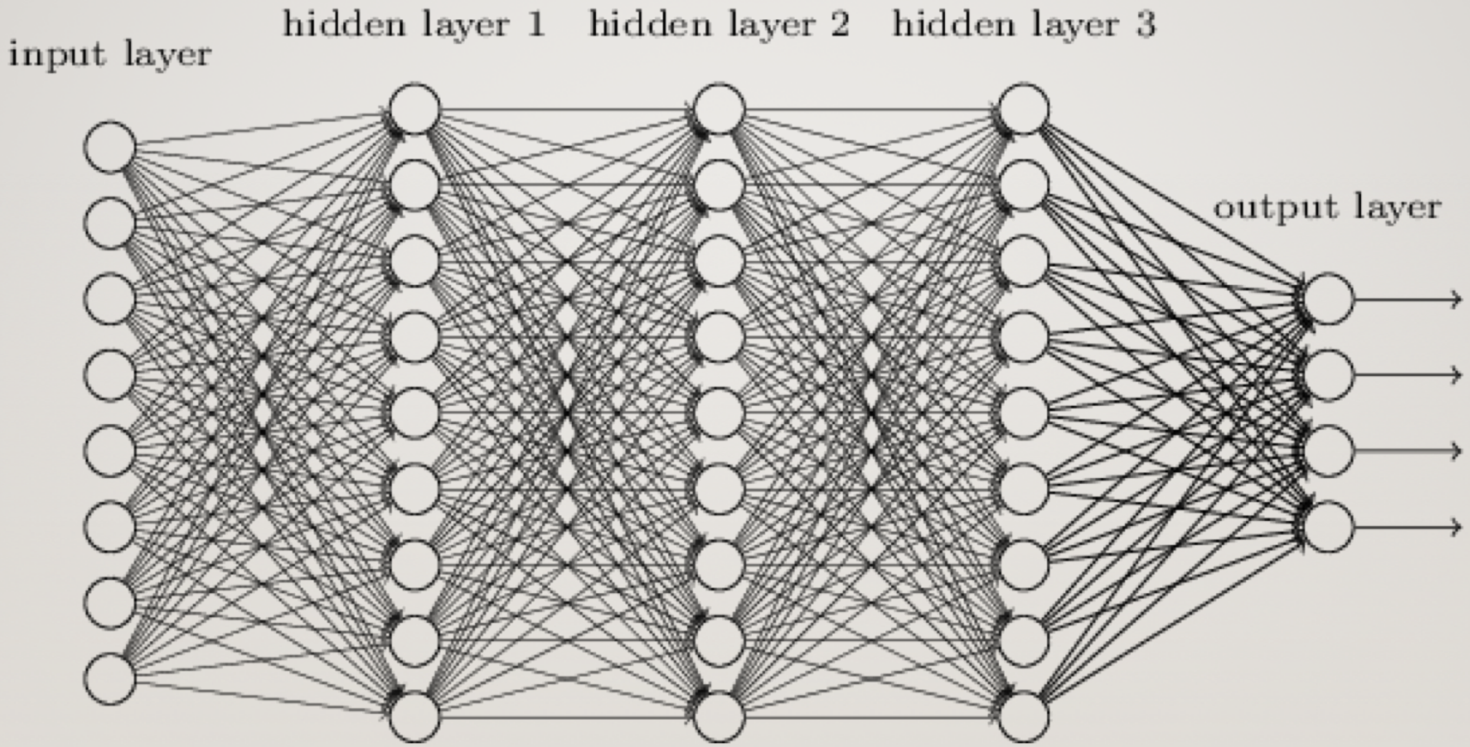
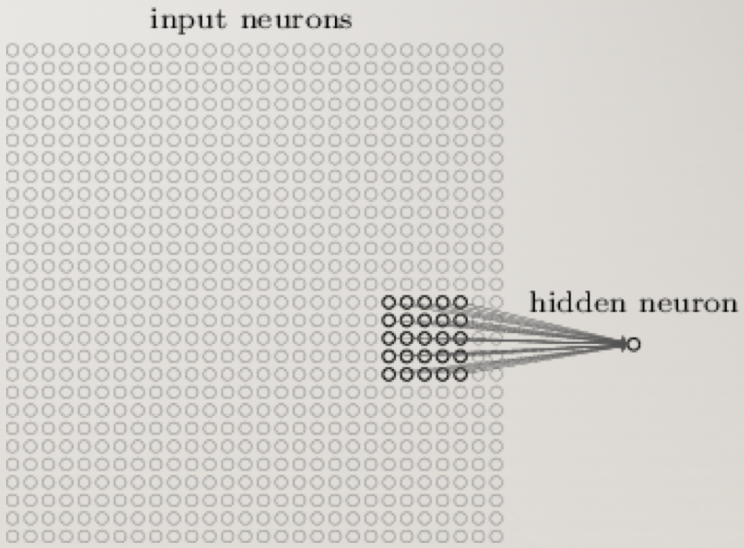
CHAPTER 6

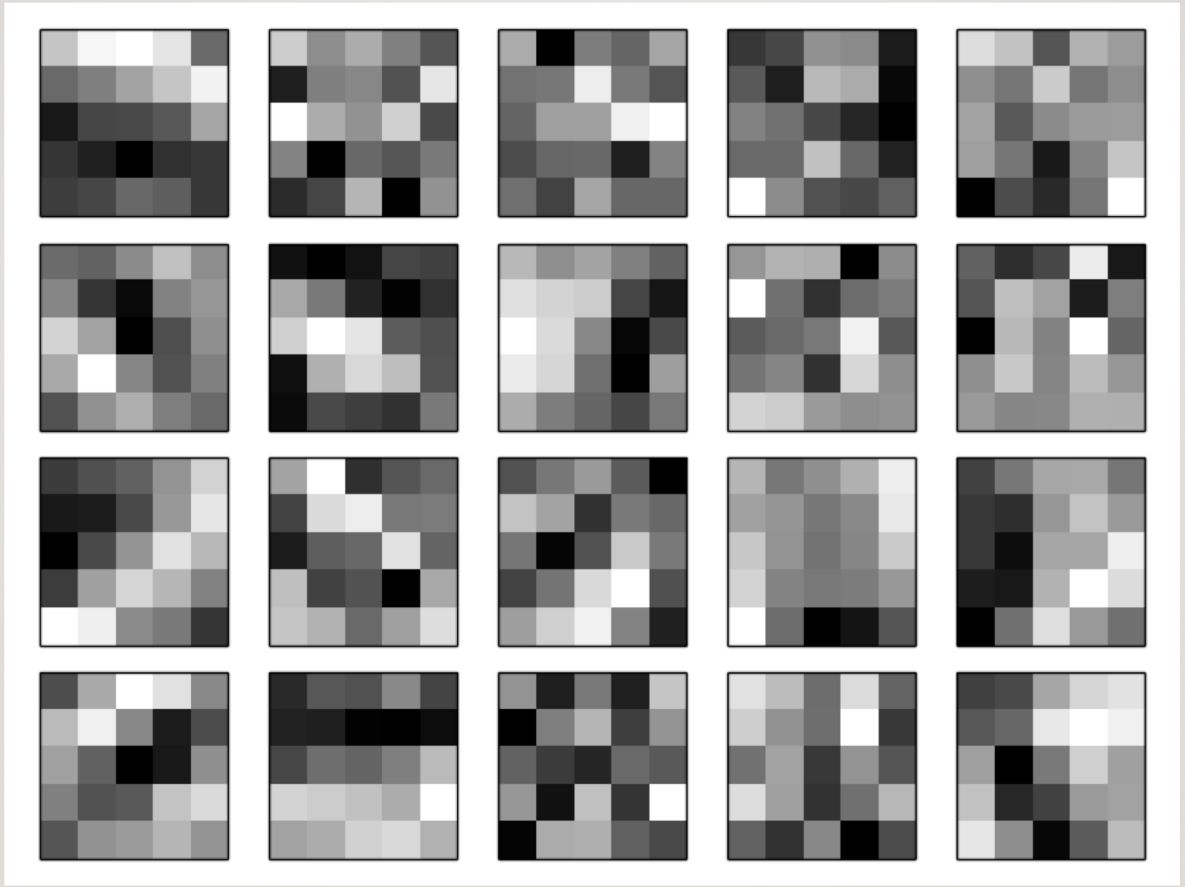
DEEP LEARNING


input layer
hidden layer 1
hidden layer 2
hidden layer 3
output layer

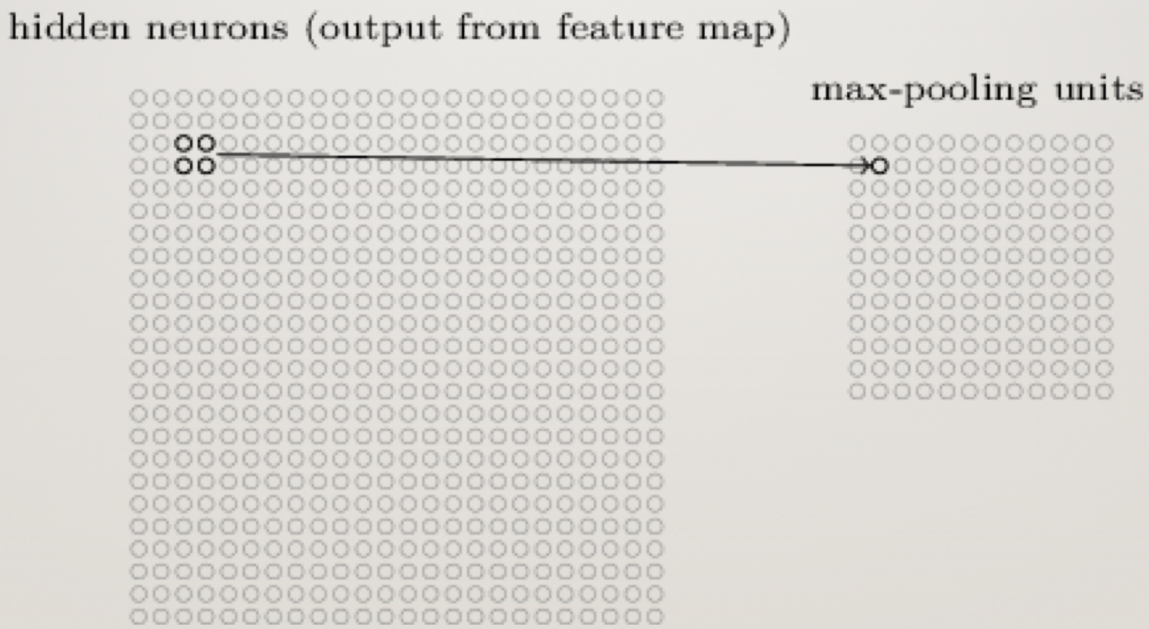
CNNS

- Convolutional Neural Nets
- Spatial understanding
- Local receptive fields (filter, kernel)
- Translation invariance
- Less parameters than MLP
  - 20 5x5 kernels -> total of 20×(5x5+1)=520 parameters defining the convolutional layer
  - 784 input neurons, 30 hidden neurons
    - total of 784×30, +30 biases, for a total of 23,550 parameters



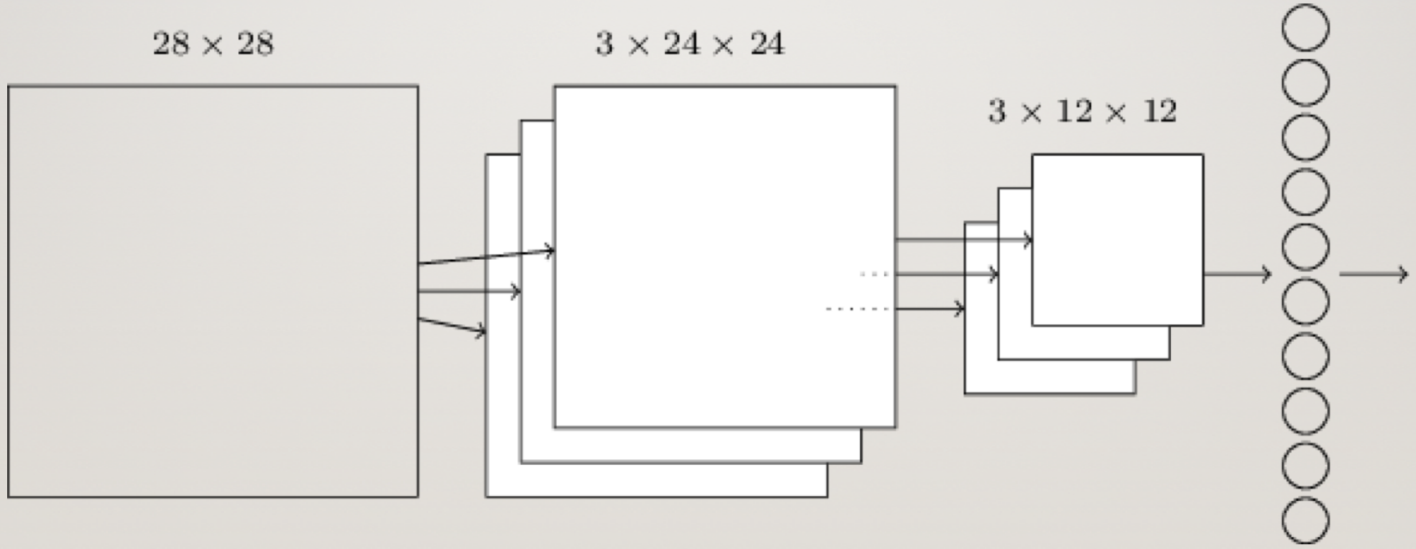


POOLING



MNIST

FORWARD PROP



RESULTS/NOTES

- More convolutional layers and more hidden layers yielded better accuracy
- Later convolutions are harder to interpret

MAKING IMPROVEMENTS

ACTIVATION FUNCTIONS

- sigmoid
  - The og
- Tanh
  - Trains faster, similar results
- ReLU
  - Higher accuracy! we are pretty much clueless as to why

MORE DATA

- Expanded MNIST
  - 250,000 more images
- You can make more data with the data you have
  - Rotation, translation, skewing
  - Elastic distortion
  - GANs

OTHER IMPROVEMENTS

- Dropout
  - Not required for convolutional layers, since they're resistant to overfitting
- Ensemble learning (kinda not really)
- Deep, Big, Simple Neural Nets Excel on Handwritten Digit Recognition, by Dan Claudiu Cireșan, Ueli Meier, Luca Maria Gambardella, and Jürgen Schmidhuber (2010)
  - hella big network, trained for a hella long time with a hella fat GPU

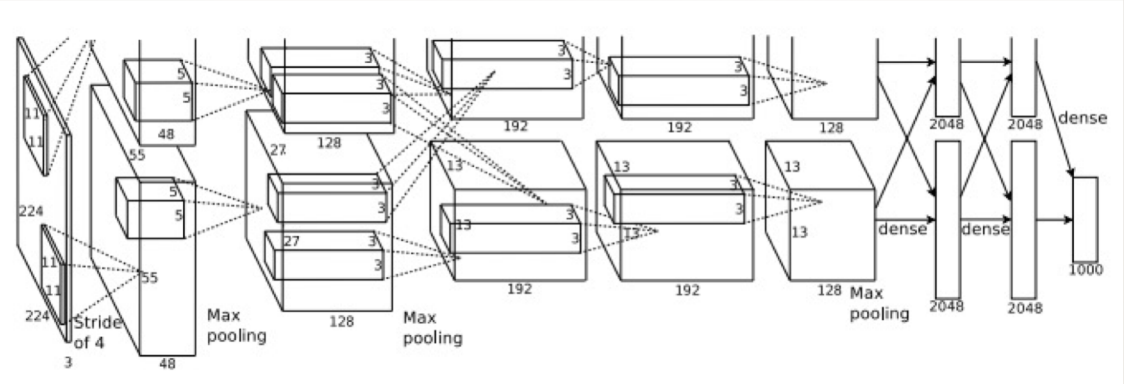
BREAKTHROUGHS

LMRD (2012)

- Group of researchers from Stanford and Google
- Used a neural network to tackle ImageNet
  - 16 million+ images, 20,000+ classifications
- 9.3% to 15.8%

KSH (2012)

- DCNN for a restricted subset of ImageNet in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC)
  - 84.7% for top-5, 63.3% for restrictive metric
  - Split on 2 GPUs



KSH MODEL ARCHITECTURE

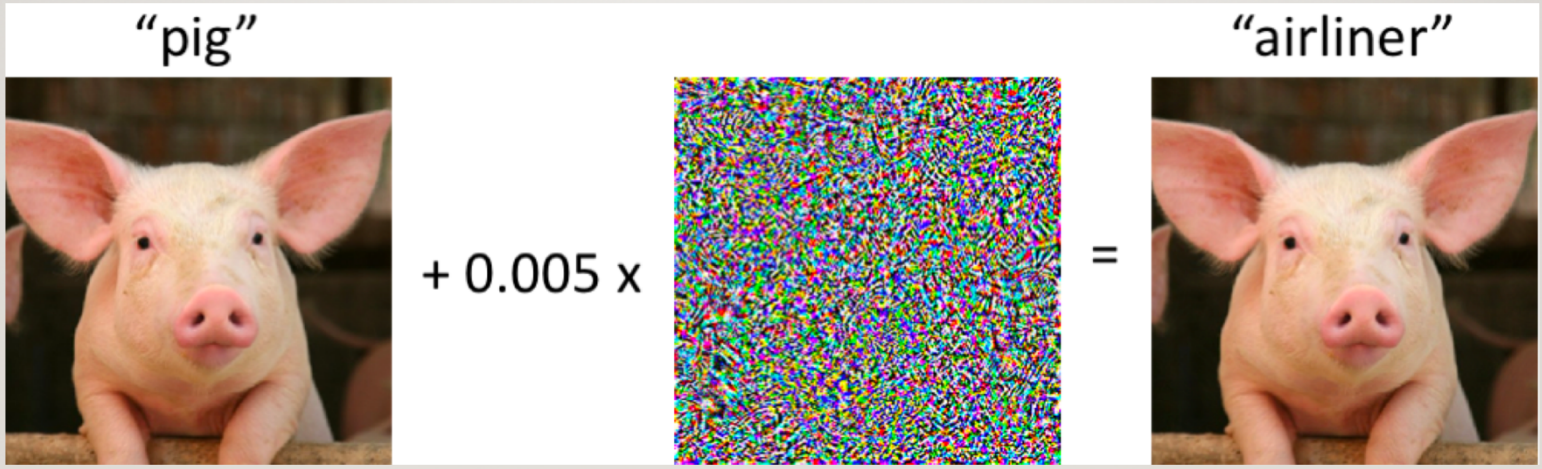
- Input is 3 x 244 x 244
  - Resize to 256 x 256, take 3 random 244 x 244 crops
- First layer: 96 11x11 kernels, stride of 4, 3x3 max pooling with stride 2
- Second layer: 256 5x5 kernels with max pooling
- Third, Fourth, Fifth: More convolutions without pooling
- Sixth, Seventh: Fully connected layers of 4096 neurons each
- Final: 1000-unit softmax layer
- Used ReLU, L2 regularization, dropout, momentum-based minibatch SGD
- Has inspired later work

ILSVRC (2014)

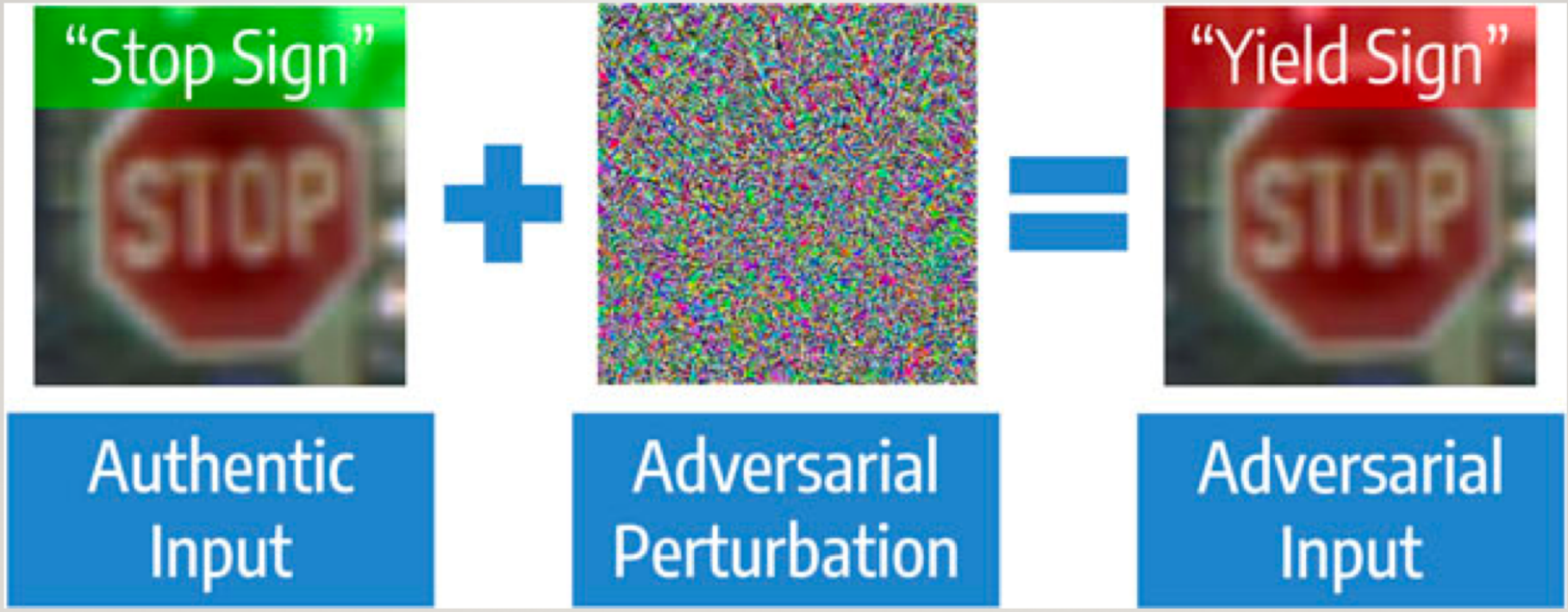
- 93.33% accuracy from team based at Google
- The dude actually sat down and labeled data
- Better-than-human vision

OTHER STUFF

ADVERSARIAL EXAMPLES

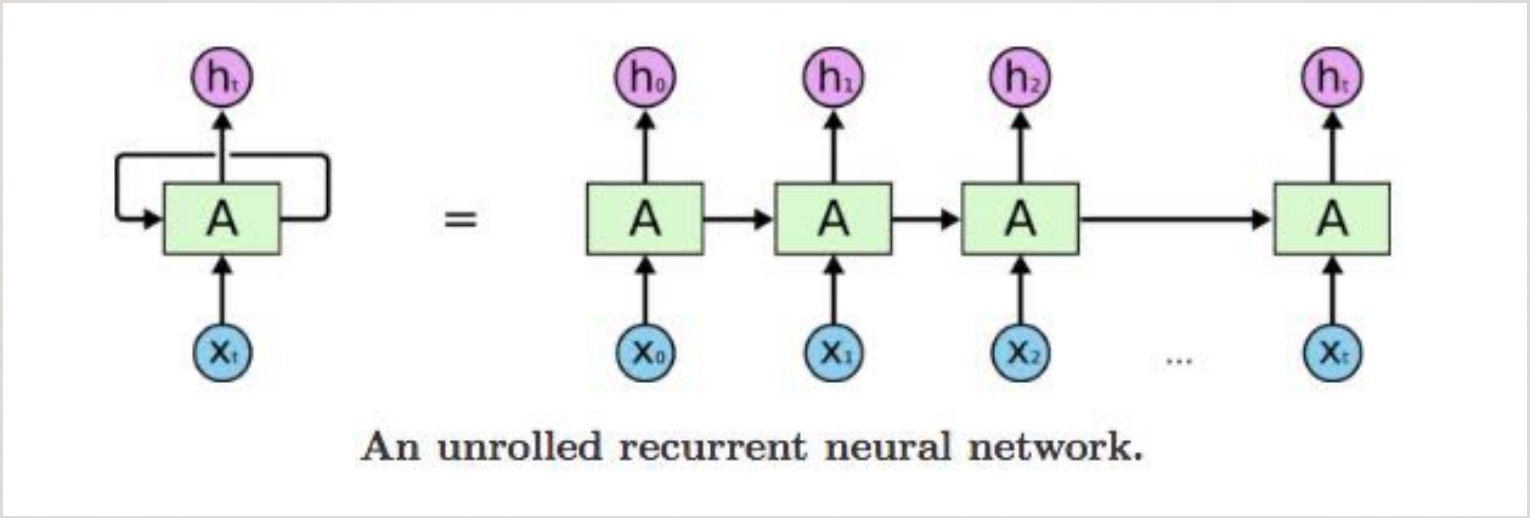



People with no idea about AI
saying it will take over the world:
My Neural Network:
Dog

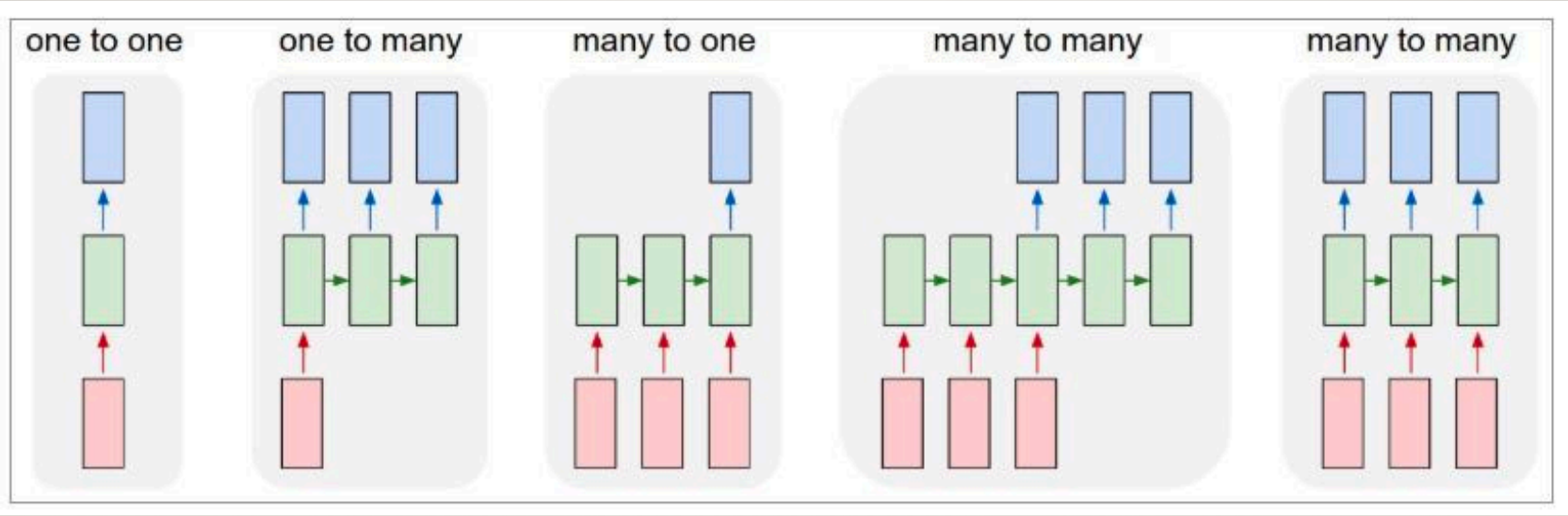

"Stop Sign"
STOP
"Yield Sign"
STOP
Authentic Input
Adversarial Perturbation
Adversarial Input

RNNS

- Just like CNNs add spatial understanding, RNNs add temporal understanding

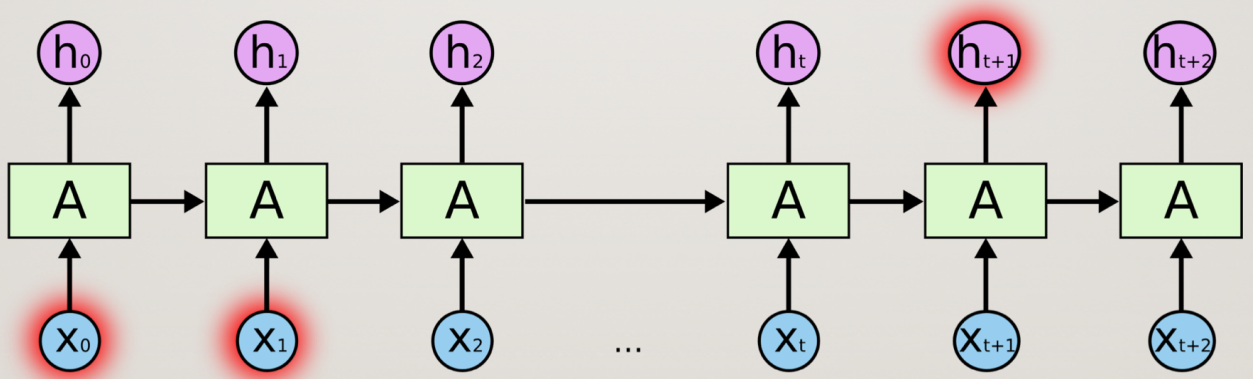


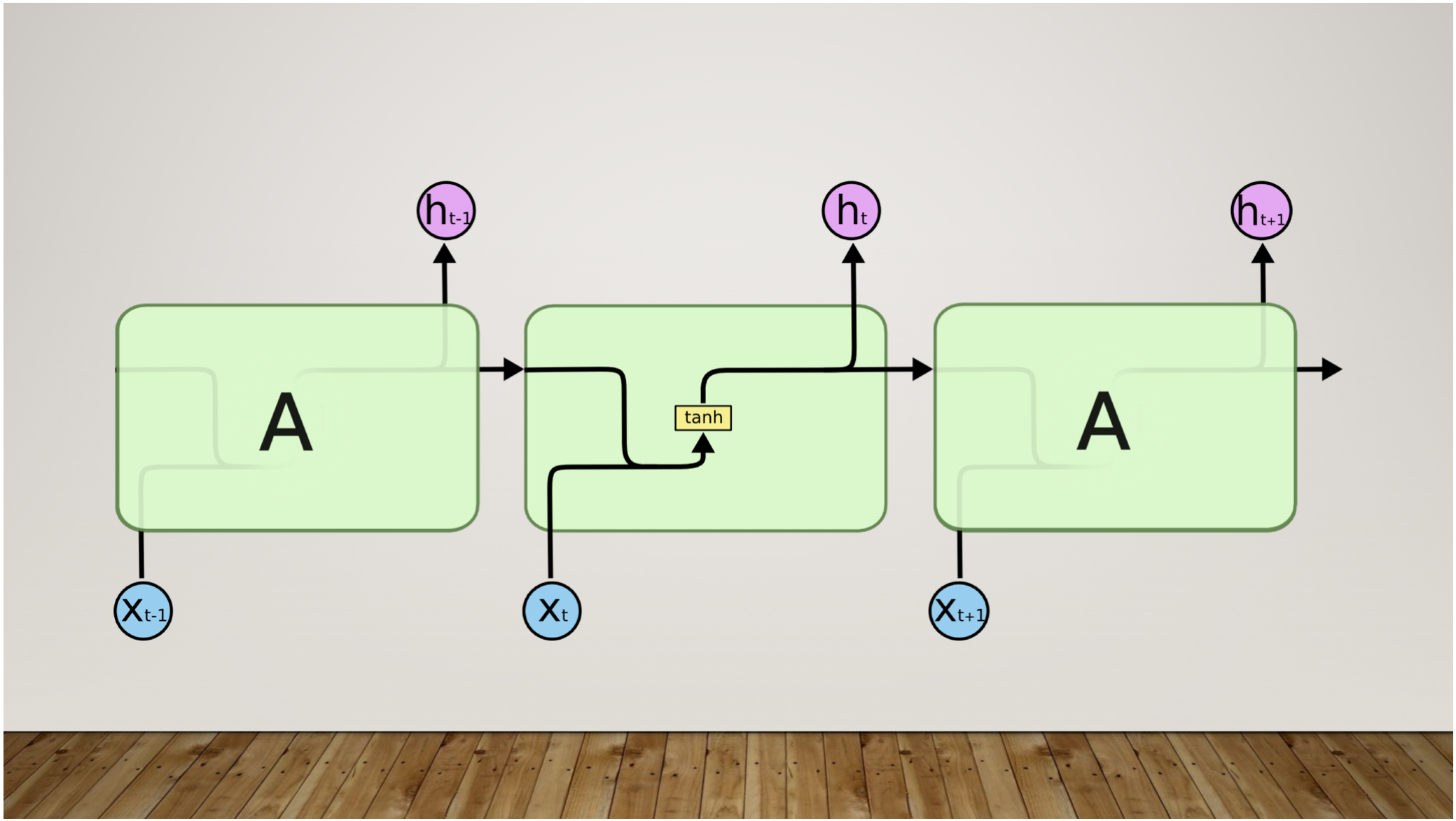
An unrolled recurrent neural network.

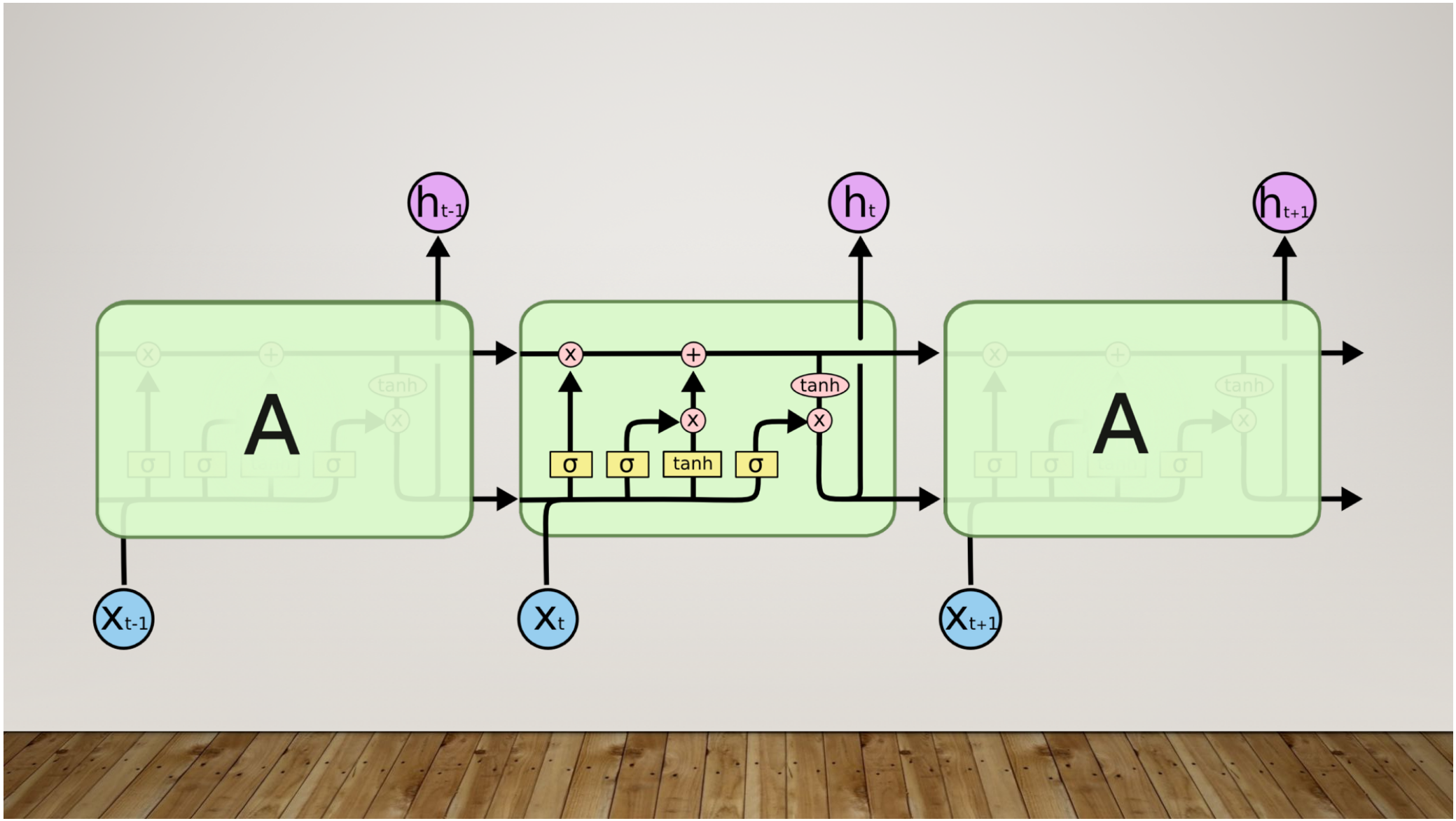

one to one
one to many
many to one
many to many
many to many

LSTMS

- RNNs take soooo long to train b/c they have unstable gradients
- "I grew up in France… I speak fluent *French*." (Long-term dependency problem)




h_{t-1}
h_t
h_{t+1}
A
tanh
A
X_{t-1}
X_t
X_{t+1}


h_{t-1}
h_t
h_{t+1}
A
A
tanh
σ
σ
tanh
σ
tanh
X_{t-1}
X_t
X_{t+1}

DBNS, GENERATIVE MODELS

- Deep Belief Network
  - Generative model
  - Unsupervised and semi-supervised
- GANs

INTENTION-DRIVEN USER INTERFACES

- Interfaces that can act on imprecision and discern the user's true intent

FUTURE OF NN?

- They've done a lot of amazing things recently
- But we don't understand them nearly well enough
  - Why is it that neural networks can generalize so well?
  - How is it that they avoid overfitting as well as they do, given the very large number of parameters they learn?
  - Why is it that stochastic gradient descent works as well as it does? How well will neural networks perform as data sets are scaled?
  - Why does pooling work?

CONWAY'S LAW

- "Any organization that designs a system... will inevitably produce a design whose structure is a copy of the organization's communication structure."
- Applies to the design and engineering of systems where we have a good understanding of the different parts
- Can't be applied directly to the development of AI because we don't know what the parts are.

FINAL THOUGHTS

- Medicine -> immunology, epidemiology, etc
- Deep learning is our "super-special weapon"
  - How powerful is it?
  - What other powerful idea will be needed for strong AI?
- We don't see a lot of specialized subfields yet, everything's built off the same ideas