Deep Learning Models – Modern Approaches in CNNs

Perez Jenny

Department of Computer Science
Earlham College

May 9, 2025

Overview

1. Introduction to CNNs
2. LeNet-5 (1998)
3. AlexNet (2012)
4. VGGNet (2014)
5. ResNet (2015)
6. EfficientNet (2019)

Introduction to CNNs

- Convolutional Neural Networks (CNNs) are specialized deep learning models for processing grid-like data (e.g., images).
- They use convolutional layers to automatically learn spatial hierarchies of features.
- Key breakthroughs in CNNs have driven the field of computer vision forward.

LeNet-5 (1998)

- **Developed by:** Yann LeCun
- **Purpose:** Digit recognition (e.g., MNIST dataset)
- **Architecture:** 7 layers (excluding input), convolution + pooling layers, followed by fully connected layers.
- **Impact:** Pioneered CNNs and showed early success in digit classification.

One of the earliest CNNs, developed for recognizing handwritten digits—like those in zip codes or bank checks. It used a simple structure of convolution and pooling layers followed by fully connected layers. It was a breakthrough because it showed that neural networks could actually recognize patterns in images

AlexNet (2012)

- **Developed by:** Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton
- **Key Innovations:** Deep architecture with 8 layers (5 convolutional + 3 fully connected)
- Used ReLU activations, dropout for regularization, and GPU acceleration.
- **Impact:** Won the ImageNet 2012 competition and revived interest in deep learning.

Much deeper than LeNet-5 and used ReLU activation instead of traditional ones. What made it special was how it used GPUs to speed up training and added dropout to prevent overfitting. It dominated the ImageNet competition and brought deep learning into the spotlight.

VGGNet (2014)

- **Developed by:** Visual Geometry Group (Oxford)
- **Main Idea:** Simplicity through uniform 3×3 convolutional filters
- Deep architecture with 16–19 layers
- **Pros:** Easy to replicate and extend
- **Cons:** Large number of parameters, computationally expensive
- **Impact:** Became a widely used CNN architecture for vision tasks.

VGGNet is known for its clean, simple architecture. It used only 3×3 filters and stacked them to make deeper networks—16 to 19 layers deep. This made it easier to reproduce and modify, but the tradeoff was that it needed a lot of computation and memory

ResNet (2015)

- **Developed by:** Microsoft Research
- **Key Innovation:** Skip (residual) connections
- Enables training of very deep networks (100+ layers)
- Solves the vanishing gradient problem in deep networks
- **Impact:** Won ImageNet 2015, revolutionized deep architecture design

ResNet introduced something called ‘skip connections,’ which help pass information from one layer to another without losing it. This was huge because it let researchers build networks with over 100 layers without running into training problems like vanishing gradients. It also won the ImageNet 2015 competition.

EfficientNet (2019)

- **Developed by:** Google AI
- **Key Concept:** Compound model scaling (depth, width, and resolution)
- Uses MBConv blocks for better performance
- **Results:** State-of-the-art accuracy with fewer parameters
- **Impact:** Efficiency + power for real-world deep learning applications

EfficientNet, developed by Google, focused on scaling models smartly. Instead of just making networks wider or deeper, it scales width, depth, and resolution in a balanced way. It's very accurate while using fewer resources, which makes it great for real-world applications