AI CAMP

school of ai Algiers

Hello and Welcome to AI Camp

AI CAMP

school of ai Algiers

Intro to computer vision

Adjal Mehdi Zakaria

Definition of Computer Vision:

Computer vision is a multidisciplinary field that focuses on enabling machines to interpret and understand visual information from the world. In simpler terms, it aims to replicate the human visual system's ability to process and make sense of visual data, such as images and videos, using computational techniques

Applications of Computer Vision:

Computer vision has a wide range of applications across various industries and domains:

- Medical Imaging
- Autonomous Vehicles
- Industrial Automation
- and more

Definition of Computer Vision:

Computer vision encompasses a wide range of sub-fields and tasks that focus on different aspects of visual data analysis and interpretation. Here's a breakdown of some prominent sub-fields and tasks within computer vision :

- Image Classification
- Object Detection
- Image Segmentation
- Image generation
- and more

Understanding Image Data

Computer vision is a multidisciplinary field that focuses on enabling machines to interpret and understand visual information from the world. In simpler terms, it aims to replicate the human visual system's ability to process and make sense of visual data, such as images and videos, using computational techniques

Understanding Images

AI CAMP

school of ai Algiers

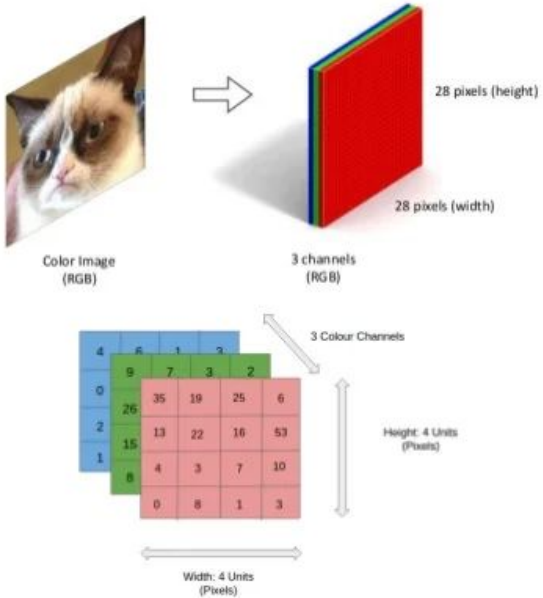
Understanding Images

In the realm of Computer Vision (CV), an image is a two-dimensional array of numerical values that represent visual data captured or generated by cameras or other imaging devices. These numerical values correspond to the intensity levels of pixels arranged in rows and columns, forming a visual representation of a scene, object, or pattern.

A pixel (short for "picture element") is the smallest unit of a digital image, representing a single point in the image grid. Each pixel contains numerical values corresponding to its color, brightness, and possibly other attributes, depending on the image type (e.g., grayscale, RGB).

Understanding Images

color image is 3rd-order tensor



AI CAMP

school of ai Algiers

Image Preprocessing in the Context of Computer Vision

AI CAMP

school of ai Algiers

Image preprocessing

Image preprocessing is a critical step in Computer Vision (CV) workflows that involves applying various techniques to enhance, transform, and prepare raw image data for subsequent analysis, feature extraction, or machine learning tasks. Preprocessing generally involves the following steps :

- **Image Resizing :** Adjusts the dimensions (width x height) of images to meet specific requirements, such as model input size, storage constraints, or computational efficiency.
- **Normalization:** Scales pixel values within a standardized range, typically between 0 and 1, to ensure consistent data distribution and facilitate model convergence.

Image preprocessing

Image preprocessing is a critical step in Computer Vision (CV) workflows that involves applying various techniques to enhance, transform, and prepare raw image data for subsequent analysis, feature extraction, or machine learning tasks. Preprocessing generally involves the following steps :

- **Image Resizing :** Adjusts the dimensions (width x height) of images to meet specific requirements, such as model input size, storage constraints, or computational efficiency.
- **Normalization:** Scales pixel values within a standardized range, typically between 0 and 1, to ensure consistent data distribution and facilitate model convergence.

Image preprocessing

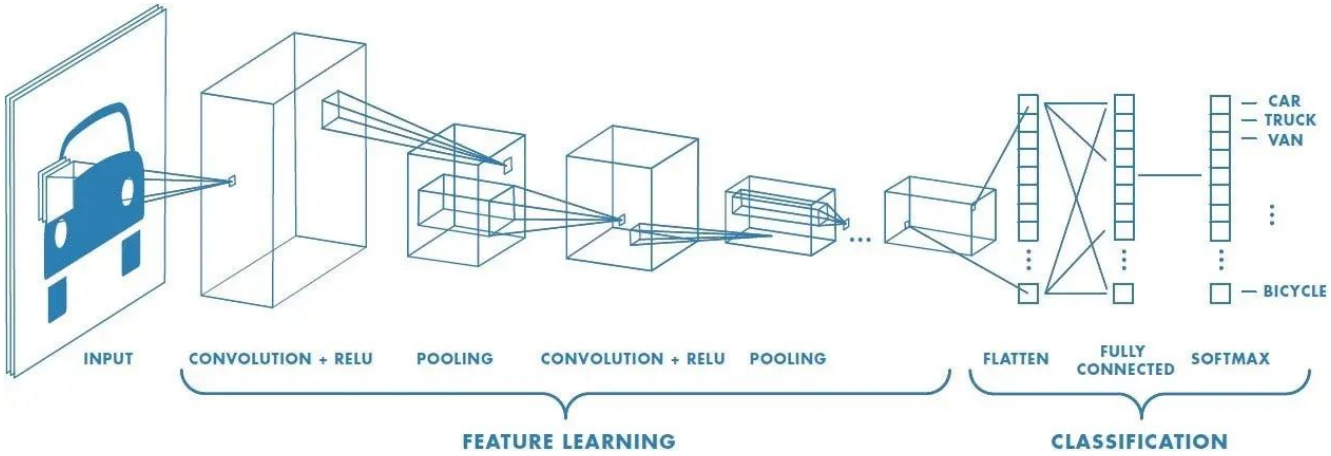
Image preprocessing is a critical step in Computer Vision (CV) workflows that involves applying various techniques to enhance, transform, and prepare raw image data for subsequent analysis, feature extraction, or machine learning tasks. Preprocessing generally involves the following steps :

- **Image Resizing :** Adjusts the dimensions (width x height) of images to meet specific requirements, such as model input size, storage constraints, or computational efficiency.
- **Normalization:** Scales pixel values within a standardized range, typically between 0 and 1, to ensure consistent data distribution and facilitate model convergence.
- **Flattening**: Typically refers to the process of converting a multidimensional data structure such as images into a one-dimensional array.

Convolutional Neural Netowrks (CNN)

Understanding Image Data

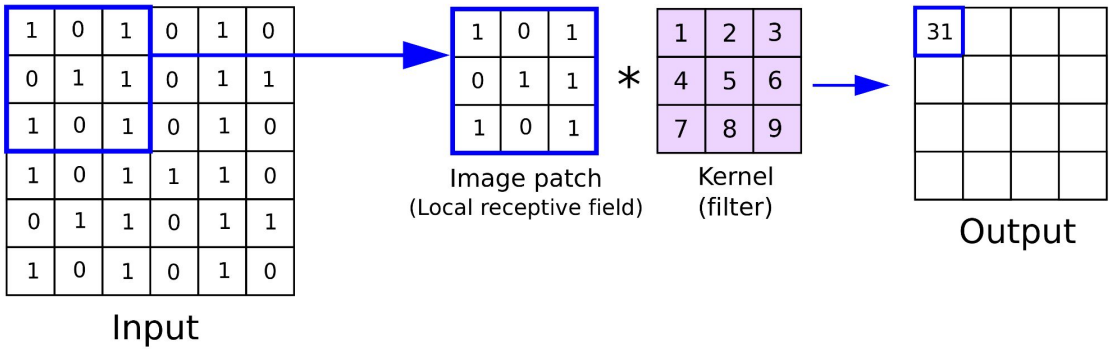
A Convolutional Neural Network (CNN) is a specialized type of deep neural network designed primarily for processing and analyzing visual data, such as images and videos. CNNs are inspired by the human visual system's architecture, emphasizing feature extraction and hierarchical learning.



Key concepts in CNNs : Convolutional layer

These layers apply convolution operations to input images, extracting local features using small, learnable filters or kernels.

Features like edges, textures, and patterns are detected through convolution operations, enabling the network to recognize more complex patterns in deeper layers.



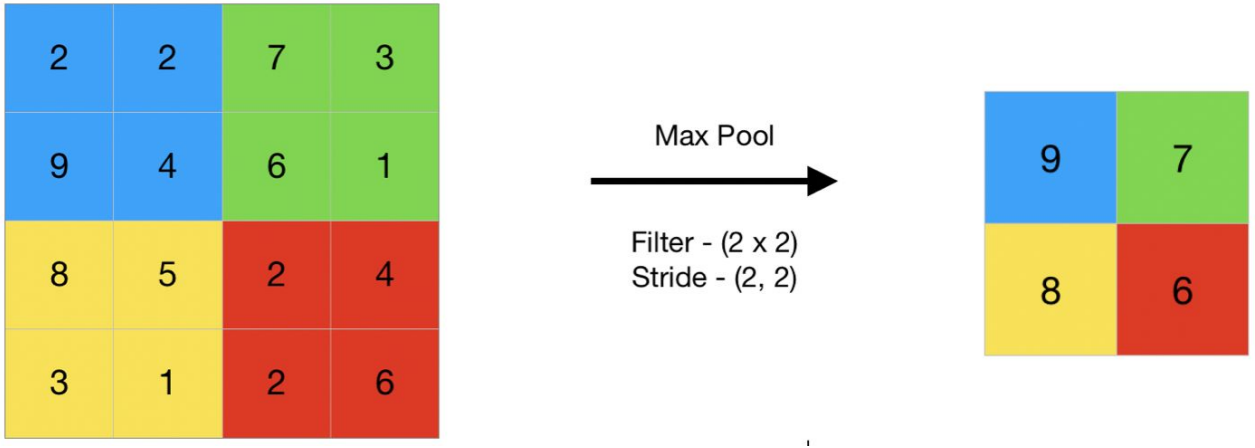
AI CAMP

school of ai Algiers

Key concepts in CNNs : Pooling layer

Pooling layers reduce spatial dimensions (width x height) of feature maps, preserving essential information while reducing computational complexity and preventing overfitting.

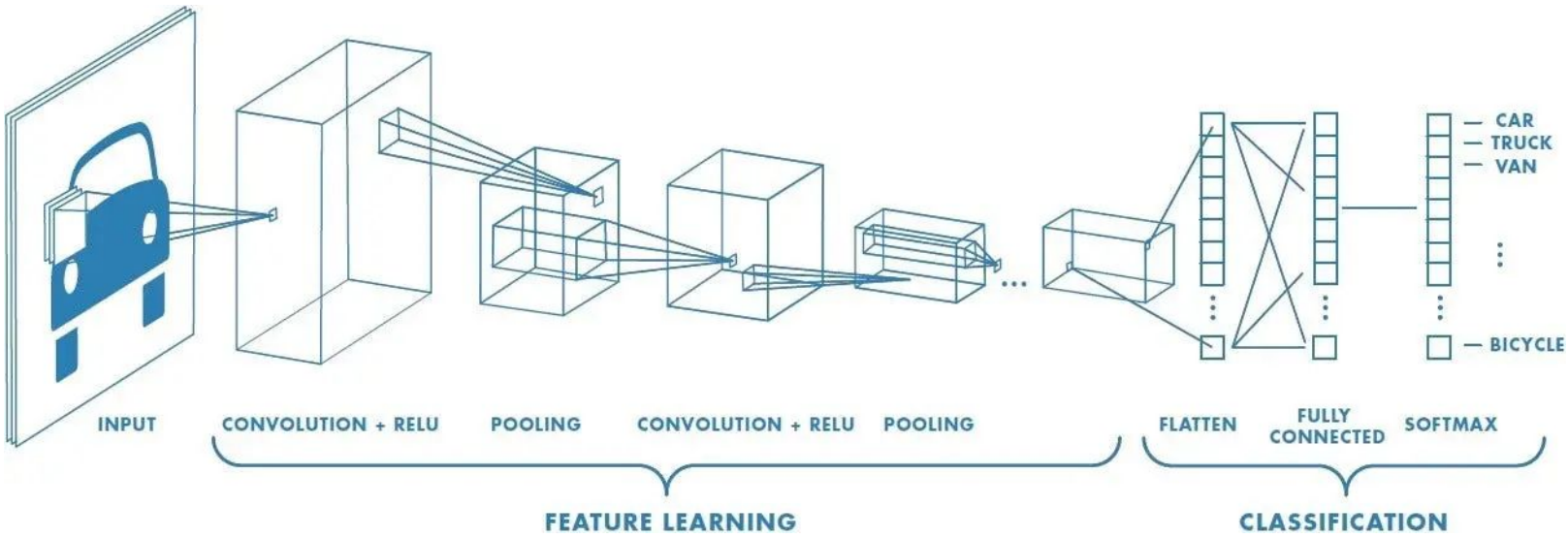
Common pooling techniques include max pooling, average pooling, and min pooling, extracting dominant features and spatial hierarchies from feature maps.



AI CAMP

school of ai Algiers

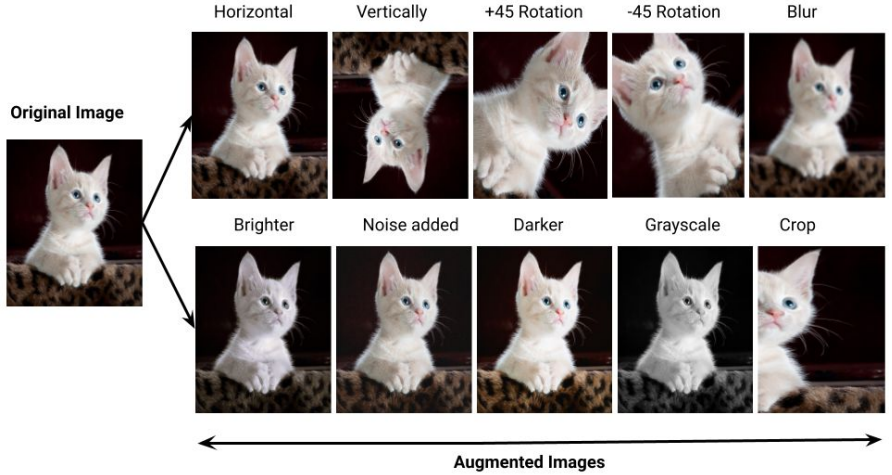
CNNs


CAR
TRUCK
VAN
BICYCLE
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
FEATURE LEARNING
CLASSIFICATION

Data augmentation

Data augmentation

Data augmentation refers to a set of techniques used to artificially expand or diversify a training dataset by applying various transformations, modifications, or enhancements to existing images or samples. These transformations introduce variations in the data, such as changes in appearance, orientation, lighting, or perspective, to improve model generalization, robustness, and performance in computer vision tasks.



AI CAMP

school of ai Algiers

Why data augmentation

Convolutional Neural Networks (CNNs) are powerful deep learning models widely used in computer vision tasks like image classification, object detection, and segmentation. However, one common challenge faced when training CNNs is overfitting.

By applying various transformations, such as rotations, translations, flips, and color adjustments, data augmentation increases the diversity, complexity, and richness of the training dataset. This exposure to diverse data variations helps the model generalize better, capturing essential features and patterns while ignoring noise, anomalies, or irrelevant details that may lead to overfitting

AI CAMP

school of ai Algiers

Thank you for attending

any questions ?