🧠 ResNet50: Deep Residual Learning for Image Recognition

📌 Overview

- **ResNet50** is a 50-layer deep convolutional neural network introduced by **Kaiming He et al.** in the landmark 2015 paper *"Deep Residual Learning for Image Recognition"*.
- It won the **ImageNet Large Scale Visual Recognition Challenge (ILSVRC)** in 2015, outperforming previous architectures like VGG and GoogLeNet.

🔍 Motivation

- **Problem**: As neural networks get deeper, they suffer from **vanishing gradients** and **degradation**—accuracy gets worse as layers increase.
- **Solution**: Introduce **residual connections** (or skip connections) that allow gradients to flow directly through the network.

📐 Architecture Highlights

- **Total Layers**: 50
- **Building Block**: Bottleneck residual blocks
- **Key Components**:

- **Conv1**: 7×7 convolution, 64 filters
- **MaxPooling**
- **Conv2_x to Conv5_x**: Series of residual blocks with 1×1, 3×3, and 1×1 convolutions
- **Average Pooling**
- **Fully Connected Layer**: 1000 classes (ImageNet)

🔁 Residual Block Structure

Input → 1×1 Conv → 3×3 Conv → 1×1 Conv → Add(Input) → ReLU

- The **identity shortcut** allows the input to bypass the convolutional layers and be added directly to the output.

📊 Performance

Metric	Value
Top-1 Accuracy	~75.3%
Top-5 Accuracy	~92.2%
Parameters	~25.6 million
File Size	~102 MB

🧪 Applications

- Image classification
- Object detection (used in Faster R-CNN)
- Feature extraction for transfer learning
- Medical imaging, satellite image analysis, and more

🧬 Variants

- **ResNet18/34**: Simpler versions with fewer layers
- **ResNet101/152**: Deeper versions for more complex tasks
- **ResNeXt, ResNetV2**: Enhanced versions with grouped convolutions and improved training

ResNet-50 is a deep convolutional neural network (CNN) architecture, developed by Microsoft Research in 2015, that is 50 layers deep. It is widely used for image recognition and classification, and is known for its use of residual connections, or shortcut connections, which address the vanishing gradient problem in very deep networks, allowing for more effective training. A pre-trained version of ResNet-50 is available, having been trained on the massive ImageNet dataset to recognize over 1000 object categories.

Key Components and Features

- **Depth:**

  The "50" in ResNet-50 indicates that the model has 50 weight layers, enabling it to learn complex patterns in images.

Residual Blocks:
The key innovation is the use of residual blocks, which contain "shortcut connections" or "skip connections".
Skip Connections:
These shortcuts allow the model to bypass certain layers. Instead of forcing information through every layer, the residual block learns a "residual" (a small change or leftover information), adding it to the original input.
Benefit of Residual Blocks:
This approach helps mitigate the vanishing gradient problem, allowing gradients to flow more easily during backpropagation and enabling the training of very deep networks without a significant drop in accuracy.
ImageNet Pretraining:
Many ResNet-50 models are pre-trained on the ImageNet dataset, which contains millions of images and over 1000 object categories. This makes the model excellent for various image classification tasks, such as identifying objects, animals, or medical images.

Applications

- **Object Recognition:** Classifying objects in images.

Feature Extraction: The learned features can be used for other computer vision tasks beyond classification, such as object detection and segmentation.
Custom Datasets: The ResNet architecture can be fine-tuned and used to train models on custom datasets for specific tasks.