Convolutional Neural Networks

CPSC 383: Explorations in Artificial Intelligence and Machine Learning
Fall 2025

Jonathan Hudson, Ph.D.
Associate Professor (Teaching)
Department of Computer Science
University of Calgary

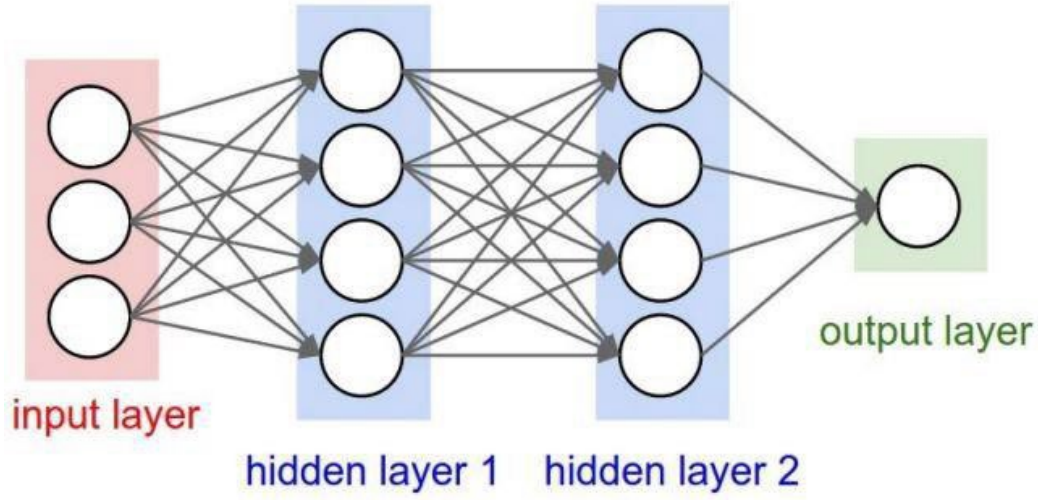
August 27, 2025

UNIVERSITY OF CALGARY

Basics

So far, we have designed and discussed a **fully connected feed-forward network**.

It is fully connected, as every neuron in the k^{th} layer is connected to every neuron in the $(k+1)^{th}$ layer.



ImageNet

ImageNet

ImageNet is a database of labelled images, based on the WordNet taxonomy of concepts.

- https://image-net.org/index.php

It currently contains over 14,000,000 human-labeled images.

https://arstechnica.com/ai/2024/11/how-a-stubborn-computer-scientist-accidentally-launched-the-deep-learning-boom/

Organism

Animal

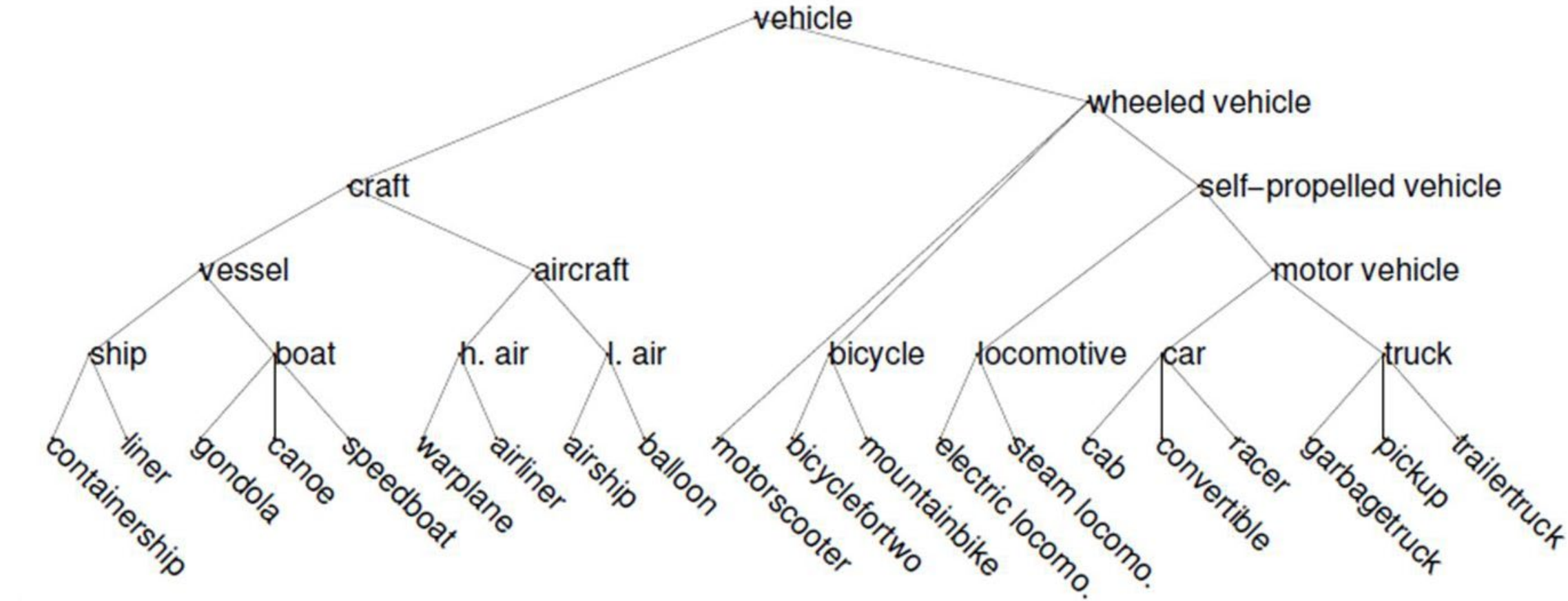
Chordate

Vertebrate

Amphibian

Frog

Tree frog

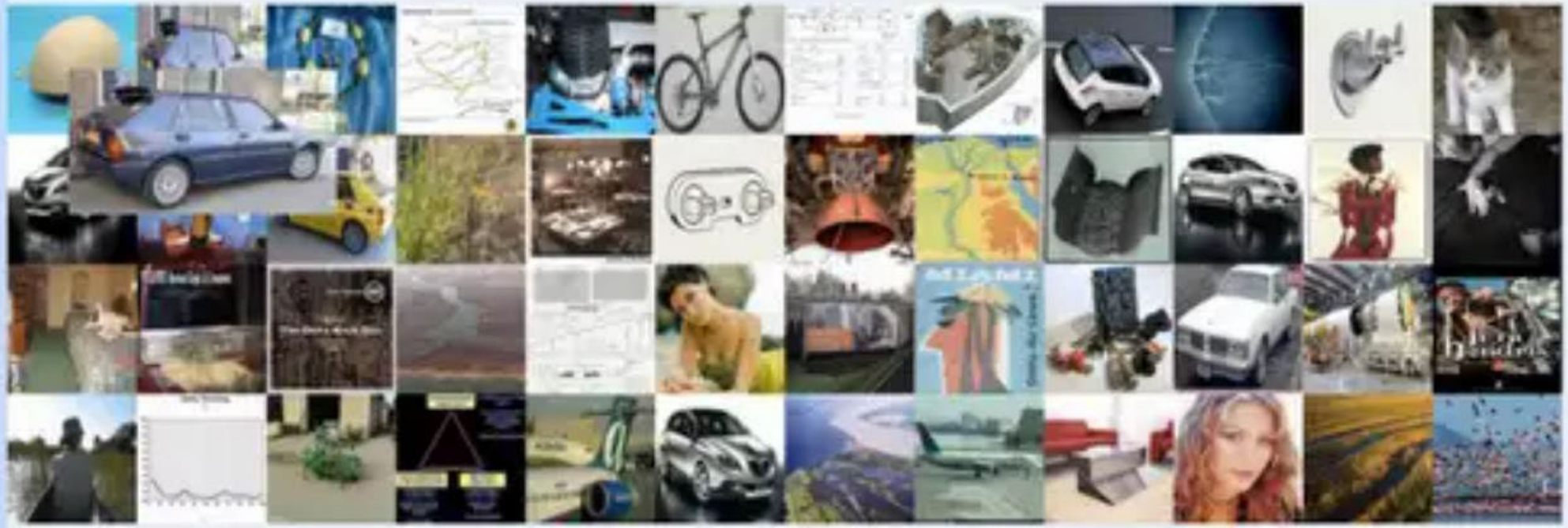


https://vision.cs.utexas.edu/projects/tom/imagenet.png

Main | Instructions | Unsure? Look up in Wikipedia | Google | [*Additional input*] No good photos? Have expertise? comments? Click here!

First time workers please click here for instructions.

Click on the photos that contain the object or depict the concept of : **delta**: **a low triangular area of alluvial deposits where a river divides before entering a larger body of water; "the Mississippi River delta"; "the Nile delta"** .(PLEASE READ DEFINITION CAREFULLY)
Pick as many as possible. ***PHOTOS ONLY, NO PAINTINGS, DRAWINGS, etc***. It's OK to have other objects, multiple instances, occlusion or text in the image.
Do not use back or forward button of your browser. OCCASIONALLY THERE MIGHT BE ADULT OR DISTURBING CONTENT.



Below are
selected
they will
to other p

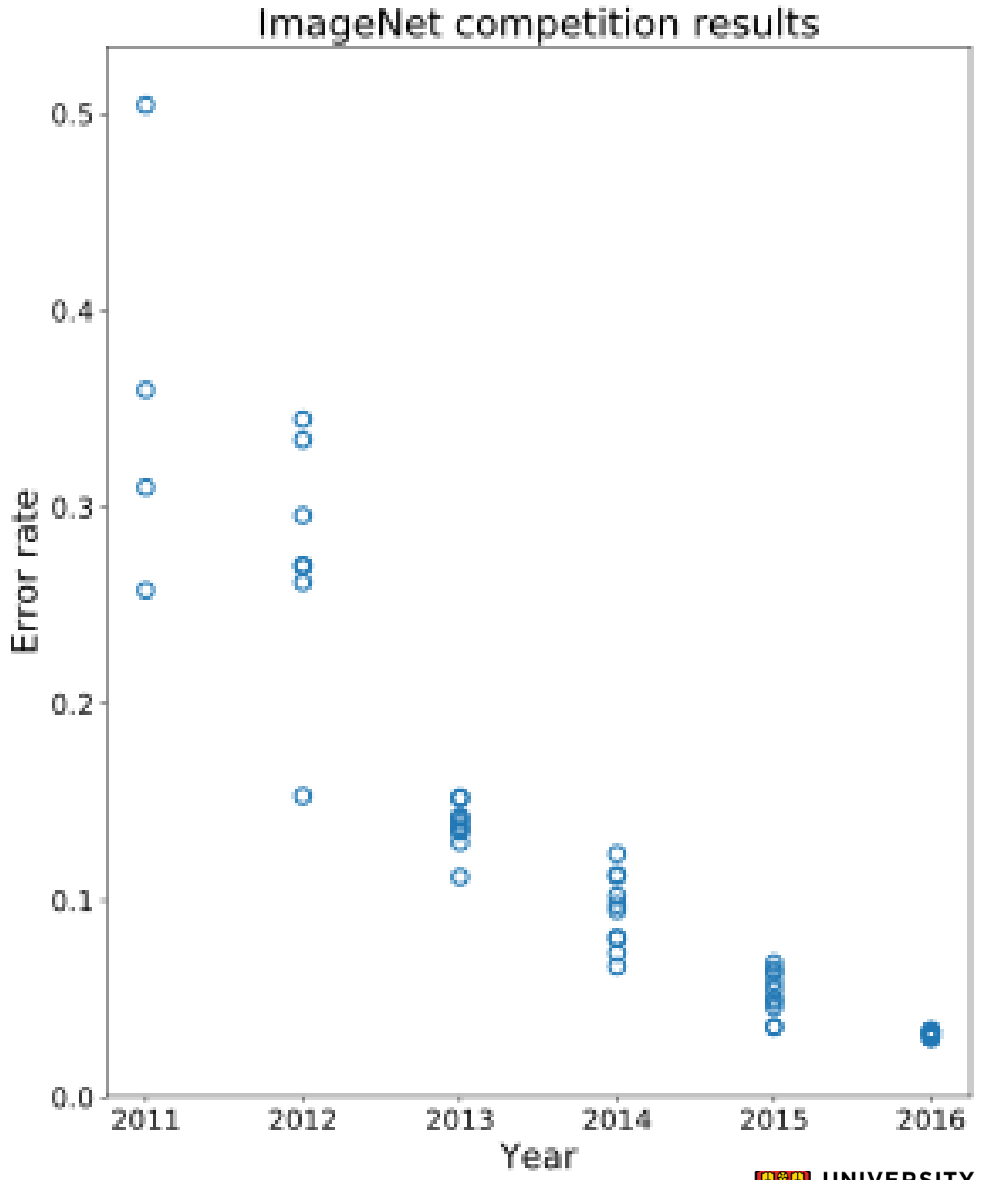
what's this? | select all | deselect all | < | page 1 of 6 | > | Submit | PREVIEW MODE. TO WOR
THIS HIT, ACCEPT IT FIR

ImageNet competition

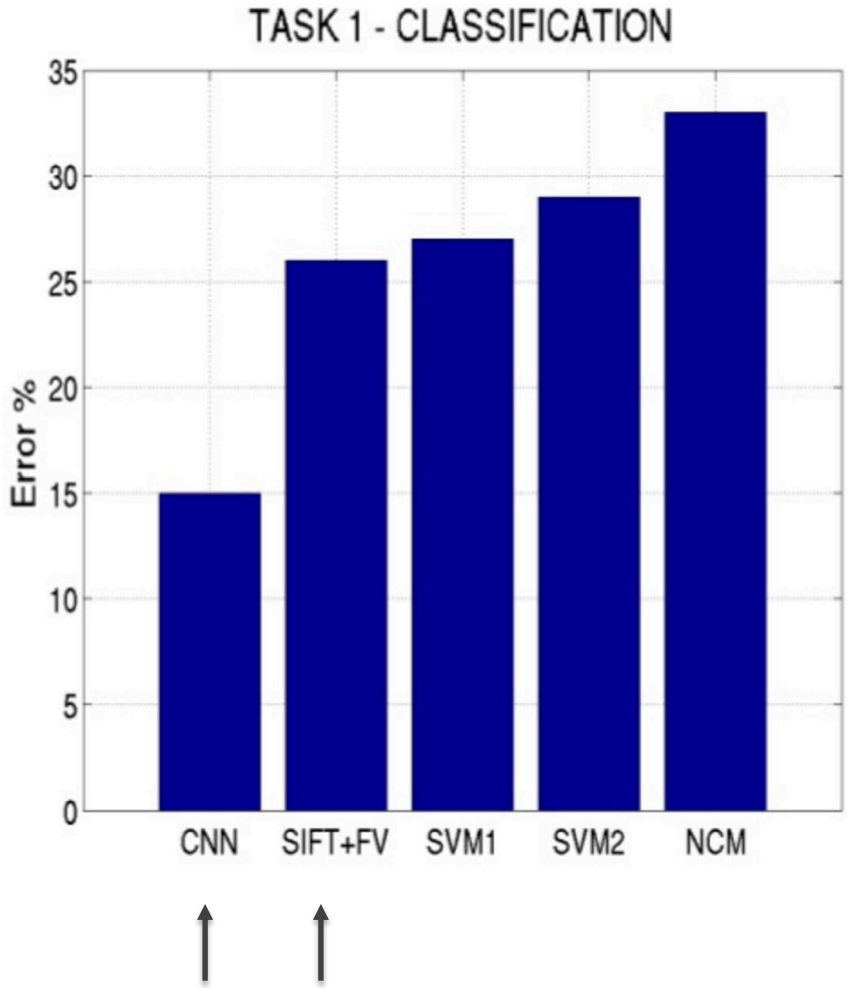
The ImageNet dataset was the basis for the ILSVRC, an annual AI competition that ran from 2010-2016.

In 2012, a neural net calls **AlexNet** created by researchers at the University of Toronto won by a large margin.

Combined CUDA on graphics cards, Convolution in deep networks, and data augmentation



Top-5 error on this competition (2012 when things changed)



conv. neural network

symbolic

mite	container ship	motor scooter	leopard
mite	container ship	motor scooter	leopard
black widow	lifeboat	go-kart	jaguar
cockroach	amphibian	moped	cheetah
tick	fireboat	bumper car	snow leopard
starfish	drilling platform	golfcart	Egyptian cat

grille	mushroom	cherry	Madagascar cat
convertible	agaric	dalmatian	squirrel monkey
grille	mushroom	grape	spider monkey
pickup	jelly fungus	elderberry	titi
beach wagon	gill fungus	ffordshire bullterrier	indri
fire engine	dead-man's-fingers	currant	howler monkey

AlexNet and Convolution

Idea

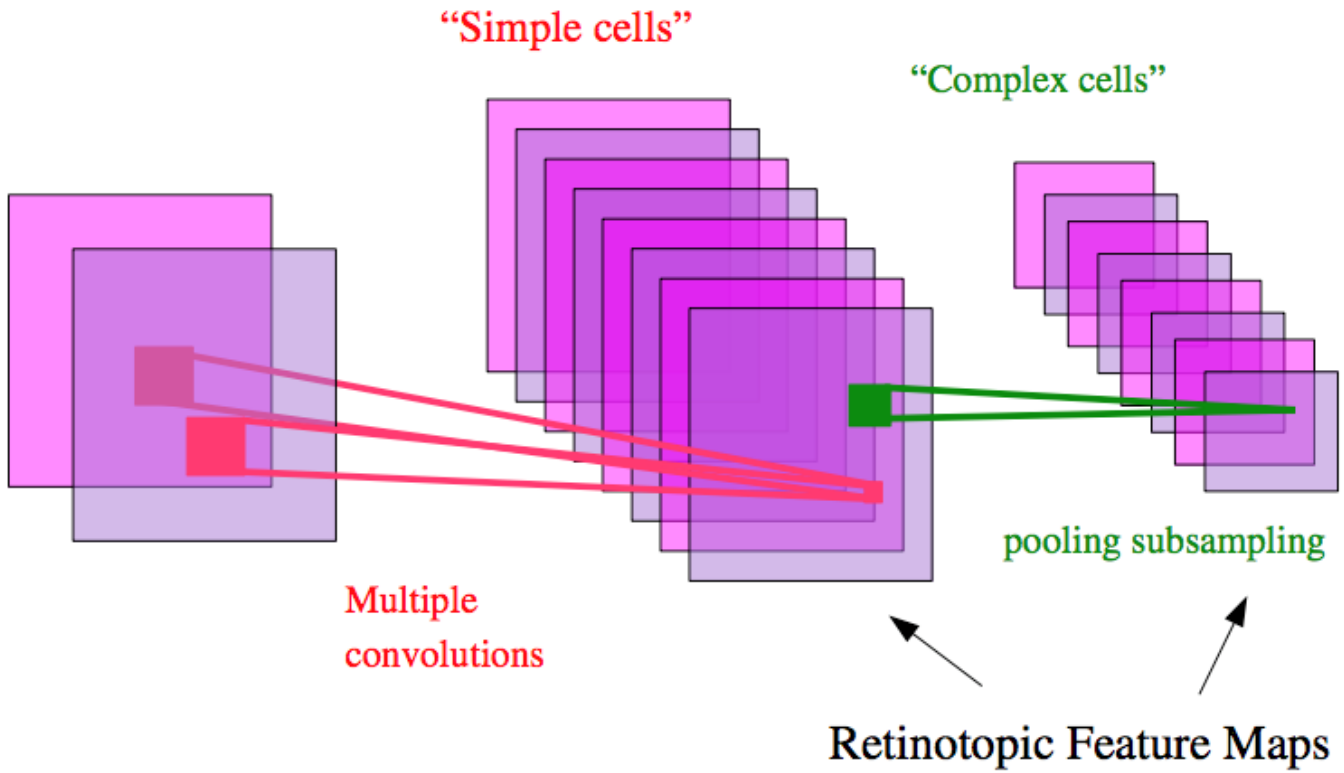
Usually if we know something about our problem, it is a good idea to build that into our network design.

- e.g. encoding data, or taking advantage of invariant properties

Convolution is a good example of this.

Model of vision in animals

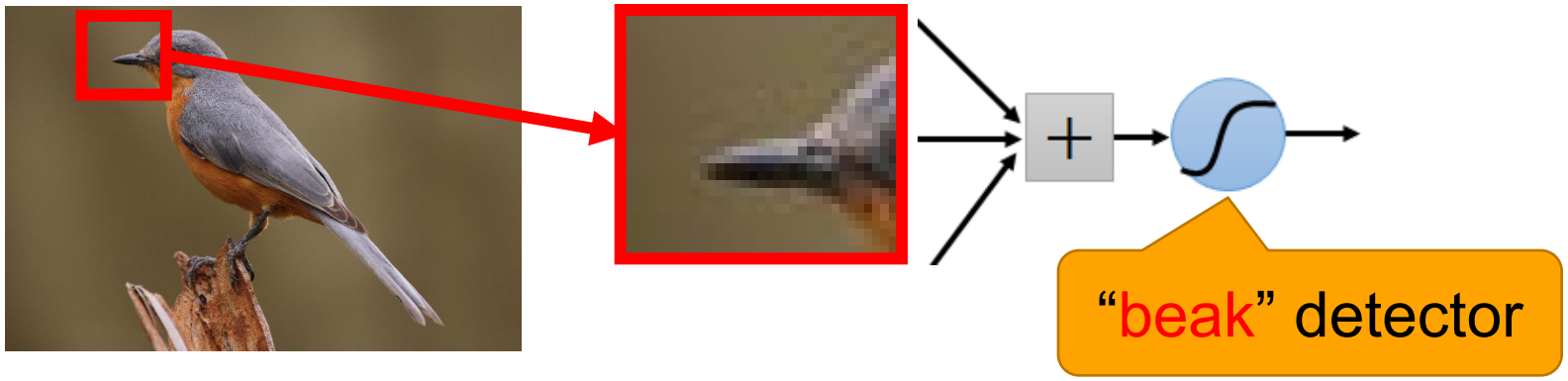
- **[Hubel & Wiesel 1962]:**
  - simple cells detect local features
  - complex cells "pool" the outputs of simple cells within a retinotopic neighborhood.



Consider learning an image:

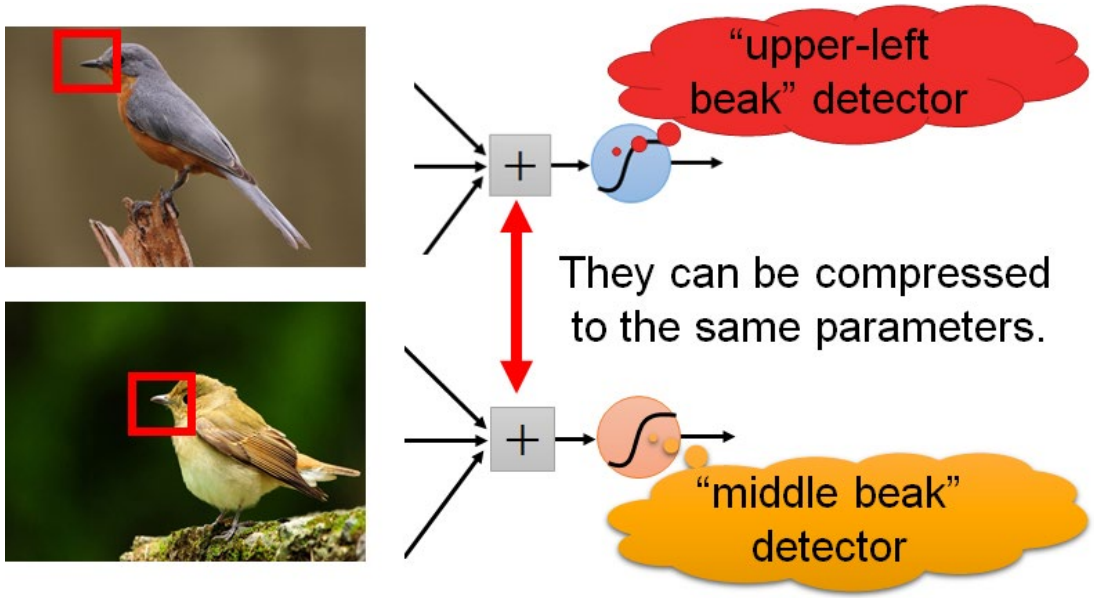
- Some patterns are much smaller than the whole image

Can represent a small region with fewer parameters



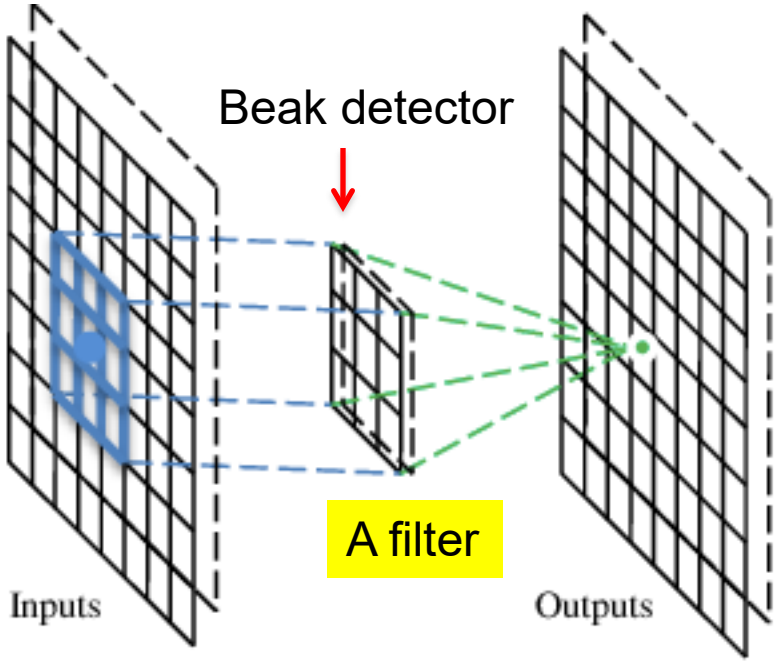
Detectors

- Same pattern appears in different places:
  They can be compressed!
  What about training a lot of such “small” detectors
  and each detector must “move around”.

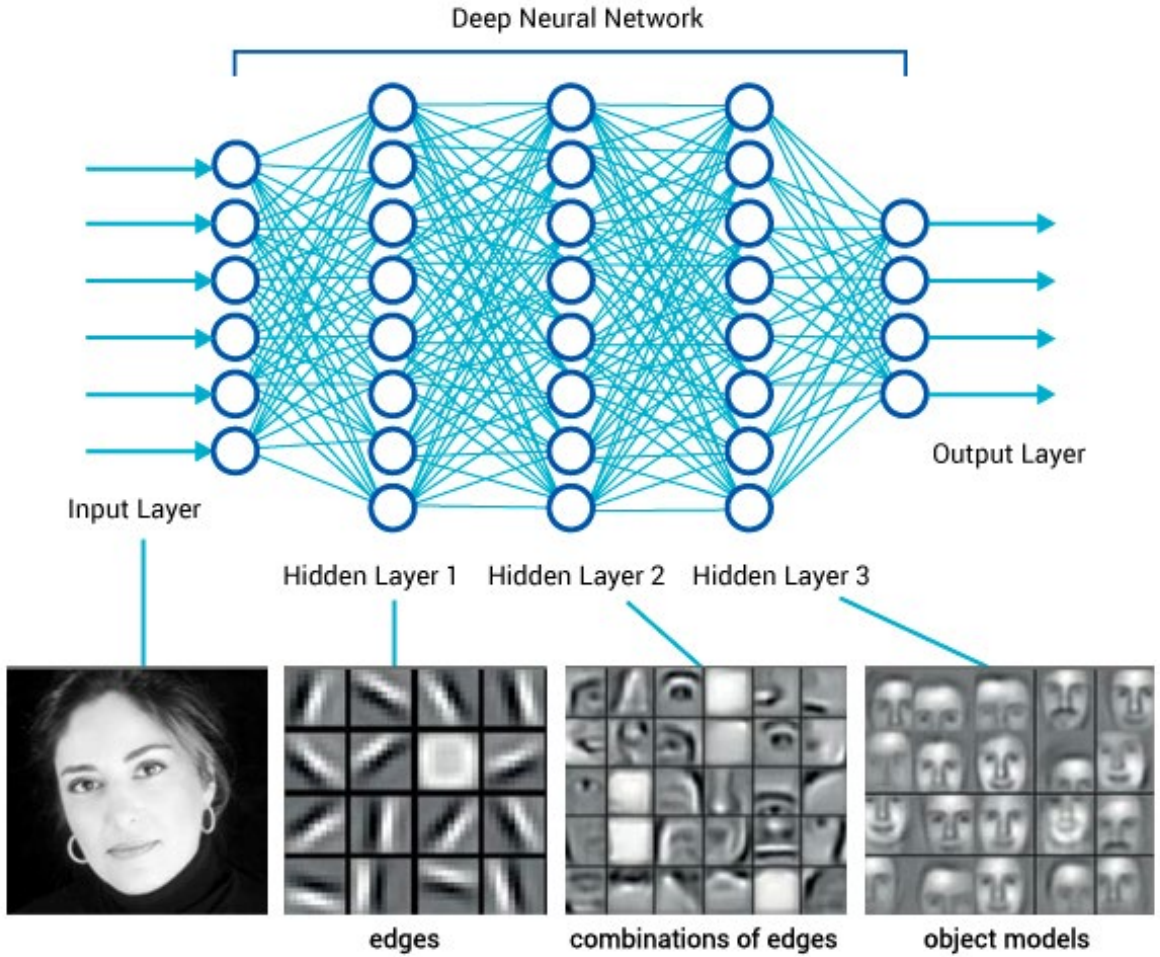


A convolutional layer

- A CNN is a neural network with some convolutional layers (and some other layers). A convolutional layer has a number of filters that does convolutional operation.



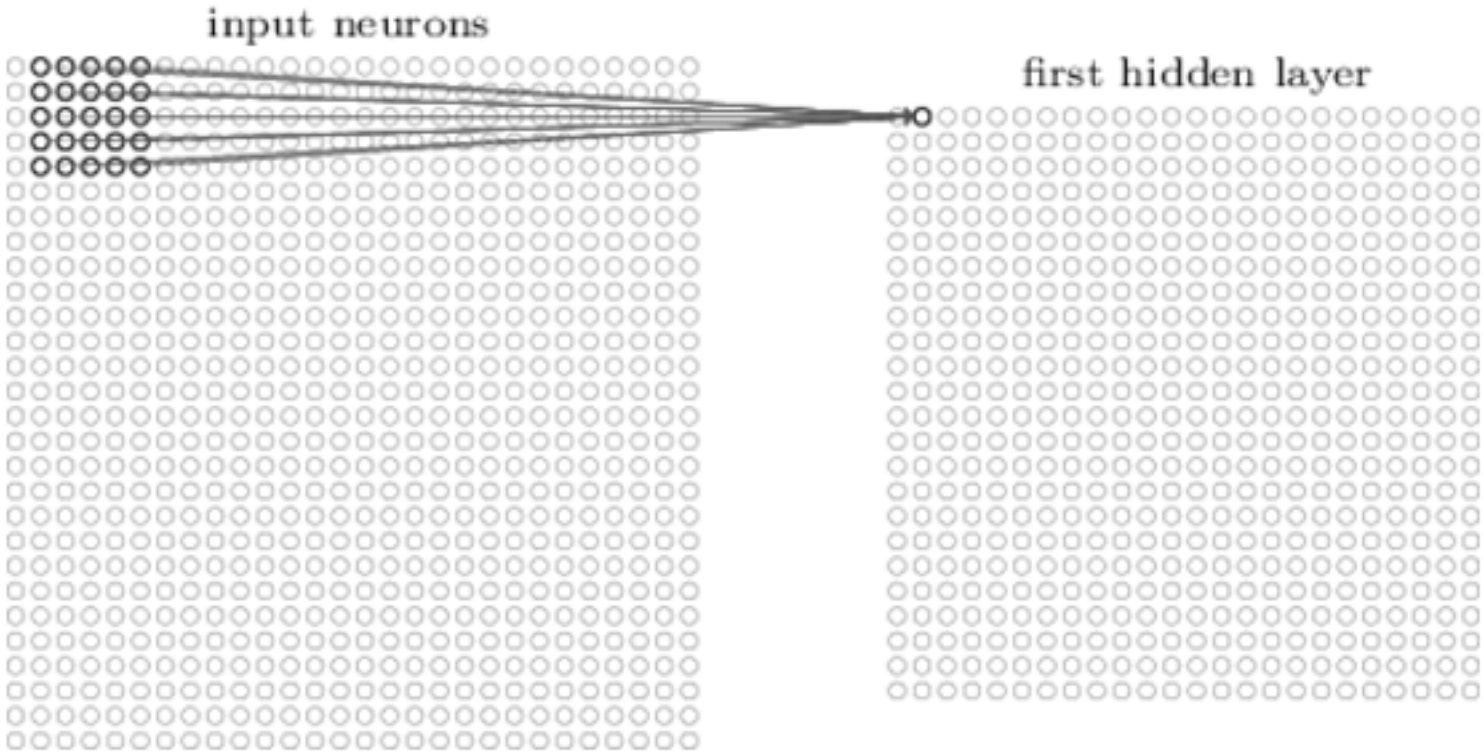
What is happening?



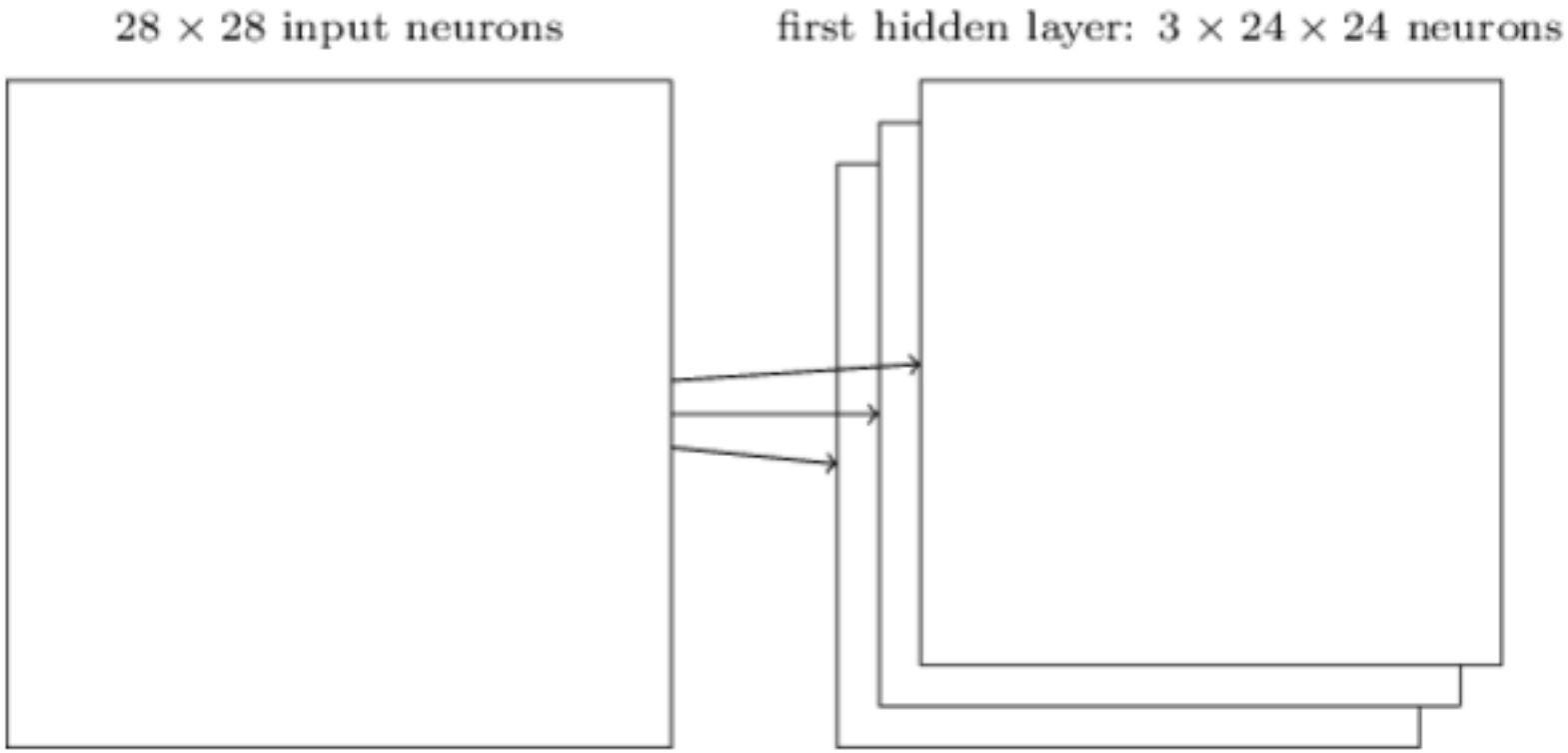
https://www.saagie.com/fr/blog/object-detection-part1

How do we convolve an image with an ANN?

Note that the parameters in the matrix defining the convolution are **tied** across all places that it is used



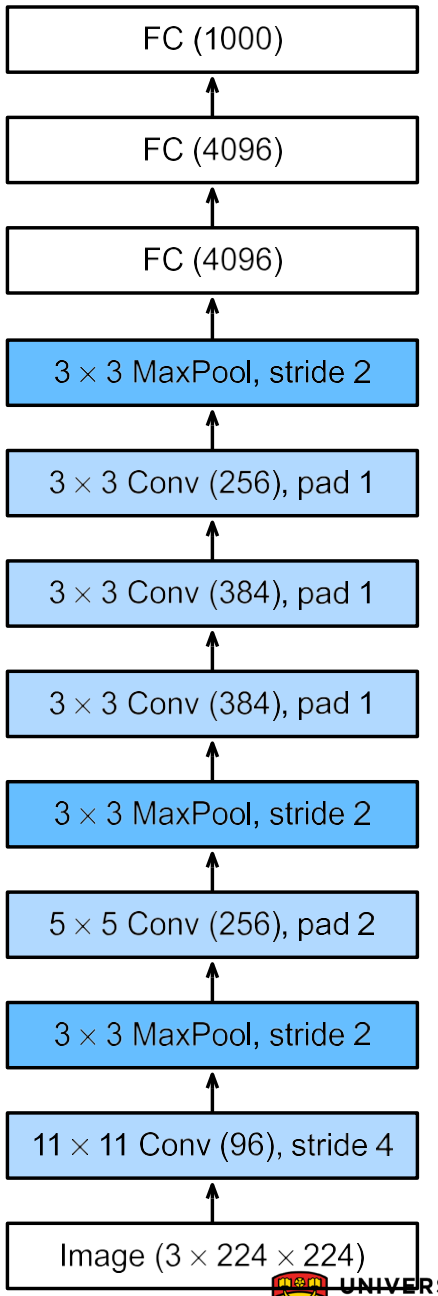
How do we do many convolutions of an image with an ANN?



AlexNet

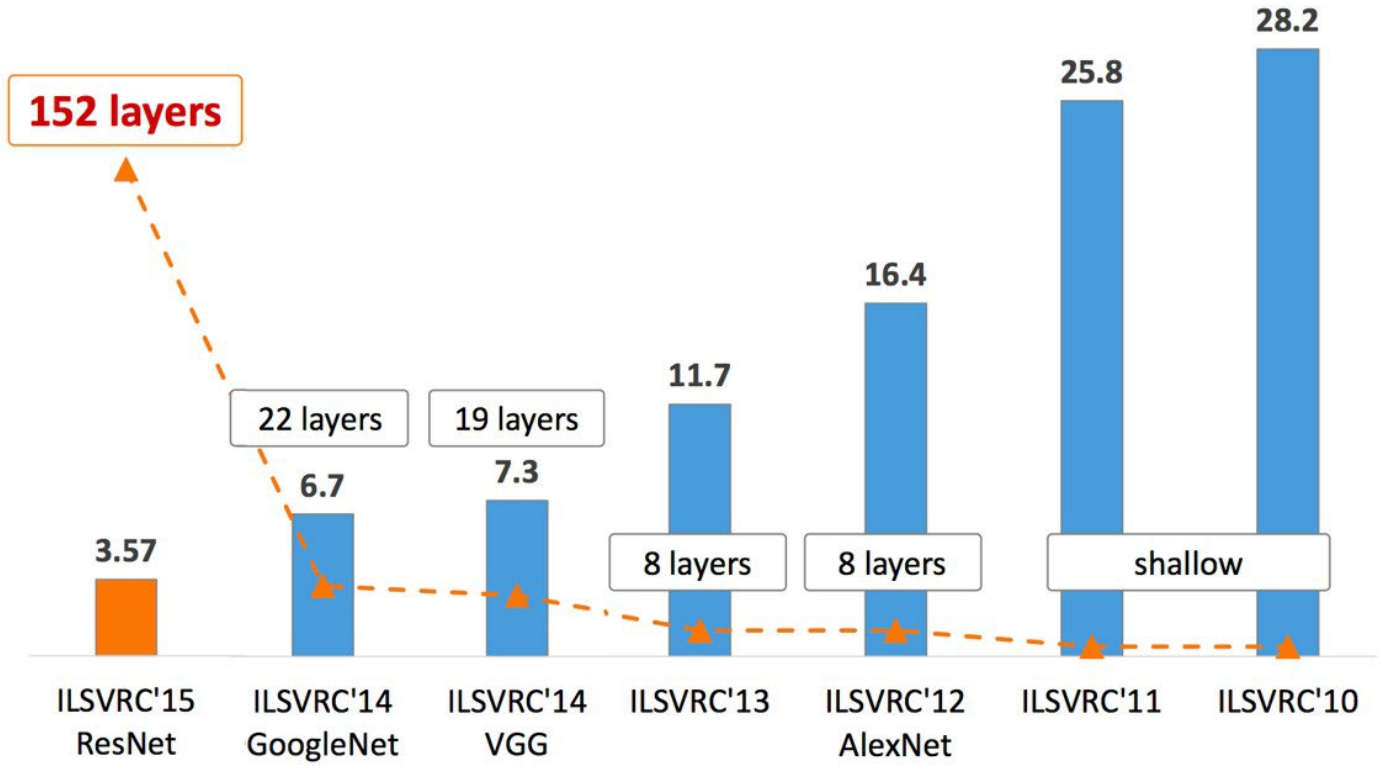
AlexNet is a **convolutional neural network** (**CNN**), consisting of the following layers:

- Fully connected (FC) layers
- Convolutional layers (Conv)
- Max pooling layers (MaxPool)



Depth

- As processors have become more powerful, neural nets have been able to get deeper and deeper:



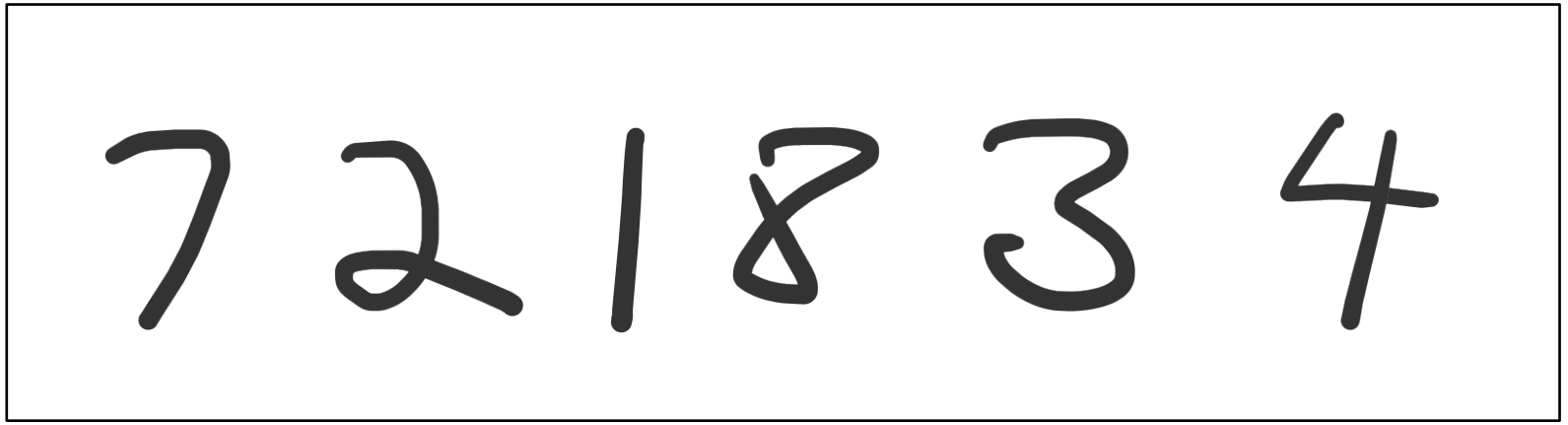
Jiang, Zichao. (2019). A Novel Crop Weed Recognition Method Based on Transfer Learning from VGG16 Implemented by Keras. IOP Conference Series: Materials Science and Engineering. 677. 032073. 10.1088/1757-899X/677/3/032073.

Convolution

Example

So far, we can identify 28x28 images of digits by flattening and inputting the image by pixels.

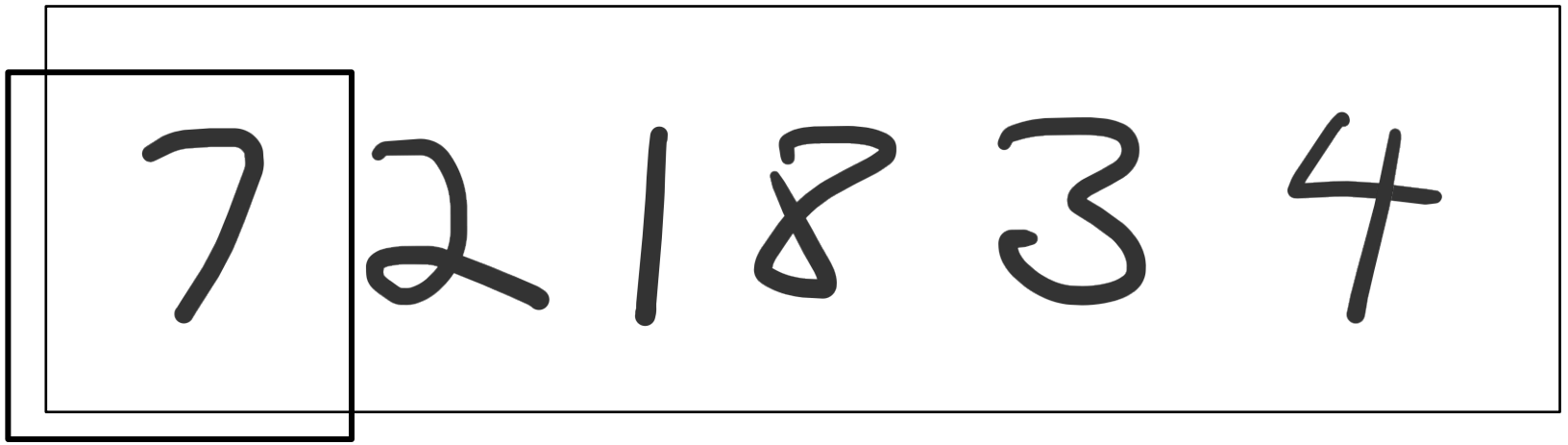
What if we want to read a longer number that is less cleanly processed?



Idea

One way we could imagine doing this is by applying a neuron to an image using a shifting window.

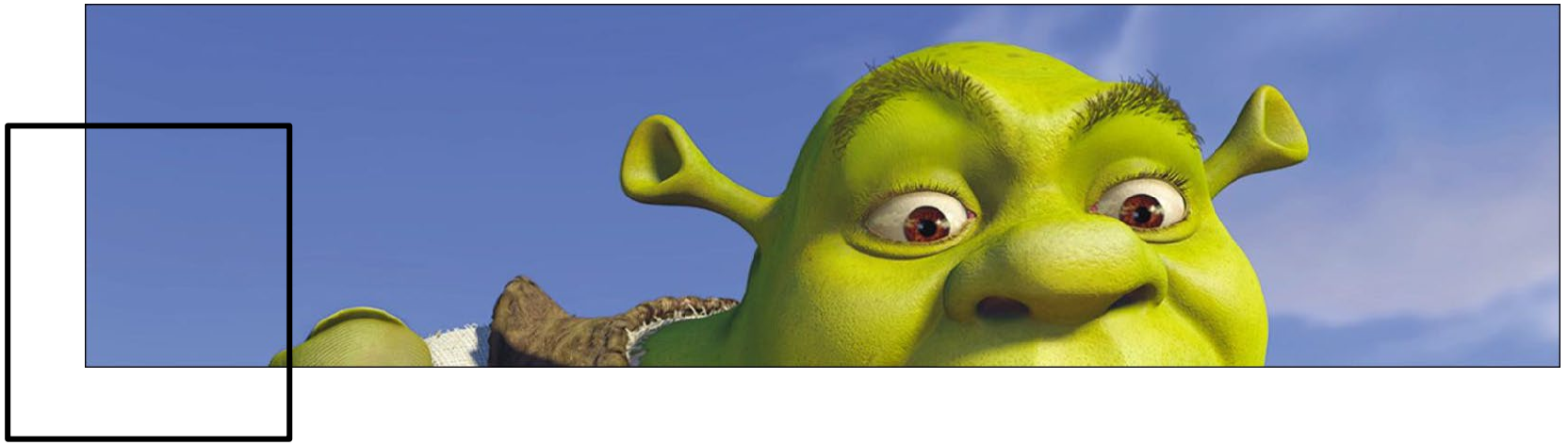
- When the image lines up with something the neuron recognizes, this will be reflected in the outputs



Convolution

Convolution is the process of taking a neuron called a **filter** or **kernel** and measuring its output at regular locations in the original input.

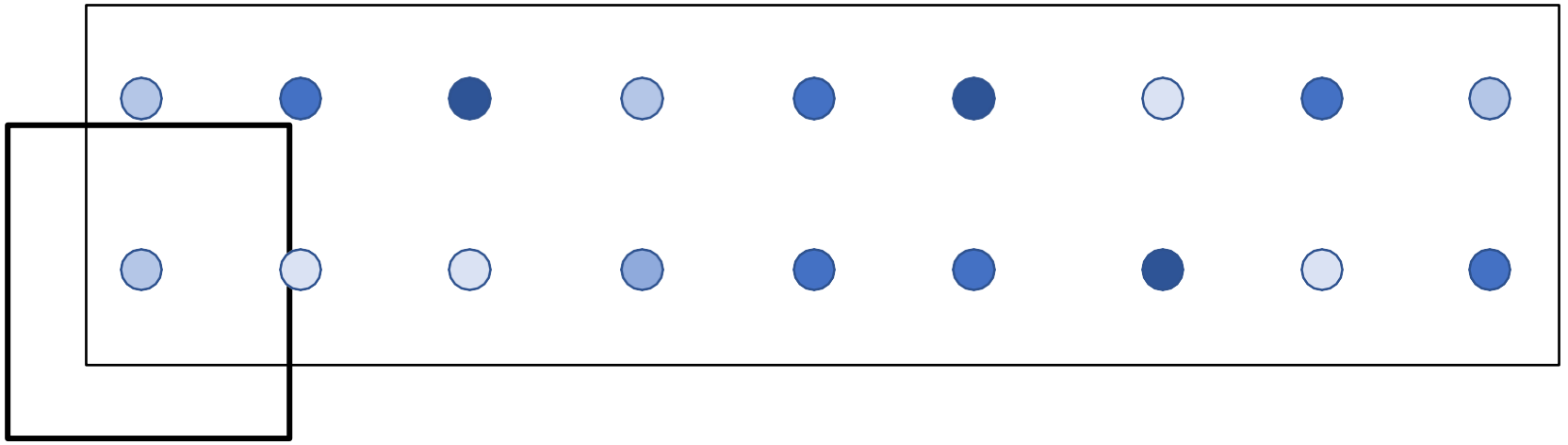
- The distance the location shifts each step is called the **stride**.



Convolution

For each location of the frame, the set of neurons gives an output value.

- Then you can use these values as inputs to the next layer
- Often repeated for many different filters (https://convviz.netlify.app/)



Example

What is the output of convolving the following input array with the given filter, assuming a stride of 1 and a bias of 0?

Input:

0	1	1	0	1	0	0	0	1

Filter:

-1	1	-1

Example

What is the output of convolving the following input array with the given filter, assuming a stride of 2 and a bias of 0?

Input:

0	1	1	0	1	0	0	0	1

Filter:

1	-1	-1

Example

What is the output of convolving the following input array with the given filter, assuming a stride of 1 and a bias of 0?

Input:

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

Filter:

-1	1
-1	-1

Pooling

Why Pooling

- Subsampling pixels will not change the object

bird

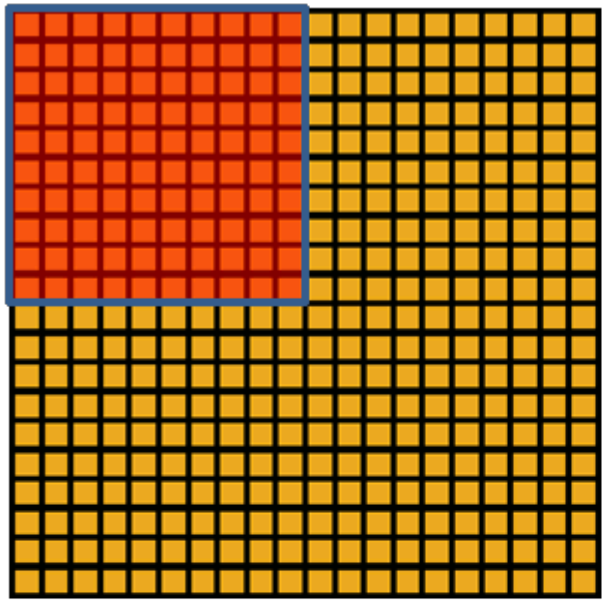
Subsampling

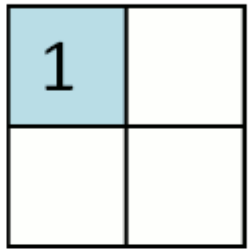
bird

We can subsample the pixels to make image smaller

fewer parameters to characterize the image

Pooling



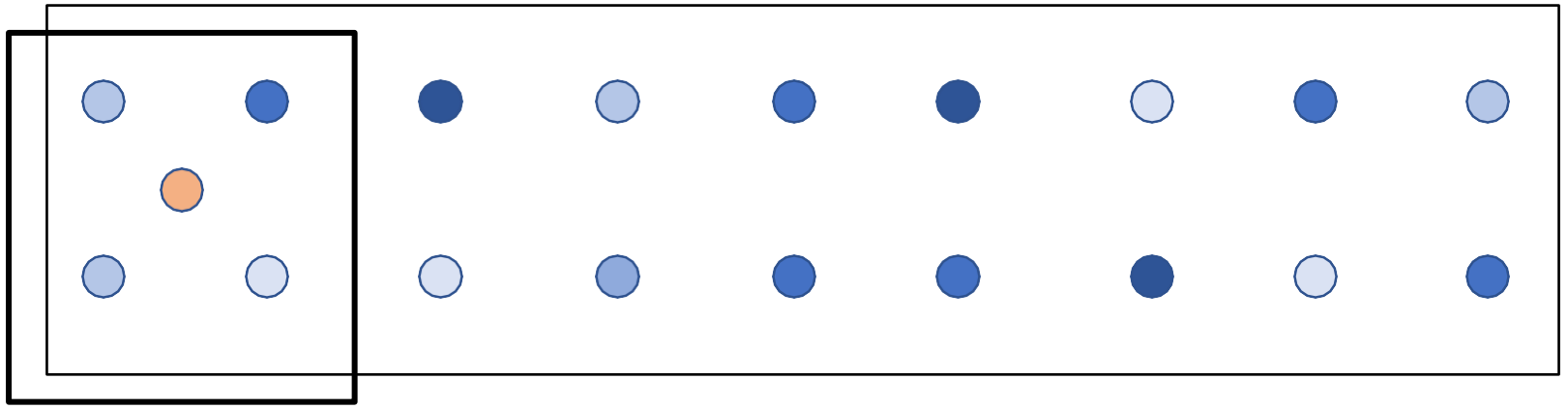


Convolved feature

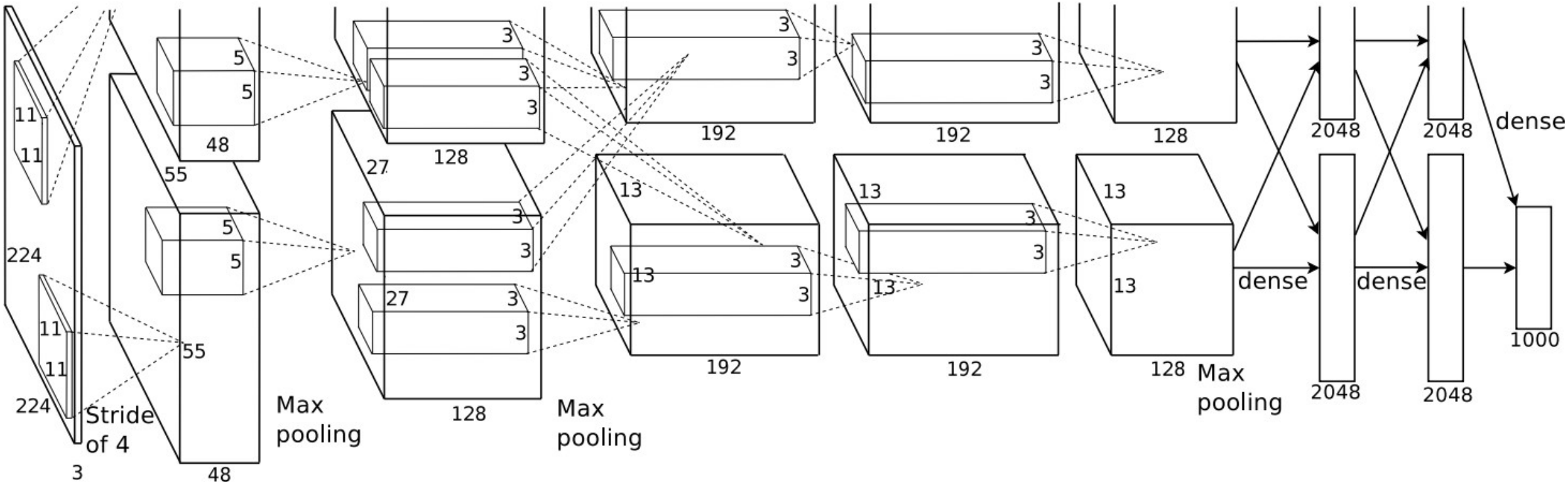
Pooled feature

Convolution

- **Max pooling** is the process of taking the maximum value in a shifting window, usually with stride > 1.
- Helps to reduce image size while maintaining information about what was found in the convolutional step

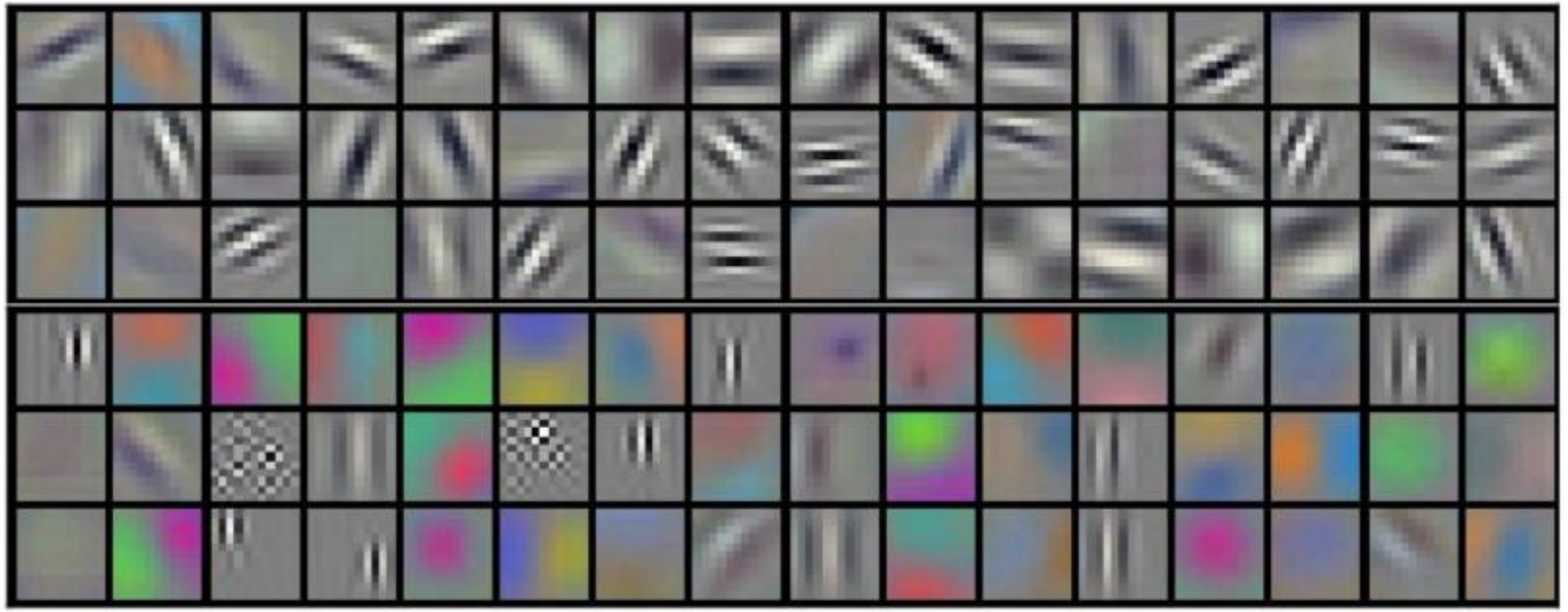


AlexNet (revisited)



Filters

- We can take the trained filters from the first convolutional layer and see what input patterns caused them to activate the most:



Dropout

Other ideas

AlexNet has 60 million parameters in total and 1000 classes of images from ILSVRC to train on.

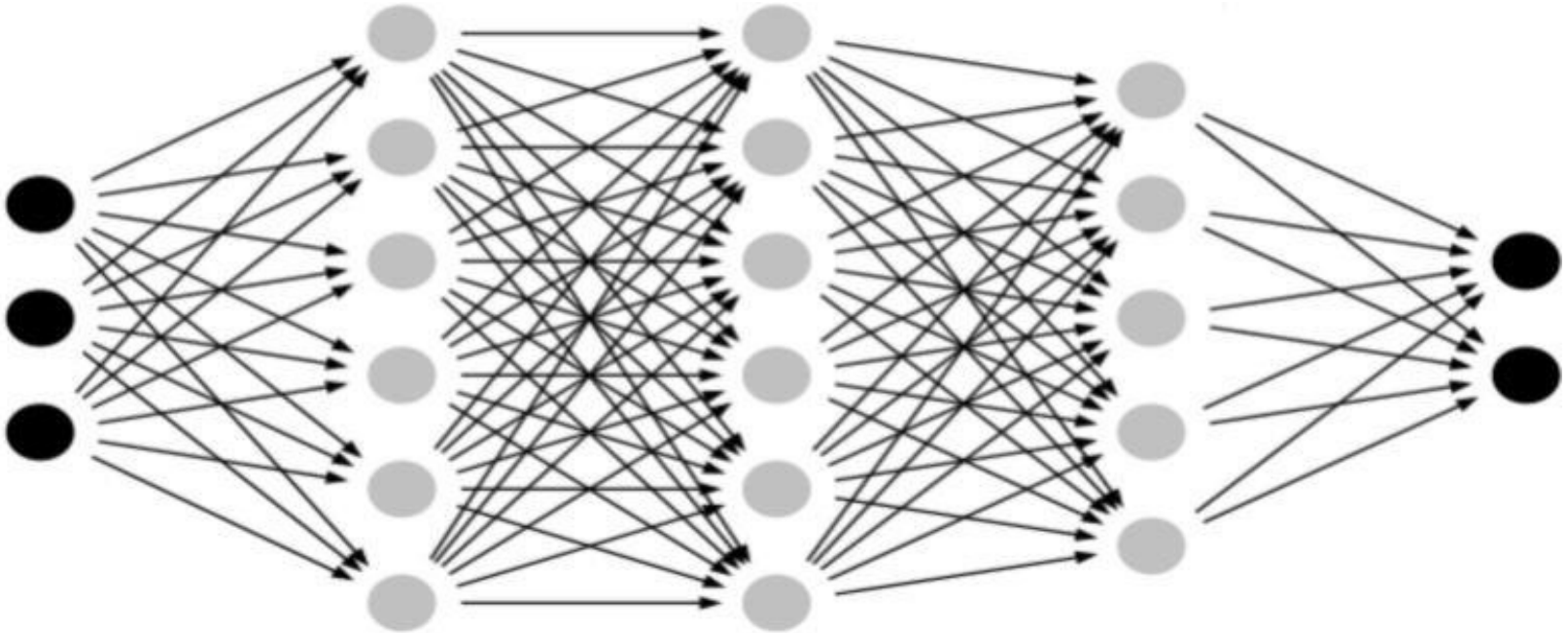
- Only about 1000 examples of each class

This makes it challenging to avoid overfitting.

Dropout

In **dropout**, for each training sample, certain hidden neurons will have their inputs set to 0 with probability p.

```
tf.keras.layers.Dropout(0.2)
```



Dropout

Dropout helps to prevent over-reliance on individual neurons.

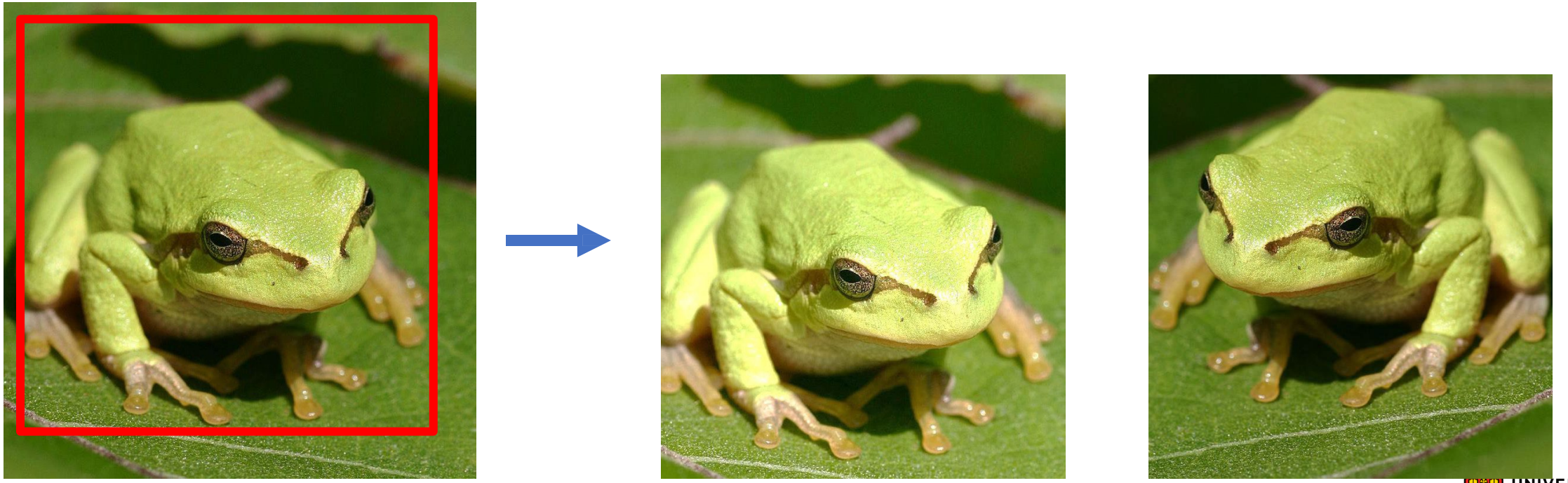
- Also helps avoid getting stuck in local optima

However, this increases the training time, since not all weights/biases are updated every step.

Data augmentation

Data augmentation

AlexNet also uses **data augmentation** to increase the amount of training data it has available to learn from.



Data augmentation

The original images in ImageNet were 256 x 256 pixels, but AlexNet operated on images with 224 x 224 pixels.

These new images were obtained from the originals by randomly cropping the borders and/or flipping the original image.

- This increased the training set size by a factor of 32 x 32 x 2 = 2048

They also randomly adjusted the colour intensity each time a sample was used for training.

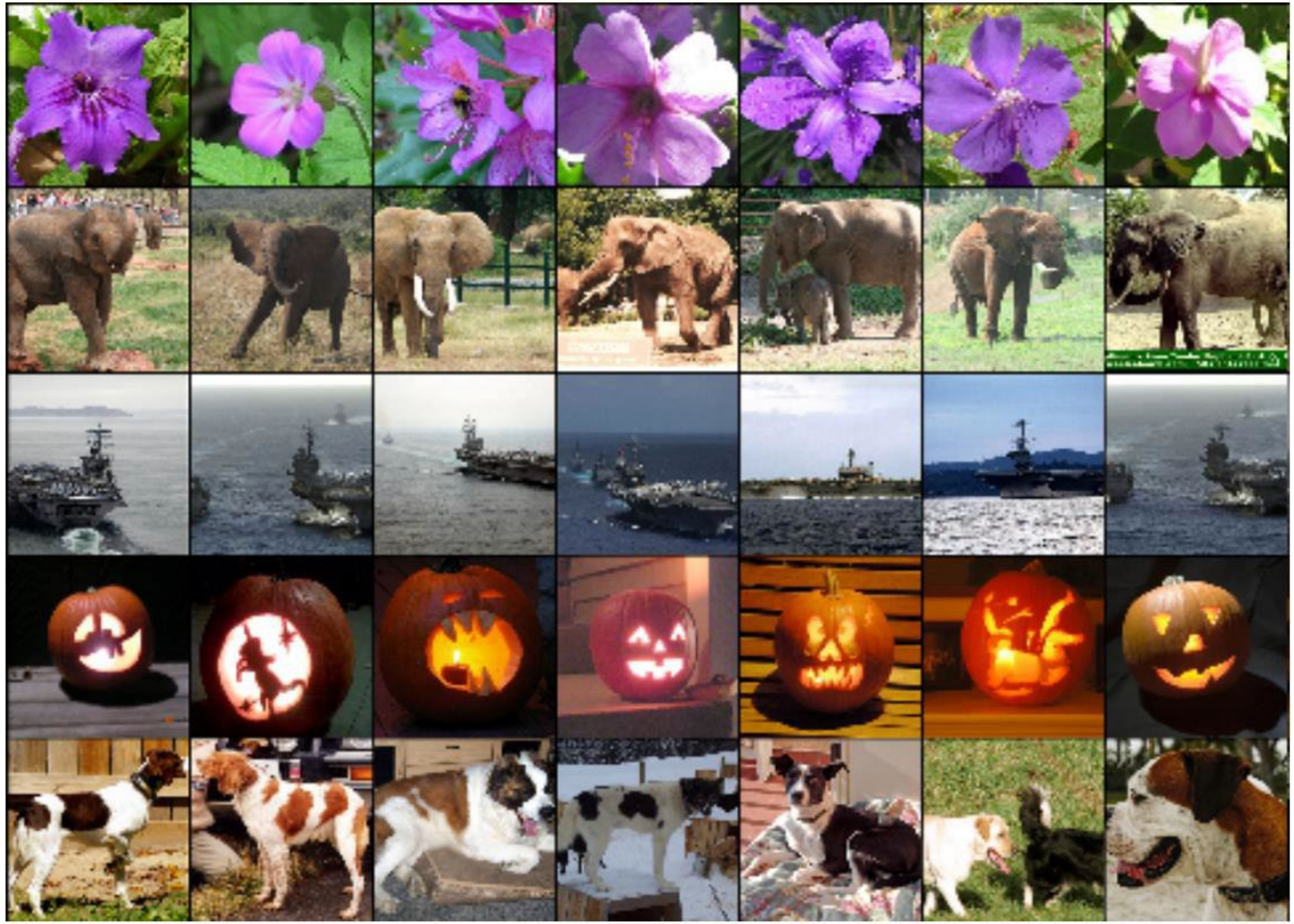
Exploration

Similarity test

One way we can look at the performance of a neural net is to see which training images produce outputs that are “close” to the output on a particular image.

This gives us some insight into what types of patterns the neural net is learning.

https://convnetplayground.fastforwardlabs.com/#/



Note

These neural nets can “see”, but not in the same way we do.

For example, humans are able to learn based on very few examples, while neural nets need hundreds or thousands for each image class.

- Difference is understanding of context and the real world

Adversarial

Adversarial examples

Neural nets behave reasonably well on inputs that resemble the training data.

However, they don't perform well in an **adversarial** setting.

- i.e. we can easily design inputs for which things go horribly wrong

This happens even for the "good" neural networks, and is based on exploiting how they work.

ballplayer 69.22%

anemone_fish 92.48%

African_elephant 89.94%

forklift 98.95%

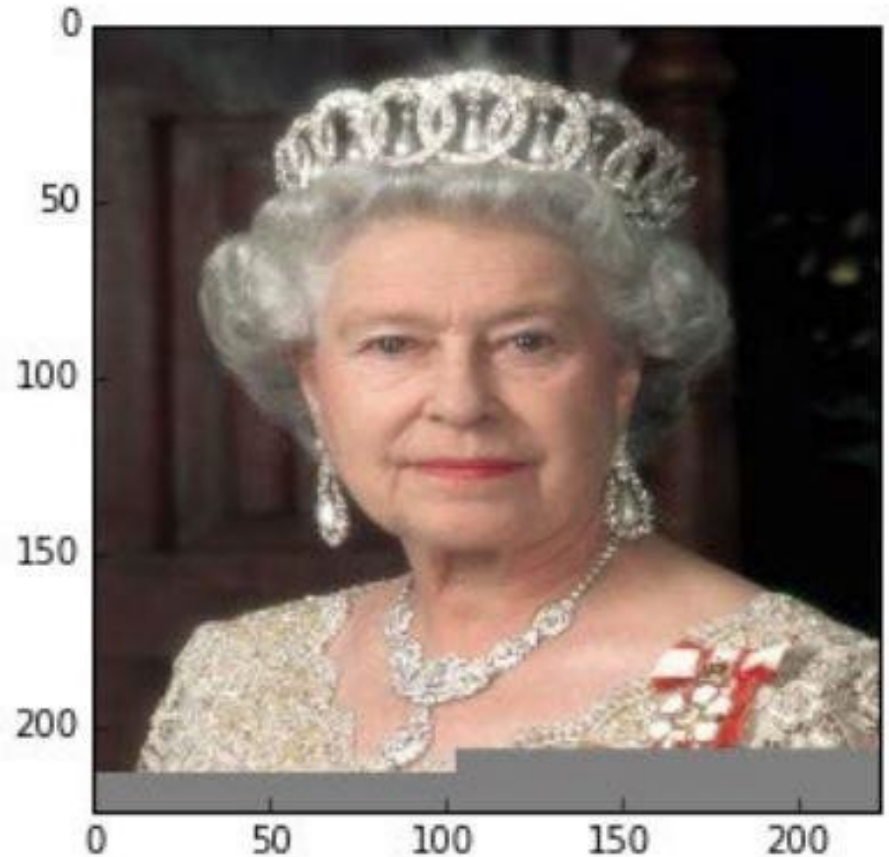
ice_cream 99.60%

lemon 97.06%

magnetic_compass 97.08%

ice_bear 84.80%

```
class: 793
label: n04209133 shower cap
certainty: 99.7%
```

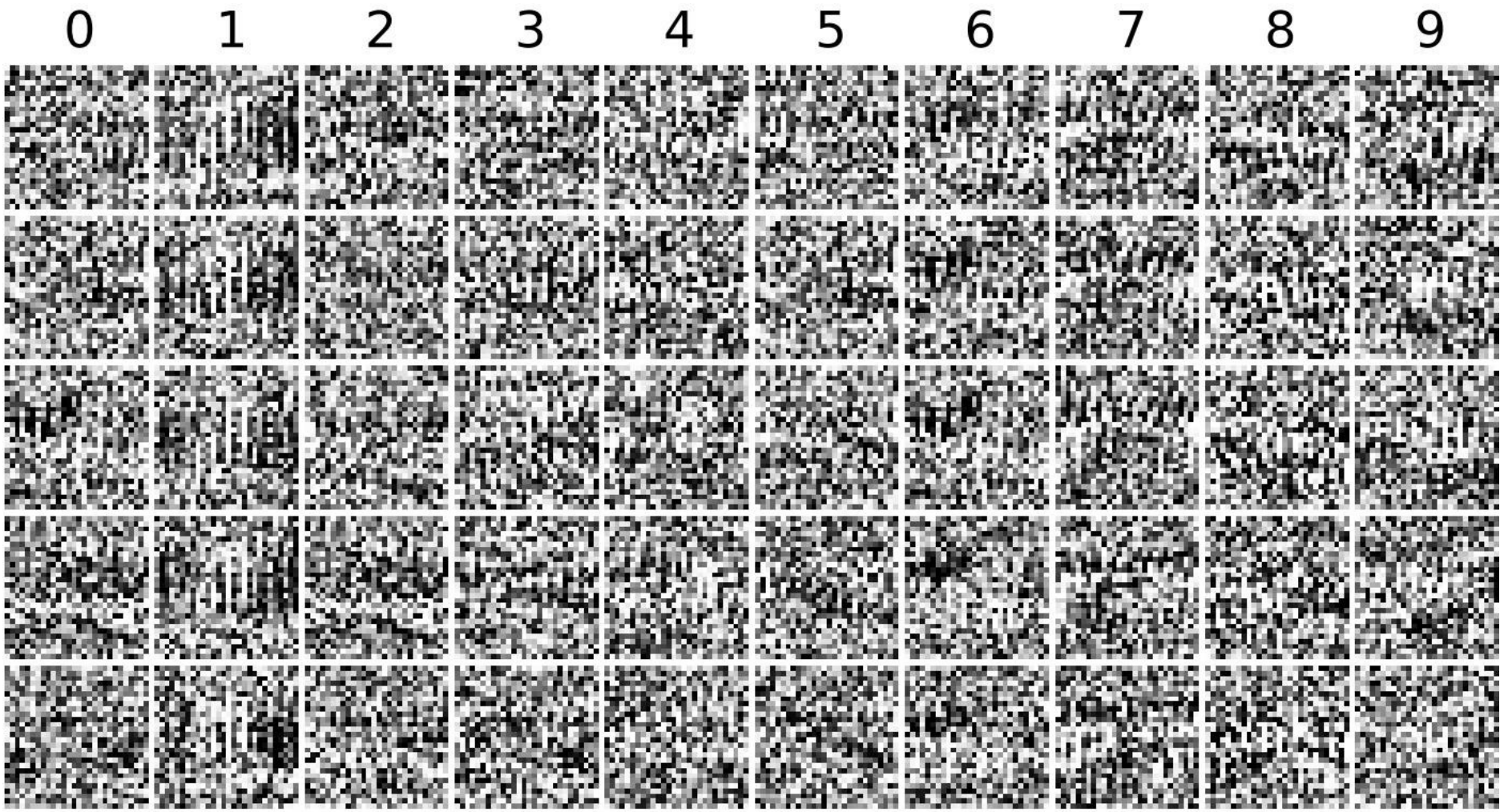


Creating images

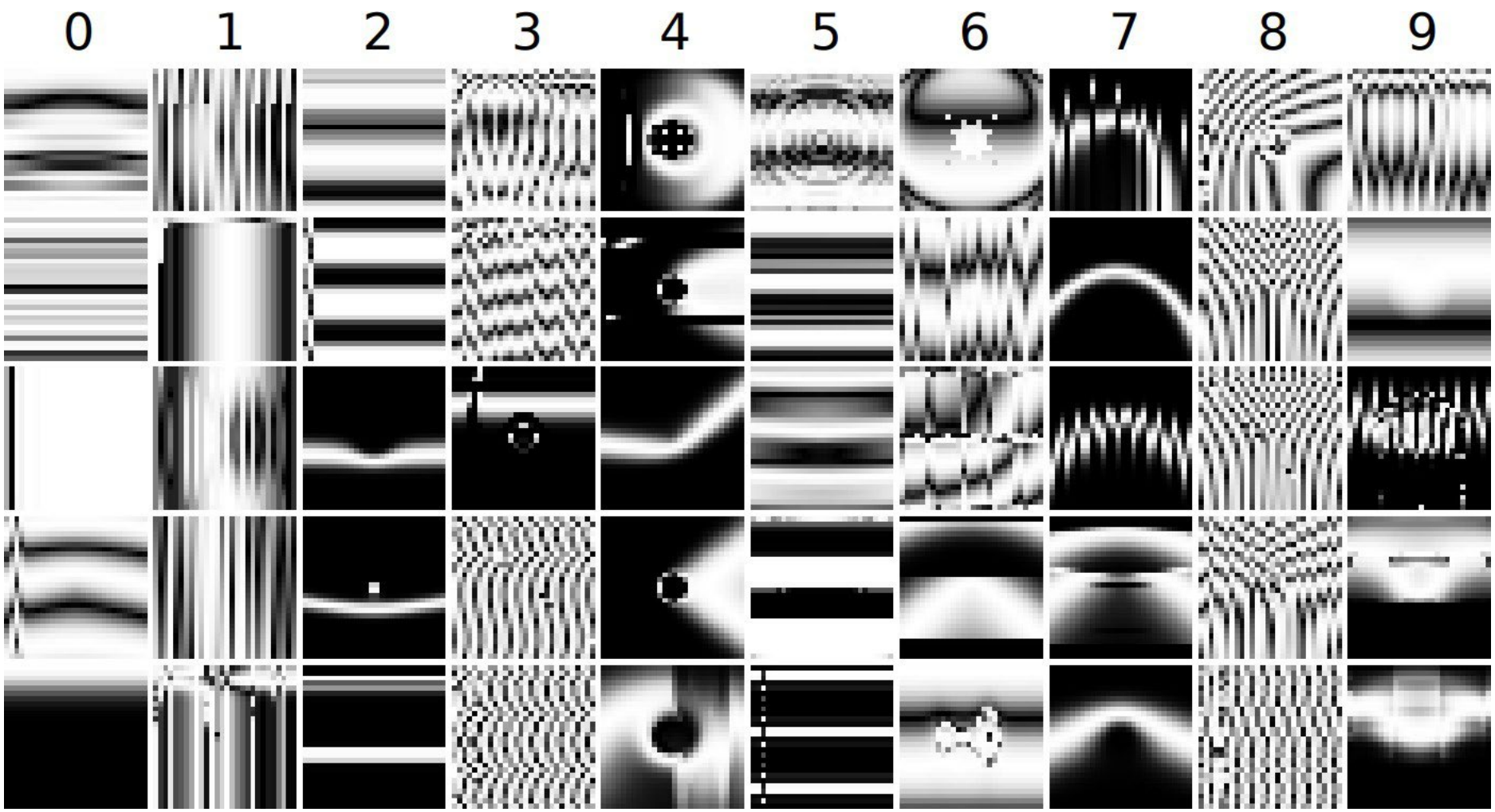
One way in which we can generate images that fool a network is with a constructive approach.

- e.g. genetic algorithms, gradient ascent, or GANs

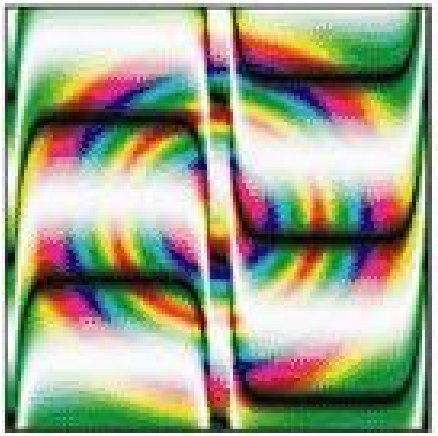
We start with an image of random noise and keep adjusting it in ways that improve the network's confidence that it is a certain target class.



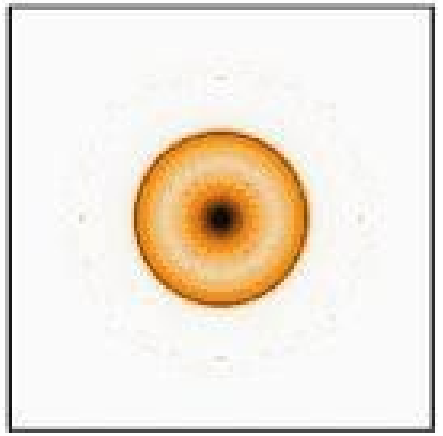
A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 427-436



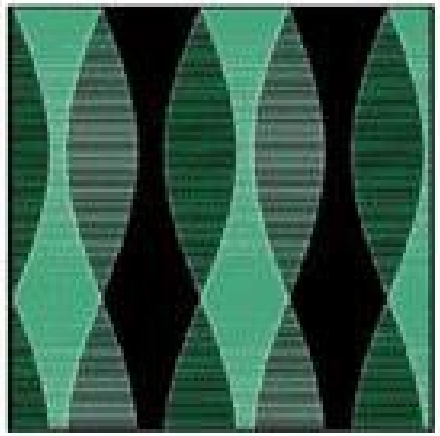
A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 427-436



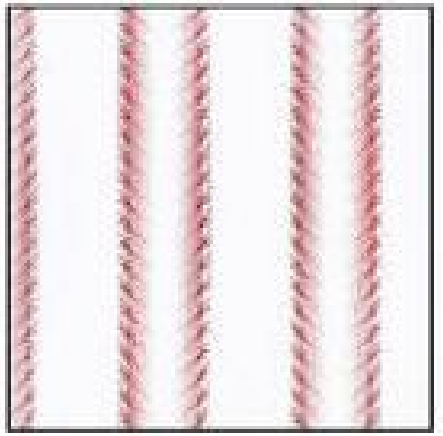
Pinwheel



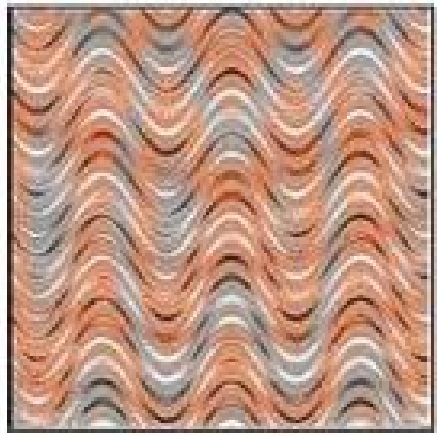
Bagel



Paddle



Baseball



Tile roof

Armadillo

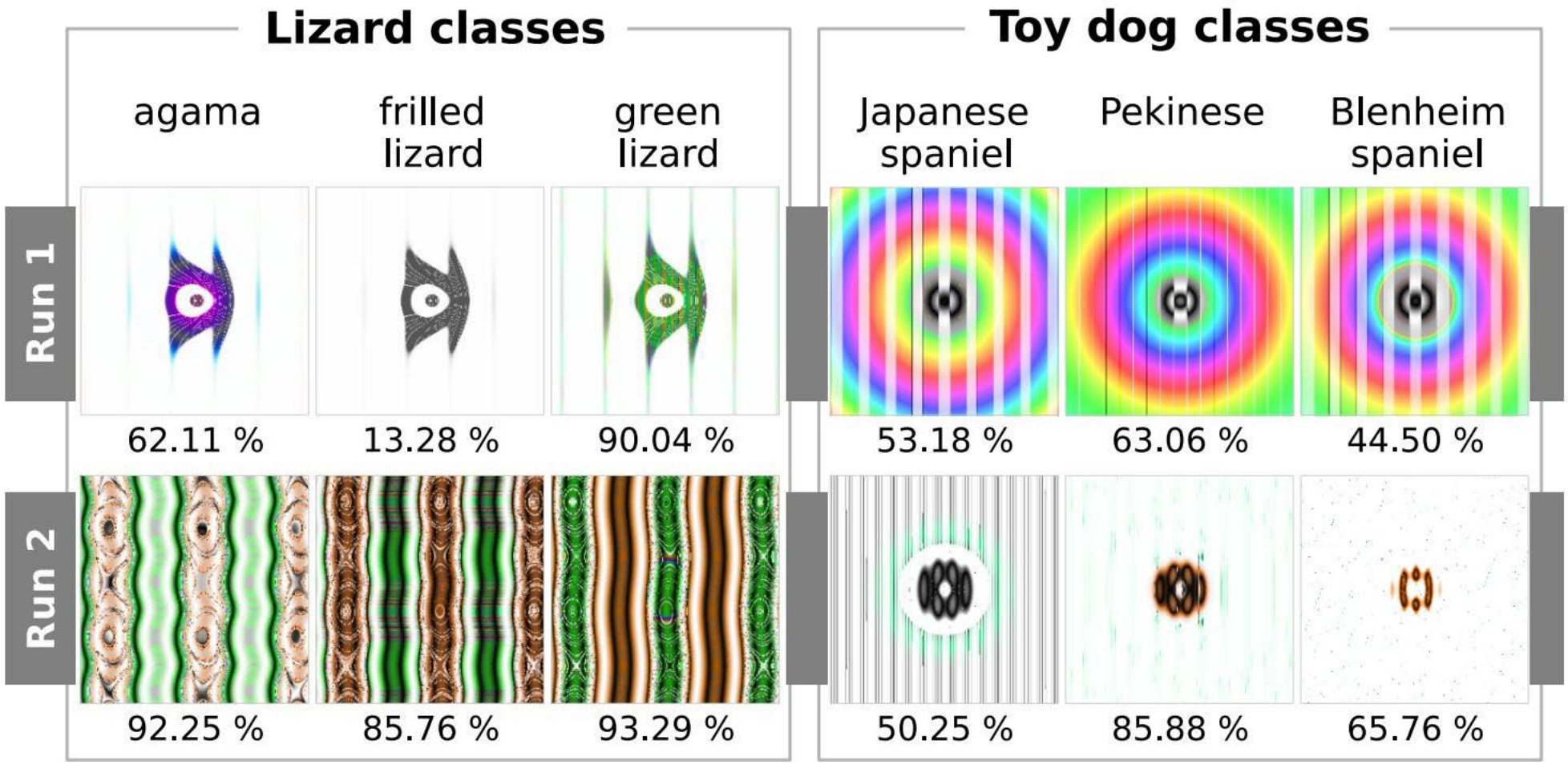
Bubble

Centipede

Jackfruit

Robin

A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 427-436

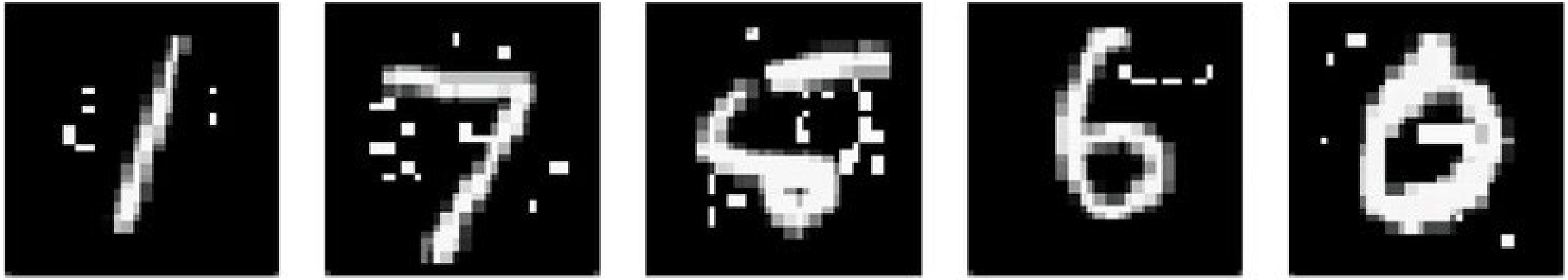


A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 427-436

Creating images

We can also start with images the neural net does recognize and alter them in ways that “tricks” the net into thinking it is a different image.

Part of the reason this works is that the model seems to care about certain pixels more than others, so by adjusting those particular pixels we can cause it to change its label.



1 *7* *5* *6* *0*

4 **9** **0** **5** **9**

A. Elsayed, D. Krishnan, H. Mobahi, K. Regan, and S. Bengio, "Humans can decipher adversarial images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 160-169.

Milk can

Baseball

Muzzle

Tree frog

Jaguar

Green lizard

Hard disk

Sand viper

Power drill

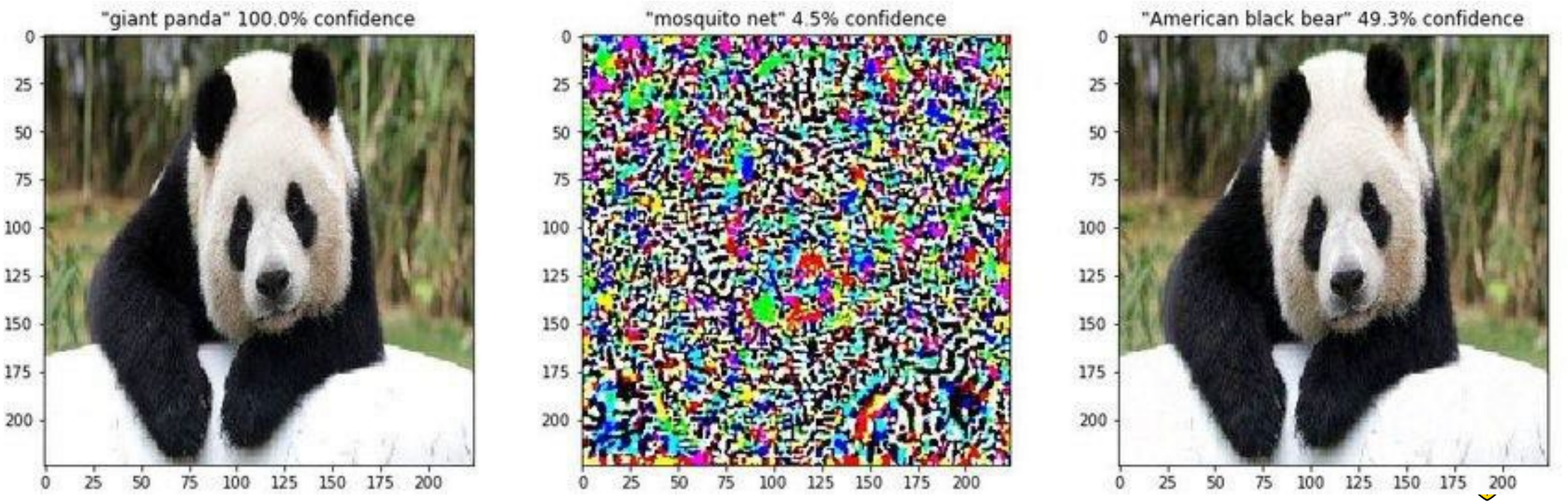
Jigsaw puzzle

A. Elsayed, D. Krishnan, H. Mobahi, K. Regan, and S. Bengio, "Humans can decipher adversarial images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 160-169.

Adding noise

A **fast gradient signed method** (**FGSM**) attack is based on altering pixels of an image in a way that maximizes its loss on a trained model.

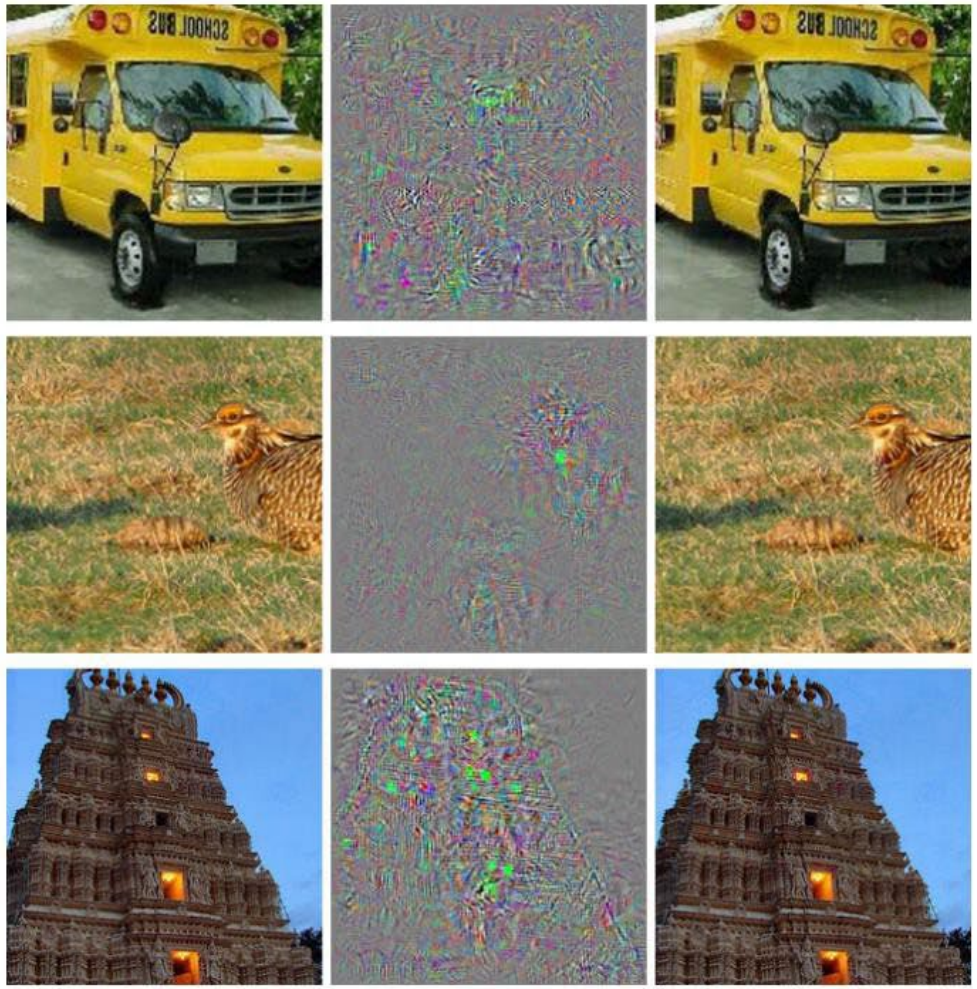
- https://www.tensorflow.org/tutorials/generative/adversarial_fgsm



Zhang, Weijia. (2019). Generating Adversarial Examples in One Shot With Image-to-Image Translation GAN. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2946461.

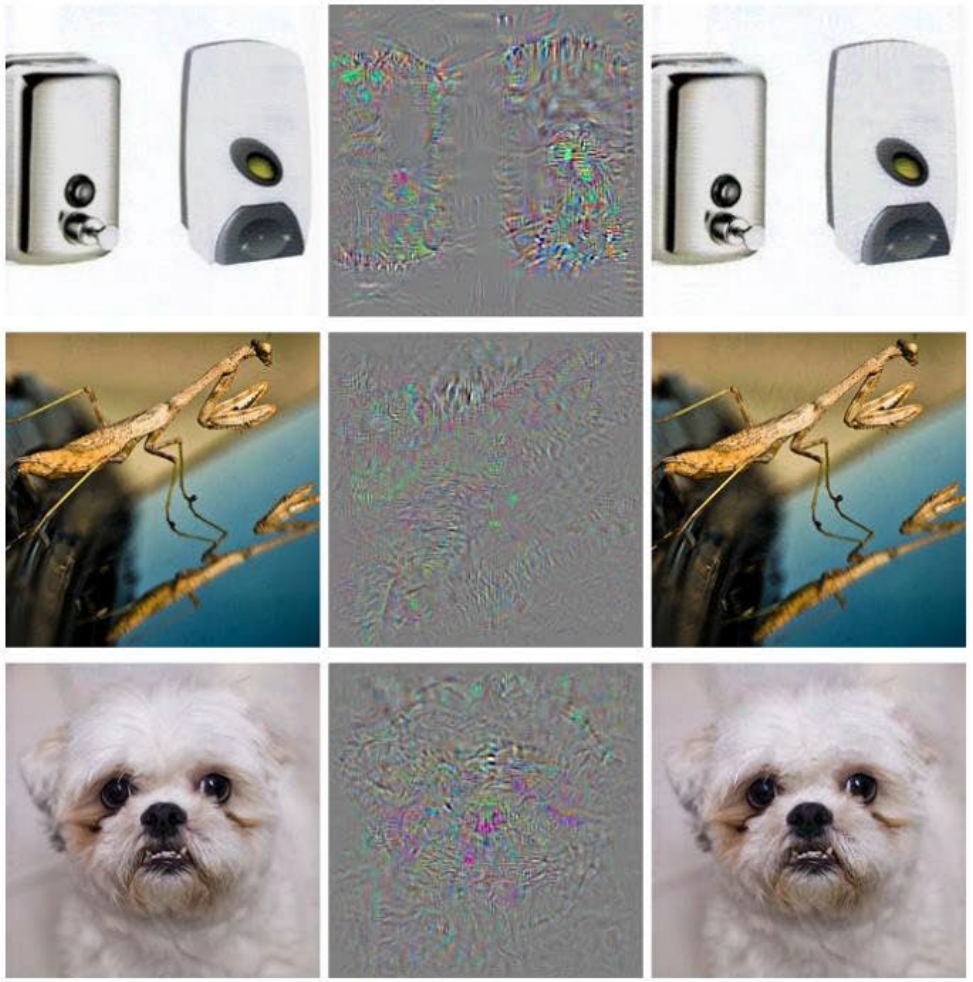
Not an Ostrich

Ostrich



Not an Ostrich

Ostrich



C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," arXiv preprint arXiv:1312.6199, 2014.

Summary

Convolutional neural networks are amazing at solving certain types of problems.

However, they don’t see things the way we do, and they can still be tricked by exploiting how they work.

Finding these adversarial images also helps us improve neural networks, which is why they are important to look for.

Next…auto-encoders

Jonathan Hudson, Ph.D.
jwhudson@ucalgary.ca
https://cspages.ucalgary.ca/~jwhudson/

UNIVERSITY OF CALGARY