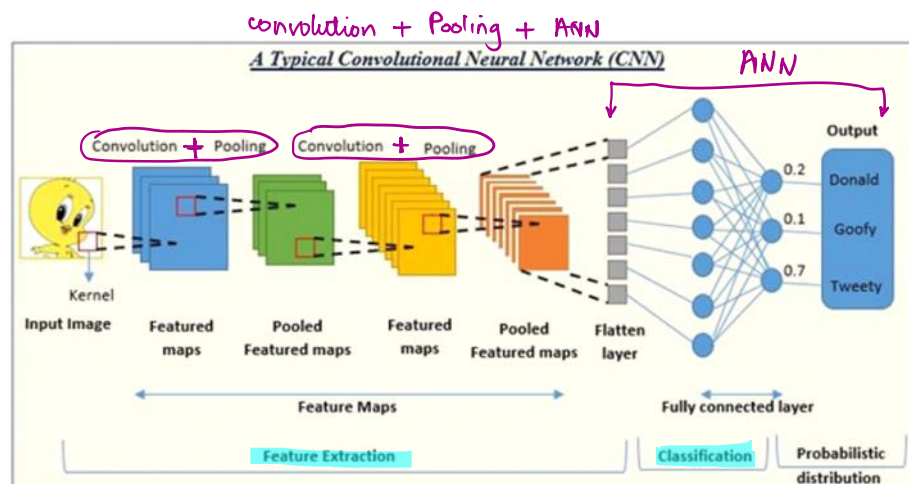
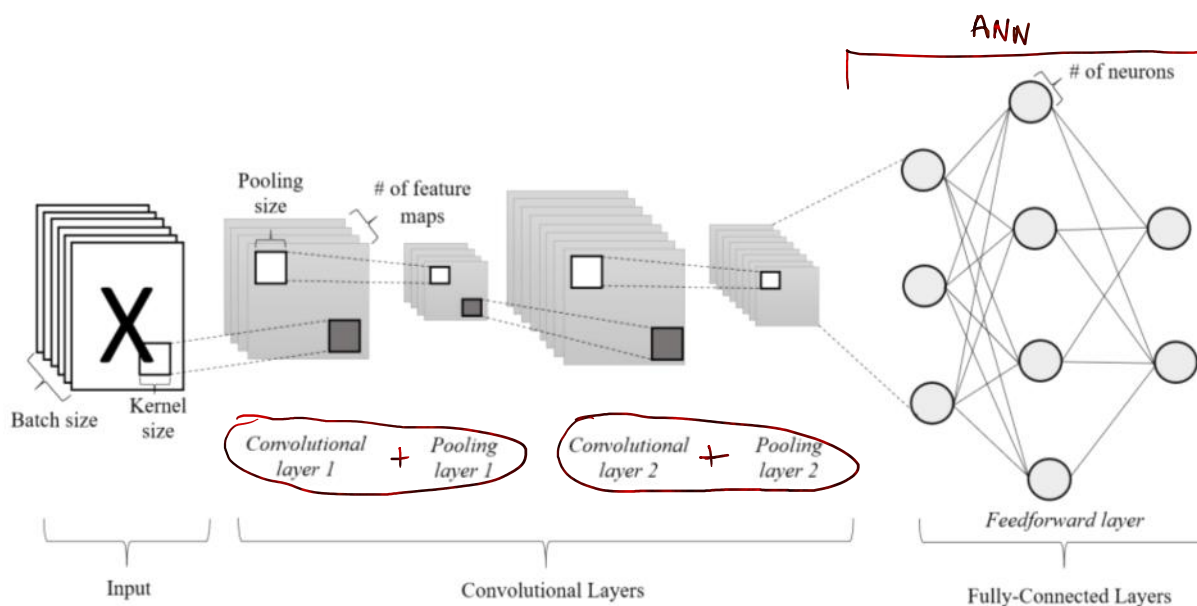
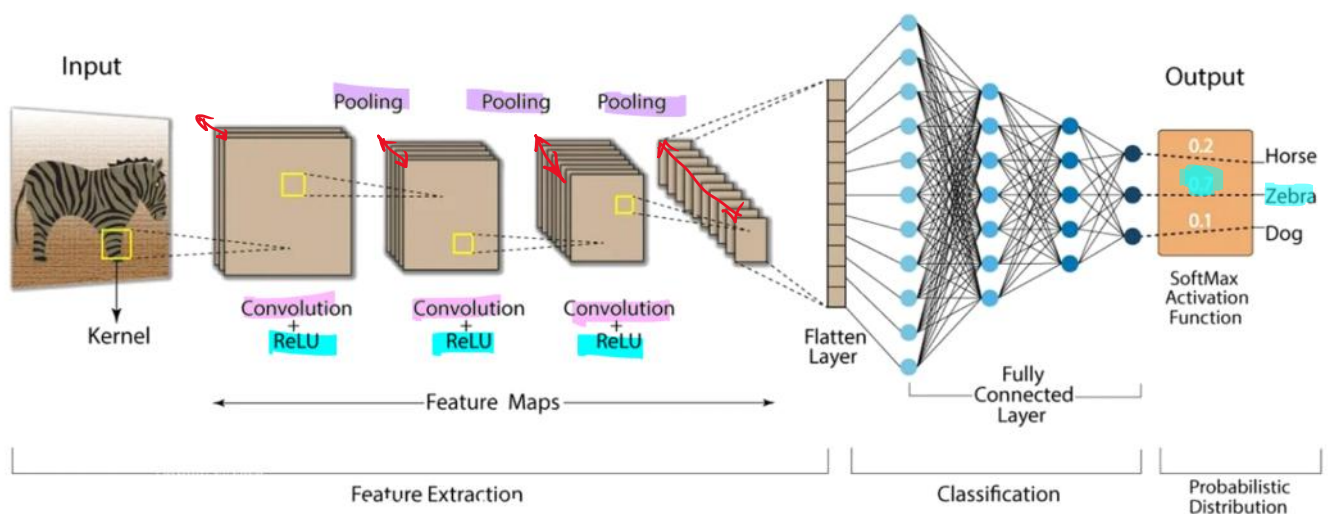


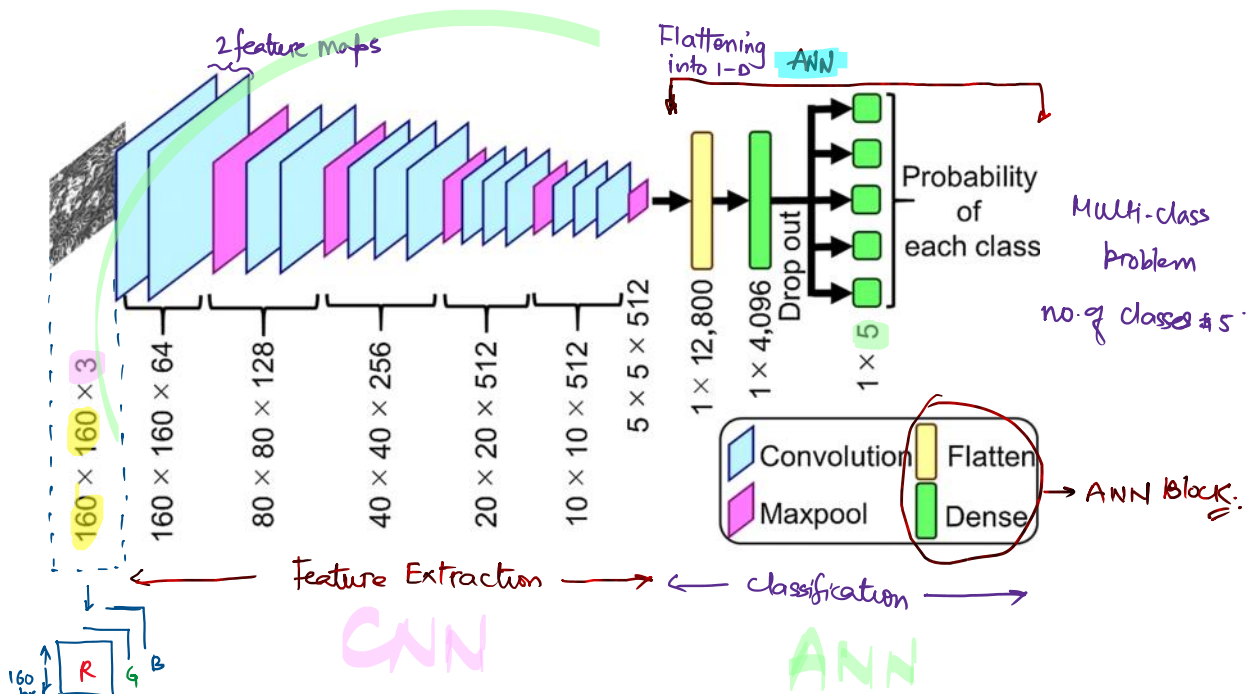
CNN Model Architecture

09 November 2025 11:31



Convolution Neural Network (CNN)





Two key terms are:

- ① CONVOLUTION
- ② POOLING

1. INPUT LAYER

Input layer of a CNN model receives the raw image data in the form of multi-dimensional array called TENSORS

For grayscale image

MNIST Use-case

Input Tensor

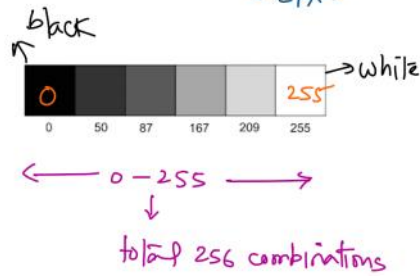
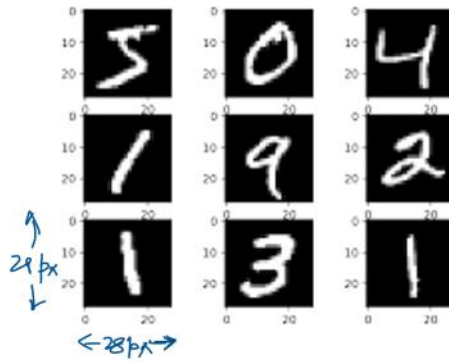
For each image: $(H \times W \times 1)$

H: Height of the image (no. of pixels vertically)

W: width of the image (—"—" horizontally)

1: single color channel: grayscale

a grayscale image has no color but only intensity information



For color images: **RGB**

Input Tensor: $(H \times W \times 3)$

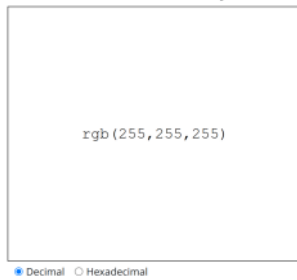
↳ represents three color channels
 ↳ Red
 ↳ Green
 ↳ Blue

<https://www.csfieldguide.org.nz/en/interactives/rgb-mixer/>

RGB Colour Mixer - Used by Screens



RGB Colour Mixer - Used by Screens



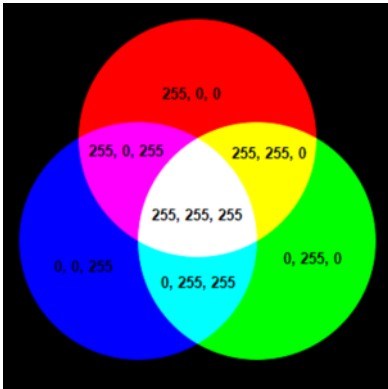
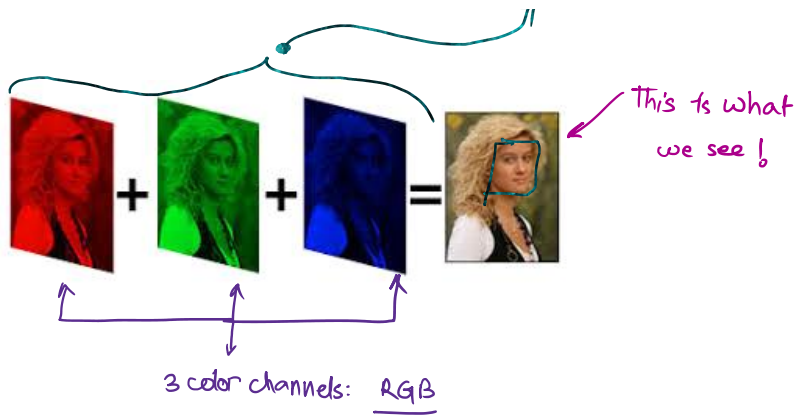
RGB Colour Mixer - Used by Screens



RGB Colour Mixer - Used by Screens



↳ This is what



Note: CNNs are naturally designed to work with multi-channel inputs

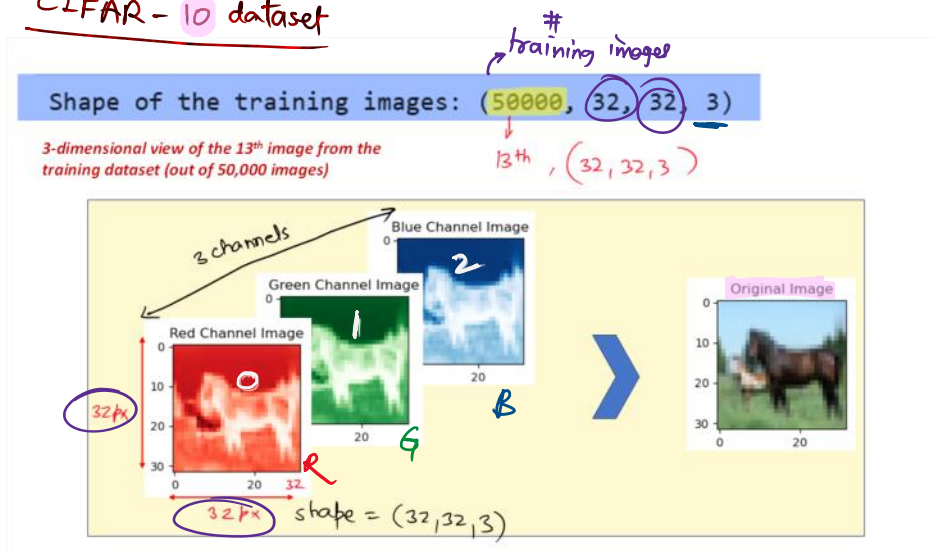
[batch-size, 3, height, width] \rightarrow 4-dimensional array



\downarrow \downarrow

(50,000) images no. of channels

In-training set.

CIFAR-10 dataset \rightarrow 10 classes \rightarrow <https://www.cs.toronto.edu/~kriz/cifar.html>



	Grayscale	RGB	(In CNN model)
① MNIST Handwritten Digits Recognition	✓	✗	shapes / patterns matter not color.
② Medical Imaging (X-rays, CT scan)	✓	✗	shapes / patterns matter Intensity is enough.
③ Need extra information / compare colors	✗	✓	stop sign → Red  Pass/Go → Green 

Conclusion: Build the baseline model using 'grayscale' and switch over to RGB if needed.

<https://imagecolorpicker.com/>



model will have tough time!



(model will struggle)


Yes, blurring does hurt CNN models.

2. CONVOLUTION (FILTERING) LAYER

convolution is at the heart of CNN algorithm.

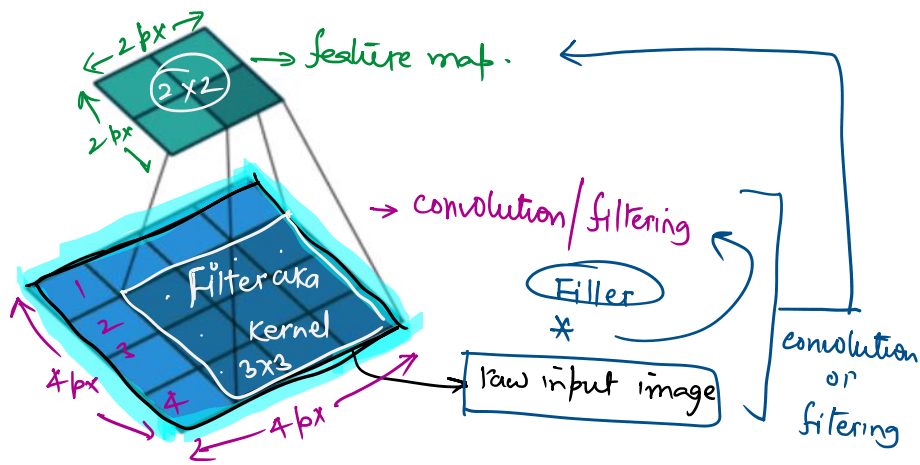
f_3 $f_1 \times f_2$ Convolutional Neural Network

$$y(t) = x(t) * h(t)$$

Input  Filter → feature map
image

The process is called convolution.

Convolution is a mathematical operation - that blends two functions to produce a third function expressing how one modifies or shapes the other.



https://maucher.home.hdm-stuttgart.de/Pics/gif/no_padding_no_strides.gif

https://adamharley.com/nn_vis/cnn/2d.html

