A Convolutional Neural Network (CNN) is a type of deep learning algorithm, widely used for image recognition, classification, and other visual data tasks. CNNs are inspired by the structure of the human visual cortex and are designed to automatically and adaptively learn spatial hierarchies of features from images.

The core idea behind CNNs lies in their convolutional layers, where small filters (or kernels) slide over input data (such as images) to extract important features like edges, textures, and shapes. These layers apply convolution operations, transforming the input data into feature maps. The algorithm uses pooling layers to downsample the feature maps, reducing their spatial dimensions and making the network computationally efficient.

CNNs consist of multiple layers that progressively capture higher-level features. Early layers detect simple patterns, while deeper layers identify more complex structures like objects or facial features. This hierarchical learning enables CNNs to achieve impressive results in various applications, including object detection, medical image analysis, and even video processing.

Training a CNN typically requires large datasets and is powered by backpropagation to minimize error and improve accuracy. As CNNs evolve with more advanced architectures, such as ResNet or EfficientNet, they have revolutionized fields like computer vision, making them indispensable in modern AI applications.