AlexNet

Introduction

- AlexNet revolutionized deep learning
- Developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton
- Won the ImageNet 2012 competition
- First CNN to show significant improvement in image classification

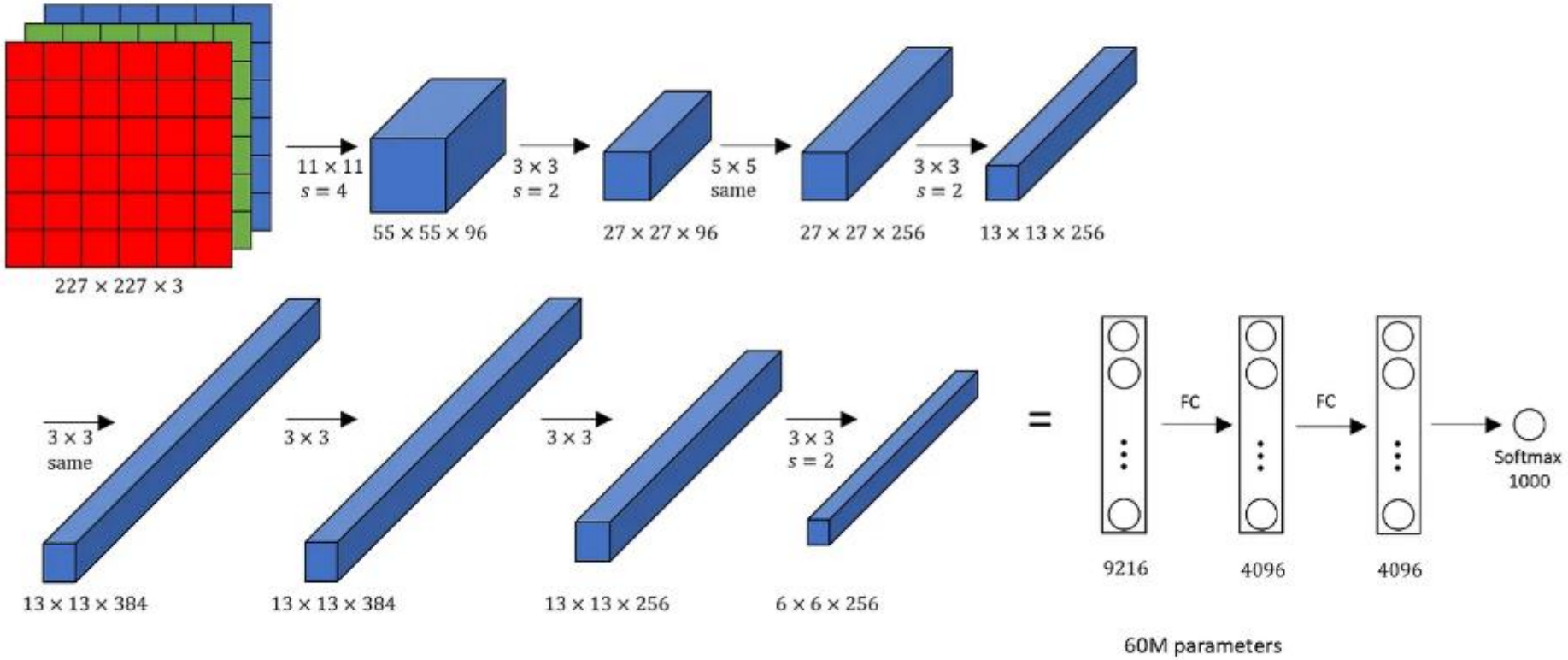
Key Features of AlexNet

- Deep CNN with 8 layers (5 convolutional, 3 fully connected)
- Uses ReLU activation function
- Employs GPU acceleration for training
- Implements Dropout to reduce overfitting
- Uses Data Augmentation for improved generalization

AlexNet Architecture Overview

- **Input Layer:** 227x227 RGB image
- **Convolutional Layers:** Feature extraction
- **Max-Pooling Layers:** Reducing spatial dimensions
- **Normalization Layers:** Enhancing generalization
- **Fully Connected Layers:** Classification
- **Softmax Layer:** Final output probabilities

AlexNet Architecture



Key features

- This was the first architecture that used GPU to boost the training performance. AlexNet consists of 5 convolution layers, 3 max-pooling layers, 2 Normalized layers, 2 fully connected layers and 1 SoftMax layer.
- Each convolution layer consists of a convolution filter and a non-linear activation function called “ReLU”.
- The pooling layers are used to perform the max-pooling function and the input size is fixed due to the presence of fully connected layers.
- The input size is mentioned at most of the places as 224x224x3 but due to some padding which happens it works out to be 227x227x3.
- Above all this AlexNet has over 60 million parameters.

AlexNet Summary

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	227x227x3	-	-	-
1	Convolution	96	55 x 55 x 96	11x11	4	relu
	Max Pooling	96	27 x 27 x 96	3x3	2	relu
2	Convolution	256	27 x 27 x 256	5x5	1	relu
	Max Pooling	256	13 x 13 x 256	3x3	2	relu
3	Convolution	384	13 x 13 x 384	3x3	1	relu
4	Convolution	384	13 x 13 x 384	3x3	1	relu
5	Convolution	256	13 x 13 x 256	3x3	1	relu
	Max Pooling	256	6 x 6 x 256	3x3	2	relu
6	FC	-	9216	-	-	relu
7	FC	-	4096	-	-	relu
8	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

Innovations in AlexNet

- **ReLU Activation Function:** Faster training
- **Overlapping Max-Pooling:** Reducing spatial size while retaining features
- **Local Response Normalization (LRN):** Improves generalization
- **GPU Parallelization:** Training deep networks efficiently

AlexNet Performance

- Achieved 16.4% top-5 error rate in ImageNet 2012
- Outperformed previous models significantly
- Proved CNNs can surpass traditional methods in computer vision

Impact of AlexNet

- Inspired deeper architectures like VGG, GoogLeNet, and ResNet
- Demonstrated the power of deep learning for real-world applications
- Sparked advancements in GPU-accelerated deep learning

Conclusion

- AlexNet set the foundation for modern deep learning
- Introduced key concepts still used today
- A milestone in AI and computer vision

LeNet and AlexNet: A Comparison of CNN Architectures

LeNet and AlexNet are both convolutional neural network (CNN) architectures, but they were developed in different eras and for different levels of complexity in image recognition tasks.

Comparison Table: LeNet vs. AlexNet

Feature	LeNet-5	AlexNet
Year	1989	2012
Developed By	Yann LeCun	Alex Krizhevsky & Team
Dataset	MNIST (Handwritten Digits)	ImageNet (1.2M Images)
Number of Layers	7	8
Activation Function	Tanh, Sigmoid	ReLU
Pooling Type	Average Pooling	Max Pooling
Regularization	None	Dropout
Computational Complexity	Low	High
GPU Usage	No	Yes

1. LeNet as a Pre-trained Model

- **Availability**: Available in deep learning libraries like TensorFlow and PyTorch
- **Pre-trained on**: Typically trained on **MNIST** (handwritten digits)
- **Use Case**: Recognizing simple patterns, digits, and small-scale classification tasks
- **Transfer Learning**: Not widely used for transfer learning due to its small size and limited dataset

2. AlexNet as a Pre-trained Model

- **Pre-trained on**: **ImageNet** (1.2 million images across 1000 classes)
- **Use Case**: Large-scale image classification, object detection, and feature extraction
- **Transfer Learning**: Widely used for transfer learning by fine-tuning the later layers on new datasets

- **LeNet** is a lightweight CNN model, mainly for small datasets.
- **AlexNet** is a robust pre-trained model, commonly used for feature extraction and transfer learning.