1.

Basic Components of a Digital Image and Representation in a Computer

A digital image is composed of pixels, which are the smallest unit of an image. Each pixel has a numerical value representing its intensity or color.

- **Representation in a computer**:
    - An image is stored as a matrix of pixel values.
    - For grayscale images, the values represent brightness (e.g., 0 for black, 255 for white in an 8-bit image).
    - For color images, each pixel has three values, representing the intensity of red, green, and blue (RGB) channels.

Difference Between Grayscale and Color Images

- **Grayscale Images**:
    - Contain only intensity information (single channel).
    - Each pixel is represented by one value (e.g., 0 to 255 in 8-bit representation).
    - Simpler and smaller in size.
- **Color Images**:
    - Contain information about colors (three channels: RGB or others like YUV).
    - Each pixel is represented by three values, one for each channel.
    - More complex and larger in size

2

 Definition: CNNs are a type of deep learning neural network specifically designed to process structured grid-like data such as images. They leverage convolutional layers to automatically and adaptively learn spatial hierarchies of features from data.

 Role in Image Processing:

- Extracts features such as edges, shapes, textures, and complex patterns.
- Useful for tasks like image classification, object detection, segmentation, and enhancement

key Advantages of CNNs Over Traditional Neural Networks

- **Local Connectivity**: CNNs use filters to focus on small local regions of an image, which helps in capturing spatial features effectively.
- **Parameter Sharing**: Filters are applied across the entire image, reducing the number of parameters compared to traditional neural networks.
- **Translation Invariance**: The learned features (e.g., edges or patterns) are invariant to their location in the input image.

3.

 Purpose: Convolutional layers are designed to extract spatial features (e.g., edges, textures) by applying filters (kernels) to the input image.

 Filters in Convolution Operation:

- Filters (or kernels) are small matrices (e.g., 3×3 or 5×5) applied to the input image.
- They detect features like edges, corners, and patterns by computing dot products between the filter and the input region.

 Padding:

- Adds extra rows/columns of zeros around the input matrix.
- Purpose: Preserve spatial dimensions and prevent loss of information at edges.

 Strides:

- Refers to the step size by which the filter moves across the input.
- Larger strides reduce the output size, while smaller strides retain more detail.

4.

 Purpose:

- Reduce the spatial dimensions of the feature maps, decreasing computational load.
- Help in extracting dominant features, making the network more robust to spatial variations.

 Comparison of Max Pooling and Average Pooling:

- **Max Pooling**:
    - Outputs the maximum value in each patch of the feature map.
    - Preserves dominant features, making it useful for tasks requiring sharp distinctions.
- **Average Pooling**:
    - Outputs the average of values in each patch.
    - Retains more contextual information but might dilute prominent features.