Underwater Image Enhancement

Sonia Cromp (cromp)

5 April 2022

1 Introduction

Underwater image enhancement (UIE) is the task of improving the visual quality of images that were captured underwater. A common issue faced by prior UIE approaches is the scarcity of paired data: while there are many sources of un-enhanced images and a respectable quantity of enhanced image sources, there are very few datasets that provide one image scene in both un-enhanced and enhanced forms. Image translation is the task of learning a function to translate an image from a source to a target domain. Because there has been reasonable success in the area of unpaired image translation (where there are no paired examples between the source and target domains), a few UIE approaches have been inspired by unpaired image translation. In this project, I intend to survey pre-existing unpaired image translation UIE approaches and (time permitting) propose a novel solution combining paired and unpaired datasets.

There are many reasons that we may wish to see underwater, from archaeology and marine biology research to piloting military submarines. A large proportion of the ocean is as yet unexplored, and documenting these areas along with the effects of mankind and climate change on the deep sea is crucial before any such information may be permanently lost. Further, from a computer science research perspective, underwater image enhancement is a challenging but interesting problem because it allows for the application of image style translation techniques from elsewhere in computer vision and machine learning.

2 Current project status

Since submitting the proposal, I have obtained compute resources, procured datasets and code for several state-of-the-art image enhancement architectures and planned out more of the architecture for the newly proposed approach.

Just after spring break (23 March), I requested an account on the Euler cluster in order to have compute resources for training models. I was under the impression that account creation is an automatic process when laying out my original plan for the project, but I gained access to Euler for the first time yesterday (4 April). However, I was still able to do some work on the project while I waited. Specifically, I have procured several standard UIE datasets (over 60GB total while compressed) recommended by [1] and learned more about and decided to focus on 2-3 architectures that are designed for unpaired data. I had time to read more about each approach and prepare to run or implement them.

CycleGAN[10] is designed to learn a mapping from images in a source domain X (such as un-enhanced underwater) to images in a target domain Y (such as enhanced underwater). To do so, the model is composed of a forwards and a backwards generator to learn functions $F : X \to Y$ and

$G : Y \to X$ respectively, along with source and target discriminators. For some image x in the source domain, the source discriminator predicts whether x was generated by function G or was drawn from the provided training dataset. The target discriminator has an analogous role for the target data distribution. The discriminators enforce that generated data in domains X and Y look similar to the provided training data. Although our goal is to learn F, we also learn G in order to enforce that the translation of the translation of an image match with the image itself, i.e. $G(F(x)) = x$ and $F(G(y)) = y$ for images x and y from source and target domains respectively. This combination of translations is a "cycle", for which CycleGAN[10] is named, and acts as CycleGAN's[10] approach to learning from unpaired datasets. CycleGAN[10] is a popular image translation architecture (even outside of underwater image enhancement) and serves as a base for many later works in underwater image enhancement.

MCycleGAN[6] draws heavily from CycleGAN[10], but adds a degree of domain knowledge to the system by combining with the dark channel prior (DCP) algorithm[3]. DCP[3] estimates the transmission map of an underwater image, which encodes the depth information of distances between the camera and the objects in the scene. This information is more useful for underwater images, because objects become unclear more quickly with respect to distance underwater than on land. The transmission map information is added to the usual CycleGAN[10] inputs and an additional loss function is introduced to enforce depth information.

FUnIE-GAN[1][4] is another approach inspired by CycleGAN[10]. The generator is a simplifed U-Net[9], which is another popular general image translation model, and the discriminator is taken from the Markovian Patch-GAN[5]. The Markovian Path-GAN[5] discriminator only discriminates on the basis of patch-level information, thus ensuring that generated images capture local texture and style images of the target domain. To learn from unpaired data, FUnIE-GAN[4] uses the cycle technique pioneered by CycleGAN[10]. FUnIE-GAN[4] also has a variation for learning more effectively from paired data, which features a normal (non-cyclical) GAN[2] training format and adds losses to enforce global style similarity and image content similarity. One interesting advantage of FUnIE-GAN[4] is that its architecture was intentionally designed to be lightweight, in order to provide real-time image enhancement for underwater robots.

3 Next steps

I have obtained the code and datasets for both CycleGAN[10] and FUnIE-GAN[4]. The unpaired data version of FUnIE-GAN[4] is only available as a TensorFlow[7] implementation, so I intend to implement a PyTorch[8] version. Time permitting, I would also like to implement MCycleGAN[6] by adding the DCP algorithm[3] to CycleGAN[10], but may instead choose to focus on implementing the novel approach.

The novel approach will fuse training with unpaired data and paired data, in principle offering the advantages of both. Paired data is easier for the model to learn from, while unpaired data is available in much greater quantities. As such, I will modify FUnIE-GAN[4] to alternate between training on paired and on unpaired data. I also hope to explore the best way to combine these two paradigms: taking turns on paired and unpaired data is simplest, but perhaps it is more effective to begin training primarily with paired data and gradually mix in unpaired training as the model becomes more advanced. Overall, the goal will be to use as little paired data as possible while still attaining acceptable performance.

[1] Name seemingly derived from "**fu**lly co**n**volutional conditional **i**mage **e**nhancement **GAN**", which is funie indeed.

4 Conclusion

Now that I have gathered resources and formed a more explicit plan, I should be able to make steady progress for the rest of the semester. There are about four weeks remaining, so my goal timeline is as follows:

- One week from now: Have CycleGAN[10] working
- Two weeks from now: Have paired data version of FUnIE-GAN[4] working
- Three weeks from now: Finish implementing unpaired data version of FUnIE-GAN[4]
- Four weeks from now: Implement novel approach of combining paired and unpaired FUnIE-GAN[4], make presentation and website

These dates are deadlines, so I plan to have finished each of these items before the given date. Exactly four weeks from today (5 April) is 3 May, which reserves three days to finish up the website and presentation. For experiments, I will save the checkpoints and experiment results from these models all trained on the same unpaired dataset. Hopefully, doing so can minimize the amount of experiments I need to run right at the end of the semester, but I will have checkpointed models I can use in the event that I do need more information.

References

[1] Saeed Anwar and Chongyi Li. "Diving deeper into underwater image enhancement: A survey". In: *Signal Processing: Image Communication* 89 (2020), p. 115978.

[2] Ian Goodfellow et al. "Generative adversarial nets". In: *Advances in neural information processing systems* 27 (2014).

[3] Kaiming He, Jian Sun, and Xiaoou Tang. "Single image haze removal using dark channel prior". In: *IEEE transactions on pattern analysis and machine intelligence* 33.12 (2010), pp. 2341–2353.

[4] Md Jahidul Islam, Youya Xia, and Junaed Sattar. *Fast Underwater Image Enhancement for Improved Visual Perception.* 2019. DOI: `10.48550/ARXIV.1903.09766`. URL: `https://arxiv.org/abs/1903.09766`.

[5] Phillip Isola et al. "Image-to-image translation with conditional adversarial networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2017, pp. 1125–1134.

[6] Jingyu Lu et al. "Multi-scale adversarial network for underwater image restoration". In: *Optics Laser Technology* 110 (2019). Special Issue: Optical Imaging for Extreme Environment, pp. 105–113. ISSN: 0030-3992. DOI: `https://doi.org/10.1016/j.optlastec.2018.05.048`. URL: `https://www.sciencedirect.com/science/article/pii/S003039921830690X`.

[7] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* Software available from tensorflow.org. 2015. URL: `https://www.tensorflow.org/`.

[8] Adam Paszke et al. "Pytorch: An imperative style, high-performance deep learning library". In: *Advances in neural information processing systems* 32 (2019).

[9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention.* Springer. 2015, pp. 234–241.

[10] Jun-Yan Zhu et al. *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*. 2017. DOI: 10.48550/ARXIV.1703.10593. URL: https://arxiv.org/abs/1703.10593.