CIS 678 Machine Learning

Convolutional Neural Networks (CNNs)

Comparison of Deep Learning models

ChatGPT-4: 1.8 trillion parameters

DeepSeek R1: 671 billion parameters

Can you design a model exceeding the number of parameters of these two models ?

Comparison of Deep Learning models

ChatGPT-4: 1.8 trillion parameters

DeepSeek R1: 671 billion parameters

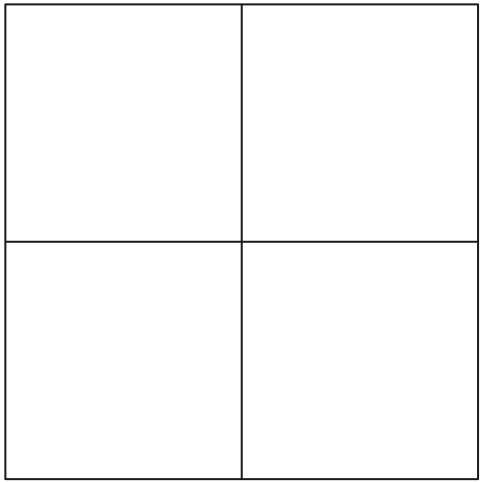
Can you design a model exceeding the number of parameters of these two models ?

It's challenging, but we will do it

Convolutional Neural Networks (CNNs)

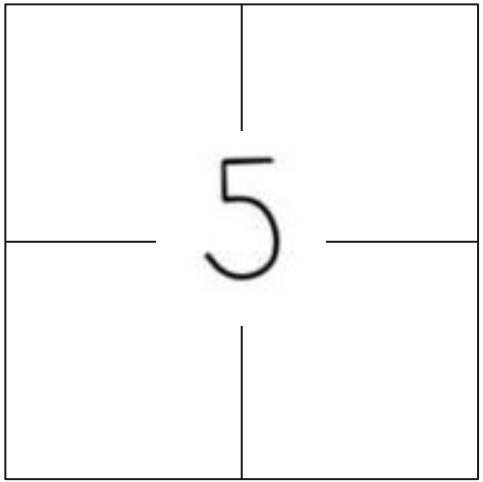
Why FF Networks don’t work well for Image Data?

- We ask you to write a “5”



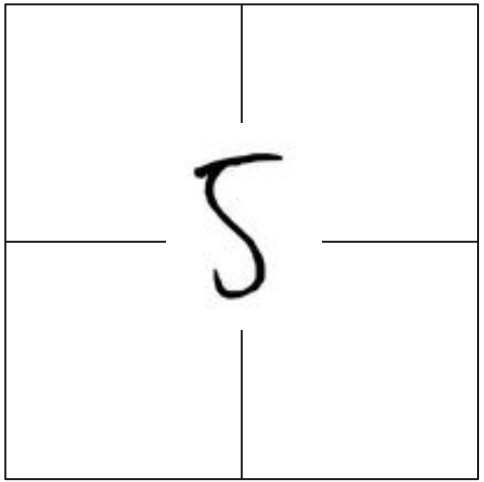
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- **A perfect "5" at the center of the coordinate**



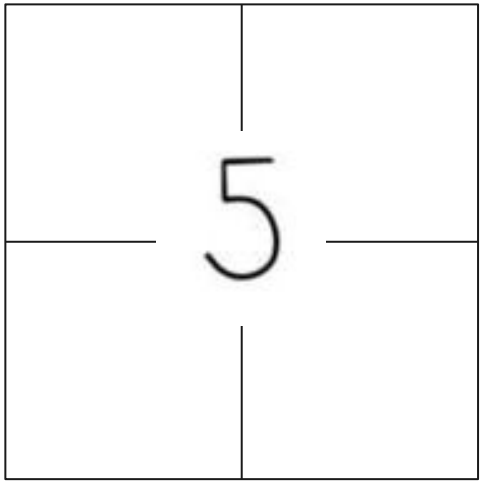
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- **This is my "5", not good as yours but still at the center of the coordinate.**



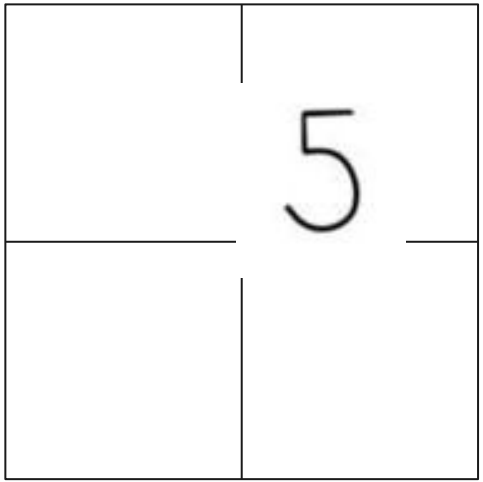
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations



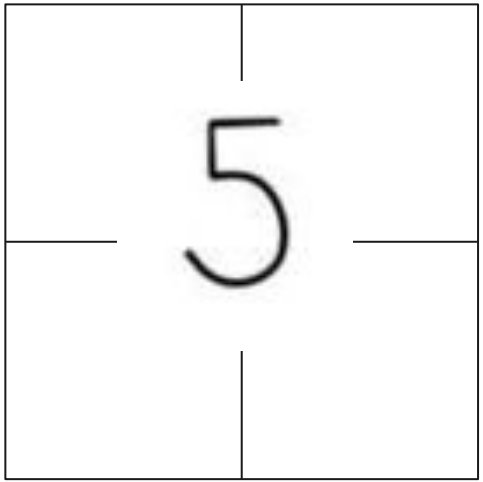
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations
- **Translated "5"**



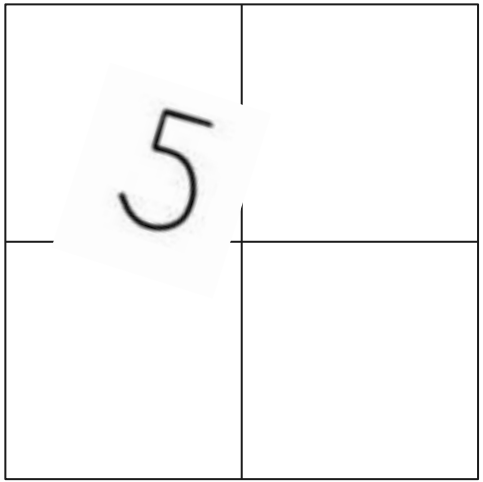
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations
- **Scaled "5"**



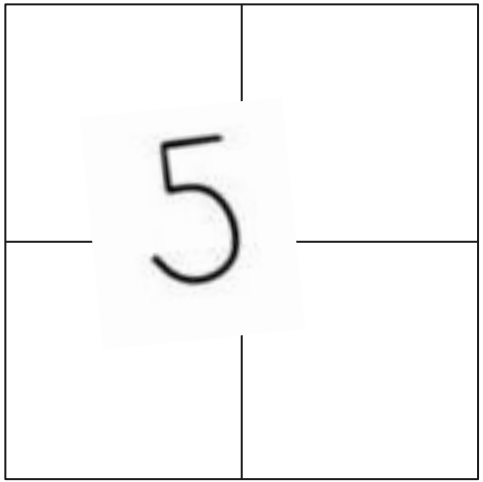
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations
- **Translated and Rotated "5"**



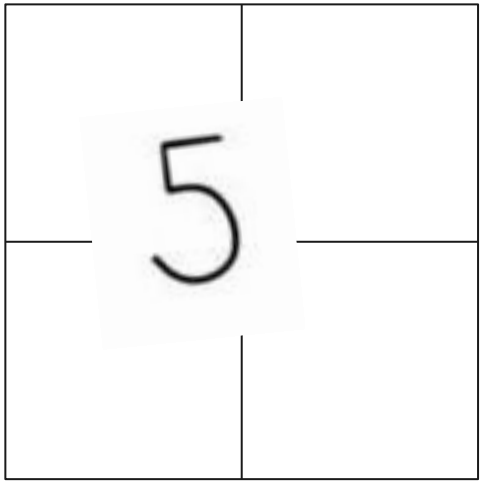
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations
- **Translated, Rotated and Scaled "5"**



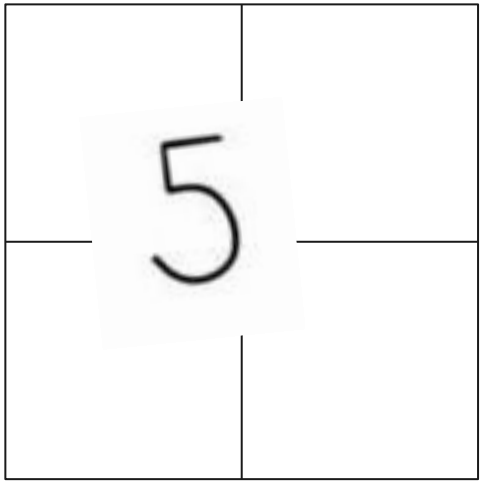
Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- **Your "5" still will have a lots of variations and each definition is unique in the vector space**



Why FF Networks don't work well for Image Data?

- We ask you to write a "5"
- A perfect "5" at the center of the coordinate
- Your "5" still will have a lots of variations and each definition is unique in the vector space
- **How many different "5"s we will need provide the model to learn a perfect model?**



Convolutional Neural Networks (CNNs)

- State of the Art for CV and some other problems
- Filters/Convolutional Kernels

Convolutional NNs

- State of the Art for CV and some other problems
- Filters/Convolutional Kernels

Examples:

- Alexnet
- VGG
- ResNet
- GoogLeNet
- ..

ref

Convolution

- Is a mathematical operation

$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau) g(t - \tau)\, d\tau.$$

Convolution

- Is a mathematical operation

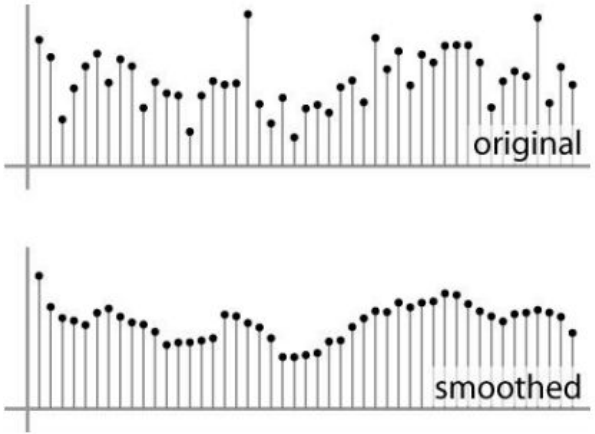
$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau) g(t - \tau)\, d\tau.$$

- Let’s learn through a practical example

Convolution

Moving average:

- Replaces each pixel with an average of all the values in its neighborhood
- Moving average in 1D: **[1, 1, 1, 1, 1] / 4**

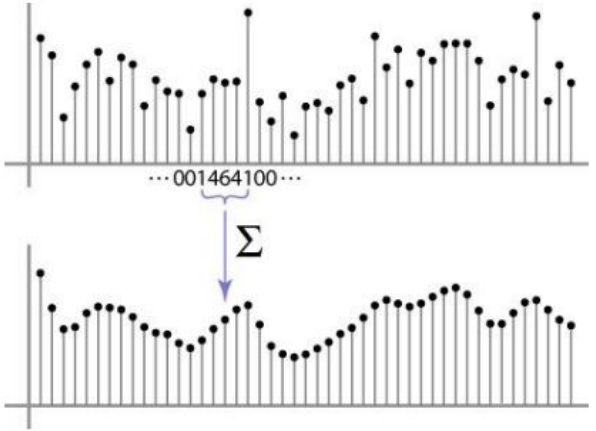


ref

Convolution

Weighted Moving average:

- Non-uniform weights **[1, 4, 6, 4, 1] / 16**



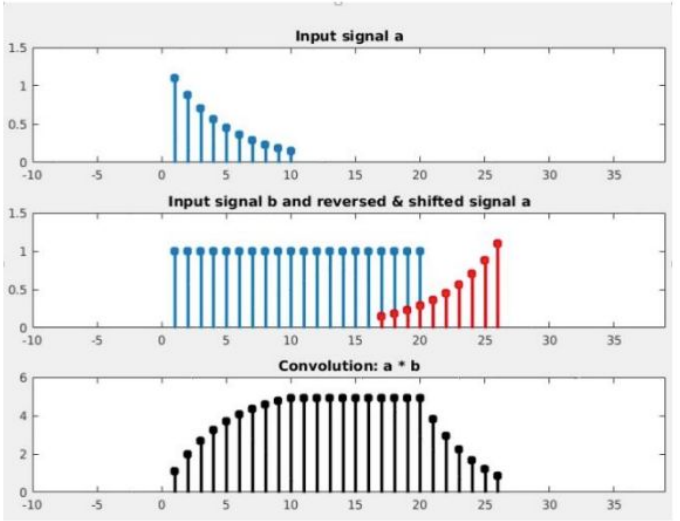
ref

Convolution

This operation is call Convolution

Example of convolution of two sequences (or “signals”)

- One of the sequences is flipped (right to left) before sliding over the other
- Notation: a * b
- Nice properties: linearity, associativity, commutativity, etc.



ref

Convolution

This operation is call Convolution

Example of convolution of two sequences (or “signals”)

- One of the sequences is flipped (right to left) before sliding over the other
- Notation: a * b
- Nice properties: linearity, associativity, commutativity, etc.

$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau) g(t - \tau)\, d\tau.$$

ref

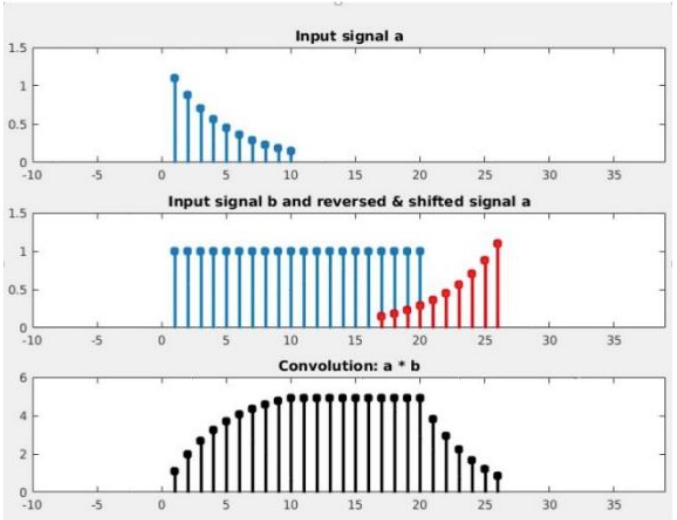


Image Filters

Moving Average in 2D

- 2D convolution operation

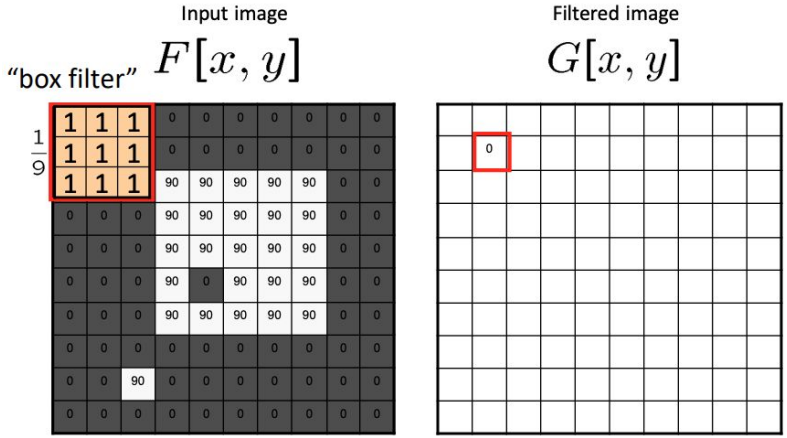


Image Filters

Moving Average in 2D

- 2D convolution operation

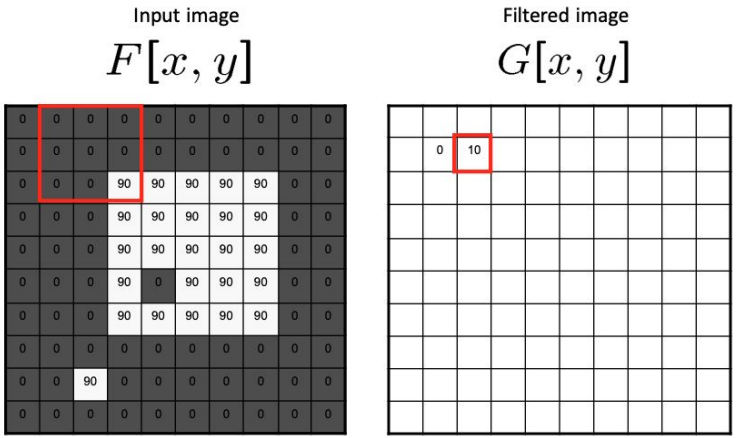


Image Filters

Moving Average in 2D

- 2D convolution operation
- Practical application (smoothing)

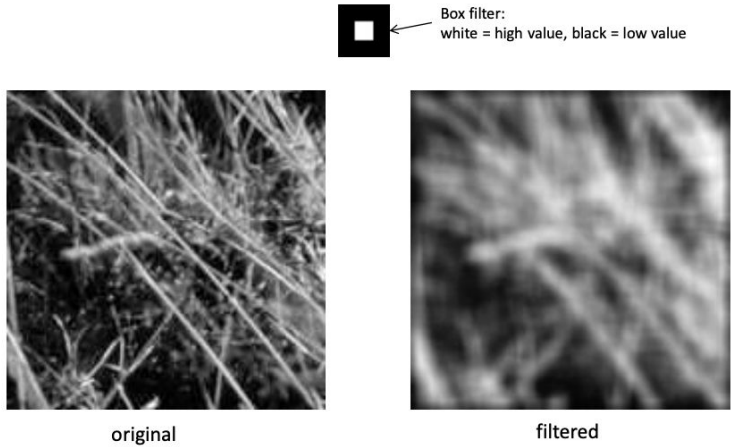


Image Filters

High Pass Filter

“If you take an image and blur it you only keep the "low frequencies". High pass makes the opposite, it only lets the "high frequencies" pass, or what most people call "**the details**". Any image can be deconstructed into these two components.”[ref]

−1	−1	−1
−1	8	−1
−1	−1	−1

Note: Only for demonstration

Image Filters

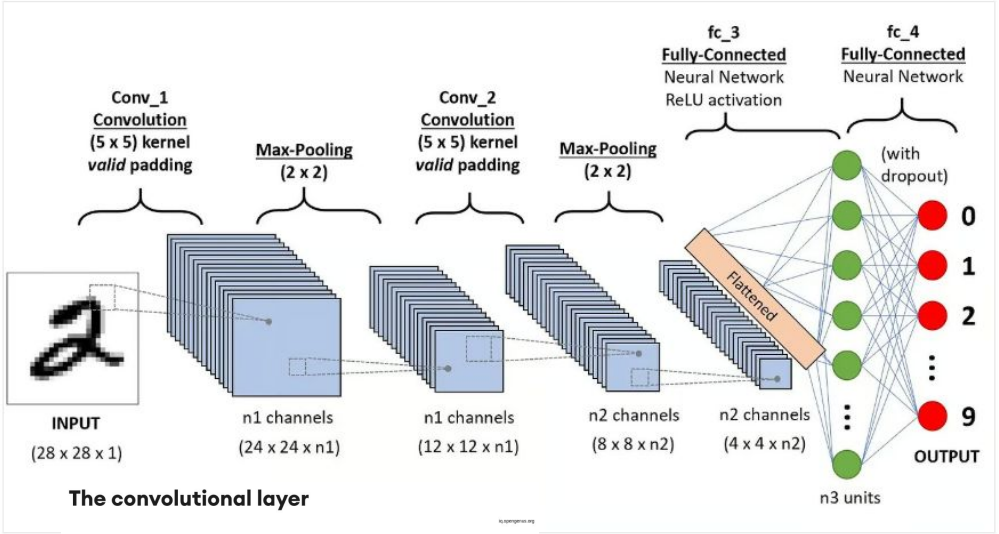
Other Filters:

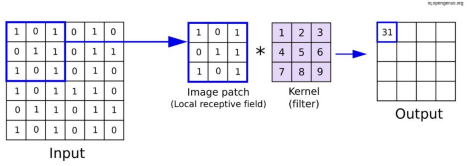
- Essentially these are convolution functions, we also call as convolution kernels

Prewitt filter $\mathbf{G_x} = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix}$ and $\mathbf{G_y} = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix}$

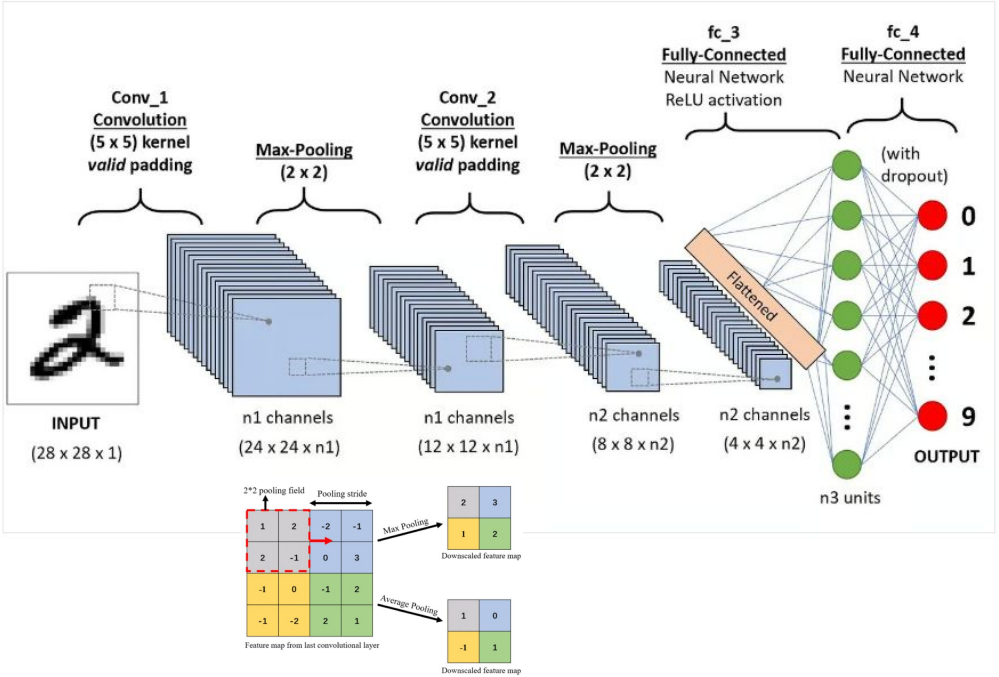
Sobel filter $\mathbf{G_x} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}$ and $\mathbf{G_y} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$

Convolutional NNs





Convolutional NNs



Coding!

Image classification (Convolutional Neural Network)