2025 ASEE Northeast Section Conference, March 22, 2025, University of Bridgeport, Bridgeport, CT, USA.

Assessing the Impact of Data Augmentation on Underwater Image Super Resolution

Nusrat Zahan
School of Engineering and Computing
Fairfield University
Fairfield, CT
nusrat.zahan@student.fairfield.edu

Sidike Paheding
School of Engineering and Computing
Fairfield University
Fairfield, CT
spaheding@fairfield.edu

***Abstract*—Image super-resolution (SR) refers to increasing the resolution of a low-resolution image to its corresponding high-resolution version. In recent years, underwater images have become increasingly difficult to super-resolve due to various noise types, low contrast, uniform illumination, monotonous color, and complex underwater backgrounds. These challenges significantly affect the performance of SR models, making it difficult to extract accurate information. Data augmentation is a key strategy to address these issues, involving deliberate adjustments to a dataset to improve its diversity. Such adjustments include image rotation, flipping, s caling, b rightness, c ontrast, a nd saturation. Data augmentation plays a significant r ole, e specially i n deep learning applications with sparse training data, such as image classification a nd s uper-resolution, b y r educing o verfitting and enhancing model generalization. In this study, we evaluate the performance of various deep learning-based data augmentation techniques on underwater images, given that data augmentation has been widely used in deep learning applications to improve model generalizability and robustness. Using a variety of color and geometric augmentation approaches, the study aims to investigate the significance o f t hese s trategies i n t erms of image reconstruction quality and overall model training caliber. The results of this study demonstrate that data augmentation enhances the performance of super-resolution (SR) models in underwater conditions. The SR methods analyzed include Super-Resolution Convolutional Neural Network (SRCNN), Super-Resolution Deep Residual Multiplier (SRDRM), and Deep Self-Attention Enhanced Super-Resolution (DEEP SESR), evaluated on the UFO-120 and USR-248 datasets. Notably, random flip augmentation achieved the highest PSNR of 26.3814 dB on the UFO-120 dataset using SRDRM and 26.3403 dB on the USR-248 dataset using SRCNN. Additionally, random saturation contributed to the highest SSIM values, reaching 0.7456 with SRCNN on UFO-120 and 0.7570 with DEEP SESR on USR-248.**

***Keywords*—Image Super-Resolution, Data Augmentation, Deep Learning, PSNR, and SSIM.**

I. Introduction

Image super-resolution (SR) is a key challenge in computer vision that reconstructs high-resolution (HR) images from their equivalent low-resolution (LR) images [1].

Although SR has made great strides in recent years with the emergence of deep learning, its application to underwater images is still a challenging problem. Underwater scenario imaging is exposed to degradation like noise, uneven illumination, less color variety, complex backgrounds, and low contrast, which could affect downstream tasks such as object detection [2]. Such conditions impede the ability of the SR models to precisely reconstruct finer scale details, which, in turn, not only degrades the reaching quality of the reconstructive images but also the extraction of valuable data, as discussed by Islam et al. [3], indicating how much SR performance underwater can be impacted by these distortions.

Data augmentation (DA) has been proven as a successful solution to these challenges. In other words, data augmentation augments the robustness and potential of deep learning models, which in turn increases the diversity of the training dataset and converts it artificially using transformations like image rotation, image flipping, image scaling, brightness, contrast, and saturation changes. This approach has been especially advantageous in such contexts with little or sparse training data, as noted in the study of Shorten et al. [4], which trains on the importance of the solution in mitigating overfitting for several tasks in vision. The influence of data augmentation on underwater image SR needs to be examined directly due to the unique distortions in underwater environments. For instance, the spatial relationship is important for SR tasks, which suggests that augmentation strategies should be designed more carefully, which is even more important in the underwater environment, according to Yoo et al. [5].

A number of recent approaches have been proposed to address this problem by gathering real-world datasets [6]–[8]. In many cases, acquiring such data on a large scale is often time-consuming and expensive. However, that is where DA can play an important role, but only a few studies have been conducted [9], [10]. Radu et al. [10] investigated several methods for performance enhancement for example-based single-image super-resolution (SISR), where one of them was data augmentation. They found consistent improvements across models and datasets, using rotation and flipping. However, they only evaluated geometric distortions using traditional SR models [11], [12]. Feng et al. [9] tested a method called (Mixup [13]) on the example-based SISR problem. However, the authors were only able to provide a general observation employing a single U-Net-like architecture and evaluated the method with the NTIRE2019 Real SR [7] dataset. In contrast, our work is further distinguished through specific underwater datasets (UFO-120 and USR-248), illustrating the various degradation characteristics of aquatic environments, and by evaluating different SR models. Unlike the narrow focus on geometric augmentations in Radu et al. [10] or Feng et al. [9] that functioned on the single model, we present color augmentations (brightness, contrast, saturation) along with a range of geometric (mainly flipping) methods aimed at tackling the complex distortion behavior of underwater images to provide a relatively wide picture of DA impacts on the performance of SR.

In this study, we evaluate the impact of data augmentation on underwater image SR by assessing three different deep learning based models, such as Super-Resolution Convolutional Neural Network (SRCNN) [14], Super-Resolution Deep Residual Multiplier (SRDRM) [3], and Deep Self-Attention Enhanced Super-Resolution (DEEP SESR) [15]. Dong et al. [14] introduced SRCNN, a pioneering model leveraging convolutional neural networks to map LR to HR images, achieving superior performance over traditional methods by learning end-to-end feature extraction. However, its simplicity limits its ability to handle severe underwater degradation. Extending residual learning by Islam et al. [3], the SRDRM introduced deep residual multipliers for the enhancement of feature recovery and showed a better performance for the reconstruction of underwater images with the USR-248 dataset but was still sensitive to noise and variance in illumination. More recently, Islam et al. [15] proposed DEEP SESR, which incorporates self-attention mechanisms to attend to critical areas in the image, significantly enhancing detail recovery in complex scenarios, though less investigated in underwater settings.

We evaluate the effect of color-based and geometric augmentation techniques on underwater SR tasks with an emphasis on reconstruction quality and training robustness. Our results show that data augmentation is able to improve underwater image super-resolution. The results emphasize that augmentation can be a pragmatic preprocessing step to leverage existing domain-independent SR models to make them tuned for an underwater environment, opening new avenues for the design of next-generation methods in this domain.

II. Methodology

This section focuses on the systematic methodology used to analyze how data augmentation influences SR of underwater images. We then describe the data enhancement techniques used to increase the diversity and robustness of the data sets, the models used, and the performance metrics used to quantitatively capture the quality of the reconstructions and the effectiveness of the models.

A. Data Augmentation Analysis

Our experimental dataset comes from underwater capturing scenarios that have common problems like noise, low contrast, and uneven illumination due to the wide variety of underwater capturing environments, so data augmentation is an important preprocessing step to enrich our dataset diversity. We will detail our application specifics for each of these augmentation presets (brightness, contrast, saturation, random flip, and combining all) to boost superficial performance in SRCNN, SRDRM, and DEEP SESR.

1) Augmentation Techniques: We apply four different data augmentation techniques to mimic real-world variations in underwater images while keeping their meaningful content intact. These methods include:

- Brightness Adjustment: In order to replicate the varied lighting conditions typical of underwater scenes, brightness modulation modifies the intensity of pixel values in the image. Let $I(x, y)$ represent the original pixel intensity at coordinates (x, y), and let α be the brightness scaling factor, where $\alpha \in [-1, 1]$. The augmented pixel intensity $I_{\text{bright}}(x, y)$ is computed as [4]

$$I_{\text{bright}}(x, y) = I(x, y) \cdot (1 + \alpha) \tag{1}$$

  where α is randomly sampled from a uniform distribution as $\alpha \sim U(-0.2, 0.2)$. This approach ensures that the brightness variations remain within realistic bounds, effectively simulating natural illumination changes.
- Contrast Enhancement: Contrast adjustment enhances the difference between pixel intensities, helping to counteract the low contrast commonly found in underwater images. The adjusted pixel intensity $I_{\text{contrast}}(x, y)$ is defined as [16]

$$I_{\text{contrast}}(x, y) = I(x, y) \cdot \beta + (1 - \beta) \cdot \mu \tag{2}$$

  where μ represents the mean intensity of the image. β is the contrast scaling factor, randomly sampled from the uniform distribution expressed as $\beta \sim U(0.8, 1.2)$. This ensures a balance between enhancement and detail preservation, preventing excessive contrast distortions.
- Saturation Modification: Saturation adjustment modifies the vibrancy of colors in an image, addressing the often dull and monotonous color palettes found in underwater environments. Given the saturation channel S and a saturation scaling factor γ, the adjusted saturation S_{sat} is computed as [16]

$$S_{\text{sat}} = S \cdot \gamma \tag{3}$$

  where γ is randomly sampled from the uniform distribution as $\gamma \sim U(0.8, 1.2)$. This ensures that saturation adjustments remain within realistic bounds, avoiding excessive color distortion while still enhancing visual clarity.
- Random Flipping: Random horizontal or vertical flipping introduces geometric diversity, helping to simulate different underwater perspectives. This augmentation enhances model robustness by allowing it to learn from varied orientations of underwater objects. Given an image I of

size $W \times H$, the flipping operations are defined as follows [4]

$$H_{I_{\text{flip}}(x,y)=I(W-x,y)} \tag{4}$$

$$V_{I_{\text{flip}}(x,y)=I(x,H-y)} \tag{5}$$

Here, $H_{I_{\text{flip}}}$ is Horizontal flip and $V_{I_{\text{flip}}}$ is Vertical flip. Each flip is applied with a 50% probability, ensuring that the augmentation introduces sufficient variation without over-distorting the training data.

- Combination of Techniques: To maximize dataset diversity, we employ a combined augmentation strategy that sequentially or concurrently applies brightness, contrast, saturation, and random flipping. The overall transformation, denoted as $I_{\text{combined}}(x, y)$, is expressed as

$$I_{\text{combined}}(x, y) = T_f(T_s(T_c(T_b(I(x, y))))) \tag{6}$$

where T_b, T_c, T_s, T_f represent the brightness, contrast, saturation, and flip transformations, respectively.

B. Model Analysis

This study examines how data augmentation affects underwater image super-resolution (SR) on three popular models, namely SCRNN, SRDRM, and DEEP SESR. SRCNN is highly effective due to its simple and lightweight architecture, making it an ideal choice for IoT-based devices. Its efficiency enables rapid SR processing, particularly beneficial for marine applications. SRDRM, a deep residual network-based generative model, is specifically designed for underwater image super-resolution, making it well-suited for autonomous underwater robots. Unlike SRCNN, SRDRM employs adversarial training and its objective function evaluates global similarity, perceptual loss, and image content loss, enhancing the quality of generated images. DEEP SESR is a residual-in-residual network-based generative model designed for underwater robotic vision, offering an efficient solution for near real-time applications. It integrates residual dense blocks, a feature extraction network, and an auxiliary attention network to produce high-resolution images with enhanced perceptual quality.

C. Experiment and Assessment

The proposed approach is evaluated on three state-of-the-art (SOTA) SR models, namely SRCNN, SRDRM, and DEEP SESR. Our models are trained and tested over two publicly accessible underwater image datasets, UFO-120 and USR-248. Islam et al. [15] provided a dataset named UFO-120 intended for Simultaneous Enhancement and Super-Resolution (SESR) tasks, which comprises 120 synthetic test images and 1,500 synthetic training images. In contrast, the USR-248 dataset is designed for SISR tasks and consists of 560 low-resolution images of multiple scales (80 × 60, 160 × 120, and 320 × 240) along with 1,060 paired images for training and 248 reference images for an extensive evaluation of the model, as indicated by Islam et al. [3]. In this paper, we focus only on 4X upsampling, where we upscale the low-resolution USR-248 dataset (160 × 120) to 640 × 480 along with the UFO-120 dataset, allowing us to directly visualize the comparative performance of all three algorithms. The models are trained separately on the training sets of UFO-120 (1,500 images) and USR-248 (1,060 paired images) for 130 epochs using the Stochastic Gradient Descent (SGD) optimizer with a learning rate of 0.0001, and tested on the respective testing sets, which consist of 120 images from UFO-120 and 248 reference images from USR-248. For quantitative evaluation of the upscaling performance, we follow the standard with two full-reference metrics, Peak Signal-to-Noise Ratio (PSNR) [17] and Structural Similarity Index (SSIM) [18]. Standard practices for SR evaluation compute these metrics to determine the fidelity between the super-resolved images and their high-resolution ground truth, where larger values imply better reconstruction quality.

III. Results And Discussion

A. Quantitative Analysis

In the analysis of the UFO 120 and USR 248 datasets, data augmentation is found to be effective in enhancing PSNR and SSIM over models.

TABLE I
QUANTITATIVE EVALUATION OF DIFFERENT DATA AUGMENTATION TECHNIQUES AT ×4 UPSCALING ON THE UFO-120 DATASET. THE BEST VALUES ARE MARKED IN BOLD.

Augmentation Type	SRCNN		DEEP SESR		SRDRM	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Without Data Augmentation	26.3444	0.7432	25.8940	0.7318	25.7659	0.7345
Random brightness	26.3730	0.7429	**26.1144**	0.7453	26.1361	0.7363
Random contrast	**26.3811**	0.7446	25.3001	0.7434	25.7867	0.7192
Random saturation	26.3627	**0.7456**	24.9079	0.7431	26.2387	0.7380
Random Flip	26.3746	0.7444	25.9431	0.7412	**26.3814**	**0.7405**
Combining ALL	26.3644	0.7440	25.7752	**0.7456**	26.1774	0.7374

TABLE II
QUANTITATIVE EVALUATION OF DIFFERENT DATA AUGMENTATION TECHNIQUES AT ×4 UPSCALING ON THE USR-248 DATASET. THE BEST VALUES ARE MARKED IN BOLD.

Augmentation Type	SRCNN		DEEP SESR		SRDRM	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Without Data Augmentation	26.2673	0.7411	24.6241	0.7391	22.6721	0.7108
Random brightness	26.3386	0.7432	25.8136	0.7469	25.7462	0.7224
Random contrast	26.3031	0.7411	26.0062	0.7449	25.1678	0.7255
Random saturation	26.2996	0.7421	**26.2996**	**0.7570**	26.2746	0.7410
Random Flip	**26.3403**	**0.7439**	26.1415	0.7488	**26.2962**	**0.7419**
Combining ALL	26.3194	0.7423	26.2082	0.7503	26.1341	0.7323

We employ random contrast in SRCNN for the UFO 120 and achieve a top PSNR of 26.3811 dB, and for random saturation, we achieve a top SSIM value of 0.7456 dB. In Table I, it is clearly shown that DEEP SESR obtains the highest PSNR of 26.1144 dB by applying random brightness, while using all augmentation simultaneously achieves the best SSIM performance of 0.7456 dB. On the other hand, SRDRM obtains the highest performance from Random Flip, leading to the highest value of PSNR 26.3814 dB and SSIM 0.7405 dB.

For USR 248, Random Flip greatly favors SRCNN with the best PSNR of 26.3403 dB and SSIM of 0.7439 dB, thus

Fig. 1. Visual Comparisons for $\times 4$ upsampling on Underwater Image sampled from UFO-120 dataset.

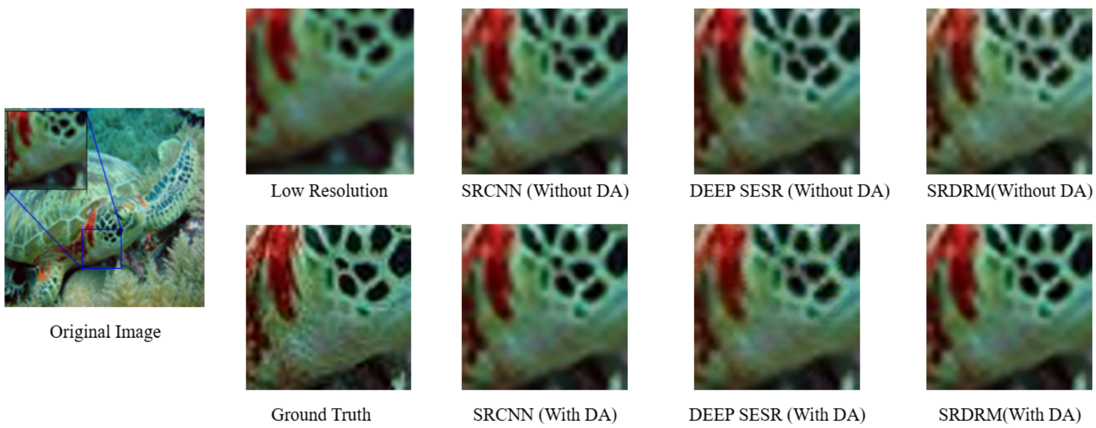

Fig. 2. Visual Comparisons for $\times 4$ upsampling on Underwater Image sampled from USR-248 dataset.

TABLE II showing that this data augmentation technique helps in the case of this model. Among all the transformations, random saturation works best in DEEP SESR as it obtains the highest PSNR of 26.2996 and the SSIM of 0.7570 dB, demonstrating that color changes contribute to feature extraction based on our experiments. In terms of SRDRM, Random Flip yields the greatest PSNR of 26.2962 dB, whereas most improvements in SSIM of 0.7255 dB instead come from using random contrast (Buffer). In conclusion, Random Flip and Random Saturation seem to be the most useful techniques, increasing accuracy by a great margin for both SRCNN and SRDRM, whereas random saturation additionally enhances DEEP SESR performance. After consistent data augmentation, the PSNR and SSIM values are improved in contrast to basic training without data augmentation.

B. *Qualitative Analysis*

Figures 1 and 2 show a qualitative comparison of superresolution results in the UFO-120 and USR-248 datasets. Links to each dataset are also accompanied by an original and low-resolution image, its corresponding ground truth high-resolution image, and the outputs from each of the SRCNN, DEEP SESR, and SRDRM models without and with data augmentation. Significantly, based on DA, the super-resolved images have better-defined edges, finer texture features, and more uniform colors compared to those of the outputs without DA, in which the blurriness and distortions are more dominant. Among the models, SRDRM and SRCNN (with DA) produce better visually pleasing results, particularly in retaining intricate textures and reducing artifacts and color contrast. The most notable improvement across all datasets is USR-248, with which complex texture details such as turtle shells are sharply defined and structured where augmentation-free training is not. This qualitative comparison indicates that data augmentation is crucial to improving performance, preserving details, and increasing perceived quality.

IV. Conclusion

In this work, we study the effect of data augmentation on underwater image SR on three state-of-the-art models, namely SRCNN, SRDRM, and DEEP SESR. Our approach was to expand the diversity of the data set employing image augmentation techniques that involved adjustments in brightness, contrast, and saturation, along with random flipping and multiple combinations of these methods. We demonstrate both qualitatively and quantitatively that data augmentation improves super-resolution performance to some extent. Our augmentation methods have proven successful at improving reconstruction quality and providing robustness for underwater image models. We plan to investigate more sophisticated augmentation techniques, such as generative adversarial networks [19] and adaptive transformations [20], with the goal of applying these methods to underwater super-resolution systems across various aquatic environments.

References

[1] Y. Wang, W. Yang, X. Chen, *et al.*, "Sinsr: Diffusion-based image super-resolution in a single step," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 25 796–25 805.

[2] A. Awad, N. Zahan, E. Lucas, T. C. Havens, S. Paheding, and A. Saleem, "Underwater simultaneous enhancement and super-resolution impact evaluation on object detection," in *Pattern Recognition and Tracking XXXV*, SPIE, vol. 13040, 2024, pp. 67–77.

[3] M. J. Islam, S. S. Enan, P. Luo, and J. Sattar, "Underwater image super-resolution using deep residual multipliers," in *2020 IEEE international conference on robotics and automation (ICRA)*, IEEE, 2020, pp. 900–906.

[4] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.

[5] J. Yoo, N. Ahn, and K.-A. Sohn, "Rethinking data augmentation for image super-resolution: A comprehensive analysis and a new strategy," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 8375–8384.

[6] A. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1692–1700.

[7] J. Cai, H. Zeng, H. Yong, Z. Cao, and L. Zhang, "Toward real-world single image super-resolution: A new benchmark and a new model," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 3086–3095.

[8] X. Zhang, R. Ng, and Q. Chen, "Single image reflection separation with perceptual losses," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4786–4794.

[9] R. Feng, J. Gu, Y. Qiao, and C. Dong, "Suppressing model overfitting for image super-resolution networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.

[10] R. Timofte, R. Rothe, and L. Van Gool, "Seven ways to improve example-based single image super resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1865–1873.

[11] S. Gu, W. Zuo, Q. Xie, D. Meng, X. Feng, and L. Zhang, "Convolutional sparse coding for image super-resolution," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1823–1831.

[12] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Computer Vision–ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part IV 12*, Springer, 2015, pp. 111–126.

[13] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," *arXiv preprint arXiv:1710.09412*, 2017.

[14] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2015.

[15] M. J. Islam, P. Luo, and J. Sattar, "Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception," *arXiv preprint arXiv:2002.01155*, 2020.

[16] S. Raveendran, M. D. Patil, and G. K. Birajdar, "Underwater image enhancement: A comprehensive review, recent trends, challenges and applications," *Artificial Intelligence Review*, vol. 54, pp. 5413–5467, 2021.

[17] J. Korhonen and J. You, "Peak signal-to-noise ratio revisited: Is simple beautiful?" In *2012 Fourth international workshop on quality of multimedia experience*, IEEE, 2012, pp. 37–38.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[19] A. Aghelan and M. Rouhani, "Underwater image super-resolution using generative adversarial network-based model," in *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, IEEE, 2023, pp. 480–484.

[20] P. Ajanya and S. Meera, "An effectual underwater image enhancement framework using adaptive trans-resunet++ with attention mechanism (aug, 10.1080/0952813x.2024.2383659, 2024)," *JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE*, 2024.