Real-Time GAN-Based Model for Underwater Image Enhancement

Danilo Avola[1(✉)], Irene Cannistraci[1], Marco Cascio[1,2], Luigi Cinque[1], Anxhelo Diko[1], Damiano Distante[2], Gian Luca Foresti[3], Alessio Mecca[1], and Ivan Scagnetto[3]

[1] Department of Computer Science, Sapienza University of Rome, Via Salaria 113, 00198 Rome, Italy
{avola,cannistraci,cascio,cinque,diko,mecca}@di.uniroma1.it

[2] Department of Law and Economics, University of Rome UnitelmaSapienza, Piazza Sassari 4, 00161 Rome, Italy
{marco.cascio,damiano.distante}@unitelmasapienza.it

[3] Department of Mathematics, Computer Science and Physics, University of Udine, Via delle Scienze 206, 33100 Udine, Italy
{gianluca.foresti,ivan.scagnetto}@uniud.it

Abstract. Enhancing image quality is crucial for achieving an accurate and reliable image analysis in vision-based automated tasks. Underwater imaging encounters several challenges that can negatively impact image quality, including limited visibility, color distortion, contrast sensitivity issues, and blurriness. Among these, depending on how the water filters out the different light colors at different depths, the color distortion results in a loss of color information and a blue or green tint to the overall image, making it difficult to identify different underwater organisms or structures accurately. Improved underwater image quality can be crucial in marine biology, oceanography, and oceanic exploration. Therefore, this paper proposes a novel Generative Adversarial Network (GAN) architecture for underwater image enhancement, restoring good perceptual quality to obtain a more precise and detailed image. The effectiveness of the proposed method is evaluated on the EUVP dataset, which comprises underwater image samples of various visibility conditions, achieving remarkable results. Moreover, the trained network is run on the RPi4B as an embedded system to measure the time required to enhance the images with limited computational resources, simulating a practical underwater investigation setting. The outcome demonstrates the presented method applicability in real-world underwater exploration scenarios.

Keywords: Underwater image enhancement · GAN · Underwater exploration

1 Introduction

In the last few years, visual information analyzing tools have become increasingly attractive for solving heterogeneous perception-based tasks such as video surveil-

G. L. Foresti et al. (Eds.): ICIAP 2023, LNCS 14233, pp. 412–423, 2023.
https://doi.org/10.1007/978-3-031-43148-7_35

lance [2,8,14,31], biomedical imaging [6,9,32], environmental modeling [1,19], gesture recognition [5,11], or human behavioral analysis [3,4,7,13,27]. Among the others, underwater imaging helps to develop new technologies and techniques to discover underwater environments. However, this process is still challenging due to limited visibility and light-related phenomena, i.e., absorption, scattering, and refraction [17]. Indeed, underwater image processing is characterized by degradation issues, including loss of color information and other optical artifacts in captured images leading to decreasing performance of visual-related tasks, e.g., segmentation, detection, or classification. Despite the promising results in the literature on terrestrial imagery, there are still plenty of chances to enhance the perceptual quality of underwater images. Therefore, quality enhancement and detailed restoration strategies are crucial for better understanding the underwater environment, identifying new species, and monitoring marine ecosystems. To this end, the literature proposes several physics- and learning-based strategies [10,12,15,28]. The former is appropriate for color correction and dehazing; however, other than being computationally expensive, it requires scene depth and assessing water quality based on the interaction of light with water, which is not always available in automated applications. Instead, the latter is appropriate to enhance the overall perceptual image quality from large data collection. This paper presents an underwater image enhancement strategy based on a novel GAN-based architecture to improve the visual perception of the given poor-quality image. Motivated by the good results in [18], the proposed model leverages U-Net [26] and PatchGAN [18] networks for learning the mapping between poor- and good-quality underwater images. The effectiveness of the proposed network architecture is evaluated on the Enhancing Underwater Visual Perception (EUVP) [18] dataset, an underwater image collection suitable for adversarial model training. Finally, to simulate and test the feasibility of the presented method in real-world underwater exploration scenarios, usually characterized by inspection robots with limited computational resources, the trained model is run on an embedded system suitable for the industrial sector. This simulation measures the time required to enhance the perceptual quality of underwater images, which is crucial for practical applications. In summary, the main contributions of this paper are as follows:

- Designing an innovative GAN model for underwater image quality enhancement, suitable for oceanic explorations and human-robot collaborative experiments;
- Achieving of State-Of-the-Art (SOTA) performance on the EUVP benchmark, a real-world underwater image dataset containing paired collections of poor- and good-quality images for supervised learning;
- Performing tests on an embedded system to measure the time required for image quality enhancement, simulating a generic and vision-based practical underwater investigation setting with limited computational resources.

2 Related Work

Underwater image enhancement is gaining ever-increasing interest in literature thanks to the wide range of applications in developing, exploring, and protecting ocean resources. According to the literature, proposed solutions can be classified into physics- and learning-based approaches.

2.1 Physics-Based Approaches

Traditional physics-based strategies comprise the definition of formal models, including mathematical or simplified underwater image formation models. In fact, [25,29] proposed methods relying on the Beer-Lambert law to recover better details of underwater images considering the light attenuation related to the material properties through which the light travels. Instead, the authors in [22] manipulate poor-quality image color channels to dehaze an image that is later combined with an enhanced version, obtained through a color correction algorithm, to achieve the final good-quality result. Also, [24] presents an underwater image dehazing method manipulating the RGB color space. First, the authors estimate the background light using the quad-tree subdivision iteration algorithm. Afterward, the RGB color space dimensionality prior is compressed to the UV color space by clustering the pixels into a hundred haze-lines and setting a haze-free boundary to compute the dehazed version of the image accurately. In [10], the authors propose a method considering the water type in the enhancement process. Given the low-quality image, they first estimate the veiling-light, and the color restoration strategy is performed for multiple water types having different properties. Finally, the best enhanced version is selected automatically based on the gray-world assumption. However, previous methods require prior knowledge that may not always provide a reliable solution relying on local and global color distribution. Therefore, [12] proposes an approach for terrestrial and underwater image quality restoration without any prior information. A multiband decomposition solution extracts the base and detail layers for intensity and Laplacian modules involved in restoring the image.

2.2 Learning-Based Approaches

Most recently, Deep Learning (DL) techniques have been applied to underwater image enhancement thanks to their capability of automatically learning the mappings between two domains from large data collection. In literature, the GAN model proved effective in improving the perceptual quality of underwater images. In [28], the authors propose the Class-conditional Attention GAN (CA-GAN) for underwater image enhancement in which the class label guides the generation of the good-quality image version. Differently, in [15], a CycleGAN-based [33] architecture is used to create a paired dataset of underwater poor- and good-quality images through the style transfer property; therefore, a fully convolutional encoder-decoder is trained for underwater image enhancement on such synthesized data. Instead, in [20], the CycleGAN model is used as the backbone for a network architecture learning the cross-domain mapping between underwater and terrestrial images. Also, in [23], the authors exploit a different GAN model that uses terrestrial images and depth maps to learn the color correction of monocular underwater images in an unsupervised fashion. In [30] is proposed a stacked conditional GAN consisting of a haze detection sub-network and color correction sub-network. In detail, the former produces a hazing detection mask from the underwater image given as input, while the latter corrects the image color by exploiting the previously predicted mask. Instead, [16] proposes a multiscale dense GAN combining residual learning, dense concatenation, and multiscale operation to correct color casts and restore image details. In [18], the authors propose a fully-convolutional conditional GAN-based model with a multi-modal objective function to evaluate the perceptual quality of the given underwater image considering global content, color, local texture, and style information. Finally, in [21], multiple inputs and a GAN architecture are combined to solve color casts, low contrast, and haze-like issues. Specifically, the model generator component comprises main and auxiliary sub-networks. Initially, the main module extracts the features from raw underwater images, whereas the auxiliary component extracts the features from the fusion-based enhanced version obtained through SOTA methods. Afterward, the two sub-networks outputs are merged to decode the restored image.

3 Method

To enhance underwater images by increasing their perceptual quality, the GAN-based architecture depicted in Fig. 1 was designed to find the non-linear mapping between poor- and good-quality underwater images. The proposed model expands the traditional GAN network by leveraging a Convolutional Neural Network (CNN) architecture to handle visual data by extracting the low-dimensional feature representation of underwater images used to improve the input images visually. To this end, the network training follows a supervised fashion, allowing to map the underwater poor-quality image subspace to the ground truth good-quality image subspace. With such a supervised training paradigm, the underwater image structural information and scene details are maintained while the network learns to generate its restored version.

3.1 Underwater Image Enhancement

Recently, learning-based approaches have shown to be effective for underwater image enhancement achieving promising results. The GAN-based models can successfully approximate a mapping function of a given input data distribution to a target data distribution by generating fake samples as if they were drawn from the target distribution itself. Indeed, the proposed network relies on this property using a CNN-based GAN architecture comprising generator (G) and discriminator (D) components, where the former produces fake images

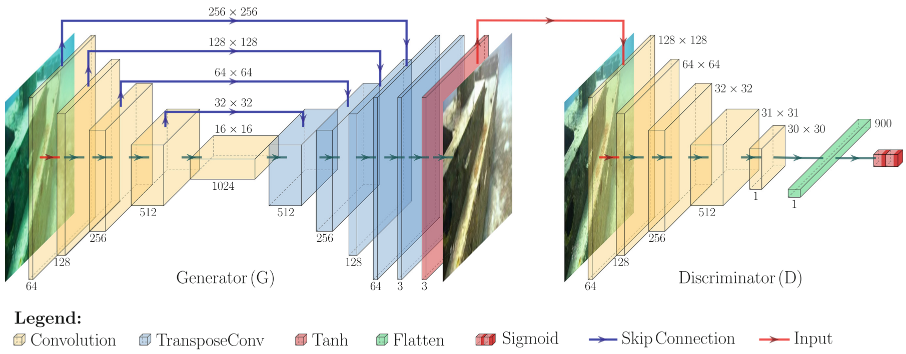

Fig. 1. The proposed GAN-based architecture. Given the poor-quality underwater image as input, the network is trained to restore good perceptual quality, obtaining a more precise and detailed image.

trying to learn the target distribution of the enhanced images while the latter discriminates against all real and fake enhanced images. Specifically, motivated by the success of skip connections in generative adversarial models for image-to-image mapping and quality enhancement, G is a customized version of the fully-convolutional U-Net architecture. In detail, the encoder part consists of five convolutional layers with 3×3 filter size and a stride of 2, each followed by batch normalization and Leaky Rectified Linear Unit (LeakyReLU) activation function. Alongside, the decoder component follows a reverse structure of the encoder with transposed convolutional layers rather than convolutions and Rectified Linear Unit (ReLU) instead of LeakyReLU activation functions to stabilize the training process. Finally, after the last transposed convolution, the decoder uses the hyperbolic tangent function to reconstruct the restored images. Precisely, given as input the RGB underwater image, with size $256 \times 256 \times 3$, the encoding part of the network learns feature maps of shape $16 \times 16 \times 1024$, representing its low-dimensional feature representation. Afterward, the decoder utilizes these latent features to generate an enhanced version of the image with the final shape $256 \times 256 \times 3$. In addition, to enable G for accurately reproducing features such as local texture and style, the discriminator is a PatchGAN-based network discriminating on patch-level information. Specifically, D comprises five padded convolutional layers with 4×4 filters and a stride of 2, except for the last layer set to 1, followed by batch normalization and LeakyReLU activation. Given the RGB image as input, the discriminator extracts feature maps of size 30×30 on which a linear layer and the sigmoid function are applied for discriminating between real or fake good-quality images. Following this adversarial training strategy, the generator G learns the significant low-dimensional features to maintain from underwater data; indeed, when D classifies both original and generated enhanced images as reals, it implies that the low-dimensional representation is very informative to the point of fooling the discriminator with the restored image. Formally, the underwater-to-enhance image mapping leverages the min-max game between G and D models, defined as:

$$L_{GAN}(D, G) = \min_{G} \max_{D}(E_Y[logD(Y)] + E_{X,Y}[log(1 - (D(G(X))))]), \quad (1)$$

where X represents the underwater images as input, Y the ground truth good-quality images, and $G(X)$ the restored image from the input, $D(Y)$ and $D(G(X))$ indicate the estimated probabilities of given good-quality and enhanced images being real. Since the L_{GAN} is designed for learning to approximate the target distribution, it does not enforce the generator G to maintain important aspects like global content, color, style, and local structures. To this end, in order for G to generate an enhanced version of the underwater input image that is consistent with the corresponding good-quality ground truth, the Mean Absolute Error (MAE) and Mean Standard Error (MSE) between the restored and original ground truth images are defined as:

$$MAE_G = \frac{1}{(W \times H)} \sum_{i=1}^{W} \sum_{j=1}^{H} [X(i,j) - Y(i,j)], \quad (2)$$

$$MSE_G = \frac{1}{(W \times H)} \sum_{i=1}^{W} \sum_{j=1}^{H} [X(i,j) - Y(i,j)]^2, \quad (3)$$

where $X(i, j)$ and $Y(i, j)$ indicate a pixel in the enhanced and original good-quality images, respectively. Thus, the training objective of the overall network is to minimize the loss function, defined as follows:

$$L(D, G) = w_{GAN} L_{GAN}(D, G) + w_{MAE} MAE_G + w_{MSE} MSE_G. \quad (4)$$

where w_{GAN}, w_{MAE}, and w_{MSE} are weighting parameters adjusting the impact of individual losses.

4 Experiments

This section presents a comprehensive evaluation of the proposed GAN-based model on the EUVP dataset for the underwater image enhancement task. Specifically, it describes significant implementation details and quantitative and qualitative evaluations. Moreover, it also reports the image enhancement time of the trained network running on the Raspberry Pi 4 model B (RPi4B) as an embedded system, simulating a generic practical underwater investigation setting with limited computational resources.

4.1 EUVP Dataset

The EUVP dataset [18] is the publicly available benchmark focused on underwater image enhancement that contains poor and good perceptual quality underwater image samples suitable for models trained in a supervised fashion. Specifically, it contains 8670 images with a size of 256×256 comprising 3700 training

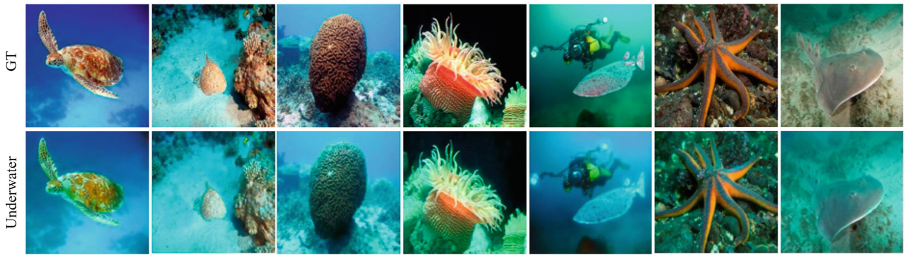

Fig. 2. EUVP dataset paired image samples. In the first row GT good-quality images; in the second row the corresponding underwater poor-quality images.

pairs and 1270 validation samples. Data augmentation is applied to training data, including horizontal flip, vertical flip, and random crop, increasing to 10882 the training pairs. For paired data, the ground truth good-quality images were generated by using the method proposed in [15]. Figure 2 illustrates some paired training image samples.

4.2 Implementation Details

The proposed architecture design has been developed using the PyTorch framework. The experiments were performed on two GPUs, i.e., $\times 2$ NVIDIA GeForce RTX 2080 Ti with 11 GB of RAM. In detail, the network was trained following a supervised setting for 50 epochs by exploiting the original EUVP data split and training protocol. For the model training, Adam was used as the optimizer with a learning rate set to 0.0002, an ϵ parameter of 1e−8, a weight decay set to 1e−2, a first β_1 and second β_2 momentum initial decay rate of 0.5 and 0.999, respectively. Finally, the weight parameters adjusting the loss functions within Eq. (4) were set to $w_{GAN} = w_{MSE} = 1$ and $w_{MAE} = 10$. For the proposed method evaluation, the standard metrics Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) were used to compare the enhanced images with the corresponding ground truth quantitatively. Precisely, the PSNR estimates the reconstruction quality of the restored image, while the SSIM compares the two images considering three properties: luminance, contrast, and structure.

4.3 Underwater Image Enhancement Evaluation

Regarding the underwater image enhancement, the quantitative results are reported measuring the PSNR and the SSIM metrics between the proposed model enhanced image version and the available respective ground truth. With the aim to approximate the image reconstruction accuracy of the restored image, the PSNR is defined as follows:

$$PSNR(X, Y) = 10 \log_{10}[255^2 / MSE(X, Y)], \tag{5}$$

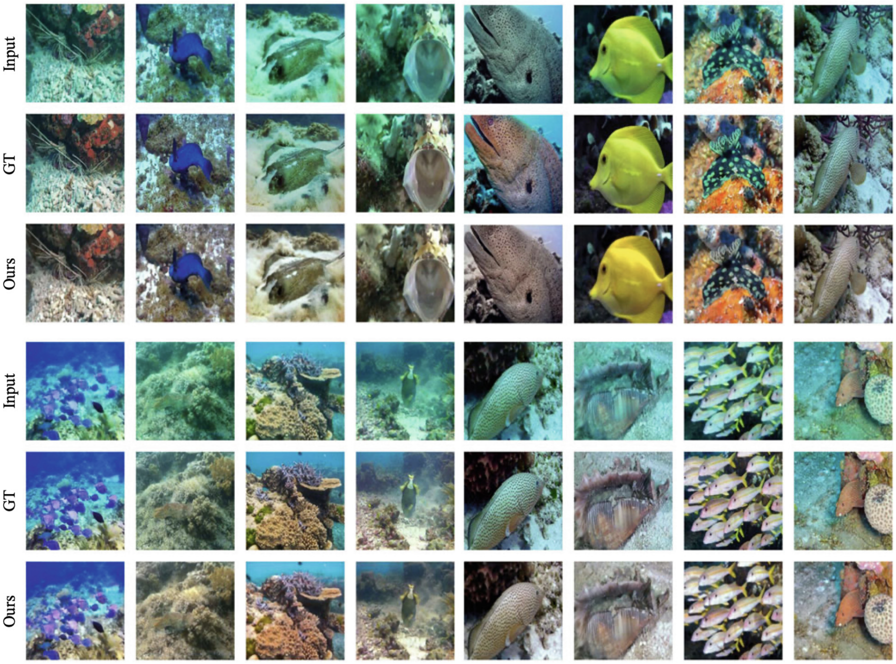

Fig. 3. Qualitative results for underwater image enhancement. In the first and fourth rows, the input underwater images; in the second and fifth rows, the corresponding good-quality GT images; in the third and sixth rows, the proposed network enhanced images.

4.4 Embedded System Evaluation for Practical Underwater Exploration Simulation

Concluding the proposed method evaluation, the trained model is run on commodity hardware with limited computational resources to simulate a practical underwater investigation setting. A generic vision-based oceanic exploration process is characterized by a visual sensor mounted on the automated underwater vehicle capturing poor-quality images. Therefore, an embedded system (e.g., RPi4B) can be used to enhance the captured image quality for underwater image processing tasks in real-time, e.g., during inspections of pipelines, dams, and offshore platforms, to enable safer and more efficient maintenance and repair operations in the industrial sector. In many applications, time plays a crucial role. The proposed neural model is tested on an RPi4B, a system based on a Quad-core Cortex-A72 (ARM v8) 64-bit SoC @1.5 GHz processor with 2 GB of RAM board, to simulate the described investigation setting. Since we use trained weights for the model, image enhancement performance does not change with respect to the reported in Sect. 4.3. Therefore, the critical factor is examining the time required to obtain a good-quality image. In detail, given a 256×256 image as input, the prediction time for a single image enhancement is about 0.029 s,

Table 1. Performance evaluation on the EUVP dataset and comparison with key literature physics- and learning-based methods.

Model	Type	$PSNR \uparrow$	$SSIM \uparrow$
Uw-HL [10]	Physics-based	18.85 ± 1.76	0.7722 ± 0.066
Mband-En [12]	Physics-based	12.11 ± 2.55	0.4565 ± 0.097
Res-GAN [18]	Learning-based	14.75 ± 2.22	0.4685 ± 0.122
Res-WGAN [18]	Learning-based	16.46 ± 1.80	0.5762 ± 0.014
LS-GAN [18]	Learning-based	17.83 ± 2.88	0.6725 ± 0.062
Pix2Pix [18]	Learning-based	20.27 ± 2.66	0.7081 ± 0.069
UGAN-P [15]	Learning-based	19.59 ± 2.54	0.6685 ± 0.075
CycleGAN [18]	Learning-based	17.14 ± 2.17	0.6400 ± 0.080
FUnIE-GAN-UP [18]	Learning-based	21.36 ± 2.17	0.8164 ± 0.046
FUnIE-GAN [18]	Learning-based	21.92 ± 1.07	0.8876 ± 0.068
Ours	**Learning-based**	**22.09 ± 1.02**	**0.9002 ± 0.059**

where X and Y are the enhanced and ground truth images, respectively. Instead, to measure the perceptual restored image quality, the SSIM is defined as:

$$SSIM(X,Y) = \left(\frac{2\mu_X\mu_Y + c_1}{\mu_X^2 + \mu_Y^2 + c_1}\right)\left(\frac{2\sigma_{XY} + c_2}{\sigma_X^2 + \sigma_Y^2 + c_2}\right) \tag{6}$$

where X and Y are always the enhanced and ground truth images being compared, μ_X and μ_Y are the pixel sample means of X and Y, respectively, σ_X and σ_Y are the standard deviations, σ_{XY} is the covariance of x and y; finally, c_1 and c_2 are constants used to stabilize the division when the denominator is close to zero. Table 1 reports the obtained results on test images and summarizes comparisons with key literature physics- and learning-based works. As can be observed, the learning-based methods generally perform significantly better than physics-based approaches. However, the proposed GAN-based network achieves remarkable and increased performances in reconstruction quality, i.e., PSNR, and restoring perceptual quality, i.e., SSIM. Regarding the qualitative evaluation, the obtained results are depicted in Fig. 3. As can be observed, the presented approach generates enhanced underwater images, restoring optical quality and successfully removing the typical blue or green tint in underwater images. Notice that, in some cases, the enhanced color in the restored images generated by the proposed method is closer to the true color rather than in the ground truths.

thus handling roughly 34FPS in real-time. This result is also confirmed by evaluating the entire test set. Therefore, the model can be executed on commodity hardware at a considerable amount of FPS, thus demonstrating its applicability in real-world scenarios.

5 Conclusions

This paper presents a novel GAN-based architecture for underwater image enhancement to obtain a precise and detailed scene, restoring good perceptual quality. By learning the low-dimensional feature representation of underwater images, the proposed network maintains the structural information and details while discovering the mapping between poor- and good-quality images following a supervised training fashion. More importantly, the proposed method achieved state-of-the-art performance on a real-world underwater image dataset containing paired collections of poor- and good-quality images and demonstrated real-time capabilities on commodity hardware for up to 34FPS.

Acknowledgements. This work was supported by "Smart unmannEd AeRial vehiCles for Human likE monitoRing (SEARCHER)" project of the Italian Ministry of Defence (CIG: Z84333EA0D); and "A Brain Computer Interface (BCI) based System for Transferring Human Emotions inside Unmanned Aerial Vehicles (UAVs)" Sapienza Research Projects (Protocol number: RM1221816C1CF63B); and "DRrone Aerial imaGe SegmentatiOn System (DRAGONS)" (CIG: Z71379B4EA); and Departmental Strategic Plan (DSP) of the University of Udine - Interdepartmental Project on Artificial Intelligence (2020–25); and "An Integrated Platform For Autonomous Agents For Maritime Situational Awareness (ARGOS)" project of the Italian Ministry of Defence (PNRM 2022); and the MICS (Made in Italy - Circular and Sustainable) Extended Partnership and received funding from Next-Generation EU (Italian PNRR - M4 C2, Invest 1.3 - D.D. 1551.11-10-2022, PE00000004). CUP MICS B53C22004130001.

References

1. Avola, D., Bernardi, M., Cinque, L., Foresti, G.L., Massaroni, C.: Adaptive bootstrapping management by keypoint clustering for background initialization. Pattern Recognit. Lett. **100**, 110–116 (2017). https://doi.org/10.1016/j.patrec.2017.10.029
2. Avola, D., et al.: A novel GAN-based anomaly detection and localization method for aerial video surveillance at low altitude. Remote Sens. **14**(16), 4110 (2022). https://doi.org/10.3390/rs14164110
3. Avola, D., Cascio, M., Cinque, L., Fagioli, A., Foresti, G.L.: Affective action and interaction recognition by multi-view representation learning from handcrafted low-level skeleton features. Int. J. Neural Syst. 2250040 (2022). https://doi.org/10.1142/s012906572250040x
4. Avola, D., Cinque, L., De Marsico, M., Fagioli, A., Foresti, G.L.: LieToMe: preliminary study on hand gestures for deception detection via fisher-LSTM. Pattern Recognit. Lett. **138**, 455–461 (2020). https://doi.org/10.1016/j.patrec.2020.08.014

5. Avola, D., Cinque, L., Fagioli, A., Foresti, G.L., Fragomeni, A., Pannone, D.: 3d hand pose and shape estimation from RGB images for keypoint-based hand gesture recognition. Pattern Recognit. **129**, 108762 (2022). https://doi.org/10.1016/j.patrec.2017.10.029
6. Avola, D., Cinque, L., Fagioli, A., Foresti, G., Mecca, A.: Ultrasound medical imaging techniques: a survey. ACM Comput. Surv. **54**(3), 1–38 (2021). https://doi.org/10.1145/3447243
7. Avola, D., Cinque, L., Foresti, G.L., Pannone, D.: Automatic deception detection in RGB videos using facial action units. In: International Conference on Distributed Smart Cameras, pp. 1–6 (2019). https://doi.org/10.1145/3349801.3349806
8. Avola, D., Foresti, G.L., Martinel, N., Micheloni, C., Pannone, D., Piciarelli, C.: Real-time incremental and geo-referenced mosaicking by small-scale UAVs. In: Battiato, S., Gallo, G., Schettini, R., Stanco, F. (eds.) ICIAP 2017. LNCS, vol. 10484, pp. 694–705. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68560-1_62
9. Avola, D., Petracca, A., Placidi, G.: Design of a framework for personalised 3d modelling from medical images. Comput. Methods Biomech. Biomed. Eng. Imaging Vis. **3**(2), 76–83 (2015). https://doi.org/10.1080/21681163.2013.853622
10. Berman, D., Levy, D., Avidan, S., Treibitz, T.: Underwater single image color restoration using haze-lines and a new quantitative dataset. IEEE Trans. Pattern Anal. Mach. Intell. **43**(8), 2822–2837 (2021). https://doi.org/10.1109/TPAMI.2020.2977624
11. Budzan, S., et al.: Using gesture recognition for AGV control: preliminary research. Sensors **23**(6), 3109 (2023). https://doi.org/10.3390/s23063109
12. Cho, Y., Jeong, J., Kim, A.: Model-assisted multiband fusion for single image enhancement and applications to robot vision. IEEE Robot. Autom. Lett. **3**(4), 2822–2829 (2018). https://doi.org/10.1109/LRA.2018.2843127
13. Dzedzickis, A., Kaklauskas, A., Bucinskas, V.: Human emotion recognition: review of sensors and methods. Sensors **20**(3), 592 (2020). https://doi.org/10.3390/s20030592
14. Elhoseny, M.: Multi-object detection and tracking (MODT) machine learning model for real-time video surveillance systems. Circuits Syst. Signal Process. **39**, 611–630 (2020)
15. Fabbri, C., Islam, M.J., Sattar, J.: Enhancing underwater imagery using generative adversarial networks. In: IEEE International Conference on Robotics and Automation, pp. 7159–7165 (2018). https://doi.org/10.1109/ICRA.2018.8460552
16. Guo, Y., Li, H., Zhuang, P.: Underwater image enhancement using a multiscale dense generative adversarial network. IEEE J. Ocean. Eng. **45**(3), 862–870 (2020). https://doi.org/10.1109/JOE.2019.2911447
17. Islam, M.J., Ho, M., Sattar, J.: Understanding human motion and gestures for underwater human-robot collaboration. J. Field Robot. **36**(5), 851–873 (2019). https://doi.org/10.1002/rob.21837
18. Islam, M.J., Xia, Y., Sattar, J.: Fast underwater image enhancement for improved visual perception. IEEE Robot. Autom. Lett. **5**(2), 3227–3234 (2020). https://doi.org/10.1109/LRA.2020.2974710
19. Kang, Z., Yang, J., Yang, Z., Cheng, S.: A review of techniques for 3d reconstruction of indoor environments. ISPRS Int. J. Geo-Inf. **9**(5), 330 (2020). https://doi.org/10.3390/ijgi9050330
20. Li, C., Guo, J., Guo, C.: Emerging from water: underwater image color correction based on weakly supervised color transfer. IEEE Signal Process. Lett. **25**(3), 323–327 (2018). https://doi.org/10.1109/LSP.2018.2792050

21. Li, H., Zhuang, P.: DewaterNet: a fusion adversarial real underwater image enhancement network. Signal Process. Image Commun. **95**, 116248 (2021). https://doi.org/10.1016/j.image.2021.116248
22. Li, H., Zhuang, P., Wei, W., Li, J.: Underwater image enhancement based on dehazing and color correction. In: IEEE International Conference on Parallel Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computer and Networking, pp. 1365–1370 (2019). https://doi.org/10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00196
23. Li, J., Skinner, K.A., Eustice, R.M., Johnson-Roberson, M.: WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. IEEE Robot. Autom. Lett. **3**(1), 387–394 (2018). https://doi.org/10.1109/LRA.2017.2730363
24. Liu, Y., Rong, S., Cao, X., Li, T., He, B.: Underwater single image dehazing using the color space dimensionality reduction prior. IEEE Access **8**, 91116–91128 (2020). https://doi.org/10.1109/ACCESS.2020.2994614
25. Petit, F., Capelle-Laize, A.S., Carre, P.: Underwater image enhancement by attenuation inversion with quaternions. In: IEEE International Conference on Acoustic Speech Signal Process, pp. 1177–1180 (2009). https://doi.org/10.1109/ICASSP.2009.4959799
26. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
27. Sharma, P., et al.: Student engagement detection using emotion analysis, eye tracking and head movement with machine learning. In: Reis, A., Barroso, J., Martins, P., Jimoyiannis, A., Huang, R.YM., Henriques, R. (eds.) TECH-EDU 2022. CCIS, vol. 1720, pp. 52–68. Springer, Cham (2023). https://doi.org/10.1007/9783031229183_5
28. Wang, J., et al.: CA-GAN: class-condition attention GAN for underwater image enhancement. IEEE Access **8**, 130719–130728 (2020). https://doi.org/10.1109/ACCESS.2020.3003351
29. Xiong, J., Zhuang, P., Zhang, Y.: An efficient underwater image enhancement model with extensive Beer-Lambert law. In: IEEE International Conference on Image Processing (ICIP), pp. 893–897 (2020). https://doi.org/10.1109/ICIP40778.2020.9191131
30. Ye, X., Xu, H., Ji, X., Xu, R.: Underwater image enhancement using stacked generative adversarial networks. In: Pacific Rim Conference on Multimedia (PCM), pp. 514–524 (2018). https://doi.org/10.1007/9783030007645_47
31. Zhou, J.T., Du, J., Zhu, H., Peng, X., Liu, Y., Goh, R.S.M.: AnomalyNet: an anomaly detection network for video surveillance. IEEE Trans. Inf. Forensics Secur. **14**(10), 2537–2550 (2019). https://doi.org/10.1109/TIFS.2019.2900907
32. Zhou, S.K., et al.: A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. Proc. IEEE **109**(5), 820–838 (2021). https://doi.org/10.1109/JPROC.2021.3054390
33. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE International Conference on Computer Vision (ICCV), pp. 2223–2232 (2017)