International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST218428

Underwater Image Enhancement with a Deep Residual Framework

Anuja Phapale[1], Puja Kasture[2], Keshav Katkar[2], Omkar Karale[2], Atal Deshmukh[2]
[1]Professor, BE Information Technology, A.I.S.S.M.S's Institute of Information Technology, Pune, India
[2]Student, BE Information Technology, A.I.S.S.M.S's Institute of Information Technology, Pune, India

Article Info
Volume 8, Issue 4
Page Number : 200-206

Publication Issue
July-August-2021

Article History
Accepted : 06 July 2021
Published : 13 July 2021

ABSTRACT

This paper focuses on framework developed with the goal to enhance the quality of underwater images using machine learning models for the Underwater Image enhancement system. It also covers the various technologies and language used in the development process using Python programming language.
The developed system provides two major functionality such as feature to provide input as image or video and returns enhanced image or video depending upon user input type with focus on more image quality, sharpness, colour correctness etc.
Keywords – Asynchronous Training, Edge Difference Loss, Residual Learning, Underwater Image Enhancement.

I. INTRODUCTION

Remote-Operated vehicles (R.O.V's) and vision-targeted Autonomous Underwater Vehicles (A.U.V's) have eventually impacted the exploration of marine life recently. For such vision-targeted explorations, clear underwater images are essential prerequisites. Although, the raw underwater images that we have obtained currently have colour distortion, low contrast, and blurred details as due to absorption, refraction of light, and scattering of light because of suspended particles in the water.

The absorption of red light is higher with respect to the absorption rate of green and blue light is better which states that the absorption of light by water has different effect throughout the visible spectrum. As a result, raw underwater images are mostly are blue or greenish as compared to that of an in-air image. Scattering of light can be divided into two types:

1. Forward Scattering and
2. Backward Scattering.

Forward scattering light usually contributes to the blurred texture features of underwater objects as it comes from the object. Backward scattering segregates the underwater image and causes noise as the light is reflected back before reaching the target object.

These challenges bring obstacles to tasks like tracking, segmentation, and vision-based navigation system. Therefore, for underwater vision tasks like deep sea exploration, underwater monitoring with the use of underwater vehicle vision, etc. Image restoration plays a very vital role. The reliability of underwater vision tasks is promoted by underwater image enhancement by increasing the underwater image colour contrast and reduction in the degradation caused by attenuation and scattering.

Thus, to deal with these challenges, this paper has proposed an underwater image enhancement solution using a deep residual framework. We provide a machine learning-based framework which improves the underwater image enhancement performance & aims to build a deeper network unlike other deep learning based underwater enhancement approaches that focus on the relation between weakly supervised learning and generative adversarial networks.

II. EXISTING SYSTEMS

Our enhancement framework has always been dependent on real physical images. There are existing methods that utilize various assumptions to achieve efficient solutions for underwater image enhancement. However, these methods share the common limitation although in some cases of scenes the embraced assumptions may fail.

However, after multiple researches about underwater image enhancements, existing systems avoided the techniques to reduce the issue of noise as seen in the output images of the existing algorithms of haze removal and it imbalances the colour of the input image. To tackle this, platforms have implemented the better solution. However, their accuracy and reliability is good as per existing systems.

III. SYSTEM DESIGN BASICS

Our Machine Learning Frame work is trained using Google Colab, which is tool made available for free By Google, it is easily accessible with a google account. The front-end for the system is developed using Pykinter.

This includes

- GUI For Selecting Files (Images to be enhanced)
- Utility Tool (Frame Extraction in case if Video is provided to the Algorithm)
- User account home (For Google Colab), and dashboard.
- User specific feature pages, etc.

The back-end is built using the Python, and Various Python Libraries (Tensorflow, PIL etc)

The Image Enhancement Framework consists of following features implemented so far:

- Under Water Image Enhancement
- Under Water Video Enhancement

The complete system design will be explained in detail as we move forward.

IV. RESEARCH AND LEARNING CURVE

Great development has been done so far in fields of underwater image enhancement with some following steps:

A. Setting Up The Training Data Set By CycleGAN:

Managed learning strategies have accomplished amazing outcomes in numerous fields of PC vision; however they have infrequently been utilized in submerged picture improvement. Since the obscured submerged pictures have no relating clear pictures as the ground truth, incredible managed learning techniques like CNN (Convolution Neural Networks), which need a ton of ''matched'' preparing information, can't be applied in this field.

To address this absence of preparing information, CycleGAN is utilized to create preparing information.

CycleGAN is a variation of a GAN, which can take in planning starting with one information conveyance then onto the next information appropriation without matched preparing information. CycleGAN comprises of discriminators DX and DY and generators F and G. The discriminator DX learns the highlights of the in-air pictures to decide whether the yield result is the in-air picture; the discriminator DY learns the highlights of the submerged pictures to judge whether the yield result is the submerged picture. The generator G learns the planning from the in-air pictures to the submerged pictures; the generator F takes in the planning from the submerged pictures to the in-air pictures to finish the shared change between the in-air pictures and the submerged pictures. Roughly 4500 in-air pictures were gathered from public informational indexes, for example, urban100 and bsd100 as space X, what's more, around 5000 submerged pictures were gathered from Flickr as space Y. CycleGAN is utilized to get familiar with the planning from X to Y. Consequently, CycleGAN is utilized to change clear in-air pictures to submerged style pictures, which produces matched preparing information for an incredible administered learning model. CycleGAN can likewise take in the process from Y to X, that is, the submerged pictures can be reestablished to in-air picture quality.

The picture sets comprising of the in-air pictures and the created ''engineered submerged pictures'' are utilized as the preparing set for the amazing administered learning models. Roughly 4000 sets were created.

B. Using VDSR for Underwater Image Enhancement:

The difficulties of submerged picture improvement are comparable to that of super-goal reproduction. Contrasted with an in-air picture, a crude submerged picture's tone is mutilated. Also, submerged pictures are obscured, and subtleties, such as the edges of submerged articles, are quieted or lost. In this manner, submerged picture upgrade requires shading deviation rectification and detail rebuilding. Additionally, super-goal remaking expects to reestablish pictures' subtleties. In the wake of providing preparing information for an incredible regulated learning model with CycleGAN, the profound super-goal (VDSR) model was acquainted with the submerged picture improvement task.

The VDSR model has 20 convolution layers. Every convolution layer utilizes 3∗3 size channels, with a step of 1 and zero-cushioning with 1 pixel. Such boundary settings guarantee that the goal of the info picture is indistinguishable from that of the yield picture. With the exception of the first and the last layers, every convolution layer has 64 channels. The main layer gets three-channel picture information as info, produces 64-channel highlight maps, and communicates them to the primary body of the organization. The last layer is the reproduction layer. It gets 64-channel include guides and yields three-channel leftover pictures. The leftover pictures are added to the info pictures to produce the last reestablished pictures. At the point when the VDSR model is utilized for super-goal remaking, the information picture is a high-goal picture created by bicubic addition of a low-goal picture, to such an extent that the info picture and the yield picture are a similar size. In this manner, when the VDSR model is applied to submerged picture reclamation, the size of information and yield pictures shouldn't be changed, and neither does the organization structure. Just proper preparing information are required for the organization to gain proficiency with the contrast among submerged and in-air pictures.

C. Edge Difference Loss:

Most picture to-picture interpretation models use per pixel distinction misfortune capacities, for example, the MSE or L1 misfortune work. The first VDSR model utilized MSE Loss, wherein it attempts to make the VDSR model become familiar with the pixel level contrast between the two pictures. Utilizing MSE Loss, the model can accomplish a higher pinnacle signal-to-commotion proportion (PSNR) score, be that as it may, the produced pictures don't give great enhanced visualizations; MSE Loss midpoints the distinctions at the pixel level and neglects to take more significant level data, like a generally structure, into account. Consequently, the MSE Loss works tends to normal the arrangement and make the picture subtleties excessively smooth, which isn't helpful for the upgrade of high-recurrence data. Because of the huge detail loss of submerged pictures, particularly concerning edge data, a punishment term is proposed called edge distinction misfortune (EDL). By punishing the models with EDL, the subtleties of created pictures are elevated to a more significant level.

D. Asynchronous Training Mode:

The organization needs to prepare each bunch twice. In the first round, EDL is utilized to compute the slopes and execute back engendering to refresh the loads of the organization; in the second round, the slope is determined utilizing MSE misfortune, what's more, engendered back to refresh the loads of the organization. Consequently, each clump is prepared twice, and the loads of the network are refreshed twice for each cluster. The non-concurrent preparing mode is received and EDL utilized in the principal preparing round. This technique can exploit EDL's capacity to give edge data and in this manner helps the network in reestablishing edge data and subtleties. Be that as it may, EDL's effect on the organization is confined constantly preparing, which restricts the organization to zero in on the distinction of the pixel level between the yield and name pictures, hence stifling the intensification impact of the Laplacian administrator on commotion.

Furthermore, if the two pieces of misfortune work are prepared with various loads by building a multi-term misfortune work, as the conventional multi-misfortune preparing model does, the suitable misfortune loads distribution should be recognized by means of countless analyses. This presents a test in deciding the ideal loads. Nonetheless, the loads' assurance is regularly unalterable and can forfeit the power of the model. Then again, non-concurrent preparing is identical to learning the k worth in condition by the organization; wherein the extent of the two sections in the misfortune work is naturally changed in accordance with accomplish the ideal arrangement. VDSR-P in non-concurrent preparing mode is indicated as VDSR-P-A.

E. Underwater Resnet (UResnet):

The VDSR model adequately accomplishes submerged pictures upgrade; notwithstanding, the VDSR model is a moderately shallow model, with 20 convolution layers and just one skip association. It is notable that the least complex way to upgrade the exhibition of the CNN model is stacking more layers. By and large, the more profound CNNs have more boundaries and better potential to deal with complex undertakings. Nonetheless, the profound CNNs are hard to prepare. ResBlocks and skip-associations can facilitate the preparation of profound CNNs. The VDSR model has one skip association. Anyway it does not use ResBlocks, which restricts the profundity of VDSR model. To additionally improve submerged picture upgrade, a more profound model is proposed, signified Underwater Resnet (UResnet). The proposed UResnet is a leftover learning model. It is made out of ResBlocks, which add the contribution of one convolution layer to the yield of the following convolution layer. Usage of ResBlocks guarantees that the data from the past layer can be completely communicated to the following layers. Stacking ResBlocks permits further organizations to be prepared.

UResnet is motivated by the super-goal recreation models EDSR and SRResnet. The proposed UResnet model is contained three primary areas: a head, body, and tail. Motivated by VDSR and EDSR, a significant distance skip association is incorporated from the head area yields to the body segment yields. The significant distance skip association adds the element data of the information layer to the yield layer of body, which obliges ResBlock modules to become familiar with the distinction between mark pictures what's more, input pictures. The head contains one convolution layer. Considering the tedious of preparing, the body part stacks 16 ResBlocks orchestrated in the accompanying request: [ConvBN-ReLU-Conv-BN]. The tail contains one convolution layer. In entirety, there are 34 convolution layers.

In the organization, a 3×3 convolution is utilized with a step of 1 pixel and a zero-cushioning of 1 pixel to keep up the state of highlight maps, which permits UResnet to acquire contributions with self-assertive shapes. In UResnet, the proposed EDL could be incorporated, and the non-concurrent preparing mode could be presented. Since of the reverberating impact of the BN layers on the submerged picture upgrade task, UResnet was planned with BN layers.

V. DETAILED SYSTEM WORKING

The entire system is based on the two main input types, i.e. underwater image and underwater video. This has been implemented using tkinter the python GUI library using one to one relationship to the User model itself. (Check tkinter with python online for more information.)
Their roles are as follows-

A. Image

- For image enhancement, we just need to give the folder path which consists of images.
- It will automatically retrieve image from that folder and it'll be provided to the CycleGAN and not to the utility tool that we made to convert it into an Image.
- After processing in CycleGAN, it'll give the result in new folder which will consist of the resultant enhanced image.

B. Video:

- For video processing, enhancing, the video path will be provided to the GUI Application.
- Video to be Enhanced can be of any format which supported by cv2 python library.
- Selected Video will be fetched by the utility tool and it'll automatically generate a new folder. The extracted images will be stored in the folder, all the frames from the video.
- 24fps is the conversion rate for frame extraction.
- Then the path of extracted features that is folder containing frames is provided to the CycleGAN.
- In CycleGAN will process the data and it'll process each and every frame that is present in the frames folder.
- Then these Enhanced frames are stored in the newly created folder. Then the folder which is contains the newly enhanced frames is provided to the utility tool again for to convert it into video.
- And then it'll create the video and it'll be provided to the GUI again.

VI. ADVANTAGES

Simple and user friendly yet eye catching UI, variety of features, and the enhancement system can be termed as the jewels of the system. Our system supports both image and video enhancement. The system itself has moderate system load, and is able to give better enhancement results with improved quality.

The system tests were performed on a laptop with following specifications-

- Core i5 7th Gen.
- Min. 4GB Ram
- 1 TB HDD
- NVIDIA 1660ti 6GB Graphics Card

During testing, while enhancing the video, the implemented models were able to give 24 fps.

VII. CONCLUSION

This paper proposes a submerged picture upgrade arrangement by a profound leftover system. Initially, CycleGAN was utilized to produce engineered submerged pictures as preparing information for the CNN models. Furthermore, the super-goal remaking model VDSR was brought into the field of submerged picture improvement, and the leftover learning model, Underwater Resnet (UResnet) was proposed. Moreover, the misfortune capacity and preparing mode were improved; a multi-term misfortune work was framed with the proposed edge contrast misfortune (EDL) and MSE misfortune lists. An offbeat preparing mode was likewise proposed to improve the execution of the multi-term misfortune work. The exploratory outcomes show the adequacy of the proposed strategies for submerged picture reclamation. EDL and the non-concurrent preparing mode can improve the exhibitions of CNN models on the submerged picture improvement task. The proposed UResnet-P-A model accomplished the best presentation with respect to both shading remedy and detail upgrade than different strategies we thought about, trailed by the proposed UResnet and VDSR-P-A (BN) models. It has additionally been shown that BN layers, however hurtful to super-goal recreation, are useful in the submerged picture upgrade task. BN layers can speed up union in preparing. Moreover, the consideration of BN layers can aid further reestablishing subtleties and improving picture contrast.

The proposed strategies can fundamentally improve the visual impacts of submerged pictures, which are useful to the execution of vision-based submerged assignments, like division and following. Besides, we consider applying the proposed strategies to the comparative spaces, like picture dehazing and super-goal remaking to test the over-simplification of the proposed techniques. We leave these to our future work.

VIII. REFERENCES

[1]. P.Liu, "Underwater Image enhancement using deep residual framework," 2019.
[2]. H.Qi, "Underwater Image enhancement using deep residual framework," 2019.
[3]. C.Zhenge, "Underwater Image enhancement using deep residual framework," 2019.
[4]. Z.Yu, "Underwater Image enhancement using deep residual framework," 2019.
[5]. P. Luo, "Underwater Image super-resolution using deep residual multipliers," 2019.
[6]. Y. Wang, "A deep CNN Method for the Underwater image enhancement," 2017.
[7]. J. Sattar, "Underwater image super-resolution with deep residual multipliers," 2019.
[8]. H.Lu, "Underwater Image super-resolution by descattering and fusion," 2017.
[9]. R.Hummel, "Image enhancement by histogram transformation,Compute Graph,Image Process," 2017.
[10]. S.Corchs, "Underwater Image processing. State of the art of restoration and image enhancement methods," 2014.

Cite this article as :

Anuja Phapale, Puja Kasture, Keshav Katkar, Omkar Karale, Atal Deshmukh, "Underwater Image Enhancement with a Deep Residual Framework", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 4, pp. 200-206, July-August 2021. Available at doi : https://doi.org/10.32628/IJSRST218428
Journal URL : https://ijsrst.com/IJSRST218428