Introduction to Convolutional Neural Network

Sammie Omranian

June 2025

UWM | College of Engineering & Applied Science

What is a CNN?

- A **Convolutional Neural Network (CNN)** is a type of neural network that processes data with a grid-like structure, such as images.

- CNNs help recognize patterns in images like edges, shapes, and objects.

Why Use CNNs?

- Traditional methods for image recognition required manual feature extraction.
- CNNs learn **features automatically** from the data.

Applications:

- Face recognition
- Self-driving cars
- Medical imaging
- Image search engines

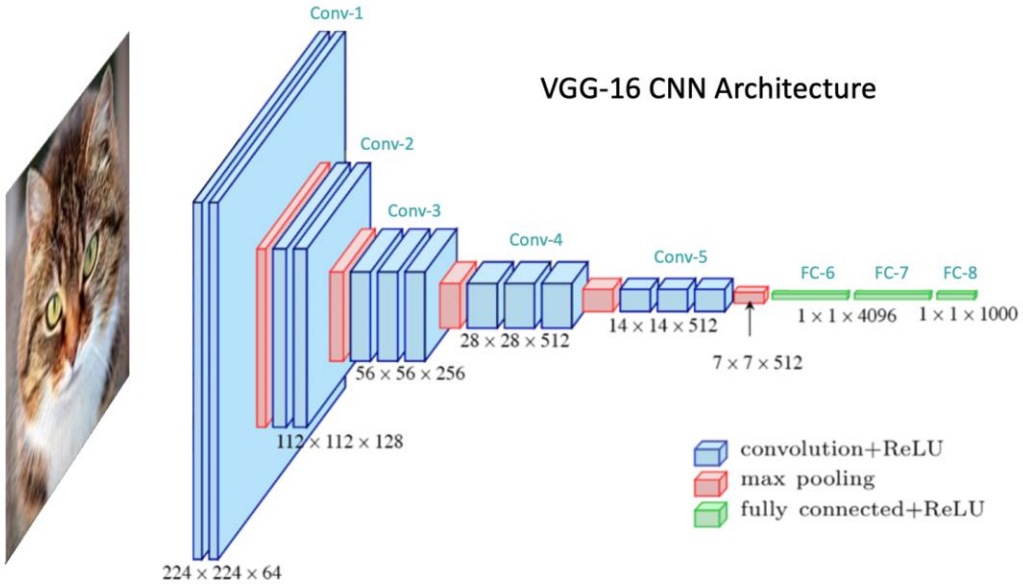


Image credit

How Does a CNN Work?

- **Input Image:** The image is broken down into numbers (pixels).
- **Convolution:** Small filters scan the image for patterns (e.g., edges, textures).
- **Pooling:** Simplifies the image by keeping only the important features.
- **Fully Connected Layers:** Combines all the features to classify the image.

How Does a CNN Work?
Convolution

- **What happens in convolution?**
  - Filters (small grids) slide over the image to detect patterns like edges and lines.
  - The result is a new "**feature map**" that highlights where patterns are found.

Like looking at an image through a magnifying glass to spot specific details.

How Does a CNN Work?
Pooling

What happens in pooling?

- Reduces the size of the image while keeping important features.

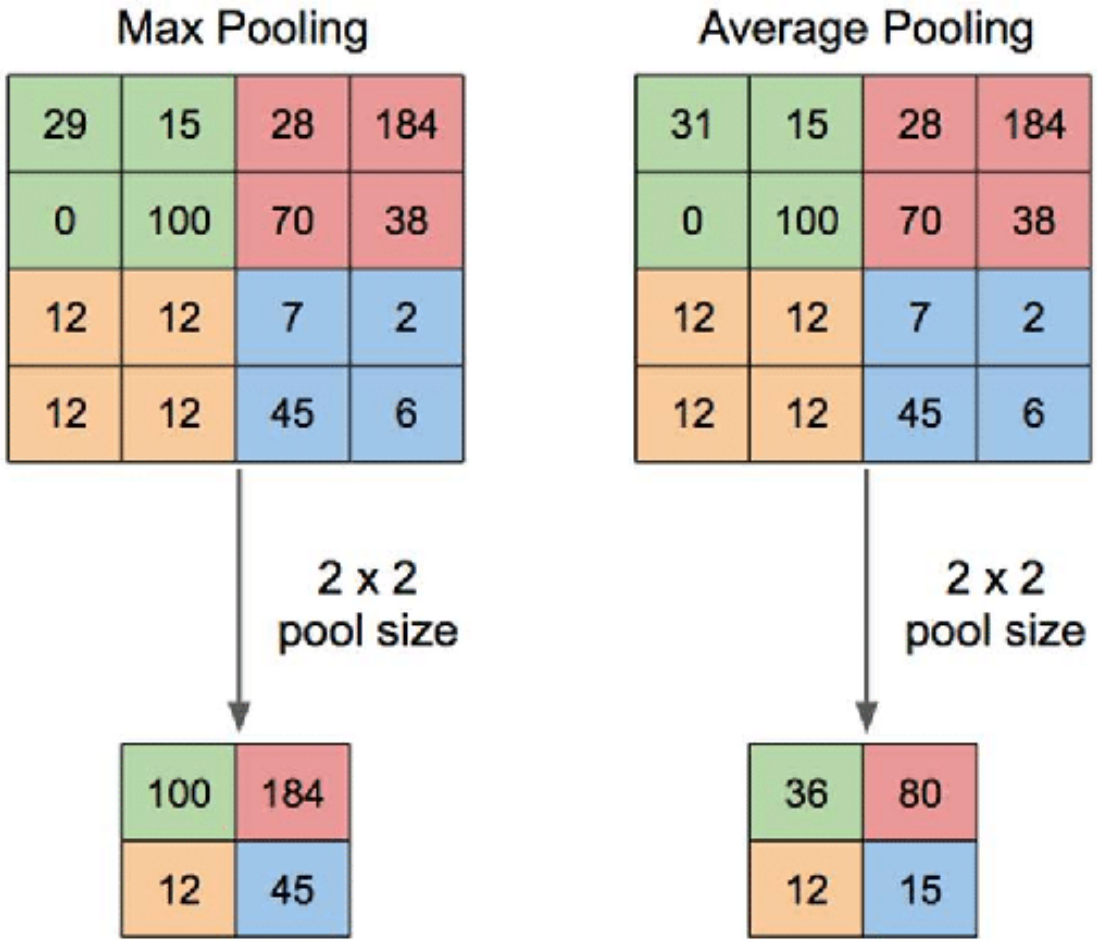
- Common type: **Max Pooling** (takes the largest value in a region).

Example:

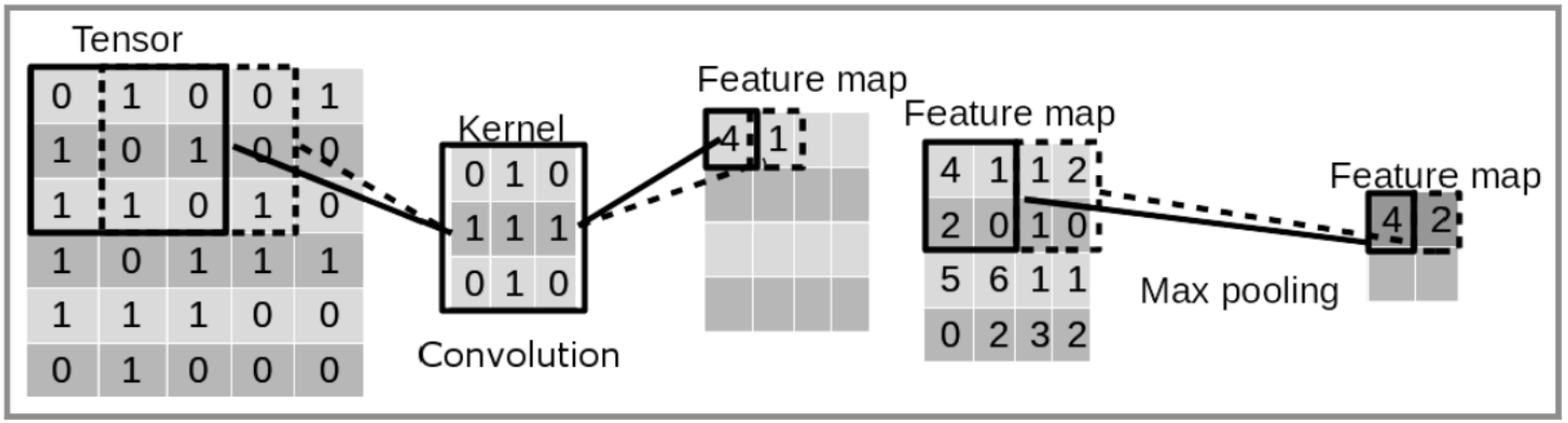
- From a 4x4 grid → 2x2 grid.

How Does a CNN Work?

Pooling



How Does a CNN Work?



Example for convolution and max pooling operation.

source

How Does a CNN Work?
Fully Connected Layers

After extracting features using convolution and pooling:

- **Fully Connected Layers** classify the image.
- Every feature is connected to every possible output class (e.g., cat, dog, car).

Like collecting puzzle pieces and putting them together to see the full picture!

UWM College of Engineering & Applied Science

Example - Recognizing a Dog

1. **Input:** A dog image.
2. **Convolution:** Detects edges, shapes (e.g., ears, eyes, nose).
3. **Pooling:** Simplifies features while keeping "dog-like" details.
4. **Fully Connected Layers:** Combines features and classifies it as "Dog."

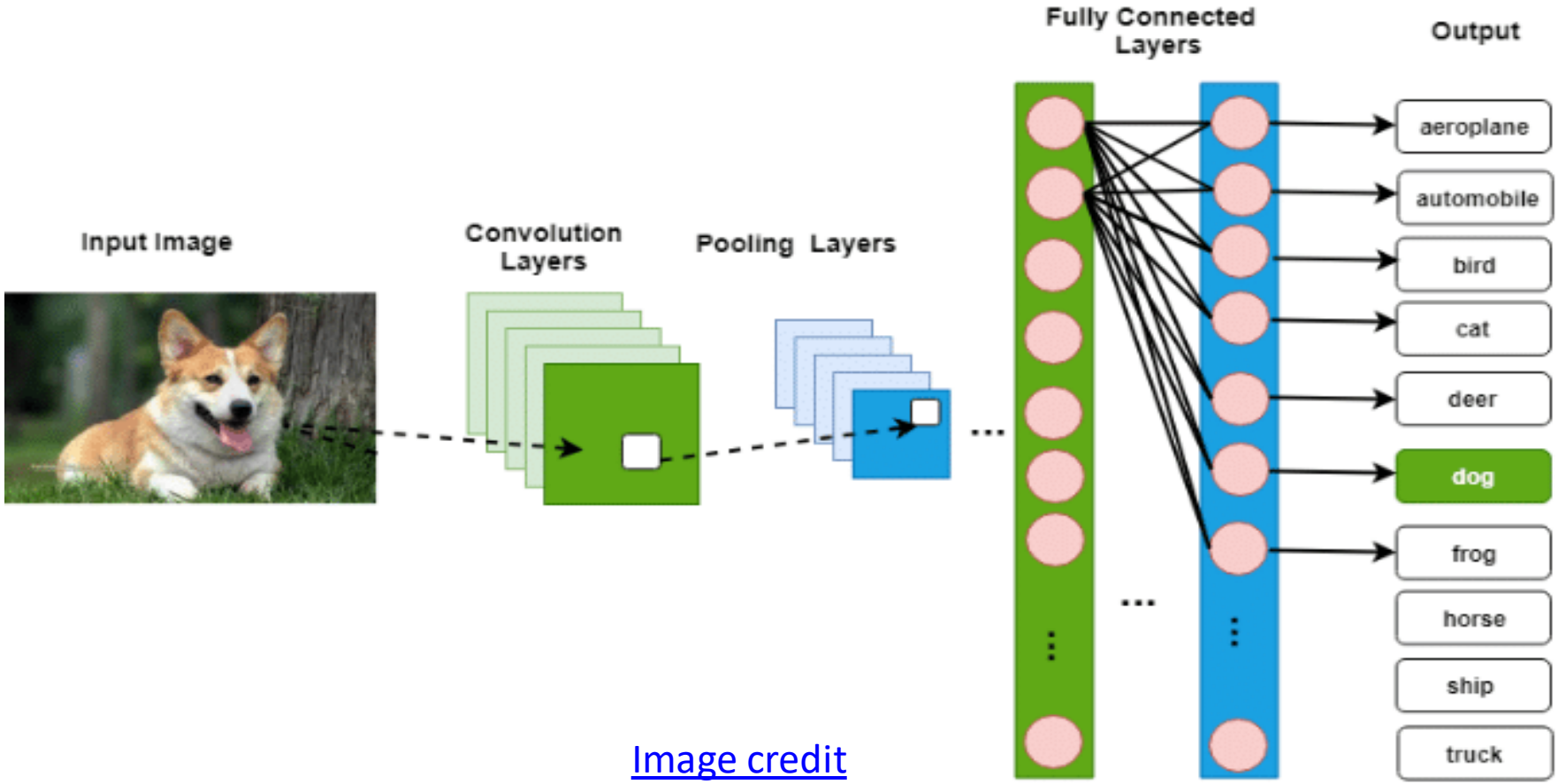


Image credit

Training a CNN

- CNNs learn by adjusting weights through **training**:
  1. Feed labeled images (e.g., cat, dog).
  2. CNN makes a prediction.
  3. Calculate the error (difference between prediction and true label).
  4. Update weights to improve predictions (using **backpropagation**).