PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
University of 8 Mai 1945 – Guelma
Faculty of Science and Technology
Department of Electro-technical and Automation Engineering

Ref:……………/2024

DISSERTATION

Presented for obtaining the **Academic MASTER diploma**

Domain: Science and Technology

Branch: Automation

Specialty: Automation and Industrial Computer Science

By: REZAIGUIA Oussama

Theme

Study and Implementation of a Mobile Robot for Object Detection Based on VLC Technology

Publicly defended, on 23/06/2024, in front of the jury composed of:

Mr. BABOURI Abdesselam	Professor	Univ.8Mai-Guelma	President / supervisor
Mme.BOUCERREDJ Leila	MCA	Univ.8Mai-Guelma	Examiner Principal
Mr. DEBECHE Mehdi	MAA	Univ.8Mai-Guelma	Examiner
Mr. CHAABNA Ameur	MCB	Univ.USTHB-Alger	Co-Supervisor
Mr. CHOUABIA Halim	Dr	Univ.8Mai-Guelma	Co-Supervisor

University Year : 2023/2024

ACKNOWLEDGMENTS

The beginning is praise and thanks to Allah, for without His success, I would not have been able to accomplish this work.

I extend my heartfelt thanks to my supervisor, **Pr. Babouri Abdesselam** for proposing this topic, and for providing me with the means, which enabled me to complete this study, and for his valuable advice and encouragement.

I would also like to express my gratitude to **Dr.Chaabna Ameur** and **Dr.Chouabia Halim**, for guiding and assisting me in my dissertation research.

I would like to extend my heartfelt thanks to my parents, my brother, my sister and all my family especially **Abdessalam Medjelekh**.

I would like to express my sincere gratitude and appreciation to my friends **Islem Belhadef** and **Mohamed Oussama Bounab** for their assistance as I wish them all the success and excellence in their academic journey.

I am grateful to everyone who has assisted me, particularly the professors of the Computer Science department, and I must express my sincere appreciation for my English teacher, Mr. **Bilal Hasnaoui**.

Abstract

In this work, we studied and addressed artificial intelligence for robots. We used advanced algorithms based on deep learning and neural networks with the Python language and its libraries to train our model for fire detection. We used a Raspberry Pi as a small-embedded processor in the robot for operation, and Arduino as a microcontroller unit to establish communication with the robot through visible light communication technology, enabling us to send detection results and control the robot movement.

Keywords: Artificial Intelligence, Robot, Algorithms, Deep learning, Neural Network, Fire Detection, Processor, Microcontroller, Visible Light Communication.

Résumé

Dans ce travail, nous avons étudié et traité l'intelligence artificielle des robots. Nous avons utilisé des algorithmes avancés basés sur l'apprentissage profond et les réseaux neuronaux avec le langage Python et ses bibliothèques dédiées pour entraîner notre modèle à détecter les incendies. Nous avons aussi utilisé le Raspberry Pi comme un petit processeur intégré dans le robot pour l'exécution, et l'Arduino comme un microcontrôleur précis pour établir une connexion avec le robot via la technologie de communication par lumière visible nous permettant d'envoyer les résultats de détection et de contrôler le mouvement du robot.

Mots-clés: Intelligence Artificielle, Robot, Algorithmes, Apprentissage Profond, Réseau de Neurones, Détection d'incendie, Processeur, Microcontrôleur, Communication par Lumière Visible.

الملخص

في هذا العمل، قمنا بدراسة ومعالجة الذكاء الإصطناعي للروبوتات و ذلك باستخدام خوارزميات متقدمة تعتمد على التعلم العميق والشبكات العصبية مع لغة بايثون ومكتباته في تدريب النموذج الخاص بنا على كشف الحرائق. واستعملنا جهاز الراسبيري باي كمعالج صغير مدمج في الروبوت للتشغيل، و الأردوينو كوحدة تحكم دقيقة لإجراء اتصال مع روبوت عن طريق تكنولوجيا الإتصال بالضوء المرئي لنتمكن من إرسال نتائج الكشف والتحكم في حركة الروبوت.

الكلمات المفتاحية: الذكاء الإصطناعي، الروبوت، خوارزميات، التعلم العميق، الشبكات العصبية، كشف الحرائق، معالج، وحدة تحكم دقيقة، الإتصال بالضوء المرئي.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI: Artificial intelligence

ANN: Artificial neural networks

CNN: Convolutional Neural Network

CPS: Cyber Physical Systems

DL: Deep Learning

DNN: Deep Neural Network

GPS: Global Positioning System

GSM: Global System for Mobile communication

HDMI: High-Definition Multimedia Interface

IDE: Integrated Development Environment

IOU: Intersection over Union

IP: Internet Protocol

LED: Light-Emitting Diode

LI-FI: Light Fidelity

LOS: Line Of Sight

MIMO: Multiple Input Multiple Output

MISO: Multiple Input Single Output

ML: Machine Learning

NLOS: Non-Line Of Sight

NMS: Non-Maximum Suppression

OS: operating system

RF: Radio Frequency

RGB: Red, Green, Blue LEDs

RNN: Recurrent neural networks

SD: Secure Digital

SIMO Single Input Multiple Output

SISO Single Input Single Output

SNN: simulated neural network

SNR: Signal-to-Noise Ratio

TV: Television

VLC: Visible Light Communication

VLP: Visible Light Positioning

VNC: Virtual Network Computing

V2V: Vehicle-To-Vehicle

WLAN: Wireless Local Area Network

YOLO: You Only Look Once

General introduction

In this era, the world is experiencing a significant technological evolution led to the growth of needs, the doubling of efforts and the increase in risks and difficulties. For this reason, humans have sought an invention that can emulate the human mind in its pattern of thinking to solve problems and make decisions, thus creating the world of robots with artificial intelligence that has invaded all sectors. Thanks to deep learning and artificial neural networks, it is now possible to recognize images or even find objects within a picture or video. Object detection technology has become important in many fields, especially in the field of safety, such as fire detection.

With the continuous advancement in the field of artificial intelligence and robotic engineering, the integration of other technological fields becomes necessary to achieve greater developments; among these fields is wireless communications. Most wireless technologies rely on radio waves/radio frequencies (RF) that suffer from interference and high latency issues. As a solution to this problem, an alternative technology has been proposed, which is Visible Light Communication (VLC), offering high-speed data transfer rates, high security, and low interference.

This thesis primarily focuses on studying and creating a mobile robot capable of detecting fires in real-time and sending the detection results to the control room using Visible Light Communication (VLC) technology. In the first chapter, we will discuss the different types of artificial intelligence, mobile robots, and wireless communications, with a focus on VLC technology. In the second chapter, we will talk about motion detection and tracking methods in videos, as well as object detection algorithms, their structure, and principles of operation. In the third chapter, we will conduct a practical application and work on creating a mobile robot capable of performing fire detection tasks, using a Raspberry Pi and a webcam, and we will use Arduino to send the results via VLC technology.

Chapter I : Introduction about Artificial Intelligence, Robot, Wireless Communication and Visible Light Communication

I.1. Introduction

Artificial Intelligence (AI) forms a comprehensive system, where machine learning is considered a branch of it, and deep learning represents a subfield of machine learning. Neural networks form the backbone of deep learning algorithms. In addition, robots rely on artificial intelligence, and with the increase in human activities, robots, especially those that are mobile, have become more important in our lives. On the other hand, wireless communications are widely popular because of their flexibility, and Visible Light Communication (VLC) technology is a significant field of study that attracts scientists' interest due to its importance and potential applications.

I.2. Artificiel intelligence

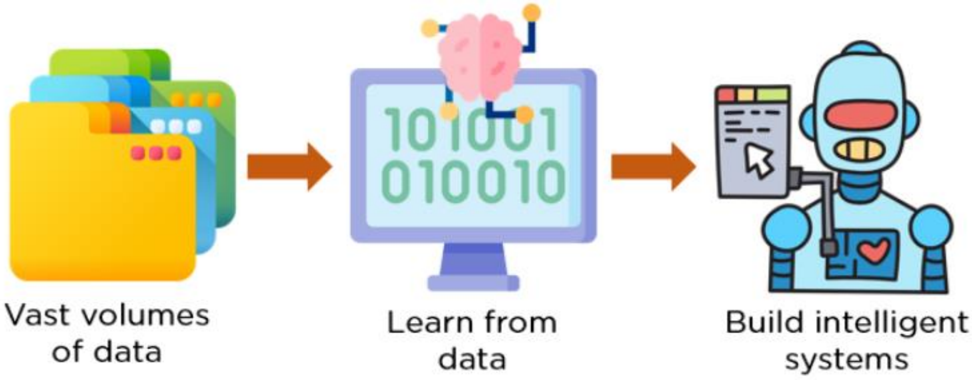

Figure I.1: Artificial intelligence [1].

I.2.1. Artificial intelligence history

In the 1950s, a group of scientists, mathematicians, and philosophers began exploring the notion of artificial intelligence. Among them was Alan Turing, who proposed that humans use available information and reason to solve problems and make decisions, so why cannot machines do the same?

John McCarthy and Marvin Minsky convened the Dartmouth Research Project on Artificial Intelligence (DSRPAI) in 1956, which played a significant role in the advancement of AI research. Over the next few years [2].

I.2.2. What is artificial intelligence

Artificial intelligence is a term used to refer to machines that simulate human intelligence and cognitive functions such as problem solving and learning [3]. That is based on the development and use of algorithms in a dynamic computer environment. Its goal is to enable computers to think and behave like humans. To do this, three components are required:

- Computer systems

- Data with management systems

- Advanced AI algorithms (code) [4].

AI has become a general term for every application that completes complex tasks that already need human assistance [5].

I.2.3. Types of artificial intelligence

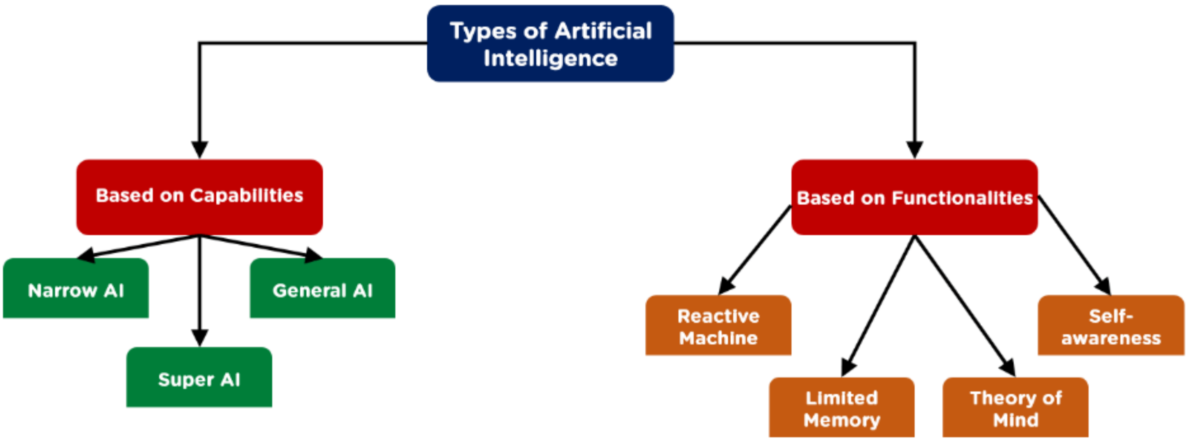

Figure I.2: Type of artificial intelligence [6].

There are two major types of Artificial Intelligence [6] [7]:

Artificial Intelligence based on capabilities which has three types:

• **Narrow AI,** also called weak AI, is focused on specific tasks and has limitations in its capacity.

• **General AI** or strong AI can comprehend and learn any intellectual task that humans can perform.

• **Super AI** is an advanced form of Artificial Intelligence, and it is smarter than humans and good at everything. It is capable of independent thinking, puzzle solving, and decision-making.

Artificial Intelligence based on functionalities that contain four main types:

• **Reactive machines** are designed to perform specific tasks and lack the capability to utilize past data or memory for future actions. They operate solely based on immediate input. For example, the computer simply watches the moves in a game of chess and then chooses the optimal move to win.

• **Limited memory:** this category of machine has the ability to collect previous data and add it to its memory. These machines can make decisions, but they have limited memory. This type is usually used in self-driving cars.

• **Theory of mind:** this type of artificial intelligence can understand emotions and thoughts and interact socially, but it is still not possible to build machines based on this type of technology.

• **Self-aware:** they are the next generation of artificial intelligence, and they will be smart, conscious, and capable. However, it is a theoretical concept that does not exist in practice yet.

I.2.4. Machine learning

Is a subset of artificial intelligence (AI) focused on developing computer algorithms that imitate human learning. Simply, machine learning allows computers to read data and make decisions with a view to improving their accuracy without the need for explicit programming [8].

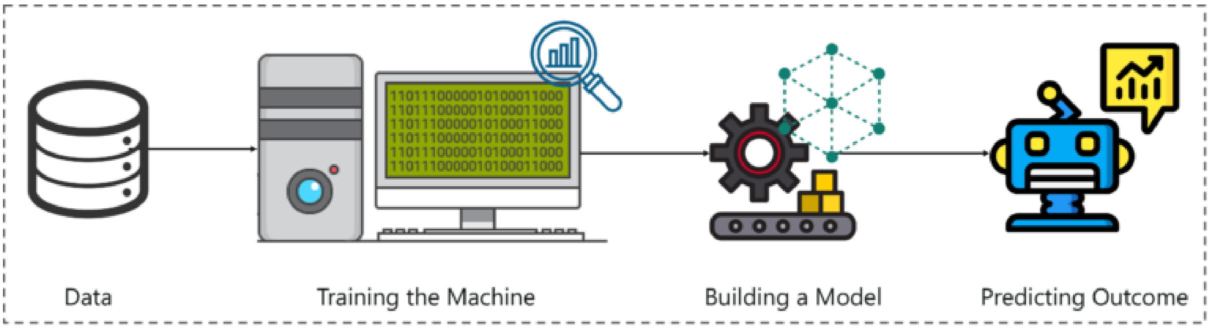

Figure I.3: What is machine learning [9].

Machine learning can be classified into four types based on the nature of the learning system and the data available: supervised learning (where the expected output for the input is known thanks to labelled data sets), unsupervised learning (where the expected outputs are not known because of the use of unlabelled data sets), semi-supervised learning (which usually has a lot of unlabelled data and very little labelled data), and reinforcement learning (where an agent learns to make decisions through interaction with the environment) [10][11][12].

There are three main categories of problems that can be solved using Machine Learning: Regression, Classification, and Clustering [9].

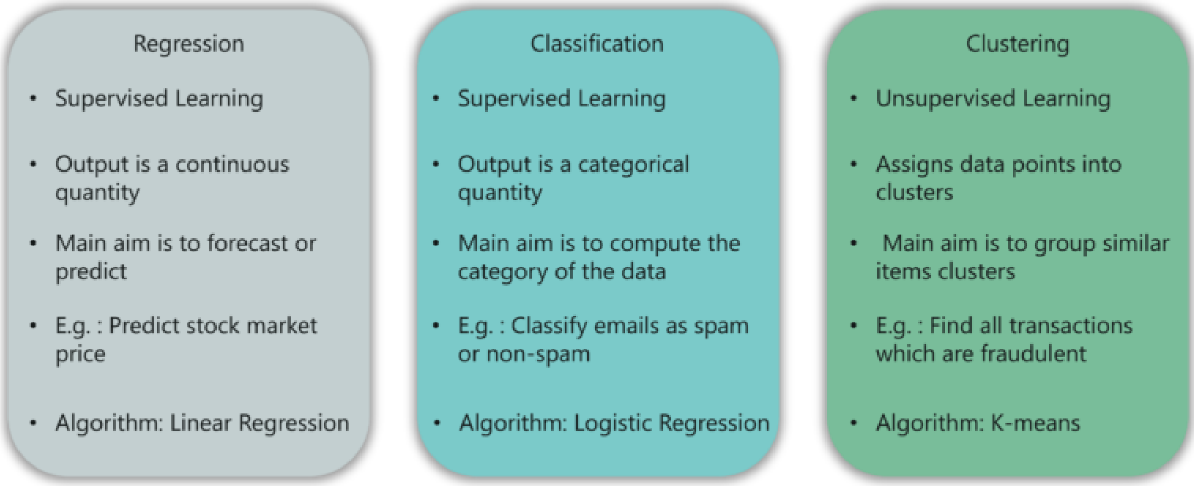

Figure I.4: Regression vs Classification vs Clustering [9].

I.2.5. Neural network

Neural Networks, also known as Artificial Neural Networks (ANNs) or Simulated Neural Networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms [3]. ANN is a computational model that imitates the way nerve cells biological work in the human biological brain [13].

An artificial neural network (ANN) consists of layers of nodes, including an input layer, one or more hidden layers, and an output layer. Every node and artificial neuron is linked to other nodes and artificial neurons, and they are all coupled with thresholds and weights [11].

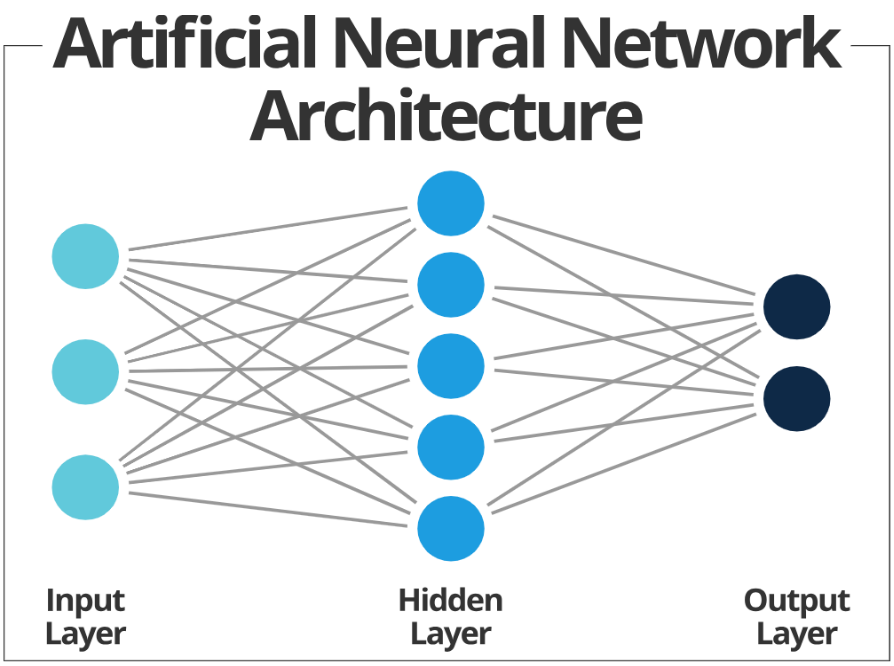

Figure I.5: Artificial neural network architecture [14].

- **Input layer:**

The artificial neural network gets information from the outside through the input layer. Input nodes handle the data, study it, or group it, then send it to the next layer.

- **Hidden layer:**

These layers get information from the input layer or other hidden layers. There can be many hidden layers. The output from the previous layer is analyzed by each hidden layer, processed, and then transferred to a different layer.

- **Output layer:**

The output layer is the result of any data processing. It may consist of one or more nodes. For example, if the problem is binary (yes/no), the output layer will contain one output node that will return a value of 1 or 0. On the other hand, the output layer may have many output nodes if the problem is one of multi-class categorization [11].

- **Recurrent neural networks**

Recurrent neural networks (RNNs) consist of recurrent units connected in a directed cycle. Each unit combines the current input with the previous hidden state, producing output and updating the hidden state for the next step. This enables RNNs to capture patterns and dependencies in sequential data, making them useful for tasks like speech recognition, language modeling, and translation [15].

- **Convolutional neural networks**

The convolutional neural network (CNN) has many hidden layers in addition to input and output layers. CNN's hidden layers usually include a series of convolution layers that convince duplication or other dot products [16].

There are four main types of CNN layers: Convolution Layer, ReLu Layer, Pooling Layer, and Fully Connected Layer [11].

I.2.6. Deep neural network

Deep neural networks (DNN) consist of multiple hidden layers linked together, each one is designed to perform specific transformations on the data. The additional hidden layers in a deep neural network enable it to learn more complex patterns. As a consequence, training a deep neural network takes more time and computer power. (It needs millions of training data examples). [17]

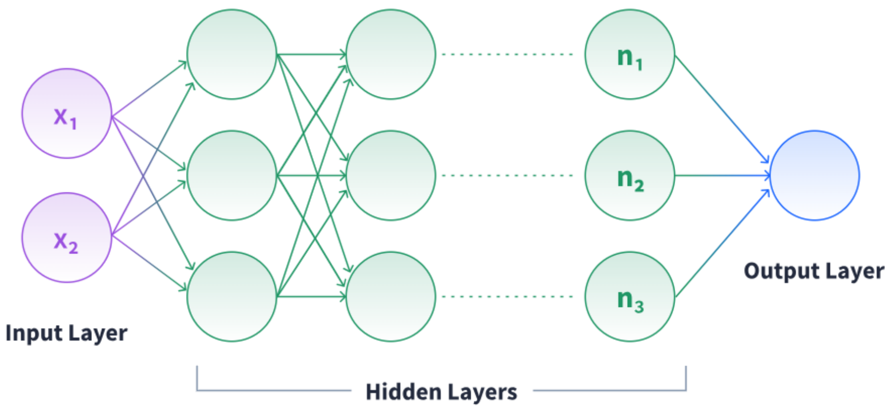

Figure I.6: Architecture of a deep neural network with n-hidden layers [17].

I.2.7. Deep learning

Deep learning (DL) is a subset of machine learning also called hidden layers that trains a computer to perform human-like tasks and uses advanced algorithms to enable an AI system to train itself to perform tasks by exposing multilayered neural networks to vast amounts of data [11][13][18].

I.2.8. ML vs DL vs NN

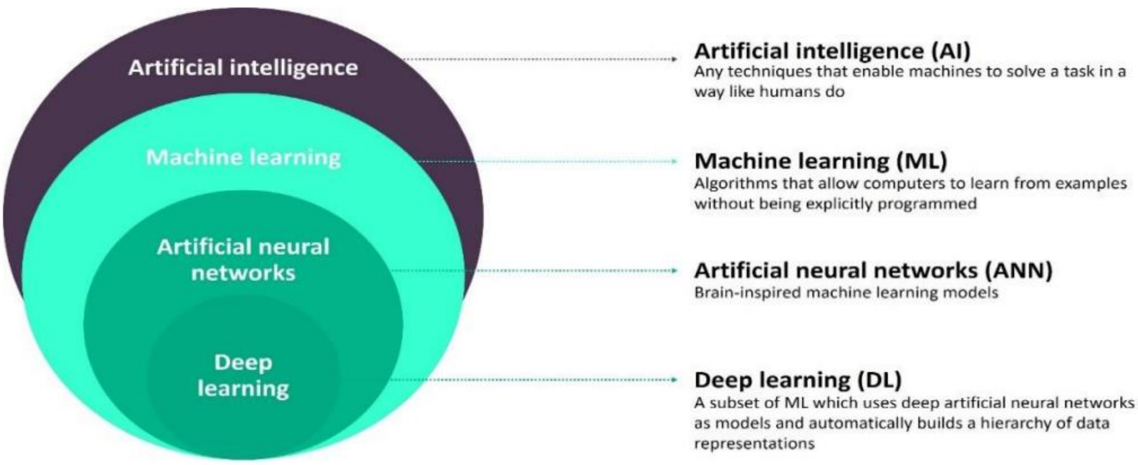

Figure I.7: AI vs ML vs ANN vs DL [8].

The major difference between deep learning and traditional machine learning is the type of data required and the methods applied during the learning process.

ML algorithms are programmed to use structured data for prediction-making, while deep learning algorithms can deal with unstructured data without having the features predefined (features are automatically extracted through some preprocessing steps and analysis before being organized internally into an orderly format) [19].

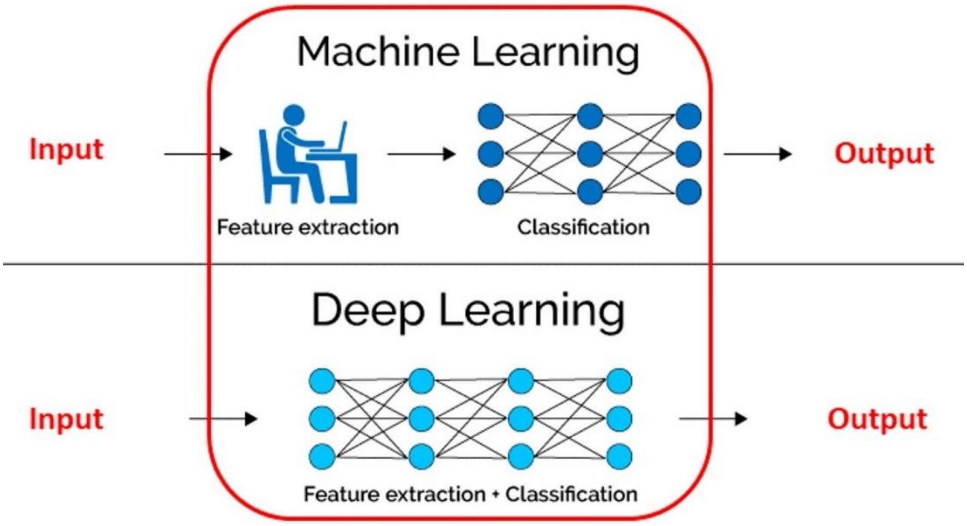

Figure I.8: The difference between ML and deep learning in terms of feature extraction [19].

The impact of data size on performance is another major difference between deep learning and ML: ML methods do not always get better when you give them more data. However, deep learning methods usually become more accurate as they get more data, making them scalable in performance [19].

A neural network with multiple hidden layers and multiple nodes in each hidden layer is known as a deep learning system or a deep neural network.

It gives high performance and takes more time to train them because it is very complicated and they have higher accuracy compared with the Neural Networks [20].

Most deep neural networks are feed-forward, meaning they only flow in one direction from input to output [3].

I.3. Robots

I.3.1. What is robotics?

Robotics is the science that studies the set of techniques allowing the design and creation of automatic machines, or robots. It is an interdisciplinary field that incorporates elements of electronics, computer science, and mechanics [21].

I.3.2. What is a robot?

The word "robot" comes from the Czech word "robota," which means "work".

In contrast to robotics, a robot is specifically described as a complex configuration of electrical and mechanical parts that are all managed by artificial intelligence.

A robot is also understood as a device that, with the assistance of an automatic control system based on a microprocessor, carries out a specific task for which it was built in the realm of industry, science, or home [21].

I.3.3. Types of robots

There are two categories of robots:

I.3.3.1. Robot manipulators (fixed base robots)

The mechanical structure of a robot manipulator consists of a few rigid bodies (links) connected by articulations (joints); a manipulator is characterized by a mobile arm. They enable the safe handling of hazardous materials in areas that are inaccessible, facilitating sensitive research and development while lowering risks. They also enable numerous applications in a variety of fields, including heavy load transportation, robotic surgery in medicine, and the handling and monitoring of radioactive materials in the nuclear industry [12][22].

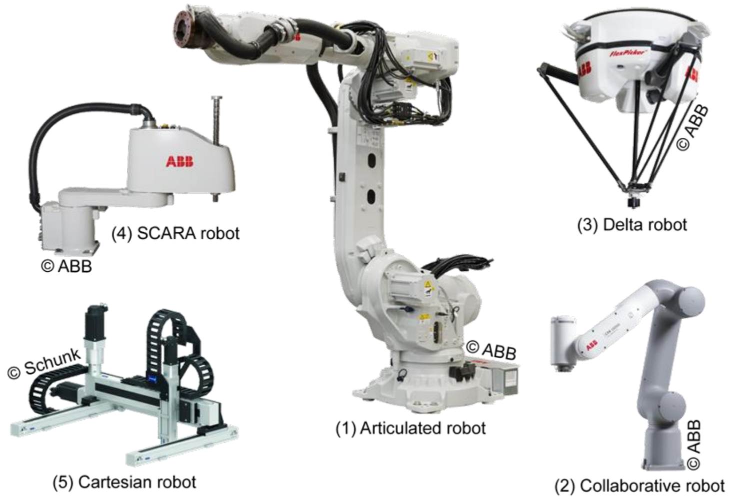

Figure I.9: Examples of different robot manipulators.

I.3.3.2. Mobile robots

A mobile robot is a machine controlled by software that uses sensors and other technology to identify its surroundings and move around its environment. Mobile robots are able to operate on the basis of a combination of AI and mechanical robotic components, e.g., wheels, tracks, or feet. Across different sectors of the economy, mobile robots are growing in popularity. They are used for tasks that are impossible or dangerous to human workers, such as helping with work processes [23].

Figure I.10: Examples of different mobile robot [24].

I.3.4. Types of mobile robots

I.3.4.1. Wheeled mobile robots

Wheeled mobile robots are one of the most widely selected, utilized, and implemented robots because of their simple bending form and structure which depend on the number, design, and function of the wheels. They move in all directions with a certain acceleration and notable speed because of the size and configuration of the wheels. They have the ability to use new technologies to occasionally alter direction and overcome challenges like climbing stairs [22] [25].

Figure I.11: Examples of wheeled robots [26].

I.3.4.2. Tracked mobile robots

When the ground is rough or of poor quality, the wheels face many problems, such as loss of efficiency and interaction with the terrain they interact with. In order to perform tasks, the use of tracks offers the advantages of good ground adherence and the ability to overcome obstacles. Generally, they are used for military purposes [22][25].

Figure I.12: Example of tracked mobile robot [21].

I.3.4.3. Flying robot

Flying robots, also known as drones, are aircraft or flight devices that can be operated by remote control and fly without a pilot or passenger. This type of robot must follow specific regulations. It is characterized by autonomy, speed, and contribution to information after analysis through various sensors. Because of its growing significance in both military and civilian uses, it can reduce pilots' exposure to dangerous environments on certain missions [22] [27].

Figure I.13: Drone.

I.3.4.4. Walking robots

Walking robots are designed to perform various tasks where access to the site is difficult, dangerous, or impossible for humans [28].

There are many mobile walking robots. They differ in their shape and the number of legs and on this basis, it can be classified [22].

- **Monopods**

This type of robot, equipped with a single leg or foot, hops to move due to its sole point of contact with the ground. Dynamic stability is considered to be the most important aspect of balancing the center of gravity and body mass [22].

Figure I.14: Example of monopodial walking robots [26].

- **Bipeds robot (humanoid)**

A humanoid robot is designed with a structure similar to the human form. This design is practical for tasks that involve using tools or navigating environments made for humans, and it is also useful for research, such as understanding how two-legged walking works. Generally, these robots are equipped with a head, torso, two arms, and two legs for functionality. [28]

The development of bipedal robots presents a series of obstacles for researchers, because of the complexity of mathematical models, the computation costs that machines require, and the materials that they use for manufacture [22].

Figure I.15: Example of bipedal walking robots.

- **Quadruped robots**

Their structure is inspired by four-legged animals; they can move easily since their legs function independently of one another. This type can be used for both commercial and military use and is more stable than bipeds [22].

Figure I.16: Example of quadruped walking robots [22].

- **Hexapod robot**

A hexapod robot is a mechatronic device that can be stationary or mobile and uses three pairs of legs to move.

Compared to bipedal and quadruped robots, hexapods tend to be more stable. Therefore, they do not depend on real-time controllers to keep them stable and walking [28].

Figure I.17: Example of hexapod walking robot [24].

- **Crawling robots**

The techniques used are similar to the locomotion methods of crawling animals (like snakes) and are used in tunnel-like environments or confined spaces, and their system consists of a set of modules with multiple mobilities [28].

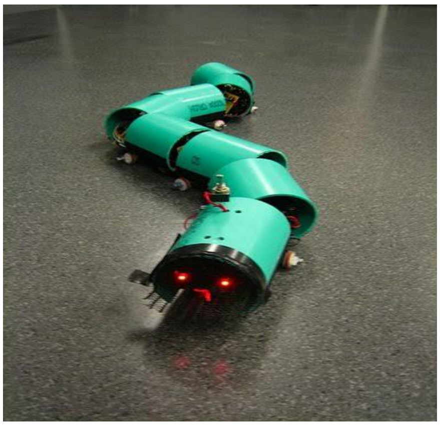

Figure I.18: Example of crawling robot [28].

I.3.4.5. Autonomous underwater vehicle (AUV)

An autonomous underwater vehicle (AUV) is an autonomous mobile robot that can move beneath the water. AUV has a key role to play in the exploration and development of deep-sea marine environment resources. Unlike drones, AUVs have six degrees of freedom, and the physical conditions of the aquatic environment (e.g., viscosity, Archimedean buoyancy) make it difficult to model and control them. In military applications, AUVs are also known as unmanned underwater vehicles (UUVs) [27].

Figure I.19: Sea-explorer underwater robot [27].

I.4. Wireless communication

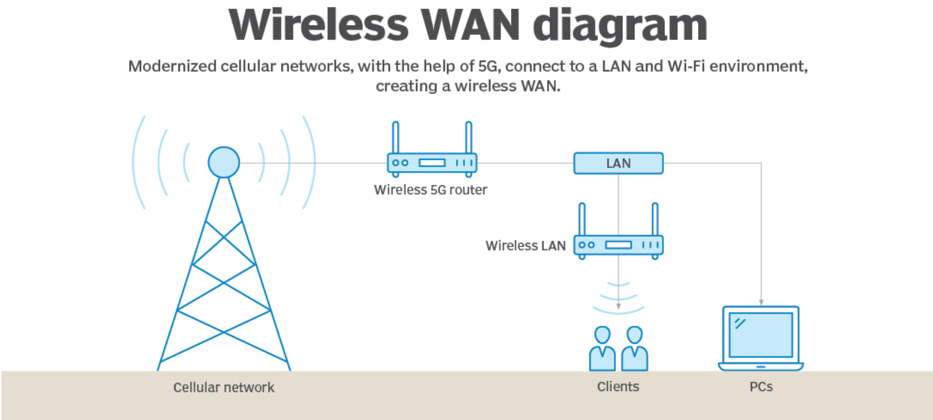

Figure I.20: Wireless system [29].

I.4.1. What is wireless communication?

Wireless communication is the transmission of information between two or more points using an electromagnetic wave, not a physical electrical conductor, as a channel, and it is a fundamental part of our lives. Some of the most common wireless communication systems that we use on a daily basis are mobile phones, GPS, remote control, Bluetooth, WiFi, etc. [30]

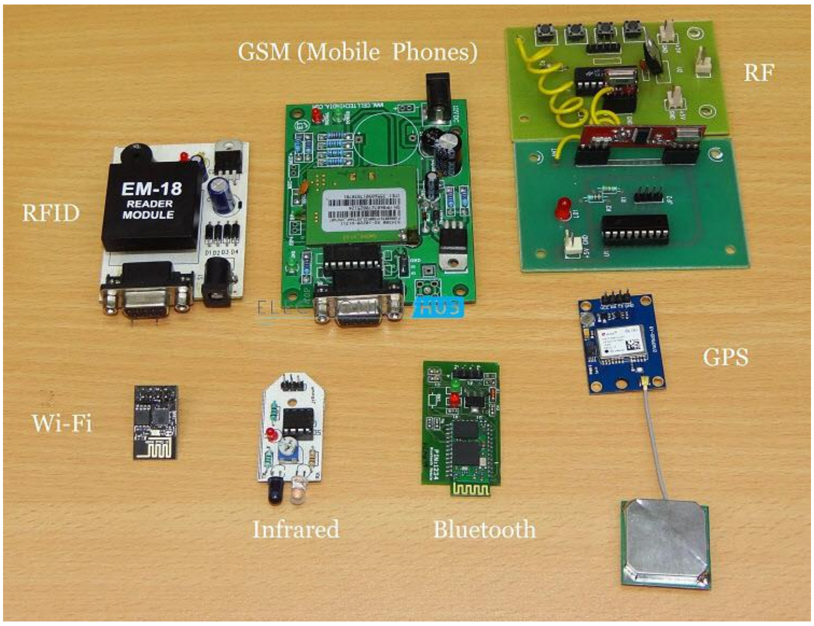

Figure I.21: Wireless type [31].

I.4.2. What is electromagnetic wave?

Electromagnetic waves, or EM waves, are waves produced by vibrations between electric and magnetic fields. [11]

Electromagnetic waves can be classified and arranged according to their various wavelengths/frequencies; this classification is known as the electromagnetic spectrum [33].

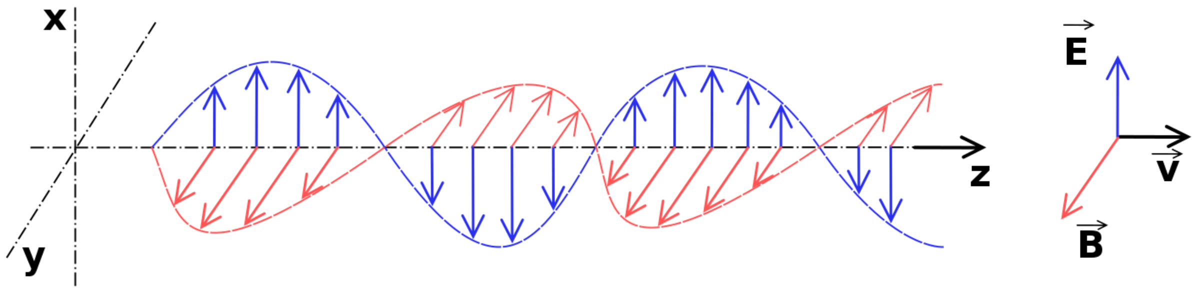

Figure I.22: A diagram showing the electric (blue) and magnetic (red) fields of an electromagnetic wave [32].

I.5. Visible light communications

I.5.1. What is VLC?

VLC is a wireless communications technology that uses light visible between colors blue and red. It is a protocol to convert Wi-Fi into VLC through the LED Light (Light Emitting Diode). While Wi-Fi uses the radio part of the electromagnetic spectrum, VLC uses the optical part and it is based on the sending of data by modulation according to a well-defined and standardized protocol [34].

I.5.1.1. Electromagnetic spectrum

VLC fully leverages the use of the visible light spectrum, which ranges between 384 and 789 THz, adding 400 THz to the available bandwidth for wireless communications [35].

The use of visible light as a medium for data allows VLC to be completely safe for human health. However, Radio Frequency (RF) are currently classified as possible causes of cancer in humans according to the World Health Organization [35].

Radio waves, microwaves, infrared radiation, visible light, ultra-violet radiation, X-rays, gamma rays and cosmic rays in the increasing order of frequency and decreasing order of wavelength give the entire range (electromagnetic spectrum) [36].

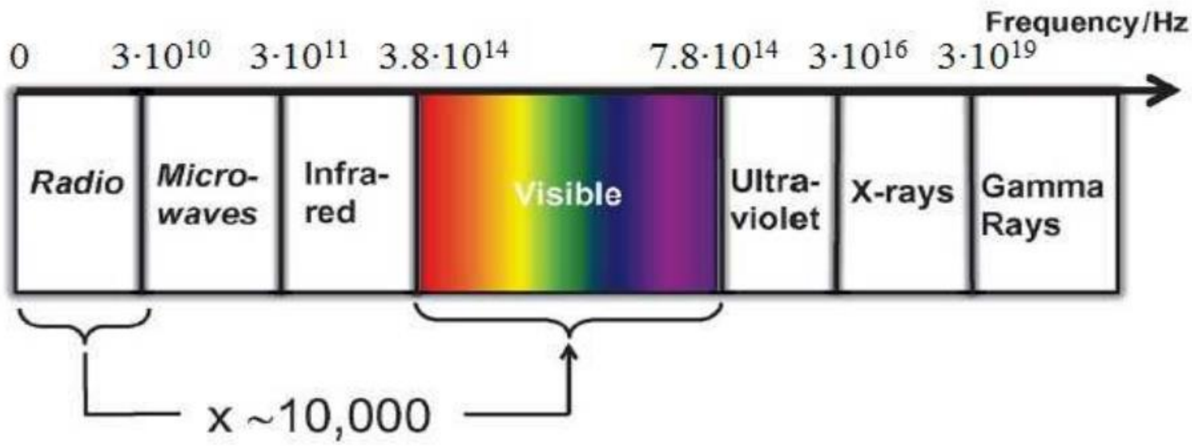

Figure I.23: Distribution of the electromagnetic spectrum and visible spectrum [37].

I.5.2. VLC architecture

A VLC system primarily includes a VLC transmitter, which modulates the light emitted by LEDs, and a VLC receiver, which uses a photosensitive element (photodiode) to extract the modulated signal from the light. The transmitter and receiver are physically separate but communicate through the VLC channel. For VLC systems, maintaining a line-of-sight (LOS) is essential [37].

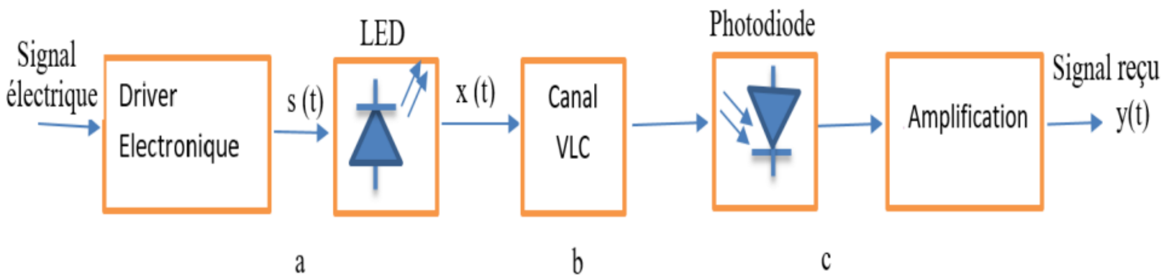

Figure I.24: Architecture of a VLC system: a Transmitter; b. Chanel; c. Receiver [35].

I.5.3. Transmitter

A VLC transmitter is an electro-optical device that transmits data using visible light waves over a wireless channel. The performance parameters of the VLC transmitter are primarily constrained by the characteristics of the LEDs [38][39].

The LED converts the electrical signal into a light signal, which is used for both lighting and communication. Before turning into a light signal, the data is handled and modulated online by the DAC by changing the amplitude or any other quality of the LED light [38].

- **LED**

LED light bulbs are now the primary medium for Visible Light Communication and the most popular due to its characteristics. [12]

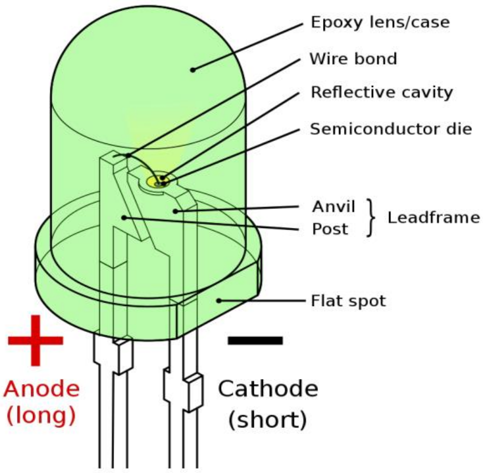

Figure I.25: Basic representation of a standard LED [12].

When a LED is on, it transmits a 1 bit; when it is off, it transmits a 0 bit. The frequency changes occur so rapidly that the human eye cannot detect them, perceiving only continuous light without flicker. This allows for data transmission speeds of up to 1 Gbps, significantly faster than the approximately 100 Mbps offered by Wi-Fi. [12]

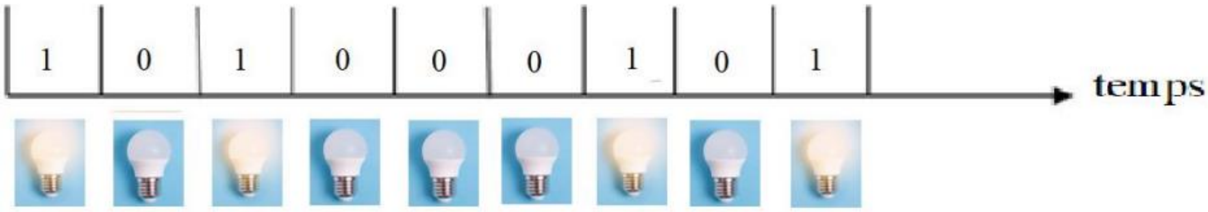

Figure I.26: Basic representation of data transfer using an LED [35].

I.5.3.1. Transmission channel

There are two types of link configurations on VLC: directed line-of-sight (LOS), no directed LOS (NLOS) as shown Figure I.27:

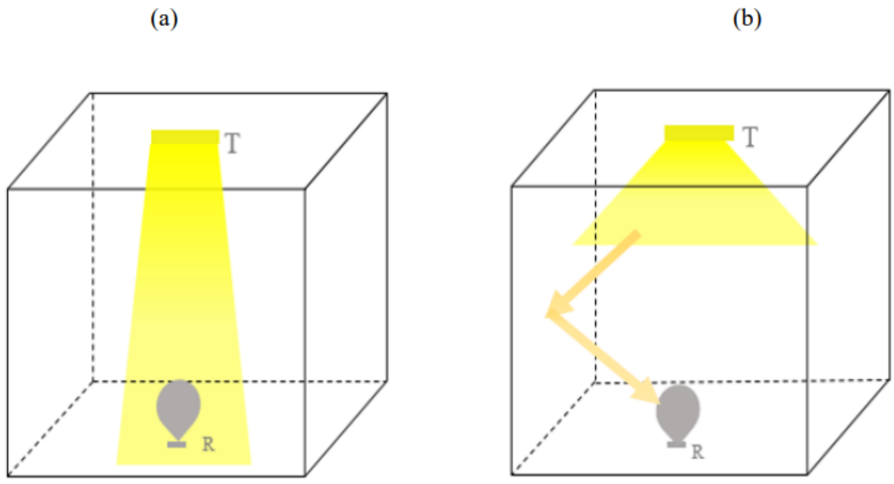

Figure I.27: Link configuration on VLC: (a) is LOS and (b) is NLOS, T is a transmitter and R is a receiver [40].

- **Line-of-sight (LOS)**

The LOS (Line of Sight) link corresponds to the path that directly connects the transmitter to the receiver. Consequently, it is inversely proportional to the square of the transmitter-to-receiver distance (d). The photodetector collects optical signals and converts them into electrical signals [34] [40].

- **Non-line-of-sight (NLOS)**

NLOS communication systems rely on multipath propagation and typically occur in indoor environments. They are more challenging to predict than LOS links because NLOS optical links communicate through signals reflected off walls, furniture, people, and other obstacles [40]. Additionally, factors like the position and orientation of the transmitter and receiver also influence NLOS communication [35].

I.5.3.2. Receiver

The VLC receiver takes data from the modulated light beam and turns it into an electrical signal, which is then demodulated and decoded by the built-in decoder. The performance of the VLC system depends on the VLC receiver, and it can be affected by other light sources, both natural and artificial, causing interference. Using an optical filter to block unwanted spectrum components can improve the VLC receiver's performance [39].

- **Photodiode**

As a type of photodetector, a photodiode is a device that converts light into electricity (current or voltage). Unlike the photodiode which works on the principle of photo-conduction, LED works on another principle called electro-luminance [41].

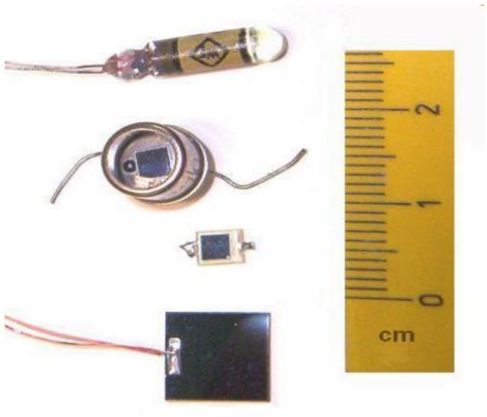

Figure I.28: Photodiode icon and examples [12].

I.5.4. Light-Fidelity Li-Fi

Li-Fi is a two-way, high-speed wireless communication technology that utilizes the visible portion of the electromagnetic spectrum. Similar to Wi-Fi, Li-Fi enables devices to connect to the internet and transfer data. However, the key difference between these technologies is in their data transmission methods. Li-Fi relies heavily on the development of LEDs (Light Emitting Diodes), as they are the primary light sources used [35].

The Li-Fi system main components are as follows:

a) A high brightness white LED which acts as transmission source.

b) A silicon photodiode with good response to visible light as the receiving element [42].

VLC technology uses the visible part of the light spectrum, while LiFi technology can use any part of the light spectrum for communication [43].

VLC systems can use the SISO system, the SIMO system, the MISO system, and the MIMO system. In reality, a VLC system can be considered a Li-Fi system if it supports multi-user communications and the mobility of the UE. On the other hand, a Li-Fi system is equivalent to a VLC system if only the VL spectrum is used in both up-link and down-link [11].

I.5.4.1. Combination WiFi and LiFi

Based on the advantages and disadvantages of LiFi technology, there are proposals to integrate it with WiFi technology. The hybrid system offers several new features, including enhanced security, high data transfer rates, broader coverage area, and improved indoor positioning. The first three features come from LiFi technology, while the coverage feature comes from WiFi technology [44].

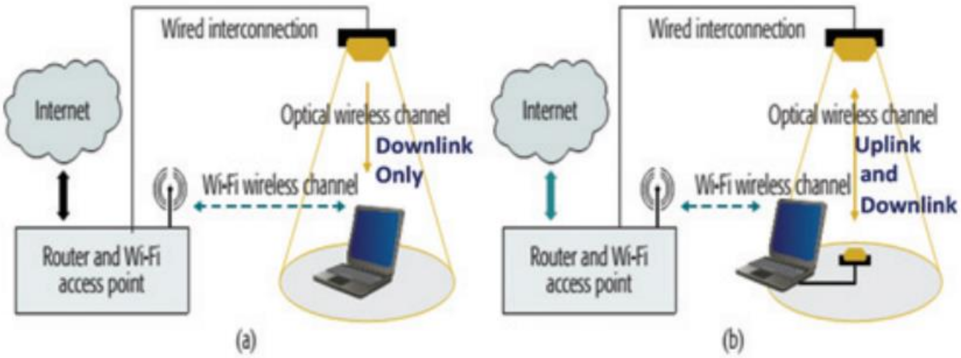

Figure I.29: Configurations of the a) hybrid technique, and b) the aggregated technique [44].

I.5.5. VLC application

I.5.5.1. Indoor

- **Mobile connectivity (Li-Fi)**

Li-Fi (Light Fidelity) is one of the most important applications of VLC, created in 2011 by Harald Haas. The data is transmitted by LEDs, which are widely used in lighting systems today. Therefore, lighting systems provide a suitable infrastructure for Li-Fi systems. LEDs can switch quickly and use a wide range of visible light, allowing them to send lots of data quickly. Although Li-Fi is an emerging technology, it is well-suited for short-range communication [35].

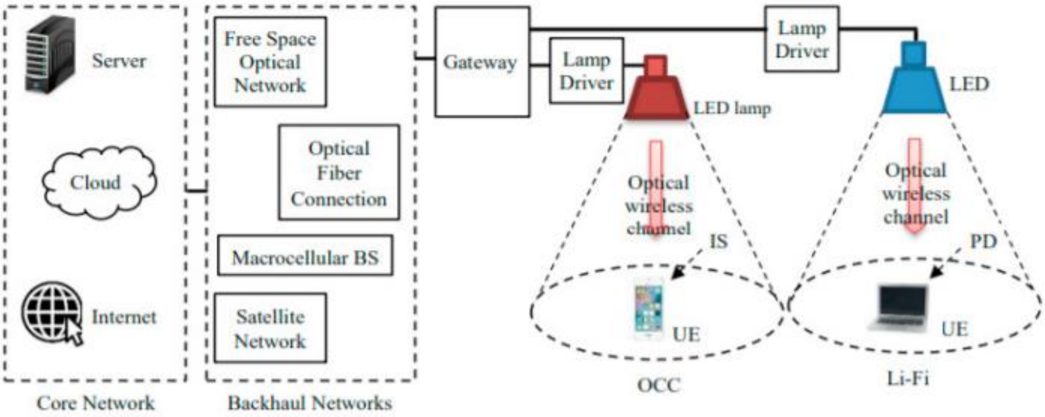

Figure I.30: Architecture of a Li-Fi system [45].

- **Creating smart places**

Developing smart locations with geolocation capabilities, like museums, is possible with VLC. Internal lighting allows the user's smartphone or tablet to display information [35].

Figure I.31: Deployment of VLC in smart homes and museums [35].

- **Places where using RF is limited or restricted**

Because the light signal is safe and available everywhere, it can be used to provide wireless communication in places where the use of RF signals is limited due to the potential dangers. Especially in hospitals and airplanes, the radio-electromagnetic transmissions may interfere with medical devices and sensitive airplane devices and cause problems. However, by incorporating LiFi technology in such places, wireless communications and accessing the Internet will be easy and safe for all people inside.

It is also prohibited in sensitive plants, such as power plants, petrochemical plants, and nuclear plants, because of the ignition risks in these environments. Nevertheless, since the light is safe in these places, the LiFi system can provide easy coverage and data transmission in such environments [44].

Figure I.32: The use of VLC in hospitals and airplanes [34].

- **Positioning**

Global Positioning System (GPS) technology does not function well indoors. Therefore, we have turned to visible light communication (VLC) technology using LED lights. VLC-based positioning systems are less affected by multipath transmission and offer much better location accuracy than RF-based indoor positioning systems [35]. With visible light communication, it is possible to identify the user’s position within a range of 1–2 m. This technology can also assist users identify which floor they are located on [46].

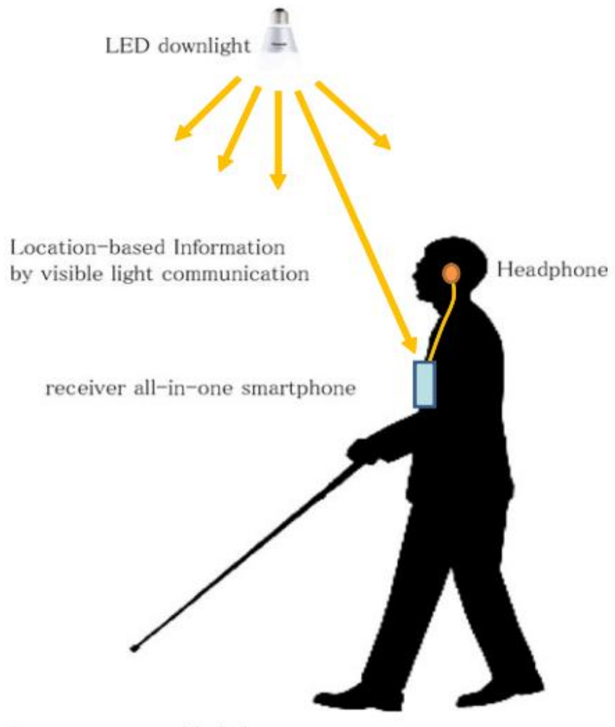

Figure I.33: Positioning using visible light communication [46].

- **Wireless industrial domain**

The fourth industrial revolution or Industry 4.0 represents a new era of industry characterized by increased operational efficiency and the emergence of new products and business models. Although Industry 4.0 is in its early stages, continuous development and research, particularly in visible light communication (VLC) technology are essential for its adoption in the industry. The integration of cyber-physical systems (CPS) is at the heart of this revolution, enabling the interconnection of end-to-end industrial processes via the Internet. To achieve the goals of Industry 4.0, it is crucial to have a fast, low-latency, and reliable data communication connection. Wireless optical communication technologies, such as VLC, are ideal for these applications because they offer high bandwidth without requiring a licensed spectrum, while being secure and free from RF interference. Channel measurements of a distributed 8 X 6 VLC system with multiple inputs and outputs (MIMO) intended for use by robots in a manufacturing cell have been reported. The results indicate that link availability was achieved with a sufficient signal-to-noise ratio (SNR) for LOS transmission. However, the movement of robots can affect the connection. To address this issue, spatial diversity has been suggested. Furthermore, a visible light positioning (VLP) technique has been proposed to improve localization services in Industry 4.0, using active receivers and low-cost infrastructure [35].

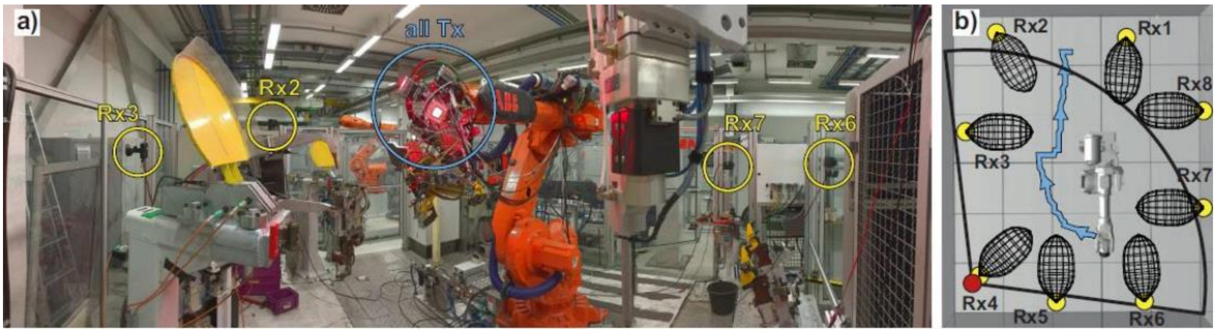

Figure I.34: (a) Manufacturing cell used for experiments in the BMW robot testing facility; and (b) the trace of the movement of the robot arm studied in [35].

I.5.5.2. Outdoor

- **Outdoor access to the internet**

Currently, cellular networks are the sole method for online connectivity in outdoor settings. WiFi's utility is confined to indoor spaces and is generally impractical for outdoor use. But LiFi technology will use outdoor lamps to make Internet access easy. So, if you are walking down the street, sitting on the beach, or playing in the park, you are going to have easy access to the Internet if you have got lamps around you [44].

- **Vehicles and transportation**

A LiFi network can be created between cars if the front and rear lights of the cars are replaced with LED lights. This would allow cars to communicate with each other (V2V) to reduce the chances of car accidents. Additionally, streetlights can be equipped with cameras to monitor the roads, detect any congestion or emergencies, and then send the information directly through LiFi technology to the traffic management center for quick action [44].

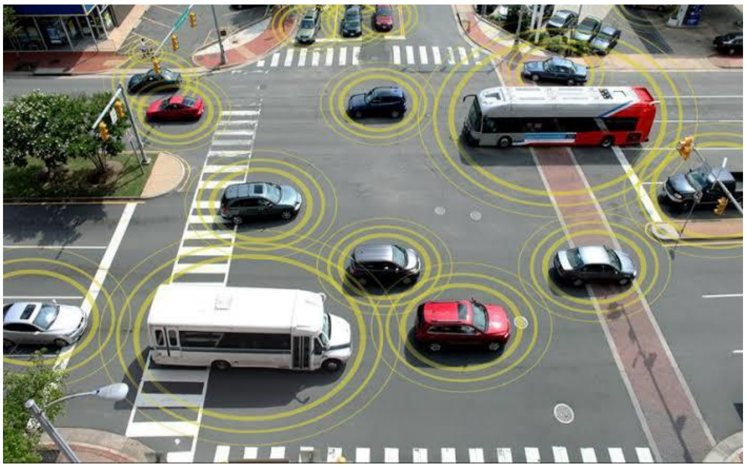

Figure I.35: The communication between cars to reduce chances of accidents [44].

- **Underwater communication**

RF waves do not propagate effectively in seawater because of its high conductivity. Divers can utilize Li-Fi technology for both illumination and data transfer simultaneously. Therefore, it is possible to communicate from diver to diver, diver to mini-sub, diver to drilling rig, etc. [47] [48]

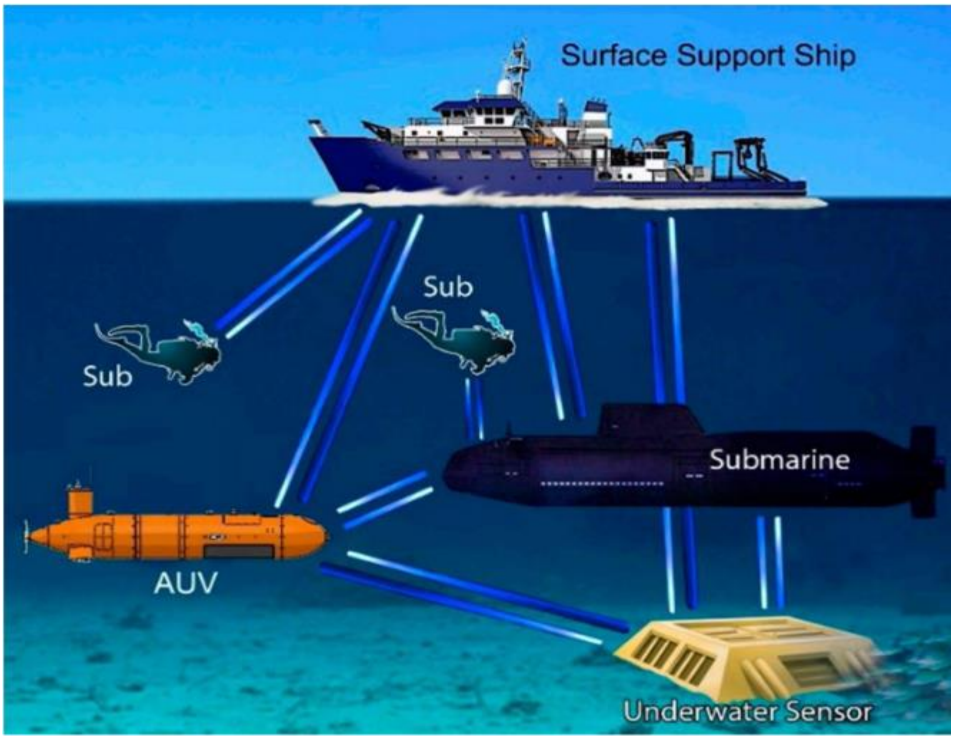

Figure I.36: Underwater communication [35].

I.6. Conclusion

In this chapter, we provide a brief overview of artificial intelligence and its history as we discuss its most important branches, including machine learning, neural networks, and deep learning, and the differences between them. Additionally, we consider robotics, especially mobile robots, and their types. We also talk about wireless communications due to their importance in our current times, focusing specifically on visible light communication (VLC), its structure, and its most significant implementations.

Chapter II: Automatic Detection and Tracking of Movements

II.1. Introduction

Object detection techniques are the foundation for the artificial intelligence field as it is also important in many other fields. Motion detection and tracking are fundamental steps in object detection. These deep learning algorithms revolutionize the way we use artificial intelligence since they have the ability to process large amounts of data and create accurate models.

II.2. Image sequence

A video can be considered a sequence of images, each containing a static view of the scenes unfolding within it. These images, or frames, are taken at regular and very short intervals (30 frames per second). The frames are two-dimensional images represented by X = (x, y), with the temporal axis't' representing the third dimension. A figure is an illustration of the compact representation of a video, with an overlay of the frames that compose it [49].

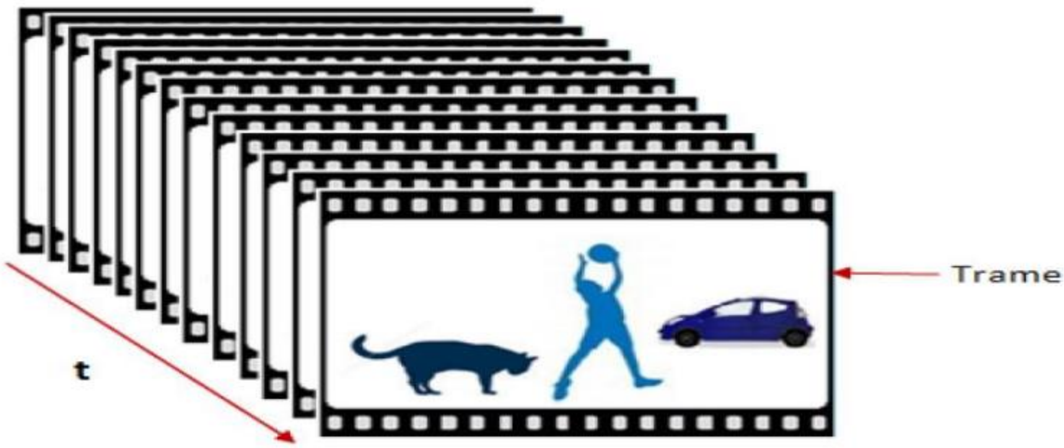

Figure II.1: Compact representation of a video, T represents the temporal axis [49].

Each frame of this video can be used in preprocessing operations. These preprocessing operations aim to detect moving objects within the video. The detection of moving objects in the video employs motion segmentation methods to label the regions of frames corresponding to the moving objects [49].

II.3. Motion detection

Motion detection is the initial step for every object tracking algorithm. It involves determining which pixels in the current image belong to the background of the scene and which represent moving objects. This is a critical and challenging step as it must be robust to variations in scene brightness as well as the presence of shadows [50].

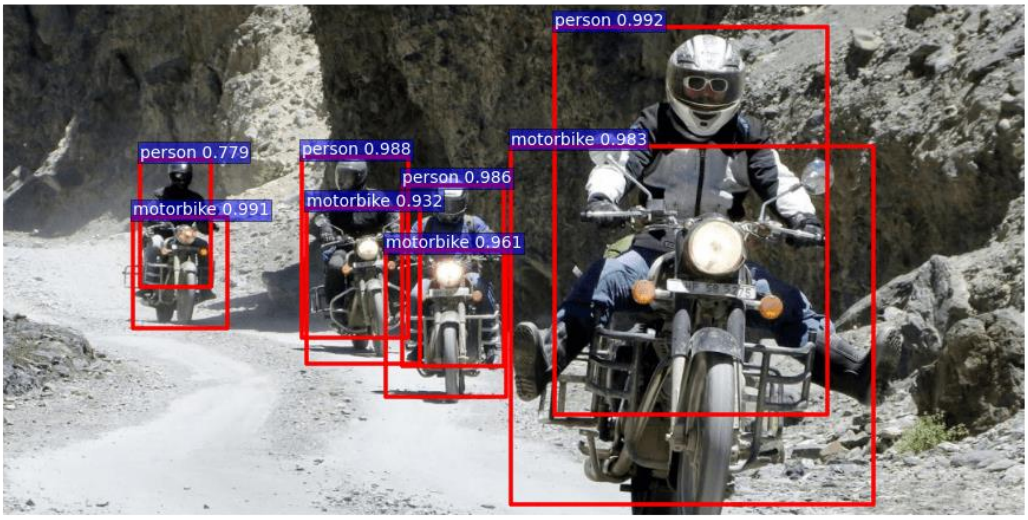

Figure II.2: Example of recognizing moving objects in an image [51].

II.3.1. Types motion detection

II.3.1.1. Detection without background modeling

These methods consist of detecting movement by calculating a mathematical quantity at every point of an image based on the intensity or color of all the pixels, which should reflect the entire visible movement in the scene without considering background information. Three approaches can be cited: the temporal derivative, the optical flow, and background subtraction [52].

- **Optical flow**

The projection of real movement on the image plan defines a field of two-dimensional vector that is the optical flow [52].

In video processing and analysis, optical flow is defined as a 2D inter-pixel vector field to transition from one image in a video sequence to another over time. The estimation of 2D motion through apparent motion is a very complex task because apparent motion is very sensitive to noise and changes in brightness. The methods for extracting optical flow are categorized into two main families:

- Differential methods.

- Region matching algorithms [53].

Figure II.3: Example of optical flow estimation [53].

- **The temporal derivative**

The principle of the temporal derivative method is to measure the change in appearance of pixels between two consecutive frames in a video stream [50].

This method adapts to the change of the scene, but it is less effective because over a period of time Δt, perhaps, only a part of an object is detected, for example: the hand, the head, etc. The void (the region where the object has moved to another place) is also detected [53].

- **Subtracting the background**

It involves detecting the region of movement by subtracting, pixel by pixel, the current image from the background image. The latter is created in several ways, such as:

The average of the first N frames of the video stream.

- The background image that has little change or no movement for a long duration.

- The first image in the video stream.

Regardless of the method used to create the background image, the background subtraction method remains very advantageous because it detects the object completely and is used in several applications [50].

Figure II.4: Example of background subtraction [54].

II.3.1.2. Local background modeling

Every pixel of the image is associated with a function or a value, thus allowing the modeling of the background appearance of that pixel. Only the observations that have taken place at a point result in the background appearance model at that point. The other pixels of the image do not intervene. The modeling can be static or probabilistic [52].

II.3.1.3. Semi-local background modeling

The only difference that characterizes this method compared to the previous one is that the modeling at a point depends on the observations that took place in the region of the image to which it belongs [52].

II.3.1.4. Global background modeling

These methods use all the observations at each moment to construct a model of the entire background [50].

II.4. Object tracking

The tracking of moving objects can be defined as the spatiotemporal localization of an object in motion during a video sequence where an object refers to a region within the image that can be modeled using contours, silhouettes, geometric primitives (such as a bounding rectangle around the object of interest), or even a central point [49].

The tracking of moving objects allows for the analysis of their trajectories and behaviors by continuously and reliably determining their positions throughout the video stream [55].

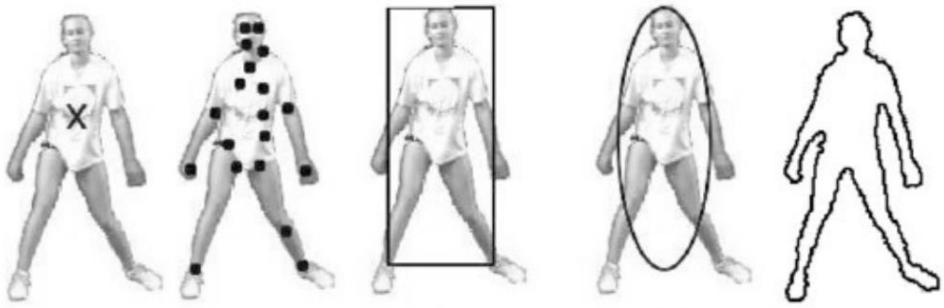

Figure II.5: Object representation. (a) Centroid (b) set of points, (c) rectangular bounding box, (d) elliptical bounding box, (e) contour [50].

II.4.1. Silhouette tracking

The representation of an object through a silhouette allows for precise consideration of the object’s shape (It effectively works on designing complex-shaped objects). The goal of silhouette-based tracking methods is to estimate the silhouette of objects of interest for each frame in a video. This approach is also known as image segmentation [49].

- **Explicit approaches**

The proposed method aims to model the silhouette through a series of parameters, followed by an appropriate filtering technique. These parameters form the state model used by the filtering algorithm. The process takes place in several stages: first, the prediction of the parameters describing the silhouette in the current image, then the correction based on the results of the prediction. The advantage of these explicit methods lies in their ability to track an object while having a low computational cost [55].

- **Implicit approaches**

Unlike previous methods, methods based on implicit representation are robust to changing the topology of objects by minimizing an energy function that allows you to track a contour or region despite topology changes. However, minimizing these functions is generally more costly, and convergence towards a global minimum is not necessarily ensured [50].

II.4.2. Kernel tracking

Kernel tracking involves tracking an object represented by a basic geometric shape, for example, a rectangular model or an elliptical shape. This estimated movement is generally parametric (translation, rotation, etc.). The object being tracked is represented by a weighted histogram calculated over an elliptical region containing the object.

The most commonly used kernel tracking algorithm is “Mean-Shift”, which leverages a color histogram density function for the object and a candidate window.

The main advantage of this method is that it allows estimating the position of the object in a limited number of iterations [49] [55].

II.4.3. Point tracking

The object is represented by a point, which is its centroid, or by a set of points. Generally, this point-based representation is suitable for tracking objects that occupy small regions in an image. In cases where all visible targets can be reliably and rapidly detected at each moment, the tracking problem can be defined as a detection-matching problem between successive images. In this category, approaches can be deterministic or probabilistic [49].

- **Deterministic approach**

In the deterministic approach, tracking is performed by minimizing a distance calculated based on certain object characteristics. These characteristics include appearance (shape similarity and/or photometric and/or motion content). Object models based on appearance can take the form of densities (color or contour histograms), a contour map (open or closed contour of the object), or a combination of these models [50].

- **Probabilistic approach**

The Probabilistic Approach addresses the issue of noise by introducing uncertainty into the object model and the models of potential targets. Tracking is achieved through filtering methods (Kalman filtering and particle filtering).These probabilistic techniques allow for robust tracking even in noisy environments [55].

II.5. Object detection

II.5.1. What is object detection?

Object detection is a technique used in computer vision for the identification and localization of objects within an image or a video. It is an important part of many applications, such as self-driving cars, robotics, and video surveillance [56] [57].

Object detection can be used to verify whether an object is present in an image or video [58].

Object detection models require training; they identify object types based on the database on which they were trained [58].

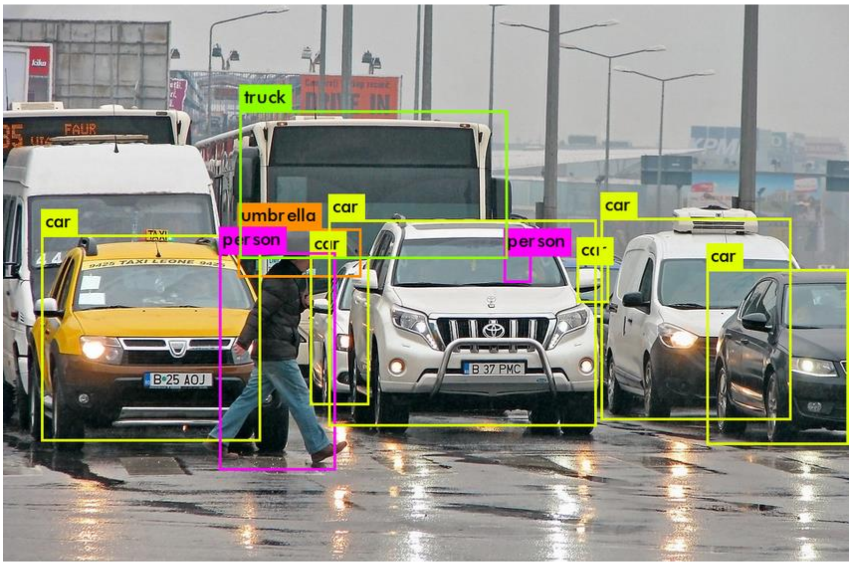

Figure II.6: example of object detection [59].

II.5.2. Object detection vs image classification

The image classification considers the whole image and does not tell you where the object appears, unlike the detection that is more advanced where it informs the object. In short, detection is a classification plus location.

Classification is useful for tags that have no physical boundaries, like “blurry” or “sunny,” while object detection excels at detecting physical objects like cars [60].

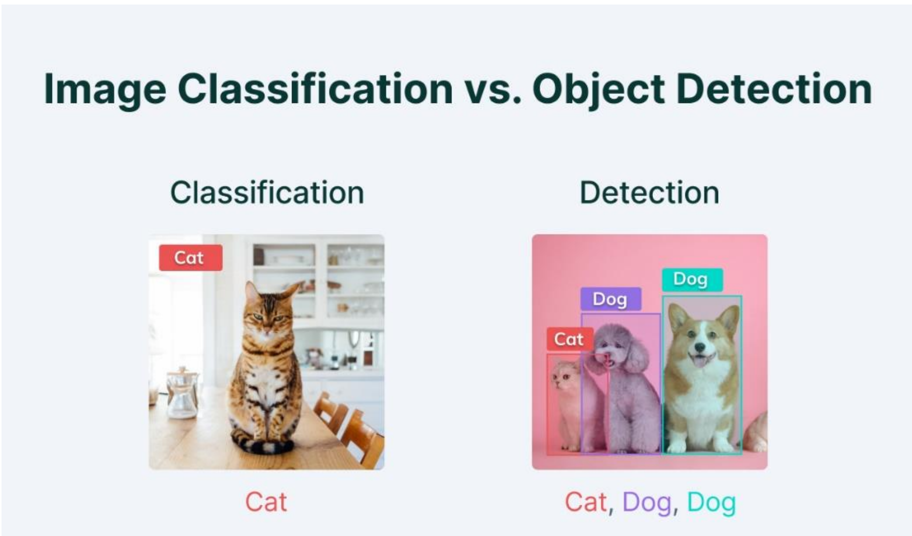

Figure II.7: Object detection vs image classification [60].

II.5.3. Object detection vs image segmentation

The segmentation process is to determine which pixels of an object class are found in an image.

Semantic image segmentation marks all pixels belonging to that tag but does not define the boundaries of each object.

Conversely, object detection locates each object instance precisely using bounding boxes without segmenting the object itself.

When semantic segmentation and object detection are combined, it results in instance segmentation. Which first detects the object instances, and then segments each within the detected boxes (known in this case as regions of interest). [60]

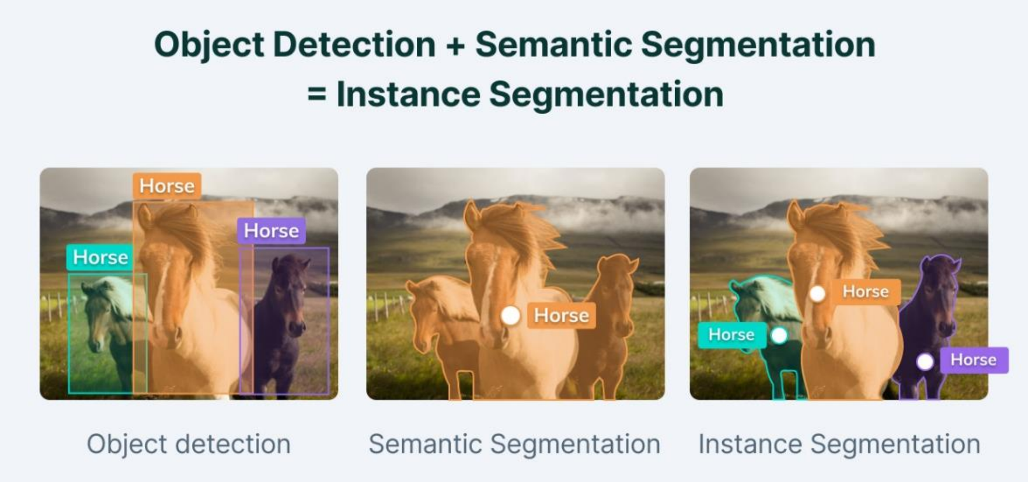

Figure II.8: Object detection vs image segmentation [60].

II.5.4. Object detection algorithms

Object detection algorithms are broadly divided into two categories based on the number of times the same input image passes through the network: one-stage object detectors and two-stage object detectors [61].

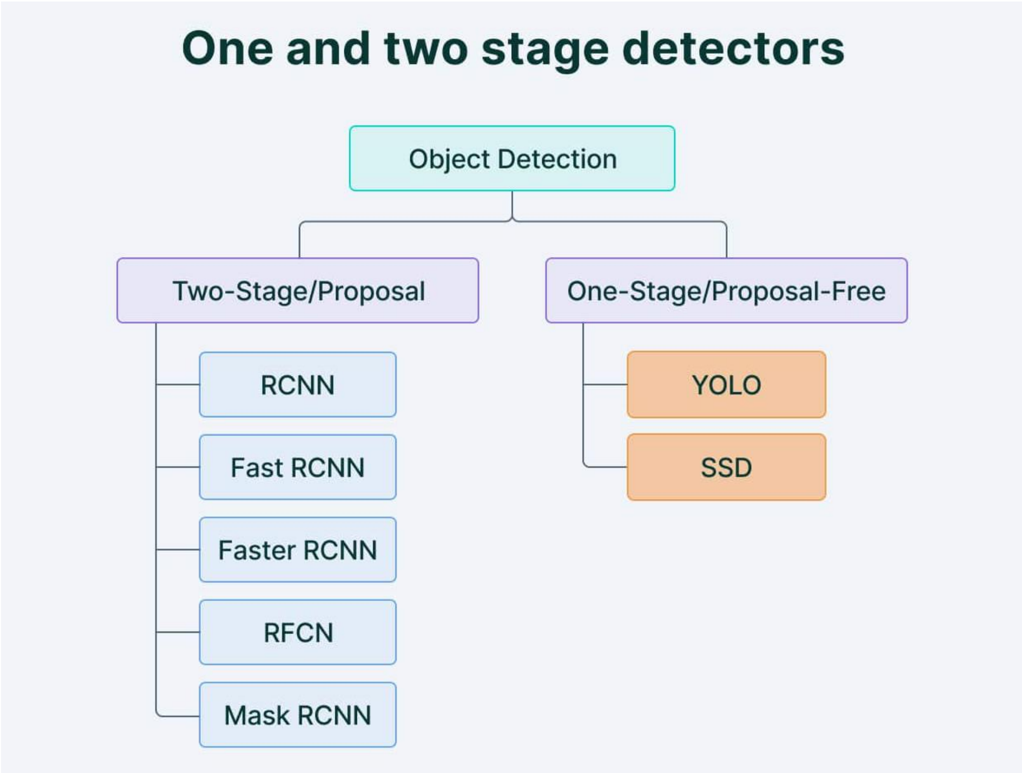

Figure II.9: One and two stage detectors [61].

These are not all the object detection algorithms shown in the previous figure; there are many of them.

- **One-stage object detection**

In single-stage object detection, the location and presence of objects in the input image are predicted after a single pass through the image. They are computationally efficient since they process a whole image in a single pass.

- **Two-stage object detection**

The two-stage object detection method makes predictions about the position and existence of objects through processing the input image through two passes. The initial pass produces a range of suggestions or possible locations for objects, while the subsequent pass refines these suggestions and produces the final predictions. The computational cost of this method is higher even if it is more effective than single-shot objects detection [61].

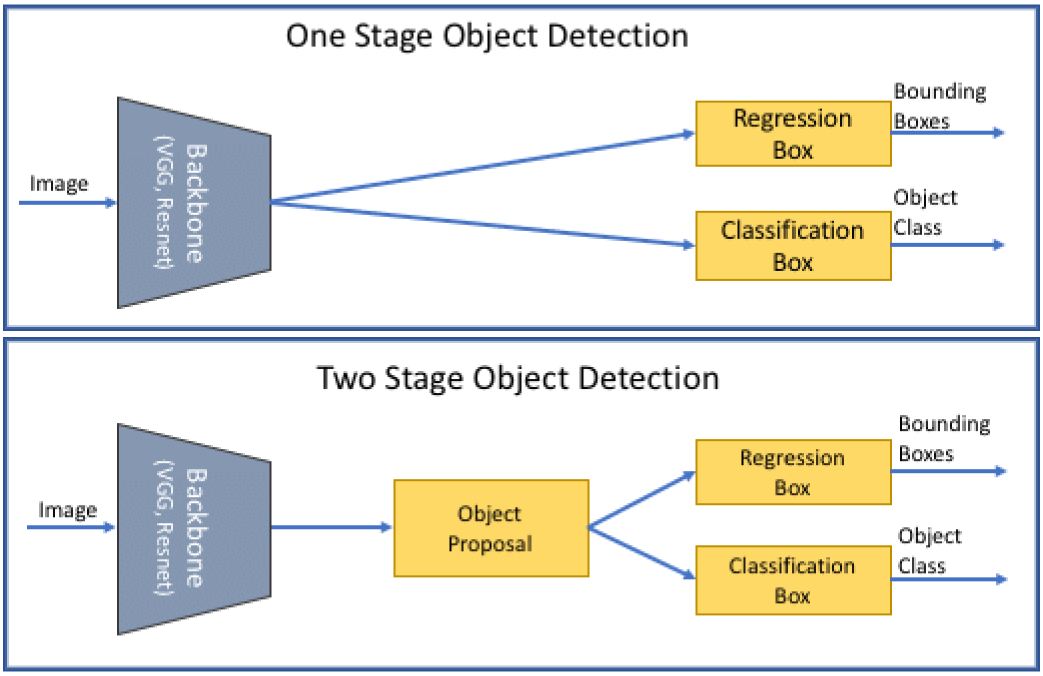

Figure II.10: One stage vs two stage object detection models [62].

Generally, one-stage detector (single-shot object detection) is better suited for real-time applications, while two-stage detector (two-shot object detection) is better for applications where accuracy is more important [61].

II.5.4.1. You Only Look Once (YOLO)

You Only Look Once (YOLO) is a state-of-the-art, real-time object detection algorithm. YOLO was first introduced by Joseph Redmon and AL in an article published in 2015. It applies the neural network to the full image, where predictions for bounding boxes and class probabilities are made all at once [56][63].

- **YOLO version**

YOLOv1

YOLOv2

YOLOv3

YOLOv4

YOLOv5

PP-YOLO, PP-YOLOv2, and PP-YOLOE

Scaled-YOLOv4

YOLOR

YOLOX

YOLOv6

YOLOv7

DAMO-YOLO

YOLOv8: YOLOv8 provided five scaled versions: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large) and YOLOv8x (extra-large) [64].

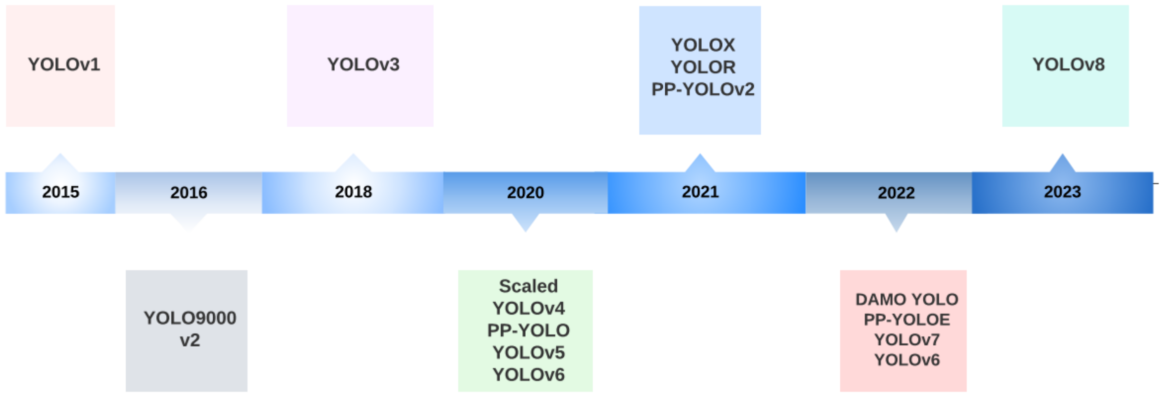

Figure II.11: A timeline of YOLO versions [64].

- **What makes YOLO popular for object detection?**

The significant factors contributing to YOLO's prominence in the competition are: [56]:

- Speed.
- Detection accuracy.
- Good generalization.
- Open source.

- **YOLO architecture**

The original YOLO architecture consists of 24 convolution layers, followed by two fully connected layers.

Before training, the original YOLO network uses 224×224 pixels and then uses 448×448 pixels for detection.

The model adapts to image classification when it is shifted from a classification model to a detection model [65].

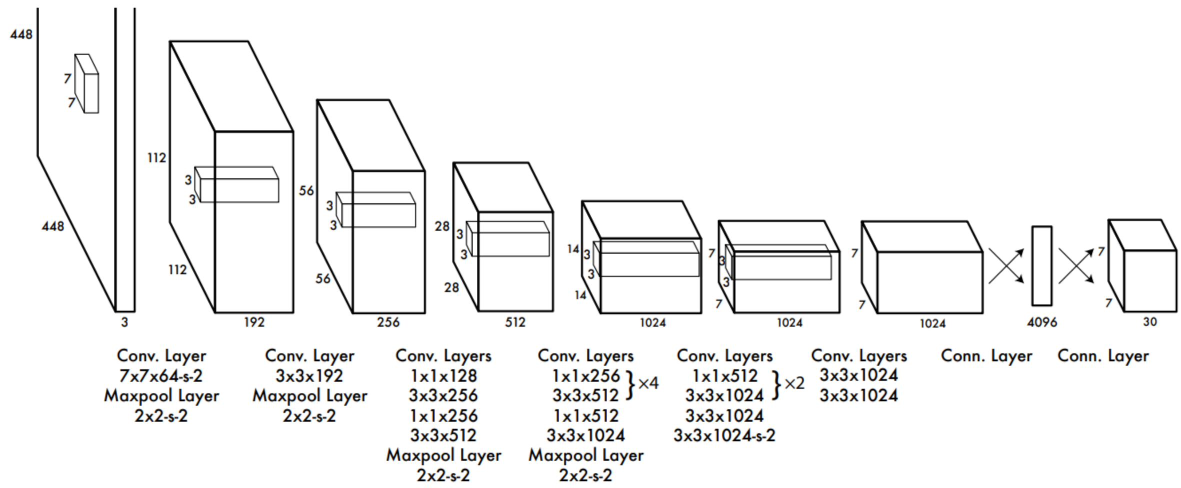

Figure II.12: YOLO architecture [66].

• The operating principle of YOLO

The system divides the input image into an S × S grid. If the center of an object falls into a grid cell, that grid cell is responsible for detecting that object. The next step is to identify bounding boxes B that outline all objects in an image. There can be many bounding boxes, one for each object (B=2 means that a grid cell can contain at most 2 objects) [56][59][66].

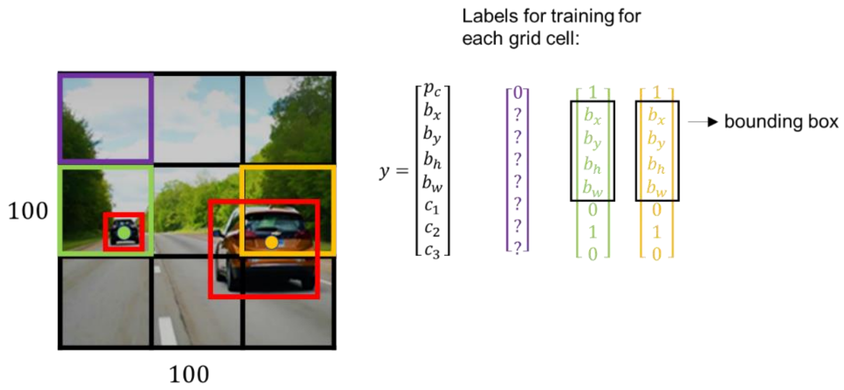

Figure II.13: Principle of YOLO model for detection [59].

- Y = the final vector representation for each bounding box.
- pc = boolean indicating whether the cell contains an object.
- (bx, by) = coordinates of the object's center.
- (bh, bw) = the height and width of the bounding box (in red).
- (c1,c2 et c3) =classes. [56][59][66]

These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor [66].

In most cases, a single object in an image can have multiple grid box candidates for prediction, although not all of them are relevant. IOU allows discarding them to keep only the relevant boxes [56].

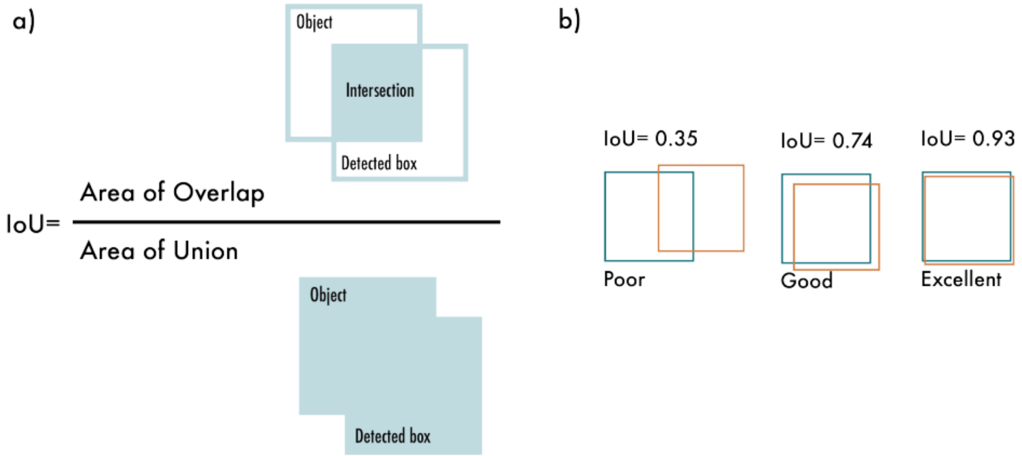

Figure II.14: Intersection over Union (IOU). a) The IOU is calculated by dividing the intersection of the two boxes by the union of the boxes; b) examples of three different IOU values for different box locations [64].

The object may contain a large number exceeding the limit of boxes that contain IOUs, and leaving them may cause noise. Therefore, NMS filters out redundant and irrelevant bounding boxes, keeping only the most accurate ones (highest probability score) [56][64].

Figure II.15: Non-Maximum Suppression (NMS) [64].

II.6. Conclusion

In this chapter, we discuss methods of motion detection since a video is a series of images. We also talk about real-time tracking methods, where tracking that uses boxes surrounding the object of interest is considered. Additionally, we consider object detection as more than just classification that informs about the presence of an object in the image; but rather it includes its location. We talk about YOLO as one of the best and most famous real-time detection algorithms. It applies neural networks to the entire image to predict both the bounding boxes and the class probabilities at the same time. Therefore, we will work on a model based on it for our project in the next chapter.

Chapter III: Hardware, Software and Implementation Analysis

III.1. Introduction

In this chapter, we will present the software and the hardware aspects related to our project, and we will discuss the methods used to detect and recognize objects, describe the technologies employed, detail the stages of project development, and clarify the results obtained through practical experiments. Moreover, we will work on designing a mobile robot using VLC technology capable of performing these tasks.

III.2. Hardware parts

III.2.1.Raspberry Pi

The Raspberry Pi is a small, low-cost single board-computer, the size of a credit card that connects to a computer monitor. It has a dedicated processor, memory, and a graphics driver, just like a PC. It also comes with its operating system, Raspberry Pi OS and a modified version of Linux. It is used globally for educational purposes, teaching programming and computer science. [67]

In our work, we will use the Pi 4 model B. It is equipped with a Quad core 64-bit ARM-Cortex A72 running at 1.5GHz and comes with 4GB of RAM.

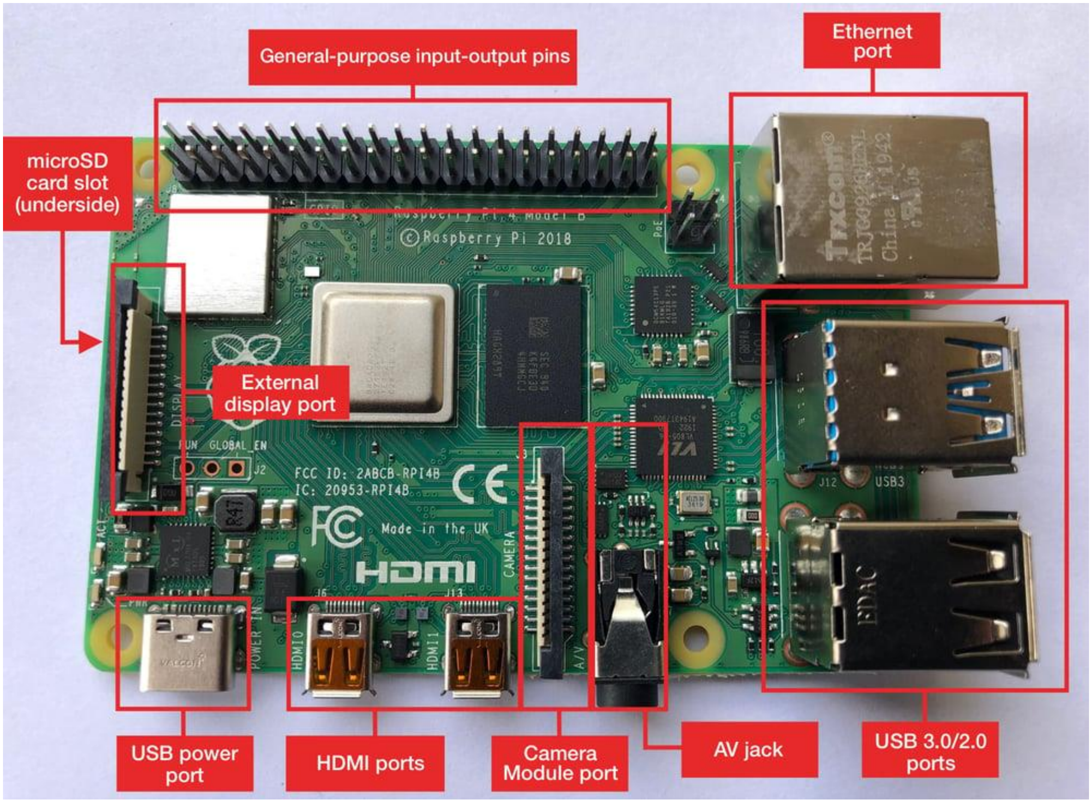

Figure III.1: Raspberry Pi 4 Model B [68].

III.2.2. Transmitter part

• **Arduino Uno:** Uno is a microcontroller board based on the ATmega328P. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz quartz crystal, a USB connection, a power jack, an ICSP header and a reset button.

Figure III.2: Arduino Uno.

• **LED**

A LED or a Light-Emitting Diode is a semiconductor device that emits light due to the electroluminescence effect. A LED is a PN Junction Diode, which emits light when forward biased.

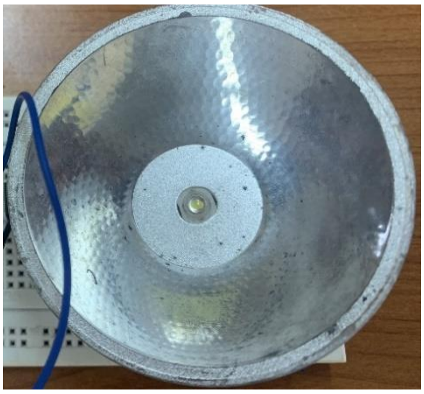

Figure III.3: LED.

• **Diode Laser (LD)**

The laser diode is a semiconductor light source with two terminals, functioning as a PN junction diode that emits light when activated.

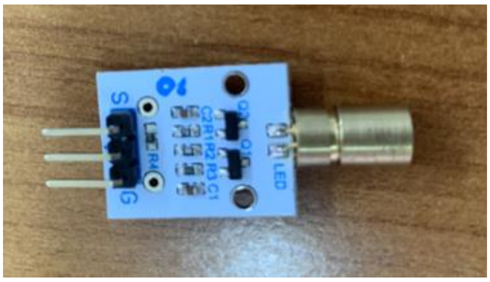

Figure III.4: Diode Laser (LD).

III.2.3.Receiver part

• **Arduino Mega:** The Mega 2560 is a microcontroller board based on the ATmega2560. It has 54 digital input/output pins (of which 15 can be used as PWM outputs), 16 analog inputs, 4 UARTs (hardware serial ports), a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header and a reset button.

Figure III.5: Arduino MEGA.

• **Photodiode PD (BPW34)**

BPW34 is a PIN photodiode with high speed and high radiant sensitivity in a miniature, flat, top view, clear plastic package. It is sensitive to visible and near infrared radiation.

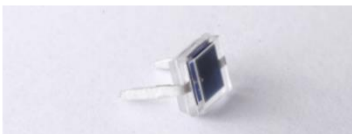

Figure III.6: BPW34 Photodiode.

• Solar cell

A solar cell is an optoelectronic device that directly converts the energy of light rays into electricity through a process called photovoltaic conversion. Its structure includes a wafer of semiconductor material with a P-N junction that acts as the main generator. Metal contacts are present on the front and back of the cell to collect the photo-generated carriers.

Figure III.7: Solar cell.

III.2.4.Robot car

We use a standard kit commonly employed in numerous robot car projects, which is well-suited for our purposes. It includes the necessary frame and four DC gear motors.

Figure III.8: 4WD Robot car.

III.3. Software parts

III.3.1.SOLIDWORKS

The SOLIDWORKS software is a volumetric modeler that allows the creation of complex 3D parts. These parts can then be used to create 2D drawings and assemblies of multiple parts. SOLIDWORKS is a dimension-driven system. One can specify dimensions and geometric relationships between elements. A change in dimension results in a change in the size of the part while preserving the design intent. A SOLIDWORKS model consists of parts, assemblies, and drawings and they display the same model in different documents.

Figure III.9: LED transmitter.

Figure III.10: Laser transmitter.

Figure III.11: Photodiode receiver.

Figure III.12: Solar cell receiver.

III.3.2.Arduino IDE

The Arduino Integrated Development Environment (IDE) is a cross-platform application (for Windows, macOS, Linux) that is written in functions from C and C++. It is used to write and upload programs to Arduino compatible boards.

III.3.3.Python

Python is the world's fastest growing and most popular programming language. Its versatility allows for a range of applications, including data analysis, visualization, artificial intelligence, machine learning, and automation.

Python’s ability to build applications across platforms (Windows, Mac, Linux, etc.), coupled with its simple syntax, which helps solve complex problems efficiently. Even though other languages can do a lot too, Python is easy to use and looks nice, making it very popular with employers and useful for both experienced programmers and beginners. [69]

III.3.4.Installation and configuration

III.3.4.1. Installing the operating system

To obtain the operating system on the Raspberry Pi:

- It is necessary to insert a memory card into the computer, and then download the raspberry pi imager application.
- Select the "Choose OS" option in Raspberry Pi imager. (In our case, we use 64-bit version).
- Go to the settings, select the name of the Raspberry Pi, the username, and the password.
- Choose "enable SSH" and select "configure wireless WAN" if you want to connect the Raspberry Pi to WiFi.
- After that, click on” Choose SD Card” to select the memory card that you have inserted into your computer.
- Start the operation in the SD card
- When the download is complete, remove the memory card from your computer and insert it into your Raspberry Pi.

To facilitate remote control of the Raspberry Pi and transfer files, there is a VNC (Virtual Network Computing) technique where the Raspberry Pi interface is opened as a window on my laptop. This is very convenient since we do not need a screen, mouse, or keyboard, just a power supply. This method is particularly effective because we will place our Raspberry Pi inside the robot.

To enable this technique, we activate the VNC option in the Raspberry Pi settings and install the VNC application on our device. Then, we enter the command "ifconfig" to obtain the Raspberry Pi's IP address.

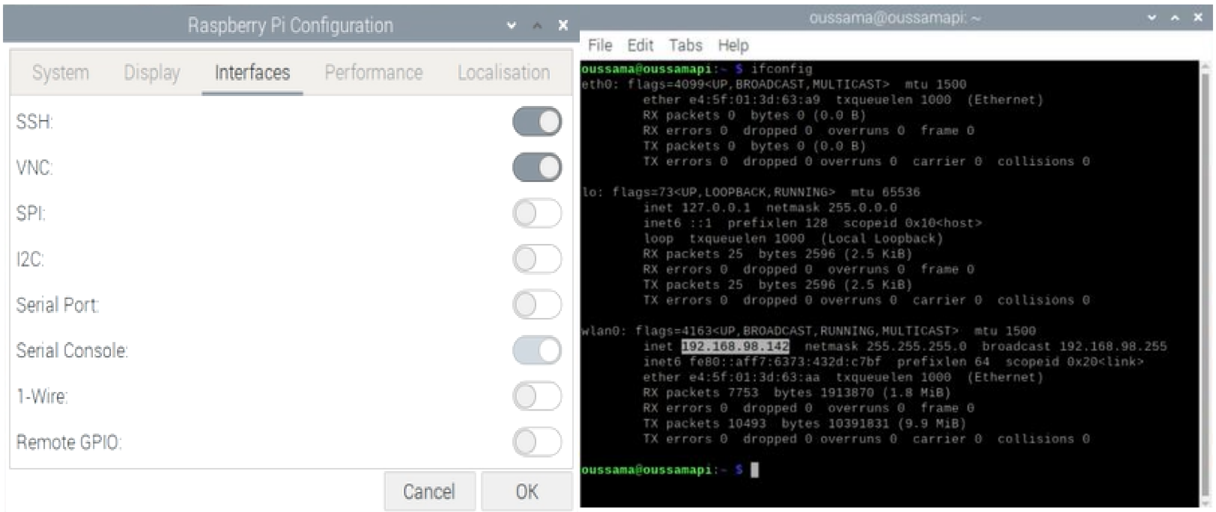

Figure III.13: Configuration and IP address of the Raspberry Pi.

Then, we open the VNC application and enter the IP address. It asks us for the username and password, The Raspberry Pi interface then opens directly.

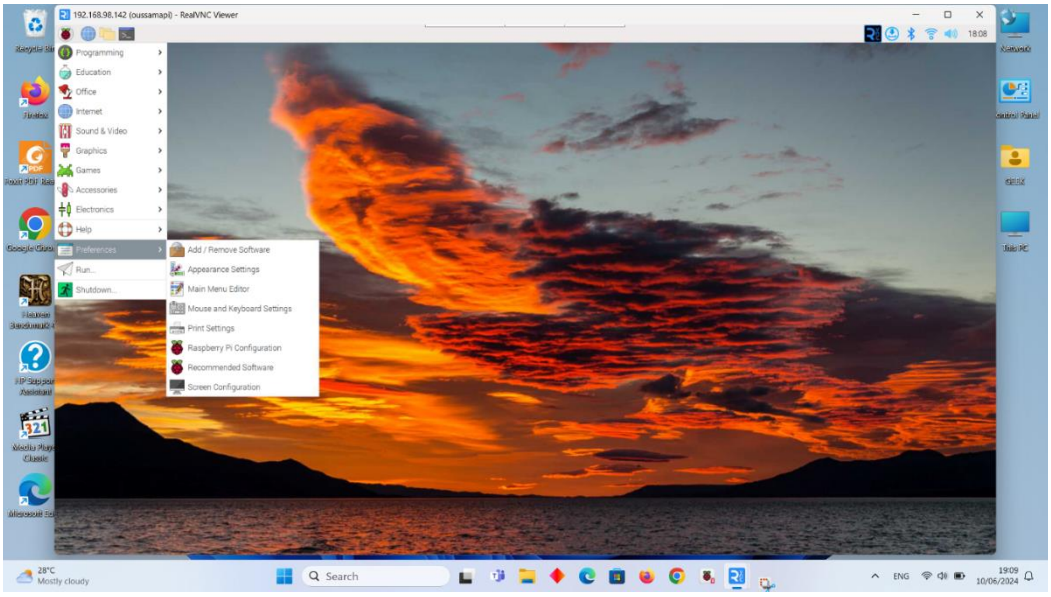

Figure III.14: VNC interface.

Sometimes, the IP address changes by itself. We will just need to search for it again and re-enter it.

III.3.4.2. Configuring a camera on a Raspberry Pi

- Insert the camera into the dedicated port on the Raspberry Pi (in our case, the USB port).
- Enter the command "sudo raspi-config", then go to Interfacing Options, Choose "Camera", and select "enable".
- Exit the configuration and restart your Raspberry Pi.
- Enter the command: "sudo apt install fswebcam". After the installation is complete, enter the command: "lsusb" to display the name of our webcam.

Figure III.15: Our Raspberry pi + webcam.

III.3.4.3. Library

- **Ultralytics**

Ultralytics is a company focused on the development of software solutions and tools for computer vision and artificial intelligence. They are known for their open-source library.

Ultralytics also offers other tools and libraries for object recognition such as YOLO.

In our work, we use the version YOLOv8n Installation is done by command: "sudo pip3 install ultralytics". Later, we discovered that the Windows commands also work on Linux. So, you can download it using the command: "pip install ultralytics".

- **Open CV (Open Source Computer Vision Library)**

Open CV (Open-Source Computer Vision Library) is an open-source library that includes several hundred computer vision algorithms. It has a modular structure, which means that the package includes several shared or static libraries.

We will use the latest version of Open CV for object detection and recognition in our work. It can be installed using the command: "pip install opencv-python".

III.4. VLC test

We will implement a wireless data transmission system based on Li-Fi technology for sending texts.

III.4.1.LED transmitter and photodiode receiver

In this model, standard components such as LEDs, photodiodes, resistors... etc. are used. A microcontroller is integrated for encoding and decoding data, ensuring the security and reliability of VLC communication. Additionally, a cone is utilized to focus the light emitted by the LED, thereby improving the system's range and accuracy. Furthermore, a 1 MΩ resistor is added in parallel to the photodiode to eliminate unwanted interference and enhance data reception.

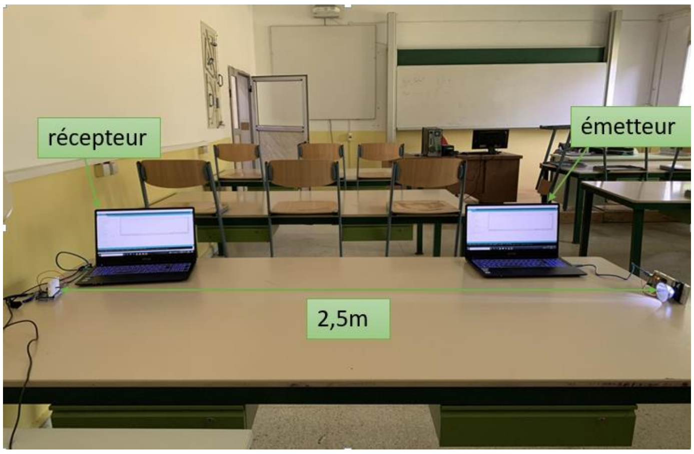

Figure III.16: Prototype of LED to photodiode.

III.4.2.LED transmitter and solar cell receiver + I2C LCD display

The LCD display is equipped with an I2C (Inter-Integrated Circuit) interface for communication between the Arduino and LCD. This is to display the transmitted texts on it as well.

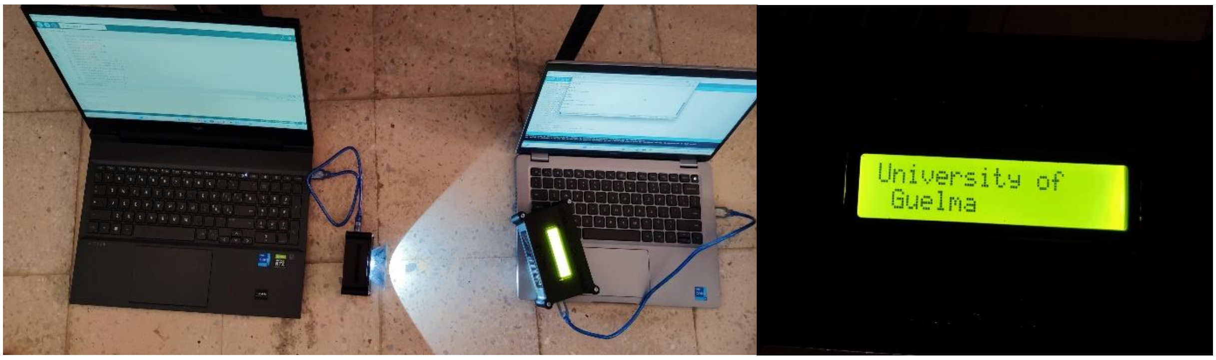

Figure III.17: Prototype of LED to solar cell.

III.4.3.Laser transmitter and solar cell receiver + I2C LCD display

For this prototype, we use a laser diode (LD) on the emitter side and a solar cell on the receiver side. The solar cells convert the visible light from the laser into electricity by absorbing photons with a semiconductor material, usually silicon, through the photovoltaic effect. This creates electron-hole pairs that are then separated by the structure of the solar cell. The electrons and holes are collected by metal contacts, thus generating a usable or storable electric current.

Figure III.18: Prototype of Diode laser to solar cell.

III.5. Implementation

III.5.1.Our scenario

In this particular section, we focus on our specific project. After controlling the robot using VLC technology to enter hard-to-reach areas and perform the task of fire detection and assuming successful results, we can send fire or smoke detection reports via VLC technology to the control room for necessary action.

III.5.2.About this application

As part of our object detection project, we downloaded a pre-trained YOLOv8n model along with its code from the GitHub platform.

The model and its code needed to be placed in one file so that the model could be loaded and recognized correctly by inputting the model’s name into this code.

At the beginning, the code did not work, because it had many errors so we worked on fixing them and it took us too much time. Eventually, the code ran successfully, opening the camera window and detected the objects that we could later test it at the university and in public places, and the results were very satisfactory.

Figure III.19: Result of detection in public places.

This model was trained on the COCO database.

COCO is a database containing 80 object classes, such as cars, persons, televisions, etc. It provides a substantial amount of data for training models.

We also tested it on the Raspberry Pi, and it worked, but it was not smooth. The frame rate in the video was about 5 fps because of the large database of this model, the limited processor of the Raspberry Pi device, and its lack of an external graphics card.

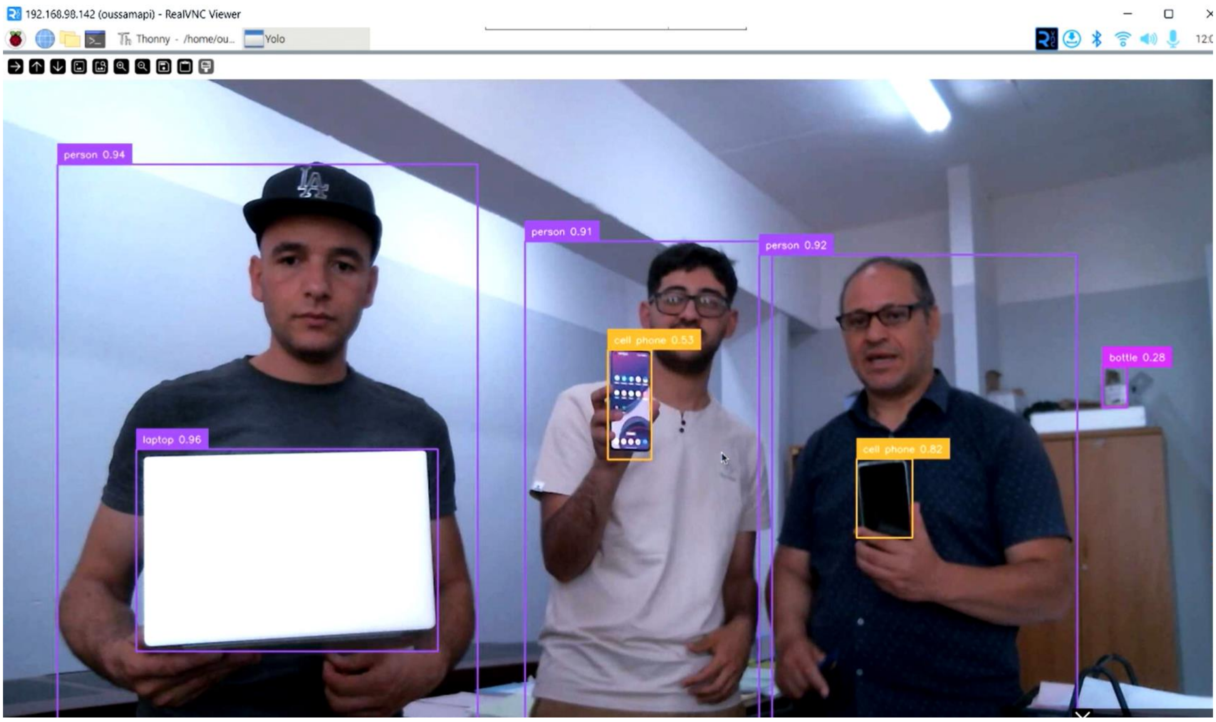

Figure III.20: Result of detection on Raspberry Pi.

III.5.3.Fire and smoke detection

The training process of our model is as follows:

- **Data preparation:**

Downloading the annotated database that contains three classes: fire, default, smoke.

The database is organized into three folders: train (training), test (testing), valid (validation).

The database is divided as follows:

  - 70% for training: These data help the model to learn and develop.
  - 20% for testing: Used to evaluate the model's accuracy after training.
  - 10% for validation: Used to adjust hyper-parameters and prevent overfitting.

A data.yaml configuration file that contains the path to the associated folders (train, test, validation) and the described classes.

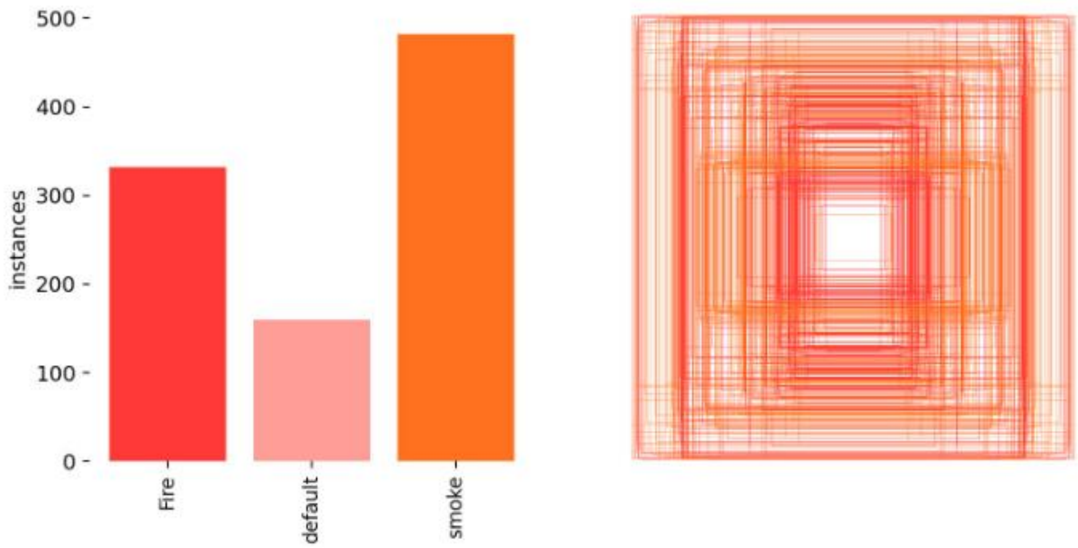

Figure III.21: The number of images and all the boxes used.

- **Development environment:**

Using Google Colab for training the model.

- **Pre-training of the model:**

Yolo was trained on COCO to have a pre-trained model. Then we took this model and trained it on the custom database (the one containing images of fire, default, smoke) so that in the end we have our own model.

In fact, we used the power of Yolo to create our own model.

- **Training parameters:**

Model used: YOLOv8n.

Number of epochs: 75.

Choosing the number of epochs:

The choice of the number of epochs during the training of a machine-learning model is crucial because it directly influences the model's performance on new data. Here is why:

- Avoid underfitting: With too few epochs, the model does not learn enough from the training data, which can result in poor performance.

- Avoid overfitting: Conversely, too many epochs can lead the model to not learn the details properly, reducing its ability to generalize well on new data.

- **Training launch:**

Initiating training with the specified parameters.

Obtaining error (loss), accuracy (precision), and recall curves towards the end of the training.

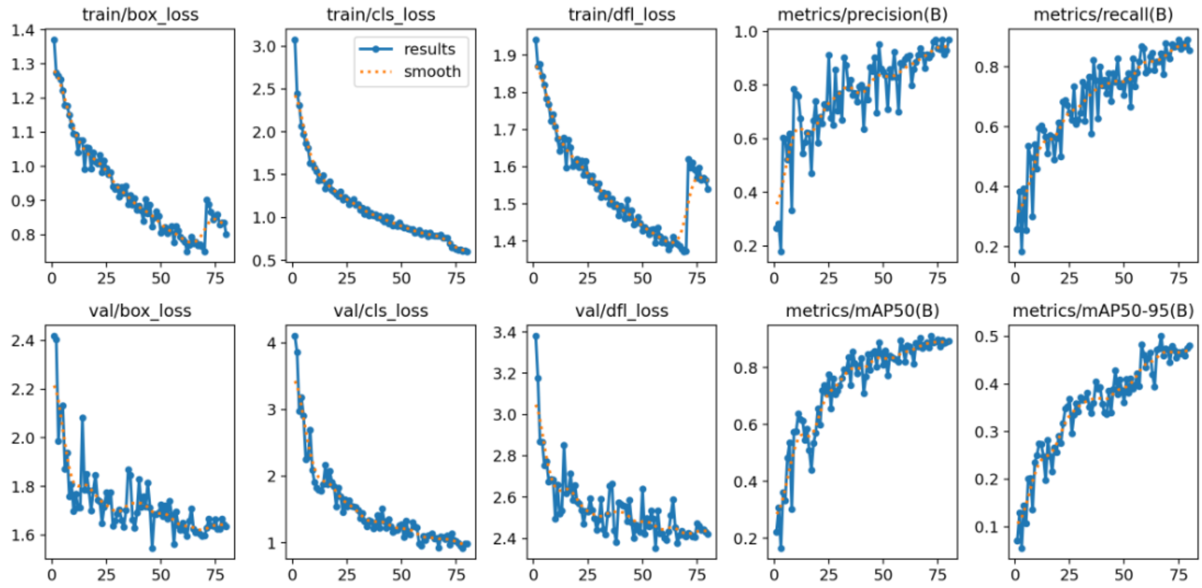

Figure III.22: Model training result.

Due to the low quality of the camera, in our code we set the confidence threshold to 60%, means that the model will only consider predictions with a confidence of 60% or higher. For example, an object identified at 40% confidence will be disregarded, while one at 70% confidence will be accepted.

We tested the model on many fire images, and it successfully recognized them. We also lit a real fire, and the model recognized it with very high confidence levels.

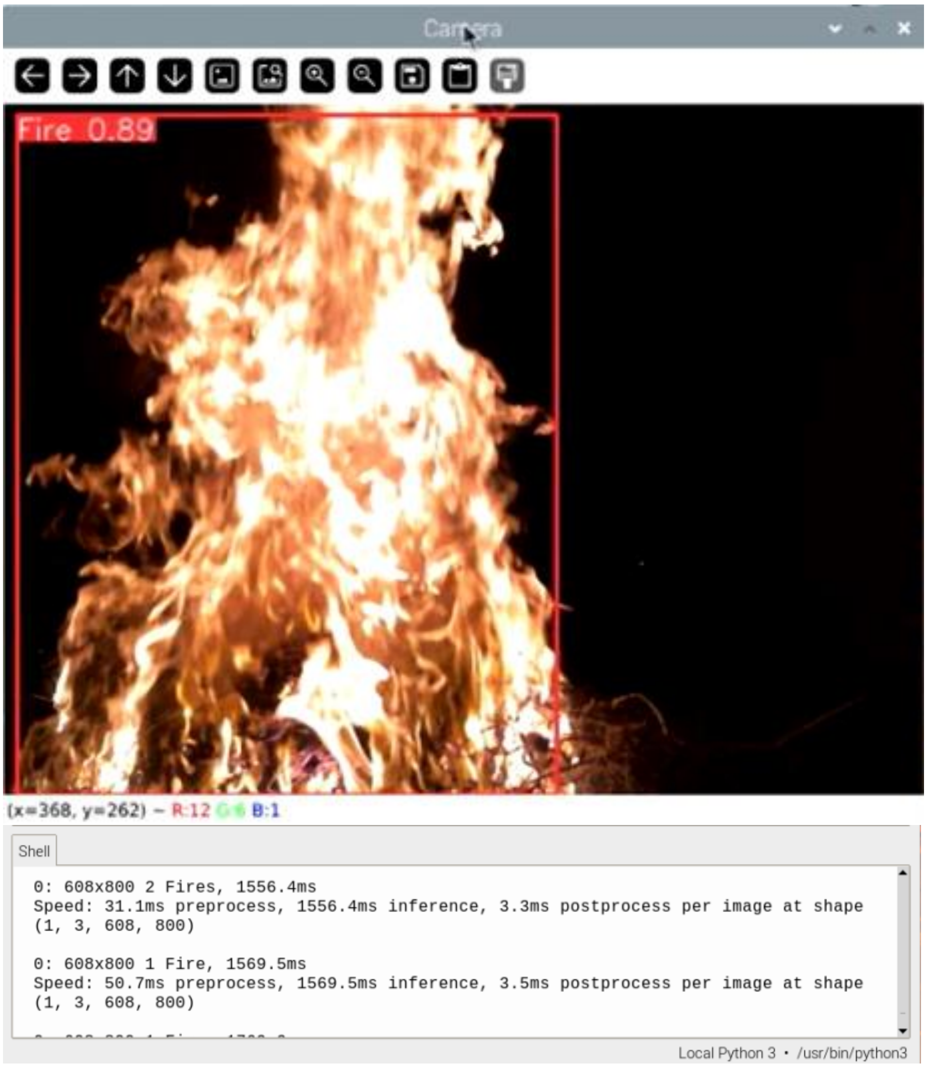

Figure III.23: Fire detection result.

The smoke detection test results were good. We tried real smoke in addition to images and videos, and it successfully recognized them with high accuracy.

Figure III.24: smoke detection result.

When it detects default, this means there is no fire nor smoke, but instead there are other things.

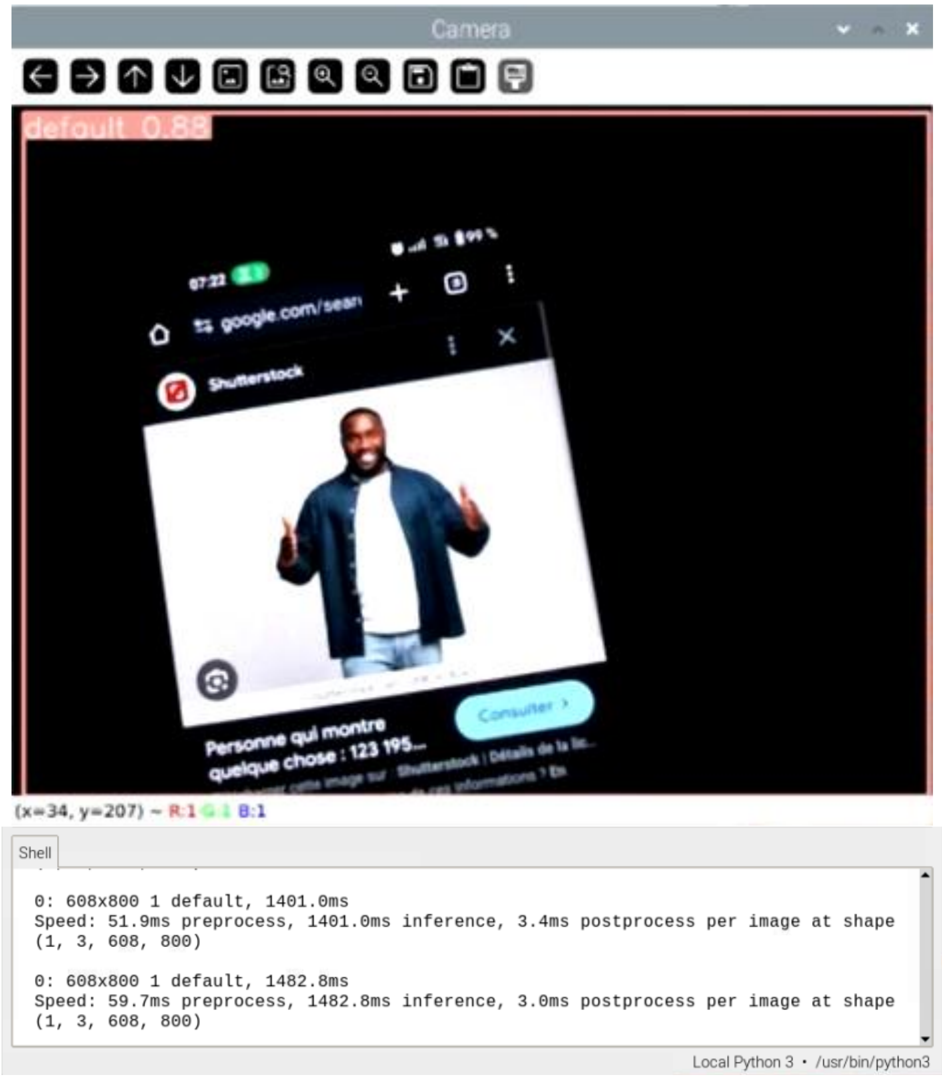

Figure III.25: default detection result.

III.5.4. Sending the results via VLC technology to the control room (PC)

At the beginning, we work on sending the result as a text from the raspberry pi to the Arduino, and then we send it using VLC technology.

III.5.4.1. Serial communication Raspberry Pi / Arduino (sending data to Arduino Tx)

It is possible to make a connection between the Raspberry Pi and Arduino and transfer data from the Raspberry Pi to the Arduino and vice versa. There are two methods of connection [70]:

- Serial communication via GPIO pins

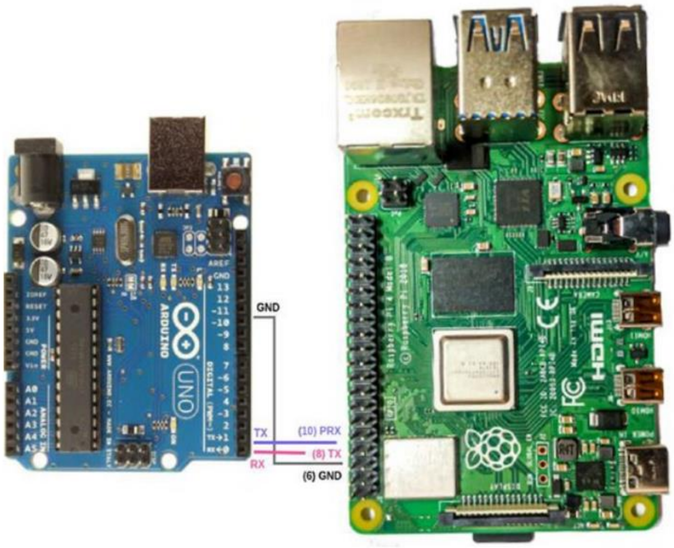

Figure III.26: Connection between Arduino and Raspberry Pi via GPIO pins.

- Serial communication via USB

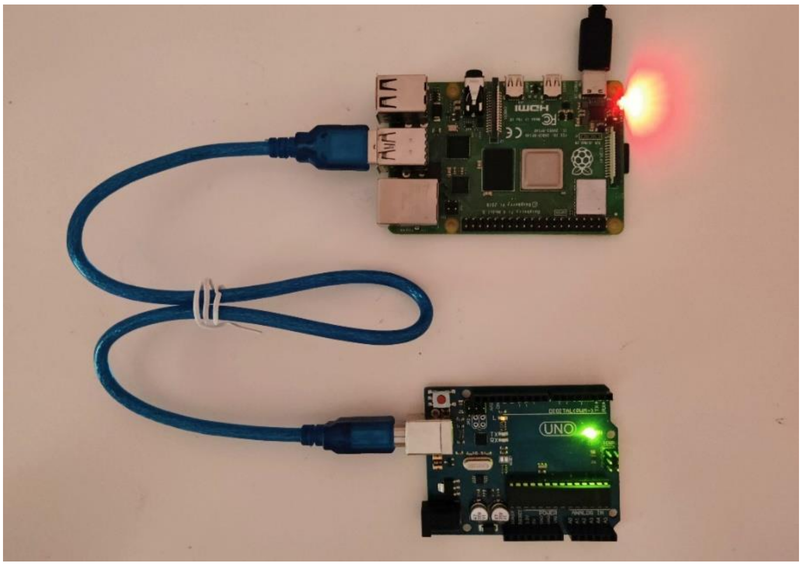

Figure III.27 Connection between Arduino and Raspberry Pi via USB cable.

When connecting the Arduino to the Raspberry Pi via a USB cable, the serial port must be specified by either installing the Arduino application on the Raspberry Pi or by entering the command " ls /dev/tty* " in the Raspberry Pi terminal to find the port file for the Arduino Uno. The port we found is "/dev/ttyACM0".

To send the detection result to the Arduino, we first display the class names of the items from the result:

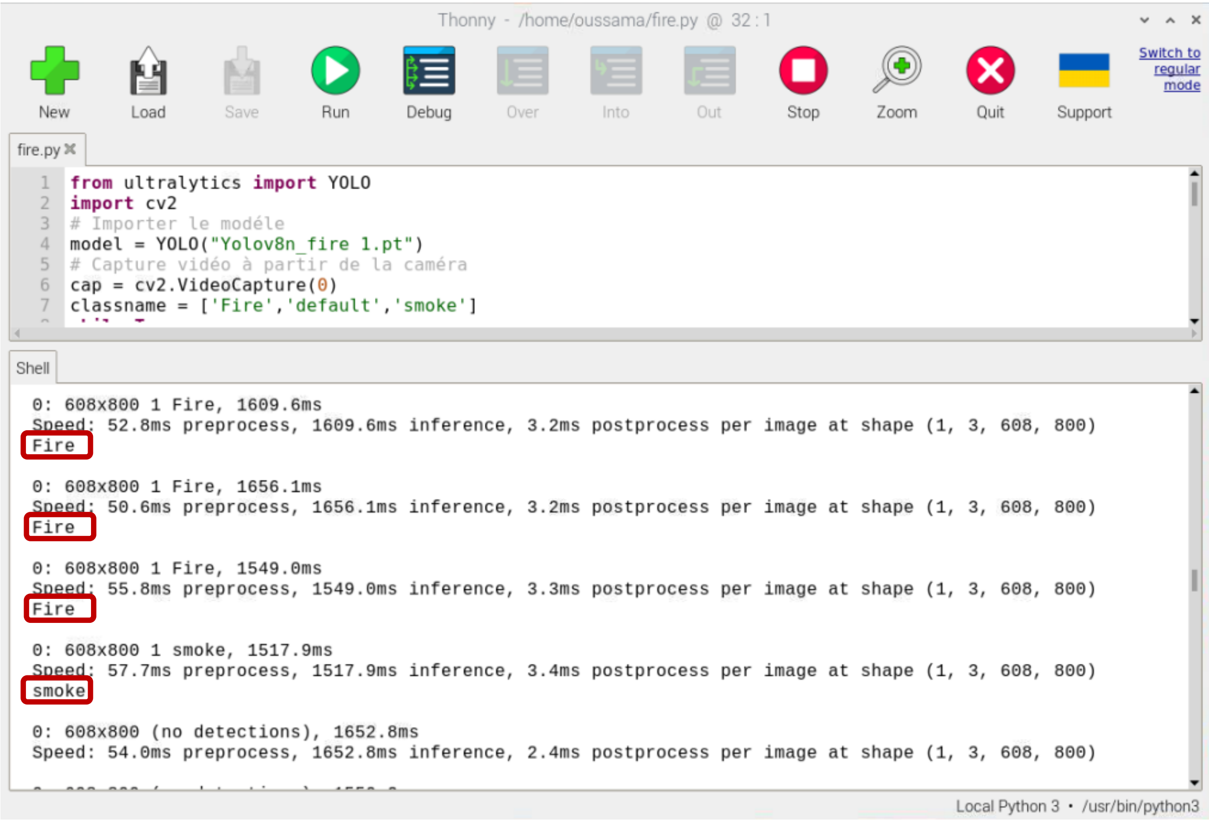

Figure III.28: An image showing the display of class names from the result.

The pySerial Python library creates a serial interface between Python and a serial port.

We use this line to open the serial connection through the /dev/ttyACM0 port at a speed of 115200 baud:

```
ser = serial.Serial('/dev/ttyACM0', 115200)
```

Sending to Arduino:

```
ser.write(name.encode())
```

ser: represents the serial port.

write: is the method used to send data.

name: is the text string you want to send.

encode (): converts the text string into bytes, because serial ports deal with bytes rather than text.

To receive the result as a text in Arduino and display it on the serial monitor:

Start the serial communication at a speed of 115200:

```
Serial.begin(115200);
```

Read the incoming data and return it as text until the end of the line:

```
incomingData = Serial.readStringUntil('\n');
```

Print the incoming data:

```
Serial.println(incomingData);
```

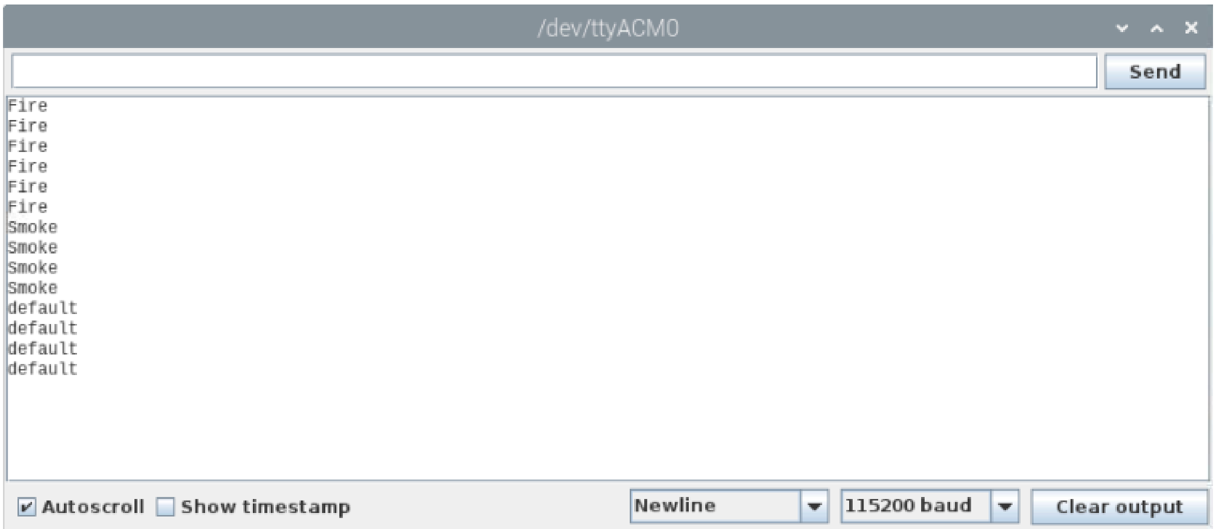

Figure III.29: Serial monitor of the Arduino Tx.

III.5.4.2. VLC communication (Arduino TX to Arduino RX)

In our work, we want to send the labels of detected items via LiFi technology after each detection. The Arduino LiFi transmitter code we used previously sends texts via LiFi but requires opening the serial monitor and manually entering the text for transmission.

We modify the LiFi transmitter code to receive detection results from a Raspberry Pi and send them automatically. The result appears immediately after each detection on the receiver's serial monitor.

III.5.4.3. Implementation on the robot

Now, we combine all the work we have done with VLC car:

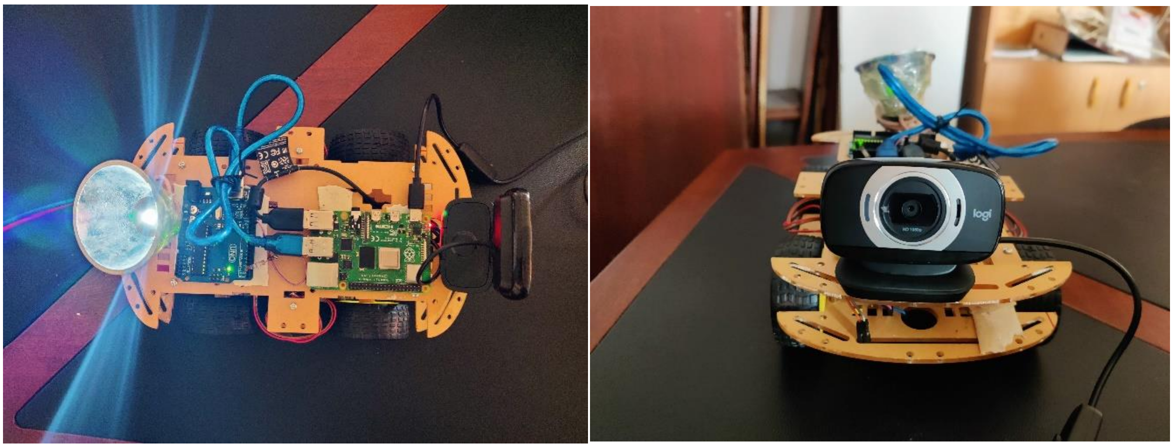

Figure III.30: VLC-car.

To implement our idea in reality, we placed a video containing fire and smoke in front of the robot’s camera. After the detection, the result is sent as a text directly via LiFi technology to the receiving Arduino, and then it appears on the serial monitor.

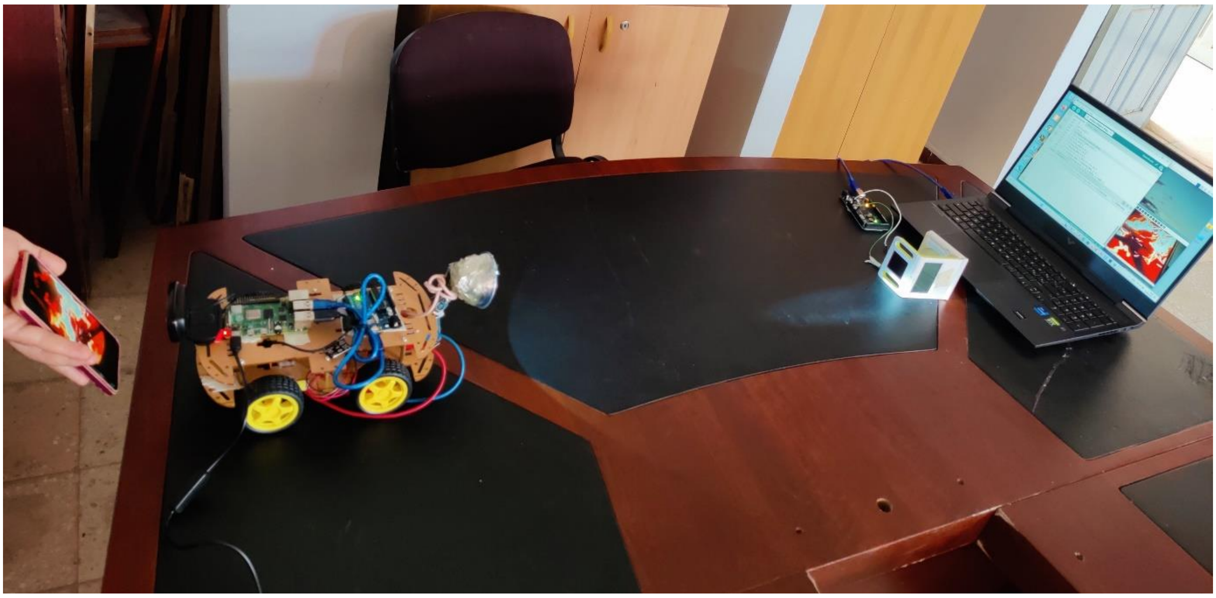

Figure III.31: Implementation of our idea.

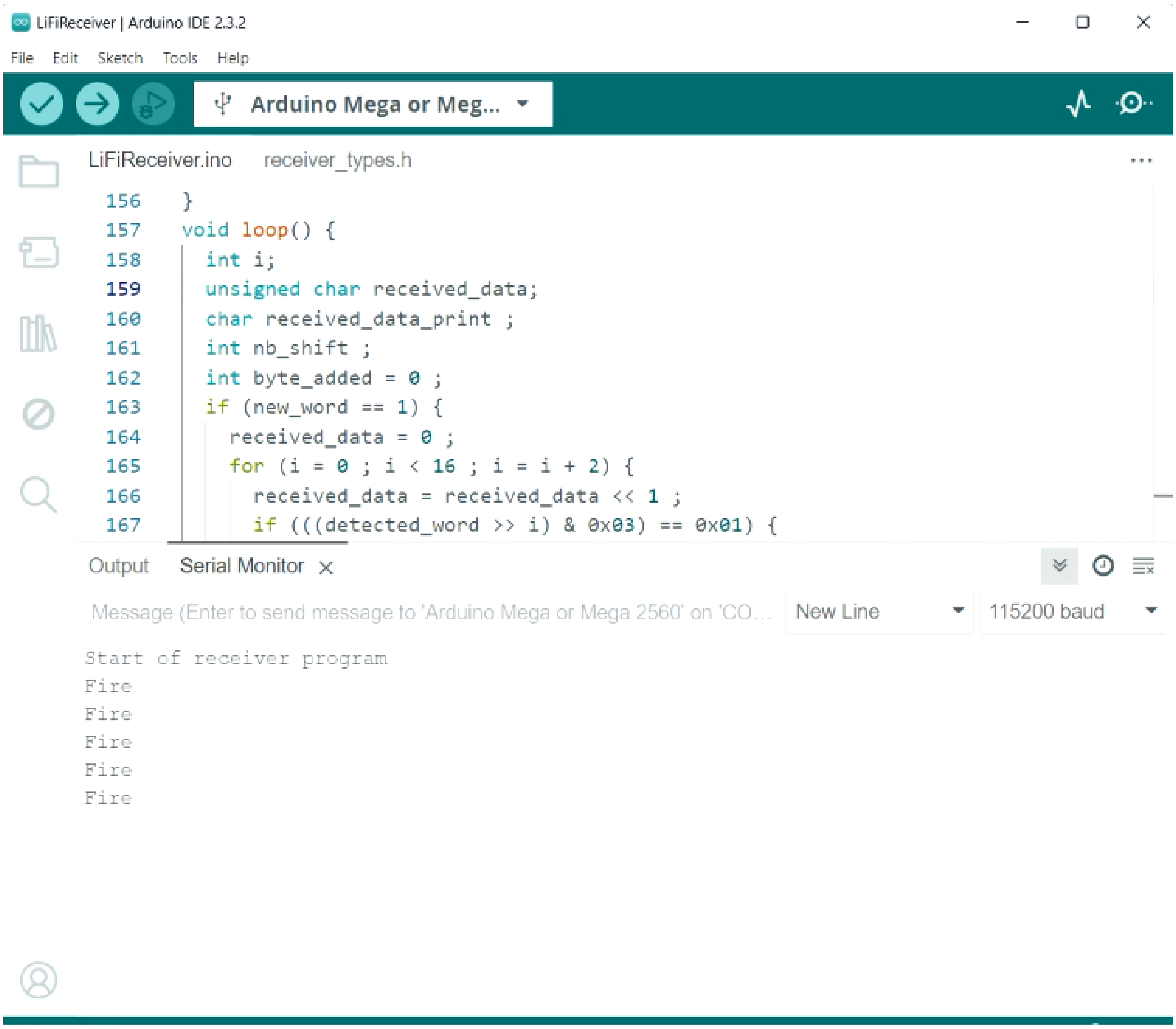

Figure III.32: Serial monitor of the receiver module.

III.6. Conclusion

In this chapter, we studied and configured the hardware and software related to our project. We implemented a wireless data transmission system based on Li-Fi technology for sending texts. We worked on a general object detection project and a specific fire detection project using a Raspberry Pi and a webcam. The results are sent to an Arduino (transmitter/receiver) serially connected to the Raspberry Pi embedded in the robot, and then the results are transmitted via LiFi technology to the receiving Arduino, displaying the results upon each detection.

General conclusion

In light of the rapid developments in the field of artificial intelligence and robotics, visible light communication (VLC) technology can also play a crucial role. VLC technology has emerged as an innovative solution that goes beyond lighting to include efficient data transmission. It is characterized by its reliance on LED lamps, which makes it widespread and significantly reduces implementation costs. Additionally, it offers features such as high bandwidth, no interference with radio waves in electromagnetically sensitive areas, and no health risks. Some examples of its applications include Li-Fi, hospitals, underwater communications, and sensitive stations such as power plants, petrochemical plants, and nuclear stations.

In the first chapter, we began by providing an overview of artificial intelligence and its branches, as well as robotics, specifically mobile robots. We discussed wireless communications with a focus on visible light communication, its structure, and its key applications.

In the second chapter, we covered the methods used for motion detection and tracking, along with the most popular object detection algorithms.

In the third chapter, we learned how to train the robot to detect fires in real-time using a Raspberry Pi 4 and a webcam. We used Arduino as a microcontroller to send the detection results via VLC technology to the control room (PC). We obtained good results despite the limited devices available to us.

This project represents an innovative step towards utilizing modern technology in safety and security fields. The VLC-equipped robot will contribute to improving emergency response and reducing fire losses through early detection and immediate communication.

Here are some points that could form perspectives for our future research:

- Determining the coordinates and location of the fire with using navigation.
- Adding a water tank to the robot for immediately intervention.
- Remote control for operating the water pump.

REFERENCES

[1] Biswal, A. "Types of Artificial Intelligence." Simplilearn. Accessed March 2024. Available at: https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/types-of-artificial-intelligence

[2] Anyoha, Rockwell. "The History of Artificial Intelligence." Science in the News, Harvard University, 28 Aug. 2017. Accessed March 2024. Available at: https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

[3] IBM. "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks." IBM Blog. Accessed March 2024. Available at: https://www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks/

[4] NetApp. "What is Artificial Intelligence?" NetApp. Accessed March 2024Available at: https://www.netapp.com/fr/artificial-intelligence/what-is-artificial-intelligence/

[5] Oracle. "What is AI?" Oracle. Accessed March 2024. Available at: https://www.oracle.com/fr/artificial-intelligence/what-is-ai/

[6] Web3Cafe. "Types of Artificial Intelligence You Should Know in 2023." Web3Cafe, 6 June 2023. Accessed 17 June 2024. Available at: https://www.web3cafe.in/artificial-intelligence/story/types-of-artificial-intelligence-you-should-know-in-2023-571325-2023-06-06

[7] PraxiLabs. "Role of Artificial Intelligence in Virtual Labs." PraxiLabs Blog, 19 April 2023. Accessed 17 June 2024. Available at: https://praxilabs.com/en/blog/2023/04/19/role-of-artificial-intelligence-in-virtual-labs/

[8] DataCamp. "What is Machine Learning?" DataCamp Blog. Accessed March 2024. Available at: https://www.datacamp.com/blog/what-is-machine-learning

[9] Edureka. "Artificial Intelligence with Python." Edureka Blog. Accessed March 2024. Available at: https://www.google.com/amp/s/www.edureka.co/blog/artificial-intelligence-with-python/amp/

[10] Ayodele Taiwo Oladipupo. "Types of machine learning algorithms." New advances in machine learning 3.19-48 (2010): 5-1.

[11] Benkirat Mohamed Achraf and Bouguerne Omar. "Integration of VLC technology into the smart factory and industry (4.0/5.0)." Master’s thesis. University of 8 Mai 1945. (2023).

[12] Menai Aymen. "Système intelligent de monitoring et de commande d'un système MIMO a base la technologie VLC." Master’s thesis. University of 8 Mai 1945. (2022).

[13] Azamat Abdoullaev, BBN Times. "Artificial Intelligence vs. Machine Learning vs. Artificial Neural Networks vs. Deep Learning." BBN Times. Accessed March 2024. Available at: https://www.google.com/amp/s/www.bbntimes.com/science/artificial-intelligence-vs-machine-learning-vs-artificial-neural-networks-vs-deep-learning/amp

[14] Diana Ramos, "Neural Network Applications." Smartsheet. October 17, 2018 Accessed March 2024. Available at: https://www.smartsheet.com/neural-network-applications

[15] Amazon Web Services. "The Difference Between Deep Learning and Neural Networks." AWS. Accessed March 2024. Available at: https://aws.amazon.com/fr/compare/the-difference-between-deep-learning-and-neural-networks/

[16] Pawar, P. M., Balasubramaniam, R., Ronge, B. P., Salunkhe, S. B., Vibhute, A. S., & Melinamath, B. (Eds.). (2020). Techno-Societal 2020: Proceedings of the 3rd International Conference on Advanced Technologies for Societal Applications—Volume 1.

[17] Dataquest. "Introduction to Deep Learning." Dataquest Blog. Accessed March 2024. March 2024. Available at: https://www.dataquest.io/blog/tutorial-introduction-to-deep-learning/

[18] Djemaa Mahir. "Classification des images par CNN." Master’s thesis, Université 8 Mai 1945 Guelma, 2023.

[19] Fatai Anifowose, Society of Petroleum Engineers (SPE). "Deep Learning and Its Applications in the Energy Industry." SPE Journal of Petroleum Technology. Accessed March 2024. Available at: https://jpt.spe.org/deep-learning-and-its-applications-in-the-energy-industry?gad_source=1&gclid=CjwKCAiA2pyuBhBKEiwApLaIO1KtHbwSBJoK0QXHPo2c3lyeUMz8vvsXViT2OVH_TNUaZz95zhRFDhoCzN4QAvD_BwE

[20] GeeksforGeeks. "Difference Between a Neural Network and a Deep Learning System." GeeksforGeeks. Accessed March 2024. Available at: https://www.geeksforgeeks.org/difference-between-a-neural-network-and-a-deep-learning-system/

[21] Ahcene Hamoudi and Karim Berkani. "Conception et réalisation d'un robot mobile autonome." Master's thesis. Université Mouloud Mammeri, TIZI-OUZOU . (2016).

[22] Boudjelkha Ichrak . "Conception et réalisation d'un robot mobile fennec camouflé basé sur l'IoT pour des applications militaires." Master's thesis. Université Mohamed Khider – BISKRA. (2022).

[23] Kate Brush, TechTarget. "Mobile Robotics." IoT Agenda. March 2024. Available at https://www.techtarget.com/iotagenda/definition/mobile-robot-mobile-robotics

[24] David Filliat. "Robotique Mobile." Ecole Nationale Supérieure de Techniques Avancées Paris Tech. (2012).

[25] Slimane Noureddine. " Système de localisation pour robots mobiles." Doctoral dissertation. Université de Batna 2. (2008) .

[26] Devopedia. "Mobile Robot." Devopedia. Accessed 17 June 2024. Available at: https://devopedia.org/mobile-robot

[27] Guebli Imed Eddine. "Navigation et commande d'un robot mobile à 4 roues sous ROS." Master's thesis, Université 8 Mai 1945 Guelma. (2020).

[28] Zekhref Oussama, Nouioua Med Islam. " Localisation et navigation d'un robot mobile par webcam. " Master's thesis, Université Blida 1. (2015).

[29] Moozakis, C., & Kerner, S. M. TechTarget. "Wireless Communication." Search Mobile Computing, (2023). Accessed March 2024. Available at: https://www.techtarget.com/searchmobilecomputing/definition/wireless

[30] Sellaoui Raid Houssem Eddine. "Commande d'un robot à base de la technologie VLC." Master's thesis , Université 8 Mai 1945 Guelma. (2021).

[31] Teja, R. Electronics Hub. "Wireless Communication: Introduction, Types, Applications." Electronics Hub. Accessed March 2024. Available at: https://www.electronicshub.org/wireless-communication-introduction-types-applications/

[32] Wikimedia Commons. "Electromagnetic Spectrum." Wikimedia Commons. March 2024. Available at: https://commons.wikimedia.org/wiki/File:Onde_electromagnetique.svg

[33] Khan Academy. "Light and the Electromagnetic Spectrum." Khan Academy. Accessed March 2024. Available at: https://www.khanacademy.org/science/physics/light-waves/introduction-to-light-waves/a/light-and-the-electromagnetic-spectrum

[34] Chaabna Ameur. ”Système de Communication par la Lumière Visible (VLC) : Architecture simplifiée et efficacité énergétique d'un système de positionnement." Doctoral dissertation, Université 8 Mai 1945 Guelma. (2019).

[35] Chouabia Halim. ”Système MIMO multifonctionnel dans le concept VLC. ” Doctoral dissertation, Université 8 Mai 1945 Guelma. (2023).

[36] BYJU'S. "Electromagnetic Spectrum and Electromagnetic Waves." BYJU'S. Accessed March 2024. Available at: https://byjus.com/jee/electromagnetic-spectrum-and-electromagnetic-waves/

[37] Alin-Mihai Cailean. ”Study, implementation and optimization of a visible light communications system. Application to automotive field. ” Doctoral dissertation, l’Université de Versailles Saint-Quentin-en-Yvelines . (2014).

[38] Hamada Louiza, ”Conception d’une architecture supportant la technologie LI-FI. ” Doctoral dissertation. Université de Haute Alsace-Mulhouse. (2022).

[39] Affoune Oussama, Naidja Abderaouf, ” Visible light communication : Etude des performances d’un système MIMO. ” Master’s thesis, Université 8 Mai 1945 Guelma. (2020).

[40] Ertunç, Ezgi, ” Non-line of Sight Visible Light Communications, ” Master's thesis. České vysoké učení technické v Praze. Vypočetní a informační centrum. (2018).

[41] Durgesh Gujjari, ” VISIBLE LIGHT COMMUNICATION. ” Master degree of Applied Science Dalhousie, University Halifax, Nova Scotia. (2012).

[42] Sharma Rahul R, Raunak, and Akshay Sanganal,“Li-Fi Technology: Transmission of data through light”. International Journal Computer Technology & Applications (IJCTA) 5.1 (2014): 150-154.

[43] RF Wireless World. "LiFi vs. VLC." RF Wireless World. Accessed March 2024. Available at: https://www.rfwireless-world.com/Terminology/LiFi-vs-VLC.html

[44] Alfattani, Safwan. "Review of LiFi technology and its future applications." *Journal of Optical Communications* 42.1 (2021): 121-132.

[45] Diambeki, D. D., et al. "Securing the light escaping in a Li-Fi network environment." Procedia Computer Science 201 (2022): 684-689.

[46] Nakajima Madoka and Shinichiro Haruyama, "New indoor navigation system for visually impaired people using visible light communication." *EURASIP Journal on Wireless Communications and Networking* 2013 (2013): 1-10.

[47] Khan Latif Ullah, "Visible light communication: Applications, architecture, standardization and research challenges." *Digital Communications and Networks* 3.2 (2017): 78-88.

[48] Singh Sukhvir, Gholamreza Kakamanshadi and Savita Gupta, "Visible light communication-an emerging wireless communication technology." *2015 2nd International conference on recent advances in engineering & computational sciences (RAECS)*. IEEE, 2015.

[49] Sadoun abdelbaki and Ouellabi yacine, "Détection et suivi d'objets mobiles. Application dans un environnement de foule", Master's thesis, Université Echahid Hamma Lakhdar D'El Oueud, (2016).

[50] Nouria Hammad and Hocini Oulhadj, "*Etude et realisation d'un système de detection et suivi d'un objet en mouvement* ", Master's thesis, Université Mouloud Mammeri, (2017).

[51] Saagie. "Object Detection: Part 1." Saagie Blog. Accessed April 2024. Available at: https://www.saagie.com/en/blog/object-detection-part1/

[52] Ouazani Yazid, "Etude et réalisation d'un système de détection et de suivi d'une personne en mouvement", Doctoral dissertation. Université Mouloud Mammeri, Tizi-Ouzou, (2018).

[53] Belhaddad Samir and Menai Mohammed, "Détection de suivi automatique des véhicules en trafic routier et autoroutier", Master's thesis, Université de 8 Mai 1945 Guelma, (2017).

[54] Viso AI. "Object Tracking with Deep Learning." Viso AI Blog. Accessed April 2024. Available at: https://viso.ai/deep-learning/object-tracking/

[55] Yasmine Amara and Akliouat Lynda, " *Détection et suivi automatiques d'objets en mouvement* ", Master's thesis, Université Mouloud Mammeri, Tizi-Ouzou, (2015).

[56] Zoumana keita, "YOLO: Object Detection Explained." DataCamp Blog. Accessed April 2024. Available at: https://www.datacamp.com/blog/yolo-object-detection-explained

[57] Andrey Germanov, free "How to Detect Objects in Images using YOLOv8." freeCodeCamp, MAY 4, 2023. Accessed April 2024. Available at: https://www.freecodecamp.org/news/how-to-detect-objects-in-images-using-yolov8/

[58] James Gallagher, "Object Detection." Roboflow Blog, AUG 22, 2023 Accessed April 2024. Available at: https://www.google.com/amp/s/blog.roboflow.com/object-detection/amp/

[59] Blent. "Detection of Images with YOLO in TensorFlow." Blent AI Blog Publié le 8 juin 2022. Accessed April 2024. Available at: https://blent.ai/blog/a/detection-images-yolo-tensorflow

[60] Alberto Rizzoli, V7 Labs. "Object Detection Guide." V7 Labs Blog. Accessed April 2024. Available at: https://www.v7labs.com/blog/object-detection-guide

[61] Rohit Kundu, "YOLO: Object Detection." V7 Labs Blog, January 17, 2023. Accessed May 2024. Available at: https://www.v7labs.com/blog/yolo-object-detection

[62] Encord. "Object Detection." Encord Blog. Accessed May 2024. Available at: https://encord.com/blog/object-detection/

[63] Détection d'objet en temps réel par Deep learning (YOLOv4) avec OpenCv et Python YouTube. Accessed May 2024. Available at: https://youtu.be/M37artO8elY?si=QbLrFHpgeMjmzN9U

[64] Terven Juan and Diana Cordova-Esparza. "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond." *arXiv preprint arXiv:2304.00501* (2023).

[65] Jiang, Peiyuan, et al, "A Review of Yolo algorithm developments." *Procedia computer science* 199 (2022): 1066-1073.

[66] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. (2016).

[67] Ravikiran A S. "What is Raspberry Pi?" Simplilearn Tutorials. Accessed May 2024. Available at: https://www.simplilearn.com/tutorials/programming-tutorial/what-is-raspberry-pi#what_is_raspberry_pi

[68] Real Python. "Python on Raspberry Pi: A Beginner's Guide." Real Python. May 2024.. Available at: https://realpython.com/python-raspberry-pi/

[69] What is python? Why python is so popular YouTube. Accessed May 2024. Available at: https://youtu.be/Y8Tko2YC5hA?si=B47j2TugdRDlYz_9

[70] Abdul Kadhar, K. M., & Anand, G. (2021). Data Science with Raspberry Pi: Real-Time Applications Using a Localized Cloud (pp. 79-96). Apress. ISBN: 9781484268254.