remote sensing

MDPI

Article

Understanding the Influence of Image Enhancement on Underwater Object Detection: A Quantitative and Qualitative Study

Ashraf Saleem [1], Ali Awad [1,*], Sidike Paheding [2], Evan Lucas [3], Timothy C. Havens [4] and Peter C. Esselman [5]

[1] Department of Applied Computing, College of Computing, Michigan Technological University, Houghton, MI 49931, USA; ashraf@mtu.edu
[2] Department of Computer Science and Engineering, Fairfield University, Fairfield, CT 06824, USA; spaheding@fairfield.edu
[3] Institute of Computing and Cybersystems, Michigan Technological University, Houghton, MI 49931, USA; eglucas@mtu.edu
[4] Department of Computer Science, College of Computing, Michigan Technological University, Houghton, MI 49931, USA; thavens@mtu.edu
[5] U.S. Geological Survey Great Lakes Science Center, Ann Arbor, MI 48105, USA; pesselman@usgs.gov
* Correspondence: aawad@mtu.edu

Abstract: Underwater image enhancement is often perceived as a disadvantageous process to object detection. We propose a novel analysis of the interactions between enhancement and detection, elaborating on the potential of enhancement to improve detection. In particular, we evaluate object detection performance for each individual image rather than across the entire set to allow a direct performance comparison of each image before and after enhancement. This approach enables the generation of unique queries to identify the outperforming and underperforming enhanced images compared to the original images. To accomplish this, we first produce enhanced image sets of the original images using recent image enhancement models. Each enhanced set is then divided into two groups: (1) images that outperform or match the performance of the original images and (2) images that underperform. Subsequently, we create mixed original-enhanced sets by replacing underperforming enhanced images with their corresponding original images. Next, we conduct a detailed analysis by evaluating all generated groups for quality and detection performance attributes. Finally, we perform an overlap analysis between the generated enhanced sets to identify cases where the enhanced images of different enhancement algorithms unanimously outperform, equally perform, or underperform the original images. Our analysis reveals that, when evaluated individually, most enhanced images achieve equal or superior performance compared to their original counterparts. The proposed method uncovers variations in detection performance that are not apparent in a whole set as opposed to a per-image evaluation because the latter reveals that only a small percentage of enhanced images cause an overall negative impact on detection. We also find that over-enhancement may lead to deteriorated object detection performance. Lastly, we note that enhanced images reveal hidden objects that were not annotated due to the low visibility of the original images.

Keywords: underwater image enhancement; underwater object detection; underwater dataset

Academic Editors: Haopeng Zhang and Pedro Melo-Pinto

Received: 8 November 2024
Revised: 24 December 2024
Accepted: 25 December 2024
Published: 7 January 2025

Citation: Saleem, A.; Awad, A.; Paheding, S.; Lucas, E.; Havens, T.C.; Esselman, P.C. Understanding the Influence of Image Enhancement on Underwater Object Detection: A Quantitative and Qualitative Study. *Remote Sens.* **2025**, *17*, 185. https://doi.org/10.3390/rs17020185

1. Introduction

Underwater computer vision is a critical area of research with important applications in many domains ranging from marine biology and environmental monitoring to underwater exploration and the maintenance of submerged infrastructure [1–5]. Detecting objects in underwater environments is essential for tasks such as marine life tracking [6], inspecting underwater pipelines [7], mapping of coral reefs [8], and many other recent applications [9–11]. However, underwater environments present unique challenges for computer vision algorithms, primarily due to the way water absorbs and scatters light. These phenomena cause substantial image degradation, leading to issues such as reduced contrast and color distortion [12].

To address these challenges, a variety of image enhancement techniques have been developed over the years. Early methods focused on improving image contrast and brightness through techniques like histogram equalization and dark channel prior, whereas more recent approaches [13–17] leverage the power of deep learning using convolutional neural networks (CNNs) and generative adversarial networks (GANs) to restore and enhance underwater images [18]. These advancements have been driven by the need to improve the visibility of underwater scenes, with the assumption that clearer images would naturally lead to better performance in high-level tasks, such as object detection. However, contrary to this assumption, numerous studies have found that image enhancement often translates into similar or degraded detection performance [19–21].

For example, Wang et al. [21] found that in the underwater domain, image enhancement suppresses object detection performance. They hypothesized that this is attributed to increased interference in the image background in the form of noise, edge blur, texture corruption, or color introduction. A specialized evaluation toolbox for object detection [22] was used to be able to obtain the false positive rate and other important parameters, concluding that enhancement leads to a higher false positive rate mainly due to the objects' edge deformation after enhancement. Similarly, [20] delved deeper into the detection–enhancement relationship and conducted their experiments using multiple enhancers and detectors on the large Real-world Underwater Object Detection (RUOD) dataset [20], which is used in our work as well. The study presented in [20] concluded that cascaded enhancement and detection lead to slightly deteriorated detection performance regardless of the enhancer or detector type. This is also extended to Single-Image Super-Resolution (SISR) enhancement algorithms, where Awad et al. [23] investigated the effect of different SISR algorithms on the detection performance and concluded that restored high-resolution images cannot match or exceed the detection performance of the original images.

Based on a synthetic approach, the interactions between image enhancement and other image tasks have also been studied by synthetically degrading images and then restoring them. For instance, Pei et al. [19] used a standard set of image effects, including a simulated 'underwater' effect; they found that restoring the degrading images did not restore the classification performance on those images. This conclusion could be extended to object detection, although this is not explicitly stated in the work presented in [19]. Arriving at a similar conclusion, Chen et al. [24] split the Underwater Robotic Picking Contest (URPC2018) dataset [25] into three domains containing different environments and analyzed the effects of enhancement on detection on those environments. Although the researchers in [24] reached the same conclusion about the adversarial effect of enhancement on classification, they conducted further analyses that involved training and testing with different environments and concluded that enhancement could lead to more reliable performance rather than higher one. However, this study warrants validation by utilizing different types of enhancement algorithms.

Contradicting the already small body of research in this field, Alawode et al. [26] found that underwater image enhancement improved object tracking performance. They claim a 5% improvement in performance but do not include an analysis of statistical significance to support this claim. Zhang et al. [27] also found that object detection performance on an enhanced underwater image dataset improved slightly over an unenhanced one. They noted that the image quality metrics do not change in a statistically significant way and were unable to attribute the slight increase in detection performance to any specific factor.

An emerging and promising research direction in this field is the simultaneous consideration of both enhancement and detection through a combined architecture. For instance, Fu et al. [20] have attempted to combine the tasks of enhancement and detection by using a shared loss function that is fed backward from the detector to the enhancer, encouraging the latter to produce enhanced images that are favored by the detector rather than the human eye. In the same direction, Wang et al. [28] developed a reinforcement learning approach to tune an image enhancement pipeline to minimize object detection loss. This is carried out by configuring enhanced images as states, enhancement types as actions, and the detector's loss as the reward. Such joint learning frameworks in which the enhancer and the detector are being trained simultaneously have shown promising results. The findings from the work presented in this paper may will help researchers with the design consideration of such a joint detection-enhancement architecture.

As presented, the connection between image enhancement and object detection has been studied by many researchers in the underwater domain. However, despite the growing body of work, more detailed, generalized, and reliable analyses of the enhancement–detection relationship are still needed. To the best of our knowledge, previous enhancement–detection studies evaluate their results on a per-dataset basis rather than a per-image basis, which helps only in drawing general conclusions. This underscores the need for a better understanding of the interaction between image enhancement and object detection.

This study aims to address these gaps by conducting a comprehensive analysis of the factors that influence the effectiveness of image enhancement in underwater object detection. Unlike previous works, our study evaluates individual images before and after enhancement. This helps in identifying the conditions under which enhancement either supports or hinders detection performance on a granular level. In addition, this can guide future research in developing enhancement–detection strategies specifically tailored for improved underwater object detection. This work can help facilitate the way for the next generation of enhancement models that not only improve visual quality but also enhance the operational effectiveness of detection systems in challenging underwater environments [29].

This study makes key contributions to the field of underwater image enhancement and object detection through a unique approach and observations. Specifically, we introduce a novel approach to analyze the impact of underwater image enhancement on object detection by categorizing images based on their detection performance before and after enhancement. This is conducted by calculating the mean Average Precision (mAP) for each individual image, allowing us to compare the performance of each image before and after enhancement. We empower this approach with very detailed analyses highlighting a negative relationship between over-enhancement and detection and introducing mixed sets of original and enhanced images that outperform the original set.

The remainder of this paper is organized as follows: Section 2 details our experimental setup, including the models used for image enhancement and object detection, the datasets used, the metrics applied, and the evaluation procedures. In Section 3, we present and discuss the results of our experiments, analyzing the impact of enhancement on detection

performance by providing a quantitative and qualitative evaluation. Section 4 concludes the paper by summarizing our findings and discussing potential directions for future research.

2. Materials and Methods

In this section, we outline our experimental setup, listing the models used for image enhancement and object detection, the dataset used, and the metrics applied. In addition, an outline of our evaluation procedures is presented later in this section.

2.1. Selected Models, Metrics, and Datasets

To comprehensively cover underwater image enhancement, we selected nine state-of-the-art models that encompass physical, non-physical, and learning-based approaches. Non-physical methods include ACDC [30], TEBCF [31], and BayesRet [32]. In particular, the ACDC model compensates for inferior color channels (red and blue in shallow waters or red and green in deep waters) using attenuation matrices based on superior color channels. Then, it applies dual-histogram iterative thresholds to enhance global contrast and Rayleigh-distributed histograms for local contrast. The outputs are fused via multiscale fusion with weight maps balancing brightness and saliency in CIELAB and HSV spaces. Finally, a multiscale unsharp masking (MSUM) technique sharpens details and textures. The TEBCF model uses two input channels: one using RGB for dehazing and sharpness and the other leveraging CIELAB space for color correction and brightness adjustment. These are fused adaptively through multi-scale image fusion to recover contrast, saturation, and sharpness. Lastly, the BayesRet model uses Bayesian optimization to enhance underwater images by separating illumination and reflectance. It incorporates first- and second-order gradient priors to maintain natural colors and fine details. Physical methods are represented by the PCDE [33] and ICSP [34] models. Specifically, the PCDE model uses piecewise color correction to address color casts and applies dual priors for spatial and texture-based enhancement, improving both contrast and detail preservation in images. The ICSP model is used to correct non-uniform illumination. It integrates statistical priors on illumination channels with a variational framework and uses efficient algorithms for optimizing illumination balance, resulting in consistent and detailed image restoration. Finally, learning-based methods utilizing CNNs and GANs include AutoEnh [15], Semi-UIR [16], USUIR [17], and TUDA [35]. On the one hand, the AutoEnhancer model uses a Neural Architecture Search (NAS)-optimized U-Net with selectable transformer modules to enhance underwater images by learning optimal feature representation across RGB and CIELAB color spaces. On the other hand, the semi-UIR model uses a mean-teacher framework with a reliable bank for pseudo-labeling and contrastive regularization to utilize unlabeled data effectively. It avoids confirmation bias and ensures consistent image restoration results. From another perspective, the USUIR model leverages the similarity between original and degraded versions of underwater images. First, raw underwater images are used to generate three components, namely, background light, transmission map, and scene radiance. Then, by simulating re-degraded images and minimizing restoration errors, it performs unsupervised enhancement, balancing real-time processing and restoration quality. Finally, the TUDA model is designed for domain adaptation, reducing the gaps between synthetic and real underwater images. It incorporates inter-domain and intra-domain adaptation techniques and ranks image quality for adaptive enhancement. This approach balances diverse underwater conditions while maintaining robust quality. More details about the selected models can be found in [36]. All models were used in their original form, utilizing the source codes and trained models provided by their respective authors without any modifications or additional training. In addition, the models were only used in inference mode using the pre-trained models provided by their respective

authors. A GitHub link for the source code and pre-trained model of each method can be found in the corresponding references.

For object detection, using multiple object detectors will be infeasible and will cause very lengthy experiments and analyses because of the nine image enhancement models utilized. Therefore, we focus our analysis on a single object detection model and ensure selecting a superior, recent, and widely used model. Regarding this, the YOLO family has been extensively utilized in underwater scenarios, demonstrating competitive performance [37–39]. We opted to use YOLO-NAS [40] as a superior variant in the YOLO family, which has been recently used in [23,41–43], to train a total of 10 object detection models in which the models are trained and tested on each image set separately. We call the model that was trained on the original non-enhanced images as the *original detector* and the models that were trained on the enhanced images as *domain detectors*. We unified the input image *resolution* to 800×600 and set the *batch size* to 16, the *optimizer* to AdamW, the *weight decay* to 0.00001, the *model's architecture* to large, and the *pre-trained weights* to COCO. This is carried out using the SuperGradients [40] library (version 3.5) and a Linux server (Ubuntu 18.04 LTS) with a Tesla V100 GPU from Nvidia, Santa Clara, CA, USA. For testing and inference, we use a non-maximum suppression threshold of 0.7, a confidence threshold of 0.5, and an Intersection Over Union (IoU) threshold of 0.5. All models trained in this work can be found at https://github.com/Ali-Awad/Underwater-Image-Enhancement-and-Object-Detection.git, accessed on 24 December 2024. We implemented our experiments on the publicly available and recent RUOD dataset [20], comprising 14,000 high-resolution underwater images, 9800 of which are used for training and 4200 of which are used for testing with a total number of annotated instances of approximately 75,000 of 10 aquatic objects, including holothurian, echinus, scallops, starfish, fish, corals, divers, cuttlefish, turtles, and jellyfish. The use of the RUOD dataset in this study is aimed at facilitating the drawing of reliable and generalized conclusions because it is one of the largest recently published comprehensive underwater datasets composed of many smaller underwater datasets and showcasing different environments, sea objects, and color casts. We use the training–testing split provided by the authors to train and test the YOLO detector, while we use the entire dataset for inference using the pre-trained enhancement models.

Finally, we used non-reference evaluation metrics to assess enhancement performance and standard metrics for detection, as detailed below:

- The Underwater Image Quality Metric (UIQM) [44], which comprises color, contrast, and sharpness indices and is given by the following equation:

$$UIQM = \zeta_1 \times UICM + \zeta_2 \times UISM + \zeta_3 \times UIConM, \quad (1)$$

  where ζ_1, ζ_2, and ζ_3 are the weights of the Underwater Image Colorfulness Metric (UICM), Underwater Image Contrast Metric (UIConM), and Underwater Image Sharpness Metric (UISM), respectively. The UICM is calculated as a linear combination of the means and variances of two opponent color components in the HSV color space, namely, RG and YB. The significance of colors in object detection is debatable, with some studies [45,46] indicating a small to vital role for colors in the detection process. Secondly, the UISM is generated by applying a Sobel edge detector and Enhancement Measure Estimation (EME). Finally, the UIconM is directly generated by applying the logAMEE measure [47]. Higher UIQM, UICM, UIConM, and UISM scores indicate better image quality.

- CCF [48], which represents a linear regression model of colorfulness, contrast, and fog density indices and is expressed as follows:

$$CCF = \omega_1 \times Color + \omega_2 \times Contrast + \omega_3 \times Fogdensity \quad (2)$$

where ω_1, ω_2, and ω_3 are the weights of the colorfulness, contrast, and fog density features, respectively. The colorfulness index is calculated in a similar fashion to the UICM but through the CIELAB color space. Secondly, the contrast is calculated by dividing the image into blocks, applying a Sobel edge filter, and summing the edge values of all blocks after subtracting the total intensity of the image from each individual pixel. Finally, the fog density is based on the difference in statistical features of foggy and natural images. These distinct features are fitted into two multivariate Gaussian models, and the Mahalanobis distance between the two models is then calculated. A higher CCF score indicates better image quality.

- Precision (Pr), which indicates the ratio of the number of correct positive predictions to the total number of positive predictions made by the model. The Pr is given by the equation below:

$$Precision\ (Pr) = \frac{TP}{TP + FP} \quad (3)$$

where TP is the true positive, i.e., a prediction matching the ground truth; FP is the false positive, i.e., a background is considered as an object; and FN is the false negative, i.e., an object is considered as a background.

- Recall (Rc), which indicates the ratio of the number of correct positive predictions to the total number of actual instances. The Rc is given by the equation below:

$$Recall\ (Rc) = \frac{TP}{TP + FN} \quad (4)$$

- Average Precision (AP), which is the area under the precision–recall curve. There are multiple ways to calculate the mAP, out of which, the COCO evaluator [49] is adopted in this work. The mAP is produced here by ordering the predictions from all images by confidence scores and calculating the accumulative Pr and Rc at each threshold. This is achieved by dividing the recall axes into 101 equal segments from [0, 101], as defined in the COCO evaluator [49] and combining the calculated areas under each segment, expressed as

$$Average\ Precision\ (AP) = \frac{1}{101} \sum_{Rc_i} Pr(Rc_i) \quad (5)$$

Practically, the mean Average Precision (mAP) is used, which is the AP average across all classes based on the set Intersection Over Union (IoU) threshold, where the IoU represents the ratio of the overlap area (intersection) between the predicted bounding box and the ground truth bounding box to the area of their combined region (union). The calculation of the mAP in the superGradients library used in this paper is based on the COCO evaluator [40].

All of the metrics used in this study were applied without any modifications, using the original weights and parameter values as specified in the original papers. GitHub links to the source codes of those metrics can be found in their corresponding references.

2.2. Evaluation Procedures

In this section, we define our evaluation schemes, concepts, and experiments. Before delving further into this section, we provide a few definitions:

- *Image Sets*: The 4200 testing images of the RUOD dataset are enhanced separately using the nine selected image enhancement models, producing nine enhanced image sets in addition to the original set.
- *Per-Image mAP*: The mAP for each image is calculated separately using only the inferred bounding boxes of that particular image. This way, we can assess the detection performance for each enhanced image and compare it with the corresponding original image.
- *Image Groups*: A group is a queried subset of an image set based on a condition. We write the query in a superscript format. For instance, $ACDC^{\geq Org}$ refers to the subset of the enhanced ACDC set that achieved a higher mAP than its corresponding original subset. All image groups generated in this work can be found at https://github.com/Ali-Awad/Underwater-Image-Enhancement-and-Object-Detection.git, accessed on 24 December 2024.
- *Mixed Image Sets*: These sets are generated by combining images from the original set and one enhanced set. Specifically, the mAP of each enhanced image is compared with the mAP of its corresponding original image; then, the highest mAP image is selected. This generates a new, equally sized mixed set but with a higher or equal total mAP than both initial sets.
- *Change in Metrics*: To be able to compare each enhanced group with its corresponding original group, we present the results based on the change in the metrics instead of presenting absolute metrics. This is conducted as follows:

$$\Delta Metric = Metric_{(Enh)} - Metric_{(Org)} \tag{6}$$

where Δ refers to the change, *Metric* refers to any of the selected metrics for this study, *Enh* refers to the enhanced image group, and *Org* refers to the original image group.

Based on the previously declared definitions, we generated the enhanced image sets using the selected enhancement models and trained a separate YOLO-NAS model on each image set, including the original set, to produce the domain and original detectors. After training, the *per-image mAP* was calculated for each image in each testing set using its respective domain detector. Then, each set was divided into two groups: (1) images with a *per-image mAP* higher or equal to their corresponding original images and (2) images with a *per-image mAP* lower than their corresponding original images. For each group, the standard total mAP for the entire group was calculated in the standard way. Subsequently, we generated the mixed sets by replacing the groups of lower mAPs of each enhanced set with the corresponding original groups. The results were then evaluated quantitatively and qualitatively.

Quantitative Evaluation: An evaluation of quality and detection performance for each set and group is presented in the quantitative evaluation. For image sets, we presented the results using absolute metrics values, and for groups, we presented the results using the *change in metrics*. This helps us make better comparisons due to the fact that the grouping is based on comparison with the original set by definition. Because our goal is to assess the overall enhancement impact on detection (i.e., the change in the TP, FP, and FN), we used mAP_{50} for grouping and evaluation because it is not very sensitive to minor changes in the IoU of the predicted bounding boxes with the ground truth. Furthermore, we conducted this evaluation based on three *comparison levels*, including the following:

1. Enhanced groups against corresponding original groups.
2. Enhancement sets produced by each model against each other.
3. Most importantly, per-set group comparison.

Qualitative Evaluation: To produce an effective qualitative evaluation, we first conducted an overlap analysis between image sets and produced three unique groups of images where all domain detectors achieved consistent detection performance, as follows:

- Enhanced images of unanimous superior detection performance (Enh > Org).
- Enhanced images of unanimous equal detection performance (Enh = Org).
- Enhanced images of unanimous inferior detection performance (Enh < Org).

Although we present images from the TECBF and ICSP models in the qualitative evaluation, they are excluded from the overlap analysis because they produce severely distorted images that might lead to inaccurate conclusions. This is especially true for the ICSP model, for which we verified its code and parameters and found it to produce overexposed images even when other researchers implemented it [50]. All unanimous groups generated in this work can be found at https://github.com/Ali-Awad/Underwater-Image-Enhancement-and-Object-Detection.git, accessed on 24 December 2024.

3. Results and Discussion

In this section, we present an extensive quantitative and qualitative evaluation of the produced enhanced sets and groups. In addition, we discuss the results of the resulting group overlap study.

3.1. Quantitative Evaluation

This section includes image quality evaluation using the image quality metrics discussed in Section 2.1 and object detection evaluation using the detection metrics also described in the same section. In addition, this section includes the results of the mixed set presented in Section 2.2.

Image Quality Evaluation: After producing the enhanced image sets and dividing them into groups, as discussed in Section 2, we present the results obtained by the selected quality metrics, as shown in Table 1. The individual components of the UIQM, namely, the UICM, UIConM, and UISM representing the color, contrast, and sharpness, respectively, further show how the selected models performed on a finer level. Furthermore, the change in metric values is generally positive, indicating that all enhancement models produced higher-quality images compared to the original images, except for the PCDE and ICSP models. The ACDC, TEBCF, and BayesRet models achieved the highest positive change in the UIQM, while learning-based methods, including the AutoEnh, Semi-UIR, USUIR, and TUDA, achieved, on average, a substantially lower UIQM. More contrasting performances are shown with the CCF metric, where the TEBCF had the highest value by a large margin compared to all other models, while the PCDE and TUDA models produced negative values, indicating a worse quality than the original images. Other models observed a similar positive trend as in the UIQM. The UICM, UIConM, and UISM models show how some models are better at enhancing colors, such as BayesRet and PCDE, while others are better at enhancing contrast, such as ACDC, and, finally, how others are better at enhancing sharpness, such as the TEBCF model. On the other hand, comparing the two groups comprising each set demonstrates a remarkable observation; that is, the first group of each set represented as $Enh^{\geq Org}$ always achieved lower quality compared to the second group represented as $Enh^{<Org}$. We conclude that *the utilized enhancement metrics are designed to produce visually pleasing images for human perception, not for machine perception. As a result, 'high quality' by these metrics does not necessarily translate to improved object detection performance. This highlights the need to develop new quality metrics that account for both human*

and machine vision to bridge this gap. In addition, the high-quality values of the underperforming group $Enh^{<Org}$ *could be associated with over-enhancement, thus leading to degraded detection performance.* Although this comparison gives us an indication of the type of enhancement–detection relationship, it is important to note the size difference between the two groups. In particular, the size of the $Enh^{\geq Org}$ group at nearly 80% is mostly four times larger than the $Enh^{<Org}$ group regardless of the utilized enhancement model.

Table 1. Quality evaluation of the generated image groups using the mean of the UIQM [44], UCIQE [51], CCF [48], UICM [44], UIConM [44], and UISM [44] on the RUOD [20] dataset. Δ refers to the change in the metric after enhancement, with higher change values indicating better performance across all metrics. The query by which each group is generated is shown as a superscript.

Groups	Size	Change in Quality Metrics				
		$\Delta UIQM$	ΔCCF	$\Delta UICM$	$\Delta UIConM$	$\Delta UISM$
$ACDC^{\geq Org}$	78%	2.41	4.09	41.37	0.20	1.77
$ACDC^{<Org}$	22%	2.89	6.42	51.47	0.24	2.01
$TEBCF^{\geq Org}$	80%	2.25	10.06	33.92	0.19	2.11
$TEBCF^{<Org}$	20%	2.54	11.03	39.03	0.21	2.32
$BayesRet^{\geq Org}$	81%	2.42	5.17	48.73	0.16	1.55
$BayesRet^{<Org}$	20%	2.97	7.53	56.92	0.22	1.95
$PCDE^{\geq Org}$	80%	1.11	−6.28	47.34	−0.06	−0.03
$PCDE^{<Org}$	20%	1.52	−4.92	58.63	−0.05	0.17
$ICSP^{\geq Org}$	68%	−0.16	5.13	−4.70	−0.05	0.46
$ICSP^{<Org}$	32%	0.07	4.28	0.56	−0.03	0.58
$AutoEnh^{\geq Org}$	84%	1.37	1.16	34.72	0.07	0.49
$AutoEnh^{<Org}$	16%	1.90	2.57	49.39	0.08	0.69
$Semi\text{-}UIR^{\geq Org}$	80%	1.60	6.09	32.67	0.11	0.97
$Semi\text{-}UIR^{<Org}$	20%	2.27	6.77	48.51	0.14	1.32
$USUIR^{\geq Org}$	82%	1.25	4.08	27.89	0.08	0.64
$USUIR^{<Org}$	18%	1.84	6.44	40.22	0.12	1.00
$TUDA^{\geq Org}$	79%	1.90	−0.65	40.51	0.13	0.95
$TUDA^{<Org}$	21%	2.68	2.06	54.27	0.20	1.46

Object Detection Evaluation: First, we evaluated the detection performance over entire sets using the selected detection metrics, as shown in Table 2. This table provides a general understanding of the effect of enhancement on the detection performance before a further detailed group performance analysis, which will be presented later. It is clearly observable that enhancement did not increase the total number of correctly detected objects represented by the TP. The only special case is the AutoEnh with approximately 100 more correctly detected objects compared to the original. This could be attributed to the Neural Architecture Search utilized by the AutoEnhancer, which allows it to optimize its architecture for underwater conditions. Most other models had a similar close performance to the original, with ICSP being the exception at 14,713 TPs compared to 16,602 TPs of the original detector. ICSP heavily relies on sparsity prior to design, which may lead to suboptimal performance in scenarios of mixed or uniform lighting, such as those of the RUOD dataset. Moving on to FP, we find that most enhancement models notably increased FPs, reaching a number as high as 9252 by the TEBCF model compared to only 6309 FPs by the original detector. The fusion process utilized by the TEBCF model combines features from multiple scales, which can introduce artifacts if the weighting maps are not accurately

estimated. This results in a halo effect in the resultant images, destroying the object's edges and leading to very high FP. Of note, the two models that previously achieved the worst in terms of the quality metrics, namely, PCDE and ICSP, here achieved lower FPs than the original image. This might be attributed to the fact that these two models degrade the quality of the images, introducing noise, and possibly destroying the edges of potential objects. The FN results are very similar across all models, except for AutoEnh and TEBCF, which are the only two models that were able to better avoid detecting false objects by a small margin of 128 and 72, respectively, compared to the original. Looking at the Pr column, the highest precision is achieved by PCDE, with a value of 0.73, indicating that this model has the best performance in minimizing false positives relative to the number of correct detections. This could be attributed to the use of dual priors in the PCDE architecture, namely, spatial priors for the base layer and texture priors for the detail layer, thus allowing for the effective decomposition of images into layers and enabling precise adjustments for contrast and detail enhancement. This layered processing ensures better preservation of the fine textures necessary for detection while improving global contrast. The AutoEnh model and Original have a precision of 0.72, which is the second-highest, indicating strong but slightly lower performance in terms of precision compared to PCDE. The model with the lowest precision is TEBCF at 0.64, indicating a higher rate of false positives compared to other models. Analyzing the Rc column, the highest recall values are achieved by Original, AutoEnh, and TEBCF, each with a value of 0.73. This indicates that these models are the best at capturing the majority of true positives, minimizing missed detections. The lowest recall is recorded by ICSP at 0.65, indicating that it misses a relatively higher number of true positives compared to other models. Finally, we notice that the Original detector achieved the highest mAP at 0.85 along with two other domain detectors, namely, PCDE and AutoEnh, while all other models had a slightly decreased performance, with ICSP exhibiting the lowest mAP at 0.8. We conclude that *different enhancement algorithms affect the mAP in different ways, with most models achieving higher FPs than the original images. In addition, enhancing all images of a set results in decreased detection performance or, at best, similar detection performance compared with the original set in terms of the mAP. Furthermore, we notice that some enhancement models have the ability to achieve a higher TP or lower FP compared with the original images while achieving the same mAP, indicating a positive impact of enhancement on detection.*

Table 2. Detection evaluation of the generated image sets using the mean of the TP, FP, FN, Pr, Rc, and mAP on the RUOD [20] dataset. The highest scores are highlighted in **bold**.

Sets	TP	FP	FN	Pr	Rc	mAP
Original	16,602	6309	6248	0.72	**0.73**	**0.85**
ACDC	16,268	7600	6533	0.68	0.71	0.84
TEBCF	16,599	9252	6176	0.64	**0.73**	0.84
BayesRet	16,284	7291	6525	0.69	0.71	0.84
PCDE	16,404	**5916**	6440	**0.73**	0.72	**0.85**
ICSP	14,713	6256	8064	0.70	0.65	0.80
AutoEnh	**16,705**	6516	**6120**	0.72	**0.73**	**0.85**
Semi-UIR	16,292	6844	6530	0.70	0.71	0.84
USUIR	16,425	6616	6412	0.71	0.72	0.84
TUDA	16,262	7035	6574	0.70	0.71	0.84

Mixed Sets: Based on the observation that some enhanced images increase detection performance while others decrease it, we attempted to find the potential of image enhancement to increase the detection performance by mixing enhanced images with original images to comprise new sets. In particular, for each enhanced set, we used the same

$Enh^{\geq Org}$ group from Table 1 and replaced the $Enh^{<Org}$ group with a new group defined as $Org^{>Enh}$, meaning that we replaced underperforming enhanced images with their corresponding original images. The results of the newly generated mixed sets containing both enhanced and original images are shown in Table 3. We first noticed that around 80% of each mixed set is composed of enhanced images, while original images comprise only 20% of images, indicating the positive role of enhancement in the majority of cases. The mix ratio slightly varies from one set to another based on how the images of each enhancement model performed. The AutoEnh images were able to achieve a contribution as high as 84%, while ICSP achieved a contribution as low as 68%, aligning with its low quality images and detection performance. All new mixed sets had slightly increased TPs, slightly reduced FPs, and substantially reduced FNs compared to the non-mixed sets in Table 2. In addition, all new mixed sets had a slight increase in Pr and a relatively greater increase in Rc. PCDE achieved the highest precision in both tables at 0.73, indicating its effectiveness in decreasing FPs while maintaining good TPs. Most importantly, an mAP improvement was achieved by most mixed sets. Different mixed sets excel in distinct areas of detection performance. For example, AutoEnh achieved a higher number of True Positives (TPs), while PCDE and TEBCF are more effective in minimizing False Positives (FPs) and False Negatives (FNs), respectively. We conclude that *enhancement has the potential to improve object detection quantitatively, especially since all enhancement models improved detection, which reduces the possibility that these improvements stem from statistical errors. In addition, mixing enhanced images with originals can increase the mAP in different ways, i.e., by increasing TP or decreasing FP or FN, depending on the utilized enhancement algorithm. Finally, we believe that some enhancement processes may have altered the characteristics of the objects, making them harder to detect. This is evidenced by a reduction in false negatives (FNs) when the enhanced images were replaced with their original counterparts.*

Table 3. Detection evaluation of the generated mixed image sets using the mean of the TP, FP, FN, Pr, Rc, and mAP on the RUOD [20] dataset. The mix ratio represents the size distribution of the groups comprising each set. This ratio is based on the detection performance of each set of enhanced images. The highest scores are highlighted in **bold**. The query by which each group is generated is shown as a superscript.

Mixed Groups	Mix Ratio		Detection Metrics					
	Enh.	Org	TP	FP	FN	Pr	Rc	mAP
Original (Org)	-	100%	16,602	6309	6248	0.72	0.73	0.85
$ACDC^{\geq Org}$ + $Org^{>ACDC}$	78%	22%	17,043	7219	5792	0.70	**0.75**	**0.87**
$TEBCF^{\geq Org}$ + $Org^{>TEBCF}$	80%	20%	17,121	8463	**5695**	0.67	**0.75**	0.86
$BayesRet^{\geq Org}$ + $Org^{>BayesRet}$	81%	19%	16,999	7124	5847	0.70	0.74	0.86
$PCDE^{\geq Org}$ + $Org^{>PCDE}$	80%	20%	16,993	**6159**	5865	**0.73**	0.74	**0.87**
$ICSP^{\geq Org}$ + $Org^{>ICSP}$	68%	32%	16,584	6544	6248	0.72	0.73	0.86
$AutoEnh^{\geq Org}$ + $Org^{>AutoEnh}$	84%	16%	**17,127**	6522	5710	0.72	**0.75**	**0.87**
$Semi\text{-}UIR^{\geq Org}$ + $Org^{>Semi\text{-}UIR}$	80%	20%	16,927	6820	5920	0.71	0.74	**0.87**
$USUIR^{\geq Org}$ + $Org^{>USUIR}$	82%	18%	17,005	6617	5849	0.72	0.74	**0.87**
$TUDA^{\geq Org}$ + $Org^{>TUDA}$	79%	21%	17,000	6857	5850	0.71	0.74	**0.87**

3.2. Qualitative Evaluation

In this section, we visually inspect the inference results and quality of three unique groups of images. In particular, we check for overlap between the generated enhanced sets to group images of unanimous superior, equal, and inferior detection performance compared to the original images, as previously described in Section 2.2. Unanimous here means that an enhanced image exclusively belongs to one of the groups, regardless of the

enhancement model. The qualitative analysis is illustrated in Figure 1, where two samples are presented from each group. Furthermore, the quantitative results of the overlap analysis are presented in Table 4, showing statistics about the three unanimous groups found in the RUOD dataset containing images of matching detection performance across all enhancers. Through this analysis, we noticed that there are some images that have consistent detection performance across all domain detectors compared to the original detector. Although the extreme outperforming and underperforming groups are small, at an average of 2.7%, the fact that there are two groups where enhanced images always result in higher or lower mAP compared to non-enhanced images regardless of the enhancement algorithm is very notable. Moreover, having enhanced and original images performing equally for almost half of the testing dataset tells us that *enhancement, for a reasonable amount of cases, does not affect the mAP regardless of the enhancement model being used*. This is even more true if the conditions for generating those groups are more flexible, e.g., generating those groups if only three or more enhancement algorithms achieved unanimous performance. Visually analyzing and understanding the detection performance of every image produced by each enhancement algorithm in each unanimous group is an unfeasible task because every domain detector is trained on the enhanced images produced by its corresponding enhancement algorithm. Nonetheless, by thoroughly visually inspecting those groups and presenting representative samples in Figure 1, we noticed that in most of the cases where there are too many objects in an image, the detection performance of those images is highly likely to increase or decrease after enhancement. We hardly noticed any difference between the three unanimous groups when visually inspecting the images, as scenes, objects, and casts overlap. Furthermore, we hardly noticed any difference in quality between the unanimous groups as the colors, contrast, and sharpness of the images appear to be similar. We observed how the mAP is affected in each group and noticed an increase in the IoU threshold and TP and a decrease in the FN in the Enh > Org group, which was expected. However, this group still has a high false positive rate. This means that *enhancement still introduces false positives even when improving the detection performance.* One example of this is the falsely detected scallop object by the ACDC version of Sample #2. Furthermore, some of these FPs are actually correctly detected objects but were miss-annotated by human annotators due to the low visibility of the original images. Sharing the same trend, the Enh < Org group also has a high number of objects in each image. This could be explained by the type of background present in an image. For example, most enhanced versions of Sample #2 of this group made fish look like the shadow of a rock in a rocky background. We also noticed that vague objects are more likely to blend with the background after enhancement, as shown in Sample #1 of the Enh < Org group. Finally, we noticed a relatively smaller number of objects in the Enh = Org group, with the enhancement almost achieving similar true positives all the time. Another observation about the images of this group is that they mostly contain clear, distinct, and non-occluded objects, resulting in consistent performance between original and enhanced images. We conclude that, *firstly, images with complex backgrounds or a high number of objects are more likely to have diverging detection performance when using or omitting enhancement. Secondly, image quality does not have a direct relationship with detection performance because images of similar qualities belong to different unanimous groups. Lastly, we noticed very low annotation accuracy where many real objects are missed in all groups due to severely degraded original images, which hinders researchers from drawing reliable conclusions and reveals a promising role for enhancement earlier in the enhancement–detection pipeline, i.e., during the annotation process.*

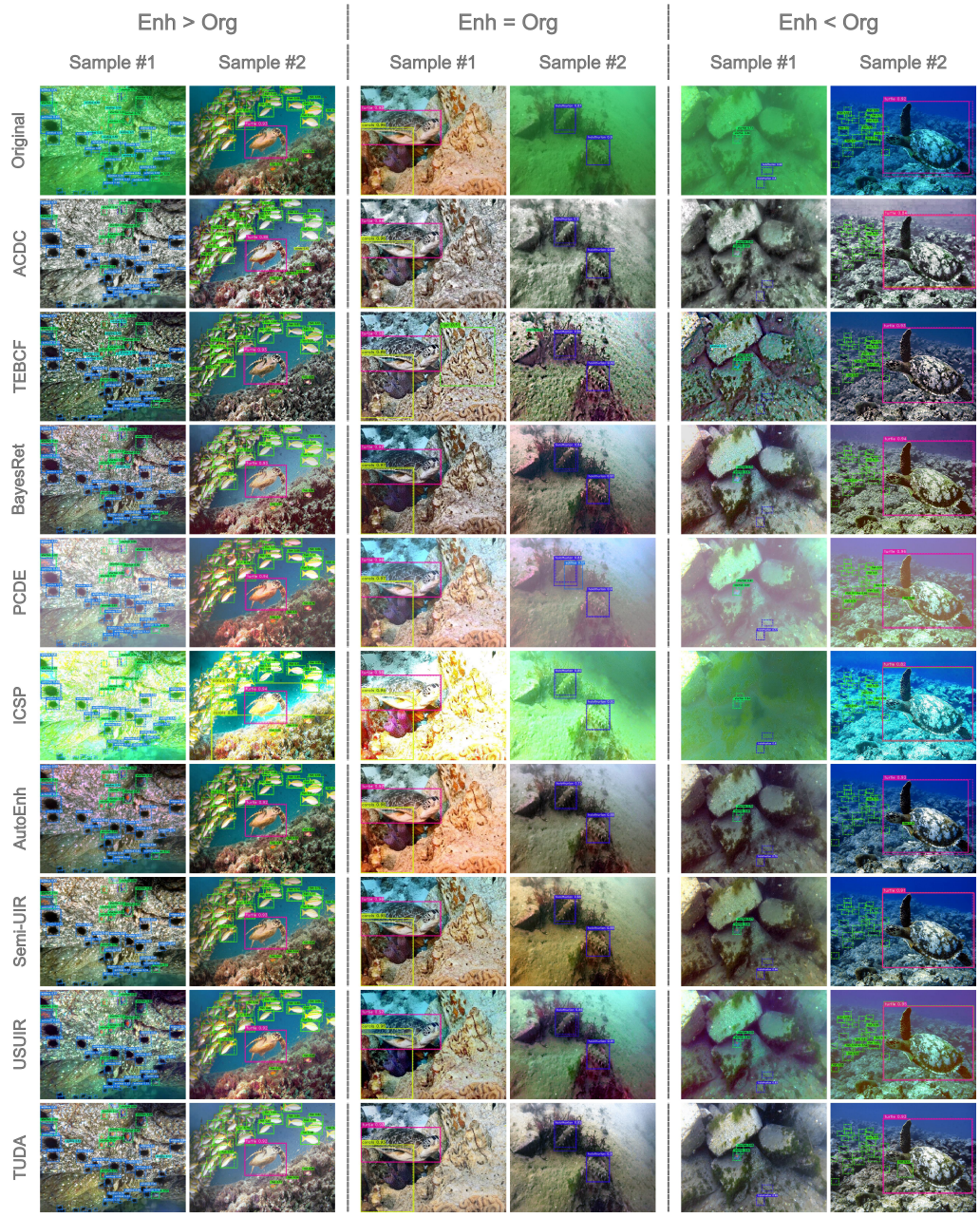

Figure 1. Samples of visualized groups where enhanced images (Enh) achieved higher, equal, and lower mAP compared to the original images (Org) regardless of the enhancement algorithm being used.

Table 4. Statistics for three groups where enhanced images achieved higher, equal, or lower mAP compared to the original images, using any domain detector. Percentages are calculated based on the total 4200 images in the RUOD test set.

Unanimous Groups	Total Images	% of RUOD Test Set
Enh (by any model) $^{>\text{Org}}$	109	2.60%
Enh (by any model) $^{=\text{Org}}$	1880	44.76%
Enh (by any model) $^{<\text{Org}}$	118	2.81%

4. Conclusions

In this paper, we conducted a detailed analysis of the enhancement–detection relationship using a per-image detection evaluation rather than a per-dataset detection evaluation.

This type of evaluation allowed for several types of analysis presented in this work, including categorizing images into groups based on their detection performance and evaluating the quality of the resultant groups, generating mixed sets of enhanced and original images that outperform the original image set, and generating unique groups of images of featured significance through specialized queries. Moreover, we checked for images that have identical detection performance across different enhancement models to rule out conclusions that are dependent on the enhancement algorithm and draw conclusions about enhancement in general. Furthermore, two small groups of enhanced images were found to consistently perform better or worse in terms of detection compared with the original images, regardless of the enhancement model. Those groups could grow much larger if the criteria for the number of enhancement algorithms utilized in this analysis are loosened. Based on the conducted experiments, we present the following findings:

- The image groups of lower detection performance achieved higher quality in terms of the utilized quality metrics. This indicates that those metrics are developed for human perception, not for machine perception. Higher values for those metrics could also indicate over-enhancement, which negatively impacts detection.
- Different enhancement algorithms impact the overall detection performance in different ways, e.g., lower TP or higher FP and FN.
- Image enhancement has a neutral or positive detection effect on the majority of images and a negative effect on a small percentage of images. This relatively small percentage of underperforming enhanced images is what causes the overall detection performance of enhanced sets to be lower than the original set.
- Enhancement, for nearly half the time, does not affect the mAP regardless of the enhancement model.
- Enhancement reveals hidden objects that are missed by human annotators due to the severely degraded original images.

Based on those conclusions, being able to predict such cases where enhancement will increase the detection performance could play a vital role in the development of future enhancement–detection algorithms. In addition, we think that enhancement may be used as early as during the annotation process to produce more reliable datasets. Further studies could confirm the positive impact of enhancement on detection when mixing enhanced and original images, particularly in scenarios with more robust annotations, diverse datasets, and different detectors.

Author Contributions: The individual contributions for this work are as follows: conceptualization, A.S., A.A. and S.P.; methodology, A.A., A.S. and E.L.; software, A.A.; analysis validation, A.A., A.S., S.P. and E.L.; investigation, A.A. and E.L.; resources, T.C.H. and P.C.E.; writing—original draft preparation, A.A., A.S. and E.L.; writing—review and editing, E.L., S.P. and T.C.H.; visualization, A.A. and E.L.; supervision, A.S., S.P., T.C.H. and P.C.E.; project administration, T.C.H. and P.C.E.; funding acquisition, T.C.H. and P.C.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by United States Geological Survey (USGS) (grant number: G23AS00029). The computing hardware was funded by the National Science Foundation (grant number: 2215734).

Data Availability Statement: All data generated in this work can be found at: https://github.com/Ali-Awad/Underwater-Image-Enhancement-and-Object-Detection.git, accessed on 24 December 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

Disclaimer: Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

1. Ditria, E.M.; Lopez-Marcano, S.; Sievers, M.; Jinks, E.L.; Brown, C.J.; Connolly, R.M. Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. *Front. Mar. Sci.* **2020**, *7*, 429. [CrossRef]
2. Li, X.; Zhuang, Y.; You, B.; Wang, Z.; Zhao, J.; Gao, Y.; Xiao, D. LDNet: High Accuracy Fish Counting Framework using Limited training samples with Density map generation Network. *J. King Saud-Univ.-Comput. Inf. Sci.* **2024**, 102143. [CrossRef]
3. Rahman, M.A.; Barooah, A.; Khan, M.S.; Hassan, R.; Hassan, I.; Sleiti, A.K.; Hamilton, M.; Gomari, S.R. Single and Multiphase Flow Leak Detection in Onshore/Offshore Pipelines and Subsurface Sequestration Sites: An Overview. *J. Loss Prev. Process Ind.* **2024**, 105327. [CrossRef]
4. Spectral Graph-Based Networks for Mooring Line Failure Detection on FPSO. In Proceedings of the ASME 2024 43rd International Conference on Ocean, Offshore and Arctic Engineering, Volume 1: Offshore Technology, Singapore, 9–14 June 2024. Available online: https://asmedigitalcollection.asme.org/OMAE/proceedings-pdf/OMAE2024/87783/V001T01A051/7360340/v001t01a051-omae2024-136899.pdf (accessed on 24 December 2024). [CrossRef]
5. Drap, P. Automatic Deep-Sea Amphorae Detection Using Optimal 2D Ultralytics Deep Learning. *Int. J. Comput. Digit. Syst.* **2024**, *17*, 1–11.
6. Chuang, M.C.; Hwang, J.N.; Ye, J.H.; Huang, S.C.; Williams, K. Underwater fish tracking for moving cameras based on deformable multiple kernels. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *47*, 2467–2477. [CrossRef]
7. Yin, F. Inspection robot for submarine pipeline based on machine vision. In *Proceedings of the Journal of Physics: Conference Series*; IOP Publishing: Beijing, China, 2021; Volume 1952, p. 022034.
8. Burns, C.; Bollard, B.; Narayanan, A. Machine-learning for mapping and monitoring shallow coral reef habitats. *Remote Sens.* **2022**, *14*, 2666. [CrossRef]
9. Chen, G.; Mao, Z.; Tu, Q.; Shen, J. A Cooperative Training Framework for Underwater Object Detection on a Clearer View. *IEEE Trans. Geosci. Remote Sens.* **2024**. [CrossRef]
10. Bajpai, A.; Tiwari, N.; Yadav, A.; Chaurasia, D.; Kumar, M. Enhancing Underwater Object Detection: Leveraging YOLOv8m for Improved Subaquatic Monitoring. *SN Comput. Sci.* **2024**, *5*, 1–10. [CrossRef]
11. Song, G.; Chen, W.; Zhou, Q.; Guo, C. Underwater Robot Target Detection Algorithm Based on YOLOv8. *Electronics* **2024**, *13*, 3374. [CrossRef]
12. Yuan, X.; Guo, L.; Luo, C.; Zhou, X.; Yu, C. A survey of target detection and recognition methods in underwater turbid areas. *Appl. Sci.* **2022**, *12*, 4898. [CrossRef]
13. Hambarde, P.; Murala, S.; Dhall, A. UW-GAN: Single-image depth estimation and image enhancement for underwater images. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [CrossRef]
14. Panetta, K.; Kezebou, L.; Oludare, V.; Agaian, S. Comprehensive underwater object tracking benchmark dataset and underwater image enhancement with GAN. *IEEE J. Ocean. Eng.* **2021**, *47*, 59–75. [CrossRef]
15. Tang, Y.; Iwaguchi, T.; Kawasaki, H.; Sagawa, R.; Furukawa, R. AutoEnhancer: Transformer on U-Net Architecture Search for Underwater Image Enhancement. In Proceedings of the Asian Conference on Computer Vision, Macao, China, 4–8 December 2022; pp. 1403–1420.
16. Huang, S.; Wang, K.; Liu, H.; Chen, J.; Li, Y. Contrastive semi-supervised learning for underwater image restoration via reliable bank. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 18145–18155.
17. Fu, Z.; Lin, H.; Yang, Y.; Chai, S.; Sun, L.; Huang, Y.; Ding, X. Unsupervised underwater image restoration: From a homology perspective. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 28 February and 1 March 2022; Volume 36, pp. 643–651.
18. Saleem, A.; Paheding, S.; Rawashdeh, N.; Awad, A.; Kaur, N. A non-reference evaluation of underwater image enhancement methods using a new underwater image dataset. *IEEE Access* **2023**, *11*, 10412–10428. [CrossRef]
19. Pei, Y.; Huang, Y.; Zou, Q.; Zhang, X.; Wang, S. Effects of image degradation and degradation removal to CNN-based image classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 1239–1253. [CrossRef]
20. Fu, C.; Liu, R.; Fan, X.; Chen, P.; Fu, H.; Yuan, W.; Zhu, M.; Luo, Z. Rethinking general underwater object detection: Datasets, challenges, and solutions. *Neurocomputing* **2023**, *517*, 243–256. [CrossRef]
21. Wang, Y.; Guo, J.; He, W.; Gao, H.; Yue, H.; Zhang, Z.; Li, C. Is Underwater Image Enhancement All Object Detectors Need? *IEEE J. Ocean. Eng.* **2023**, *49*, 606–621. [CrossRef]
22. Bolya, D.; Foley, S.; Hays, J.; Hoffman, J. Tide: A general toolbox for identifying object detection errors. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part III 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 558–573.
23. Awad, A.; Zahan, N.; Lucas, E.; Havens, T.C.; Paheding, S.; Saleem, A. Underwater simultaneous enhancement and super-resolution impact evaluation on object detection. In Proceedings of the Pattern Recognition and Tracking XXXV, Baltimore, MD, USA, 15–16 April 2019; SPIE: National Harbor, MD, USA, 2024; Volume 13040, pp. 67–77.

24. Chen, X.; Lu, Y.; Wu, Z.; Yu, J.; Wen, L. Reveal of domain effect: How visual restoration contributes to object detection in aquatic scenes. *arXiv* **2020**, arXiv:2003.01913.
25. Available online: www.urpc.org.cn/ (accessed on 2 July 2024).
26. Alawode, B.; Dharejo, F.A.; Ummar, M.; Guo, Y.; Mahmood, A.; Werghi, N.; Khan, F.S.; Javed, S. Improving underwater visual tracking with a large scale dataset and image enhancement. *arXiv* **2023**, arXiv:2308.15816.
27. Zhang, J.; Zhu, L.; Xu, L.; Xie, Q. Research on the correlation between image enhancement and underwater object detection. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 5928–5933.
28. Wang, H.; Sun, S.; Bai, X.; Wang, J.; Ren, P. A reinforcement learning paradigm of configuring visual enhancement for object detection in underwater scenes. *IEEE J. Ocean. Eng.* **2023**, *48*, 443–461. [CrossRef]
29. Han, D.; Zhang, J.; Han, F.; Su, Z.; Yang, J.; Zhao, W. Enhancing DeepLabV3+ for Underwater Equipment Detection and Segmentation. In Proceedings of the ISOPE International Ocean and Polar Engineering Conference, Rhodes, Greece, 16–21 June 2024; ISOPE: Mountain View, CA, USA 2024; p. ISOPE-I.
30. Zhang, W.; Wang, Y.; Li, C. Underwater Image Enhancement by Attenuated Color Channel Correction and Detail Preserved Contrast Enhancement. *IEEE J. Ocean. Eng.* **2022**, *47*, 718–735. [CrossRef]
31. Yuan, J.; Cai, Z.; Cao, W. TEBCF: Real-world underwater image texture enhancement model based on blurriness and color fusion. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–15. [CrossRef]
32. Zhuang, P.; Li, C.; Wu, J. Bayesian retinex underwater image enhancement. *Eng. Appl. Artif. Intell.* **2021**, *101*, 104171. [CrossRef]
33. Zhang, W.; Jin, S.; Zhuang, P.; Liang, Z.; Li, C. Underwater image enhancement via piecewise color correction and dual prior optimized contrast enhancement. *IEEE Signal Process. Lett.* **2023**, *30*, 229–233. [CrossRef]
34. Hou, G.; Li, N.; Zhuang, P.; Li, K.; Sun, H.; Li, C. Non-uniform illumination underwater image restoration via illumination channel sparsity prior. *IEEE Trans. Circuits Syst. Video Technol.* **2023**, *34*, 799–814. [CrossRef]
35. Wang, Z.; Shen, L.; Xu, M.; Yu, M.; Wang, K.; Lin, Y. Domain adaptation for underwater image enhancement. *IEEE Trans. Image Process.* **2023**, *32*, 1442–1457. [CrossRef]
36. Awad, A.; Saleem, A.; Paheding, S.; Lucas, E.; Al-Ratrout, S.; Havens, T.C. Evaluating the Impact of Underwater Image Enhancement on Object Detection Performance: A Comprehensive Study. *arXiv* **2024**, arXiv:2411.14626.
37. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
38. Zhang, M.; Xu, S.; Song, W.; He, Q.; Wei, Q. Lightweight underwater object detection based on yolo v4 and multi-scale attentional feature fusion. *Remote Sens.* **2021**, *13*, 4706. [CrossRef]
39. Liu, K.; Peng, L.; Tang, S. Underwater object detection using TC-YOLO with attention mechanisms. *Sensors* **2023**, *23*, 2567. [CrossRef] [PubMed]
40. Aharon, S.; Louis-Dupont; Masad, O.; Yurkova, K.; Fridman, L.; Lkdci; Khvedchenya, E.; Rubin, R.; Bagrov, N.; Tymchenko, B.; et al. Super-Gradients. 2021. Available online: https://zenodo.org/records/7789328 (accessed on 24 December 2024).
41. Hamzaoui, M.; Aoueileyine, M.O.E.; Romdhani, L.; Bouallegue, R. An Efficient Method for Underwater Fish Detection Using a Transfer Learning Techniques. In Proceedings of the International Conference on Advanced Information Networking and Applications, Kitakyushu, Japan, 17–19 April 2024; Springer: Berlin/Heidelberg, Germany, 2024; pp. 257–267.
42. Ercan, M.F. Gesture Recognition for Human and Robot Interaction Underwater. In Proceedings of the 2023 IEEE 11th Conference on Systems, Process & Control (ICSPC), Malacca, Malaysia, 16 December 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 13–16.
43. Saleem, A.; Awad, A.; Paheding, S.; Marcarelli, A. Multi-class plant type detection in great lakes region using remotely operated vehicle and deep learning. In Proceedings of the Pattern Recognition and Tracking XXXIV, Orlando, FL, USA, 3–4 May 2023; SPIE: Bellingham, WA, USA, 2023; Volume 12527, pp. 34–40.
44. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [CrossRef]
45. Singh, A.; Bay, A.; Mirabile, A. Assessing the importance of colours for cnns in object recognition. *arXiv* **2020**, arXiv:2012.06917.
46. Funt, B.; Zhu, L. Does colour really matter? Evaluation via object classification. In Proceedings of the Color and Imaging Conference, Society for Imaging Science and Technology, Scottsdale, AZ, USA, 13–17 November 2022; Volume 26, pp. 268–271.
47. Panetta, K.; Agaian, S.; Zhou, Y.; Wharton, E.J. Parameterized logarithmic framework for image enhancement. *IEEE Trans. Syst. Man Cybern. Part B (Cybernetics)* **2010**, *41*, 460–473. [CrossRef] [PubMed]
48. Wang, Y.; Li, N.; Li, Z.; Gu, Z.; Zheng, H.; Zheng, B.; Sun, M. An imaging-inspired no-reference underwater color image quality assessment metric. *Comput. Electr. Eng.* **2018**, *70*, 904–913. [CrossRef]
49. Sharma, A. Mean Average Precision (mAP) Using the COCO Evaluator. In *PyImageSearch*; Chakraborty, D., Chugh, P., Gosthipaty, A.R., Huot, S., Kidriavsteva, K., Raha, R., Thanki, A., Eds. 2022. Available online: https://pyimg.co/nwoka (accessed on 24 December 2024).

50. Du, D.; Li, E.; Si, L.; Xu, F.; Niu, J.; Sun, F. A Physical Model-Guided Framework for Underwater Image Enhancement and Depth Estimation. *arXiv* **2024**, arXiv:2407.04230.
51. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [CrossRef]