Democratic and People's Republic of Algeria
Ministry of Higher Education and scientific research
Larbi Tébessi University - Tébessa
Faculty of Exact Sciences and Natural and life Sciences
Department: Mathematics and Computer Science

MASTER THESIS

Domain: Mathematics and Computer Science

Sciences field: Computer science

Option : Information System

Smart Waste Underwater Segmentation: Deep Learning Based Approach

Defended by:

BENDIB Mohamed Dhia

Supervised by:

Dr. BOUROUGAA Salima

Supported publicly in front of the jury:

SLIMI Hamda	**MCB**	**Larbi Tébessi University**	**President**
DHIFALLAH Salwa	**MAA**	**Larbi Tébessi University**	**Examiner**

College Year :

2023/2024

Appreciation

First of all, we thank ALLAH for giving us the strength and courage we needed to carry out this work.

We would like to express our gratitude to our supervisor Dr. Bourougaa Salima for supporting our choices, for advising us, guiding us and helping us throughout this project.

We thank Dr. HAOUEM Yacine and Dr. BENDIB Issam who helped us during the entire period of preparation of our memory and who contributed to our training.

And finally, a big thank you to our families and friends who encouraged, supported and followed us throughout this project.

Dedications

I dedicate this work to my dear and precious parents to have encouraged me, to help during all my years of schooling and to have supported my choices. to Nizar and Abdelhak who encouraged me, to my dear brother and sister. And to my whole family

Abstract

Waste sorting is a crucial process for efficient waste management, with automation being a significant factor for waste companies. Advancements in robotics, artificial intelligence, and autonomous driving have enabled underwater waste cleaning robots to accurately locate and recognize underwater waste. The image segmentation method, compared to deep-learning-based target detection, provides a more refined and accurate approach to target recognition, improving environmental perception and efficiency in underwater waste sorting. To simplify the process, we proposed two trained models (YOLOv8and Mask RCNN). We used them to segment underwater waste and then facilitate the process of sorting the waste into different types such as mask, glass bottles, plastic bottles, electronics, metal, tire, plastic bags and waste. The proposed system is tested on Underwater_debris dataset. Our YOLOv8 model achieved 84% of average precision after 200 training epochs and an average precison that exceeds 83.3% for our Mask RCNN model after 100 training epochs.

The underwater waste separation process is made quicker, efficient, and more intelligent by our proposed waste segmentation approach.

Keywords: underwater waste sorting, YOLO, Mask RCNN, Deep learning, segmentation.

Résumé

Le tri des déchets est un processus crucial pour une gestion efficace des déchets, l'automatisation étant un facteur important pour les entreprises de gestion des déchets. Les progrès de la robotique, de l'intelligence artificielle et de la conduite autonome ont permis aux robots de nettoyage des déchets sous-marins de localiser et de reconnaître avec précision les déchets sous-marins. La méthode de segmentation d'image, par rapport à la détection de cible basée sur l'apprentissage profond, fournit une approche plus raffinée et plus précise pour la reconnaissance de cible, améliorant la perception environnementale et l'efficacité dans le tri des déchets sous-marins. Pour simplifier le processus, nous avons proposé deux modèles entraînés (YOLOv8 et Mask RCNN). Nous les avons utilisés pour segmenter les déchets sous-marins puis faciliter le processus de tri des déchets en différents types tels que masque, bouteilles en verre, bouteilles en plastique, électronique, métal, pneu, sacs en plastique et déchets. Le système proposé est testé sur l'ensemble de données Underwater_debris. Notre modèle YOLOv8 a atteint 84% de précision moyenne après 200 périodes d'entraînement et une précision moyenne qui dépasse 83,3% pour notre modèle Mask RCNN après 100 périodes d'entraînement.

Le processus de séparation des déchets sous-marins est rendu plus rapide, efficace et plus intelligent par notre approche de segmentation des déchets proposée.

Mots-clés : tri des déchets sous-marins, YOLO, masque RCNN, apprentissage approfondi, segmentation.

ملخص

يعد فرز النفايات عملية حاسمة لإدارة النفايات بكفاءة، مع كون الأتمتة عاملاً مهمًا لشركات النفايات. مكنت التطورات في مجال الروبوتات والذكاء الاصطناعي والقيادة الذاتية روبوتات تنظيف النفايات تحت الماء من تحديد موقع النفايات تحت الماء والتعرف عليها بدقة. توفر طريقة تقسيم الصور، مقارنة باكتشاف الأهداف القائمة على التعلم العميق، نهجًا أكثر دقة ودقة للتعرف على الأهداف، وتحسين الإدراك البيئي والكفاءة في فرز النفايات تحت الماء. لتبسيط العملية، اقترحنا نموذجين مدربين (YOLOv8 و Mask RCNN). استخدمناها لتقسيم النفايات تحت الماء ثم تسهيل عملية فرز النفايات إلى أنواع مختلفة مثل الأقنعة والزجاجات والزجاجات البلاستيكية والإلكترونيات والمعادن والإطارات والأكياس البلاستيكية والنفايات. يتم اختبار النظام المقترح على مجموعة البيانات Underwater_debris. حقق طرازنا YOLOv8 84٪ من متوسط الدقة بعد 200 حقبة تدريبية ومتوسط مسبق يتجاوز 83.3٪ لنموذج Mask RCNN بعد 100 حقبة تدريبية.

أصبحت عملية فصل النفايات تحت الماء أسرع وفعالة وأكثر ذكاءً من خلال نهج تقسيم النفايات المقترح.

الكلمات الرئيسية: فرز النفايات تحت الماء، YOLO، Mask RCNN، التعلم العميق، التجزئة.

List of content

List of figures

List of tables

Glossary

EPA	Environmental Protection Agency
MSW	municipal solid waste
EU	European Union
AUVs	autonomous underwater vehicles
ROVs	remotely controlled vehicles
AI	Artificial intelligence
ML	Machine Learning
DL	Deep Learning
DNNs	deep neural networks
LSTM	long short-term memory
GAN	Generative Adversarial Neural Network
CNN	Convolutional Neural Network
FC	fully connected
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
CV	computer vision
RPN	Region Proposal Network
ROI	region of interest
RCNN	Region-based Convolutional Neural Network
YOLO	You Only Look Once
SSD	Single Shot MultiBox Detector
U-NET	U-shaped Network
MAP	Mean Average Precision
GPU	Graphics Processing Unit
VIA	VGG Image Annotator
JSON	JavaScript Object Notation
FPN	Feature Pyramid Network
TP	TRUE POSITIVE
TN	TRUE NEGATIVE
FP	FALSE POSITIVE
FN	FALSE NEGATIVE
P	PRECISION
R	RECALL

GENERAL INTRODUCTION

General Introduction

Globally, it is expected that annual solid waste will exceed 2.2 billion tons by 2025, resulting in waste management costs of $375.5 billion. Poor waste management can devastate the economy, public health, and the environment. The Environmental Protection Agency (EPA) has identified recycling of municipal solid waste (MSW) as the second most environmentally friendly method for managing urban waste [1].

Recycling waste optimally benefits both the economy and the environment. It allows for the recovery of raw resources, saves energy, reduces greenhouse gas emissions, decreases water pollution, and prevents the creation of new landfills [1]

The process of separating waste is called sorting, which aims to facilitate waste recycling. When waste is sorted, the amount of waste sent to landfills decreases significantly, resulting in lower levels of air and water pollution. It is important to remember that waste sorting should be done based on the type of waste and the most appropriate treatment and disposal methods.[1]

The use of multiple waste sorting methods, including the integration of artificial intelligence, is facilitated by waste sorting. Specifically, waste segmentation based on deep learning (DL) is necessary to ensure an optimal and rapid strategy for waste reduction and the valorization of recyclable waste.[1]

Manual waste separation has a high cost. Recent advances in deep learning have led to extraordinary breakthroughs in the field of computer vision. The convolutional neural network (CNN) is one of the most well-known deep learning techniques and is proposed in this literature to perform waste segmentation.[1]

In our project, we aim to segment waste to subsequently classify it into several categories: mask, metal, glass bottles, plastic bottles, plastic bags, electronics, tires, and waste. By using deep learning methods, when waste is deposited in underwater environments, we have used **underwater_debris** dataset and then use **Mask RCNN** and **YOLOv8** to segment, recognize, and classify the waste. Our models reached a maP of 83.3% for Mask RCNN and 84% for Yolov8.

The structure of the thesis:

The structure of the thesis is organized as follows:

▪ **Chapter I**: We begin our thesis with the first chapter, which introduces the concept of underwater waste management.

▪ **Chapter II**: The second chapter introduces the basic concepts of deep learning, and various computer vision tasks, and concludes with an overview of the most significant work based on deep learning.

▪ **Chapter III**: In the final chapter, we present the design and implementation of our deep learning-based approach for smart underwater waste segmentation.

Chapter 1: Underwater Waste Management

I. Chapter 1: Underwater Waste Management

1. Introduction

Waste things we throw out into the environment plays a big part in smart cities and other intelligent environments. Underwater waste management and recycling are key elements in the fight against climate change. However, keeping several containers and handling waste properly can be difficult if you don't have enough time, money, or understanding. Thus, we need to think of smart bins that allow for waste sorting and recycling if we want to protect the environment, fight pollution, and save time, money, and knowledge. The current solid waste management policies and methods are becoming more challenging due to the exponential increase in urban population.

The issue is more complicated in cities that are expanding quickly since there is a rise in the amount of waste produced and managing it with the current infrastructure is difficult. Reducing the amount of solid waste dumped in landfills requires fundamental solutions including prevention, recycling, reuse, and recovery of waste, particularly in rapidly expanding cities where more sustainable management practices must be implemented.

Waste must be sorted by classification into numerous groups. To manage underwater waste effectively, sorting needs to be done well. As soon as it is practical, separation should be carried out to lessen the chance that other materials will contaminate the underwater waste. This can be achieved and waste management companies' workload reduced by automating the process.

This chapter will introduce underwater waste management through a number of topics, including wastes definition, hierarchy, and history, underwater waste management, sources and risks of underwater wastes, and finishing by presenting underwater wastes sorting, reducing and valorization

2. Waste Definition:

Most nations have defined waste, which is typically associated with the idea of disposal.

1) "Waste" are objects or substances that are disposed of, intended to be disposed of, or required to be disposed of by national legislation, according to Article 5 of the Basel Convention [3].
2) Council on the Management of Transboundary Waste Movements Designed for Recovery Operations Decision: Wastes are substances or objects, excluding radioactive materials covered by other international agreements, that are either being recovered or disposed of, ii) planned to be recovered or disposed of, or iii) required to be recovered or disposed of by national-law requirements.[4]
3) Article 3 of the Waste Framework Directive of the European Union defines "waste" as "any substance or object which the holder discards, intends to discard, or is required to discard."[5]

3. History of Waste

The development of waste management has always been inextricably linked to its impact on the environment and human health, both positively and negatively. With recycling and other advancements, the waste management business of today is positioned to make much more progress [2].

Here is a chronology of some noteworthy advancements in the history of waste management [2]:

- **Ancient Times**

3,000 B.C. | The first landfill was established in Knossos, Crete, Greece, where waste was dumped into big pits excavated in the ground.

2,000 B.C. | In China, easy techniques of composting and recycling are created and exploited, notably for bronze to be used later.

500 B.C. | To maintain the city's aesthetic appeal and avoid disease, Athens, Greece, passed legislation mandating that waste be disposed of at least a mile outside of the city.

- **Middle Ages**

1388 | Waste disposal in ditches and public rivers is outlawed by the English Parliament.

1551 | Packaging was first used in written form by German papermaker Andreas Bernhart, who starts putting his paper in covers marked with his name and address.

1657 | The first anti-littering rule is passed in New Amsterdam, which is now Manhattan in New York City. It forbids throwing or leaving waste on the streets.

- **Industrial Revolution**

I. Chapter 1: Underwater Waste Management

1757 | Ben Franklin launches the first street cleaning business and encourages people to dispose of their rubbish by excavating pits in the ground.

1875 | The English Parliament passes the Public Health Act of 1875, granting authorities for waste collection and prohibiting mass scavenging. Additionally, the initial idea for a transportable container is developed.

1885 | America constructs its first incinerator at New York City's Governors Island.

- 20th Century

1914 | In the US and Canada, there are currently close to 300 incinerators in use. Landfills start to expand.

1921 | With the development of the back loader waste truck, waste pickup is now more efficient.

1934 | Municipal waste cannot be dumped into the ocean, according to the US Supreme Court.

1945 | Most places forbid burning waste in backyards or open-air dumping sites.

1947 | Because plastics are so inexpensive and convenient, consumerism reaches an all-time high, leading to a 67% rise in packaging waste.

1963 | It passes the Clean Air Act.

1965 | Government research on resource recovery and landfill research is authorized by the Solid Waste Disposal Act.

1968 | The percentage of recyclable and nonrecyclable municipal waste in the United States is over 33%. Recycling initiatives like newspaper curb recycling and pay-per-can are implemented in many states.

1970 | Established as an autonomous government organization, the Environmental Protection Agency (EPA) was created to oversee the security and safety of the country's natural environment.

1976 | Waste management, recycling, and conservation are all planned for under the Resource Conservation and Recovery Act. Recycling is governed by laws in 26 states.

1991 | In all 50 states, there are more than 3,000 household hazardous waste programs in place.

1994 | Purchasing and utilizing recycled goods is mandated for federal entities. The directive's enforcement is handled by the Office of Federal Environmental Executive.

- **21st Century**

In the 2000s | Pay-as-you-throw programs, which charge people depending on the amount of rubbish each family or building tossed away, were implemented by more than 5,000 U.S. localities.

The 2010s | Modern waste trucks can now pack and transport up to three times as much waste as their predecessors thanks to technological advancements.

Today | Despite making up less than 5% of the world's population, Americans produce over 250 million tons of waste annually, accounting for roughly 20% of the global waste problem.

4. Waste Management

Reducing waste's detrimental impacts on the environment, public health, planetary resources, and aesthetics is the goal of waste management, from the point of creation until disposal. It covers both direct and indirect health problems brought on by human activity, such as the intake of food, water, and soil.[6]

The goal of waste management is to reduce the negative effects that municipal solid waste—produced by commercial, industrial, and residential activities—has on the environment and public health.[7] According to a survey, between 20% and 50% of municipal revenues are usually devoted to efficient waste management, which calls for the development of integrated systems that are resilient, long-lasting, and socially useful.[8]

The majority of waste management strategies focus on managing municipal solid waste, which is the main type of waste generated by households, companies, and enterprises.[9] The IPCC projects that by 2050, the world's output of municipal solid waste will have surpassed 3.4 gigatons; but, by changing laws and policies, waste production can be decreased in different cities and areas. [10]

Integrated import/export techno-economic systems, efficient disposal locations, circular economy management, and environmentally friendly product design are some examples of waste management techniques.[11] The "7R" approach—refuse, reduce, reuse, repair, repurpose, recycle, and recover is essential to efficient waste management. The first two emphasize cutting back on consumption and steering clear of pointless purchases in order to prevent waste. "Recover," which entails energy recovery, is the least effective tactic.[12]

5. Waste Hierarchy

In order to reduce end-of-product waste, the waste hierarchy is a waste management system that emphasizes product reduction, reuse, and recycling. It is arranged in a pyramidal fashion to promote legislation that forbids the creation of waste. Three stages make up the hierarchy: material recovery, waste-to-energy, recycling, and prevention. The last phase is disposal, which can be either landfilling or burning without energy recovery. The last phases of a product's life cycle are represented by the hierarchy.[13]

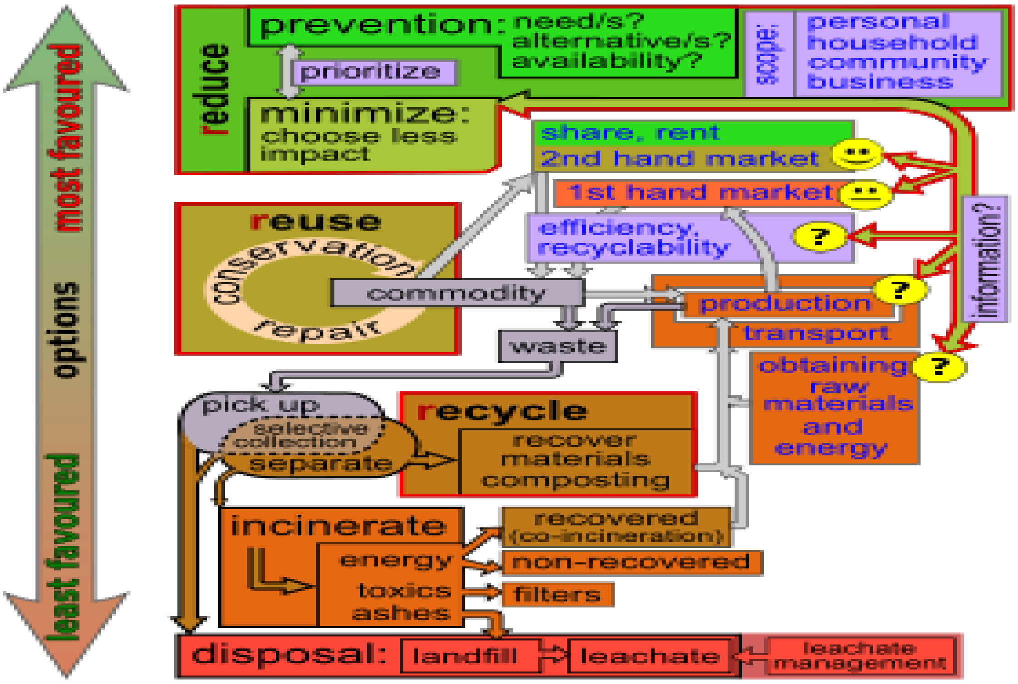

Figure 1: waste hierarchy [13]

5.1. Waste Life Cycle

The life cycle begins with design, proceeds through primary use, distribution, and manufacturing, and concludes with the waste hierarchy's reduce, reuse, and recycle stages. At each phase of the product life cycle, there is an opportunity for policy action, including reevaluating the product's need, revamping it to cut down on waste, and increasing its application. With the help of product life-cycle analysis, waste may be avoided and the planet's limited resources can be used to their fullest potential.[14]

5.2. Resource Efficiency

Resource efficiency reflects the understanding that current patterns of production and consumption cannot sustain ongoing global economic growth and development. More resources are used by humanity than the planet can sustainably produce on a global basis. Resource efficiency is the process of minimizing the environmental impact of creating and using these products, from the extraction of the ultimate raw material to their eventual use and disposal.[14]

6. Waste Reduction and Sorting

The environment and general public’s health are seriously threatened by the careless or ineffective disposal of waste. Hazardous waste is currently disposed of via surface impoundments or landfills, which exposes the bodies of people and animals and puts the environment and public health at serious danger. People are becoming more and more aware of how tough it may be to manage and get rid of rubbish.[15]

6.1. Reduce

The most important waste control tactic is to focus on needs and make fewer purchases overall. We are able to concentrate on waste reduction rather than waste disposal because this technique reduces the requirement for raw materials, production, shipping materials, and resource allocation. This strategy reduces the requirement for waste management.[15]

6.1.1. Reduce: What Can We Do?

People are a contributing factor to waste management problems since they waste money and resources and frequently replace technology with new models. Water bottles, reusable cups, and silverware are examples of high-quality products that are more affordable and long-lasting than disposable ones. It's also very important to reduce the amount of packaging, including bags, boxes, packing peanuts, and plastic wrappers. Bring your own shopping bags to avoid using plastic ones, as paper bags degrade more quickly than plastic ones. Reusable bags are frequently offered for sale at the register, and some businesses even give out plastic bags for patron use. People can significantly lessen their environmental effect by evaluating their waste management practices and cutting back on waste.[15]

6.2. Reuse

Most of us keep wasting because we can`t come up with a better way to use items like the phone book from the previous year, short curtains, or a Reusing begins with the idea that the materials we consume daily can be repurposed as resources instead of being discarded. We can learn to view the items we discard as resources that can be used again to meet requirements and find solutions to daily issues. However, many of us haven't even started to utilize the resources found in our waste. Once you've decided to recycle waste, you may start coming up with ideas and organizing a brainstorm.[15]

6.2.1. Reuse: What Can We Do?

Reusing materials for different purposes is a sustainable approach known as reuse. For instance, containers can be used again at home or for school projects. In a similar vein, extra clothing might be given to friends or a worthy cause. Appliances that break can be given to charitable organizations or trade schools for use in art programs or repairs. Old towels and sheets can be repurposed as dust cloths, and single-sided paper can be utilized for writing notes. Donations of books and periodicals can be made to schools, nursing homes, and public libraries. You may repurpose old tires for play yards and gardens. Use tote bags rather than plastic bags and make sure they are cleaned after every use to help reduce waste.[15]

6.3. Recycle

Recycling has several advantages, such as boosting the economy, creating jobs, and cutting expenses. Compared to landfilling, which is expected to yield six jobs, recycling 10,000 tons of debris would generate 36 jobs, according to the EPA. Additionally, communities have operated job-training partnerships, worked with disabled workshop operators, and hired jobless individuals through recycling initiatives. The industry has been hampered by the idea of profitability, though, since some recyclables might not be advantageous for particular uses.[15]

6.3.1. Collecting Recyclables

Paper, newspapers, corrugated cardboard, aluminum, steel cans, glass, plastic, motor oil, natural waste, and scrap metal are just a few of the many materials made from human waste that can be recycled. Families should separate their recyclables from their garbage, and used goods can be disposed of at recycling drop-off locations. Additionally, buyback facilities buy different metals and commodities from waste.[15]

7. Waste Management by Region

The complex issue of waste management affects both mankind and the environment. According to World Bank projections, there will be a 70% rise in the amount of waste produced globally by 2050, making managing waste a more complicated problem. The main factor contributing to the 2 billion metric tons of solid waste produced each year is population growth. By 2050, the amount of waste produced worldwide is predicted to increase from 3.4 billion tons to over 3 billion tons. Municipal solid waste is the third-largest generator of methane in the world. Figure 2 indicates a positive correlation between the amount of waste produced in the area and the level of wealth. [16]

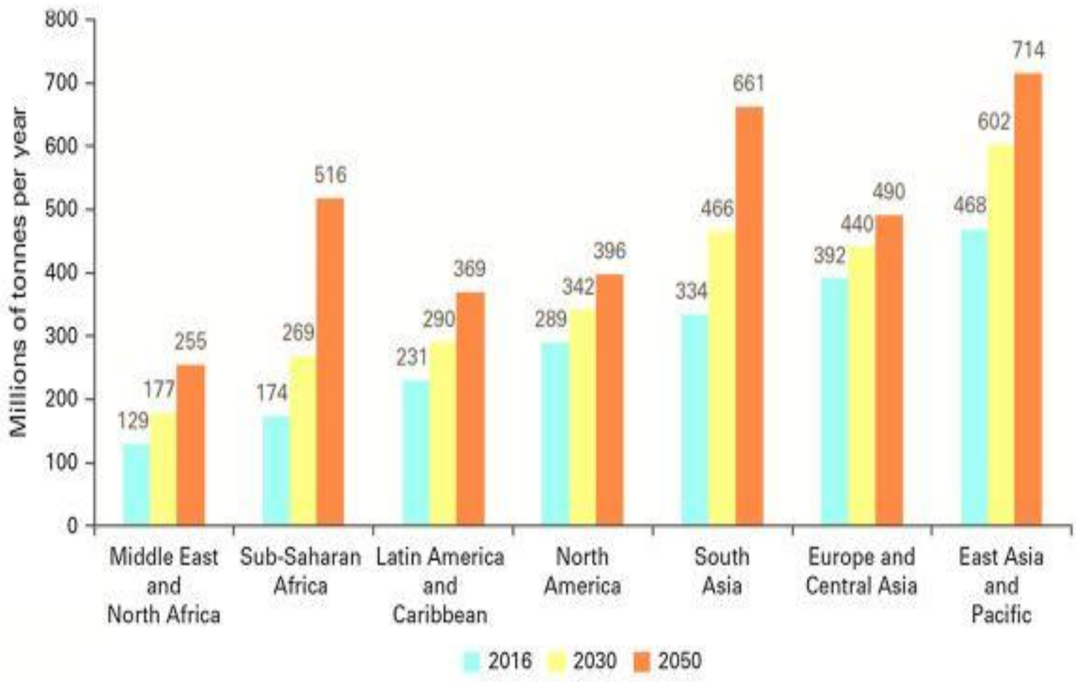

Figure 2:Waste generation in different regions [16]

8. Smart Waste Management

The goal of smart waste management is to use intelligent systems to address the solid waste management issues that were previously described. [17]

8.1. What is Smart Waste Management?

A creative approach to controlling and getting rid of waste is smart waste management. Based on artificial intelligence and Internet of Things technology, smart waste management provides data on waste producing trends and behavior. This makes it possible for towns, municipalities, and waste collectors to boost sustainability, streamline waste operations, and make more intelligent economic decisions.[17]

8.2. The Old Way of Doing Waste Management

Roughly 80% of waste collections take place during inappropriate hours. Overflowing trash cans, dirty surroundings, complaints from the public, illegal disposal, and higher cleaning and collection expenses are all consequences of delayed garbage collections. Early disposal of waste increases carbon emissions, increases operational expenses, and creates greater traffic jams. The outdated approach to waste management is completely ineffective. Adopting an innovative and data-driven approach is the only way to win in today's ever-technological environment. [17]

Municipalities and waste management companies used to have a predetermined pickup route and schedule. This suggests that waste collection trucks would take out every waste container along their path of collection, even if there was no need to do so. This leads to high labor and fuel costs, which the locals ultimately bear. This is also an unsustainable way of operating, as more cars on the road gathering pointless stuff equates to more carbon emissions released into the environment.[17]

8.3. The Future of Waste Management

In today's unstable and constantly evolving environment, the waste sector needs an event-driven waste collection system. Authorities in charge of managing waste have to stop focusing on outdated waste trends and patterns that don't fit with our modern way of life. The waste industry requires a system that utilizes real-time data to guarantee that waste is only collected when necessary. And this is when using a waste disposal with judgment becomes helpful.[17]

8.4. The Benefits of Smart Waste Management

- **Optimized Resources:** Smart waste management systems enable data-driven decision-making, real-time data allocation, and enhanced accuracy in meeting waste criteria in order to give better resource allocation and waste management.[17]
- **Reduced expenses:** Smart waste management may result in lower operational expenses and cash flow by utilizing data, identifying unnecessary spending, and allocating funding to recycling or zero waste efforts.[17]
- **Cleaner Streets:** Smart waste management solutions help communities cut trash in dirtier regions and respond quickly to overflowing bins by leveraging changing waste trends.[17]
- **Better Working Environments**: Sensible trash management enhances working conditions for waste collectors, boosts productivity, and encourages a safer, more hygienic workplace for efficient waste disposal.[17]
- **Reduced Carbon Emissions:** To deal with overflowing waste, towns usually use scheduled collection rounds, however this is inefficient and harmful to the environment. Productivity can be raised by using data to guide decisions about waste management.[17]
- **Increased Recycling Rates:** Data can increase recycling involvement and streamline waste management plans, leading to a better understanding of recycling behaviors and ongoing campaign improvement, even though it is hard to manage without measurement.[17]

9.Underwater waste Definition

waste that winds up in seas, oceans, or other big bodies of water is known as underwater waste. There are numerous ways that this man-made waste enters the ocean. Waste is frequently dumped into the ocean from boats or offshore structures like oil rigs, or it is left on beaches. Litter occasionally finds its way from the land into the ocean. Storm drains, canals, and rivers transport this detritus. Waste from landfills and other locations may even be blown into the lake by the wind. Ships may sink or lose cargo due to storms and maritime mishaps.[18]

Figure 3: underwater waste [20]

9.1. What We Can Do About underwater waste

There are numerous ways that this man-made waste enters the ocean. Waste is frequently dumped into the ocean from boats or offshore structures like oil rigs, or it is left on beaches. Litter occasionally finds its way from the land into the ocean. Storm drains, canals, and rivers transport this detritus. Waste from landfills and other locations may even be blown into the lake by the wind. Ships may sink or lose cargo due to storms and maritime mishaps.[18]

The large amount of the waste and the difficulties in reaching it once it has sunk to the ocean floor make cleaning up underwater waste a difficult task. Governments and international organizations are enacting regulations against ocean dumping as part of environmentalists' efforts to stop additional waste from ending up in the oceans. Public awareness campaigns on the risks associated with ocean littering are being carried out by groups such as the National Geographic Society and the Algalita Marine Research Foundation. David de Rothschild, a National Geographic Emerging Explorer, works with Algalita to lessen underwater waste, particularly plastics. Everyone should refrain from littering and practice reduce, reuse, and recycle as even waste that has been put on land can wind up in waterways. The ecology will absorb less waste the less of it is produced.[18]

Figure 4: what happen to your waste [19]

10. Underwater waste sources and risks

10.1. Sources

Here are some key sources of underwater waste:

The European Union and the world economy are facing serious challenges as a result of the growing quantity of waste produced by humans that ends up in seas and oceans. This covers consumer products, industrial procedures, and other human activities. Underwater waste is also influenced by land-based waste, such as that which ends up in storm drains, rivers, or wind turbines. Under international and EU law, ships, including those engaged in regular operations, are being managed appropriately. Water damage and leaks have resulted from the World War II deposition of chemical weapons into the ocean.[19]

Since 1993, it has been illegal for humans to dispose of radioactive waste, which has long been thrown into maritime environments. The production and management of waste from the ocean have also been impacted by the COVID-19 epidemic. In conclusion, there are many different sources of underwater waste, from past operations like the disposal of chemical weapons and radioactive waste to direct and indirect inputs from human activity on land and at sea.[19]

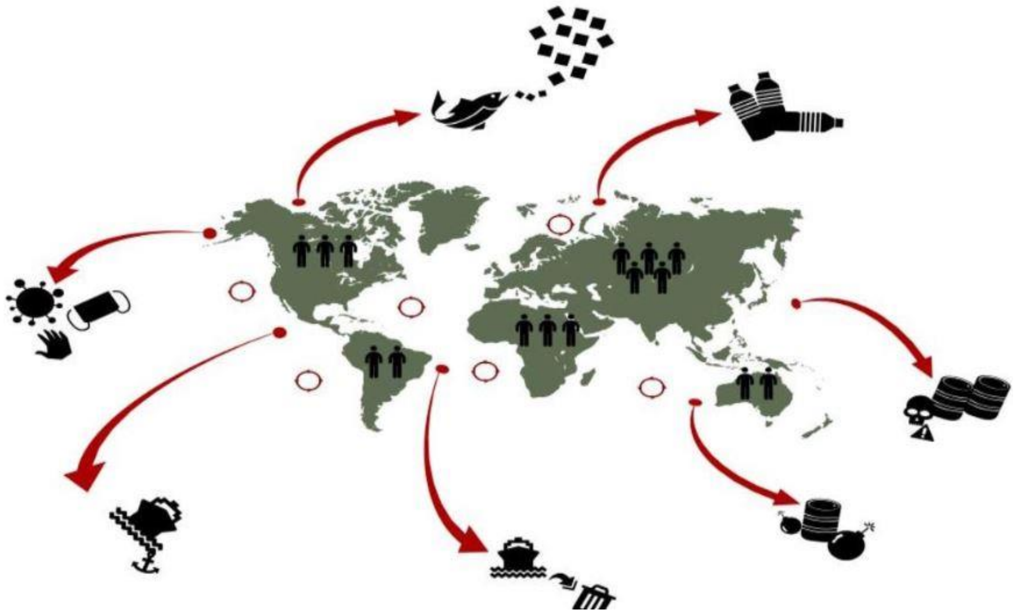

Figure 5: underwater waste sources [19]

10.2. Risks

Underwater waste poses a major threat to human health, ecosystems, and marine life. The ocean is made up of about 5.25 trillion pieces of plastic waste, and including 269,000 tons float on the surface and four billion plastic microfibers per square kilometer in the deep sea. After World War 2, chemical and radioactive waste spilled into the sea, contaminating the water. Since 1993, the disposal of radioactive waste has been strictly prohibited. [20]

Nearly all oceanic species groups are negatively impacted by plastic pollution, which has an adverse effect on 90% of examined species. Additionally, it has an impact on the productivity of mangroves and coral reefs, two of the most significant maritime ecosystems in the world. Underwater waste also poses a risk to human health because contaminated beach water can harm people and eating fish can cause toxins and microplastics to enter the food chain. Therefore, in order to reduce these risks, efficient waste management techniques are desperately needed.[21]

11. Underwater waste reduction and sorting

As part of marine conservation efforts, underwater waste reduction and sorting—which entails gathering, classifying, and processing waste items found in underwater environments—are essential. Divers can do this procedure manually, or autonomous underwater vehicles (AUVs) or remotely controlled vehicles (ROVs) can be used. After that, the waste is categorized according to its kind, such as abandoned fishing gear or plastic waste. Non-recyclable items are disposed of in an environmentally responsible manner, while recyclable ones are transported to recycling facilities. [22]

A number of approaches and technologies have been created to evaluate how well cleaning technologies for underwater waste work. Given that rivers account for over 80% of the waste in the ocean, it is imperative to prioritize waste collection technologies in rivers.

Divers, scientists, and conservationists must join this global campaign to stop additional pollution and advance sustainable practices.[22]

12. Underwater waste valorization

Underwater waste valorization refers to the process of adding value to waste materials found underwater, often through recycling or other forms of processing. This is a crucial aspect of marine conservation efforts, as it helps to reduce the impact of human waste on marine ecosystems.[23]

underwater waste valorization is a critical aspect of marine conservation, involving the collection, reporting, and processing of underwater waste to reduce its impact on marine ecosystems. It’s a global effort that requires the participation of divers, scientists, and conservationists around the world. It’s not just about cleaning up the oceans, but also about preventing further pollution and promoting sustainable practices.[23]

13. Conclusion

This chapter concludes by outlining the fundamentals of underwater waste management. We've covered the concepts of waste definition, history and hierarchy and how to sort and reduce waste, we have introduced smart waste management and underwater waste definition, sources, risks, also we have explained underwater waste management, sorting, reducing and valorization. In the next chapter we will see the concepts of machine learning and deep learning and the algorithms used to segment, recognize and classify underwater waste.

Chapter 2:

State of the Art

II. Chapter 2: State of the Art (SOTA)

1. Introduction

Artificial intelligence has officially become a scientific discipline that aims to break down intelligence into elementary functions in order to create machines capable of mimicking human cognitive abilities. Thinking, as well as communication and memory which is the ability to record and retain data.

In this chapter we will explore in more detail machine learning and deep learning, we will cover exactly the neural networks and its different types and architectures, we introduce the different concepts of computer vision, also we will present deep learning approaches for computer vision and finally the we present related works of our project.

II. Chapter 2: State of the Art (SOTA)

2. Artificial Intelligence

Artificial intelligence, or AI, is generation that allows computer systems and machines to simulate human intelligence and problem-fixing capabilities.[24]

On its very own or blended with different technologies (e.g., sensors, geolocation, robotics) AI can carry out responsibilities that could in any other case require human intelligence or intervention. Digital assistants, GPS guidance, self-sustaining vehicles, and generative AI tools (like Open AI`s Chat GPT) are only some examples of AI withinside each day information and each day lives.[24]

As an area of pc science, synthetic intelligence encompasses (and is frequently noted collectively with) device learning and deep learning. These disciplines contain the improvement of AI algorithms, modeled after the decision-making methods of the human brain, that can `learn` from to be had facts and make more and more correct classifications.[24]

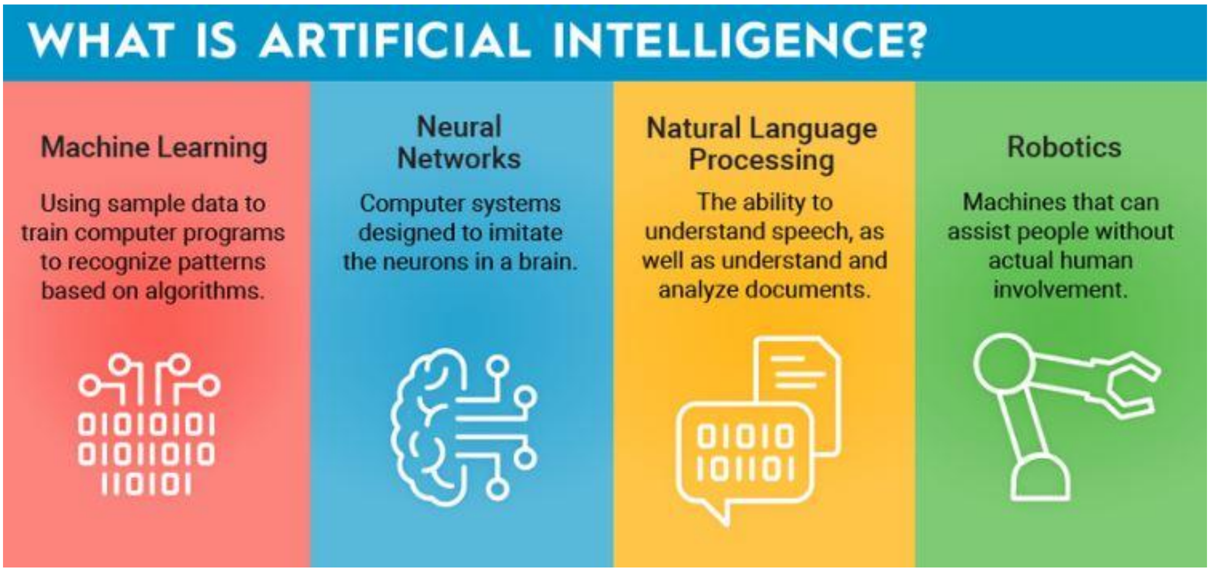

Figure 6 :What Is AI? [25]

3. Machine Learning

Without the assistance of a person, machines can learn from data and make predictions thanks to the artificial intelligence discipline of machine learning (ML). Computers that are aware of machine literacy can operate without programming and can adjust to new data. ML algorithms work better with more instances and rely on repetition to uncover patterns. A branch of machine learning called deep literacy trains computers to imitate natural traits for better performance measures.[26]

3.1.How does ML work?

Using a training dataset, ML algorithms create models, which are subsequently used to predict outputs when new input data is provided.[26]

II. Chapter 2: State of the Art (SOTA)

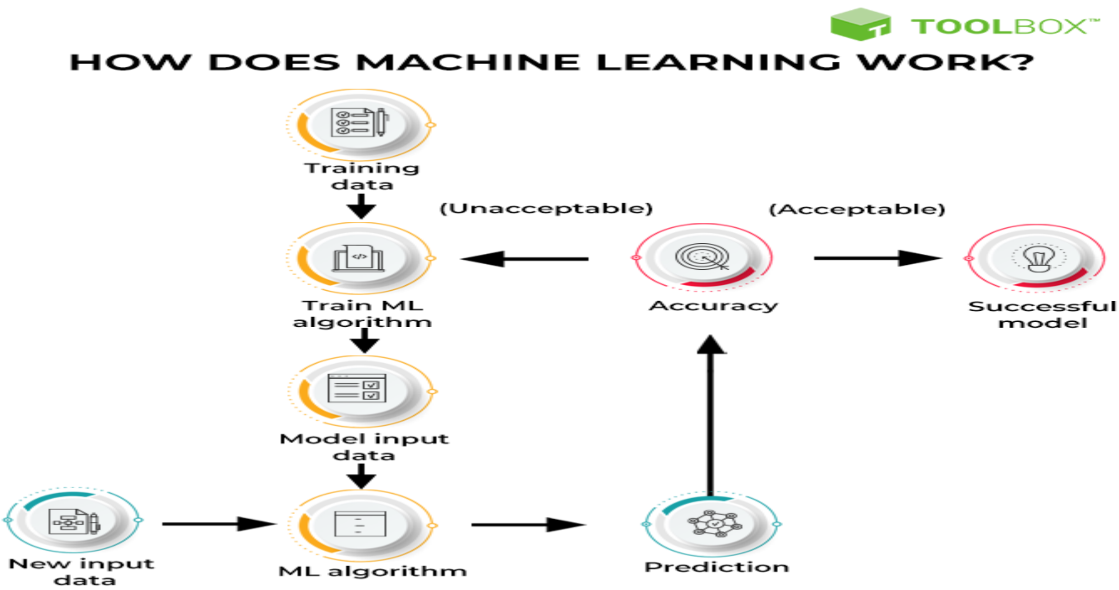

Figure 7:How ML Works [26]

Until the required accuracy has been achieved, the machine learning system is continuously trained or deployed using a better dataset.[26]

3.2.Machine Learning types

There are four types of learning styles and methodologies that can be used to train ML algorithms.[26]

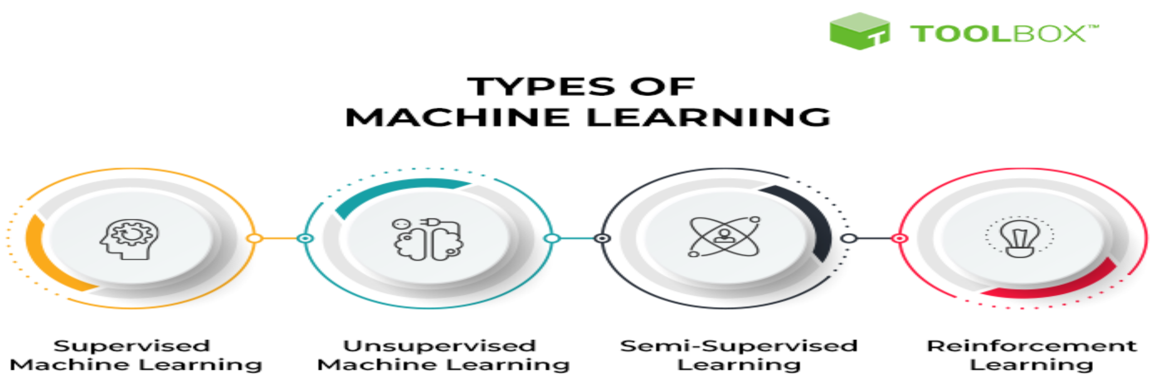

Figure 8:Types of ML [26]

3.2.1. Supervised machine learning

Using labeled datasets, machine learning involves training machines to predict outputs based on their prior knowledge. The labeled dataset serves as the basis for planning the input and output parameters. Based on the test dataset, a gadget is developed to read the outcomes. The Random Forest Algorithm, Support Vector Machine Algorithm, Lasso Regression, Decision Tree Regression, and Simple Linear Regression are popular supervised techniques.[26]

3.2.2. Unsupervised machine learning

Unsupervised learning is a technique where a computer learns on its own, making autonomous predictions, from an unlabeled dataset. It uses the input's similarities, differences, and styles to classify the unsorted dataset. Apriori algorithms, Principal Component, DBSCAN, K-Means clustering, Mean-Shift, and Independent Component analysis are a few examples.[26]

3.2.3. Semi-supervised learning

Both supervised and unsupervised gadget mastering capabilities are found in semi-supervised mastering. To educate its algorithms, it combines datasets with and without labels. Semi-supervised mastering addresses the constraints of the formerly defined answers through utilizing each kind of datasets.[26]

3.2.4. Reinforcement learning

The manner of reinforcement Learning of is feedback-driven. Here, the AI factor acts, learns from mistakes, and complements overall performance via way of means of robotically assessing its surroundings via the hit-and-trial method. The factor is rewarded foreach a success motion and punished for every unsuccessful one. Therefore, via way of means of doing well, the reinforcement learning seeks to maximize rewards.[26]

4. Deep Learning

A kind of machine learning called "deep learning" makes use of multi-layered neural networks to mimic the intricate decision-making capabilities of the human brain. Large volumes of data are used to train deep neural networks, or DNNs, to detect and categorize occurrences, spot patterns, assess options, and make predictions and choices. A deep neural network's extra layers refine and optimize results for higher accuracy, but single-layer neural networks can still produce valuable predictions and judgments.[27]

4.1. History and Applications of Deep Learning

With two significant pauses, deep learning—a computer model modeled after the neural networks in the human brain—was made possible by the "threshold logic" techniques created in 1943 by Walter Pitts and Warren McCulloch.[28]

- **The 1960s**

Henry J. Kelley's 1960 research served as the basis for Stuart Dreyfus' 1962 development of the Back Propagation Model. Back propagation was useless before to 1985. Deep learning algorithms were attempted in 1965 by Alexey Grigoryevich Ivakhnenko and Valentin Grigorňevich Lapa utilizing statistical analysis, polynomial activation functions, and forwarding best features.[28]

- **The 1970s**

Due to unfulfilled expectations and insufficient financing for deep learning and AI research, the first AI winter appeared in the 1970s. Renowned convolutional neural network designer Kunihiko Fukushima invented the first artificial neural network, Neocognitron, in

1979. Recurrent activation in multiple layers was used to teach the machine to recognize patterns in visual stimuli.[28]

Top-down connections and innovative learning strategies have enabled the building of various neural networks. Neocognitrons are now able to identify patterns with partial information and add missing details to images to complete them. Following the 1970 introduction of a FORTRAN code for back propagation by Seppo Linnainmaa, the field of back propagation saw significant evolution.[28]

- **The 1980s and 90s**

In 1989, Yann LeCun used convolutional neural networks at Bell Laboratories to show backpropagation. While research on neural networks and deep learning was negatively harmed by the second AI winter (1985–1990s), a number of individuals persisted in these areas, producing important breakthroughs like the support vector machine and LSTM for recurrent neural networks. With the development of GPUs in 1999, neural networks were able to outperform support vector machines and achieve higher processing rates.[28]

- **2000-2010**

Neural networks' condensed input and narrow output ranges have prevented them from learning features since the Vanishing Gradient Problem in 2000. This problem was unique to gradient-based learning methods and was solved by layer-by-layer pre-training and long short-term memory growth. The amount and speed of data sources are increasing, as demonstrated by a 2001 study by META Group. In 2009, AI researcher Fei-Fei Li created ImageNet to train neural nets with the goal of revolutionizing machine learning and enabling learning through data.[28]

- **2011-2020**

Convolutional neural networks were able to train without the need for layer-by-layer pre-training in 2011 because to the GPU speed improvement, which paved the way for the creation of deep learning algorithms like AlexNet. This resulted in global competitions in 2011 and 2012. In 2012, the Cat Experiment looked at repeating patterns in unlabeled data to investigate the difficulties associated with unsupervised learning. 2014 saw the introduction of the Generative Adversarial Neural Network (GAN), which pits neural networks against one another in a game of photo emulation to trick the other network into believing it to be real. Con artists have utilized GAN to improve their products.[28]

4.2.Convolutional Neural Network (CNN)

A deep learning method called convolutional neural networks (CNNs) is intended for use in image processing and recognition applications. They can extract hierarchical point representations from raw pictures with less preparation than bracket models. Convolutional layers in CNNs are especially good in describing color objects and characteristics in images by using contaminants to represent original patterns. CNN connection patterns are inspired by the visual cortex of the brain, which facilitates effective recognition of spatial relationships and patterns in images. By layering convolutional and pooling layers, CNNs may learn more complicated features and achieve high accuracy in tasks like image classification, object recognition, and segmentation.[29]

4.2.1. Convolutional Neural Network Architecture Model

Processing text, audio, and visual data is a highly effective use of convolutional neural networks. A fully connected (FC) layer, pooling layers, and convolutional layers make up their three main layers. The machine learning model can be made more sophisticated and accurate by using many pooling and convolutional layers. The model becomes more adept at recognizing patterns and objects in the data with each additional layer.[29]

The Convolutional Layer

Convolutional layers in a network carry out the majority of its computations and are mainly utilized by feature detectors to filter input data. A portion of a two-dimensional image is represented as a two-dimensional array of weights by the filter, which computes the dot product between pixels. In order to create nonlinearity, the CNN applies the ReLU adjustment to each feature map following each convolution.[29]

Pooling and convolutional layers make up a convolutional block. Multiple convolution blocks are added after the initial block to form a hierarchical structure. A CNN model, for instance, may identify every aspect of a car as a low-level pattern, which the neural network would then merge to produce a high-level pattern.[29]

The Pooling Layers

A down sampling or pooling layer lowers the input's dimensionality. Pooling operations, which do not employ weights, employ a filter to sweep the entire input image, just like convolutional operations do. Instead, the filter utilizes an aggregation function to fill the output array with values from the receptive field. [29]

Two categories of pooling exist:

Mean pooling: When the filter scans the input, it determines the average value of the receptive field.

Max pooling: To fill the output array, the filter delivers the pixel with the highest value. Compared to average pooling, this strategy is more popular.

Despite the fact that they can lead to some information loss, pooling layers are crucial because they make CNNs simpler and more effective. It lowers the possibility of overfitting as well. [29]

The Fully Connected Layer

A CNN's fully linked layer is its last layer. The features that were extracted by the earlier layers and filters are used by the FC layer to carry out classification tasks. The FC layer usually employs a SoftMax function, which more accurately classifies inputs and generates a probability score between 0 and 1, in place of ReLU functions.[29]

4.2.2. Types of Convolutional Neural Network Algorithms

There are many CNN’s algorithms, such as:

- **LeNet**

LeNet is a cutting-edge CNN created to identify handwritten characters. In the late 1990s, Patrick Haffner, Yoshua Bengio, Leon Bottou, and Yann LeCun all made the proposal. LeNet is made up of a fully connected layer, a SoftMax classifier, and several convolutional and pooling layers. It was one of the earliest instances of deep learning being successfully used to computer vision. Banks have used it to recognize numbers in grayscale input images written on checks.[29]

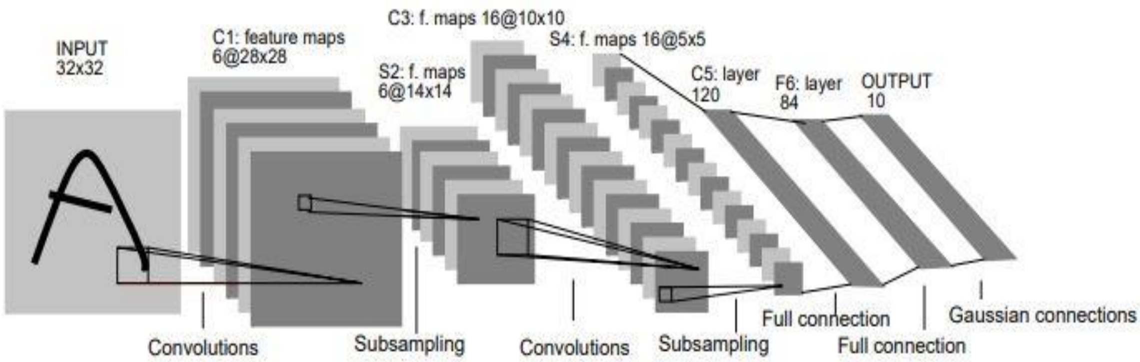

Figure 9:Architecture of LeNet [34]

- **VGG**

In addition to additional convolutional layers and pooling layers, VGG 16 boasts a 16-layer architecture. For nearby pixels, 3x3 convolutions are used. Because it has fewer parameters, VGGNet is a deeper network with smaller filters, which permits more layers and modest filters. One 7x7 convolutional layer is equal to its effective receptive field.[30]

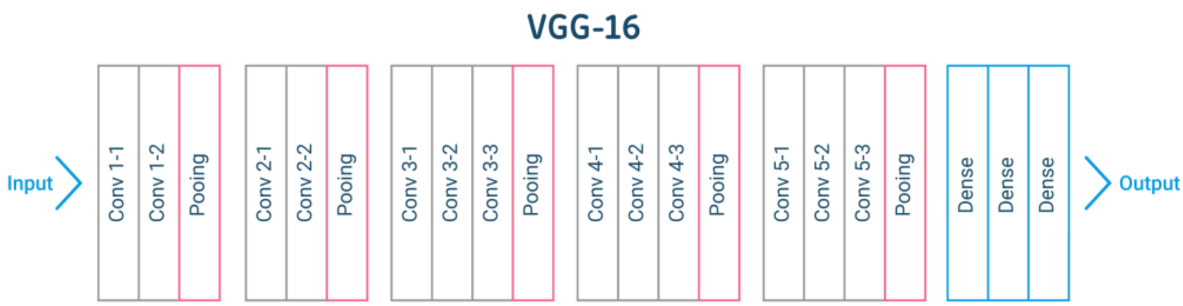

Figure 10:VGG[30]

- **AlexNet**

With a 15% margin of error reduction, the deep neural network AlexNet emerged victorious in the 2012 ImageNet Large Scale Visual Recognition Challenge. The University of Toronto's Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton developed it. It makes use of three FC layers and five convolutional layers with max pooling, 11x11, and 3x3 filters, as well as ReLU activation functions. Because of AlexNet's success, deep learning methods are receiving more attention, and GPUs are becoming more commonly used for deep neural network training.[29]

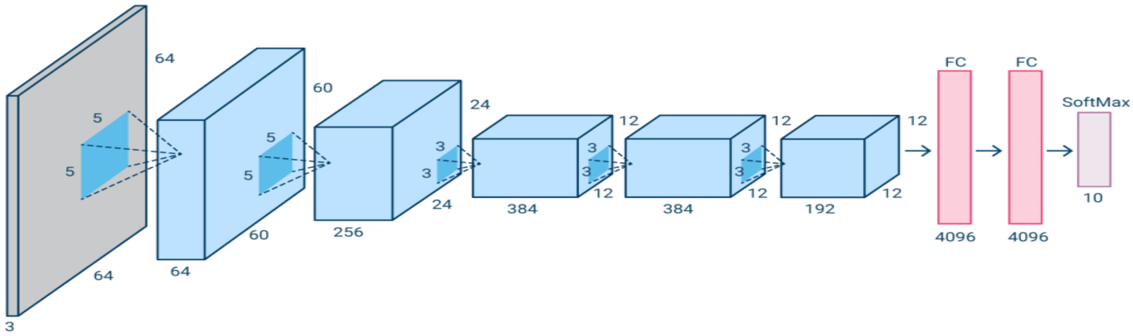

Figure 11: AlexNet [30]

- **ResNet**

In order to handle fading gradients in extremely deep networks, ResNet is a deep convolutional neural network. Through the usage of residual blocks, gradients can spread straight across the network. An activation function, convolutional layers, and a shortcut link that appends the original input to the output following the activation function make up ResNet. The network can then learn residual functions that reflect the variation between input and output, which lessens the issue of vanishing gradients in very deep networks.[29]

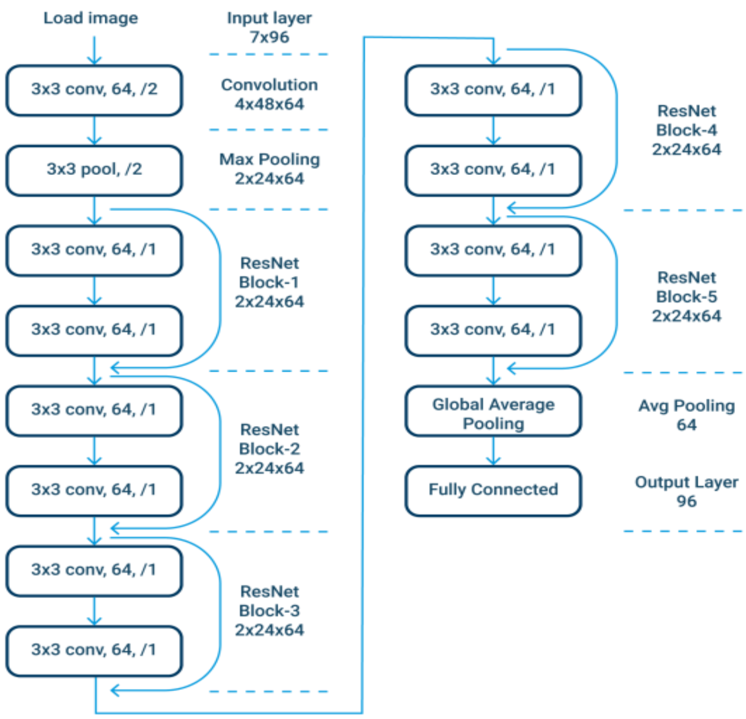

Figure 12:ResNet[30]

- **GoogLeNet**

With a top-five mistake rate of 6.67%, Google's deep convolutional neural network, GoogLeNet, took first place in the ILSVRC in 2014. The network can learn features at various resolutions and sizes because to its Inception module, which lowers computing costs. It is composed of numerous convolutional layers, filter sizes, pooling layers, and output concatenation. Auxiliary classifiers are also included at intermediate levels in order to prevent overfitting and promote discriminative learning. Compared to earlier convolutional neural networks, GoogLeNet is a more advanced and complex network.[29]

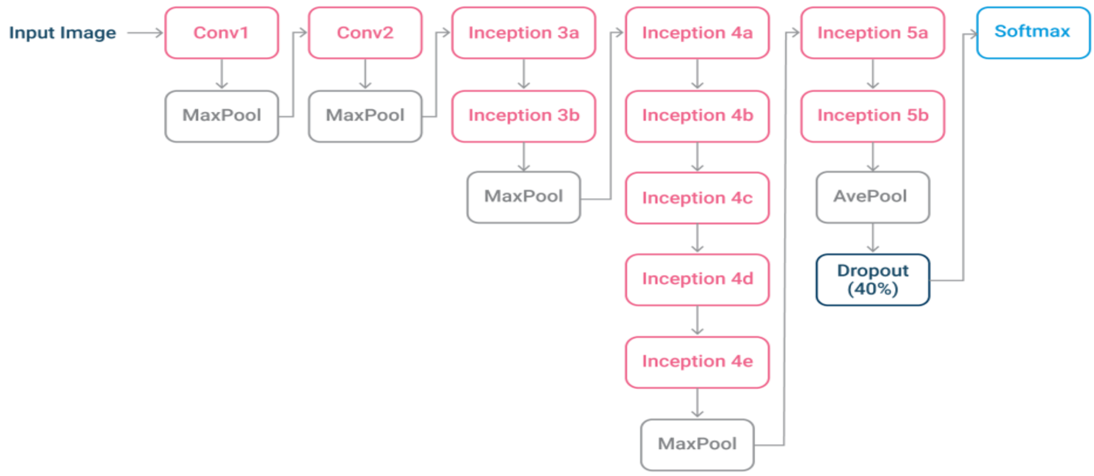

Figure 13: GoogLeNet [30]

5. DL for Computer Vision Tasks

The scientific discipline of computer vision (CV) describes how machines decipher the meaning of images and movies. Following an analysis of specific parameters in photos and videos, computer vision algorithms apply their interpretations to tasks involving prediction or decision-making.[30]

Deep learning methods are now mostly applied to computer vision. The various applications of deep learning for computer vision are examined in this article. You will gain specific knowledge regarding the benefits of employing convolutional neural networks (CNNs), which offer a multi-layered design that enables neural networks to concentrate on the image's most pertinent aspects.[30]

5.1.Uses of Deep Learning in Computer Vision

The advancement of deep learning technology has made it possible to construct computer vision models that are more intricate and accurate. The integration of computer vision applications is becoming increasingly beneficial as these technologies advance. Here are some examples of how computer vision is being enhanced using deep learning.[30]

5.1.1. Image Classification

Labeling of the visual object class can be done based on the likelihood of its presence. The majority of deep learning techniques use bags of visual words, which at first. After obtaining a quantized visual word histogram, we go on to the classification step. Sparse coding is typically used to recover lost information.[31]

5.1.2. Object detection

Two popular methods of object detection using computer vision algorithms are as follows: [30]

- **Two-stage object detection:** A Region Proposal Network (RPN) is needed for the first phase, which yields a number of potential regions that might include significant items. The second phase involves sending area recommendations to a neural classification architecture, such as Fast RCNN's region of interest (ROI) pooling or an RCNN-based hierarchical grouping algorithm and Mask RCNN. These methods can be very slow, but they are highly accurate.[30]
- **One-step object detection:** In response to the requirement for real-time object detection, architectures for one-step object detection, including YOLO, SSD, and RetinaNet, have been developed. Through the regressing of bounding box predictions, these integrate the detection and classification steps. Because each bounding box is represented by a small number of coordinates, processing may be sped up and the detection and classification steps can be combined more easily.[30]

5.1.2.1.DL techniques used for detection

The methods of object detection based on deep learning are multiple, the R-CNN algorithms are among the methods for applying deep models Learning the problem of object detection. Here is someone :

a. RCNN Architecture

One kind of deep learning architecture used for object detection in computer vision tasks is the region-based convolutional neural network (R-CNN). One of the first models to combine the capabilities of region-based methods and convolutional neural networks to enhance object detection was the RCNN.[33]

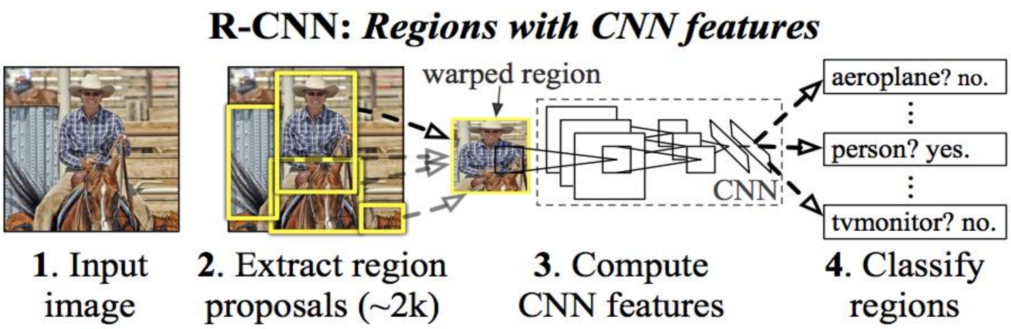

Figure 14:RCNN Architecture [33]

b. Faster R-CNN Architecture

2015 saw the release of the state-of-the-art object identification system, Faster R-CNN architecture, by Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. This unified architecture combines the advantages of deep learning, convolutional neural networks, and region proposal networks to identify objects in images and assist in their detection. Two components make up the Faster R-CNN architecture, which improves model accuracy and speed.[34]

- Region Proposal Network (RPN)
- Fast R-CNN detector

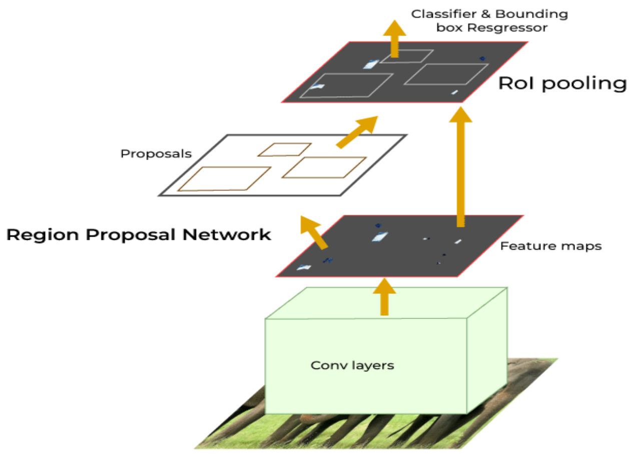

Figure 15:Faster R-CNN [34]

c. Architecture YOLO:

GoogLeNet and YOLO architecture are comparable. It has two fully connected layers, four max-pooling layers, and 24 convolutional layers total, as shown below.[35]

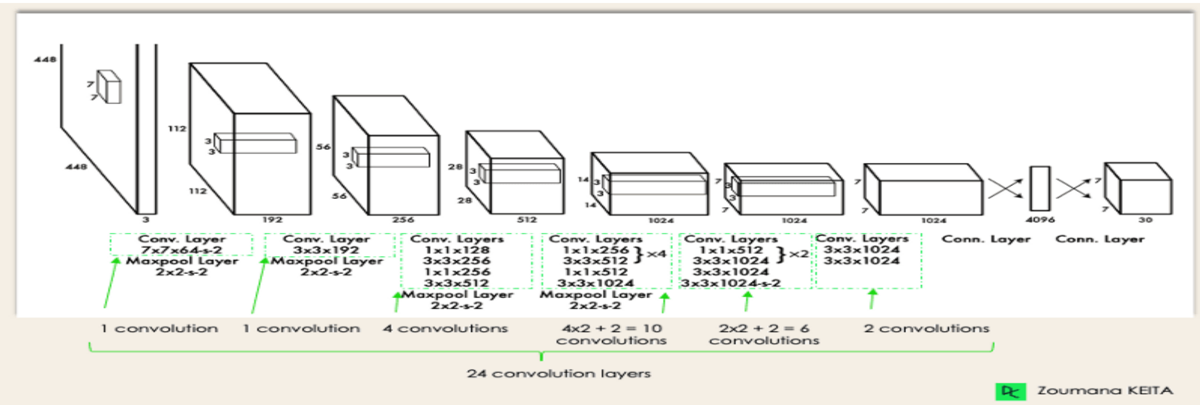

Figure 16:YOLO Architecture [35]

The input image is resized by the architecture to 448 by 448 before being processed by a neural network. After 1x1 convolution, 3x3 convolution produces a cuboidal result. Save for the last layer, the underlying activation function is ReLU. The model is regularly adjusted and overfitting is avoided by the use of batch normalization and dropout techniques.[35]

5.1.3. Semantic segmentation

Object segmentation, or semantic segmentation, is comparable to object detection but relies on the particular pixels associated with an object. This eliminates the need for boundary boxes and allows image objects to be more precisely specified. Typically, U-Nets or fully convolutional networks (FCN) are used for semantic segmentation. Semantic segmentation is widely used in autonomous vehicle training. Researchers can utilize these photos of streets or throughways with clearly defined object boundaries.[35]

5.1.3.1.DL techniques used for semantic segmentation

- **U-Net**

Two paths make up the semantic segmentation architecture U-Net: a contracting path and an expanding one. Recurring 3x3 convolutions, rectified linear units (ReLU), and a 2x2 max pooling operation for down sampling are all part of the contracting path. Up sampled feature maps, concatenation with the contracting path, 2x2 and 3x3 convolutions are all part of the expanding path. Cropping is inevitable since boundary pixels are lost after each convolution. The last layer uses a 1x1 convolution to distribute 64-component feature vectors to the necessary number of classes.[36]

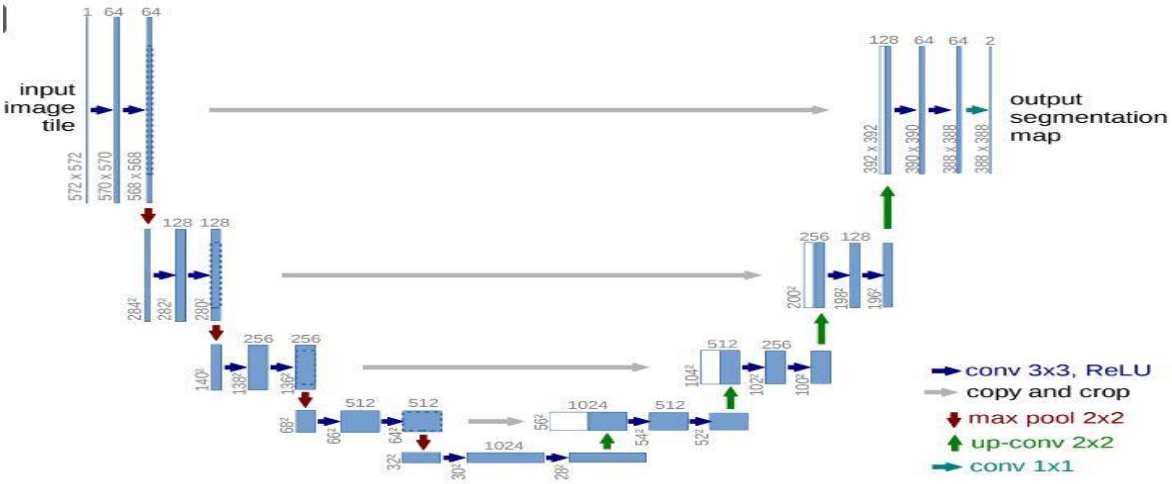

Figure 17: U-Net architecture [36]

5.1.4. Instance Segmentation

More detailed than semantic segmentation, instance segmentation recognizes every instance of the same thing. Instance segmentation, for instance, treats each of the three elephants in an image as a single instance and will detect and highlight them individually.[32]

5.1.4.1.DL techniques used for instance segmentation

- **Mask RCNN**

A deep learning model called Mask R-CNN creates precise segmentation masks for every item it detects by fusing object detection and instance segmentation. Its capacity to perform instance segmentation pixel-by-pixel is its primary novelty. Feature Pyramid Network (FPN) and ROIAlign are two significant advancements that enhance the precision of segmentation and feature extraction for small objects. These modifications increase the model's comprehension of object context.[37]

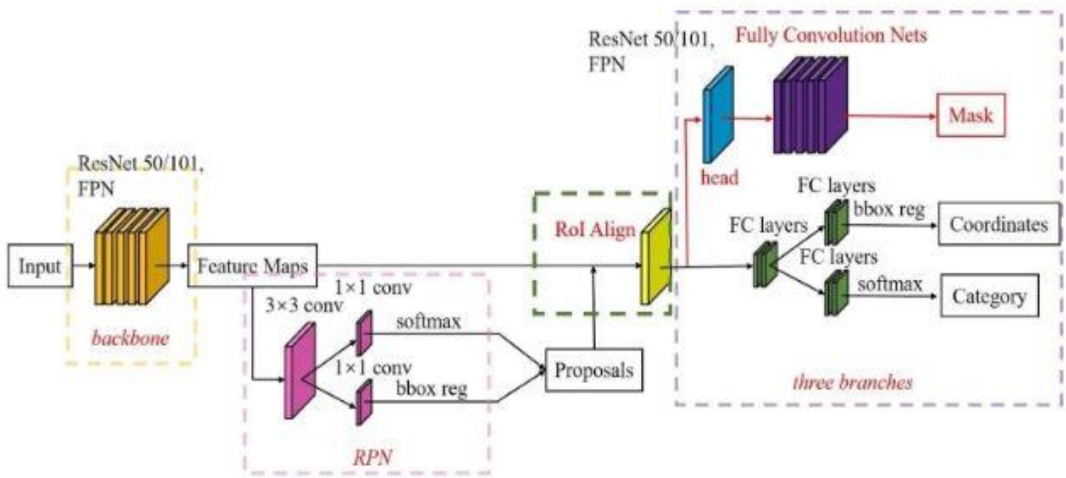

Figure 18: Mask RCNN architecture [37]

5.2.Object Detection vs. Classification vs. Segmentation

Object detection models identify the existence and placement of objects in pictures or videos, identifying several unique objects in a single image or video. While object detection models provide information about the object's location, image classification assigns a single label. A computer vision approach called image segmentation aids in precisely locating an object down to the pixel level. Nevertheless, in spite of its possible advantages, it necessitates lengthier projections and accurate label completion. Figure 19 below shows the difference between object classification, detection and segmentation.[38]

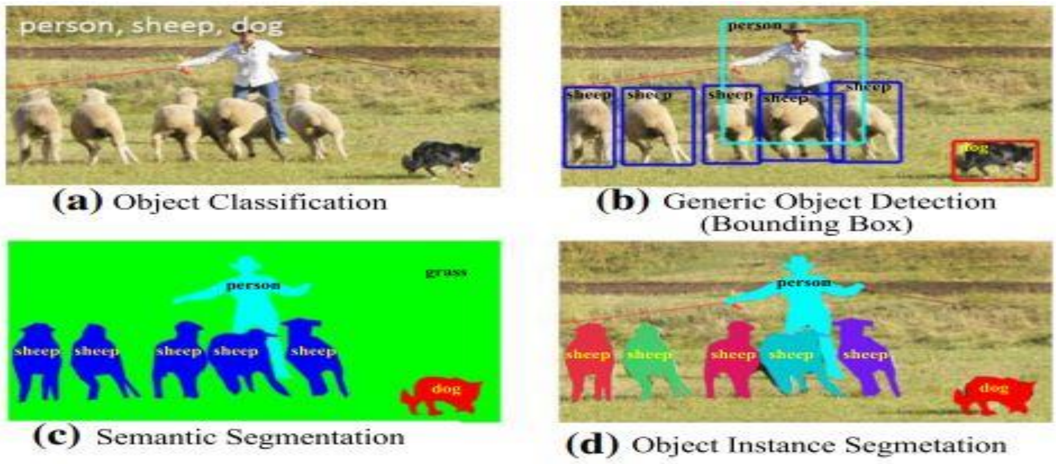

Figure 19 :Object Detection vs. Classification vs. Segmentation [39]

The following table brings together each DL model with its domain:

Table 1: DL Models

DL Model	Domain
RCNN	Object detection
Faster RCNN	Object detection
U-Net	Segmentation

Mask RCNN	Segmentation +Object detection
YOLO	Segmentation +Object detection

5.3.Computer Vision and Human Vision

Here's a comparison between computer vision and human vision:[40]

Table 2:Computer Vision vs Human Vision [40]

Computer Vision	Human Vision
Like human vision, computer vision enables a computer to sense and recognize objects in its environment.	**People see things as they are and remember what they recognize, putting it away deep in their brains for when those things come up again.**
Computer vision recognizes, separates, and classes objects using machine learning methods and algorithms.	**The process by which the eyes recognize patterns in light and work with the brain to convert that light into images is known as human vision.**
One of the hardest issues in computer vision is object recognition.	**People can identify objects in a scene very easily and have little trouble explaining them.**

Is computer vision better than human vision?

While computer vision performs exceptionally well in some scenarios, it is less effective in others, such as pattern identification and fraud detection. A critical ability unique to the human brain is invariant object identification, which enables people to recognize objects in complicated settings with ease. It is clear that further study is necessary to fully understand how the brain forms invariant representations of objects, even with a great deal of research having been done.[40]

6. Existing Studies

The study of the existing is a basic step, allowing to elaborate a synthesis bibliographic of theoretical knowledge on the subject. In the following, we will present some work based on deep learning applied to underwater waste management.

- **Robotic Detection of Marine Litter Using Deep Visual Detection Models [41]**

Fulton et al. (2019) used the LabelImg program to label 5,720 images from video data and assessed four algorithms in real time: Tiny-YOLO, Faster-RCNN, YOLOv2, and SSD. Faster-RCNN performed better with an 81% mAP, while YOLOv2 and Tiny-YOLO had a superior speed-accuracy balance. However, their analysis is more accurate when considering plastic detection average precision at 70.3% and 82.3%, compared to Faster RCNN's 83.3% plastic precision score. Each network processes three more films in addition to analyzing the data on the test set; these are shown in Figure 20.[41]

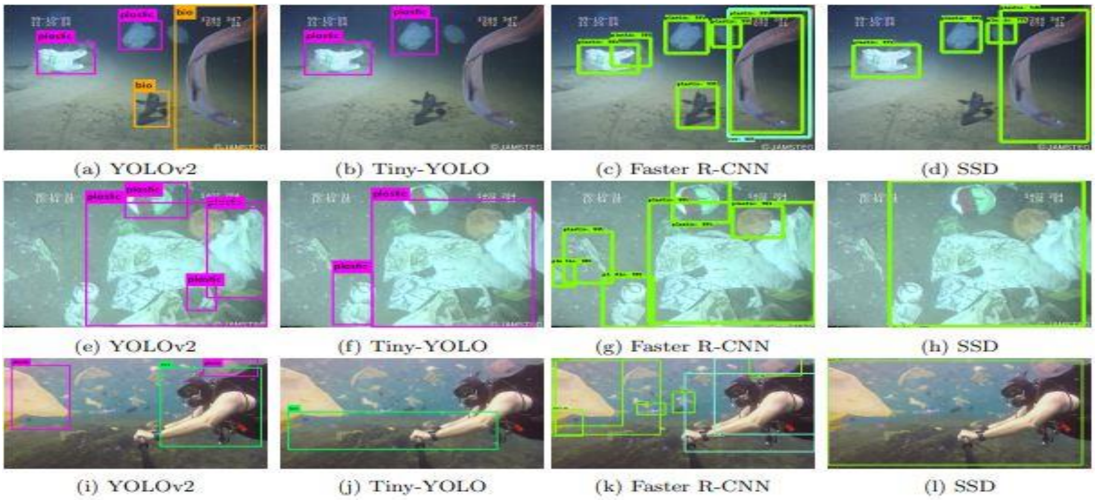

Figure 20: Detection results on two images.[41]

- **Underwater and airborne monitoring of marine ecosystems and debris [42]**

Watanabe et al. (2019) used foggy conditions to capture underwater trash images. They downloaded images from Google Open Images and used three cameras to capture beach detritus. They tested 1,127 photos and trained 6,908 images for fish, turtles, and jellyfish. They used a distinct detection algorithm for marine biodiversity and trash. Due to limited underwater data, they separated debris into plastic bottles, bags, and driftwood. Using YOLOv3, they achieved a Sea life mAP of 69.6% and Debris mAP of 77.2%. Example results of marine debris detection for manually entered footage at beaches are displayed in Figure 21.[42]

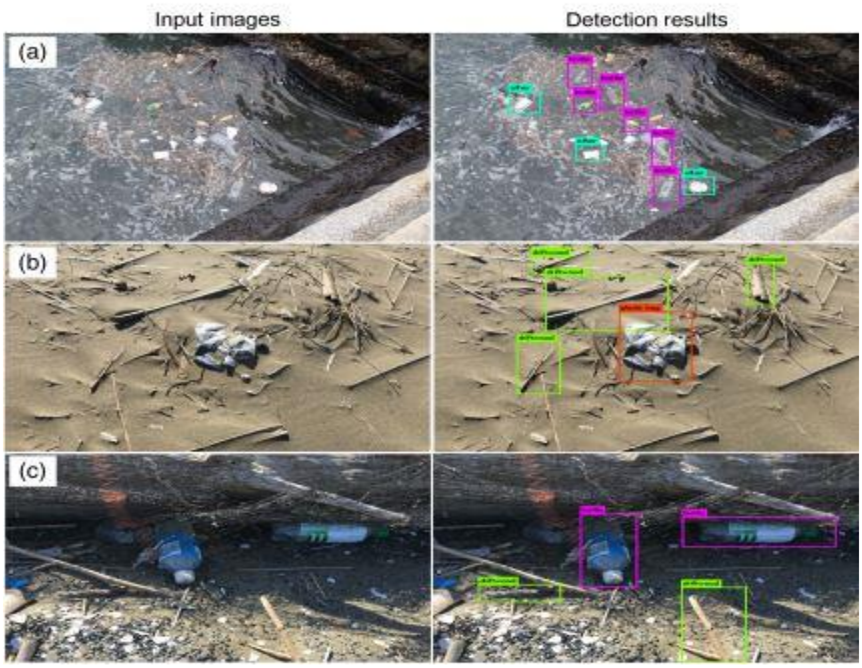

Figure 21:Example detection results of debris [42]

- **AquaVision: Automating the detection of waste in water bodies using deep transfer learning [43]**

Proenca and Simoes (2020) developed TACO, an open-source database for urban litter detection. Panwar et al. (2020) modified this database to create the AquaVision collection, which includes 369 underwater annotated photos of marine debris. The images were categorized into glass, metal, paper, and plastic, and with RetinaNet, achieved 68.7 precision. Figure 22 highlights the projected results that were obtained.[43]

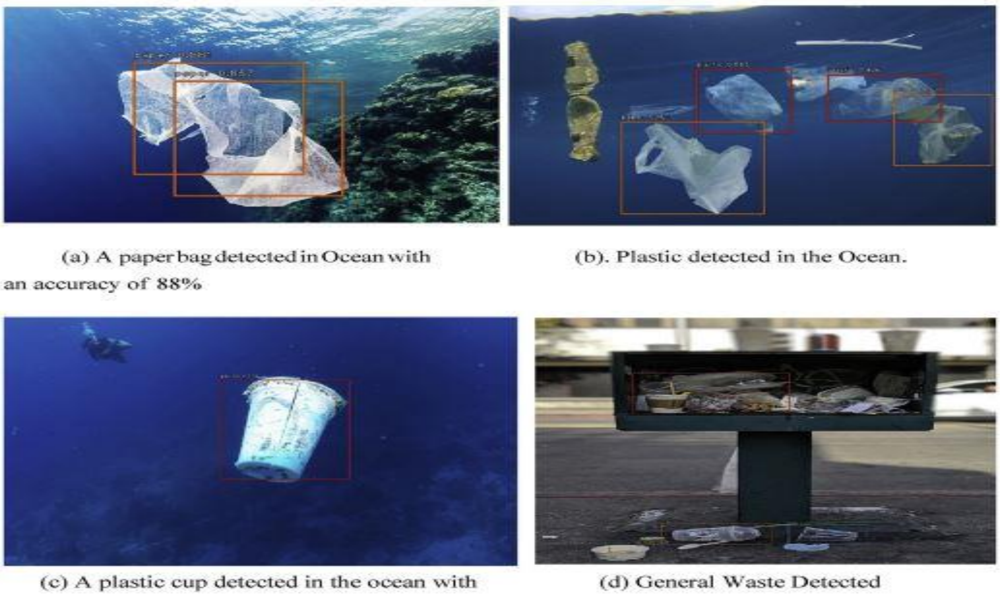

Figure 22: test of detection [43]

- **Deep-sea Debris Identification Using Deep Convolutional Neural Networks [44]**

With a mean Average Precision (mAP) of 83%, the authors of the other paper (Xue et al., 2021b) introduced a one-stage network called ResNet50-YOLOV3. They found that their results outperformed earlier detection networks in all seven classes with 10,000 images. A few debris photos from each dataset type are shown in Figure 23.[44]

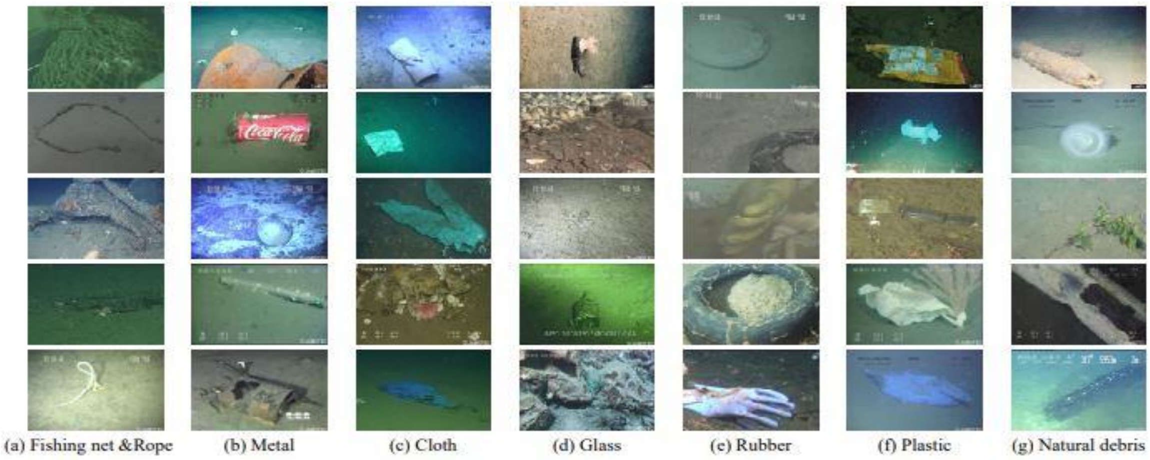

Figure 23: Examples of 7 types of deep-sea debris.[44]

II. Chapter 2: State of the Art (SOTA)

- **An Embeddable Algorithm for Automatic Garbage Detection Based on Complex Marine Environment [45]**

Deng et al.'s study focuses on instance segmentation and maritime debris detection using dilated convolution in the Feature Pyramid Network. They improved feature extraction and instance segmentation accuracy using the Trashcan dataset with 16 classes, achieving a mAP50 of 59.2%, 2.5% higher than the standard Mask R-CNN. In instance segmentation, SOLO (55.9%) and CondInst (57.1%) underperformed compared to other algorithms. In object detection, enhanced lateral connectivity increased the mAP50 to 65.2%, 9.5% higher than the original structure, outperforming rival models like FC0S, Retina-NET, and Faster-RCNN. The efficiency of the enhanced network in this work for the tasks of marine waste identification and instance segmentation is shown in Figure 24.[45]

Figure 24: model predictions [45]

- **Detection of River Plastic Using UAV Sensor Data and Deep Learning [46]**

The YOLOLabel tool was utilized by Maharjan et al. (2022) to document the bounding box of each plastic piece found in photos of river debris. They used georeferenced ortho-imagery and deep learning to classify the items into plastics and non-plastics. Two databases of plastic garbage found in rivers were created, and the YOLOv2, YOLOv3, YOLOv4, and YOLOv5 algorithms were used to train them. The accuracy rates for recognizing river plastics with an autonomous unmanned vehicle were 77%, 81%, 83%, and 83%, respectively. Transfer Learning from One Location to Another The results of the transfer learning experiments are shown in Table 3.[46]

Table 3:Performance comparison based on mAP [45]

YOLO Family	Best Model (Pre-Trained)	Evaluation Dataset	Mean Average Precision (mAP)			
			Training from Scratch	Pretraining on COCO; No Transfer Learning	Transfer from	Pretraining on COCO + Transfer
YOLOv5	YOLOv5s	HMH	0.74	0.81	TT	0.83
		TT	0.53	0.61	HMH	0.62
YOLOv4	YOLOv4	HMH	0.76	0.80	TT	0.83
		TT	0.54	0.60	HMH	0.61
YOLOv3	YOLOv3-spp	HMH	0.59	0.79	TT	0.81
		TT	0.39	0.57	HMH	0.59
YOLOv2	YOLOv2	HMH	0.58	0.72	TT	0.77
		TT	0.37	0.49	HMH	0.51

- **An experimental study on marine debris location and recognition using object detection [47]**

In a follow-up study, Sanchez-Ferrer et al. (2023) examined 1,223 photos from seventeen distinct classes using the CleanSea dataset. To arrange the set, they employed an 80:20 split with 10% validation. They used Mask RCNN to augment the data in order to improve the findings, obtaining Material mAP of 65.2% and Instance mAP of 63.5%. The goal of the study was to improve the images' classification accuracy. Figure 25 shows Examples of scenes extracted from the e-CleanSea corpus.[47]

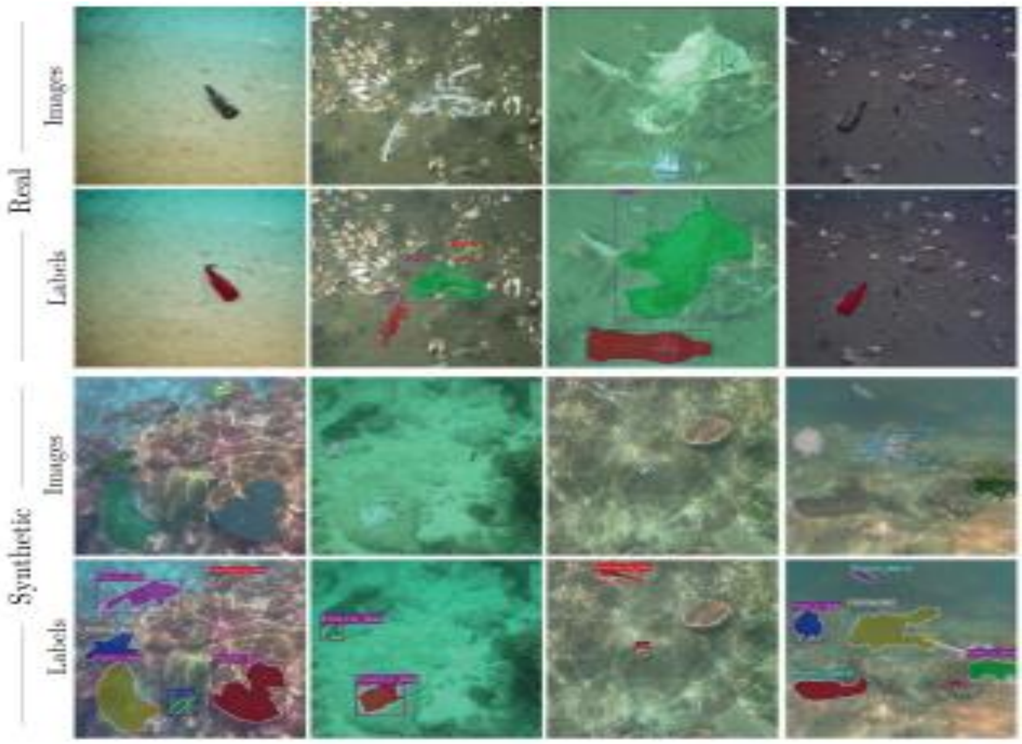

Figure 25:Examples of scenes extracted from the e-CleanSea corpus [47]

II. Chapter 2: State of the Art (SOTA)

- **Marine Debris Segmentation using Capsule Network (SegCaps) [48]**

This study introduces SegCaps, an object segmentation method based on capsule networks that reduces network depth layer complexity and produces accurate object segmentation, making it better than traditional CNN-based methods. The Trashcan dataset, which comprises 7,212 object images from three main categories with mask-level annotations, is used in the study. With a lower score of 26.25, the SegCaps model performs better than the MaskRCNN model, which obtained a mAP of 55.3. When compared to Mask RCNN, the SegCaps model's deeper instance segmentation design yields a lower mAP score. A process flow diagram for all the processes that are mostly carried out during the network modeling stage is shown in Figure 26.[48]

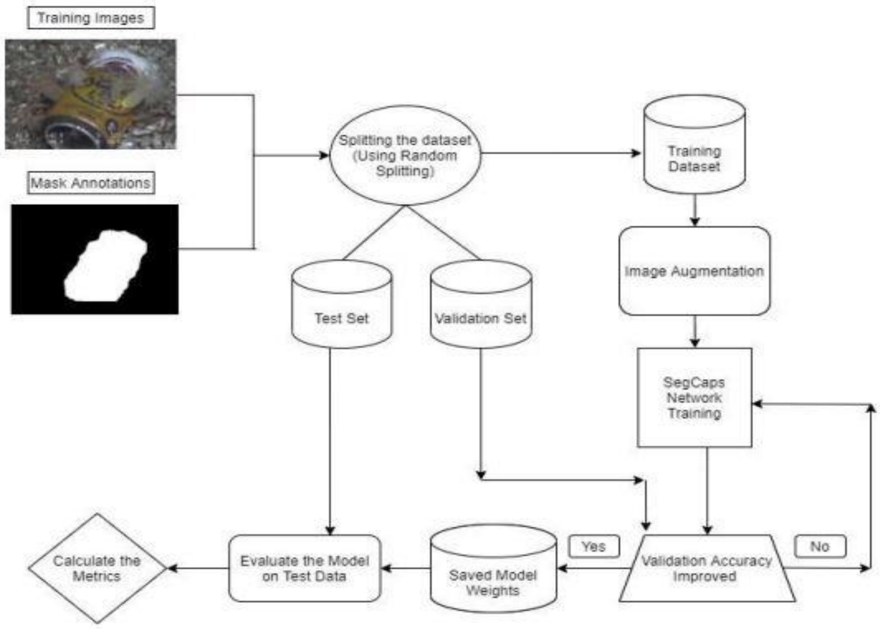

Figure 26: Modeling process flow [48]

Below we pass a summary of all the work mentioned before for the management of Underwater wastes:

Table 4: State of the Art

Domain	Method	Project	Model	Data	Classes number	maP(%)
		Robotic Detection of Marine Litter Using Deep Visual Detection Models	Tiny YOLO	5,720	3	82.3%
			YOLOv2			70.3%
			, Faster RCNN			81%

Deep Learning	Object Detection	Underwater and airborne monitoring of marine ecosystems and debris	YOLOv3	189 (debris) 8,036 (bio)	7	Debris 77.2% Sea life 69.6%
		Aqua Vision: Automating the detection of waste in water bodies using deep transfer learning	RetinaNet (Resnet50 backbone & FPN)	1,272	4	68.7%
		Deep-sea Debris Identification Using Deep Convolutional Neural Networks	ResNet50-YOLOv3	10,000	7	83.4%
		Detection of River Plastic Using UAV Sensor Data and Deep Learning	YOLOv2	7,212 >500 Tiles	1	77%
			YOLOv3			81%
			YOLOv4			83%
			YOLOv5s			83%
	Segmentation	An Embeddable Algorithm for Automatic Garbage Detection Based on Complex Marine Environment	Mask RCNN	7,212	16	65.2%
		Marine Debris Segmentation using Capsule	Mask R-CNN	7212	16	55.3%

		Network (SegCaps)				
		An experimental study on marine debris location and recognition using object detection	Mask R-CNN	1,223	17	Instance 63.5%; Material 65.2%

7. Conclusion

In this chapter, we have addressed the theoretical aspect of artificial intelligence, machine learning and deep learning with its history and fields of application, while explaining the models that deep learning can offer, then a presentation of segmentation and object detection models with their architectures to conclude with a comparison between computer vision tasks and a presentation of an overview of most important work based on deep learning. In the next chapter we will present our models with evaluation and predictions results, and we will compare our model with relative works in our domain.

Chapter 3:

Underwater Waste Segmentation, Design, Implementation, and Experimentation

III. Chapter 3: Underwater Waste Segmentation, Design, Implementation, and Experimentation

1. Introduction

Process automation in all aspects of underwater waste management has attracted more attention in computer science in recent years.

Numerous deep learning methods for sorting underwater waste have been put out in the literature to address this. Typically, there are two types of deep learning techniques: segmentation techniques and object detection techniques.

In order to achieve the best sorting possible, we have suggested two segmentation techniques in this chapter that deal with the segmentation and object detection of underwater waste. These techniques allow us to classify the wastes into eight categories: electronics, masks, plastic bottles, plastic bags, glass bottles, tires, metal, and waste.

This study's primary goal is to categorize the various waste categories. The performance of the proposed models is assessed using underwater-debris dataset from roboflow. Ultimately, we examined our results and compared them with another research done in the same field.

2. Overall Architecture

The photos were obtained from the underwater-debris dataset available on Roboflow. the next step was the data preprocessing (annotation, normalization, and data augmentation), the pre-trained model was downloaded and the training parameters (learning rate, weight decay, image size, batch size, etc.) were set up. The training (transfer learning) phase of our model is the following step. Lastly, we evaluated our model by generating predictions on test images. Figure 27 show the overall architecture of our work.

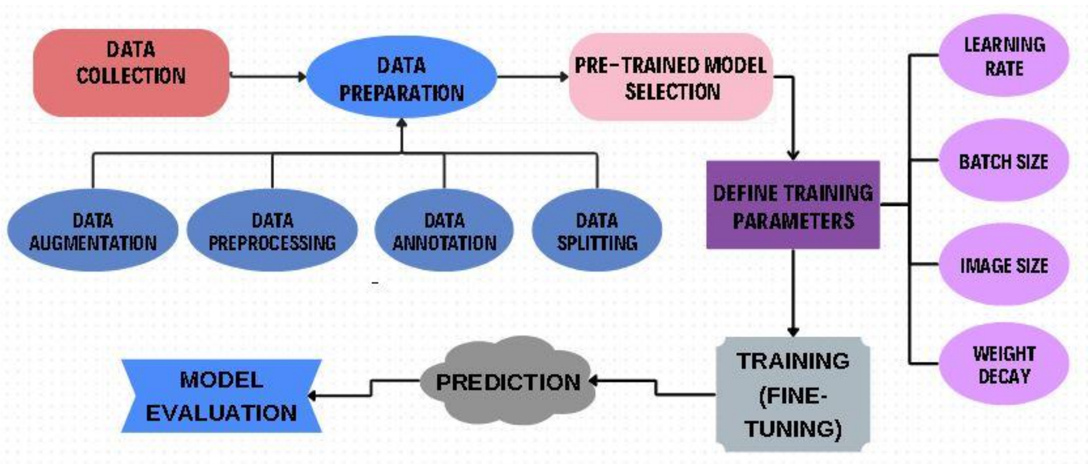

Figure 27:Overall Architecture

We addressed underwater waste segmentation in this work. We suggest two methods for this, Mask RCNN and YOLOv8, both of them are based on instance segmentation (segmentation + object detection).

3. Data Presentation

We will present the preprocessing and annotation done on the public dataset that we used for our research.

3.1. Underwater-Debris dataset

The Sri Sai Ram Engineering College has released a dataset on roboflow that has 5127 photos of underwater waste categorized into 15 different categories. No preprocessing or augmentations were applied in the creation of this dataset. Figure 28 show the BibTeX of dataset in roboflow.[50]

```
@misc{
underwater_debris_dataset,
title = { Underwater_Debris Dataset },
type = { Open Source Dataset },
author = { Sri Sai Ram Engineering College },
howpublished = { \url{ https://universe.roboflow.com/sri-sai-ram-engineering-college-
hnimv/underwater_debris } },
url = { https://universe.roboflow.com/sri-sai-ram-engineering-college-hnimv/underwater_debris },
journal = { Roboflow Universe },
publisher = { Roboflow },
year = { 2023 },
month = { oct },
```

Figure 28 :dataset BibTeX [50]

3.2. Image Annotation

This task consists of creating the label part of the dataset. Our dataset will be used for underwater waste segmentation.

- **Mask RCNN Annotation**

An annotation based on bbox is necessary for DL-based models created instance segmentation; in other words, the mask is represented as a polygon. VGG Annotator is the best annotation tool for instance segmentation; figure 29 provides an example of annotation using a JSON file for the picture that looks like this:

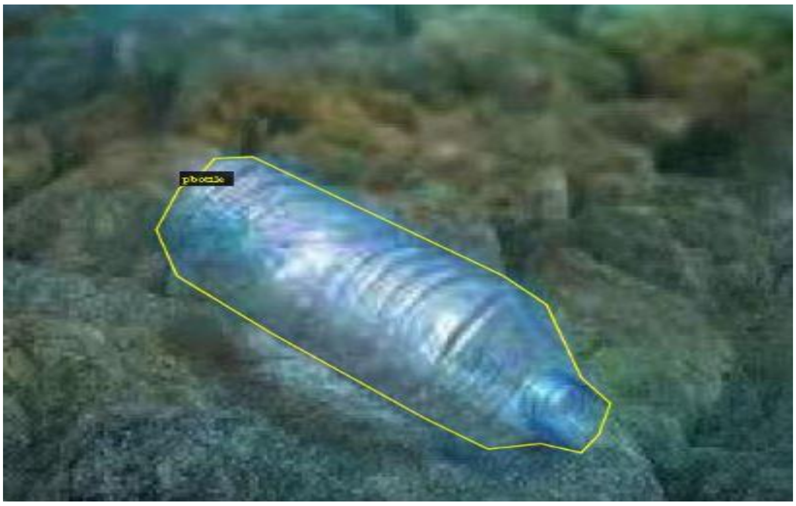

Figure 29: Mask RCNN annotation using VGG annotator

underwater waste data will be stored in a JSON file, as seen in figure 30:


```
{
    "cellphone (1).jpg33056":
{
    "filename":"cellphone (1).jpg",
"size":33056,
"regions":
[{
    "shape_attributes":
    {
        "name":"polygon",
        "all_points_x":[171,155,251,253],
        "all_points_y":[49,286,295,60]},
        "region_attributes":
        {
            "object":"electronics"}}
            ,{
                "shape_attributes":
                {"name":"polygon",
                "all_points_x":[147,66,82,155],
                "all_points_y":[287,413,412,297]},
                "region_attributes":{"object":"metal"}}
```


Figure 30: VGG annotation

✓ All hand annotated images = 1543 images

The attributes of the JSON file:

- ✓ shape attributes: relate to the characteristics that define a specific shape.
- ✓ all_points_x: refers to all of the points that define a shape's x-coordinates.
- ✓ all_points_y: refers to all of the points that define a shape's y-coordinates.
- ✓ region attributes: For descriptions, use text labels to define the object.

- **Yolov8 Annotation**

the mask is represented as a polygon. Roboflow Annotate is the best annotation tool for instance segmentation; figure 31 provides an example of annotation using a JSON COCO file for the picture that looks like this:

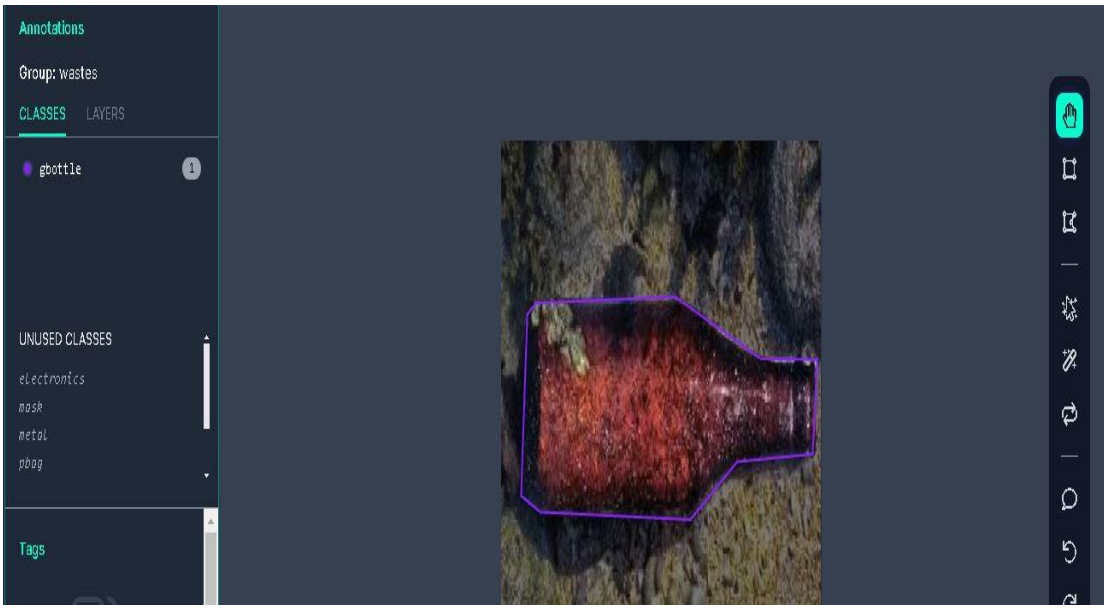

Figure 31: YOLOv8 annotation using Roboflow Annotate

Figure 32 illustrates how underwater waste data would be kept in a JSON file:


```
{"id":7,"image_id":1,
"category_id":2,
"bbox":[3,274,121.538,66.154],
"area":8040.237,
"segmentation":
[[33.846,273.846,32.308,290.769,4.615,287.692,3.077,316.923,27.692,316.923,40,335.385,124.615,340,121.538,296.923,33.846,273.846]
],"iscrowd":0},
```


Figure 32:Roboflow Annotate annotation

The attributes of the JSON file:

- ✓ Id: a special identification within the dataset for this particular annotation
- ✓ image_id: relates this annotation to a particular dataset image.
- ✓ category_id: refers to the category to which this object belongs.
- ✓ Bbox: This generates a bounding box around the images's object.
- ✓ Area: This is an indication of the bounding box's calculated area.
- ✓ Segmentation: This gives the object a more accurate segmentation mask.
- ✓ Iscrowd: This is a flag that indicates whether the item is in a crowd. It is a single, distinct object if the value is 0.

3.3. Data Preprocessing

The process of preparing raw data and fitting it to a deep learning model is known as data preprocessing. This is a crucial first step in creating a deep learning model.

Certain treatments are required to increase the effectiveness of our machine learning models due to the heterogeneity of the dataset data.

- **Mask RCNN**
  - ✓ **Mask conversion**

```
mask = np.zeros([info["height"], info["width"], len(info["polygons"])],
                dtype=np.uint8)
for i, p in enumerate(info["polygons"]):
    # Get indexes of pixels inside the polygon and set them to 1
    rr, cc = skimage.draw.polygon(p['all_points_y'], p['all_points_x'])
    rr = np.clip(rr, 0, info["height"]-1)
    cc = np.clip(cc, 0, info["width"]-1)
```

Figure 33:loading mask in Mask RCNN

✓ **Resize**

Before we can feed our DL model, we need to make sure that the images all the same size. Therefore, we defined standard size 512×512.

- **YOLOv8**
  - ✓ **Mask Conversion**

. We will use the Roboflow library to load the dataset and annotations. Then mask will be converted via Roboflow.

```
import roboflow as rf
# Load the dataset
project = rf.workspace('maskrcnn').project('wastes')
dataset = project.version(1).download('yolov8')
```

Figure 34: Roboflow load data

✓ **Resizing and normalisation**

We've applied auto-orient for our images, 640×640 resizing and auto-adjusting contrast.

3.4. Data augmentation

The table below presents data augmentation applied in our models.

Table 5: Data augmentation

Model	Mask RCNN	YOLOv8
Augmentation	No data augmentation applied.	90° Rotation applied

3.5. Data splitting

we devised the dataset, 70% and for training and 30% for validation respectively for the Mask RCNN model, and concerning YOLOv8 80% for training, 14% for validation and 6% for testing:

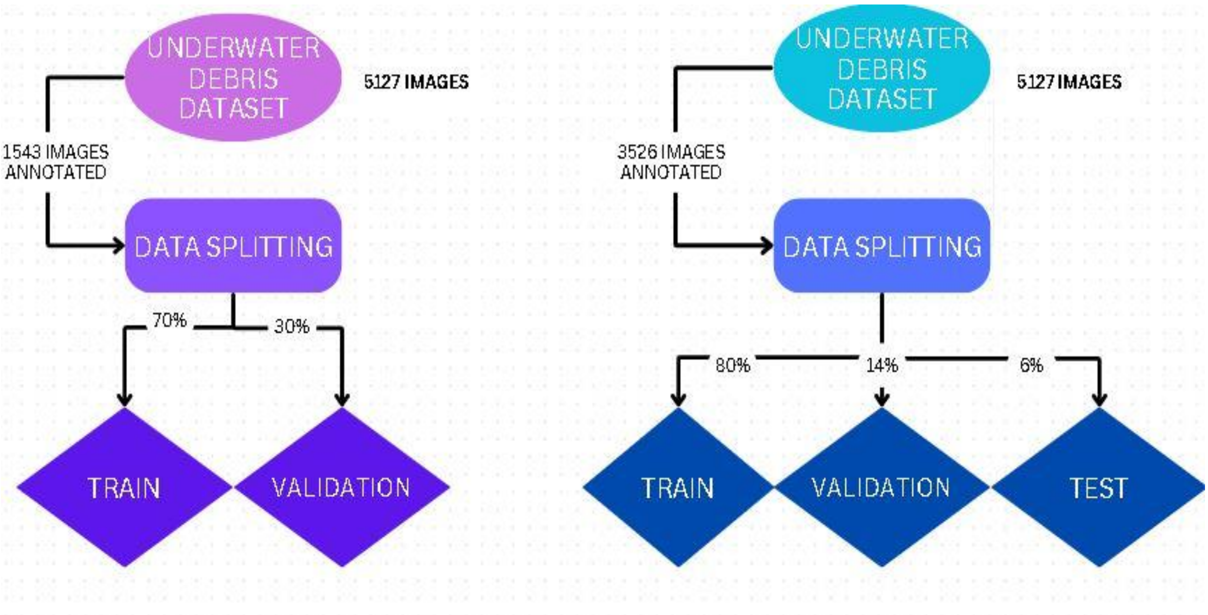

Figure 35:Data Splitting

4. DL Models Proposed for Segmentation

In order to address the concern of underwater waste segmentation, we propose two Deep Learning models based on instance segmentation on our dataset, MASK RCNN and YOLOv8.

4.1.Our model Mask RCNN

Remember that chapter 2 introduced the Faster R-CNN principle, which provides two classes—one for labeling boxes and another for candidates—for each item. The Mask R-CNN, in fact, suggests the mask of the item of interest and extends the Faster R-CNN with a third branch. The object's location and segmentation on the images are made by this model. In order to categorize each pixel into a predetermined set of categories, the objective is to locate each item using bounding boxes and segmentation.

III. Chapter 3: Underwater Waste Segmentation, Design, Implementation, and Experimentation

Our proposed method's operation mechanism can be summarized following the two steps below:

- The very first step consists of two basic networks that use the Resnet 101 and the RPN. These networks perform one task per image in order to propose a set of regions, the proposals being regions of the map of entities that include the objects.
- During the second stage, the network provides for the box and the object class of each of the regions proposed by Resnet 101 and the RPN in the first stage. All regions can be represented by a different size, while fully connected layers in networks still require a size vector fixed to make the prediction. This is why the size of the proposed regions will be fixed through ROI pools.

The following figure depicts our Mask RCNN model's architecture:

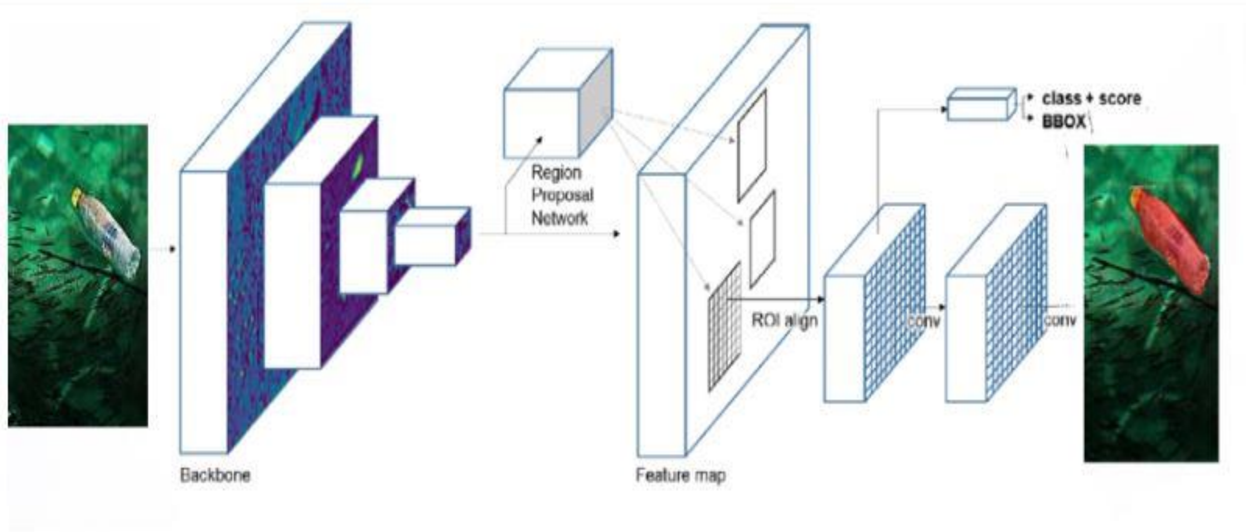

Figure 36: Our mask RCNN model architecture

4.2.YOLOv8 Architecture

The head, neck, and backbone make up the architecture. A pre-trained Convolutional Neural Network (CNN) serves as the main component and is used to extract low-, medium-, and high-level feature maps from an input image. Using path aggregation blocks such as the Feature Pyramid Network (FPN), the neck combines these feature maps. It transfers them to the brain, where it makes predictions about bounding boxes and classes objects. One-stage or dense prediction models, like YOLO or Single-shot Detector (SSD), can make up the head. As an alternative, it may include sparse or two-stage prediction algorithms, such as the R-CNN series. [49]

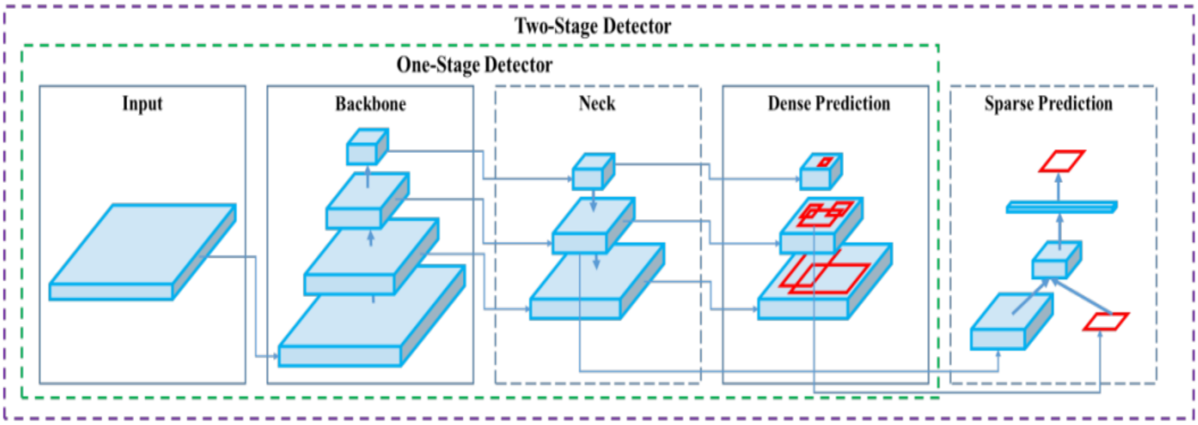

Figure 37: YOLOv8 Architecture [49]

We didn't change anything in the architecture of YOLOv8, we did use Roboflow to train our model and it doesn't change anything in the original architecture of YOLOv8.

5. Evaluation Metrics

We employed multiple metrics to evaluate our models:

- **precision:** The number of true positives (TP) divided by the total number of true positives plus false positives (FP) is known as precision (P).

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

- **Recall:** Recall (R) is defined as the number of true positives (TP) over the number of true positives plus the number of false negatives (FN).

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

- **Mean Average Precision (mAP):** Over recall levels ranging from 0 to 1, the average precision (AP) value is computed.[51]

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{the AP of class } k$$

$$n = \text{the number of classes}$$

Figure 38: how to calculate maP [51]

6. Implémentation and Expérimentation

The working environment, programming languages, libraries, and framework will all be defined in this section first. Subsequently, we will showcase the results of each approach by evaluating the performance of the models employed.

6.1.Environment and Work Tools

6.1.1. Jupiter notebook

Jupyter is a broad project that encompasses a wide range of software products and resources, such as the well-known web-based notebook creation and editing tools, Jupyter Notebook and JupyterLab. The goal of the Jupyter project and its affiliated projects is to provide computational notebooks with tools (and standards) for interactive computing.[52]

6.1.2. Python

Python is a high-level, interpreted, object-oriented language with dynamic semantics. Its dynamic typing and dynamic binding, along with its high-level built-in data structures, make it an appealing language for Rapid Application Development and for usage as a scripting or glue language to join existing components. Because of its straightforward, basic syntax, Python promotes readability, which lowers software maintenance costs. Python's support for packages and modules promotes code reuse and program modularity. The large standard library and the Python interpreter are freely distributable and accessible for free on all major platforms in source or binary form.[53]

the Python libraries most used in our approach are:

6.1.3. TensorFlow

Google created the open-source TensorFlow library, mostly for use in deep learning applications. It is also compatible with conventional machine learning. Originally designed for large-scale numerical computations, TensorFlow does not take deep learning into consideration. Nevertheless, Google made it publicly available after discovering that it was also highly helpful for deep learning research.[54]

6.1.4. Keras

Google created Keras, a high-level deep learning API for building neural networks. It is used to simplify the implementation of neural networks and is written in Python. Multiple backend neural network computation is also supported. Because Keras offers a high-level, abstract Python frontend with the flexibility to use numerous back-ends for computation, it is comparatively simple to understand and use. As a result, Keras is far more user-friendly for beginners yet slower than other deep learning frameworks.[55]

6.1.5. PyTorch

The Python programming language and the Torch library serve as the foundation for the open-source ML framework PyTorch. Torch is an open-source ML package built in the Lua programming language that is used to build deep neural networks. It's a favored platform for research in deep learning. The purpose of the framework is to accelerate the research prototyping to deployment process.[56]

6.1.6. Scikit-learn

The most reliable and practical Python machine learning library is called Scikit-learn, or Sklearn. Through a Python consistency interface, it offers a range of effective tools for statistical modeling and machine learning, including as regression, clustering, classification, and dimensionality reduction.[57]

6.1.7. NumPy

A Python library called NumPy is used to work with arrays. It also includes functions for working with matrices, the Fourier transform, and linear algebra. In the year 2005, Travis Oliphant founded NumPy. You are free to use it as it is an open-source project. Numerical Python is referred to as NumPy.[58]

6.1.8. Matplotlib

Matplotlib is a low-level Python graph charting toolkit that functions as a tool for visualization. The creator of Matplotlib is John D. Hunter. Matplotlib is available for free and without restriction. The majority of Matplotlib is written in Python; but, for platform compatibility, a small portion is also written in C, Objective-C, and JavaScript.[59]

6.2.Implementation Steps of DL Models

6.2.1. Loading data:

In this step, we have performed a data load, here is the code to perform this operation.

- **Mask RCNN**

```
assert subset in ["train", "val"]
dataset_dir = os.path.join(dataset_dir, subset)
annotations = json.load(open(os.path.join(dataset_dir, "via_region_data.json")))
annotations = list(annotations.values())  # don't need the dict keys
annotations = [a for a in annotations if a['regions']]
for a in annotations:
    if type(a['regions']) is dict:
        polygons = [r['shape_attributes'] for r in a['regions'].values()]
        objects = [s['region_attributes']['object'] for s in a['regions'].values() ]
    else:
        polygons = [r['shape_attributes'] for r in a['regions']]
        objects = [s['region_attributes']['object'] for s in a['regions'] ]
    name_dict = {"mask": 1,"electronics": 2,"gbottle": 3,"metal": 4,"pbag": 5,"pbottle": 6,"tire": 7,"waste": 8}
    num_ids = [name_dict[a] for a in objects]
    image_path = os.path.join(dataset_dir, a['filename'])
    image = skimage.io.imread(image_path)
    height, width = image.shape[:2]
    self.add_image(
        "object",
        image_id=a['filename'],  # use file name as a unique image id
        path=image_path,
        width=width, height=height,
        polygons=polygons,
        num_ids=num_ids)
```

Figure 39: data loading Mask RCNN

- **YOLOv8**

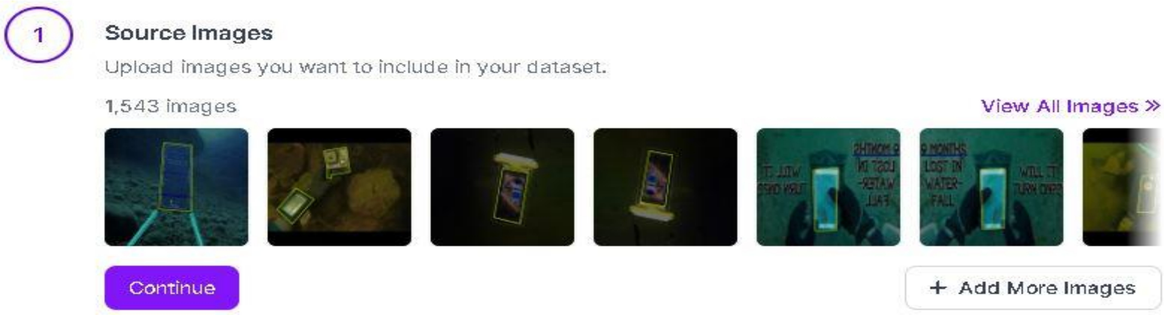

Figure 40 : YOLOv8 loading data in Roboflow

6.2.2. Model Training Phase

We use the train function to train Mask RCNN and YOLOv8

- **Mask RCNN**

We must first download the pre-trained model for Mask RCNN (“mask_rcnn_coco.h5”), we will fine-tune our model with the pre-trained model. We have used learning rate of 0.001, batch size=4, weights decay= 0.00005 and image size = 512*512.

```
model.train(dataset_train, dataset_val,
            learning_rate=config.LEARNING_RATE,
            epochs=100,
            layers='all')
```

Figure 41: Mask RCNN training

- **YOLOv8**

We also begin from weights of coco pre-trained model than in training we will update weights and bias to get our model.

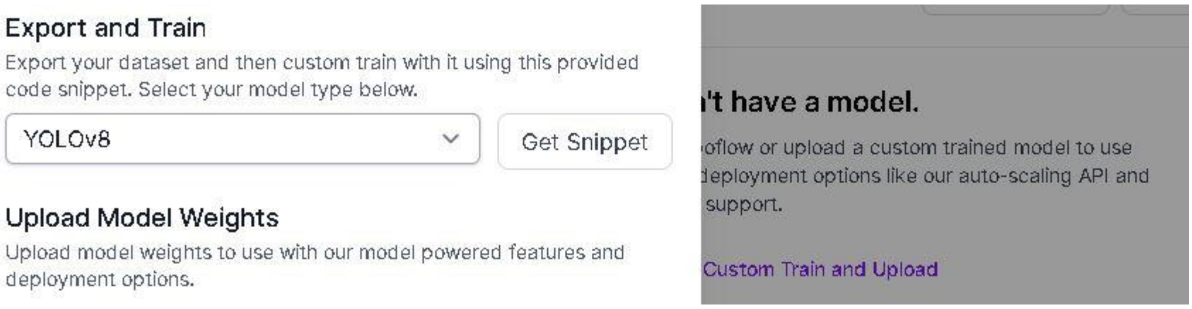

Figure 42: YOLOv8 training in Roboflow

7. Experimentation Results

7.1.Performance of Mask RCNN Segmentation Model

The following graphs represent, "Loss" of training, "Mask loss" , "Class loss" and "bounding box loss".

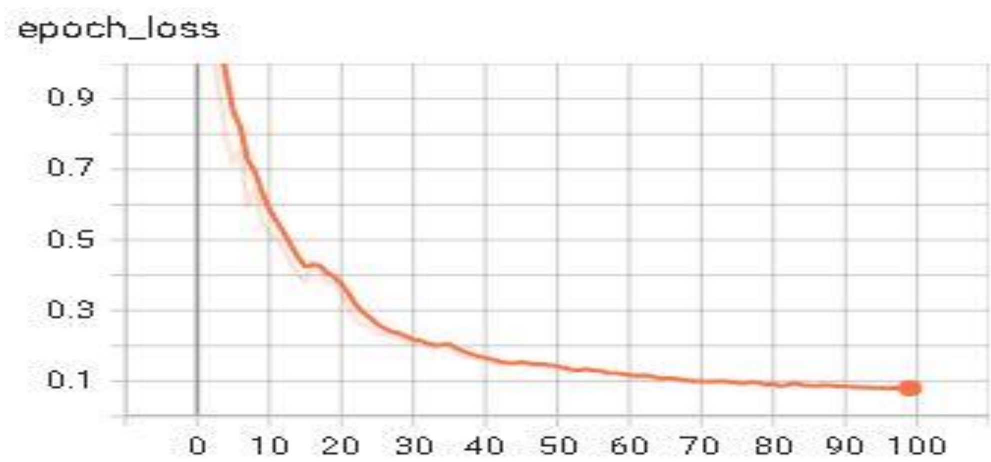

Figure 43: Mask RCNN training loss

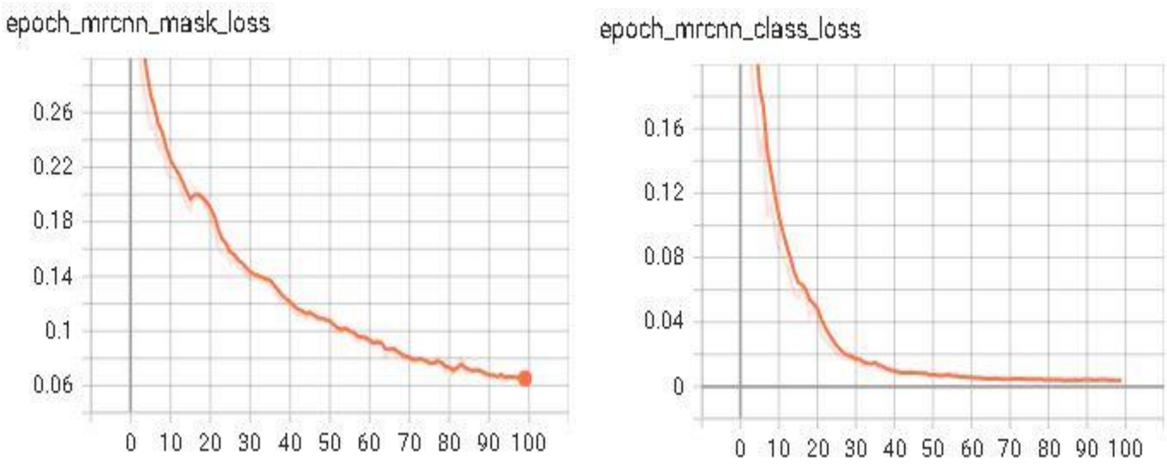

Figure 44:Mask RCNN Mask and Class loss

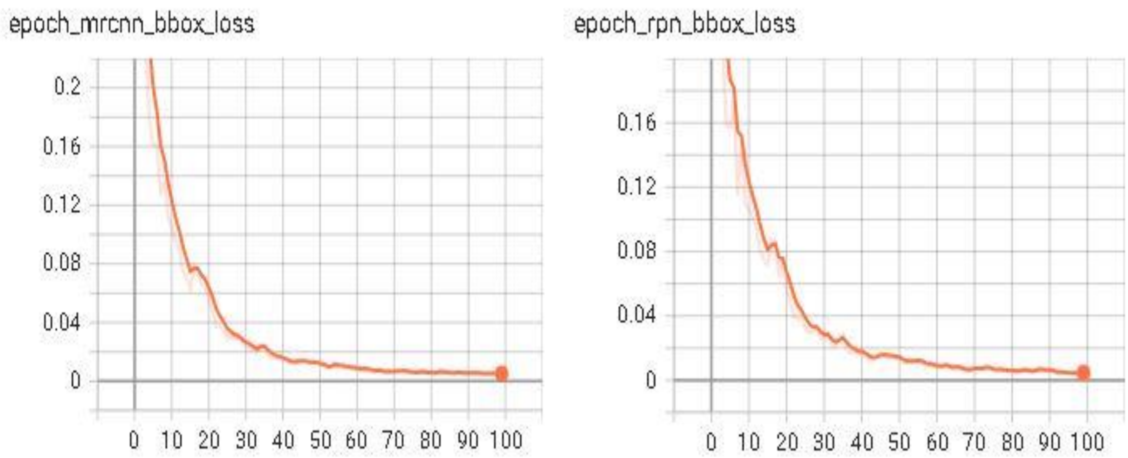

Figure 45: bounding box loss

In order to train our Mask RCNN model with different parameters, the following table represents the best results achieved:

Table 6 : Mask RCNN performances

Metrics / Model	LOSS	maP (%)
Mask RCNN	0.07	83.3%

7.2.Performance of YOLOv8 Segmentation Model

The “train and validation loss”, “Mean precision," "precision," and "recall" graphs shown below:

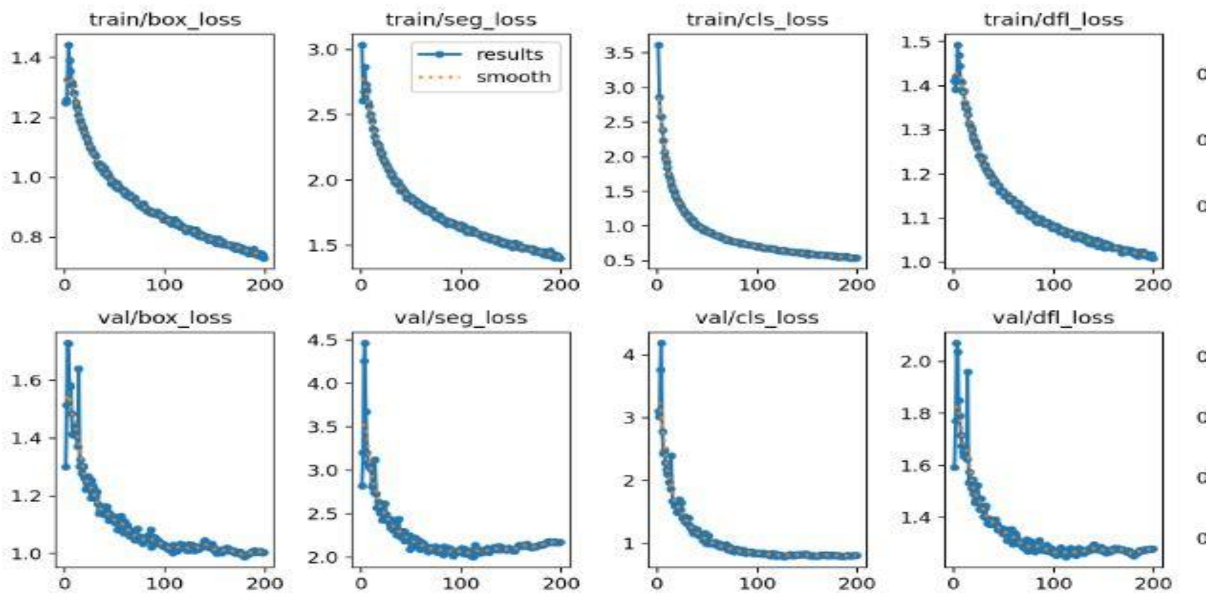

Figure 46: Yolov8 loss

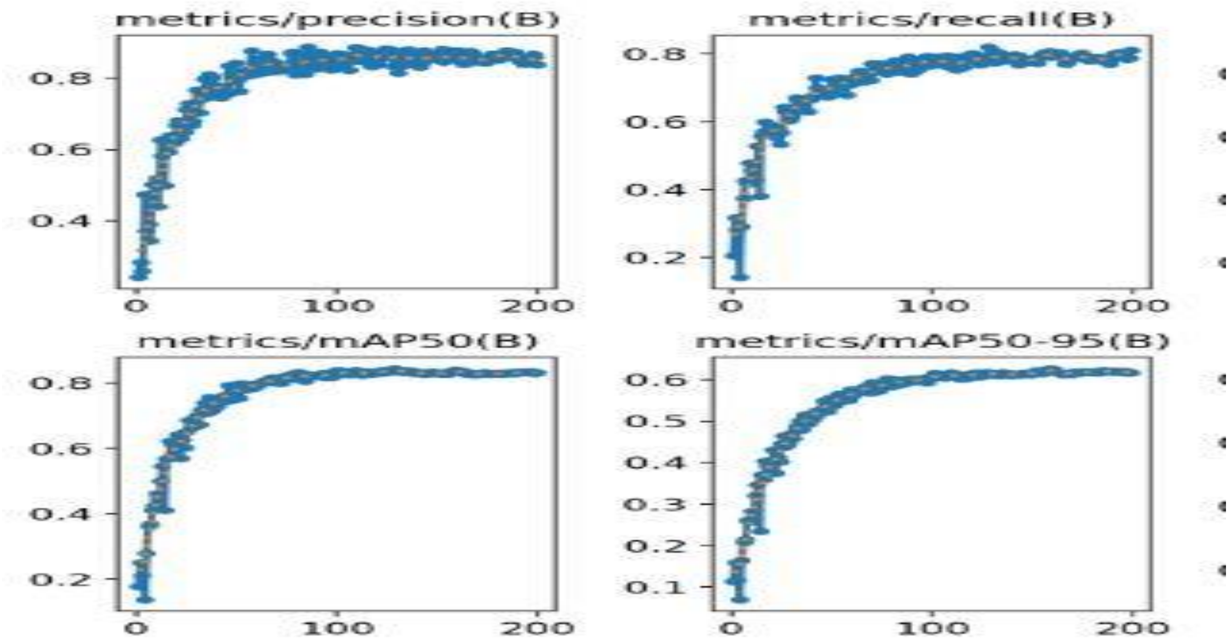

Figure 47: YOLOv8 precision, recall and maP metrics

In order to train our YOLOv8 model with different parameters, the following figure represents the best results achieved:

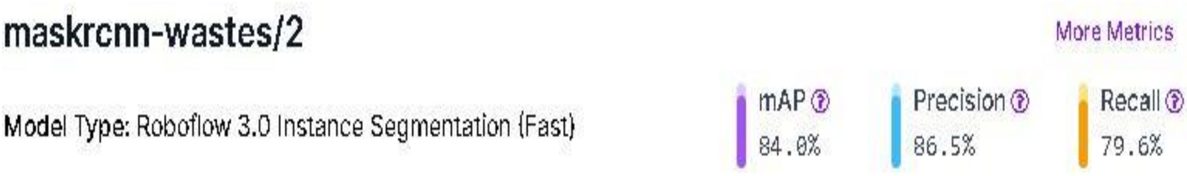

Figure 48: YOLOv8 performance

7.3. Comparison Between Our Models and Models Mentioned in the State-of-the-Art

Table 6 presents a comparative analysis of the Mask RCNN model and the state-of-the-art work, specifically focusing on underwater waste segmentation and object detection. The results indicate that our model performs better than the most recent published work on underwater waste segmentation.

Table 7 :comparative analysis of the Mask RCNN model and the state-of-the-art Works

Model	Article	Dataset	Classes	maP (%)
RetinaNet (Resnet50 backbone & amp; FPN)	[43]	Custom Dataset	4	68.7
Faster RCNN	[41]	Custom Dataset	3	81
YOLOv3	[42]	Custom dataset	7	77.2
Mask R-CNN	[47]	CleanSea	17	63.5
Our Model Mask RCNN		Underwater_debris	8	83.3

We compared the approaches based on instance segmentation and state-of-the-art underwater waste detection, with our YOLOv8 model which is based on segmentation of instances:

Table 8:comparative analysis of the YOLOv8 model and the state-of-the-art Works

Model	Article	Dataset	Classes	maP (%)
Tiny-YOLO	[41]	Custom Dataset	3	82.3
ResNet50-YOLOv3	[44]	Custom Dataset	7	83.4
Mask RCNN	[45]	Trashcan	16	65.2
YOLOv5s	[46]	River-debris dataset	1	83
Our Model YOLOv8		Underwater-Debris	8	84

7.4. Advantages and Disadvantages of YOLOv8 and Mask R-CNN

The advantages and disadvantages of each object detection model must be carefully considered before selecting one. Let's examine the advantages and disadvantages of Mask RCNN and YOLOv8 for object detection. Table 9 below shows advantages and disadvtanges of yolov8 and Mask RCNN.[60]

Table 9: Advantages and Disadvantages of YOLOv8 and Mask R-CNN[60]

Model	Advantages	Disadvantages
Mask RCNN	**Precise instance segmentation:** Mask R-CNN is a great option for exact instance segmentation in applications like medical and autonomous driving analysis since it excels at precisely assigning masks to images. **Robustness:** Mask R-CNN captures comprehensive spatial information for dependable identification and segmentation, enabling it to handle occlusions and complex scenes with effectiveness.	**Computationally intensive:** When working with huge datasets or real-time requirements, the complex two-stage Mask R-CNN, which frequently requires more processing resources, can slow down inference times. **Complex implementation:** Because Mask R-CNN's training and design are more complex than YOLOv8, implementation and fine-tuning will be more difficult, which will take longer to complete.
YOLOv8	**Real-time performance**: Because of its reputation for real-time identification and picture analysis, YOLOv8 is perfect for applications that need quick reactions. **Efficiency:** By directly detecting objects, cutting down on inference times, and optimizing resource use, YOLOv8 makes effective use of resources. **Accuracy:** YOLOv8 is a quick and precise object detection system that performs well in situations with multiple objects.	**Lower localization precision:** For real-time applications, YOLOv8's performance usually exceeds this drawback, even though its design may lead to slightly less accurate object bounding when compared to two-stage models like Mask R-CNN. **Less suitability for instance segmentation:** While YOLOv8 is an excellent choice for tasks requiring precise instance segmentation, its absence of direct per-pixel segmentation makes it less useful for object classification and localization.

Analyze the specific demands of our project effectively while comparing Mask R-CNN with YOLOv8 for object detecting tasks.[60]

7.5.Predictions and Testing

7.5.1. Mask RCNN predictions

Figure 48 below shows Mask RCNN predictions on test images

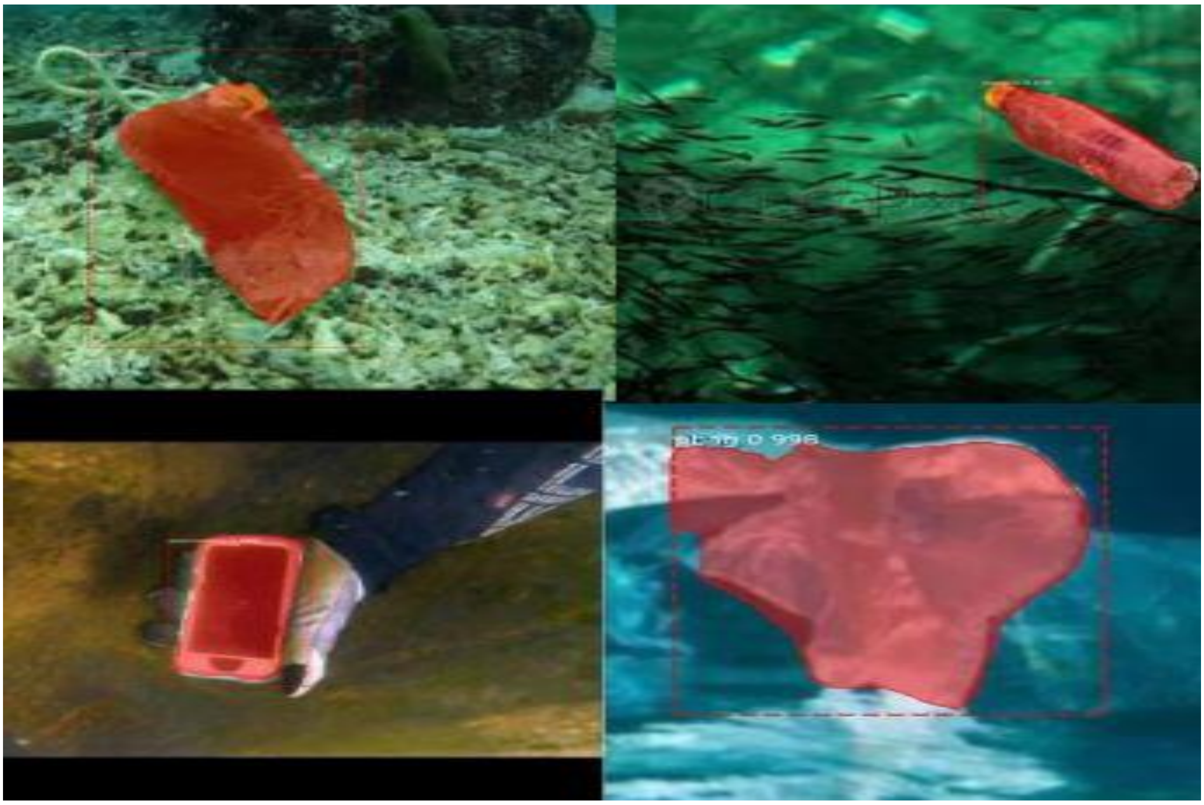

Figure 49 : Mask RCNN predictions

7.5.2. YOLOv8

Predictions of our Model YOLOv8 are shown in figure 49 below

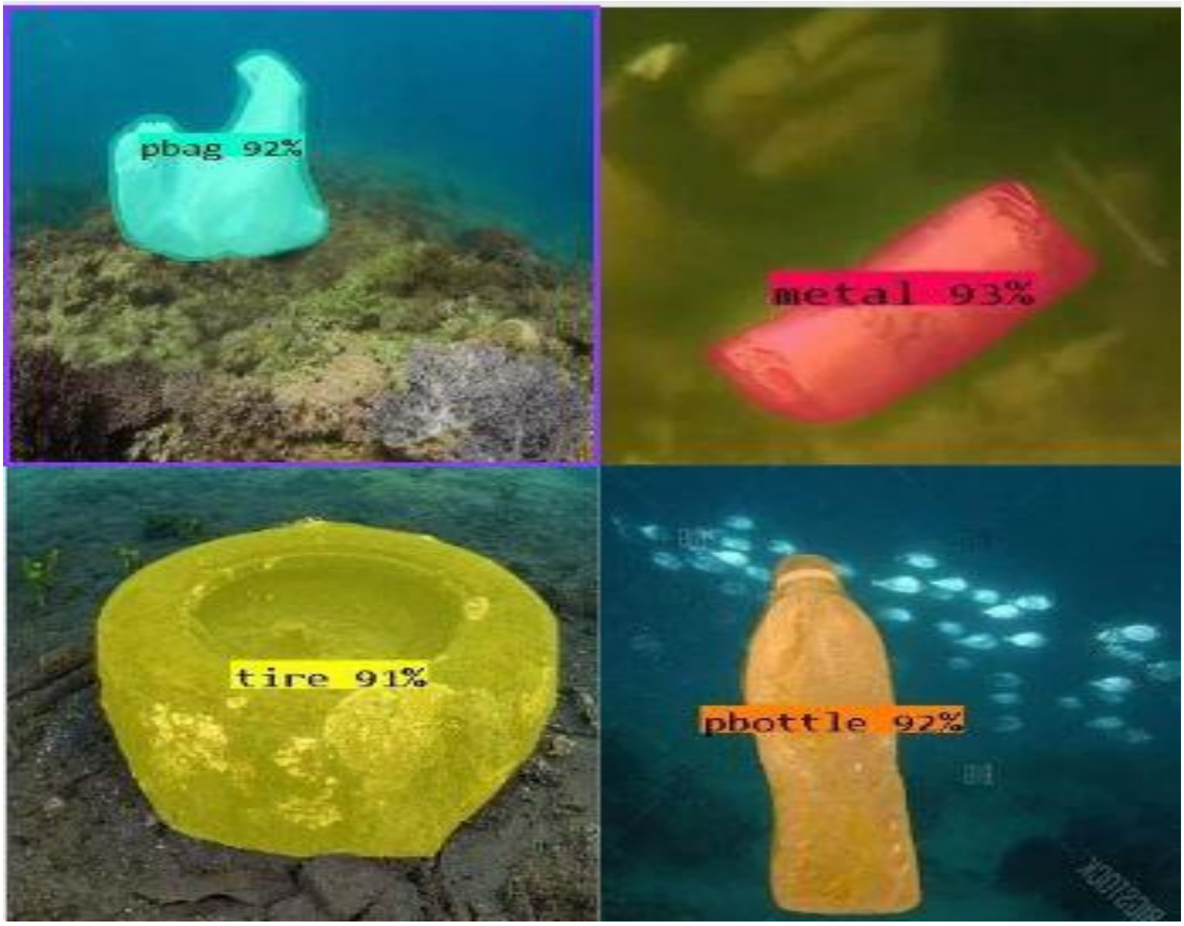

Figure 50: YOLOv8 predictions

8. Conclusion

Our contribution to the field of deep learning-based intelligent underwater waste segmentation is presented in this chapter. To increase the accuracy of the underwater waste segmentation procedure, we suggested two deep learning-based algorithms. We were able to achieve the highest percentage in this field—mAP=83.3 for Mask RCNN and maP=84 for YOLOv8—by using an algorithm that used Mask RCNN with 100 learning epochs and YOLOv8 with 200 learning epochs. This indicates that our dataset for training, validation, and testing are very good.

General Conclusion

General Conclusion

In order to reduce the environmental impact of inappropriate underwater waste disposal, we offer an automated system based on deep learning and traditional methodologies that aim to correctly segment and separate waste into categories: mask, glass bottles, plastic bags, plastic bottles, tire, metal, electronics and waste. The results suggest that the approach Mask RCNN is an effective strategy for this problem, with the best scenario reaching maP=83.3% using 100 epochs and an average precision that exceeds 84% with the YOLOv8.

Since this proportion is more than every other one in the same field of study, our stated objective of recognizing underwater wastes has been accomplished.

Nevertheless, RCNN approaches need more powerful computers and are computationally more expensive than other methods. Furthermore, future research will look into how different techniques like augmentation and having higher-quality images might increase the accuracy of RCNN systems. In addition, the RCNN tends to perform better in situations where there is a greater amount of data.

the future prospects of our project

In future research, we will work to develop our models to:

- ❖ We want to work on a greater number of classes and enhance our segmentation on various underwater waste types.
- ❖ In order to have the best model of underwater wastes recognition, we will continue to develop on several approaches and compare them with our Mask RCNN and YOLOv8 Results.
- ❖ be able to work in the robotics field and integrate it into the procedure for sorting underwater waste.

Bibliography

[1] D. Hoornweg and P. Bhada-Tata, What a Waste: A Global Review of Solid Waste Management, World Bank, Washington, DC, USA, 2012: Retrieved 19 april 2024

[2] A Brief History Of Waste Management : https://www.commercialzone.com/a-brief-history-of-waste-management/ Retrieved 19 april 2024

[3] Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal (187 signatory countries as of February 2019). Available at: http://www.basel.int/TheConvention/Overview/TextoftheConvention/tabid/1275/Default.aspx Retrieved 19 april 2024

[4] Decision of the Council on the Control of Transboundary Movements of Wastes Destined for Recovery Operations, OECD/LEGAL/0266. Available at: https://legalinstruments.oecd.org/public/doc/221/221.en.pdf Retrieved 19 april 2024

[5] Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32008L0098 Retrieved 19 April 2024

[6] Giusti, L. (1 August 2009). "A review of waste management practices and their impact on human health". Waste Management. 29 (8): 2227–2239. Bibcode:2009WaMan..29.2227G. doi:10.1016/j.wasman.2009.03.028. ISSN 0956-053X. PMID 19401266. Archived from the original on 25 November 2018. Retrieved 20 April 2024

[7] "Wastes". U.S. Environmental Protection Agency. 2 November 2017. Retrieved 19 August 2023. Retrieved 20 April 2024

[8] "Solid Waste Management". World Bank. Archived from the original on 30 September 2020. Retrieved 20 April 2024

[9] Glossary of environmental and waste management terms". Handbook of Solid Waste Management and Waste Minimization Technologies. Butterworth-Heinemann. 2003. pp. 337–465. doi:10.1016/B978-075067507-9/50010-3. ISBN 9780750675079.

[10] "Climate Change 2022: Mitigation of Climate Change". www.ipcc.ch. Retrieved 21 April 2024.

[11] Gollakota, Anjani R. K.; Gautam, Sneha; Shu, Chi-Min (1 May 2020). "Inconsistencies of e-waste management in developing nations – Facts and plausible solutions". Journal of Environmental Management. 261: 110234. doi:10.1016/j.jenvman.2020.110234. ISSN 0301-4797. PMID 32148304. S2CID 212641354. Archived from the original on 20 September 2021. Retrieved 21 April 2024

[12]Sankar, Ajith (2015). Environmental Management. New Delhi: Oxford University Press: https://archive.org/details/environmentalman0000sank Retrieved 21 April 2024

Bibliography

[13] "14.6: Waste Management". Medicine LibreTexts. 30 August 2021. Retrieved 22 April 2024

[14]] Guidelines for National Waste Management Strategies Moving from Challenges to Opportunities (PDF). United Nations Environmental Programme. 2013. Archived from the original (PDF) on 4 March 2016. Retrieved 21 April 2024

[15] Reduce, Reuse, Recycle: Alternatives for Waste Management: https://pubs.nmsu.edu/_g/G314/index.html Retrieved 22 April 2024

[16] Techno-economic assessment of energy generation through municipal solid waste: a case study for small/medium size districts in Pakistan: Retrieved 22 April 2024 : https://www.researchgate.net/publication/347711215_Techno-economic_assessment_of_energy_generation_through_municipal_solid_waste_a_case_study_for_smallmedium_size_districts_in_Pakistan

[17] The Ultimate Guide to Smart Waste Managementhttps://nordsense.com/the-ultimate-guide-to-smart-waste-management/

[18]Marine Debris : https://education.nationalgeographic.org/resource/marine-debris/

[19] Marine Waste—Sources, Fate, Risks, Challenges and Research Needs: https://www.mdpi.com/1660-4601/18/2/433

[20] IMPACTS OF PLASTIC POLLUTION IN THE OCEANS ON MARINE SPECIES, BIODIVERSITY AND ECOSYSTEMS:https://wwfint.awsassets.panda.org/downloads/wwf_impacts_of_plastic_pollution_on_biodiversity.pdf

[21] Ocean Trash: 5.25 Trillion Pieces and Counting, but Big Questions Remain: https://education.nationalgeographic.org/resource/ocean-trash-525-trillion-pieces-and-counting-big-questions-remain/

[22] Cleaning technology for marine debris: A review of current status and evaluation

https://link.springer.com/article/10.1007/s13762-022-04373-8

[23] Valorization of Marine Waste: Use of Industrial By-Products and Beach Wrack Towards the Production of High Added-Value Products : https://www.frontiersin.org/articles/10.3389/fmars.2021.723333/full

[24] What is artificial intelligence (AI)? : https://www.ibm.com/topics/artificial-intelligence Retrieved 27 April 2024

[25] What Is Artificial Intelligence? : https://www.fool.com/terms/a/artificial-intelligence/ Retrieved 27 April 2024

[26] What Is Machine Learning? Definition, Types, Applications, and Trends for 2022 : https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/ Retrieved 28 April 2024

[27] What is deep learning? https://www.ibm.com/topics/deep-learning Retrieved 28 April 2024

[28] A Brief History of Deep Learning By Keith D. Foote on February 4, 2022 : https://www.dataversity.net/brief-history-deep-learning/ Retrieved 28 April 2024

[29] Convolutional Neural Network: Benefits, Types, and Applications : https://datagen.tech/guides/computer-vision/cnn-convolutional-neural-network/ Retrieved 29 April 2024

[30] Deep Learning for Computer Vision : https://www.run.ai/guides/deep-learning-for-computer-vision Retrieved 29 April 2024

[31] Deep Learning For Computer Vision Tasks: A review: https://www.researchgate.net/publication/324472206_Deep_Learning_For_Computer_Vision_Tasks_A_review Retrieved 29 April 2024

[32] Computer Vision Tasks (Comprehensive 2024 Guide): https://viso.ai/deep-learning/computer-vision-tasks/ Retrieved 29 April 2024

[33] What is R-CNN?: https://blog.roboflow.com/what-is-r-cnn/ Retrieved 29 April 2024

[34] Faster R-CNN | ML: https://www.geeksforgeeks.org/faster-r-cnn-ml/ Retrieved 29 April 2024

[35] YOLO Object Detection Explained : https://www.datacamp.com/blog/yolo-object-detection-explained Retrieved 29 April 2024

[36] U-Net: https://paperswithcode.com/method/u-net#:~:text=**U%2DNet**%20is,architecture%20of%20a%20convolutional%20network. Retrieved 29 April 2024

[37] What is Mask R-CNN? The Ultimate Guide : https://blog.roboflow.com/mask-rcnn/ Retrieved 29 April 2024

[38] What is Object Detection? The Ultimate Guide.: https://blog.roboflow.com/object-detection/ Retrieved 29 April 2024

[39] Li Liu1,2 · Wanli Ouyang3 · Xiaogang Wang4 · Paul Fieguth5 · Jie Chen2 · Xinwang Liu1 · Matti Pietikäinen2 , Deep Learning for Generic Object Detection: A Survey / Accepted: 26 September 2019 Retrieved 29 April 2024

Bibliography

[40] The Difference Between Computer Vision and Human Vision : https://visionaisuite.net/blog/the-difference-between-computer-vision-and-human-vision Retrieved 29 April 2024

[41] Michael Fulton1 , Jungseok Hong2 , Md Jahidul Islam3 , Junaed Sattar4 : Robotic Detection of Marine Litter Using Deep Visual Detection Models∗ September 24, 2018 Retrieved 04 May 2024

[42] Jun-ichiro Watanabe,* Yang Shao, and Naoto Miura , Underwater and airborne monitoring of marine ecosystems and debris Oct. 24, 2019 Retrieved 04 May 2024

[43] Harsh Panwar a , P.K. Gupta a , Mohammad Khubeb Siddiqui b,* , Ruben Morales-Menendez b , Prakhar Bhardwaj a , Sudhansh Sharma a , Iqbal H. Sarker c : AquaVision: Automating the detection of waste in water bodies using deep transfer learning 15 July 2020 Retrieved 04 May 2024

[44] Bing Xue, Baoxiang Huang, Member, IEEE, Ge Chen, Haitao Li, Weibo Wei: Deep-sea Debris Identification Using Deep Convolutional Neural Networks Retrieved 04 May 2024

[45] Hongjie Deng, Daji Ergu *, Fangyao Liu, Bo Ma and Ying Cai : An Embeddable Algorithm for Automatic Garbage Detection Based on Complex Marine Environment : 24 September 2021 Retrieved 04 May 2024

[46] Nisha Maharjan 1 , Hiroyuki Miyazaki 1,2 , Bipun Man Pati 3 , Matthew N. Dailey 1 , Sangam Shrestha 4 and Tai Nakamura : Detection of River Plastic Using UAV Sensor Data and Deep Learning 25 June 2022 Retrieved 04 May 2024

[47] Alejandro Sánchez-Ferrer, Jose J. Valero-Mas, Antonio Javier Gallego∗ , Jorge Calvo-Zaragoza: An experimental study on marine debris location and recognition using object detection 26 December 2022 Retrieved 05 May 2024

[48] Palash Shinde, Marine Debris Segmentation using Capsule Network (SegCaps) 23/09/2021 Retrieved 05 May 2024

[49] https://universe.roboflow.com/sri-sai-ram-engineering-college-hnimv/underwater_debris

[50] A Guide to YOLOv8 in 2024: https://viso.ai/deep-learning/yolov8-guide/ Retrieved 09 May 2024

[51] Mean Average Precision (mAP) Explained: Everything You Need to Know : https://www.v7labs.com/blog/mean-average-precision#:~:text=The%20mAP%20is%20calculated%20by,over%20a%20number%20of%20classes.&text=The%20mAP%20incorporates%20the%20trade,metric%20for%20most%20detection%20applications. Retrieved 10 May 2024

[52] Project Jupyter Documentation: https://docs.jupyter.org/en/latest/ Retrieved 10 May 2024

[53] What is Python? Executive Summary: https://www.python.org/doc/essays/blurb/ Retrieved 10 May 2024

[54] What is TensorFlow? Deep Learning Libraries & Program Elements: https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-tensorflow Retrieved 10 May 2024

[55] What Is Keras: The Best Introductory Guide To Keras : https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-keras Retrieved 10 May 2024

[56] PyTorch : https://www.techtarget.com/searchenterpriseai/definition/PyTorch Retrieved 10 May 2024

[57] Scikit Learn – Introduction: https://www.tutorialspoint.com/scikit_learn/scikit_learn_introduction.htm Retrieved 10 May 2024

[58] NumPy Introduction: https://www.w3schools.com/python/numpy/numpy_intro.asp Retrieved 10 May 2024

[59] Matplotlib Tutorial: https://www.w3schools.com/python/matplotlib_intro.asp Retrieved 10 May 2024

[60] YOLOv8 vs Mask R-CNN: In-depth Analysis and Comparison : https://www.labelvisor.com/yolov8-vs-mask-r-cnn-in-depth-analysis-and-comparison/ Retrieved 10 May 2024