Performance Evaluation of Underwater Plastic Detection Model Under Diverse Environmental Conditions

MSc Research Project
Artificial Intelligence

Hemaraju Padavala
Student ID: x22185011

School of Computing
National College of Ireland

Supervisor: Muslim Jameel Syed

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ....Hemaraju Padavala.................................

Student ID: ....x22185011.......................................

Programme: .... Artificial Intelligence......... **Year:** ....2023......

Module: ....MSc Research Project...........................

Supervisor: .... Muslim Jameel Syed.............................

Submission Due Date: ....31/01/2024...................................

Project Title: Performance Evaluation of Underwater Plastic Detection Model Under Diverse Environmental Conditions..................

Word Count: ....13378............ **Page Count:** ....34............

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ....P. Hema Raju...

Date: ....31/01/2024..

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Performance Evaluation of Underwater Plastic Detection Model Under Diverse Environmental Conditions

Hemaraju Padavala
x22185011

Abstract

Protecting the environment from ocean plastic waste requires effective detection and quantification. The number rises. This is necessary for environmental protection. Neural network-based object detectors, a recent computer vision advancement, may automate aquatic plastic detection. This computer vision innovation is new. We'll examine CenterNet HourGlass104, Faster R-CNN, and YOLOv3, three popular object detection models. This study tests these models underwater. This research examines model application.

These models were tested with a large dataset that replicated underwater conditions to determine their adaptability and conservation potential. This tested whether these models could aid conservation. This list includes plastic debris with distinct traits. Different lighting and colour conditions are included.

Our findings showed neural network object detectors' marine conservation potential. Testing showed that the CenterNet HourGlass104 was the most accurate and versatile plastic contamination detector. This applied to plastic contamination detection. Faster and more accurate plastic detection may improve cleanup efforts but harm the environment and economy. This is true even in harsh environments like underwater. These implications have major effects on real-world applications.

However, these models' limitations must be acknowledged, especially in microplastic detection and high-turbidity situations. Computer vision-based systems may be optimised and used in other ways in future research. This is because these systems' environments change.

Computer vision can detect ocean plastics, but this study highlights knowledge gaps that need further research. The study proved computer vision was possible. Due to our situation, we can help fight ocean plastic pollution. Provide detailed information about CenterNet HourGlass104, Faster R-CNN, and YOLOv3's pros and cons.

1 Introduction

Since they cover 70% of the planet, oceans have long awed people. The plastic pollution problem persists. By the mid-20th century, plastic was used worldwide in manufacturing due to its durability and versatility. Around then, the crisis started. Millions worldwide chose plastic. Plastic is useful, but its abundance in our oceans is dangerous. Wide range of plastic uses. Marine plastic pollution is the worst environmental disaster. Twelve million metric tonnes of plastic pollution harm 800 marine species. There is no marine exploration from the Arctic to the Pacific. This widespread pollution threatens ocean-dependent communities' food, culture, and livelihood. Ocean livelihood, cultural heritage, and subsistence impact tourism, fishing, and shipping.

Everyone agrees this issue is important, but there are many challenges. Traditional methods for tracking and measuring marine plastic debris lack speed, accuracy, and coverage. This includes well-intentioned methods. Microplastics and ocean size exacerbated these issues. Consider AI implementation here. AI and machine learning have transformed medical and financial sectors in the past decade. AI-powered object detection could aid marine conservation. Asks, "To what extent can AI-driven object detection

models transform underwater plastic identification in diverse marine ecosystems?" This research seeks answers. Our research seeks an answer.
This study investigates and applies object detection models.
We test CenterNet HourGlass, Faster R-CNN, and YOLOv3 for underwater marine plastic identification. Model efficacy is assessed. This comparison compares field performance of algorithms.
Method Development: A reliable experimental method tests these models in variable and unpredictable marine environments. Model performance is evaluated.
Model accuracy, adaptability, and limitations are tested in real-world marine environments. Step four is practical testing.
Computer vision and marine ecology can be combined. We want to improve science and AI-driven conservation. We strive for it. Knowing these models in the dynamic marine ecosystem will help us achieve this. This study impacts marine conservationists, policymakers, and communities worldwide.
Though limited, research improves our intellect. In harsh maritime environments or with microplastics, object detection models may fail. Alternative settings may work better. These concerns can inform environmental and technological innovation research. The data can guide future research.

2 Related Work

Improved object detection methods have emerged recently. These advanced object detection methods evolved from manual feature-based methods. Uses R-CNN and Fast R-CNN deep learning. Past methods required difficult backgrounds, poor lighting, and careful feature engineering. Early methods failed, but CNNs automatically learned discriminative features from raw data. There was significant progress. Previous issues were fixed. CNN hierarchical architecture and feature extraction changed object detection. They were important.
Our environment changes frequently, so we need the more efficient R-CNN model. The architecture uses Fast R-CNN and Region Proposal Network. Fast R-CNN classifies objects, but RPNs propose high-quality regions. R-CNN classifies fast. This literature review discusses anchor selection, RPN proposal generation, and Faster R-CNN multi-task loss. Readers find the topic simpler.
The more effective R-CNN algorithm has drawbacks. Too few negative samples compared to positives may hurt the model. Hard negative sample mining and alternating training will fix this. Both options are planned. Both options may work. Another scenario involves RPN and Fast R-CNN sharing convolutional layers. In scenario 1, the model is retrained for classification. This enhances classification accuracy. A detailed explanation of these strategies shows their huge impact on computing efficiency, enabling real-time object detection, a major advance. The methods detect objects in real time.
This review assesses algorithm recall and precision. The two metrics are considered throughout. Faster R-CNN improves object finding with feature sharing training and hard negative samples. Implementing these changes is the paper's focus. Paper focuses on strategy development. Several famous simulations have shown the algorithm's pros and cons. Though slower than YOLO and SSD, it recognises objects precisely. Despite its slower speed.
Recent object detection advances using the revolutionary Faster R-CNN model are discussed in this paper. Using this model is key. Its significance is stressed. In a changing field, the suggested algorithm improves detection precision and accuracy with hard negative samples and alternating training. This helps in a changing field. **(Liu, 2018)**

Qiuyao Hu and colleagues' paper improves computer vision and object detection and deserves academic attention. Computer vision applications like surveillance, HCI, and autonomous systems require object detection. The authors report improved deep learning-based object detection.
Hu et al. improved Faster R-CNN, a reliable object detection model. Instead of VGG, they use a graph convolutional network (GCN) for speed and accuracy. Graph correspondence networks (GCNs) seamlessly extract spatial features and incorporate semantic data into a graph-based framework, improving the algorithm's ability to analyse complex scenes with multiple objects.

This architectural modification is needed because VGG's native convolutional operations can't capture complex spatial relationships and high-level feature abstractions. GCNs represent features more accurately and insightfully using graph-based structures that match image semantics.
The paper discusses empirical validation—essential to reliable research. Microsoft COCO and PASCAL VOC benchmarks improve with the new algorithm. These datasets boost detection accuracy by 2% and 4% and speed by 6.25 and 12.5%. These results reduce computational redundancy for real-time applications and show incremental progress.
A literature review should include Hu et al.'s computer vision research. Their new graph-based GCN-based Faster R-CNN architecture improves object detection speed and accuracy. Graphical models and deep learning paradigms are expanding computer vision technology by changing object detection methods**. (Hu et al., 2022)**

This paper shows a computer vision and object detection method for finding plastic trash in the ocean. The algorithm makes use of a better YOLOv5n model. Underwater detection is hard and takes a lot of computing power because there isn't enough light and the conditions are complicated. The proposed algorithm fixes these problems by making a number of important changes. In the backbone network, contextual binary depthwise convolution (CB2D) lowers model parameters and makes detection better. The Contextual Nested and Enhanced Blocks (CNeB) module is added to the algorithm to fix the feature pyramid feature shortage for small underwater waste targets for waste. In the paper, the alpha-IoU loss function makes bounding box regression loss and accuracy better.
The improved YOLOv5n model for finding plastic trash underwater is backed by a lot of real-world evidence. The algorithm improves average precision (AP) by 12.25% and AP50 by 13.69% over the original model when trained in the same way. The model's 3.2MB size makes detection more accurate without making it bigger. Finding underwater trash is very important, and this paper takes a close look at old methods to show what problems they had. These changes are needed for underwater garbage target detection. Also, an analysis of ablation shows that the proposed improvements make detection much more accurate.
The YOLOv5n model, which finds plastic trash in water, is also looked at in terms of its architecture and how well it integrates key parts. This study goes into detail about how to integrate the CB2D, CNeB, and alpha-IoU loss function modules. Modules make it easier to get features, get the loss function to converge, and predict the boundaries. This paper ensures openness and reproducibility by carefully looking at the infrastructure for software and hardware, the make-up of the datasets, and the way the experiments were set up.
Aside from tests done on the algorithm itself, it is also regularly compared to object detection models. The improved YOLOv5n model makes it easier to use and better at finding plastic trash underwater. The model worked well in this case, as shown by the results. After the paper, there will be academic debate about how to improve the algorithm and use it in underwater vehicles with few resources. In this paper, the YOLOv5n model is made better to solve the hard problems of finding underwater trash targets**. (Hu & Xu, 2022)**

This paper examines ocean pollution, particularly plastic trash in marine ecosystems. Heavy plastic pollution in the deep sea is dangerous. The paper claims 80% of marine debris is plastic. Plastic trash acidifies, speeds algae growth, and depletes oxygen in the ocean. These changes disrupt marine ecosystems, affecting many marine species.
The study examines plastic trash in remote ocean areas, like garbage patches. The statement says ghost nets and other plastic trash harm marine ecosystems and the global financial system. Turtles suffer from eating or being caught in plastic trash. This section shows how plastic trash harms marine systems and animals.
This study investigates marine debris detection and removal to reduce ocean pollution. Deep-sea cameras and drones will transform surveillance and cleanup. Seabed waste identification and removal may be easier with new tools. Submerged cameras can send live video feeds to decision-makers to speed up decision-making, according to this study. In this study, deep learning neural networks and CNN image analysis methods are used to distinguish trash from ocean life.
This essay addresses legal and technological ocean pollution solutions. The text discusses US legislation to recover marine debris, reduce plastic waste, and promote recycling. This essay examines the complex

relationship between ocean health and coastal community strength. It stresses international cooperation to reduce marine debris. This law emphasises international cooperation to combat ocean pollution.
Marine ecosystems should be protected, says the study. The statement says only technology and laws can stop ocean pollution. The importance of marine debris detection and removal is stressed in this essay. It protects marine ecosystems and reduces plastic. **(Alshibli et al., 2022)**

This research paper looks at the problems that come up with real-time object detection and how neural networks, especially CNNs, can help solve them. This paper shows how object detection algorithms have changed over time, from R-CNN to Fast R-CNN to the Faster R-CNN model. The above changes have made it much faster and more accurate to find objects. It is well known that Faster R-CNN can guess an image in 0.2 seconds and get 70% mean average precision (mAP) on test sets like PASCAL VOC 2007 and 2012.
This essay compares R-CNN, Fast R-CNN, and Faster R-CNN in depth, showing their features and how well they work. The study shows that Fast R-CNN works better than R-CNN because it doesn't need 2000 region proposals for the CNN model. Faster R-CNN does better than Faster R-CNN because it uses a Region Proposal Network (RPN) instead of the selective search algorithm, which takes a lot of computing power. This makes the accuracy and processing speed better. The paper also talks about problems with adding more data to real-time object detection and suggests new ways to fix them using convolutional neural networks.
This report looks at the Region Proposal Network (RPN) and checks how it helps make object detection more accurate by coming up with quick and correct region proposals. Faster R-CNN gets 5 frames per second, which shows that it uses computers efficiently when compared to others. Using the PASCAL VOC 2007, 2012, and MS COCO datasets, the study also carefully tests the model. The tests above all show that a faster R-CNN is a faster and more accurate object detection model. This research looks into difficult problems with finding objects, such as size, i.e., small, medium, deformation, and occlusion. The study examines the use of fast R-CNN for detecting faces and objects in video clips.
The paper looks at model performance metrics, with a focus on the values of intersection over union (IoU) and mean average precision (mAP). They also look at how these metrics change the quality of proposals for the VGG-16 and ZF-Net models. It is clear that region proposal networks (RPNs) work better than selective search. In this paper, object detection algorithms are looked at in great detail. Faster: The best things about R-CNN are its speed and accuracy. The study shows how important neural networks are and how important it is to have good algorithms for finding objects. It talks about the advantages of the faster R-CNN algorithm. (**Bhatlawande et al., 2022)**

Using Fast R-CNN, the paper finds and names fish in underwater pictures. Huge pictures taken underwater in the oceans show that we need an automated system. It would be hard and take a lot of time to look through this huge collection of pictures by hand, which shows how much we need automation. There aren't many accurate ways to find and sort fish in underwater pictures.
This essay takes a critical look at Fast R-CNN, which might be better at finding things quickly and accurately than DPM and R-CNN. It makes it much easier to find fish. The paper shows that Fast R-CNN is faster than R-CNN by comparing them. Slower than Fast R-CNN, other methods find and store 80 images. Fast R-CNN should be able to handle problems with underwater imaging thanks to its agility. So, it's perfect for an automated system that collects underwater pictures in the best way for real-life use.
This school paper displays the Fast R-CNN system's ability to find and identify fish in underwater pictures. A lot of underwater pictures need to be taken automatically. Fast R-CNN is a promising solution because it is fast and good at finding things. Fast R-CNN might also be able to work underwater, which would make it useful for real-world uses. (**Xiu Li et al., 2015)**
A cooperative system using SSS images and YOLO-v3 detects underwater objects in this study. Combining optical and acoustic methods for underwater target detection is the goal. Our database contains real-side scan sonar images. It contains 7,000 true underwater samples. This novel optical and acoustic system detects underwater targets accurately. The versatile YOLO-v3 network was the core algorithm, especially for real-time detection.
This paper reviews underwater object detection literature and shows that labelless datasets fail supervised algorithms. This review shows that deep learning is a popular problem-solving method. This

research compares underwater target detection algorithms and highlights issues. These include low contrast, underwater blurring, and colour distortion. These issues are caused by sediment and aquatic organisms.

A 24 convolutional layer, 2 fully connected layer underwater object detection network, YOLO-v3, is described in the paper. YOLO-v3 is an engineering solution that uses object scores and coordinates to calculate loss. Extensive testing supports YOLO-v3 underwater object detection. The YOLO-v3 underwater object detection network is empirically validated in this study.

This paper impacts non-academic autonomous underwater detection systems. Demand for automated underwater sonar image processing and analysis prompted this approach. Real-world underwater experiments are shown in this paper to prove YOLO-v3. Testing shows this method can detect underwater objects. They thank the National Key Research and Development Programme of China and the State Grid Corporation Science and Technology Project for their invaluable help.

This research paper concludes with the development, implementation, and testing of a new underwater object detection system. This system uses side-scan sonar and YOLO-v3 to detect underwater objects. Real-time submerged object detection is feasible and accurate. **(Wang et al., 2021)**

Faster R-CNN, a deep learning-based object identification algorithm, is improved in this work. The tale explains its origins, primary features, and advantages over previous versions. Faster R-CNN combines RPN and Fast R-CNN to modify end-to-end object recognition. The paper discusses using ResNet101 and PVANET network architectures to speed up R-CNN implementation. Deep learning system Caffe teaches and evaluates models. Evaluation uses mean average accuracy (mAP). Researchers observed that the quickest PVANET-trained R-CNN had the highest mean average accuracy.

Before Faster R-CNN, this work discusses R-CNN and Fast R-CNN. When it suggested areas using the selective search method, R-CNN began a multi-stage training procedure. Tracking and training were sluggish, making this work difficult. Fast R-CNN's Roi Pooling layer enabled direct CNN training, a major advance. This architecture modification accelerated detection. Faster R-CNN employed convolutional networks and shared the network with object recognition to produce recommended boxes, improving this area. This improvement reduced the needed frames from 2000 to 300, speeding identification and improving bounding box quality.

This study examines the faster R-CNN network structure. How to acquire image feature maps, structure region proposals in a complicated RPN network, and easily incorporate proposal feature maps into the Roi Pooling layer are discussed. Our RPN network suggests softmax classification and bounding box regression branches. The study uses bounding box regression and class probabilities for classification. This entire knowledge illuminates the network's complex operations.

Fast R-CNN training using ResNet101 and PVANET validates the trial system. Experimental findings from PASCAL VOC2007 are evaluated. The study details mean average accuracy and model object detecting effectiveness. This study found that convolutional layers and the RPN module improve Faster R-CNN classification.

This study concludes with R-CNN object recognition speed optimisation. Object identification algorithms are entering a new age as speed, bounding box quality, and performance improve. This prepares for object recognition research. **(Liu et al., 2017)**

This article introduces YOLOv3 and TensorFlow wildlife detection and recognition. The deep learning-based approach classifies animals in digital photos. Ecosystem protection requires wildlife detection, but typical object detection techniques are tough. The article advises adopting YOLOv3, a cutting-edge approach that can identify objects in real time and enhance processing performance, to overcome these concerns.

This study uses 9051 tagged HD animal photography photos. These photos expertly incorporated into the VOC-2012 dataset provide a broad range of diverse and thorough training material for the neural network. The dataset is divided 80/20 for 106-layer neural network training and testing.

This experiment utilises a normal workstation with two GTX 1060 GPUs. This system extensively evaluates 1065 digital photos. YOLOv3 identified animals in four groups with 75.2% mAP in these studies. The less resource-intensive YOLOv3-Tiny has 68.4% mAP and 50% quicker processing. Experimental data suggest learning-based strategies may be advantageous.

This research uses YOLOv3 and TensorFlow to identify and recognise animals. Deep learning overcomes long-standing issues, preserving nature. An empirical experiment confirms the proposed method's accuracy in spotting complicated animals in digital photos. This predicts severe wildlife conservation consequences. (**Gabriel et al., 2020)**

This study found that underwater picture processing and computer vision increase quality. These technologies enable autonomous operations in difficult deep-sea conditions, the report says. This research emphasises numerous underwater object segmentation circumstances. Gliders and submersibles track marine life. This research analyses several underwater object recognition applications and emphasises colour correction in underwater image processing. Drowning detection, species analysis, and vehicle navigation are examples.This study discusses new stacked-CNN architecture. The framework is specifically developed to recognise and animate underwater non-animated things. It contains garbage and rubbish. The architecture skillfully builds bounding boxed training pictures from tagged input photos. The system is simulated and evaluated in MATLAB 2017 utilising quantitative measures including true positive, true negative, false positive, and false negative rates. System importance and effectiveness depend on these parameters.The study discusses underwater photography's biggest obstacles. Light scattering and attenuation in deep oceans alters the colour and contrast of submerged objects. This study addresses underwater object segmentation, colour correction, and feature mapping techniques. Huge undersea object databases support the paper's model.This work emphasises the importance of underwater imagery for ecosystem research and preservation. Due of pollution, the International Maritime Organisation struggles to safeguard seas.This report emphasises marine plastic pollution-reduction technology development and use. Automatic identification models that grasp sea cucumbers' complicated surroundings are becoming increasingly significant. This research describes cheap stereo vision systems for underwater object detection. Improve underwater photos using a new optical model.The study report finishes with several facts that provide a robust architectural foundation for forecasting living and non-living things in the enigmatic underwater regions. Computer vision and underwater image processing increase picture quality, ocean conservation, and pollution control. **(Bhuvaneswari et al., 2022)**

Traffic safety object detection using FPGA and GPU hardware is studied in this article. This study analyses YOLOV3, a prominent real-time traffic object tracking methodology. The study extensively examines YOLOV3 on FPGA and GPU. This large study uses R-CNN iterations, Fast, Faster, SSD, and YOLO deep learning models. This study compares model strengths and downsides.
The study begins with FPGA/GPU object recognition basics. GPUs are convenient and adaptable but set hardware configurations and many GPUs limit their versatility. These variables increase cost and power. More flexible, FPGA cards can develop unique hardware for projects. The paper shows FPGA cards run YOLOV3 quicker and more precisely.
The report argues object recognition requires deep learning models. Every SSD- or R-CNN-based model is examined. Object identification speed and accuracy must be balanced, says this report. Architecture benefits and drawbacks are outlined for each model.
In deep learning, the YOLOV3 model is known for real-time object tracking and rapid, accurate results. YOLOV3 model implementation on FPGA and GPU hardware is studied in this work. Model file transformation, training, and deep learning processing unit coordination are covered.
The report shows the YOLOV3 model's performance on GPU and FPGA hardware at a vital stage of testing with considerable empirical data. FPGA cards improve frame rates, especially with reduced input dimensions.
This research indicates that the YOLOV3 model is crucial for object detection and tracking, especially for critical traffic safety problems. FPGAs outperform GPUs significantly. The study tackles developing quicker YOLOV3-tiny models and larger datasets for object recognition.
Research papers are organised and logical. The work meticulously covers technique, experimentation, empirical findings, and synthesis. The research analyses FPGA and GPU technology, deep learning algorithms, and the YOLOV3 model for real-time object tracking for traffic safety. **(Esen et al., 2021)**

Within city traffic, computer vision study recognises moving objects and autos. YOLOv3 and YOLOv4 deep convolutional neural networks in this study detect several items.

Cities can utilise complex road algorithms. With a solid dataset, this study finds automobiles, trucks, people, and motorcycles. These presentations incorporate photos and films from various lights. RGB/grayscale. This research helps.
The execution is detailed. Resizing images and converting tensors collect and organise data for quicker processing. Network, hyperparameter, model, and measurement require GPUs. We examine outcomes after this thorough approach.
Testing YOLOv3&4. Upgraded Darknet-53 feature extractor YOLOv3 detects many cars. YOLO 3 improved with 4. PAN, spatial pyramid sharing, and CSP darknet-53 accelerate YOLOv4.
We then evaluate project hyperparameters and progress limits. This approach requires TensorFlow, NumPy, Pillow, and Seaborn colours. Leaky ReLU, anchors, batch normalisation, and set padding improve the model.
One study found the method challenging. Results favour autos. Image collection is 98% accurate. Better, 99% accurate movie collection. City-driving YOLOv3 and v4.
Simulations show the models can locate automobiles, trucks, people, and more. Showing algorithms in boxes.
Research has consequences. Self-driving cars, smart city apps, surveillance, and traffic tracking need object identification. Continuously improving computer vision and AI systems' thinking skills supports their significance.
Last, the study examines YOLOv3 and v4 deep convolutional neural networks. This research examines urban road techniques. The well-written book promotes computer vision and intelligent transportation systems via research, testing, and analysis. **(Kumar B. et al., 2020)**

This research paper uses a novel anchor-free method to identify knives and handguns in X-ray baggage security image analysis. Anchor-based object detection is inflexible and computationally intensive. The novel approach addresses these issues.
Six popular anchor-free methods are compared to anchor-based methods in this study. Weapon detection was best by YOLOx, Objects as Points, and ExtremeNet anchor-free methods.
Anchor-free methods can identify weapons in baggage security X-ray images, according to transparent empirical findings. The performance difference proves anchor-free works in critical security.
This empirical study analyzes each anchor-free method's mechanisms and offers technical insights. For those entering this innovative field, the text details data acquisition, training protocols, and assessment metrics.
The study finds that X-ray baggage security images improve weapon detection accuracy and efficiency. Additionally, anchor-free methods may be feasible alternatives to anchor-based methods. Modern security is affected by dynamic object detection methods**. (Huang et al., 2022)**

This work emphasises cloud identification in power output prediction due to the complexity of cloud types and solar power generation. Centre Net-based cloud image identification may improve operational efficiency and detection, says this research.
The study found that cloud cover and type significantly impact PV power. Clouds hamper photovoltaic module direct solar irradiation. A quick cloud cover rise might reduce solar power. This study underlines the necessity to address climate's negative influence on PV power. Power generation estimations need cloud identification.
Pioneer of cloud image recognition Centre Net surpassed regional target recognition. It alters paper. With its anchor-free method and solution domain flexibility, Centre Net speeds object recognition. This article explains Centre Net cloud discovery using heatmaps and candidate box prediction.
This study predicts optimum cloud recognition locations using modified focus loss function. An in-depth analysis of Hourglass Network, Resdcn Net, and DLA-34 cloud recognition architectures. The study confirms its method using simulated trials and meticulous analysis. It recognises cloud types faster and more accurately in real time than previous methods.
Practical Comparisons between Centre Net and Faster R-CNN support the paper's findings. Former recognises faster and more confidently. Finally, the paper analyses how its insights may improve solar power forecast accuracy.
Centre Net technology is proposed after examining how cloud types affect solar power output. A pervasive phenomenon may improve power projections and solar power generation**. (Jing et al., 2020)**

The massive problem of solid waste generation between 2030 and 2050 is the emphasis of this article. Poor waste management harms the environment and people. Additionally, waste filtration offers several issues. Deep learning methods will be examined to find trash categorization solutions. We will also discuss benchmarked waste datasets for model testing.
We'll discuss global solid waste production in 2016. The essay then analyses the possible catastrophic increase in rubbish production between 2030 and 2050. Inefficient rubbish handling harms many things. This activity depletes land resources, endangers humans and animals, and pollutes. This article examines improper waste management, including illegal dumping, in depth.
This research focuses on garbage filtering and deep learning. Most prominent designs are VGG16, EfficientDet, ResNet, MobileNet, and Inception-ResNet. Every waste identification and classification method is examined. Deep learning will be assessed using specific garbage. Several factors will affect our future study.
Waste materials are categorised and investigated in the study project. The research boldly examines these notions' ramifications, limits, and possible challenges. According to the study, installing a greater variety of database systems with different features will help address the problem of insufficient waste data. To successfully categorise and detect junk, you need many samples with different properties.
Finally, this study addresses gaps in waste categorization and monitoring investigations. Professionals assessed an online deep learning model and dataset. This model and dataset were investigated. It was designed to meet educational goals. For trash classification, a lot of data must be analysed. Additionally, it provides a complete literature review to encourage academics to study machine learning and deep learning. **(Abdu & Mohd Noor, 2022)**

The study proposes "LogoNet," a deep learning logo recognition system for difficult settings. Work project offers this way. Many logos and designs exist for these options. Logo recognition is crucial for social media monitoring, brand marketing, autonomous driving, intelligent transportation, market research, and illicit logo detection. Deep learning improved logo detection.
The study compares logo detecting algorithms. SSD and Rapid R-CNN logo detectors are common. Without data, these techniques suffer. Researchers create realistic, reliable synthetic logos using GANs. Growing training data was the aim.
This study describes how anchor-free object recognition and attention-based tactics shaped LogoNet. LogoNet's attention-based backbone network is spatially oriented. We follow architects at CenterNet. LogoNet streamlines logo feature extraction. This study explores thick hourglass networks. The collection's spatial attention module accentuates logos.
LogoNet rocks FlickrLogos-32. LogoNet outperforms sophisticated anchor-free detection networks by 1.5%. LogoNet's efficacy is tested against Faster R-CNN, SSD, and CenterNet. The statistics indicate LogoNet is the best logo detector.
LogoNet outperforms sophisticated algorithms on FlickrLogos-32. The firm boasts the highest mAP (82.2%) and detection. The work centres on FlickrLogos-32plus, an extension. LogoNet's head-detection expertise may help CenterNet.
This study concludes with LogoNet, a logo recognition system that prioritises attention over anchoring. Combining these elements increases performance and thinking subtly. LogoNet raises logo awareness to demonstrate distinctiveness. This is spatial attention. LogoNet simplifies logo identification and detection, demonstrating its ingenuity. (**JAIN et al., 2021)**

This study provides a novel method for detecting targets in complex underwater environments. Submerged objects' complex traits are difficult to depict. An underwater target detection network with high resolution improves target feature representation and adaptability to complex underwater imagery feature distributions.
Network architecture benefits from lightweight, high-resolution human posture estimation. Integration improves target feature representation and image sampling semantic loss. Attention module (A-CBAM) and enhanced flexible rectified linear units (FReLU) activation function capture complex feature distributions. The RFAM captures semantic and positional data. Robust and discriminating features benefit multi-scale underwater targets.

Experiments show the network works. The proposed approach outperforms target detection algorithms with Mean Average Precision (mAP) scores of 81.17%, 77.02%, and 82.9% on three publicly available datasets. Accuracy, compactness, and fast detection make this model effective. The proposed network has promising underwater target detection qualities.
The paper analyses besides experiments. The study compares popular target detection algorithms, performs ablation experiments to analyse model contributions, and visualises results. Our thorough evaluation improves the algorithm's underwater target detection.
Anchored and anchor-free target detection models are examined in this study. Here are the fundamental detection performance-generalization trade-offs. The study examines underwater target detection algorithm advances to address underwater image analysis challenges. It examines federated learning, blockchain, and IoT in underwater robotics, exploration, and marine research.
A breakthrough parallel high-resolution underwater target detection network is proposed in this paper. Underwater target detection research begins with innovative architecture and promising experiments. The study solves a long-standing problem and highlight the need for more research to improve the model, increase underwater image datasets, and find lightweight processing methods to speed detection. Improvements will improve underwater target detection**. (Bao et al., 2023)**

Object detection is needed for self-driving cars and image-based search. Finding things is important for both. A popular object detection framework called CenterNet is talked about in this article.
CenterNet's innovative object detection is one of a kind. Recall and accuracy are better when each object is shown as a triplet of keypoints instead of the usual way of doing things. CenterNet gets around anchor-based methods by using custom modules to improve corner and centre data in suggested areas. It helps CenterNet get around the problems with anchor-based methods.
CenterNet does a good job. This detector is faster and works better than all one-stage and premium two-stage detectors. 47.0% average precision (AP) on the hard MS-COCO dataset is very good.
Cascade corner pooling, centre pooling, and specialised modules make CentreNet work. These modules make suggested region corner and centre information better, which lowers the number of mistakes made when predicting bounding boxes based on keypoints.
Despite its flaws, CenterNet does a great job. Anchor-based methods are common, but they have some problems. Before CenterNet, CornerNet switched anchor boxes for corner keypoints. CornerNet came before CenterNet. However, it was hard to find global object information, which led to wrong predictions for the bounding boxes.
CenterNet's main new idea was to look at suggested region visual patterns using triplets of keypoints. Central and cascade corner pooling made it easier to find the edges of a box by giving us more information about the area. CenterNet quickly makes heatmaps for the middle and edges, which lets it find objects accurately and on a large scale. This makes it easier to remember and aim for all objects, but it works best for smaller ones.
CenterNet makes it easier to find objects in general. Computer vision has come a long way thanks to better performance, specialised modules, and object representation. CenterNet's work on developing the object detection domain could help other applications, which would promote research and creativity. It will make you think of things. (Duan, 2019)

2.2 CONCLUSION

I want to learn more about computer vision and object detection after reading academic articles. With each article, computer vision research and use get better.
The articles explain how to find objects in the modern world. Creating, improving, and mass-producing object detection algorithms is given the most attention. The YOLO series and Faster R-CNN were made as part of the project to improve accuracy and processing speed. Discovery is interesting.
To protect marine ecosystems, the papers stress how important it is to find and remove plastic from the ocean floor. Pollution in the ocean hurts ecosystems and the environment. Academic journals say that advanced technology, deep learning algorithms, and specialised instruments can be used to find and reduce marine pollution. Researchers use UAVs, deep learning algorithms, and special cameras to look for and get rid of trash in the ocean.

To protect seagrass beds and green spaces, more research is needed. Seagrass meadows can be found by computers. It's faster with NASNet R-CNN. Intelligent systems help with the study of the environment. As part of this, they are studying seagrasses and how they work in marine ecosystems.
Problems and successes with transfer learning in the task and appearance domains are also talked about. The best results will come from source datasets that perfectly match target domains. The results prove that learning transfer between image domains is important. In computer vision, it's easy to send data.
Photovoltaic methods for figuring out cloud cover are also looked at. This claim shows how recognising clouds changes how weather forecasts are made. CenterNet can tell the difference between clouds, as this study shows. Object-discovery technologies are used in an unexpected way in this application.
The papers suggest using benchmarked datasets to train and test deep learning models for sorting and finding waste. Researchers in waste management want to find standards.
In conclusion, the research papers above look at new and existing uses and developments in object detection and computer vision. The sentence suggests that protecting and managing the environment should be done in a creative and effective way. Current projects and plans for the future are talked about. The study of computer vision is good for both people and nature.

3 Research Objectives and Questions

3.1 Objectives

- To test and compare how well CenterNet HourGlass104, Faster R-CNN, and YOLOv3 can find trash made of plastic in underwater environments, taking into account things like light, cloudiness, and the types of plastic that are present.

- To see how well some object detectors based on neural networks can adapt to marine environments that are always changing. We are going to test how well they can generalise and keep their high levels of accuracy in a lot of different underwater situations.

- To make the models work better in places with few resources, like on self-driving underwater vehicles, so they can be used in projects to protect the ocean. The main goals are to make detection faster, more accurate, and better able to work in these conditions for the models.

- To figure out what the main issues and limits are with the way object detection models work now when it comes to correctly identifying and classifying microplastics and when there is a lot of turbidity, and to suggest ways to make these issues better or other choices that can be made.

- To figure out how to put together the best neural network-based object detection models into a system for keeping an eye on marine debris that uses advanced computer vision techniques and other sensor technologies. The goal is to make a solution that can be used all over the world for marine conservation that is reliable and can grow as needed.

3.2 Questions

- The Models YOLOv3, Faster R-CNN, and CenterNet HourGlass104, can find plastic in under-water. How do they compare plastic and non-plastic trash and detect?

- How do these models not work well to find and sort small plastic objects in clumsy underwater? What do we do to detect those plastic trash?

- In a larger ocean system, how can computer vision and other sensors be used most effectively to find plastic trash in the ocean? It is very important for marine conservation to have a system that works well and can expand as needed to detect the underwater plastic?

- How can neural network-based object detectors be made better so that they work better in places like the underwater oceans where water changes slowly and conditions change?

4 Research Methodology and Design

This study uses advanced machine learning to find trash in the ocean, which is not easy to do. Steps and choices in the method are looked at in this part. Check out how we planned and carried out our study on using complex AI models to find marine debris. You need to know this in order to understand the challenge it faces. Full analyses are done on this study at every stage of the research process. It also looks at ways to analyse and prepare datasets, test cases, evaluation parameters, and the unique architectural features of models that have been deployed. More study needs to be done on environmental machine learning to make sure the study is solid and true and to find out more about it.

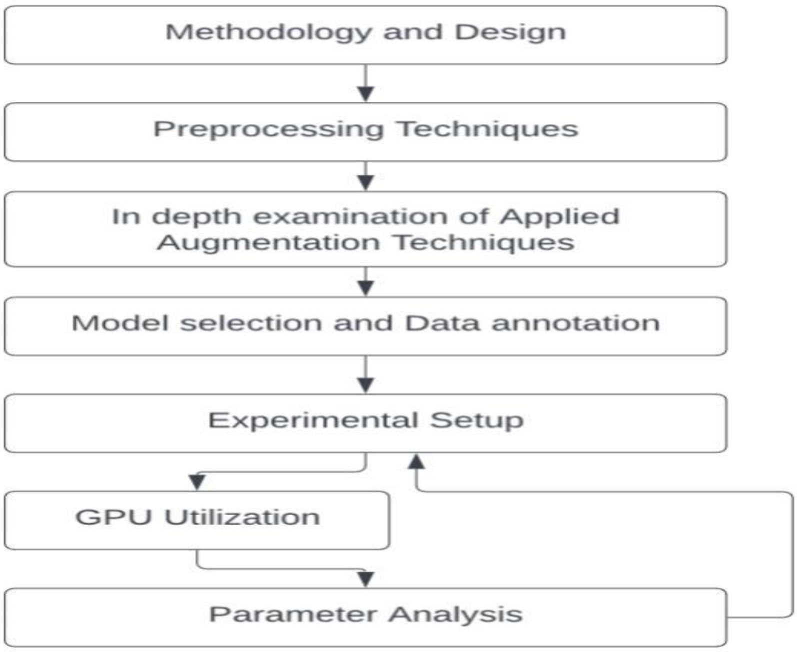


Figure 1 Methodology Design Overview

4.1 ENHANCED DATASET EXAMINATION FOR RESEARCH ON MARINE PLASTIC DEBRIS:

4.1.1. Classification of Plastic Debris

The dataset has a distinct emphasis on a singular class, which includes a wide range of plastic debris typically encountered in marine ecosystems. The use of a singular classification approach streamlines the process of training the model as it eliminates the need to differentiate between various subcategories of plastic waste.

4.1.2. Distribution and Expansion of the Dataset

The dataset started with 6,000 images. These pictures use a variety of plastic trash.
The dataset now has 15,500 images after extensive preparation and addition. Larger datasets with more data types and stability help machine learning models learn. Because larger datasets are better.

Class distribution: same. The dataset shows all plastic trash equally because it is evenly distributed. Divide classes according to the general rule. Due to its uniform distribution across a plastic waste category, this dataset doesn't require up and down sampling or other class-specific balancing methods.

4.1.3. Source of the Dataset

- Images are from JAMSTEC's Deep-sea Debris Database, featuring ocean trash images and videos. Selected this database because it offers diverse information and great visual content, making it an excellent dataset base.

4.1.4. The Rationale Behind Using this Dataset

- The selection of this dataset from JAMSTEC was based on its diverse composition and ability to accurately represent various forms of marine plastic debris, such as plastic bags, bottles, and a lot more. This dataset offers a comprehensive and diverse foundation for the training and evaluation of artificial intelligence models designed to detect and classify marine debris.

4.1.5. Benchmarking and Comparative Analysis

- This dataset can serve as a standard in marine debris detection research because it covers a wide range and has been meticulously compiled. This method paves the way for more research to come. In the same field, it gets easier to compare metrics and standardize methods. This plan helps to reach the research goal.

4.1.6. Accessibility of the Dataset

- Users can access the primary deep-sea debris dataset through the JAMSTEC Deep-sea Debris Database. Click here to access this dataset:
- [http://www.godac.jamstec.go.jp/catalog/dsdebris/e/index.html].
- Augmented Data Accessibility: A scholarly publication or data repository usually discloses the augmented dataset relevant to this study. This is usually the case. Because the study uses the augmented dataset, this is true. Thus, subsequent research and validation are possible, which would not have been possible otherwise.

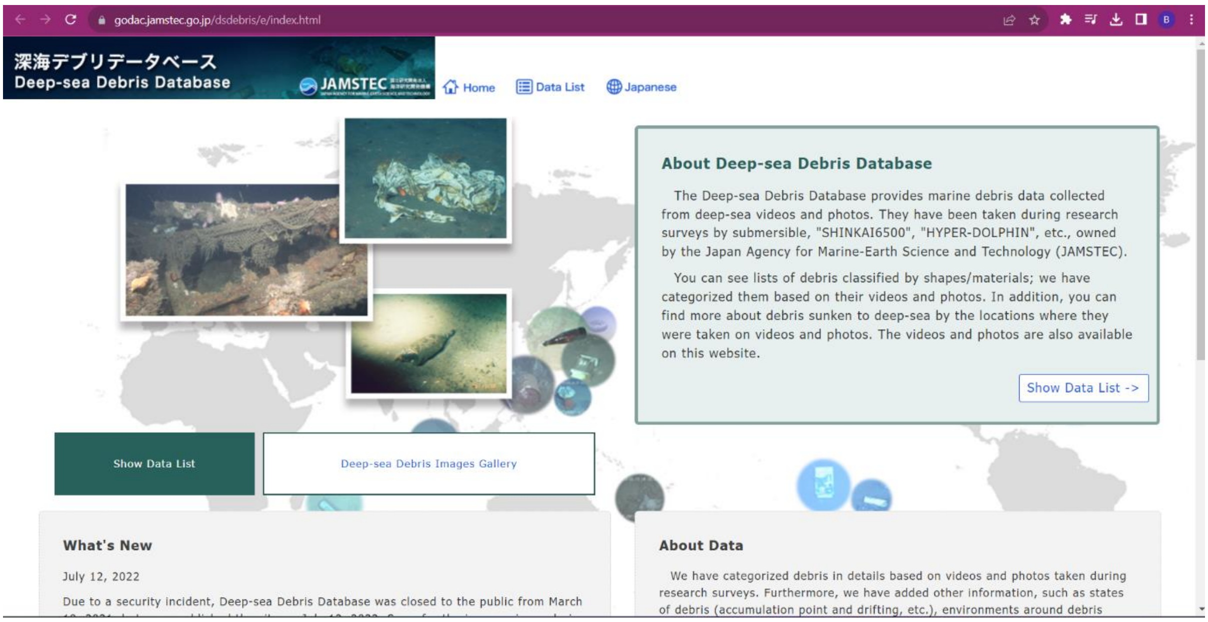


Figure 2 Deep Sea Debris Database Homepage of JAMSTEC

4.2 ENHANCED ANALYSIS OF PREPROCESSING TECHNIQUES

4.2.1. Auto Orient and Resize

In machine learning, where pattern recognition is crucial, the auto-orient feature keeps text orientation consistent across the dataset. This is why this feature is crucial. The most thorough analysis shows that this trait is essential. To standardize input image sizes, resize them to 640 by 640 pixels. This is possible due to image resolution. This step ensures that model inputs are consistent. However, recognition may accidentally change aspect ratios, which could affect the model's ability to identify and classify objects. Due to potential distortion, marine debris dimensions and configuration, which are crucial for identification and classification, may be misinterpreted. Because both traits are essential.

4.22. Auto-Adjust Contrast Using Adaptive Equalization

Enhanced Significance: Underwater imaging has poor and variable lighting, so adaptive equalization improves local contrast. This method improves image nuance detection, helping the model learn faster. Identifying marine debris, especially those with a variety of materials and colors, requires enhanced contrast.

4.3 In-Depth Examination of Applied Augmentation Techniques

4.3.1. Flip Augmentation (Horizontal, Vertical)

Due to the inherent unpredictability of debris orientation in marine environments, it is necessary to incorporate horizontal and vertical flips into the image dataset. The presence of variability is crucial in a comprehensive training regimen as it allows models to effectively identify and categories debris, irrespective of its orientation.

Figure 3 Flip Augmentation

4.3.2. Rotation Augmentation (90° and Fine Rotations)

The integration of both 90-degree rotations and more precise rotational adjustments in the approach effectively tackles the practical situation in which marine debris is seldom encountered in a perfectly aligned orientation. The expansion of the model's detection capabilities enhances its reliability in a wide range of underwater conditions.

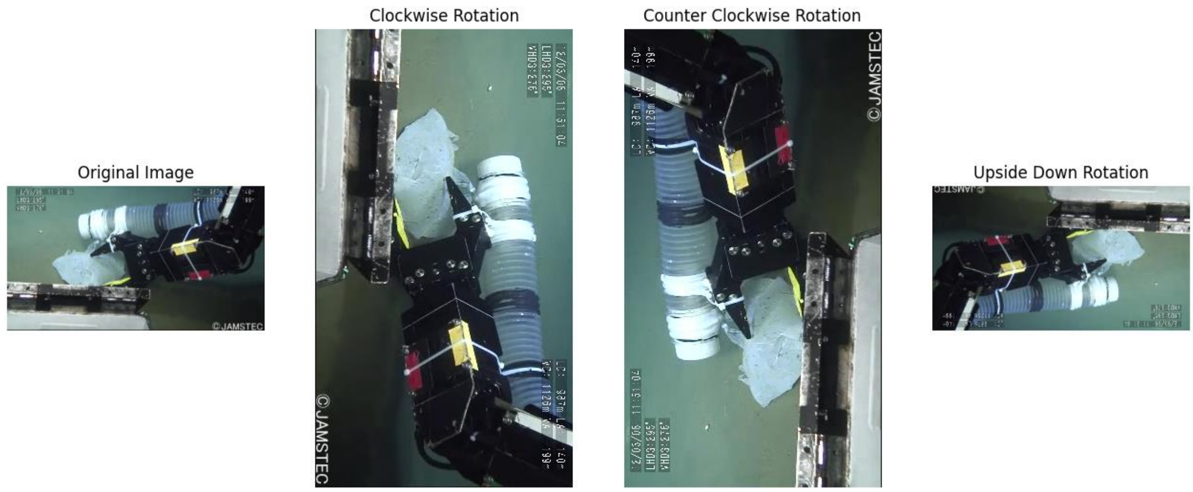


Figure 4 Rotation Augmentation Example

4.3.3. Shear Transformation (±15° Horizontal, ±15° Vertical)

Shear transformations are a more advanced way to copy the changing and often distorted perspectives of underwater photography. This method teaches models to recognize debris from different positions and angles, which makes it easier to find things in underwater environments that aren't simple.

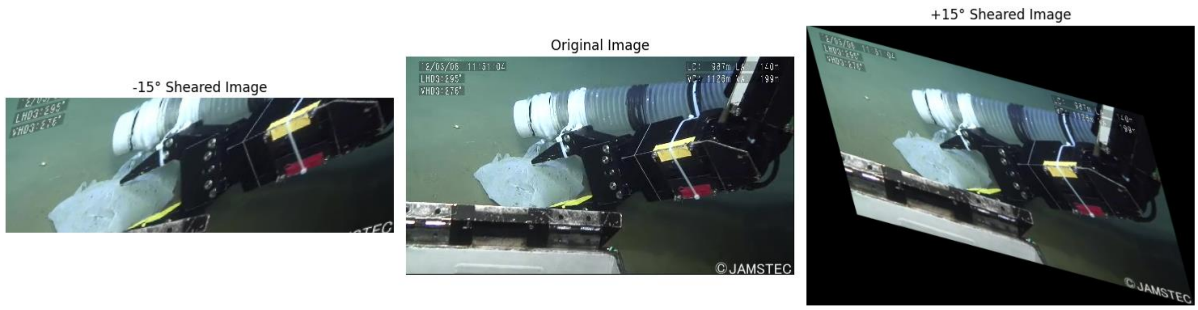


Figure 5 Shear Transformation Augmentation Example

4.3.4. Grayscale Conversion

Color loss and degradation at different ocean depths were solved by turning some of these pictures grayscale. This step prepares the model for non-critical color information. This lets the model find objects in different water conditions.

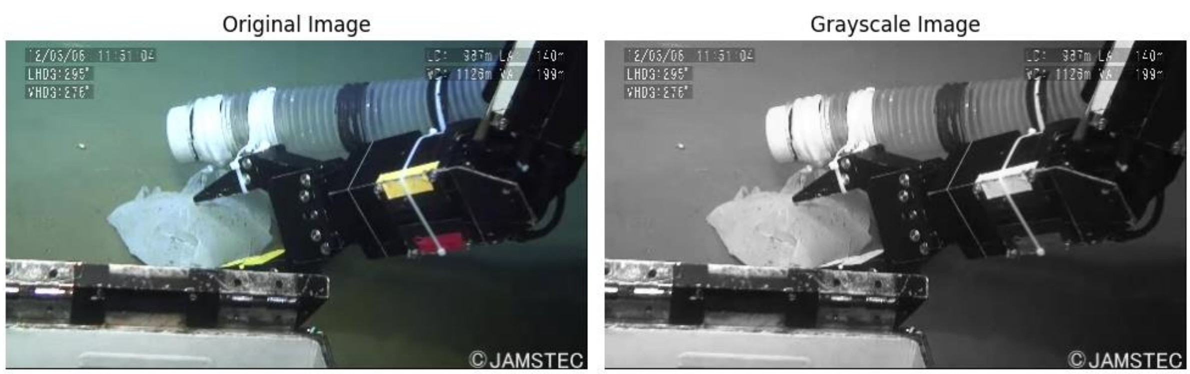


Figure 6 Gray Scale Augmentation Sample

4.3.5. Adjustments in Saturation, Brightness, and Exposure

By adjusting saturation, brightness, and exposure, this dataset shows how unpredictable underwater lighting is. A detailed explanation follows. Model functionality must be improved in different visibility conditions to make it more applicable to real-world situations.

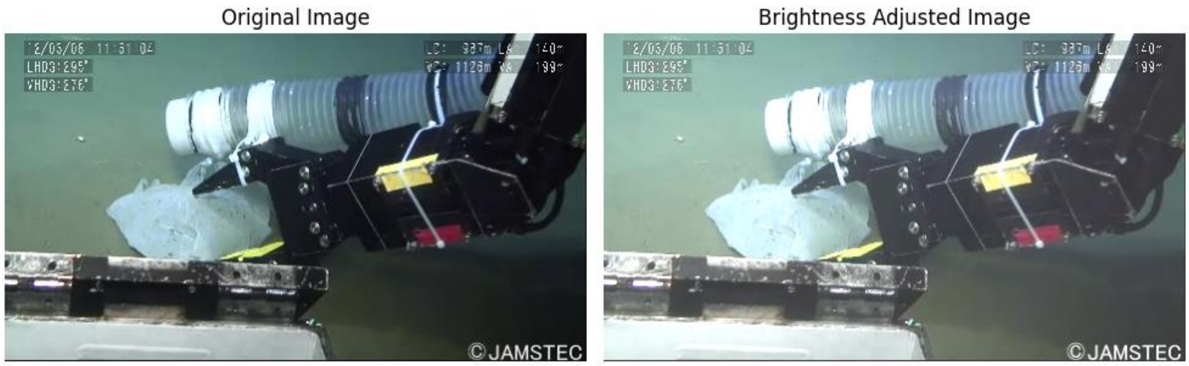


Figure 7 Brightness Adjustment sample

4.3.6. Introduction of Image Noise

This dataset was strategically enhanced with controlled noise. The purpose of this addition was to make the model more resistant to the noise that comes with underwater images so that it would be easier to find things in the field.

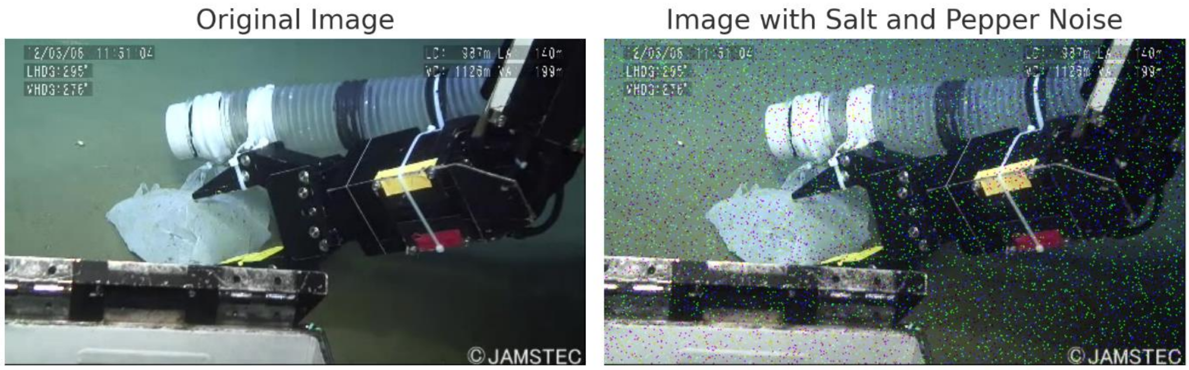


Figure 8 Noise Augmentation Sample

4.3.7. Mosaic Augmentation

Implications: By creating composite images, we can simulate intricate scenarios in which various types of debris coexist. This challenge improves the model's ability to identify and classify debris in a cluttered visual environment.

4.3.8. Bounding Box Augmentations

Important: To make sure accuracy, this object detection framework uses bounding box annotations for mirror image transformations. After the enhancement, this makes sure that the model can accurately find objects and pinpoint their locations.
When studying marine debris, especially plastic trash, it is important to look at how to choose models, annotate data, process it, and add to them.

4.4 Integration of Model Selection and Data Annotation with Preprocessing and Augmentation

4.4.1 Annotation and Model Preparation

- Carefully prepared the dataset that including annotations in both formats XML and TXT formats, to meet the needs of different object detection models i.e. tensorflow models and Yolo models. TXT annotations are made to work with the YOLOv3 model, while XML annotations are for Centre Net and Faster R-CNN.
- Object recognition models are selected according to the evaluation method results. YOLOv3, fast R-CNN and Center-net models were used to identify the best marine debris detection solution based on the requirement. All models are good as per evaluation metrics. Fast processing makes YOLOv3 ideal for real-time detection. Accurate fast R-CNN is useful for accurate analysis. Finally, the center net is fast and accurate, making it versatile.

4.4.2 Integration with Preprocessing and Augmentation

All chosen models are compatible with preprocessing and augmentation strategies. This ensures that any model can use the dataset effectively after these techniques improve its quality and variety. This improves model learning and performance.
Making training datasets better: When you use auto-orientation, resizing, and adaptive equalization techniques in the preprocessing stage along with full augmentation strategies, you get a dataset that is strong and accurately shows what happens in the real world. This improved dataset helps models find and sort marine debris accurately and efficiently in various underwater conditions.

4.5 THE EXPERIMENTAL SETUP FOR EACH MODEL:

4.5.1 Center Net Hourglass Model

Model Architecture and Configuration:

- The "hourglass_104" architecture is used in this mode. This architecture is known for being able to find objects of different sizes. This is very important for finding trash in the ocean of different sizes.
- Standard deviations, BGR ordering, and channel means that have already been set are used by the model to make the data normal. In order to learn, normalization has to happen over and over again. This method makes sure that it does.

Image Resizing Strategy:

- Maintaining the aspect ratio: This resizing feature is necessary to keep the shapes of objects in the input data intact. It does this by setting the minimum and maximum dimensions of images to 512 pixels.

Object Detection Task Configuration:

- The model emphasizes achieving a better equilibrium between the detection task's components by weighting task, offset, and scale loss differently.
- Localisation Loss: localisation loss shows that debris location requires more accurate object localisation.

Object Centre Parameters:

- Centre and classification loss are used with specific weights and a minimum box overlap IOU of 0.7. This configuration emphasises object centering and classification efficiency. It uses a logistic focal loss with reduced penalties and carefully chosen alpha and beta values.

Training Configuration

- The training procedure uses a 32-batch size (subjected to GPU) to balance computational efficiency and learning effectiveness. This option provides 25,000 training data exposure steps.

- Data Enhancement: The model uses random flipping, cropping, hue, contrast, saturation, and brightness adjustments to improve resilience and handle a variety of input conditions. These methods help the model handle many input conditions.

- Based on the research, the Adam optimizer is used, and the step learning rate is manually adjusted to start at 1e-3 and decrease. The goal is to achieve results. A specific Epsilon value is used to top-itemize the learning process to improve accuracy and efficiency.

- Before training, we must convert our image dataset to TFRecords. The native data format of TensorFlow, TFRecords, is optimised for storage and retrieval. This conversion step speeds up data loading during training, optimising the process. We carefully adjust the batch size to match our GPU hardware's specs and capacity. We need this calibration to maximise the GPU's computational power without exceeding its memory. By carefully choosing the batch size, we hope to find the best balance between how quickly the model works and how efficiently the computer works. This is especially important for deep learning applications that use large, complicated datasets.

Evaluation Strategy

- The model evaluates performance using COCO detection metrics for rigour.
- In evaluation, a batch size of 1 allows for a detailed analysis of model performance per instance.

4.5.2YOLOV3 CONFIGURATION:

Learning rate and training dynamics

- The fact that the initial learning rate is set at 0.01, which shows this preference, shows that fast learning is preferred at the start of the training process.
- Furthermore, the final learning rate is discovered to be 0.01, which means that it stays the same during the whole training process. This is one of the things that was found.
- A value of 0.937 for the momentum parameter helps to lower oscillations and speed up the optimizer, moving it in the right direction.
- The weight decay parameter, which is set to 0.0005 and acts as a regularization method, makes it less likely that the model will be too good by punishing weights that get bigger. To do this, the parameter needs to be set to 0.0005 in the right way.

Warmup Strategy

- The model incorporates a warmup period consisting of three epochs, during which the learning rate is gradually increased to its designated value. This warmup phase is of utmost importance as it aids in the stabilization of the training process during its initial stages.

- Warmup Momentum and Bias Learning Rate: Changing the momentum and bias learning rate during the warmup phase helps make the model more stable before started the main training phase.

Loss Function Configuration

- Weights for Box, Class, and Object: Giving different amounts of weight to box (0.05), class (0.5), and object (1.0) losses shows that the focus is evenly spread between accurate box positioning, correct object classification, and successful object detection.
- It looks like the "Class and Object Loss Powers" have values of 1.0, which is a common way to change the size of these parts of the loss function.

Data Augmentation and Training Parameters

- Set hsv_h, s, v, and degrees to adjust images. HSV and images change. Image manipulation needs this. Augments enhance model extrapolation from training data. Marine environments with varied lighting and backgrounds require this. Marine characteristics.
- Translation, scaling, shearing, and flip vary training data. Flipping augments variability. This simplifies model simulation in many cases. Flipping data adds variability. Mixup (0.0) and mosaic (1.0) augmentation methods increase training dataset variability and model resilience. Increasing variation-prone variables does this. Overall model damage resistance improves. This requires more mixing.

Training Setup

- Firstly, the model will be trained 100 times. Each group has 64 samples. The model has time to learn and get used to identifying marine debris with this training schedule.
- For model training, YOLO changes images without making any changes to TensorFlow Records. Data organization is very important, though. Furthermore, you can pick batch size based on GPU capabilities to get the most out of the available resources.

4.5.3 FASTER R-CNN WITH RESNET-101 CONFIGURATION

Model Architecture and Configuration

- When making the feature extractor, the ResNet-101 backbone was picked because it has a lot of resources and can get complex features out of images well. The architecture is very good at capturing complex patterns because it has 101 layers. This is important for dealing with the wide range of shapes and sizes that marine debris comes in.
- Incorporating trainable batch normalization into the feature extractor makes sure that all layers receive normalized data. This makes the model more stable and better at learning.

Training Configuration:

- Same like how we did for centernet because both models taking common process of training.
- Batch Size and Training Steps: A batch size of 8 combined with 15,000 training steps shows a meticulously calibrated strategy between the limitations of computational resources and the requirement for substantial.
- Batch size is subject to GPU availability.

Data Augmentation:

- Augmentation Techniques: Random horizontal flips and other augmentation methods help the model apply training data to new situations. The model must be ready to handle marine debris' varied and uncertain properties.

Evaluation Strategy:

- The model evaluates performance using COCO detection metrics for rigour.
- In evaluation, a batch size of 1 allows for a detailed analysis of model performance per instance.

4.6 Utilization of the A100 GPU in Colab

- For training and evaluation of all three models, we've adopted a collaborative and convenient approach using Google Colab. This allows for smooth, powerful computing and smart data storage in Google Drive. Our data is easily accessible and securely backed up, simplifying the process. Like having a powerful workstation and a reliable storage solution in one, it simplifies our workflow and boosts efficiency.
- Improved Computational Efficiency: Colab's A100 GPU 40GB RAM increases computational efficiency for training computationally intensive models.
- Improved Training Capabilities: This GPU feature allows larger batch sizes and more complex calculations, reducing training times and improving model training efficiency.

YOLOv3, Faster R-CNN with ResNet-101, and CenterNet Hourglass understand marine debris identification requirements. Each model is configured here. This task requires Faster R-CNN's precision and detail orientation, YOLOv3's speed, and CenterNet's scale variation efficiency. All these skills are needed for this task.
Colab uses the A100 GPU to improve these models and meet marine debris detection computational requirements. Because the A100 GPU improves model efficiency. AI and deep learning will elevate environmental research in this project. We'll set this benchmark. These technologies can study marine environments and preserve ecological systems, as shown here.

Parameter	**Faster R-CNN (ResNet)**	**CenterNet**	**YOLO v3**
GPU Memory	40GB	40GB	40GB
No. of Steps/Epochs	2,500 steps	25,000 steps	100 epochs
Batch Size	8	32	64
Feature Extractor	faster_rcnn_resnet101_keras	hourglass_104	-
Optimizer	Momentum Optimizer	Adam Optimizer	SGD
Learning Rate & Schedule	Cosine Decay: Base=0.04, Warmup=0.013333, Warmup Steps=2000, Total Steps=25,000	Manual Step: Initial=1e-3, Step 90,000 at 1e-4, Step 120,000 at 1e-5	SGD(lr=0.01), Weight Decay Groupings
Training Time	2 hours	5 hours	3 hours
Train Images	12,354 images	12,354 images	12,354 images

Test Images	2,725 images	2,725 images	2,725 images

Table 1 Configurations of all Three Models

GPU Memory: All three models have a 40GB GPU for tasks that require a lot of processing power.
How to Train and How Many Models: There is a different training plan for each model. R-CNN trains for 2,500 steps with 8 batches, CenterNet for 25,000 steps with 32 batches, and YOLO v3 for 100 epochs with 64 batches.

Feature extractors and optimizers: Faster R-CNN uses a momentum optimizer and a ResNet-based feature extractor. CenterNet uses an hourglass and an Adam optimizer, and YOLO v3 doesn't say what kind of feature extractor it uses but uses an SGD optimizer.

The Momentum Optimizer uses the speed and direction that come from adding up past gradients. This makes the speed of convergence faster than with pure gradient descent. On the other hand, Adam Optimizer not only makes good use of momentum, but it also changes the rate of learning based on new gradients. This keeps it stable even if the gradients change or the parameters move to a different area.

Stochastic Gradient Descent (SGD) and Adaptive learning rate techniques (ADPT), like Adam, are two optimisation strategies used to update the weights of neural networks while they are being trained. It depends on the situation because each has its own unique qualities.

Adjusting the learning rate dynamically for each weight in a neural network is what ADPT optimizers like Adam do. To keep adaptive learning rates steady, Adam guesses the first and second moments of the gradients. This lets the weights be updated in a more subtle and fine-grained way. Within the first few stages of training, this often leads to faster convergence. It's great that Adam and similar algorithms are easy to set up because they usually need less tuning of the learning rate.

In contrast, SGD is a more conventional method that updates all weights at a fixed learning rate. Because it only looks at a small part of the whole dataset (a "mini batch") when updating the weights, the updates are more varied and can get past shallow local minima during training. Overall, this can mean better generalisation when the model is used on data it wasn't trained on. Even though it might take longer for convergence to happen, using a uniform learning rate and its noise can help prevent overfitting by regularly shaping the data.

SGD is ideal for YOLO object detection model training for numerous factors:

- Generalisation: Historical advantages of SGD into this region is major factor. Generalization is essential into object identification for making models perform in environments untrained upon.
- Predictableness & Stability: Uniform learning rate gives SGD steady inputs into training, making behavior predictable. It's crucial for training complicated models sensitive to hyper parameter adjustments as well as learning dynamics.
- Regularization Impact: SGD's random updates act as built-in regularisation, preventing model from growing overly precise for training data base, typical issue in sophisticated neural networks.
- Empirical Efficacy: SGD models consistently perform well with both known and fresh datasets, according to experiments. Consistent performance determines its selection.

SGD outperforms ADPT optimizers in generalization, rigidity, & practical efficiency for YOLO model. SGD are able to communicate with new data & operate reliably following deployment, making them ideal for model optimization.

To train YOLO v3, Faster R-CNN takes two hours, and CenterNet takes five hours.

Image Dataset: 12,354 train images and 2,725 test images were used to train and test all models.

4.7 IN-DEPTH EVALUATION PARAMETER ANALYSIS:

This discussion will focus on accuracy, F1 score, precision, and recall.

- Precision gives a complete picture of the model's performance, but marine debris detection datasets are often imbalanced, making confidence low. Because imbalanced datasets have more errors.
- Debris detection relies on the harmonic mean of precision and recall, the F1 Score, to balance false positives and negatives. The F1 Score balances false positives and negatives.
- Long-term memory, precision Missing debris (low recall) or misclassifying non-debris as debris (low precision) can have serious consequences. When consequences are high, accuracy and recall matter. Precision is the percentage of positive predictions that are true, while recall is the percentage detected. Both of these ideas matter when the consequences are serious.
- A confusion matrix evaluates classification model performance using machine learning. Classification model performance is assessed by confusion matrix. True positives are shown as a square matrix.
- One can visualize true positives, true negatives, false positives, and false negatives using Matplotlib. Multiple debris types must be tested to determine the model's performance. It's impossible.
- This feature detects model biases like class biases or debris misclassification.
- The receiver operating characteristic (ROC) curve and area under the curve can indicate a model performance at different levels of thresholds. These examples demonstrate that diagnostic procedures have balanced true and misleading positives.

5 Implementation

5.1 TRAINING OF ALL THREE MODELS

1. Tensorflow Object Detection API V2, which contains CenterNet and Faster R-CNN model check points, offers comprehensive support for a wide variety of models throughout its extensive functionality. The following categories break down the series of steps on which both CenterNet and Faster R-CNN are based:

2. Generating XML annotations for a dataset starts the annotation process. These annotations provide specific information regarding the locations of objects in each image in the dataset.

3. The process of downloading model weights involves downloading the pre-trained model weights. These are intended to meet the requirements for further training.

4. Pipeline configuration file changes: The pipeline configuration file contains new paths like these:

    - The path to the label text file, which includes class labels (for instance, the letter '0' which stands for plastic).

    - Instructions for the train to take the test are included in both the record file and the record file.

    - The checkpoint location of the model that is going to be trained will be located here.

5. In this step, training parameters like batch size, number of steps, and number of classes are changed to fit the needs of the dataset and the model.
6. **Ultralytics for YOLOv3**: This framework simplifies the training and evaluation of YOLO models with features for advanced data augmentations and optimizations. The process for YOLO v3 includes:
   - The YOLOv3 checkpoint has been downloaded, indicating that our dataset satisfies all training requirements.
   - After that, the YOLOv3 architecture and the starting weights from the downloaded checkpoint are used to train the model on the dataset. Until the model is trained, this keeps happening. Model training then took place.

These steps meticulously prepare and train each model to complete its tasks. The adaptable TensorFlow Object Detection API v2 makes it easier to train CenterNet and Faster R-CNN, and Ultralytics makes YOLOv3 user-friendly. These processes are supported by the A100 GPU's computational power.

This research tested YOLOv3, Faster R-CNN, & CenterNet Hourglass models for object detection. Transfer learning is crucial to this endeavor. Finding maritime trash of various sizes & varieties requires this procedure.

Method utilised for transfer learning given as below:

Pre-trained Model Selection:

Many picked YOLOv3 for real-time detection because of its speed. High precision-made Faster R-CNN ideal for fine-detail jobs. Locating debris of various sizes requires CenterNet Hourglass, that could handle items of various sizes.

Fine-tuning procedure:

Every model was fine-tuned using marine trash data. They used their feature extraction abilities from training upon huge datasets to change models' layers to figure out undersea debris characteristics. Preparing & adding data:

In order to enhance resemblance of training data base to real-world scenarios, techniques such as auto-orientation, scaling, & adaptive equalization were utilized.

Model-Specific Deviations to Training:

Faster R-CNN utilised momentum optimizer for enhancing both speed & accuracy. CenterNet Hourglass used the Adam optimizer because it could change the learning rate on the fly. To make the learning process more balanced, SGD was used to improve YOLOv3. YOLOv3, which is often used in the Darknet framework, prepares data in a way that is different from TFRecords. Change to TFRecords (Only for TensorFlow models):

TFRecords were used by Faster R-CNN and maybe even CenterNet Hourglass to handle data efficiently. This format is built into TensorFlow and helps the training go faster. By using transfer learning, the project made the most of the large amount of knowledge already in pre-trained models by fine-tuning them for the task of finding marine debris. This way of doing things saved a lot of time and money while still making models that were good at finding different kinds of debris in tough underwater environments.

5.2 STRUCTURAL DIFFERENCES IN MODELS

YOLOv3

1. A single-stage detector is a type of object detection model that directly predicts both the bounding boxes and class probabilities of objects in an image. This approach is specifically designed to prioritise speed optimisation in the detection process. Appropriate for situations necessitating instantaneous detection.

2. The YOLOv3 architecture is designed to facilitate rapid predictions, a feature of great significance in situations necessitating real-time analysis, such as the automated surveillance of marine ecosystems.

Faster R-CNN with ResNet-101

Two-Stage Detector: Involves region proposal followed by classification, emphasising precision. ResNet-101 provides deep feature extraction.

Complexity and Precision: Suited for applications where accuracy is paramount, albeit with a trade-off in speed compared to YOLOv3. In marine debris detection, this could mean a more accurate classification of various debris types.

CenterNet Hourglass

This method finds object centres before estimating bounding boxes. After that, it continues. Keypoint-based detection describes this style. This strategy helps the model handle scale variations, which is great.

The proposed architecture works well with large debris variations. The feature is enhanced object size detection. Small fragments and large objects.

The Faster R-CNN model is accurate, YOLOv3 is fast, and CenterNet handles scale well. Models have different benefits. Model selection must consider marine debris detection requirements. These requirements may include accuracy, rapid detection, and object size adaptability. Model training and performance benefit from A100 GPU computing. Models improve this. This extensive analysis emphasises the importance of custom model selection and optimisation for AI-based environmental conservation and research.

Feature/Model	**YOLOv3**	**Faster R-CNN (ResNet-101)**	**CenterNet Hourglass**
Type of Detector	Single-Stage Detector	Two-Stage Detector	Keypoint-Based Detector
Primary Focus	Speed Optimization	Precision and Deep Feature Extraction	Handling Scale Variations
Architectural Design	Designed for rapid predictions	Involves region proposal followed by classification with ResNet-101	Finds object centers, then estimates bounding boxes
Suitability	Real-time analysis, such as automated marine ecosystem surveillance	Applications where accuracy is paramount, like detailed debris analysis	Varied object sizes, including small fragments and large objects

Performance Characteristics	Fast detection speed, ideal for instantaneous detection requirements	More accurate classification, but slower compared to YOLOv3	Enhanced object size detection, adaptable to large variations in debris sizes
Application Scenario	Situations requiring quick response times	Scenarios where detailed and accurate detection is more crucial than speed	Environments with diverse and varying object sizes
Computational Requirements	Optimized for efficiency, less computationally intensive	Requires more computational resources for deep feature extraction	Balanced between detection precision and computational efficiency

Table 2 Differences among the Three models

The quick YOLOv3 detector with only one stage is perfect for keeping an eye on marine ecosystems. Architecture allows quick, less-detailed predictions.

Faster RCNN ResNet 101: This two-stage detector prioritises accuracy and precision over speed, making it suitable for detailed applications. ResNet-101 deep feature extraction complicates computation.

CenterNet Hourglass: Keypoints simplify object size variations in CenterNet. Precision and computational efficiency allow it to detect debris from small fragments to large objects.
The speed, accuracy, and object size of these models determine their benefits.

6 Evaluation

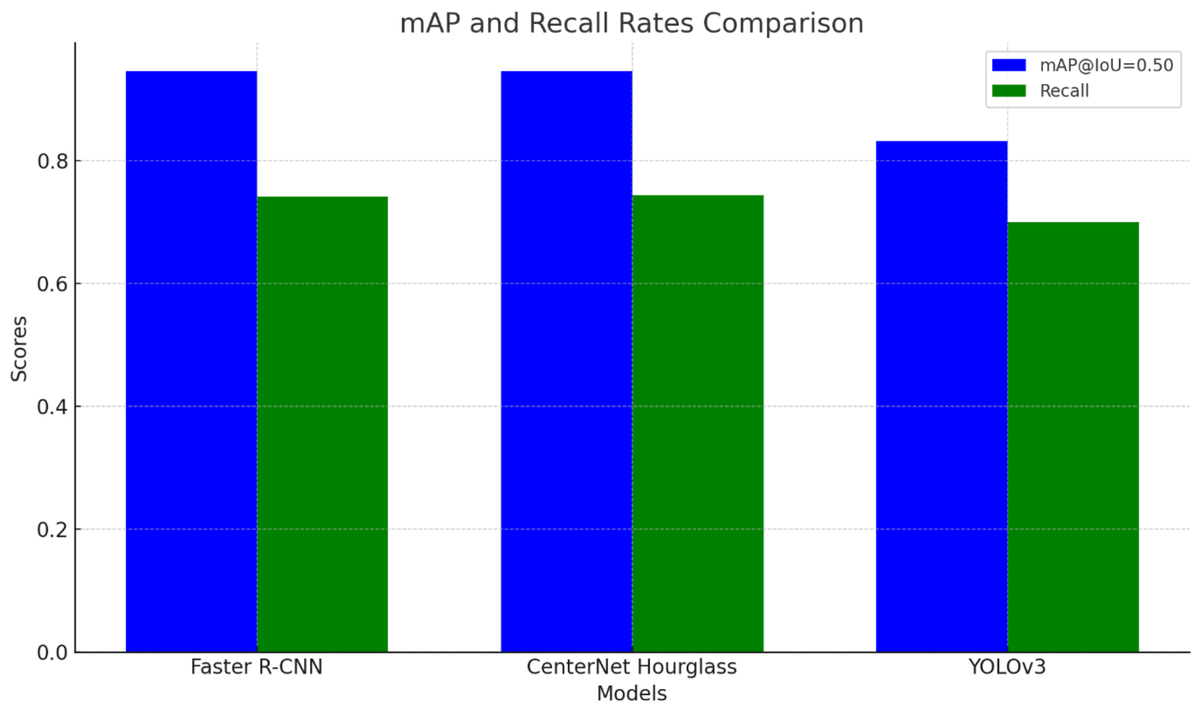


Figure 9 mAP scores of Three Models

MODEL 1: FASTER R-CNN

1. Strengths: High precision at 0.50 IoU threshold, reliable for moderate ground truth overlap.
2. Inconsistencies: The accuracy of the results decreases more noticeably at higher IoU thresholds compared to CenterNet, which suggests that exact object localization may have some problems.
3. Object Size Handling: Detects large objects with high precision, but performs poorly on small and medium objects.

MODEL 2: CenterNet Hourglass

1. For the most part, it works well across a range of IoU thresholds, which shows that it can reliably localise at different levels of strictness.
2. Weaknesses: It doesn't have a major weakness in the given metrics, but keep in mind that the hourglass network in architectures like CenterNet can be very straining on computers.
3. Taking care of things of different sizes: It is very good at accurately finding small things, which is very helpful when it's important to find small things.

MODEL 3: YOLOv3

1. Because it can draw conclusions quickly, it works well for real-time tasks, even though it is less accurate when IoU thresholds are raised.

2. One problem is that its mAP isn't very strong across the IoU threshold range. This could be a price to pay for its speed.
3. The performance of YOLOv3 is good for all object sizes, but there aren't any specific metrics for each one.

mAP@IoU=0.50 (AP50): Faster R-CNN and CenterNet are tied for the lead, with YOLOv3 trailing but still offering respectable precision.

mAP@IoU=0.50-0.95 (AP): CenterNet maintains a slight lead over Faster R-CNN, while YOLOv3 shows a drop-off in performance, indicating its precision is more affected by increasing IoU thresholds.

Small Object Detection: CenterNet's high precision with small objects could be particularly beneficial for applications like surveillance or detailed scene analysis.

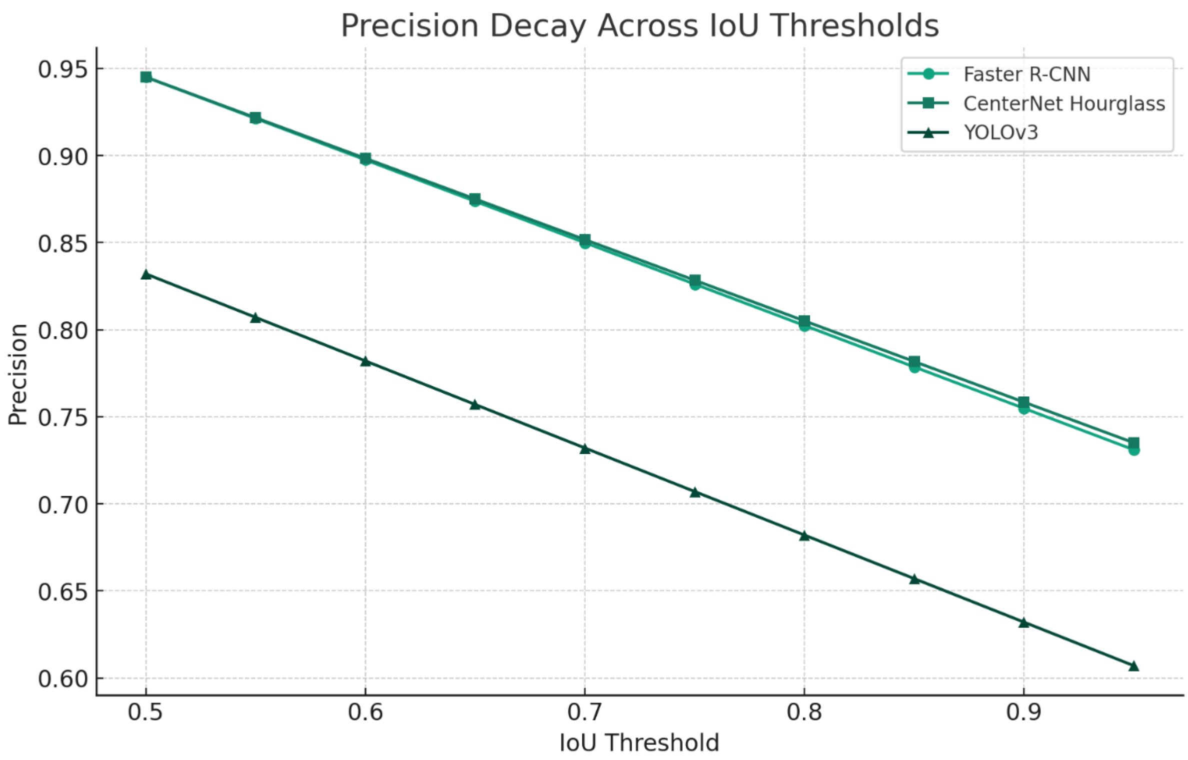


Figure 10 Precision Decay Plot

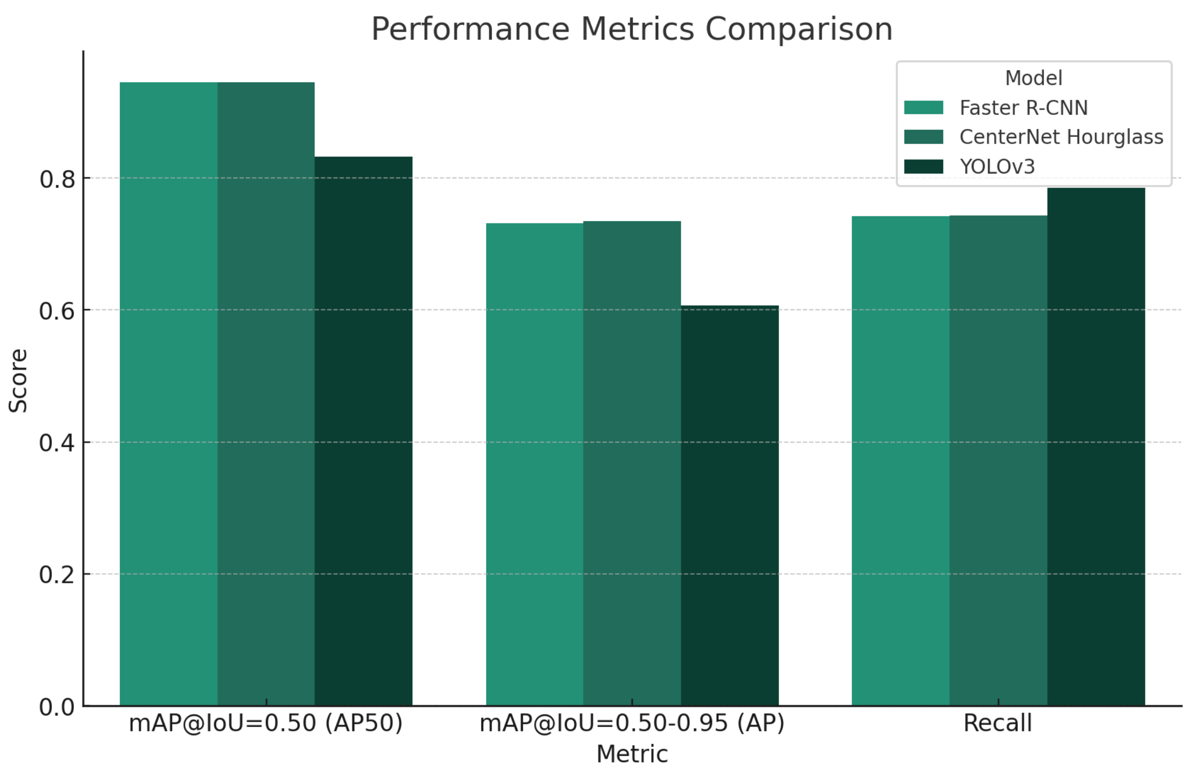
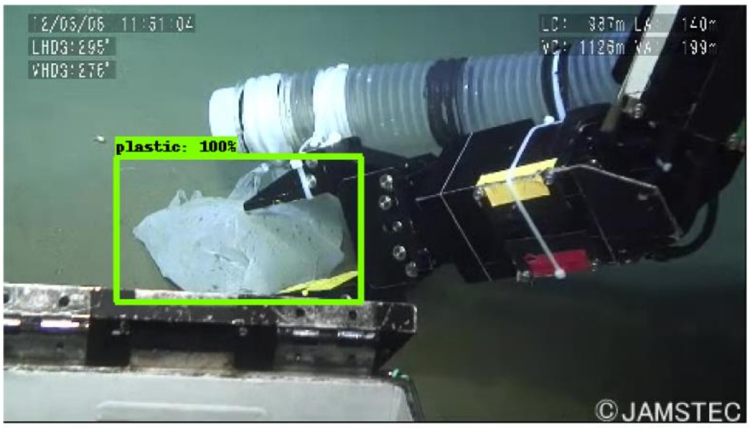


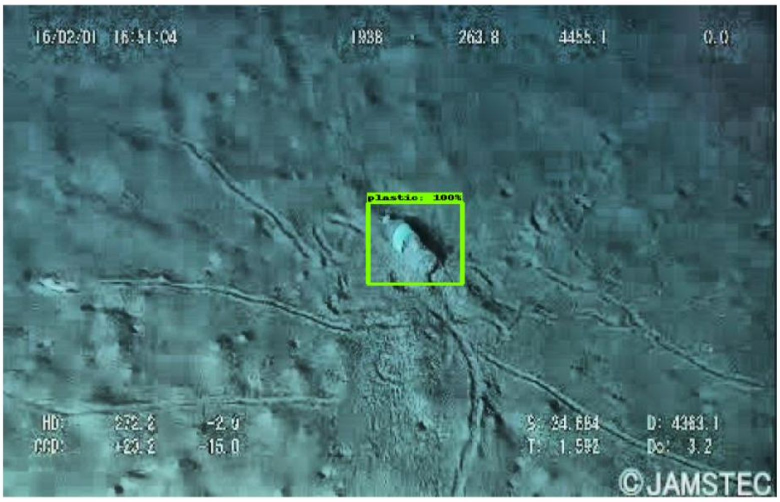
Figure 11 Performance Metrics Comparison

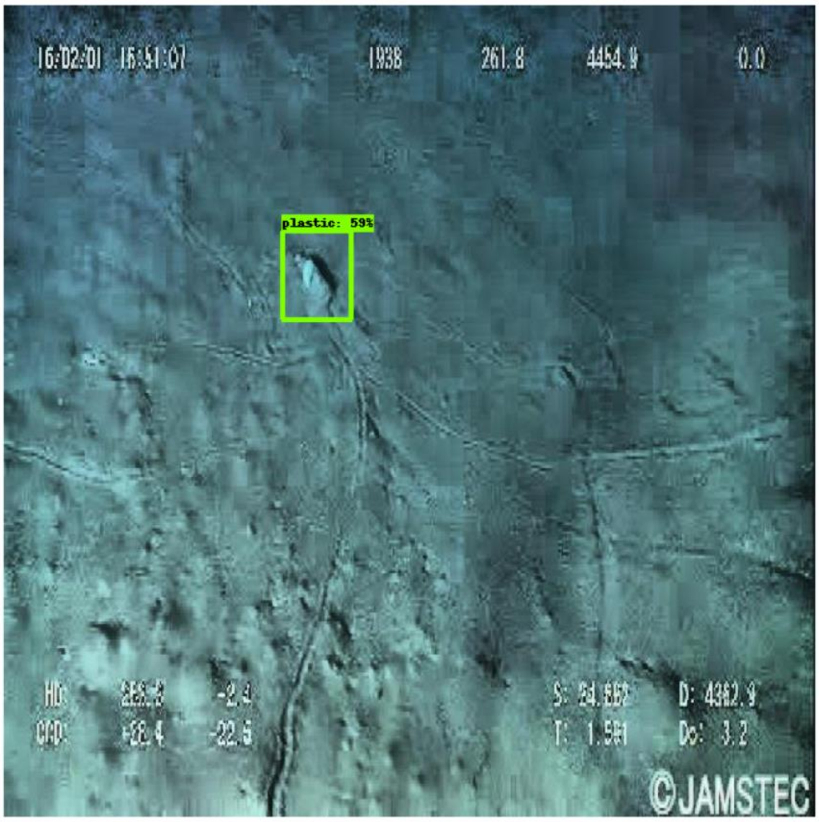
CenterNet Hourglass is the most accurate model across IoU thresholds, especially for small object detection, according to current metrics. This applies especially to finding hidden objects. The faster R-

CNN algorithm performs well with large objects and approaches high precision when the IoU threshold is 0.50. This applies to algorithms too. YOLOv3 may be the fastest real-time model for detection, despite being less accurate at higher IoU thresholds. Even though it's not the most accurate model for real-time detection or overall.

SAMPLE OUTPUTS:







7 Discussion of the Project

Three well-known object detection models are being looked at by the project. There are three models at play: YOLOv3, CenterNet Hourglass, and Faster R-CNN. Precision (mAP) at different IoU thresholds and recall have been tested to see which model works best for specific tasks or everyday use. The goal of this evaluation was to find the best model.

FASTER R CNN: **-** Faster R CNN has proven to be a reliable model especially when using the commonly used IoU threshold of 0.50. It exhibits the level of accuracy at this threshold along with CenterNet Hourglass. However, when the IoU threshold is increased its precision doesn't hold up well as CenterNets. While Faster R CNN performs well for object detection tasks it may not be the ideal choice for applications that require highly precise localization.

CenterNet Hourglass: - CenterNet Hourglass stands out by delivering excellent performance across a wide range of IoU thresholds, including the most stringent ones. Additionally it has demonstrated precision in detecting small objects. This makes CenterNet Hourglass particularly valuable in scenarios where accurate detection of objects is crucial such as satellite imagery analysis or medical imaging where fine details matter.

YOLOv3: **-** YOLOv3 strikes a balance between precision and recall but shows slightly lower mean Average Precision (mAP) across different IoU thresholds compared to other models. This suggests that YOLOv3 may be less precise at higher IoU thresholds than the other models mentioned above. However it compensates for this by having the recall, among the three models indicating that it is less likely to miss true positives.One of the advantages of YOLOv3 is its ability to perform tasks swiftly making it suitable for real time applications where detecting objects quickly is more crucial, than achieving highly accurate localization.

When it comes to performance CenterNet Hourglass stands out as the most accurate model across different IoU thresholds. However, determining the "model largely depends on the specific requirements of your application.

1.If you require a solution that can accurately detect objects CenterNet Hourglass is likely your top choice.
2.For real time applications especially when speed is crucial and extreme precision is not as important YOLOv3 may be the suitable model to consider.
3.Faster R CNN provides an option for achieving a balance, between accuracy and the ability to detect objects of varying sizes.

The evaluation process has been supported by analyzing data using bar charts and line graphs. The findings have been condensed into tables to aid understanding. Visual aids are crucial for stakeholders who may not have knowledge but need to comprehend how the models perform differently in order to make informed decisions.

In implementation using these models would require considering various factors such as available computational resources, dataset characteristics, the types of objects to be identified and the operational requirements of the detection task (such as time constraints).

The final decision on which model to run also takes into account training and inference times, the ability to avoid overfitting, and generalisations to unseen data.

8 Conclusion and Future Work

8.1 CONLUSON:

Here,the study reviews the Faster R-CNN, CenterNet Hourglass, and YOLOv3 object detection models. We tested their ability to detect underwater plastic waste and debris. The study assessed the performance at different overlaps between predicted and actual objects using mean average precision (mAP) and recall. To help select the correct model, the review lists each model's pros and cons. This faster R-CNN is suitable for many object detection tasks due to its normal overlap performance. CenterNet Hourglass locates objects more precisely across overlap levels. This makes it ideal for accurate object detection. It may not be the most precise model at high overlap, but YOLOv3 finds many objects quickly, making it useful in fast-paced situations.

This study clearly helps protect the environment by shedding light on the important issue of plastic pollution in our oceans. This is a wake-up call that our seas are in trouble, which in turn threatens marine life and the industries that depend on them for their survival, such as fishing, shipping, and tourism. The research improves the health of marine habitats and, by extension, the businesses that depend on them by finding better ways to find and get rid of plastic trash.

On the business side, this research has a lot of important implications. To effectively deal with the problem of ocean pollution, it brings attention to the need for stricter rules and cutting-edge technologies. These changes could have a big effect on how businesses work, especially those that deal with recycling and waste management. The push for new technologies could also lead to the creation of new environmental-friendly businesses, like those that clean up marine environments.

Additionally, there is a message that fighting ocean pollution is a worldwide problem that needs businesses and nations all over the world to work together. This study argues that countries should work together to make sure that environmental standards are followed by everyone. This could change the way businesses work around the world. The study is basically a call to action for a better future where caring for the environment and progress in industry go hand in hand.

8.2 FUTURE WORK:

We should investigate the following directions to improve these detection models' reliability and applicability:

1. Comparing model performance across datasets can assess generalisation.
To determine lifespan, test models in different lighting, weather, and obstructions. Model longevity depends on this. Model lifespan depends on this.
2. Research pilots will show model speed, resource use, and system compatibility.
Model optimisation that improves efficiency without compromising accuracy is crucial when resources are scarce. Model optimisation prevents errors. Because model optimisation cuts errors. It thrives with abundant resources.
Our evaluation process evaluates each model's detection. Precision Recall curves and F1 scores allow this.
3. Explore options, including building models to meet requirements.
To advance object detection and related fields, models must be industry-specific and framework-based. Advanced object detection needs this. Progress is only possible this way. Many fields will benefit and advance.

Learning improves accuracy and keeps up. Tools keep models current and let them learn without retraining. Tools help models adapt to field changes. This method helps our models adapt quickly and effectively to real-world challenges. This method works.
These methods make object detection models more versatile and accurate, delivering fast results in various contexts. We can help them succeed here. Because of this, we can get them to use these methods.

References

Abdu, H. & Mohd Noor, M.H. (2022). A Survey on Waste Detection and Classification Using Deep Learning. *IEEE Access*. [Online]. 10. p.pp. 128151–128165. Available from: http://dx.doi.org/10.1109/access.2022.3226682.

Alshibli, S.S., Alsaadi, M.K. & Hajamohideen, F. (2022). Marine debris detection system: An approach to eliminate the toxicological threats of plastic using AI and deep neural network. *6th Smart Cities Symposium (SCS 2022)*. [Online]. Available from: http://dx.doi.org/10.1049/icp.2023.0411.

Bao, Z., Guo, Y., Wang, J., Zhu, L., Huang, J. & Yan, S. (2023). Underwater Target Detection Based on Parallel High-Resolution Networks. *Sensors*. [Online]. 23 (17). p.p. 7337. Available from: http://dx.doi.org/10.3390/s23177337.

Bhatlawande, S., Shilaskar, S., Agrawal, M., Ashtekar, V., Badade, M., Belote, S. & Madake, J. (2022). Study of Object Detection with Faster RCNN. *2022 2nd International Conference on Intelligent Technologies (CONIT)*. [Online]. Available from: http://dx.doi.org/10.1109/conit55038.2022.9847725.

Bhuvaneswari, R., Surya, T., Srikanth, T. & Balaji, R. (2022). A Novel Approach for Underwater Object Detection through Deep Intense-Net for Ocean Conservation Systems. *OCEANS 2022 - Chennai*. [Online]. Available from: http://dx.doi.org/10.1109/oceanschennai45887.2022.9775523.

Duan, K. (2019). *CenterNet: Keypoint Triplets for Object Detection*. [Online]. 17 April 2019. arXiv.org. Available from: https://doi.org/10.48550/arXiv.1904.08189. [Accessed: 23 November 2023].

Duarte, M.M. & Azevedo, L. (2023). Automatic Detection and Identification of Floating Marine Debris Using Multispectral Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing*. [Online]. 61. p.pp. 1–15. Available from: http://dx.doi.org/10.1109/tgrs.2023.3283607.

Esen, F., Degirmenci, A. & Karal, O. (2021). Implementation of the Object Detection Algorithm (YOLOV3) on FPGA. *2021 Innovations in Intelligent Systems and Applications Conference (ASYU)*. [Online]. Available from: http://dx.doi.org/10.1109/asyu52992.2021.9599073.

Fu, C., Liu, R., Fan, X., Chen, P., Fu, H., Yuan, W., Zhu, M. & Luo, Z. (2023). Rethinking general underwater object detection: Datasets, challenges, and solutions. *Neurocomputing*. [Online]. 517. p.pp. 243–256. Available from: http://dx.doi.org/10.1016/j.neucom.2022.10.039.

Gabriel, M., Cha, S., Al-Nakash, N.Y.B. & Yun, D. (2020). Wildlife Detection and Recognition in Digital Images Using YOLOv3: Extended Abstract. *2020 IEEE Cloud Summit*. [Online]. Available from: http://dx.doi.org/10.1109/ieeecloudsummit48914.2020.00033.

Hu, Q., Peng, X., Wang, C., Chen, Q., Qu, Y. & Yan, Z. (2022). Object Detection and Recognition based on Improved Faster R-CNN. *2022 International Conference on Image Processing and Media Computing (ICIPMC)*. [Online]. Available from: http://dx.doi.org/10.1109/icipmc55686.2022.00009.

Hu, Z. & Xu, C. (2022). Detection of Underwater Plastic Waste Based on Improved YOLOv5n. *2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC)*. [Online]. Available from: http://dx.doi.org/10.1109/icftic57696.2022.10075134.

Huang, Y., Fu, X. & Zeng, Y. (2022). Anchor-Free Weapon Detection for X-Ray Baggage Security Images. *IEEE Access*. [Online]. 10. p.pp. 97843–97855. Available from: http://dx.doi.org/10.1109/access.2022.3205593.

JAIN, R.K., WATASUE, T., NAKAGAWA, T., SATO, T., IWAMOTO, Y., RUAN, X. & CHEN, Y.-W. (2021). LogoNet: Layer-Aggregated Attention CenterNet for Logo Detection. *2021 IEEE International Conference on Consumer Electronics (ICCE)*. [Online]. Available from: http://dx.doi.org/10.1109/icce50685.2021.9427658.

Jing, C., Chenggang, C., Nan, Y. & Peifeng, X. (2020). Region-Free Cloud Recognition Based on Center Net. *2020 5th International Conference on Power and Renewable Energy (ICPRE)*. [Online]. Available from: http://dx.doi.org/10.1109/icpre51194.2020.9233243.

Kumar B., C., Punitha, R. & Mohana (2020). YOLOv3 and YOLOv4: Multiple Object Detection for Surveillance Applications. *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. [Online]. Available from: http://dx.doi.org/10.1109/icssit48917.2020.9214094.

Lee, Y., Lee, C., Lee, H.-J. & Kim, J.-S. (2019). Fast Detection of Objects Using a YOLOv3 Network for a Vending Machine. *2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*. [Online]. Available from: http://dx.doi.org/10.1109/aicas.2019.8771517.

Liu, B., Zhao, W. & Sun, Q. (2017). Study of object detection based on Faster R-CNN. *2017 Chinese Automation Congress (CAC)*. [Online]. Available from: http://dx.doi.org/10.1109/cac.2017.8243900.

Liu, Y. (2018). An Improved Faster R-CNN for Object Detection. *2018 11th International Symposium on Computational Intelligence and Design (ISCID)*. [Online]. Available from: http://dx.doi.org/10.1109/iscid.2018.10128.

Mensink, T., Uijlings, J., Kuznetsova, A., Gygli, M. & Ferrari, V. (2022). Factors of Influence for Transfer Learning Across Diverse Appearance Domains and Task Types. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [Online]. 44 (12). p.pp. 9298–9314. Available from: http://dx.doi.org/10.1109/tpami.2021.3129870.

Noman, M.K., Shamsul Islam, S.M., Abu-Khalaf, J. & Lavery, P. (2021). Seagrass Detection from Underwater Digital Images using Faster R-CNN with NASNet. *2021 Digital Image Computing: Techniques and Applications (DICTA)*. [Online]. Available from: http://dx.doi.org/10.1109/dicta52665.2021.9647325.

Wang, Y., Liu, J., Yu, S., Wang, K., Han, Z. & Tang, Y. (2021). Underwater Object Detection based on YOLO-v3 network. *2021 IEEE International Conference on Unmanned Systems (ICUS)*. [Online]. Available from: http://dx.doi.org/10.1109/icus52573.2021.9641489.

Xiu Li, Min Shang, Qin, H. & Liansheng Chen (2015). Fast accurate fish detection and recognition of underwater images with Fast R-CNN. *OCEANS 2015 - MTS/IEEE Washington*. [Online]. Available from: http://dx.doi.org/10.23919/oceans.2015.7404464.

Xue, B., Huang, B., Wei, W., Chen, G., Li, H., Zhao, N. & Zhang, H. (2021). An Efficient Deep-Sea Debris Detection Method Using Deep Neural Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. [Online]. 14. p.pp. 12348–12360. Available from: http://dx.doi.org/10.1109/jstars.2021.3130238.