

Definition of a Convolutional Neural Network (CNN)

A **Convolutional Neural Network (CNN)** is a type of deep learning model specifically designed to process structured grid-like data such as images or videos. CNNs are highly effective in identifying spatial patterns through convolution operations, which detect local features like edges, textures, and shapes.

Core Components

1. Convolutional Layer:

- Applies convolution operations to extract feature maps.
- Uses kernels (filters) to learn local patterns.

$$(I * K)(x, y) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} K(i, j) \cdot I(x + i, y + j)$$

2. Pooling Layer:

- Reduces the spatial dimensions of the feature maps while retaining the most significant information (e.g., max pooling).

$$P(x, y) = \max_{i,j \in F} (I(x + i, y + j))$$

3. Fully Connected Layer:

- Flattens feature maps and connects every input to every output neuron.
- Performs classification tasks using outputs from convolution and pooling layers.

4. Activation Function:

- a. Introduces non-linearity (e.g., ReLU or Softmax).

Our paper, "*Face Recognition with ESP-32 CAM using Neural Networks*", explains CNNs as the core architecture for face recognition. The system processes visual inputs, extracts features via convolutional kernels, and matches features to identities.

Key characteristics:

- Transfer learning with pretrained models.
- High efficiency on resource-constrained devices like ESP-32 CAM.
- Robust to varying lighting conditions and partial obstructions.