Journal of Theoretical and Applied Information Technology
15th June 2024. Vol.102. No. 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ELEVATING PLASTIC LITTER SURVEILLANCE: DRONES AND DEEP LEARNING FOR EFFICIENT DETECTION

MIMOUN YANDOUZI[1], MOUNIR GRARI[2], MOHAMMED BERRAHAL[3], AISSA KERKOUR ELMIAD[4], MOHAMMED BADAOUI[5], ALAEEDDINE BARKAOUI[6], YASSINE ZARHLOULE[6]

[1]LSI Lab., ENSAO, Mohammed First University, Oujda, Morocco

[2]MATSI Lab., ESTO, Mohammed First University, Oujda, Morocco

[3]LMC Lab., PFS, Cadi Ayyad University, Safi, Morocco

[4]LARI Lab., FSO, Mohammed First University, Oujda, Morocco

[5]LaMSD Lab., ESTO, Mohammed First University, Oujda, Morocco

[6]2GPMH Lab., FSO, Mohammed First University, Oujda, Morocco

E-mail: m.yandouzi@ump.ac.ma

ABSTRACT

In our contemporary world, environmental issues and the ever-present threat of plastic pollution endanger not only the health of our planet but also that of its inhabitants, underscoring the urgency of action. The need to closely monitor these destructive phenomena and develop effective detection systems is imperative to preserve our fragile ecosystem. Fortunately, the emergence of cutting-edge technologies has revolutionized our ability to monitor and detect environmental threats with unprecedented precision. The combined use of drones and artificial intelligence, particularly deep learning, yields promising results, leveraging drones' unique capabilities to cover vast areas and the power of deep learning to analyze collected data swiftly and accurately. Our study focuses on optimizing the utilization of drones and object detection algorithms through deep learning for effective detection and supervision of plastic litter. We will explore the performance of two major families of object detection models, namely single-pass and double-pass, using drone images captured at varying heights. The overarching objective is to identify the optimal performance-to-resource conditions, maximizing efficiency in our detection and supervision endeavors. This research is crucial in addressing the pressing environmental concerns posed by plastic pollution, offering innovative solutions to mitigate its impact and safeguard the health of our planet for future generations.

Keywords: *Plastic Litter, Object Detection, Deep Learning, Drones/UAV, Faster R-CNN, YOLO*

1. INTRODUCTION

Environmental problems facing our planet are becoming increasingly numerous and concerning. Ecosystem degradation, biodiversity loss, climate change, and pollution are among the most urgent and complex challenges humanity faces today. Among these challenges, the rise of plastic pollution stands out as one of the most pressing and intricate issues of our time [1].

Plastic, valued for its convenience, hygiene, and affordability, has gradually become a dominant part of our daily lives, displacing traditional materials such as paper, glass, wood, and metal in many applications. According to projections from the Organization for Economic Co-operation and Development (OECD), plastic production is expected to increase from 460 million tons in 2019 to 490 million tons in 2023. Estimates suggest it could reach up to 975 million tons by 2050 [2]. However, only 9% of the plastic produced is recycled [3]. The vast majority, 79%, ends up in landfills or disperses into the environment as litter. Over time, much of this litter unfortunately finds its way into our oceans, becoming an unwanted final repository. Plastics, which are resistant to degradation, break down into smaller microplastics. These micro fragments can lead to entanglement or ingestion by marine organisms, causing serious harm, even death. In addition to the environmental impact, the presence of microplastics in oceans also raises concerns about their potential to enter the food chain and impact human health. Studies have shown that microplastics can accumulate in seafood, raising questions about the long-term effects on human health.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The production figures for plastic are staggering, with millions of tons produced annually worldwide. However, the plastic recycling rate remains very low in comparison, exacerbating the magnitude of the problem. The consequences of this plastic pollution are severe and manifold: they affect the health of marine and terrestrial ecosystems, threaten biodiversity, pollute water resources, and have detrimental effects on human health [4].

In particular, plastic pollution poses a serious threat to marine life. Plastic litter fragments over time into microplastics, thus contaminating the oceans and waterways. Marine organisms ingest these microplastics, which can lead to severe health problems and even death. Moreover, toxic chemicals present in plastics can bioaccumulate in the food chain, eventually reaching humans who consume seafood [5].

In the face of this environmental crisis, the detection and monitoring of plastic pollution are crucial to implement effective prevention and cleanup strategies. Traditional techniques such as manual surveys and satellite tracking systems are often used, but they have limitations in terms of spatial coverage, cost, and resolution [6].

However, rapid advances in emerging technologies offer new perspectives in the fight against plastic pollution. The use of drones equipped with specialized sensors offers a promising solution to efficiently monitor areas affected by plastic pollution [7]. Drones can cover vast expanses of territory, including hard-to-reach areas, and provide detailed, real-time data on the extent and distribution of plastic litter [8].

Drones offer several advantages for environmental monitoring. Their agility and ability to fly at different altitudes enable them to collect precise data over large areas. Moreover, their flight autonomy and capacity to carry various payloads make them suitable for a range of surveillance missions [9].

Furthermore, advances in computer vision through deep learning have revolutionized the ability to analyze and interpret data collected by drones. Deep learning algorithms can be trained to automatically recognize plastic litter in images and videos, thereby facilitating the process of mapping and tracking plastic pollution with increased accuracy [10].

By combining the capabilities of drones with advances in deep learning, it is possible to create sophisticated and effective surveillance systems to combat plastic pollution. This integrated approach offers considerable potential to improve plastic litter management and protect our marine and terrestrial ecosystems [11]. By harnessing emerging technologies and collaborating globally, we can work together to preserve our environment for future generations [12].

Drones present a multitude of qualities that make them exceptionally well-suited for tackling the issue of plastic pollution, particularly in marine environments. Firstly, their aerial surveillance capabilities enable efficient monitoring of vast oceanic expanses and coastlines. Their versatility and ability to maneuver at various altitudes allow for coverage of remote and inaccessible areas, which are often difficult to reach using conventional methods [13]. Equipped with specialized sensors, drones can capture high-resolution imagery and data, providing comprehensive insights into the extent and distribution of plastic litter. Furthermore, their real-time data transmission capabilities facilitate prompt response and intervention in areas identified as plastic pollution hotspots [14]. Additionally, drones offer cost-effective and environmentally friendly alternatives to manned aircraft or boats, reducing operational expenses and minimizing carbon emissions. In summary, leveraging drones in the fight against plastic pollution offers a flexible and effective approach to monitoring, identifying, and mitigating this global environmental challenge [15].

In this paper, we delve into the study of various deep learning models for object detection, with the specific aim of identifying plastic litter on beaches. Our ultimate goal is to develop an effective solution that can be deployed by drones to map and identify this waste, thereby contributing to the preservation of our coastal ecosystems.

A major challenge in our work lies in determining which detection model is best suited to our use case, taking into account the specificities of plastic litter and beach environments. We evaluate different algorithms and deep learning architectures to find the one that offers the best performance in terms of precision and detection speed. The deep learning models under study include YOLOv6, YOLOv7, and YOLOv8, as well as Faster R-CNN with the backbones ResNet50, VGG16, and VGG19.

Another crucial dimension of our research involves determining the optimal drone flying height. By adjusting the flight altitude, we aim to optimize the resolution of captured images while ensuring adequate coverage of the target area. This requires a thorough analysis of trade-offs between spatial resolution, field of view, and operational efficiency. Our work is part of a broader project aimed at combating plastic pollution on beaches. Our technology demonstration site is located in Saidia, Morocco.

2. BACKGROUNDS

2.1 Computer Vision Techniques

Computer vision stands at the forefront of artificial intelligence, striving to equip machines with the ability to perceive and interpret visual cues from their surroundings. This interdisciplinary field aims to empower computer systems to not just see, but comprehend the visual data they encounter, ultimately enabling them to make informed decisions based on this information [16].

The scope of applications for computer vision is vast and continuously expanding, encompassing diverse domains such as object recognition, facial detection, image segmentation, optical character recognition, video surveillance, augmented reality, autonomous navigation, medical diagnostics, and virtual environments. Its impact extends across various sectors including healthcare, security, manufacturing, scientific research, transportation, and entertainment [17].

In healthcare, computer vision aids in medical imaging analysis, assisting doctors in diagnosing diseases and guiding surgical procedures with greater precision. In security, it enhances surveillance systems, enabling real-time monitoring and threat detection in public spaces and sensitive installations. Industries leverage computer vision for quality control, automating inspection processes to ensure product integrity and consistency [18]. Research endeavors benefit from its capabilities for data analysis and visualization, accelerating scientific discoveries and innovation [19].

Moreover, computer vision plays a pivotal role in the evolution of transportation, facilitating the development of autonomous vehicles equipped with vision-based perception systems for navigation and obstacle avoidance. It also enriches user experiences through immersive technologies like augmented and virtual reality, blurring the lines between the physical and digital realms [20].

As computer vision continues to advance, fueled by breakthroughs in deep learning and sensor technologies, its transformative potential across various domains becomes increasingly apparent [21]. By harnessing the power of visual intelligence, we pave the way for a future where machines seamlessly interact with and interpret the visual world, revolutionizing how we perceive, understand, and interact with technology.

2.2 Deep Learning

Drawing inspiration from the intricate workings of the human brain, deep learning empowers computers to autonomously acquire knowledge. Despite being a relatively recent advancement, its impact has been profound, particularly in tasks like visual content recognition, speech comprehension, and natural language processing [22].

However, deep learning faces challenges, notably its dependence on vast amounts of data for effective learning. Nevertheless, these hurdles haven't dampened the remarkable results this captivating field of machine learning has achieved. One of the most intriguing aspects of deep learning is its ability to discern complex patterns and structures within data, enabling it to perform tasks previously deemed exclusive to human capabilities. For instance, in healthcare, deep learning is employed to analyze medical images, assisting doctors in diagnosing diseases with heightened precision [23].

Furthermore, deep learning is evolving rapidly, with new architectures and emerging optimization techniques consistently enhancing its performance. Convolutional Neural Networks (CNNs) have revolutionized image processing, while transformers have significantly boosted models' capacity to understand and generate natural language. Moreover, deep learning holds immense potential across various domains such as autonomous driving, finance, security, and beyond [24]. Its impact is still unfolding, and we are likely only scratching the surface of its applications and implications for the future of technology and society.

2.3 Object Detection

Object detection stands as a pivotal pillar in the realm of computer vision, dedicated to pinpointing and categorizing objects within images or videos. Its utility spans across a spectrum of applications, encompassing realms like video surveillance, automatic license plate recognition, and advancements in medical imaging. The advent of deep learning has heralded a transformative era in object detection, supplanting traditional techniques with the prowess of deep neural networks. These neural architectures possess the ability to autonomously discern intricate features, thereby facilitating superior generalization and heightened performance across a myriad of contexts [25].

Within the panorama of object detection methodologies, two primary paradigms hold sway (refer to Figure 1): single-shot models typified by exemplars such as YOLO, and two-shot models epitomized by Faster RCNN. Single-shot models boast rapidity as their hallmark trait, rendering them ideal for real-time applications like urban surveillance. Conversely, two-shot models, albeit more intricate, furnish a finer degree of precision, rendering them apt for tasks necessitating meticulous object identification. The swift evolution witnessed in object detection, propelled by the engine of deep learning, portends continual progress. Researchers perpetually tread uncharted territory, devising novel architectures geared towards bolstering accuracy, speed, and model adaptability. This innovative momentum charts a course towards increasingly sophisticated applications, spanning from the realms of autonomous driving to the early detection of ailments, thereby sculpting the trajectory of computer vision's future [26].

Moreover, the integration of object detection algorithms with emerging technologies such as edge computing and Internet of Things (IoT) devices opens up new avenues for real-time, context-aware applications. This fusion enables the deployment of object detection models directly on devices, minimizing latency and enhancing privacy by processing data locally. Additionally, advancements in hardware acceleration, such as the utilization of specialized processing units like GPUs and TPUs, further catalyze the speed and efficiency of object detection systems, making them more accessible and scalable across various domains [27].

Furthermore, the incorporation of multi-modal information, including depth data from LiDAR sensors or thermal imaging, enriches the perceptual capabilities of object detection models, enabling them to operate robustly in challenging environmental conditions such as low light or adverse weather. This multi-sensory approach not only enhances detection accuracy but also augments the resilience of these systems to real-world variability, thus broadening their applicability in domains ranging from autonomous vehicles to industrial automation [28].

In essence, the synergistic amalgamation of cutting-edge methodologies, technological advancements, and interdisciplinary collaborations propels object detection into a realm of unparalleled innovation and practical utility, promising transformative impacts across a multitude of sectors and heralding a future characterized by unprecedented advancements in computer vision.

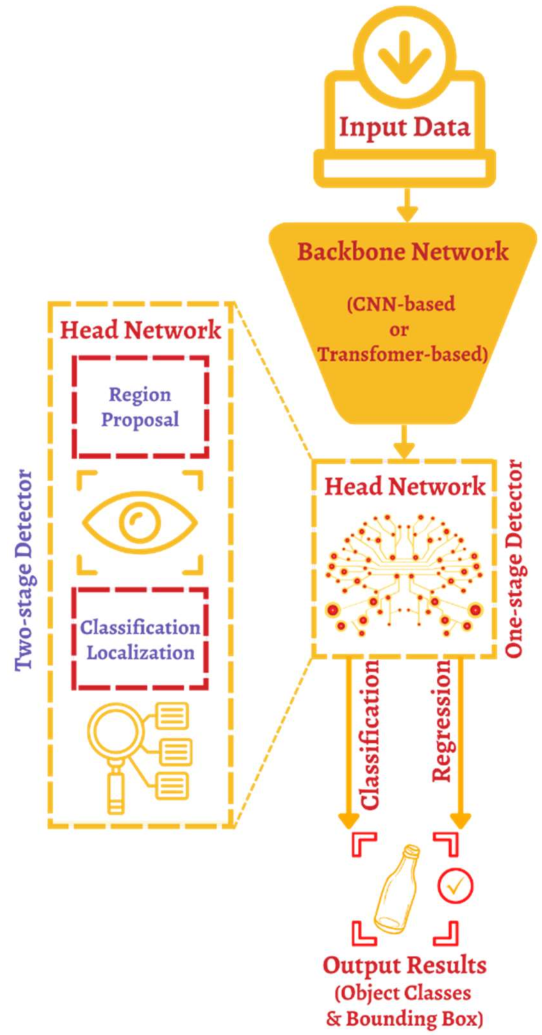

Fig. 1. The Two Primary Paradigms of Object Detection Techniques

2.3.1 Regions with Convolutional Neural Networks (R-CNN) and its variants

The evolution of object detection techniques has seen significant strides with the development of R-CNN, Fast R-CNN, and Faster R-CNN. These models have progressively improved the accuracy and efficiency of identifying objects within images. At the heart of R-CNN lies the segmentation of images into regions of interest (RoI), followed by the application of convolutional neural networks (CNNs) to extract features from each region [29].

Fast R-CNN [30] introduced a pivotal advancement by integrating the region proposal process directly into the network architecture. Unlike its predecessor, Faster R-CNN [31] takes this integration further by introducing a dedicated network, the "Region Proposal Network" (RPN).

This specialized network streamlines and accelerates the generation of region proposals, leading to expedited and more precise object detection.

These successive models represent a significant leap forward in object detection efficiency, thanks to innovations like computational resource optimization, dedicated network integration, and the refinement of region proposal techniques [32]. They showcase the continuous evolution and refinement in the field, pushing the boundaries of what is achievable in object detection tasks.

2.3.2 You Only Look Once (YOLO) variants

The YOLO architecture stands as a pioneering convolutional neural network model within the domain of real-time object detection [17]. Its distinguishing feature lies in its ability to conduct object detection in a single pass through the network, a departure from traditional methodologies that necessitate multiple steps. Introduced in 2016, the inaugural version, YOLOv1, showcased commendable efficiency; however, it encountered limitations in accuracy, particularly concerning smaller objects. Over time, subsequent iterations such as YOLOv2 (also known as YOLO9000), YOLOv3, and YOLOv4 emerged, each bringing substantial advancements in both precision and computational speed.

YOLOv2 introduced the concept of multi-scale detection and expanded the capacity to recognize a diverse array of object classes. YOLOv3, on the other hand, fine-tuned the architecture to achieve heightened precision [18]. The unveiling of YOLOv5 in 2021 marked a significant milestone in the evolution of the series. Following its release, subsequent versions including YOLO v6 [33] and v7 [34] (both launched in 2022), and the most recent iteration, v8 [35] introduced in 2023, have continued to push the boundaries of performance within the YOLO framework.

These successive releases have not only bolstered precision and processing speed but also signify a steadfast dedication to refining the architecture. Notably, the latest iterations have introduced groundbreaking enhancements, including the integration of a novel segmentation pipeline. This forward-looking approach extends the capabilities of YOLO beyond mere object detection, promising a future where the architecture becomes even more versatile and impactful.

2.4 Drone Technology for the Environment

Drones, or unmanned aerial vehicles, stand as a pivotal technological advancement in aviation and data collection. Their ingenious design boasts several distinctive qualities [36]:

- ✓ **Maneuverability:** Drones are often equipped with advanced stabilization systems, enabling precise flight and adaptability to diverse environments.
- ✓ **Versatility:** Their modular design and ability to be equipped with various sensors make them versatile, suitable for a multitude of tasks and environments.
- ✓ **Accessibility:** With their compact size and relatively affordable cost, drones are accessible to a wide range of users, from researchers to governmental agencies to private enterprises.
- ✓ **Data Collection:** Drones serve as efficient tools for data collection, offering unique aerial perspectives and the ability to cover vast areas quickly and accurately.

Drones are widely utilized for environmental applications, providing innovative solutions to various ecological challenges. Here are some of their uses in this field [37]:

- ✓ **Wildlife Monitoring:** Drones are employed to monitor animal populations, particularly in hard-to-reach habitats, providing valuable data for species conservation and management.
- ✓ **Fire Detection and Monitoring:** Using thermal cameras and specialized sensors, drones can detect wildfires and monitor their progression effectively, assisting firefighting teams in making informed decisions.
- ✓ **Ecosystem Mapping and Monitoring:** Drones are used to map and monitor terrestrial and marine ecosystems, providing valuable data on biodiversity, land use changes, and environmental impacts.
- ✓ **Pollution Monitoring:** Drones equipped with special sensors can detect and monitor sources of pollution, such as chemical spills or industrial discharges, facilitating the management and prevention of environmental pollution.
- ✓ **Reforestation and Planting:** Some drones are capable of dropping seeds or seed pods in deforested or degraded areas, contributing to reforestation efforts and ecosystem restoration.

In summary, drones offer considerable potential to enhance environmental surveillance, management, and conservation, paving the way for new innovative and sustainable approaches to protecting our planet.

2.5 REMEDIES Project

The REMEDIES project [38], funded by Horizon Europe, is an innovation program dedicated to creating innovative solutions and technologies to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

monitor, collect, prevent, and valorize (micro)plastics from our oceans. It adopts a collaborative and innovative approach focusing on several key aspects. REMEDIES project consists of 3 technical Work Packages (WP1 Monitoring plastic litter, WP2 Collection of plastic litter, WP3 Circular solutions for prevention of plastic litter) and 4 nontechnical (WP4 Sustainability assessment and optimization of implemented measures, WP5 Scaling & Replication, WP6 Community engagement, dissemination & communication, WP7 Project Management) work packages:

- ✓ **Plastic litter Monitoring (WP1):** Led by the National Institute of Chemistry, this component aims to enhance monitoring protocols for marine debris by developing innovative technological tools to detect and monitor plastic litter on land and at sea.
- ✓ **Plastic litter Collection (WP2):** Directed by Clera One, this WP focuses on collecting marine debris from identified plastic pollution sites, with the goal of cleaning up at least 85% of these sites using technologies that minimize plastic pollution in the sea.
- ✓ **Circular Solutions for Plastic litter Prevention (WP3):** Under the guidance of Alchemia-nova Greece, this aspect aims to develop strategies for valorizing plastic litter, emphasizing circular solutions to reduce plastic leakage into the environment.
- ✓ **Sustainability Assessment and Optimization of Measures (WP4):** Led by F6S, this WP evaluates the impacts of measures implemented in other WPs on the marine environment and establishes a digital platform to consolidate and communicate the results obtained.
- ✓ **Scaling Up and Replication (WP5):** Under the direction of Impact Hub Athens, this component aims to extend the innovations developed in the REMEDIES project to other Mediterranean regions by establishing acceleration networks to facilitate their adoption.
- ✓ **Community Engagement, Outreach, and Communication (WP6):** Led by the National Institute of Chemistry, this WP seeks to raise awareness and engage local communities in the fight against plastic pollution using participatory practices and communication activities.
- ✓ **Project Management (WP7):** Also led by the National Institute of Chemistry, this WP ensures effective coordination and management of the REMEDIES project, ensuring that partners progress in accordance with established plans.

By working on these various components, the partners of the REMEDIES project aim to develop innovative and sustainable solutions to reduce plastic pollution in the Mediterranean oceans, while mobilizing local communities and promoting a more circular blue economy. Our research endeavors fall within the scope of Work Package 1 dedicated to monitoring efforts, ensuring comprehensive data collection and analysis for informed decision-making.

3. RELATED WORKS

Pan and colleagues [39] endeavored to detect waste in the Asahi River, Japan, using two object detection models, YOLOv5 and RetinaNet. Their study yielded highly precise results in waste recognition at the study site. By employing a large collection of PET (Positron Emission Tomography) images gathered from the internet as the training dataset and experimenting with the aforementioned object detection models with various parameters (Batch size, Epochs), the addition of a PET dataset for training, with a Ground Sampling Distance (GSD) similar to that of the test dataset, led to an improvement in recall value. However, without combining the original dataset collected at the study site, the authors encountered difficulties in detecting PET using only the supplementary dataset. They concluded that combining the original dataset with the supplementary dataset is a relatively better method for enhancing the recall value of PET detection.

Iordache et al. [40] deployed remotely piloted aircraft equipped with multispectral cameras over polluted areas to detect waste. Their approaches utilize classification algorithms based on random forests to distinguish four classes of soil types and five classes of waste. The results show that detecting different types of waste is indeed feasible in the proposed scenarios, with machine learning algorithms achieving accuracies exceeding 85% for all classes in test data. Furthermore, the study explored error sources, the effect of spatial resolution on extracted maps, and the applicability of the designed algorithm to detect floating waste.

Gonçalves et al. [41] utilized images captured at very low altitudes, collected using a low-cost RGB camera mounted on a drone flying over a sandy beach to characterize the presence of macro debris. They developed an object-oriented classification strategy to automatically identify marine macro debris based on drone orthomosaic images. Three automated object-oriented machine learning techniques were

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

compared: Random Forest (RF), Support Vector Machine (SVM), and k-Nearest Neighbors (KNN). The detection results obtained were satisfactory for all three techniques, with average F-scores of 65%. Specifically, the average F-scores were 65% for KNN, 68% for SVM, and 72% for RF. A comparison with manual detection showed that the RF technique was the most accurate macro debris detector, providing the best overall detection quality (F-scores) with the fewest false positives. Additionally, KNN is considered the simplest classifier with only one parameter, as the three generated abundance maps exhibited similar correlation with the manually produced abundance map.

Jakovljevic et al. [42] sought to determine the relevance of deep learning algorithms for the automatic extraction of floating plastics from drone orthophotos, testing the ability to differentiate between plastic types and exploring the relationship between spatial resolution and detectable plastic size. Their aim was to define a methodology for drone-based monitoring to map floating plastic. Two study areas and three datasets were used to train and validate the models. A semantic segmentation algorithm based on the U-Net architecture using ResUNet50 provided the highest accuracy in mapping different plastic materials (F1-score: Oriented Polystyrene (OPS): 0.86; Nylon: 0.88; polyethylene terephthalate (PET): 0.92; plastic (general): 0.78), demonstrating its ability to identify plastic types. Classification accuracy decreased with decreasing spatial resolution, with optimal performance achieved at a resolution of 4 mm for all plastic types. The model also provided reliable estimates of plastic surface area and volume, crucial information for cleanup campaigns.

Taddia et al. [43] investigated strategies for mapping anthropogenic marine debris on beaches using various ground resolutions and supported by elevation and multispectral data to create RGB orthomosaics. Operators with different levels of expertise and coastal environmental knowledge manually mapped debris along four to five transects using a range of photogrammetric tools. The initial study was repeated a year later, during which beach debris samples were collected and analyzed in the laboratory. Operators assigned three levels of confidence when identifying and describing debris. Preliminary validation of the results showed that items were identified with high confidence and were almost always correctly classified. The approach of assessing items in terms of surface area instead of simple counts significantly increased the percentage of mapped debris compared to those collected. The proposed methodology is a practical solution for mapping beach debris using RGB imagery and a spatial resolution of approximately 200 pixels/m for detecting macroplastics. This approach is feasible, fast, convenient, and sustainable for evaluating and monitoring potential sources of microplastics.

4. ARCHITECTURE AND RESEARCH METHODOLOGY

The architectural proposal presented herein signifies a groundbreaking advancement in environmental beach monitoring, aimed at substantially enhancing efficiency. At its core lies a strategy that seamlessly integrates cutting-edge interfaces, leveraging drone imagery and deploying sophisticated deep learning techniques for the detection of plastic litter, with a specific emphasis on combating this prevalent issue. This ambitious endeavor hinges on the crucial implementation of meticulously crafted deep learning algorithms, adept at discerning and identifying the characteristic indicators of plastic litter amidst the vast dataset captured by aerial drones. These drones, equipped with state-of-the-art high-resolution cameras, ensure precise visual data acquisition, capturing intricate details crucial for early plastic litter detection.

Upon completion of the data collection phase, a meticulously designed pipeline will be activated, utilizing pre-trained object detection models to automate the identification of emerging signs of "Plastic Litter", as depicted in Figure 2. The culmination of this global initiative promises to revolutionize environmental monitoring, proactively mitigating the risks posed by plastic litter and refining beach management strategies.

The proposed methodology consists of five crucial steps, as shown in Figure 3. The steps are presented in detail below.

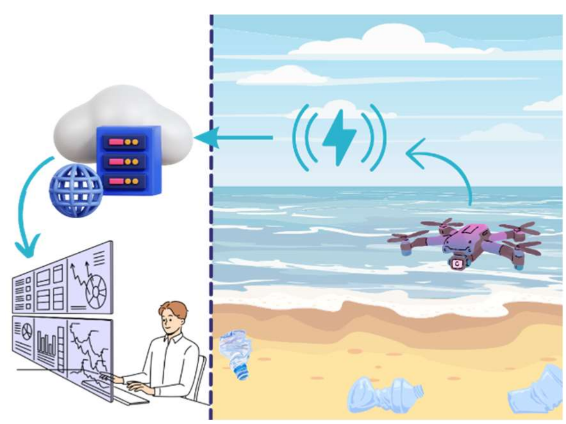

Fig. 2. Proposed Architecture

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

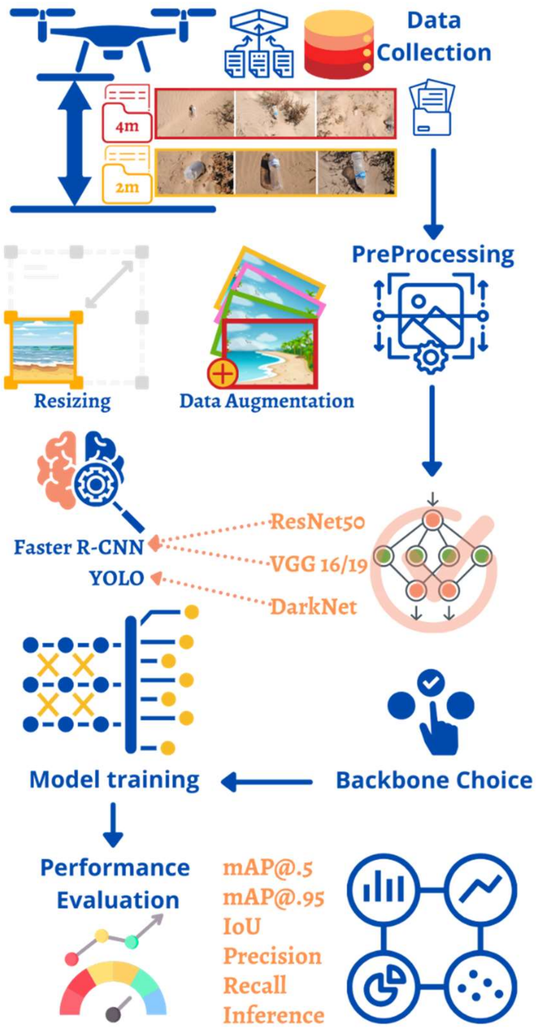

Fig. 3. Research Methodology

4.1 Data Collection

The initial phase of our proposed methodology involves gathering a comprehensive array of images within a beach area, encompassing plastic litter and areas free from any visible debris (Figure 4). These images will be captured by drones flying at two different altitudes: 2 meters and 4 meters.

Fig. 4. Collecting Images Featuring Plastic Litter and Areas Free from Debris from the Database

The selected site for collecting images for training and testing purposes is the SIBE (Site of Biological and Ecological Interest) of the Moulouya River Mouth, located in eastern Morocco (Figure 5). This region is renowned for its ecological diversity, encompassing forest, steppe, and wetland ecosystems. The SIBE of the Moulouya River Mouth represents one of these wetland ecosystems in the eastern Mediterranean region, harboring exceptional floristic and faunistic biodiversity, including numerous endemic species. We have chosen this site for collecting and training our plastic litter detection models for several crucial reasons. Firstly, its exceptional biodiversity underscores the negative impact of plastic litter on fragile ecosystems, enhancing awareness of the urgency of waste management and biodiversity conservation. Secondly, the diversity of local ecosystems ensures a representative and robust image database for training our models. Lastly, by selecting this site in the eastern Mediterranean region, we broaden the geographical scope of our study, contributing to a more comprehensive understanding of plastic pollution in Mediterranean coastal ecosystems. This approach may lead to more targeted policy and environmental recommendations at regional and international levels.

Fig. 5. Image Collection and Testing Site: Moulouya River Mouth, Saidia, Morocco

We collected images of plastic litter in the study area using drones during multiple site visits at various times of the year. Two distinct datasets were formed, each containing 3000 images after applying data augmentation techniques. The first dataset was captured at a height of 4 meters by drone, while the second was taken at a height of 2 meters. This approach allows us to have a comprehensive and detailed representation of plastic litter in the area of interest from different perspectives. This diversity significantly enriches our dataset and enhances the quality of our analysis. Furthermore, it enables us to study the optimal flight conditions for image collection, thus refining our research methodology.

Figure 6 illustrates the prototype of the plastic litter counting architecture we are

Journal of Theoretical and Applied Information Technology
15th June 2024. Vol.102. No. 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

considering, for which we aim to find the appropriate deep learning model and optimal flight parameters. The test area measures 100m x 100m, and drones are required to navigate it in a maze-like pattern, detecting and counting the waste. During this phase, we are striving to identify the best parameters for effective detection.

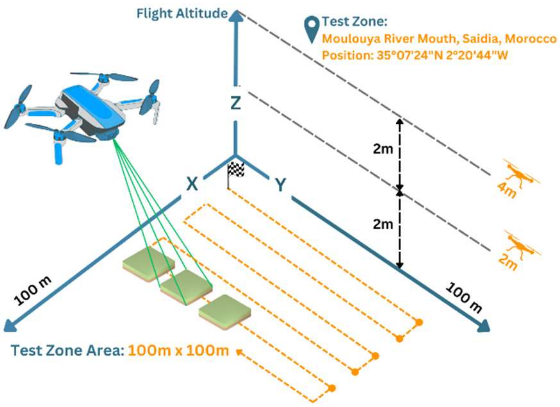

Fig. 6. Prototype Testing for Plastic Litter Detection and Counting Architecture on Beaches

4.2 Preprocessing

The second step in our methodology "image data processing" is a pivotal stage, essential for enhancing the quality and consistency of the dataset. Its primary goal is to refine the data by eliminating unwanted elements like noise and artifacts, which could potentially hinder the model's efficacy. This phase encompasses a range of operations, including intensity normalization, distortion correction, and image scaling. Through these processes, preprocessing ensures a standardized and clean input, thereby facilitating smoother convergence during the model's learning phase [44].

Moreover, a crucial aspect of preprocessing lies in image data augmentation. This involves applying diverse transformations to existing images, such as rotations, flips, zooms, and other geometric alterations with randomized parameters. The core aim of augmentation is to inject greater diversity into the dataset, enriching it with a broader spectrum of variations. By doing so, the model becomes adept at recognizing patterns amidst various real-world scenarios, thus bolstering its robustness and adaptability [45].

By synergizing these two steps, the overall preprocessing of image data culminates in the creation of an optimized training set. This optimized set not only maximizes the model's performance but also equips it with the resilience needed to navigate through varying conditions effectively.

4.3 Backbone Choice

In our endeavor to enhance Plastic Litter detection in drone-captured images of beaches designated for monitoring purposes, we are embarking on an investigation utilizing two distinct methodologies: single-pass models and double-pass models, representing the third phase of our study.

Regarding single-pass models, we will delve into the efficacy of YOLOv6, YOLOv7, and YOLOv8 architectures, all leveraging the Darknet backbone. Conversely, for double-pass models, we will analyze the performance yielded by Faster R-CNN, utilizing backbones such as ResNet50, VGG16, and VGG19. The selection of these convolutional neural network architectures as backbones is informed by previous research indicating their effectiveness in analogous scenarios [22].

Moreover, to ensure comprehensive evaluation, all models will be tested at varying drone heights of 2m and 4m above ground level. This height variation consideration is crucial as it can significantly impact detection accuracy, especially in the context of plastic litter on beaches.

This comparative approach is designed to facilitate the identification and selection of the most suitable model tailored to our specific objective of Plastic Litter detection in maritime environments. Additionally, it aims to mitigate any potential biases and ensure robustness in our model selection process.

ResNet50. a variant of the Residual Network architecture, stands as a pivotal milestone in the realm of deep learning, spearheaded by Microsoft Research in 2015. The essence of its innovation lies in its approach to addressing the inherent challenges posed by training extremely deep neural networks. At its core, ResNet50 introduces the concept of residual learning, a paradigm shift in neural network design. This groundbreaking concept revolves around the integration of shortcut connections, also known as skip connections, which enable the flow of information to bypass certain layers during forward propagation [23]. By doing so, ResNet50 effectively mitigates the notorious vanishing gradient problem, a common hurdle encountered in training deep networks, thereby facilitating the training process with improved efficiency and accuracy. ResNet50 has garnered widespread acclaim for its remarkable performance in various computer vision tasks, particularly in image recognition and object detection. Its ability to harness the advantages of deep neural networks while overcoming traditional training obstacles has solidified its position as a cornerstone model in the field [46]. With its robust architecture and proven track record, ResNet50 continues to serve as a beacon of innovation, driving advancements in deep learning research and application.

VGG. short for Visual Geometry Group, stands as a prominent ensemble of convolutional neural networks (CNNs) celebrated for their simplistic yet powerful design. Within this family, VGG16 and VGG19 emerge as notable figures, boasting 16 and 19 layers respectively. Originating from the laboratories of the University of Oxford, these models are characterized by a series of convolutional layers interlaced with densely connected layers. Their straightforward architecture belies their remarkable effectiveness, particularly evident in their prowess in image classification endeavors. Through their contributions, VGG16 and VGG19 have left an indelible mark on the landscape of deep learning within the realm of computer vision, cementing their legacy and enduring relevance. VGG16 and VGG19 are distinguished not only by their layer depths but also by their uniform kernel size of 3x3 throughout the entire network, contributing to their consistent performance across various datasets and tasks [47].

DarkNet. a pioneering neural network framework pioneered by Joseph Redmon, the mastermind behind the YOLO algorithm, is renowned for its agility and effectiveness in the realm of computer vision. Engineered to excel in real-time object detection tasks, this lean architecture strikes a remarkable equilibrium between swiftness and precision, making it a preferred choice for a wide array of applications. What sets DarkNet apart is its versatility, supporting computations on both CPU and GPU platforms without compromising performance. Its open-source nature and modular design have fostered widespread acceptance within the deep learning community, where it seamlessly integrates into various projects, adapting effortlessly to diverse requirements and scenarios. This adaptability and ease of integration have solidified DarkNet's position as a go-to framework, elevating the standards of efficiency and efficacy in computer vision applications [48].

4.4 Model Training

In our quest to refine our object detection models, the training step takes precedence, representing the crucial fourth phase in our methodology. We engage in meticulous adjustment and fine-tuning of our models, leveraging the preprocessed and augmented dataset to enhance their performance.

Our dataset is meticulously divided into training, validation, and testing subsets, with allocations of 70%, 20%, and 10%, respectively. This segmentation ensures robust training while facilitating effective evaluation of model performance.

Our arsenal comprises a range of cutting-edge models including YOLOv6, YOLOv7, and YOLOv8, alongside various iterations of Faster R-CNN such as those based on ResNet50, VGG16, and VGG19 architectures. These models are built upon a solid foundational backbone, ensuring versatility and adaptability to our specific task.

During the training process, we harness the power of labeled data, where objects of interest are meticulously outlined with bounding boxes and associated with class information using state-of-the-art Open-Source Data Labeling software. This step ensures the model's ability to accurately identify and classify objects within images.

Our training pipeline involves diverse data formats tailored to the requirements of each model architecture. For instance, Faster R-CNN models utilize TensorFlow record (TFRecord) files, while YOLO models rely on TXT annotations and YAML config files. This adaptability allows us to optimize the training process for each model, maximizing efficiency and effectiveness.

Ultimately, the objective of this training phase is to refine models capable of achieving unparalleled accuracy and reliability in detecting plastic litter. This endeavor aims to equip stakeholders with powerful tools for effectively managing and mitigating the issue at hand.

4.5 Performance Evaluation

To assess the performance of our models, we proceed to the Performance Evaluation step by meticulously analyzing all gathered data. This evaluation, integral to our process, allocates 10% of the test set to measure the average accuracy and inference speed of each model. This testing phase is of paramount importance as it allows us to gauge the models' performance on unseen data, shedding light on their overall effectiveness in detecting plastic litter and their potential real-time utilization by drones. To ensure a comprehensive evaluation, we plan to employ multiple metrics to capture various dimensions of our models' performance.

5. RESULTS ANALYSIS AND DISCUSSION

In the forthcoming section dedicated to the analysis and discussion of results, we will embark on elucidating the hardware specifications instrumental in our exploration of object detection utilizing a drone-mounted camera. This delineation will provide a foundational understanding of the technological underpinnings crucial for our study.

Subsequently, we will introduce and explicate the array of evaluation metrics employed to meticulously gauge the accuracy and efficacy of the developed system. These metrics encompass a multifaceted approach to comprehensively evaluate the performance of our models.

To further explore the object detection results, we will conduct a comparative analysis using two separate datasets obtained from actual plastic litter sites. These datasets will be examined at two different flying altitudes, 2m and 4m, to ensure a thorough evaluation of detection capabilities across different conditions.

Concluding this section, we will present a thorough discourse on the outcomes attained through the proposed methodology. This discussion will not only highlight the successes but also critically analyze any limitations or challenges encountered during the course of our investigation.

Moreover, it's imperative to note that multiple metrics will be employed to assess and validate the performance of our models comprehensively. These metrics will encompass traditional evaluation measures such as precision, recall, and F1-score, alongside domain-specific metrics tailored to the intricacies of plastic litter detection in maritime environments. Such a comprehensive evaluation framework will ensure a nuanced understanding of our models' capabilities and limitations.

5.1 Hardware Specifications

In our investigation, the hardware specifications played a pivotal role in ensuring the efficacy of our experimental setup. Our drone of choice was the Mavic Air model by DJI, which boasted a high-resolution camera crucial for capturing detailed images of the beach environment under scrutiny [49].

To train and test our object detection models, we leveraged a robust infrastructure centered around a DELL PowerEdge R740 server. This server was equipped with an Intel Xeon Silver 4210 2.2G processor, providing substantial processing power essential for handling the complexities of deep learning algorithms [50]. Complementing the processor, the server boasted an impressive 80GB of RAM, facilitating efficient data processing and model training tasks.

Furthermore, to expedite the computational workload associated with training our models, the server was augmented with two NVIDIA RTX A5000 GPUs. These GPUs, each endowed with 24GB of graphics memory, significantly accelerated the training process by parallelizing computations and harnessing the power of advanced GPU architecture.

The amalgamation of these hardware components ensured a robust and efficient experimental setup, capable of handling the computational demands inherent in training and testing sophisticated object detection models. Additionally, this hardware configuration enabled us to achieve optimal performance and accuracy in our Plastic Litter detection endeavors, ultimately contributing to the success of our study.

5.2 Model Evaluation Metrics

When evaluating the effectiveness of object detection algorithms, a range of metrics is available, each shedding light on different aspects of model performance.

5.2.1 Precision, Recall and F1 score

Precision, recall, and F1 score are fundamental metrics for evaluating object detection algorithms. Precision measures the proportion of correctly identified instances out of all instances labeled as a particular class, including false positives. Recall, on the other hand, assesses the ability of the model to detect all actual occurrences of the target class. F1 score, which is the harmonic mean of precision and recall, provides a balanced assessment of the algorithm's performance.

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (1)$$

5.2.2 Average Precision (AP)

Average Precision offers a comprehensive overview of the detector's performance across all classes, including the specific class of interest. By constructing precision-recall curves and averaging precision values at various recall levels, AP captures the overall effectiveness of the detector. However, it may lack granularity when evaluating individual classes [51].

The formula for AP is:

$$AP = \sum_{k=0}^{k=n-1} [Recalls(k) - Recalls(k+1) * Precisions(k)] \quad (2)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where Recalls(n)=0, Precisions(n)=1, and n=Number of thresholds

5.2.3 Mean Average Precision (mAP)

Mean Average Precision is a sophisticated metric that considers precision and recall for each class, providing a detailed perspective on model performance. Calculated at different confidence thresholds, mAP offers insights into the robustness of the model. Despite its complexity, mAP enhances evaluation by offering detailed insights into overall model effectiveness. The formula for mAP is:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} APi \qquad (3)$$

5.2.4 Intersection over Union (IoU)

Intersection over Union measures the overlap between a detected object and its ground truth, ensuring accurate object localization. Despite its complexity, IoU provides detailed insights into the model's ability to precisely align detected objects with their actual references. IoU complements metrics like AP and mAP by offering specific information on object localization accuracy.

$$IoU = \frac{(Area\ of\ Intersection)}{(Area\ of\ Union)} \qquad (4)$$

$$Area\ of\ Intersection = (min(x2, x4) - max(x1, x3)) * (min(y2, y4) - max(y1, y3)) \quad (5)$$

$$Area\ of\ Union = (x2 - x1) * (y2 - y1) + (x4 - x3) * (y4 - y3) - Area\ of\ Intersection \qquad (6)$$

5.2.5 Inference Time

Inference time refers to the duration it takes for a deep learning model to analyze an input image and generate predictions. Influenced by factors like model architecture and computational resources, optimizing inference time is crucial for real-time applications while maintaining sufficient accuracy. Efficient inference time ensures timely responses, particularly in applications requiring swift decision-making processes. Inference time, a critical metric in deep learning model performance evaluation, is heavily contingent upon the underlying hardware infrastructure utilized during computation. Influenced by factors like model architecture and computational resources, optimizing inference time is crucial for real-time applications while maintaining sufficient accuracy. Efficient inference time ensures timely responses, particularly in applications requiring swift decision-making processes.

5.3 Findings and Discussion

This study aims to assess the performance of two families of object detection models in identifying plastic waste, with the intention of integrating this drone-based surveillance capability into a comprehensive monitoring platform. Additionally, we seek to determine the optimal flight altitude of our drone to enhance detection performance and reduce errors during waste counting operations. We conducted an assessment of two-stage Faster-RCNN models, each utilizing different backbone networks whose selection was justified in a previous study, along with various versions of the one-stage YOLO model (v6, v7, and v8). These models were trained using two distinct datasets consisting of images captured at two different altitudes directly by a drone at a work and test site located in the northeast of Morocco. We trained each model for 100 epochs to ensure convergence. The models underwent training and evaluation using various metrics, including mAP at IoU thresholds of 0.5 and 0.95, recall, precision, and F1 score. Additionally, we measured the inference time (milliseconds per image) using two Nvidia RTX A5000 GPUs. The F1 score, a metric that combines precision and recall in object detection models, provides a valuable balance between avoiding false detections and effectively capturing real objects. mAP@0.5 is chosen as it provides a balanced assessment of object detection models across various object sizes and levels of detection confidence. In disciplines such as medicine, precision holds paramount importance. The metric mAP@0.95 stands out for its ability to deliver dependable outcomes, effectively minimizing instances of false positives.

Table 1 and Figure 7 comprehensively illustrate the outcomes yielded by various models throughout both the complete testing phase and the training process, respectively. An initial observation reveals that all models successfully identify plastic litter with precision, and notably, YOLO models require less training time. Regardless of the dataset used, YOLO models consistently achieve convergence significantly earlier compared to Faster RCNN models. The latter typically require nearly forty epochs to converge.

The initial observation is that for all the models studied, the results obtained using the dataset of images collected at a 2-meter drone flight height are significantly better than those achieved with the dataset collected at 4 meters. Additionally, processing times are also slightly shorter. Even during the training phase, all the models studied

Journal of Theoretical and Applied Information Technology
15th June 2024. Vol.102. No. 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

require fewer epochs to converge when the images are taken from closer distances. Therefore, in our scenario, it is preferable to minimize the flight height while ensuring that the drone battery autonomy is sufficient to cover the entire study area.

Table 1: Performance Results of Implemented Models on the Test Set

Deep Learning Model	Flight Altitude	Inference time (ms / frame)	Precision %	Recall %	F1 Score %	IoU %	mAP @0.5 %	mAP @0.95 %
Faster-RCNN (RS50)	4m	~71.91	87,81	88,14	87,97	88.06	87.34	75.91
	2m	~70.95	93,93	94,06	93,99	93.55	93.32	79.59
Faster-RCNN (VGG19)	4m	~97.92	85,83	86,08	85,95	85.74	85.76	75.77
	2m	~96.14	92,65	92,94	92,79	92.77	92.75	78.54
Faster-RCNN (VGG16)	4m	~87.77	86,73	86,91	86,82	86.31	86.19	75.63
	2m	~87.16	92,22	91,73	91,97	91.32	91.22	78.42
YOLOv6	4m	~2.49	79,61	79,83	79,72	79.26	79.12	73.92
	2m	~2.32	88,14	88,32	88,23	88.17	88.02	76.39
YOLOv7	4m	~3.54	81,41	81,56	81,48	81.23	81.29	74.17
	2m	~3.51	89,14	89,92	89,53	89.33	89.28	76.57
YOLOv8	4m	~1.46	83,61	83,78	83,69	83.76	83.53	74.56
	2m	~1.35	90,51	91,01	90,76	90.65	90.44	76.81

Among the various models evaluated, Faster-RCNN emerged as the top performer regardless of flight altitude, delivering exceptional results. Specifically, the Faster-RCNN model utilizing ResNet50 (RS50) backbones demonstrated outstanding performance at a drone flight altitude of 2 meters, achieving an F1 score of 93.99% and a mean average precision (mAP) of 93.32% at an intersection over union (IoU) threshold of 0.5, along with a mAP of 79.59% at an IoU of 0.95. At a flight altitude of 4 meters, performance remained high with an F1 score of 87.97% and a mAP of 87.34% at an IoU of 0.5, and a mAP of 75.91% at an IoU of 0.95. These findings unquestionably underscore the effectiveness and dependability of models configured in this manner. Additionally, this specific model touts the swiftest inference time among Faster-RCNN variants, averaging between 70.95 and 71.91 milliseconds per image, contingent upon flight altitude. While YOLO models have demonstrated slightly lower performance compared to Faster-RCNN models, YOLOv8 achieves notable results, with a mAP@0.5 of 90.44% and a mAP@0.95 of 76.81% for dataset 2m, and a mAP@0.5 of 83.53% and a mAP@0.95 of 74.56% for dataset 4m. However, where YOLO models truly excel is in inference time. Specifically, YOLOv8 averages only 1.35 milliseconds per image for a flight height of 2 meters.

It is noteworthy that YOLO v6, v7, and v8 are not sequential versions, meaning that one is not necessarily newer than the other. Instead, they represent results from distinct research endeavors. This positions YOLO, especially YOLOv8, as an optimal choice for real-time applications, such as drone data collection, where fast processing speed is crucial. While the Fast-RCNN model (RS50) also yielded satisfactory results, its higher processing speed compared to YOLO models makes it the preferred choice in an architecture where the drone sends images to a ground station responsible for Plastic Litter object recognition.

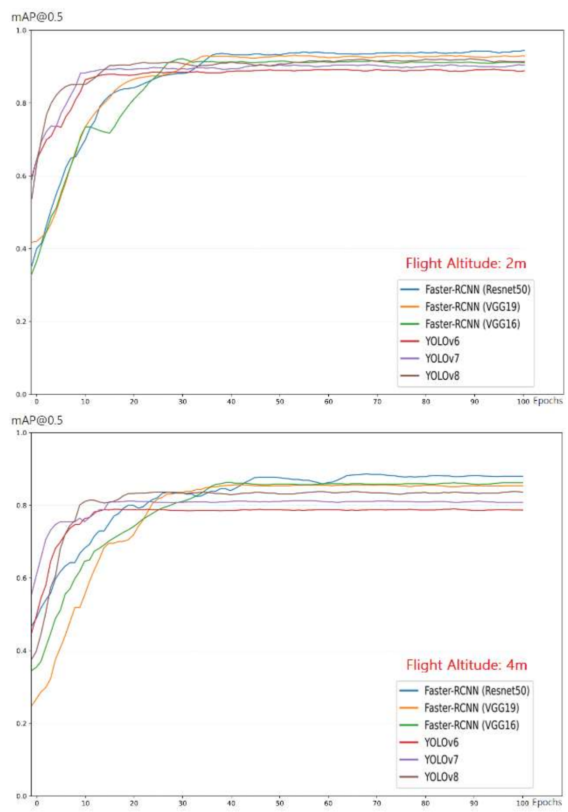

Fig. 7. Progression of mAP@0.5 Over Epochs for Examined Models (Validation Dataset) in both Flight Altitude Scenarios

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The decision between YOLO and Faster RCNN models for detecting plastic litter hinges on the desired balance between precision and processing speed. For tasks prioritizing high accuracy, Faster RCNN models, especially the one utilizing Resnet50 as its backbone, would be the favored choice. Conversely, for real-time applications demanding swift processing speeds, YOLO models, particularly YOLOv8, emerge as the preferred option (refer to Figure 6).

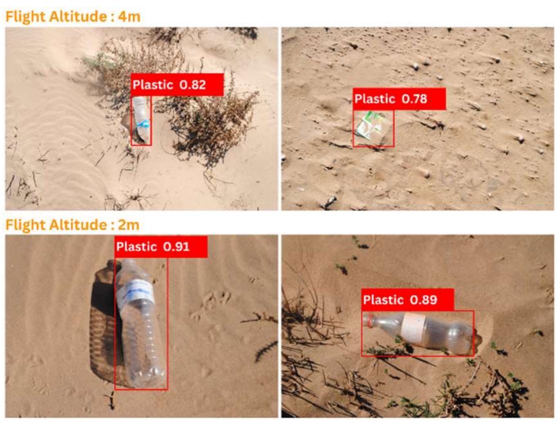

Fig. 8. Examples of Plastic Litter Detection Using YOLOv8

6. CONCLUSION

The advancement of technology has spurred a revolution in environmental monitoring, fundamentally transforming our capacity to safeguard and oversee our planet with unparalleled precision and efficiency. Integrating drones and AI, particularly through deep learning techniques, presents a formidable toolkit, empowering real-time data collection and analysis, thereby revolutionizing environmental monitoring practices as never seen before. Moreover, deep learning has revolutionized object detection, a cornerstone of computer vision, thereby enabling its application in the detection and monitoring of plastic litter, thus bolstering our ability to combat environmental pollution more effectively.

This research contributes a comparative study of deep learning object detection models, examining their potential for detecting 'plastic litter' using drones. Additionally, the study endeavors to explore optimal drone flight behaviors to enhance the efficacy of environmental surveillance efforts. The research endeavors to assess and compare various methodologies to ascertain the most efficacious approach for early and precise detection, thus laying the foundation for pioneering solutions in environmental monitoring. In our investigation, we delved into two distinct categories of object detection models: the single-pass YOLO models (versions v6, v7, and v8), and the two-stage Faster R-CNN models, utilizing three backbone variants: ResNet50, VGG16, and VGG19. Our findings reveal encouraging outcomes for the detection and monitoring of plastic litter, suggesting its potential as a valuable asset within this domain. Selecting the most suitable model for drone images requires striking a balance between accuracy and processing speed. For optimal precision, leveraging the faster RCNN model with ResNet50 as its backbone is recommended. On the other hand, for real-time applications where speed is crucial, the YOLO model, particularly YOLOv8, stands out as the preferred option. It boasts a mAP@0.5 surpassing 90.44% and a mAP@0.95 exceeding 76.81%, with an impressive inference time of approximately 1.35 milliseconds per image.

In our scenario, optimizing model performance and reducing processing times involves minimizing the drone's flight height. This strategy consistently yields superior results compared to higher altitudes across all studied models, with shorter convergence epochs observed during training. Thus, it is crucial to minimize the flight height while considering the importance of ensuring sufficient drone battery autonomy to cover the entire study area.

This study illuminates not just the capabilities of deep learning models in detecting plastic litter early on with drones, but also emphasizes the pivotal role of integrating innovative technology in propelling forward environmental surveillance. Through the evaluation and comparison of diverse deep learning methodologies, this research enriches the expanding realm of knowledge aimed at refining the precision and effectiveness of disease detection within precision environmental monitoring systems.

As future work, extending the scope of this research to include plastic litter counting is paramount. Additionally, further investigation could focus on fine-tuning deep learning models specifically for detecting and monitoring plastic litter in diverse environmental conditions, potentially incorporating multi-sensor data fusion techniques to enhance detection accuracy. Additionally, exploring the integration of unmanned aerial vehicles (UAVs) equipped with advanced sensors beyond visual imagery, such as hyperspectral or LiDAR, could provide complementary insights for comprehensive environmental monitoring and management strategies. Lastly, efforts to develop automated decision-making frameworks leveraging machine learning algorithms to prioritize and allocate resources for environmental cleanup and mitigation efforts could significantly contribute to addressing plastic pollution challenges on a broader scale.

ACKNOWLEDGEMENTS

We wish to express our gratitude to Mohammed First University for their support through the REMEDIES project. Heartfelt thanks go to the project contributors for their generous funding and support via the HORIZON Innovation Actions - Prevent and eliminate pollution of our oceans, seas, and waters (HORIZON-MISS-2021-OCEAN-03). We sincerely appreciate the dedicated efforts of all project partners, researchers, and collaborators in co-creating and implementing effective strategies for both valorizing and preventing plastic litter.

REFERENCES:

[1] M. Wolf et al., "Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q)," Environ. Res. Lett., vol. 15, no. 11, p. 114042, Nov. 2020, doi: 10.1088/1748-9326/ABBD01.

[2] OECD Environment Statistics (Dataset), "OCDE (2024), « Global Plastics Outlook : Plastics use by region - projections »."

[3] OECD, "Global Plastics Outlook: Plastic waste collected for recycling," 2023. .

[4] B. Carrera, V. L. Piñol, J. B. Mata, and K. Kim, "A machine learning based classification models for plastic recycling using different wavelength range spectrums," J. Clean. Prod., vol. 374, p. 133883, Nov. 2022, doi: 10.1016/J.JCLEPRO.2022.133883.

[5] J. Lorenzo-Navarro et al., "Deep learning approach for automatic microplastics counting and classification," Sci. Total Environ., vol. 765, p. 142728, Apr. 2021, doi: 10.1016/J.SCITOTENV.2020.142728.

[6] D. Mohanty, A. Kasliwal, B. Udupa, and P. Seth, "Sailing Through Spectra: Unveiling the Potential of Multi-Spectral Information in Marine Debris Segmentation." Accessed: Mar. 30, 2024. [Online]. Available: https://www.esa.int/SPECIALS/Eduspace_EN/.

[7] K. Topouzelis, A. Papakonstantinou, and S. P. Garaba, "Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018)," Int. J. Appl. Earth Obs. Geoinf., vol. 79, pp. 175–183, Jul. 2019, doi: 10.1016/J.JAG.2019.03.011.

[8] Y. Zarrouk, M. Yandouzi, M. Grari, M. Bourhaleb, M. Rahmoune, and K. Hachami, "REVOLUTIONIZING POTATO LATE BLIGHT SURVEILLANCE: UAV-DRIVEN OBJECT DETECTION INNOVATIONS," J. Theor. Appl. Inf. Technol., vol. 15, no. 7, 2024, Accessed: Apr. 22, 2024. [Online]. Available: www.jatit.org.

[9] M. Yandouzi et al., "Review on forest fires detection and prediction using deep learning and drones," J. Theor. Appl. Inf. Technol., vol. 100, no. 12, pp. 4565–4576, 2022.

[10] J. Delgado, A. Barkaoui, M. Petelin, A. Palatinus, and M. Velimirovic, "AI-driven aerial drones and monitoring app: New developments to facilitate citizen science initiatives on plastic pollution monitoring and clean-ups on beaches," EGU24, Mar. 2024, doi: 10.5194/EGUSPHERE-EGU24-7368.

[11] T. Malche, P. Maheshwary, P. K. Tiwari, A. H. Alkhayyat, A. Bansal, and R. Kumar, "Efficient solid waste inspection through drone-based aerial imagery and TinyML vision model," Trans. Emerg. Telecommun. Technol., p. e4878, 2023, doi: 10.1002/ETT.4878.

[12] M. Yandouzi et al., "Investigation of Combining Deep Learning Object Recognition with Drones for Forest Fire Detection and Monitoring," Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 3, pp. 377–384, 2023, doi: 10.14569/IJACSA.2023.0140342.

[13] M. Yandouzi et al., "A Lightweight Deep Learning Model for Forest Fires Detection and Monitoring," pp. 697–705, 2023, doi: 10.1007/978-981-19-6223-3_71.

[14] S. O. Fakayode et al., "Microplastics: Challenges, toxicity, spectroscopic and real-time detection methods," Appl. Spectrosc. Rev., 2024, doi: 10.1080/05704928.2024.2311130.

[15] G. Gonçalves, U. Andriolo, L. M. S. Gonçalves, P. Sobral, and F. Bessa, "Beach litter survey by drones: Mini-review and discussion of a potential standardization," Environ. Pollut., vol. 315, p. 120370, Dec. 2022, doi: 10.1016/j.envpol.2022.120370.

[16] M. Berrahal et al., "Investigating the effectiveness of deep learning approaches for deep fake detection," Bull. Electr. Eng. Informatics, vol. 12, no. 6, pp. 3853–3860, Dec. 2023, doi: 10.11591/EEI.V12I6.6221.

[17] M. Grari et al., "Forest Fire Detection and Monitoring through Environment Sound Spectrum using Deep Learning," J. Theor. Appl.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Inf. Technol., vol. 15, p. 19, 2023, Accessed: Jan. 17, 2024. [Online]. Available: www.jatit.org.

[18] M. Grari et al., "IoT-Based Approach for Wildfire Monitoring and Detection," pp. 205–213, 2023, doi: 10.1007/978-3-031-35251-5_19.

[19] Y. Hammoudi, I. Idrissi, M. Boukabous, Y. Zerguit, and H. Bouali, "Review on maintenance of photovoltaic systems based on deep learning and internet of things," Indones. J. Electr. Eng. Comput. Sci., vol. 26, no. 2, May 2022.

[20] Y. Zerguit, Y. Hammoudi, and M. Derrhi, "ENERGY EFFICIENCY IN INDUSTRY 4.0: IMPLEMENTING REAL-TIME POWER FACTOR CORRECTION IN SMART MANUFACTURING," J. Theor. Appl. Inf. Technol., vol. 101, no. 1, pp. 359–373, Jan. 2024.

[21] M. Grari et al., "Using IoT and ML for Forest Fire Detection, Monitoring, and Prediction: a Literature Review," J. Theor. Appl. Inf. Technol., vol. 100, pp. 5445–5461, 2022.

[22] M. Yandouzi et al., "Forest Fires Detection using Deep Transfer Learning," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 8, pp. 268–275, Oct. 2022, doi: 10.14569/IJACSA.2022.0130832.

[23] M. Berrahal et al., "Enhancing the medical diagnosis of COVID-19 with learning-based decision support systems," Bull. Electr. Eng. Informatics, vol. 13, no. 2, pp. 1316–1324, Apr. 2024, doi: 10.11591/eei.v13i2.6293.

[24] S. Kardakis, I. Perikos, F. Grivokostopoulou, and M. Paraskevas, "Emotion Recognition using Deep Learning Techniques and Word Embeddings," Proc. - 2021 IEEE/ACIS 6th Int. Conf. Big Data, Cloud Comput. Data Sci. BCD 2021, pp. 31–36, Sep. 2021, doi: 10.1109/BCD51206.2021.9581748.

[25] A. Berroukham, K. Housni, M. Lahraichi, and I. Boulfrifi, "Deep learning-based methods for anomaly detection in video surveillance: a review," Bull. Electr. Eng. Informatics, vol. 12, no. 1, pp. 314–327, Feb. 2023, doi: 10.11591/EEI.V12I1.3944.

[26] L. Cheng, X. Liu, L. Li, L. Jiao, and X. Tang, "Deep Adaptive Proposal Network for Object Detection in Optical Remote Sensing Images," Jul. 2018, Accessed: Mar. 31, 2024. [Online]. Available: https://arxiv.org/abs/1807.07327v1.

[27] M. Maity, S. Banerjee, and S. Sinha Chaudhuri, "Faster R-CNN and YOLO based Vehicle detection: A Survey," Proc. - 5th Int. Conf. Comput. Methodol. Commun. ICCMC 2021, pp. 1442–1447, Apr. 2021, doi: 10.1109/ICCMC51019.2021.9418274.

[28] J. Kim, J. Kim, and J. Cho, "An advanced object classification strategy using YOLO through camera and LiDAR sensor fusion," 2019, 13th Int. Conf. Signal Process. Commun. Syst. ICSPCS 2019 - Proc., Dec. 2019, doi: 10.1109/ICSPCS47537.2019.9008742.

[29] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "R-CNN: Regions with CNN features," Proc. ieee Conf. Comput. Vis. pattern Recognit., 2014, [Online]. Available: http://gwylab.com/pdf/rcnn_chs.pdf.

[30] "R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms | by Rohith Gandhi | Towards Data Science." https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e (accessed Jan. 16, 2024).

[31] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2015, doi: 10.48550/arxiv.1506.01497.

[32] M. C. Roh and J. Y. Lee, "Refining faster-RCNN for accurate object detection," Proc. 15th IAPR Int. Conf. Mach. Vis. Appl. MVA 2017, pp. 514–517, Jul. 2017, doi: 10.23919/MVA.2017.7986913.

[33] C. Li et al., "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," Sep. 2022, Accessed: Mar. 22, 2023. [Online]. Available: https://arxiv.org/abs/2209.02976v1.

[34] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv Prepr. arXiv2207.02696, 2022.

[35] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8," 2023. https://github.com/ultralytics/ultralytics (accessed Mar. 30, 2023).

[36] M. Miloš and P. Mirosavljević, "Analysis of the Performance and Kinematics of the Movement of UAV," FME Trans., vol. 51, no. 4, pp. 627–636, Jan. 2023, doi: 10.5937/fme2304627M.

[37] M. Hassanalian and A. Abdelkefi, "Classifications, applications, and design challenges of drones: A review," Prog. Aerosp. Sci., vol. 91, pp. 99–131, May 2017, doi: 10.1016/J.PAEROSCI.2017.04.003.

[38] "Home - Remedies." https://remedies-for-ocean.eu/ (accessed Apr. 01, 2024).

[39] S. PAN, K. YOSHIDA, A. S. BONEY, and S. NISHIYAMA, “the Application of Drone-Assisted Deep Learning Technology in Riverbank Garbage Detection,” J. Japan Soc. Civ. Eng. Ser. B1 (Hydraulic Eng., vol. 78, no. 2, p. I_133-I_138, 2022, doi: 10.2208/jscejhe.78.2_i_133.

[40] M. D. Iordache et al., “Targeting Plastics: Machine Learning Applied to Litter Detection in Aerial Multispectral Images,” Remote Sens., vol. 14, no. 22, 2022, doi: 10.3390/rs14225820.

[41] G. Gonçalves, U. Andriolo, L. Gonçalves, P. Sobral, and F. Bessa, “Quantifying marine macro litter abundance on a sandy beach using unmanned aerial systems and object-oriented machine learning methods,” Remote Sens., vol. 12, no. 16, pp. 1–19, 2020, doi: 10.3390/RS12162599.

[42] G. Jakovljevic, M. Govedarica, and F. Alvarez-Taboada, “A deep learning model for automatic plastic mapping using unmanned aerial vehicle (UAV) data,” Remote Sens., vol. 12, no. 9, 2020, doi: 10.3390/RS12091515.

[43] Y. Taddia, C. Corbau, J. Buoninsegni, U. Simeoni, and A. Pellegrinelli, “Uav approach for detecting plastic marine debris on the beach: A case study in the po river delta (italy),” Drones, vol. 5, no. 4, pp. 1–18, 2021, doi: 10.3390/drones5040140.

[44] A. Aghasi, A. Rai, and Y. Xia, “A Deep Learning and Image Processing Pipeline for Object Characterization in Firm Operations,” https://doi.org/10.1287/ijoc.2022.0260, Dec. 2023, doi: 10.1287/IJOC.2022.0260.

[45] X. Hao, L. Liu, R. Yang, L. Yin, L. Zhang, and X. Li, “A Review of Data Augmentation Methods of Remote Sensing Image Target Recognition,” Remote Sens. 2023, Vol. 15, Page 827, vol. 15, no. 3, p. 827, Feb. 2023, doi: 10.3390/RS15030827.

[46] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.

[47] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015.

[48] “Darknet: Open Source Neural Networks in C.” https://pjreddie.com/darknet/ (accessed Jan. 02, 2024).

[49] “Mavic Air - Product Information - DJI.” https://www.dji.com/global/mavic-air/info#specs (accessed Jan. 18, 2024).

[50] “Serveur au format rack Dell PowerEdge R740 : serveurs | Dell France.” https://www.dell.com/fr-fr/shop/cty/pdp/spd/poweredge-r740/per7403r (accessed Apr. 01, 2024).

[51] “Average Precision | CloudFactory Computer Vision Wiki.” https://wiki.cloudfactory.com/docs/mp-wiki/metrics/average-precision (accessed Jan. 05, 2024).