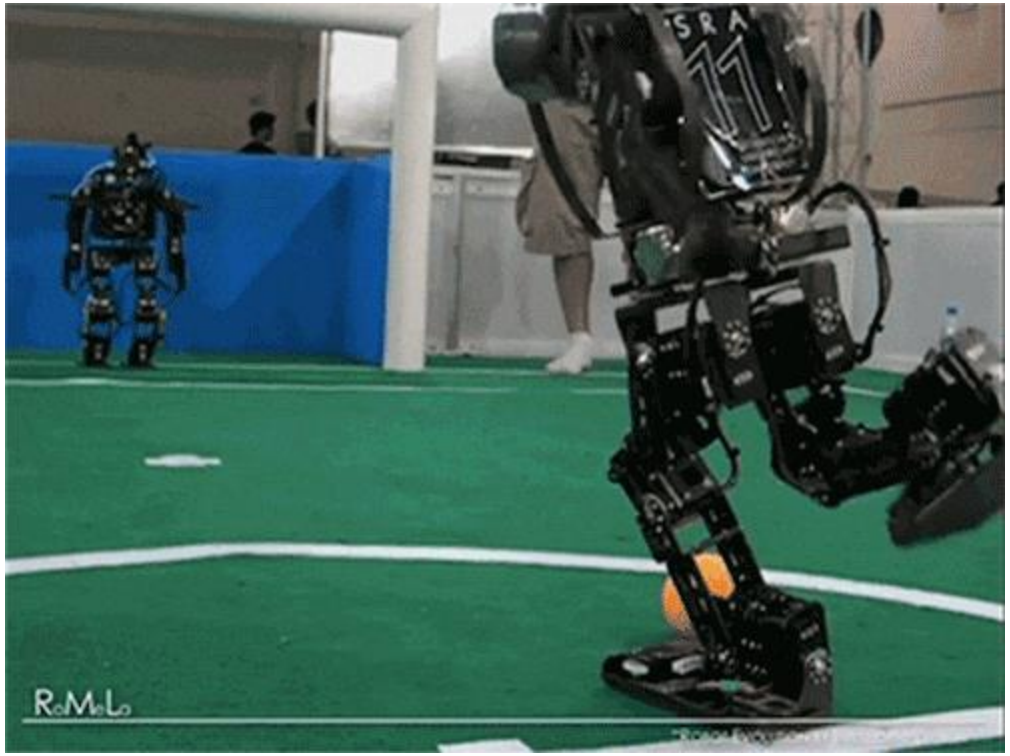
10 – Classical Computer Vision for Robotics

Robotics and Computer Vision
BPC-PRP

Ing. Petr Šopák
Brno University of Technology
2025

Motivation

"Imagine living without vision – that’s how robots experience the world without computer vision."



What will we learn today?

- **The Standard Pipeline of Computer Vision**
- **Camera as a Sensor**
- **Introduction to Images and Preprocessing**
- **Segmentations and Feature Description**
- **From vision to Action**

Next Lecture:

- **Advance Computer Vision** – CNN and more

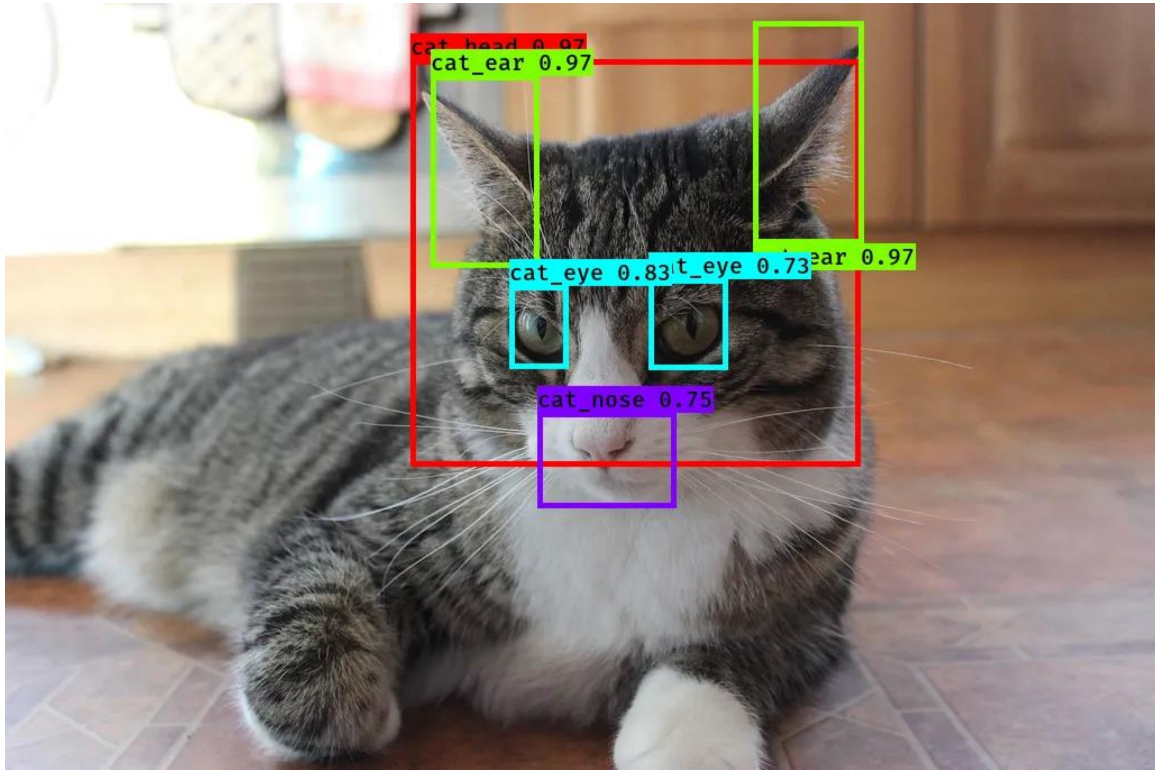
What is Computer Vision?

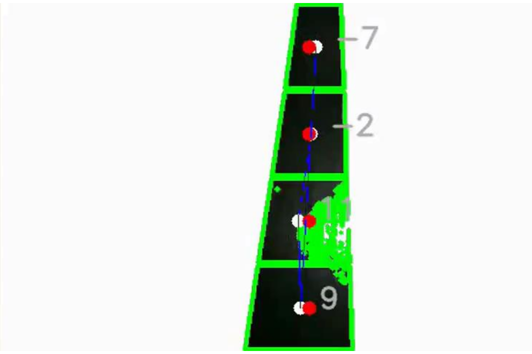
Definition of the term computer vision, its applications and limitations. Then, a discussion of the standard pipeline.

What is Computer Vision?

- A field of AI that enables machines to "see" and understand visual data
- Tries to mimic the Human vision
- Transforms images or videos into useful information

- **The Goal**: recognize, interpret, and react to what is seen

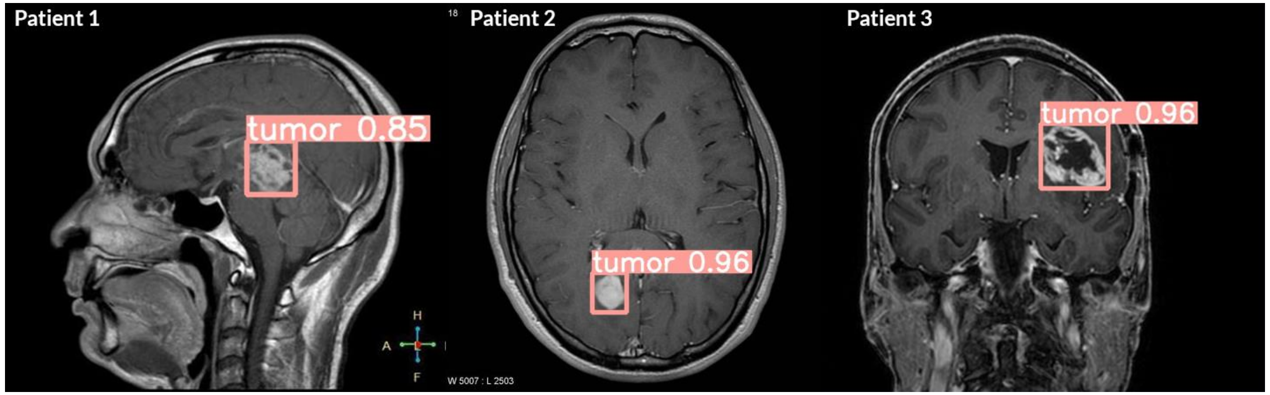




1. Xie, E.. „*Using Computer Vision to Find the Best Cat Photo*“. Medium, 2020. [online]. Available: [Using Computer Vision to Find the Best Cat Photo | by Emily Xie | TDS Archive | Medium]
2. CRM-UAM, „*VisionRace*“. GitHub Repository, 2019. Available: https://github.com/CRM-UAM/VisionRace

Computer Vision - Uses

- **Example Uses:**
  - Autonomous navigation
  - Healthcare image processing
  - Augument Reality
  - OCR
  - Quality Inspection
  - Barcode and OQ code recognition
  - Security systems
  - Generative Applications
  - 3D Reconstruction
  - More ...

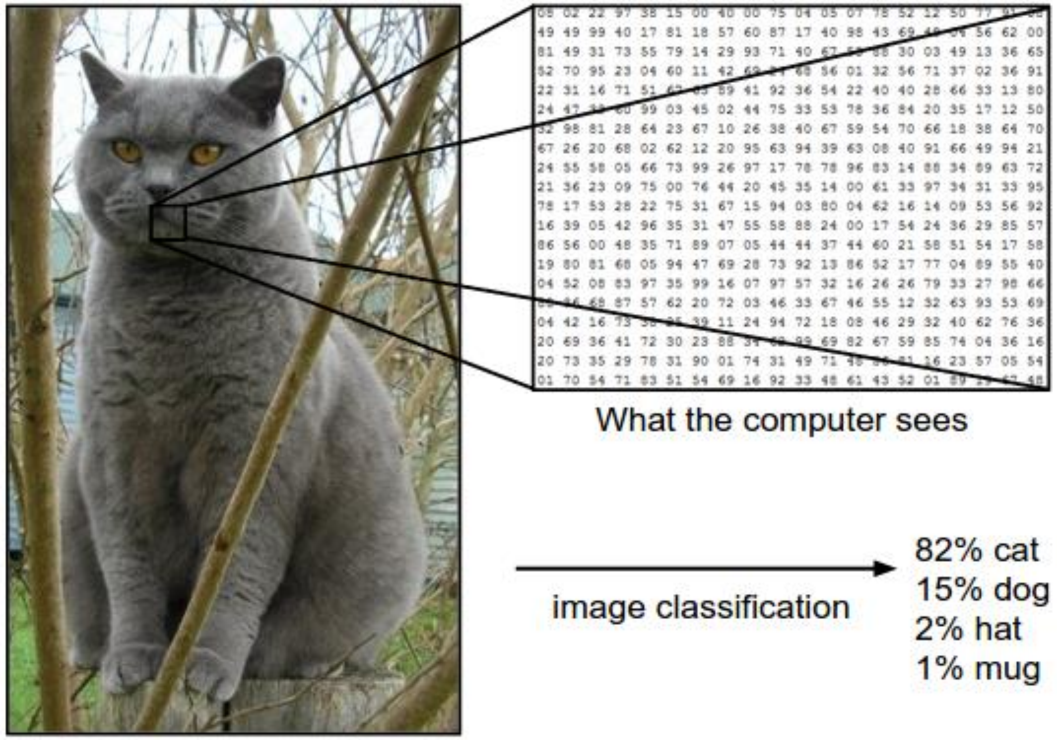


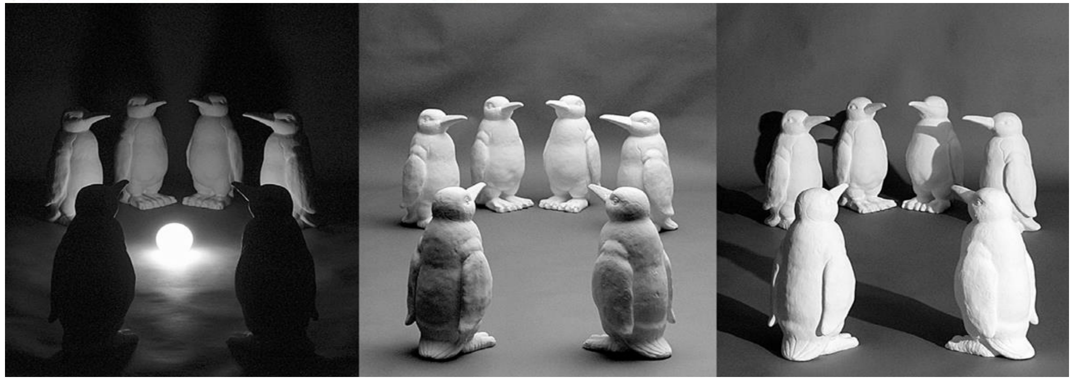
1. Giuseppebrb. „ *BrainTumorDetection*“. Github Repository, 2022. [online]. Available: giuseppebrb/BrainTumorDetection: Brain tumor detection using computer vision
2. Junyanz, „*CycleGAN*“. GitHub Repository, 2019. Available: junyanz/CycleGAN: Software that can generate photos from paintings, turn horses into zebras, perform style transfer, and more.

Computer Vision - Challenges

- **Why Computer Vision is hard?**

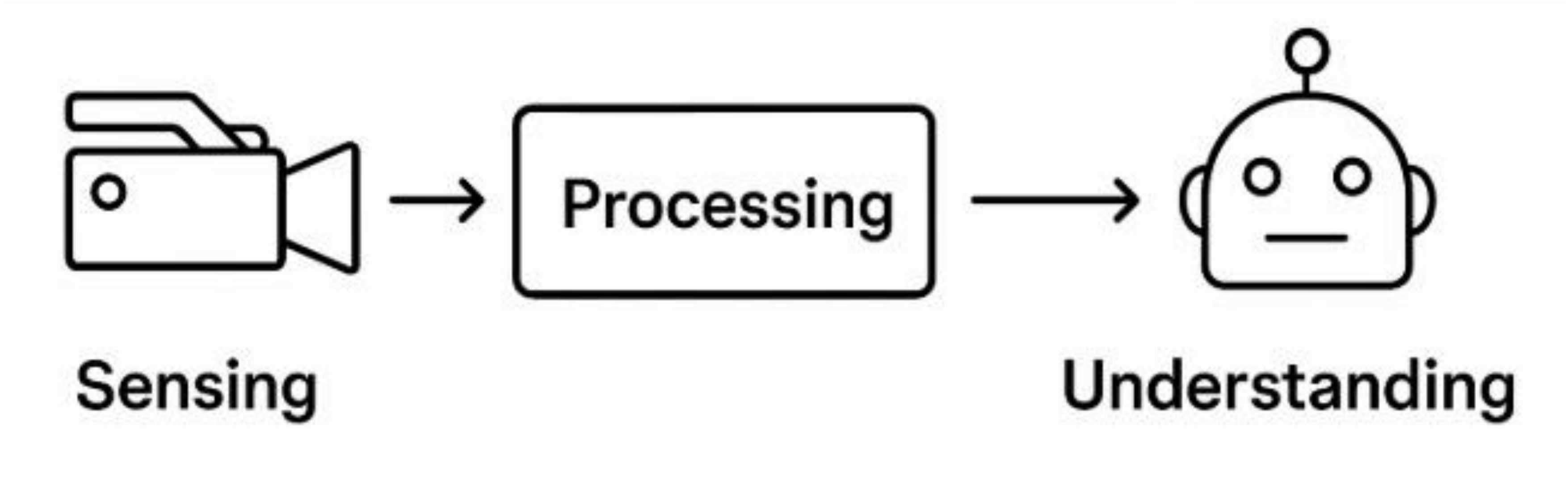
- System sees pixels, not objects
- Sensitive to lighting, angle and occlusion
- Struggles with Generalization
- Requires huge dataset (Labeled)
- Vulnerable to noise
- Intra-class variation
- Image Stitching

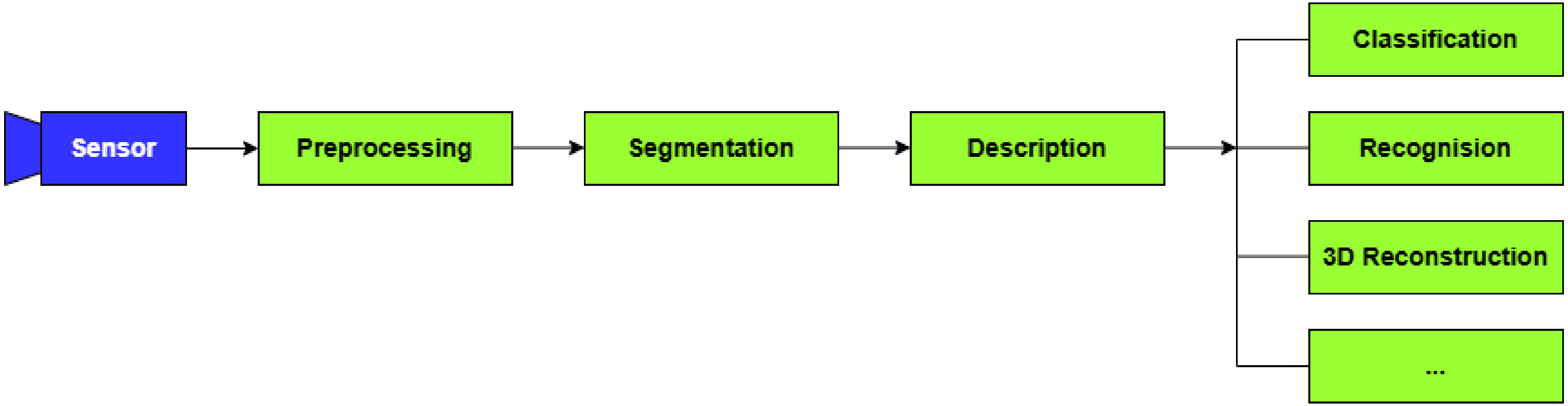




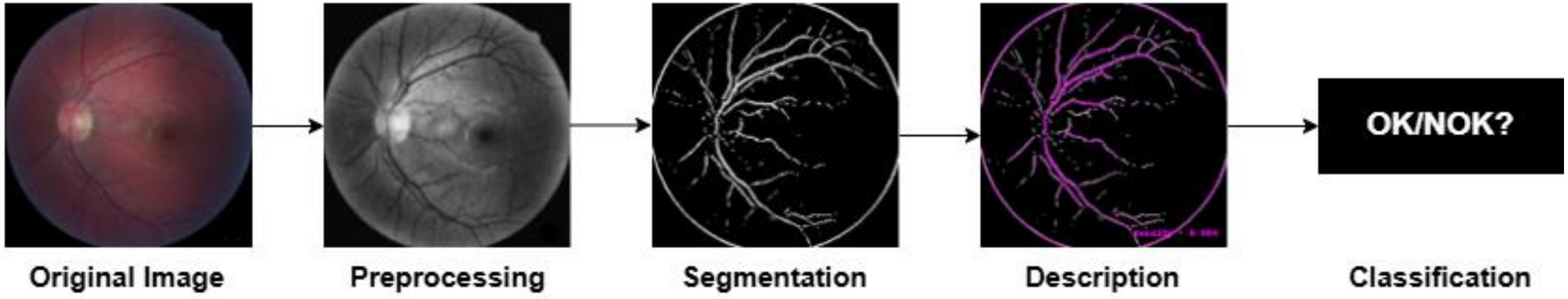
1. Bansac, F., „ *Computer Vision and convolutional* neural networks". Ailephant, 2018. [online]. Available: Computer vision and convolutional neural networks – AILEPHANT
2. J. Koederink

What is Perception?




Sensor
Preprocessing
Segmentation
Description
Classification
Recognision
3D Reconstruction
...

Example:

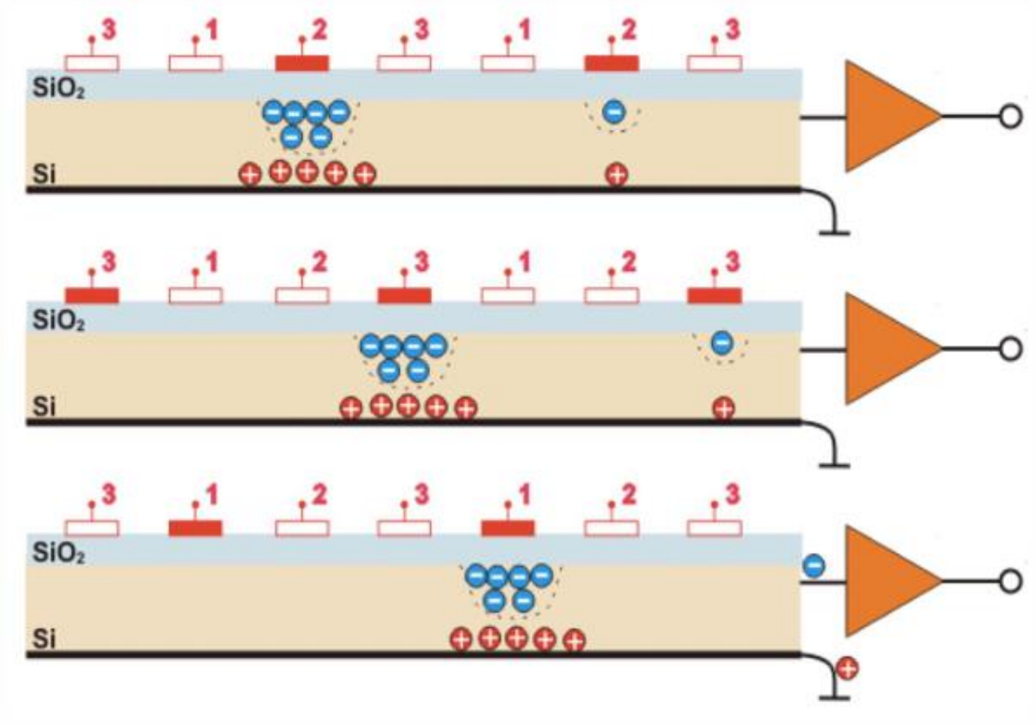

OK/NOK?
Original Image
Preprocessing
Segmentation
Description
Classification

Camera as a Sensor

Introduction to how machines capture visual information using cameras. Cover camera types, important parameters, and limitations that affect image-based processing.

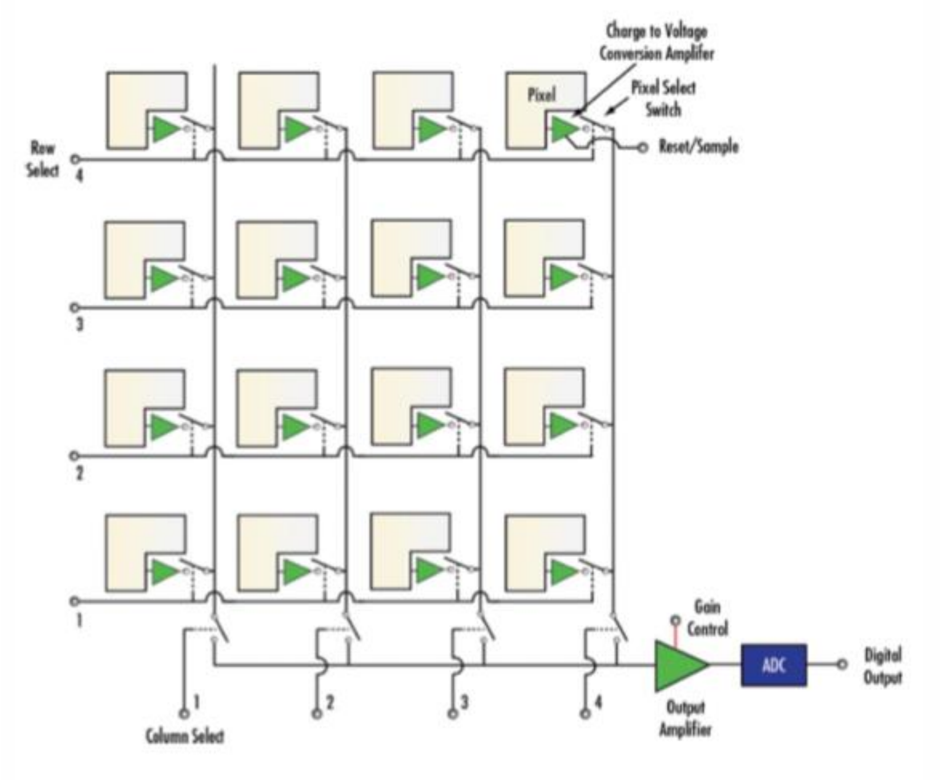
CCD – Charge-Coupled Device

- Photoelectric effect
- Photons are converted into electron-proton pair
- Charge shifters move charge from each pixel one by one to the ADC
- The ADC is typically very precise – lower noise, higher dynamic range



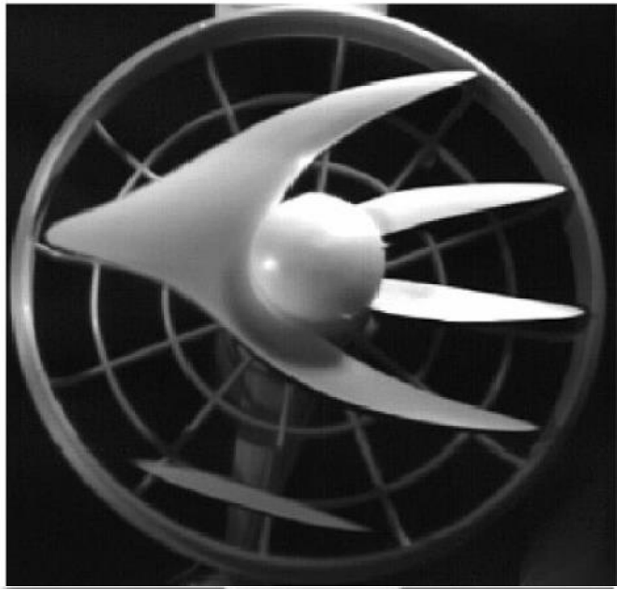
CMOS – Complementary Metal-Oxide-Semiconductor

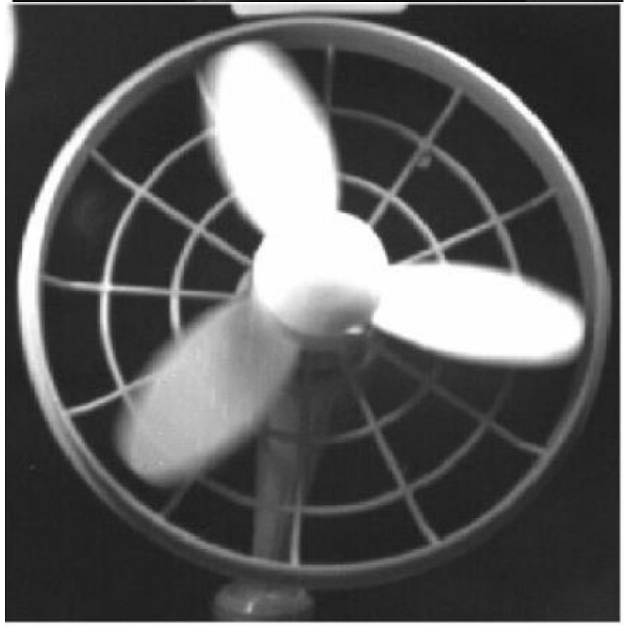
- **Every pixel has own photodiode and amplifier**
- To get the pixel value we measure the current that the photodiode generates
- Each pixel value can be measured in a different time (subwindowing, rolling shutter)
- Large number of accompanying electronics reduce the sensing space
- Example: **Rasberry Pi Camera**



CMOS – Complementary Metal-Oxide-Semiconductor

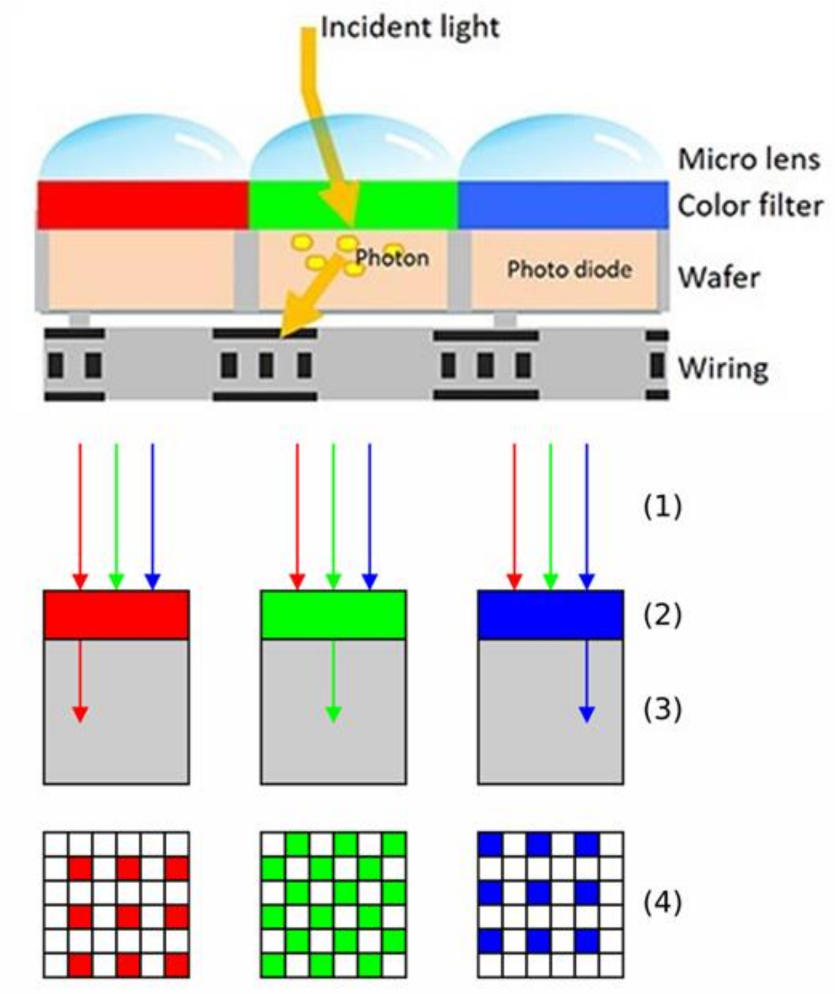
- **Every pixel has its own photodiode and amplifier**
- To get the pixel value we measure the current that the photodiode generates
- Each pixel value can be measured in a different time (subwindowing, rolling shutter)
- Large number of accompanying electronics reduce the sensing space
- **Rasberry Pi Camera**
  - **RGB CMOS sensor**
  - **Rolling shutter**
  - **Bayer mask**
  - **Demosaicing**





Sensor Structure and Pixel-Color Filtration

- Every pixel is a potential well that traps excited electrons till they’re read out
- To maximize sensitivity, each pixel has its own micro lens that focuses light
- Color differentiation is achieved through the filtering
- Our eyes are most sensitive to green color, so Bayer filters have 2 green filters for every red and blue

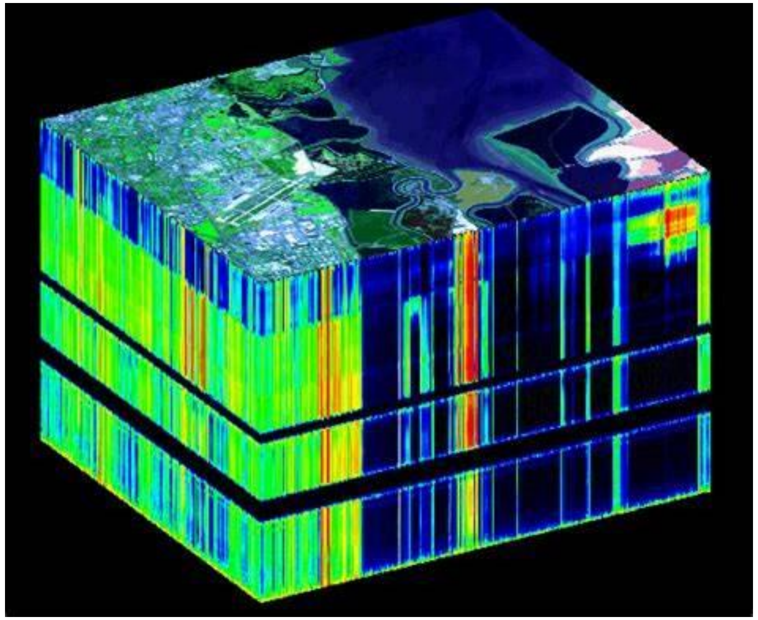


CMOS vs CCD

	CCD	CMOS
Price	High	Low
Power con.	High	Low
Quality	High	Low
Noise	Low	High
Dynamic Range	High	Low
Resolution	High	Mid
Sensing Speed	Mid	High
Fill Factor	High	Mid

Other Types of Camera Sensors

- RGB cameras
- Monochrome Cameras
- IR Cameras (NIR or thermal)
- Depth Cameras
- Hyperspectral / Multispectral Cameras
- Line Scan Camera



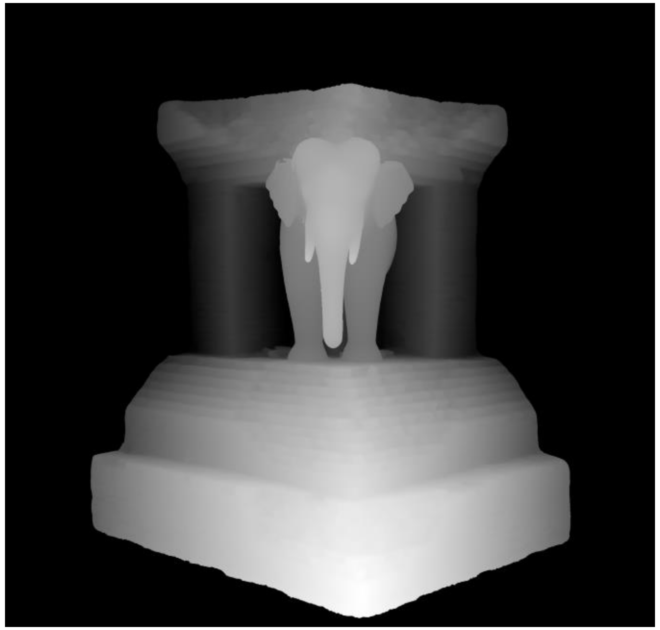
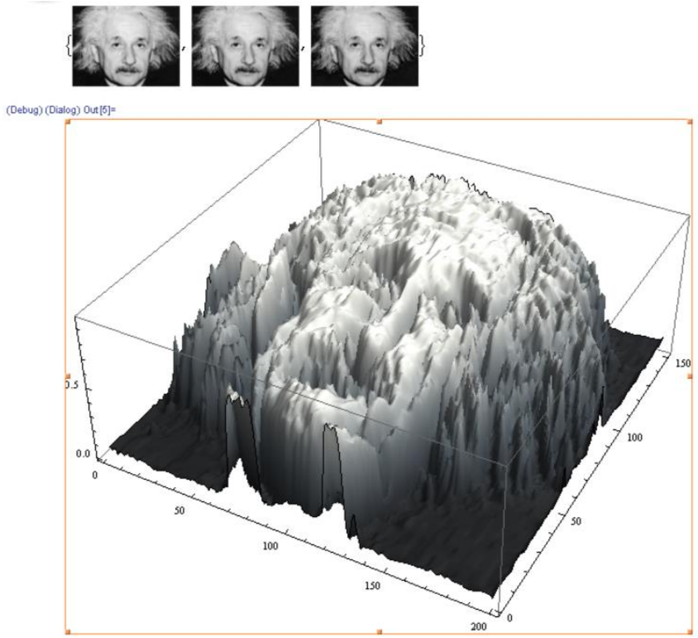


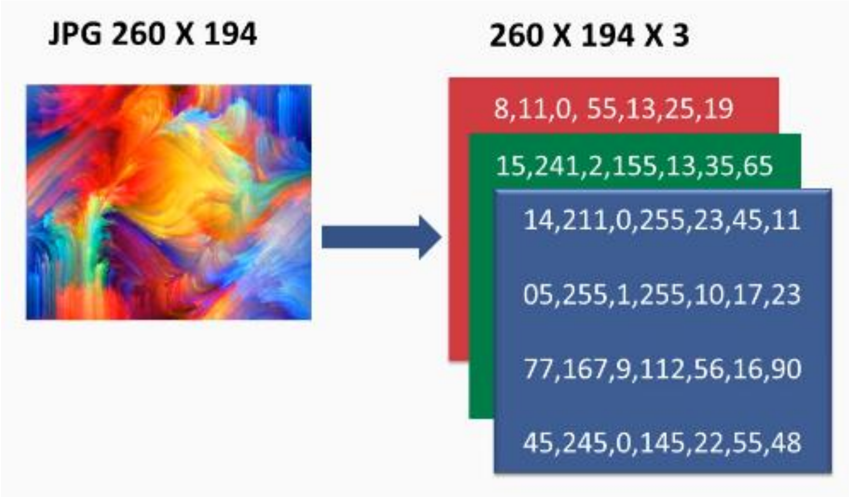
Image Preprocessing

Understanding what an image is and how we can operate on it — the foundation for understanding image preprocessing.

What is an Image?

- An image is 2D array of pixels
- Each pixel stores intensity or color information
- Basic Types of Images:
  - Grayscale Image
  - RGB Image
  - Binary Image
  - HSV / HSL Image
  - Thermal Image
  - And more...





- Raw images are often **too noisy, too detailed, or inconsistent**
- Preprocessing helps us:
  - Standardize input (size, color format)
  - Enhance relevant information
  - Remove noise and distortions
  - Simplify the image for analysis

Image Preprocessing - examples

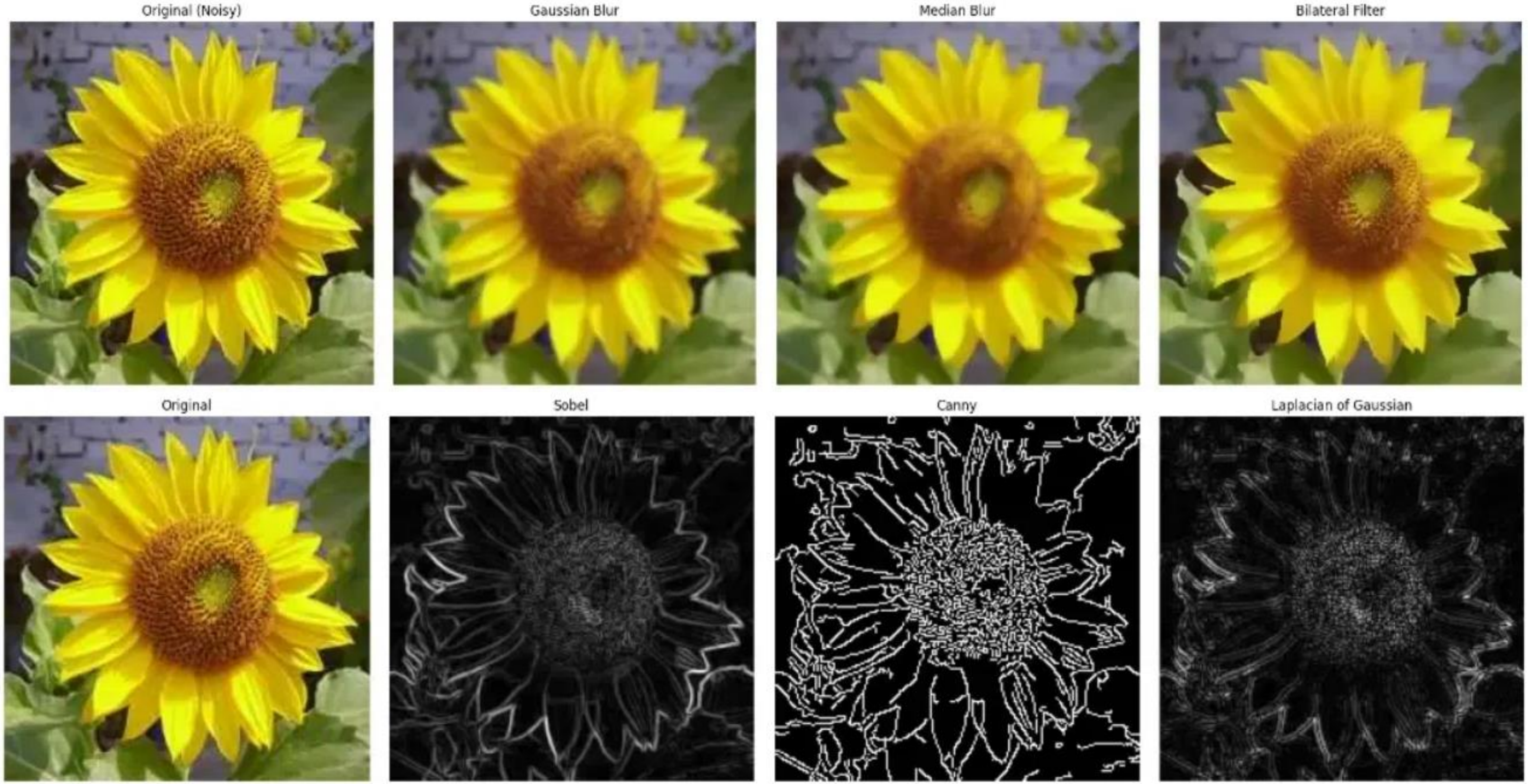
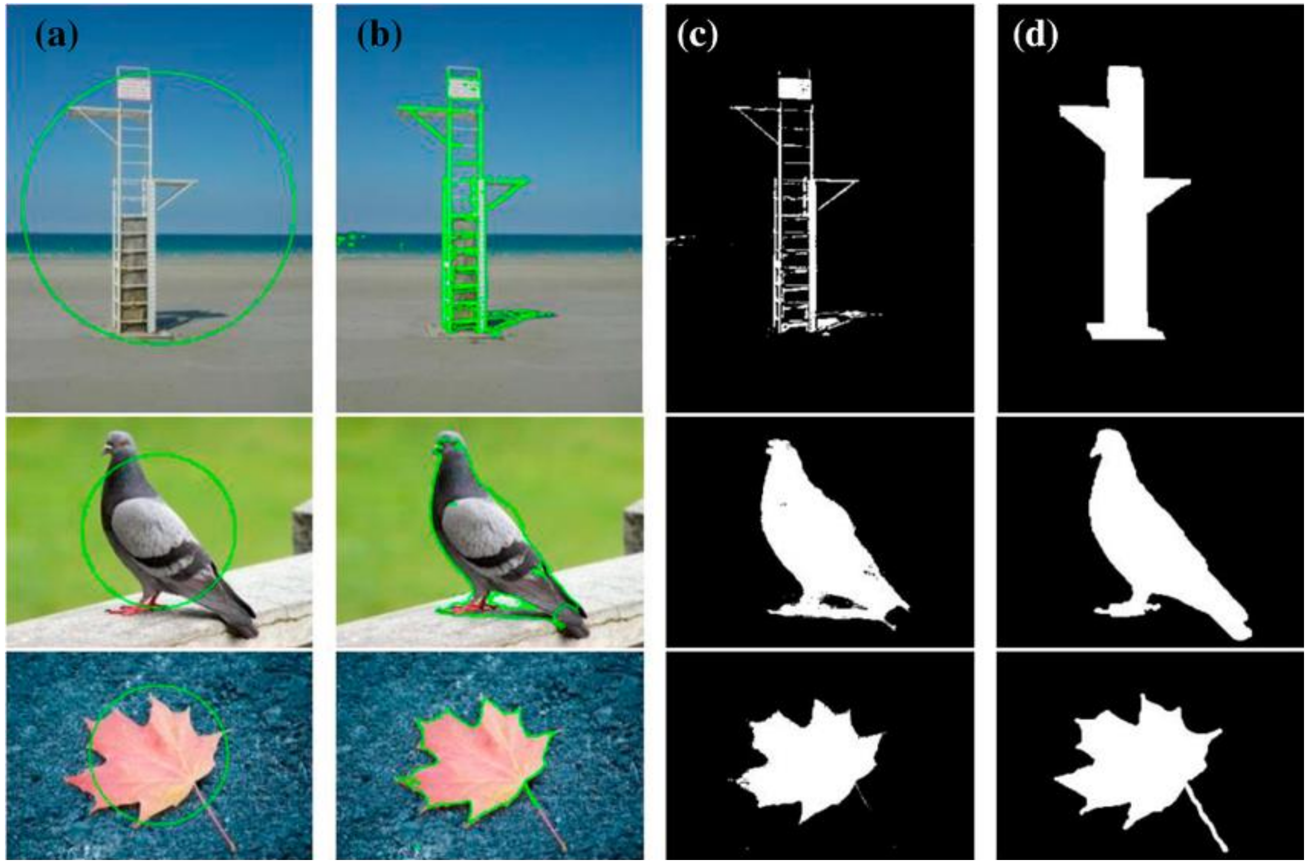


Image Segmentation

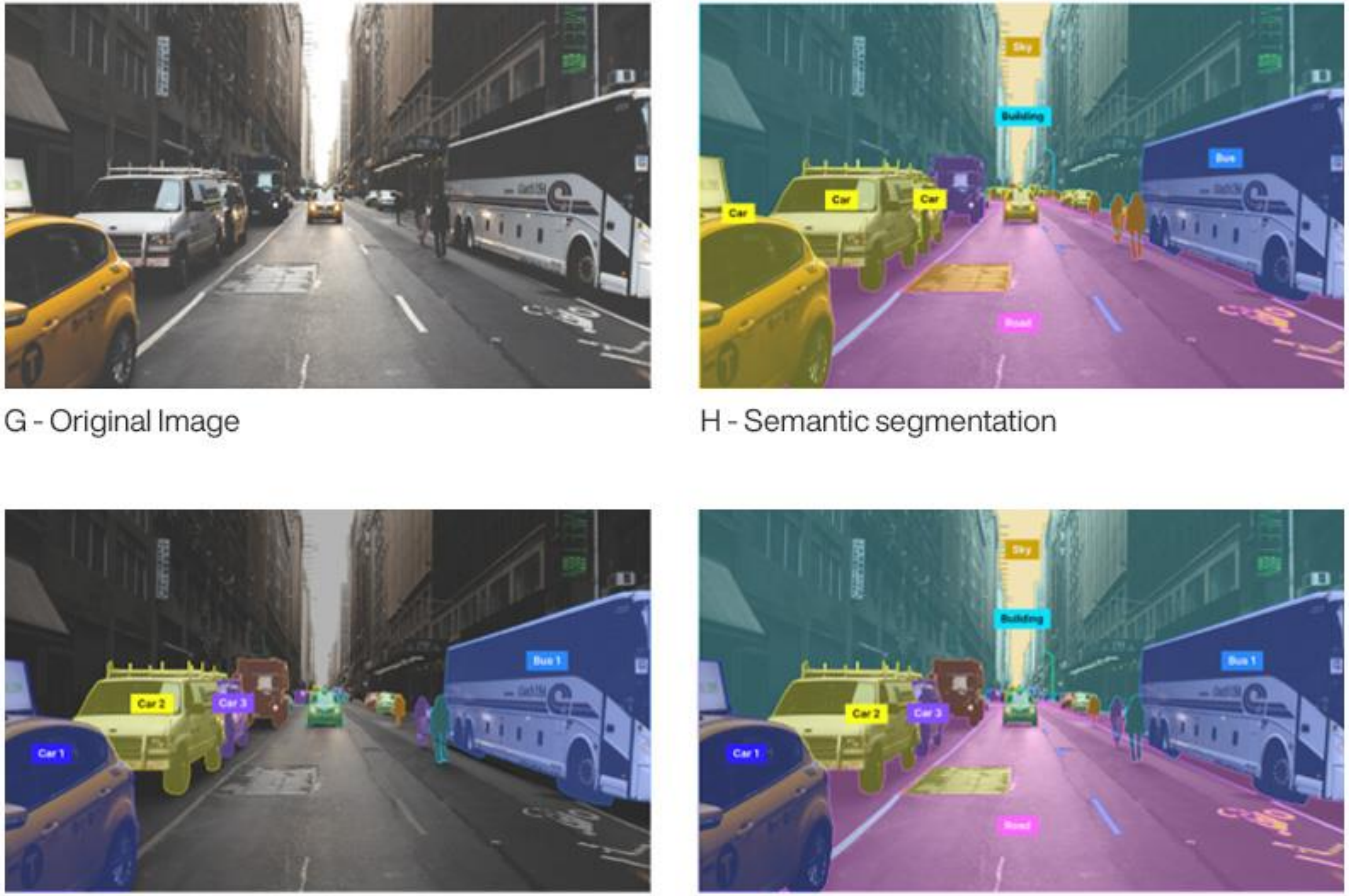
Dividing the image into meaningful parts — the first step toward understanding structure and content.

- Segmentation means **dividing an image into meaningful parts**
- Helps isolate objects, regions, or features
- Simple methods: **thresholding, contour detection, region growing**
- Advanced: **semantic or instance segmentation** (deep learning)



1. MING, D., ZHIHENG, Z. ed. „Image Segmentation Using Level Set Driven by Generalized Divergence“. *Circuits Syst Signal Process, 2021*. [online]. Available: 10.1007/s00034-020-01491-x

Image Segmentation – Sematic and Instance



G - Original Image

H - Semantic segmentation

I - Instance segmentation

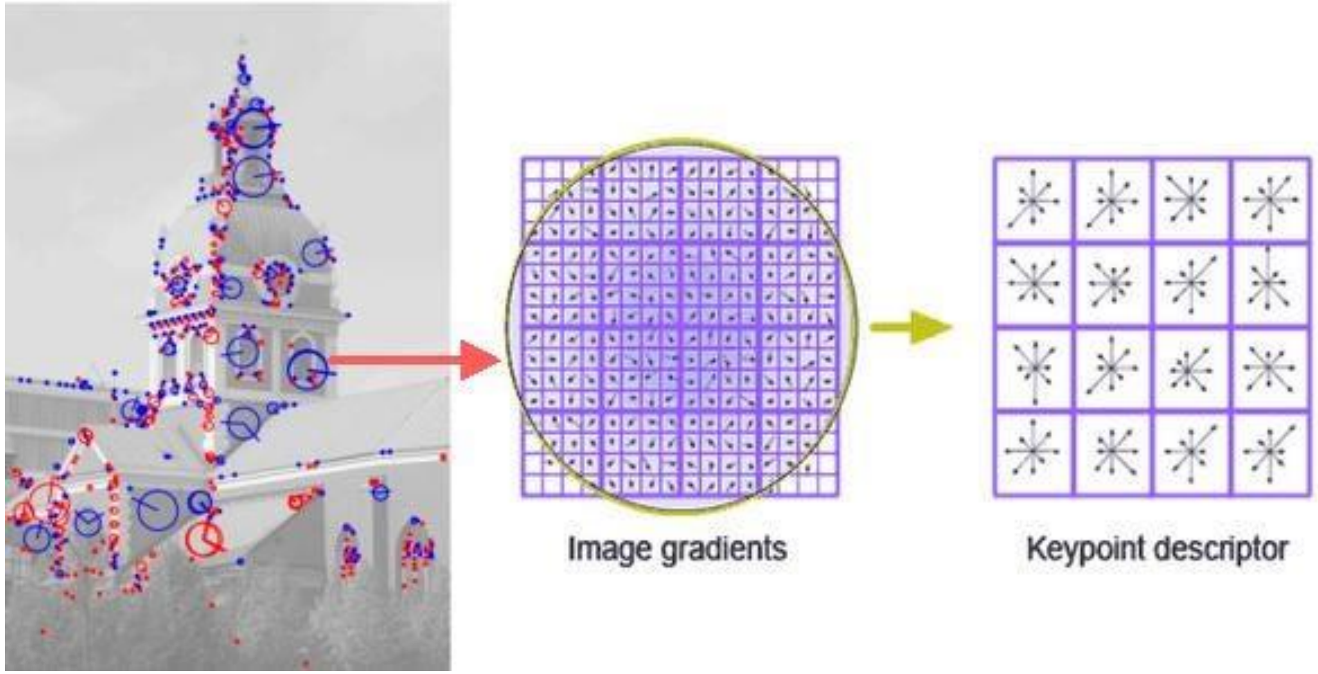
J - Panoptic segmentation

1. *VINA. A.„Using AI for crack detection and segmentation*“. UltraLytics YOLO, 2024. [online]. Available: [AI Crack Detection & Segmentation with YOLOv8]

Feature Extraction and Description

Identifying and describing visual patterns — the key to recognition and matching

- Extracts **distinctive information** from the image
- Transforms image regions into **numerical descriptors**
- Types of features:
  - **Local**: corners, blobs (e.g. ORB, SIFT)
  - **Global**: color histograms, HOG, shape descriptors
- Used for **matching, tracking, recognition**



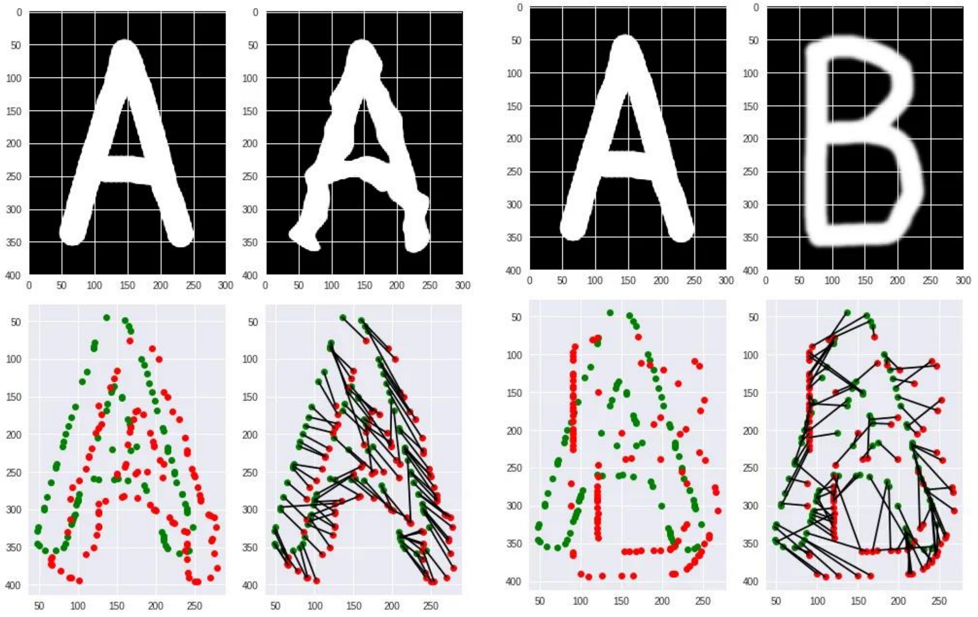
1. Nikishaev, A., *„Feature extraction and similar image search with OpenCV for newbies“*. Medium, 2018. [online]. Available: [Feature extraction and similar image search with OpenCV for newbies | by Andrey Nikishaev | Machine Learning World | Medium]

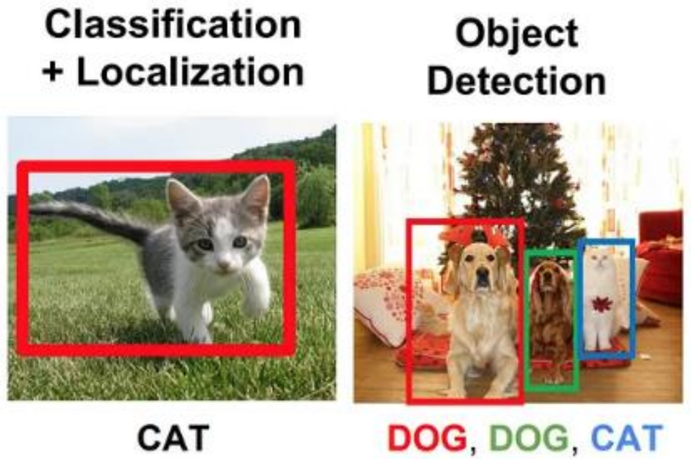
What Happens After Description?

From understanding to outcome — the final step of vision processing.

Final step – What Happens now?

- The standard computer vision pipeline often ends with:
  - **Classification** – What is it?
  - **Identification** – Who is it? / Which specific object is this?
  - (**Verification** – Is it the same as..?)
  - **Matching** – Is it similar to a known pattern?
  - **Localization** – Where is it in the image/space?





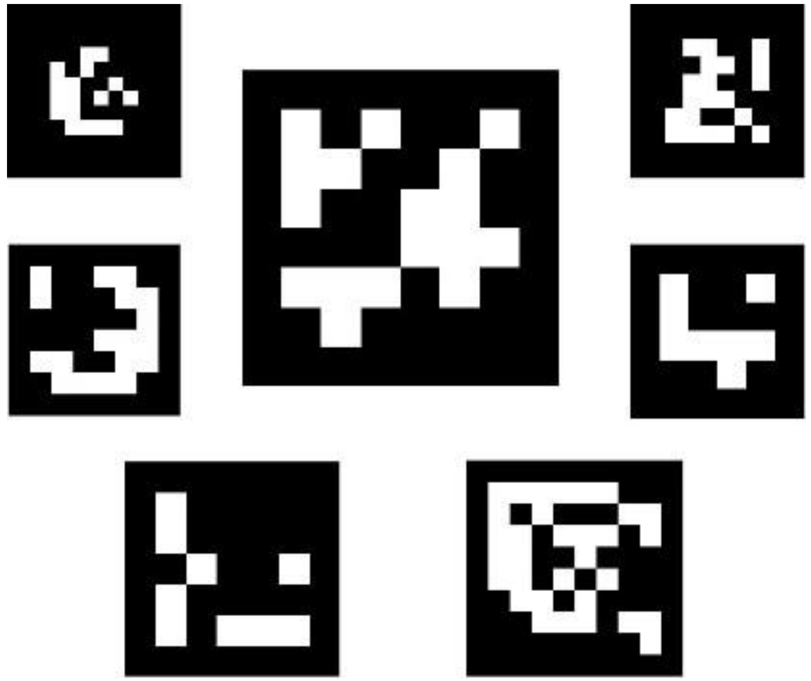
1. NIKISHAEV, A.„Shape Context descriptor and fast character recognision“. *Medium, 2018*. [online]. Available: Shape Context descriptor and fast characters recognition | by Andrey Nikishaev | Machine Learning World | Medium
2. PATEL, A. „What is Object Detection?“. *Medium, 2020*. [online]. Available: What is Object Detection?. Computer Vision Object detection… | by Ashish Patel | ML Research Lab | Medium

Practical Application: ArUco Marker Detection

Applying the vision pipeline to detect, identify, and locate square binary markers in the real world

What is ArUco Marker

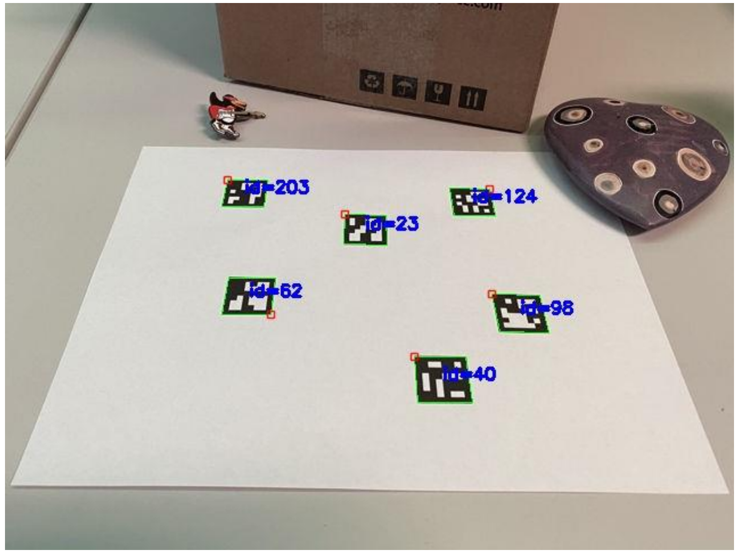
- **Black-and-white square** designed for easy detection by computer vision systems
- Each marker contains:
  - A **black border** (used for shape detection)
  - A **binary inner grid** encoding a unique ID (4x4, 5x5, 6x6 bits)
- Markers are grouped into **dictionaries** (e.g. *DICT_5X5_100)*

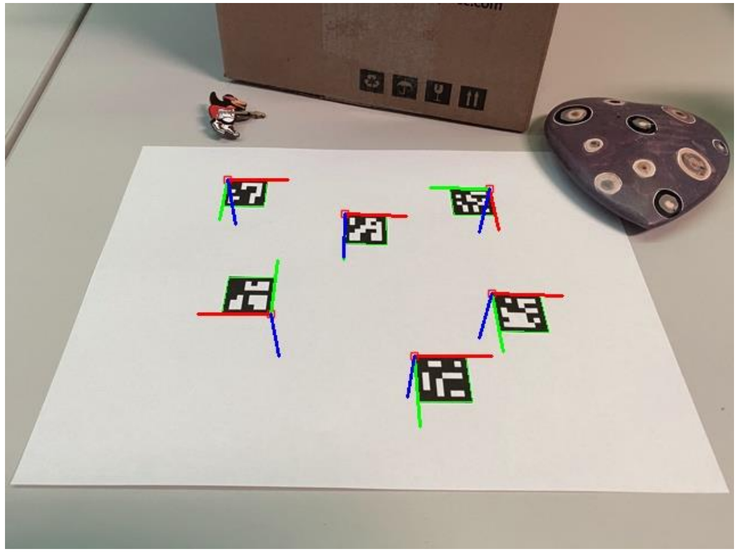


- **Advantages:**
  - Simple and lightweight
  - Fast to detect
  - Robust to perspective, lighting and partial occlusion
  - Provide **Object ID**, and **Precise location and orientation**

ArUco Detection Process

- Step-By-Step Detection Process:
  1. Capture Image
  2. Preprocess (convert to grayscale, apply thresholding)
  3. Find contures – locate square shapes
  4. Check corner order – make sure it‘s properly oriented
  5. Warp the square – to get front-facing view
  6. Split the marker into cells – read the binary values
  7. Compare agains dictionary – decode the marker ID
  8. (Optional) Estimate pose





ArUco Detection Process – OpenCV example

- Step-By-Step Detection Process:
  1. Load the input image or camera frame
  2. Convert the image to grayscale

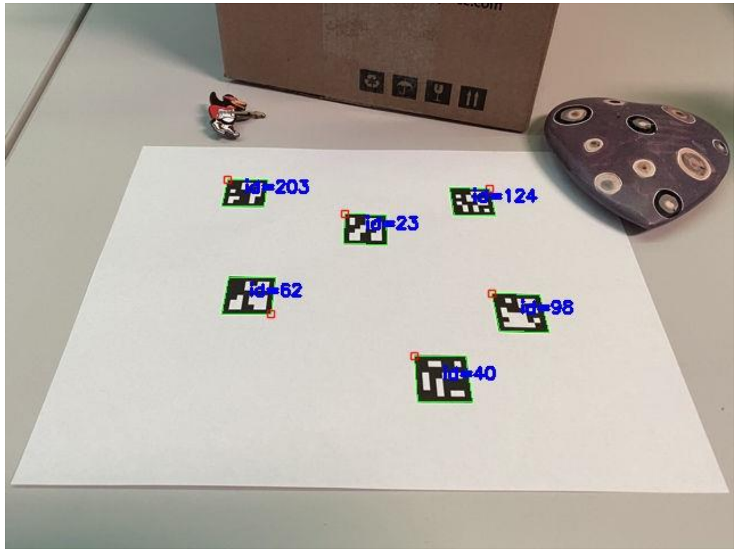
     cv::cvtColor(frame, gray, cv::COLOR_BGR2GRAY);
  3. Get the ArUco dictionary

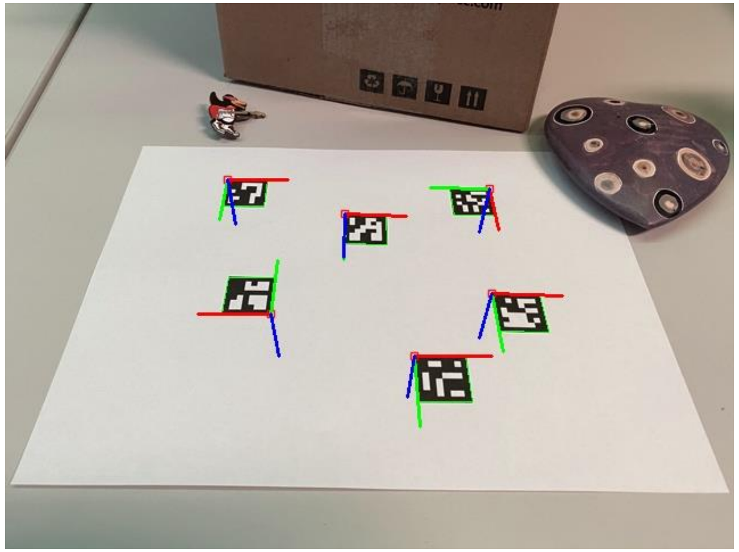
     auto dictionary = cv::aruco::getPredefinedDictionary(cv::atuco::DICT_4x4_50);
  4. Detect Markers

     cv::aruco::detectMarkers(gray, dictionary, corners, ids);
  5. Print the detected marker IDs
  6. (Optional) Draw markers on the image

     cv::aruco::drawDetectedMarkers(frame, croners, ids);
  7. (Optional) Estimate the pose of each marker

     cv::aruco::estimatePoseSingleMarkers(corners, markerLength, cameraMatric, distCoeffs, Rvecs, tvecs);





Summary

- **Computer vision – its challenges and applications**
- Standard Vision Pipeline
  - Sensors (CMOS, CCD, other types)
  - Image Preprocessing
  - Image Segmentation
  - Feature Extraction
  - Final Step
- ArUco marker - ID detection and Pose estimation

Summary

- Computer vision – its challenges and applications
- **Standard Vision Pipeline**
  - Sensors (CMOS, CCD, other types)
  - Image Preprocessing
  - Image Segmentation
  - Feature Extraction
  - Final Step
- ArUco marker - ID detection and Pose estimation

Summary

- Computer vision – its challenges and applications
- Standard Vision Pipeline
  - Sensors (CMOS, CCD, other types)
  - Image Preprocessing
  - Image Segmentation
  - Feature Extraction
  - Final Step
- **ArUco marker - ID detection and Pose estimation**

Ing. Petr Šopák
xsopak00@vutbr.cz

Brno University of Technology
Faculty of Electrical Engineering and Communication
Department of Control and Instrumentation

Robotics and AI Research Group