

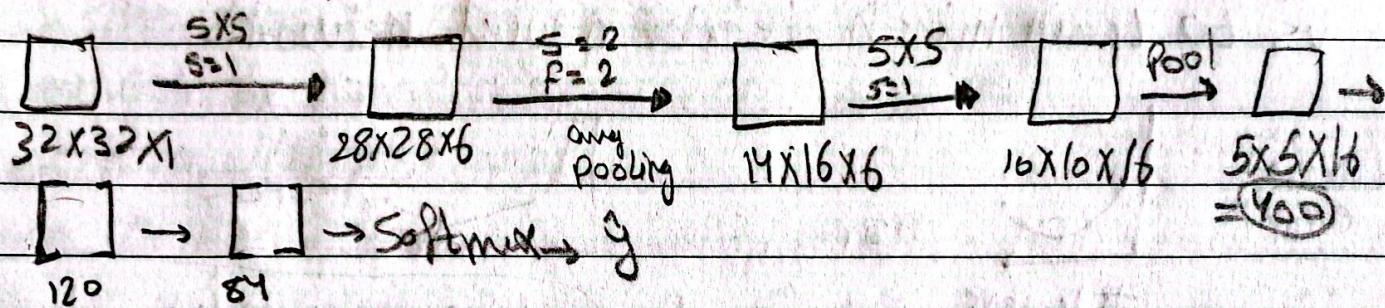
→ why Convolutions.

- pros of Conv. ① parameters sharing & space connection: means that in each layer, each output value depends only on small number of inputs

→ (week 2)

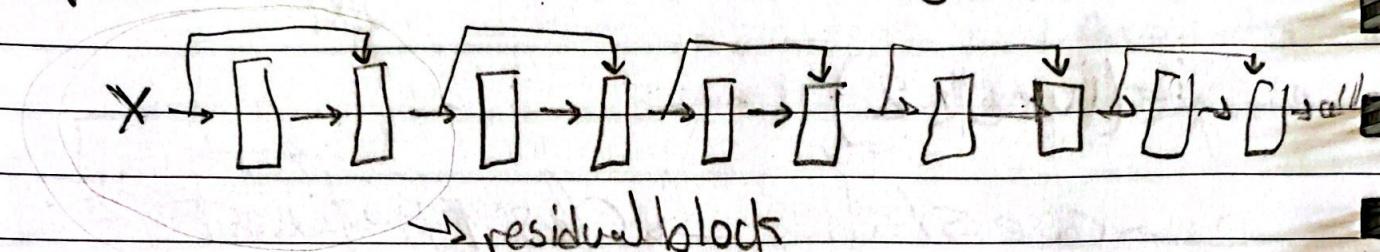
case studies → classic networks

- LeNet-5 : recognize handwritten

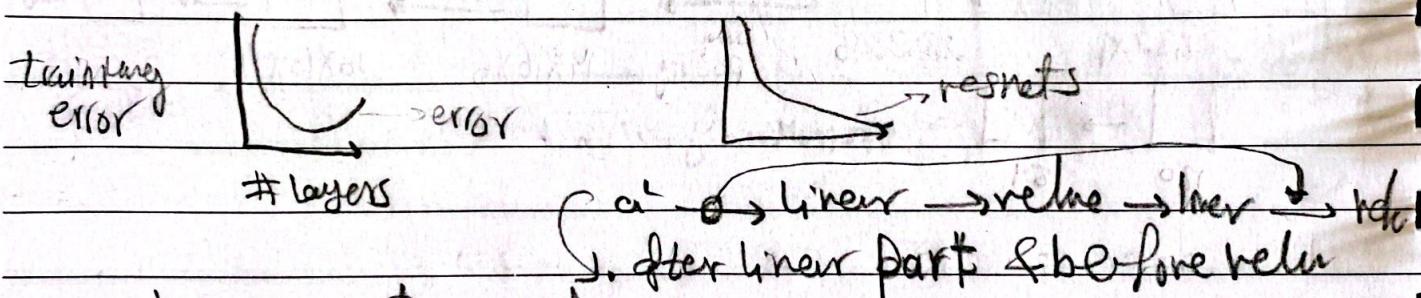


* Residual Networks (ResNets)

- we will learn about skip connections. That allows us to take the activation from one layer & feed it to another layer even much deeper in the NN. Focusing that we build ResNets.
- ResNets is built over residual blocks



- ResNets is made up of residual blocks & ResNets allows us to train much deeper neural networks



→ Why ResNets work

- Identity Function is easy for residual blocks to learn
- each residual block takes input $X \rightarrow$ series of transformations (conv. & batch normalization) then add transformation to original before act
- improve gradient flow - Avoidance of overfitting

• the output of Conv layers is added to the input & this addition enables the network to learn the residual mapping rather than the original mapping.

- original mapping: network attempts to learn from input to output through series of layers

- Residual Mapping: difference between the desired output & the original input. It allows the network to learn how to adjust

(N;N) : replace traditional conv. with multi-layer perceptrons

→ network in network & 1×1 Conv.

- techn. In NN to reduce parameters

↳ enhance feature extraction

$$\begin{array}{c} \boxed{} \\ 6 \times 6 \times 1 \end{array} * \begin{array}{c} \boxed{2} \\ \text{convolve } 1 \times 1 \end{array} = \begin{array}{c} \boxed{} \\ 6 \times 6 \end{array}$$

$$\text{if } 6 \times 6 \times 32 \xrightarrow{\text{Then}} 1 \times 1 \times 32 = 6 \times 6 \times \text{filters}$$

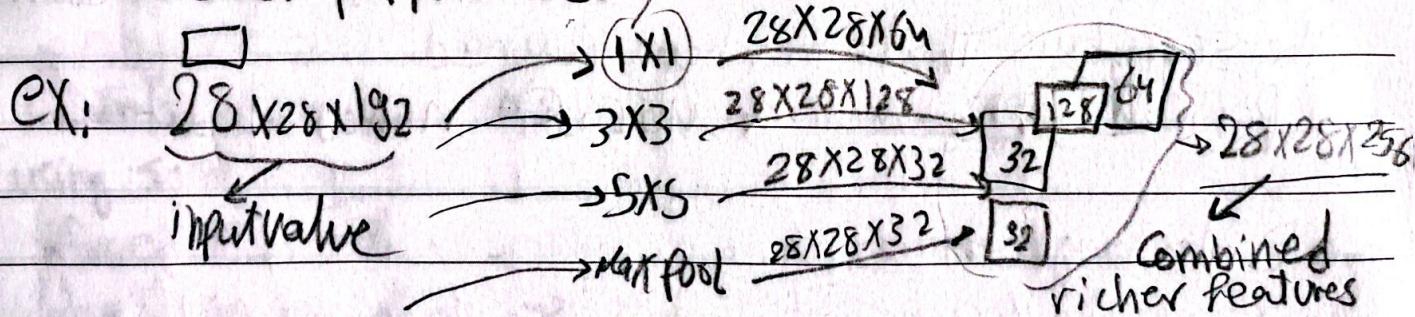
• if $(28 \times 28 \times 192)$ and we want to shrink num of channels we use 32 filter that are 1×1 . each filter will get dimension $1 \times 1 \times 192$ but using 32 filters the output will have $28 \times 28 \times 32$.

- each filter from the 32 mask sop on the 192 channels

• 1×1 conv used for dimensionality reduction & reduce channels if save computation

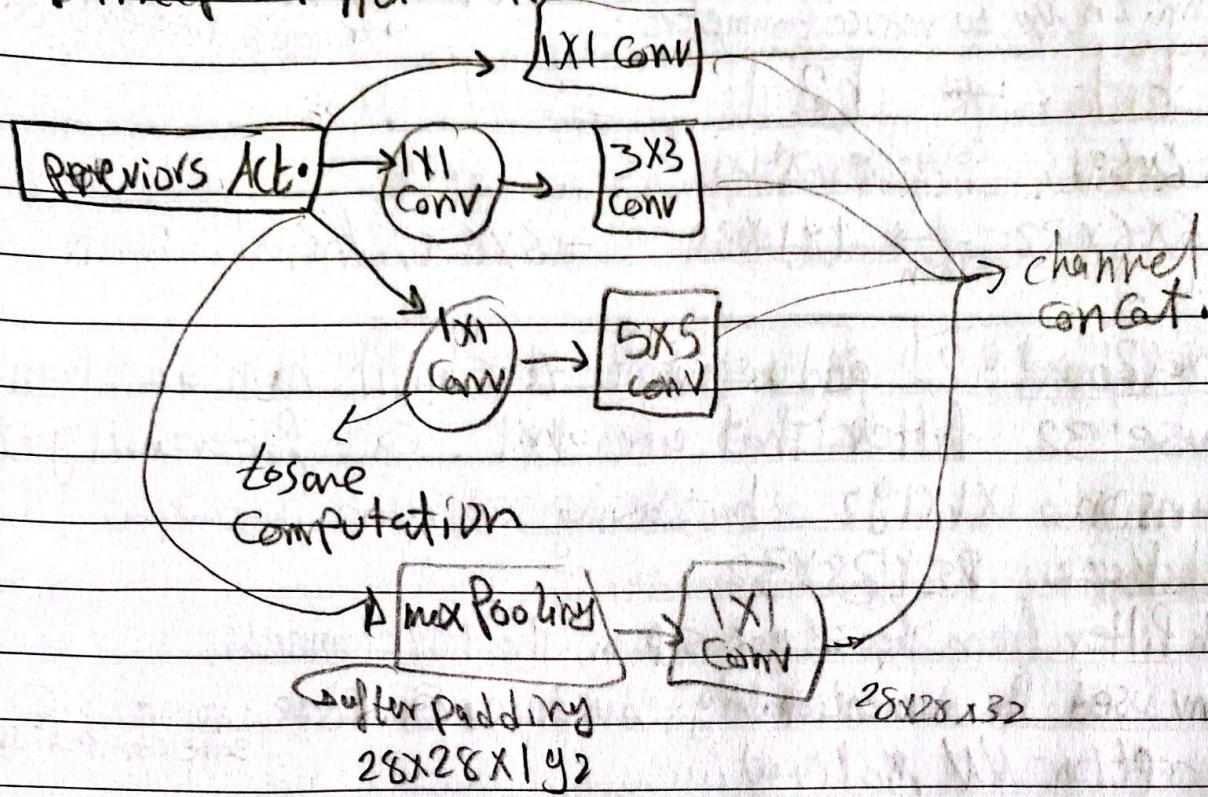
→ inception NN Motivation

• when designing the layers of CNN, we pick 3×3 or 5×5 filter or pooling layer, what inception network does is it let's use all of these, this makes the network more complicated but better performance ↗ reduce channels



• what inception network says that instead of choosing what filter size i want in conv or i want conv & pool, lets do them all

→ Inception Network:



→ Using open source implementation ✓

→ Transfer learning

- Rather than training the weights from scratch with random initialization process, it's more effective to download the structure that they have trained
- PASCAL, ImageNet & MS COCO names of datasets

→ Data augmentation

- one of the techniques that is used to improve the performance of Computer Vision

- mirroring
- random cropping
- rotation
- shearing
- local warping
- color shifting