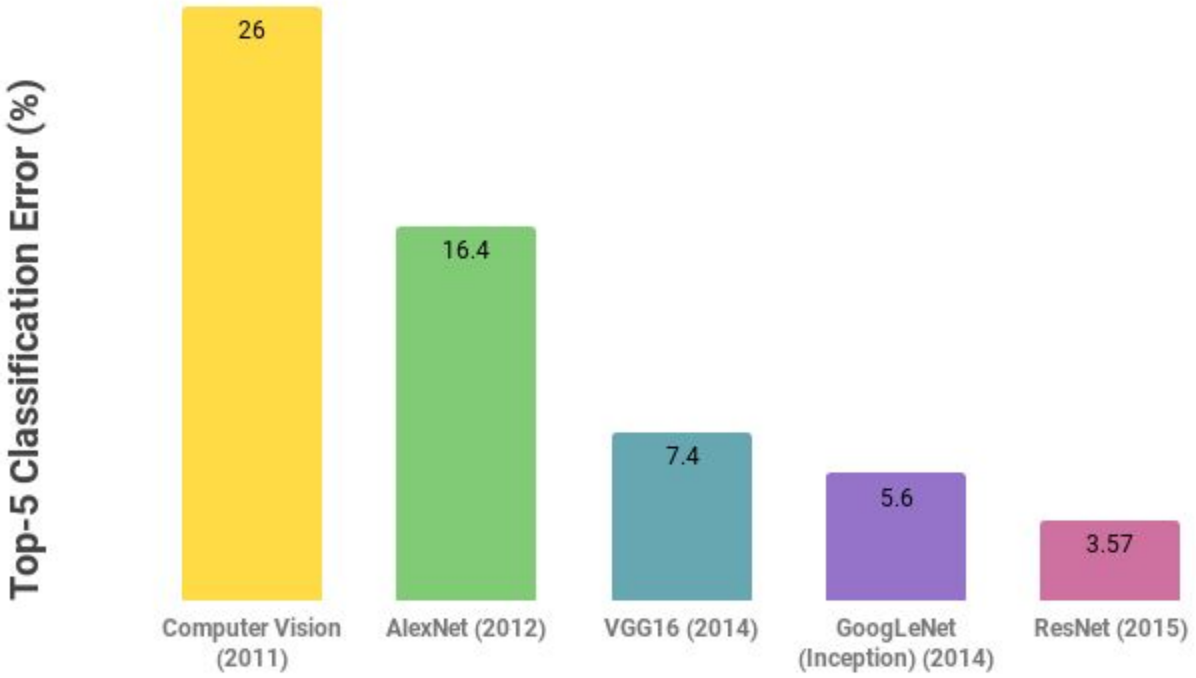


Image from Deep Learning for vision systems

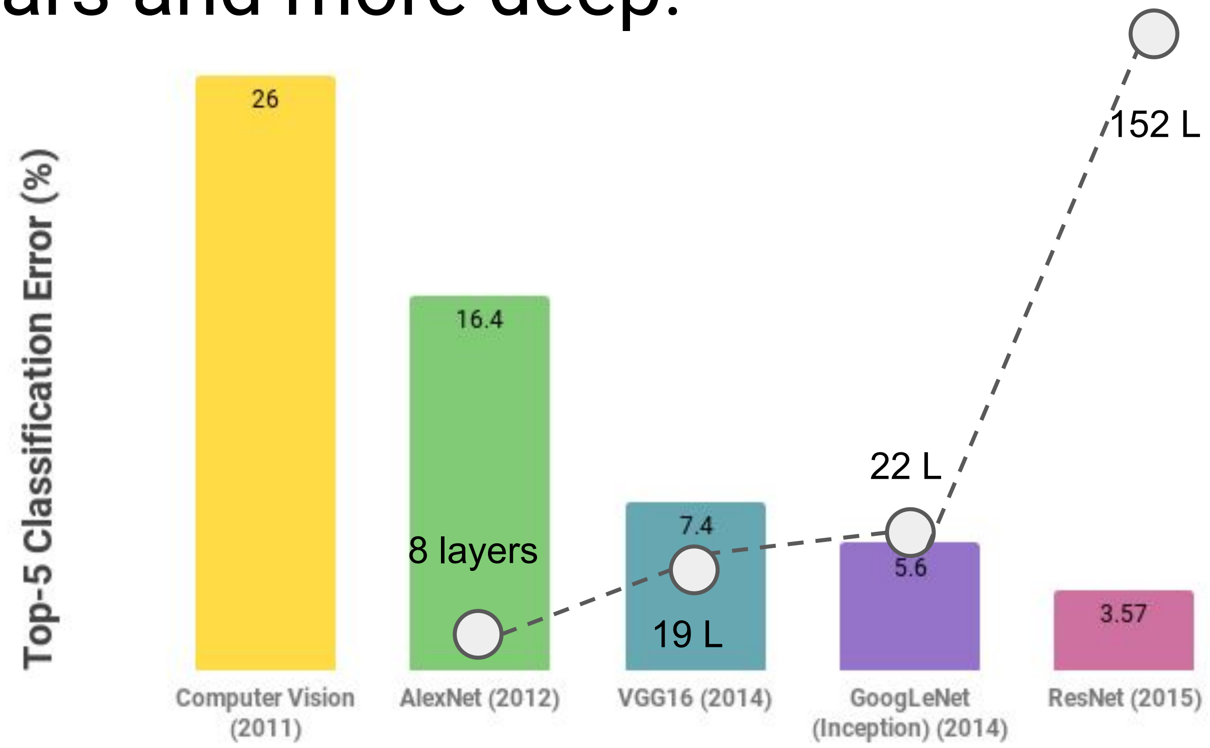
ResNet: Residual networks and the idea of skip



The networks have become more powerful over the years



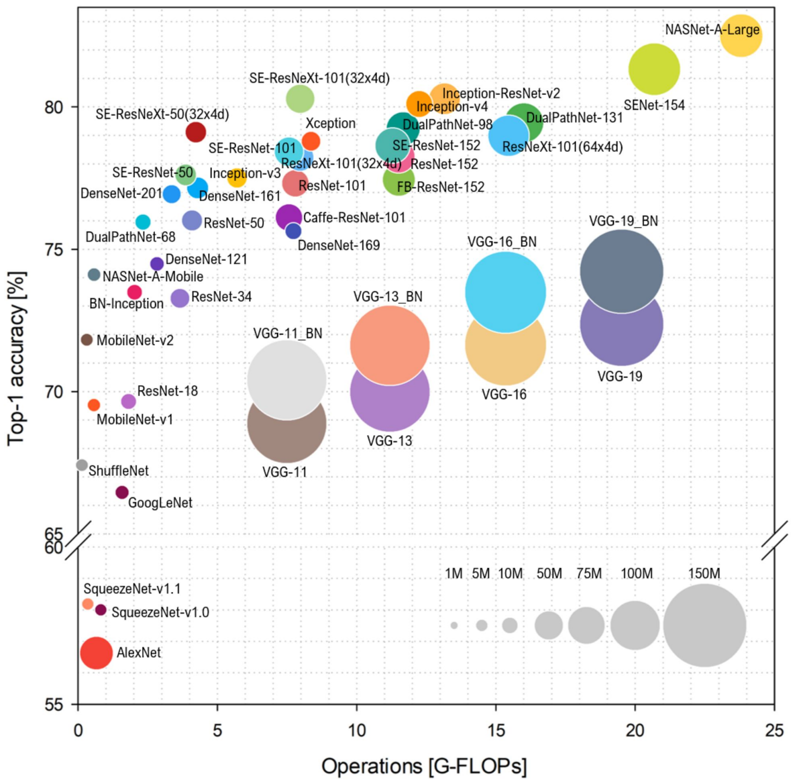
The networks have become more powerful over the years and more deep!

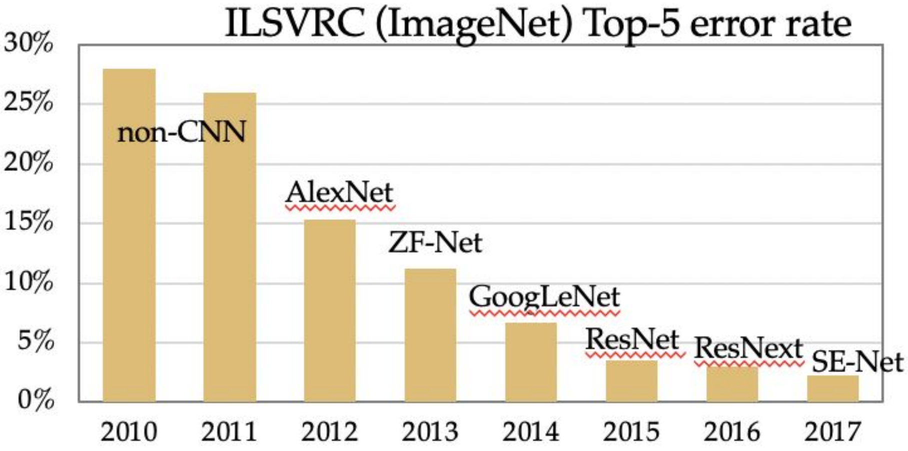


Timeline of CNN architectures

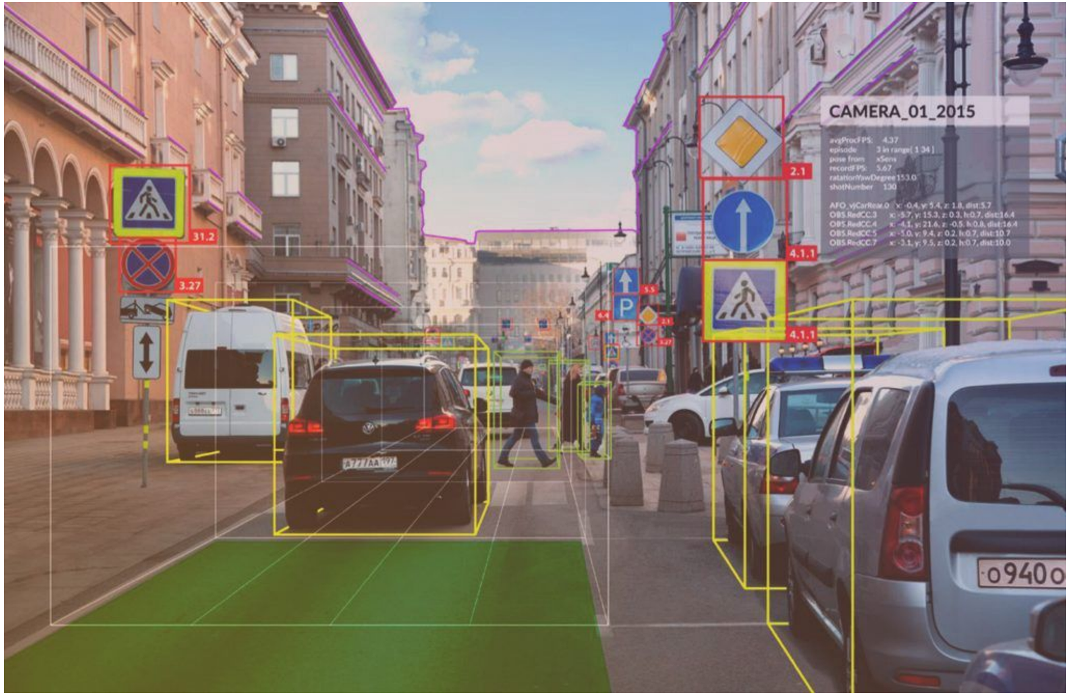
S. Bianco, et al., Benchmark Analysis of Representative Deep Neural Network Architectures, arXiv:1810.00736v2 [cs.CV] 19 Oct 2018

Timeline of CNN architectures





P5.
CNN applications today!



Machine Learning for Pneumonia Detection on Lung Ultrasound

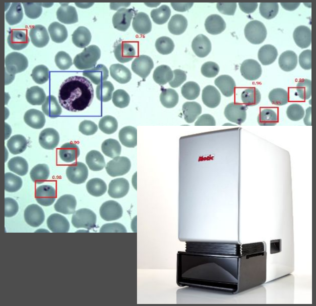
18 November 2021

Courosh Mehanian, Research Associate Professor

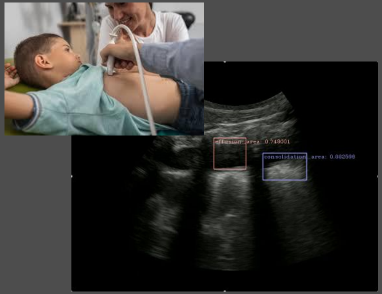
Examples of Global Health projects

GH+ Labs

AI-powered Malaria microscope



AI-powered ultrasound for pneumonia detection



Automated Visual Evaluation for cervical precancer screening

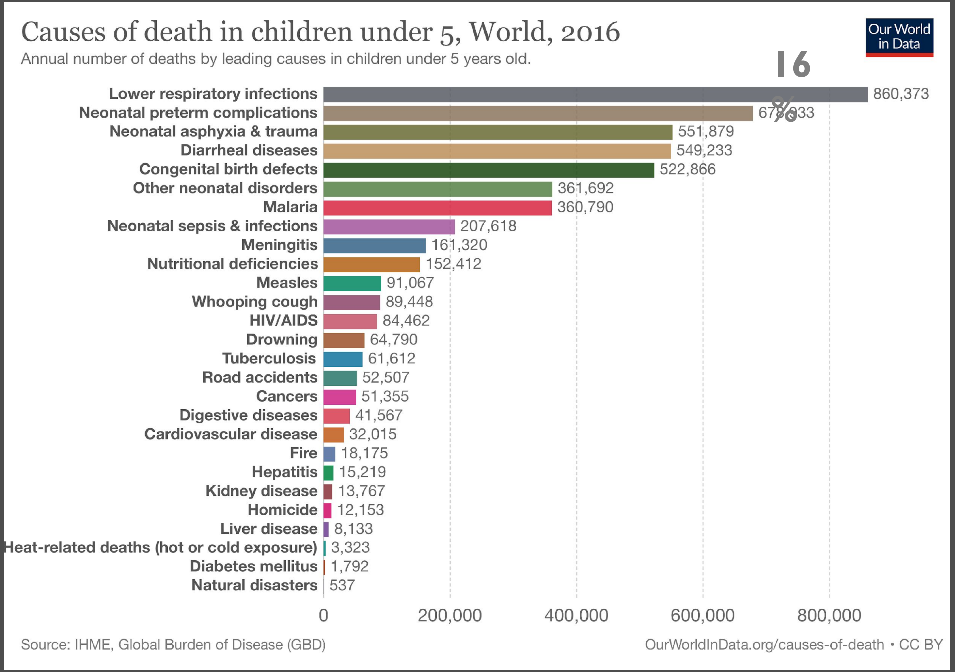
Arktek for cold-chain transport

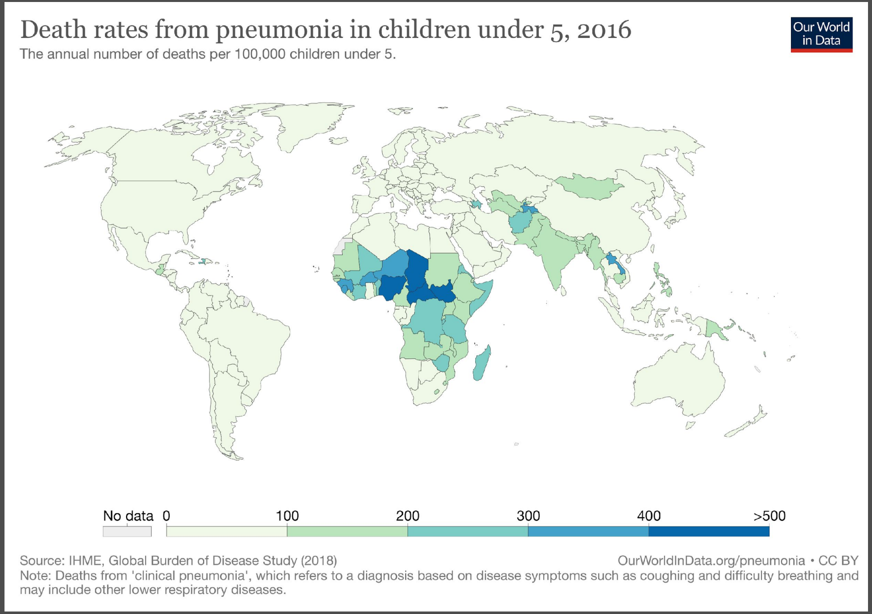
Our mission

We build tools and technologies to address unmet needs in primary healthcare centers and last-mile service delivery in low- and middle-income countries (LMICs) around the world.

Childhood pneumonia burden of disease

How can AI assist in early diagnosis of pneumonia in LMICs?

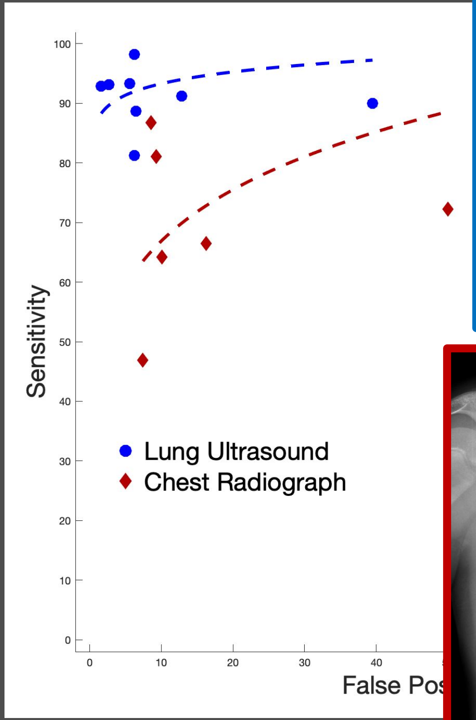


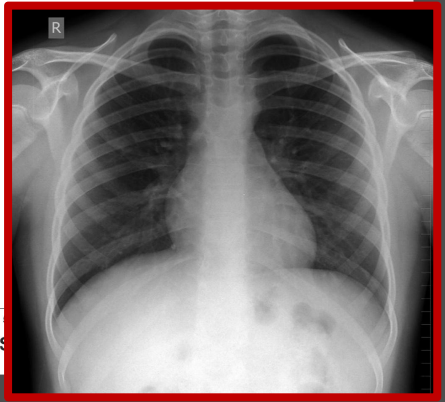


- 2,400 children per day died in 2016
- Most prevalent in South Asia and sub-Saharan Africa
- Can be detected and treated

Lung ultrasound for pneumonia diagnosis?

Ultrasound features are artifact, not anatomy





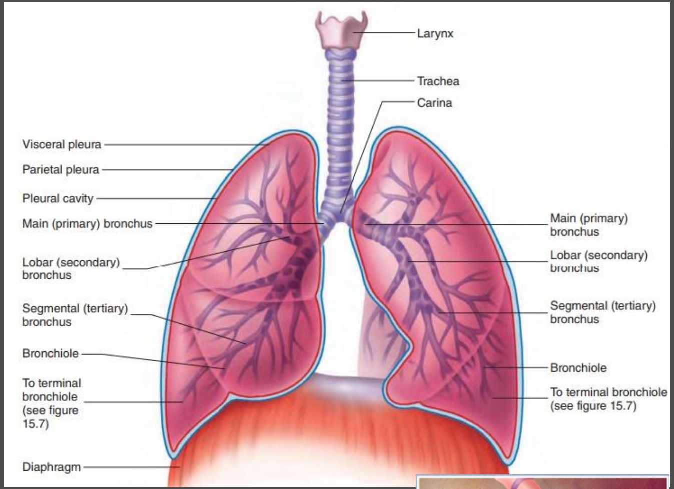
X-ray allows visualization of chest and lung

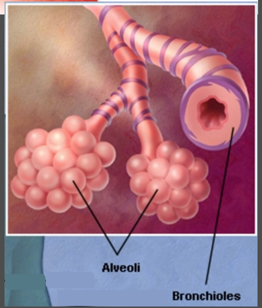
Advantages	LUS	CXR
Accuracy	good	medium
Safety	no radiation	ionizing radiation
Availability	point-of-care	needs referral
Portability	portable	usually fixed
Equipment cost	low	
Barriers	**LUS**	**CXR**
Acquisition skill	medium	low
Interpretation skill	difficult	medium
Ubiquity	not widely used	standard-of-care

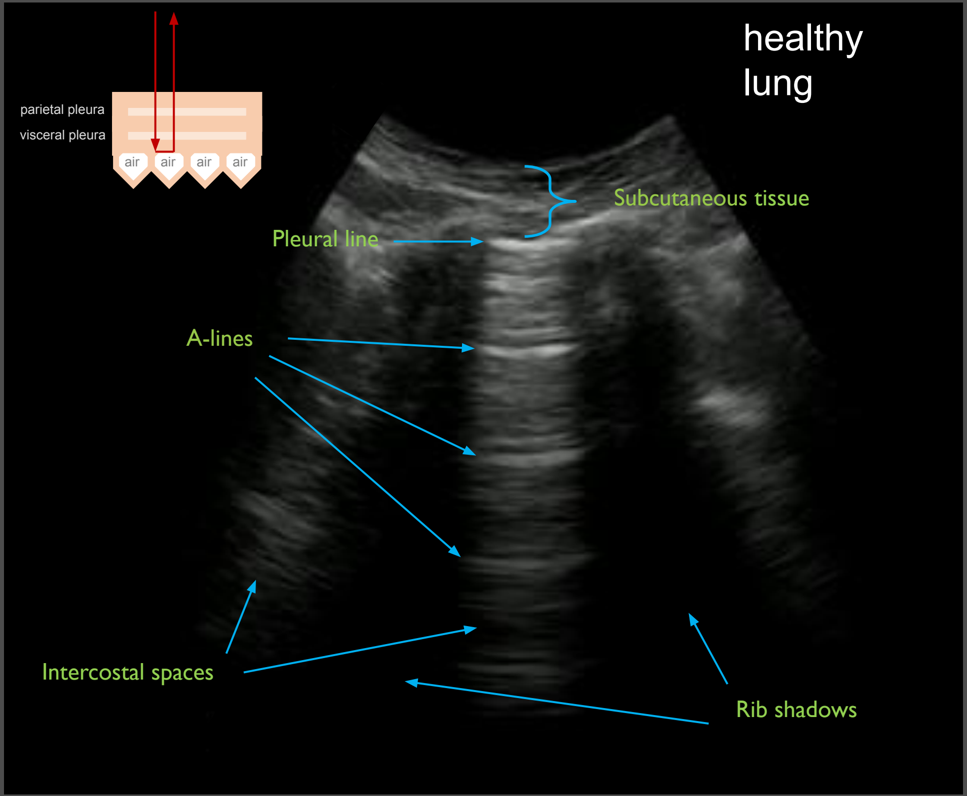
Lung ultrasound adoption could be enhanced by
AI-assisted interpretation, especially in LMICs

Lung ultrasound 101

Acoustic artifacts provide recognizable features

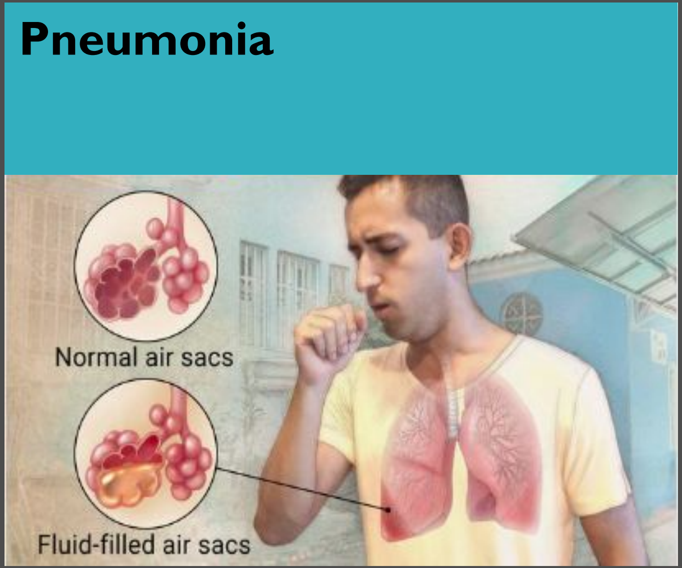


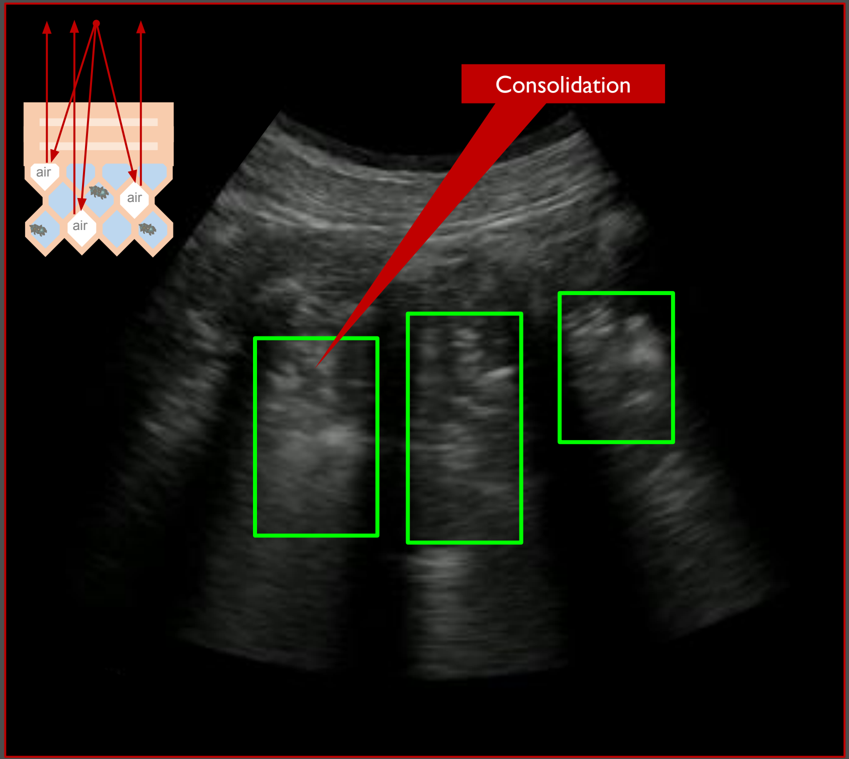




Typical pneumonia in lung ultrasound

Lung infection/inflammation partially fills lung air spaces with fluid

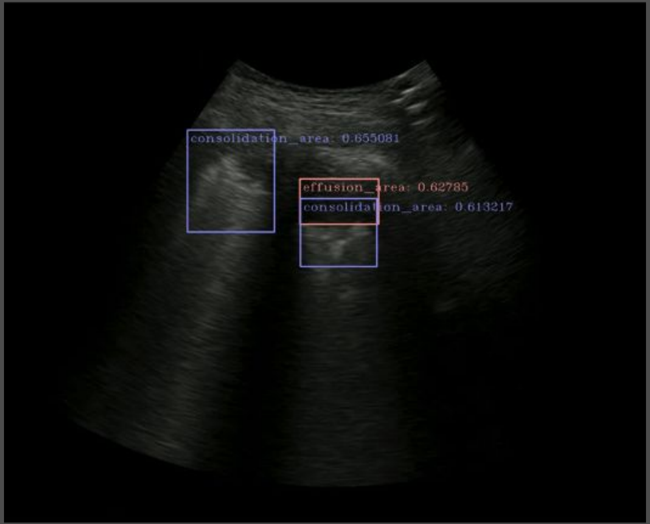




Use cases and requirements

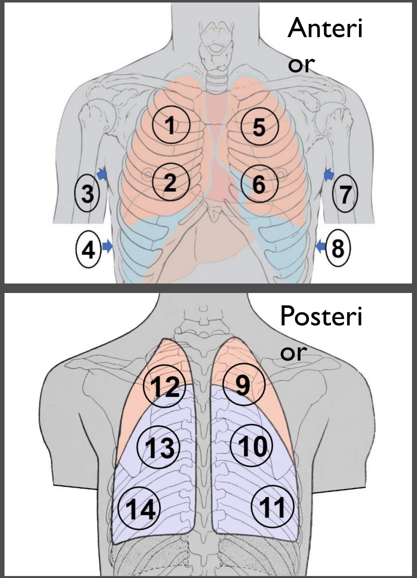
- Patient with pneumonia-like symptoms
- Evaluation includes AI-assisted ultrasound plus other clinical measurements
- Patient is released, prescribed antibiotics, or referred for follow-up

- Detect consolidation and pleural effusion
- Adults ($\geq$ 18 years old)
- Children ($\geq$ 10 kg and $<$ 18 years old)
- Body Mass Index (BMI $\leq 40\ \mathrm{kg/m^2}$)
- Sensitivity $\geq 90\%$ at video level
- Specificity $\geq 90\%$ at video level
- Draw bounding boxes around detections
- Time-to-results ≤ 1 minute, post-exam on Samsung S4 tablet
- Data collected by sonographer trained in lung ultrasound



Data collection

Image Acquisition ⟷ Image Annotation ⟷ Algorithm Development



Oregon Health & Science University

OHSU Doernbecker Children's Hospital

Columbia University Medical Center

MedStar Hospital Washington, DC

Brooke Army Medical Center, San Antonio, TX

Womack Army Medical Center, Fort Bragg, NC

Tripler Army Medical Center, Honolulu, HI

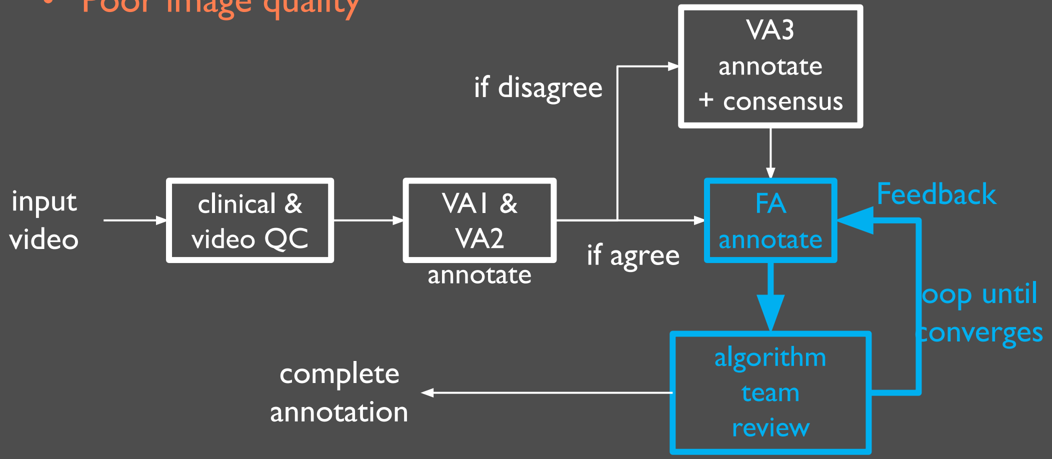
	Children	Adults
Patients	306	617
Ultrasound exams	320	1,225
Annotated videos	1,726	6,498
Annotated frames	103,560	389,880

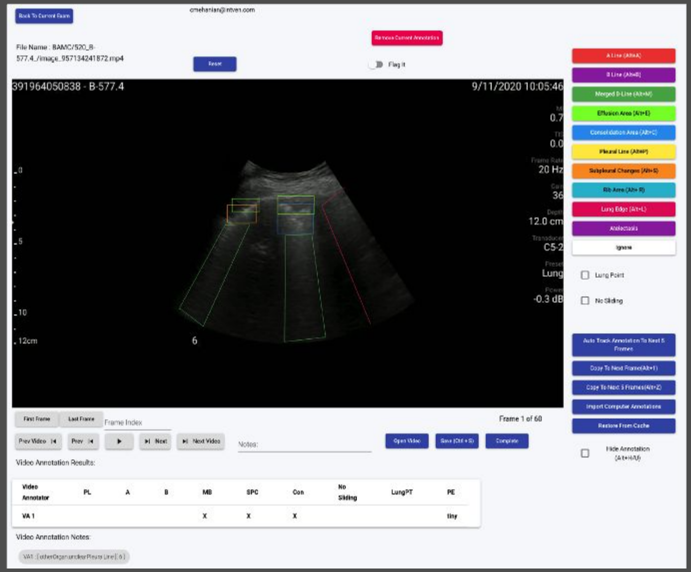
*Volpicelli, et al., International evidence-based recommendations for point-of-care lung ultrasound, Intensive Care Medicine (2012), 38:577-591

Curating high-quality annotated image library

Image Acquisition ⟷ **Image Annotation** ⟷ Algorithm Development

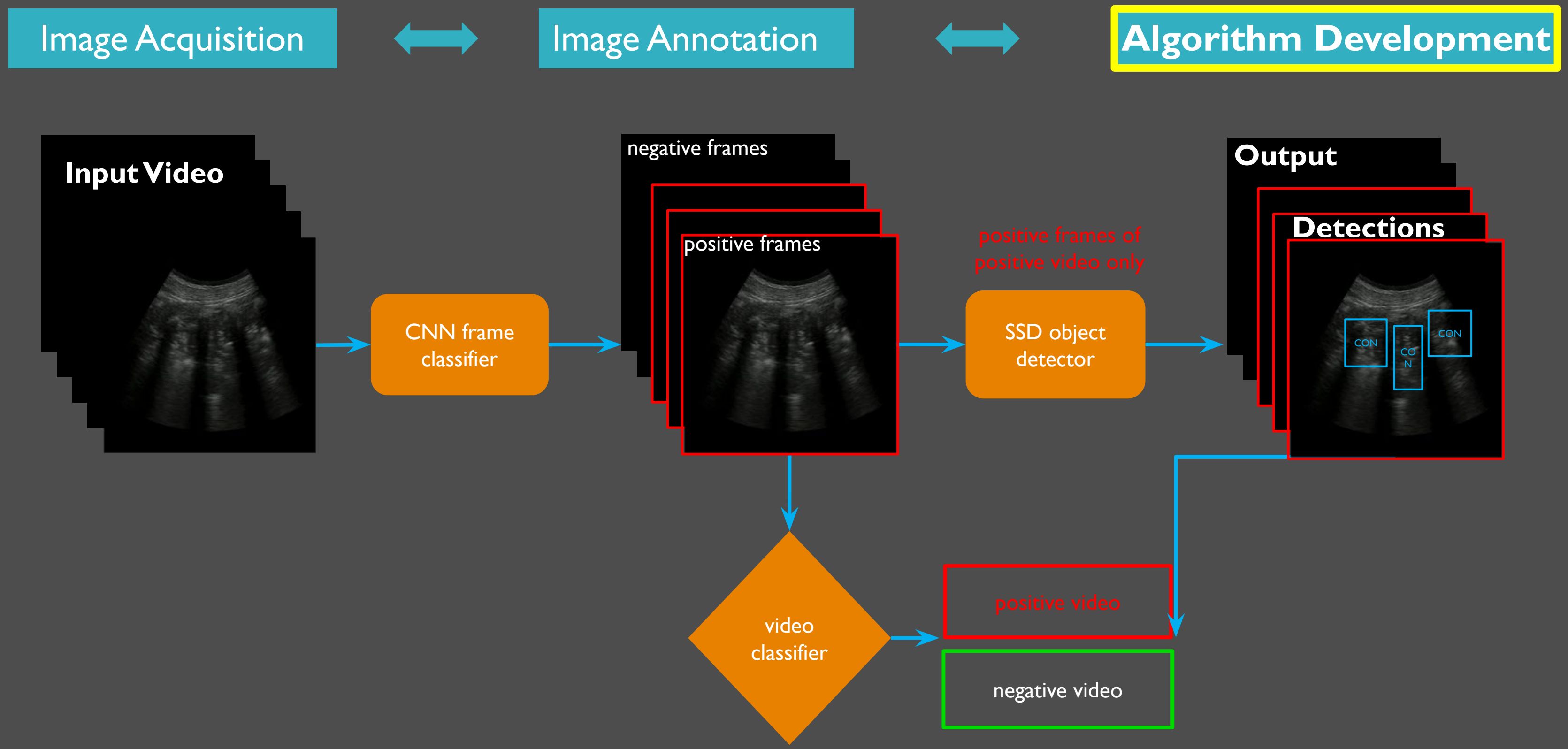
Video exclusion criteria:
- Incorrect acquisition parameters
- Excessive transducer movement
- Incorrect transducer positioning
- Poor image quality





Careful frame-by-frame annotation by expert radiologist

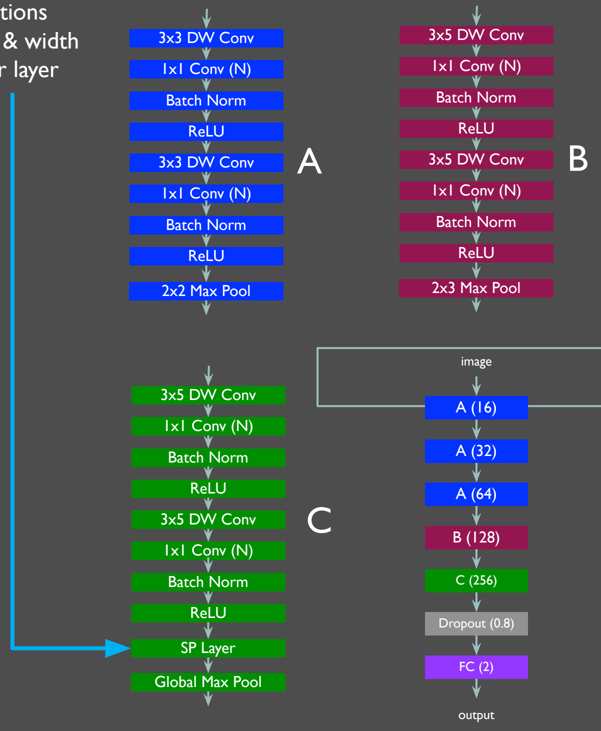
Cascade architecture for classification & localization



Classification and detection models

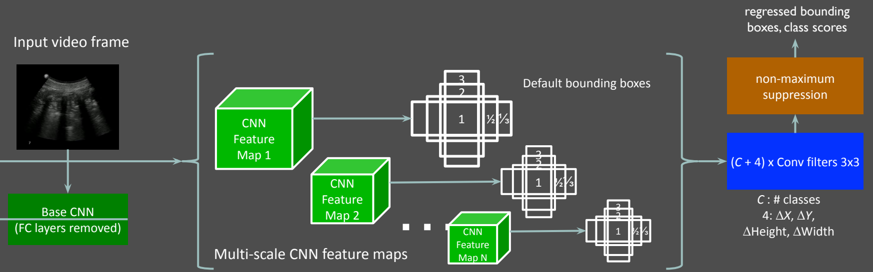
Adult consolidation classification

- VGG-like CNN architecture
- Simplify convolutions
- Optimize depth & width
- Add spatial prior layer



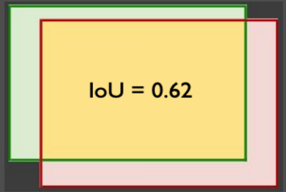
Detection models (all lung features)

- SSD architecture is accurate
- Features at multiple scales, aspect ratios
- Efficient, suitable for on-device inference



Summary of results

Specifications are close to being met on holdout sets



ADULT CONSOLIDATION

	patients	exams	neg. videos	pos. videos
TRAIN	46	102	984	358
VAL	15	21	55	66
HOLDOUT	26	31	132	132

	sensitivity	specificity	avg. IoU
HOLDOUT	0.88	0.90	0.60

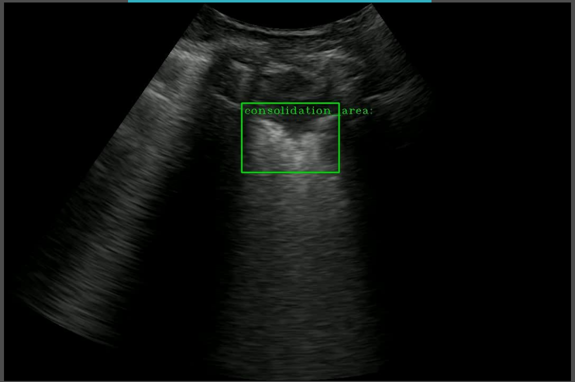
PEDIATRIC CONSOLIDATION

	patients	exams	neg. videos	pos. videos
TRAIN	55	61	466	402
VAL	26	31	178	136
HOLDOUT	27	27	127	133

	sensitivity	specificity	avg. IoU
HOLDOUT	0.88	0.89	0.62

Healthy lung

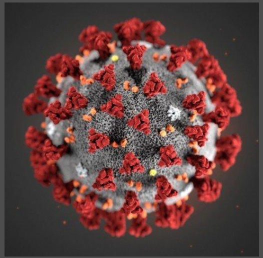
Consolidation



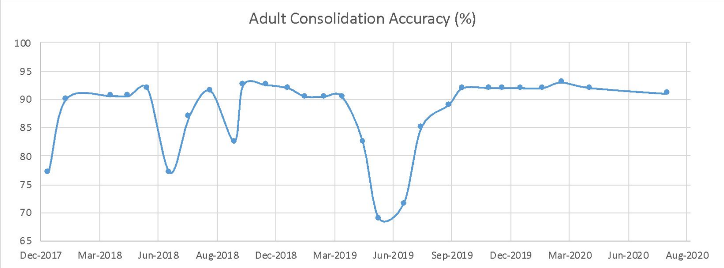
All models execute in < 30 sec/video

Articles
AAAS 2018
ASA 2018
MHSRS 2018 * 2
MICCAI 2018
MHSRS 2019 * 2
MICCAI 2019
MHSRS 2020 * 2
Critical Care 2019

Summary and Conclusions



- Developed algorithm for AI-assisted detection of pneumonia with lung ultrasound
  - Worked with a global leader in ultrasound technology
  - Algorithms may be directly applicable to smoke inhalation, COVID-19 and other ultrasound applications are in the pipeline: compartment syndrome, FAST, optic nerve-sheath diameter
- Learned
  - Importance of large, high-quality, diverse data set
  - Importance of high-quality annotations suitable for machine learning
  - Importance of collaboration between doctors and ML scientists
- Learn-as-you-go



UNIVERSITY OF OREGON

Phil and Penny Knight
Campus for Accelerating Scientific Impact

Acknowledgements

Funded by Bill and Melinda Gates Foundation Trust and DARPA

Sourabh Kulhare	Michele Griffin	Kenton Gregory	Todd Graham	Matthew Hepburn
Rachel Millin	Bruce Goldstein	Cynthia Gregory	Katelyn Hostetler	Matthew Kappes
Dipayan Banik	Davina Inslee	Meihua Zhu	Annie Cao	Paul Sheehan
Liming Hu	Doug Bradley	Hua Xie	Michael Stone	Andrea Cantu
Charles Delahunt	Sheana Creighton	Fen Wang	Andrew Hersh	Jeffrey Shupp
Zohreh Laverriere	Debra Peat	Xijun Zhang	David Kessler	Laura Johnson
Matthew Horning	Megan Bettilyon	James Jones	Cristian Madar	Peter Weimersheimer
Benjamin Wilson	Shannon Kuyper	Jack Lazar	Damon Forbes	Kristin Dwyer
Xinliang Zheng	Renée Rivers	Amber Halse	Brian Shahan	Jacob Avila

Summary of lessons learned

Worked well …	**Lung ultrasound**
Curation of high-quality, diverse, multisource data	✓
Feedback mechanism for establishing ground truth	✓
Collaboration between ML people and SMEs	OHSU Clinicians
Focus on product, patient, and use-case	Video-level localization and classification
Partnership with commercial entity	Ultrasound Co.
Early engagement with regulatory agency	U.S. FDA
Keep an eye on …	
Explore vs. exploit: Bayesian approach	Strike the right balance, hedge your bets
Bottom-up vs. top-down: Collective responsibility	Joint ownership of goals and timelines

Questions ?

Next lecture:

CNN in action

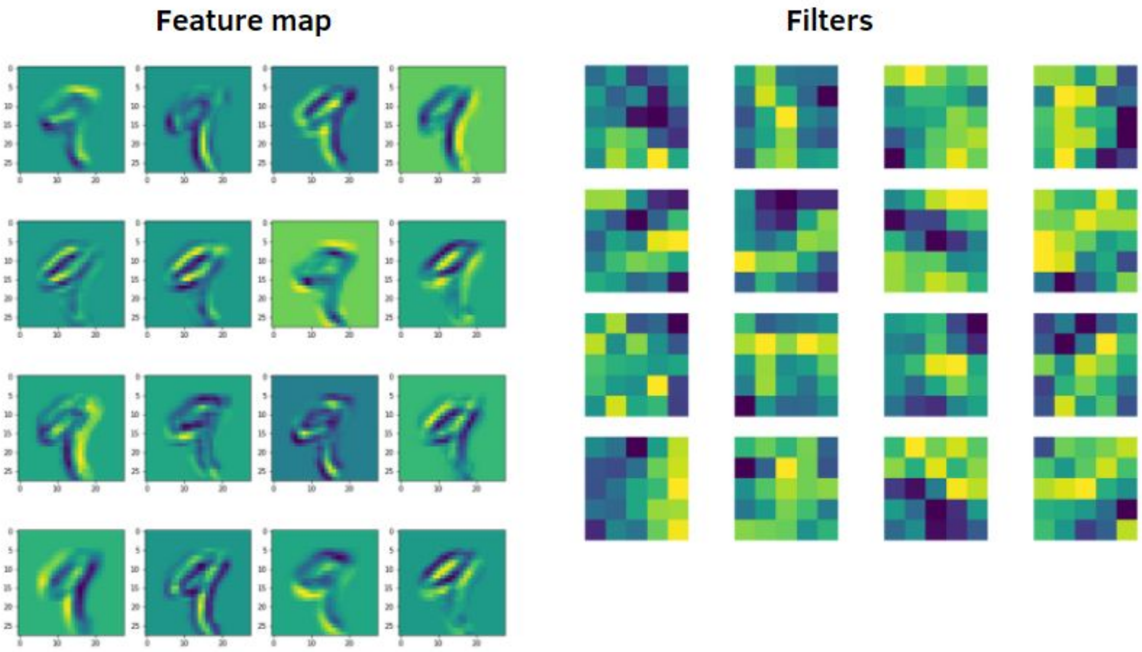


Image from Eugenia Anello, medium: visualizing feature maps and filters in CNNs