HW to Chapter 14 “Convolutional Neural Networks”

***Non-programming Assignment***

1. What is pooling layer and how it works?
2. What are three major types of layers in the convolutional neural network?
3. What is the architecture of a convolutional network?

1. **What is a Pooling Layer and How Does it Work?**

A **pooling layer** is used in a Convolutional Neural Network (CNN) to reduce the spatial dimensions (width and height) of feature maps while retaining important information. This helps to reduce computational complexity, prevent overfitting, and make the model more robust to small variations in the input.

***How Pooling Works***

The pooling operation moves a filter (also called a kernel or mask) over the feature map and applies an operation within the filter’s region. Two common types of pooling are:

* **Max Pooling**: Selects the maximum value in the covered region. This helps preserve the most significant features while reducing dimensions.
* **Average Pooling**: Computes the average of all values in the covered region, leading to a smoother representation.

***Example of Max Pooling (2x2 filter, stride=2)***

If we apply max pooling with a 2×2 filter and a stride of 2 to the following 4×4 matrix:

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 3 | 2 | 1 |
| 4 | 6 | 5 | 3 |
| 7 | 8 | 9 | 6 |
| 4 | 7 | 3 | 2 |

The output would be:

|  |  |
| --- | --- |
| 6 | 5 |
| 8 | 9 |

Since the filter moves across the matrix and selects the highest value in each 2×2 region.

1. **What Are the Three Major Types of Layers in a CNN?**

A CNN has three main types of layers that work together:

**Convolutional (CONV) Layer**

* The core building block of a CNN.
* Applies **filters** (small matrices) to detect patterns like edges, textures, and shapes.
* Uses an activation function (like ReLU) to introduce non-linearity.

**Pooling (POOL) Layer**

* Reduces the spatial dimensions while preserving important features.
* Helps make the model more invariant to translations and reduces computations.
* Two main types: **Max Pooling** (selects max value) and **Average Pooling** (computes average).

**Fully Connected (FC) Layer**

* Connects every neuron from the previous layer to the next.
* Flattens the feature maps from the convolutional and pooling layers.
* Typically used for classification at the end of the network.

1. **What is the Architecture of a Convolutional Neural Network?**

A CNN follows a structured architecture designed for feature extraction and classification. A general CNN architecture consists of the following stages:

**Input Layer**

* + Takes an image (e.g., 32×32×3 for RGB).
  + Normalization (optional) can be applied to standardize pixel values.

**Feature Extraction (Convolutional & Pooling Layers)**

* + **Convolutional layers** apply filters to detect local patterns.
  + **Pooling layers** down sample the feature maps to reduce spatial dimensions.

**Flattening & Fully Connected Layers**

* + After feature extraction, the output is flattened into a 1D vector.
  + One or more **fully connected (FC) layers** are used to make final predictions.

**Output Layer**

* + Applies a SoftMax (for multi-class classification) or sigmoid (for binary classification).
  + Outputs the probability distribution for different classes.

***Example CNN Architecture (LeNet-5)***

One of the earliest CNNs, **LeNet-5**, follows this architecture:

**Input**: 32×32 grayscale image

* + **Conv1**: 6 filters of size 5×5 → Output: 28×28×6
  + **Pool1**: 2×2 Max Pooling → Output: 14×14×6
  + **Conv2**: 16 filters of size 5×5 → Output: 10×10×16
  + **Pool2**: 2×2 Max Pooling → Output: 5×5×16
  + **Flatten** → Output: 400 neurons
  + **Fully Connected Layer (FC1)** → 120 neurons
  + **Fully Connected Layer (FC2)** → 84 neurons
  + **Output Layer** → 10 classes (digits 0-9)

This design is used in modern deep learning models with deeper architectures like **AlexNet, VGG and ResNet.**