Object Recognition – Part 1

Danna Gurari
University of Colorado Boulder
Fall 2023

Review

- Last lecture:
  - Ways of seeing: image and video acquisition
  - Evolution of computer vision (before versus after 2012)
  - Background of machine learning and neural networks
  - Training deep neural networks

- Assignments (Canvas)
  - Reading assignment due earlier today
  - Next two reading assignments due next Monday and Wednesday

- Questions?

Object Recognition: Today’s Topics

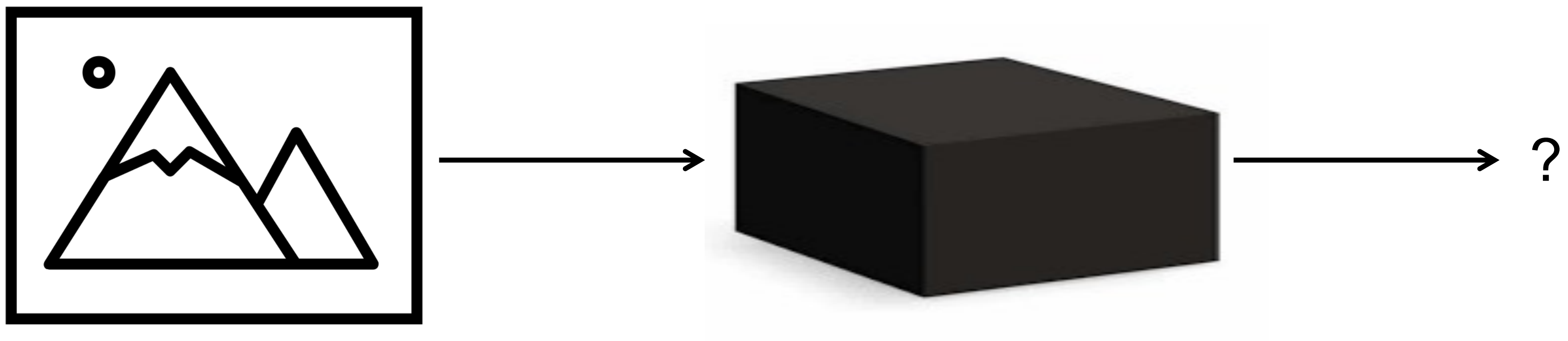
- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Networks

Object Recognition: Today's Topics

- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Networks

Object Recognition

- Given an image, indicate what object(s) are in the image



Object Recognition

e.g.,

INPUT OUTPUT

→

→ ?

Object Recognition

e.g.,

INPUT

OUTPUT

Sunflower ✓

Object Recognition

e.g.,

INPUT

OUTPUT

?

Object Recognition

e.g.,

INPUT

OUTPUT

Weasal

Object Recognition: Today’s Topics

- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Networks

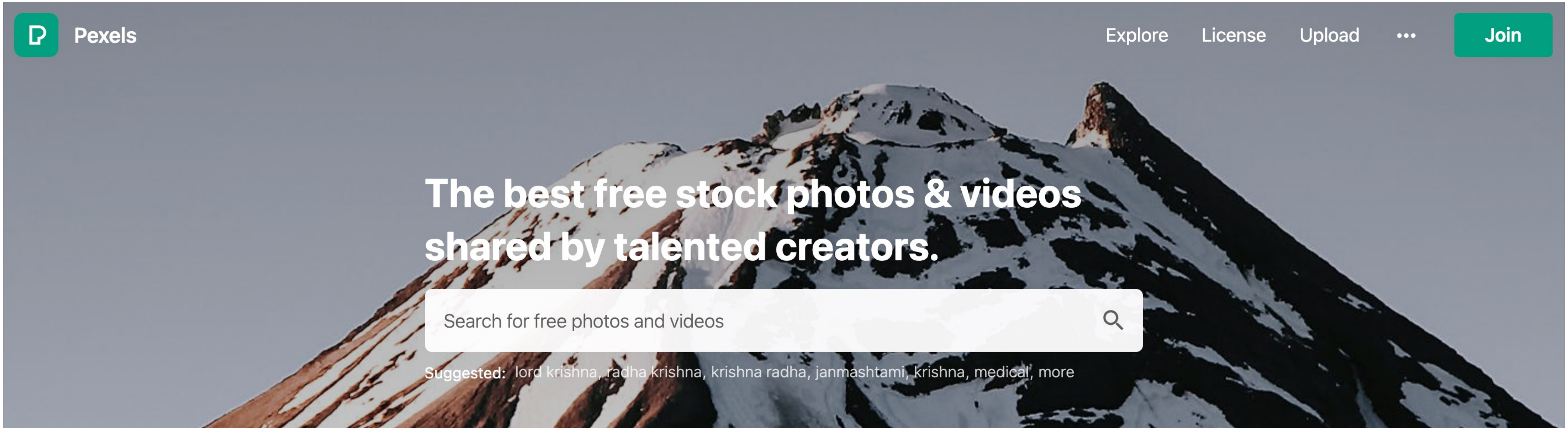
Shopping

Take a picture of an object and find where to buy it

Photo Organization

Demo: https://www.youtube.com/watch?v=aBqmWUalnho
(start video at 1:46)

Image Search



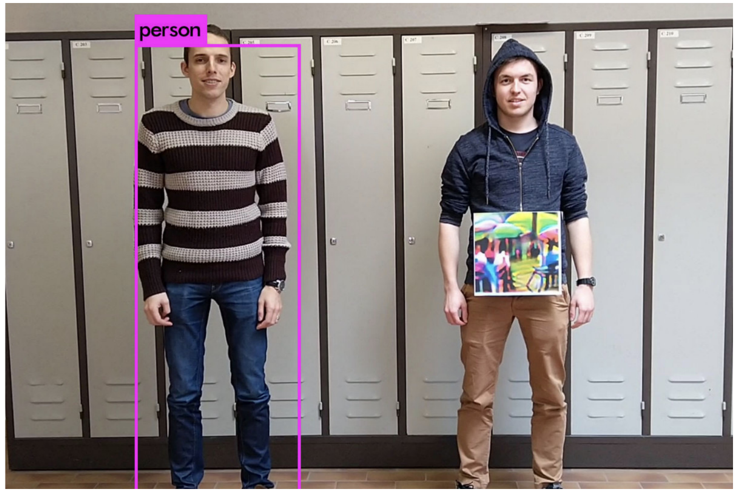
Applications Gone Wrong

- Ethical mistake: people tagged as “gorillas”

http://www.usatoday.com/story/tech/2015/07/01/google-apologizes-after-photos-identify-black-people-as-gorillas/29567465/

- Security risk: people mis-recognized or invisible when wearing special designs

https://www.cs.cmu.edu/~sbhagava/papers/face-rec-ccs16.pdf



https://www.theverge.com/2019/4/23/18512472/fool-ai-surveillance-adversarial-example-yolov2-person-detection

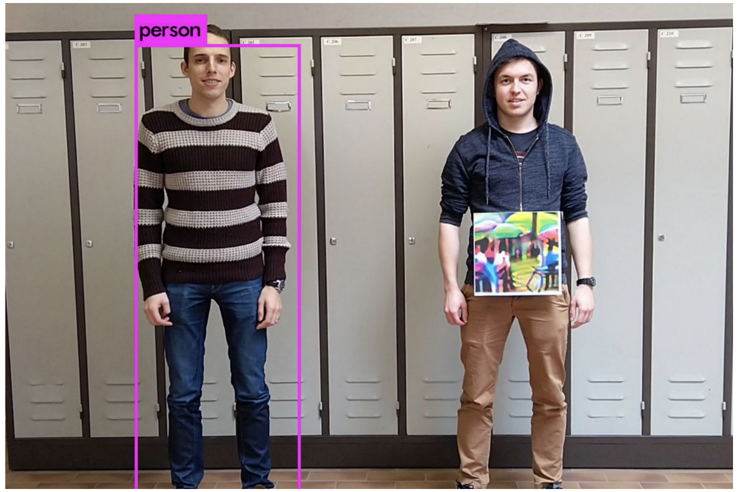
Applications Gone Wrong

1) Why might these mistakes occur?

2) If you were the CEO providing these products, how would you respond to these issues?

http://www.usatoday.com/story/tech/2015/07/01/google-apologizes-after-photos-identify-black-people-as-gorillas/29567465/

https://www.cs.cmu.edu/~sbhagava/papers/face-rec-ccs16.pdf

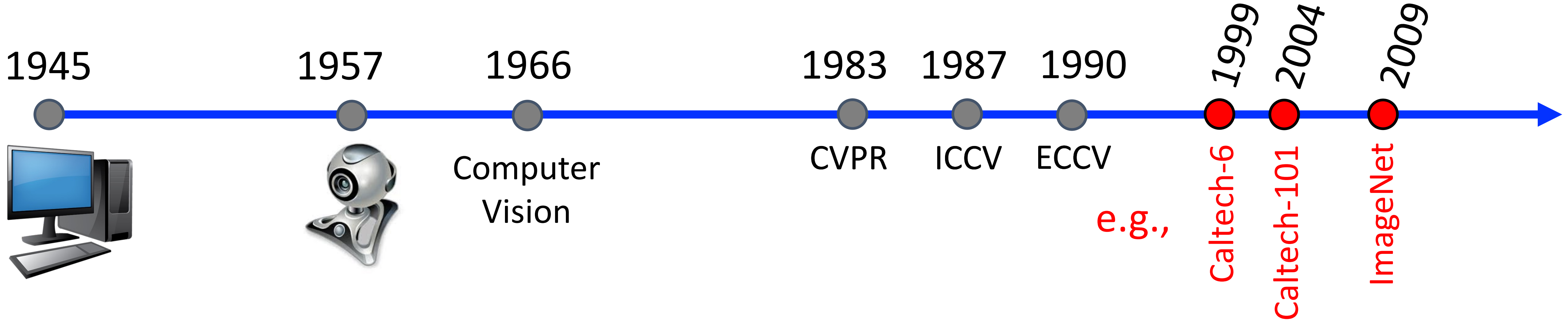


https://www.theverge.com/2019/4/23/18512472/fool-ai-surveillance-adversarial-example-yolov2-person-detection

Object Recognition: Today’s Topics

- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Networks

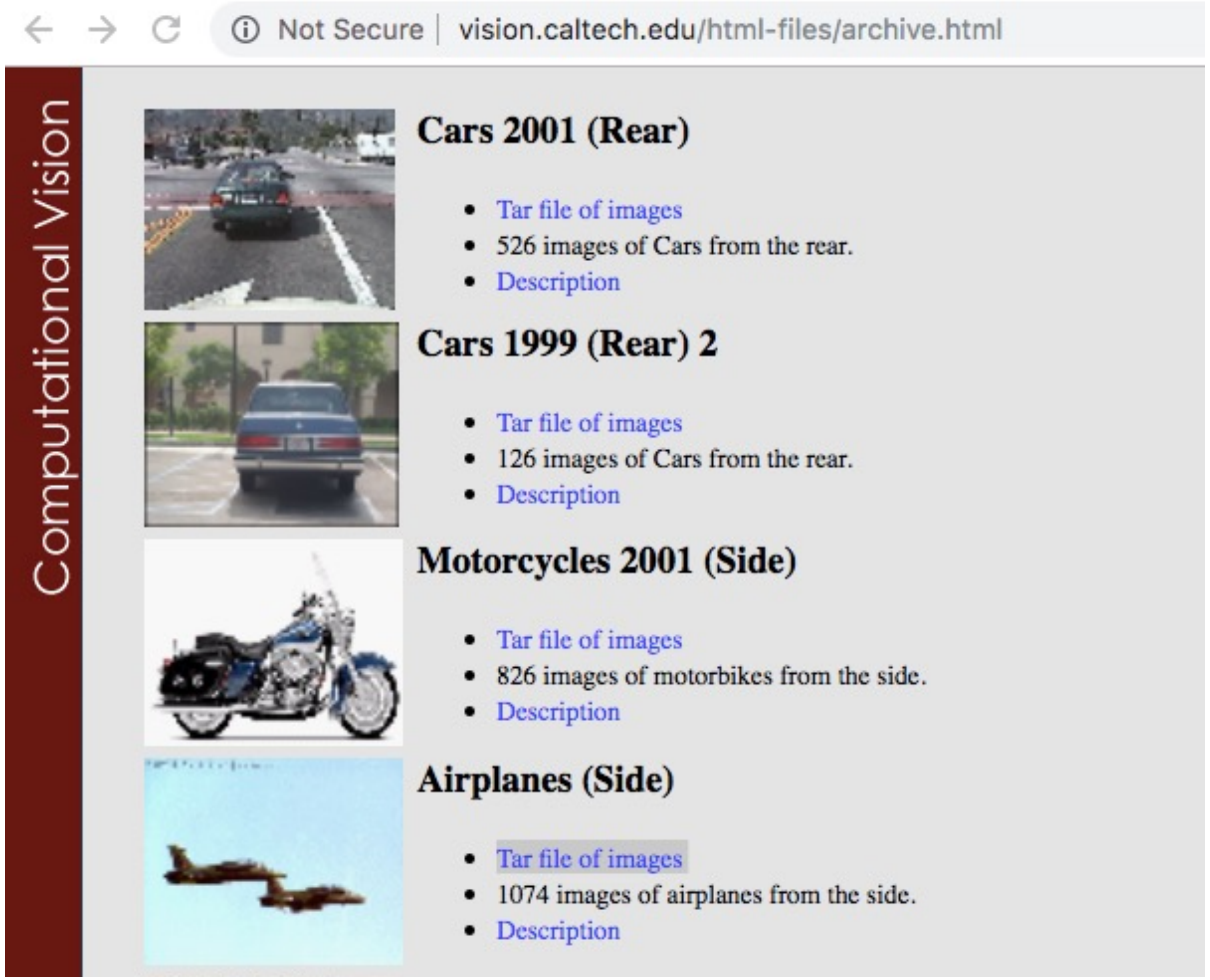
Public Datasets Since Early ~2000s



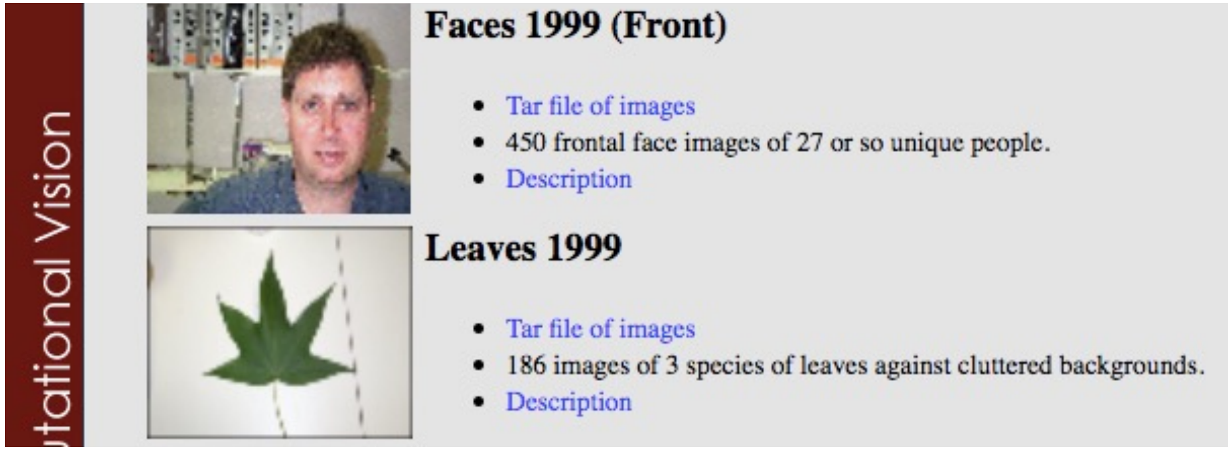
	Caltech-6	Caltech-101	ImageNet
# Categories:	6	101	1000
# Images:	3,738	9,144	1,461,406

Trend: build bigger datasets

Caltech-6



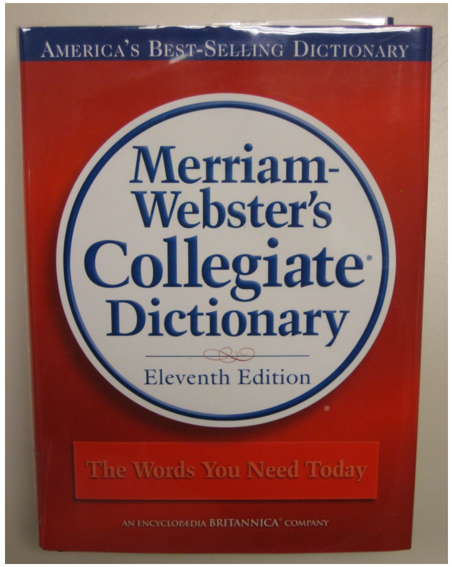
(1) Six categories selected and then (2) students took pictures or collected images from the web



http://www.vision.caltech.edu/html-files/archive.html

Caltech-101

1. Category Selection

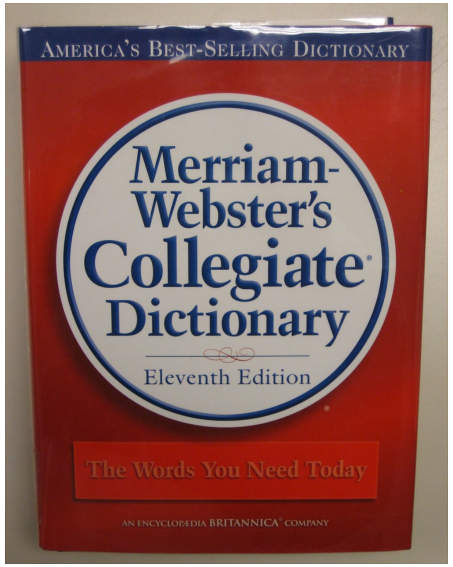


Flipped through a dictionary and chose 101 categories associated with a drawing

Li Fei Fei, Rob Fergus, & Pietro Perona. CVPR 2004.

Caltech-101

1. Category Selection



Flipped through a dictionary and chose 101 categories associated with a drawing

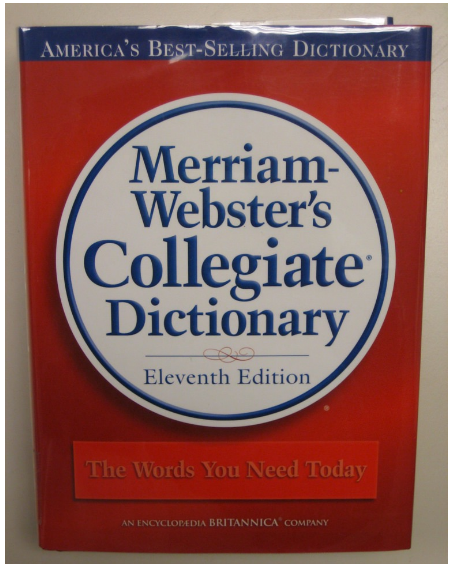
2. Image Collection

Search for each category

Li Fei Fei, Rob Fergus, & Pietro Perona. CVPR 2004.

Caltech-101

1. Category Selection



Flipped through a dictionary and chose 101 categories associated with a drawing

2. Image Collection

Search for each category

3. Human Verification

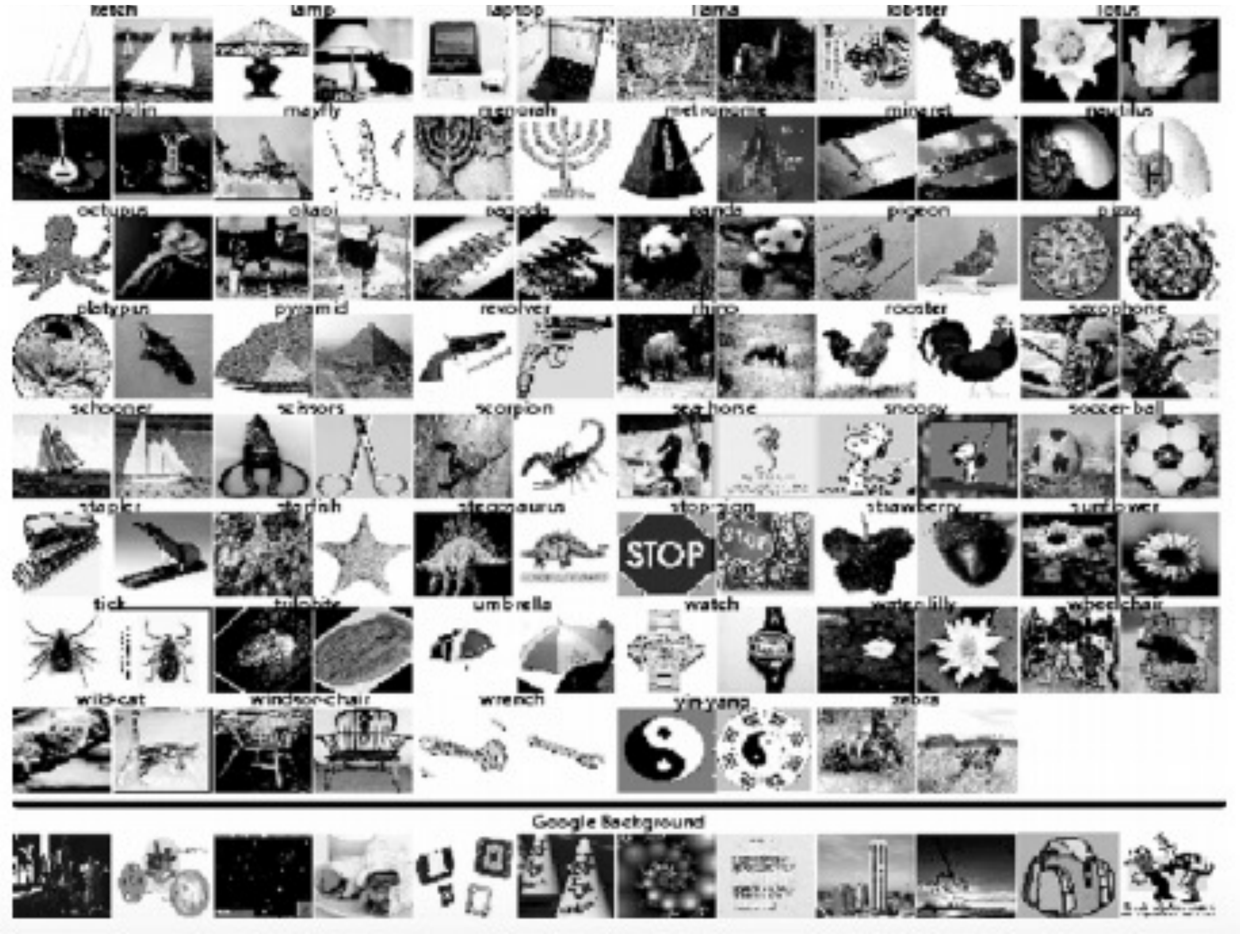
- 2 graduate students reviewed & discarded irrelevant images

- Result is 9,144 grayscale 300x200 pixel images with 45-400 images per category

Li Fei Fei, Rob Fergus, & Pietro Perona. CVPR 2004.

Caltech-101

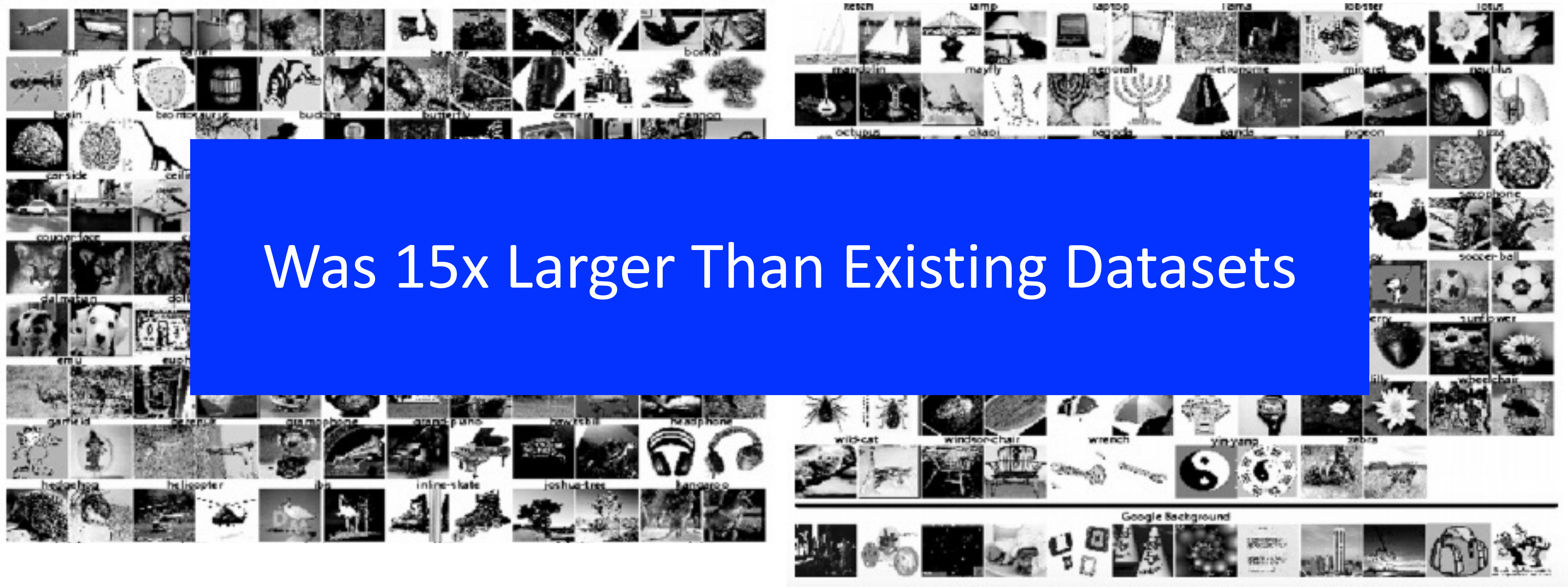
Two random samples per category



Dataset location: http://vision.caltech.edu

Caltech-101

Two random samples per category

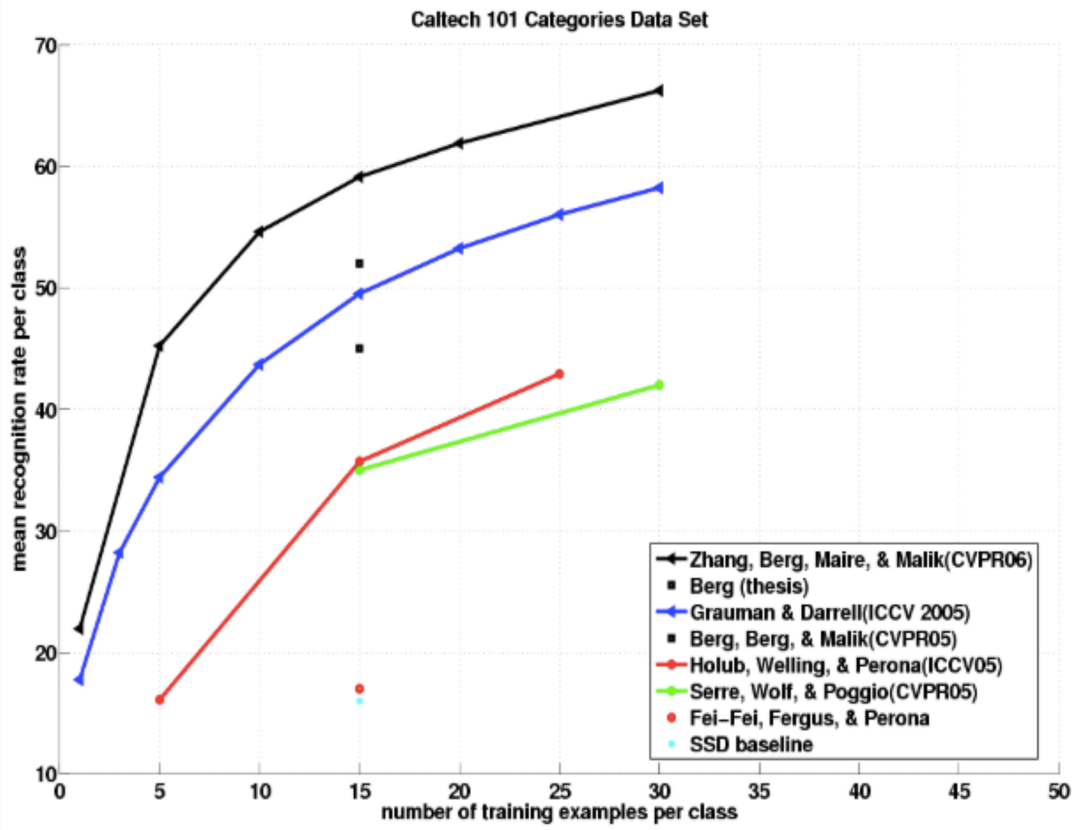


Was 15x Larger Than Existing Datasets

Dataset location: http://vision.caltech.edu

Caltech-101

Progress of algorithms charted



Latest results (March 2006) on the Caltech 101 from a variety of groups. (published results only).

If you would like to include your algorithm's performance please email us at holub@caltech.edu or greg@vision.caltech.edu with a citation and your results. Thanks!

We are also interested in the time it takes to run your algorithm. Both during the training and during the classification stage

Plot courtesy of Hao Zhang.

Update by holub, April 2006.

Li Fei Fei, Rob Fergus, & Pietro Perona. CVPR 2004.

ImageNet

After creating Caltech-101 and finishing her PhD, Fei-Fei Li began her career as an assistant professor creating ImageNet.

Hear her tell her story:

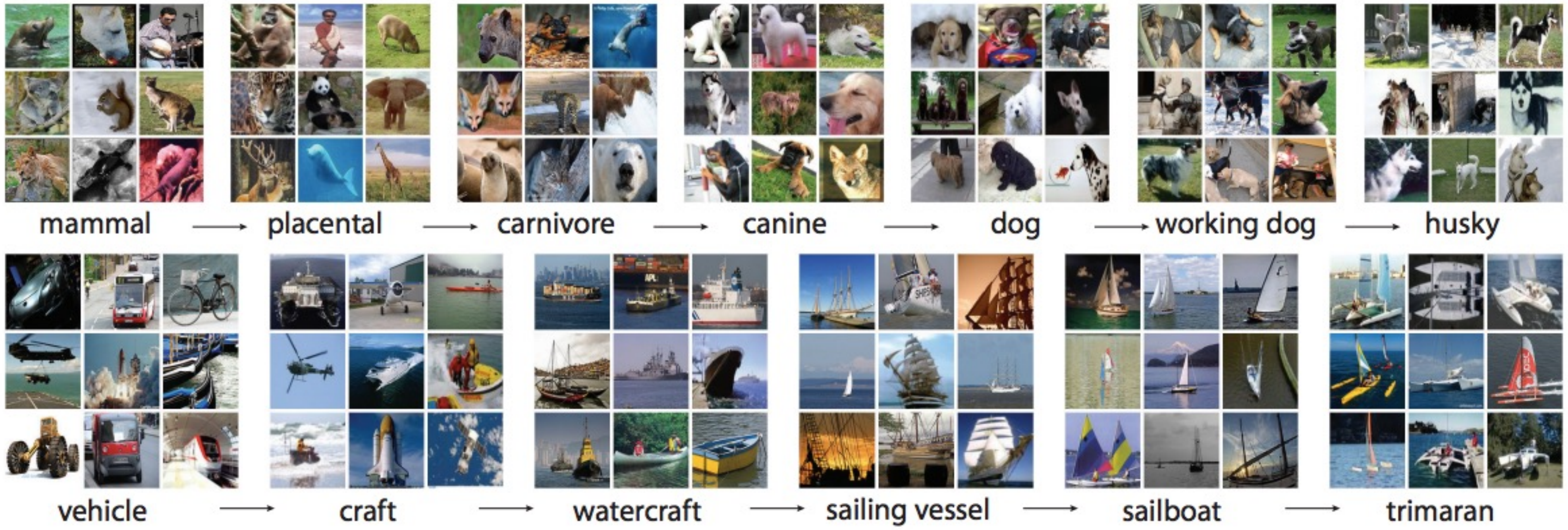
https://www.youtube.com/watch?v=40riCqvRoMs
(5:44 – 9:35)

ImageNet

1. Category Selection

~10% of concepts (synonym sets) in WordNet taxonomy

e.g., two root-to-leaf branches of ImageNet with nine examples for each "synonym set"



Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, & Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. CVPR 2009.

ImageNet

1. Category Selection

~10% of concepts (synonym sets) in WordNet taxonomy

2. Image Collection

flickr

(& more search engines)

Query expansion:
- Augment queries
- Translate queries to different languages

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, & Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. CVPR 2009.

ImageNet

Key Insight: use crowdsourcing to recruit many people to verify images

1. Category Selection

~10% of concepts (synonym sets) in WordNet taxonomy

→

2. Image Collection

flickr

(& more search engines)

Query expansion:
- Augment queries
- Translate queries to different languages

→

3. Human Verification

- Humans verify if image contains queried object

- Use majority vote decision from multiple humans to support high quality results

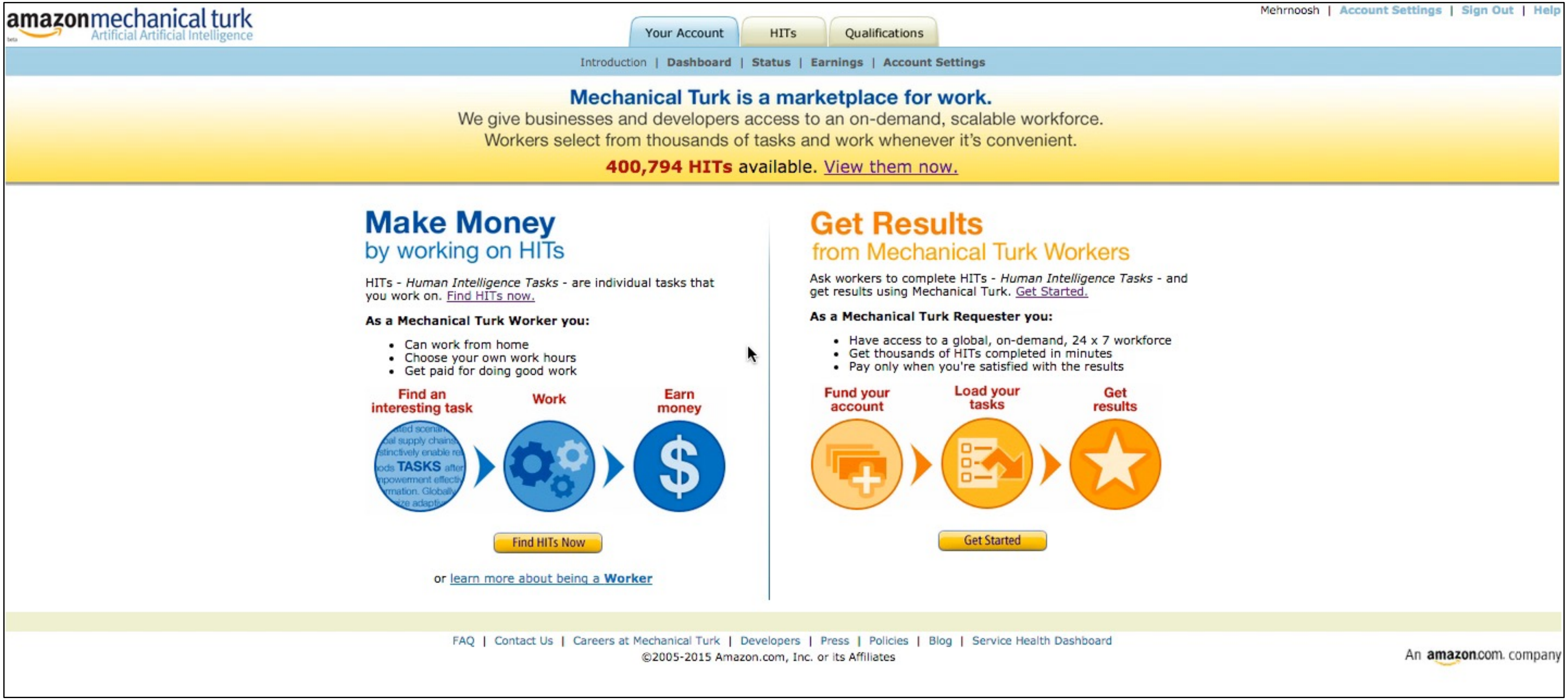
Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, & Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. CVPR 2009.

ImageNet: Crowdsourcing Task

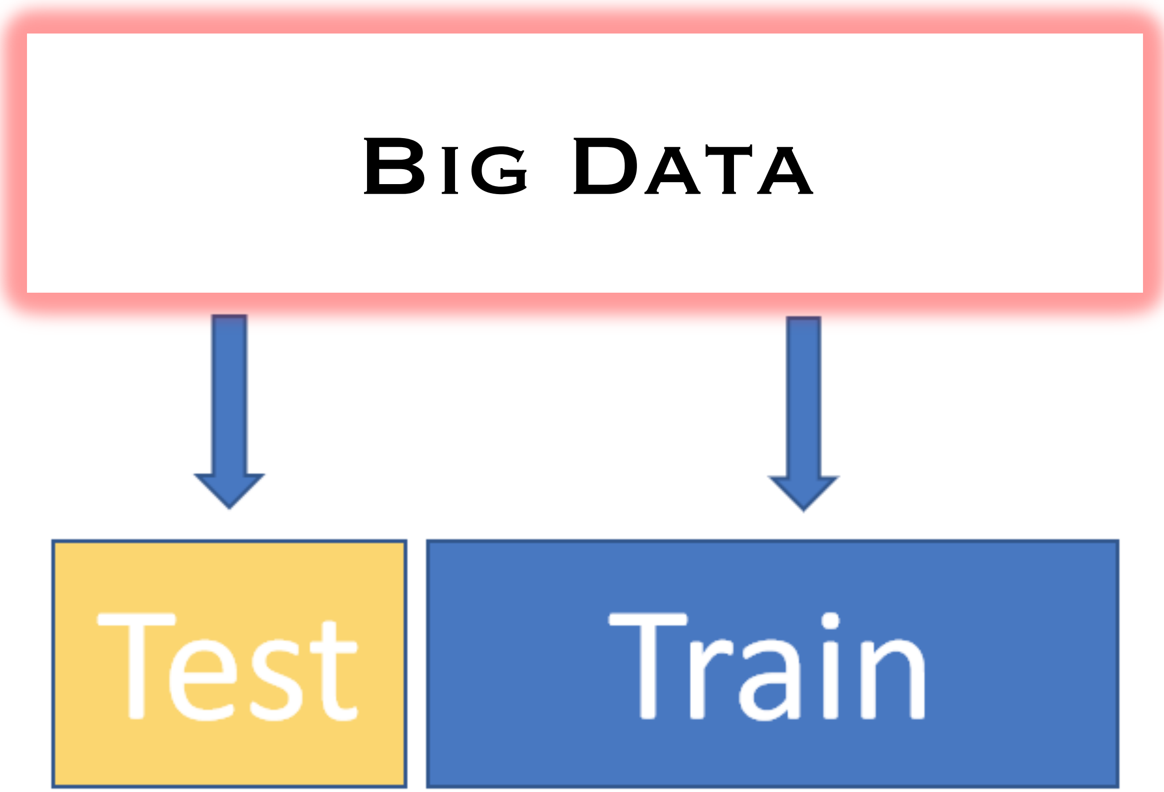
Definition of the target synonym set with link to Wikipedia.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, & Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. CVPR 2009.

ImageNet: Crowdsourcing Platform



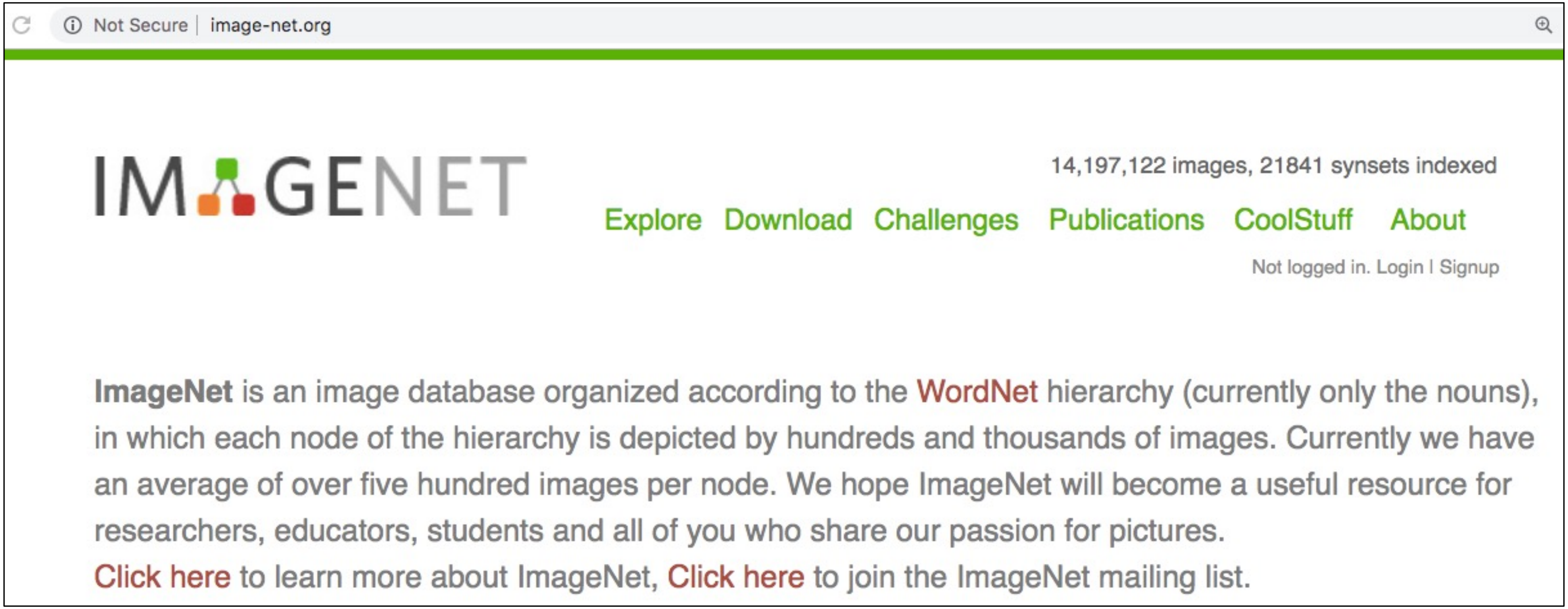
ImageNet: Challenge



(Human labels hidden) **(Human labels public)**

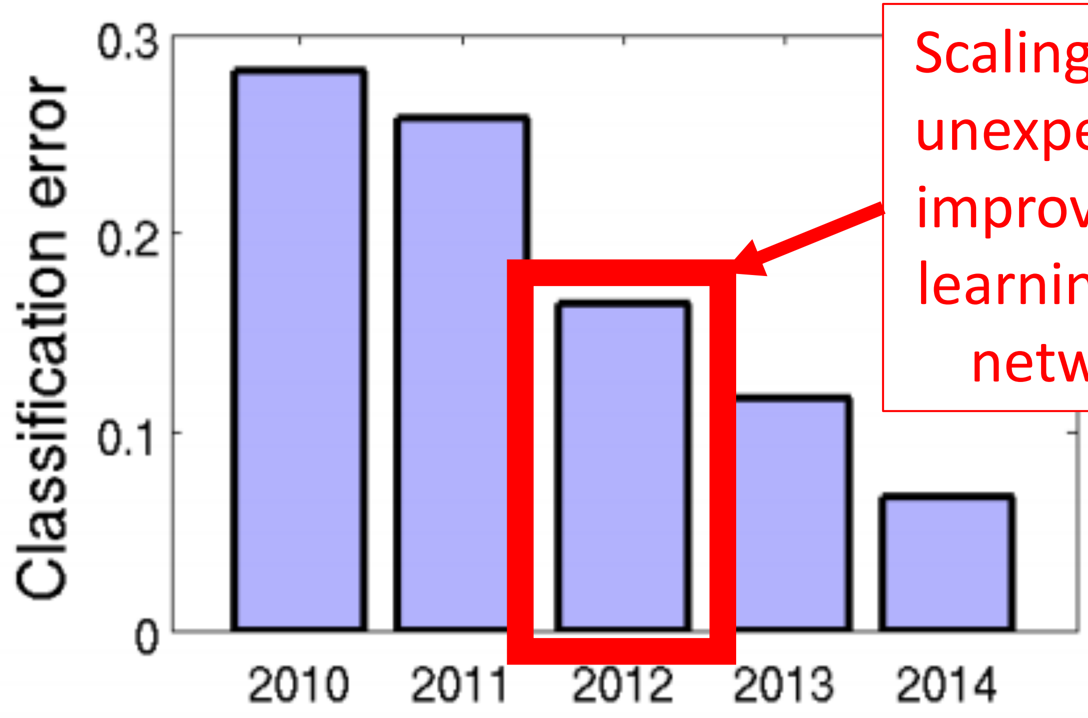
Winner: highest scoring method on the hidden test set

ImageNet: Website with Evaluation Server



ImageNet: Catalyst for Revolution

Progress of models on ImageNet

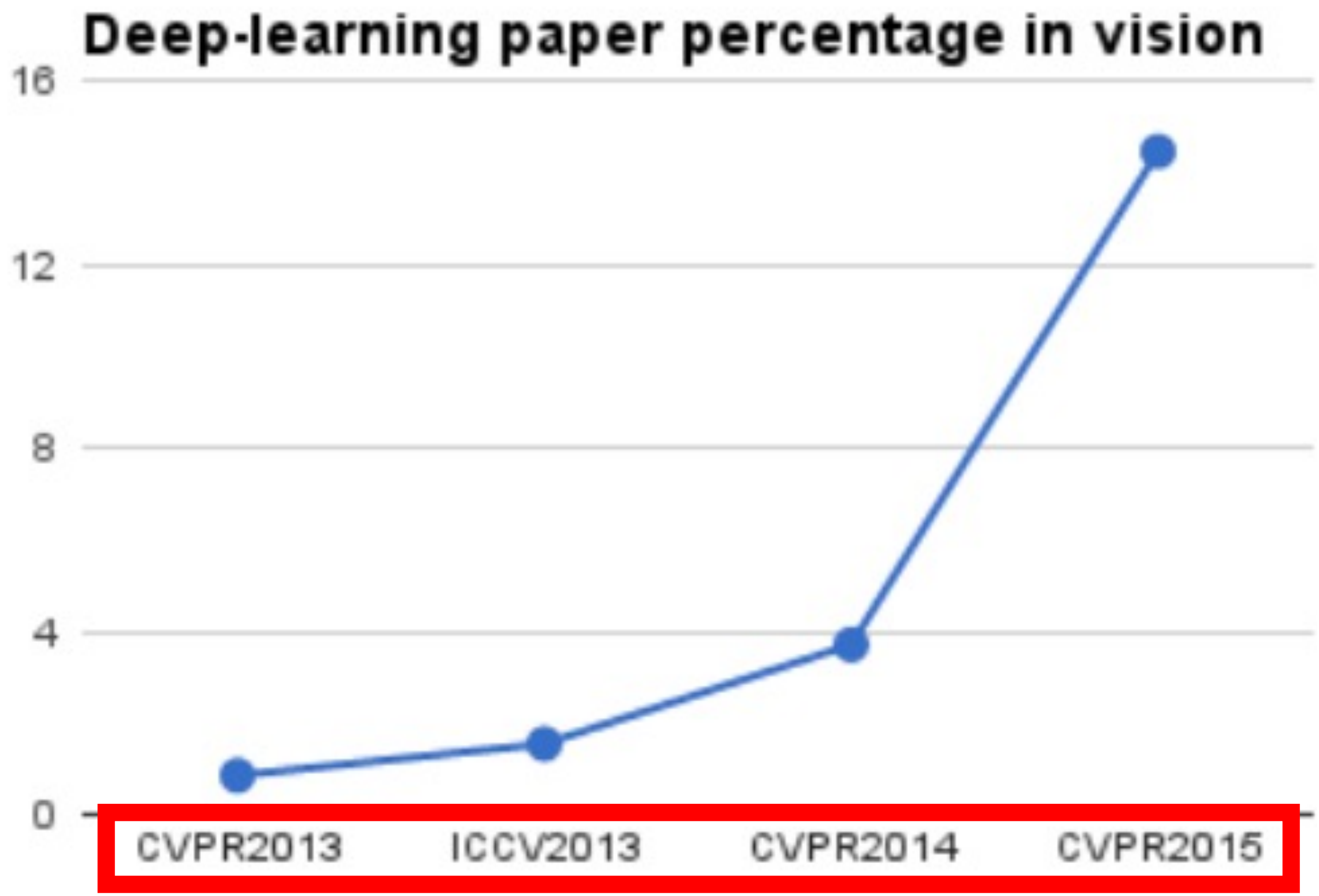


Scaling up the dataset led to unexpected, unprecedented improvements from a "deep learning" model (i.e., neural network) called AlexNet

(15.3% error, 10.8 percentage points less than that of the runner up)

Olga Russakovsky et al. ImageNet Large Scale Visual Recognition Challenge. IJCV 2015.
Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet Classification with Deep Neural Networks. NIPS 2012.

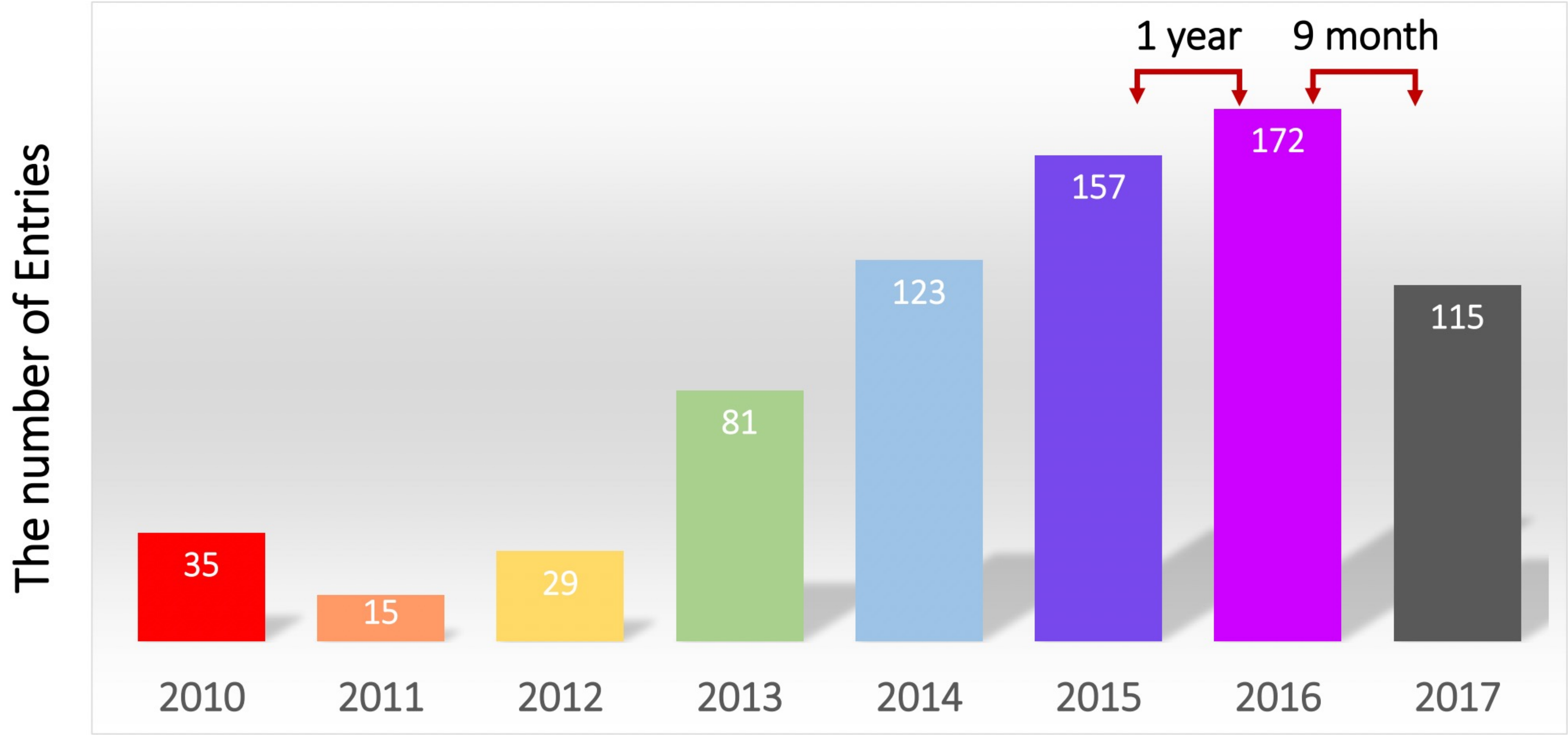
ImageNet: Catalyst for Revolution



Inspired, many more researchers in the computer vision community focused on neural networks and discovered they succeed for many more vision problems!

Figure : https://www.slideshare.net/xavigiro/saliency-prediction-using-deep-learning-techniques

ImageNet: Challenge Engagement



- 727 entries (plus an entry from Baidu that famously was kicked out in 2015 for cheating)
- Labor cost ~$110 million: assuming 3 people contribute to each entry and $50k cost per person

Source: https://image-net.org/static_files/files/ILSVRC2017_overview.pdf

ImageNet Impact Recognized

1945 1957 1966 1983 1987 1990 2009 2019

Computer Vision

CVPR ICCV ECCV ImageNet Award

PAMI Longuet-Higgins Prize

Retrospective Most Impactful Paper from CVPR 2009

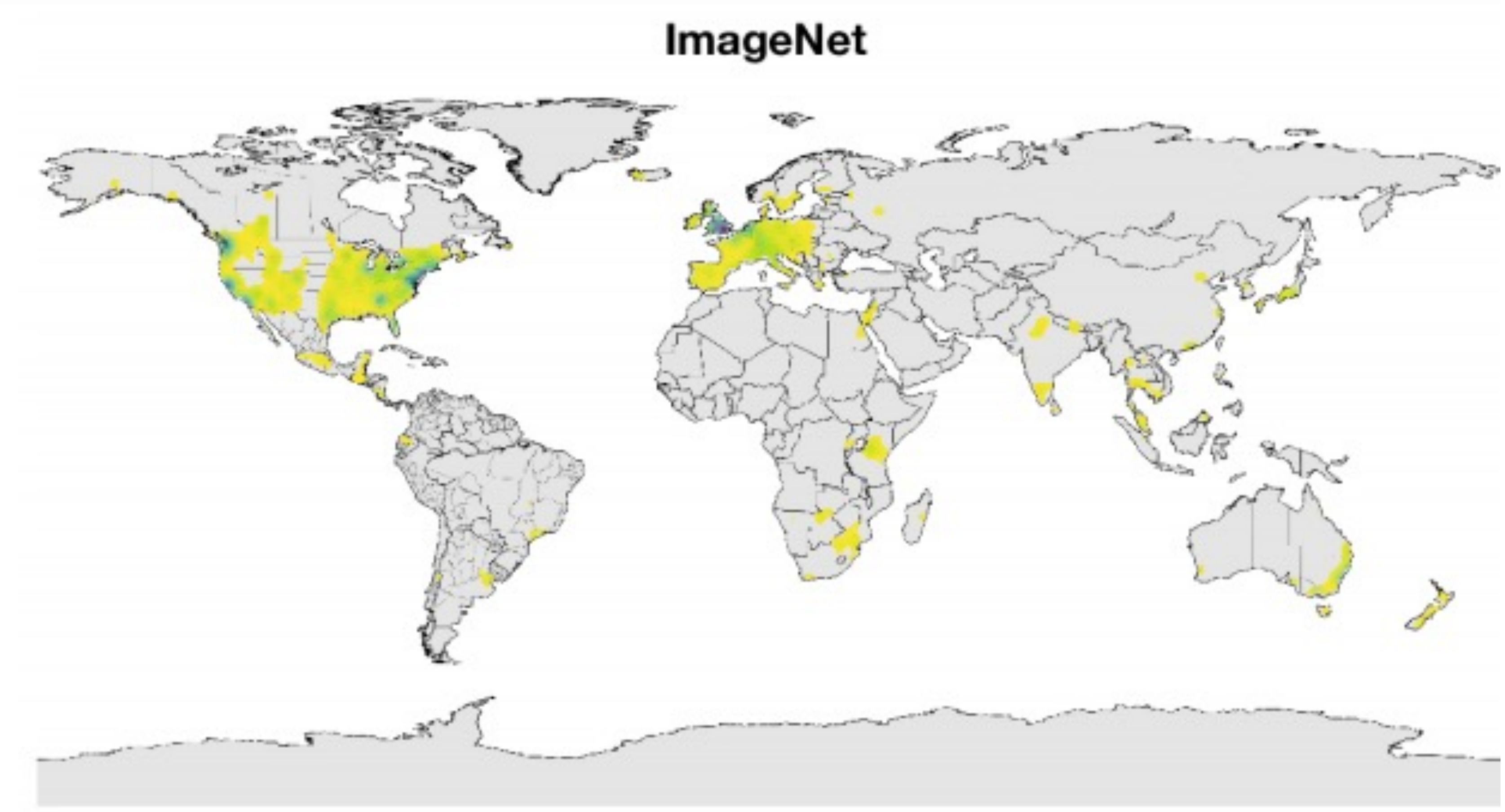
ImageNet: A large-scale hierarchical image database

Jia Deng, Wei Dong, Richard Socher,
Li-Jia Li, Kai Li, and Li Fei-Fei

https://syncedreview.com/2019/06/18/cvpr-2019-attracts-9k-attendees-best-papers-announced-imagenet-honoured-10-years-later/

ImageNet: Great Start…

Geographical origins of images in the ImageNet using Flickr metadata:



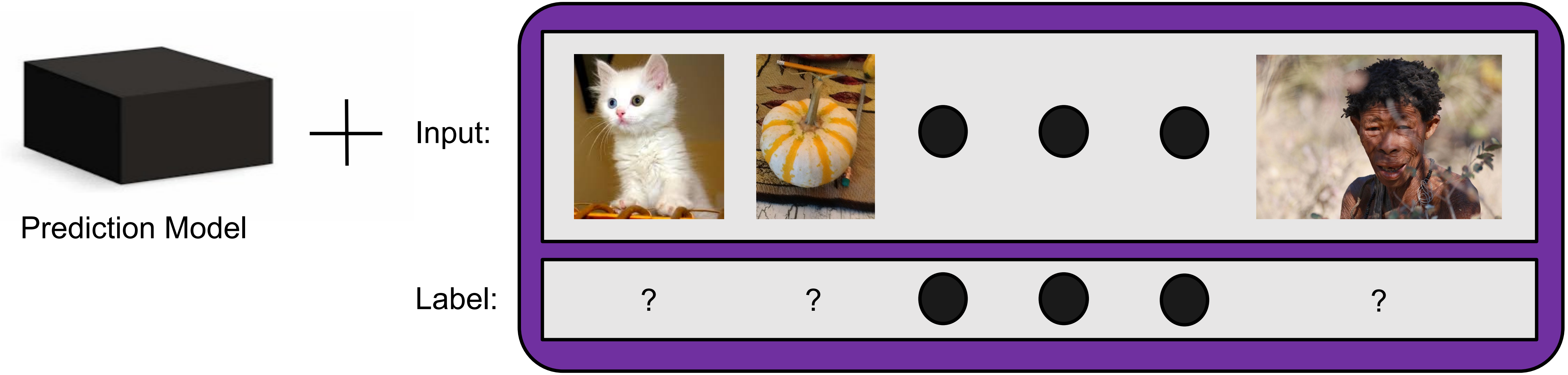
Terrance DeVries et al. Does Object Recognition Work for Everyone. CVPRW 2019.

Object Recognition: Today's Topics

- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Networks

Goal: Design Models that **Generalize** Well to New, Previously Unseen Examples

Apply model on "test set" to measure generalization error



Evaluation Metric for ImageNet Challenge

Assumption: 1 ground truth label per image

Top *N* error: average over all test images using this rule per image:

* 0 if any of *N* predictions match the ground truth
* 1 otherwise

e.g., top 5 error

Steel drum

Output:
Scale
T-shirt
<u>Steel drum</u>
Drumstick
Mud turtle

✔

Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

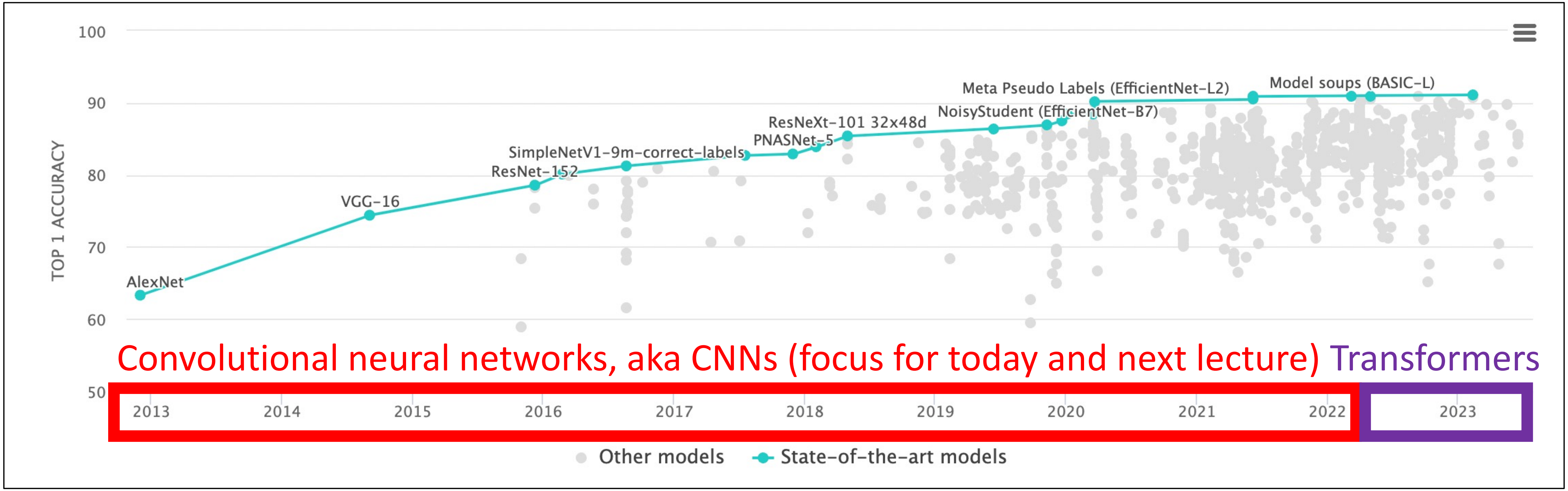
✗

Source: https://image-net.org/static_files/files/ILSVRC2017_overview.pdf

Object Recognition: Today's Topics

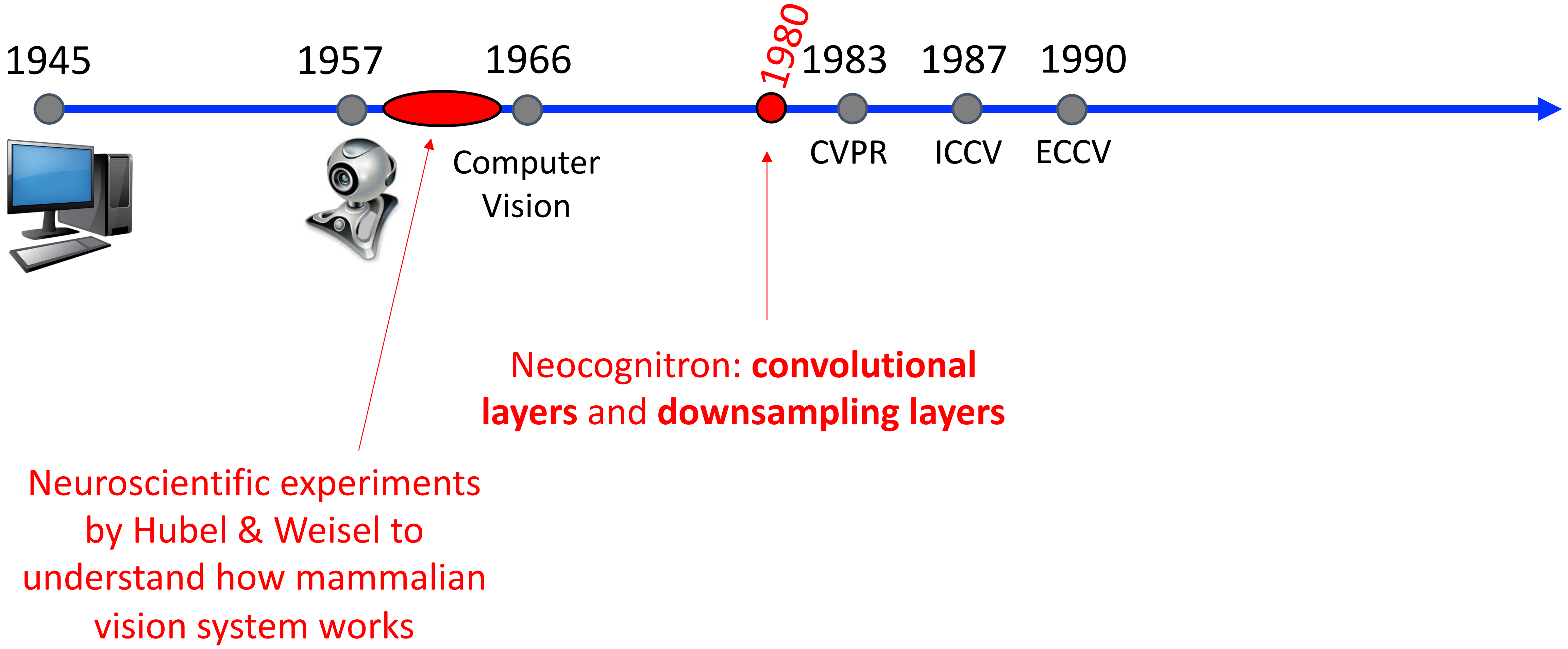
- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Network

ImageNet Challenge Winners Over Time



https://paperswithcode.com/sota/image-classification-on-imagenet?p=deepvit-towards-deeper-vision-transformer

CNN Origins


1945
1957
1966
1980
1983
1987
1990
Computer Vision
CVPR
ICCV
ECCV
Neocognitron: **convolutional layers** and **downsampling layers**
Neuroscientific experiments by Hubel & Weisel to understand how mammalian vision system works

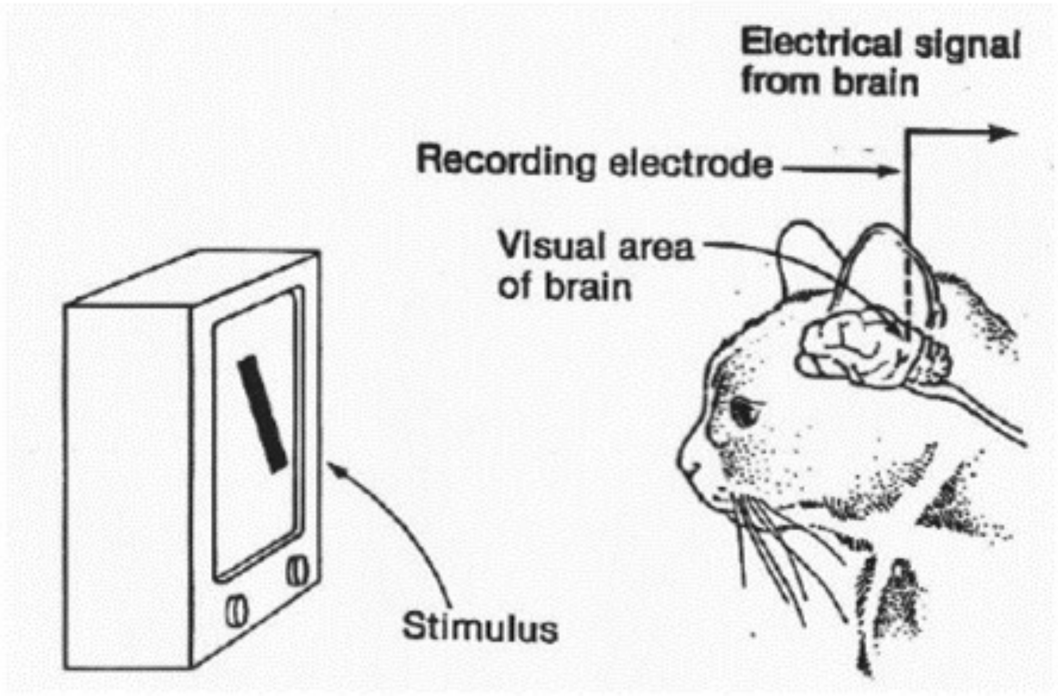
CNN Motivation: How Vision System Works

Nobel Prize winning insight from Hubel and Weisel!

Image Source: https://braintour.harvard.edu/archives/portfolio-items/hubel-and-wiesel

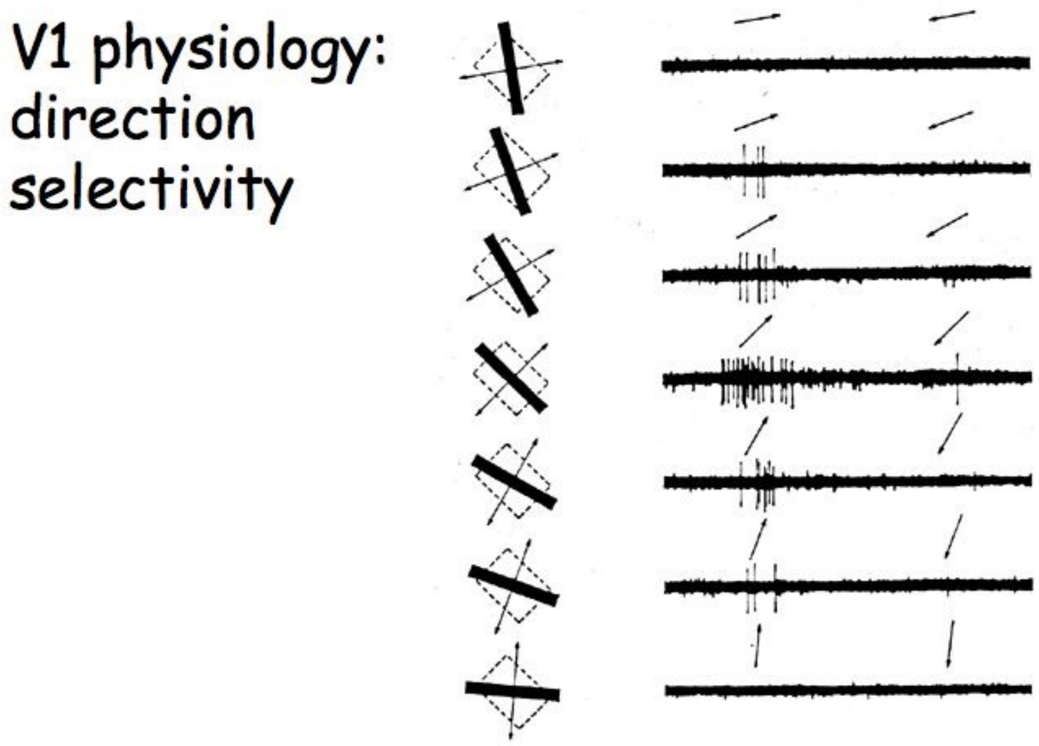
CNN Motivation: How Vision System Works

Experiment Set-up:



https://www.esantus.com/blog/2019/1/31/convolutional-neural-networks-a-quick-guide-for-newbies

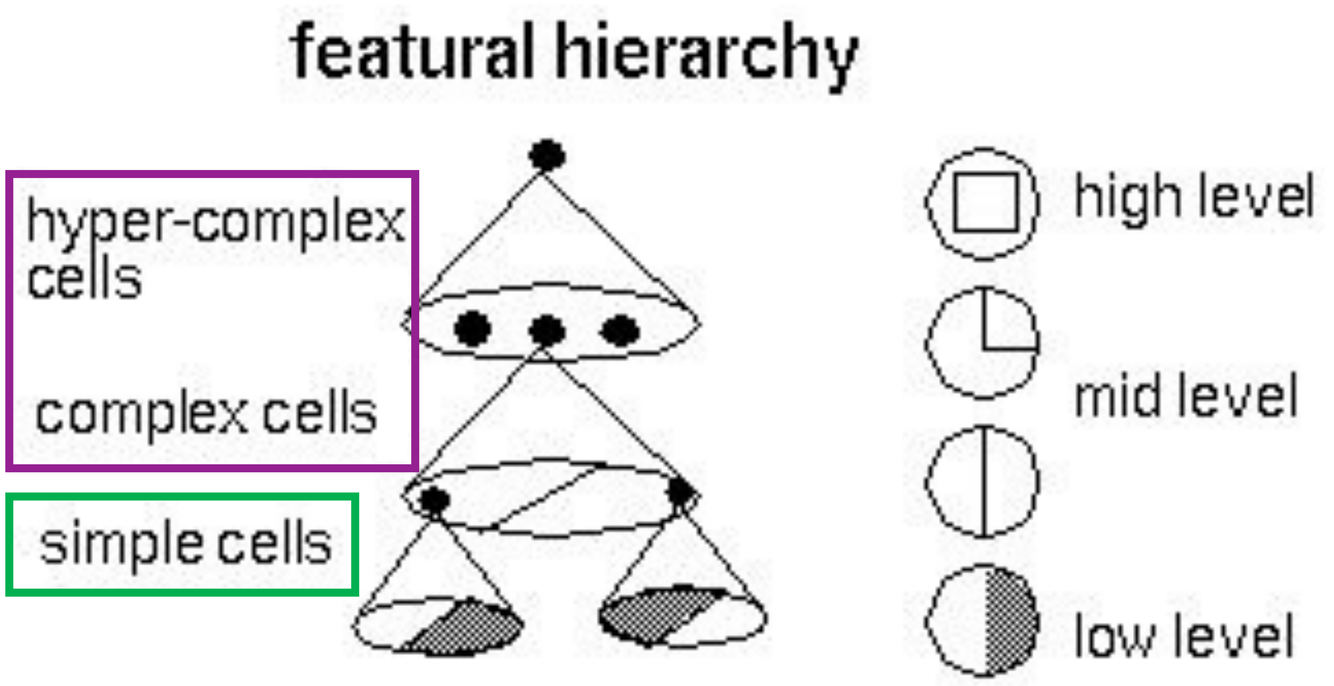
Key Finding: initial neurons responded strongly only when light is shown in certain orientations



https://www.cns.nyu.edu/~david/courses/perception/lecturenotes/V1/lgn-V1.html

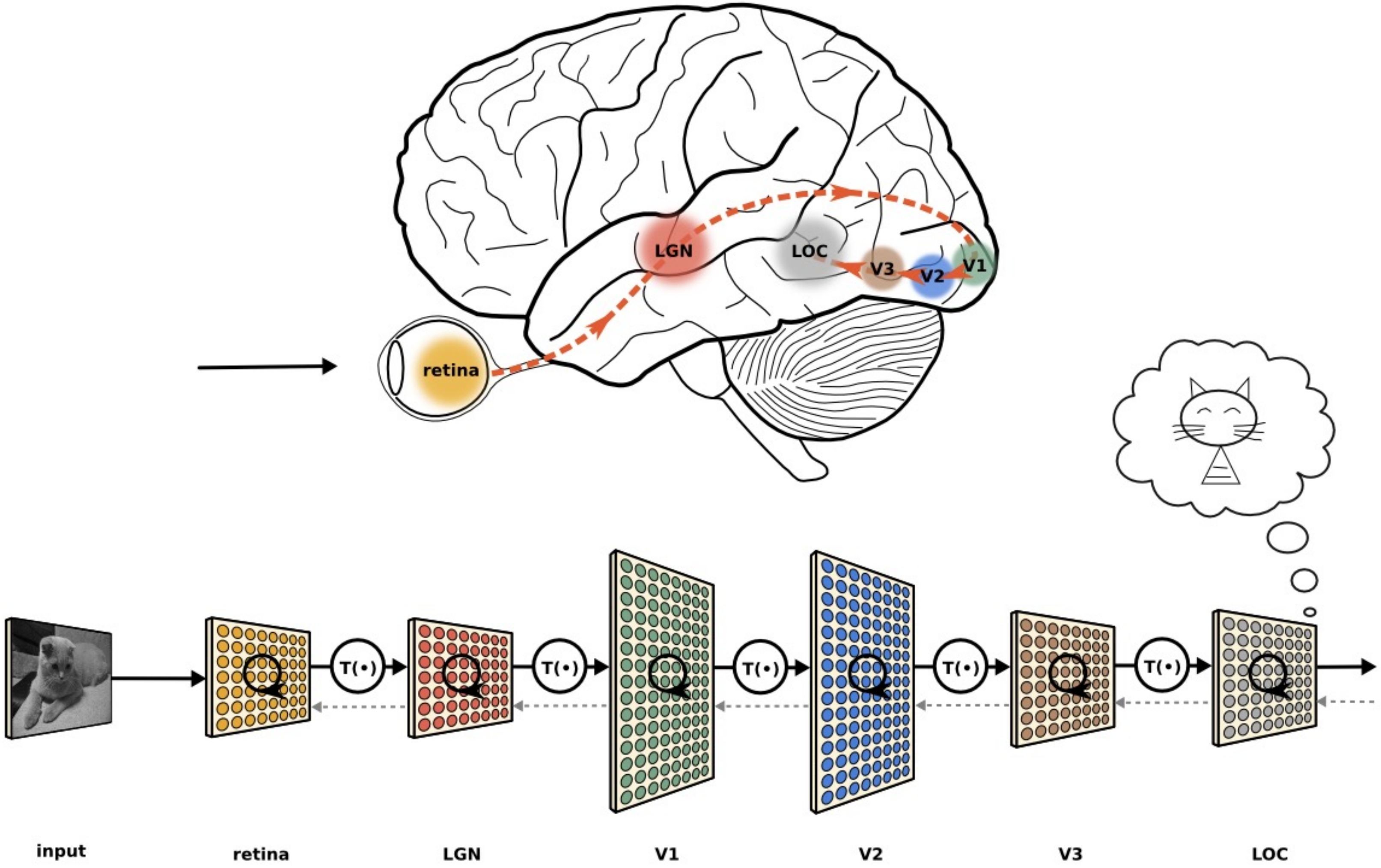
CNN Motivation: How Vision System Works

Key Idea: cells are organized as a hierarchy of feature detectors, with higher level features responding to activated patterns in lower level cells



Source: https://bruceoutdoors.files.wordpress.com/2017/08/hubel.jpg

CNN Motivation: How Vision System Works



https://neuwritesd.files.wordpress.com/2015/10/visual_stream_small.png

Neocognitron: Key Ingredients

http://personalpage.flsi.or.jp/fukushima/index-e.html

"In this paper, we discuss how to synthesize a neural network model in order to endow it an ability of pattern recognition like a human being… the network acquires a similar structure to the hierarchy model of the visual nervous system proposed by Hubel and Wiesel."

- Fukushima, Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. Biological Cybernetics, 1980.

Neocognitron: Key Ingredients

Cascade of simple and complex cells:

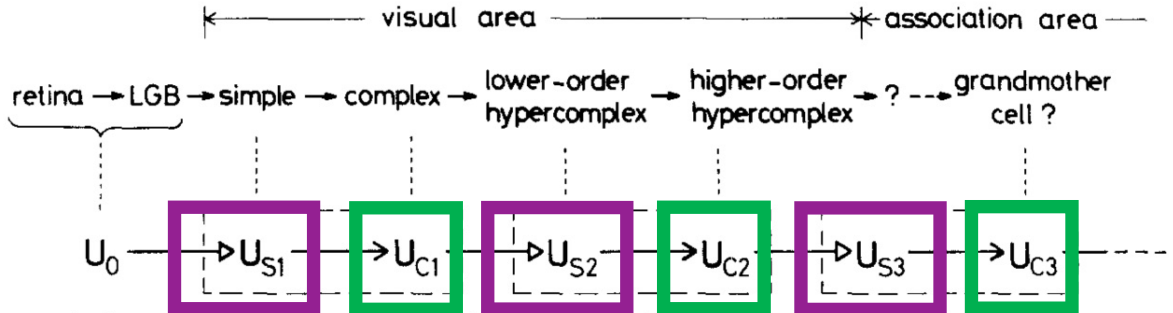


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of the neocognitron

1. Convolutional layers

—▷ modifiable synapses

—> unmodifiable synapses

2. Pooling Layers

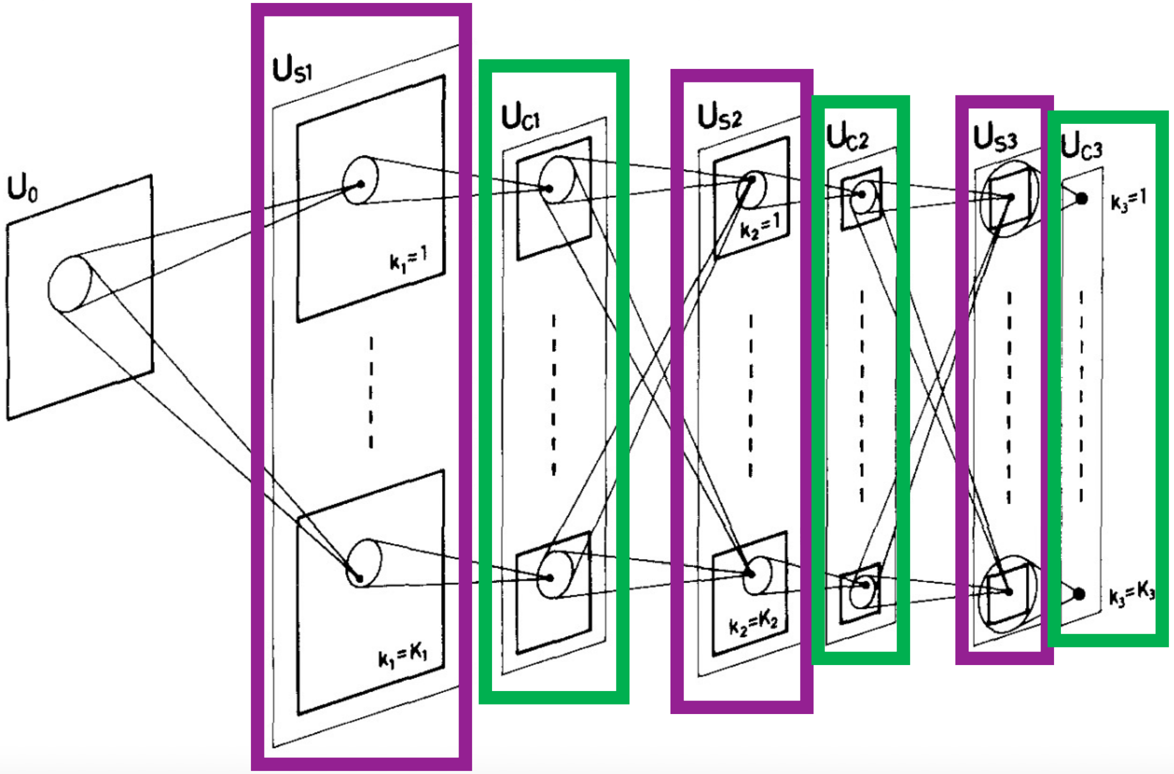
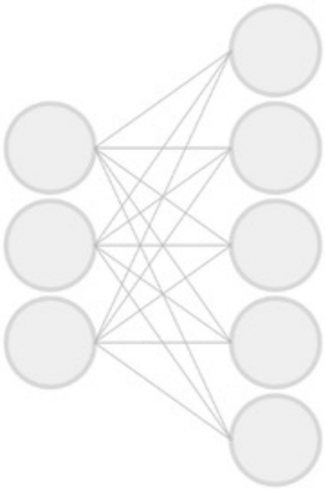


Fig. 2. Schematic diagram illustrating the interconnections between layers in the neocognitron

Fukushima, 1980.

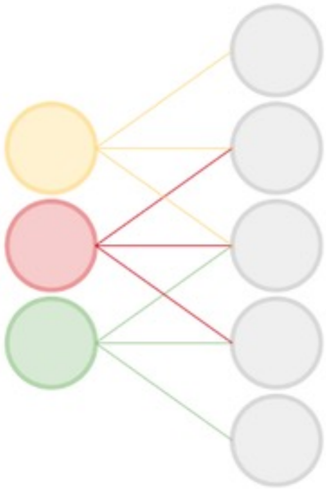
Key Ingredient 1: Convolutional Layers

Fully-connected:



Rather than have each node provide input to each node in the next layer…

Convolutional:

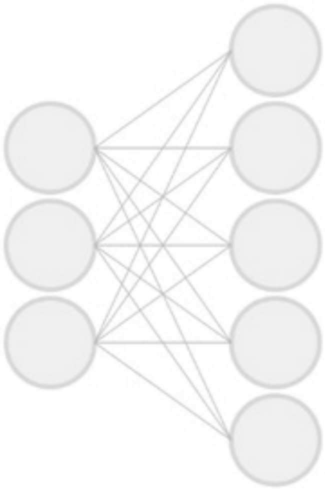


each node receives input only from a small neighborhood in previous layer (and there is parameter sharing)

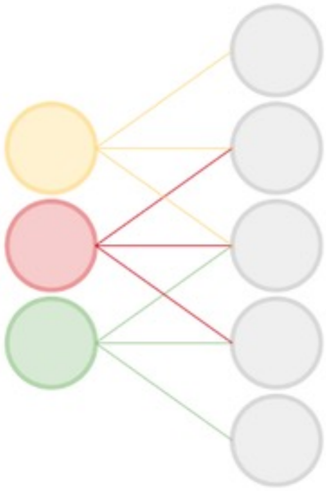
Figure Source: https://qph.fs.quoracdn.net/main-qimg-2e1f0071ca9878f7719ed0ea8aeb386d

Key Ingredient 1: Convolutional Layers

Fully-connected:



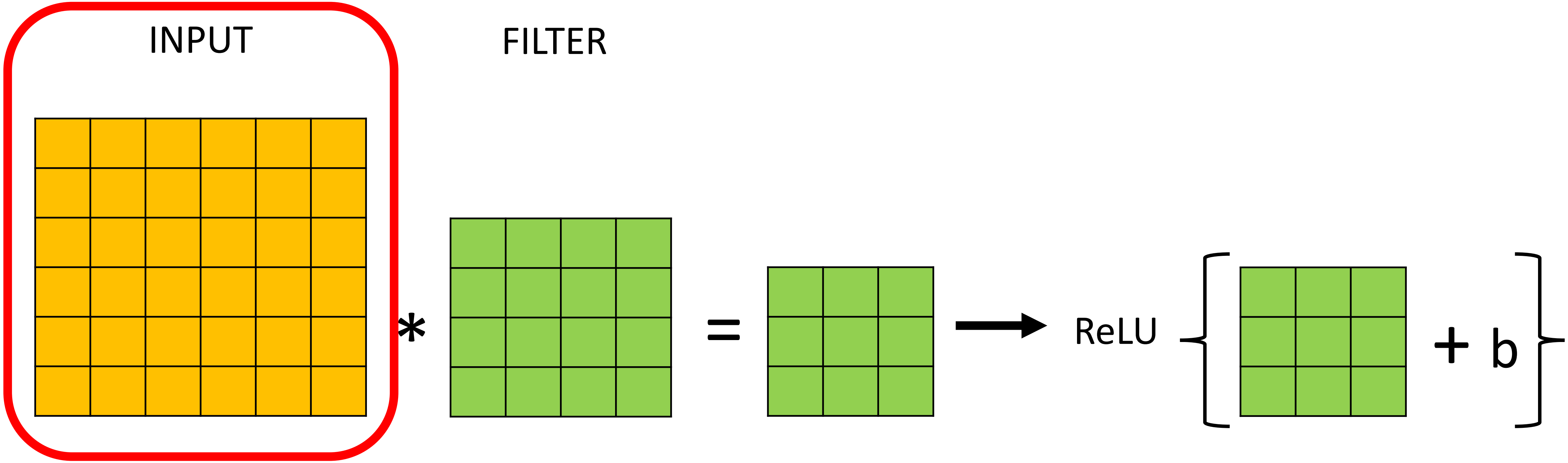
Convolutional:



Convolutional layers dramatically reduce number of model parameters!

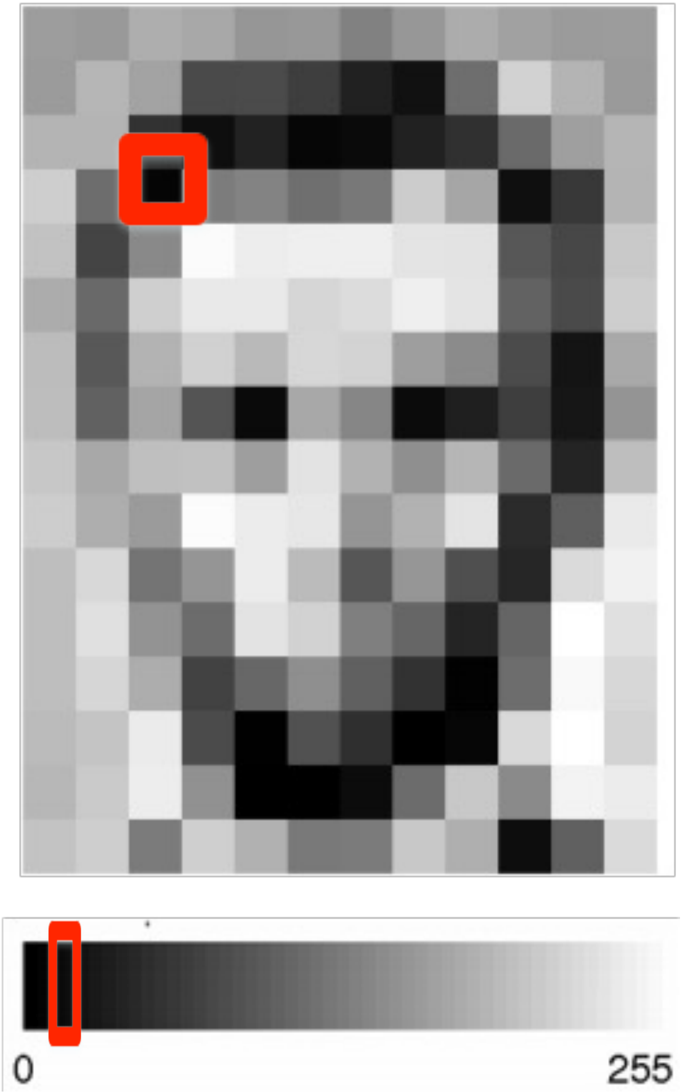
Figure Source: https://qph.fs.quoracdn.net/main-qimg-2e1f0071ca9878f7719ed0ea8aeb386d

Key Ingredient 1: Convolutional Layers



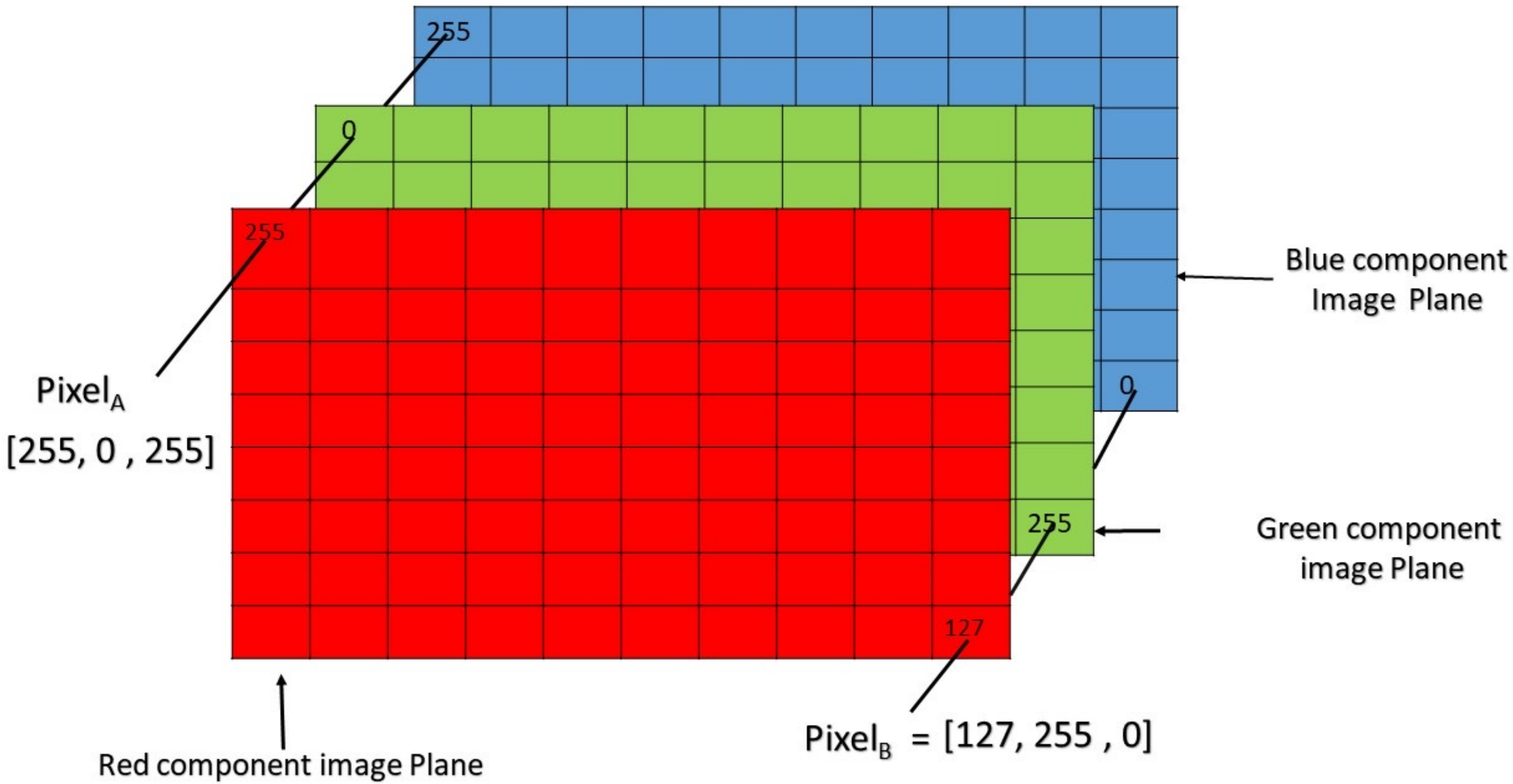
Recall: Image Representation (8-bit Grayscale)

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	10[illegible]	5	[illegible]4	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



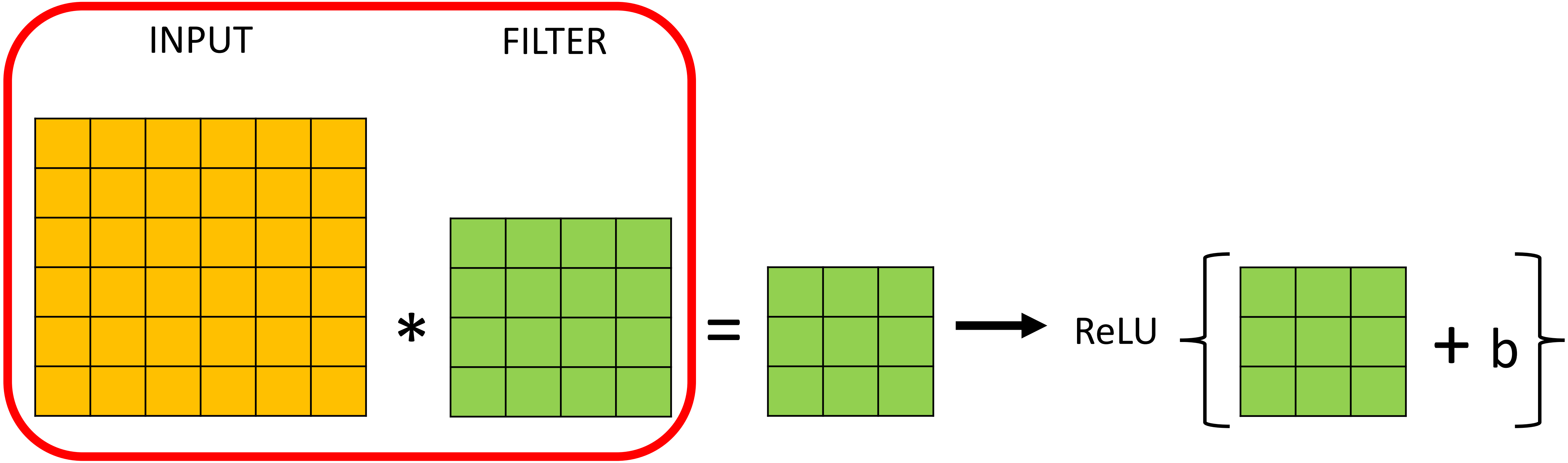
https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html

Color Images (e.g., 24-bit RGB image)

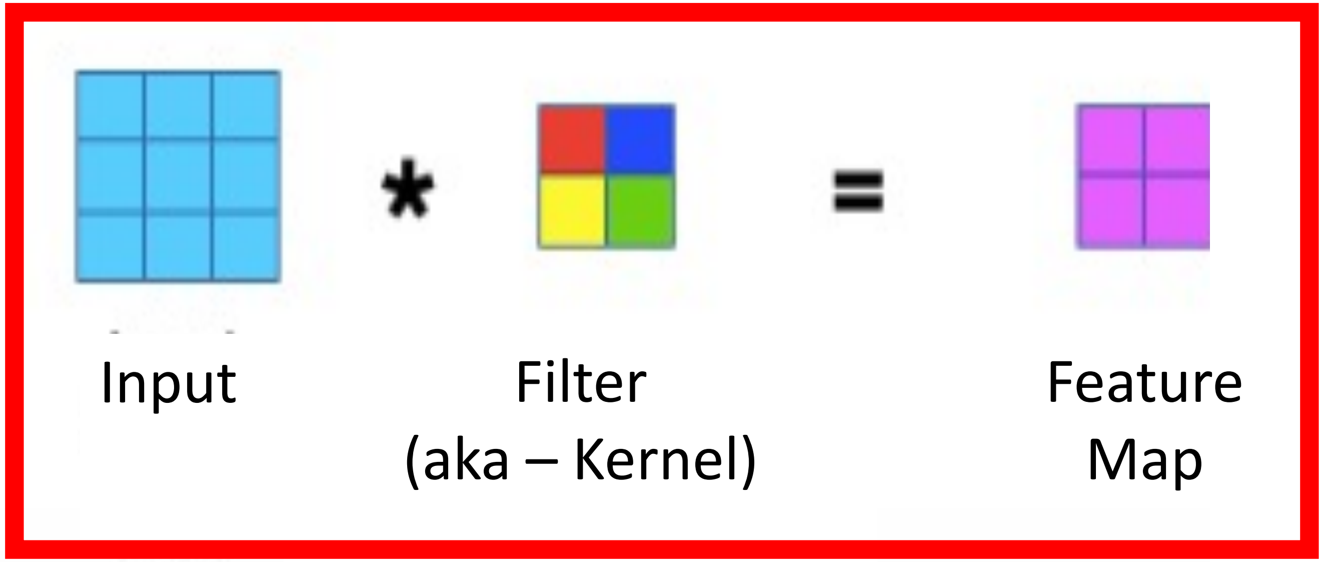


https://www.geeksforgeeks.org/matlab-rgb-image-representation/

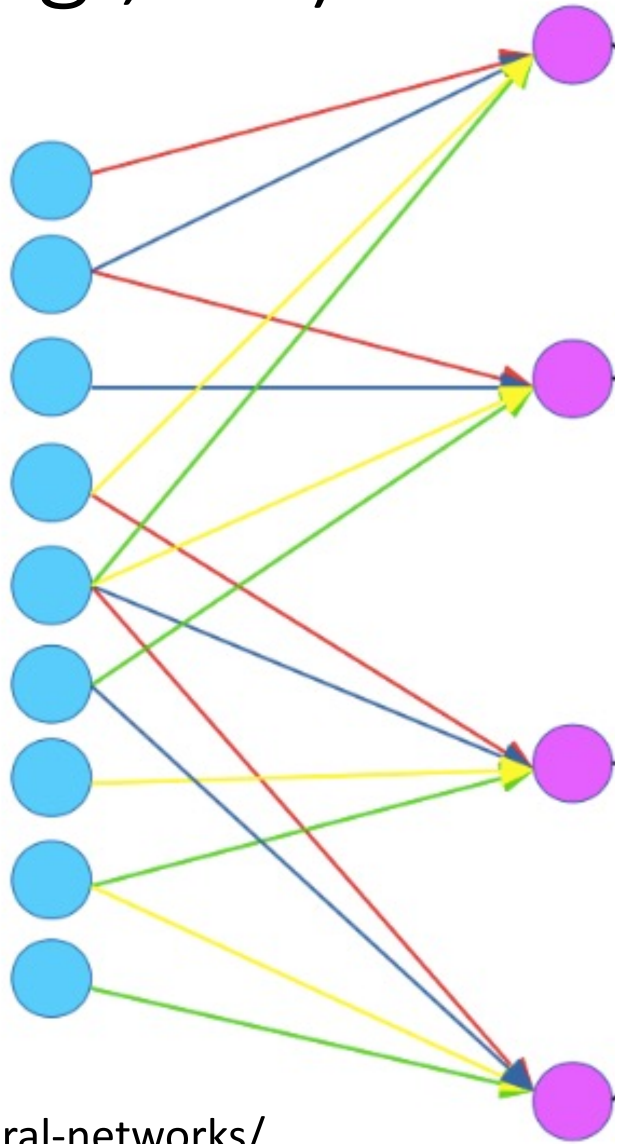
Key Ingredient 1: Convolutional Layers



Convolution: Applies Linear Filter (e.g., 2D)



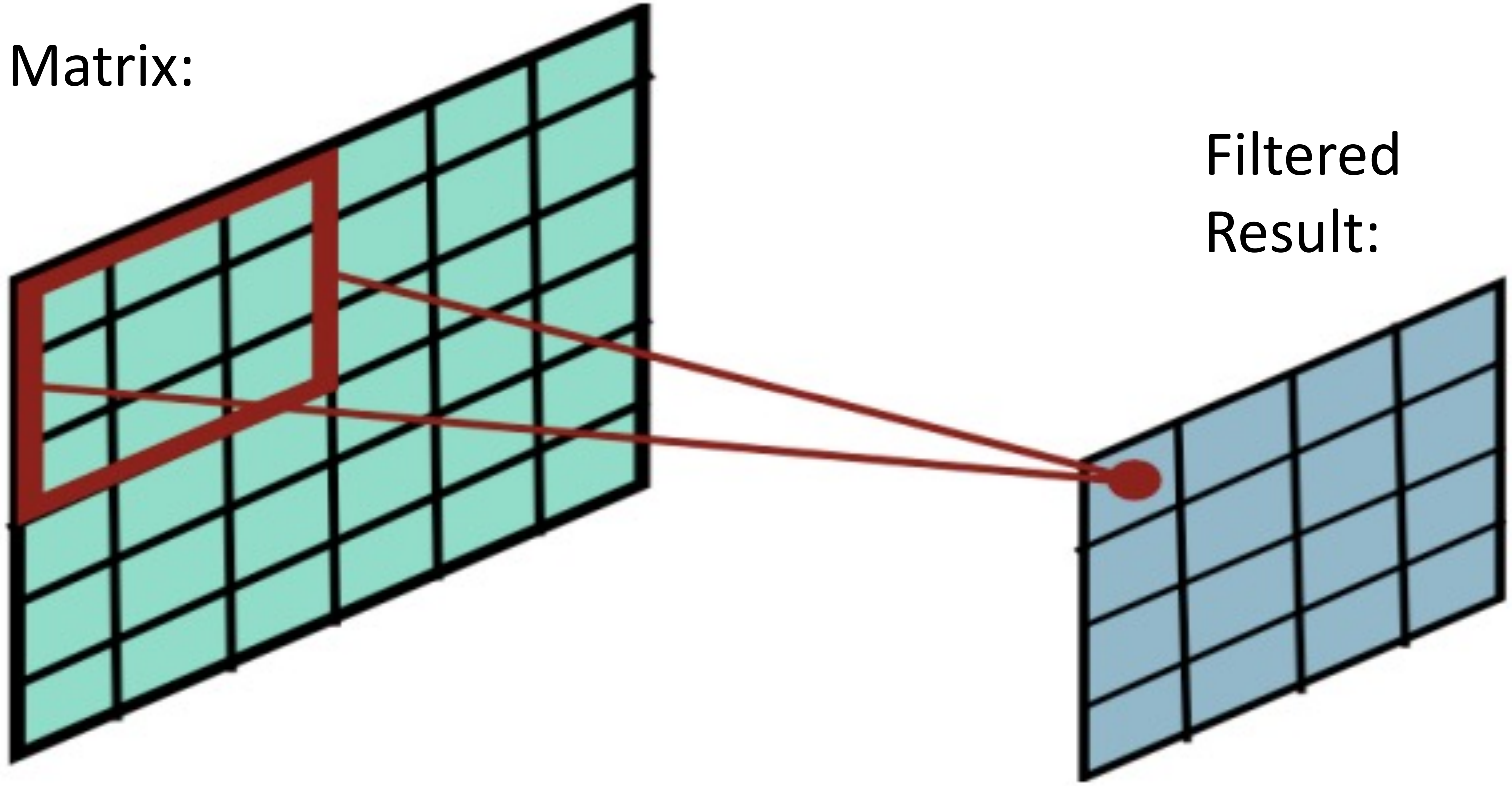
Way to Interpret Neural Network



- Compute a function of local neighborhood for each location in matrix
- A filter specifies the function for how to combine neighbors' values

https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

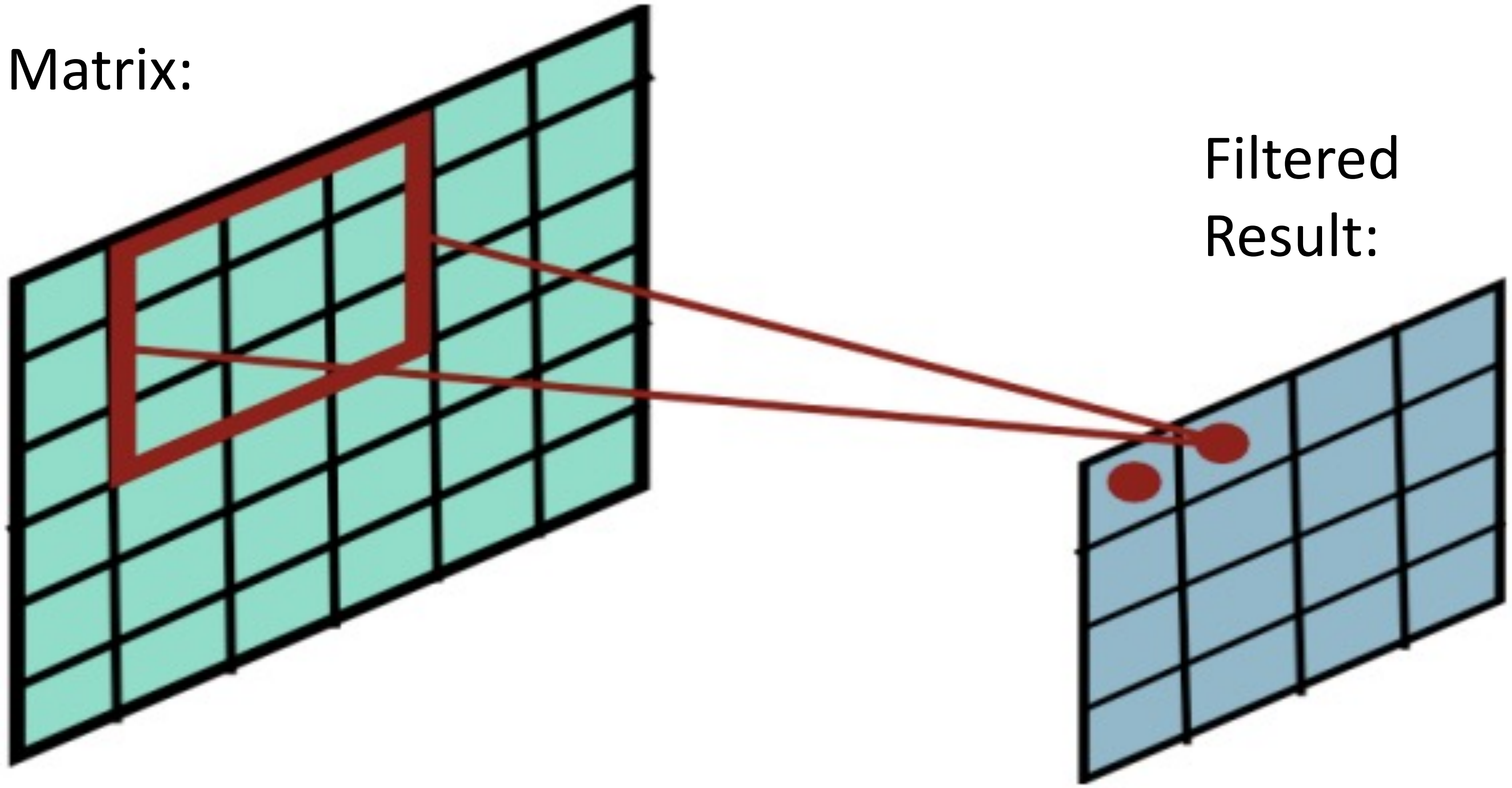
2D Filtering



Slides filter over the matrix and computes dot products

https://people.eecs.berkeley.edu/~jrs/189/lec/cnn.pdf

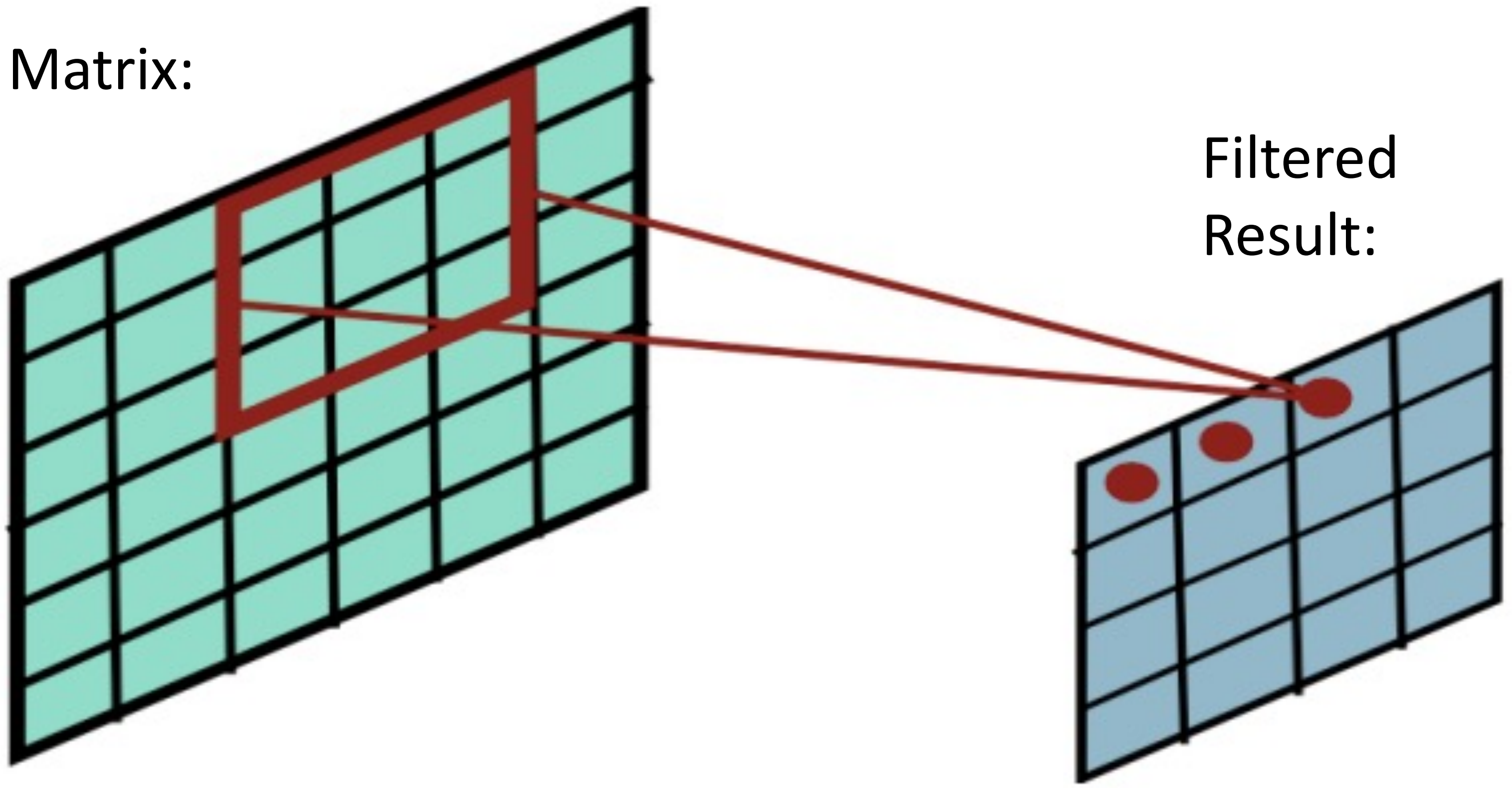
2D Filtering



Slides filter over the matrix and computes dot products

https://people.eecs.berkeley.edu/~jrs/189/lec/cnn.pdf

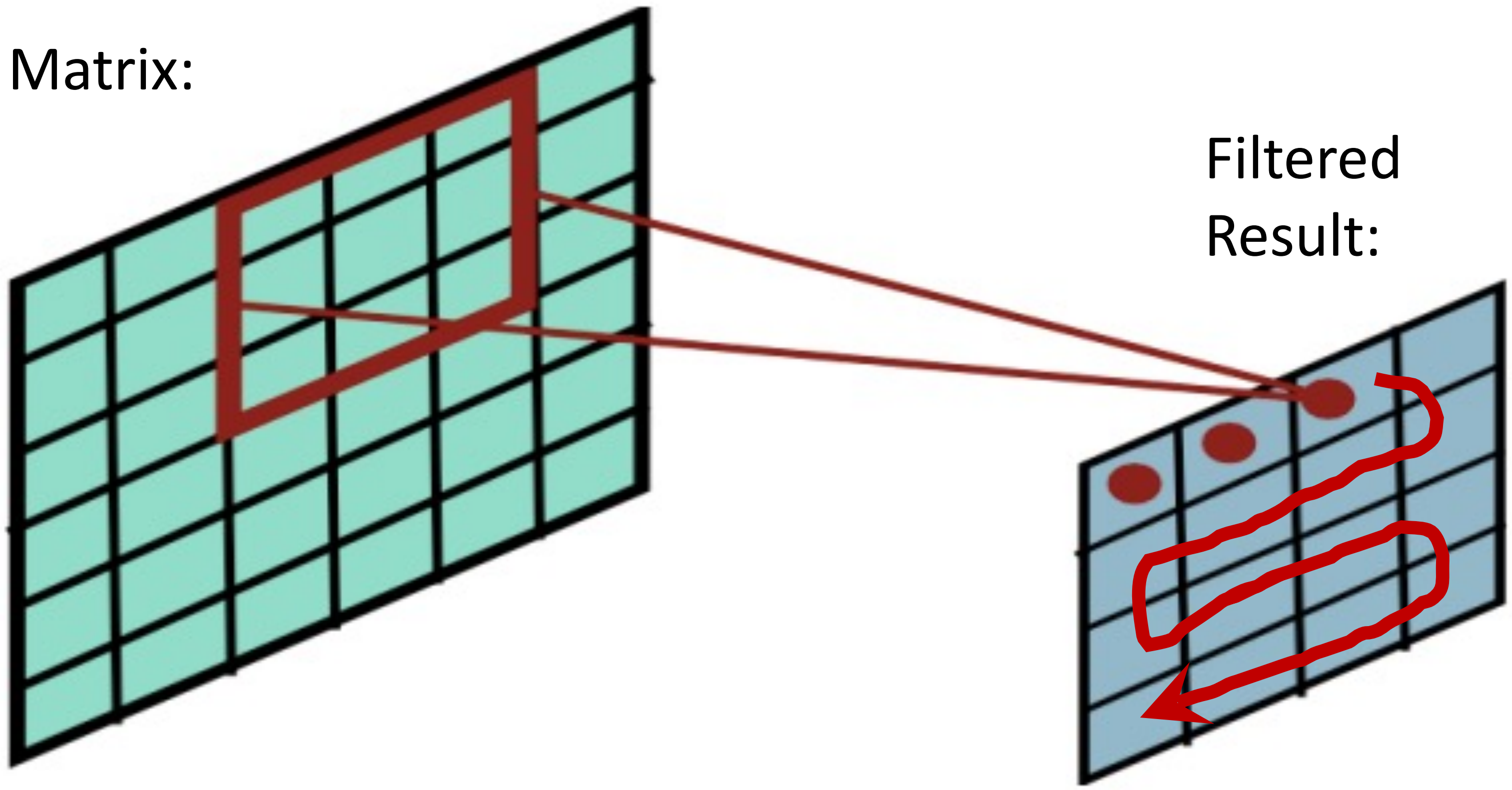
2D Filtering



Slides filter over the matrix and computes dot products

https://people.eecs.berkeley.edu/~jrs/189/lec/cnn.pdf

2D Filtering



Slides filter over the matrix and computes dot products

https://people.eecs.berkeley.edu/~jrs/189/lec/cnn.pdf

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

?	?	?
?	?	?
?	?	?

Dot Product = 1*1 + 1*0 + 1*1 + 0*0 + 1*1 + 1*0 + 0*1 + 0*1 + 0*0 + 0*0 + 1*1

Dot Product = 4

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	?	?
?	?	?
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	?
?	?	?
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
?	?	?
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
2	?	?
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
2	4	?
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
2	4	3
?	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
2	4	3
2	?	?

2D Filtering: Toy Example

Input

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

Filter

1	0	1
0	1	0
1	0	1

Feature Map

4	3	4
2	4	3
2	3	?

2D Filtering: Toy Example

Input

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

Filter

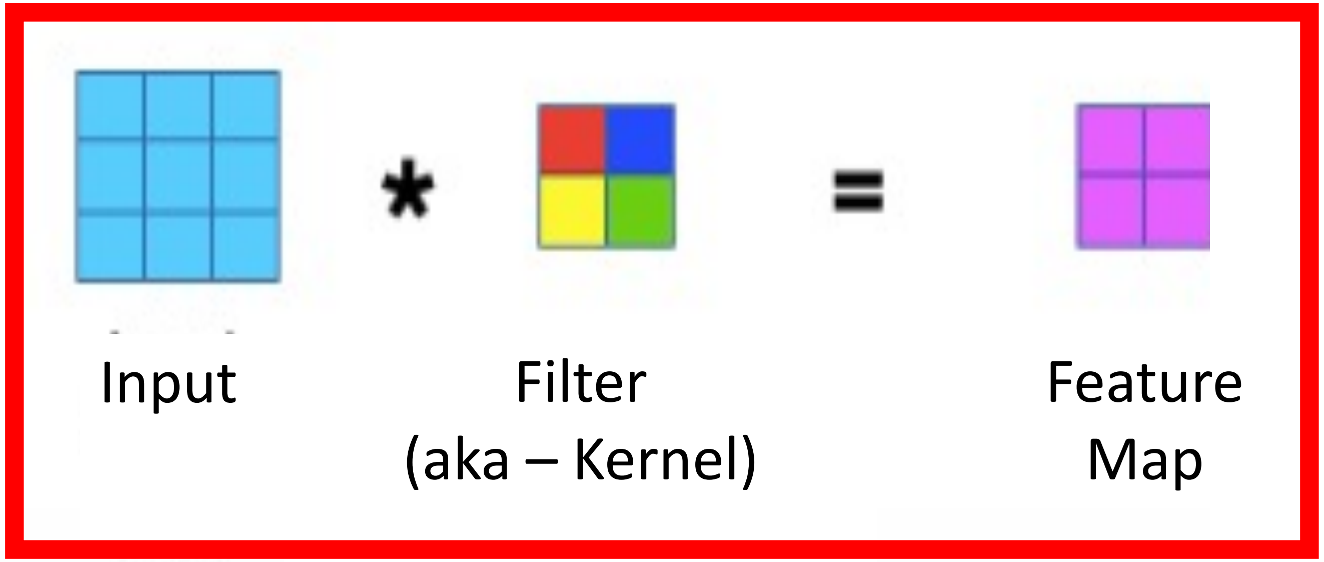
1	0	1
0	1	0
1	0	1

Feature Map

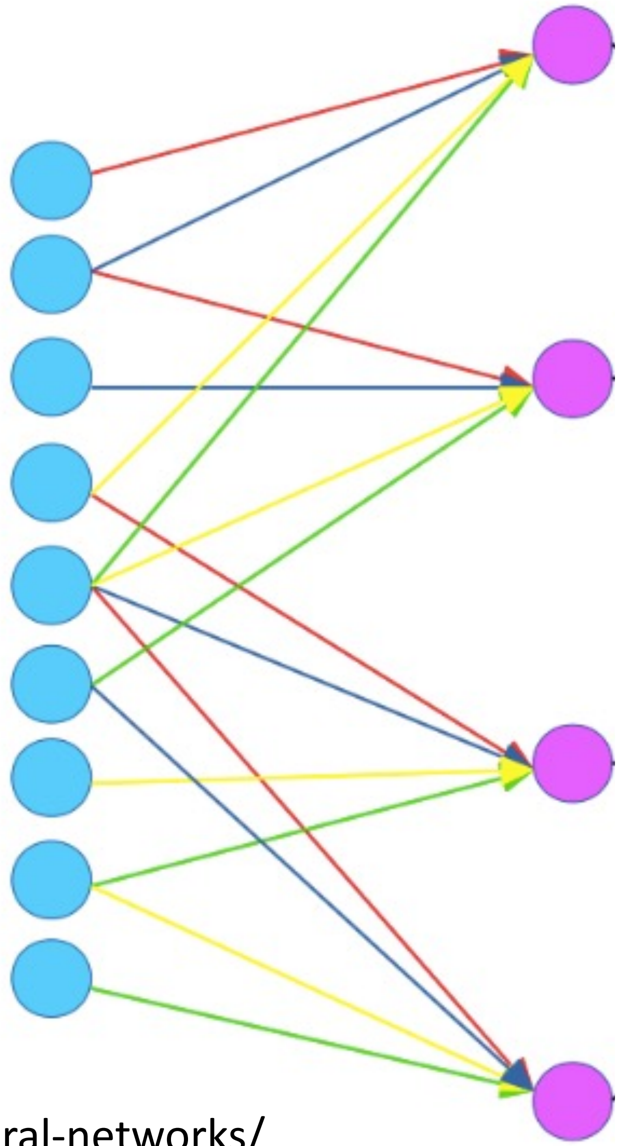
4	3	4
2	4	3
2	3	4

Convolutional Layer

- Many neural network libraries use “convolution” interchangeably with “cross correlation”; for mathematicians, these are technically different
- Examples in these slides show the “cross-correlation” function

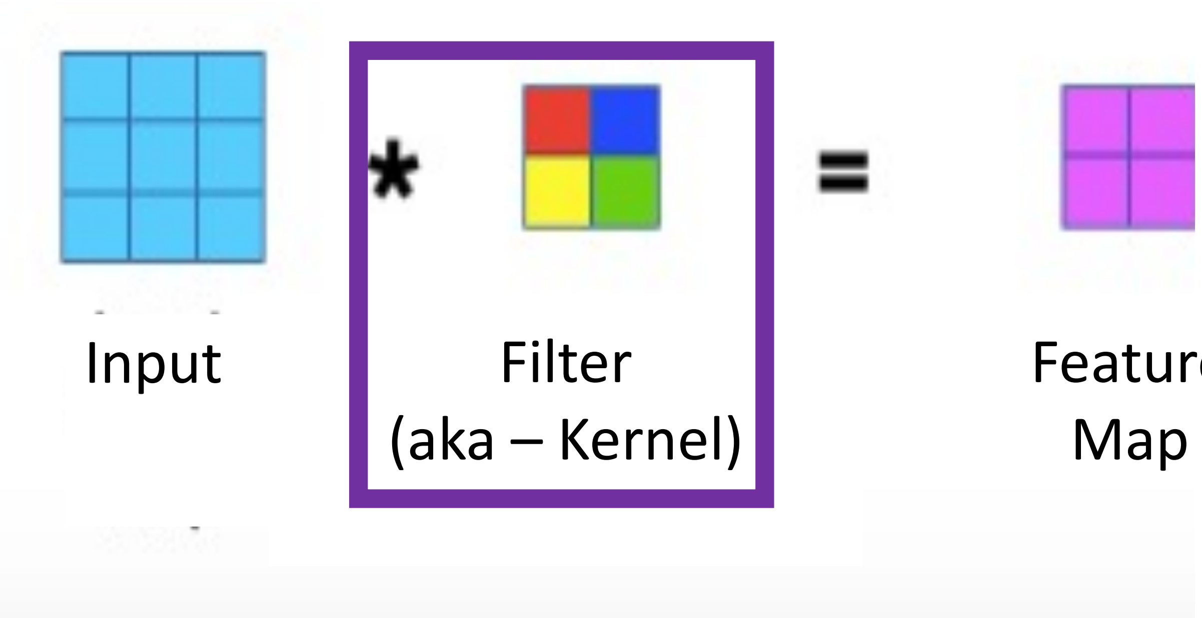


Way to Interpret Neural Network

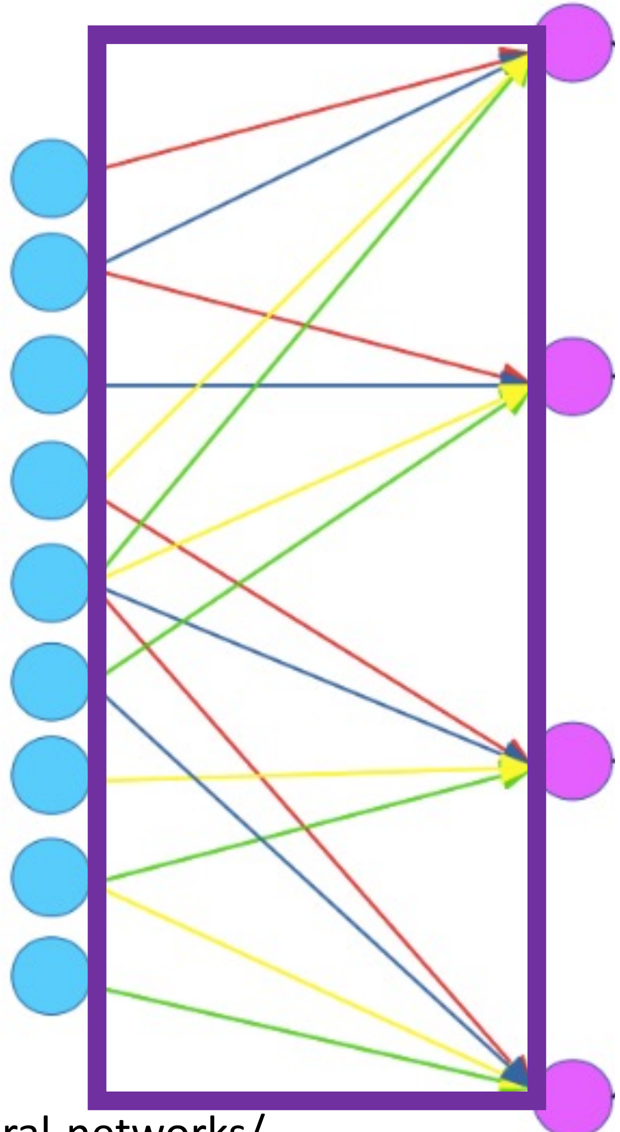


https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Convolutional Layer: Parameters to Learn



Way to Interpret
Neural Network

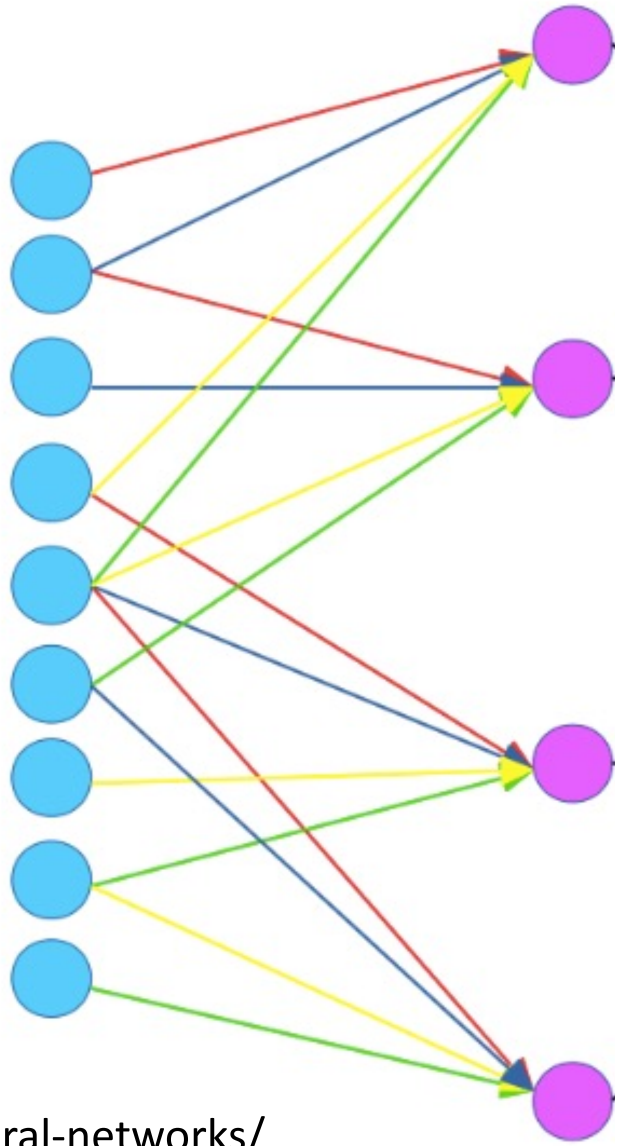


https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Convolutional Layer: Parameters to Learn

- For shown example, how many weights must be learned?
  - 4 (red, blue, yellow, and green values)

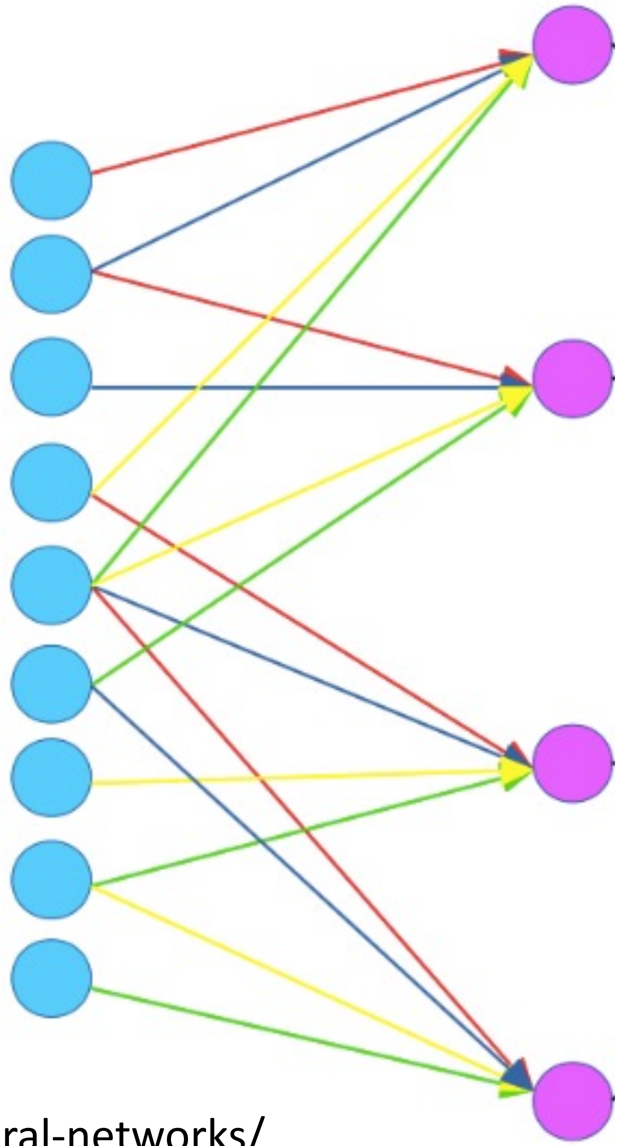
- If we instead used a fully connected layer, how many weights would need to be learned?
  - 36 (9 turquoise nodes x 4 magenta nodes)



https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

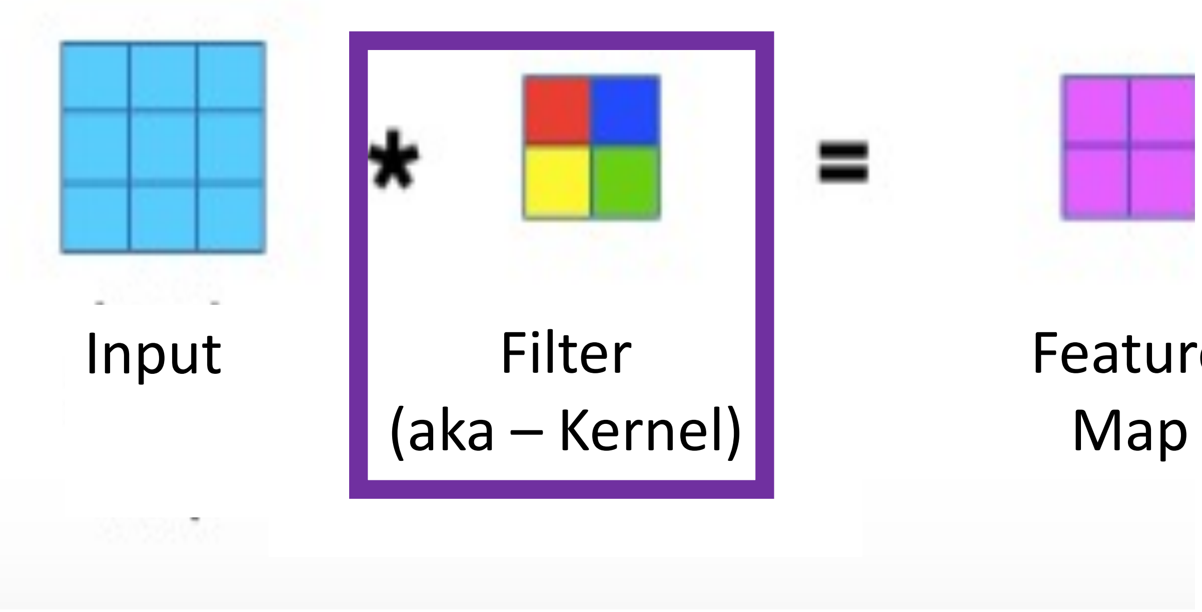
Convolutional Layer: Parameters to Learn

Neocognitron hard-coded filter values...
filter values are learned for CNNs

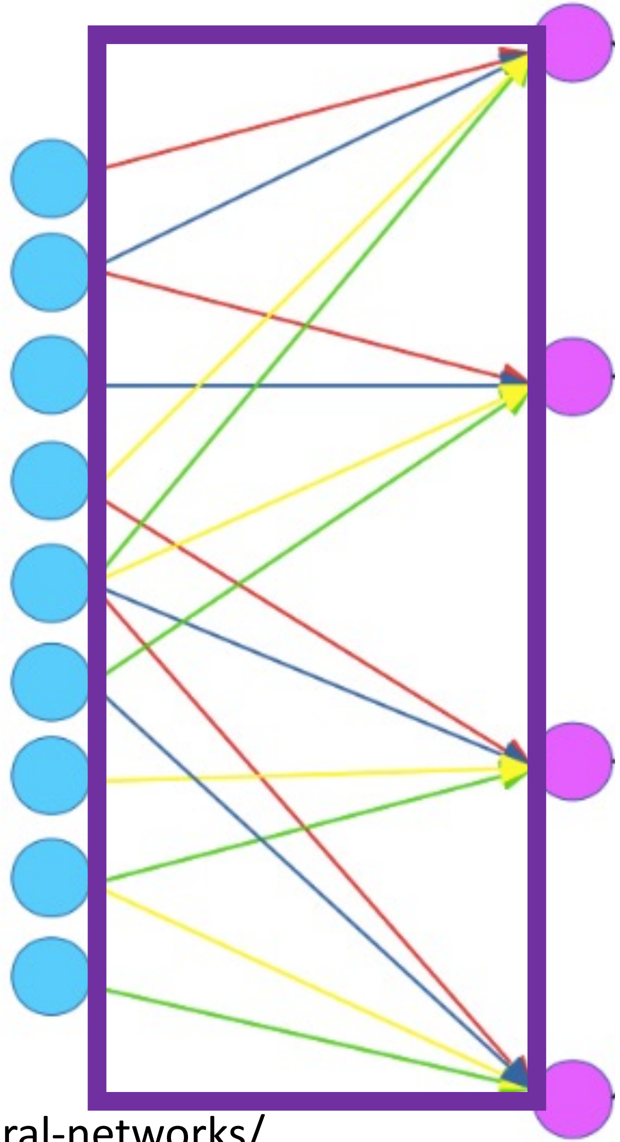


https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Convolutional Layer: What Can Filters Do?



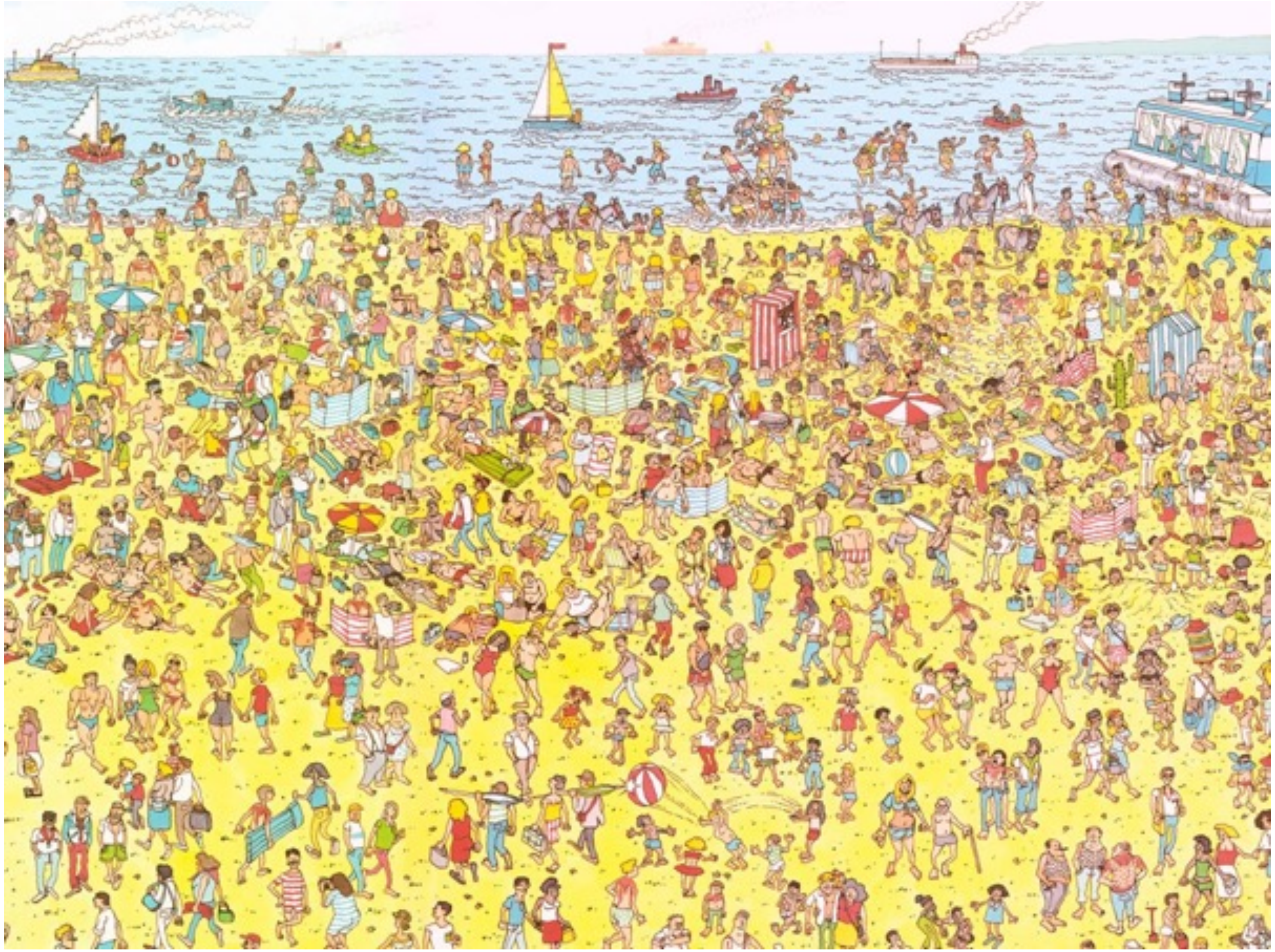
Way to Interpret
Neural Network



https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Convolutional Layer: What Can Filters Do?

Filter



Convolutional Layer: What Can Filters Do?

- e.g.,

Filter

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Visualization of Filter

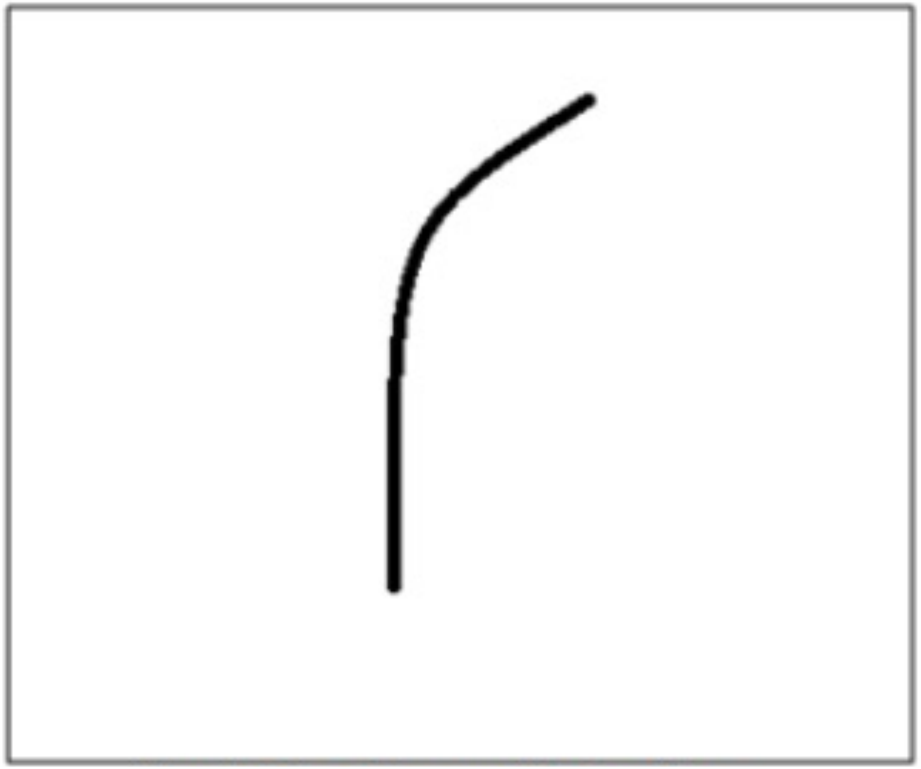
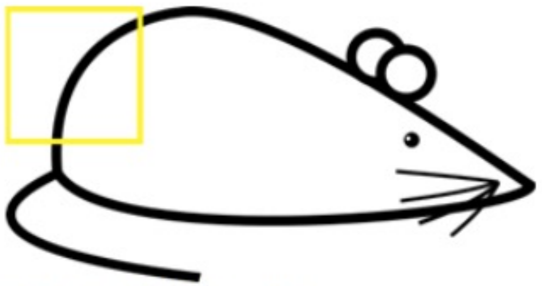


Image Credit: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

Convolutional Layer: What Can Filters Do?

- e.g.,

Filter Overlaid on Image



Image

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

*

Filter

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Weighted Sum = ?

Weighted Sum = (50x30) + (20x30) + (50x30) + (50x3) + (50x30)

Weighted Sum = 6600 **(Large Number!!)**

Image Credit: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

Convolutional Layer: What Can Filters Do?

- e.g.,

Filter Overlaid on Image

Image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

*

Filter

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Weighted Sum = ?

Weighted Sum = 0 **(Small Number!!)**

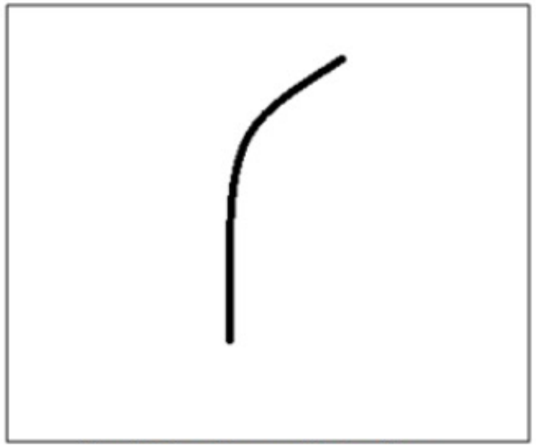
Image Credit: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

Convolutional Layer: What Can Filters Do?

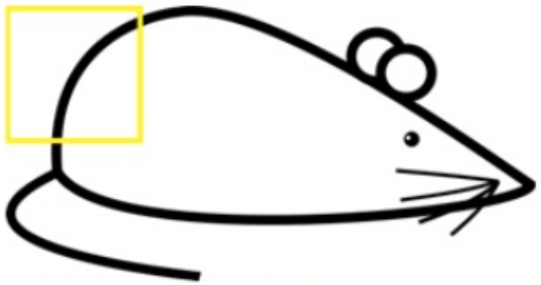
This Filter is a Curve Detector!

- e.g.,

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0



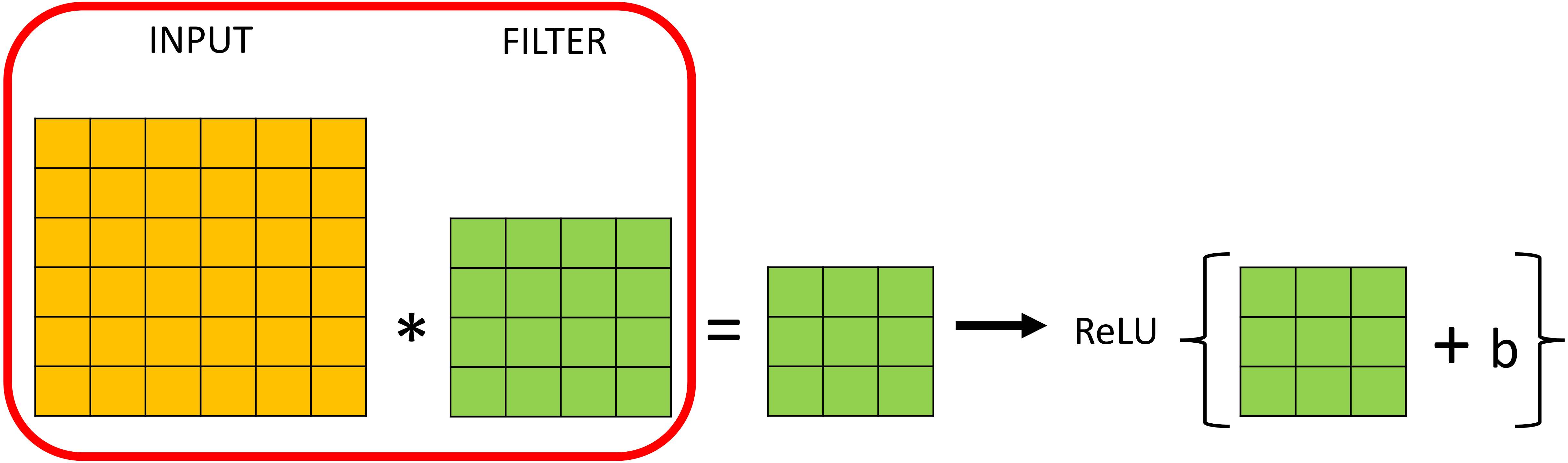
Filter Overlaid on Image (Big Response!)



Filter Overlaid on Image (Small Response!)

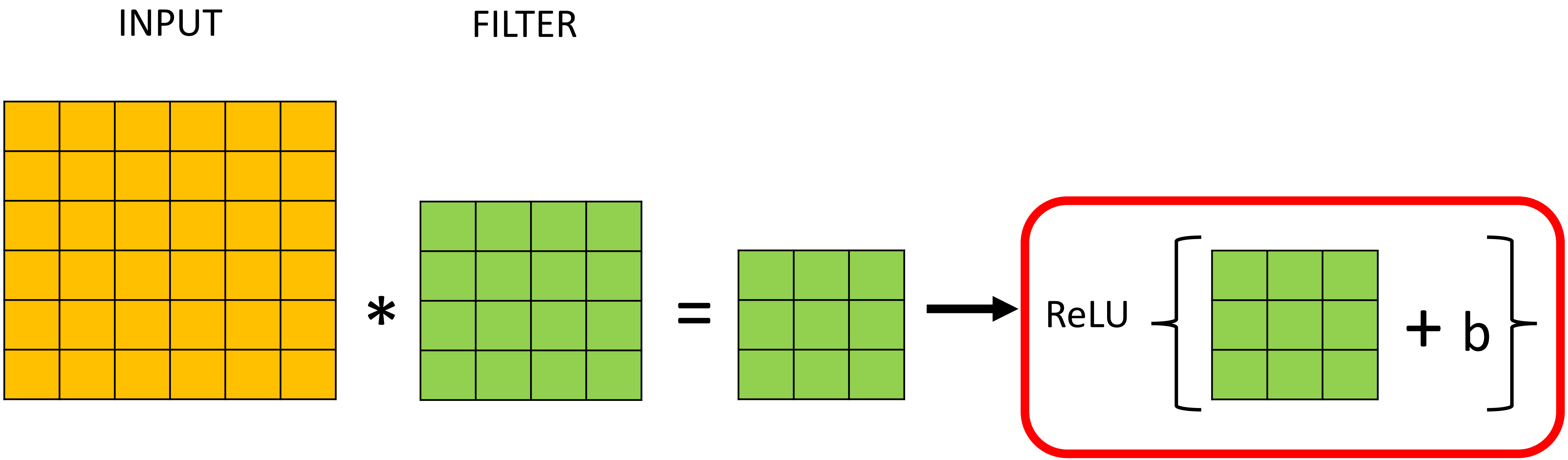
Image Credit: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

Key Ingredient 1: Convolutional Layers



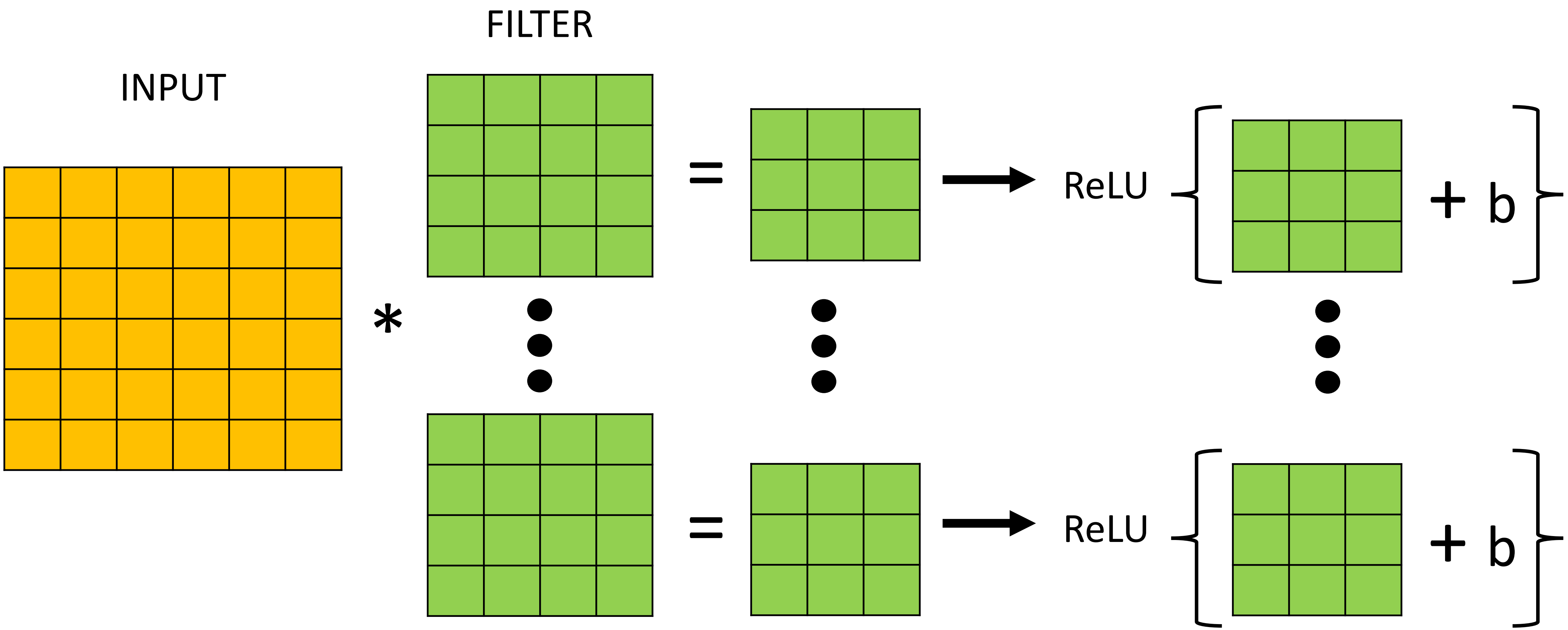
Can choose filters of any size to support feature learning!

Key Ingredient 1: Convolutional Layers



Filtered results are passed, with a bias term, through an activation function to create **activation/feature maps**

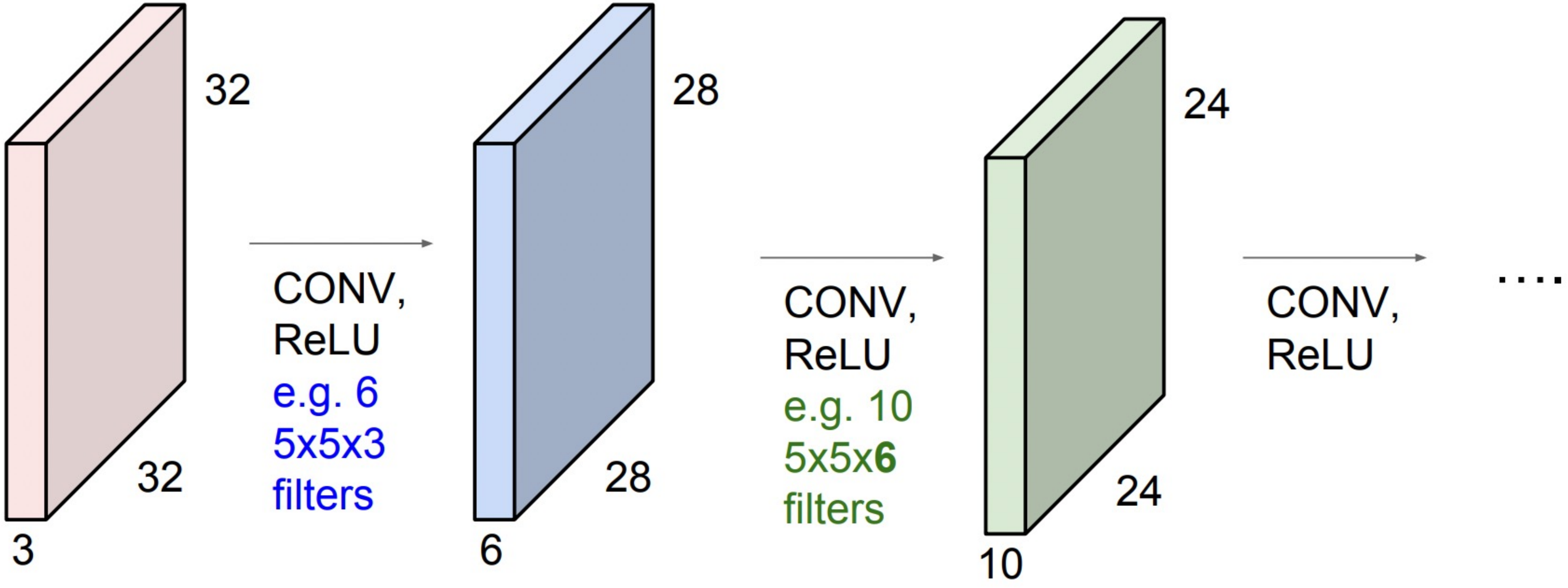
Key Ingredient 1: Convolutional Layers



Can have multiple filters (with a unique bias parameter per filter)

Convolutional Layers Stacked

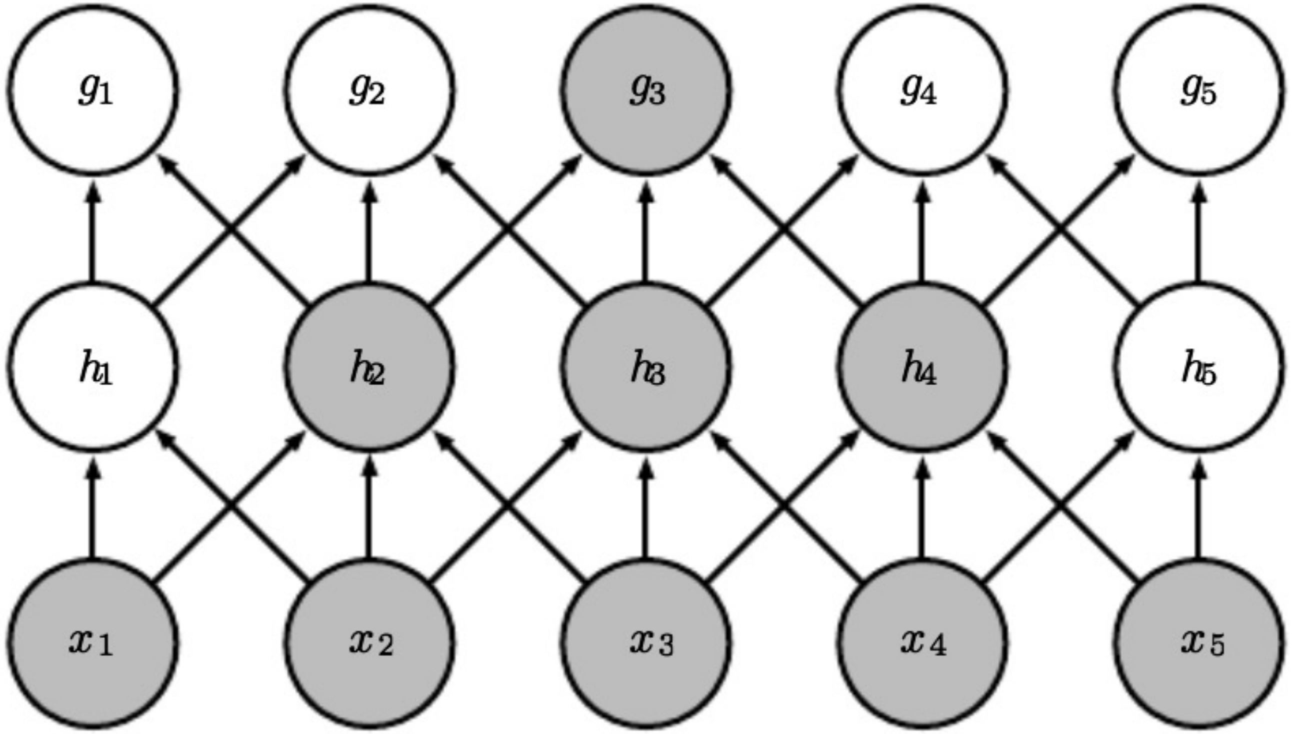
Can then stack a sequence of convolution layers; e.g.,



http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture05.pdf

Convolutional Layers Stacked

Can then stack a sequence of convolution layers, which leads to identifying patterns in increasingly **larger regions of the input (e.g., pixel) space:**

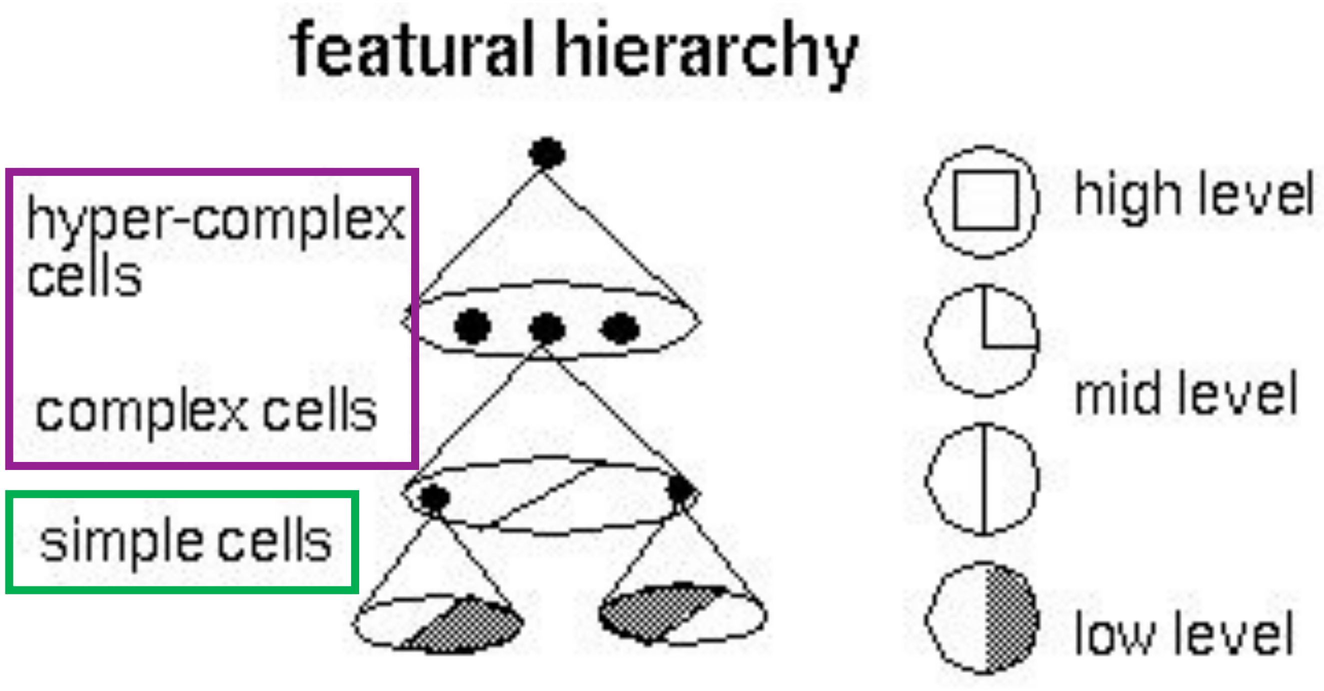


https://www.deeplearningbook.org/contents/convnets.html

Convolutional Layers Stacked

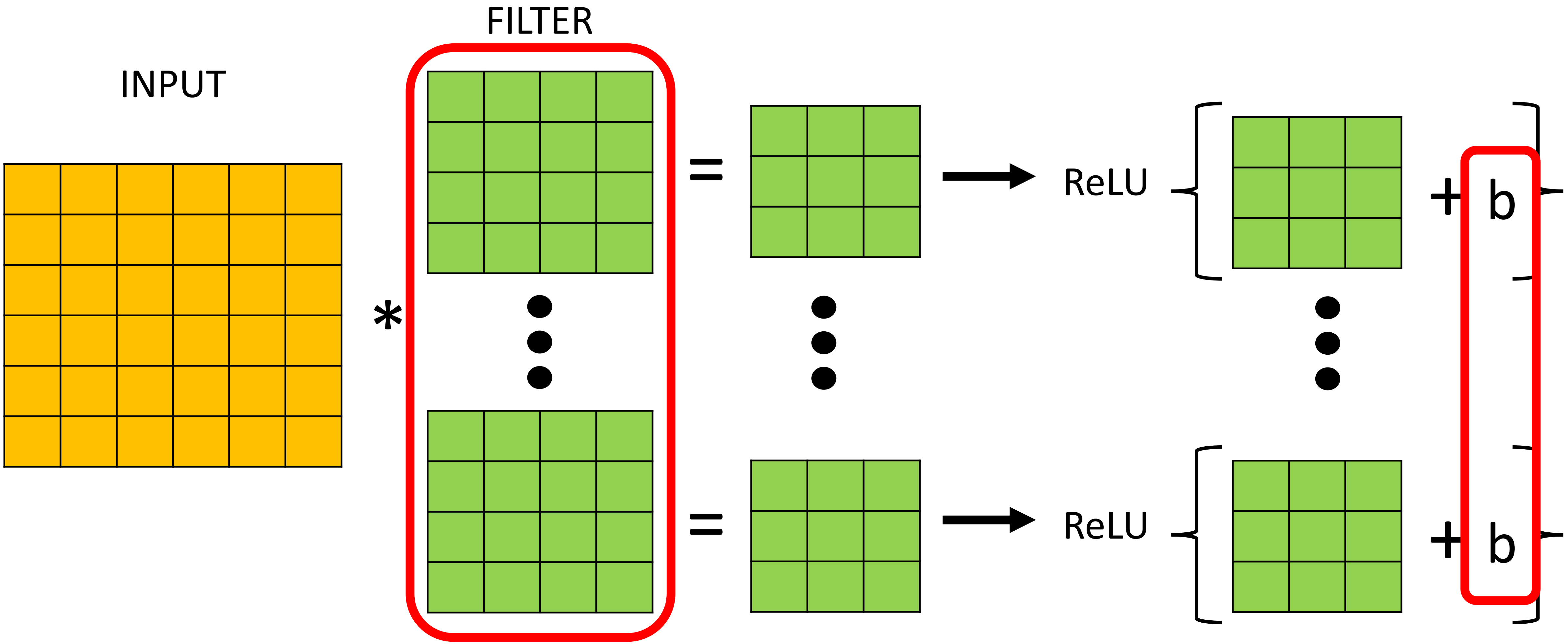
Can then stack a sequence of convolution layers, which leads to identifying patterns in increasingly **larger regions of the input (e.g., pixel) space** and **mimicking vision system**:

Higher level features are constructed by combining lower level features



Source: https://bruceoutdoors.files.wordpress.com/2017/08/hubel.jpg

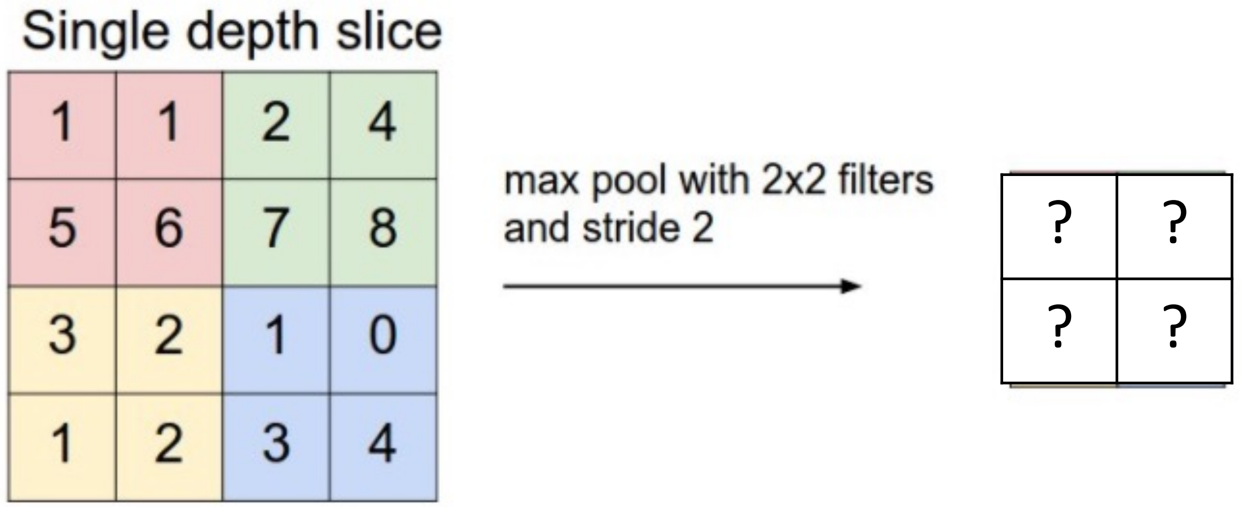
Key Ingredient 1: Convolutional Layer Summary



Neural networks learn values for all filters and biases in all layers

Key Ingredient 2: Pooling Layer

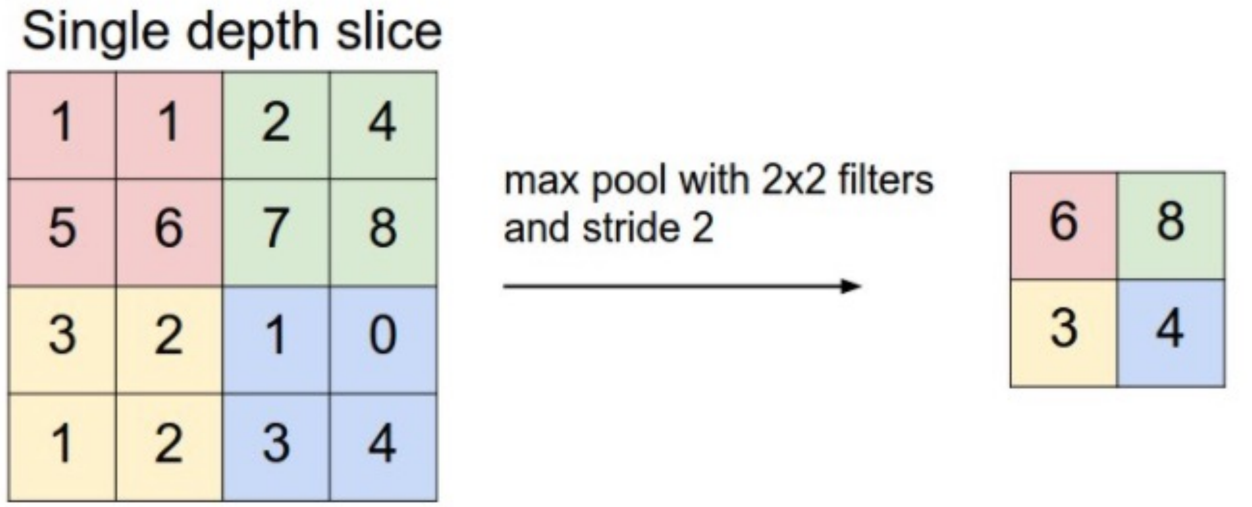
- Summarizes neighborhood; e.g., **max-pooling** partitions input into non-overlapping rectangles and outputs maximum value per chunk



http://cs231n.github.io/convolutional-networks/#pool

Key Ingredient 2: Pooling Layer

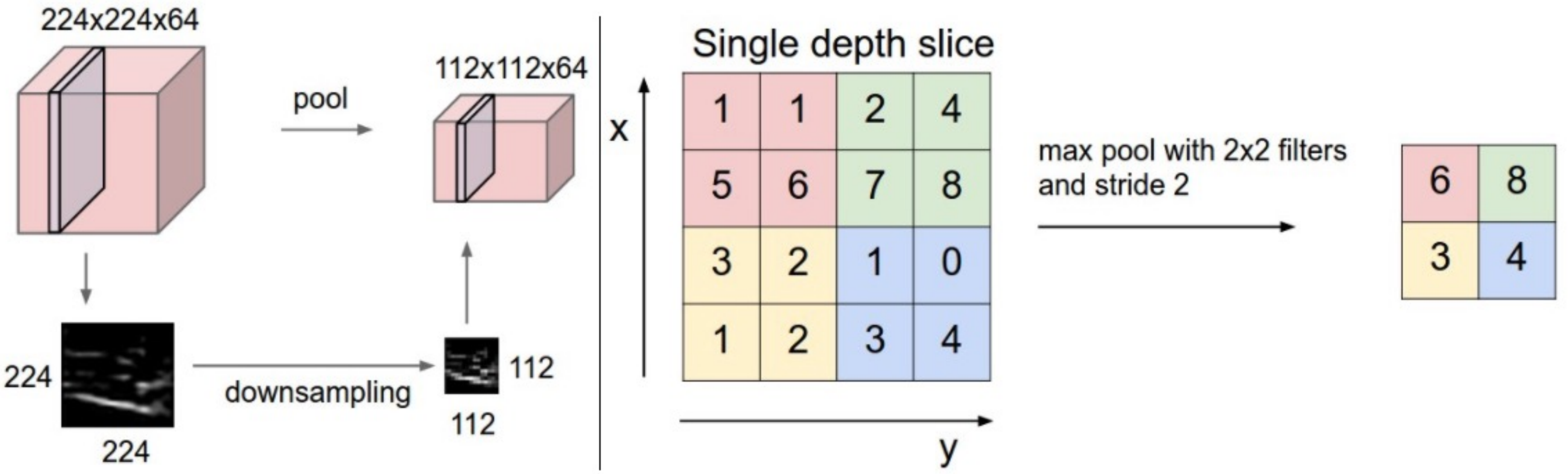
- Summarizes neighborhood**;** e.g., **max-pooling** partitions input into non-overlapping rectangles and outputs maximum value per chunk



http://cs231n.github.io/convolutional-networks/#pool

Key Ingredient 2: Pooling Layer

- Summarizes neighborhood**;** e.g., **max-pooling** partitions input into non-overlapping rectangles and outputs maximum value per chunk



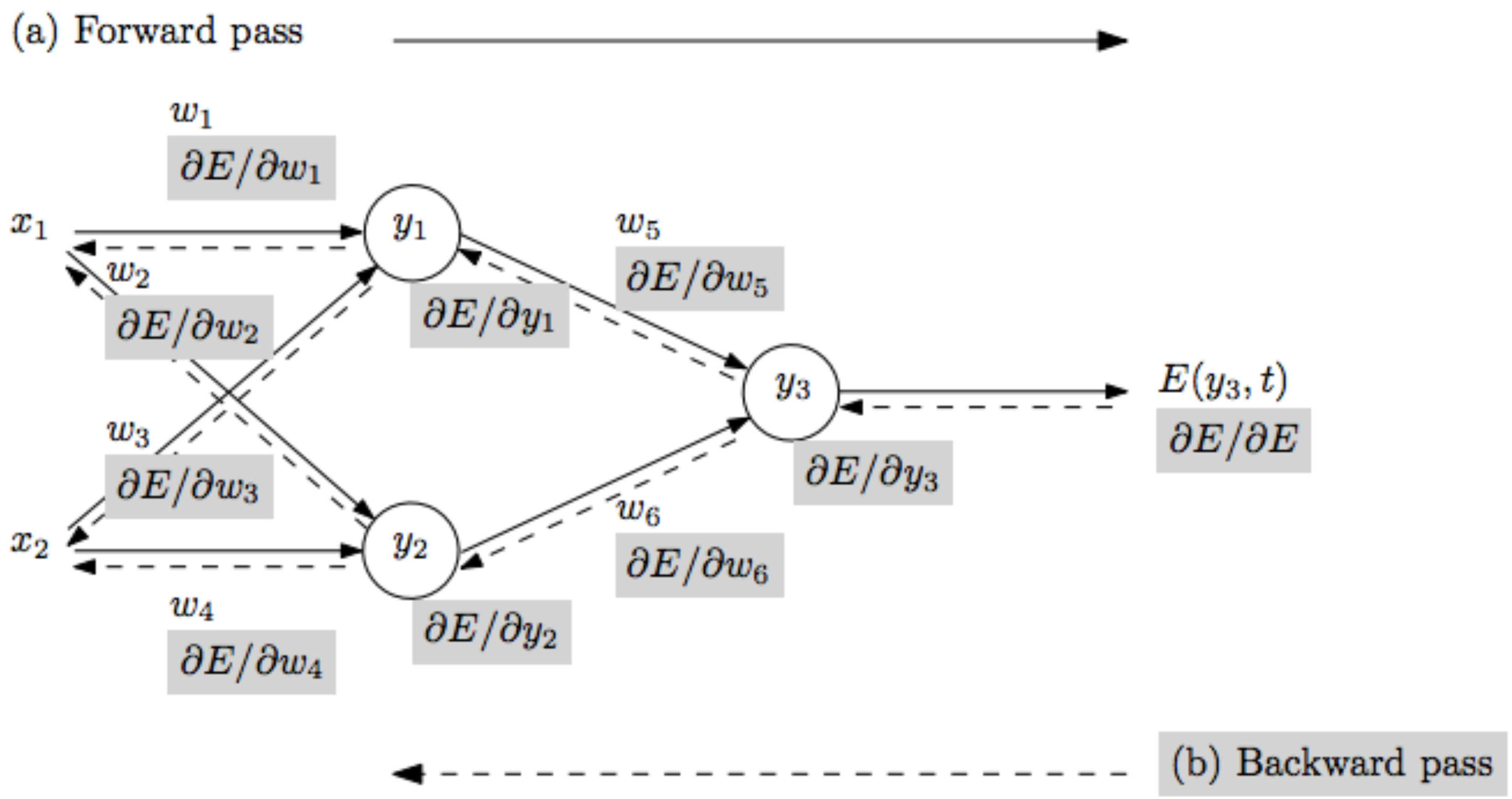
http://cs231n.github.io/convolutional-networks/#pool

Key Ingredient 2: Pooling Layer Benefits

- How many parameters must be learned?
  - None

- Benefits?
  - Reduces memory requirements
  - Reduces computational requirements

Approach to Train CNNs



- Repeat until stopping criterion met:
  1. **Forward pass**: propagate training data through model to make predictions
  2. **Error quantification**: measure dissatisfaction with a model's predictions on training data
  3. **Backward pass**: using predicted output, calculate gradients backward to assign blame to each model parameter; account for weight sharing by using average of all connections for a parameter
  4. Update each parameter using calculated gradients

https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Object Recognition: Today's Topics

- Problem
- Applications
- Datasets
- Evaluation metric
- A Popular Solution: Convolutional Neural Network


The End