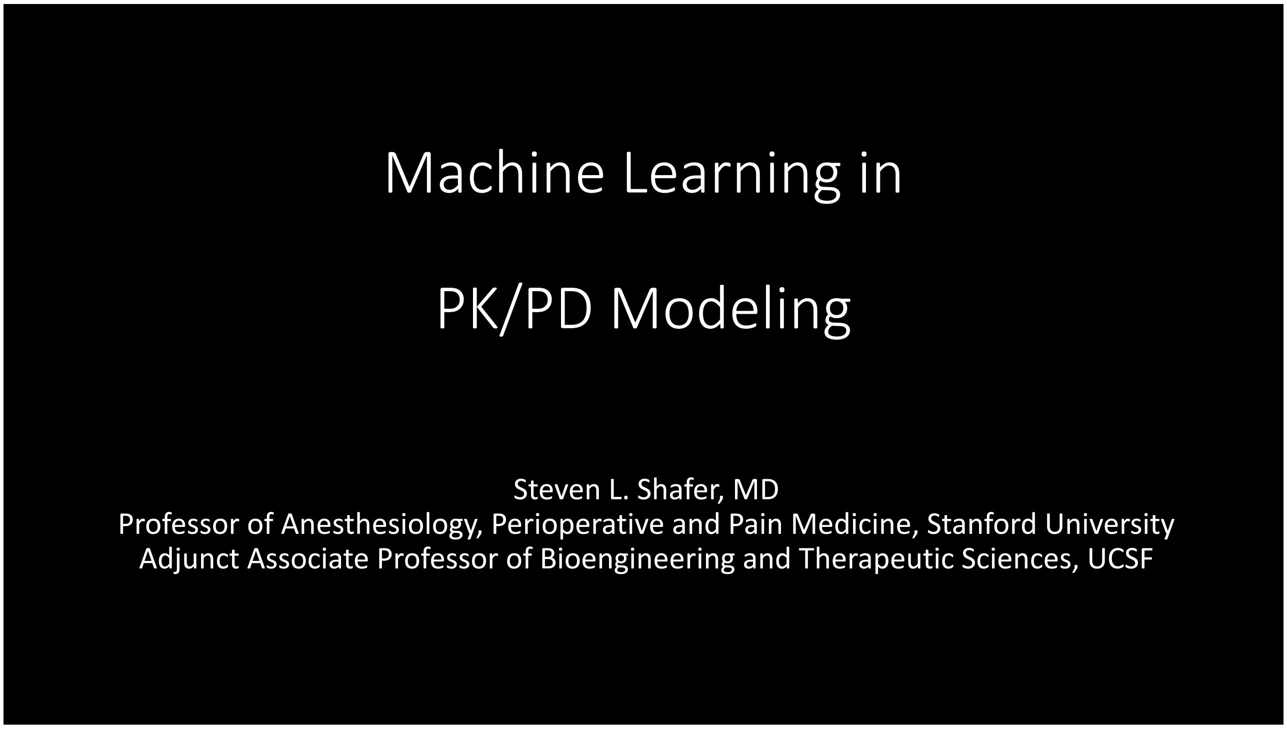
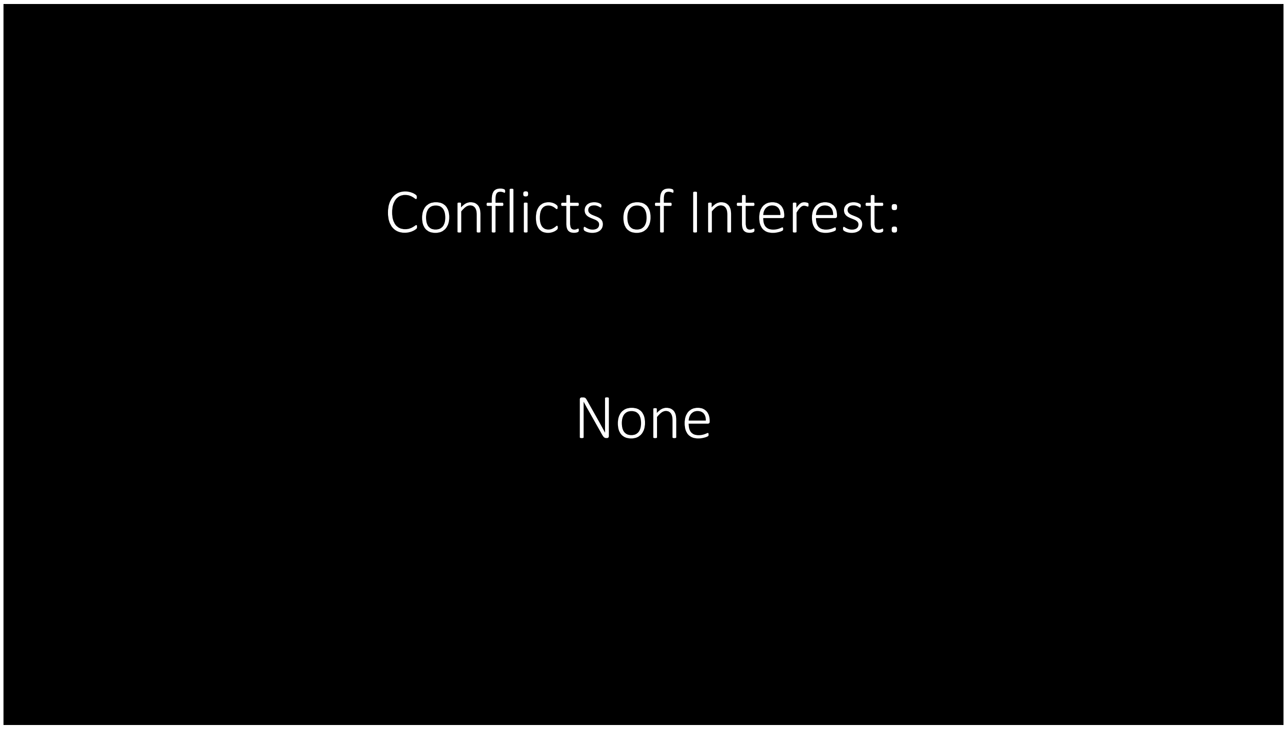


Machine Learning in

PK/PD Modeling

Steven L. Shafer, MD
Professor of Anesthesiology, Perioperative and Pain Medicine, Stanford University
Adjunct Associate Professor of Bioengineering and Therapeutic Sciences, UCSF



Conflicts of Interest:

None

MIT TECHNOLOGY REVIEW

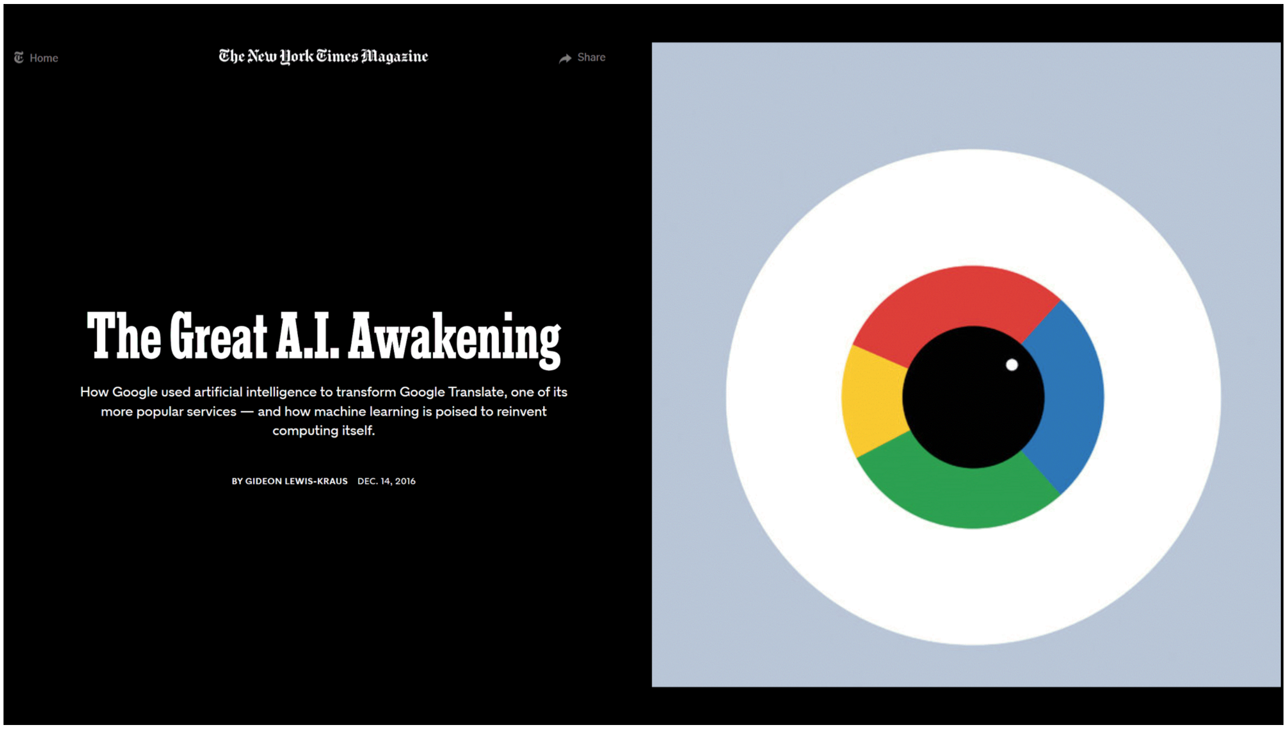
Business Impact

Google Stakes Its Future on a Piece of Software

Alphabet, number 5 on our list of the 50 Smartest Companies, thinks it can wrest the cloud computing market away from Amazon by helping companies make use of machine learning with a tool called TensorFlow.

by Tom Simonite June 27, 2017

TensorFlow



The Great A.I. Awakening

How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself.

BY GIDEON LEWIS-KRAUS DEC. 14, 2016

Prof. Jun Rekimoto, Ph.D

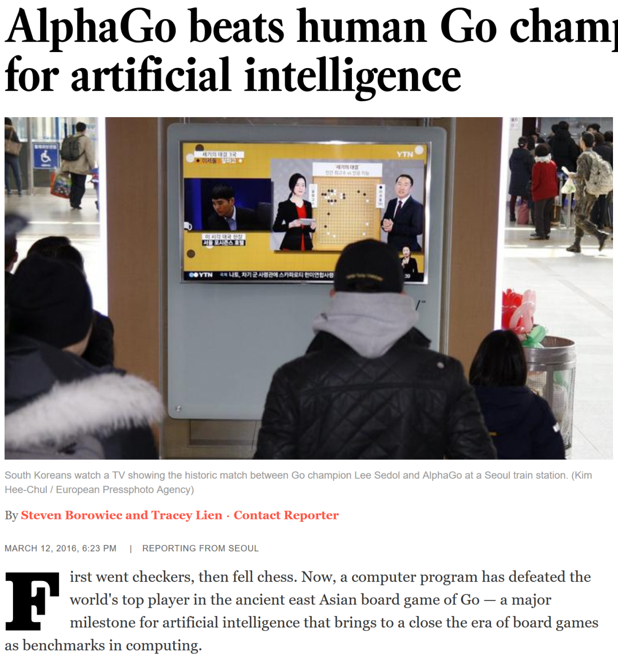
Professor, Interfaculty Initiative in Information Studies,
The University of Tokyo

Kilimanjaro is a mountain of 19,710 feet covered with snow and is said to be the highest mountain in Africa. The summit of the west is called “Ngaje Ngai” in Masai, the house of God. Near the top of the west there is a dry and frozen dead body of leopard. No one has ever explained what leopard wanted at that altitude.

Kilimanjaro is a snow-covered mountain 19,710 feet high, and is said to be the highest mountain in Africa. Its western summit is called the Masai “Ngaje Ngai,” the House of God. Close to the western summit there is the dried and frozen carcass of a leopard. No one has explained what the leopard was seeking at that altitude.

This is an example of Google Translate. I have no idea if the translation is any good, but I thought I would test it and let you be the judge. Does the language seem natural?

Dit is een voorbeeld van Google Translate. Ik heb geen idee of de vertaling goed is, maar ik dacht dat ik het zou testen en je de rechter zou laten zijn. Lijkt de taal natuurlijk?



AlphaGo beats human Go champ in milestone for artificial intelligence

South Koreans watch a TV showing the historic match between Go champion Lee Sedol and AlphaGo at a Seoul train station. (Kim Hee-Chul / European Pressphoto Agency)

By Steven Borowiec and Tracey Lien · Contact Reporter

MARCH 12, 2016, 6:23 PM | REPORTING FROM SEOUL

First went checkers, then fell chess. Now, a computer program has defeated the world's top player in the ancient east Asian board game of Go — a major milestone for artificial intelligence that brings to a close the era of board games as benchmarks in computing.

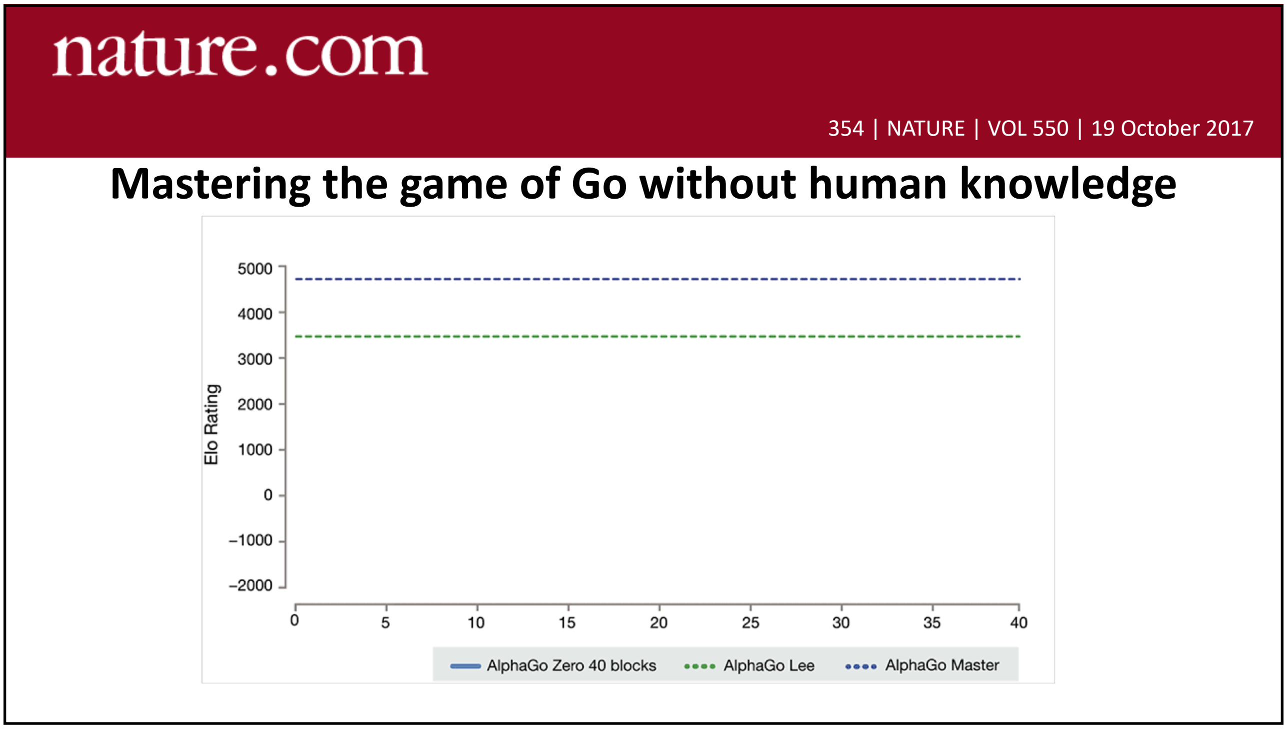
Monte Carlo tree search of neural networks.

Neural networks boot-strapped from human gameplay expertise.

Trained further by playing games against other instances of itself

AlphaGo Zero

Starting from scratch



nature.com

354 | NATURE | VOL 550 | 19 October 2017

Mastering the game of Go without human knowledge

Google's AlphaZero Destroys Stockfish In 100-Game Match

Mike Klein
Dec 6, 2017, 12:50 PM

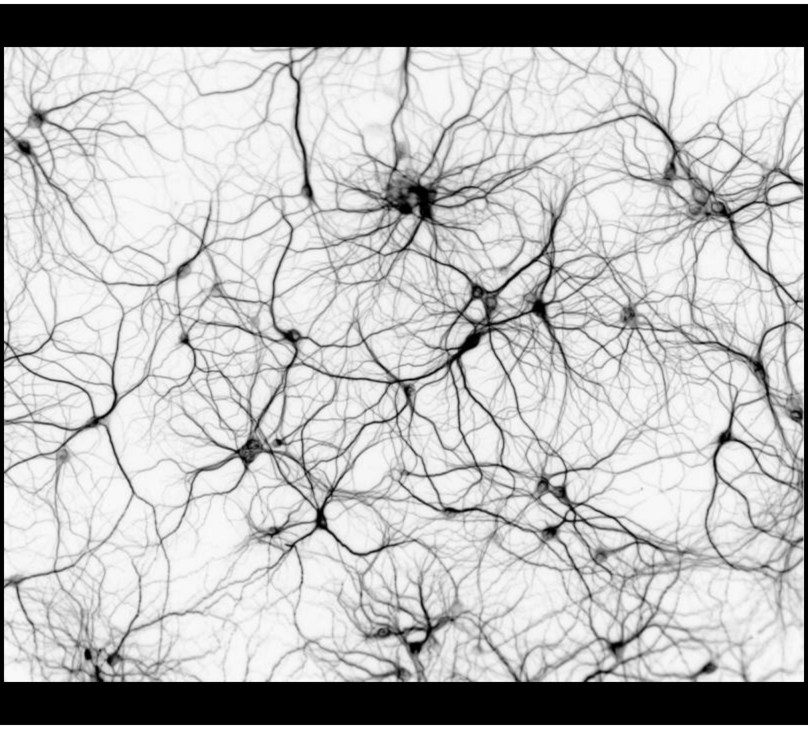
Chess changed forever today. And maybe the rest of the world did, too.

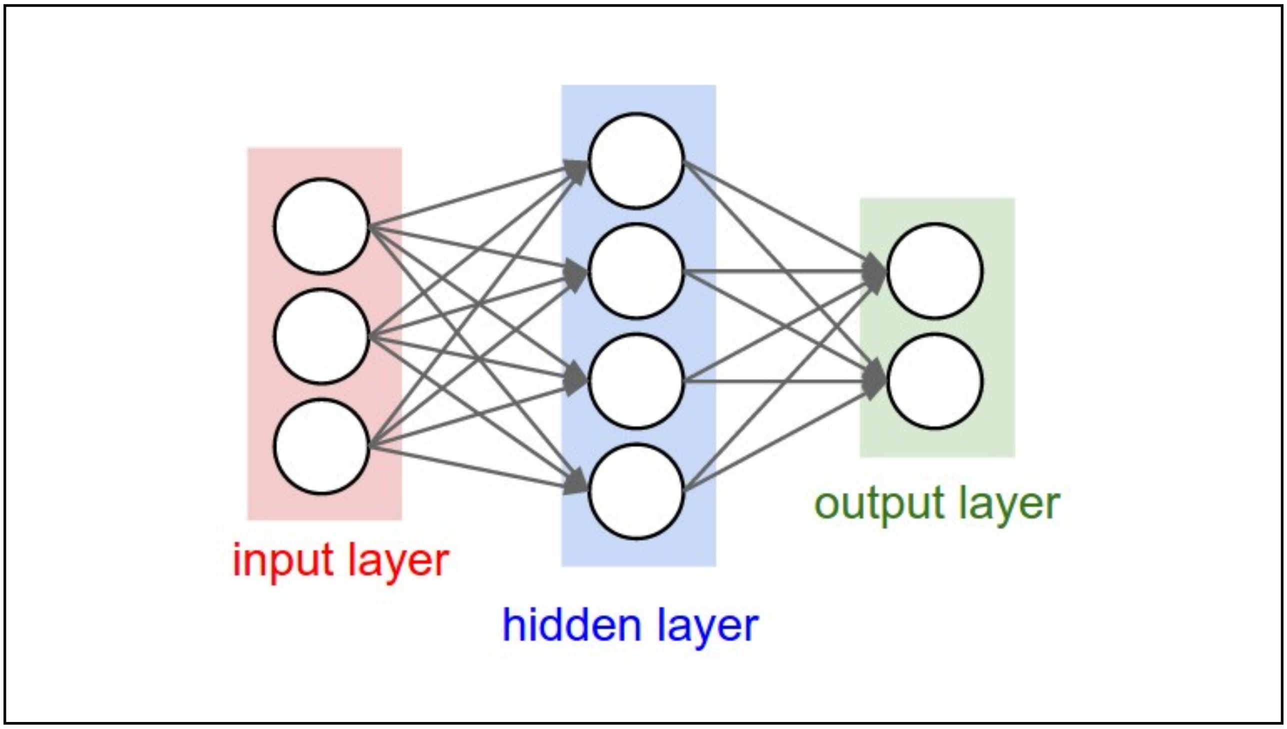
A little more than a year after AlphaGo sensationally won against the top Go player, the artificial-intelligence program AlphaZero has obliterated the highest-rated chess engine.

Stockfish, which for most top players is their go-to preparation tool, and which won the 2016 TCEC Championship and the 2017 Chess.com Computer Chess Championship, didn't stand a chance. AlphaZero won the closed-door, 100-game match with 28 wins, 72 draws, and zero losses.

Oh, and it took AlphaZero only four hours to "learn" chess. Sorry humans, you had a good run.

That's right -- the programmers of AlphaZero, housed within the DeepMind division of Google, had it use a type of "machine learning," specifically reinforcement learning. Put more plainly, AlphaZero was not "taught" the game in the traditional sense. That means no opening book, no endgame tables, and apparently no complicated algorithms dissecting minute differences between center pawns and side pawns.




input layer
hidden layer
output layer

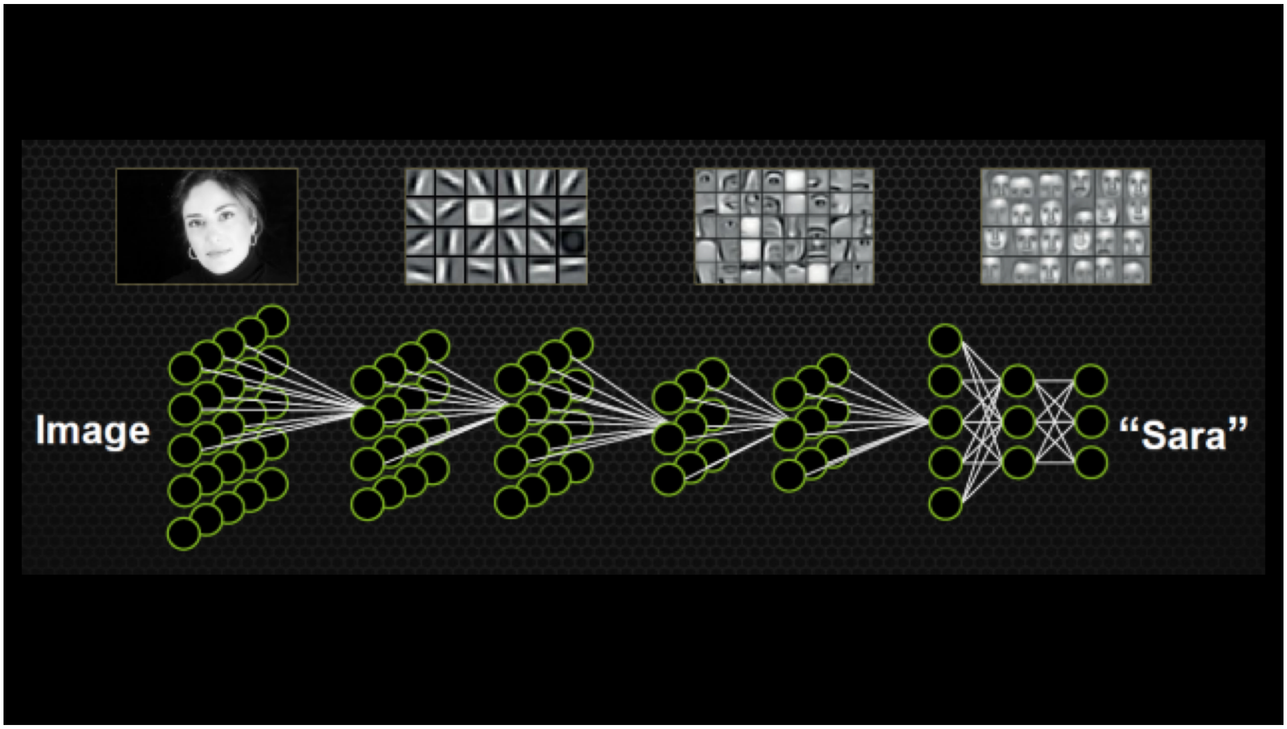
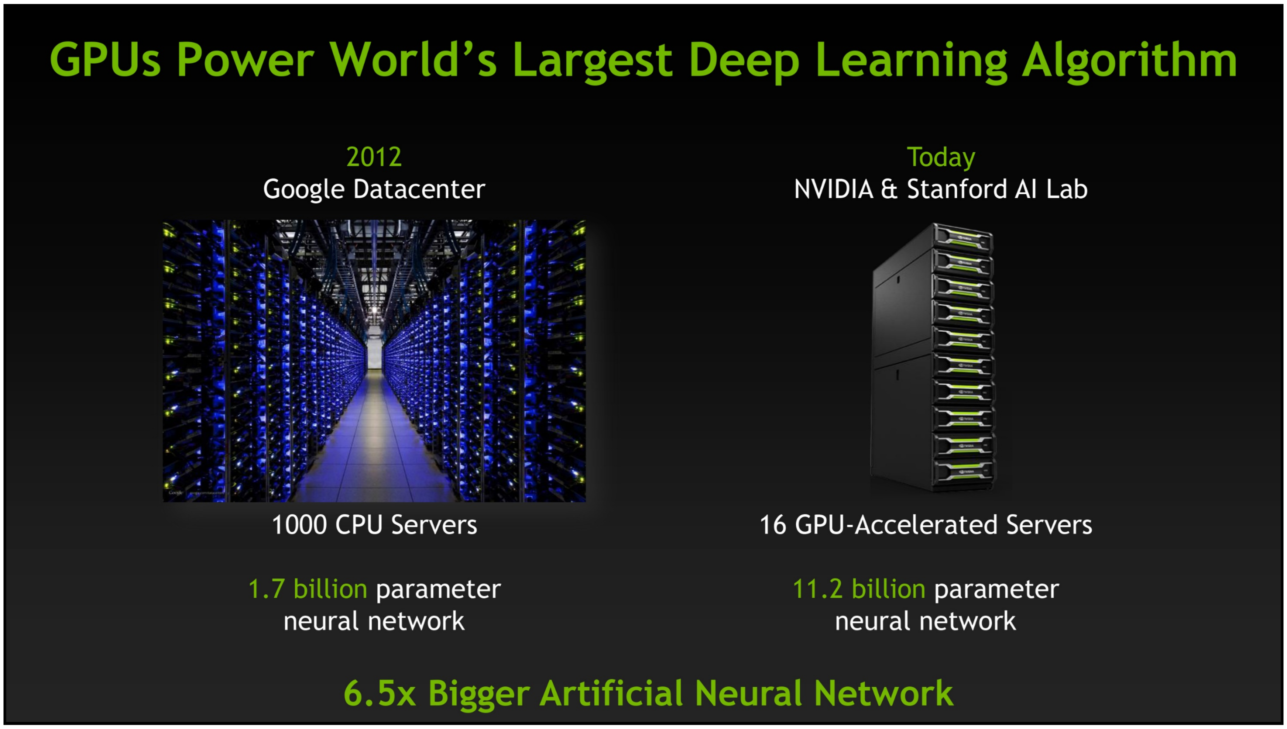


Image
"Sara"



GPUs Power World’s Largest Deep Learning Algorithm

2012
Google Datacenter

1000 CPU Servers

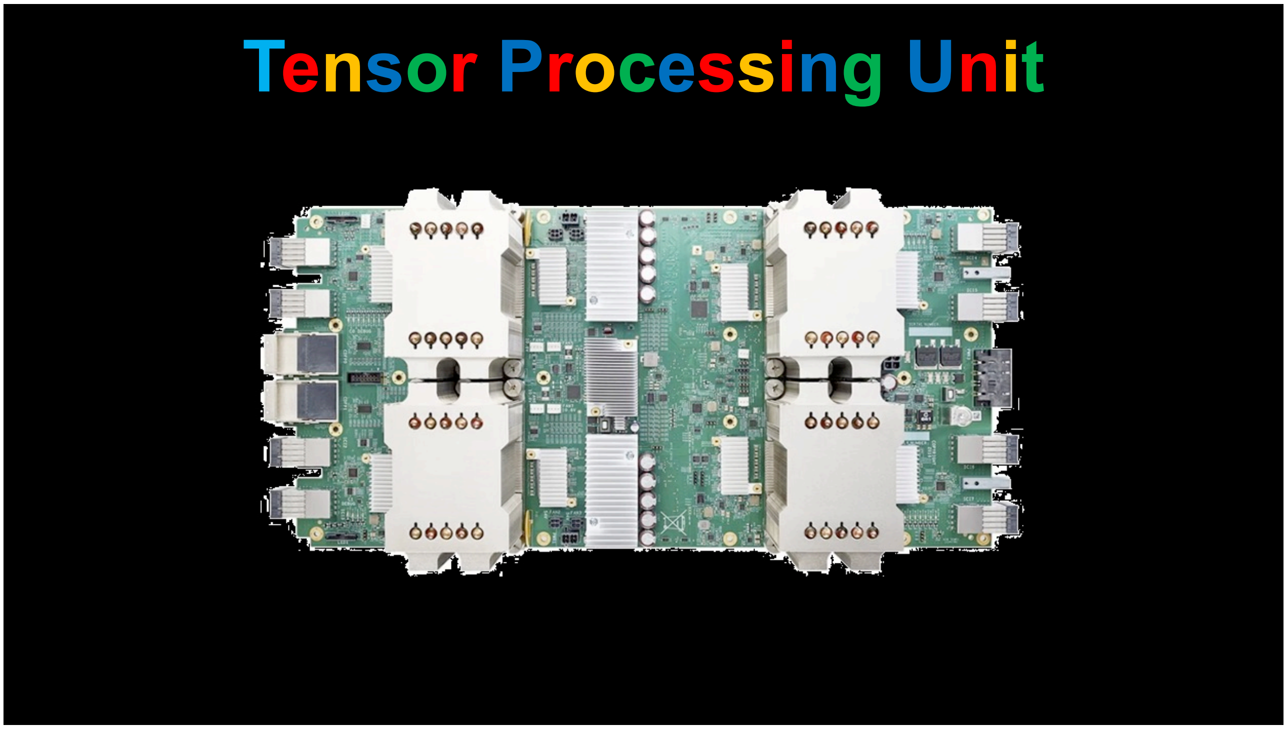
1.7 billion parameter
neural network

Today
NVIDIA & Stanford AI Lab

16 GPU-Accelerated Servers

11.2 billion parameter
neural network

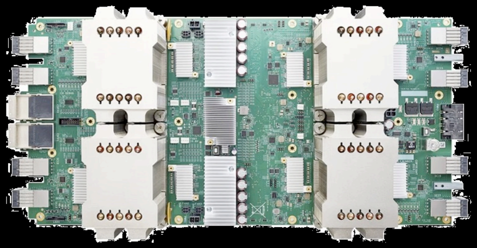
6.5x Bigger Artificial Neural Network



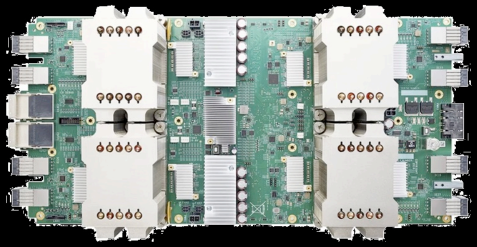
Tensor Processing Unit

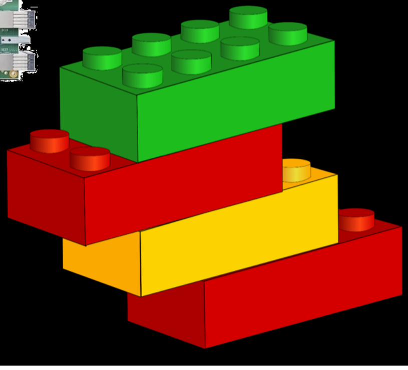
TensorFlow™

TensorFlow™



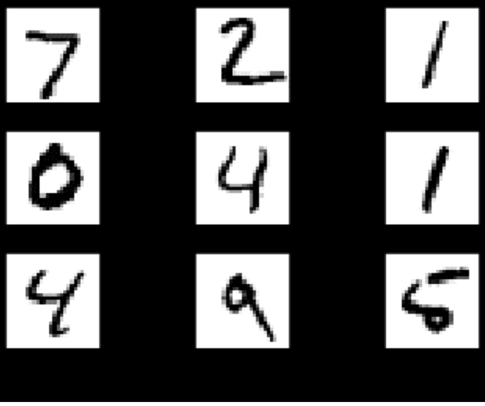
TensorFlow ™

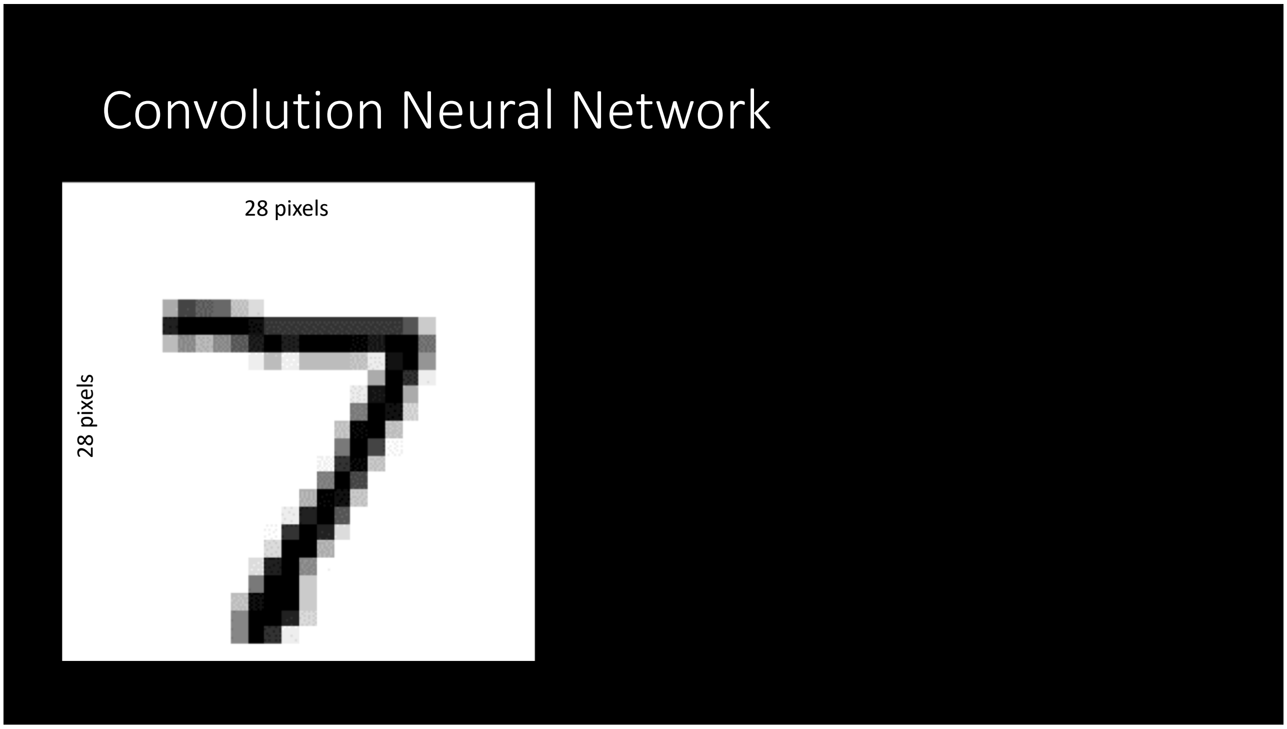




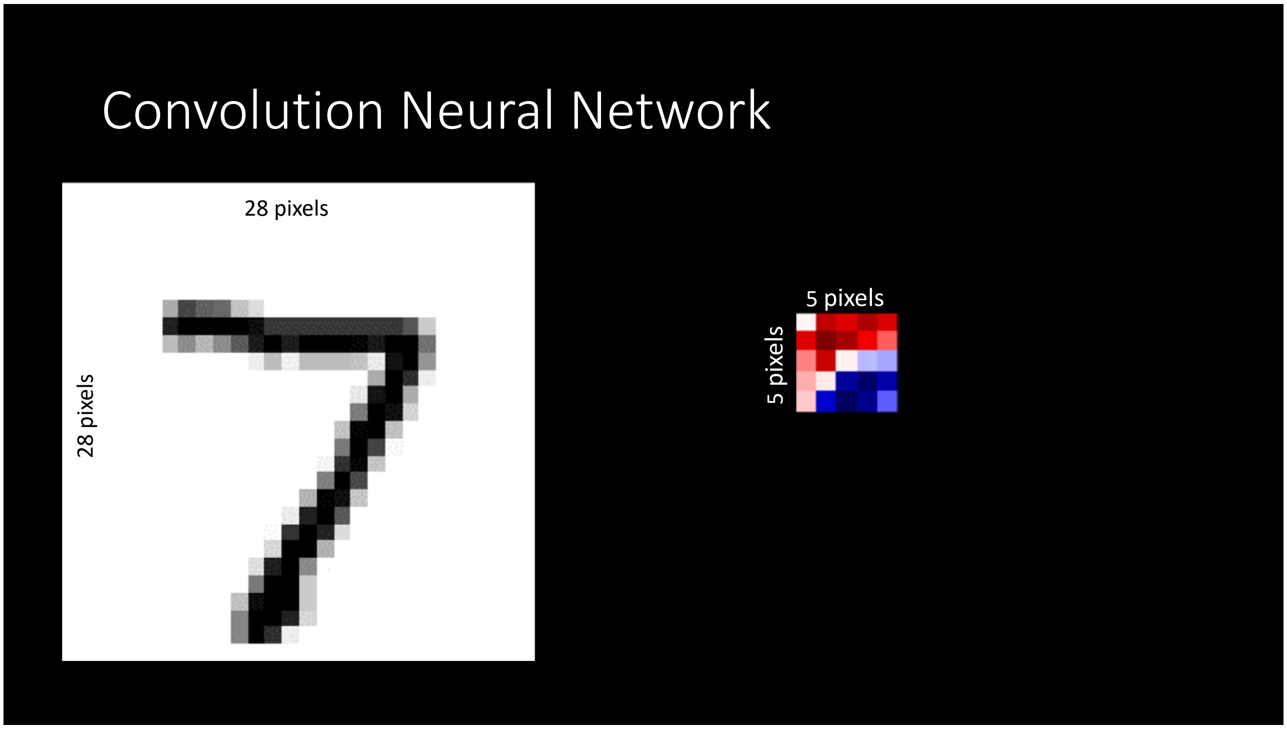
MNIST EXAMPLE

- National Institute of Standards and Technology
- 60,000 handwritten digits, 0-9



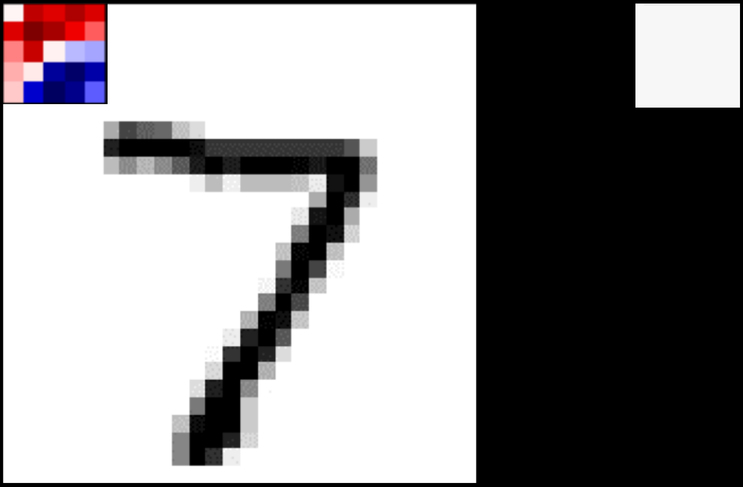


Convolution Neural Network

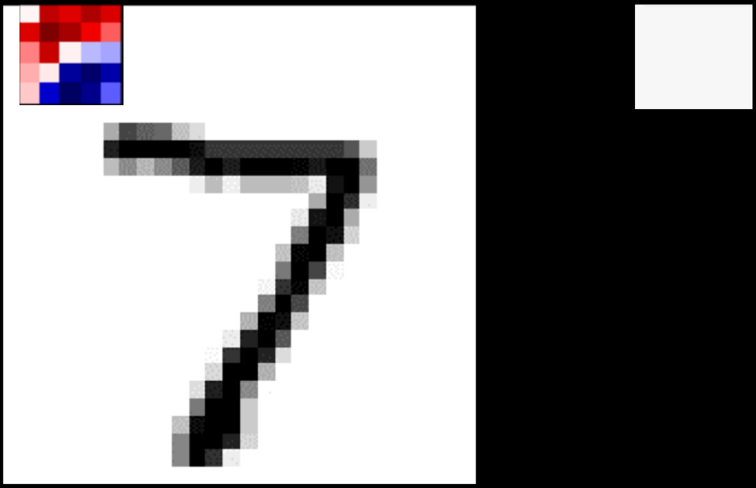


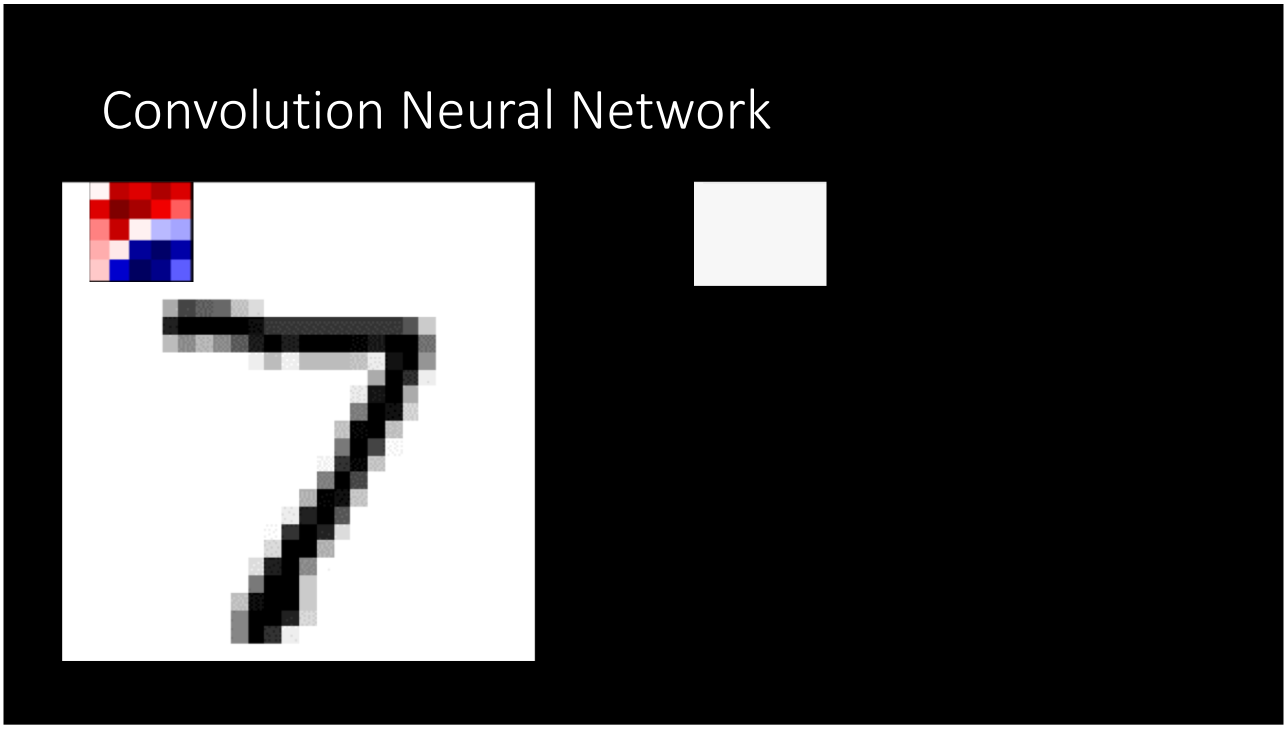
Convolution Neural Network

Convolution Neural Network

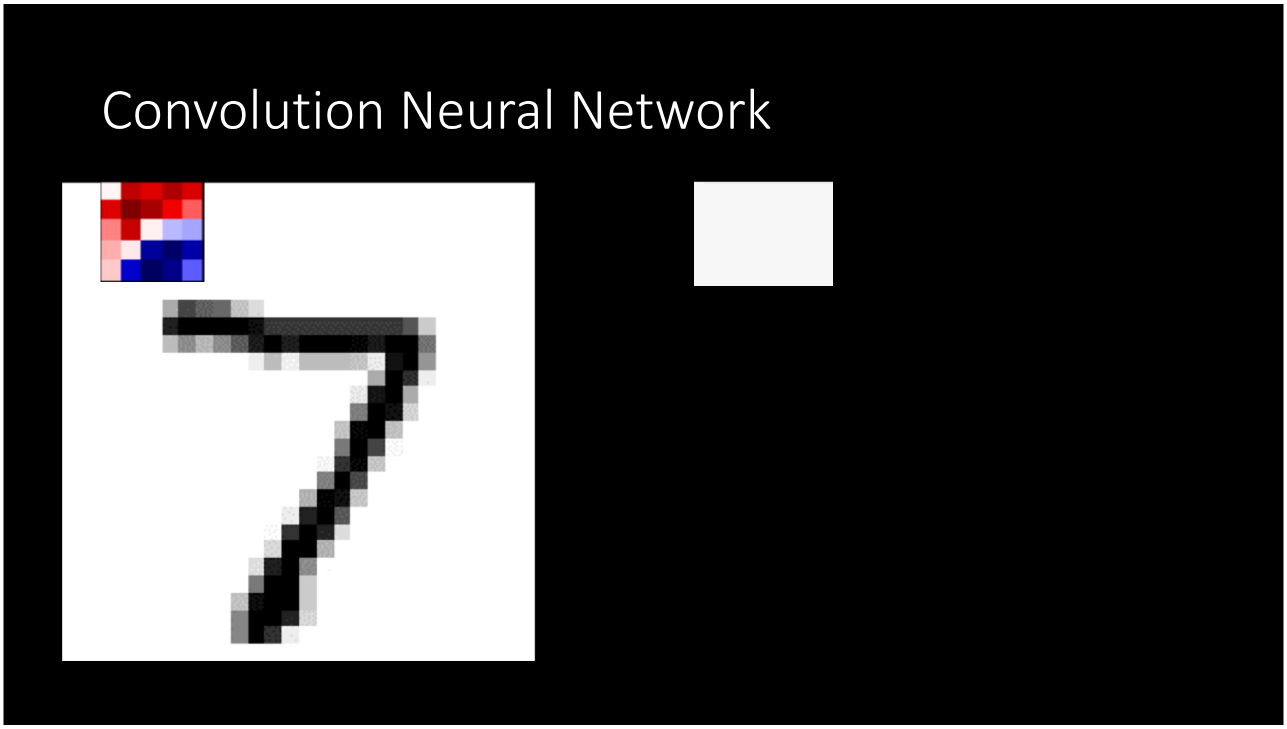


Convolution Neural Network



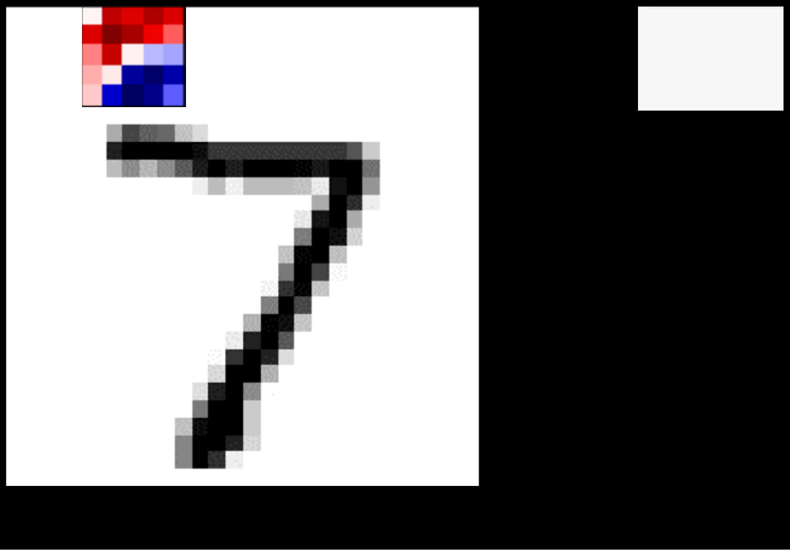


Convolution Neural Network

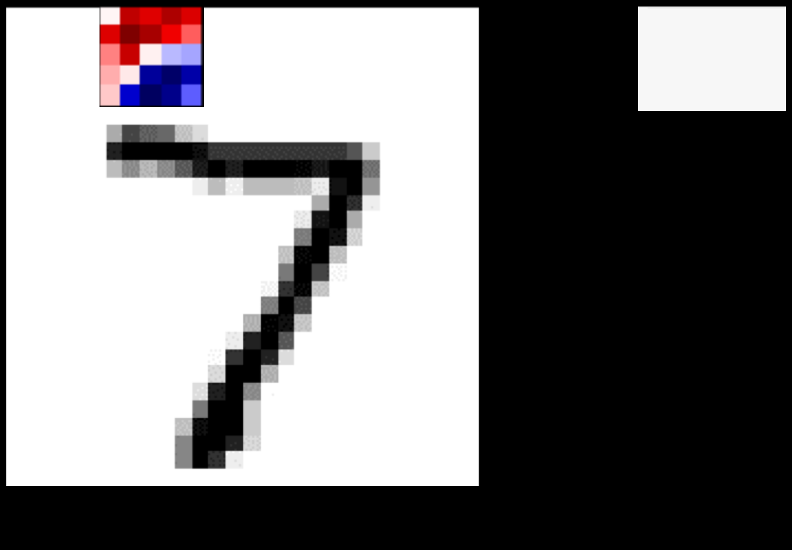


Convolution Neural Network

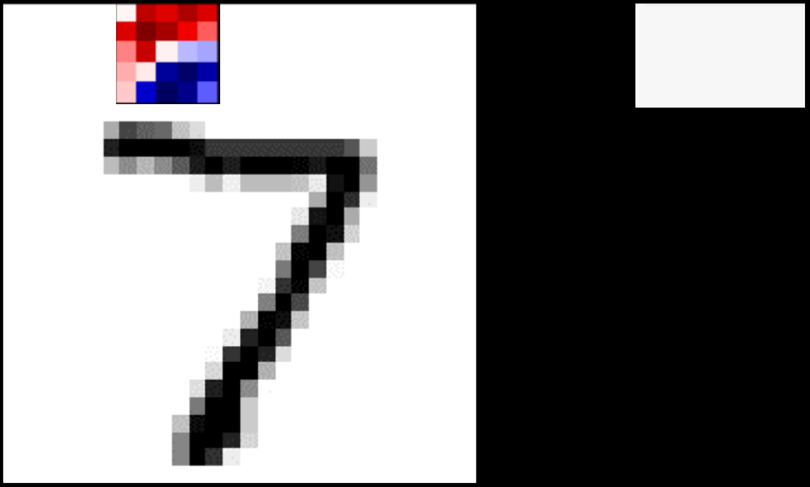
Convolution Neural Network



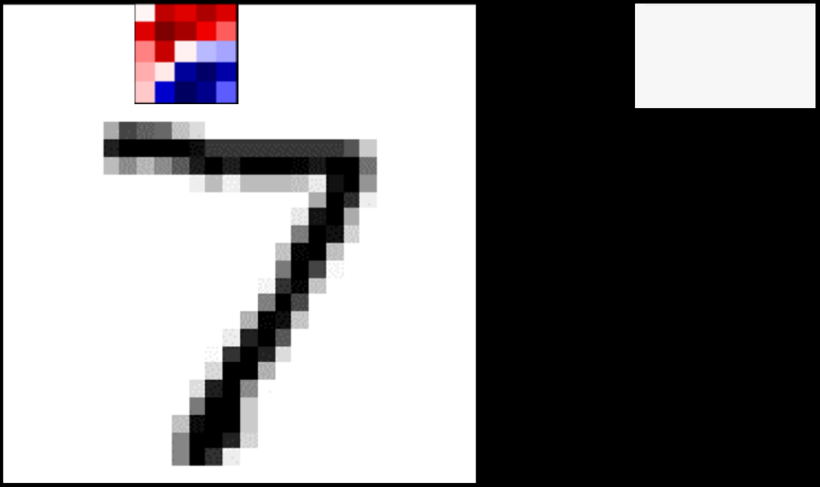
Convolution Neural Network



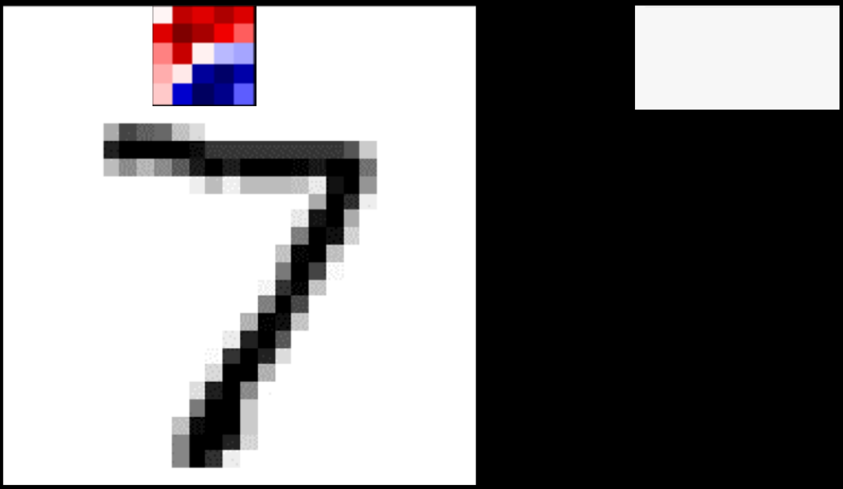
Convolution Neural Network



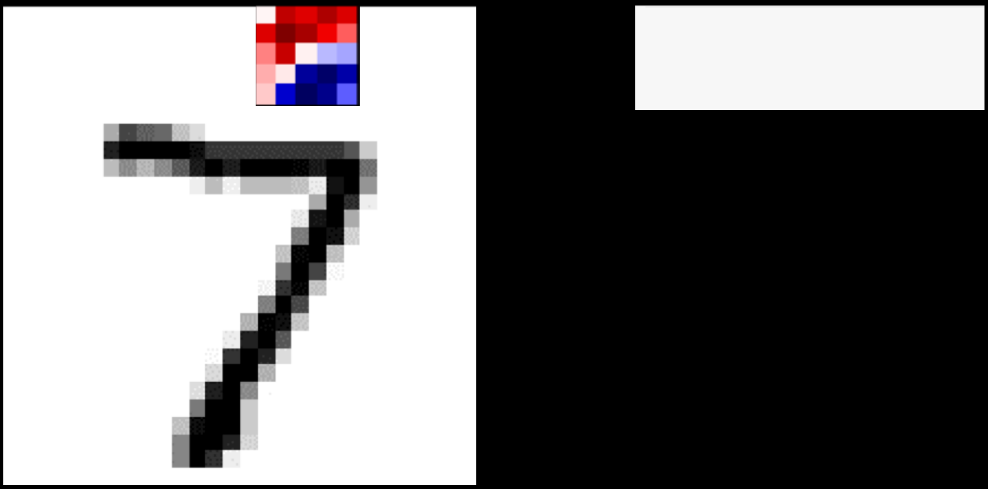
Convolution Neural Network

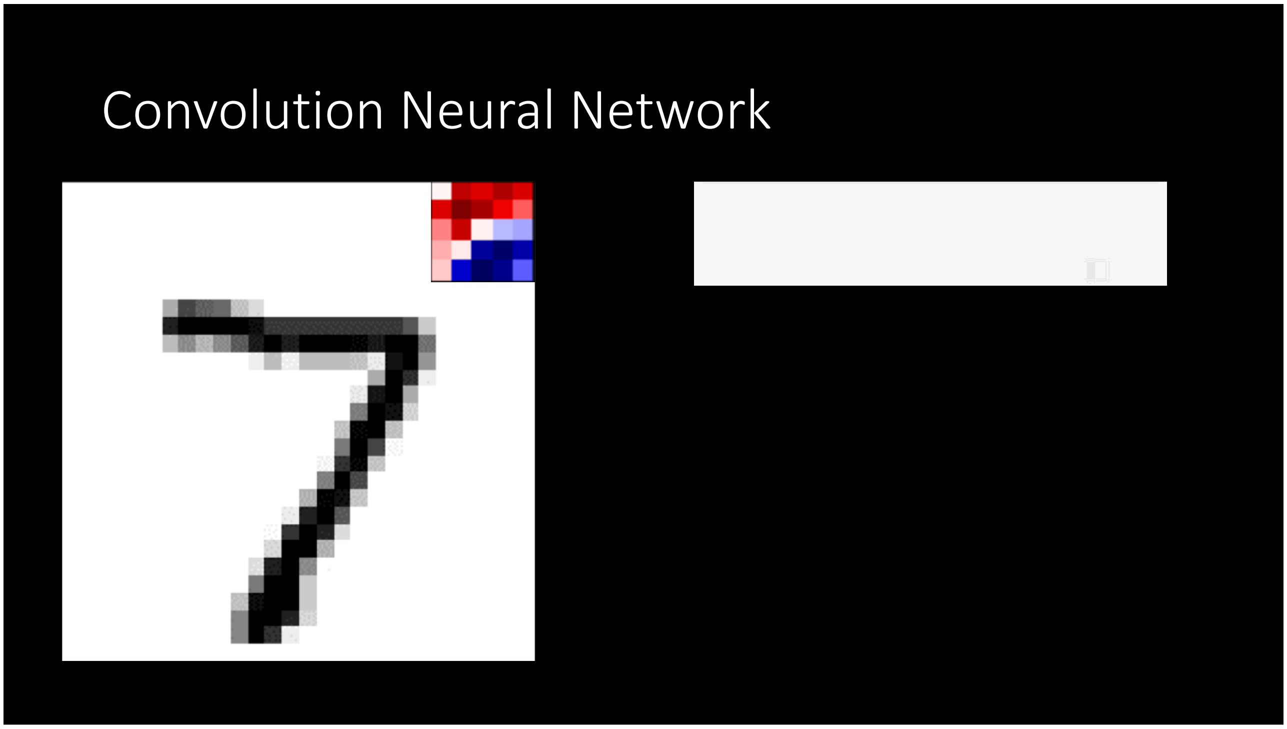


Convolution Neural Network

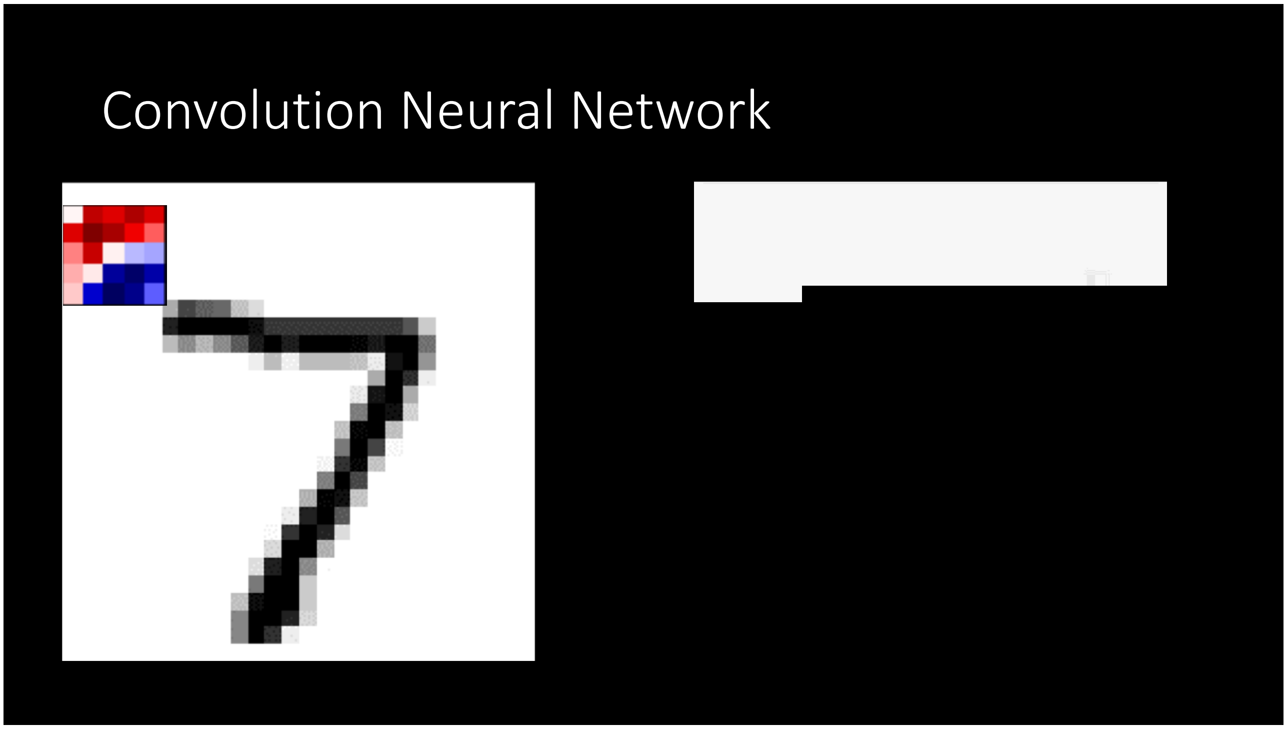


Convolution Neural Network

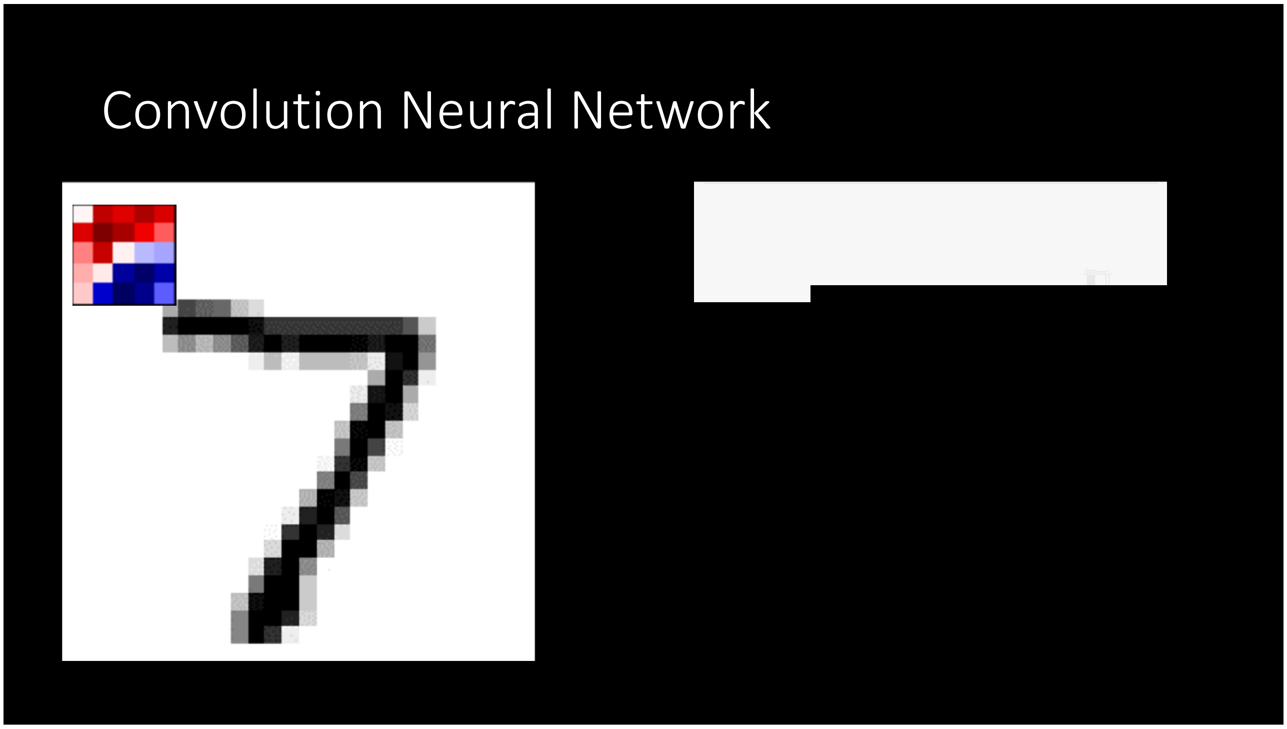




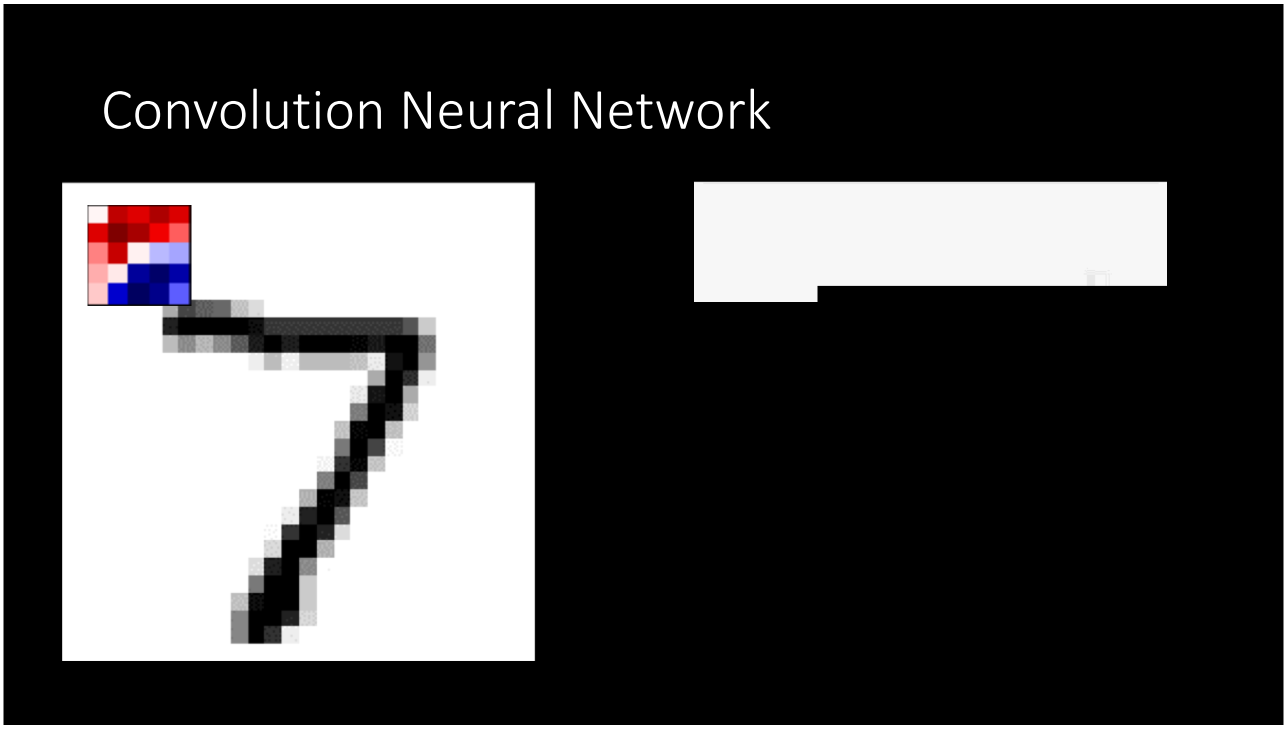
Convolution Neural Network



Convolution Neural Network

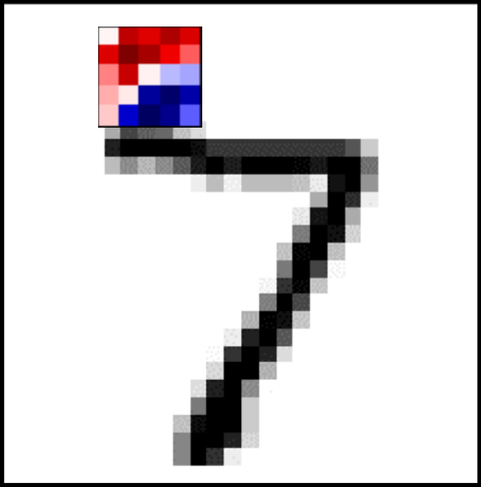


Convolution Neural Network

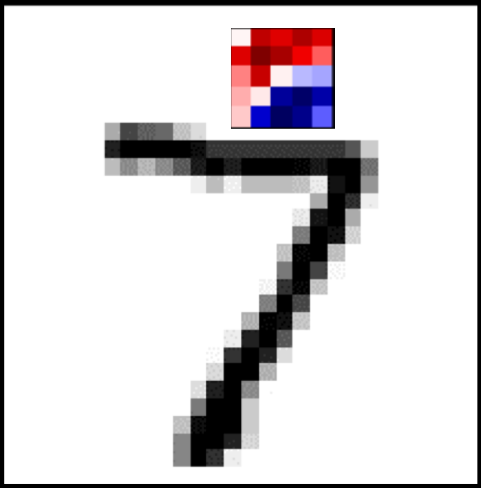


Convolution Neural Network

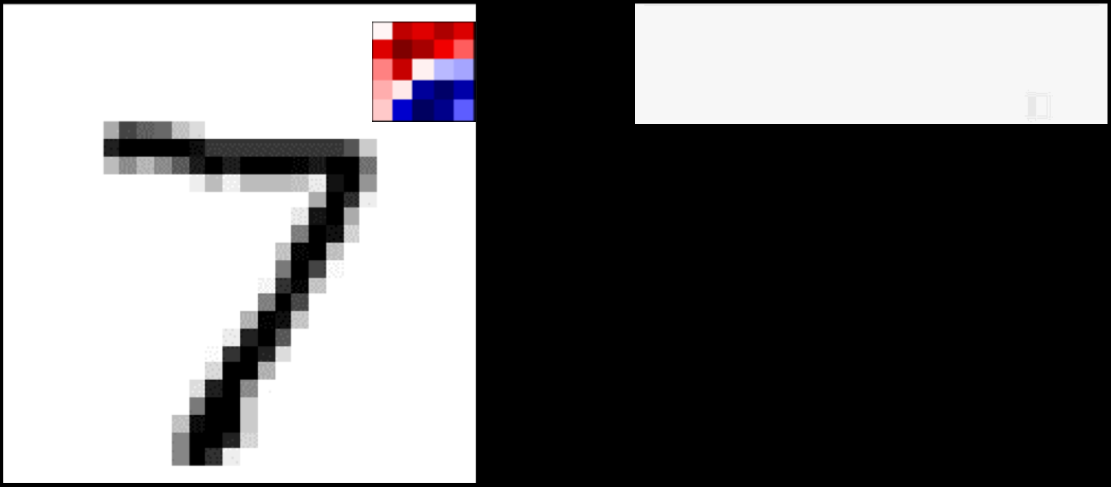
Convolution Neural Network



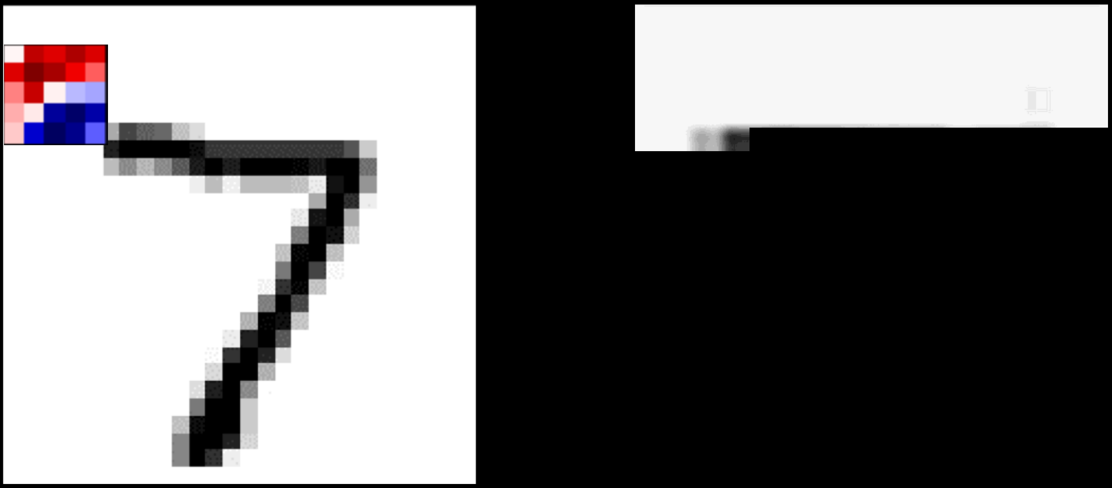
Convolution Neural Network

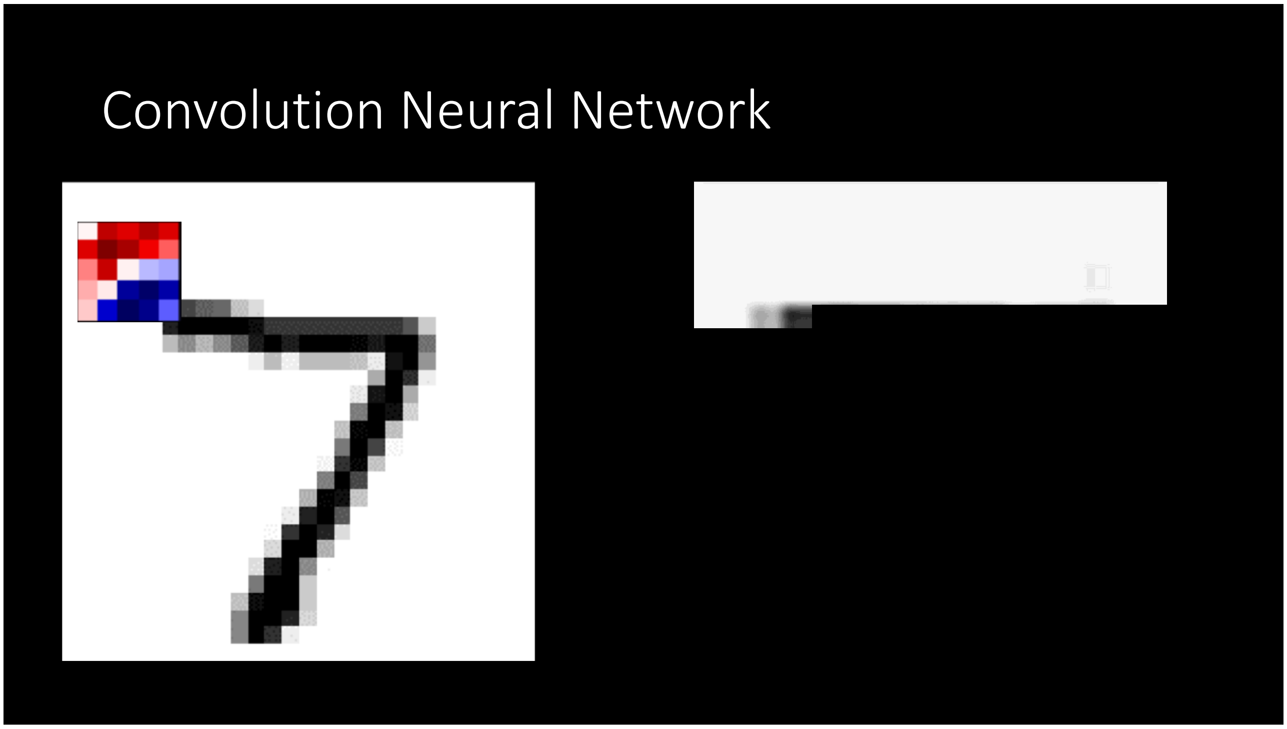


Convolution Neural Network

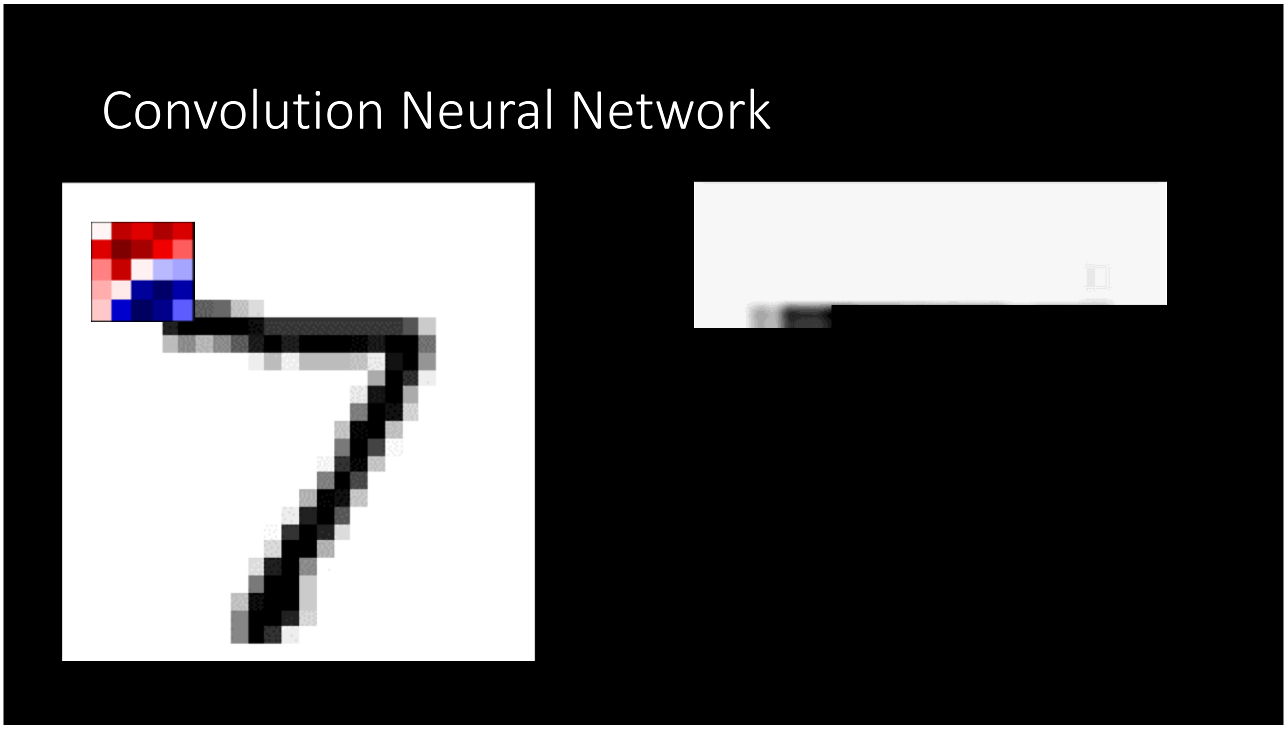


Convolution Neural Network

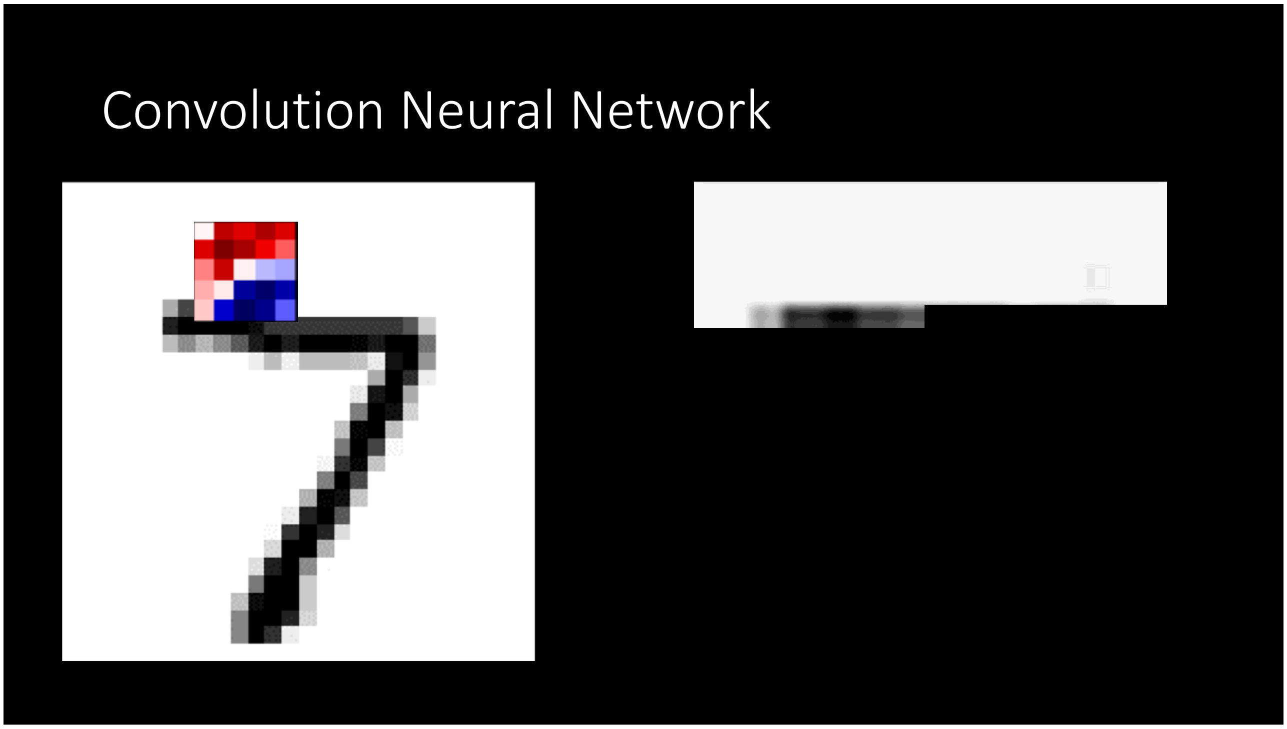




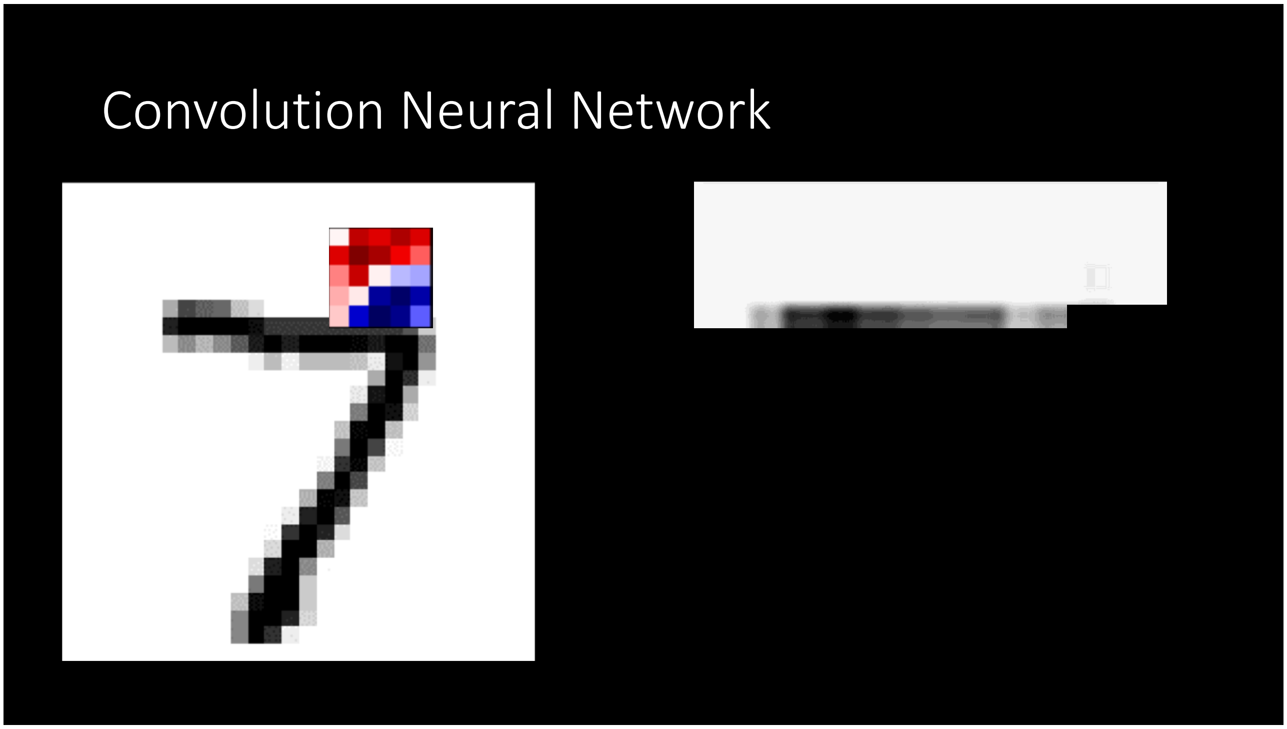
Convolution Neural Network



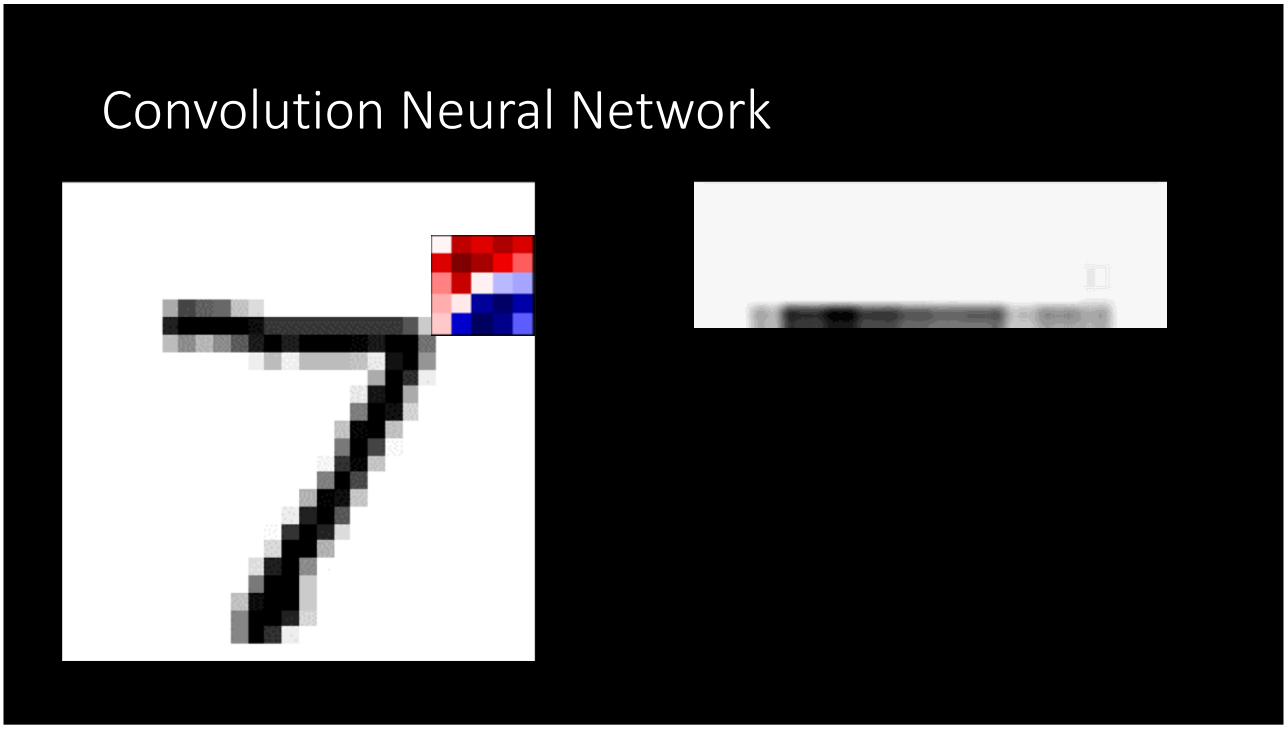
Convolution Neural Network



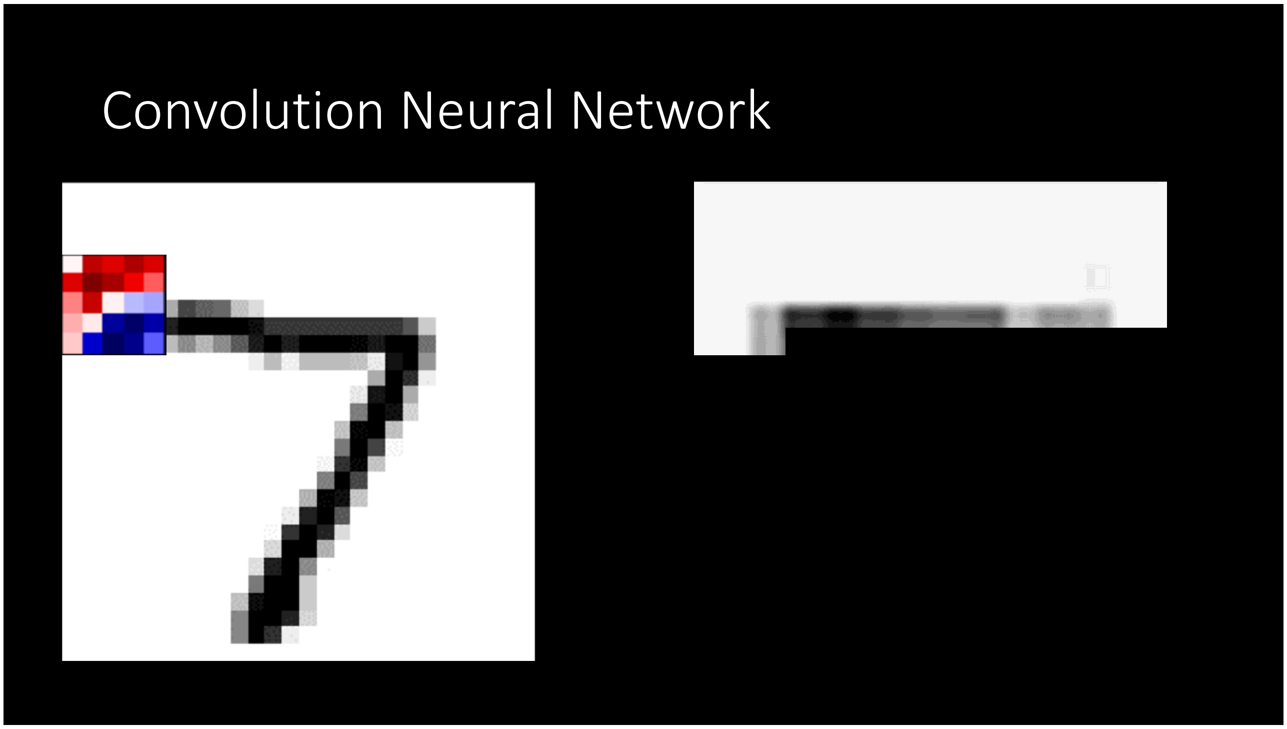
Convolution Neural Network



Convolution Neural Network

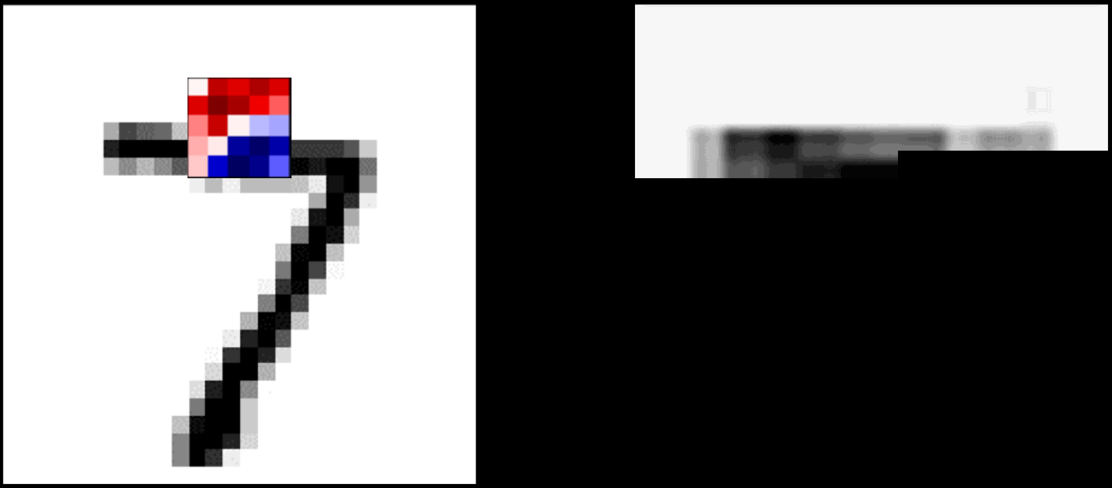


Convolution Neural Network

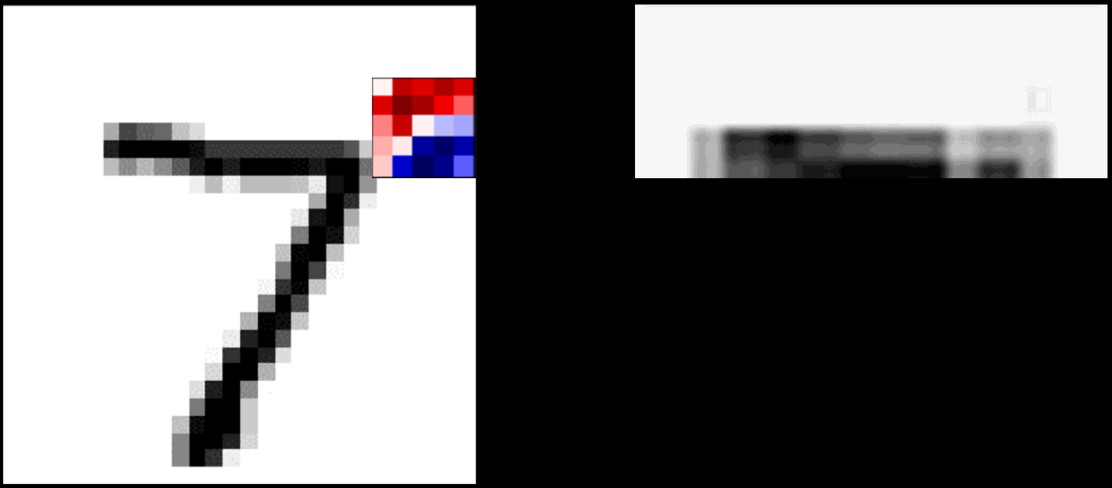


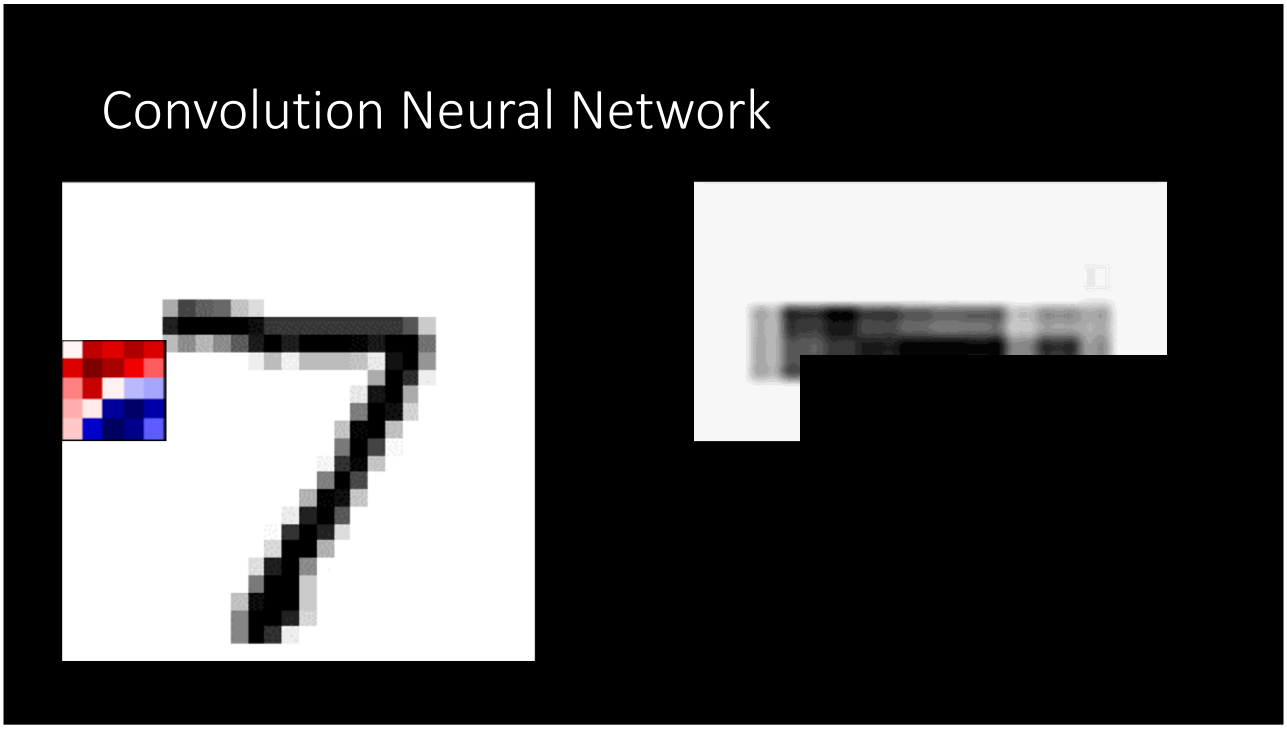
Convolution Neural Network

Convolution Neural Network

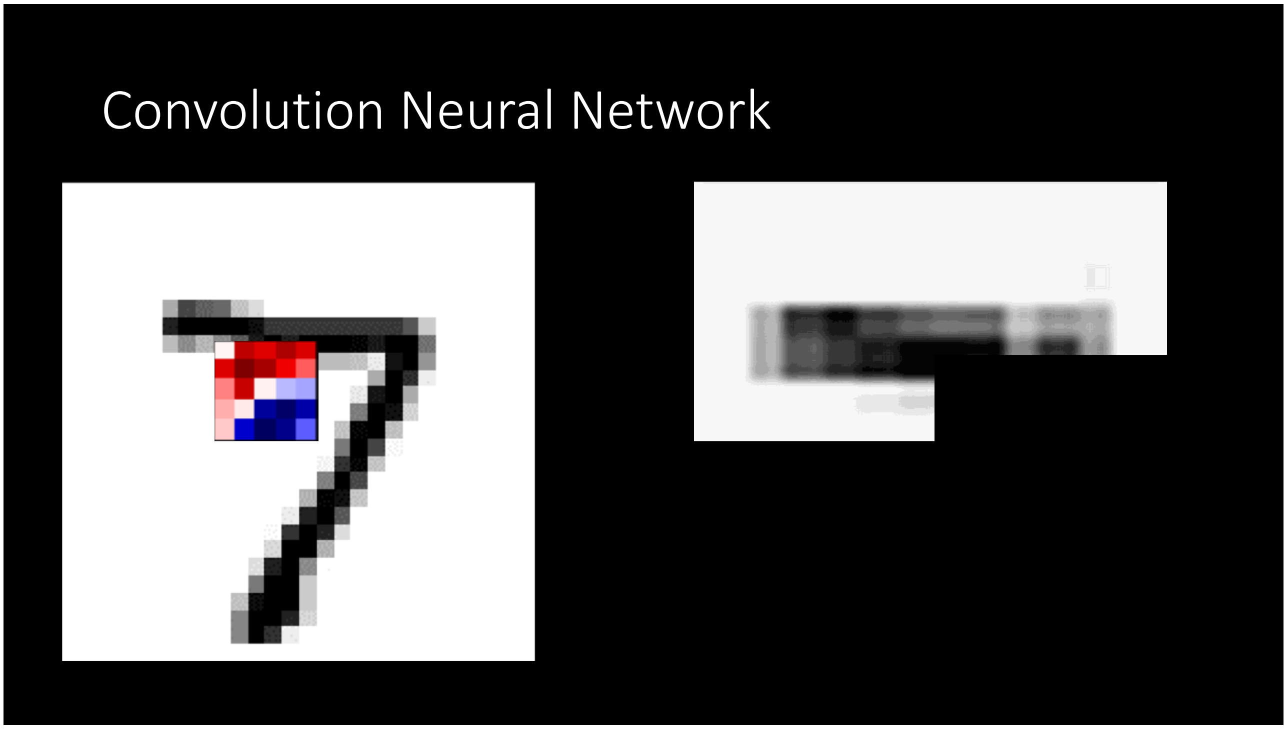


Convolution Neural Network



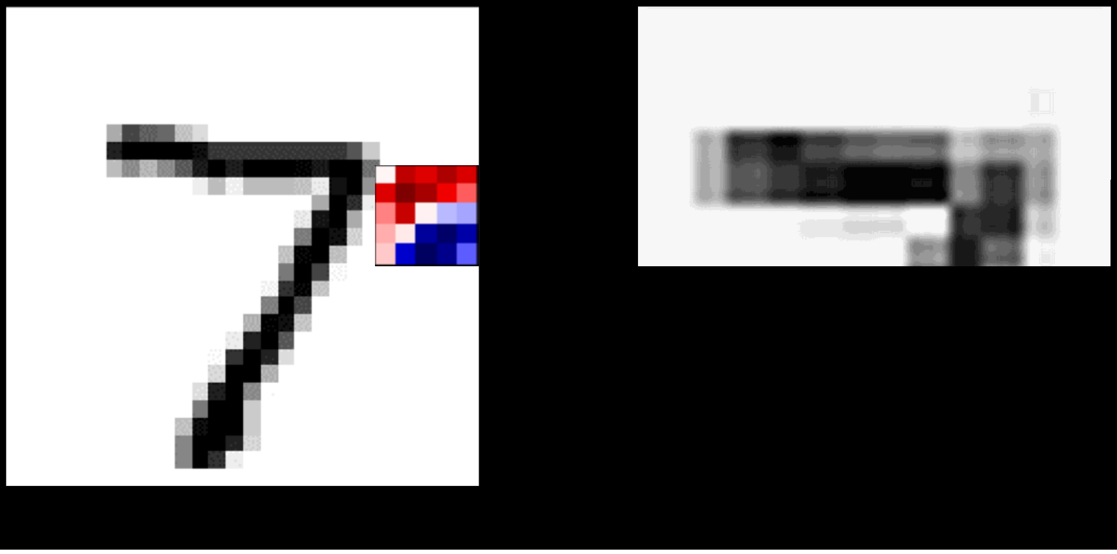


Convolution Neural Network

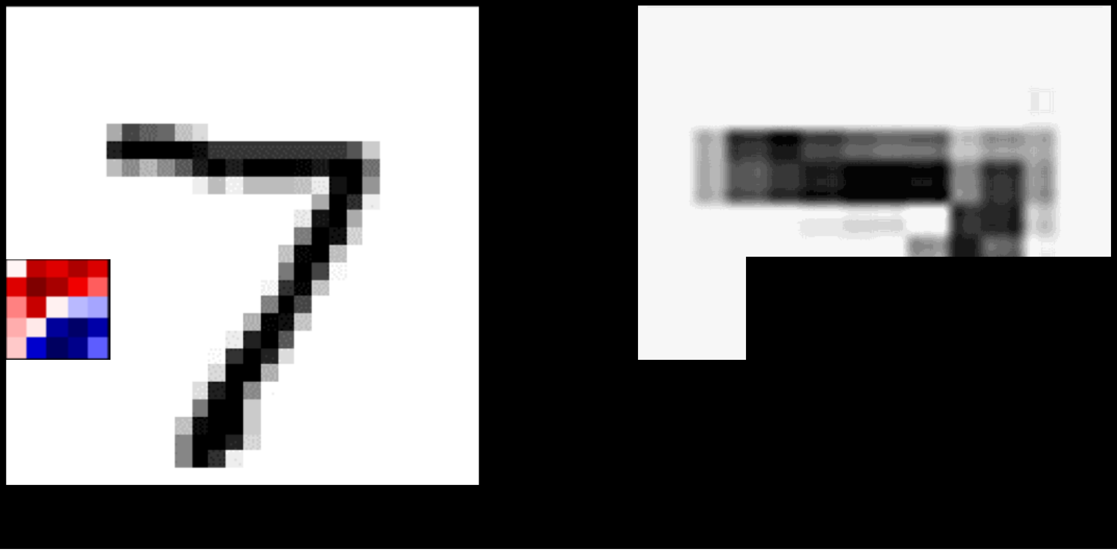


Convolution Neural Network

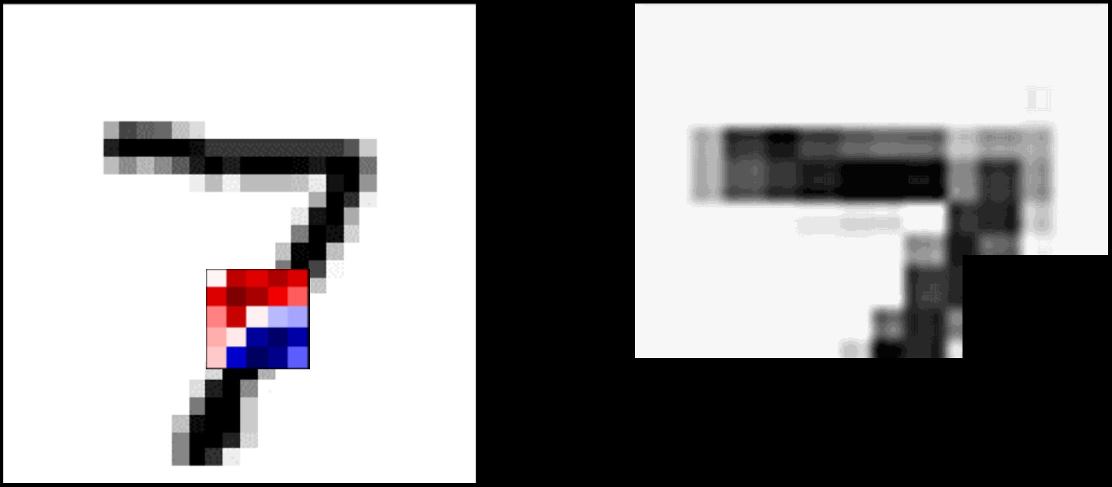
Convolution Neural Network



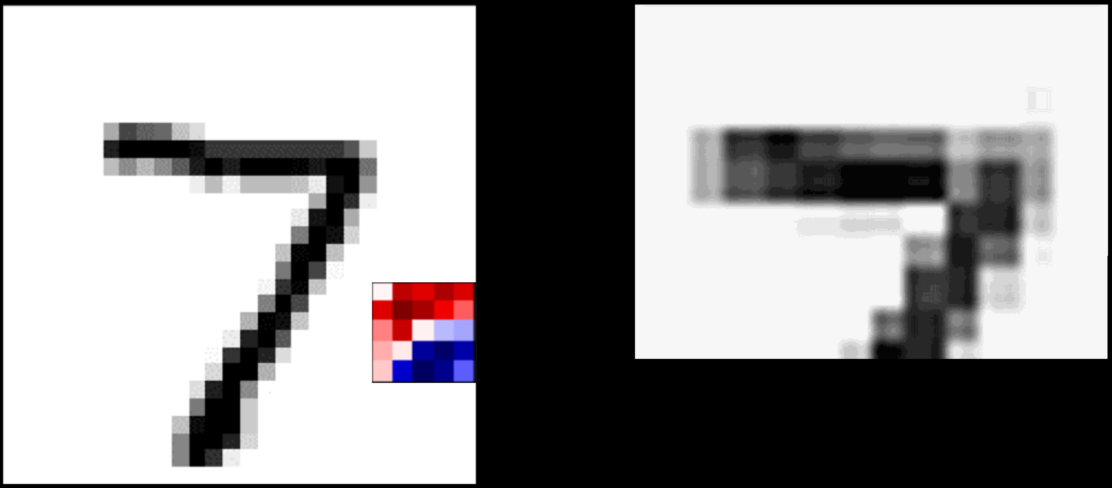
Convolution Neural Network

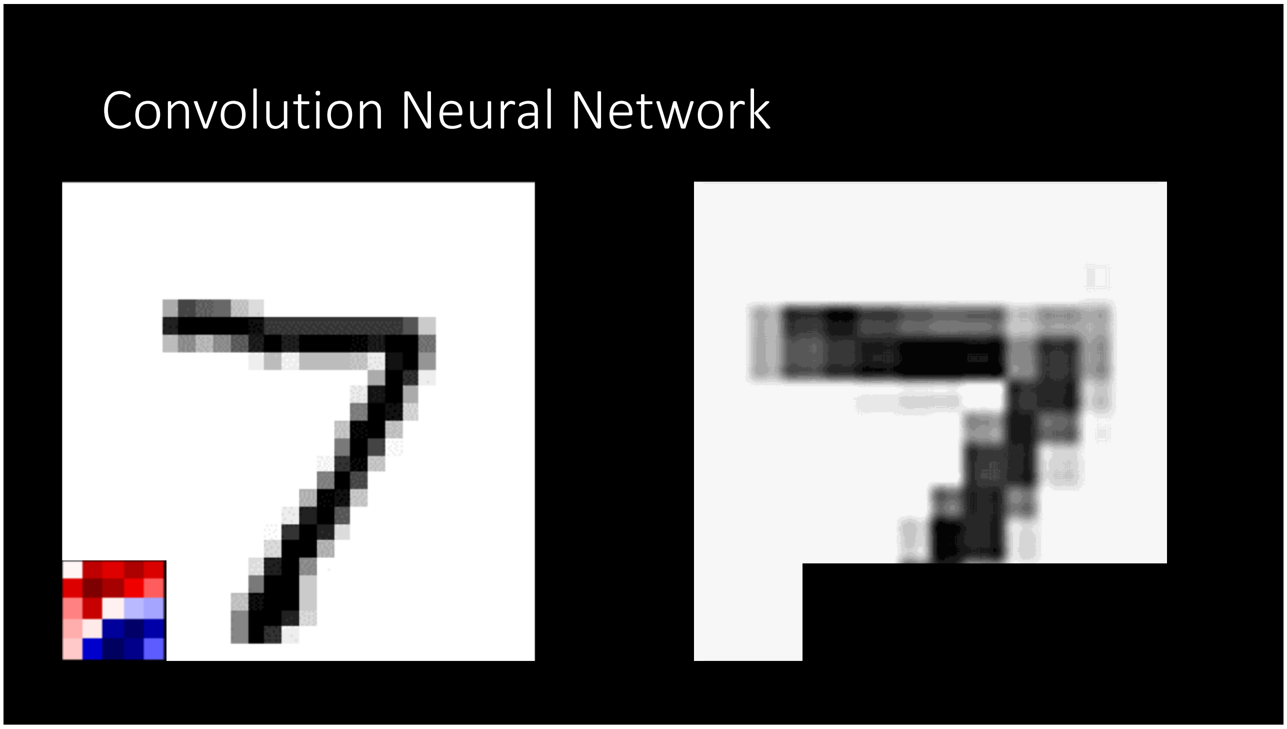


Convolution Neural Network

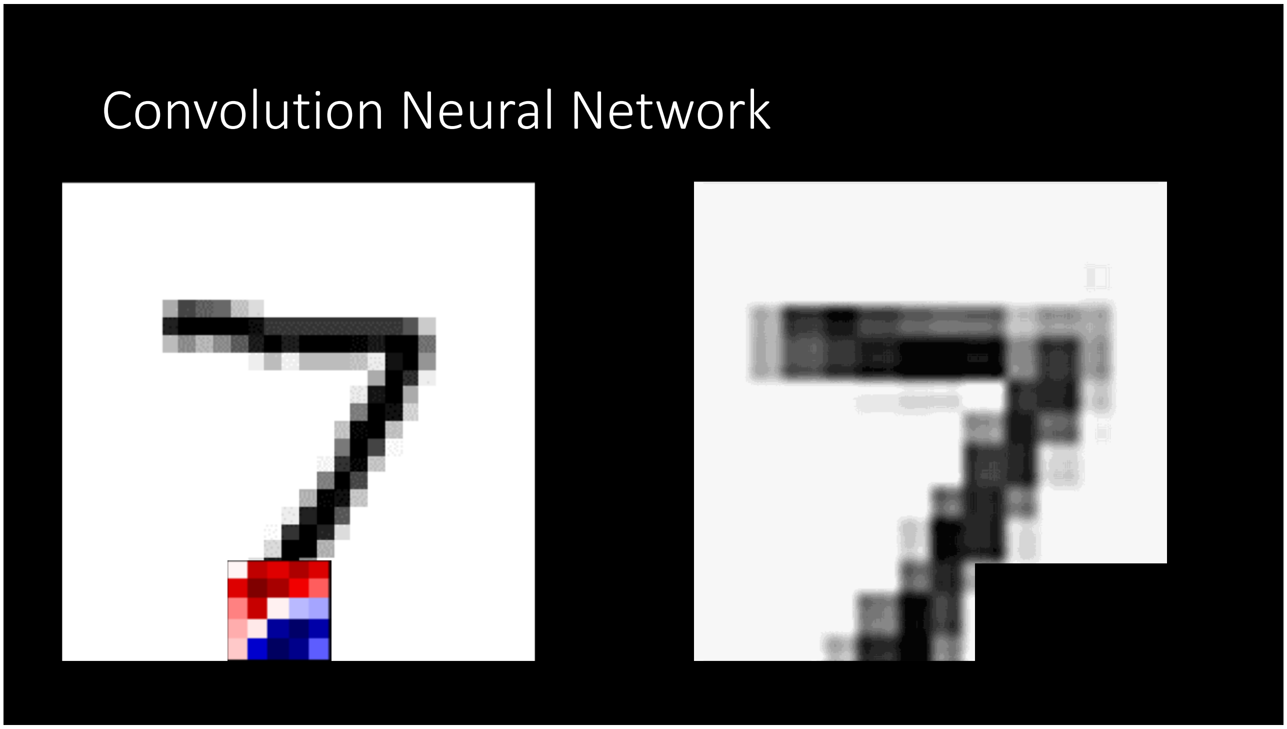


Convolution Neural Network

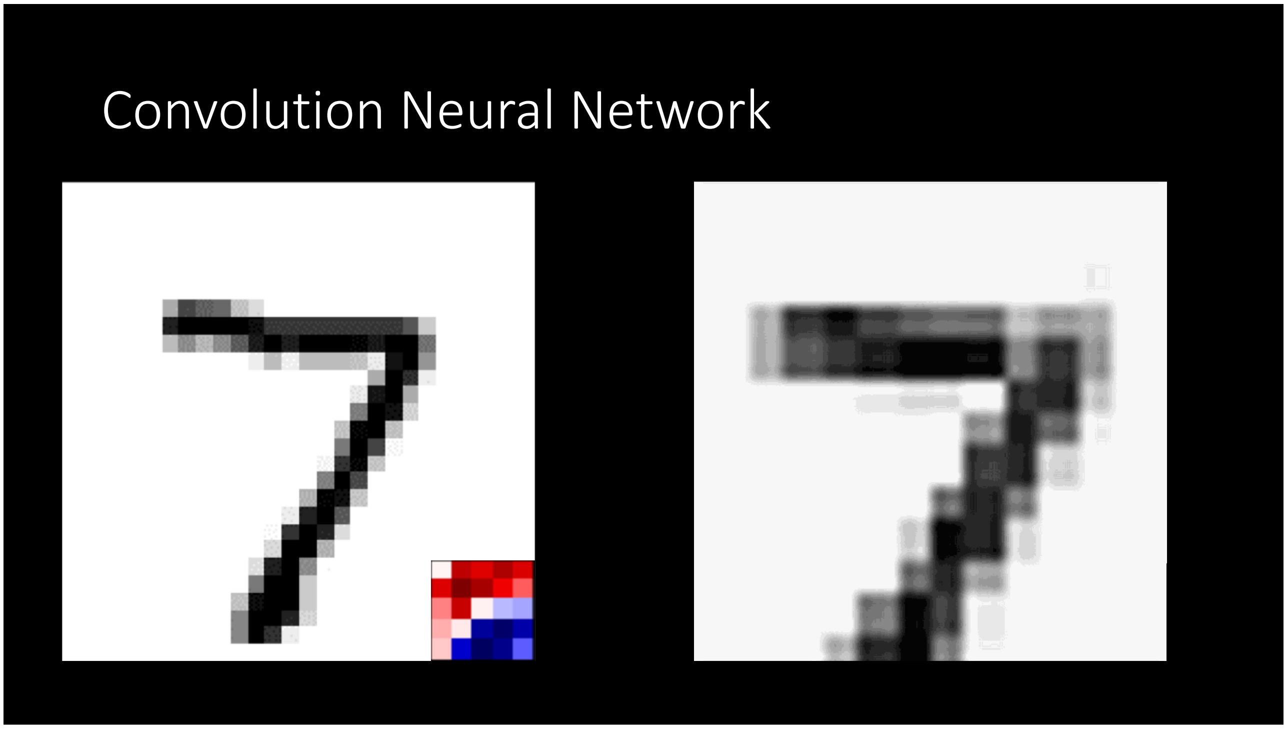




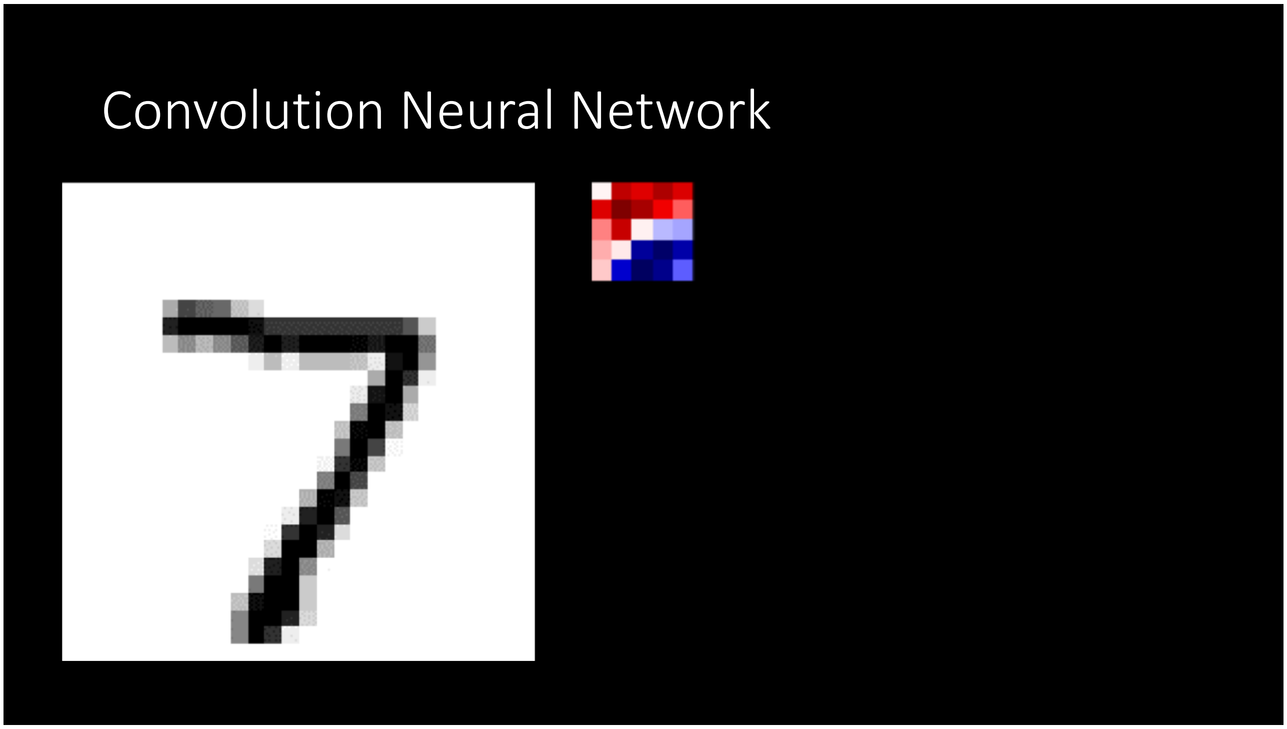
Convolution Neural Network



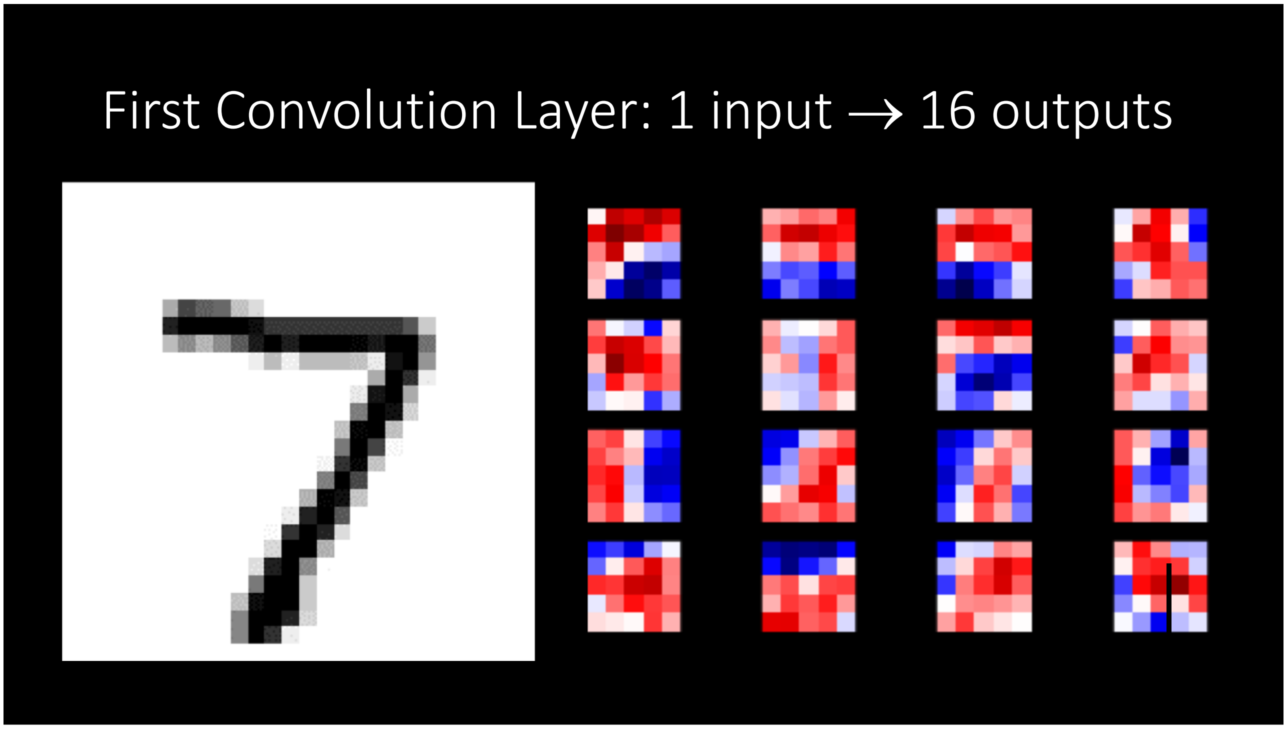
Convolution Neural Network



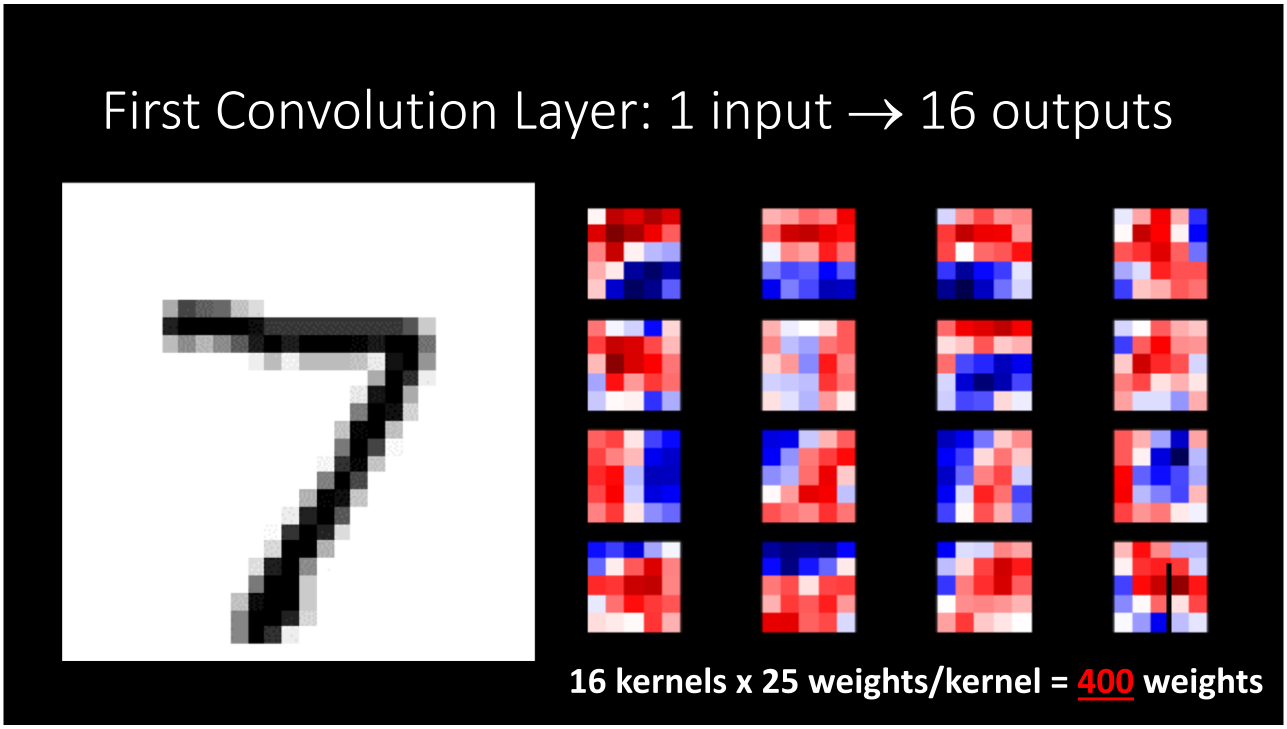
Convolution Neural Network



Convolution Neural Network

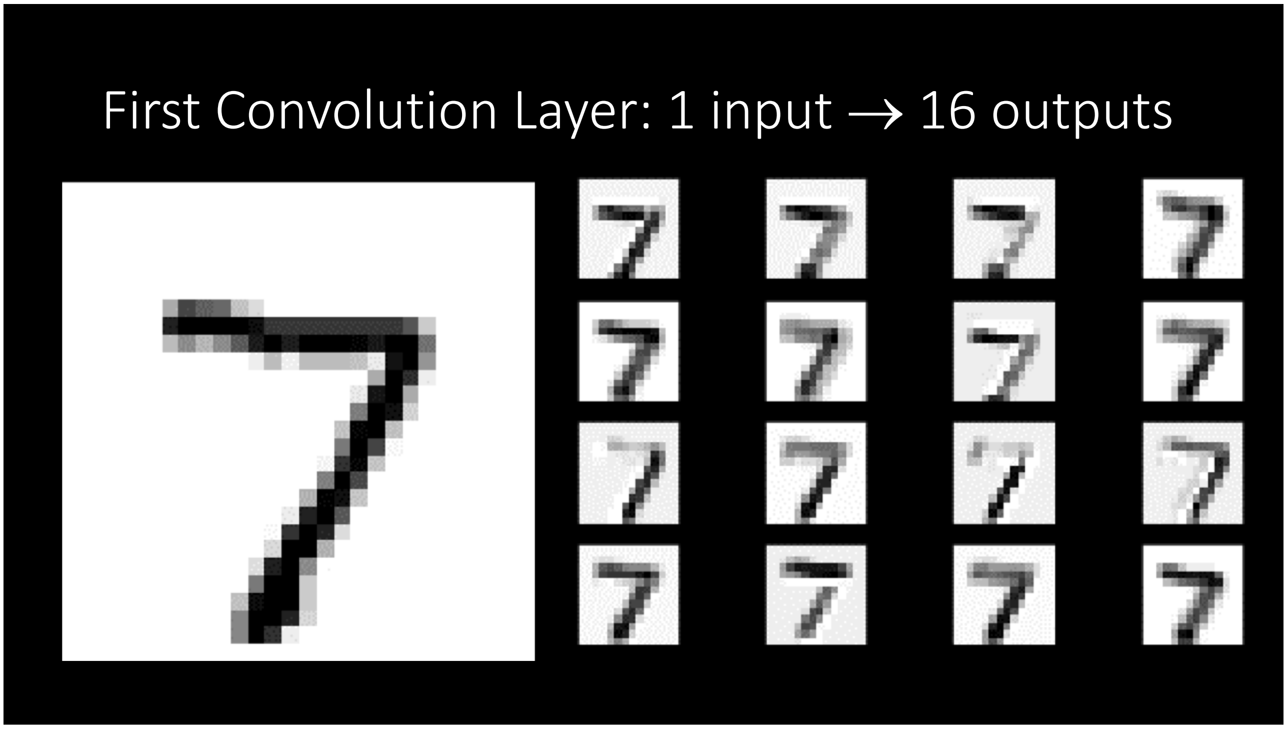


First Convolution Layer: 1 input → 16 outputs

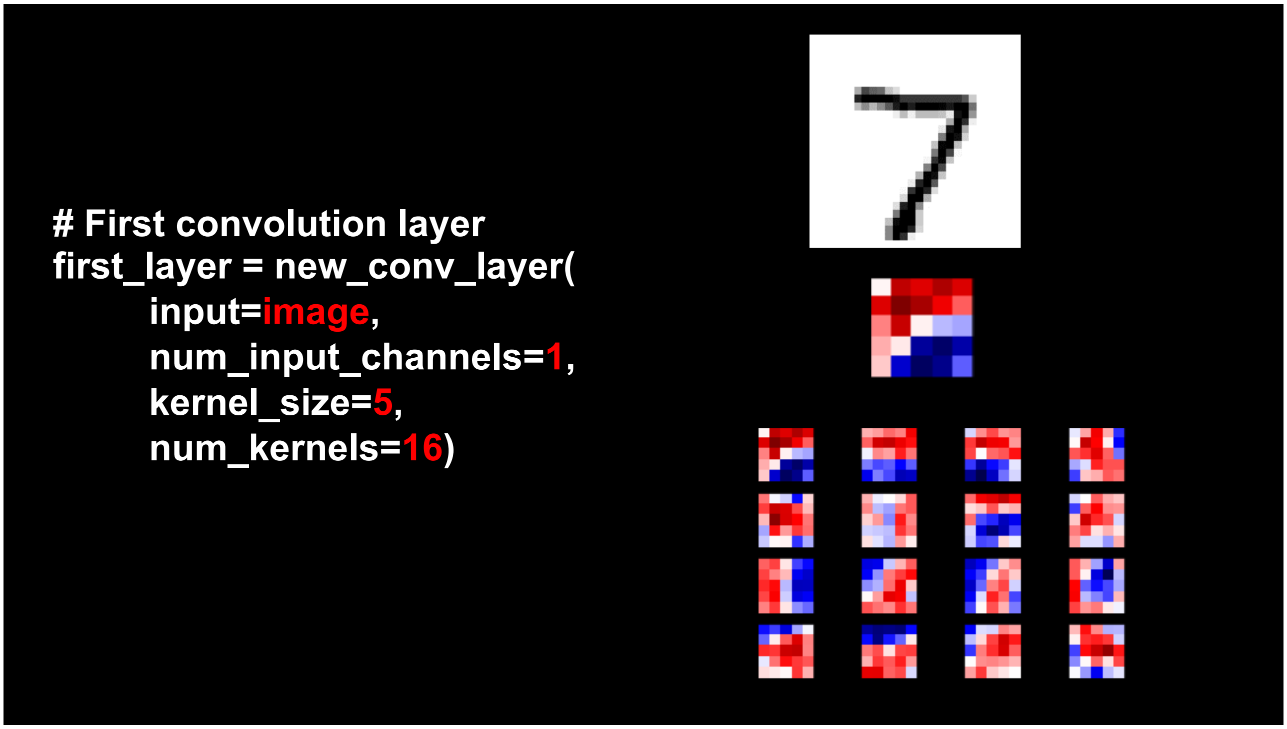


First Convolution Layer: 1 input → 16 outputs

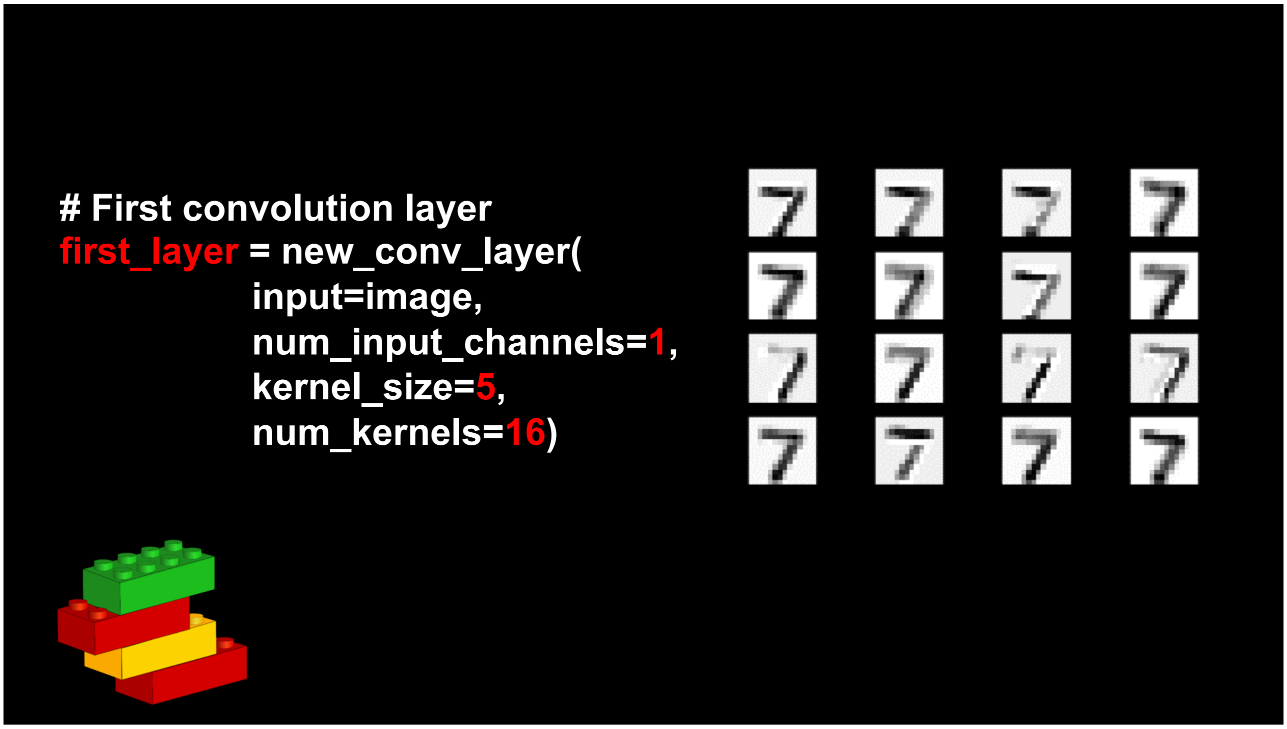
16 kernels x 25 weights/kernel = 400 weights



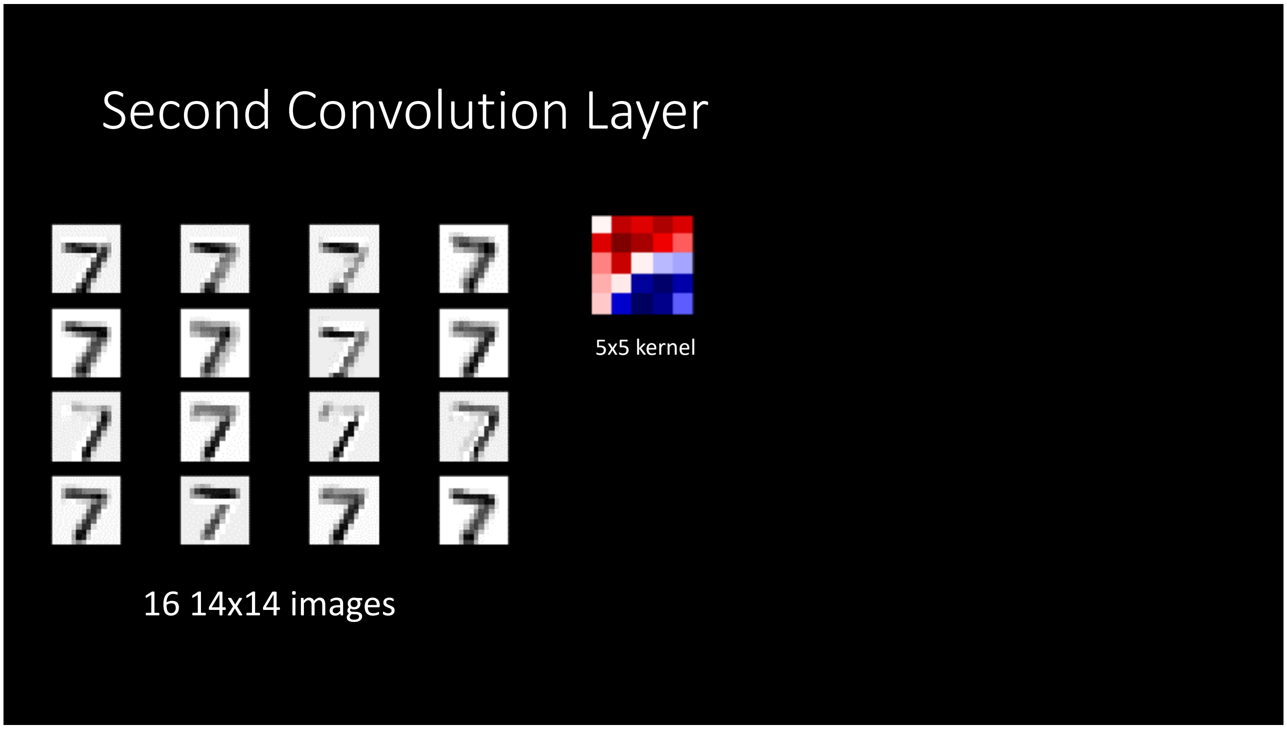
First Convolution Layer: 1 input → 16 outputs



```
# First convolution layer
first_layer = new_conv_layer(
        input=image,
        num_input_channels=1,
        kernel_size=5,
        num_kernels=16)
```



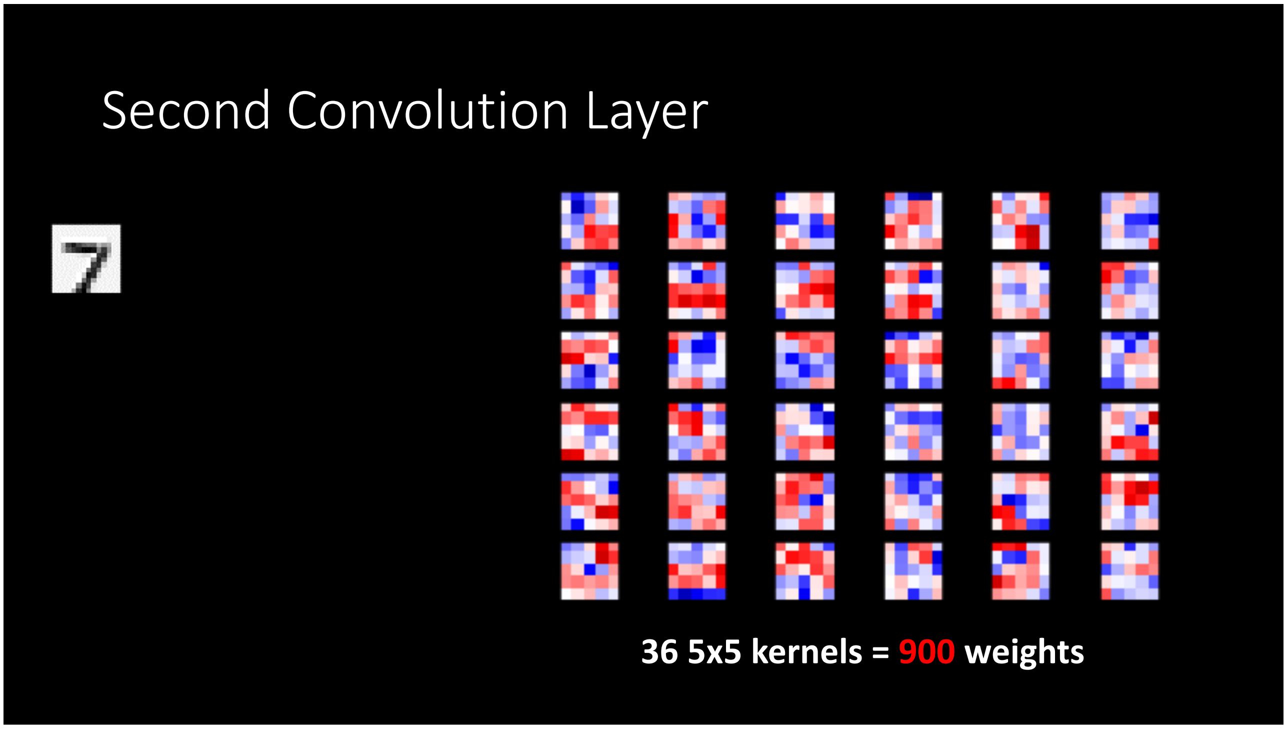
```
# First convolution layer
first_layer = new_conv_layer(
                input=image,
                num_input_channels=1,
                kernel_size=5,
                num_kernels=16)
```



Second Convolution Layer

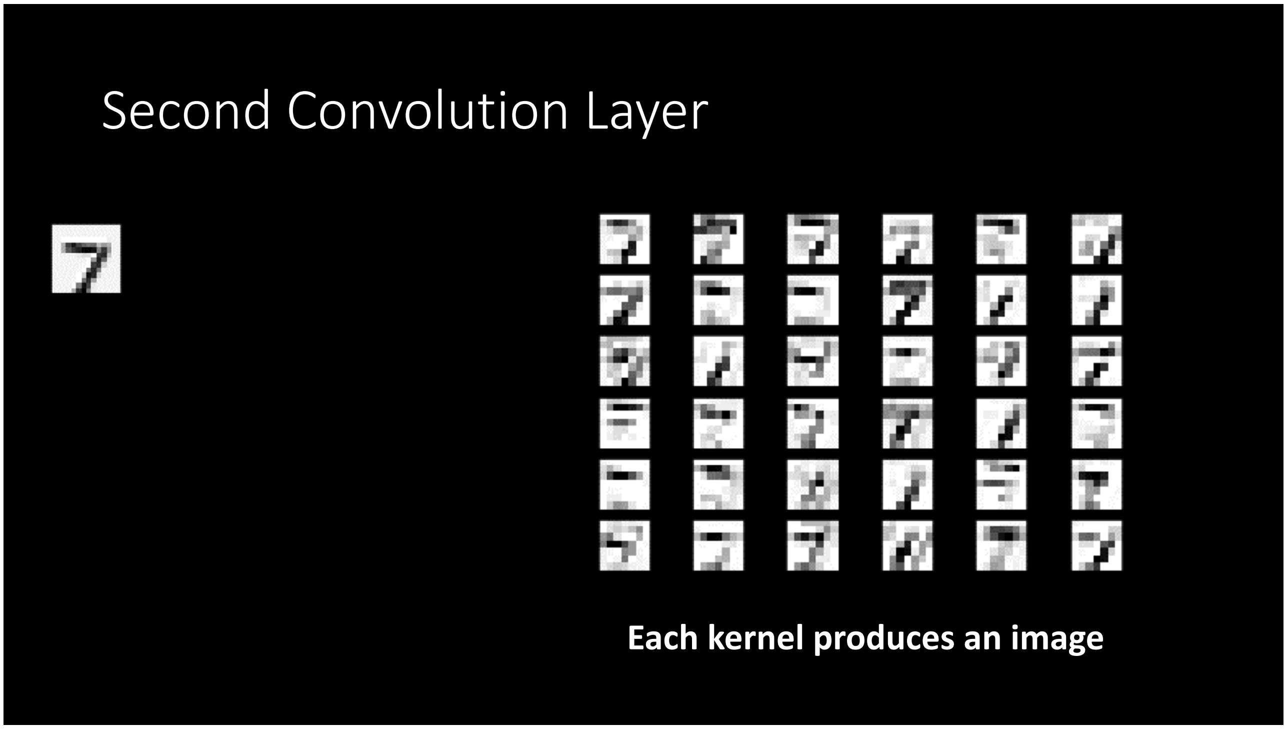
5x5 kernel

16 14x14 images



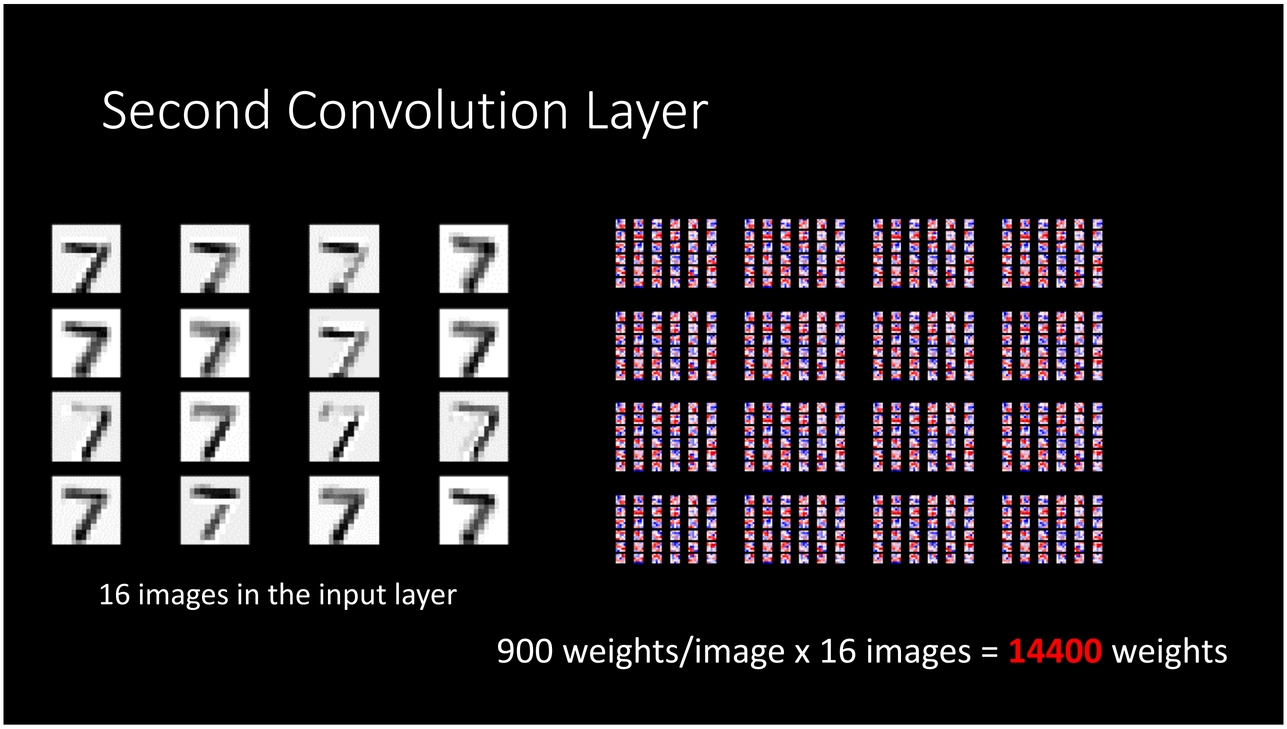
Second Convolution Layer

36 5x5 kernels = 900 weights



Second Convolution Layer

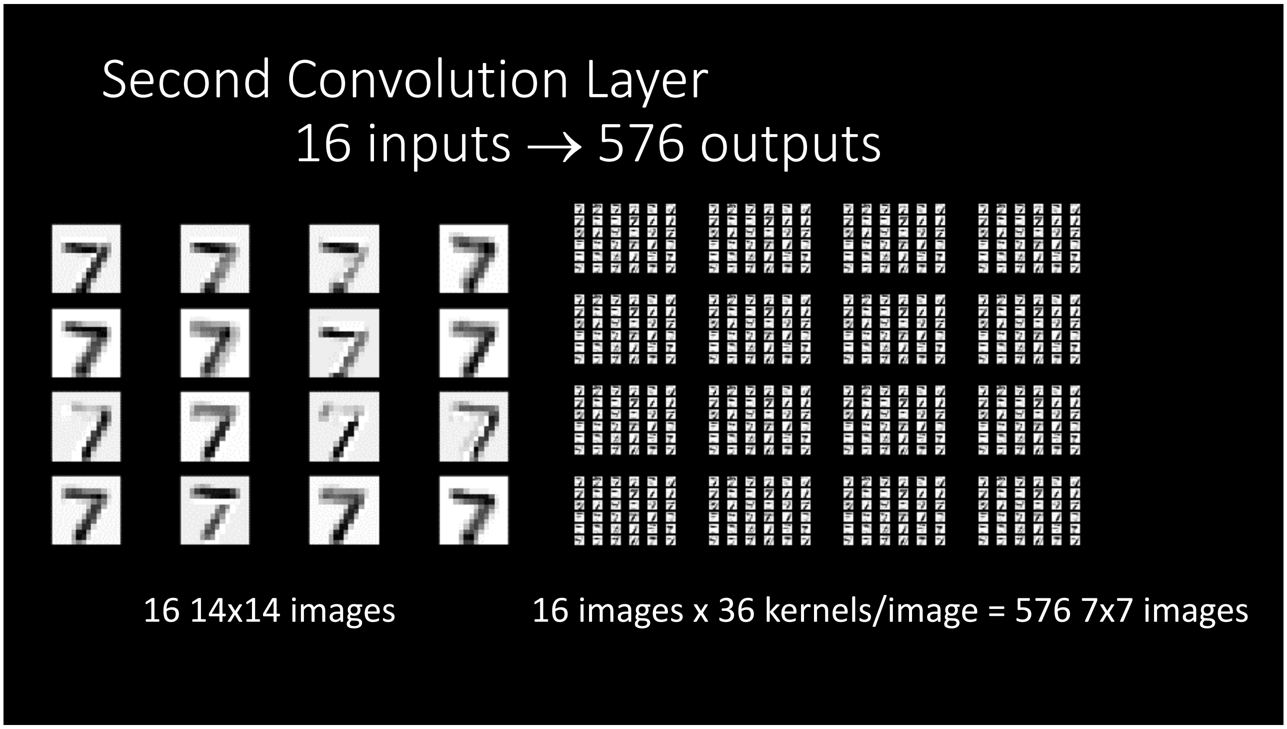
Each kernel produces an image



Second Convolution Layer

16 images in the input layer

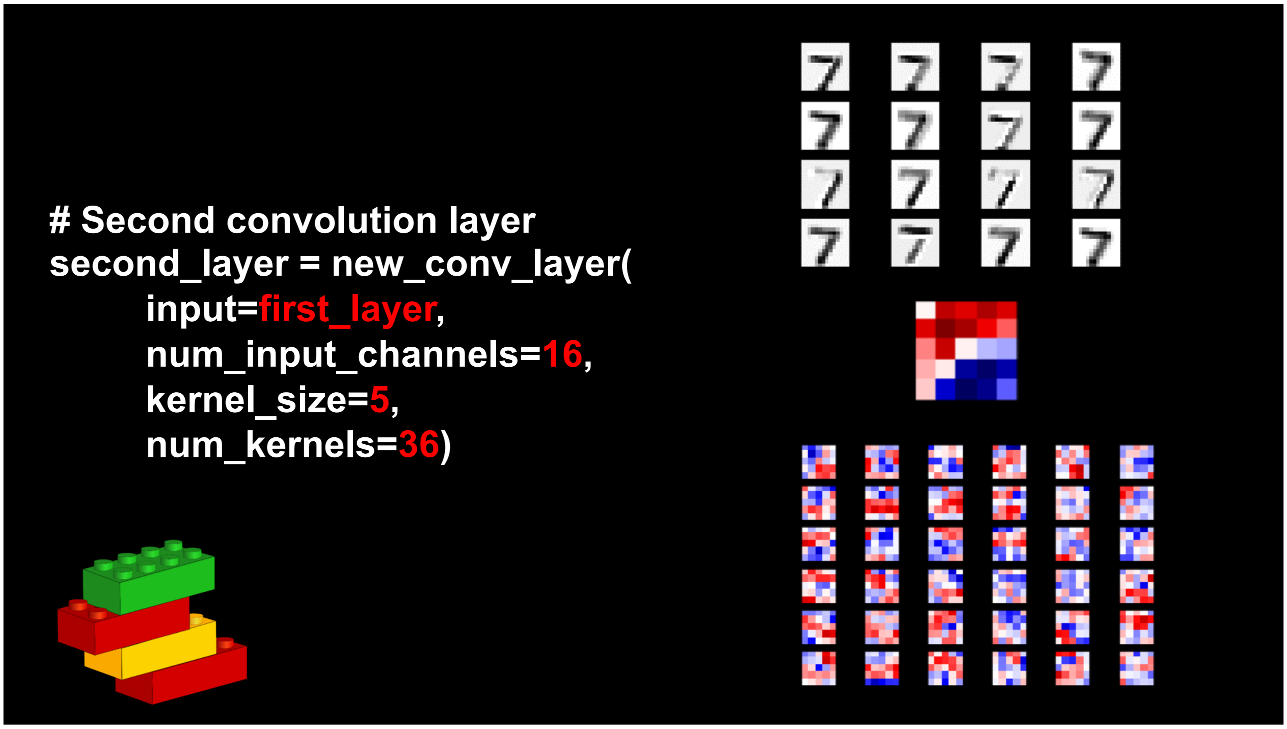
900 weights/image x 16 images = **14400** weights



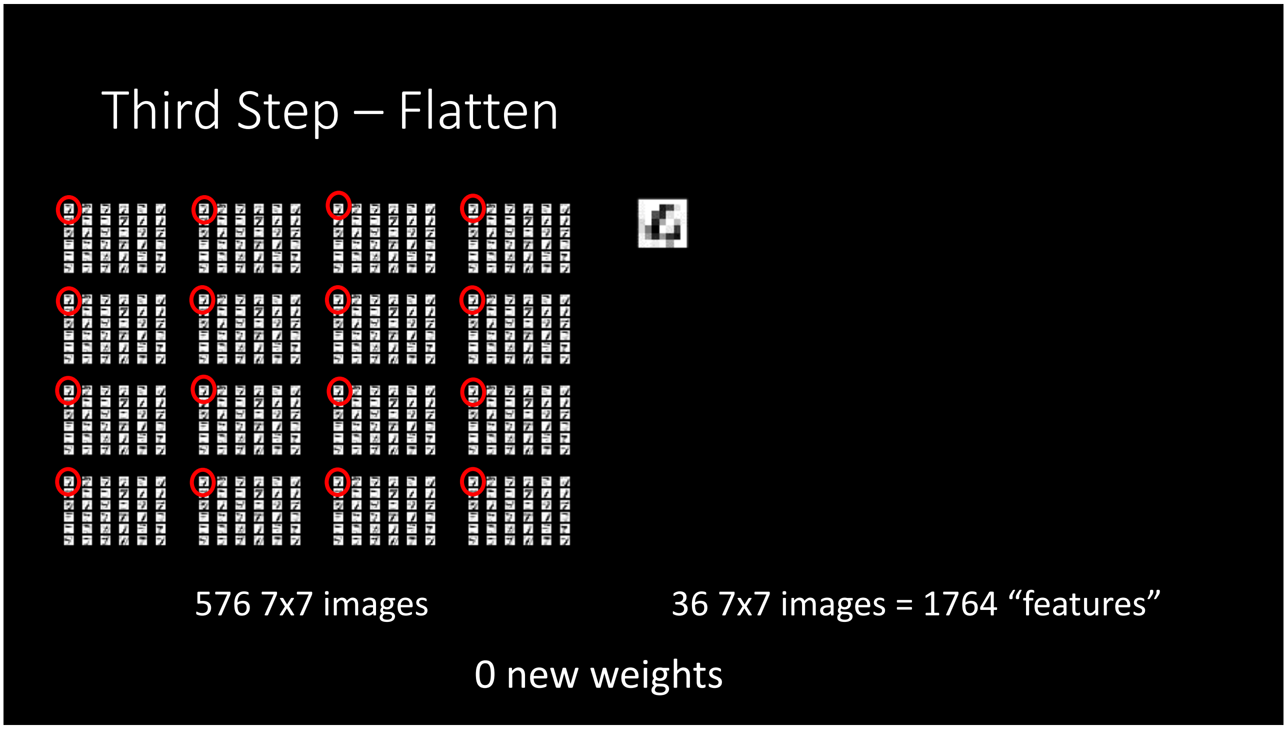
Second Convolution Layer
16 inputs → 576 outputs

16 14x14 images

16 images x 36 kernels/image = 576 7x7 images



```
# Second convolution layer
second_layer = new_conv_layer(
        input=first_layer,
        num_input_channels=16,
        kernel_size=5,
        num_kernels=36)
```

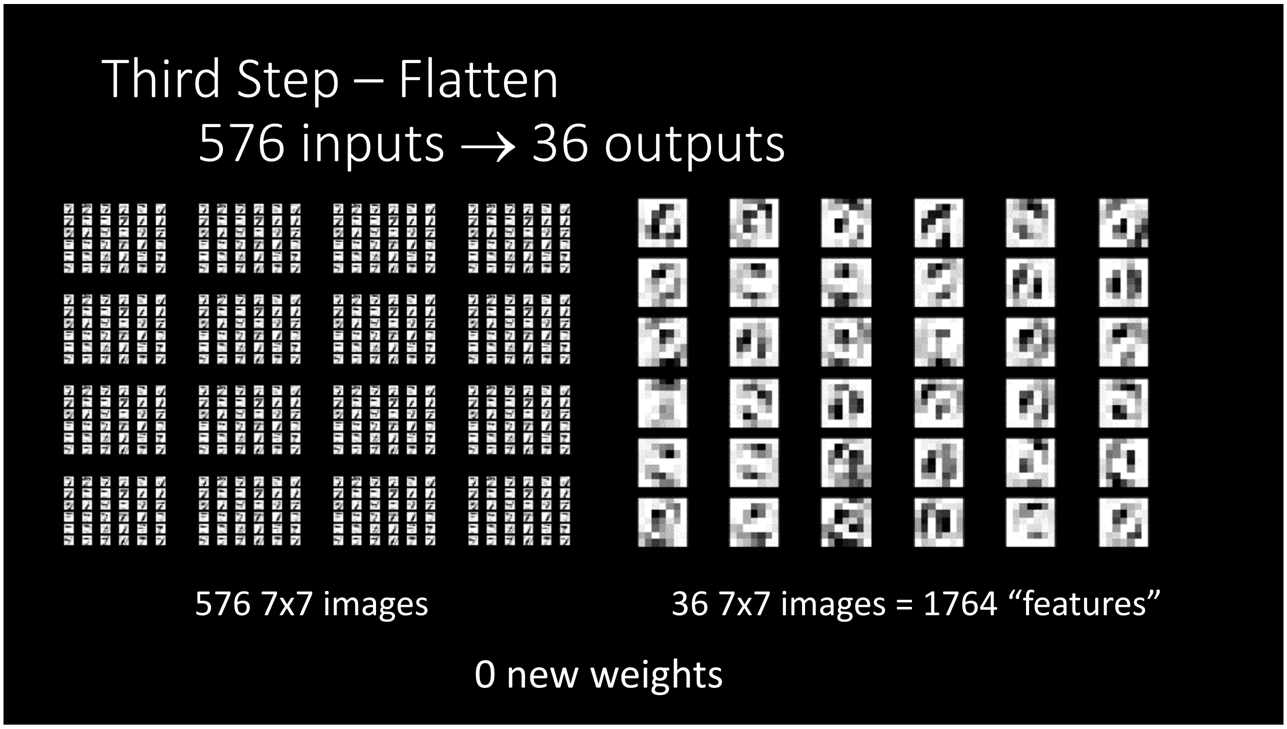


Third Step – Flatten

576 7x7 images

36 7x7 images = 1764 “features”

0 new weights



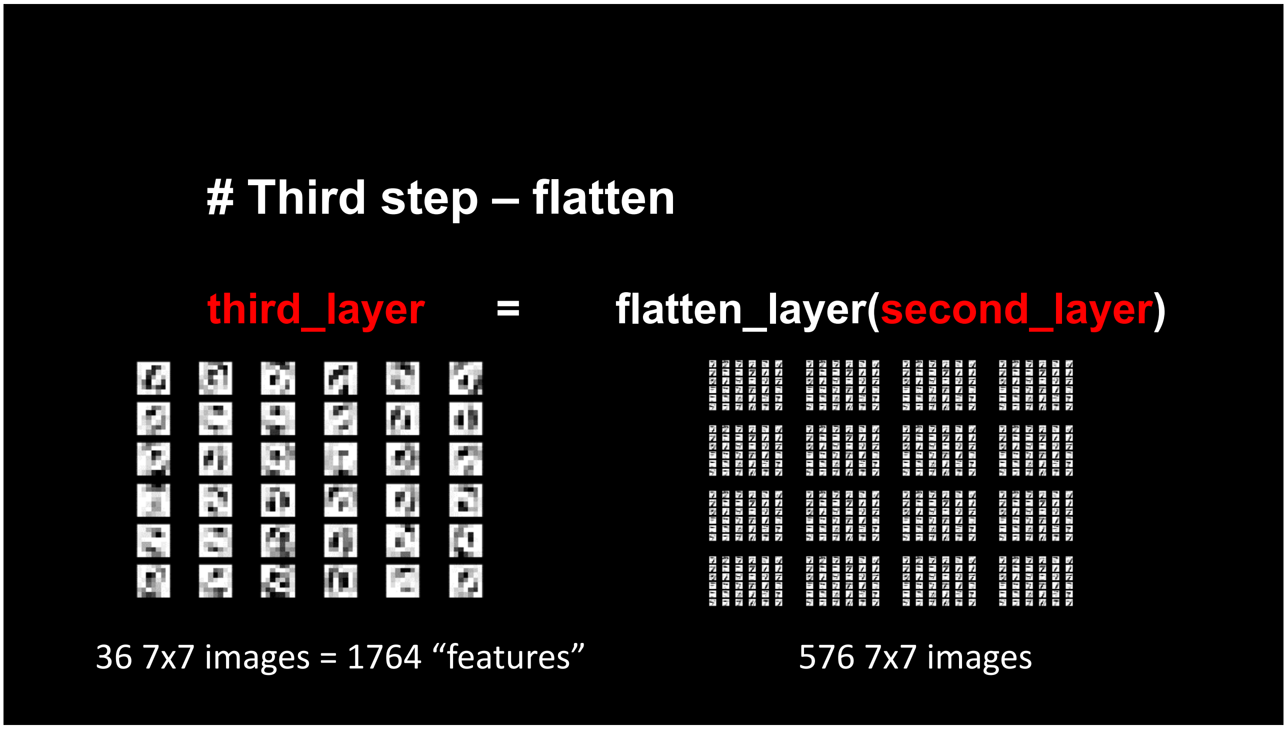
Third Step – Flatten

576 inputs → 36 outputs

576 7x7 images

36 7x7 images = 1764 “features”

0 new weights

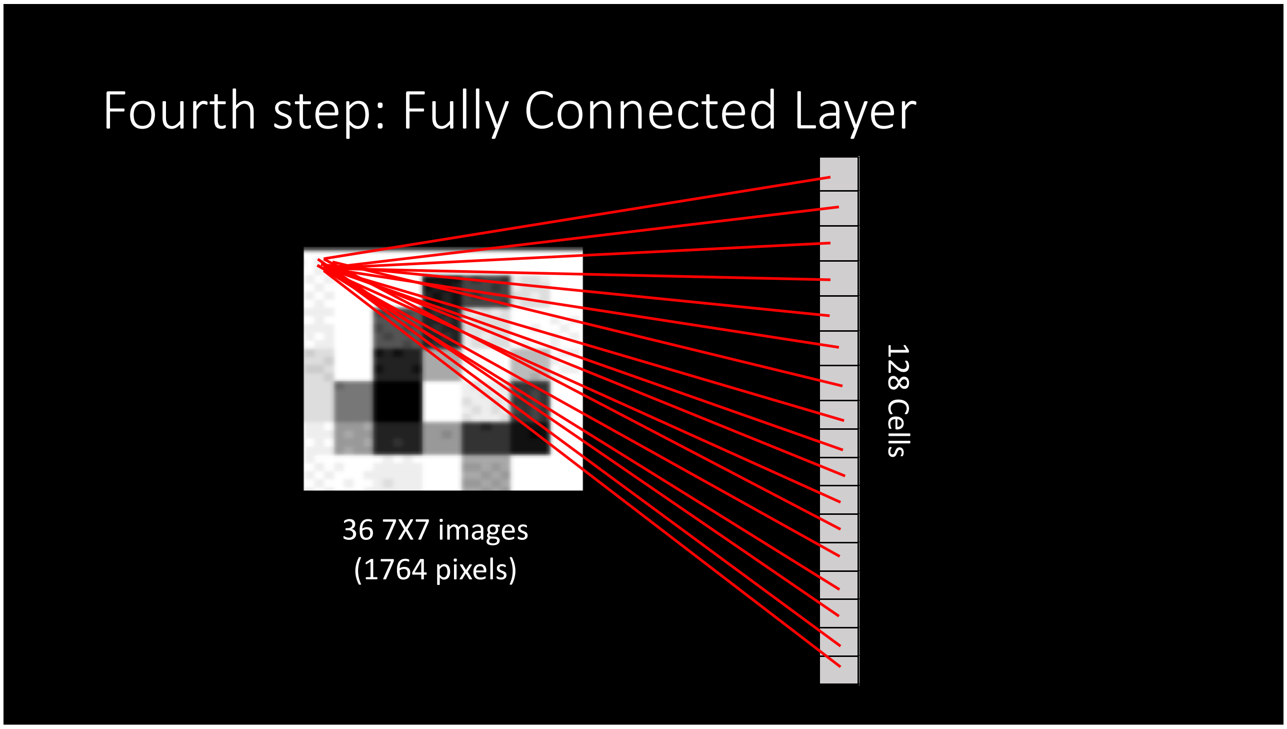


```
# Third step – flatten

third_layer    =    flatten_layer(second_layer)
```

36 7x7 images = 1764 “features”

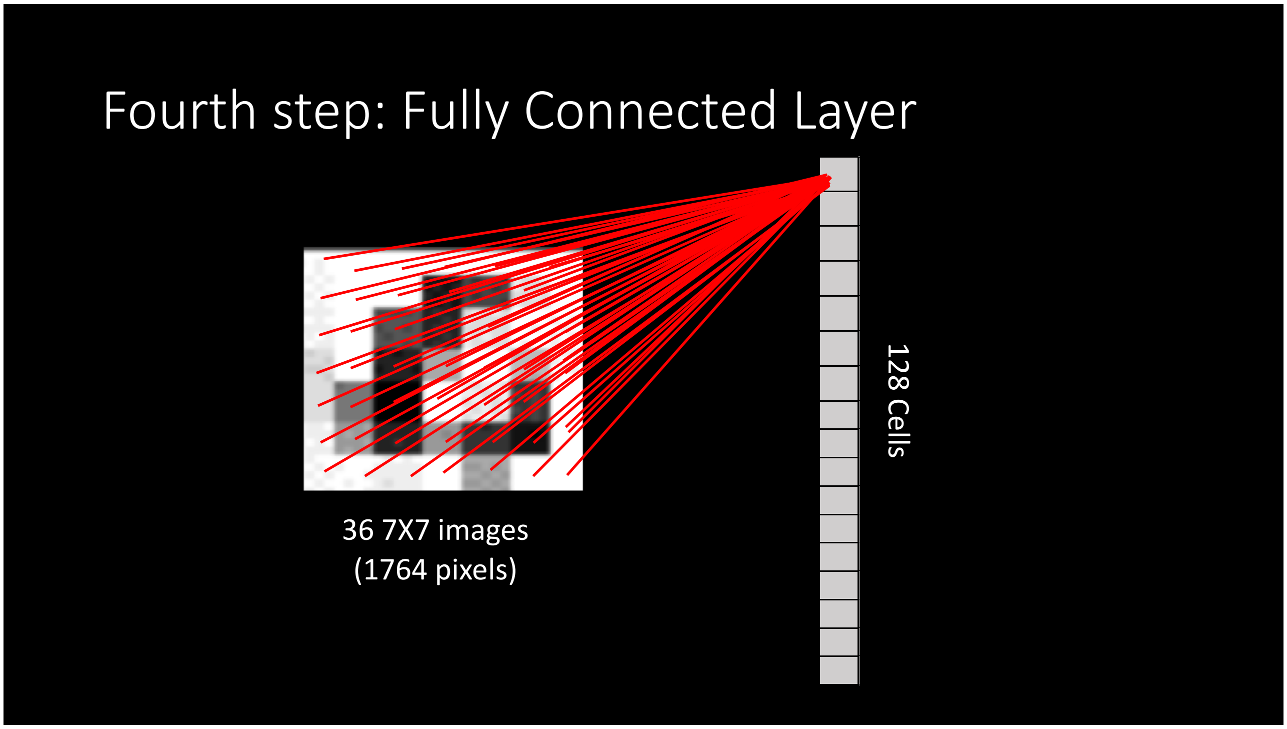
576 7x7 images



Fourth step: Fully Connected Layer

36 7X7 images
(1764 pixels)

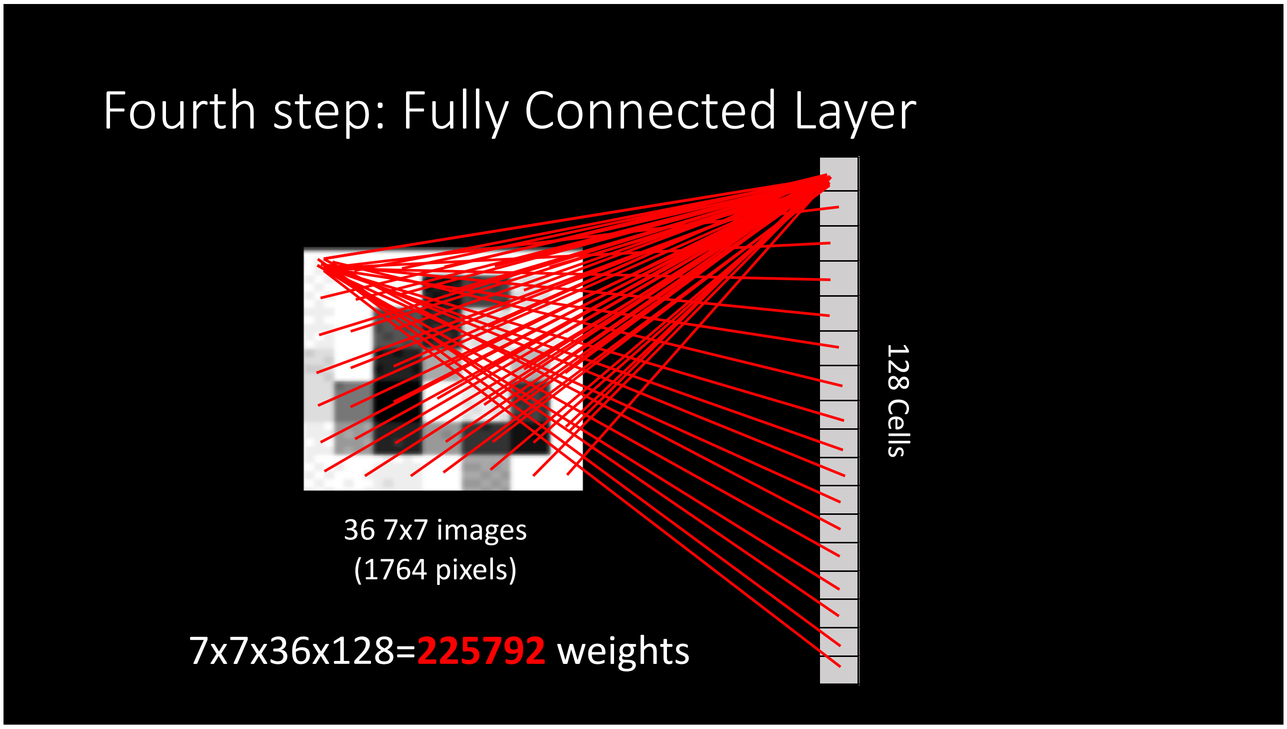
128 Cells



Fourth step: Fully Connected Layer

36 7X7 images
(1764 pixels)

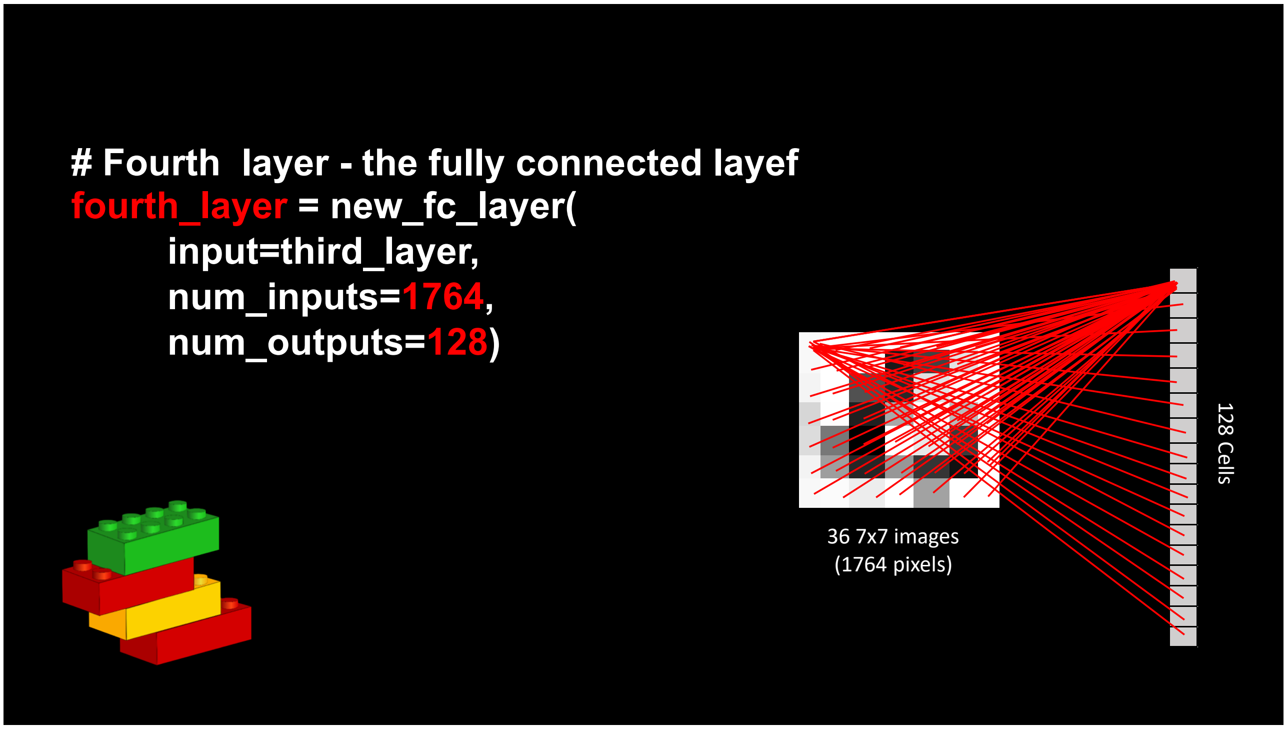
128 Cells



Fourth step: Fully Connected Layer

36 7x7 images
(1764 pixels)

7x7x36x128=225792 weights

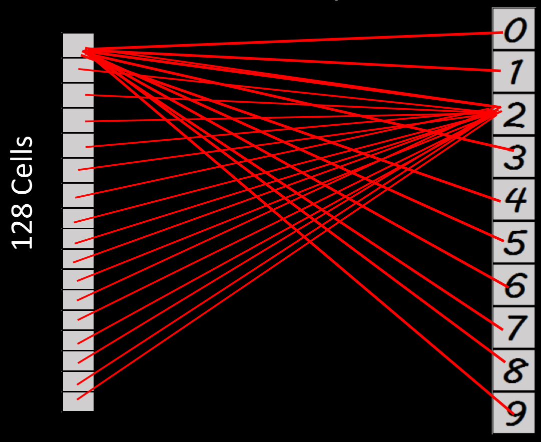


```
# Fourth  layer - the fully connected layef
fourth_layer = new_fc_layer(
        input=third_layer,
        num_inputs=1764,
        num_outputs=128)
```

36 7x7 images
(1764 pixels)

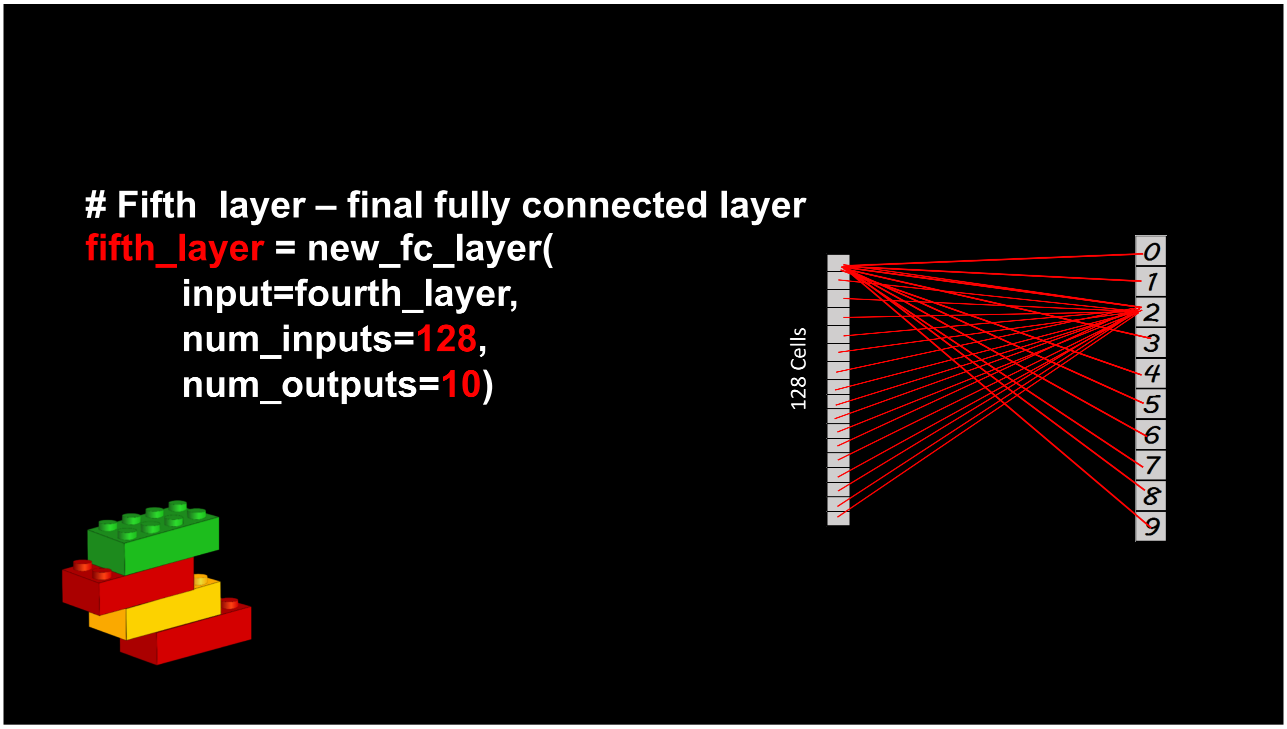
Final Fully Connected Layer

128 inputs → 10 outputs

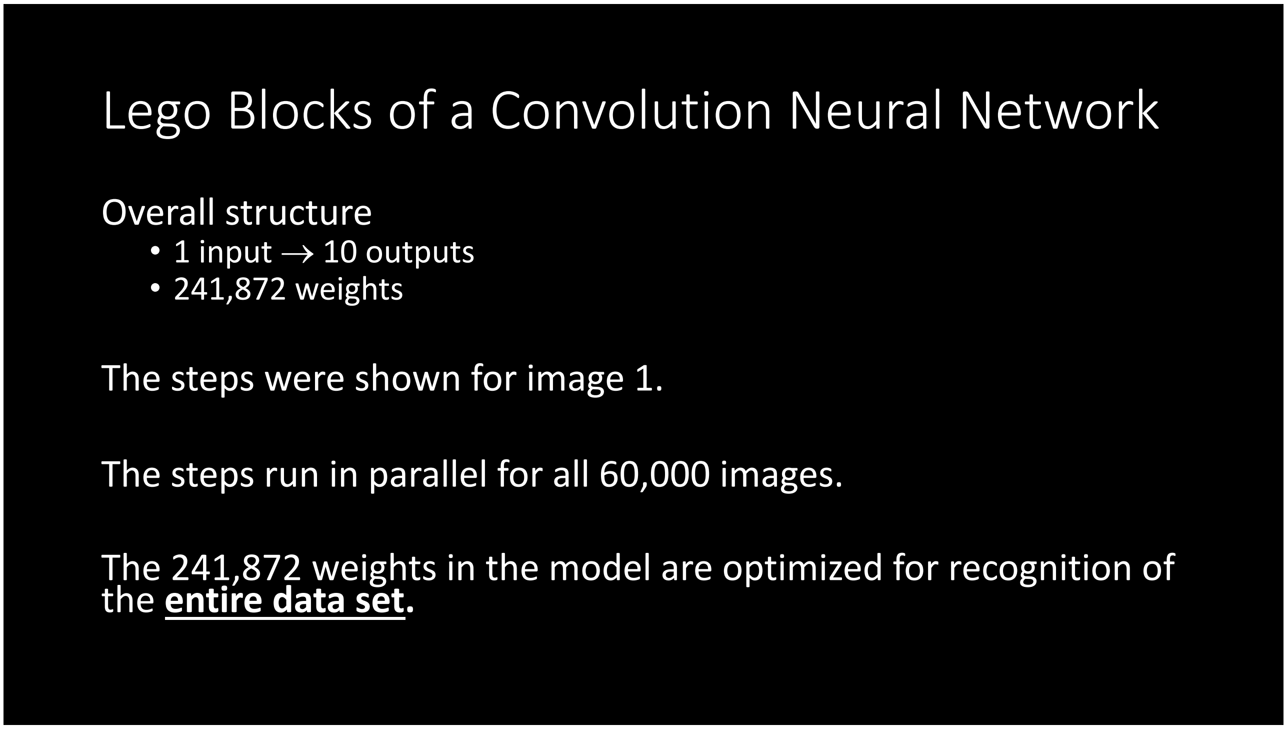


1280 weights

```
# Fifth  layer – final fully connected layer
fifth_layer = new_fc_layer(
        input=fourth_layer,
        num_inputs=128
        num_outputs=10)
```



```
# Fifth  layer – final fully connected layer
fifth_layer = new_fc_layer(
        input=fourth_layer,
        num_inputs=128,
        num_outputs=10)
```



Lego Blocks of a Convolution Neural Network

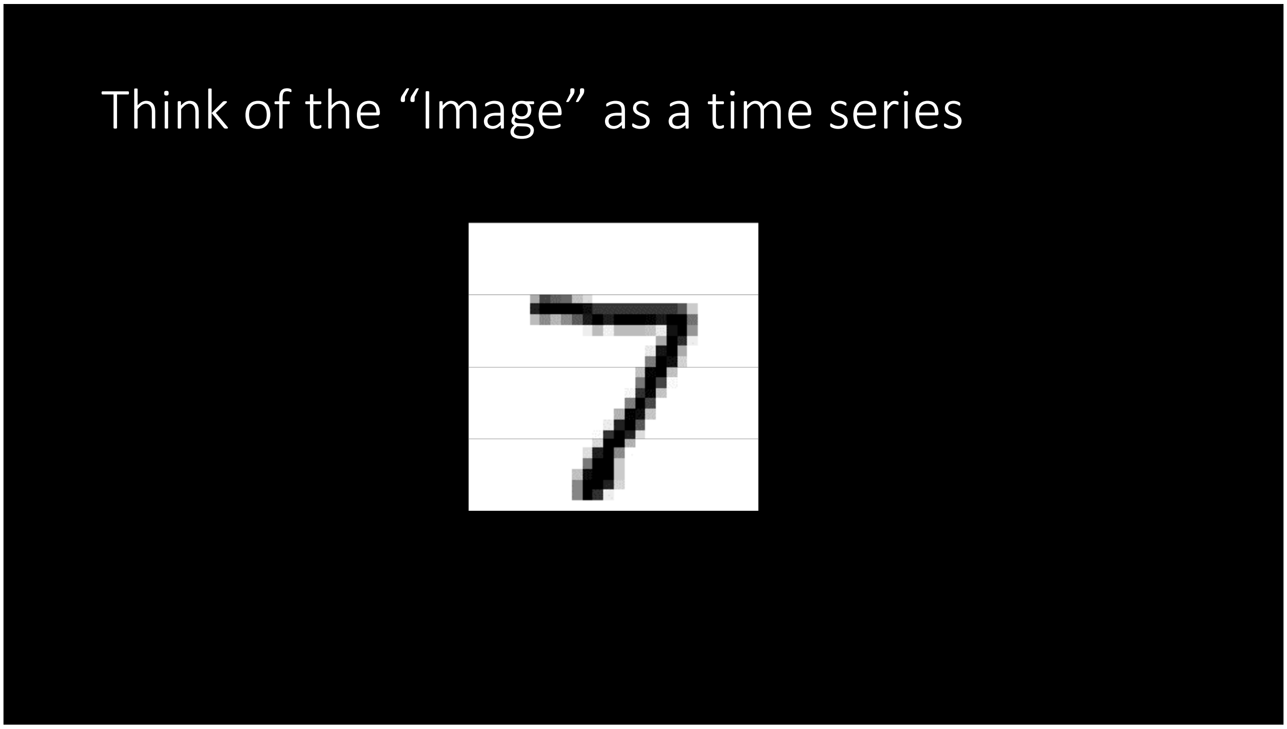
Overall structure

- 1 input → 10 outputs
- 241,872 weights

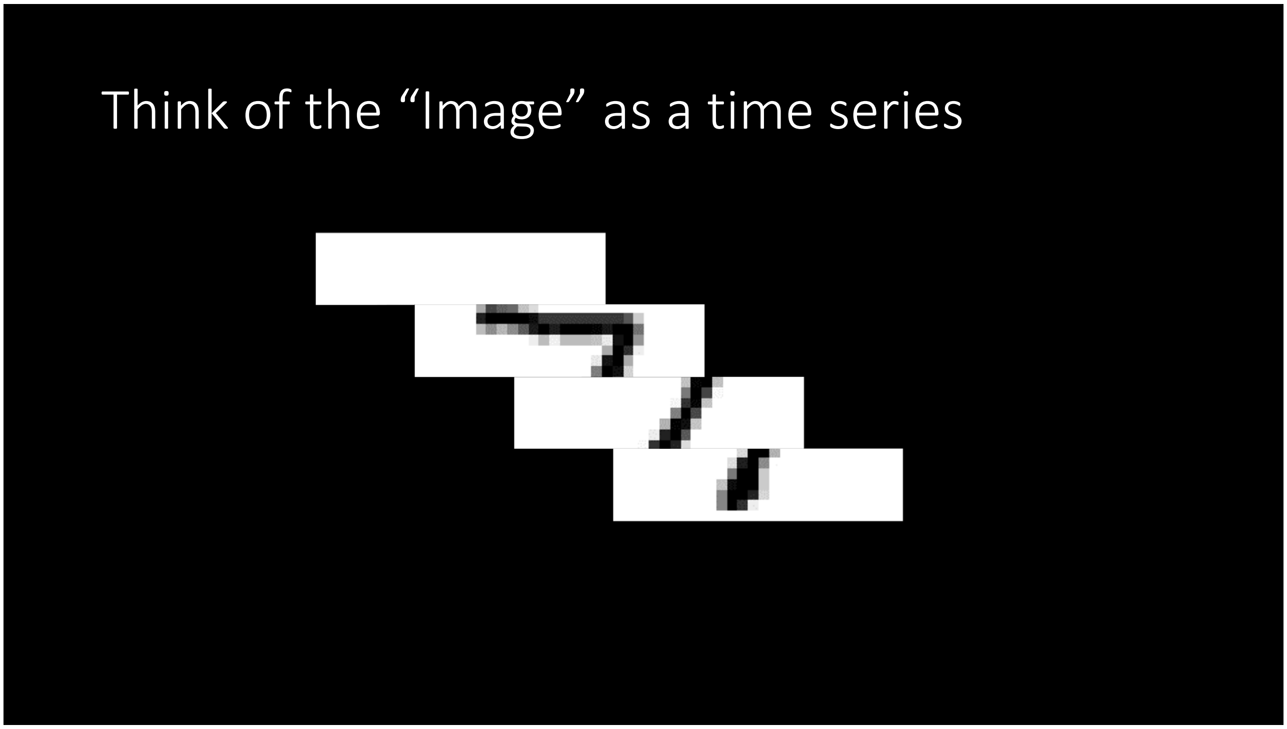
The steps were shown for image 1.

The steps run in parallel for all 60,000 images.

The 241,872 weights in the model are optimized for recognition of the **entire data set.**

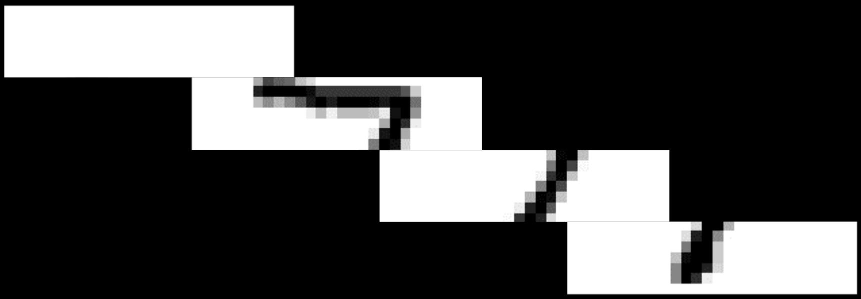


Think of the “Image” as a time series

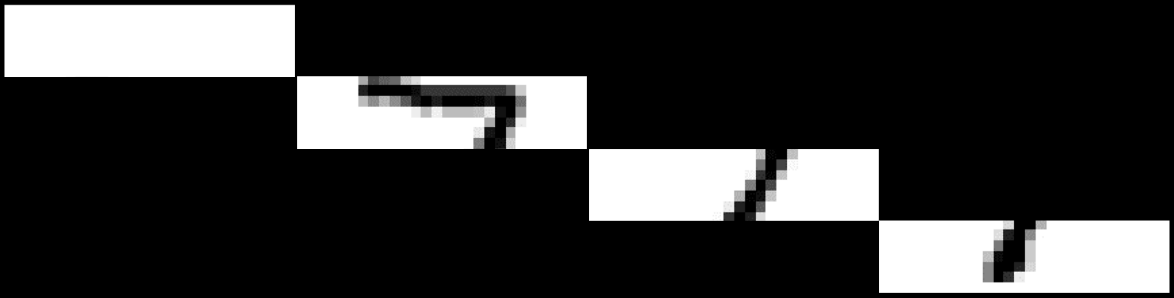


Think of the “Image” as a time series

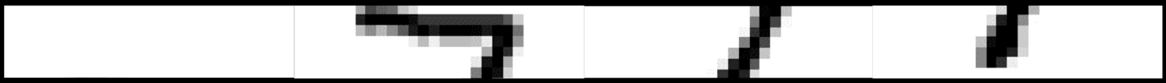
Think of the “Image” as a time series



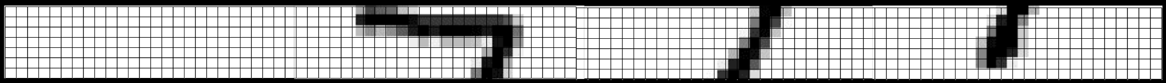
Think of the “Image” as a time series



Think of the “Image” as a time series

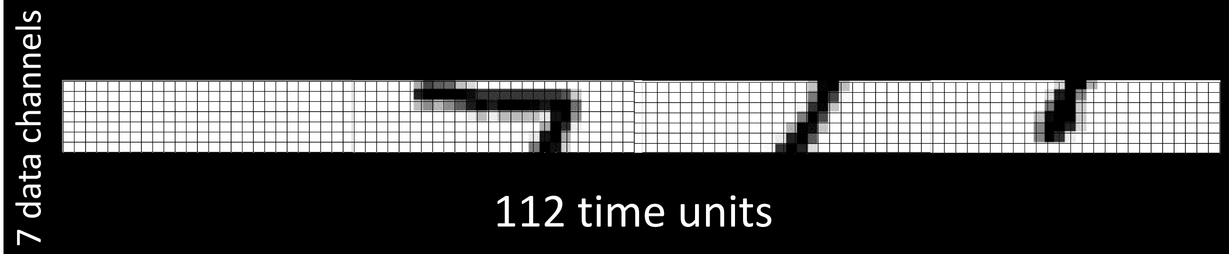


Our “image” has become a time series with 7 data channels (drug levels, vital signs, lab tests, etc)

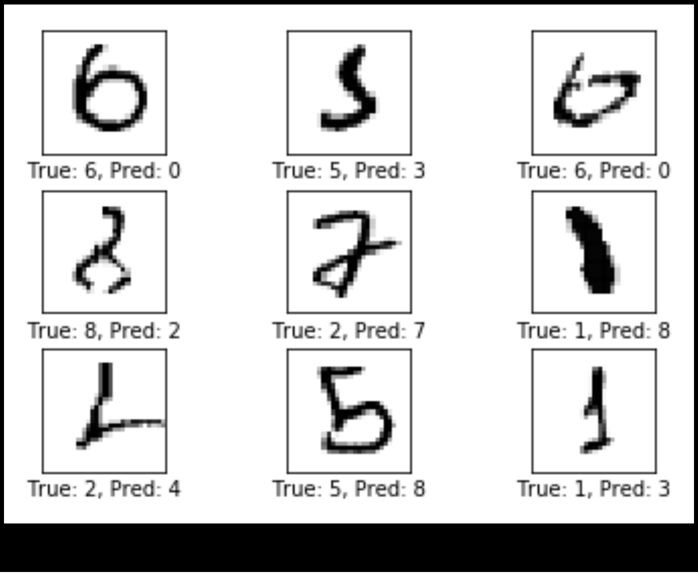


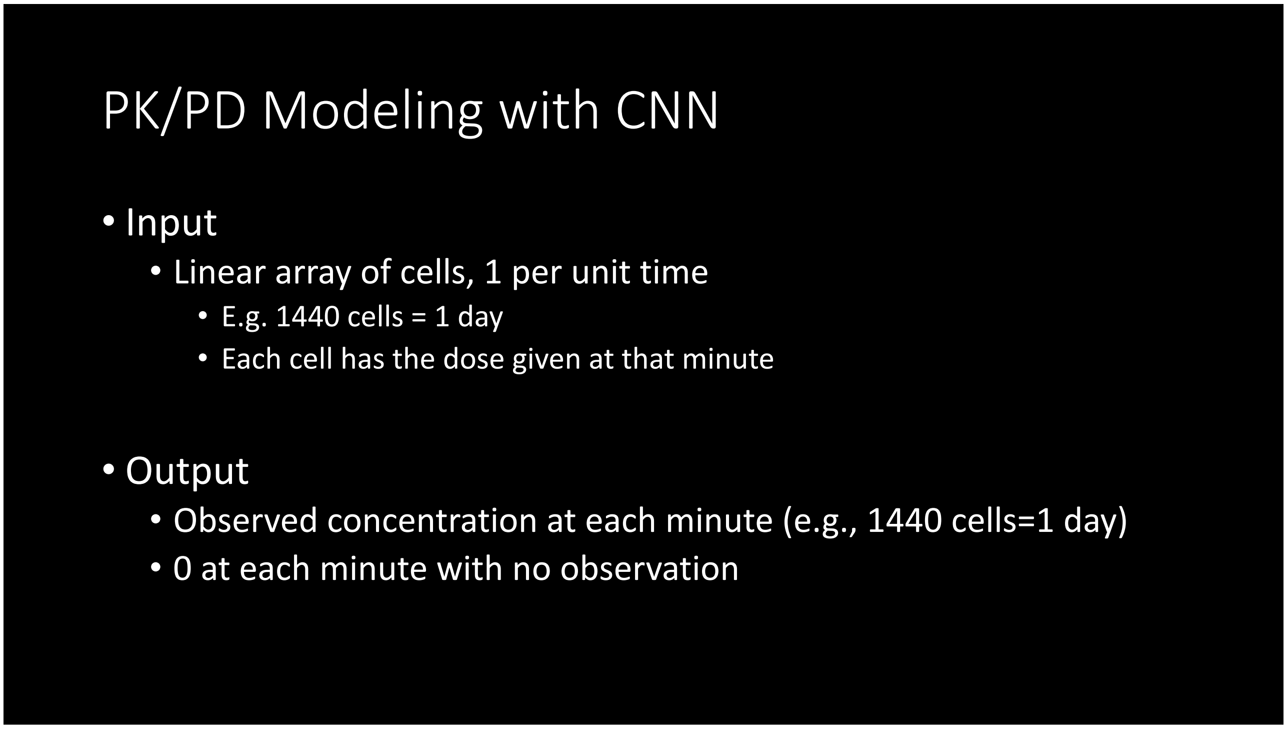
112 time units →

And 112 time units (months, clinic visits, etc)



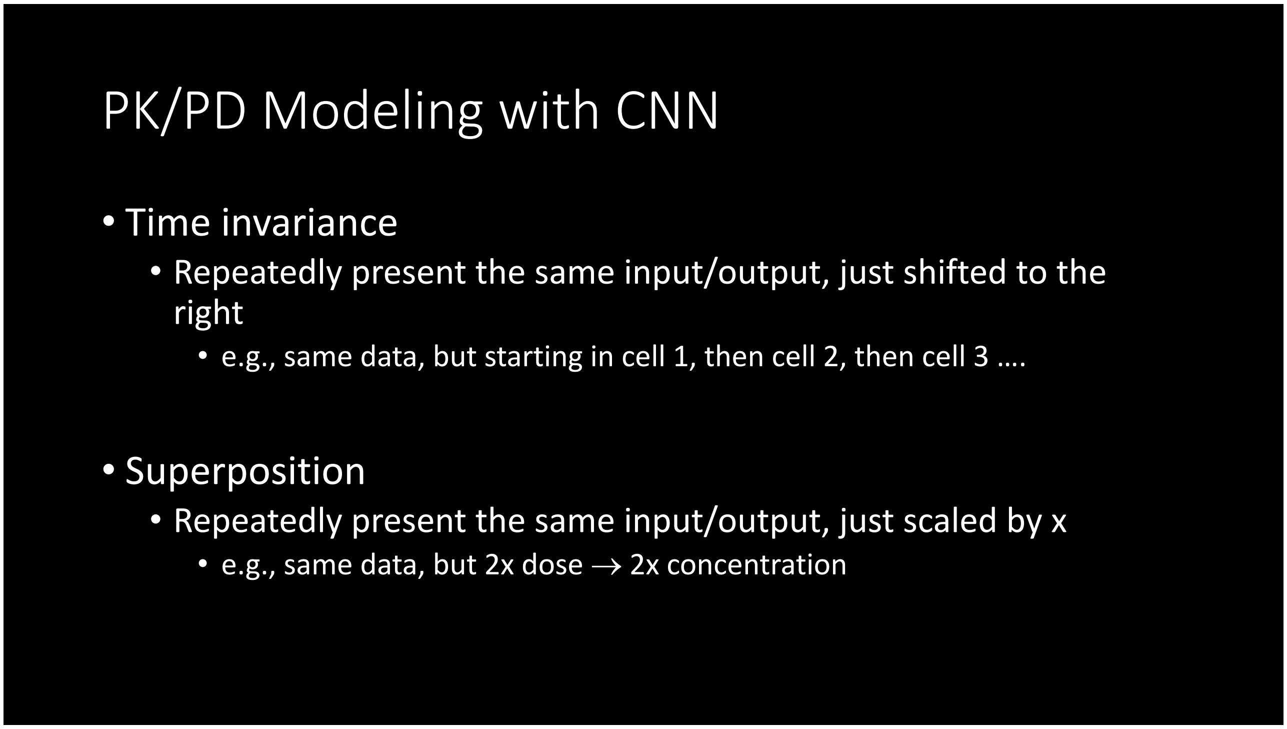
30,000 Iterations required 4 minutes
Predicted digits with 99% accuracy





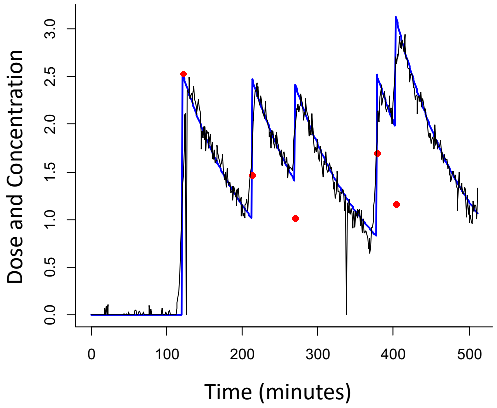
PK/PD Modeling with CNN

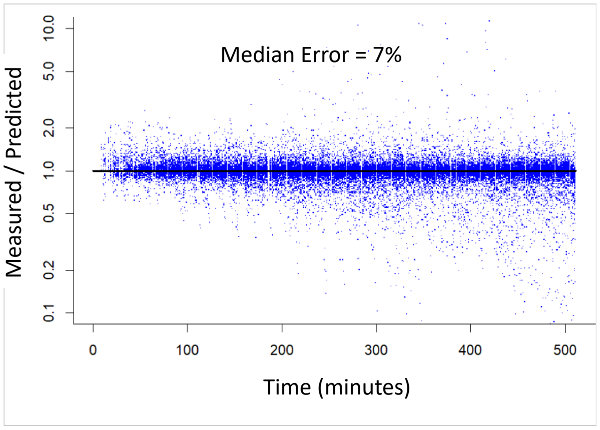
- Input
  - Linear array of cells, 1 per unit time
    - E.g. 1440 cells = 1 day
    - Each cell has the dose given at that minute
- Output
  - Observed concentration at each minute (e.g., 1440 cells=1 day)
  - 0 at each minute with no observation



PK/PD Modeling with CNN

- Time invariance
  - Repeatedly present the same input/output, just shifted to the right
    - e.g., same data, but starting in cell 1, then cell 2, then cell 3 ….
- Superposition
  - Repeatedly present the same input/output, just scaled by x
    - e.g., same data, but 2x dose → 2x concentration

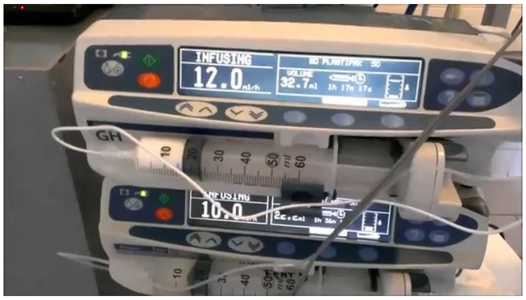




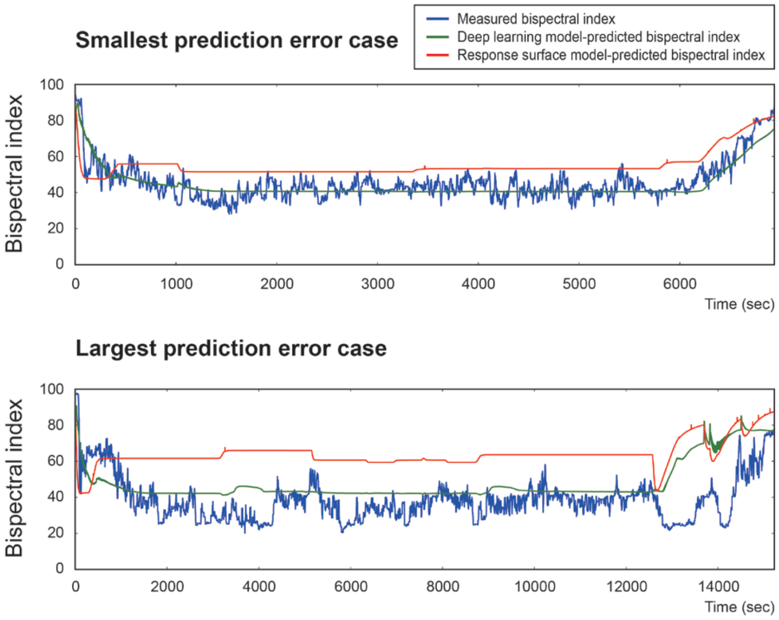
Prediction of Bispectral Index during Target-controlled Infusion of Propofol and Remifentanil

A Deep Learning Approach

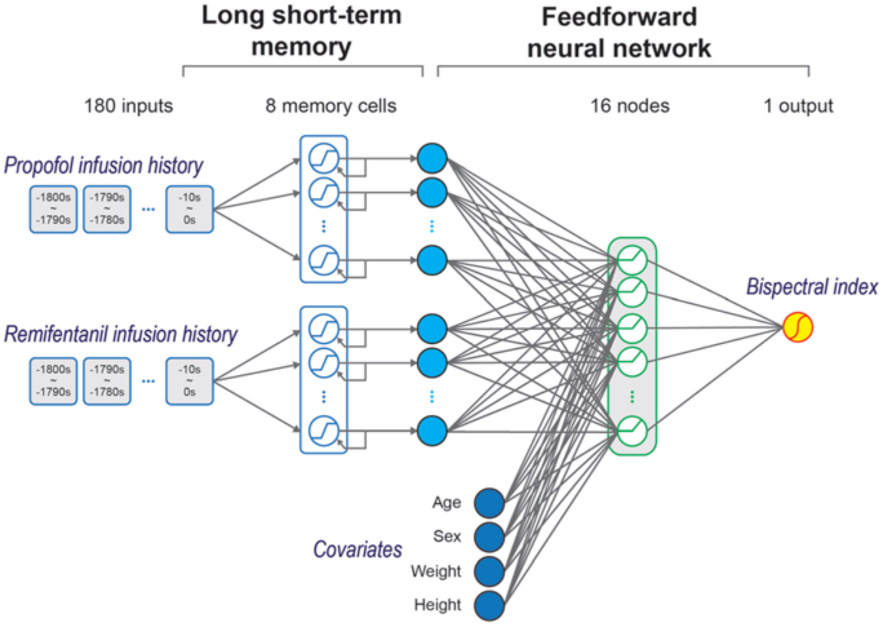
Hyung-Chul Lee, M.D., Ho-Geol Ryu, M.D., Ph.D., Eun-Jin Chung, M.D., Chul-Woo Jung, M.D., Ph.D.



(Anesthesiology 2018; 128:492-501)

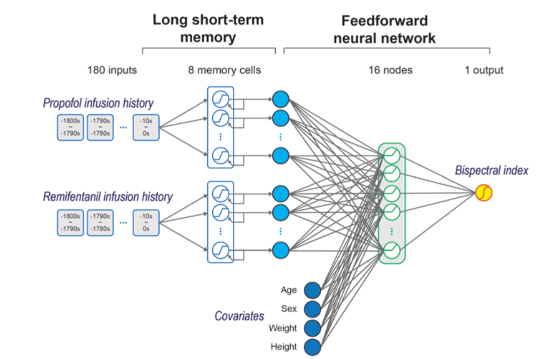


(Anesthesiology 2018; 128:492-501)



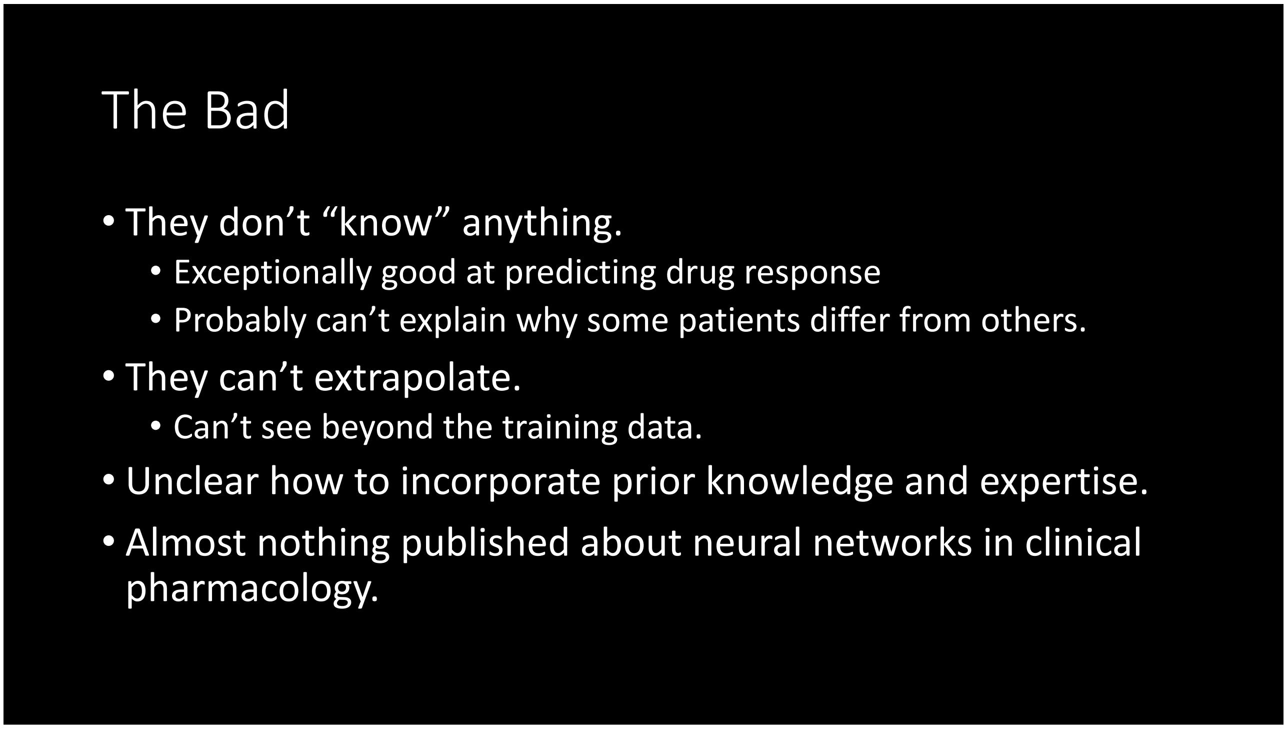
(Anesthesiology 2018; 128:492-501)

```
Propofol_NN       = LSTM(8, input=Propofol_Dose)
Remifentanil_NN   = LSTM(8, input=Remifentanil_Dose)
Combined_NN       = merge([Propofol_NN, Remifentanil_NN, Covariates])
NN2               = Dense(16, input=Combined_NN)
BIS               = Dense(1, input=NN2)
```



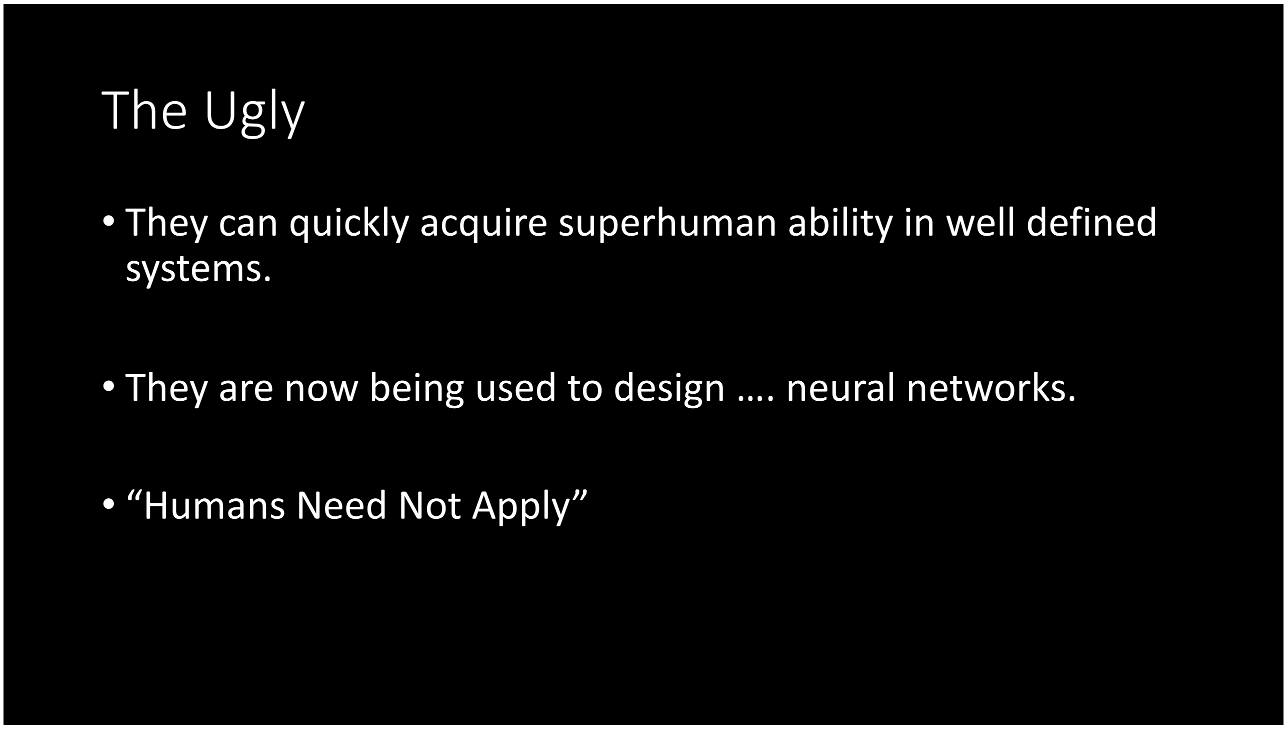
The Good

- Huge interest in many domains.
- They work.
  - If you aren’t convinced, ask Alexa
- They always work.
- They scale.
  - Google Translate, built on a 7 layer LSTM model can translate any language to any other language.
  - Google uses neural networks to recognize places, faces, and features in millions of photographs every day
- TensorFlow + GPU/TPU = tractable Neural Networks



The Bad

- They don’t “know” anything.
  - Exceptionally good at predicting drug response
  - Probably can’t explain why some patients differ from others.
- They can’t extrapolate.
  - Can’t see beyond the training data.
- Unclear how to incorporate prior knowledge and expertise.
- Almost nothing published about neural networks in clinical pharmacology.

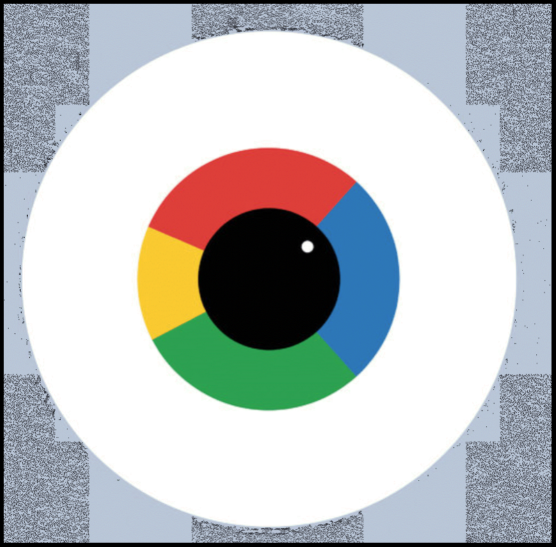


The Ugly

- They can quickly acquire superhuman ability in well defined systems.
- They are now being used to design …. neural networks.
- “Humans Need Not Apply”

Thank You

Sleep Well



Google Is Watching You