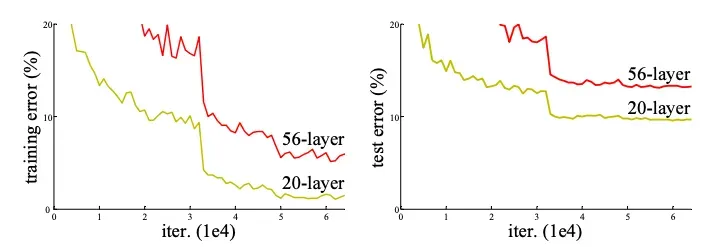
**What is ResNet-50?**

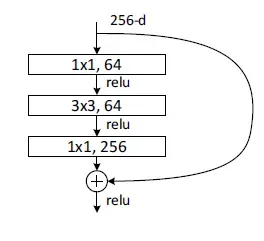
ResNet-50 is CNN architecture that belongs to the ResNet (Residual Networks) family, a series of models designed to address the challenges associated with training deep neural networks. Developed by researchers at Microsoft Research Asia, ResNet-50 is renowned for its depth and efficiency in image classification tasks. ResNet architectures come in various depths, such as ResNet-18, ResNet-32, and so forth, with ResNet-50 being a mid-sized variant.

**ResNet and Residual Blocks**

The primary problem ResNet solved was the degradation problem in deep neural networks. As networks become deeper, their accuracy saturates and then degrades rapidly. This degradation is not caused by overfitting, but rather the difficulty of optimizing the training process.

*Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer “plain” networks. The deeper network has higher training error, and thus test error.*

ResNet solved this problem using Residual Blocks that allow for the direct flow of information through the skip connections, mitigating the vanishing gradient problem.

**The residual block used in ResNet-50 is called the Bottleneck Residual Block. This block it has the following architecture:

Here's a breakdown of the components within the residual block:

**ReLU Activation**: The ReLU (Rectified Linear Unit) activation function is applied after each convolutional layer and the batch normalization layers. ReLU allows only positive values to pass through, introducing non-linearity into the network, which is essential for the network to learn complex patterns in the data.

**Bottleneck Convolution Layers**: the block consists of three convolutional layers with batch normalization and ReLU activation after each.:

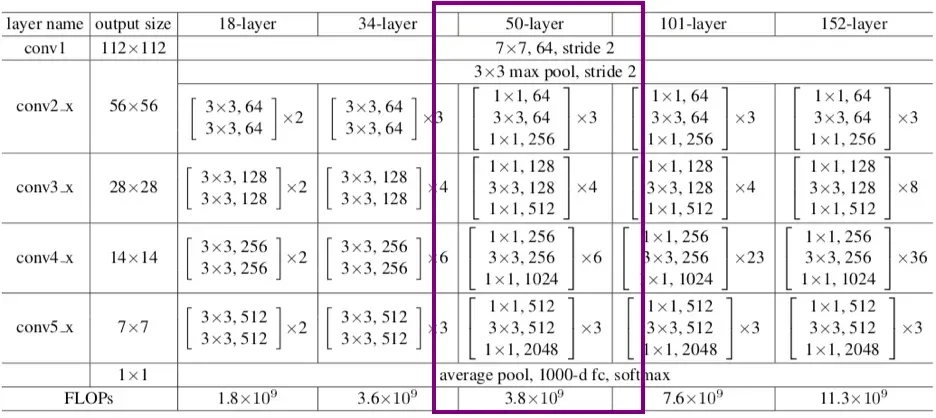
* The first convolutional layer likely uses a filter size of 1x1 and reduces the number of channels in the input data. This dimensionality reduction helps to compress the data and improve computational efficiency without sacrificing too much information.
* The second convolutional layer might use a filter size of 3x3 to extract spatial features from the data.
* The third convolutional layer again uses a filter size of 1x1 to restore the original number of channels before the output is added to the shortcut connection.

**Skip Connection**: As in a standard residual block, the key element is the shortcut connection. It allows the unaltered input to be added directly to the output of the convolutional layers. This bypass connection ensures that essential information from earlier layers is preserved and propagated through the network, even if the convolutional layers struggle to learn additional features in that specific block.

By combining convolutional layers for feature extraction with shortcut connections that preserve information flow, and introducing a bottleneck layer to reduce dimensionality, bottleneck residual blocks enable ResNet-50 to effectively address the vanishing gradient problem, train deeper networks, and achieve high accuracy in image classification tasks.

**Stacking the Blocks: Building ResNet-50**

ResNet-50 incorporates 50 bottleneck residual blocks, arranged in a stacked manner. The early layers of the network feature conventional convolutional and pooling layers to preprocess the image before it undergoes further processing by the residual blocks. Ultimately, fully connected layers positioned at the pinnacle of the structure utilize the refined data to categorize the image with precision.

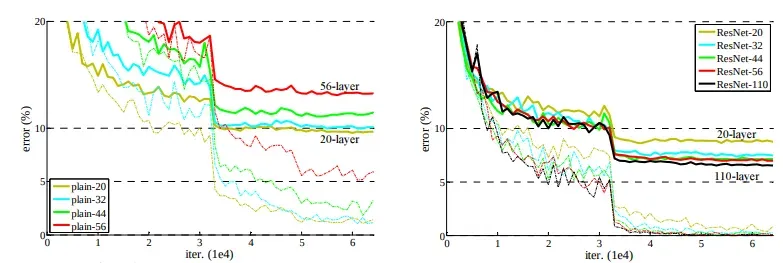
**

*ResNet Architectures. Building blocks are shown in brackets with the number of blocks stacked. ResNet-50 architecture is highlighted.*

Through the strategic integration of bottleneck residual blocks and shortcut connections, ResNet-50 adeptly mitigates the vanishing gradient issue, enabling the creation of more profound and potent models for image classification. This innovative architectural approach has opened the door to notable strides in the field of computer vision.

**ResNet Performance**

In this section, we are going to show the ResNet-20, -32, -44, -56, and -110 performance compared to plain neural networks.

**

*Comparison between plain neural networks and ResNets.* [*Source*](https://arxiv.org/pdf/1512.03385.pdf?ref=blog.roboflow.com)

The dashed lines denote training error, and bold lines denote testing error on CIFAR-10. The left chart shows the training and testing errors using plain networks. The error of plain-110 is higher than 60% and is not displayed. The right chart shows the training and testing errors using ResNets.

In essence, the charts demonstrate the advantage of using skip connections in neural networks. By mitigating the vanishing gradient problem, skip connections allow for deeper networks that can achieve higher accuracy in image classification tasks.

**Cite this Post**

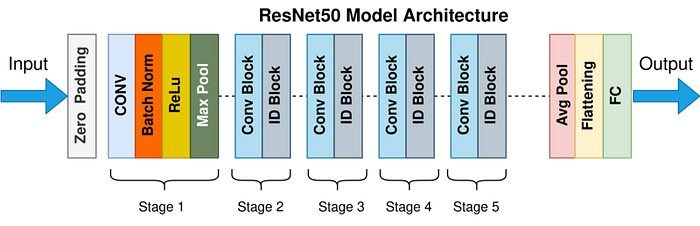
[*Petru Potrimba*](https://blog.roboflow.com/author/potrimba/)*. (Mar 13, 2024). What is ResNet-50?. Roboflow Blog: https://blog.roboflow.com/what-is-resnet-50/*

**OTHER**

ResNet50 is a powerful image classification model that can be trained on large datasets and achieve state-of-the-art results. One of its key innovations is the use of residual connections, which allow the network to learn a set of residual functions that map the input to the desired output. These residual connections enable the network to learn much deeper architectures than was previously possible, without suffering from the problem of vanishing gradients.

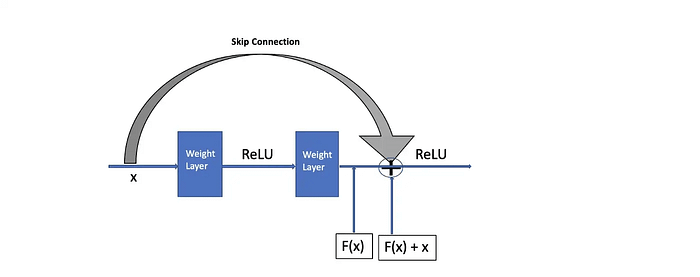
The architecture of ResNet50 is divided into four main parts: **the convolutional layers, the identity block, the convolutional block, and the fully connected layers**. The convolutional layers are responsible for extracting features from the input image, while the identity block and convolutional block are responsible for processing and transforming these features. Finally, the fully connected layers are used to make the final classification.

1. The convolutional layers in ResNet50 consist of several convolutional layers followed by batch normalization and ReLU activation. These layers are responsible for extracting features from the input image, such as edges, textures, and shapes. The convolutional layers are followed by max pooling layers, which reduce the spatial dimensions of the feature maps while preserving the most important features.
2. The identity block and convolutional block are the key building blocks of ResNet50. The identity block is a simple block that passes the input through a series of convolutional layers and adds the input back to the output. This allows the network to learn residual functions that map the input to the desired output. The convolutional block is similar to the identity block, but with the addition of a 1x1 convolutional layer that is used to reduce the number of filters before the 3x3 convolutional layer.
3. The final part of ResNet50 is the fully connected layers. These layers are responsible for making the final classification. The output of the final fully connected layer is fed into a softmax activation function to produce the final class probabilities.



ResNet50 has been trained on large datasets and achieves state-of-the-art results on several benchmarks. It has been trained on the ImageNet dataset, which contains over 14 million images and 1000 classes. On this dataset, ResNet50 achieved an error rate of 22.85% which is on par with human performance, which is an error rate of 5.1%.

**How it solved the problem of vanishing gradients:**



Skip connections, also known as residual connections, are a key feature of the ResNet50 architecture. They are used to allow the network to learn deeper architectures without suffering from the problem of vanishing gradients.

Vanishing gradients is a problem that occurs when training deep neural networks, where the gradients of the parameters in the deeper layers become very small, making it difficult for those layers to learn and improve. This problem becomes more pronounced as the network becomes deeper.

Skip connections address this problem by allowing the information to flow directly from the input to the output of the network, bypassing one or more layers. This allows the network to learn residual functions that map the input to the desired output, rather than having to learn the entire mapping from scratch.

In ResNet50, skip connections are used in the identity block and convolutional block. The identity block passes the input through a series of convolutional layers and adds the input back to the output, while the convolutional block uses a 1x1 convolutional layer to reduce the number of filters before the 3x3 convolutional layer and then adds the input back to the output.

The use of skip connections in ResNet50 allows the network to learn deeper architectures while still being able to train effectively and prevent vanishing gradients.

**Summary:**

In summary, ResNet50 is a cutting-edge deep convolutional neural network architecture that was developed by Microsoft Research in 2015. It is a variant of the popular ResNet architecture and comprises of 50 layers that enable it to learn much deeper architectures than previously possible without encountering the problem of vanishing gradients. The architecture of ResNet50 is divided into four main parts: the convolutional layers, the identity block, the convolutional block, and the fully connected layers. The convolutional layers are responsible for extracting features from the input image, the identity block and convolutional block process and transform these features, and the fully connected layers make the final classification. ResNet50 has been trained on the large ImageNet dataset, achieving an error rate on par with human performance, making it a powerful model for various image classification tasks such as object detection, facial recognition and medical image analysis. Additionally, it has also been used as a feature extractor for other tasks, such as object detection and semantic segmentation.