CS F425: Deep Learning

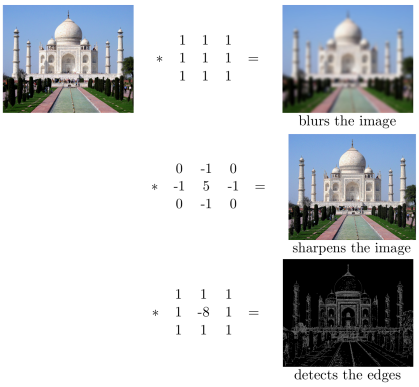
18 Convolutional Neural Network (CNN)

Dr. Kamlesh Tiwari
Assistant Professor, Department of CSIS,
BITS Pilani, Pilani Campus, Rajasthan-333031 INDIA

Mar 02, 2023 ON-CAMPUS Campus @ BITS-Pilani [Jan-May 2023]

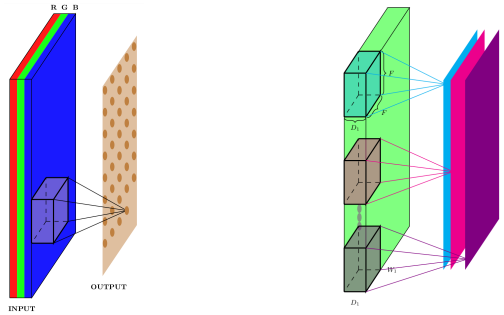
http://ktiwari.in/dl

Image is a 2D signal



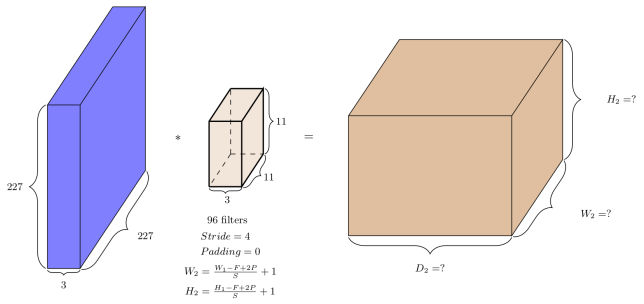
Filter in 3D

- It would refer to volume. Assume the filter extends to the depth
- Output is 2D



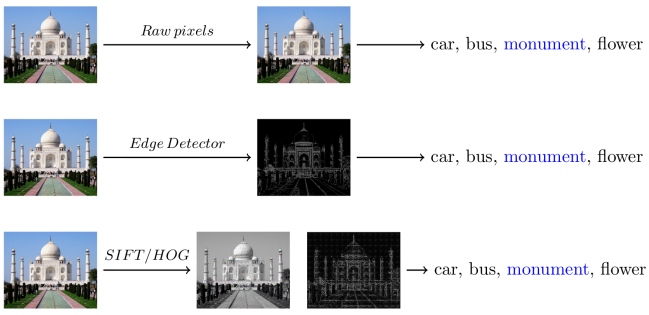
- Apply multiple filters to get multiple feature maps

Filters, Padding and Stride



$$H, W = \frac{H - F + 2P}{S} + 1, \frac{W - F + 2P}{S} + 1$$

Classification Pipeline

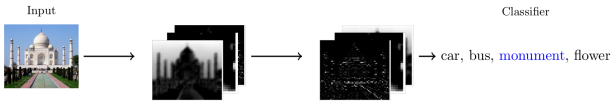


- Where is learning? (hand craft features, then learn weights for classification). One can see feature as a convolution.

Automate feature kernel discovery

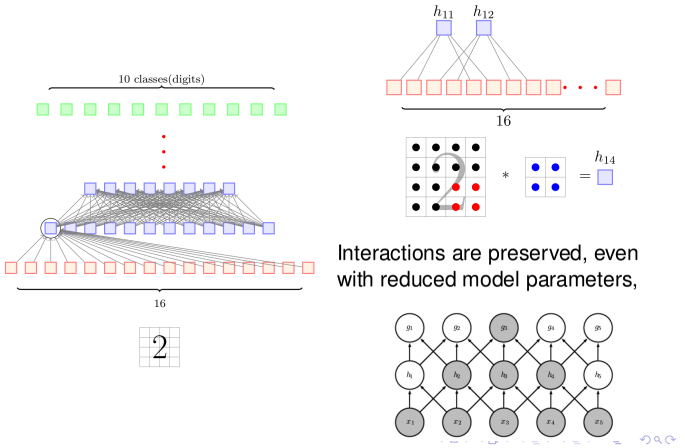
Instead of handcrafted kernels, learn filters

(Why not multiple?) learn multiple layers of kernels/filters



Treating these kernels as parameters and learning them.

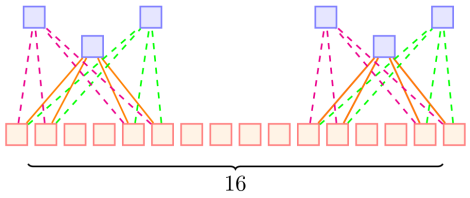
CNN has sparse connectivity with respect to NN



Interactions are preserved, even with reduced model parameters,

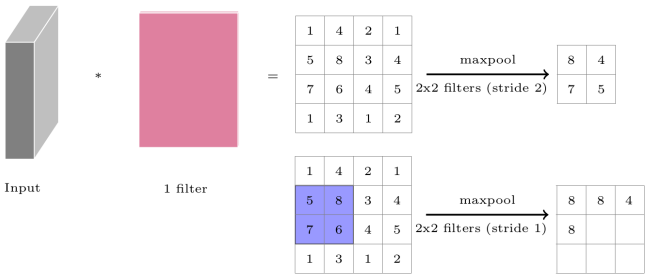
Weight sharing in CNN

One place you are extracting edge other place something else? So we do not want the kernel to be different for different portions of the image.



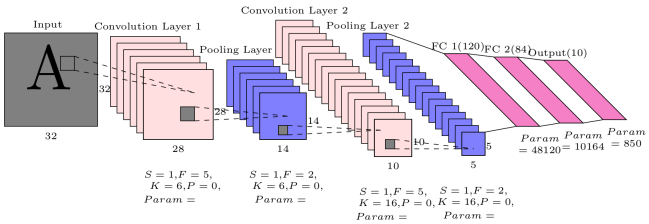
- Weight sharing in CNN makes the job of learning weights easier
- Multiple kernels help get different feature at the same level

Pooling (max, min, average)



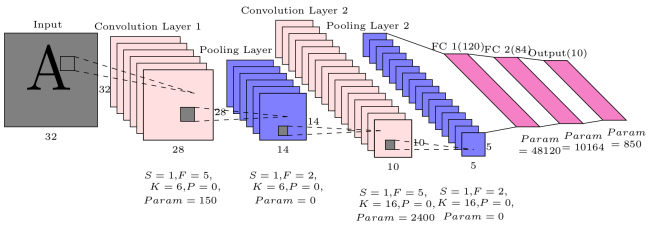
Training CNN? backpropagation!

Case-Study: LeNet-5 for handwritten character recognition [1]



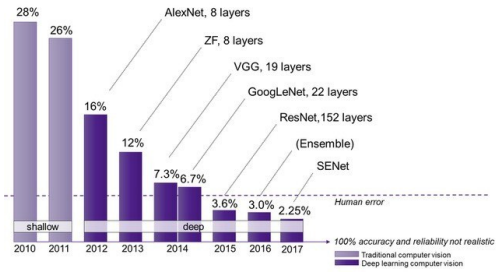
[1] Yann Lecun and Leon Bottou and Yoshua Bengio and Patrick Haffner, *Gradient-based learning applied to document recognition*, pp 2278–2324, IEEE-1998

LeNet-5 for handwritten character recognition [2]



[2] Yann Lecun and Leon Bottou and Yoshua Bengio and Patrick Haffner, *Gradient-based learning applied to document recognition*, pp 2278–2324, IEEE-1998

ImageNet ILSVRC[3]



- (2009) 22K category, 14M images
- Challenge 1000 class, 1431167 images
- HoG, LBP, SVM ...

[3] Imagenet large scale visual recognition challenge http://www.image-net.org/challenges/LSVRC/

AlexNet [4]

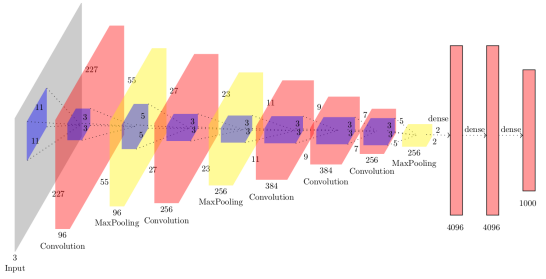


Image size 227 × 227 × 3 → [96 F=11, S=4, P=0] → [MaxPool F=3, S=2] → [256 F=5, S=1, P=0] → [MaxPool F=3, S=2] → [384 F=3, S=1, P=0] → [384 F=3, S=1, P=0] → [256 F=3, S=1, P=0] → [MaxPool F=3, S=2] → FC 4096 → FC 4096 → FC 1000

27.55 M parameters. Convolution is passes through ReLU activation.

[4] Cite: 128212, **Imagenet classification with deep convolutional neural networks**, Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey E, In: Advances in neural information processing systems pages 1097–1105, NIPS-2012

ZFNet [5]

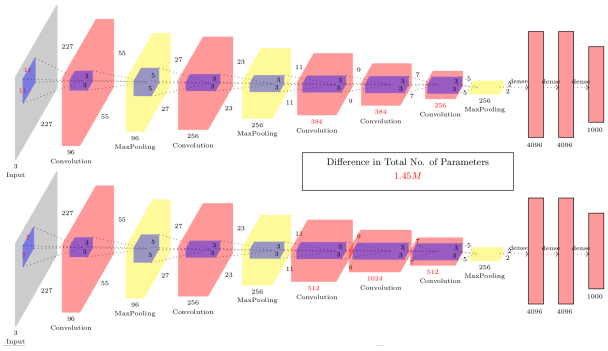


Image size 227 × 227 × 3 → [69 F=7, S=4, P=0] → [MaxPool F=3, S=2] → [256 F=5, S=1, P=0] → [MaxPool F=3, S=2] → [512 F=3, S=1, P=0] → [1024 F=3, S=1, P=0] → [512 F=3, S=1, P=0] → [MaxPool F=3, S=2] → FC 4096 → FC 4096 → FC 1000

[5] Cite 7560 **Visualizing and understanding convolutional networks**, Zeiler, Matthew D and Fergus, Rob, pages 818–833, European conference on computer vision (ECCV) Springer-2014

Thank You!

Thank you very much for your attention!