

CMSC 373 Artificial Intelligence Fall 2025 16-Convolution Networks

Deepak Kumar
Bryn Mawr College

1

Deep Learning Revolution

- **Deep Learning** refers to deep neural networks (i.e. many hidden layers)
- The “deep” in Deep learning is NOT “learning that is deep” (i.e. meaningful or sophisticated learning!!)
- The “deep” ONLY refers to the “**depth in layers**” of the neural network.
- **Convolution Networks** are a kind of Deep Neural network.

2

2

Taking Inspiration from the Visual Cortex

- Hubel & Wiesel's experiments with cats and primates
 - Visual cortex is a hierarchical series of layers of neurons. Layers communicate back and forth extensively.
 - Layers act as feature detectors (edges, shapes, objects, etc).
 - Each neuron receives input corresponding to a specific small region of the visual scene in a neuron's receptive field).
 - Neurons activate only if their receptive field contains a particular kind of edge/feature (e.g. horizontal edge, vertical edge, angular edge, etc)
 - Lower-level neurons feed into higher level layers of the visual cortex for detecting shapes, objects, faces, etc.
- This is still a gross simplification. The brain is much more complex!



Image From: <https://www.news-medical.net/life-sciences/How-do-Visual-Neurons-Work.aspx>

3

3

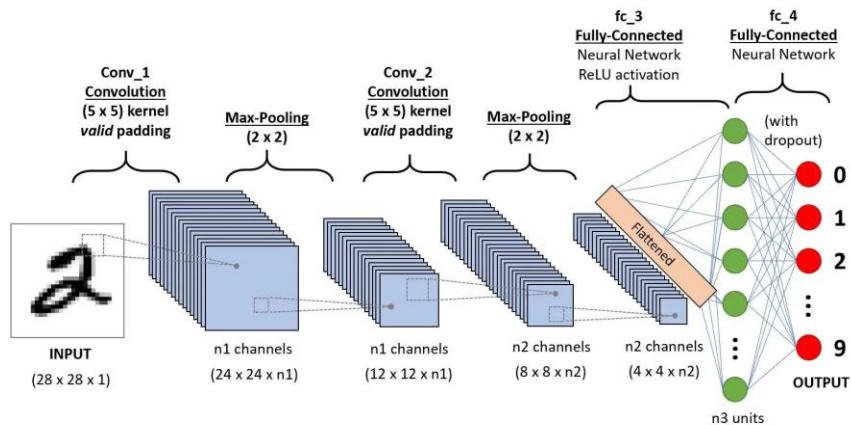
Convolution Networks – Short History

- Taking inspiration from Hubel & Wiesel...
- From Cognitron, to NeoCognitron (Fukushima, 1970s)
Showed how a hierarchical network (using Relu!) could learn using unsupervised means.
- To Convolutional Neural Networks (1989) used for hand-written zipcodes. Developed by Yann LeCun (at AT&T Bell Labs). In 1995, LeCun et al developed LeNet-5 to classify handwritten digits (32x32 pixel images). Used to recognize numbers on checks by banks.

4

4

ConvNet Architecture



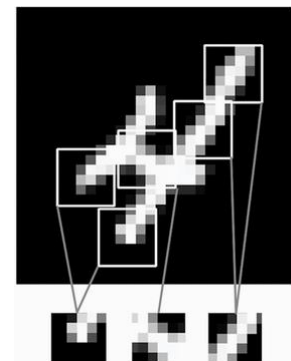
From: <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>

5

5

Dense Networks *versus* Convolution Networks

- Layers in dense networks learn **global patterns** in the input
E.g. MNIST Digit recognition: we flattened 28×28 images into 784 units and fed them into the hidden layer.
- Convolution Networks learn **local patterns** in input
e.g. They look for patterns in small 2D windows (using *patches/filters/kernels*) of input images.
- Convolution Networks learn spatial hierarchies of patterns (e.g., edges, larger patterns, etc.)

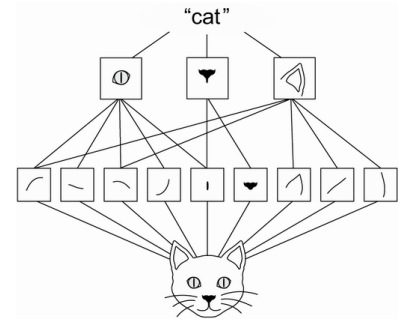


6

6

Convolution Networks: Core Idea

- Convolution Networks learn **local patterns** in input
e.g. They look for patterns in small 2D windows (using *patches/filters/kernels*) of input images.
- Convolution Networks learn **spatial hierarchies of patterns** (e.g., edges, larger patterns, etc.)



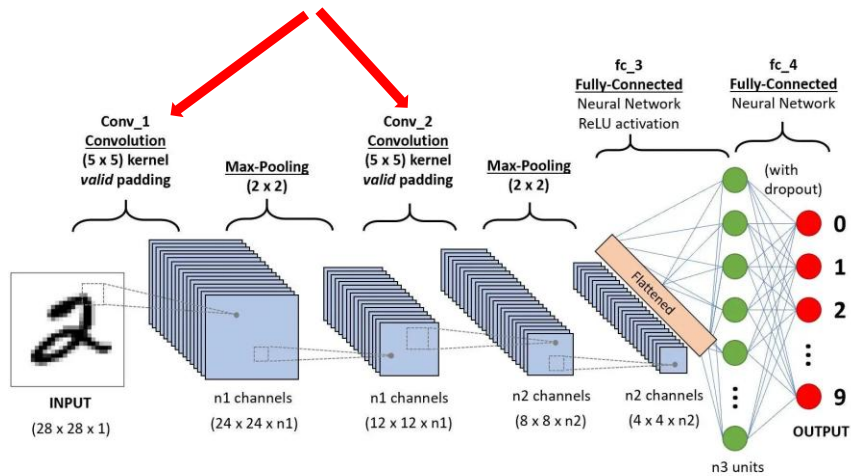
7

Convolution Networks: Basic Elements

- **Convolution** (using filters/patches/kernels)
Small patches of input image/map are filtered to recognize local features
- **Feature Maps**
Starting with an input image (e.g. 28x28x1)
Using a 3x3 filter to get K output maps (26x26xK)
- **Max Pooling**
A way of down sampling a feature map.

8

ConvNet Architecture: Convolutions

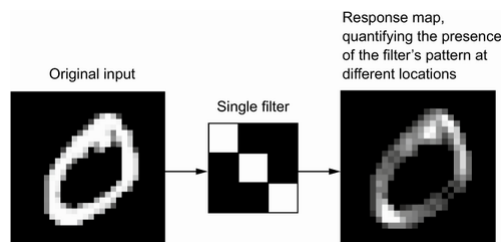
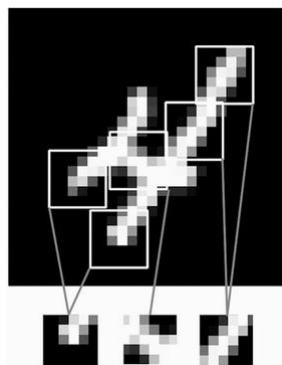


From: <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>

9

9

Filters/Patches/Kernels

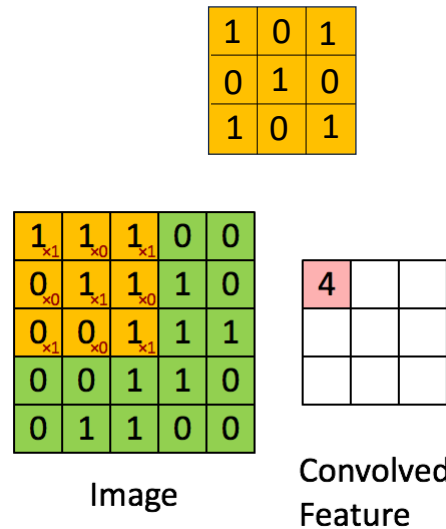
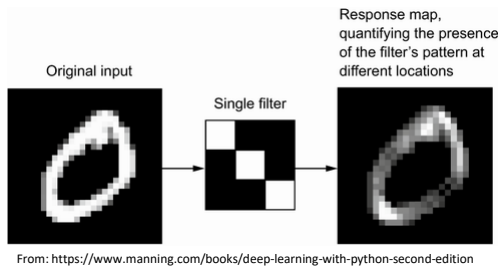


From: <https://www.manning.com/books/deep-learning-with-python-second-edition>

10

10

Filters/Patches/Kernels



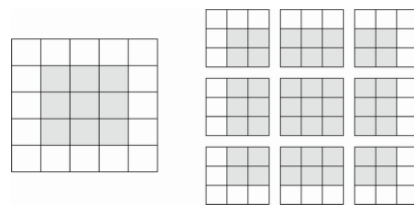
11

11

Filters: Padding

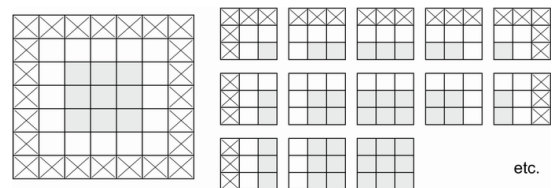
- “valid” Padding

No padding of input is done.
Reduces the width and height of resulting map.



- “same” Padding

Pads input in such a way to preserve the width and height of the resulting map.

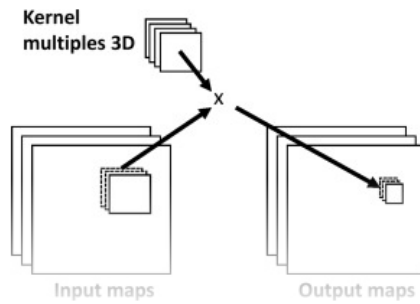


From: <https://www.manning.com/books/deep-learning-with-python-second-edition>

12

12

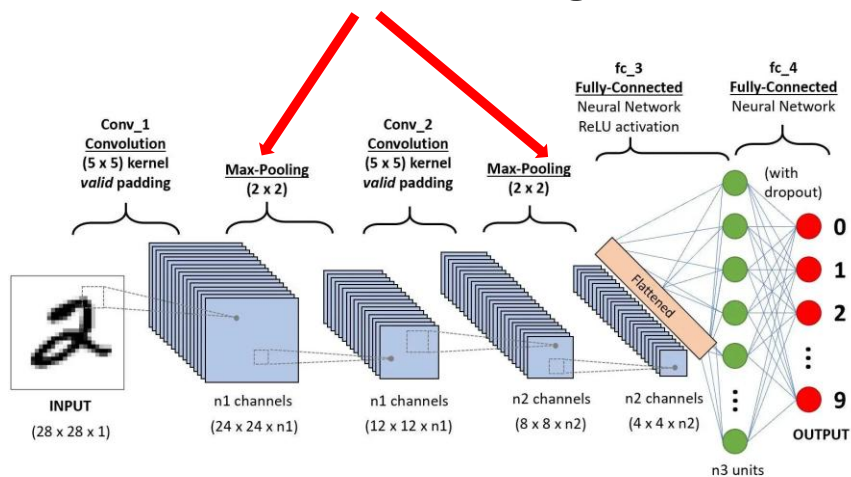
Input and Output Feature Maps



13

13

ConvNet Architecture: Pooling



From: <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>

14

14

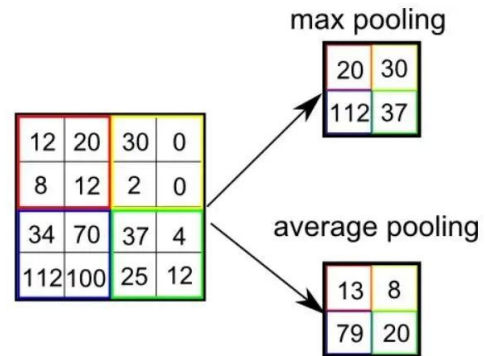
Max Pooling

- Pooling uses a kernel/patch/window

There two popular kinds: **Max pooling**,
Average Pooling

- It is purely an arithmetic operation that helps further downsample a feature map (i.e. performs dimensionality reduction useful for reducing computational load).

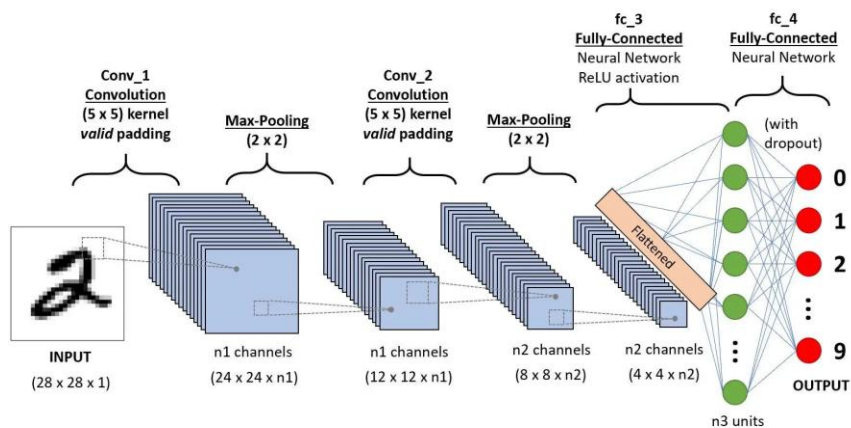
- We typically use a 2x2 window.



15

15

ConvNet Architecture

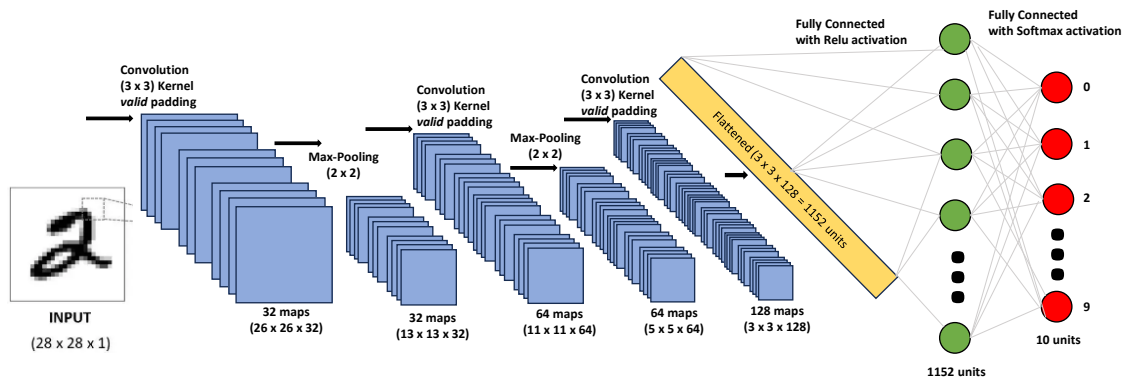


From: <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>

16

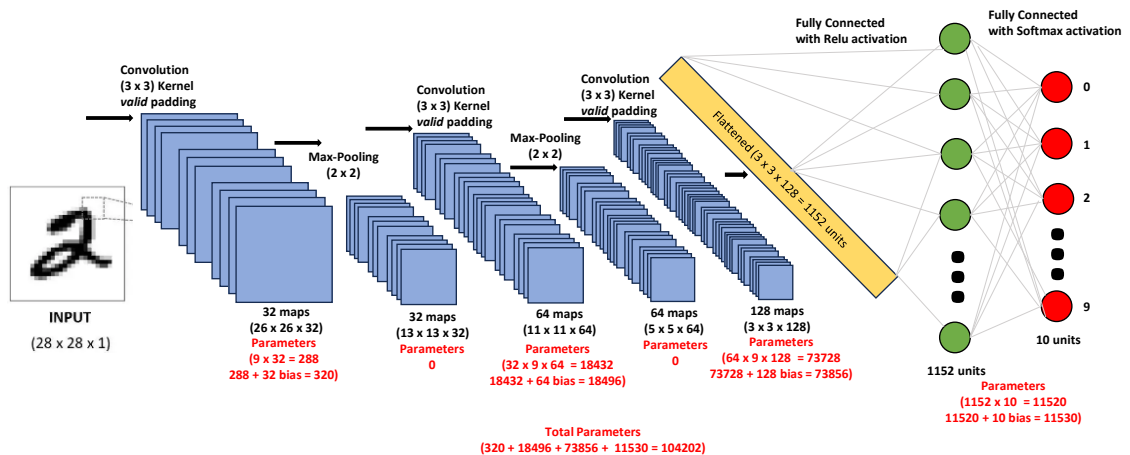
16

ConvNet for MNIST Digits Recognition



17

ConvNet for MNIST Digits Recognition



18

Over to Colab

- <https://colab.research.google.com/drive/1XqXJiz20EAgHStE2qsPiFa2rkL91uQwp?usp=sharing>

19

19

ConvNet Architecture Patterns

- **Overfitting/Dropout**

When models tend to overfit, a way to resolve overfitting is to use **dropout**. This is called a **regularization technique**.

A dropout layer can be added in between layers. Typically, a dropout layer eliminates (sets to 0) 20-50% of the outputs (dropout value can be set during model/layer specification).

- **Data/Batch Normalization**

Normalizing the values on inputs and outputs helps with gradient propagation and allows for deeper networks. Hence liberally used in very deep network architectures.

- **Data Augmentation**

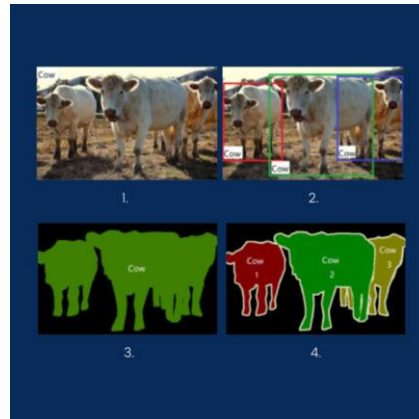
When datasets are small, “new” data can be created by **transforming** images in the dataset using flipping, rotation, zooming, etc. to augment the dataset with transformed images. Helps avoid overfitting and leads to better generalization.

20

20

ConvNets Applications

- Input Image (1)
- Image classification
- Image Segmentation (3, 4)
- Object Detection (3)



From: <https://www.smart-interaction.com/2022/07/14/computer-vision-the-ultimate-guide-on-the-4-main-tasks/>

21

21

CovNet Architectures

• AlexNet

Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton, University of Toronto, 2017.

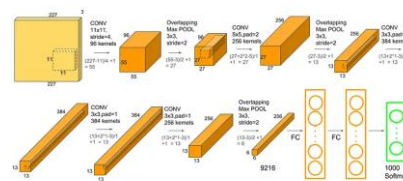
One of the early CNNs to use GPUs to speed up training (in 2012).

• VGGNet

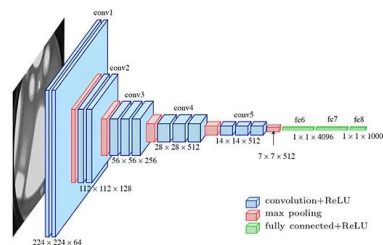
Karen Simonyan and Andrew Zisserman, University of Oxford, 2014.

Groundbreaking DL model for Object Recognition.

- There are many more.



From: <https://medium.com/analytics-vidhya/concept-of-alexnet-convolutional-neural-network-6e73b4f9ee30vgg>



From: <https://medium.com/analytics-vidhya/vggnet-convolutional-network-for-classification-and-detection-3543aaf61699>

22

22

References

- F. Chollet: *Deep Learning with Python, Second Edition*, Manning 2021.
- A. Geron: *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow 2nd Edition*. O'Reilly, 2019.
- M. Mitchell: *Artificial Intelligence: A Guide For Thinking Humans*, Farrar, Strouss, Giroux, 2019.
- M. Wooldridge: *A Brief History of Artificial Intelligence*. Flatiron Books, 2020.

23