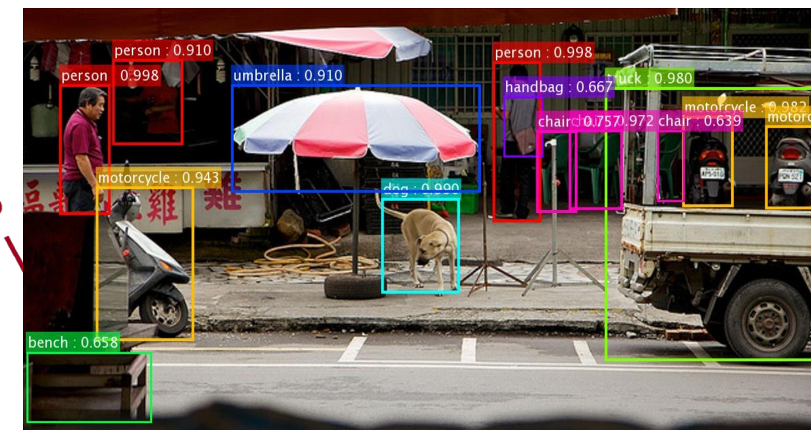
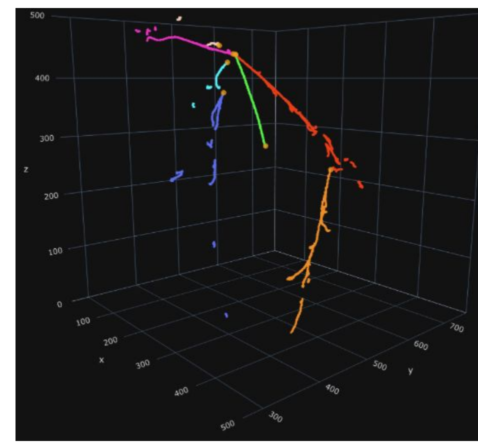
Introduction to Convolutional Neural Networks

November 17th 2020
KMI-2020
Nagoya University

Kazuhiro Terao
SLAC National Lab

Image Analysis In Computer Vision





U.S. DEPARTMENT OF ENERGY
Office of Science

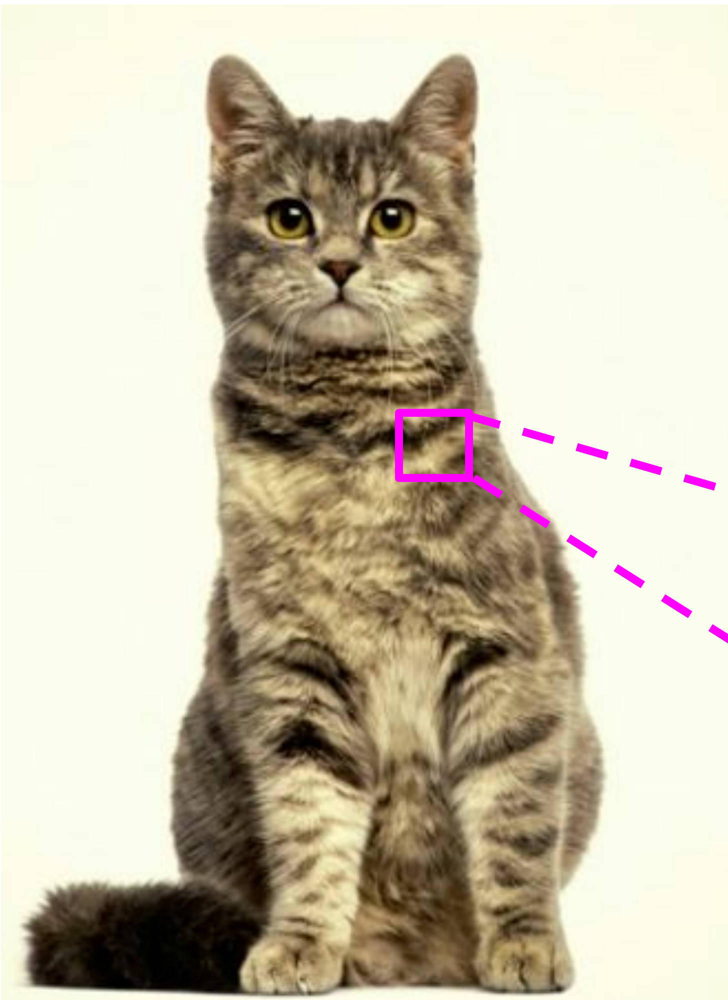
SLAC NATIONAL ACCELERATOR LABORATORY

Challenge in Computer Vision: Find a Cat

How to write an algorithm to identify a cat?

Convolutional Neural Networks

Challenge in Computer Vision: Find a Cat



How to write an algorithm to identify a cat?

... very hard task ...

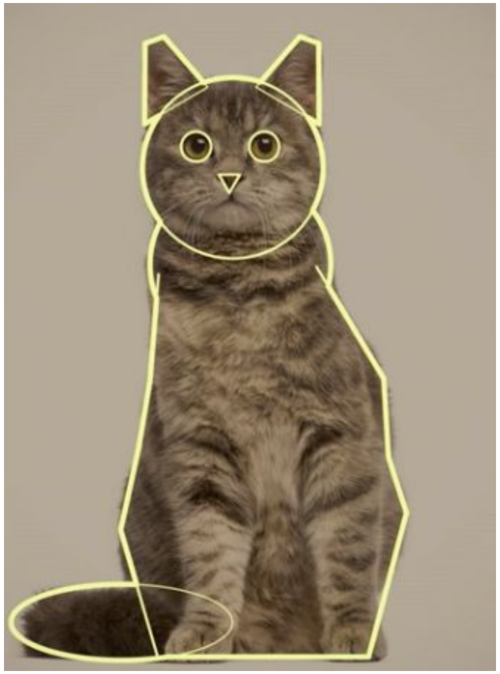
16	08	67	15	83	09
37	52	77	23	22	74
35	42	48	72	85	27
68	36	43	54	21	33
79	60	10	25	54	71
18	55	38	73	50	47

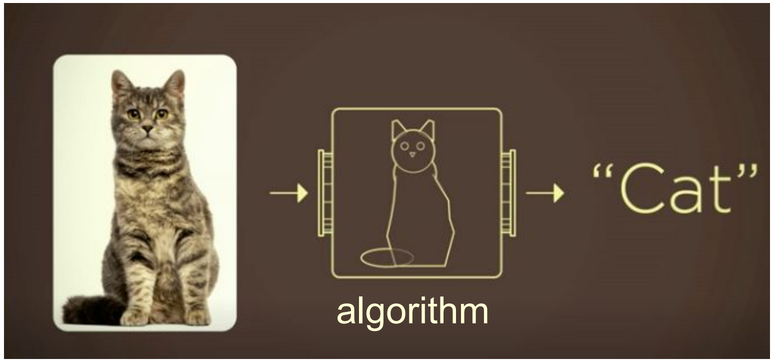
Challenge in Computer Vision: Find a Cat

SLAC

Development Workflow for non-ML algorithms

1. Write an algorithm based on some principles





A cat = collection of certain shapes
(or, a neutrino)

Convolutional Neural Networks

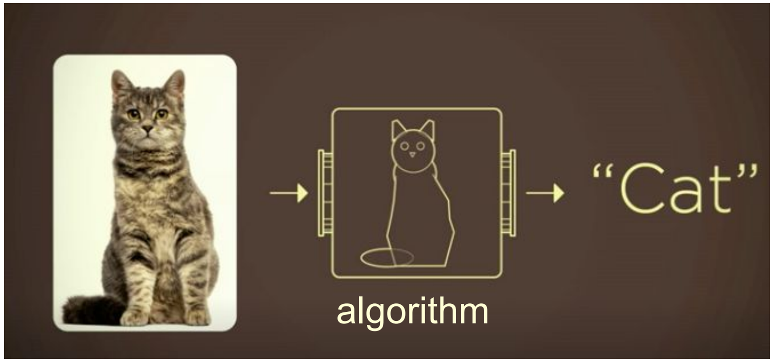
Challenge in Computer Vision: Find a Cat

Development Workflow for non-ML algorithms

1. Write an algorithm based on some principles
2. Run on data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

Partial cat
(escaping the detector)

Stretching cat
(Nuclear FSI)



A cat = collection of certain shapes
(or, a neutrino)

Challenge in Computer Vision: Find a Cat

SLAC

Development Workflow for non-ML algorithms

1. Write an algorithm based on some principles
2. Run on data samples
3. Observe failure cases, implement fixes/heuristics
4. Iterate over 2 & 3 till a satisfactory level is achieved
5. Chain multiple algorithms as one algorithm, repeat 2, 3, and 4.

Machine Learning

- **"Learn patterns from data"**
  - automation of steps 2, 3, and 4"
- **"Chain algorithms & optimize"**
  - step 5 addressed by design
- **"Deep Neural Network"**
  - de-facto solutio in computer vision

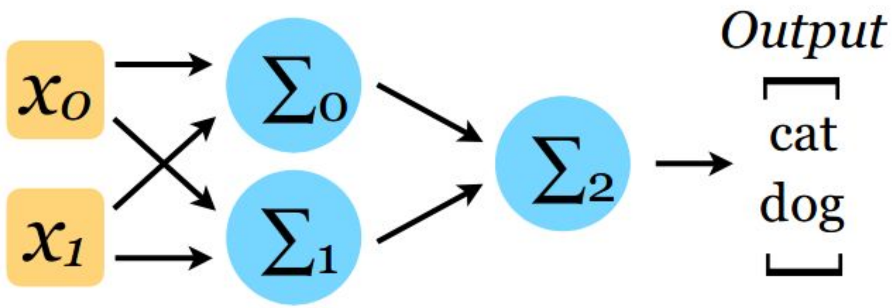
Image Analysis Using Neural Networks

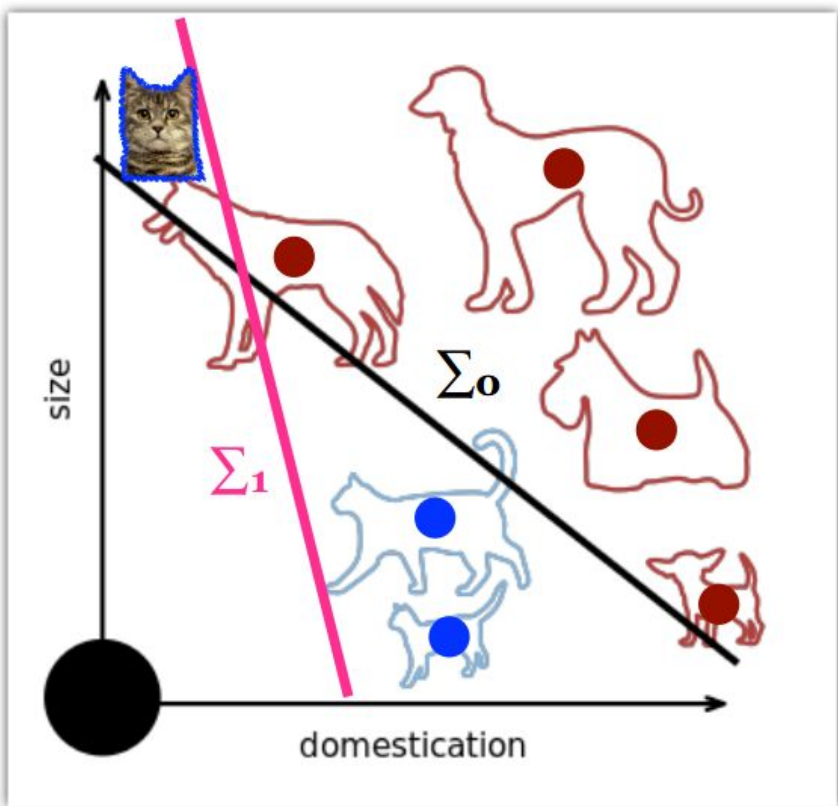
U.S. DEPARTMENT OF ENERGY
Office of Science

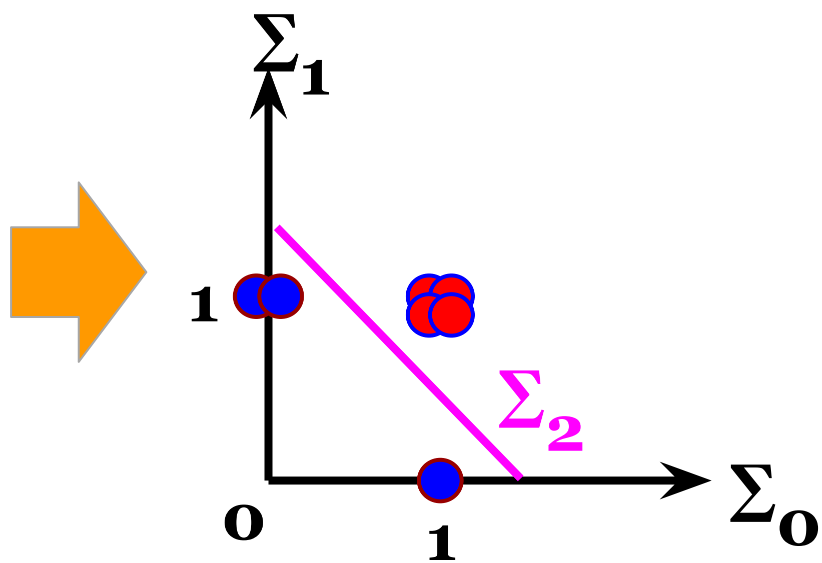
SLAC NATIONAL ACCELERATOR LABORATORY

Convolutional Neural Networks
Challenge in Computer Vision: Find a Cat

From yesterday…



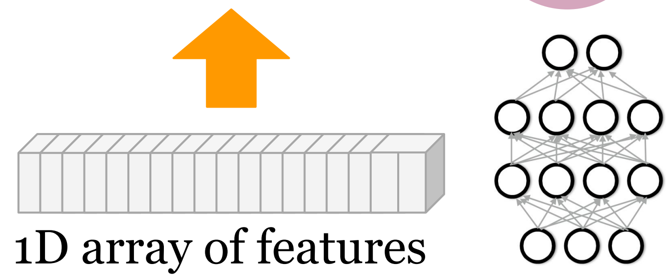


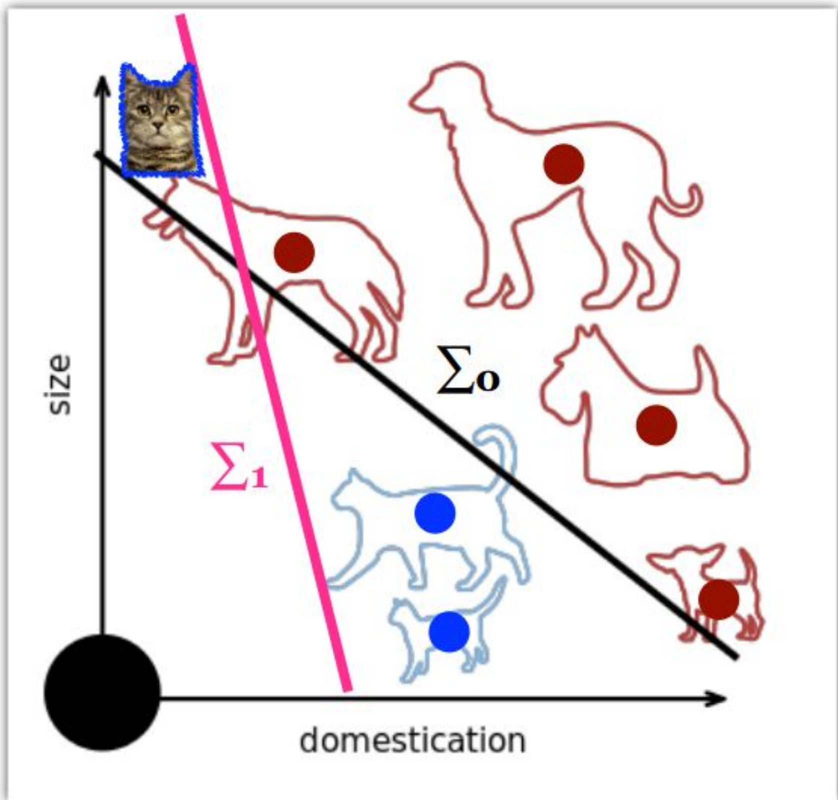


Challenge in Computer Vision: Find a Cat

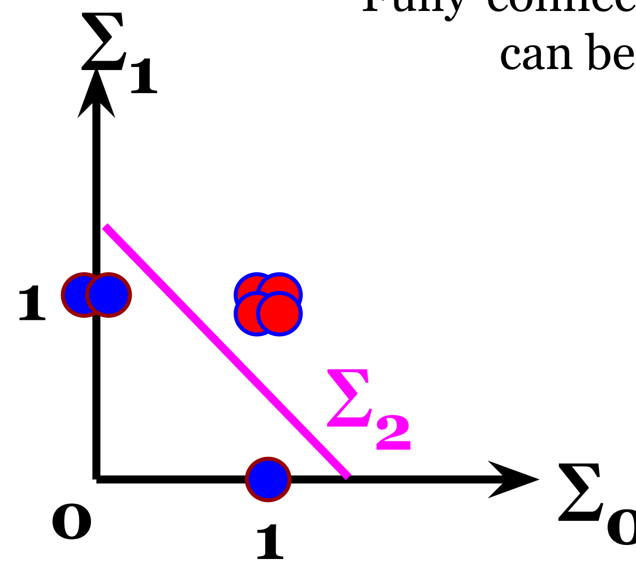
From yesterday…

Goal: Dog or Cat ?





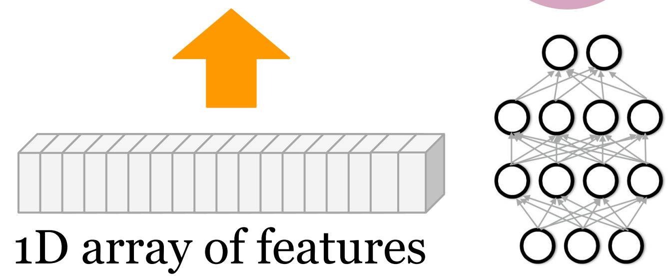
Fully-connected NN can be useful.

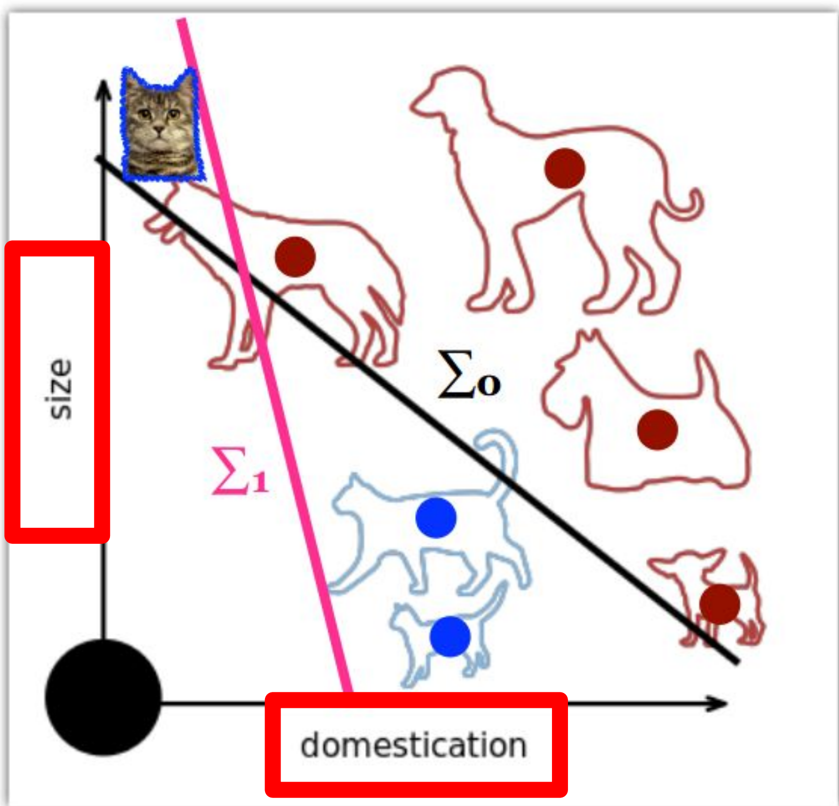


Convolutional Neural Networks
Challenge in Computer Vision: Find a Cat

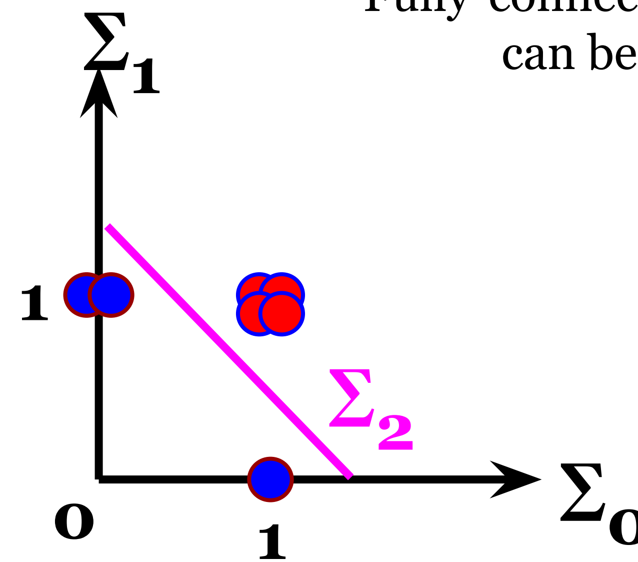
From yesterday…

Goal: Dog or Cat ?





Fully-connected NN can be useful.

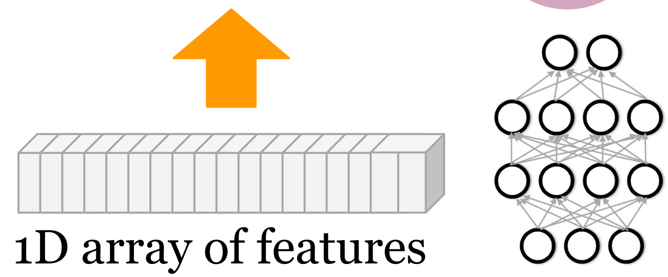


Convolutional Neural Networks
Challenge in Computer Vision: Find a Cat

For today!

How?

Goal: Dog or Cat ?



Fully-connected NN
can be useful.

How can we extract
"features" from image?

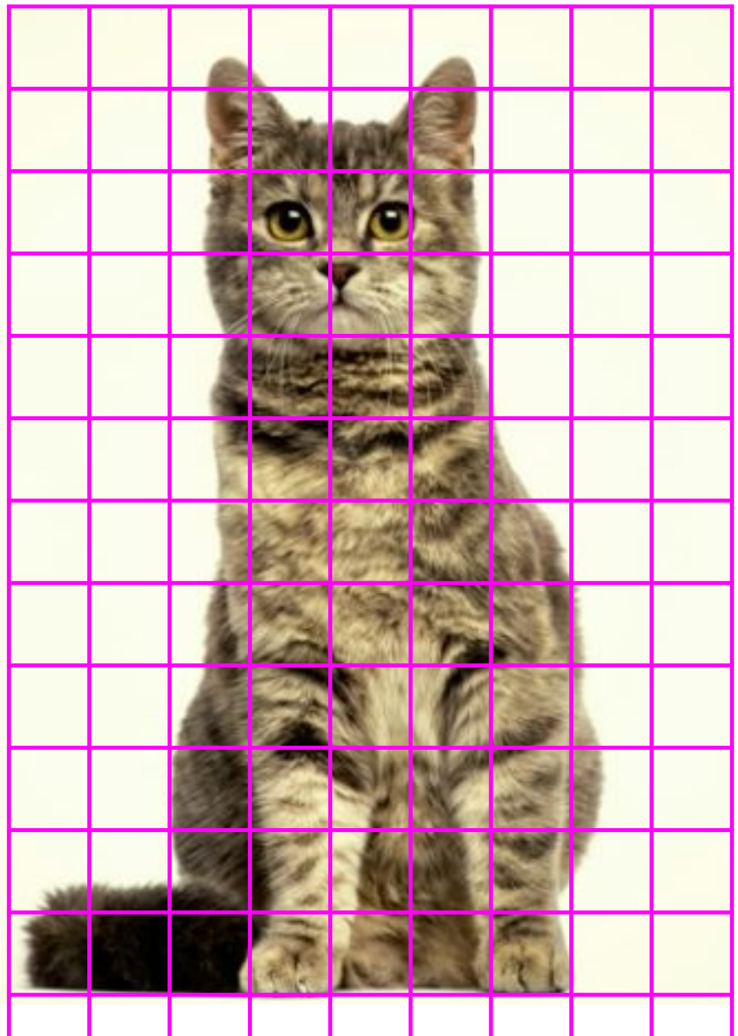
Fully-connected NN?

Convolutional Neural Networks

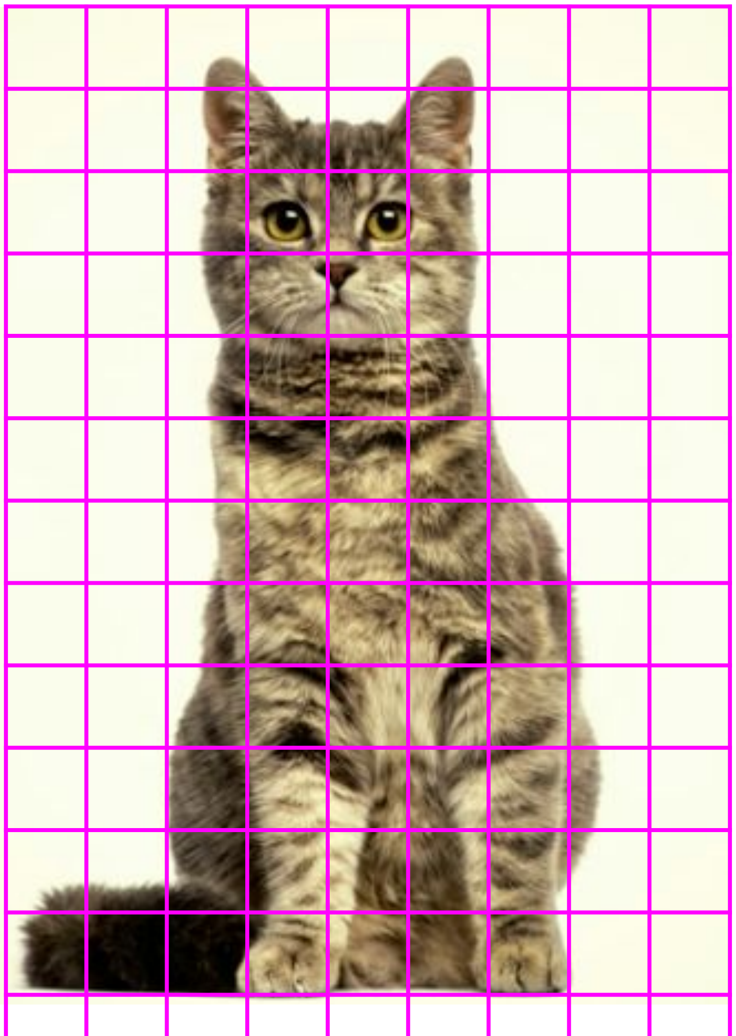
Challenge in Computer Vision: Find a Cat

SLAC

How about flattened image + MLP?



Convolutional Neural Networks
Challenge in Computer Vision: Find a Cat

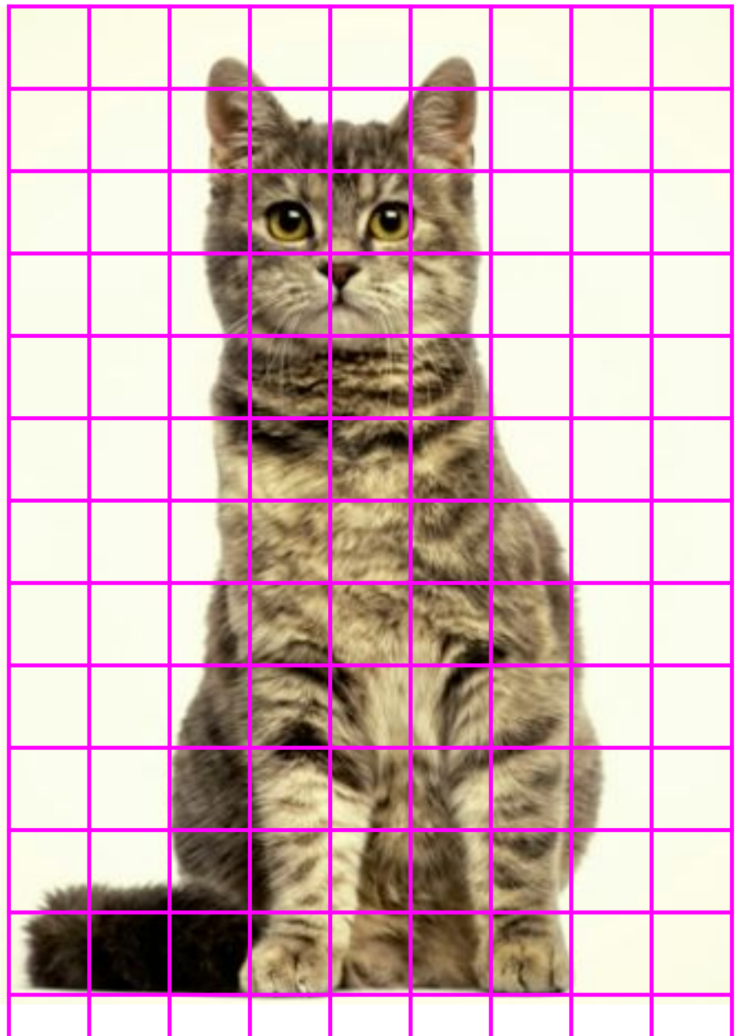


How about flattened image + MLP?

- For an input image of 100x100 pixels RGB image, how many weights does 1 neuron carry?

Convolutional Neural Networks

Challenge in Computer Vision: Find a Cat

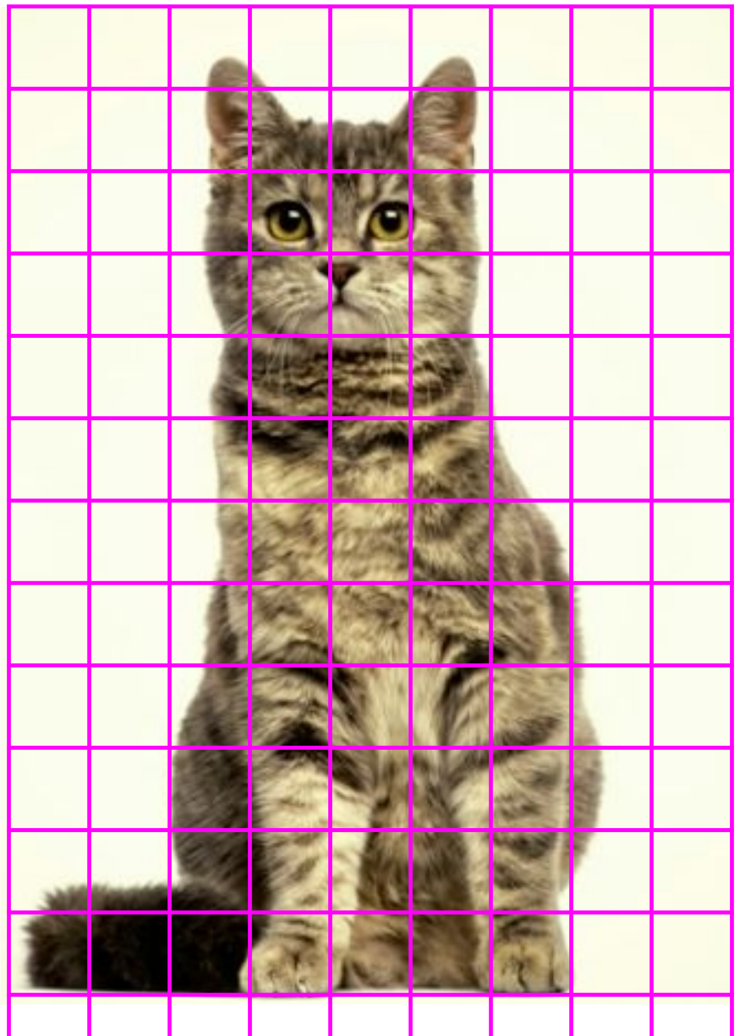


How about flattened image + MLP?

- For an input image of 100x100 pixels RGB image, how many weights does 1 neuron carry? 30,000 for just 1 neuron!

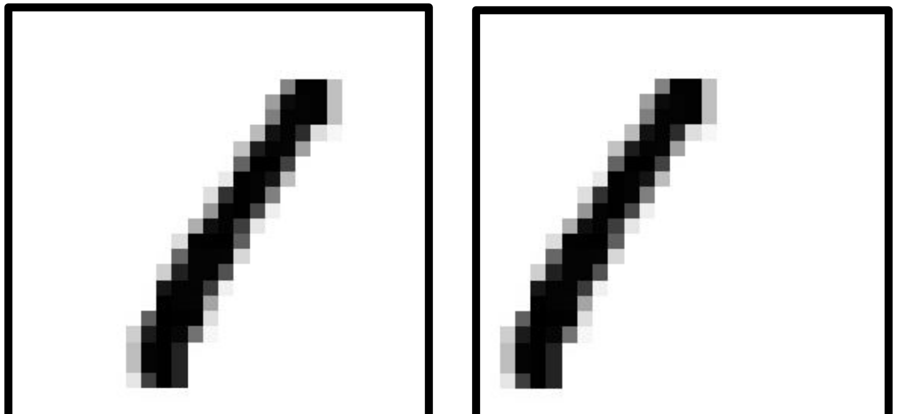
Convolutional Neural Networks

Challenge in Computer Vision: Find a Cat



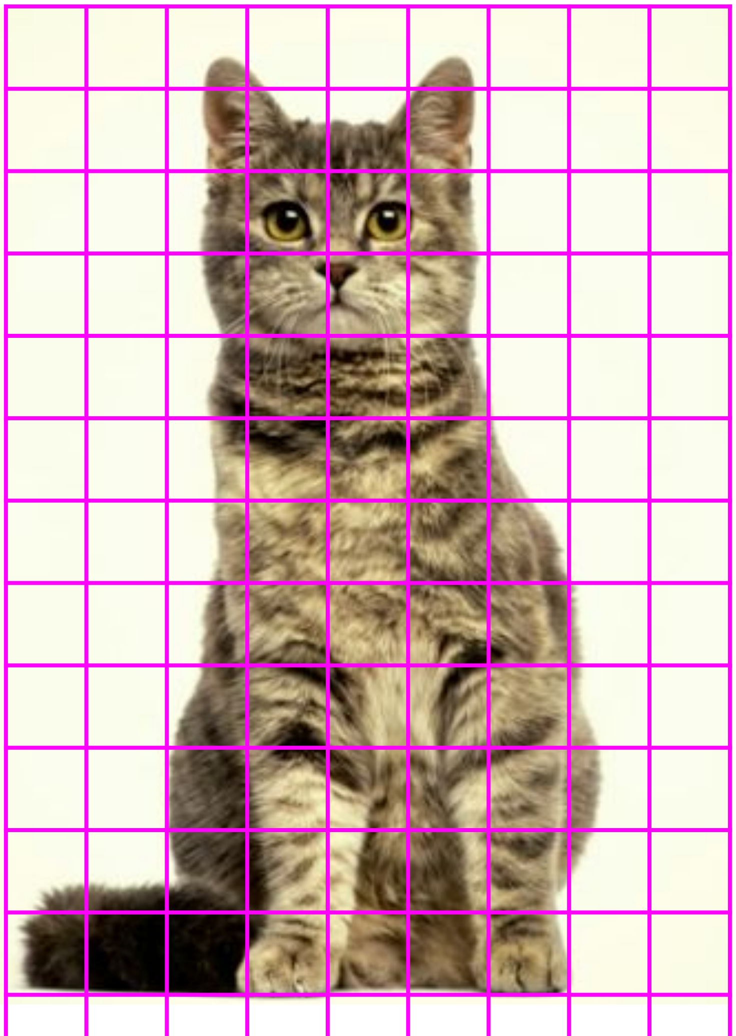
How about flattened image + MLP?

- For an input image of 100x100 pixels RGB image, how many weights does 1 neuron carry? 30,000 for just 1 neuron!
- Two image of the same cat, but in a different position w.r.t. the frame. Would neuron react the same?



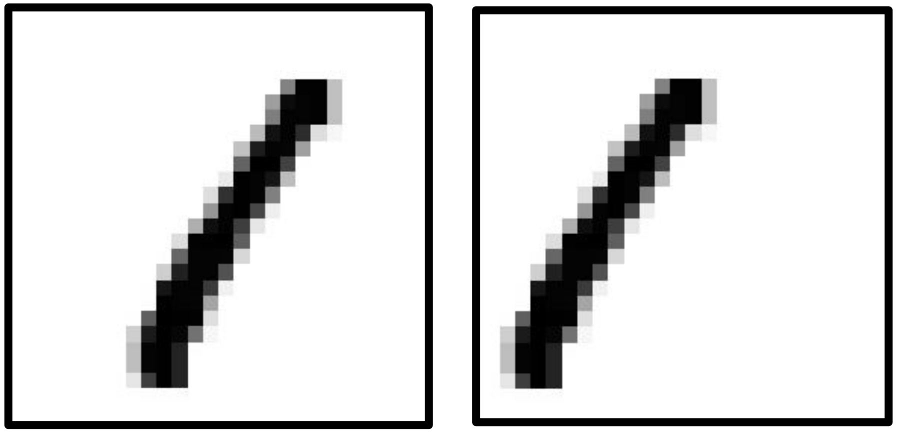
Convolutional Neural Networks

Challenge in Computer Vision: Find a Cat



How about flattened image + MLP?

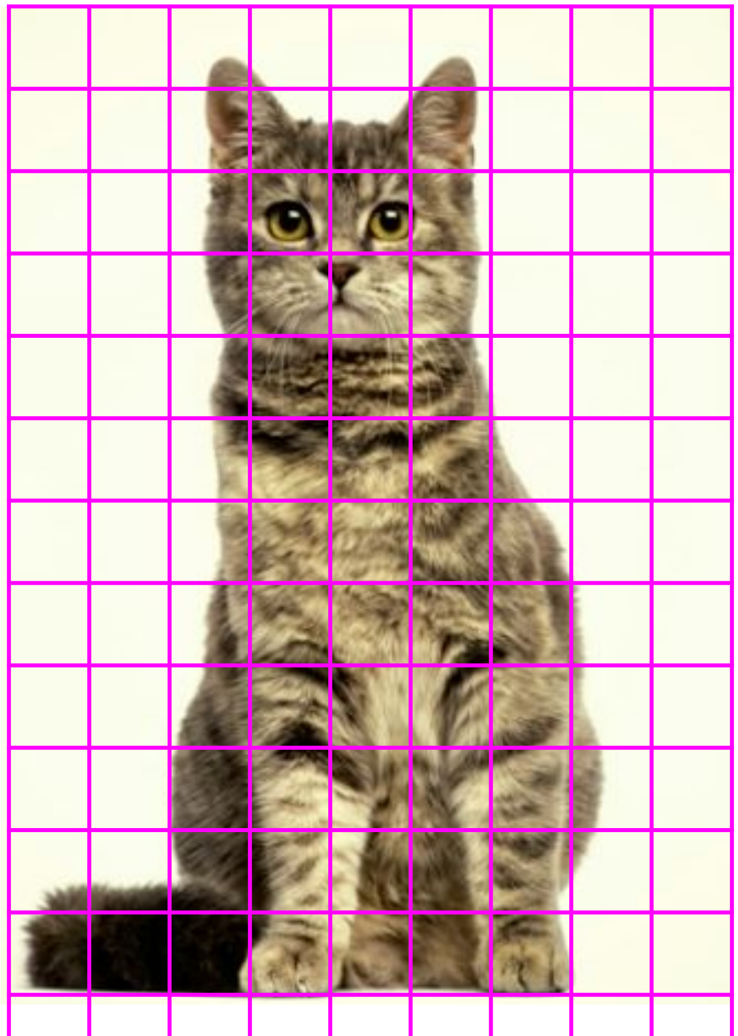
- For an input image of 100x100 pixels RGB image, how many weights does 1 neuron carry? 30,000 for just 1 neuron!
- Two image of the same cat, but in a different position w.r.t. the frame. Would neuron react the same? No! Position information is encoded!



Convolutional Neural Networks

Challenge in Computer Vision: Find a Cat

How about flattened image + MLP?



- For an input image of 100x100 pixels RGB image, how many weights does 1 neuron carry? 30,000 for just 1 neuron!
- Two image of the same cat, but in a different position w.r.t. the frame. Would neuron react the same?
  No! Position information is encoded!

solution?

Convolutional NN!

Convolutional Neural Networks

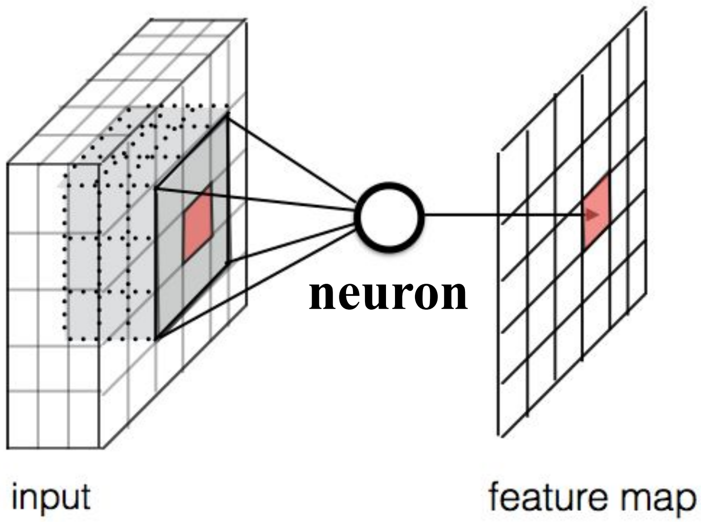
U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Convolutional Neural Networks

How does it work?

CNNs introduce a ***limitation to MLP*** by forcing a neuron to look at only local, translation invariant features



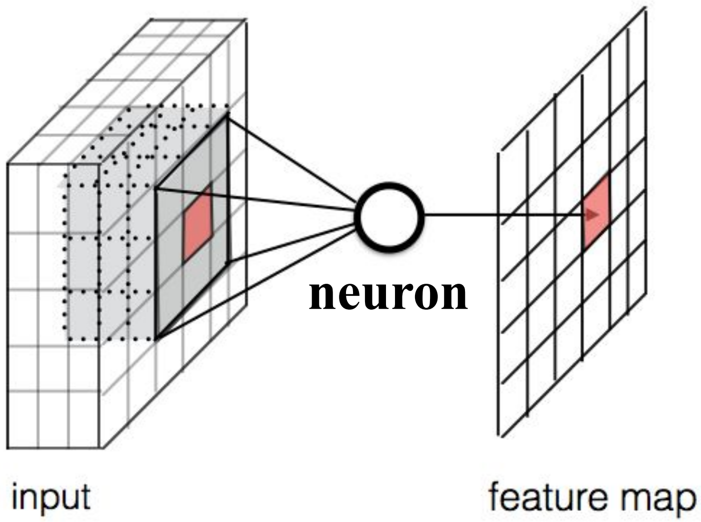
$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

Still a linear model!
Weights=matrix, output=scalar
Analyze a fixed-size, local sub-matrix from the input.

Convolutional Neural Networks

How does it work?

CNNs introduce a ***limitation to MLP*** by forcing a neuron to look at only local, translation invariant features



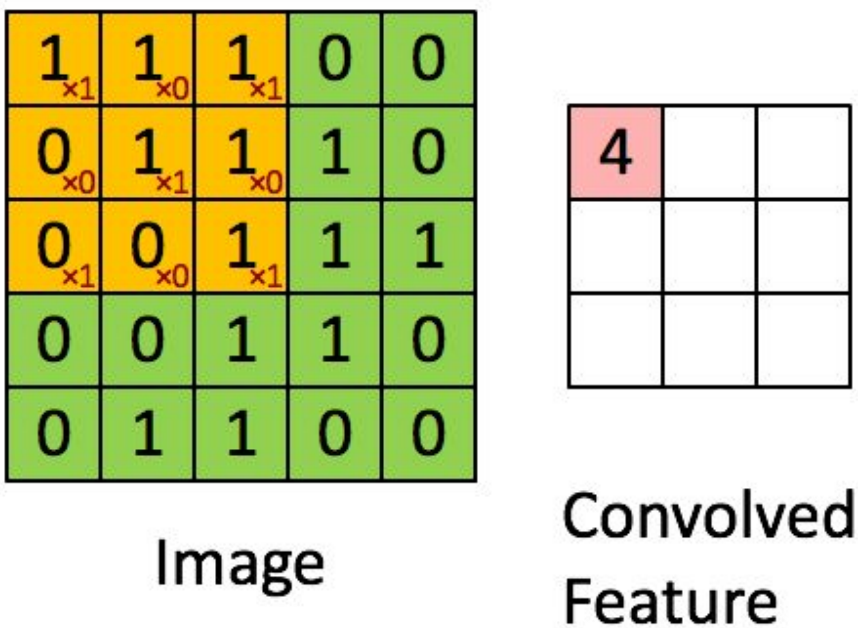
$$f_{i,j}(X) = \sigma\left(W_i \cdot X_j + b_i\right),$$

Still a linear model!
Weights=matrix, output=scalar
Analyze a fixed-size, local sub-matrix from the input.

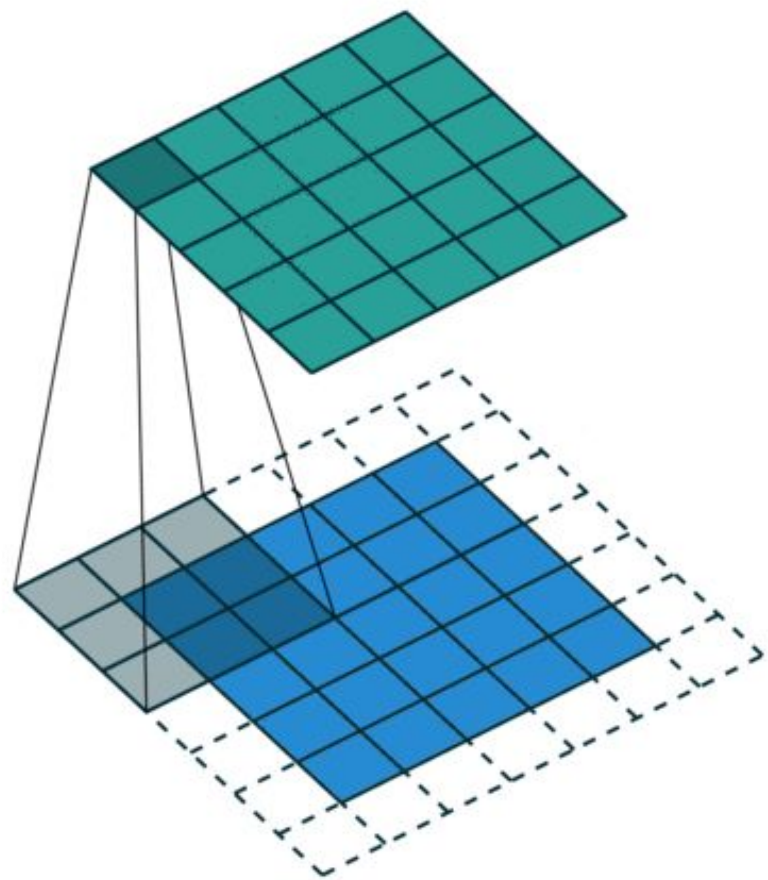
- Traverse over 2D space to process the whole input
- Neuron **learns translation-invariant features**

Convolutional Neural Networks
How does it work?

Convolution 3x3
Stride 1, no padding



Convolution 3x3
Stride 1, padding 1



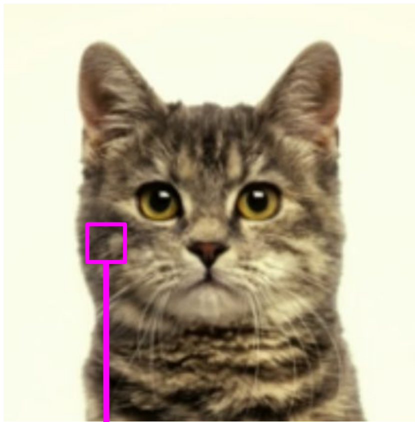
Convolutional Neural Networks

How does it work?

Goal: Dog or Cat ?

Convolutional Neural Networks
How does it work?

Goal: Dog or Cat ?



Convolution
3 x 3 “**kernel**”

0	1	0
0	2	0
0	1	0

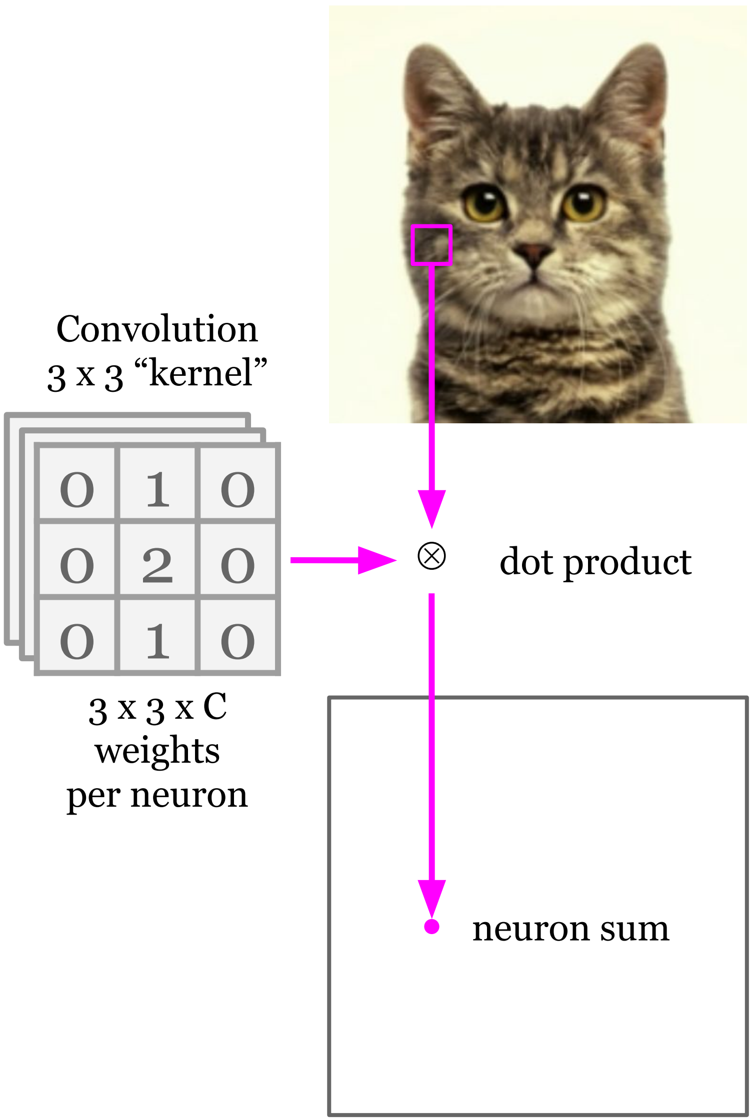
3 x 3 x C
weights
per neuron

$\otimes$ dot product

Convolutional Neural Networks

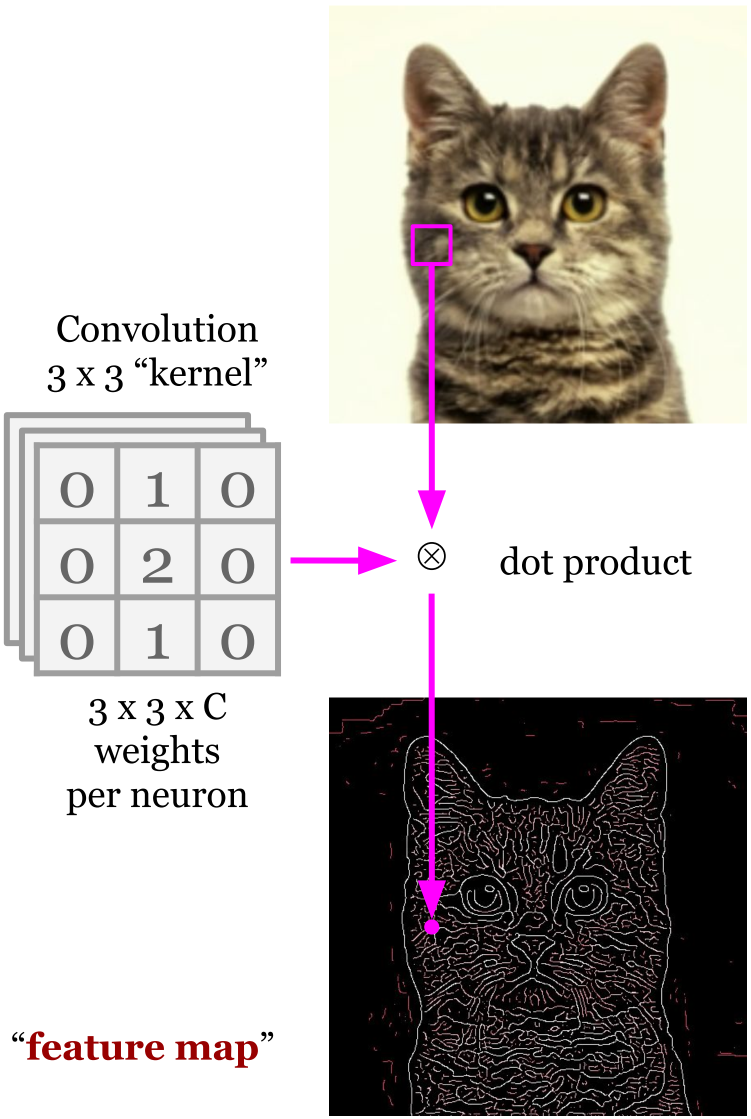
How does it work?

Goal: Dog or Cat ?



Convolutional Neural Networks
How does it work?

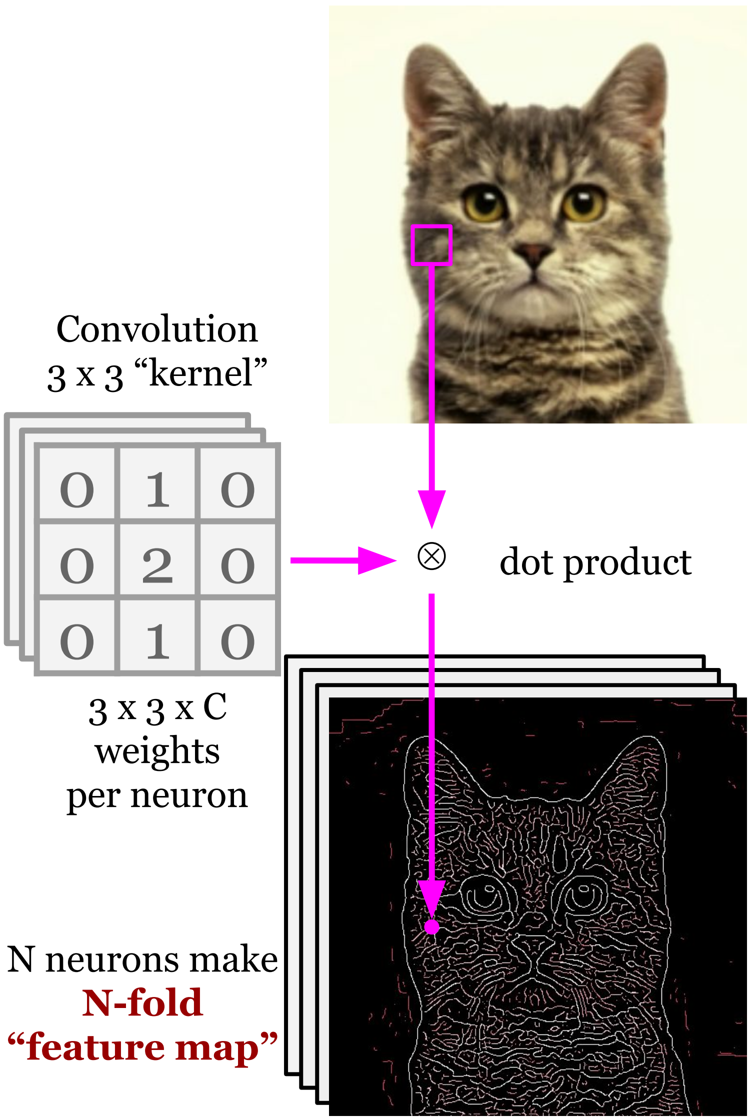
Goal: Dog or Cat ?



Convolutional Neural Networks

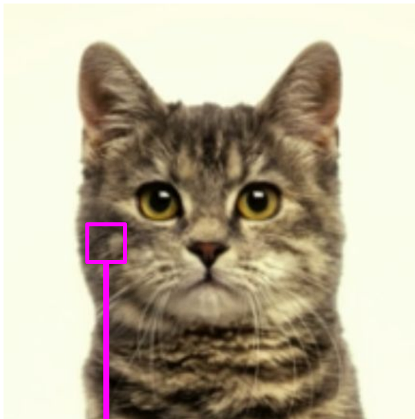
How does it work?

Goal: Dog or Cat ?

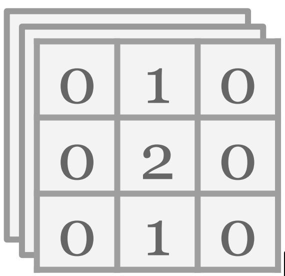


Convolutional Neural Networks
How does it work?

Goal: Dog or Cat ?

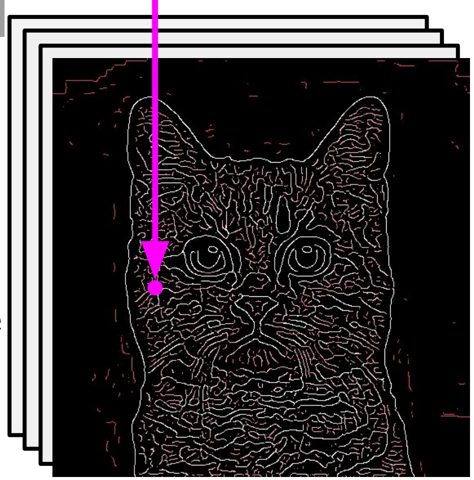


Convolution
3 x 3 "kernel"



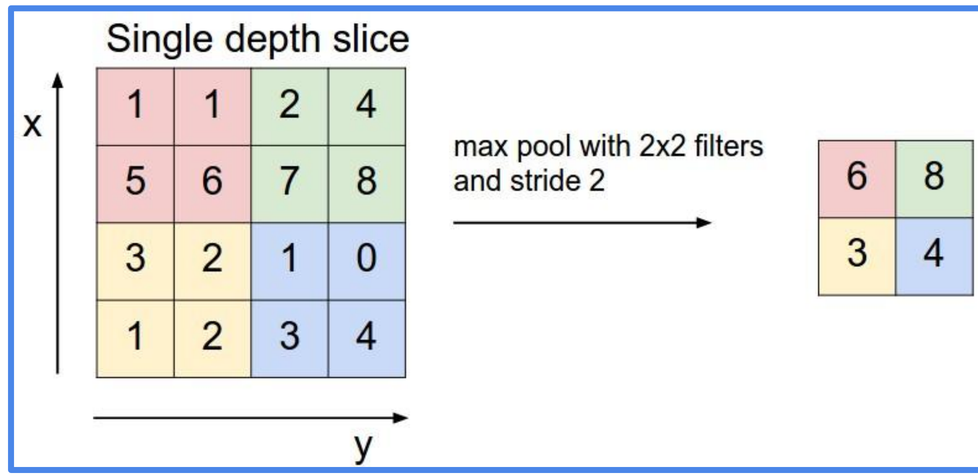
⊗ dot product

3 x 3 x C
weights
per neuron



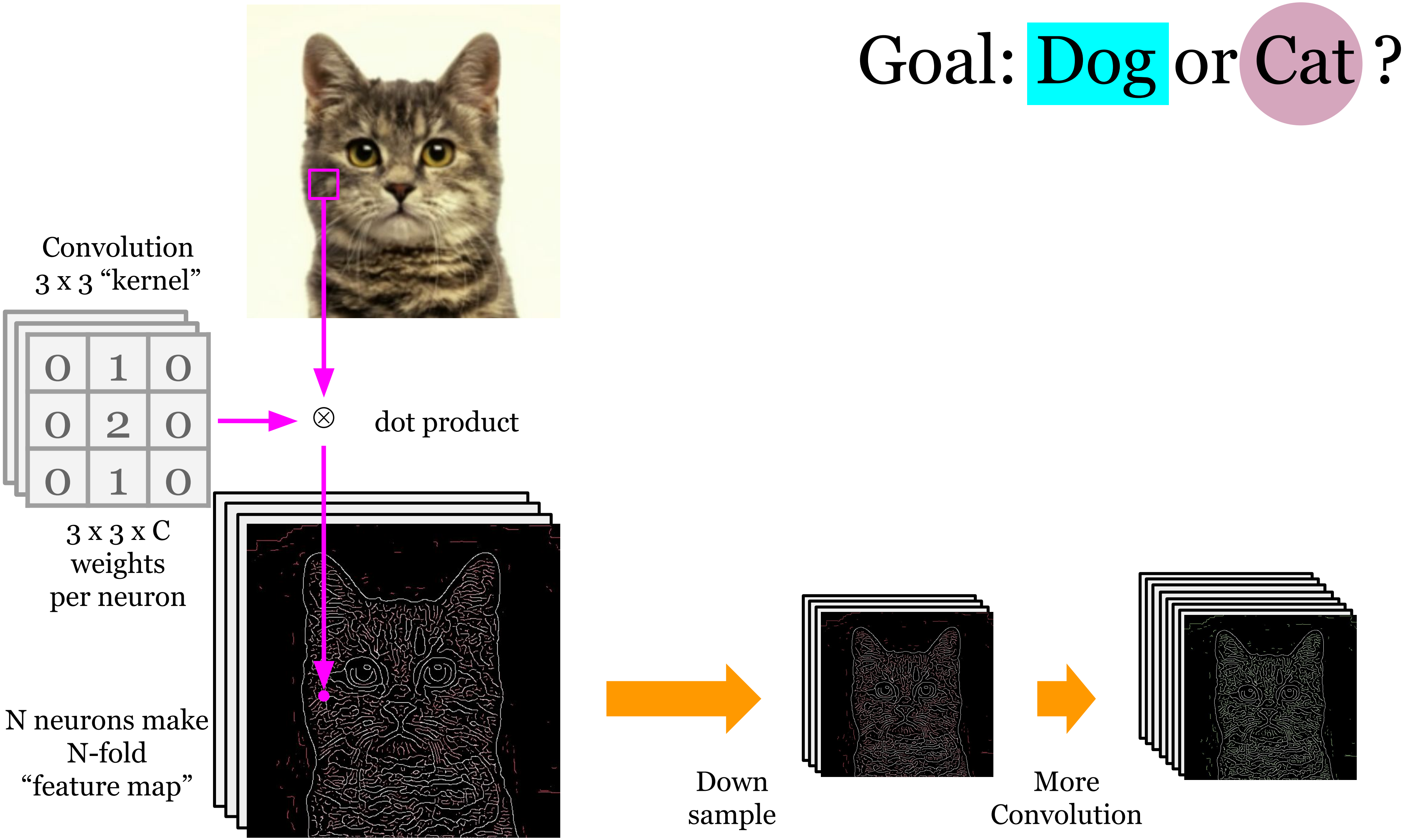
N neurons make
N-fold
"feature map"

e.g) max pooling



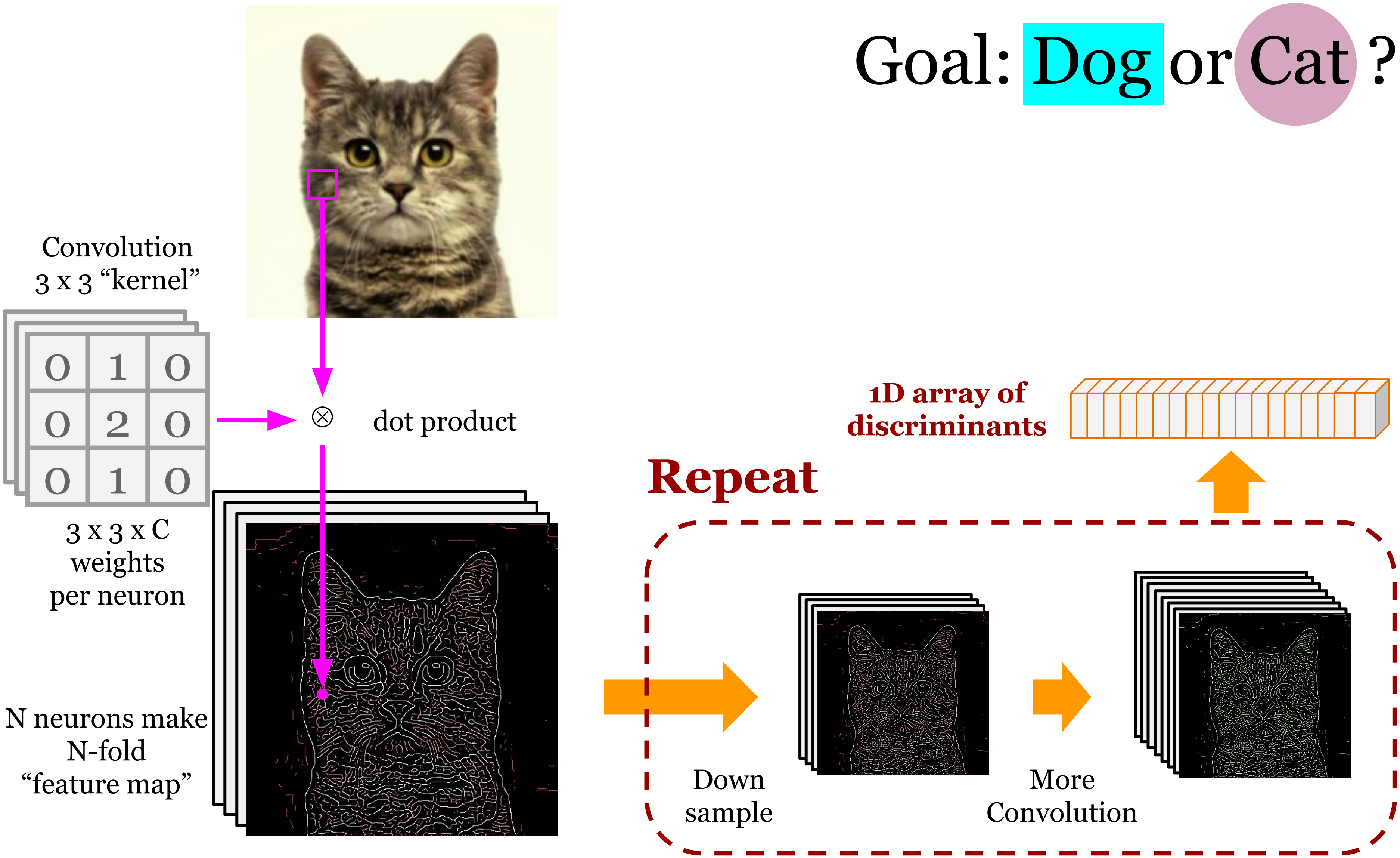
Down sample

Convolutional Neural Networks
How does it work?

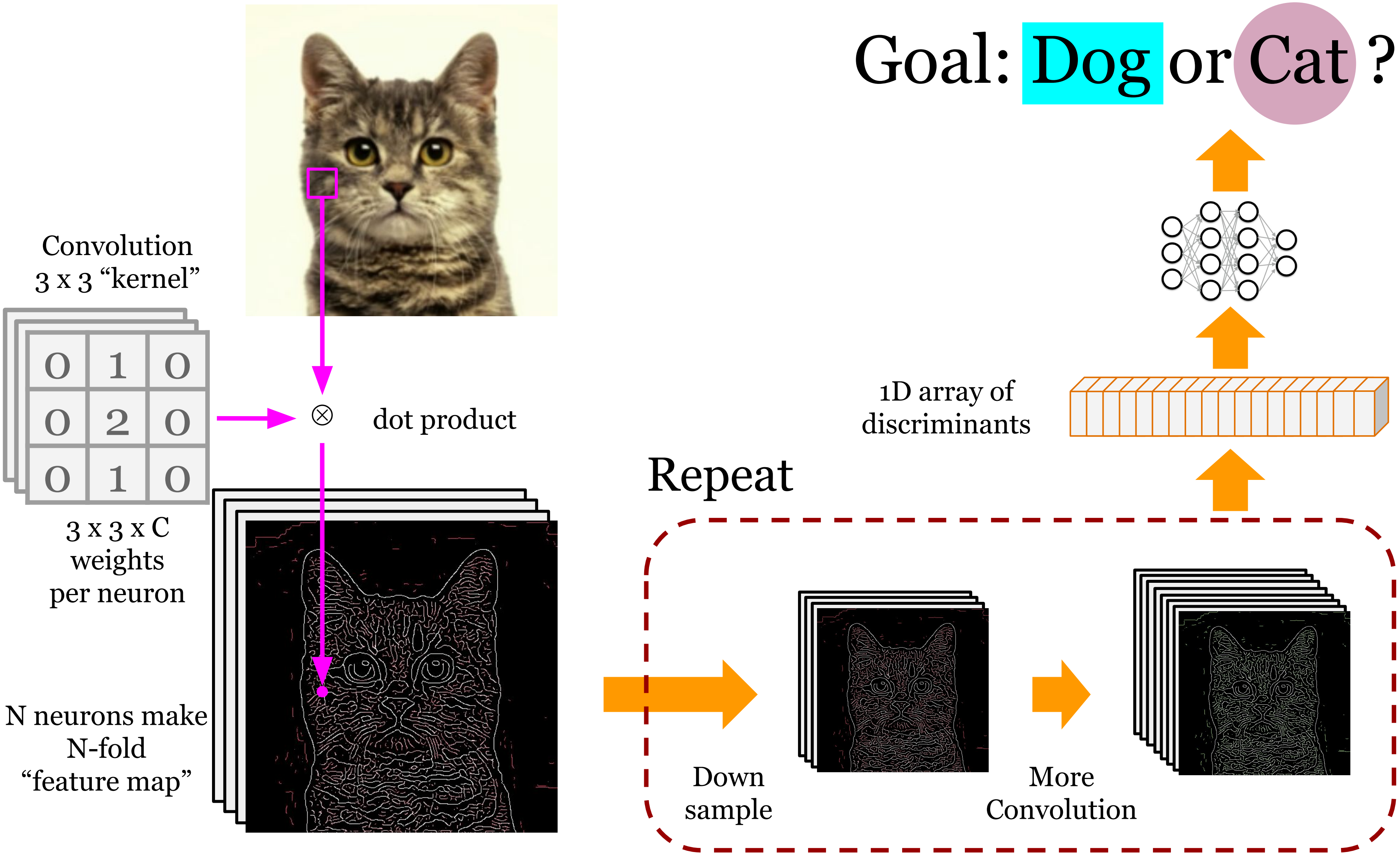


Convolutional Neural Networks

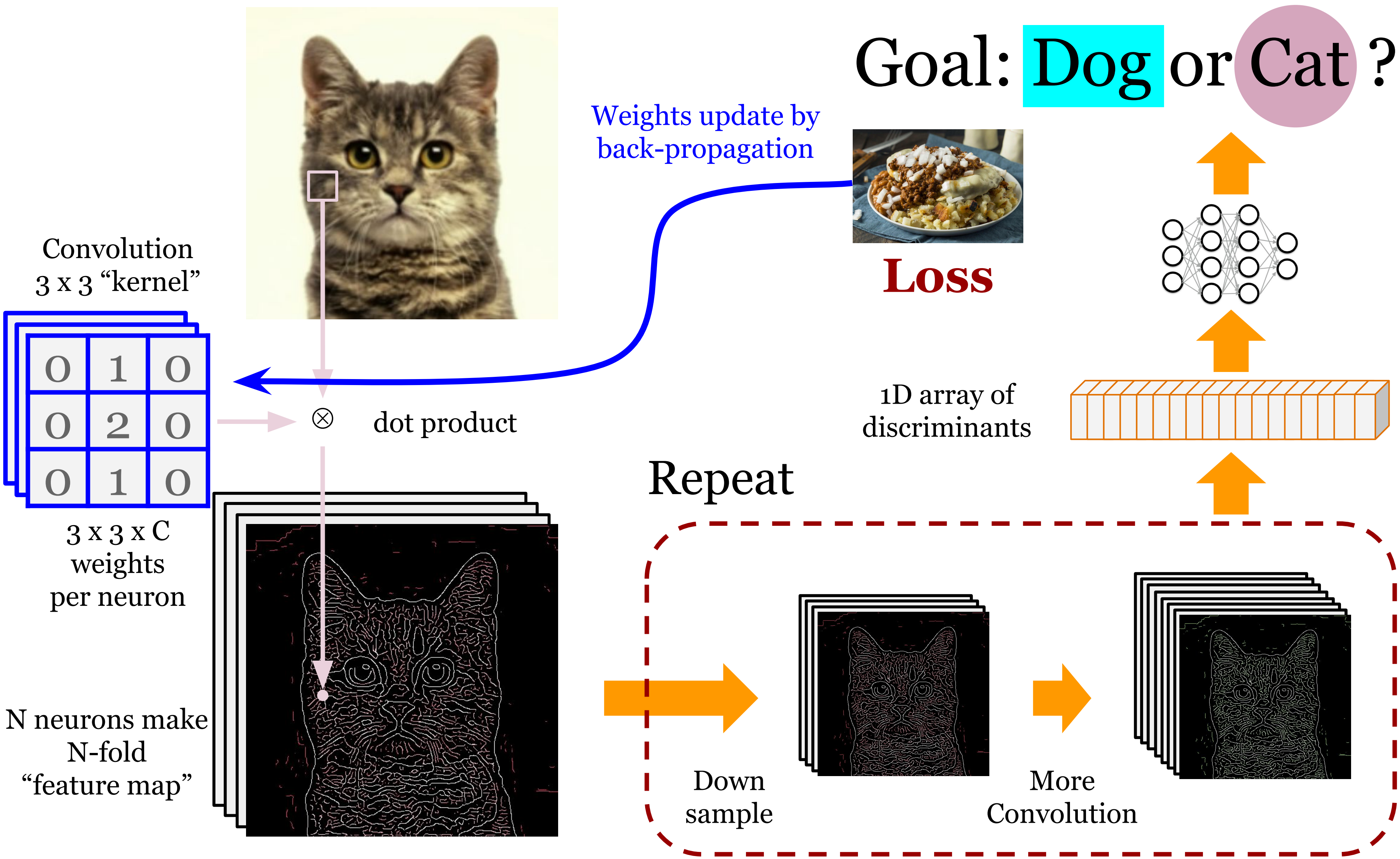
How does it work?



Convolutional Neural Networks
How does it work?

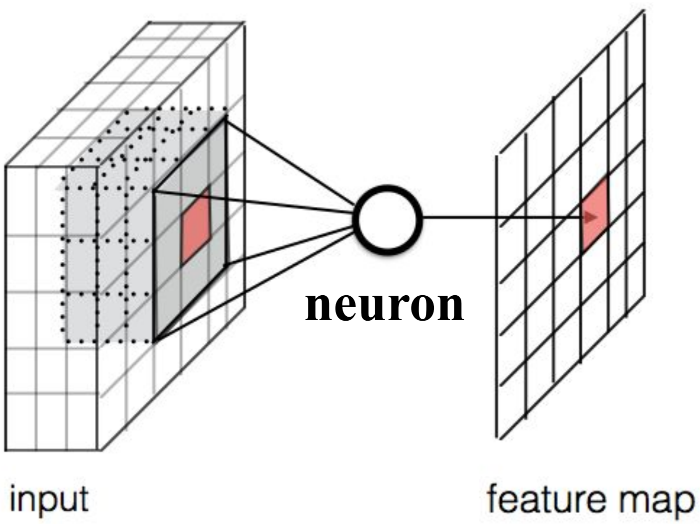


Convolutional Neural Networks
How does it work?

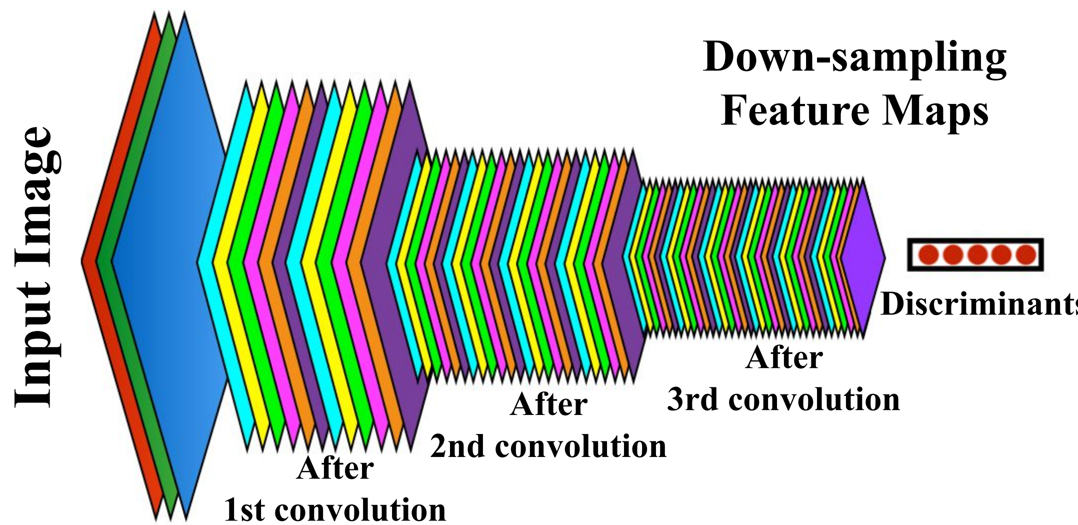


Convolutional Neural Networks
How does it work?

- **CNNs** are "**feature extraction machine**"
  - Consists of a "convolution layer" with "kernels"
  - A chain of parallelizable linear algebra operations
- **CNN** seen as a **geometrical data transformer**



Convolution Operation



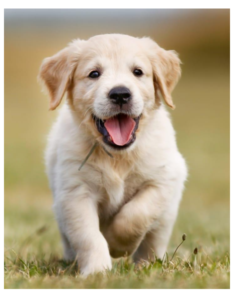
Series of convolutions + down-sampling

Convolutional Neural Networks

How does it work?

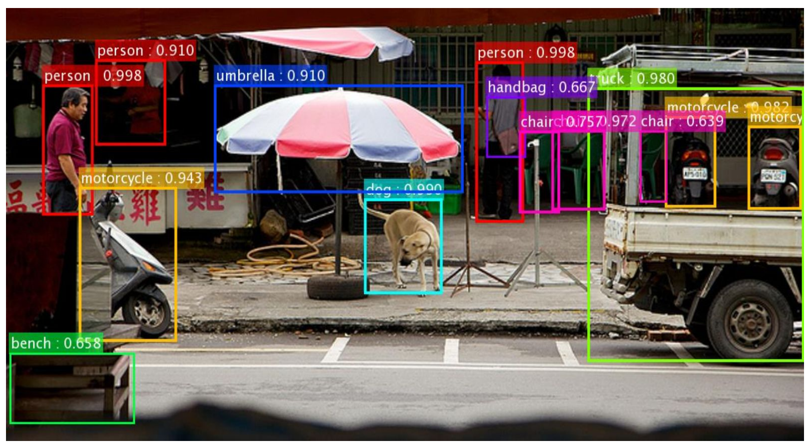
- **CNNs** are “**feature extraction machine**”
  - Consists of a “convolution layer” with “kernels”
  - A chain of parallelizable linear algebra operations
- **CNN** seen as a **geometrical data transformer**

Covered

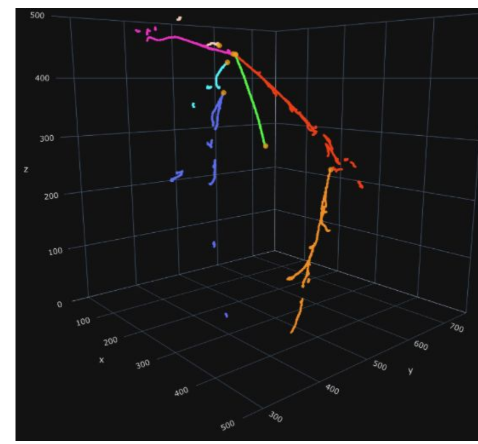


Cool v.s. Adorable

Later in this lecture



Object Detection



Pixel Classification

How to compress/extract spatial information ... depends on applications!
Study of **CNN “architecture designs”**