POOLING LAYER IN CNN

The problem with Convolution:
- Memory issues
- Translation variance

(i) Memory issues:

$228 \times 228 \times 3$ * 3×3 $\longrightarrow$ $(226 \times 226) \times 100$

input RGD   100 filters   feature map size.
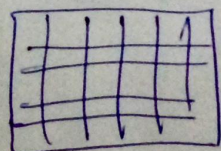image

- for output feature map if values are stores in bit values them, and with also 100 training. data   them.

$$226 \times 226 \times 100 \times 32 \times 100 = 1.5 \text{ GB storage}.$$

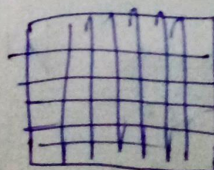- for reducing the size we can use the stride. or pooling

(ii) Translation Variance:

- in convolution operation the features detection are location dependent (translation variance)

- to resolve this issue we want a translation invariance. Solve by pooling by using down sampling operation.

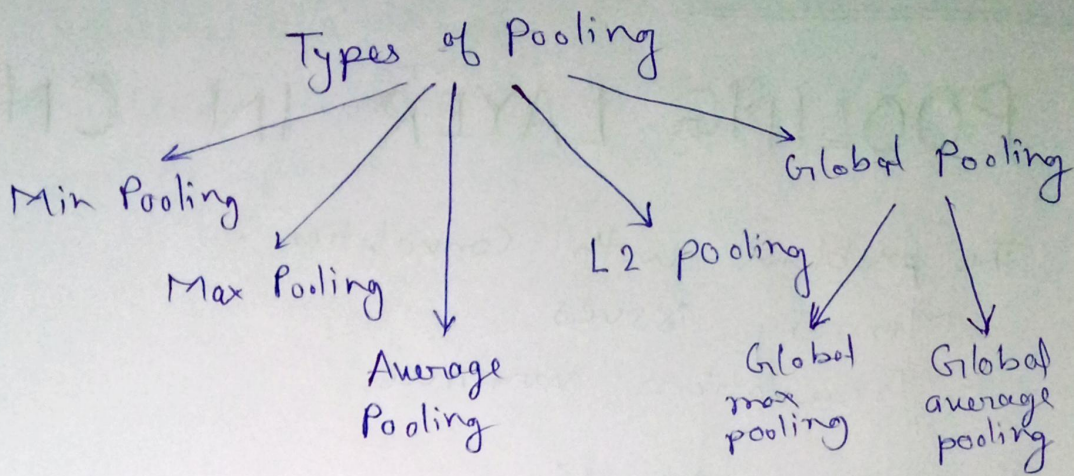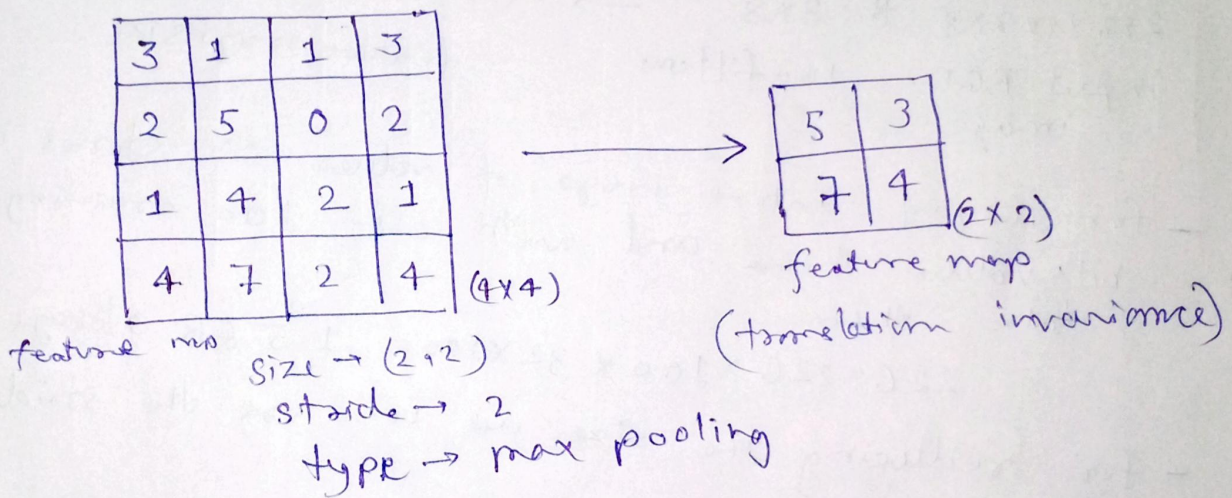- pooling is apply after the convolution operation.

feature map → ReLU → non-linear feature map → pooling

Types of Pooling

- Min Pooling
- Max Pooling
- Average Pooling
- L2 pooling
- Global Pooling
  - Global max pooling
  - Global average pooling

Example:

3	1	1	3
2	5	0	2
1	4	2	1
4	7	2	4

(4×4)

→

5	3
7	4

(2×2)

feature map
(translation invariance)

feature map

Size → (2,2)
stride → 2
type → max pooling

Advantages of Pooling:

(i) Reduce the size of feature map.

(ii) Translation invariance.

(iii) Enhanced features (only in case of max pooling)

(iv) There is no need of training (it is only a aggregate operation).

Disadvantages of Pooling:

(i) not used in image segmentation.

(ii) we loose lot of information. due to size decrease.

Types of CNN Architecture:

Types of CNN Architecture.
- LeNET
- AlexNET
- GoogleNET
- VggNET
- ResNET
- Inception module.
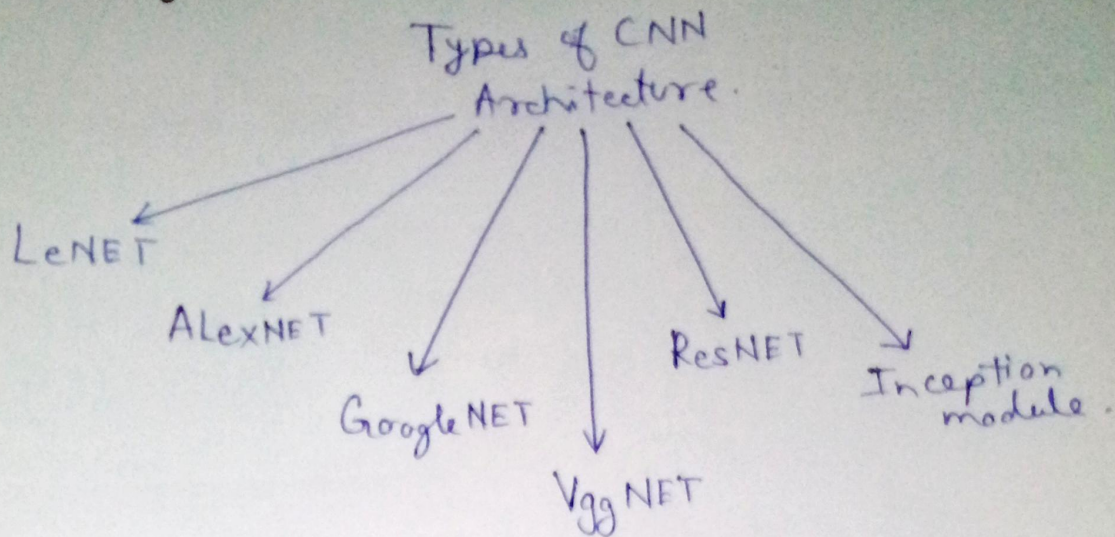
Difference between the ANN and CNN:

Aspect	ANN	CNN
Architecture	Fully connected layers (Dense Layers)	Combination of convolutional, pooling and dense layer.
input data	working well with 1D data (e.g. tabular data)	Design for 2D/3D data (e.g., image, videos).
Feature extraction	Manually engineered or no feature extraction	Automatic feature extraction using convolutional Layers.
spatial Hierarchy	No spatial awareness	capture spatial hierarchies through filters and pooling.
Translation invariance	Lacks of translation invariance	Achieves translation invariance through convolution and pooling
performance on image	poor performance on image data	Excellent performance on image data due to its ability to extract spatial features.