DRIVES

Development and Research on Innovative Vocational Education Skills

COMPUTER VISION FUNDAMENTALS

U2.E1.COMPUTER VISION OVERVIEW

Computer Vision Expert

May 2021, Version 1

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

LEARNING OBJECTIVES

The student is able to …

CVE.U2.E1.PC1	Define Computer Vision.
CVE.U2.E1.PC2	Understand the functioning of the human visual system.
CVE.U2.E1.PC3	Understand the origin, need and purpose of computer vision.
CVE.U2.E1.PC4	Understand the digital imaging process.
CVE.U2.E1.PC5	Understand how computer vision fits into artificial intelligence.
CVE.U2.E1.PC6	Identify different application domains of computer vision.

WHAT IS COMPUTER VISION?

DRIVES
Development and Research on Innovative Vocational Education Skills

"Field of study surrounding how computers see and understand digital images and vídeos"

"Field of artificial intelligence that trains computers to interpret and understand the visual world."

"It is like imparting human intelligence and instincts to a computer."

Computer Vision is a branch of computer science that specializes in developing digital systems that can process, analyze and make sense of images or videos in a similar way to human beings.
The conception of computer vision is based on the instruction of computers to process an image at a pixel level as well as understand it.

EVERY PICTURE TELLS A STORY

DRIVES
Development and Research on Innovative Vocational Education Skills

COMPUTER VISION GOAL: DEVELOP PROGRAMS THAT CAN INTERPRET IMAGES

What kind of scene?

Where are the cars?

How far is the entrance door?

COMPUTER VISION WORKS IN THREE BASIC STEPS:

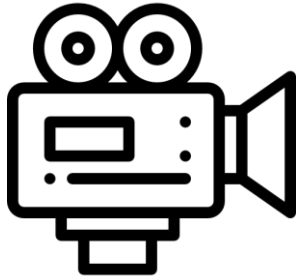


Image Acquisition

Images can be acquired in real-time through videos, photos or 3D technologies.

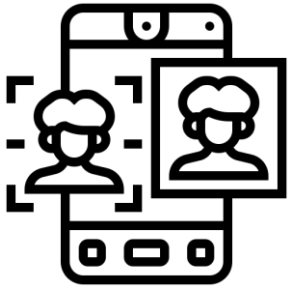


Image Processing

Use of Deep learning models or training models.

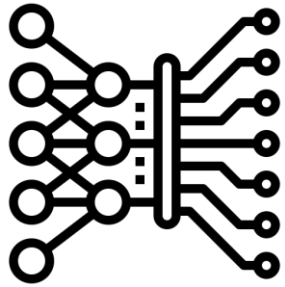


Image Understanding

Interpretation, identification or classification of an object.

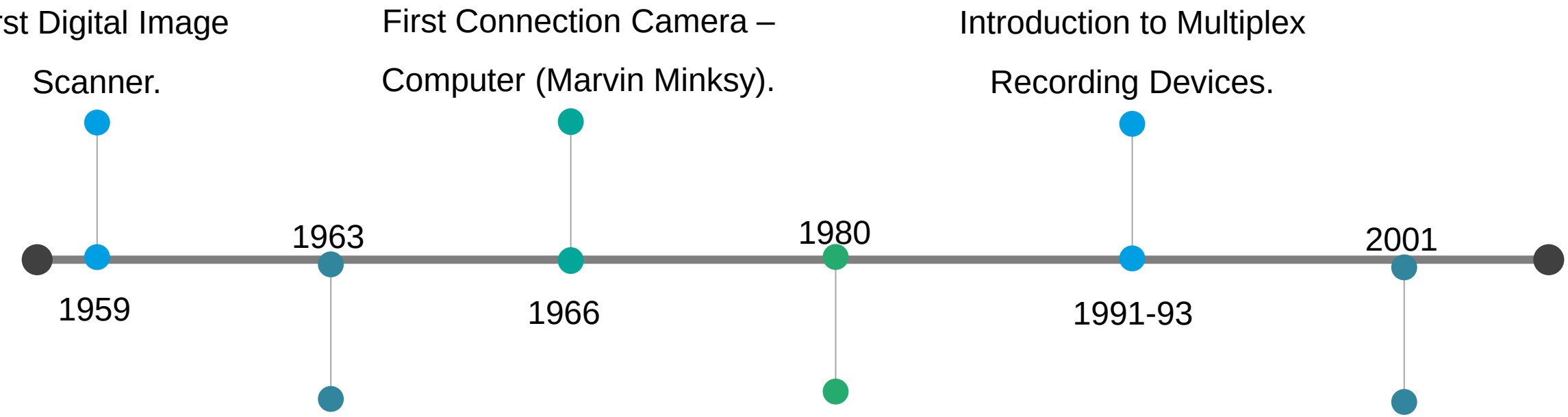
HUMAN VISION VS COMPUTER VISION

HUMAN VISION	FUNCTION	COMPUTER VISION
The luminosity control that penetrates the eye is carried out by the pupil and the iris.	**IMAGE SENSOR**	The luminosity control is done by a sensor.
The focus and projection of the image is performed by the retina.	**IMAGE TRANSFER**	A digitizer produces images from discrete intensity values.
The conversion of images to neutral signals (understood by the brain) is done by cones (color) and rods (contrast).	**IMAGE PROCESSOR**	Image Processing is carried by a algorithm.

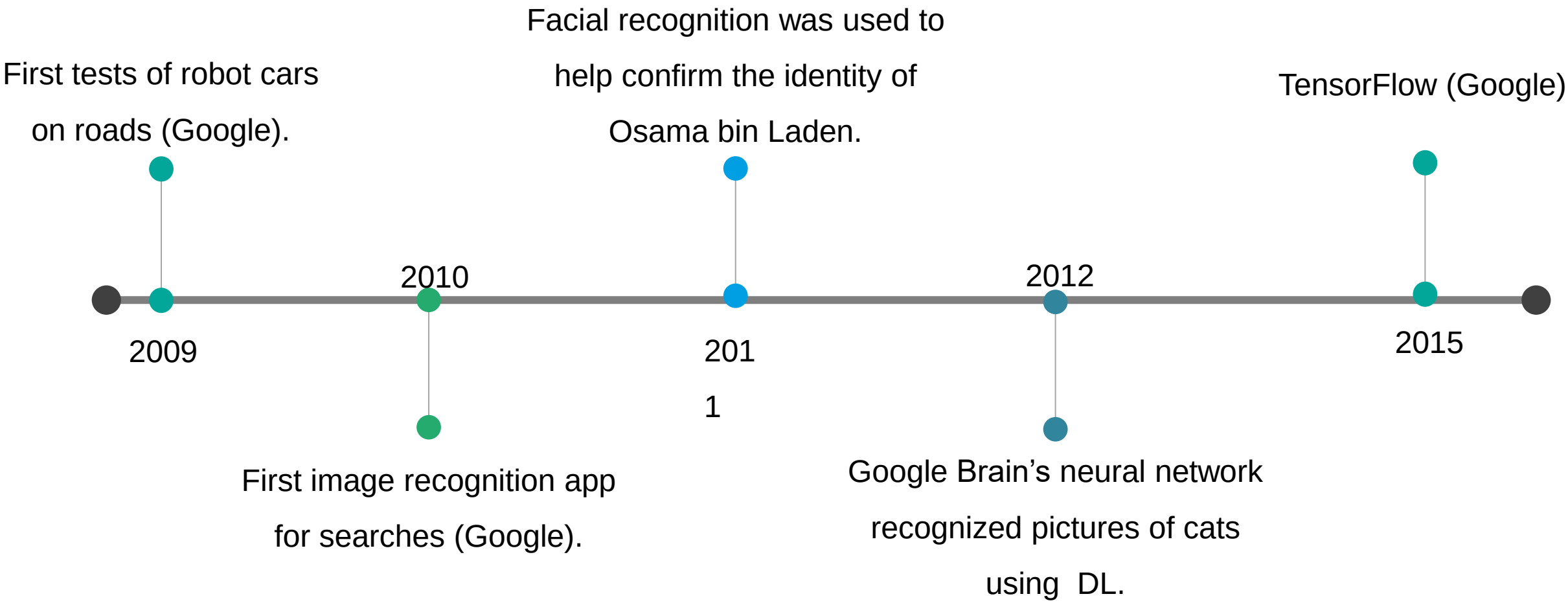
HUMAN VISION VS COMPUTER VISION

HUMAN VISION	FUNCTION	COMPUTER VISION
The front of the brain is managed in a short time and the back of the brain processes the visual information.	**STORAGE**	Frame Buffers are used short-term and archive storage for optical/magnetic disks.
The brain displays the image, but speech is necessary to express what we see.	**DISPLAY**	Images are recorded and displayed through outlets like monitors, tvs, etc.

HISTORY OF COMPUTER VISION



HISTORY OF COMPUTER VISION



HISTORY OF COMPUTER VISION

Google DeepMind's AlphaGo algorithm beat the world Go champion.

2016

Apple released the iPhone X, advertising face recognition as one of its primary new features.

2017

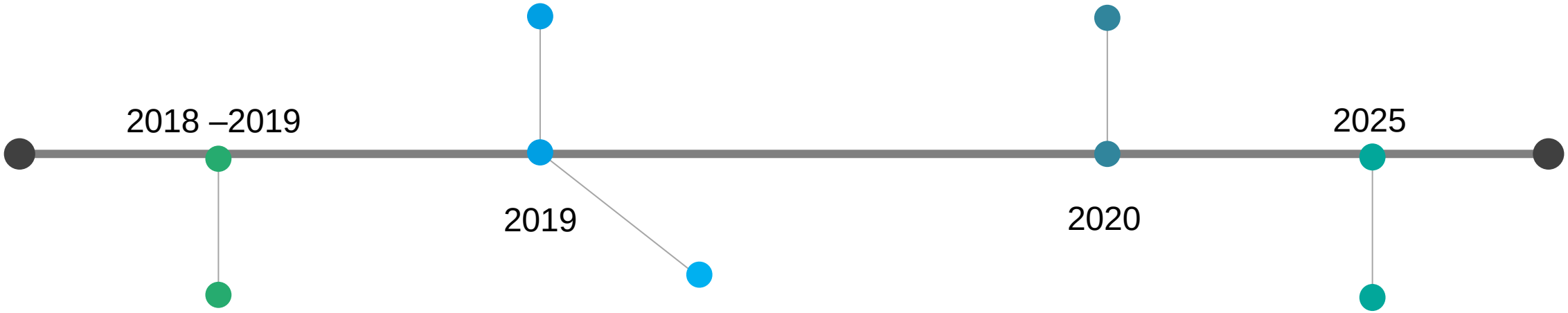
Waymos use Uber for allegedly stealing trade secrets.

Amazon sold its real time face recognition system Rekognition to police departments.

2018

Alibaba's AI model scored better than humans in a Stanford University reading and comprehension test.

HISTORY OF COMPUTER VISION



The Indian government announced a facial recognition plan allowing police officers to search images through mobile app.

The US added four of China's leading AI start-ups to a trade blacklist.

Use of automatic facial recognition technology is lawful.

Launch the Intel Xe graphics card pushing into the GPU market.

Regulation in FR will significantly diverge between China and US/Europe.

WHY WE NEED COMPUTER VISION ?

Safety

Enables self-driving cars to make sense of their surroundings

Health

Automate tasks such as detecting cancerous moles in skin images

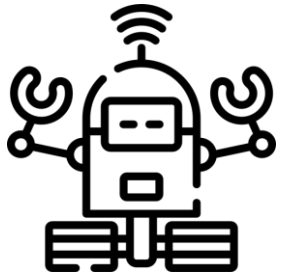
Security

Plays an important role in facial recognition

WHY WE NEED COMPUTER VISION ?

Fun

Plays an important role in augmented and mixed reality



Comfort

Enables the digital world to interact with the physical world.

The importance of Artificial Intelligence is in the problems it can solve.

Computer Vision Became an Important Manifestation of Artificial Intelligence

- It does not predict;
- It gives the outcomes depending upon the present content of images;
- It is able to gain more accuracy with the passage of time;
- Data collecting process is easier;
- The library of images and videos are growing daily;

IMAGE CLASSIFICATION

Objective: Categorize a entire image into a class such as “people”, “animals”, “outdoors”.

***Input*:** An image with a single object (example: a photograph).

***Output*:** A class label (exemple: one or more integers that are mapped to class labels).

OBJECT LOCATION

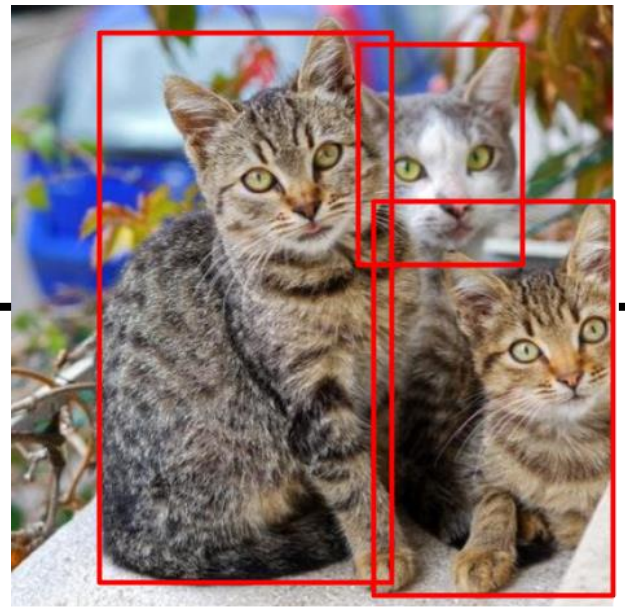
CAT

Objective: Locate the presence of objects in an image and indicate their location with a bounding box.

***Input*:** An image with one or more objects, such as a photograph.

***Output*:** One or more bounding boxes (e.g. defined by a point, width, and height).

OBJECT DETECTION

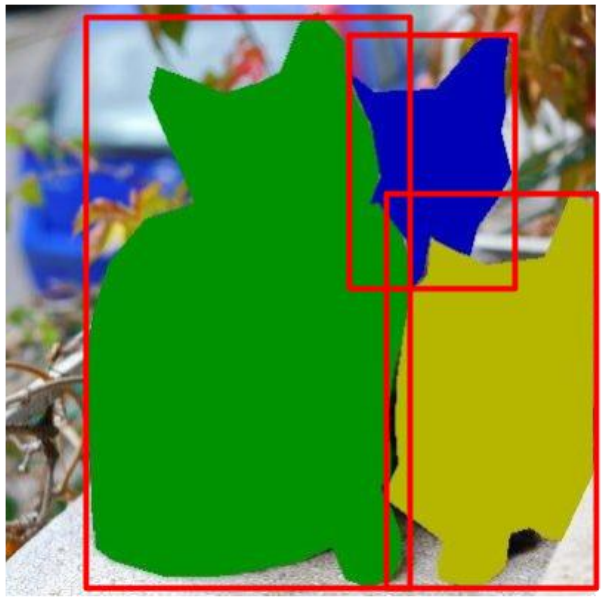


Objective: Locate the presence of objects with a bounding box and predict types or classes of the located objects in an image.

Input: An image with one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g. defined by a point, width, and height) and a class label for each bounding box.

SEGMENTATION



Objective: Identify parts of the image and understanding what object they belong to.

It is again divided into the following categories:

1. **Semantic segmentation**
2. **Instance segmentation**
3. **Panoptic segmentation**

KEY POINT DETECTION

Keypoint detection involves detecting people and localizing their key points simultaneously.

Keypoints are spatial locations or points in the image that define what is interesting or what stands out in the image. They are invariant to image rotation, shrinkage, translation, distortion, and so on.

ADVANTAGES OF COMPUTER VISION

1. **Simpler and faster processes:** You will be able to check your products faster, as protracted visual checks are replaced by fast computers.

2. **Reliability:** Contrary to a human eye, cameras and computers never get tired. The human factor is eliminated, you will not notice any fluctuations in reliability based on how your controllers slept that day or what day of the week it is.

3. **Accuracy:** Your final products will be, thanks to computer imaging, flawless.

ADVANTAGES OF COMPUTER VISION

4. **Wide range of use:** Ranging from factories, through banks to the medical industry. You can use the same system for various activities performed in your production company (from quality checks through warehouse supply tracking to counting of final deliveries in the shipping process), which reduces costs for continual new system staff training.

5. **Reduction of costs:** You will save time of people as well as machines and eliminate faulty products. You will be able to shunt your staff from the control department to other activities and improve your image by customers through the reduction of error rate in deliveries.

LIMITATIONS OF COMPUTER VISION

1. **Necessity of specialists**: there is a huge necessity of specialist related to the field of Machine Learning and Artificial Intelligence. A professional that knows how those devices work and take full advantage of Computer Vision. Also, the person can repair them when necessary. There are a lot of work opportunities after doing a Master in Artificial Intelligences. However, companies still wait for those specialists.

2. **Spoiling**: eliminate the human factor may be good in some cases. But when the machine or device fails, it doesn't announce or anticipate that problem. Whereas a human person can tell in advance when the person won't come.

LIMITATIONS OF COMPUTER VISION

3. **Failing in image processing**: when the device fails because of a virus or other software issues, it is highly probable that Computer Vision and image processing will fail. But if we do not solve the problem, the functions of the device can dissapear. It can froze the entire production in the case of warehouses.

4. **Need for regular monitoring** – What if a computer vision system breaks down or has a technical glitch? To ensure that doesn't happen, companies have to get a dedicated team onboard for regular monitoring and evaluation.

DIGITAL IMAGE PROCESSING

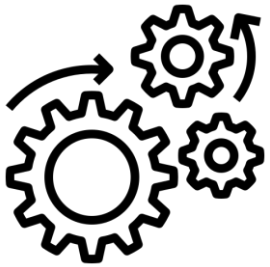


IMAGE PROCESSING

Low-Level Process:

- Reduce Noise
- Constrast Enhancement
- Image Sharpening

IMAGE ANALYSIS

Mid-Level Process:

- Segmentation
- Classification

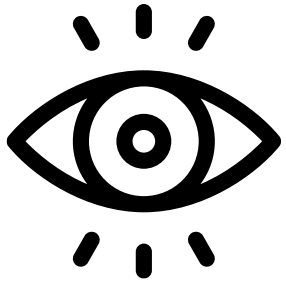
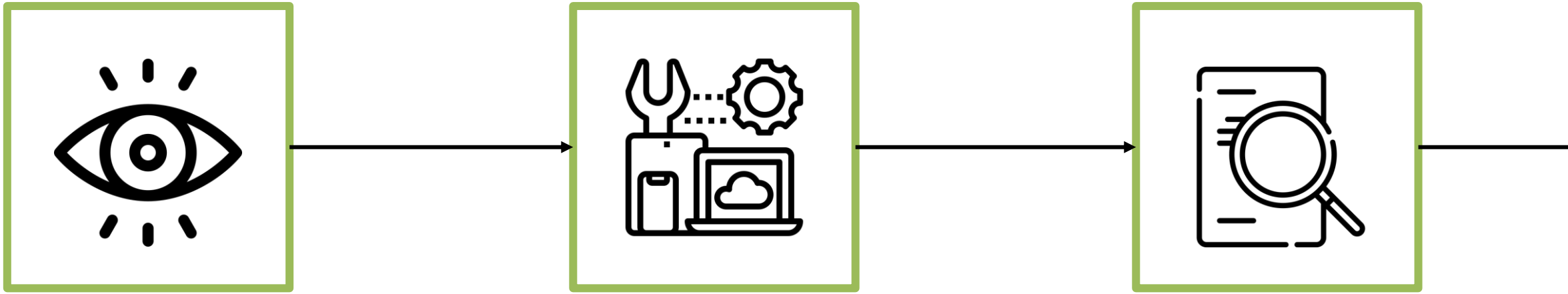


IMAGE VIZUALIZATION

High-Level Process:

- Object Recognition

The purpose of image processing is divided into 5 groups:



Visualization

Observe the objects that are not visible.

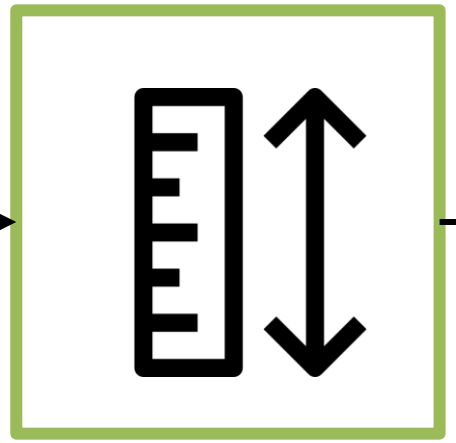
Image sharpening and restoration

To create a better image.

Image retrieval

Seek for the image of interest.

The purpose of image processing is divided into 5 groups:



Measurement of pattern

Measures various objects in an image.

Image Recognition

Distinguish the objects in an image.

DIGITAL IMAGE PROCESSING: STEP BY STEP

Image processing include the following steps:

1 Image Acquisition

This step can be as simple as acquire in an image that is already in digital form.

2 Image Enhancement

The main idea of this second step is to bring out details that are hidden.

3 Image Restoration

This step also deals with image appearance improving with mathematic models of image degradation

4 Color Image

The color image processing includes color modeling in a digital domain.

5 Wavelets and Multiresolution

Wavelets are the foundation for representing images in various degrees of resolution

6 Compression

Compression deals with techniques for reducing the storage required to save an image

DRIVES
Development and Research on Innovative Vocational Education Skills

Image processing include the following steps:

7 Morphological

Deals with tools for extracting image components that are useful in the representation and description of shape

8 Image Segmentation

Segmentation procedures partition an image into its constituent parts or objects.

9 Image Representation and Description

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself.

10 Object Recognition

The color image processing includes color modeling in a digital domain.

11 Knowledge Base

Wavelets are the foundation for representing images in various degrees of resolution

DRIVES
Development and Research on Innovative Vocational Education Skills

Computer vision is important in many different ways:

Image segmentation partitions an image into multiple regions or pieces to be examined separately.

Facial recognition is an advanced type of object detection that recognizes a human face in an image and identifies a specific individual.

Edge detection is a technique used to identify the outside edge of an object or landscape to better identify what is in the image.

Object detection identifies a specific object in an image. Advanced object detection recognizes many objects in a single image: a football field, an offensive player, a defensive player, a ball and so on. These models use an X,Y coordinate to create a bounding box and identify everything inside the box.

DRIVES
Development and Research on Innovative Vocational Education Skills

Computer vision is important in many different ways:

Pattern detection

Process of recognizing repeated shapes, colors and other visual indicators in images.

Image classification

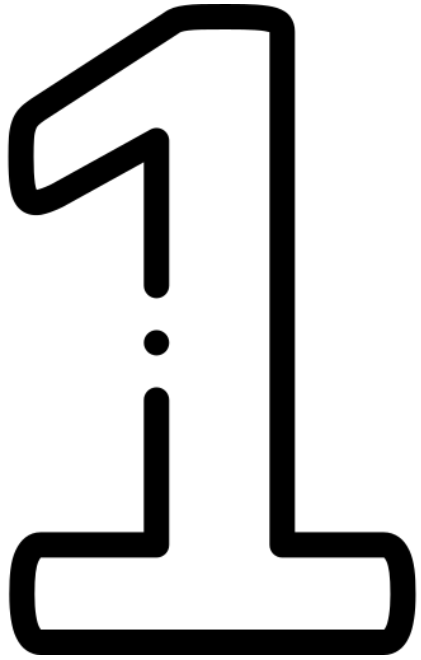
Groups images into different categories.

Feature matching

Type of pattern detection that matches similarities in images to help classify them.

Simple Applications

Simple applications of computer vision may only use one of these techniques, but more advanced uses, like computer vision for self-driving cars, rely on multiple techniques to accomplish their goal.



Retail

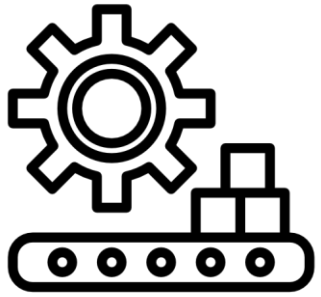
- Enhance the shopping experience;
- Increase loss prevention;
- Detect out-of-stock shelves.

Computer vision is already helping customers checkout more quickly – aiding using self-checkout machines or combining with machine learning to alleviate the checkout process completely.

Manufacturing

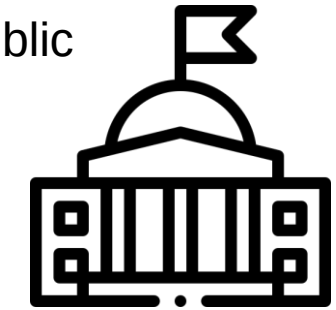
- Identify product defects in real time. A

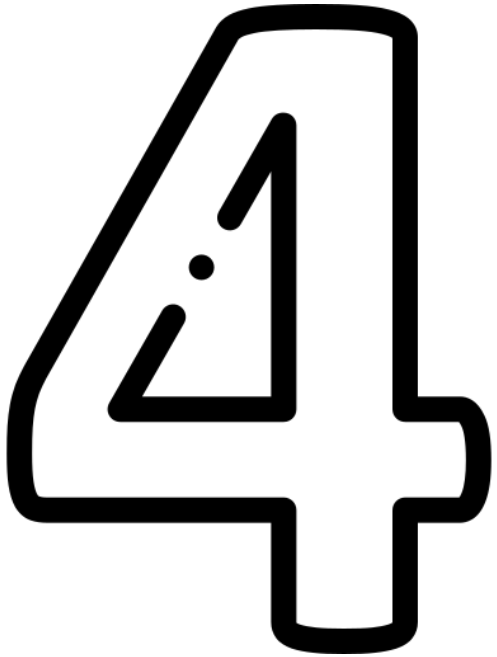
As the products are coming off the production line, a computer processes images or videos, and flags dozens of different types of defects.



Government

- Understand the physical condition of assets under their control;
- Help agencies perform predictive maintenance by analyzing equipment and infrastructure images to make better decisions on which of these require maintenance.
- Help monitor compliance with policies and regulations.
- As drones become used for more defense and homeland security needs, the use of analytics to identify and analyze critical elements from the visual feed will rise to the forefront of computer vision use cases in the public sector.





Healthcare

- Computer vision systems thoroughly examine imagery from MRIs, CAT scans and X-rays to detect abnormalities as accurately as human doctors.
- Medical professionals also use neural networks on three-dimensional images like ultrasounds to detect visual differences in heartbeats and more.

5

Insurance

- Companies use computer vision to conduct more consistent and accurate vehicle damage assessments.

The advancement is reducing fraud and streamlining the claims process.

SUMMARY

- Computer Vision is a branch of computer science that specializes in developing digital systems that can process, analyze and make sense of images or videos in a similar way to human beings.
- The goal of computer vision is develop programs that can interpret images.
- Computer Vision works in three steps: image acquisition, image processing and image understanding.
- The tasks of computer vision are image classification, object location, object detection, segmentation, and key point detection.
- The advantages are simpler and faster processes, reliability, accuracy, wide range of use and Reduction of costs.
- The limitations are necessity of specialists, spoiling, failing in image processing, and need for regular monitoring.
- Digital image processing has 11 steps.
- Computer vision can be applied in several domains like **Retail, Healthcare, Insurance, Government and manufacturing.**

REFERENCE TO AUTHORS

Regina Sousa

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0002-2988-196X

Ana Luísa Sousa

- PhD student in Information System and Tecnologies
- Research Collaborator of the Algoritmi Research Center

0000-0001-5731-3583

Diana Ferreira

- PhD student in Biomedical Engineering
- Research Collaborator of the Algoritmi Research Center

0000-0003-2326-2153

REFERENCE TO AUTHORS

António Abelha

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0001-6457-0756

José Machado

- Associate Professor with Habilitation at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-4121-6169

Victor Alves

- Assistant Professor at the University of Minho
- Integrated Researcher of the Algoritmi Research Center

0000-0003-1819-7051

REFERENCE TO AUTHORS

This Training Material has been certified according to the rules of **ECQA – European Certification and Qualification Association.**

The Training Material was developed within the international job role committee **“Computer Vision Expert”:**

UMINHO – University of Minho (https://www.uminho.pt/PT)

The development of the training material was partly funded by the EU under Blueprint Project DRIVES.

DRIVES

Development and Research on Innovative Vocational Education Skills

Thank you for your attention

DRIVES project is project under **The Blueprint for Sectoral Cooperation on Skills in Automotive Sector**, as part of New Skills Agenda.

The aim of the Blueprint is **to support an overall sectoral strategy and to develop concrete actions to address short and medium term skills needs.**

Follow DRIVES project at:

More information at:

www.project-drives.eu

Co-funded by the Erasmus+ Programme of the European Union

The Development and Research on Innovative Vocational Educational Skills project (DRIVES) is co-funded by the Erasmus+ Programme of the European Union under the agreement 591988-EPP-1-2017-1-CZ-EPPKA2-SSA-B. The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.