CNN

A CNN (Convolutional Neural Network) is a type of deep learning model mainly used for image processing, computer vision, and increasingly for audio, text, and spatial data. It is especially powerful at detecting patterns such as edges, textures, shapes, and objects from raw images.

LeNET

- Introduced by Yann LeCun in 1998 for handwritten digit recognition (MNIST).
- 7 layers with learnable parameters (excluding input), including 2 conv layers and 2 fully connected layers.
- Uses average pooling (subsampling) instead of max pooling.
- Total parameters ≈ 60,000 (specifically ~60k–62k depending on implementation).
- Activation: tanh (original version), not ReLU.
- Input size is 32×32 grayscale, giving the network slightly more border context for convolution.

AlexNet

- Introduced in 2012 by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton — first CNN to win ImageNet with huge performance gap.
- Total parameters ≈ 60 million, significantly larger than previous CNNs like LeNet.
- 8 learnable layers: 5 convolutional + 3 fully connected layers.
- Used ReLU activation, which dramatically sped up training compared to tanh/sigmoid.
- Introduced dropout (0.5) in fully connected layers to reduce overfitting.
- Uses overlapping max-pooling and local response normalization (LRN) in early layers.

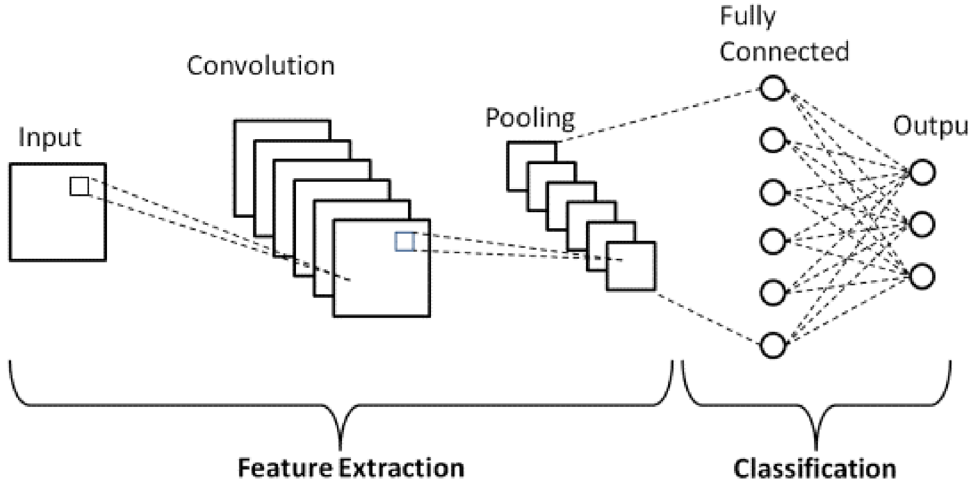
GoogleNet

- Introduced in 2014 (ImageNet winner) by Szegedy et al. using the Inception architecture.
- Very deep network — 22 layers (27 with pooling), but computationally efficient.
- Uses Inception modules, combining 1×1, 3×3, 5×5 convolutions + pooling in parallel.
- 1×1 convolutions used for dimensionality reduction, drastically lowering computation.
- Total parameters ≈ 6.8 million, far fewer than AlexNet (60M).
- Includes auxiliary classifiers to help gradient flow in deep layers.

ResNet

- Introduced in 2015 (He et al.) and won ImageNet 2015 by a large margin.
- Introduced Residual Learning using skip/shortcut connections to solve vanishing gradients.
- Extremely deep networks — ResNet-18, 34, 50, 101, 152 layers.
- Identity skip connections help gradients flow directly, enabling stable training of 100+ layers.
- Uses bottleneck blocks (1×1 → 3×3 → 1×1) in deeper versions like ResNet-50+.
- Parameters:

  ResNet-18 → ~11.7M
  ResNet-34 → ~21.8M
  ResNet-50 → ~25.6M
  ResNet-101 → ~44.5M



CNN=CONVOLUTION+POOLING

Feature Extraction

Dimension Reduction