People´s Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

University 20 August -1955-Skikda

Faculty of Science

Department of Computer Science

Dissertation Submitted in partial fulfillment of the requirements for the master´s degree in computer science + Startup Certificate under Ministerial Decree 1275.

Option: Artificial Intelligence (I, A)

THEME

Intelligent Search Submarine Using Artificial Intelligence

Presented by :

SACI Nedjla

LARAFA Anfal

BENNIKOUS Sirina

KHARFANE Mohammed Ali

Supervised by :

Dr. BOULEHOUACHE Soufiane

2025/2026

Edit with WPS Office

Acknowledgem

First and foremost, we express our deepest gratitude to Almighty God for His guidance and countless blessings throughout this journey. Without His divine support, our success would not have been possible. We also wish to thank our supervisor,

Dr. BOULAHOUACHE Sofiane, for his knowledge, experience, advice, and guidance, thanks to which we were able to produce high-quality work in such a short time. We would like to thank our families for their unwavering psychological support, encouragement, and the resources they provided, all of which played a vital role in our success. We extend our gratitude to all the teaching staff of the Computer Science Department for their dedication and work, which contributed greatly to our learning.

Edit with WPS Office

Abstract

The rapid development of Artificial Intelligence and embedded technologies has significantly contributed to the creation of intelligent systems capable of operating in complex environments. Underwater search is one of the fields that can benefit from these technologies, as underwater environments are often characterized by reduced visibility, communication difficulties, and limited human intervention.

This thesis presents the design and development of a low-cost intelligent Remotely Operated Vehicle (ROV) prototype dedicated to underwater object search. The proposed system combines embedded electronics, remote control, computer vision, and Artificial Intelligence to assist the operator during underwater search missions.

The ROV is remotely controlled by the operator, while the vision system analyzes images captured by the onboard camera and detects objects of interest using a YOLO-based object detection model. The entire system is built around a Raspberry Pi 5, which serves as the main processing unit and is responsible for camera management, motor control, video processing, and the execution of the Artificial Intelligence model. The Raspberry Pi 5 also manages communication between the ROV and the external control station.

The current prototype is designed as a proof of concept, demonstrating the feasibility of integrating Artificial Intelligence into a low-cost underwater robotic platform for object search applications. This work highlights the potential of intelligent underwater systems to support search tasks and provides a foundation for future improvements, including enhanced communication methods, improved waterproofing, more powerful detection models, and advanced real-time processing capabilities.

Keywords: Artificial Intelligence (AI), Computer Vision (CV) , Remotely Operated Vehicles (ROV) , Deep learning (DL) , Convolutional Neural Networks (CNN), Optical Character Recognition (OCR), Decision Support Systems (DSS) , YOLO (You Only Look Once).


Edit with WPS Office

Resume

Le developpement rapide de l´intelligence artificielle et des technologies embarquees a considerablement contribue a la creation de systemes intelligents capables d´evoluer dans des environnements complexes. La recherche sous-marine constitue l´un des domaines pouvant tirer profit de ces technologies, car les environnements aquatiques sont souvent caracterises par une faible visibilite, des difficultes de communication et une intervention humaine limitee.

Ce memoire presente la conception et le developpement d´un prototype de vehicule sous-marin teleopere (ROV) intelligent a faible cout destine a la recherche d´objets sous l´eau. Le systeme propose combine l´electronique embarquee, le controle a distance, la vision par ordinateur et l´intelligence artificielle afin d´assister l´operateur lors des missions de recherche sous-marine.

Le ROV est pilote a distance par un operateur, tandis que le systeme de vision analyse les images capturees par la camera embarquee et detecte les objets d´interet a l´aide d´un modele de detection d´objets base sur l´algorithme YOLO. L´ensemble du systeme repose sur une carte Raspberry Pi 5, qui constitue l´unite principale de traitement et assure la gestion de la camera, le controle des moteurs, le traitement video ainsi que l´execution du modele d´intelligence artificielle. La Raspberry Pi 5 assure egalement la communication entre le ROV et la station de controle externe.

Le prototype developpe constitue une preuve de concept (Proof of Concept) demontrant la faisabilite de l´integration de l´intelligence artificielle dans une plateforme robotique sous-marine a faible cout destinee aux applications de recherche d´objets. Ce travail met en evidence le potentiel des systemes sous-marins intelligents pour assister les operations de recherche et constitue une base solide pour de futurs developpements, notamment l´amelioration des moyens de communication, du systeme d´etancheite, des modeles de detection et des capacites de traitement en temps reel.

Mots-cles : Intelligence Artificielle (AI), Vehicules Telecommandes (ROVs), Vision par Ordinateur (CV), Apprentissage Automatique (ML), Apprentissage Profond (DL), Reseaux de Neurones Convolutifs (CNN), Reconnaissance Optique de Caracteres (OCR), Systemes d´Aide a la Decision (DSS), YOLO (You Only Look Once).

الملخص


Edit with WPS Office

ساهم التطور السريع في مجال الذكاء الاصطناعي والتقنيات المدمجة في ظهور أنظمة ذكية قادرة على العمل في بيئات معقدة. ويُعد البحث تحت الماء من المجالات التي يمكن أن تستفيد بشكل كبير من هذه التقنيات، نظراً لما تتميز به البيئات المائية من انخفاض في مستوى الرؤية، وصعوبة في الاتصالات، ومحدودية التدخل البشري.

تتناول هذه المذكرة تصميم وتطوير نموذج أولي **(Prototype)** لغواصة ذكية منخفضة التكلفة يتم التحكم فيها عن بعد**(ROV)** ، مخصصة للبحث عن الأجسام تحت الماء. ويجمع النظام المقترح بين الإلكترونيات المدمجة، والتحكم عن بعد، والرؤية الحاسوبية، وتقنيات الذكاء الاصطناعي بهدف مساعدة المستخدم أثناء عمليات البحث في البيئات المائية.

يتم التحكم في الغواصة عن بعد بواسطة المستخدم، بينما يقوم نظام الرؤية الحاسوبية بتحليل الصور الملتقطة بواسطة الكاميرا المثبتة على الغواصة، واكتشاف الأجسام المستهدفة باستخدام نموذج لا كتشاف الأجسام يعتمد على خوارزمية **YOLO**. ويعتمد النظام بالكامل على لوحة **Raspberry Pi 5** التي تمثل وحدة المعالجة الرئيسية، حيث تتولى إدارة الكاميرا، والتحكم في المحركات، ومعالجة الفيديو، وتشغيل نموذج الذكاء الاصطناعي، بالإضافة إلى إدارة الاتصال بين الغواصة ووحدة التحكم الخارجية.

ويمثل النموذج المطور إثباتاً للمفهوم**(Proof of Concept)** ، يهدف إلى إبراز إمكانية دمج تقنيات الذكاء الاصطناعي في منصة روبوتية مائية منخفضة التكلفة مخصصة للبحث عن الأجسام تحت الماء. كما يبرز هذا العمل الإمكانات الواعدة للأنظمة الذكية تحت الماء في دعم عمليات البحث، ويمثل أساساً يمكن تطويره مستقبلا ً من خلال تحسين وسائل الاتصال، وتعزيز مقاومة الماء، واستخدام نماذج ذكاء اصطناعي أكثر تطوراً، وتحقيق معالجة وعرض للنتائج في الزمن الحقيقي.

الكلمات المفتاحية: الذكاء الاصطناعي **(AI)** ، المركبات المُشغّلة عن بُعد **(ROVs)** ،الرؤية الحاسوبية **(CV)** ،تعلم الآلة **(ML)** ،التعلم العميق **(DL)** ،الشبكات العصبية الالتفافية **(CNN)** ،التعرّف الضوئي على الحروف **(OCR)** ،أنظمة دعم القرار **(DSS)**،، **YOLO (You Only Look Once)**

Edit with WPS Office

Table of content

Introduction generale	01
Chapitre 1 :	
1. Introduction	07
2.Characteristics of the Underwater Environment	07
2.1 Physical Characteristics	07
2.2 Chemical Characteristics	08
2.3 Biological and Ecological Systems	09
3. Visibility Challenges	10
4 . ROVs	11
4.1 Definition of ROVs	11
4.2 Historical Development of ROVs	11
4.3 General Architecture of ROV Systems	12
4. 4 Types of ROVs	13
5. Existing ROV Systems	13
5.1 Jason ROV	13
5.2 ROV SuBastian	14
5.3 BlueROV2	15
5.4 VideoRay ROV	16
6. Conclusion	17
Chapter 2 :	
1. Introduction	20
2. Introduction to Artificial Intelligence	21
2.1. Definition of artificial intelligence	21
2.2. Domains of Artificial Intelligence	21
2.3 Role of Artificial Intelligence in Problem Solving	22
3. Machine Learning	24
3.1. Definition of Machine Learning	24
3.2. Types of Machine Learning	24
3.3. Machine Learning Workflow	26
4. Deep Learning	27
4.1. Definition of Deep Learning	27

Edit with WPS Office

4.2. Neural Networks	27
4.3. Advantages of Deep Learning over Machine Learning	28
5. Convolutional Neural Networks (CNN)	29
5.1. CNN Architecture	29
5.1.1 Convolution Layer	29
5.1.2 Pooling Layer	30
5.1.2 Fully Connected Layer	30
6. Computer Vision	31
6.1. Definition of Computer Vision	31
6.2. Tasks of Computer Vision	31
7. Object Detection Techniques	33
7.1. Traditional Object Detection Methods	33
7.2. Deep Learning-Based Detection Methods	33
8. YOLO Algorithm	34
8.1. Introduction to YOLO	34
8.2. YOLO Architecture	35
8.3. How YOLO Works	36
8.4. Advantages of YOLO	36
9. Underwater Image Processing Challenges	37
9.1. Light Absorption	37
9.2. Color Distortion	37
9.3. Low Visibility	37
9.4. Noise	38
10. Conclusion	39
Chapter 3 :	
1. Introduction	41
2 Overview of the Proposed System	41
3 Hardware Architecture	43
3.1 3D Printed Parts	43
3.2 Acrylic Waterproof Body	44
3.3 Front and Rear End Caps	44
3.4 Propellers	45

Edit with WPS Office

3.5 DC Motors	45
3.7 Motor Driver Module	45
3.8 Raspberry Pi 5	46
3.9 MPU6050 IMU	46
3.10 Pressure Sensing Unit	46
3.11 18650 Batteries	47
3.12 DC–DC Converters	47
3.13 Syringes and Tubes	47
3.14 VL53L0X Time-of-Flight Sensor	47
3.15 Internal Arrangement of Components	48
3.16 Mechanical Operation of the ROV	49
3.16.1 Propulsion Mechanism	49
3.16.2 Buoyancy Control Mechanism	50
4 Software Architecture	52
4.1 Programming Environment	53
4.2 Video Streaming System	54
4.3 AI Detection Module	54
5 Communication System	54
5.1 Remote Control System	55
5.2 Data Transmission	56
6 System Workflow	56
Conclusion	57
Chapter 4 :	
1. Introduction	60
2. Dataset Collection	61
2.1 Selection of the Underwater Dataset	61
2.2 Characteristics of the Selected Dataset	61
2.3 Underwater Object Categories	62
2.4 Dataset Organization	63
2.5 Importance of the Dataset for the Proposed ROV	64
3. Data Preprocessing	65

Edit with WPS Office

3.1 Dataset Preparation	65
3.2 Dataset Configuration	65
3.3 Image Preparation	66
3.4 Data Augmentation	67
3.5 Contribution of Preprocessing to the Underwater Exploration System	67
4. Underwater Data Annotation and AI Configuration	68
4.1 Instance Segmentation Annotation	68
4.2 YOLOv8 Annotation Format	68
4.3 Dataset Configuration using YAML	69
4.4 Contribution of Annotation to the AI Vision Module	69
5. Development and Training of the AI Vision Module	70
5.1 Development Environment	70
5.2 Transfer Learning Strategy	70
5.3 Selection of YOLOv8 Models	71
5.4 Training Configuration	71
5.5 Training Procedure	72
5.6 Training Outputs	72
6. Performance Evaluation of the AI Vision Module	73
6.1 Evaluation Metrics	73
6.2 Precision	74
6.3 Recall	75
6.4 Mean Average Precision (mAP)	75
6.5 Analysis of the Training Curves	75
6.6 Precision–Recall Curve	76
6.7 Confusion Matrix	76
6.8 Selection of the Final AI Model	77
7. Experimental Results and Discussion	78
7.1 Analysis of the Training Results	78
7.2 Performance of Underwater Object Recognition	78
7.3 Discussion of the Obtained Results	79
7.4 Contribution of Artificial Intelligence to the Proposed ROV	79
7.5 Current Limitations	80
7.6 Future Improvements	80



Edit with WPS Office

8. Development of the Project Presentation Website	81
8.1 Objectives of the Website	81
8.2 Website Structure	81
8.3 Website Design	82
8.4 Contribution of the Website to the Project	82
9.Conclusion	83

Edit with WPS Office

Liste of figures

Chapitre 1 :	
Figure 1.1 : Ocean temperature profile showing the thermocline.	08
Figure 1.2 :	09
Figure 1.3 : The demand for underwater target optical detection.	11
Figure 1.4 : General architecture of a Remotely Operated Vehicle (ROV).	12
Figure 1.5 : Jason ROV.	14
Figure 1.6: ROV SuBastian.	14
Figure 1.7: BlueROV2.	15
Figure 1.8: VideoRay ROV.	16
Chapitre 2 :	
Figure 2.1: Major Domains of Artificial Intelligence	22
Figure 2.2: Supervised Learning.	25
Figure 2.3: Type of Machine Learning	26
Figure 2.4: Machine Learning Workflow	27
Figure 2.5: Deep Neural Network	28
Figure 2.6: CNN Architecture	29
Figure 2.7: Convolution Layer	30
Figure 2.8: Tasks of Computer Vision	32
Figure 2.9: Object Detection Techniques	34
Figure 2.10: YOLO Architecture	35
Figure 2.11: Underwater Image Before and After Enhancement	38
Chapter 3 :	
Figure 3.1: The overall architecture of the proposed system	42
Figure 3.2 hardware components	43
Figure 3.3: Internal arrangement of the hardware components inside the proposed ROV prototype.	49
Figure 3.4 : syringe based buoyancy control mechanism	52
Figure 3.5 : Software architecture of AquaSearch AI	53
Figure 3.6 : communication architecture of AquaSearch AI	55
Figure 3.7: *Operational workflow of AquaSearch AI.*	57
Chapter 4 :	
Figure 4.1 : Directory structure of the TrashCan instance segmentation Dataset.	64

Edit with WPS Office

Figure 4.2 : dataset directory organization in the file system.	65
Figure 4.3 : YAML Configuration file for dataset.	66
Figure 4.4 : Sample image from the dataset.	66
Figure 4.5 : comparison between the original image and the instance segmentation result.	68
Figure 4.6 : Example of an instance segmentation Annotation File.	69
Figure 4.7 : training batch with instance segmentation annotations (batch 0).	73
Figure 4.8 : YOLOv8 Nano instance segmentation results.	77
Figure 4.9 : YOLOv8 Small instance segmentation results after training on the underwater trash detection dataset.	77
Figure 4.10 : YOLOv8 Medium instance segmentation results	78

Edit with WPS Office

List of tables

Chapter 4 :	
Table 4.1 : Characteristics of the TrashCan instance segmentation Dataset.	62
Table 4.2 : Underwater Object Categories.	63
Table 4.3 : software environment and devlopement tools.	70
Table 4.4 :comparison of YOYOv8 segmentations models.	71
Table 4.5 :Experimental setup for model traning.	72
Table 4.7 :Structure of the project website.	82

Edit with WPS Office

INTRODUCTION GENERAL

1. Context and Motivation

The rapid advancement of Artificial Intelligence, embedded systems, and computer vision has significantly contributed to the development of intelligent solutions capable of operating in complex and challenging environments. Among these environments, the underwater domain represents one of the most difficult areas for technological intervention due to several physical and technical constraints, such as low visibility, light attenuation, water pressure, turbidity, and communication limitations.

Underwater search operations are often difficult to perform manually, especially when the target area is large, deep, or unsafe for human divers. In such situations, underwater robotic systems can provide an effective solution by allowing users to inspect aquatic environments remotely and reduce the risks associated with direct human intervention. Remotely Operated Vehicles, commonly known as ROVs, are particularly useful because they can be controlled from the surface while carrying cameras, sensors, and electronic components that support observation and search tasks.

In recent years, the integration of Artificial Intelligence and Computer Vision has improved the capabilities of robotic systems. Object detection algorithms, especially YOLO-based models, have shown promising results in identifying and localizing objects in images and videos. When applied to underwater environments, these technologies can assist the operator by detecting objects of interest and supporting decision-making during search operations.

This master´s thesis presents the design and development of a low-cost intelligent ROV prototype dedicated to underwater object search. The proposed system combines remote control, embedded electronics, video acquisition, Wi-Fi communication, motor control, and AI-based object detection. The main motivation behind this project is to provide an accessible and realistic proof of concept that demonstrates how Artificial Intelligence can support underwater search tasks using affordable components and an integrated embedded processing unit.

2. Problem Statement

Underwater environments present several challenges that make search operations difficult and sometimes inefficient. Human divers are limited by safety conditions, visibility, depth, time, and physical effort. In addition, professional underwater robotic systems are often expensive and may

Edit with WPS Office

not be accessible to universities, small research laboratories, or local institutions with limited resources.

The initial idea of this project was related to searching for drowned people. However, after studying the technical requirements of such a system, it became clear that the success of the Artificial Intelligence model depends strongly on the availability of a suitable and reliable dataset. Since sufficient underwater datasets related to drowned people were not available, the objective of the project was adjusted to focus on underwater object search in general. This makes the system more realistic and scientifically appropriate for the current stage of development.

Another important challenge concerns the design of a compact and low-cost architecture capable of combining several tasks within the same system. The ROV requires a processing unit that can handle camera acquisition, communication with the user, motor control, and the execution of an AI-based detection model. For this reason, the proposed prototype relies on the Raspberry Pi 5 as the central processing unit. It is used to manage the camera, control the motors, support Wi-Fi-based communication, and run the YOLO detection module according to the available hardware capabilities.

This master´s project specifically addresses the following issues:

• The difficulty of performing underwater search manually in environments with limited visibility and communication constraints.

• The high cost of professional underwater robotic systems, which limits their accessibility for academic and small-scale research use.

• The need for a low-cost ROV prototype that can assist users during underwater object search tasks.

• The challenge of integrating Artificial Intelligence and Computer Vision into an embedded system suitable for a small underwater prototype.

• The need to design a centralized architecture where the Raspberry Pi 5 manages video acquisition, motor control, communication, and AI processing.

• The need to select an object detection model suitable for the available hardware capabilities and underwater search conditions.

3. Objectives of the Master´s Project

Edit with WPS Office

The primary goal of this master´s thesis is to design and develop a low-cost intelligent ROV prototype for underwater object search, supported by Artificial Intelligence and Computer Vision techniques. The project aims to demonstrate the feasibility of building an accessible system that can assist users in searching for objects of interest under water.

The specific objectives are :

1. To design and build a low-cost underwater ROV prototype capable of moving under water through remote control.

2. To develop a hardware architecture based on accessible components, including Raspberry Pi 5, an underwater camera, motors, battery, and waterproof housing.

3. To use the Raspberry Pi 5 as the main processing and control unit responsible for camera handling, motor control, Wi-Fi communication, and AI-related processing.

4. To integrate a camera system capable of capturing underwater video and providing visual data for observation and object detection.

5. To study the use of Artificial Intelligence and Computer Vision for underwater object detection using a YOLO-based model.

6. To compare different YOLOv8 models, especially YOLOv8n, YOLOv8s, and YOLOv8m, in order to justify the selection of a model suitable for the available hardware.

7. To design a software architecture that links video acquisition, motor control, communication, and object detection within the same embedded system.

8. To propose future improvements, such as improving underwater communication, enhancing real-time video transmission, optimizing AI performance, and training the system on richer underwater datasets.

4. Scope and Limitations of the Master´s Project

The scope of this project includes the design, development, and partial integration of a low-cost intelligent ROV prototype dedicated to underwater object search. The project focuses on the mechanical structure, embedded electronic components, remote control mechanism, video acquisition capability, motor control, and the study of an AI-based object detection module using YOLO.

Edit with WPS Office

The current prototype is not designed to be autonomous. It is a remotely controlled system where the user is responsible for guiding the ROV and controlling its movement. The Artificial Intelligence component is intended to assist the user by detecting objects of interest according to the trained model and the available dataset.

The project also does not aim to compete with professional commercial underwater robots. Instead, it represents a proof of concept that demonstrates the possibility of developing an AI-assisted underwater search system using low-cost and accessible components.

Several limitations must be considered in the current version of the project. First, although the Raspberry Pi 5 offers better performance than previous low-cost boards, it remains limited compared to specialized AI hardware such as dedicated GPUs or embedded AI accelerators. Therefore, the choice of the YOLO model must take into account the available processing power and the need for efficient execution.

Second, underwater communication remains a technical challenge. The current prototype considers Wi-Fi-based communication within the limits of the experimental environment, but the reliability and range of this communication may vary depending on depth, water conditions, distance, and the physical structure of the ROV. Therefore, communication performance remains an important aspect to evaluate during practical testing.

Another limitation concerns the dataset and the final AI model. At this stage, the dataset and the trained model are still under development. The system´s ability to detect objects depends on the availability and quality of the underwater dataset, as well as the object classes included in the training process. Therefore, the project remains open to future adaptation depending on the target application and the available data.

5. Thesis Outline

This thesis is structured into five chapters in order to present the theoretical background, system design, implementation process, and business perspective of the proposed intelligent ROV system.

Chapter 1: Underwater Search Systems and Technologies

This chapter presents the underwater environment and its main challenges, including visibility, light attenuation, turbidity, pressure, and communication difficulties. It also introduces

Edit with WPS Office

underwater robots such as ROVs and AUVs, their applications, existing technologies, and the limitations of current underwater solutions.

Chapter 2: Artificial Intelligence and Computer Vision

This chapter provides the theoretical background related to Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Networks, Computer Vision, object detection techniques, and the YOLO algorithm. It also discusses the challenges of underwater image processing.

Chapter 3: System Design and Architecture

This chapter describes the proposed system architecture, including the hardware and software components of the ROV. It presents the role of the Raspberry Pi 5 as the central processing and control unit, as well as the underwater camera, motors, battery, waterproof housing, communication system, and the general workflow of the system.

Chapter 4: Implementation and Model Development

This chapter presents the implementation steps of the prototype and the development of the AI model. It discusses dataset collection, preprocessing, annotation, model selection, comparison between YOLOv8n, YOLOv8s, and YOLOv8m, object detection process, evaluation metrics, and experimental results according to the progress achieved.

General Conclusion

The general conclusion summarizes the work carried out, highlights the main contributions, discusses the limitations of the current prototype, and proposes future improvements for developing the system into a more advanced underwater search solution.

Edit with WPS Office

Chapter 1:
Search and Systems in Underwater Environment

Edit with WPS Office

1 Introduction:

The underwater environment is one of the most complex and least explored domains on Earth. Unlike terrestrial environments, it is characterized by high pressure, reduced visibility, light attenuation, and unpredictable physical conditions. These factors make direct human intervention difficult, risky, and sometimes impossible, especially during underwater search operations.

In recent years, technological advances have significantly improved the ability to operate in such harsh environments. Underwater robots, particularly Remotely Operated Vehicles (ROVs) have become essential tools for collecting data, observing submerged areas, and supporting search tasks in environments that are difficult or unsafe for human divers.

Moreover, The integration of Artificial Intelligence (AI) and Computer Vision (CV) has further enhanced the capabilities of ROV systems. Vision-based intelligent methods can support the detection and recognition of objects of interest, the analysis of underwater images, and the interpretation of visual information in real time. In this context, the combination of ROV, AI, and CV forms the foundation of the proposed system.

2 Characteristics of the Underwater Environment:

The underwater environment is shaped by interacting physical, chemical, and biological factors that affect marine ecosystems and the performance of submerged technologies. Understanding these characteristics is essential for designing and operating an ROV effectively. [1, 2]

2.1 Physical Characteristics:

The underwater environment is governed by several physical parameters that strongly influence both marine life and the performance of underwater systems:

- **Hydrostatic Pressure:**

  Hydrostatic pressure increases with depth due to the weight of the water column, rising by approximately 1 atmosphere every 10 meters. At greater depths, this pressure becomes a major engineering challenge for underwater vehicles, requiring pressure-resistant structures, strong materials, and reliable sealing systems. [2, 3]

Edit with WPS Office

- **Light Absorption :**

Light intensity decreases rapidly with depth because water absorbs and scatters different wavelengths. Red light disappears near the surface, while blue and green wavelengths penetrate deeper. As depth increases, optical visibility becomes limited, and underwater systems may require artificial lighting or acoustic sensors for navigation, mapping, and object detection. [4, 5]

- **Thermal Dynamics:**

Ocean temperature generally decreases with depth, forming layers separated by a thermocline where temperature changes rapidly. Deep ocean waters are typically very cold, often ranging between 0 °C and 4 °C. These temperature gradients affect water density and sound propagation, influencing requiring thermal protection for sensitive electronic components. [5]

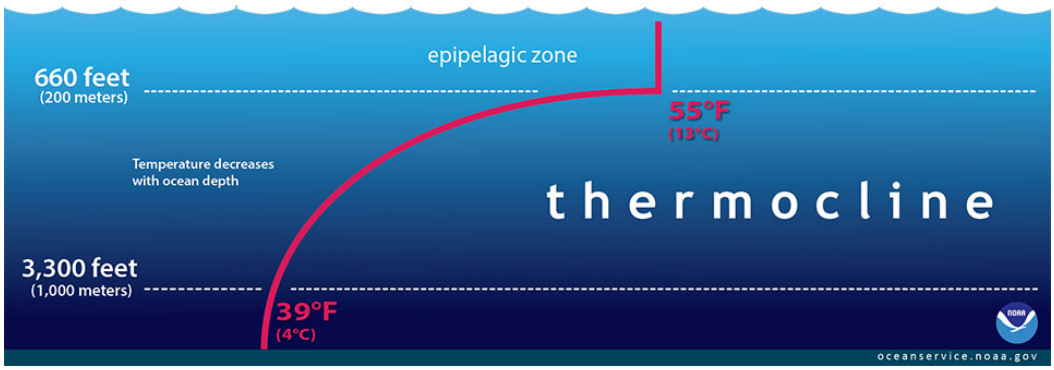


Figure 1.1: Ocean temperature profile showing the thermocline [5].

2.2 Chemical Characteristics:

The chemical composition of seawater determines the availability of essential elements for marine life and governs many geochemical processes:

- **Salinity and Ionic Composition:**

Average seawater salinity is about 35‰ and is mainly caused by dissolved sodium and chloride ions. Salinity influences water density, circulation patterns, and corrosion behavior. High salinity also accelerates material degradation, making

Edit with WPS Office

protective coatings and corrosion-resistant materials necessary for underwater systems. [6, 7]

- **Acidity (pH):**

  Average ocean pH is approximately 8.1; however, the absorption of atmospheric CO_2 contributes to gradual ocean acidification. Lower pH levels can accelerate material degradation, which requires chemically durable materials and protective design choices for underwater systems. [8]

2.3 Biological and Ecological Systems:

Biological properties describe how marine organisms interact with and adapt to the physical and chemical conditions of the underwater environment:

- **Chemosynthesis:**

  In deep regions where sunlight is absent, some organisms rely on chemosynthesis, using chemical energy from hydrothermal vents to sustain unique ecosystems. Such environments may expose underwater vehicles to high temperatures and corrosive substances that can damage sensors and structures. [9]

- **Ecological Zonation:**

  Marine ecosystems are divided into ecological zones based on depth, light availability, and biological activity. Operating conditions vary across these zones. Shallow waters may cause biofouling on vehicle surfaces, whereas deep-sea operations require caution to avoid disturbing seabed sediments and fragile habitats. [2]

Edit with WPS Office

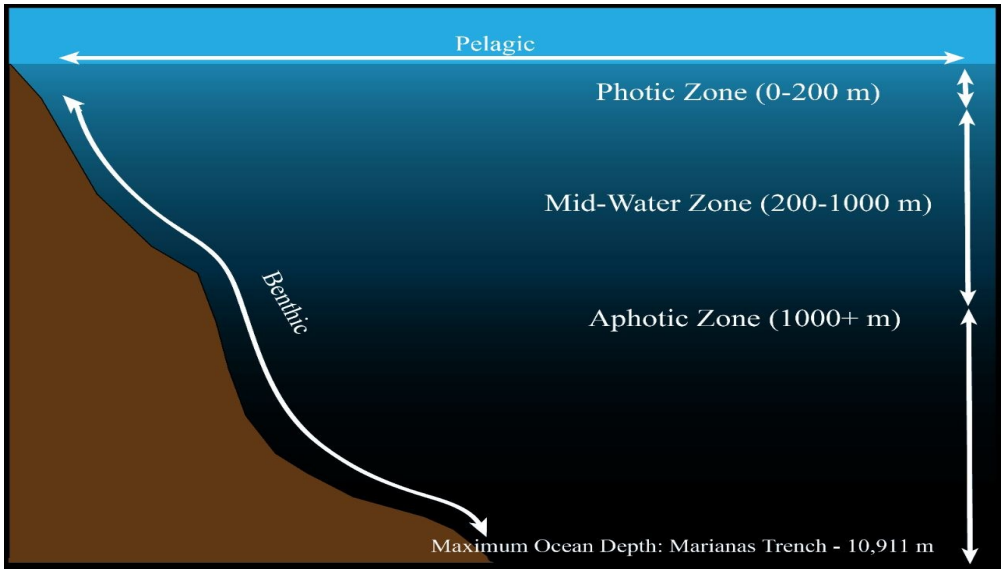


Figure 1.2: Ecological zonation of the marine environment [2].

3 Visibility Challenges:

- **Light Absorption:** Water strongly attenuates light as it travels through it. This attenuation depends on wavelength. Long wavelengths, such as red and orange, are absorbed near the surface, while shorter wavelengths, such as blue and green, penetrate farther. As depth increases, spectral information is progressively lost, causing underwater scenes to appear blue-green. This color shift affects both human perception and computer vision systems, making color correction an important topic in underwater image processing. [10, 11]

- **Light Scattering:** Light scattering occurs when suspended particles, such as sediments, plankton, and organic matter, cause light to deviate from its original path before reaching the eye or imaging sensor. Forward scattering reduces image sharpness, while backward scattering creates a veiling effect that obscures distant objects and reduces visibility. This phenomenon is one of the main causes of the hazy appearance commonly observed in underwater images. [10, 12]

Edit with WPS Office

- **Low Light Levels:** Light intensity decreases rapidly with depth due to the combined effects of absorption and scattering. Even in relatively clear water, most sunlight is attenuated within the upper layers, leaving deeper regions poorly illuminated. Artificial lighting is therefore often required, although it may introduce non-uniform illumination and increased backscatter. [10]

- **Refraction:** Refraction occurs when light passes between media with different refractive indices, such as water and air. This bending of light changes the perceived position and size of objects, making them appear closer and larger than they actually are. Such distortion affects depth perception and spatial accuracy, which are important for underwater navigation and measurement. [11]

- **Turbidity:** Turbidity refers to the cloudiness of water caused by suspended particles such as sediment or organic matter. Turbid water scatters and absorbs more light, reducing visibility and sometimes limiting it to only a few meters. Turbidity can also change over time because of currents, weather conditions, or human activity. [10, 12]

- **Low Contrast:** Underwater environments often have low contrast because light is absorbed and scattered as it travels through water. As a result, the difference between objects and their background becomes weaker, making edges and fine details harder to identify. This condition is challenging for both human operators and computer vision systems used in underwater exploration [12].

- **Visual Adaptation:** The human visual system is not naturally adapted to underwater environments. Without a diving mask, the eye cannot focus properly because the refractive indices of water and the cornea are similar. A mask creates an air layer that restores focusing ability, but vision remains affected by reduced brightness, altered color perception, and a limited field of view. [10]

Edit with WPS Office

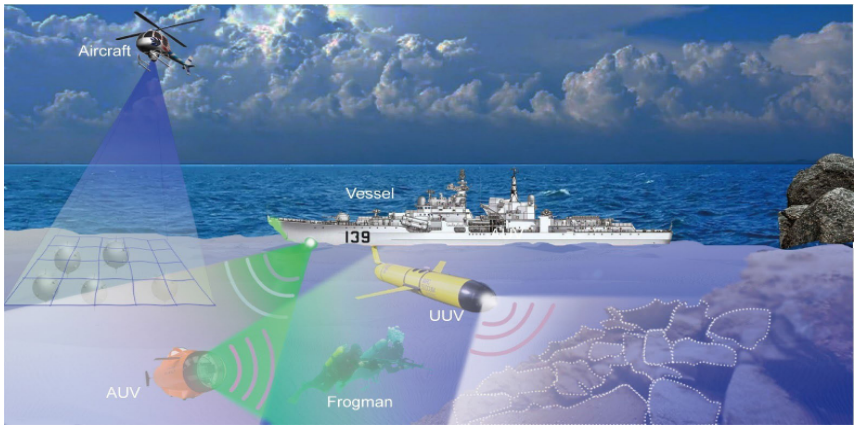


Figure 1.3: The demand for underwater target optical detection.

4 ROVs :

4.1 Definition of ROVs :

An ROV can be defined as an unmanned underwater vehicle operated remotely by a human user. It is generally equipped with cameras, lights, motors, control electronics, and sometimes additional sensors or tools depending on the mission. The operator controls the vehicle from the surface while observing the underwater environment through the video stream sent by the onboard camera. [11, 12]

4.2 Historical Development of ROVs :

ROVs were developed to overcome the limitations and risks of human diving. As underwater missions became deeper, longer, and more complex, the need for remotely operated systems increased. Early ROVs were mainly used for military, scientific, and industrial purposes, especially for underwater inspection and observation tasks. Over time, advances in electronics, cameras, materials, communication systems, and propulsion technologies made ROVs more reliable, compact, and accessible.

Today, ROVs are used in different fields, including marine research, underwater inspection, environmental observation, infrastructure maintenance, education, and object search.

Edit with WPS Office

Recent developments in embedded systems and artificial intelligence have also made it possible to integrate computer vision techniques into ROV platforms, allowing them to support operators with real-time visual analysis. [13]

4.3 General Architecture of ROV Systems:

A typical ROV system is composed of several interconnected parts that allow the vehicle to operate underwater while being controlled from the surface.

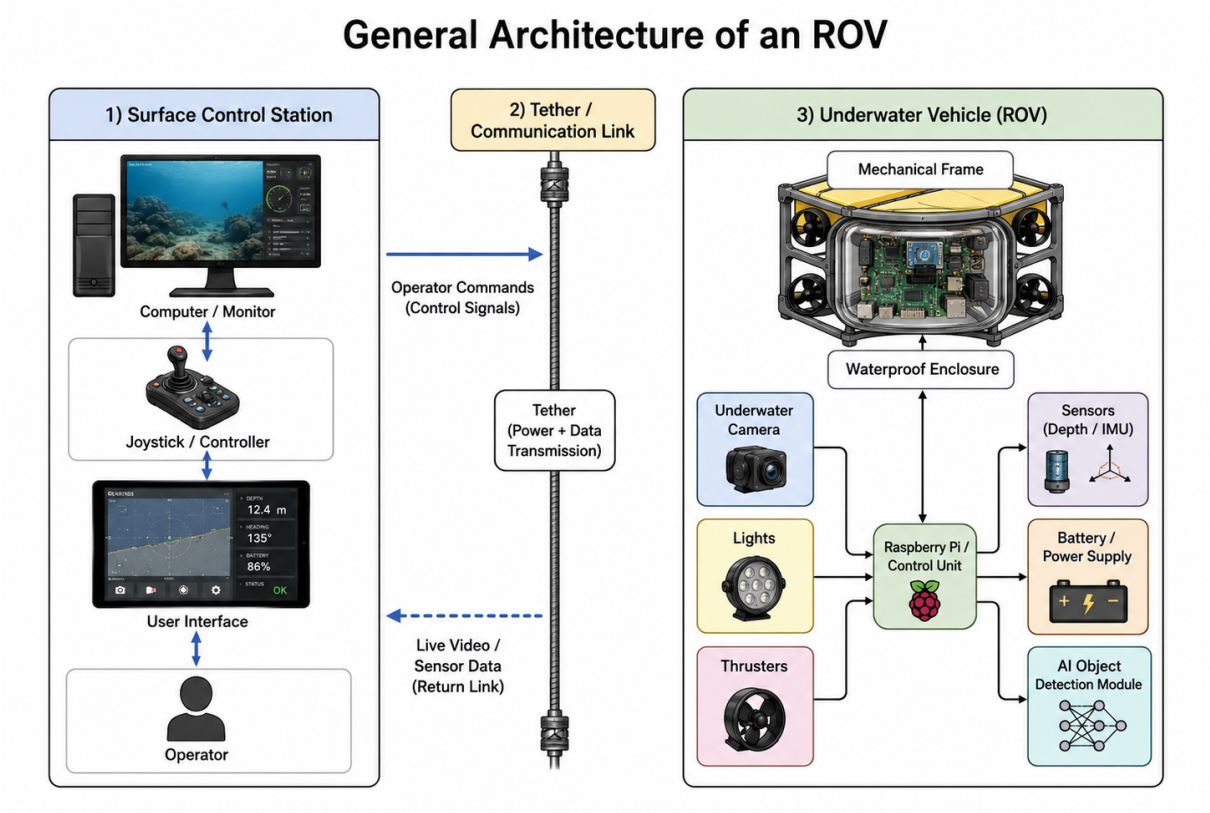


Figure 1.4: General architecture of a Remotely Operated Vehicle (ROV).

This diagram presents the general architecture of an ROV. It highlights the connection between

Edit with WPS Office

the surface control station, the tether, and the underwater vehicle. Through the tether, the operator sends control commands to the ROV, while live video and sensor data are transmitted back to the surface. This architecture is suitable for real-time underwater search operations. [11, 12]

4.4 Types of ROVs :

ROVs can be classified according to their size, capabilities, and mission type.

- **Observation-class ROVs:** are small and relatively simple vehicles mainly used for visual observation, education, and light inspection tasks. They are usually portable, easier to deploy, and less expensive than larger systems. This class is close to the objective of the proposed prototype because it focuses on visual feedback and underwater search support.[13]
- **Inspection-class ROVs:** are more advanced systems designed for underwater inspection of structures such as ship hulls, pipelines, dams, bridges, aquaculture cages, and ports. They often include better cameras, stronger lighting, navigation sensors, and sometimes sonar or measurement tools.[13]
- **Work-class ROVs:** are large and powerful vehicles used in demanding industrial operations, especially in offshore oil and gas, deep-sea intervention, and heavy underwater maintenance. They may include robotic arms, advanced sensors, powerful thrusters, and complex control systems. Although highly capable, these systems are expensive and require specialized operators and support equipment.[13]

5 Existing ROV Systems :

In the modern era, ROV technology has advanced significantly, leading to the development of sophisticated systems such as:

5.1 Jason ROV:

Edit with WPS Office

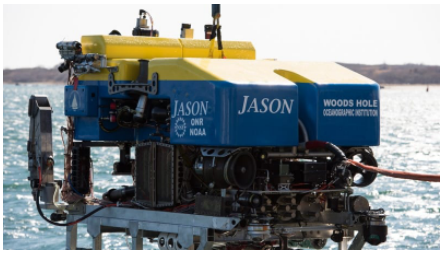


Figure 1.5 : Jason ROV.

- **Field of Application:** The Jason ROV, developed by the Woods Hole Oceanographic Institution (WHOI), supports scientific research and deep-sea exploration by enabling access to seafloor environments. It is used to study marine ecosystems, hydrothermal vents, and geological formations, and to collect samples without requiring scientists to enter deep-water environments. It has conducted numerous dives in the Pacific, Atlantic, and Indian Oceans. [14]

- **Advantages:** Jason operates at depths of several thousand meters through a robust fiber-optic tether that supplies power and transmits high-definition video and sensor data in real time. Its manipulator arms allow precise tasks such as collecting rock, sediment, and biological samples. The system is supported by advanced sonar, lighting, and imaging technologies. [14]

- **Challenges:** Jason operations require a dedicated support vessel equipped with a control room, winch, and launch-and-recovery system. Operational costs are high because of the sophisticated hardware, maintenance requirements, and specialized technical support. The tether also restricts mobility compared with untethered vehicles and must be carefully managed during deployment. [14]

5.2 ROV SuBastian:

Edit with WPS Office



Figure 1.6: ROV SuBastian.

- **Field of Application:** ROV SuBastian, operated by the Schmidt Ocean Institute, is designed for deep-sea scientific exploration, including seafloor mapping, sample collection, seawater analysis, and imaging of hard-to-reach marine life. It supports interactive experiments, equipment deployment, and telepresence operations at depths beyond 4,500 meters. [15]
- **Advantages:** SuBastian reaches a maximum operating depth of 4,500 meters and can support reconnaissance, visual observation, and physical sampling. High-precision control is enabled by navigation systems such as Inertial Navigation System

  (INS) and Doppler Velocity Log(DVL), while ultra-high-definition cameras provide detailed visual data. The vehicle can also carry scientific payloads, including sampling tools and sonar systems. [15]
- **Challenges:** Its large size and significant weight limit portability and require specialized handling equipment. High operational costs result from advanced hydraulics, power systems, scientific interfaces, and maintenance requirements. Deployment also requires a skilled technical team capable of managing the vehicle and its control software. [15]

5.3 BlueROV2:

Edit with WPS Office



Figure 1.7: BlueROV2.

- **Field of Application:** The BlueROV2 is commonly used in university research, educational projects, prototyping, and light-duty underwater inspection. It provides an accessible platform for mapping and exploration in controlled environments such as pools, tanks, and shallow waters. Its modular design makes it suitable for academic experimentation and validation of underwater robotics techniques. [16]
- **Advantages:** The BlueROV2 is relatively affordable compared with professional ROV systems and offers a six-thruster vectored configuration suitable for inspection and aquaculture applications. Its open-source software and hardware allow customization, easy transport, and deployment with a tether of up to 300 meters. These features make it useful for teaching, testing, and rapid prototyping. [16]
- **Challenges:** Its operating depth is limited to approximately 100 meters, which restricts its use in deep-sea applications. Payload capacity and accuracy may decrease in harsh environments because of its compact frame and basic sensor configuration. Its endurance and thrust capability are also limited compared with larger professional ROVs. [16]

Edit with WPS Office

5.4 VideoRay ROV :

Figure 1.8: VideoRay ROV.

- **Field of Application:** VideoRay ROVs are mainly used in professional underwater missions such as infrastructure inspection, underwater search, defense and security operations, port and harbor inspection, offshore inspection, and intervention tasks. They can also be used in environments that are difficult or dangerous for human divers, especially when real-time visual feedback and precise remote control are required.[17]
- **Advantages:** The main advantage of VideoRay ROVs is their compact and portable design, which makes them easier to deploy compared with larger work-class ROVs. They also offer strong maneuverability, high-quality visual feedback, and modular architecture. Depending on the model and mission configuration, they can support additional tools and sensors such as sonar, manipulators, navigation sensors, and other mission-specific payloads. This flexibility allows VideoRay ROVs to be adapted to different underwater tasks, from simple observation to more complex inspection and intervention missions.[17]
- **Challenges:** Despite their capabilities, VideoRay ROVs remain professional systems, which



makes them relatively expensive compared with low-cost educational or student-built prototypes. Their operation may require trained users, specific maintenance, and additional accessories depending on the mission. Like most ROVs, they may also face underwater limitations such as tether management, reduced visibility, water currents, and image degradation caused by turbidity, low light, and scattering. In addition, intelligent object detection is not always integrated by default, which highlights the interest of adding AI-based visual assistance in the proposed project.[17]

Edit with WPS Office

6 Conclusion :

In conclusion, this chapter presented the main concepts related to underwater search and ROV systems. It introduced the underwater environment as a complex and challenging domain characterized by pressure, reduced visibility, light attenuation, salinity, temperature variation, and unpredictable physical conditions. These constraints make direct human intervention difficult and highlight the importance of using Remotely Operated Vehicles to support underwater search tasks. The chapter also discussed the main visibility challenges that affect underwater imaging, such as light absorption, scattering, turbidity, low contrast, and refraction, which can impact both human operators and computer vision systems. Furthermore, it presented the definition, historical development, general architecture, and main types of ROVs, as well as examples of existing ROV systems and their limitations. This study shows the relevance of developing an AI-based ROV system capable of assisting the operator during underwater search operations through visual analysis and object detection. The next chapter will focus on the proposed system and its design, based on the concepts and challenges discussed in this chapter.

Edit with WPS Office

Edit with WPS Office

Chapter 2 :
Artificiel Intelligence and Computer Vision

Edit with WPS Office

1. Introduction

In recent years, AI has experienced a remarkable development, becoming one of the most influential fields in modern technology. Its ability to simulate human intelligence and process large amounts of data has enabled the emergence of advanced systems capable of solving complex problems in various domains. Among these domains, computer vision has gained significant attention due to its capacity to analyze and interpret visual information.

The combination of machine learning and deep learning techniques has considerably improved the performance of image analysis systems. These approaches allow computers not only to recognize patterns in data but also to make accurate predictions based on learned representations. As a result, tasks such as object detection and image classification have reached a high level of efficiency, making them suitable for real-world applications.

Despite these advancements, applying such techniques in underwater environments remains particularly challenging. The quality of visual data captured underwater is often degraded due to several factors, including poor lighting conditions, color distortion, and the presence of noise. These limitations make the detection and recognition of objects more difficult compared to standard environments.

In this context, the integration of computer vision techniques into underwater systems represents a promising solution. The project presented in this work is based on the use of an intelligent underwater vehicle equipped with a camera, where the captured video stream is analyzed in real time using deep learning algorithms. This approach aims to assist the operator in detecting objects more efficiently during underwater exploration and search operations.

This chapter provides the theoretical background necessary for understanding the proposed system. It starts with an overview of AI and Machine Learning (ML) concepts, followed by a detailed presentation of Deep Learning (DL) and Convolutional Neural Networks (CNN). Then, it introduces the fundamentals of Computer Vision and Object Detection techniques, with a particular focus on the YOLO algorithm. Finally, the chapter discusses the main challenges related to underwater image processing.

Edit with WPS Office

2. Introduction to Artificial Intelligence

2.1. Definition of artificial intelligence

AI is a major field of computer science that focuses on the development of systems capable of performing tasks that typically require human intelligence. These tasks include learning, reasoning, perception, and decision-making. The primary objective of AI is to design intelligent systems that can analyze their environment and act in a rational and autonomous manner.

AI is inherently multidisciplinary, as it integrates concepts from various domains such as computer science, data analytics, statistics, hardware and software engineering, linguistics, neuroscience, as well as philosophy and psychology. This interdisciplinary nature allows AI to address complex real-world problems from multiple perspectives.

In practice, AI enables machines to replicate certain cognitive functions of the human brain, including understanding their environment, acquiring new knowledge, and generating insights. For instance, AI is widely used in Optical Character Recognition (OCR), where it is applied to extract text and data from images and documents. This process transforms unstructured content into structured and usable information, which facilitates data analysis and supports decision-making processes.[18]

2.2. Domains of Artificial Intelligence

A wide range of abilities that allow machines to simulate human intelligence are collectively referred to as AI. It performs exceptionally well in decision-making, where it evaluates data to produce insights and suggestions, and automation, where it can handle monotonous tasks effectively and without getting tired. AI systems can also learn from data, using algorithms like machine learning to gradually improve their performance.[19]

The transformative power of AI is demonstrated by its everyday applications. With voice commands, virtual assistants like Siri, Alexa, and Google Assistant simplify daily tasks. Autonomous vehicles use AI in transportation to improve safety and navigation. AI improves healthcare through more accurate diagnosis and individualized treatment regimens. AI is used for content curation on social media and for product recommendations on retail sites like Amazon. These instances highlight how AI is pervasively changing industries and lifestyles by improving technology's responsiveness, efficiency, and user-centricity.[19]

Edit with WPS Office

Based on its capabilities, we can categorize artificial intelligence into three groups.

- Weak artificial intelligence is another name for narrow artificial intelligence. This is the stage of artificial intelligence where machines are only capable of a limited number of specialized tasks.[20]
- Strong artificial intelligence is another name for general artificial intelligence. Machines will be able to think and act similarly to humans at this point in the development of artificial intelligence.[20]
- Super Artificial Intelligence: This is the stage of AI where computers will be able to perform tasks that humans cannot.[20]

These three categories of AI allow us to categorize the fields of AI as follows:

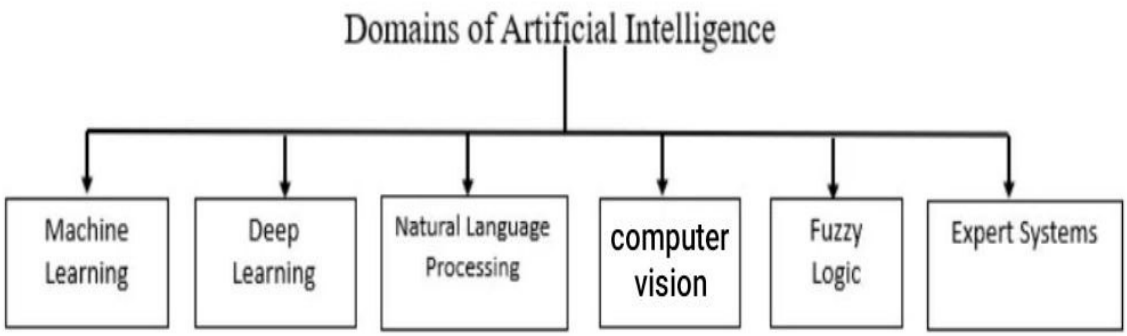


Figure 2.1: Major Domains of Artificial Intelligence

2.3 Role of Artificial Intelligence in Problem Solving

Problem solving is a central component of human thinking, enabling both individuals and organizations to address challenges, make informed decisions, and accomplish their objectives. In recent years, the emergence of AI has significantly transformed this field. AI systems now play an increasingly important role in supporting humans by facilitating the identification, analysis, and resolution of problems in a wide range of domains, including business, healthcare, and daily life activities. This section aims to examine the contribution of AI to problem solving by highlighting its main approaches, practical applications, as well as the challenges it faces and

Edit with WPS Office

the perspectives for its future development.

Through the provision of tools and technologies that augment human capabilities, artificial intelligence is revolutionizing the field of problem-solving. The following are some significant ways that AI aids in problem solving:

Data analysis and insights: AI excels at processing vast amounts of data quickly and efficiently. Machine learning algorithms can identify patterns and correlations within complex datasets that may not be apparent to human analysts. For example, in financial markets, AI can analyze historical trading data to identify trends, enabling better investment decisions. Using past data, AI systems are able to forecast future events. Predictive models in the medical field can identify patient risks, allowing for early interventions and better patient care.[21]

Decision support systems: AI-driven Decision Support Systems (DSS) are intended to help people and organizations make wise decisions. These systems simulate situations, evaluate data, and make recommendations based on predetermined standards. AI in agriculture can evaluate crop yields, soil types, and weather patterns to suggest the best planting plans, assisting farmers in increasing output and reducing waste.[21]

Automation of repetitive tasks: AI can automate repetitive problem-solving tasks, allowing humans to focus on more difficult problems. In order to free up human agents to work on more complex problems that call for empathy and critical thinking, customer service chatbots, for example, can handle routine questions.[21]

The convergence of problem solving and artificial intelligence is profoundly transforming the way individuals and organizations deal with complex situations. Through its advanced capabilities in data analysis, decision-making support, and process automation, AI contributes to strengthening human performance and fostering the development of innovative solutions. Nevertheless, the effective use of these technologies requires careful consideration of several critical issues, particularly those related to data reliability, system transparency, and ethical implications. Ensuring responsible and well-regulated use of AI is therefore essential. By promoting a collaborative approach between human intelligence and artificial intelligence, it becomes possible to better manage the growing complexity of modern challenges and to create new opportunities for progress and innovation in an increasingly connected world.

Edit with WPS Office

3. Machine Learning

3.1. Definition of Machine Learning

ML is a subfield of Artificial Intelligence that focuses on developing algorithms capable of learning from data and improving their performance without being explicitly programmed. Instead of relying on predefined rules, machine learning models analyze data, identify patterns, and make decisions based on learned knowledge.

ML systems are widely used in various applications such as image recognition, speech processing, recommendation systems, and autonomous systems. These models improve over time as they are exposed to more data, making them highly effective in solving complex problems.

Machine learning is defined as a method of data analysis that automates analytical model building and enables systems to learn from data, identify patterns, and make decisions with minimal human intervention.[22]

3.2. Types of Machine Learning

Machine learning can be categorized into three main types depending on how the data is used during the learning process.

- **Supervised Learning**

  Supervised learning is based on labeled data, where each input is associated with a correct output. The model learns to map inputs to outputs by analyzing these examples.

  This approach is widely used in tasks such as classification and object detection, where the model is trained to recognize specific patterns in the data.

  Supervised learning uses labeled datasets to train algorithms that classify data or predict outcomes accurately.[22]

Edit with WPS Office

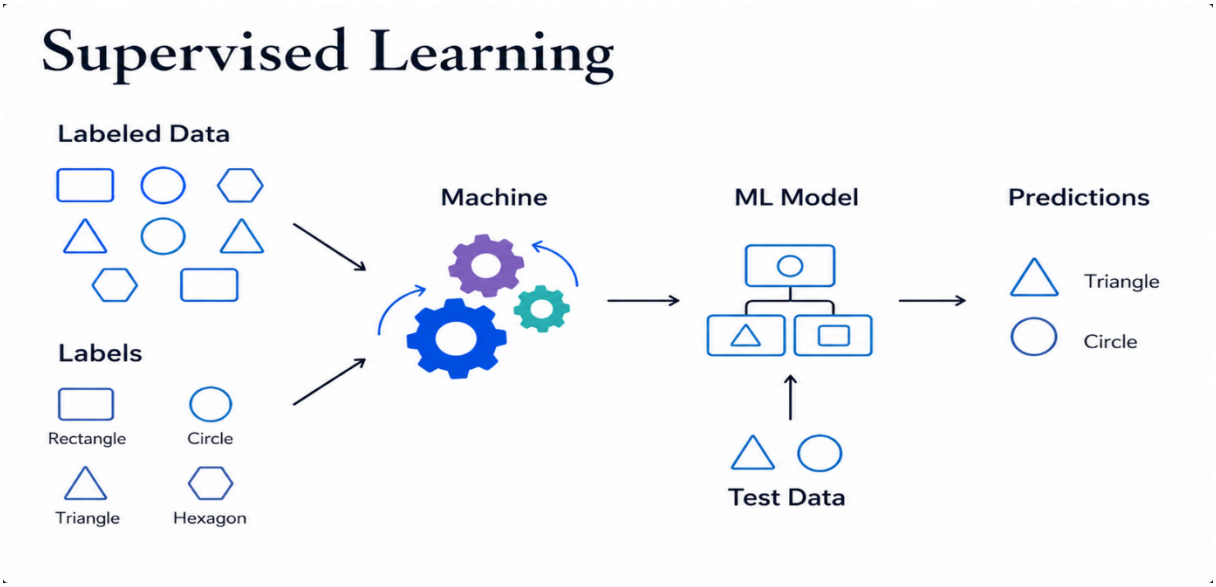


Figure 2.2: Supervised Learning.

- **Unsupervised Learning**

Unsupervised learning deals with unlabeled data, where the model tries to identify patterns or structures without prior knowledge of the correct output.

It is commonly used in clustering and data exploration tasks, where the objective is to discover hidden relationships within the data.

Unsupervised learning algorithms are used to draw inferences from datasets consisting of input data without labeled responses.[22]

- **Reinforcement Learning**

Reinforcement learning is based on an interaction between an agent and its environment. The agent learns by taking actions and receiving rewards or penalties, gradually improving its performance over time.

This type of learning is commonly used in robotics and decision-making systems.

Reinforcement learning is a type of machine learning in which an agent learns to make decisions by performing actions and receiving feedback from the environment.[22]

Edit with WPS Office

Example	Data Type	Type of Learning
Object Detection	Labeled Data	Supervised Learning
Clustering	Unlabeled Data	Unsupervised Learning
Robotics	Interaction	Reinforcement Learning

Figure 2.3: Type of Machine Learning

3.3. Machine Learning Workflow

Machine learning follows a series of steps that allow the system to learn from data and produce reliable results. The process usually begins with data collection, followed by data preprocessing to clean and prepare the data.

After that, the model is trained using the prepared dataset, and its performance is evaluated using specific metrics. Finally, the trained model is used to make predictions on new data.

A typical machine learning workflow includes data collection, data preparation, model training, evaluation, and deployment, forming a complete pipeline for building intelligent systems.[23]

Edit with WPS Office

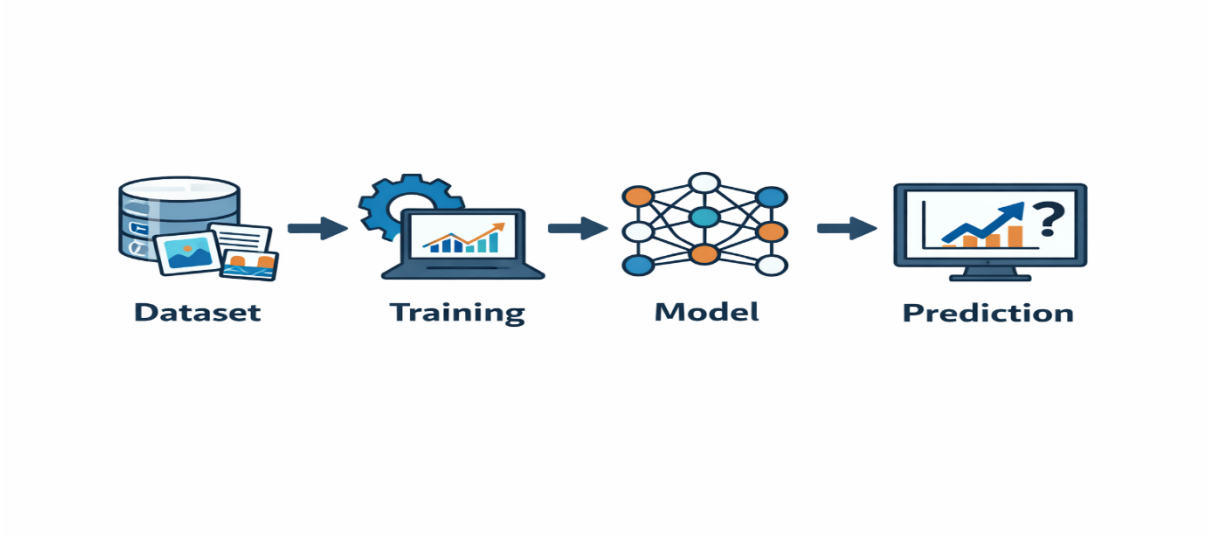


Figure 2.4: Machine Learning Workflow

4. Deep Learning

4.1. Definition of Deep Learning

Deep Learning is a subset of Machine Learning that focuses on using artificial neural networks with multiple layers to model complex patterns in data. Unlike traditional machine learning techniques, deep learning can automatically learn feature representations directly from raw data, especially in tasks involving images, audio, and video.

Deep learning has achieved remarkable success in various domains such as image recognition, natural language processing, and speech recognition, due to its ability to handle large-scale and high-dimensional data.

Deep learning is a class of machine learning techniques that uses neural networks with many layers to progressively extract higher-level features from raw input data.[23]

4.2. Neural Networks

ANNs are the fundamental building blocks of deep learning. They are inspired by the structure and functioning of the human brain, where interconnected neurons process and transmit information.

A neural network is composed of multiple layers, including an input layer, one or more hidden layers, and an output layer. Each neuron receives input data, applies a mathematical

Edit with WPS Office

transformation, and passes the result to the next layer.

These networks are capable of learning complex relationships between inputs and outputs, making them particularly effective in tasks such as image classification and object detection.

A neural network is a set of algorithms designed to recognize patterns, loosely inspired by the way the human brain operates.[22]

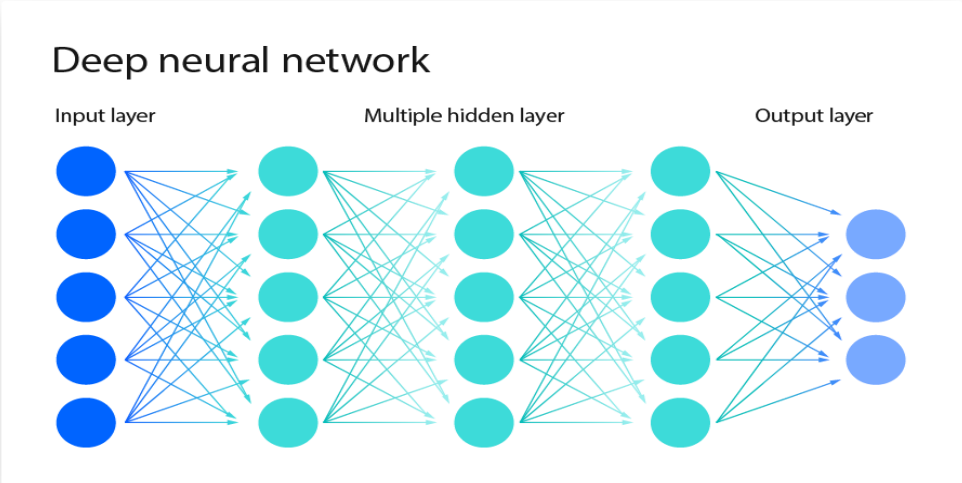


Figure 2.5: Deep Neural Network

4.3. Advantages of Deep Learning over Machine Learning

Deep Learning offers several advantages compared to traditional machine learning methods. One of its main strengths is its ability to automatically extract features from raw data, eliminating the need for manual feature engineering.

In addition, deep learning models generally achieve higher accuracy when dealing with complex data such as images and videos. They are also more scalable and can improve their performance as the size of the dataset increases.

Deep learning techniques have significantly improved the state-of-the-art in many fields, particularly in computer vision tasks where traditional machine learning approaches often fall short.[24]

Edit with WPS Office

5. Convolutional Neural Networks (CNN)

5.1. CNN Architecture

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed for processing structured grid data such as images. They are widely used in computer vision tasks due to their ability to automatically extract spatial features from images.

A typical CNN architecture consists of several layers, including convolutional layers, pooling layers, and fully connected layers, each playing a specific role in the learning process.

CNNs are deep neural networks that are particularly effective for analyzing visual imagery and are composed of multiple layers designed to automatically and adaptively learn spatial hierarchies of features.[22]

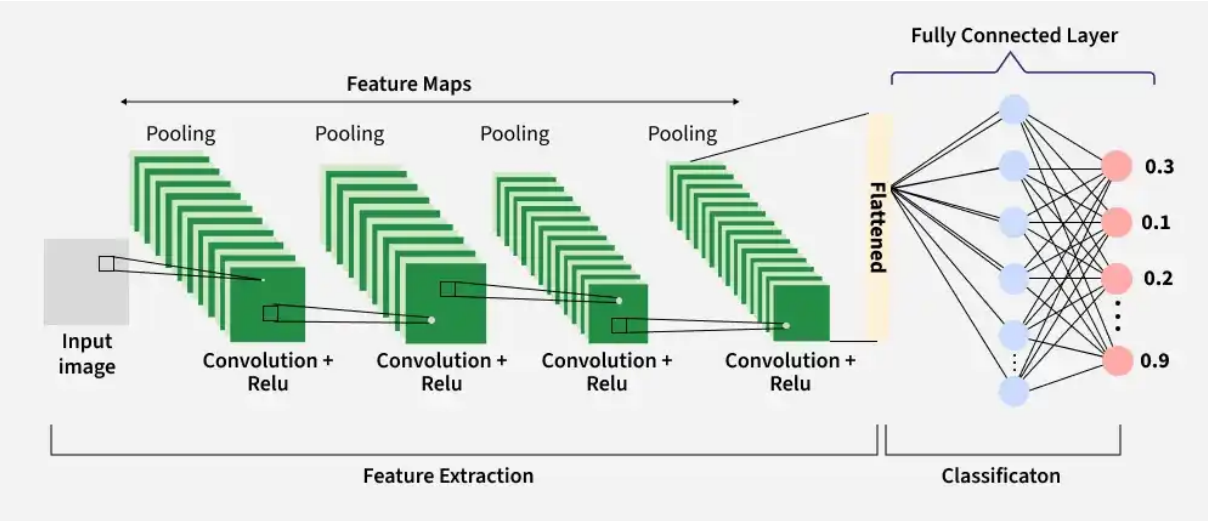


Figure 2.6: CNN Architecture

5.1.1 Convolution Layer

The convolution layer is the core component of a CNN. It applies filters (also called kernels) to the input image in order to extract important features such as edges, textures, and shapes.

Each filter scans the image and produces a feature map that highlights specific patterns. This operation allows the network to detect local features in the image.

Convolution is a mathematical operation on two functions that produces a third function expressing how the shape of one is modified by the other.[25]

Edit with WPS Office

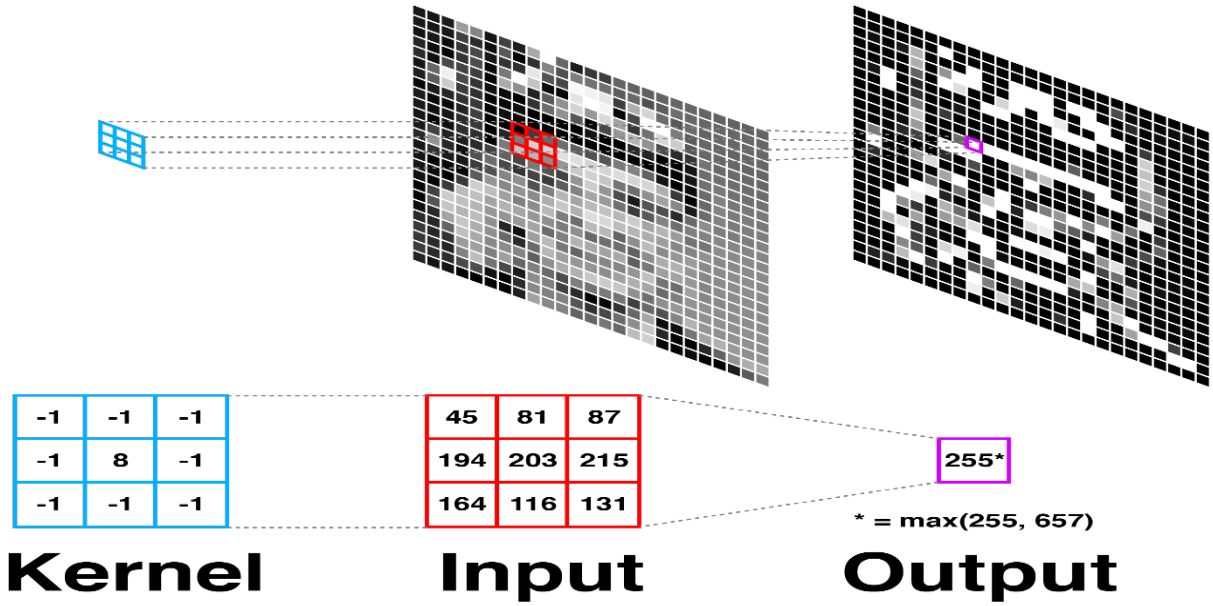


Figure 2.7: Convolution Layer

5.1.2 Pooling Layer

The pooling layer is used to reduce the spatial dimensions of the feature maps while preserving the most important information. This helps decrease computational complexity and prevents overfitting.

Common types of pooling include max pooling and average pooling, where the most significant values are selected from a region of the feature map.

Pooling layers progressively reduce the spatial size of the representation to reduce the number of parameters and computation in the network.[22]

5.1.2 Fully Connected Layer

The fully connected layer is usually placed at the end of the network. It takes the features extracted by previous layers and uses them to perform classification or prediction.

In this layer, all neurons are connected to each other, allowing the network to combine the extracted features and make final decisions.

Fully connected layers connect every neuron in one layer to every neuron in another layer, enabling high-level reasoning based on extracted features.[22]

5.2. How CNN Works on Images

CNNs process images through a sequence of transformations. First, the input image is

Edit with WPS Office

passed through convolution layers to extract low-level features such as edges and textures. As the data moves deeper into the network, more complex features such as shapes and objects are detected.

Then, pooling layers reduce the size of the feature maps while preserving important information. Finally, the fully connected layers use these features to classify the image or detect objects.

Convolutional neural networks automatically learn hierarchical feature representations from images, starting from simple patterns and progressing to more complex structures.[26]

6. Computer Vision

6.1. Definition of Computer Vision

Computer Vision is a field of Artificial Intelligence that focuses on enabling machines to interpret and understand visual information from the real world, such as images and videos. It aims to simulate the human visual system by allowing computers to extract meaningful information from visual data.

This field plays a key role in many applications, including image recognition, object detection, medical imaging, and autonomous systems.

Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world by using digital images and deep learning models.[22]

6.2. Tasks of Computer Vision

Computer Vision includes several fundamental tasks that allow machines to analyze and interpret visual data effectively.

Image Classification

Image classification consists of assigning a label to an entire image based on its content. The model predicts the category to which the image belongs.

This task is widely used in applications such as object recognition and medical diagnosis.

Image classification is the task of assigning a label to an image from a predefined set of categories.[27]

Object Detection

Edit with WPS Office

Object detection goes a step further by not only identifying objects in an image but also determining their locations using bounding boxes.

This task is essential in real-world applications where multiple objects need to be identified simultaneously.

Object detection is a computer vision technique for locating instances of objects in images or videos.[22]

Image Segmentation

Image segmentation involves dividing an image into different regions or segments, where each region represents a specific object or part of the image.

This allows for a more detailed understanding of the image compared to classification or detection.

Image segmentation is the process of partitioning an image into multiple segments to simplify its representation.[28]

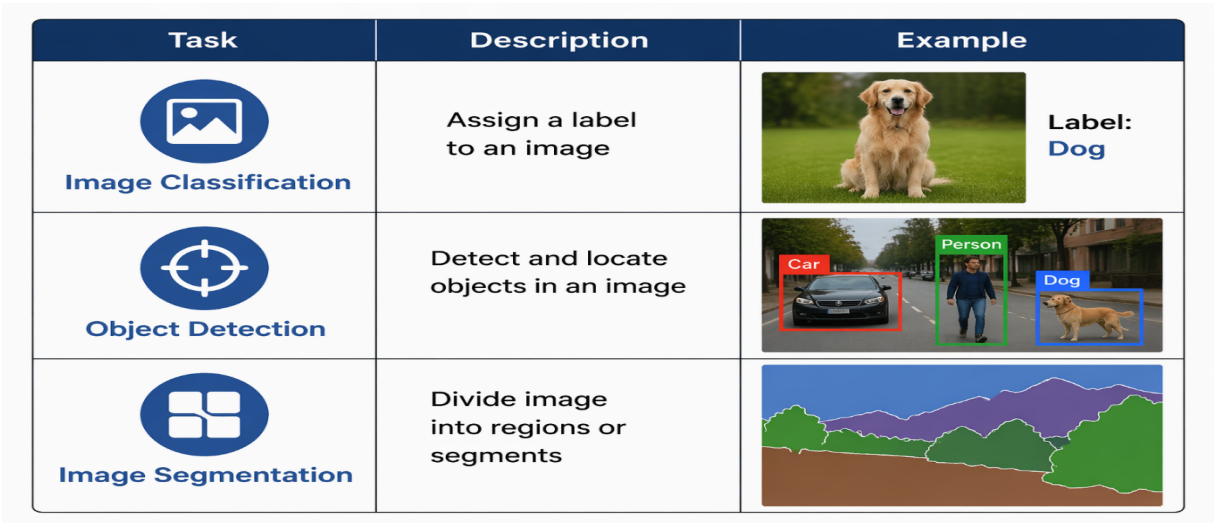


Figure 2.8: Tasks of Computer Vision

7. Object Detection Techniques

7.1. Traditional Object Detection Methods

Before the emergence of deep learning, object detection relied on traditional computer vision techniques based on handcrafted features. These methods used algorithms to extract

Edit with WPS Office

specific characteristics from images, such as edges, textures, and shapes.

Common approaches included methods like Haar Cascades and Histogram of Oriented Gradients (HOG), which were combined with classifiers such as Support Vector Machines (SVM). Although these techniques were effective in simple scenarios, they struggled with complex images and varying environmental conditions.

Traditional object detection methods rely on manually designed features and classical machine learning algorithms to identify objects in images.[29]

7.2. Deep Learning-Based Detection Methods

With the advancement of deep learning, object detection techniques have significantly improved. Modern approaches use Convolutional Neural Networks (CNNs) to automatically extract features from images, leading to higher accuracy and better performance.

These methods are generally divided into two main categories: two-stage detectors and one-stage detectors.

Deep learning-based object detection methods use neural networks to learn feature representations directly from data, resulting in improved detection accuracy.[22]

- **Two-Stage Detectors**

  Two-stage detectors first generate region proposals where objects are likely to be located, and then classify these regions.

  Examples include:

  1. R-CNN
  2. Fast R-CNN
  3. Faster R-CNN

  These models are known for their high accuracy but are relatively slow due to their complex processing.

  Two-stage detectors first propose candidate object regions and then classify them, leading to high accuracy but increased computational cost.[30]

- **One-Stage Detectors**

Edit with WPS Office

One-stage detectors perform detection in a single step by directly predicting object classes and bounding boxes from the input image.

Examples include:

1. YOLO (You Only Look Once)
2. SSD (Single Shot Detector)

These models are much faster and suitable for real-time applications.

One-stage detectors perform object detection in a single pass, making them significantly faster than two-stage methods.[31]

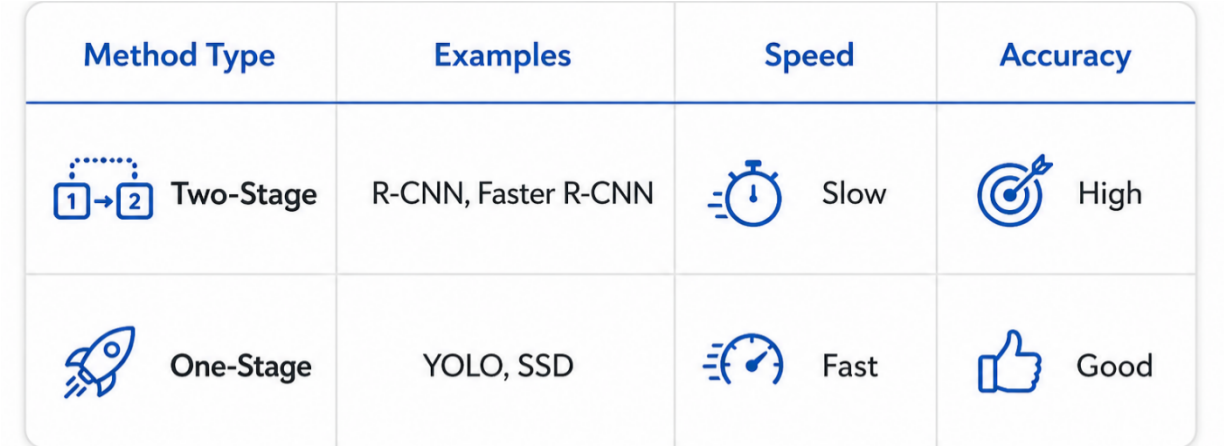


Figure 2.9: Object Detection Techniques

8. YOLO Algorithm

8.1. Introduction to YOLO

YOLO (You Only Look Once) is a state-of-the-art object detection algorithm based on deep learning. Unlike traditional methods, YOLO performs object detection in a single forward pass through the neural network, making it extremely fast and suitable for real-time applications.

The main idea behind YOLO is to treat object detection as a regression problem, where the model directly predicts bounding boxes and class probabilities from the input image.

As introduced by Joseph Redmon et al.:

We frame object detection as a regression problem to spatially separated bounding boxes

Edit with WPS Office

and associated class probabilities.[32]

This approach significantly improves detection speed compared to previous methods.

8.2. YOLO Architecture

YOLO uses a single Convolutional Neural Network (CNN) to process the entire image and predict multiple bounding boxes simultaneously. The image is divided into a grid of cells, where each cell is responsible for detecting objects whose centers fall within it.

Each grid cell predicts:

- Bounding box coordinates
- Confidence score
- Class probabilities

The YOLO architecture allows the model to analyze the entire image globally, rather than focusing on specific regions.

YOLO uses a unified architecture that applies a single neural network to the full image, enabling global reasoning and efficient detection.[32]

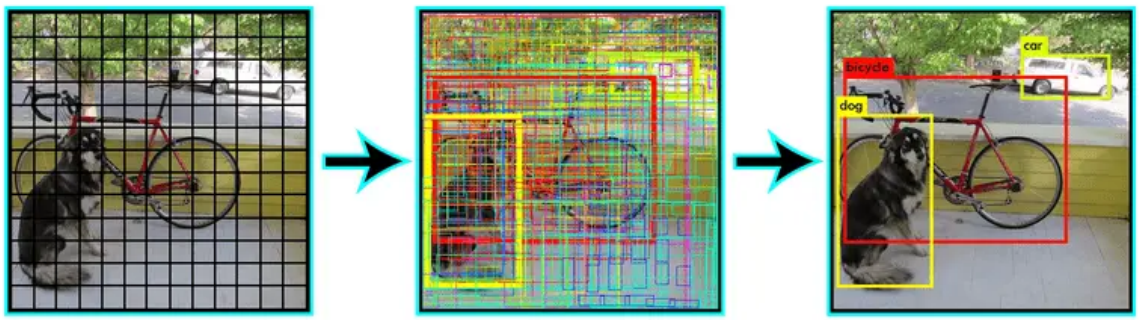


Figure 2.10: YOLO Architecture

8.3. How YOLO Works

YOLO divides the input image into an S × S grid. Each grid cell predicts a fixed number of bounding boxes along with confidence scores and class probabilities.

The confidence score indicates how likely it is that a bounding box contains an object, as well as how accurate the prediction is.

During detection, YOLO processes the image in one pass, making predictions for all objects

Edit with WPS Office

simultaneously.

Each grid cell predicts bounding boxes and confidence scores, allowing the model to detect multiple objects in a single evaluation.[32]

8.4. Advantages of YOLO

YOLO offers several advantages over traditional object detection methods. One of its main strengths is its high speed, which makes it suitable for real-time applications. In addition, YOLO processes the entire image at once, allowing it to capture contextual information more effectively.

Another advantage is its simplicity, as it uses a single neural network instead of multiple stages.

YOLO is extremely fast and can process images in real time while maintaining good accuracy compared to other detection methods.[33]

8.5. YOLO for Real-Time Detection

One of the most important features of YOLO is its ability to perform real-time object detection. This makes it particularly useful in applications that require immediate response, such as surveillance systems, autonomous vehicles, and robotics.

Due to its speed and efficiency, YOLO can process video streams and detect objects frame by frame without significant delay.

Real-time object detection is one of the key advantages of YOLO, enabling its use in time-sensitive applications.[33]

9. Underwater Image Processing Challenges

9.1. Light Absorption

Underwater environments are characterized by significant light absorption, which directly affects image quality. As light penetrates water, it is gradually absorbed, especially in deeper regions. This leads to darker images and loss of important visual details.

Different wavelengths of light are absorbed at different rates, with red light disappearing first, followed by green and blue. As a result, underwater images often appear bluish or greenish.

Edit with WPS Office

Light absorption in water significantly reduces image brightness and affects color representation, making underwater vision more challenging.[34]

9.2. Color Distortion

Color distortion is another major challenge in underwater imaging. Due to the selective absorption of light wavelengths, colors are not preserved accurately in underwater images.

Objects that appear red or orange in normal conditions may lose their original color and appear darker or completely different underwater. This distortion affects the performance of computer vision models, as color is an important feature for object recognition.

Underwater images suffer from severe color distortion due to wavelength-dependent light absorption.[35]

9.3. Low Visibility

Underwater visibility is often limited due to suspended particles, turbidity, and environmental conditions. These factors reduce the clarity of images and make it difficult to distinguish objects.

Low visibility leads to blurred images and reduces the contrast between objects and their background, which negatively impacts object detection accuracy.

Underwater visibility degradation is mainly caused by scattering and absorption of light, which reduces image contrast and clarity.[36]

9.4. Noise

Noise is commonly present in underwater images due to various environmental and technical factors. It can be caused by low lighting conditions, water particles, and limitations of imaging sensors.

Noise affects the quality of the captured images and makes it more difficult for algorithms to extract meaningful features.

Underwater images are often degraded by noise, which reduces the effectiveness of image processing and object detection techniques.[37]

Edit with WPS Office



Figure 2.11: Underwater Image Before and After Enhancement

Relation with Our Project

In this project, underwater image processing challenges have a direct impact on the performance of the object detection system. The video stream captured by the underwater camera is affected by low visibility, color distortion, and noise, which makes object detection more difficult.

To address these issues, deep learning models such as YOLO combined with CNNs are used, as they are capable of extracting robust features even in challenging conditions. This improves the accuracy and reliability of object detection in real-time underwater environments.

Edit with WPS Office

10. Conclusion

This chapter provided an overview of the fundamental concepts related to AI and CV, which form the theoretical foundation of this work. It began with a general introduction to AI and ML, highlighting their importance in developing intelligent systems capable of learning from data and making decisions.

The chapter then explored DL and CNN, emphasizing their effectiveness in processing and analyzing visual data. In addition, it presented the main tasks of CV and different object detection techniques, with a particular focus on the YOLO algorithm, which is widely used for real-time detection.

Furthermore, the challenges associated with underwater image processing were discussed, including light absorption, color distortion, low visibility, and noise. These factors significantly affect the quality of visual data and make object detection more complex in underwater environments.

Overall, this chapter established the necessary theoretical background for understanding the proposed system. The next chapter will focus on the practical implementation of the project, including the system architecture, hardware components, and the integration of the object detection model.

Edit with WPS Office

Chapter 3:

System Design and Architecture

Edit with WPS Office

1. Introduction

The success of an intelligent underwater robotic system depends on a well-designed architecture that ensures the seamless integration of hardware and software components. A robust system architecture not only guarantees reliable operation in challenging underwater environments but also facilitates communication between the different modules responsible for navigation, image acquisition, artificial intelligence, and user interaction.

In this project, **AquaSearch AI** was designed as an integrated underwater exploration platform combining robotics, computer vision, and artificial intelligence. The proposed system consists of a remotely operated vehicle (ROV) equipped with an underwater camera, propulsion mechanisms, buoyancy control, embedded computing, and an AI-based object detection module. Each subsystem was carefully selected and integrated to achieve real-time underwater monitoring and intelligent object detection while maintaining a compact and cost-effective design.

This chapter presents the overall architecture of the proposed system, followed by a detailed description of both the hardware and software components. It also explains the communication mechanism between the different modules and describes the complete workflow of the system during an underwater exploration mission

2 Overview of the Proposed System

The proposed system, **AquaSearch AI**, is an intelligent underwater remotely operated vehicle (ROV) designed to assist underwater exploration and object detection through the integration of robotics, embedded systems, and artificial intelligence. The system combines a waterproof mechanical structure, an embedded computing platform, sensing devices, and a computer vision module to perform real-time underwater monitoring and analysis.

At the core of the system is the **Raspberry Pi 5**, which serves as the main processing unit. It coordinates the operation of all hardware components, acquires video from the underwater camera, processes sensor data, and executes the **YOLOv8 Instance Segmentation** model for real-time object detection.

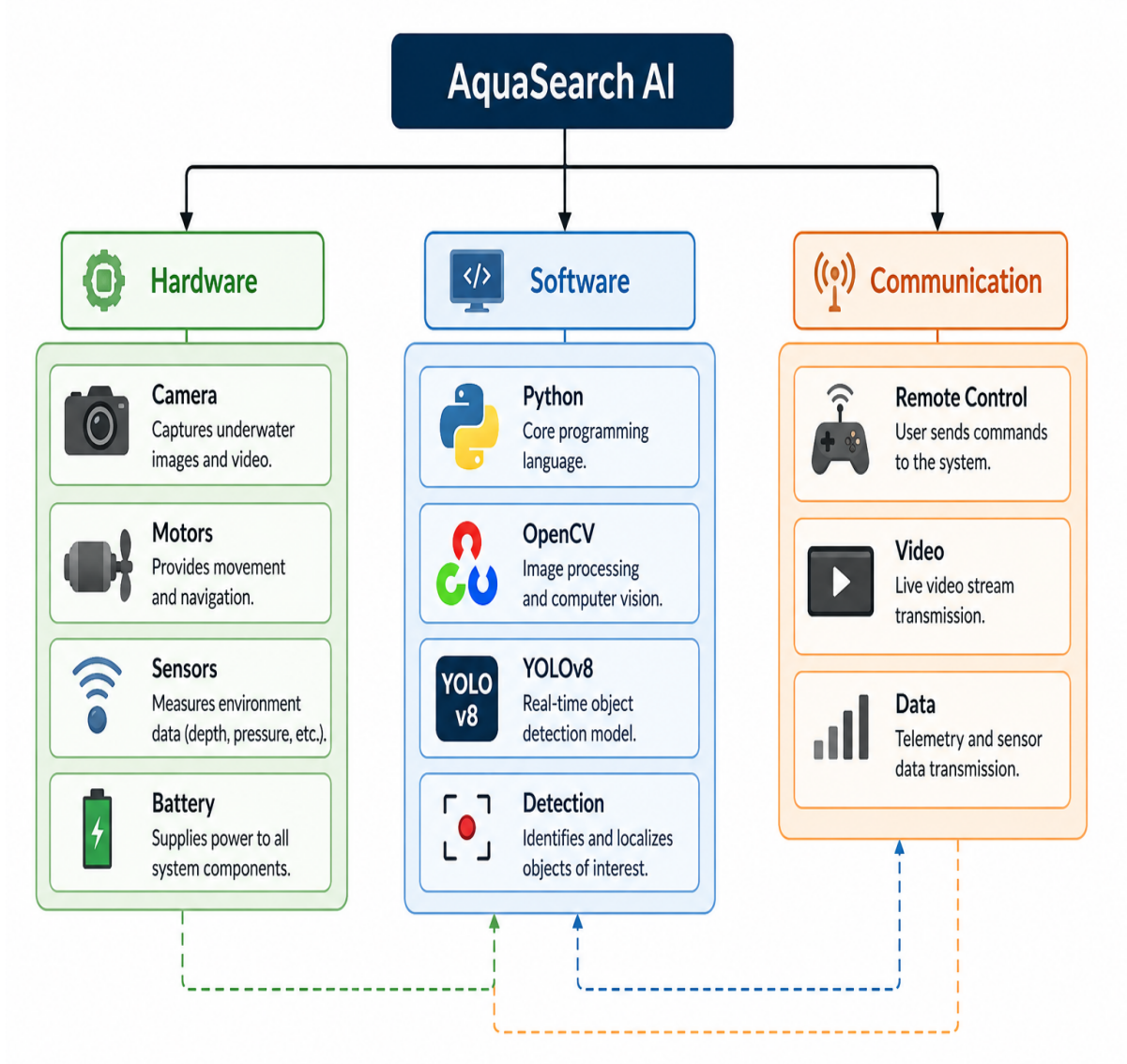
The propulsion system consists of DC motors coupled with propellers, allowing the ROV to navigate underwater. In addition, a buoyancy control mechanism based on **syringes, tubes, and N20 geared motors** enables the vehicle to perform controlled diving and ascending operations by regulating the amount of water inside the syringes.

Edit with WPS Office

To enhance environmental awareness, the prototype integrates multiple sensors, including an **MPU6050 IMU** for motion measurement, a **pressure sensing unit** for pressure monitoring, and a **VL53L0X Time-of-Flight (ToF) sensor** for short-distance obstacle detection.

The captured video stream is processed by the artificial intelligence module to identify underwater objects belonging to the trained dataset. The detection results can then be visualized by the operator, providing real-time assistance during underwater exploration missions.

The overall architecture of the proposed system is illustrated in **Figure 3.1**, while the hardware components are presented in **Figure 3.2**.



Edit with WPS Office

Figure 3.1: The overall architecture of the proposed system

3 Hardware Architecture

The hardware architecture of AquaSearch AI is composed of several interconnected components that work together to ensure reliable underwater navigation, real-time image acquisition, buoyancy control, and intelligent object detection. Each component was carefully selected according to the system requirements while considering compactness, cost, and ease of integration. Figure 3.2 illustrates the complete mechanical design and the arrangement of the hardware components inside the ROV.



Edit with WPS Office

Figure 3.2 hardware components

3.1 3D Printed Parts

The mechanical structure of the ROV relies on a set of custom-designed **3D printed parts**. These components provide the necessary support for assembling the different modules while maintaining a lightweight and compact design. They include the motor holders, internal supports, structural connectors, and mounting brackets used to securely position the electronic components inside the waterproof enclosure.

The use of additive manufacturing offers several advantages, including rapid prototyping, low production cost, and the possibility of easily modifying the design according to future project requirements. This approach also allows the prototype to be customized without the need for complex manufacturing processes.

3.2 Acrylic Waterproof Body

The main body of the ROV consists of a transparent acrylic tube that serves as the waterproof enclosure of the system. Its primary function is to protect the internal electronic components from water while allowing visual inspection of the internal assembly during testing and maintenance.

Acrylic was selected because it provides good mechanical strength, corrosion resistance, and transparency while remaining lightweight. The cylindrical structure also simplifies sealing and contributes to maintaining the integrity of the system during underwater operation.

3.3 Front and Rear End Caps

The acrylic tube is sealed at both ends using custom-designed front and rear end caps. These components ensure the watertight integrity of the enclosure while providing mechanical support for the propulsion system and the internal electronic modules.

In addition to sealing the enclosure, the end caps act as structural interfaces that connect the waterproof body to the motors, propellers, and other external components, ensuring the overall stability of the ROV during operation.

3.4 Propellers

The propulsion system uses underwater propellers attached directly to the DC motors. Their role is to convert the rotational motion generated by the motors into thrust, enabling the ROV to move through

Edit with WPS Office

the water.

The selected propellers were designed to provide sufficient thrust while maintaining stable movement and minimizing energy consumption. Their geometry contributes to improving propulsion efficiency during underwater navigation.

3.5 DC Motors

The movement of the ROV is achieved using DC motors coupled with the propellers. These motors generate the mechanical force required for forward, backward, and turning movements during underwater exploration.

The Raspberry Pi 5 sends control signals to the motor driver, which regulates the power supplied to each motor according to the desired direction and speed. This configuration allows smooth navigation while ensuring reliable propulsion throughout the mission**3.6 N20 3.6 Geared Motors**

In addition to the propulsion motors, the prototype incorporates **N20 geared motors** dedicated to the buoyancy control mechanism.

Unlike the propulsion motors, these geared motors are not responsible for moving the ROV horizontally. Instead, they drive the plungers of the syringe-based buoyancy system. Their high torque and low rotational speed make them particularly suitable for precise mechanical movements, allowing accurate control of the amount of water entering or leaving the syringes during diving and surfacing operations.

3.7 Motor Driver Module

The DC motors are controlled through a dedicated **motor driver module**, which acts as an interface between the Raspberry Pi 5 and the propulsion system. Since the Raspberry Pi cannot directly supply the current required by the motors, the motor driver amplifies the control signals and provides the appropriate power needed for motor operation.

By regulating the voltage and current delivered to each motor, the driver enables precise control of the ROV's movement, including forward and backward navigation as well as left and right turning. This configuration improves the stability and responsiveness of the propulsion system during underwater missions.

3.8 Raspberry Pi 5

Edit with WPS Office

The **Raspberry Pi 5** serves as the central processing unit of the AquaSearch AI system. It coordinates the operation of all hardware components, including the camera, sensors, motor driver, and artificial intelligence module.

One of the main reasons for selecting the Raspberry Pi 5 is its significantly improved computational performance, which allows it to execute the **YOLOv8 Instance Segmentation** model while simultaneously managing the different system processes. It also handles image acquisition, sensor communication, motor control, and the execution of the software responsible for real-time underwater object detection.

As the brain of the ROV, the Raspberry Pi 5 ensures efficient communication between all subsystems, enabling reliable and synchronized operation throughout the exploration mission.

3.9 MPU6050 IMU

The prototype integrates an **MPU6050 Inertial Measurement Unit (IMU)** to monitor the motion and orientation of the ROV during underwater navigation.

The sensor combines a three-axis accelerometer and a three-axis gyroscope, allowing the system to measure acceleration and angular velocity in real time. These measurements provide valuable information about the vehicle´s movement and orientation, which can be used to improve navigation accuracy and system stability.

3.10 Pressure Sensing Unit

A pressure sensing unit is incorporated into the prototype to measure the surrounding water pressure during underwater operation. Since water pressure increases with depth, this sensor provides useful information about the operating conditions of the ROV.

The acquired pressure measurements can be monitored by the control system and may also be used in future developments for estimating diving depth and improving autonomous underwater navigation.

3.11 18650 Batteries

The power required by the ROV is supplied by rechargeable **18650 lithium-ion batteries**. These batteries were selected due to their high energy density, compact dimensions, and ability to provide stable power for portable embedded systems.

Edit with WPS Office

They supply electrical energy to all onboard components, including the Raspberry Pi 5, motor driver, sensors, and propulsion system, ensuring continuous operation throughout the underwater mission.

3.12 DC–DC Converters

The different electronic components of the ROV require specific operating voltages. For this reason, **DC–DC converters** are used to regulate the voltage supplied by the battery pack and distribute stable power to the onboard electronics.

These converters protect sensitive electronic components from voltage fluctuations while improving the overall efficiency and reliability of the power management system.

3.13 Syringes and Tubes

One of the distinctive features of AquaSearch AI is its buoyancy control mechanism based on **syringes and flexible tubes.**

Instead of relying on conventional ballast tanks, the prototype controls buoyancy by regulating the volume of water inside the syringes. The N20 geared motors move the syringe plungers, allowing water to enter or leave the system. Increasing the amount of water increases the overall density of the ROV, causing it to descend, whereas removing water decreases its density, allowing it to rise toward the surface.

This simple mechanical solution offers a low-cost and effective method for controlling the vertical movement of the prototype while reducing the complexity of the overall design.

3.14 VL53L0X Time-of-Flight Sensor

To improve the environmental awareness of the ROV, a **VL53L0X Time-of-Flight (ToF) sensor** is integrated into the system.

The sensor measures the distance between the ROV and nearby objects by calculating the time taken for a laser pulse to travel to the target and return to the sensor. This information can assist the operator during navigation by detecting nearby obstacles and helping to avoid collisions in confined underwater environments.

Although the sensor has a relatively short measurement range, it provides accurate distance measurements suitable for close-range underwater applications.

3.15 Internal Arrangement of Components

The internal arrangement of the hardware components plays a fundamental role in ensuring

Edit with WPS Office

the stability, reliability, and overall performance of the proposed ROV. Since the available space inside the waterproof enclosure is limited, each component was carefully positioned to optimize weight distribution, simplify electrical connections, facilitate maintenance, and improve the overall balance of the vehicle during underwater operation.

The underwater camera and the distance sensor were installed at the front section of the ROV to maximize the forward field of view and enable obstacle detection during navigation. Positioning these components at the front allows the artificial intelligence module to continuously analyze the environment located directly ahead of the vehicle.

The propulsion motors were mounted symmetrically on both sides of the ROV to generate balanced thrust during movement. This symmetrical configuration minimizes undesired rotational effects and improves directional stability while navigating underwater.

The buoyancy control mechanism, composed of syringes, flexible tubes, and N20 geared motors, was installed along the upper central part of the enclosure. This location allows the buoyancy system to operate efficiently without interfering with the propulsion system or the electronic modules.

The Raspberry Pi 5 was positioned near the center of the enclosure, where it acts as the central processing unit. This central location reduces cable lengths between the processing unit, sensors, motor drivers, and camera, thereby improving communication reliability and simplifying the internal wiring.

The motor driver module was installed close to the Raspberry Pi and the propulsion motors in order to minimize power losses and reduce electrical interference. Similarly, the pressure sensor and the IMU were placed near the central body of the ROV to provide more stable and accurate measurements during underwater movement.

Finally, the battery pack and the DC–DC converters were positioned toward the rear section of the enclosure. This arrangement contributes to maintaining the longitudinal balance of the ROV while ensuring stable power distribution to all onboard electronic components.

Overall, the selected arrangement improves the compactness of the prototype, simplifies maintenance operations, and contributes to the mechanical stability required for underwater exploration.

Edit with WPS Office

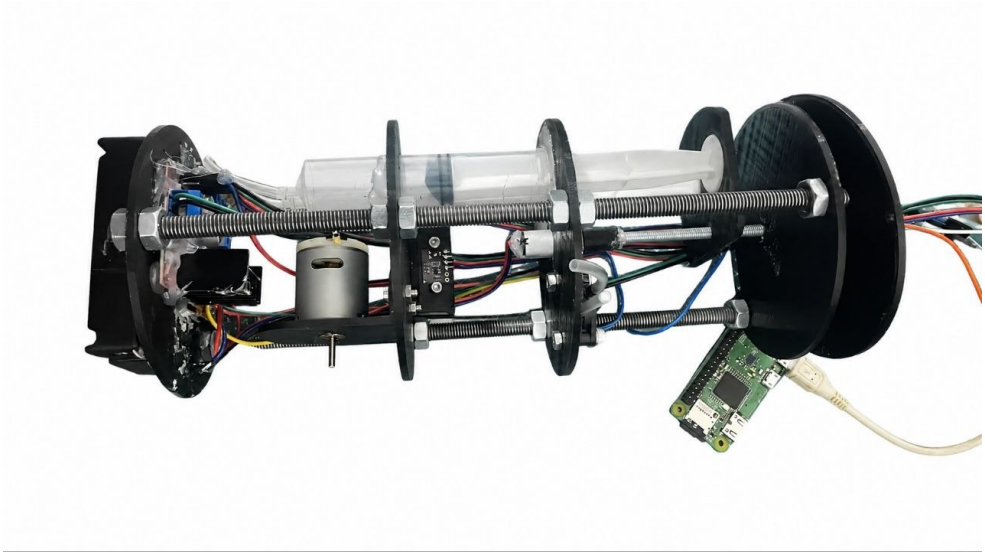


Figure 3.3: *Internal arrangement of the hardware components inside the proposed ROV prototype.*

3.16 Mechanical Operation of the ROV

The underwater locomotion of AquaSearch AI is achieved through the integration of two complementary mechanical systems: the propulsion mechanism and the buoyancy control mechanism. While the propulsion system is responsible for generating horizontal movement, the buoyancy mechanism controls the vertical displacement of the vehicle by modifying its overall density.

The combination of these two mechanisms enables the ROV to perform the basic underwater movements required during exploration missions, including forward navigation, turning, diving, and ascending.

3.16.1 Propulsion Mechanism

The propulsion mechanism is based on a set of DC motors mechanically coupled with underwater propellers. These motors receive electrical power through the motor driver module, which is controlled by the Raspberry Pi 5 according to the commands issued by the operator.

When an electric current is supplied to a DC motor, the motor shaft begins to rotate. Since the propeller is directly attached to the motor shaft, its rotation generates a flow of water opposite to the

Edit with WPS Office

desired direction of movement. According to Newton´s Third Law of Motion, this action produces an equal and opposite reaction force that propels the ROV through the water.

The propulsion system enables several movement modes.

When both propulsion motors rotate at the same speed and in the same direction, the generated thrust is balanced, allowing the ROV to move forward in a straight line.

If the rotation direction of both motors is reversed, the generated thrust also reverses, causing the vehicle to move backward.

Turning maneuvers are achieved by creating a difference in thrust between the two propulsion motors. For example, increasing the speed of the left motor while reducing the speed of the right motor produces a rotational moment that turns the vehicle toward the right. Conversely, increasing the thrust on the right side causes the ROV to rotate toward the left.

The motor driver continuously regulates the voltage delivered to each motor according to the commands received from the Raspberry Pi, allowing smooth speed control and precise maneuverability during underwater navigation.

The efficiency of the propulsion system depends on several factors, including the geometry of the propellers, motor speed, water resistance, and the hydrodynamic design of the ROV. The selected configuration provides sufficient thrust for the prototype while maintaining low energy consumption and stable underwater movement.

3.16.2 Buoyancy Control Mechanism

Unlike conventional underwater vehicles that rely on ballast tanks and water pumps, AquaSearch AI employs an original buoyancy control mechanism based on syringes, flexible tubes, and N20 geared motors.

The principle of operation is based on **Archimedes´ Principle**, which states that an object immersed in a fluid experiences an upward buoyant force equal to the weight of the displaced fluid. Consequently, the vertical movement of the ROV depends on the relationship between its weight and the buoyant force generated by the surrounding water.

The buoyancy mechanism consists of medical syringes connected to flexible tubes. Each syringe is mechanically actuated by an N20 geared motor that moves the syringe plunger with high precision.

Edit with WPS Office

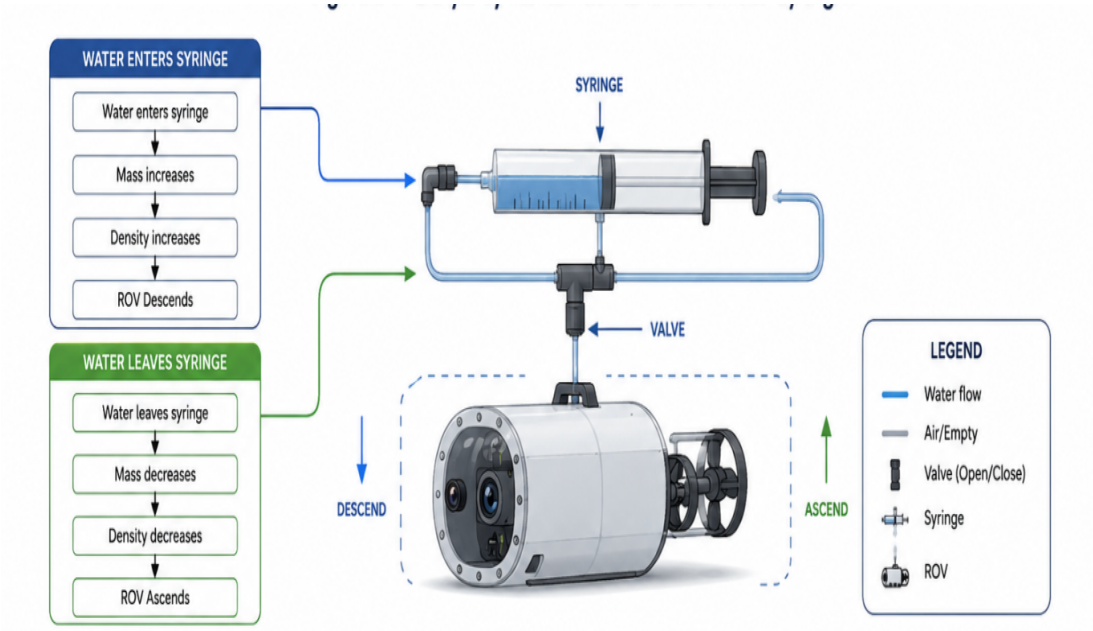
When the geared motor pushes the plunger backward, water enters the syringe through the connected tube. As additional water is stored inside the syringes, the total mass of the ROV increases while its external volume remains almost constant. This increase in average density causes the vehicle to sink gradually.

Conversely, when the geared motor pushes the plunger forward, the stored water is expelled from the syringe back into the surrounding environment. As the amount of internal water decreases, the total mass of the ROV is reduced, decreasing its average density. Consequently, the buoyant force becomes greater than the vehicle's weight, allowing the ROV to ascend toward the water surface.

The use of N20 geared motors is particularly advantageous because they provide high torque at low rotational speeds. This characteristic allows the plungers to move smoothly and accurately, ensuring gradual buoyancy adjustments rather than abrupt vertical movements.

Compared with conventional ballast systems, the syringe-based mechanism offers several advantages. It is mechanically simple, lightweight, inexpensive, easy to manufacture using commercially available components, and requires relatively low electrical power. These characteristics make it particularly suitable for prototype development and educational underwater robotic platforms.

The integration of this buoyancy control mechanism with the propulsion system enables AquaSearch AI to perform complete three-dimensional underwater navigation while maintaining mechanical simplicity and reducing the overall cost of the prototype.



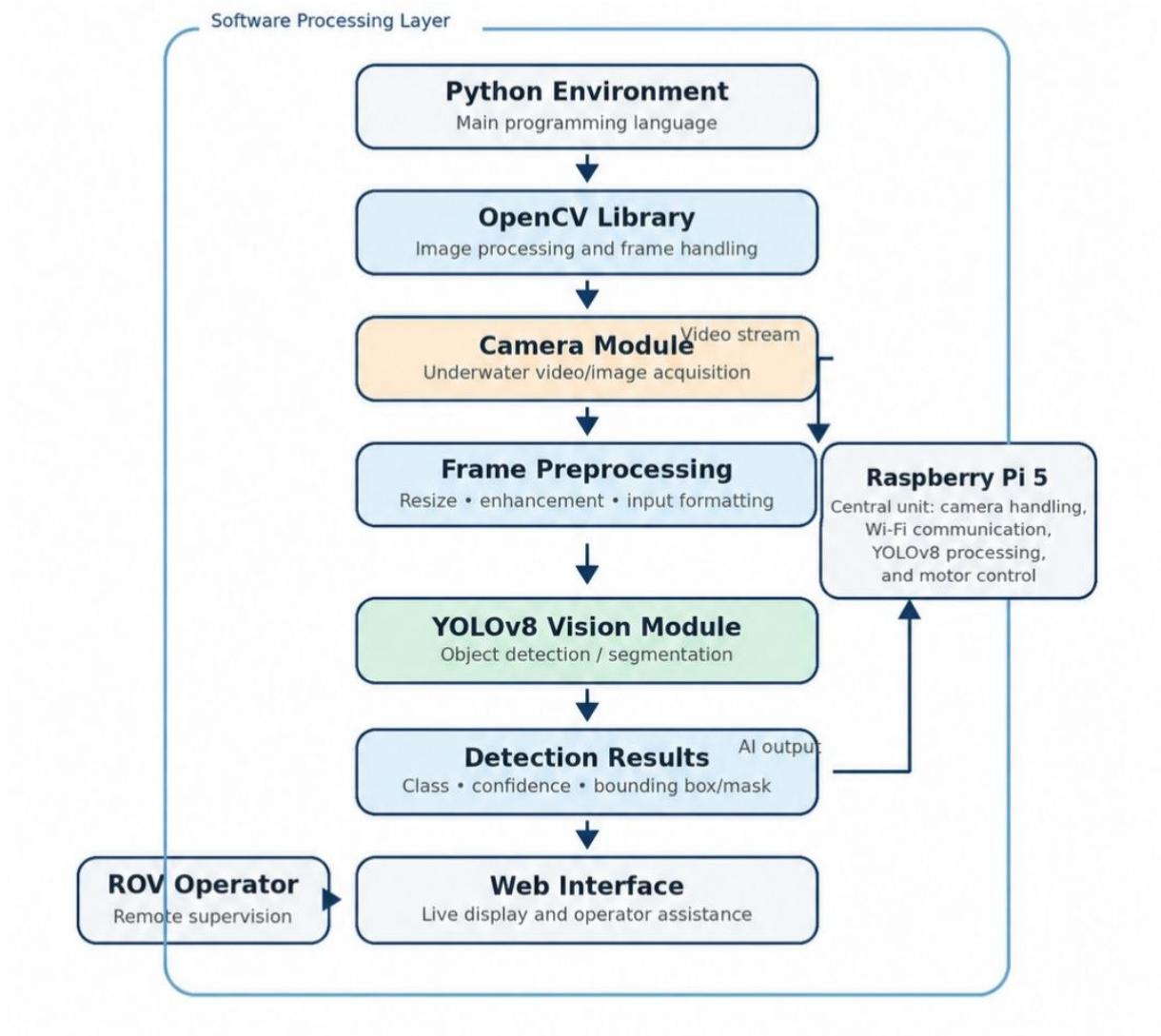
Edit with WPS Office

Figure 3.4 : syringe based buoyancy control mechanism

4 Software Architecture

The software architecture of **AquaSearch AI** is responsible for coordinating the different software components that enable system control, image processing, artificial intelligence, and communication between the hardware modules. It provides the necessary environment for acquiring video data, processing images in real time, executing the object detection model, and ensuring reliable interaction between the embedded system and the operator.

The software architecture is organized into three main components: the programming environment, the video streaming system, and the artificial intelligence detection module. Together, these components allow the ROV to perform intelligent underwater exploration while maintaining efficient and reliable operation.



Edit with WPS Office

Figure 3.5 : Software architecture of AquaSearch AI

4.1 Programming Environment

The software of AquaSearch AI was developed using the **Python** programming language due to its simplicity, flexibility, and extensive support for computer vision and artificial intelligence applications.

The development process relied on several Python libraries to control the hardware components, process video streams, communicate with sensors, and execute the deep learning model. In particular, the **Ultralytics YOLOv8** framework was employed to perform real-time instance segmentation of underwater objects.

The software was executed on the **Raspberry Pi 5**, which integrates all software modules into a single embedded platform capable of controlling the ROV while simultaneously performing artificial intelligence inference.

4.2 Video Streaming System

The underwater camera continuously captures video frames from the surrounding environment during the mission. These frames are transmitted directly to the Raspberry Pi 5, where they are processed in real time.

The streaming system serves two main purposes. First, it provides the artificial intelligence module with live image data for object detection. Second, it enables the operator to monitor the underwater environment through a continuous video feed, improving situational awareness during exploration.

The use of real-time video streaming allows the system to immediately analyze underwater scenes and provide visual feedback without significant delay.

4.3 AI Detection Module

The artificial intelligence module represents the core component of AquaSearch AI. Its primary function is to analyze the incoming video frames and identify underwater objects using the **YOLOv8 Instance Segmentation** model.

Before deployment, the model was trained on a custom underwater dataset containing multiple object classes, including marine organisms, underwater waste, and remotely operated vehicles. During operation, each captured frame is processed by the trained model, which generates segmentation masks, class labels, and confidence scores for every detected object.

Edit with WPS Office

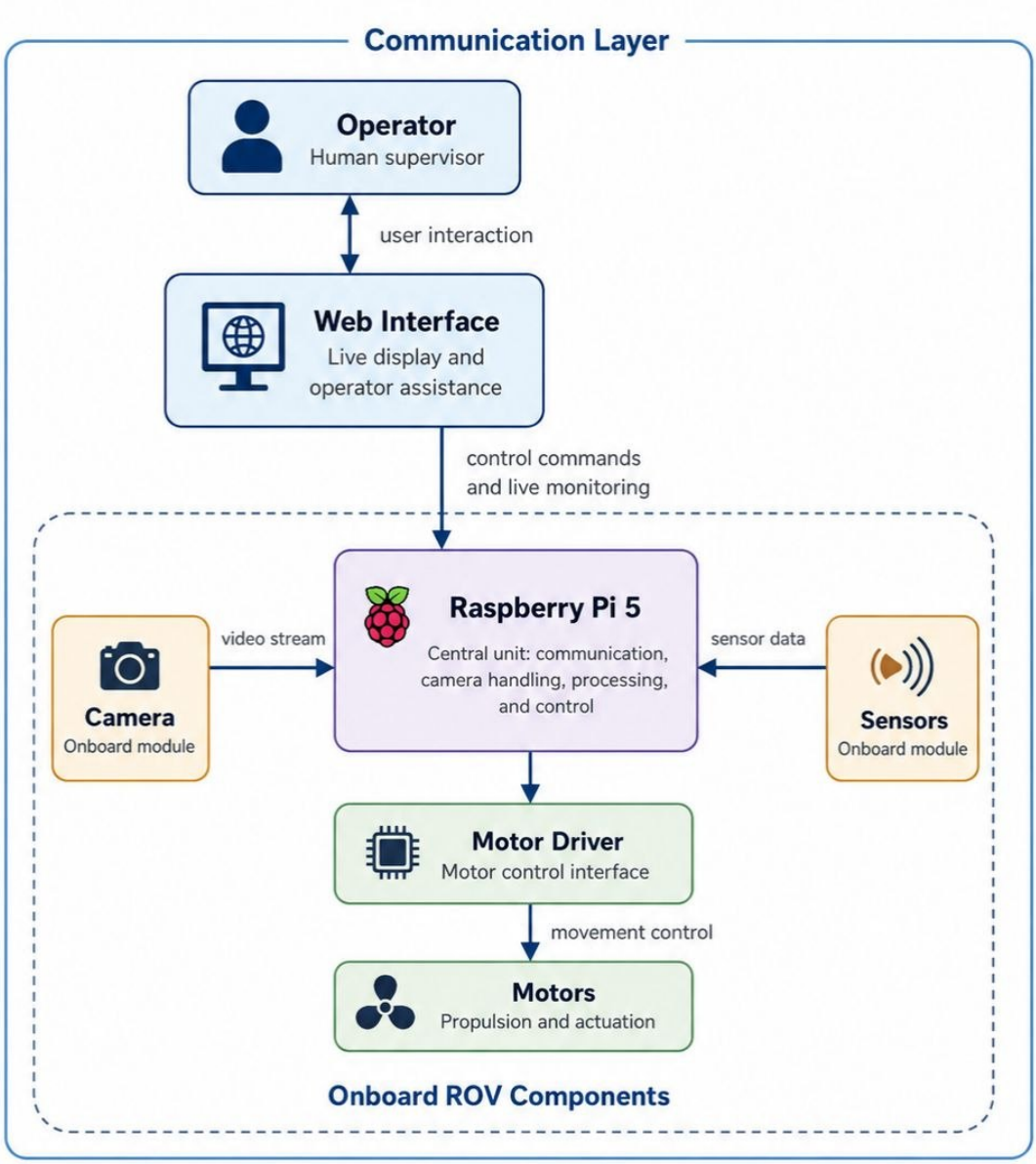
By performing real-time instance segmentation, the system assists the operator in identifying objects of interest more efficiently, thereby improving the effectiveness of underwater exploration missions.

5 Communication System

Reliable communication is essential for ensuring efficient interaction between the operator and the underwater robotic system. In AquaSearch AI, the communication system enables the transmission of control commands, sensor information, and video data, allowing the operator to monitor the underwater environment and supervise the ROV during the mission.

The communication architecture consists of two main components: the remote control system and the data transmission process.

Edit with WPS Office



Edit with WPS Office

Figure 3.6 : communication architecture of AquaSearch AI

5.1 Remote Control System

The ROV is operated remotely through a control interface that allows the user to supervise the mission and manage the movement of the vehicle. Control commands generated by the operator are received by the Raspberry Pi 5, which interprets these instructions and coordinates the corresponding actions of the propulsion system.

This approach provides the operator with direct control over the vehicle while maintaining continuous interaction with the different onboard components throughout the underwater mission.

5.2 Data Transmission

During operation, the different hardware modules continuously exchange information with the Raspberry Pi 5. The underwater camera provides a live video stream, while the onboard sensors transmit environmental and motion-related measurements to the embedded computer.

The Raspberry Pi processes these data in real time before executing the artificial intelligence model. Once the analysis is completed, the detection results are displayed to the operator, providing immediate visual feedback and improving the efficiency of underwater exploration.

The continuous exchange of control commands, video frames, and sensor data ensures the reliable operation of the entire system.

6 System Workflow

The overall operation of AquaSearch AI follows a sequential workflow that integrates hardware control, image acquisition, artificial intelligence processing, and result visualization. Figure 3.3 illustrates the main operational stages of the proposed system.

The workflow begins when the operator starts the mission and remotely controls the ROV. The Raspberry Pi 5 receives these commands and coordinates the propulsion system while simultaneously collecting video frames from the underwater camera and measurements from the onboard sensors.

The captured video is then processed by the YOLOv8 Instance Segmentation model, which identifies and segments the underwater objects contained in each frame. Finally, the detection results are displayed to the operator in real time, allowing continuous monitoring of the underwater environment and facilitating decision-making during the mission.

Edit with WPS Office

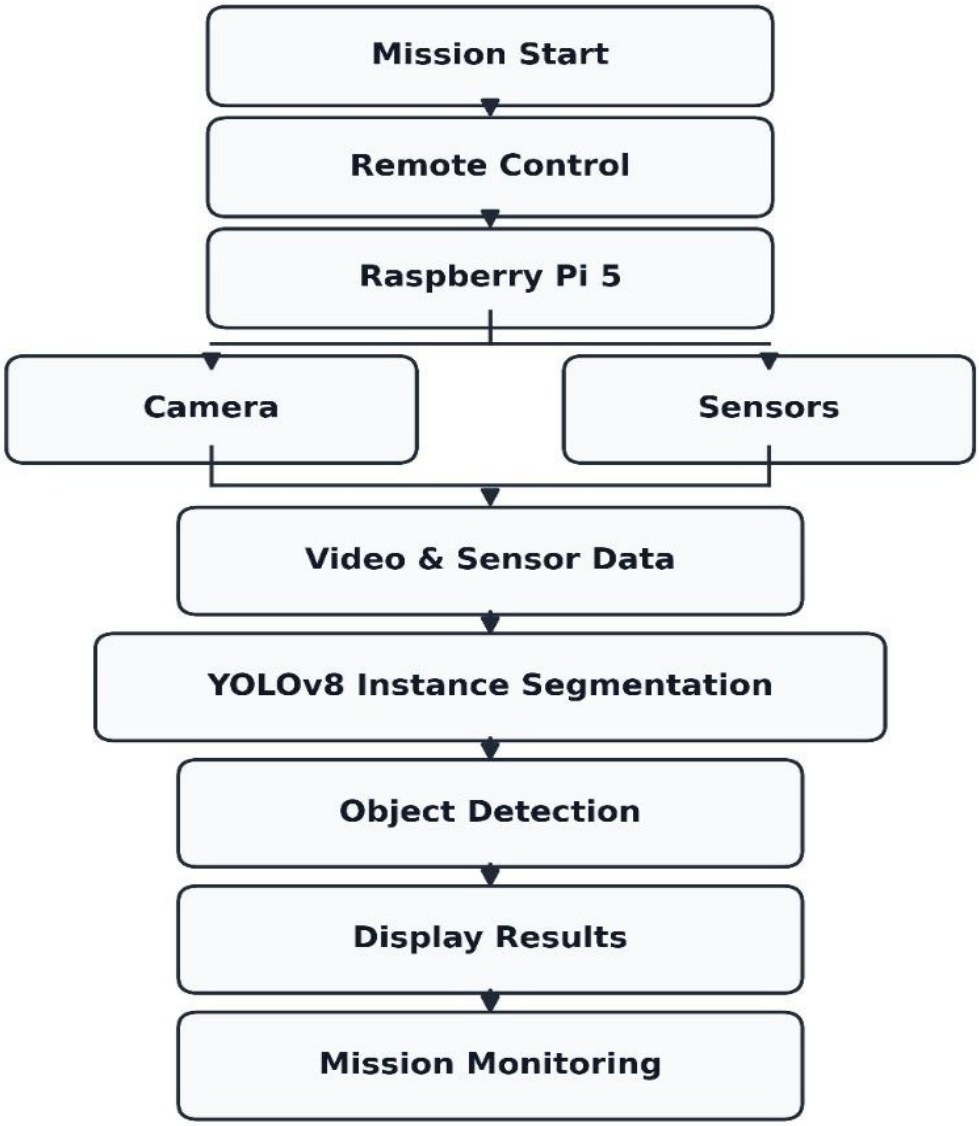


Figure 3.7: *Operational workflow of AquaSearch AI.*

Edit with WPS Office

Conclusion

This chapter presented the design and architecture of the proposed AquaSearch AI system. It described the overall organization of the platform, including its hardware components, software architecture, communication mechanism, and operational workflow.

The integration of embedded computing, underwater sensing, real-time video processing, and artificial intelligence provides a robust foundation for intelligent underwater exploration. The architecture was designed to ensure reliable operation while maintaining flexibility for future improvements and additional functionalities.

The next chapter focuses on the implementation and evaluation of the artificial intelligence module, including dataset preparation, model training, and the analysis of the experimental results obtained during the development of the system.

Edit with WPS Office

Chapter 4:

Implementation and Model Development

Edit with WPS Office

1. Introduction

The previous chapters presented the theoretical foundations of Artificial Intelligence and Computer Vision, as well as the design and hardware architecture of the proposed underwater exploration system. This chapter focuses on the practical implementation of the intelligent vision module integrated into the Remotely Operated Vehicle (ROV), describing the complete development process from data preparation to the deployment of the trained Artificial Intelligence model.

The primary objective of the proposed system is to improve underwater exploration by enabling the ROV to automatically recognize and identify objects encountered in the underwater environment. Unlike conventional underwater vehicles that simply transmit live video to the operator, the developed system incorporates an Artificial Intelligence model capable of analyzing underwater scenes and providing semantic information about the detected objects. This functionality allows the operator to better understand the surrounding environment during exploration missions.

To achieve this objective, an underwater image dataset containing various marine organisms, aquatic vegetation, artificial objects, and underwater debris was selected. These images represent the types of objects that the onboard camera of the ROV is expected to observe during real underwater exploration. The selected dataset therefore constitutes the learning foundation of the intelligent vision module.

After preparing the dataset, an Instance Segmentation model based on the YOLOv8 architecture was developed using the Ultralytics framework. A transfer learning approach was adopted to adapt a pre-trained model to underwater exploration scenarios while reducing the computational cost and training time. Several YOLOv8 segmentation models were trained and evaluated in order to determine the most suitable architecture for the proposed underwater exploration platform.

Once the training process was completed, the best-performing model was selected to serve as the visual intelligence component of the ROV. During operation, the model analyzes the images captured by the underwater camera, detects the objects present in the scene, and generates segmentation masks that accurately delineate their boundaries. This information assists the operator in recognizing marine organisms, underwater vegetation, and submerged debris, thereby improving the overall efficiency of underwater exploration missions.

In addition to the implementation of the Artificial Intelligence module, a dedicated project presentation website was developed to showcase the underwater exploration system. The website presents the objectives of the project, its technical characteristics, the adopted technologies, and the main functionalities of the developed ROV. It serves as a communication platform that promotes the project and provides visitors with a comprehensive overview of the proposed intelligent underwater exploration solution.

This chapter describes each stage of the implementation process, including dataset collection, data preprocessing, annotation, model development, performance evaluation, experimental analysis, and

Edit with WPS Office

the development of the project presentation website.

2. Dataset Collection

2.1 Selection of the Underwater Dataset

One of the main challenges encountered during the development of this project was the availability of underwater datasets specifically designed for exploration tasks. Although the initial idea was to develop an intelligent system capable of recognizing drowning victims, no sufficiently large and publicly available annotated dataset containing underwater images of drowning persons could be found.

For this reason, an alternative solution was adopted by selecting an underwater instance segmentation dataset containing a wide variety of objects commonly encountered in marine environments. The selected dataset includes marine organisms, aquatic vegetation, underwater debris, and artificial objects, making it more suitable for the primary objective of the project, which is underwater exploration and environmental observation rather than rescue operations.

This choice allows the developed ROV to assist the operator in identifying different underwater elements during exploration missions while maintaining the possibility of extending the system in the future by training it on additional datasets dedicated to specific applications.

2.2 Characteristics of the Selected Dataset

The selected dataset corresponds to the TrashCan 1.0 Material Instance Segmentation Dataset, which contains underwater images acquired in realistic marine environments. The images were collected under different lighting conditions, water visibility levels, camera viewpoints, and object scales, making the dataset representative of real underwater exploration scenarios.

Unlike conventional image datasets, underwater images present additional visual challenges such as color attenuation, suspended particles, light reflections, and partial object occlusions. Training the AI model on such conditions improves its ability to operate reliably during underwater exploration missions.

The dataset contains :

Edit with WPS Office

Property	Value
Training instances	6008
Validation instances	1204
Number of classes	16
Annotation type	Instance Segmentation
Environment	Real underwater scenes

Table 4.1 : Characteristics of the TrashCan instance segmentation Dataset.

2.3 Underwater Object Categories

To make the intelligent exploration system capable of understanding different underwater environments, the dataset includes sixteen object categories representing both natural and artificial elements.

These categories are summarized in Table 4.2.

ID	Class	Description
0	ROV	Remotely Operated Vehicle
1	Plant	Aquatic vegetation
2	Animal Fish	Fish
3	Animal Starfish	Starfish
4	Animal Shells	Shells
5	Animal Crab	Crabs
6	Animal Eel	Eels
7	Animal Etc	Other marine animals
8	Trash Etc	Miscellaneous debris
9	Trash Fabric	Fabric waste
10	Trash Fishing Gear	Fishing nets and ropes
11	Trash Metal	Metal debris

Edit with WPS Office

ID	Class	Description
12	Trash Paper	Paper waste
13	Trash Plastic	Plastic waste
14	Trash Rubber	Rubber waste
15	Trash Wood	Wood debris

Table 4.2 : Underwater Object Categories.

The diversity of these categories enables the developed AI model to recognize various underwater objects that may appear during exploration missions. Instead of detecting only a single object type, the intelligent vision module is capable of interpreting complex underwater scenes containing multiple natural and artificial elements.

2.4 Dataset Organization

Before starting the training process, the dataset was organized according to the directory structure required by the Ultralytics YOLOv8 framework. The images and their corresponding annotation files were divided into two subsets:

- **Training Set:** used to teach the AI model the visual characteristics of underwater objects.
- **Validation Set:** used to evaluate the learning progress and verify the generalization capability of the trained model.

Edit with WPS Office

The adopted folder structure is illustrated in Figure 4.1.

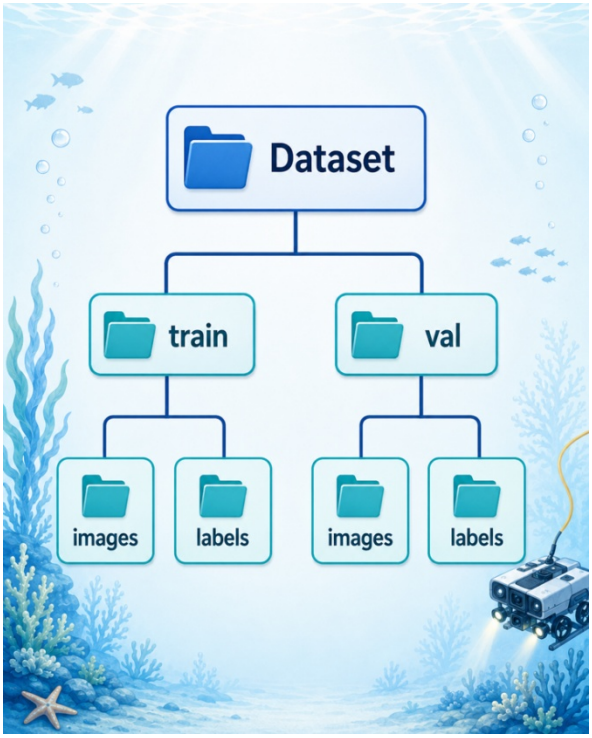


Figure 4.1 : Directory structure of the TrashCan instance segmentation Dataset.

This organization allows the training framework to automatically associate each underwater image with its corresponding annotation file during the learning process.

2.5 Importance of the Dataset for the Proposed ROV

The selected dataset constitutes the knowledge base of the Artificial Intelligence module integrated into the underwater exploration system. During training, the neural network analyzes thousands of underwater images to learn the visual characteristics of marine organisms, underwater vegetation, and submerged debris.

After the learning phase, this knowledge is embedded within the trained model and becomes available during exploration missions. Consequently, every image captured by the ROV camera can be compared with the learned visual patterns, enabling the system to recognize underwater objects automatically and provide meaningful information to the operator.

Therefore, the effectiveness of the intelligent exploration system depends largely on the quality, diversity, and realism of the selected dataset.

3. Data Preprocessing

Edit with WPS Office

3.1 Dataset Preparation

The selected underwater dataset was initially stored as a compressed archive. During the implementation process, it was extracted into the Google Colaboratory working environment, where all images and annotation files became accessible for the training pipeline.

After extraction, the directory structure was automatically verified to ensure that the dataset followed the organization required by the Ultralytics framework. This verification guarantees that each underwater image is correctly associated with its corresponding annotation file before the learning process begins.



Figure 4.2 : dataset directory organization in the file system.

3.2 Dataset Configuration

Once the dataset was organized, a configuration file in YAML format was created to establish the connection between the dataset and the YOLOv8 training framework.

This configuration file contains all the information required by the learning algorithm, including the dataset location, the paths of the training and validation images, and the list of underwater object classes. During training, the framework automatically reads this file to locate the images and their corresponding annotations.

Edit with WPS Office

Using a YAML configuration file simplifies the implementation process and ensures that the Artificial Intelligence module is trained using the correct dataset structure.

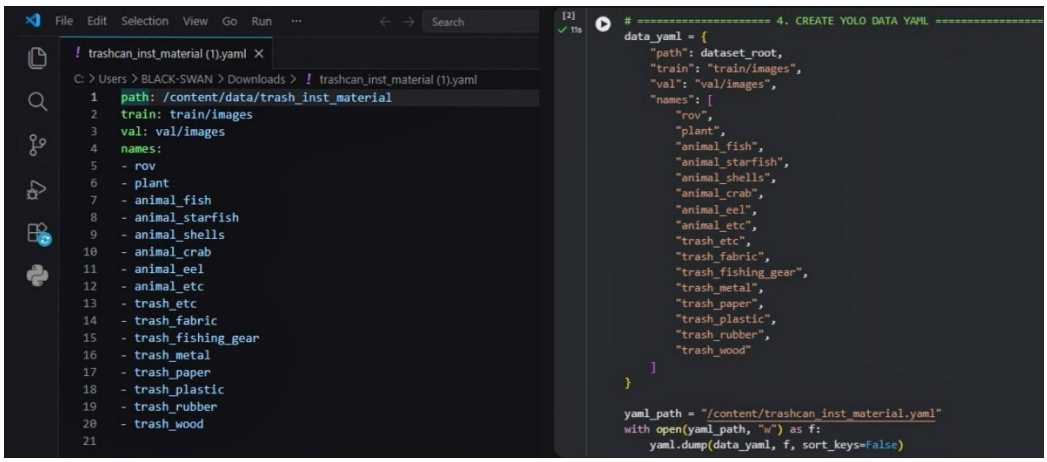


Figure 4.3 : YAML Configuration file for dataset.

3.3 Image Preparation

One of the main characteristics of underwater exploration is the variability of environmental conditions. Images acquired underwater often present reduced visibility, non-uniform illumination, suspended particles, and color degradation caused by water absorption.

Instead of applying manual image enhancement techniques, the proposed system was trained directly using realistic underwater images contained in the selected dataset. This approach allows the Artificial Intelligence model to learn the natural appearance of underwater scenes and improves its ability to operate under real exploration conditions.

Training the model using realistic underwater images contributes to the robustness of the intelligent vision module and reduces the gap between the training environment and real underwater operation.

Edit with WPS Office



Figure 4.4 : Sample image from the dataset.

3.4 Data Augmentation

To improve the robustness of the intelligent vision system, data augmentation techniques were employed during the training process.

Rather than generating new images manually, the Ultralytics YOLOv8 framework automatically applies several transformations while loading the training data. This online augmentation strategy increases the diversity of the training samples without modifying the original dataset stored on disk.

Among the augmentation operations applied during training are:

- Horizontal flipping.
- Random scaling.
- Image translation.
- Rotation.
- Perspective transformation.
- Brightness and color adjustment.
- Mosaic augmentation.

These transformations expose the Artificial Intelligence model to multiple visual situations similar to those encountered during underwater exploration, enabling the developed ROV to recognize underwater objects from different viewpoints, distances, and illumination conditions.

Edit with WPS Office

3.5 Contribution of Preprocessing to the Underwater Exploration System

Data preprocessing represents an essential stage in the development of the intelligent underwater exploration platform. Proper organization of the dataset ensures a reliable learning process, while automatic augmentation improves the generalization capability of the Artificial Intelligence model.

As a result, the developed vision module becomes more robust when analyzing underwater scenes captured during exploration missions. This contributes directly to improving the reliability of underwater object recognition and enhances the overall performance of the proposed ROV.

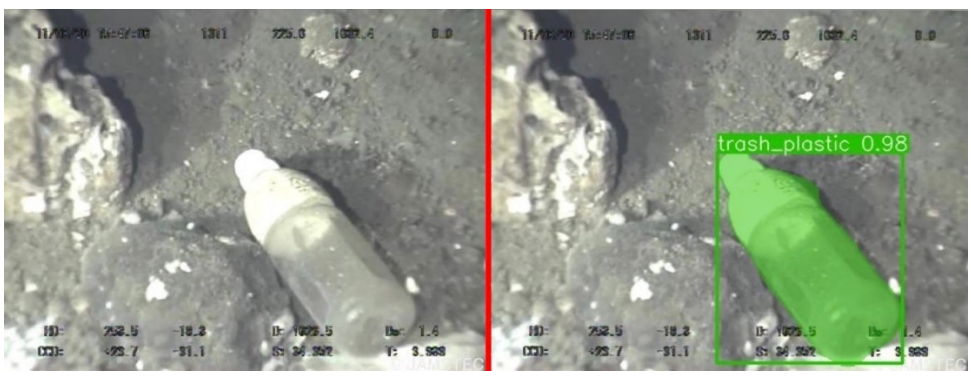
4. Underwater Data Annotation and AI Configuration

4.1 Instance Segmentation Annotation

Unlike conventional object detection, which only predicts rectangular bounding boxes around objects, the proposed system relies on **Instance Segmentation**.

This technique enables the Artificial Intelligence model to identify each underwater object individually while simultaneously determining its precise contour. As a result, the generated output contains not only the object category but also a segmentation mask accurately following its boundaries.

Such detailed information is particularly valuable in underwater exploration because many marine organisms and underwater objects have irregular shapes that cannot be represented accurately by simple bounding boxes. Instance segmentation therefore provides richer visual information that assists the operator in understanding complex underwater scenes.



Edit with WPS Office

Figure 4.5 : comparison between the original image and the instance segmentation result.

4.2 YOLOv8 Annotation Format

The selected underwater dataset was already converted into the annotation format required by the YOLOv8 framework.

Instead of storing annotations as JSON files, each image is associated with a text file containing all object annotations. Every line of the file corresponds to one object and includes:

- the object class identifier,
- the normalized bounding box coordinates,
- the polygon points defining the segmentation mask.

This annotation format allows the YOLOv8 framework to simultaneously learn object localization and pixel-level segmentation during the training process.

```
2 0.722917 0.616661 0.745833 0.608328 0.766663 0.583333 0.772917
8 0.308333 0.374993 0.347902 0.347222 0.347894 0.333333 0.354161
```

Figure 4.6 : Example of an instance segmentation Annotation File.

4.3 Dataset Configuration using YAML

To connect the underwater dataset with the YOLOv8 training framework, a YAML configuration file was created.

This file acts as the communication bridge between the dataset and the learning algorithm. It specifies:

- the dataset location,
- the training image directory,
- the validation image directory,
- the list of underwater object categories.

During the training process, the Ultralytics framework automatically reads this configuration file to access all required resources without additional programming.

4.4 Contribution of Annotation to the AI Vision Module

Edit with WPS Office

The annotation process represents one of the most important stages in the development of the Artificial Intelligence vision module.

By associating each underwater object with its corresponding class and segmentation mask, the neural network gradually learns the visual characteristics of the different underwater categories throughout the training process.

Consequently, once integrated into the proposed underwater exploration system, the trained model becomes capable of automatically recognizing underwater objects captured by the onboard camera and presenting meaningful visual information to the operator.

5. Development and Training of the AI Vision Module

5.1 Development Environment

The implementation of the Artificial Intelligence module was carried out using **Google Colaboratory**, which provides cloud-based computational resources suitable for deep learning applications.

The complete development environment includes:

Component	Purpose
Google Colab	Model development and training
Python	Programming language
Ultralytics	YOLOv8 implementation
PyTorch	Deep learning framework
Matplotlib	Visualization of training results
Google Drive	Dataset and model storage

Table 4.3 : software environment and devlopement tools.

The use of Google Colab eliminated the need for powerful local hardware while providing GPU acceleration that considerably reduced the training time.

Edit with WPS Office

5.2 Transfer Learning Strategy

Training a deep neural network from the beginning generally requires millions of images and considerable computational resources.

To overcome this limitation, the proposed underwater exploration system adopts a **Transfer Learning** strategy.

Instead of initializing the neural network with random parameters, pre-trained YOLOv8 segmentation models are used as the starting point. These models have already learned general visual features from large-scale image datasets.

During the training process, the model adapts these previously learned features to underwater exploration by learning the appearance of marine organisms, aquatic vegetation, and underwater debris.

This approach improves convergence speed while reducing the amount of training data required.

5.3 Selection of YOLOv8 Models

To determine the most appropriate architecture for the proposed underwater exploration system, three YOLOv8 Instance Segmentation models were evaluated.

Model	Characteristics
YOLOv8n-seg	Smallest model with very fast inference
YOLOv8s-seg	Balanced between speed and accuracy
YOLOv8m-seg	Higher accuracy with increased computational requirements

Table 4.4 :comparison of YOYOv8 segmentations models.

Rather than selecting the model solely based on speed, the comparison focused on obtaining the best compromise between segmentation accuracy and computational efficiency for underwater exploration applications.

5.4 Training Configuration

The training process was configured according to the implementation adopted for the proposed underwater exploration system.

The main training parameters are summarized in Table 4.3.

Edit with WPS Office

Parameter	Value
Task	Instance Segmentation
Framework	YOLOv8
Image Size	640 × 640
Batch Size	16
Epochs	100
Models	YOLOv8n, YOLOv8s, YOLOv8m
Dataset	TrashCan Instance Segmentation

Table 4.5 :Experimental setup for model traning.

5.5 Training Procedure

The training procedure followed a sequential workflow to develop the intelligent vision module.

First, the underwater dataset was loaded from Google Drive into the Colab environment. The dataset was then extracted and automatically organized according to the directory structure required by the Ultralytics framework.

Next, a YAML configuration file describing the dataset structure and object categories was created. This file enabled the training framework to access the underwater images and their corresponding segmentation annotations.

Finally, the three YOLOv8 segmentation models were trained independently using the same dataset and training configuration. During each epoch, the network gradually learned the visual characteristics of underwater objects while continuously updating its internal parameters to minimize the prediction error.

5.6 Training Outputs

At the end of the training process, the YOLOv8 framework automatically generated several output files that summarize the learning process and store the trained model.

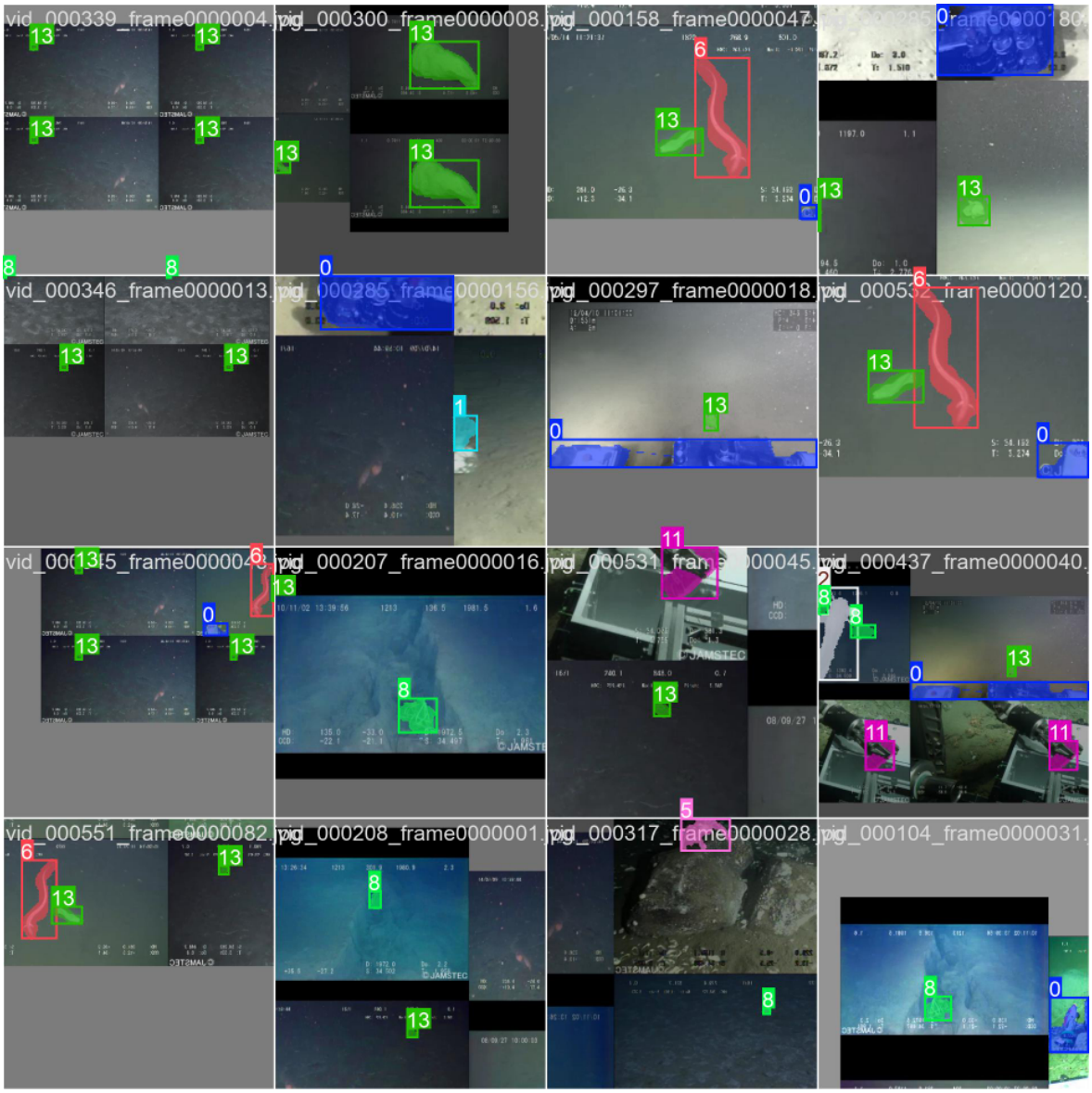
Among the generated outputs are:

- **best.pt**: the model with the highest validation performance.
- **last.pt**: the model obtained after the final training epoch.

Edit with WPS Office

- **results.png**: graphical representation of the learning curves.
- **runs/segment**: directory containing all training outputs.

The **best.pt** file represents the final Artificial Intelligence model selected for integration into the proposed underwater exploration system.



Edit with WPS Office

Figure 4.6 : training batch with instance segmentation annotations (batch 0).

6. Performance Evaluation of the AI Vision Module

6.1 Evaluation Metrics

The performance of the developed AI model was assessed using internationally recognized evaluation metrics commonly employed in object detection and instance segmentation tasks.

The selected metrics provide complementary information regarding the quality of the predictions generated by the intelligent vision system.

The principal evaluation metrics include:

- Precision
- Recall
- Mean Average Precision (mAP)
- Training Loss
- Precision–Recall Curve
- Confusion Matrix

Together, these metrics provide a complete overview of the model's capability to recognize underwater objects accurately and consistently.

6.2 Precision

Precision measures the proportion of detected underwater objects that are correctly classified by the Artificial Intelligence model.

A high Precision value indicates that most predictions generated by the intelligent vision system correspond to actual underwater objects rather than incorrect detections.

For the proposed underwater exploration platform, maintaining a high Precision is important because it minimizes false alarms and improves the reliability of the visual information presented to the ROV operator.

The Precision is calculated as follows:

$$Precision=\frac{TP}{TP+FP}$$

Edit with WPS Office

Where:

- TP = True Positives
- FP = False Positives

6.3 Recall

Recall evaluates the ability of the trained model to detect all underwater objects present in the analyzed scene.

Unlike Precision, Recall focuses on reducing missed detections. A high Recall value indicates that the intelligent vision module successfully identifies most underwater objects appearing in the camera image.

For underwater exploration, achieving a high Recall is essential because undetected marine organisms or underwater debris may lead to incomplete environmental observations.

The Recall is defined by:

$$Recall=\frac{TP}{TP+FN}$$

Where:

- TP = True Positives
- FN = False Negatives

6.4 Mean Average Precision (mAP)

Mean Average Precision (mAP) is considered the reference metric for evaluating object detection and instance segmentation models.

This metric combines Precision and Recall over different confidence thresholds to provide an overall estimation of the recognition capability of the trained model.

Within the proposed underwater exploration system, the mAP value was one of the principal criteria used to compare the three YOLOv8 segmentation models and to identify the architecture that offers the best compromise between segmentation accuracy and computational efficiency.

6.5 Analysis of the Training Curves

Edit with WPS Office

During the training process, the Ultralytics framework automatically generated several learning curves illustrating the evolution of the neural network.

These curves include:

- Box Loss
- Segmentation Loss
- Classification Loss
- Precision
- Recall
- mAP

The progressive decrease of the loss functions together with the continuous improvement of Precision and Recall indicates that the Artificial Intelligence model successfully learned the visual characteristics of the underwater dataset and progressively improved its prediction capability throughout the training process.

6.6 Precision–Recall Curve

The Precision–Recall Curve illustrates the relationship between Precision and Recall for different confidence thresholds.

A curve located near the upper-right corner indicates excellent detection performance, demonstrating that the trained model simultaneously achieves high detection accuracy while minimizing missed underwater objects.

This graphical representation provides additional confirmation that the developed Artificial Intelligence module is suitable for underwater exploration applications.

6.7 Confusion Matrix

The confusion matrix provides a detailed analysis of the predictions generated by the trained model.

Correct classifications appear along the principal diagonal, while incorrect predictions are represented outside the diagonal.

Analyzing this matrix allows the identification of underwater object categories that are easily recognized as well as those that remain more challenging because of visual similarity or underwater

Edit with WPS Office

environmental conditions.

This analysis contributes to understanding the strengths and limitations of the developed intelligent vision module.

6.8 Selection of the Final AI Model

Three YOLOv8 Instance Segmentation architectures were evaluated during the implementation phase:

- YOLOv8n-seg
- YOLOv8s-seg
- YOLOv8m-seg

The comparative analysis demonstrated that the Medium architecture achieved the highest segmentation performance while maintaining a computational complexity compatible with the requirements of the proposed underwater exploration system.

Consequently, **YOLOv8m-seg** was selected as the final Artificial Intelligence model integrated into the intelligent vision module of the developed ROV.

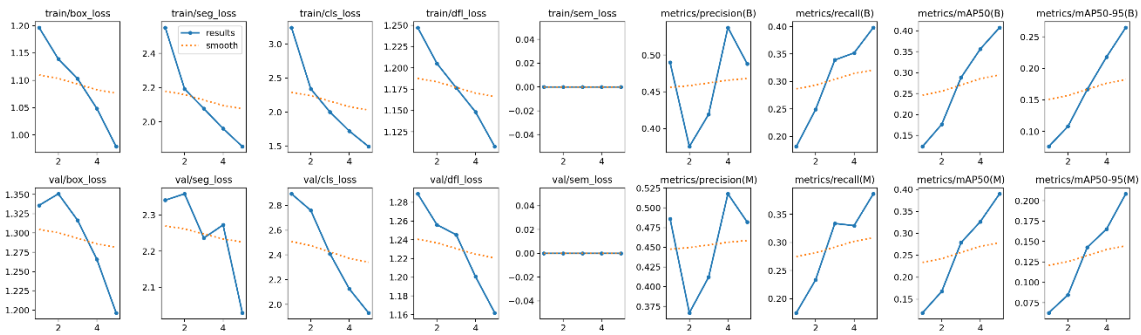


Figure 4.7 : YOLOv8 Nano instance segmentation results.

Edit with WPS Office

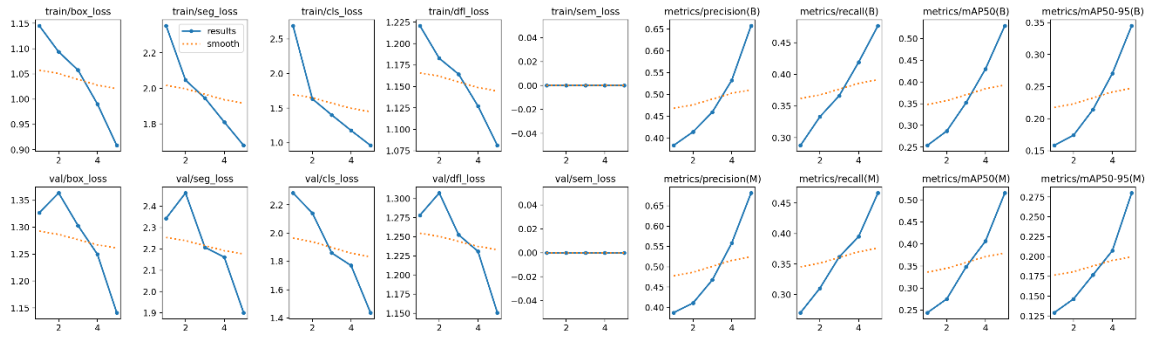


Figure 4.8 : YOLOv8 Small instance segmentation results after training on the underwater trash detection dataset.

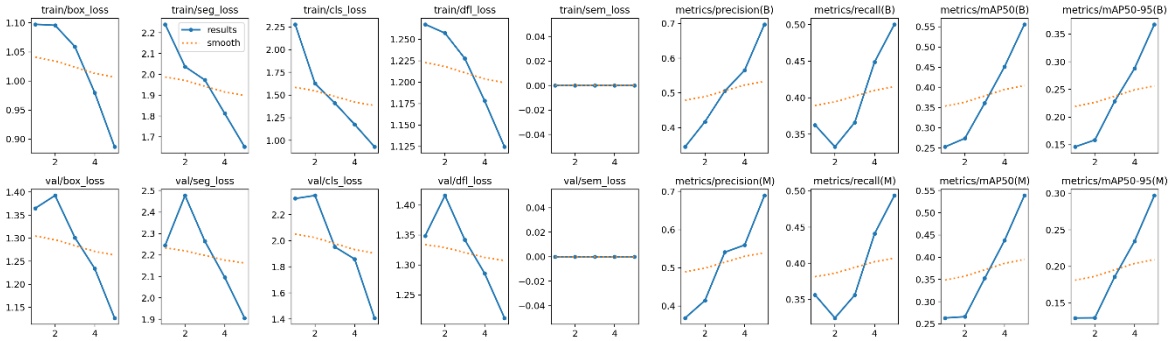


Figure 4.9 : YOLOv8 Medium instance segmentation results

7. Experimental Results and Discussion

7.1 Analysis of the Training Results

The training process demonstrated a continuous improvement in the learning capability of the developed Artificial Intelligence model.

Throughout the successive training epochs, the loss values gradually decreased while the evaluation metrics progressively improved. This behavior indicates that the neural network successfully learned the visual characteristics of underwater objects contained in the selected dataset.

Edit with WPS Office

The convergence of the training curves also confirms that the adopted transfer learning strategy was effective in adapting the pre-trained YOLOv8 model to underwater exploration scenarios without requiring excessive computational resources.

Overall, the obtained learning curves demonstrate that the developed AI module reached a stable training state and became capable of recognizing underwater objects with satisfactory performance.

7.2 Performance of Underwater Object Recognition

The prediction results demonstrate that the trained model successfully recognizes a wide variety of underwater objects belonging to different semantic categories.

During inference, the developed AI module identifies marine organisms, aquatic vegetation, underwater debris, and artificial structures while simultaneously generating precise segmentation masks describing the contours of each detected object.

Unlike traditional object detection systems that only generate rectangular bounding boxes, the adopted instance segmentation approach provides a more detailed interpretation of underwater scenes.

This additional visual information allows the ROV operator to better understand the surrounding environment and facilitates the identification of underwater objects during exploration missions.

7.3 Discussion of the Obtained Results

The experimental evaluation confirms that the proposed Artificial Intelligence module satisfies the main objective of the project, namely assisting underwater exploration through automatic visual scene interpretation.

The obtained results indicate that the model is capable of detecting multiple underwater objects simultaneously while maintaining accurate segmentation boundaries.

This capability significantly improves the amount of visual information available to the operator compared with conventional underwater video observation.

However, several challenges remain inherent to underwater environments.

Reduced visibility, suspended particles, low illumination, and the presence of partially occluded objects may occasionally decrease the prediction confidence of the model.

Similarly, extremely small objects or objects located at considerable distances from the camera

Edit with WPS Office

may be more difficult to segment accurately.

Despite these challenges, the obtained performance remains sufficiently reliable for practical underwater exploration applications.

7.4 Contribution of Artificial Intelligence to the Proposed ROV

The integration of Artificial Intelligence considerably enhances the functionality of the developed underwater exploration system.

Instead of relying exclusively on human observation, the onboard vision module continuously analyzes the underwater video stream and automatically identifies the visible objects.

This intelligent assistance enables the operator to:

- Identify marine organisms more rapidly.
- Detect underwater debris during exploration.
- Better understand complex underwater scenes.
- Improve the efficiency of underwater observation missions.
- Reduce the operator's visual workload during long exploration sessions.

Consequently, the developed AI module transforms the proposed ROV into an intelligent underwater exploration platform capable of providing semantic information in addition to conventional video transmission.

7.5 Current Limitations

Although the developed system achieved satisfactory performance, several limitations remain.

The Artificial Intelligence model was trained using a publicly available underwater dataset rather than images collected directly by the developed ROV.

Consequently, the visual characteristics of the operational environment may differ from those represented in the training dataset.

Furthermore, the current implementation focuses on underwater object recognition and environmental observation. The model was not specifically trained to recognize drowning victims because an adequately annotated underwater dataset for this application was not available during the development of the project.

Edit with WPS Office

Nevertheless, the proposed architecture remains sufficiently flexible to incorporate additional datasets in future work, enabling the development of specialized underwater recognition applications.

7.6 Future Improvements

Several improvements can further enhance the proposed underwater exploration platform.

Future work may include:

- Collecting a dedicated underwater dataset using the developed ROV.
- Training the AI model with a larger number of underwater images.
- Integrating additional underwater object categories.
- Optimizing the model for embedded execution on Raspberry Pi.
- Extending the system toward autonomous underwater navigation.
- Incorporating specialized datasets for underwater search and rescue applications.

These improvements would increase the robustness, versatility, and practical applicability of the proposed intelligent underwater exploration system.

8. Development of the Project Presentation Website

8.1 Objectives of the Website

The development of the project website aims to complement the technical implementation of the intelligent underwater exploration system by providing an online platform capable of presenting the project to a wider audience.

The main objectives of the website are:

- Present the concept of the intelligent underwater exploration system.
- Introduce the objectives and motivations of the project.
- Highlight the integration of Artificial Intelligence into underwater exploration.
- Present the hardware and software technologies used during development.
- Explain the main functionalities of the proposed ROV.
- Promote the project through a modern and responsive web interface.

By achieving these objectives, the website contributes to increasing the visibility of the project and facilitates communication with researchers, students, and anyone interested in underwater robotics.

Edit with WPS Office

8.2 Website Structure

To ensure intuitive navigation and an effective presentation of the project, the website was organized into several sections, each dedicated to a specific aspect of the underwater exploration system.

The overall structure is summarized in Table 4.7.

Section	Description
Home	General introduction to the project
About	Presentation of the underwater exploration system
Features	Main functionalities of the intelligent ROV
Technologies	Hardware and software technologies used
Artificial Intelligence	Description of the AI vision module
Contact	Project information and communication section

Table 4.6:Structure of the project website.

8.3 Website Design

The website was developed following modern web design principles in order to provide a pleasant browsing experience and facilitate access to project information.

Particular attention was given to:

- Simplicity of navigation.
- Clear organization of information.
- Responsive layout compatible with different screen sizes.
- Attractive visual presentation.
- Consistent color palette inspired by the underwater environment.

These design choices improve readability and enhance the overall presentation of the intelligent underwater exploration system.

8.4 Contribution of the Website to the Project

Edit with WPS Office

Although the website is not involved in controlling the underwater vehicle, it represents an important complementary component of the proposed project.

It allows visitors to understand the objectives of the intelligent underwater exploration system without requiring technical expertise in Artificial Intelligence or underwater robotics.

Furthermore, the website provides a professional digital identity for the project and demonstrates its practical implementation through images, technical descriptions, and project presentations.

Consequently, the website contributes to disseminating the project and promoting its technological innovation.

9.Conclusion

This chapter presented the complete implementation process of the proposed AI-based underwater exploration system. Beginning with the preparation of a representative underwater dataset, the development focused on building an intelligent vision module capable of recognizing and segmenting underwater objects using the YOLOv8 Instance Segmentation architecture.

A transfer learning strategy was adopted to efficiently adapt pre-trained models to underwater exploration scenarios. The trained models were evaluated using standard performance metrics, leading to the selection of the YOLOv8m-seg model as the most suitable solution for the proposed ROV due to its balance between segmentation accuracy and computational efficiency.

The experimental analysis demonstrated that the developed AI module can effectively support underwater exploration by providing semantic information about the surrounding environment, thereby assisting the operator during exploration missions. In addition, a dedicated project presentation website was developed to showcase the system, its objectives, and the technologies employed throughout the implementation process.

Overall, this chapter demonstrates that combining underwater robotics with Artificial Intelligence significantly enhances the capabilities of conventional remotely operated vehicles, transforming them into intelligent platforms capable of supporting underwater environmental exploration.

Edit with WPS Office

Edit with WPS Office

GENERAL CONCLUSION

This thesis has presented the design and development of a low-cost intelligent Remotely Operated Vehicle (ROV) prototype dedicated to underwater object search. The proposed project combines embedded systems, remote control, computer vision, and Artificial Intelligence in order to assist the user during underwater search operations. The main objective was not to build a professional commercial submarine, but to propose a realistic proof of concept showing how accessible technologies can be used to support underwater object detection tasks.

The project addressed several challenges related to underwater environments, such as reduced visibility, light attenuation, turbidity, water pressure, and the difficulty of direct human intervention. By adopting an ROV-based architecture, the system allows the operator to control the vehicle remotely while receiving visual information from an onboard camera. The integration of a YOLO-based object detection approach represents an important step toward making the system more intelligent and capable of identifying objects of interest according to the trained model and the available dataset.

From the hardware side, the prototype relies on accessible and low-cost components. The Raspberry Pi 5 was selected as the central processing and control unit of the system. It is responsible for handling the camera, managing video acquisition, controlling the motors, supporting Wi-Fi-based communication, and running the AI detection module. This centralized architecture simplifies the design of the prototype and allows the main functions of the ROV to be managed by a single powerful embedded board.

The study also highlighted the importance of choosing an appropriate AI model for the available hardware. Although larger YOLO models may provide better accuracy, they also require more processing power. For this reason, the selection of the detection model must consider the balance between accuracy, speed, and the processing capabilities of the Raspberry Pi 5. The final performance of the detection module remains dependent on the quality of the underwater dataset, the annotation process, the object classes used during training, and the environmental conditions during testing.

Despite the promising outcomes, the project acknowledges several limitations. Underwater environments remain difficult for computer vision systems because image quality can be affected by low light, turbidity, reflections, and unstable camera movement. In addition,

Edit with WPS Office

communication performance may vary depending on the distance, depth, water conditions, and experimental setup. Moreover, the dataset and final AI model are still considered part of the development process and may require further improvement before obtaining reliable detection results in different underwater conditions.

In conclusion, this work demonstrates the potential of low-cost AI-assisted ROV systems in supporting underwater object search tasks. It provides a solid foundation for future improvements and opens the way for developing more advanced underwater search solutions. Future work will aim to improve the mechanical design, enhance the waterproof housing, optimize motor control, improve video transmission, increase the reliability of the communication system, optimize YOLO execution on the Raspberry Pi 5, and train the system on richer and more specialized underwater datasets.

Edit with WPS Office

Bibliography

[1]:National Oceanic and Atmospheric Administration. (2024). How does pressure change with ocean depth? National Ocean Service.

https://oceanservice.noaa.gov/facts/pressure.html

[2]:National Oceanic and Atmospheric Administration. (2024). How far does light travel in the ocean? National Ocean Service.

https://oceanservice.noaa.gov/facts/light_travel.html

[3]:Hao, Y., Yuan, Y., Zhang, H., & Zhang, Z. (2024). Underwater optical imaging: Methods, applications and perspectives. Remote Sensing, 16(20), 3773.
https://doi.org/10.3390/rs16203773

[4]:National Oceanic and Atmospheric Administration. (2024).

What is a thermocline? National Ocean Service.

https://oceanservice.noaa.gov/facts/thermocline.html

[5]:U.S. Geological Survey. (n.d.). Why is the ocean salty?

https://www.usgs.gov/faqs/why-ocean-salty

[6]:Feely, R. A., Jiang, L.-Q., Wanninkhof, R., Carter, B. R., Alin, S. R., Bednaršek, N., & Cosca, C. E. (2023).

Acidification of the global surface ocean: What we have learned from observations. Oceanography, 36(2–3), 120–129.

https://doi.org/10.5670/oceanog.2023.222

[7]:Sogin, E. M., Leisch, N., & Dubilier, N. (2020).

Chemosynthetic symbioses. Current Biology, 30(19), R1137–R1142.

https://doi.org/10.1016/j.cub.2020.07.050

[8]:Gu, Z., Liu, X., Hu, Z., Wang, G., Zheng, B., Watson, J., & Zheng, H. (2023).

Underwater computational imaging: A survey. Intelligent Marine Technology and Systems, 1, Article 2.

https://doi.org/10.1007/s44295-023-00009-w

[9]:Shuang, X., Zhang, J., & Tian, Y. (2024).

Algorithms for improving the quality of underwater optical images: A comprehensive review. Signal Processing, 219, 109408.

https://doi.org/10.1016/j.sigpro.2024.109408

[10]:Vuong, P., McKinley, A., & Kaur, P. (2023).

Edit with WPS Office

Understanding biofouling and contaminant accretion on submerged marine structures. npj Materials Degradation, 7, Article 50.

https://doi.org/10.1038/s41529-023-00370-5

[11]:National Oceanic and Atmospheric Administration. (n.d.). Remotely Operated Vehicles (ROVs). NOAA Ocean Exploration.

https://oceanexplorer.noaa.gov/technology/subs-rovs/

[12]:National Oceanic and Atmospheric Administration. (n.d.). What is an ROV? NOAA Ocean Exploration.

https://oceanexplorer.noaa.gov/ocean-fact/rov/

[13]:Ghatak, A., Mukherjee, K., & Singh, Y. (2023). Classification on unmanned underwater vehicles: A review. International Journal of Vehicle Autonomous Systems, 17(3/4), 192–220.

https://doi.org/10.1504/IJVAS.2023.140519

[14]:Woods Hole Oceanographic Institution (WHOI), "ROV Jason – National Deep Submergence Facility,"

https://ndsf.whoi.edu/jason/. [Accessed: Apr. 3, 2026]

[15]:Schmidt Ocean Institute, "4500 m Remotely Operated Vehicle (ROV SuBastian),"

https://schmidtocean.org/technology/robotic-platforms/4500-m-remotely-operated-vehicle-rov/.

[16]:Blue Robotics, "BlueROV2 – Affordable and Capable Underwater ROV,"

https://bluerobotics.com/store/rov/bluerov2/.

[17]:VideoRay.(n.d.). Mission Specialist Defender ROV.https://videoray.com/products/mission-specialist-defender/

[18]:https://cloud.google.com/learn/what-is-artificial-intelligence?hl=fr

[19]:https://www.appliedaicourse.com/blog/domains-of-artificial-intelligence/

[20]:https://medium.com/@vkisho25/domains-in-artificial-intelligence-aba424350624

[21]:https://www.longdom.org/open-access/problem-solving-and-artificial-intelligence-revolutionizing-decisionmaking-10

[22]:IBM,https://www.ibm.com/topics/machine-learning

[23]:Google Cloud, "Machine Learning Workflow",https://cloud.google.com/learn/what-is-machine-learning

[24]: Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning

[25]:Wikipedia, "Convolution",https://en.wikipedia.org/wiki/Convolution

[26]:Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning.

Edit with WPS Office

[27]:Google Cloud, "Image Classification",https://cloud.google.com/vision/docs

[28]:Wikipedia, "Image Segmentation",https://en.wikipedia.org/wiki/Image_segmentation

[29]:Wikipedia, "Object Detection",https://en.wikipedia.org/wiki/Object_detection

[30]:Towards Data Science, "R-CNN Explained",https://towardsdatascience.com

[31]:Redmon YOLO research paper (YOLO Paper),https://pjreddie.com/media/files/papers/yolo.pdf

[32]:Joseph Redmon et al.,"You Only Look Once: Unified, Real-Time Object Detection",https://pjreddie.com/media/files/papers/yolo.pdf

[33]:IBM, "YOLO Object Detection",https://www.ibm.com/topics/yolo

[34]:MDPI, "Underwater Image Enhancement Review",https://www.mdpi.com

[35]:IEEE, "Underwater Image Processing",https://ieeexplore.ieee.org

[36]:Springer, "Underwater Vision Challenges",https://link.springer.com

[37]:ScienceDirect, "Underwater Image Degradation",https://www.sciencedirect.com

Edit with WPS Office

الجمهورية الجزائرية الديمقراطية الشعبية

Republique Algérienne Démocratique Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur de la Recherche Scientifique

Université 20 août 1955 Skikda

جامعة 20 أوت 1955 – سكيكدة

Faculté des Sciences
Département d'Informatique

كلية العلوم
قسم الإعلام الآلي
الرقم: .... / ق/ ا/ ا/ ك.ع2026/

Autorisation de Dépôt de Mémoire de Master

Je soussigné : Bouhouamdia Soufiane

Certifie que l'étudiant(e) : Saci Nadjib

Spécialité : AI

Ayant soutenu le projet intitulé : Intelligent Search Submarine Using Artificial Intelligence

A apporté les corrections nécessaires sur son manuscrit de Master.

Signature de l'encadreur

Scanned with CamScanner

Edit with WPS Office