Underwater Image Enhancement and Object Detection Based on GANs

Armaan C. Rao Sruthi Bijoy Vanathi G.

Supervisor: Ms. Lakshmipriya S.

Department of CSE, Sri Sivasubramaniya Nadar College of Engineering

Highlights of Proposed Model

Developed an underwater image enhancement and object detection (UIEOD) system that:

- Enhances underwater images using a modified Super Resolution Generative Adversarial Network (SRGAN).
- Detects the location of underwater organisms in the given scene using YOLOv3.
- Recognizes the detected organism as starfish, holothurian, echinus or scallop.
- Helps understand the effect of image enhancement in underwater object detection.
- Achieves 8.01% increase in average IoU of object detection from tradition method.

Functional Modules

- Data collection and pre-processing
  - Image enhancement dataset - UIEB Dataset
  - Object detection dataset - DUO Dataset
  - Extracting patches from UIEB dataset
  - Converting annotations of DUO dataset to conform to Darknet
- Train SRGAN image enhancement model on the UIEB dataset
- Enhance the DUO dataset using SRGAN model
- Train YOLOv3 object detection model on the DUO dataset and the Enhanced DUO dataset

Dataset Description

- Underwater Image Enhancement Benchmark (UIEB) Dataset
  - UIEB training set consists of 890 raw underwater images of varying sizes along with their corresponding reference images.
  - It contains a challenging set of 60 raw underwater images.
- Detecting Underwater Objects (DUO) Dataset
  - DUO contains a collection of diverse underwater images with traditional annotations.
  - DUO training set consists of 6671 underwater images.
  - The bounding box annotations of organisms in the training images are stored in a single JSON file.
  - The test set contains 1111 underwater images.

Proposed Method

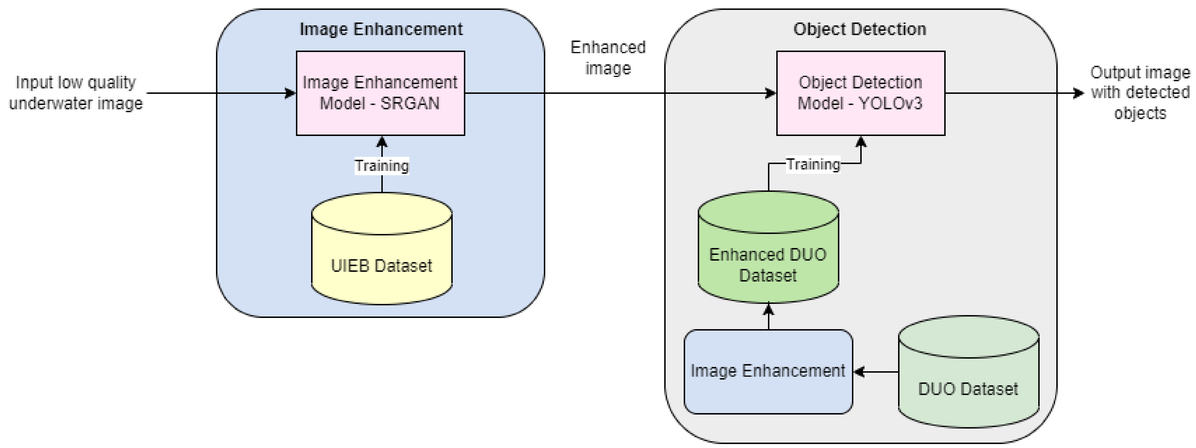

Figure 1. Architecture Diagram

Performance Analysis

SRGAN Performance

- Training time was around 2.5hr for SRGAN-2 and 8hr for SRGAN-3
- Testing time was approx. 1-2s for low res., 4-5s for medium res. and 8-10s for high res. images
- Evaluation metric used for SRGAN was Frechet Inception Distance (FID)
- SRGAN-3 performed best in terms of enhancement on images from UIEB dataset as it had least FID
- SRGAN-2 produced visually pleasing results on images from DUO dataset despite having slightly higher FID

Epochs	FID
2000	45.972
6000	41.691

Table 1. FID values for SRGANs trained on the whole dataset

YOLOv3 Performance

- YOLOv3 took 26 hours to train a single model for 8000 epochs.
- YOLOv3 took 58 seconds to identify objects on 1110 images and calculate evaluation metrics.

Metric	4000 epochs			7000 epochs		
	YOLO-v3	E2-YOLOv3	E3-YOLOv3	YOLO-v3	E2-YOLOv3	E3-YOLOv3
AP: Holothurian	81.25%	80.22%	81.55%	84.10%	81.19%	83.61%
AP: Echinus	87.37%	87.86%	87.83%	88.06%	85.66%	87.61%
AP: Scallop	49.67%	54.40%	54.89%	54.01%	56.69%	57.09%
AP: Starfish	87.91%	88.30%	86.66%	89.29%	88.36%	87.31%
TP	8510	8403	8591	8930	8160	8718
FP	1743	1479	1955	2421	1192	2080
FN	2007	2114	1926	1587	2357	1799
mAP	76.55%	77.69%	77.73%	78.87%	77.97%	78.91%
Precision	0.83	0.85	0.81	0.79	0.87	0.81
Recall	0.81	0.80	0.82	0.85	0.78	0.83
F1-score	0.82	0.82	0.82	0.82	0.82	0.82
Avg. IoU	67.49%	69.11%	66.37 %	63.80%	71.81%	65.48%

Table 2. YOLOv3, E2-YOLOv3, and E3-YOLOv3 evaluation comparison

Inferences

- General hue of images shifted from very blue to browns and grays as SRGAN was trained for more epochs.
- An increase in scallop AP was noted as SRGAN was trained for more epochs. Reason: white-grey scallops stood out better from background.
- Identified that the proposed **E2-YOLOv3** model trained for 7000 epochs gave the **best results**.
- When compared to other models, a significantly higher IoU was obtained with minimal loss in mAP.
- Hence, it had a highly improved ability to draw bounding boxes accurately (high IoU)
- The number of false positives output by E2-YOLOv3 is lower by 1229 compared to YOLOv3.
- A higher precision was also obtained, indicating more accurate classification.
- E2-YOLOv3 achieved a trade off between accurate bounding boxes and higher mAP.
- Domain quality has a strong positive correlation with the accuracy of bounding box positions.

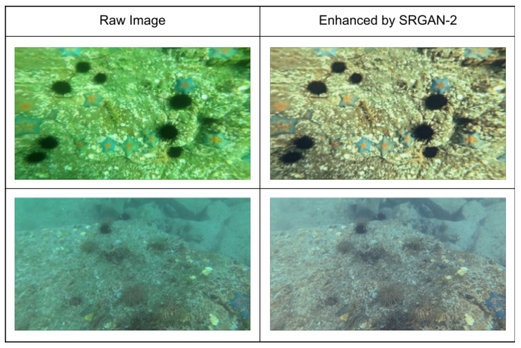

Figure 2. SRGAN-2 results on DUO dataset

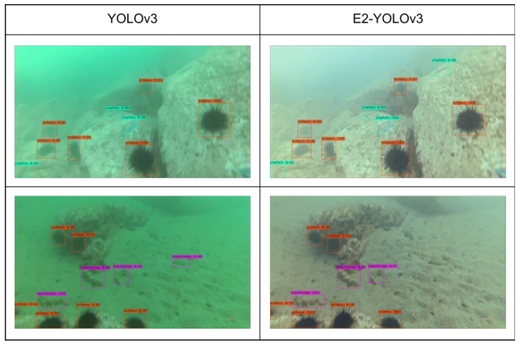

Figure 3. Object detection results for YOLOv3 & E2-YOLOv3