

## Lenet -5 And Alexnet Assignments

**Q1.** Explain the architecture of LeNet-5 and its significance in the field of deep learning.

**Ans:**

LeNet-5, developed by Yann LeCun in 1998, was one of the earliest convolutional neural networks (CNNs) designed for handwritten digit classification (MNIST dataset). It has a simple architecture consisting of convolutional and fully connected layers, and its success demonstrated the potential of CNNs in tasks like image recognition.

Significance in Deep Learning:

**Pioneering Work:** LeNet-5 laid the foundation for modern deep learning models by introducing the core concepts of convolution, pooling, and hierarchical feature extraction. **End-to-End Learning:** LeNet-5 demonstrated the effectiveness of end-to-end learning for image classification, paving the way for more complex models in computer vision. **Influence on Modern CNNs:** Many modern CNN architectures, including AlexNet, ResNet, and VGG, are built on principles established by LeNet-5.

**Q2.** Describe the key components of LeNet-5 and their roles in the network.

**Ans:**

LeNet-5 has seven layers (excluding the input layer) and is composed of both convolutional and fully connected layers:

Input Layer (32x32 grayscale image):

The input to LeNet-5 is a 32x32 pixel grayscale image. MNIST images are typically 28x28, so padding is applied to make them 32x32. C1 – First Convolutional Layer:

Applies 6 filters (kernels) of size 5x5, resulting in six feature maps of size 28x28. Role: Extract low-level features like edges and textures. S2 – First Subsampling (Pooling) Layer:

Performs average pooling (2x2) on each of the feature maps from C1, reducing their size to 14x14. Role: Reduce spatial dimensions and retain important information. C3 – Second Convolutional Layer:

Applies 16 filters of size 5x5, resulting in 16 feature maps of size 10x10. Role: Extract higher-level features by building on the lower-level ones from C1. S4 – Second Subsampling (Pooling) Layer:

Performs average pooling (2x2), reducing the size of the feature maps to 5x5. Role: Further reduce dimensionality while preserving key features. C5 – Fully Connected Convolutional Layer:

Applies 120 filters of size 5x5, resulting in 120 feature maps of size 1x1 (essentially a fully connected layer). Role: Connects the convolutional layers to the fully connected part of the network, transforming the extracted features for final classification. F6 – Fully Connected Layer:

84 neurons, fully connected to the output from C5. Role: Perform further processing before the final output. Output Layer:

10 neurons corresponding to the 10 digit classes (0-9) in the MNIST dataset. Role: Provide the classification result.

**Q3.** Discuss the limitations of LeNet-5 and how subsequent architectures like AlexNet addressed these limitations?

**Ans:**

Limitations of LeNet-5:

Small Scale: LeNet-5 was designed for small images (32x32 grayscale), and it struggles with large-scale, high-resolution images. Shallow Network: LeNet-5 has a limited number of layers, making it unsuitable for learning complex features from large datasets. Limited Dataset: LeNet-5 was trained on MNIST, a relatively simple dataset. It was not designed for more complex datasets like ImageNet. No GPU Utilization: LeNet-5 was implemented when hardware, such as GPUs, wasn't as advanced, limiting its computational efficiency. How AlexNet Addressed These Limitations:

Large Scale: AlexNet was designed to work with large RGB images (227x227) and was trained on the ImageNet dataset, which contains millions of high-resolution images. Deeper Architecture: AlexNet introduced a much deeper architecture with 5 convolutional layers and 3 fully connected layers, allowing the network to learn more complex and hierarchical features. ReLU Activation: AlexNet used the ReLU activation function, which speeds up training and helps mitigate the vanishing gradient problem, compared to the sigmoid/tanh functions used in LeNet-5. GPU Utilization: AlexNet was one of the first deep learning models to utilize GPUs for training, making it feasible to train deep networks on large datasets efficiently.

**Q4.** Explain the architecture of AlexNet and its contributions to the advancement of deep learning.

**Ans:**

AlexNet, developed by Alex Krizhevsky et al. in 2012, was a landmark model in the history of deep learning. It achieved state-of-the-art performance on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and marked a significant leap in the use of deep learning for large-scale image classification.

Architecture of AlexNet:

Input Layer: 227x227x3 RGB image. Conv1: First convolutional layer with 96 filters of size 11x11, stride 4, followed by max pooling. Conv2: Second convolutional layer with 256 filters

of size 5x5, followed by max pooling. Conv3, Conv4, Conv5: Three more convolutional layers with 384, 384, and 256 filters of size 3x3, respectively. Fully Connected Layers: Three fully connected layers with 4096, 4096, and 1000 neurons, respectively. Output Layer: 1000 classes corresponding to ImageNet categories. Contributions of AlexNet:

Breakthrough in ImageNet: AlexNet's victory in the ILSVRC competition showcased the power of deep learning for image classification on large-scale datasets. ReLU and GPU Acceleration: AlexNet popularized the use of the ReLU activation function, which significantly sped up training. It also demonstrated how GPU acceleration could be used to train deep networks on large datasets efficiently. Dropout: AlexNet introduced dropout regularization to prevent overfitting, a technique that became widely adopted in subsequent architectures. Data Augmentation: AlexNet applied data augmentation techniques such as random cropping and mirroring to improve generalization and reduce overfitting.

**Q5.** Compare and contrast the architectures of LeNet-5 and AlexNet. Discuss their similarities, differences, and respective contributions to the field of deep learning.

**Ans:**

Similarities:

Convolutional Architecture: Both LeNet-5 and AlexNet are based on the convolutional neural network (CNN) architecture, where convolutional layers are used to extract features from images. Pooling Layers: Both networks use pooling layers to reduce the spatial dimensions of feature maps, reducing computational cost and retaining important features. Fully Connected Layers: Both architectures have fully connected layers toward the end for classification purposes. Differences:

Scale: LeNet-5: Designed for small grayscale images (32x32) and simple tasks like handwritten digit recognition. AlexNet: Designed for large RGB images (227x227) and large-scale tasks like ImageNet classification with millions of high-resolution images. Depth: LeNet-5: Shallow network with 2 convolutional layers and 2 pooling layers. AlexNet: Deeper network with 5 convolutional layers and 3 fully connected layers, allowing it to learn more complex features. Activation Function: LeNet-5: Uses the sigmoid or tanh activation function, which can lead to the vanishing gradient problem. AlexNet: Uses the ReLU activation function, which mitigates the vanishing gradient problem and speeds up training. Dataset: LeNet-5: Primarily designed for the MNIST dataset (handwritten digits). AlexNet: Designed for the much larger and more complex ImageNet dataset. Contributions:

LeNet-5: Introduced key CNN concepts like convolution, pooling, and hierarchical feature extraction. It was a pioneering model but limited to small-scale tasks. AlexNet: Revolutionized computer vision by showing that deep networks, when trained with large datasets and computational resources, could achieve breakthrough performance. AlexNet's use of GPUs, ReLU, and dropout set new standards in deep learning.