

# Convolutional Neural Networks

Dev Singh

CS Seminar: Machine Learning

May 1, 2024

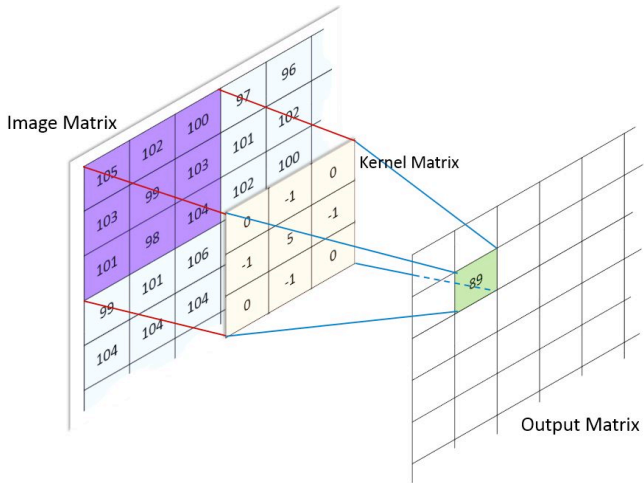
# Key Terms

- ▶ Kernel - filter that is convolved over the image
  - ▶ Kernel size -  $m \times n$  matrix  $(m, n)$
- ▶ Stride - Number of pixels to move over on each convolution
- ▶ Padding - pixel dimensions and their value to add to each image
  - ▶ In practice, the border of each image is zero-padded to not lose meaningful edge data of the image.
- ▶ 2-D convolutions: convolutions in  $\mathbb{R}^2$  over multiple input planes (such as color images which have one input plane per color channel).
  - ▶ See Conv2D layer in PyTorch/Keras

# Overview of Concepts

- ▶ Overarching goal: to extract the most important spatial features from image or image-like data, by processing through a network of convolutional filters.
- ▶ Introduced for image classification by LeCun et al. (1989).
- ▶ Filter  $w$  is convolved with the image  $X$  with chunks  $x$ , i.e., “slide the filter over chunks of the image, computing the dot products”.
- ▶ Dimensions of input tensor: [batch size  $\times$  colors  $\times$  frames  $\times$  length  $\times$  width]
- ▶ One convolution outputs  $w^T x + b$ , where  $b$  is bias.
- ▶ Used for “feature extraction” to extract important traits of the provided image.

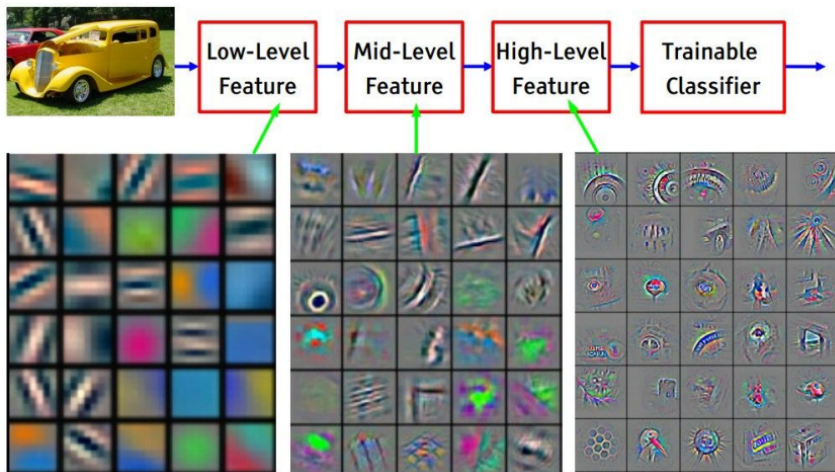
# Conv2D Visualization



# Activation Maps

- ▶ Each filter set creates an activation map that represents features in the image that the specific kernel extracted.
  - ▶ For example, if there are six  $5 \times 5$  filters convolved over images in the Fashion-MNIST dataset which are  $32 \times 32$ , the output tensor would have dimensions  $[6 \times 28 \times 28 \times 6]$ .
- ▶ The filters should reduce the sizes of the feature map ( $f: \mathbb{R}^a \mapsto \mathbb{R}^b, b \ll a$ ), but not too quickly.
- ▶ The values contained in the kernel are tuned as part of the training process.
- ▶ The kernel size, stride, and padding are hyper-parameters that are set when the model is designed.

# Feature Map Visualization



# Pooling Layers

- ▶ Used to further reduce the feature dimensionality of the activation maps
- ▶ Has the same hyperparameters as a “traditional” convolving layer.
- ▶ 2 types: max pooling vs. average pooling
  - ▶ Take the maximum of an area vs. the average of an area in the feature map.
  - ▶ Max pooling shown to perform better in most cases as it reduces noise.
- ▶ Emphasizes features that were found with the convolutional layers.

# MaxPool1D Visualization

2	2	7	3
9	4	6	1
8	5	2	4
3	1	2	6

Max Pool  
→

Filter - (2 x 2)  
Stride - (2, 2)

9	7
8	6



# Sources

- ▶ <https://pytorch.org/docs/stable/generated/torch.nn.Conv2d.html>
- ▶ <https://pytorch.org/docs/stable/generated/torch.nn.MaxPool2d.html>
- ▶ <http://yann.lecun.com/exdb/publis/pdf/lecun-99.pdf>