Atlantis: Enabling Underwater Depth Estimation with Stable Diffusion

Supplementary Material

In this supplementary material, we first showcase more underwater images generated in *Atlantis*. Then, we provide more visual results on unseen real underwater scenes from UIEB [5] and SQUID [2] datasets, as well as more underwater image enhancement results using improved depth.

1. More Examples in Atlantis

Here, we present more examples from our underwater depth dataset, *Atlantis*, along with the terrestrial depth maps used for image generation in Figure 1. These images highlight the effectiveness of our dataset in maintaining the structure of the original depth maps while introducing a diverse range of underwater textures, lighting conditions, and color casts in the generated images. The realistic and varied underwater scenes, coupled with corresponding accurate depth maps, are instrumental in training depth estimation models specifically for underwater applications and improve their performance on unseen real underwater images.

2. More Qualitative Results

We provide more qualitative results of UIEB [5] dataset in Figures 2 to 8. These results show a significant domain gap in the depth estimates produced by pretrained terrestrial models across all four methods. This gap is evident in unclear scene layouts, ambiguous depth assignments, and inaccurate distance estimation for water bodies, often accompanied by severe artifacts. For clear underwater images (*e.g.*, Seahorse images in Figure 2 and Tank image in Figure 3), some pretrained models struggle to distinguish foreground objects from the background. The challenge is more evident in dimly lit scenes (*e.g.*, the 4th image in Figure 3), where models frequently fail to produce plausible depths. In contrast, models trained on *Atlantis* show evident improvement when compared to their competitors. They accurately estimate depths, preserve scene layouts, and create distinct edges in depth maps. Water bodies are correctly identified as distant, and foreground objects are more prominently featured, which is crucial for applications like Autonomous Underwater Vehicle (AUV) navigation.

Further, we provide more visual results from the SQUID dataset [2] in Figures 9 to 11. Images in it pose greater challenges due to increased turbidity and color cast. Pretrained models generally fail to estimate depths accurately in many scenes (*e.g.*, the 1st image in Figure 9, the 2nd image in Figure 10 and most images in Figure 11). However, models trained on *Atlantis* significantly outperform their pretrained counterparts, yielding plausible depth maps with better scene layout preservation and object discrimination. Notably, in complex cases like those in Figure 11, depths estimated by models trained on *Atlantis* occasionally appear more reliable than the dataset's reference depth maps, which are derived from stereo pairs.

These results collectively underscore the effectiveness and generalization of *Atlantis* in enhancing the performance of terrestrial depth models for underwater depth estimation, offering a straightforward and effective solution for this challenging task.

3. More Enhancement Results

We provide more visual results of Sea-thru [1] enhancement utilizing the improved depth outputs by models trained on *Atlantis* in Figure 12. The appealing visual results of enchanced images confirm the effectiveness of the improved depth and the practical utility of our dataset in assisting underwater image enhancement.

4. Discussions

While *Atlantis* has brought promising improvements in underwater depth estimation, it's worth noting that this is just the initial version, featuring a relatively smaller sample size compared to KITTI [3] and NYU Depthv2 [9]. Currently, *Atlantis* comprises only 3,200 data pairs, generated using a limited set of 700 underwater images for the *Depth2Underwater* ControlNet training and 400 outdoor depth maps for underwater image generation. Despite its efficacy, there are substantial opportunities for enhancing the dataset's quality and diversity, *e.g.*, incorporating a broader range of underwater images for ControlNet pretraining and generating more varied data. Although we currently employ relative depth for its convenience and compatibility with state-of-the-art depth models, exploring the integration of absolute metric depth in the whole data generation pipeline is a promising avenue, which remains as our future work.

References

[1] Derya Akkaynak and Tali Treibitz. Sea-thru: A method for removing water from underwater images. In *CVPR*, 2019. 1

[2] Dana Berman, Deborah Levy, Shai Avidan, and Tali Treibitz. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE TPAMI*, 43(8):2822–2837, 2020. 1, 11, 12, 13

[3] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *CVPR*, 2012. 1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

[4] Praful Hambarde, Subrahmanyam Murala, and Abhinav Dhall. Uw-gan: Single-image depth estimation and image

enhancement for underwater images. *IEEE Transactions on Instrumentation and Measurement*, 70:1–12, 2021.

[5] Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. An underwater image enhancement benchmark dataset and beyond. *IEEE TIP*, 29: 4376–4389, 2019. 1, 4, 5, 6, 7, 8, 9, 10, 14

[6] Chongyi Li, Saeed Anwar, and Fatih Porikli. Underwater scene prior inspired deep underwater image and video enhancement. *PR*, 98:107038, 2020.

[7] Ce Liu, Suryansh Kumar, Shuhang Gu, Radu Timofte, and Luc Van Gool. Va-depthnet: A variational approach to single image depth prediction. In *ICLR*, 2022.

[8] Yelena Randall and Tali Treibitz. Flsea: Underwater visual-inertial and stereo-vision forward-looking datasets. *arXiv preprint arXiv:2302.12772*, 2023.

[9] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012. 1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13

[10] Nan Wang, Yabin Zhou, Fenglei Han, Haitao Zhu, and Jingzheng Yao. Uwgan: underwater gan for real-world underwater color restoration and dehazing. *arXiv preprint arXiv:1912.10269*, 2019.

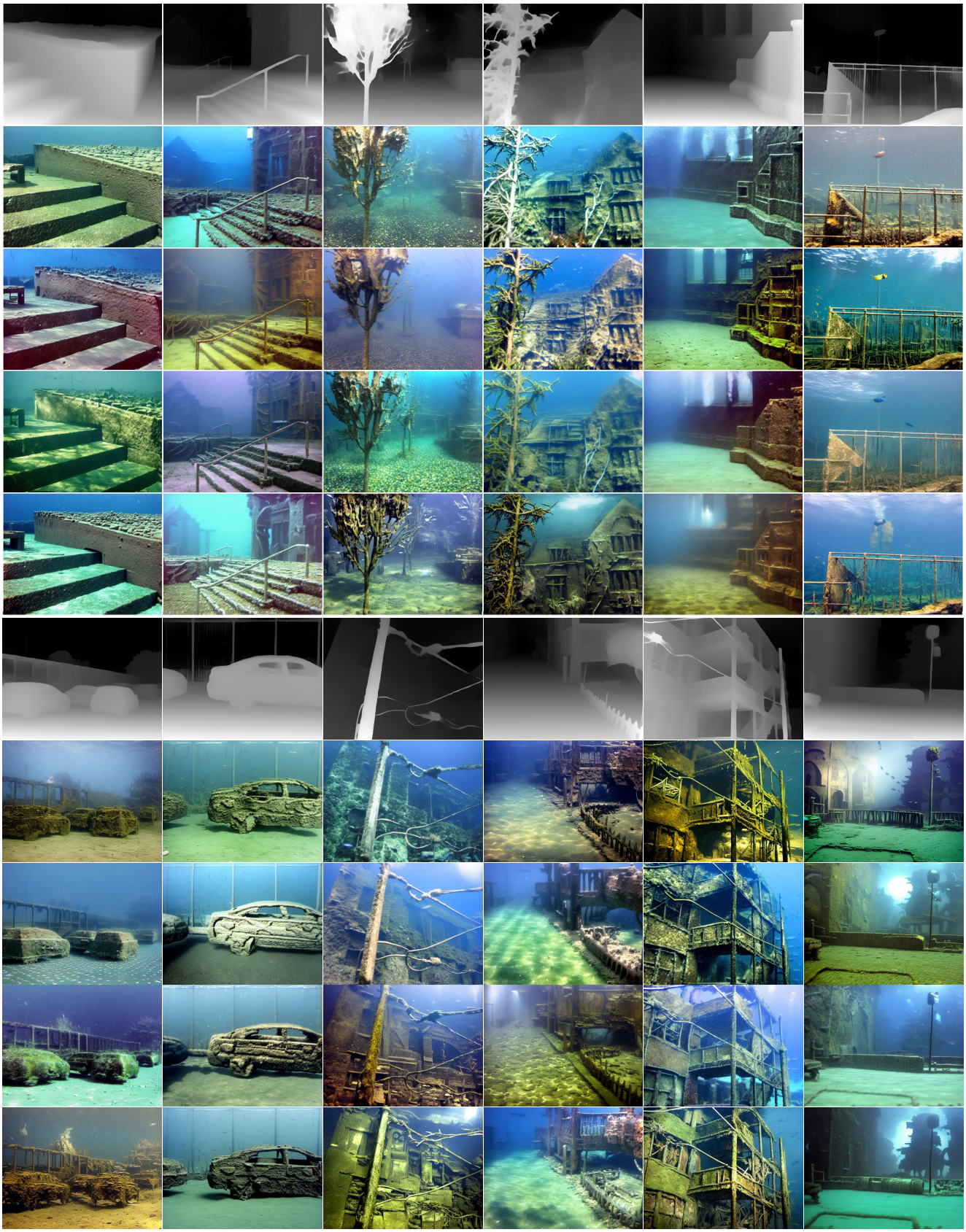

Figure 1. More examples from *Atlantis*, demonstrating diverse variations in scene content, lighting conditions, and color casts, retaining the scene layout of the terrestrial depth maps used for generation.

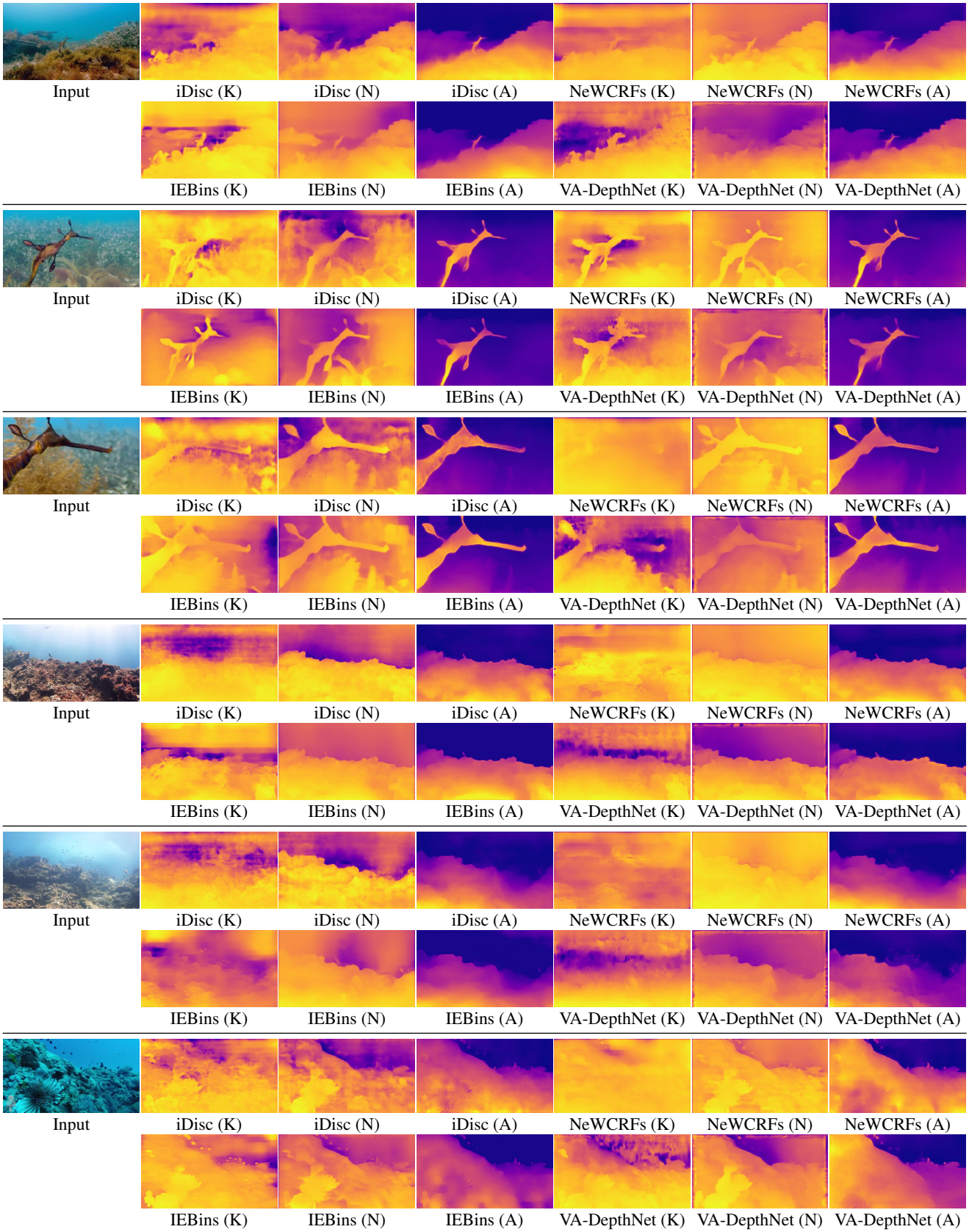

Figure 2. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

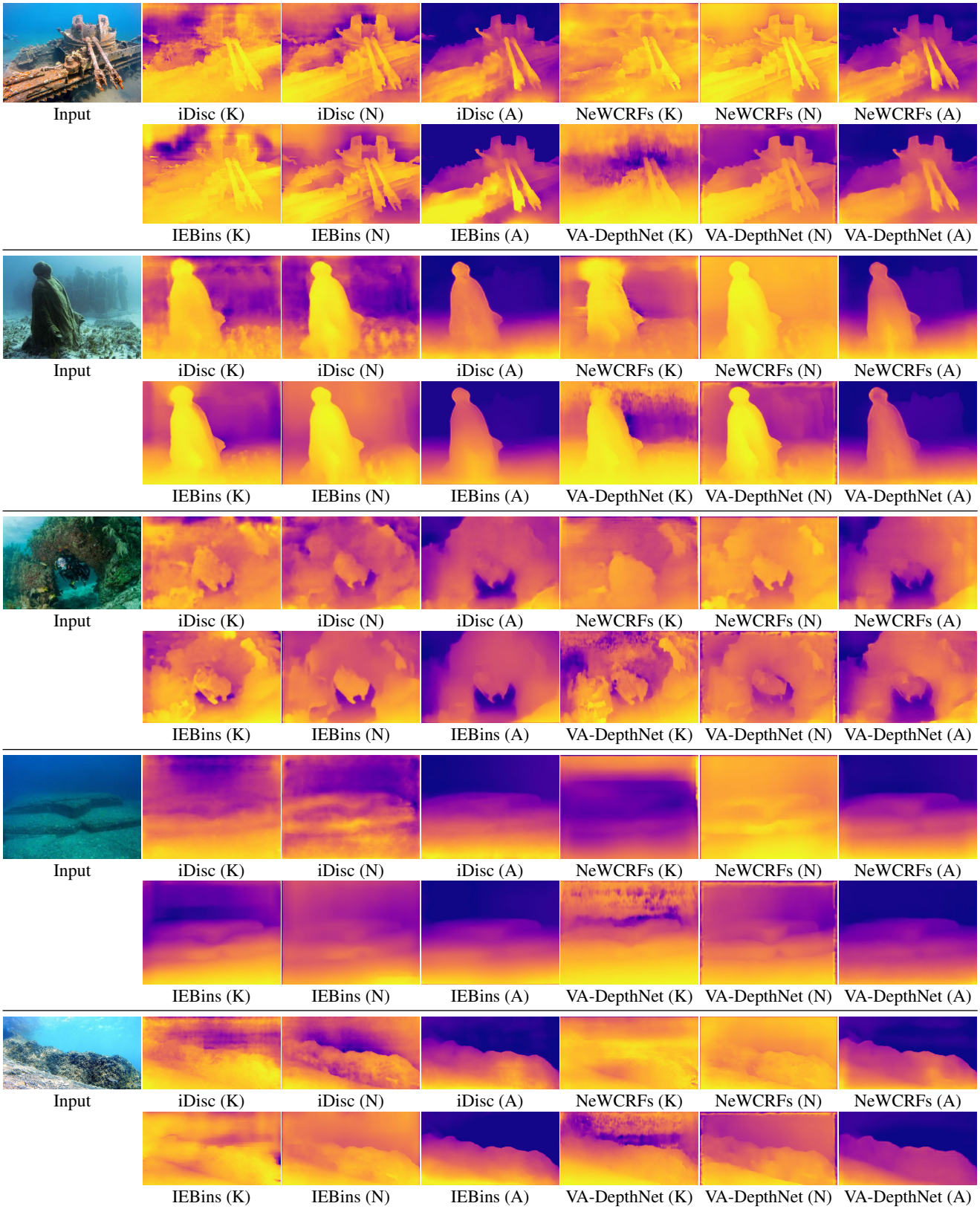

Figure 3. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

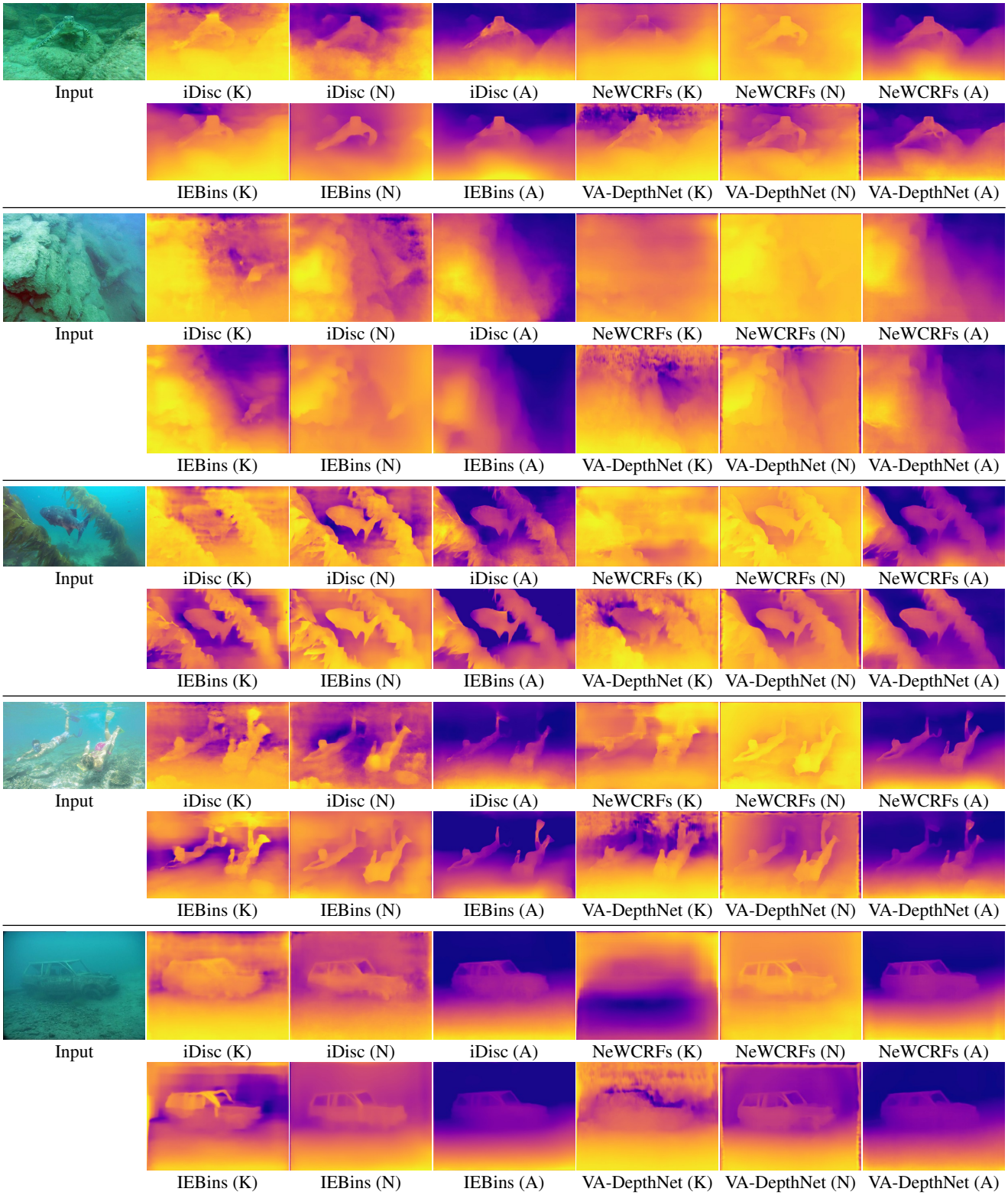

Figure 4. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

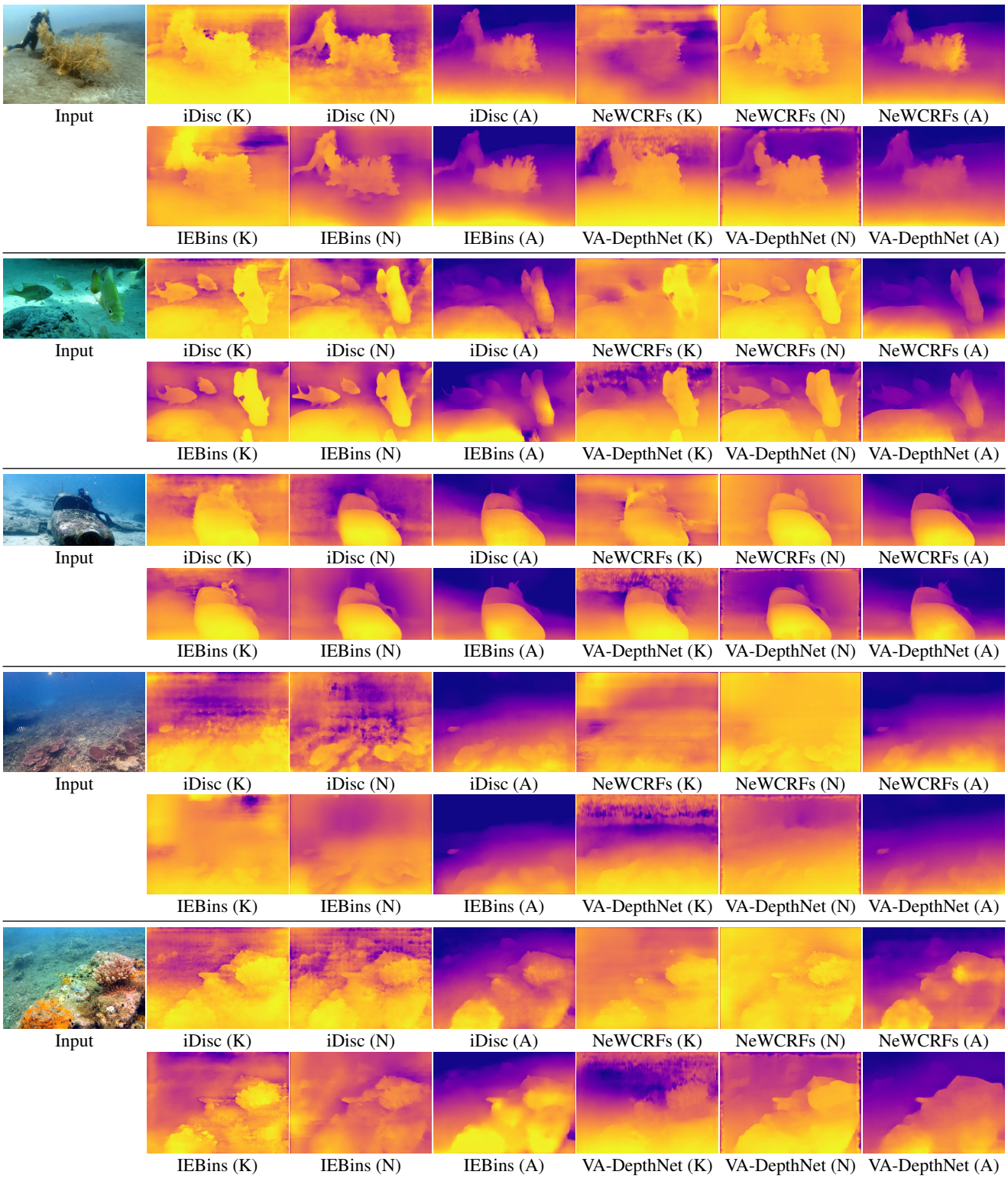

Figure 5. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

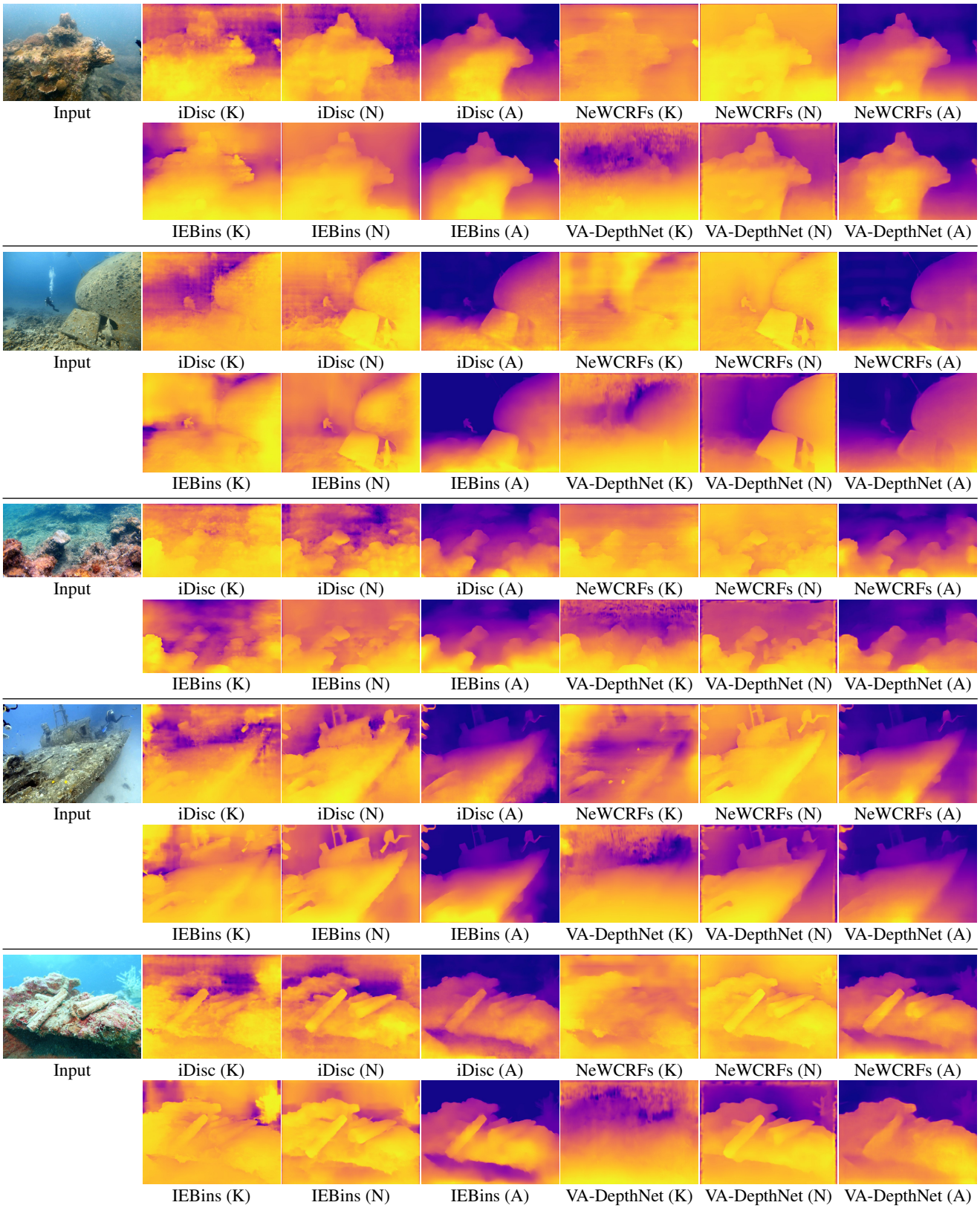

Figure 6. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

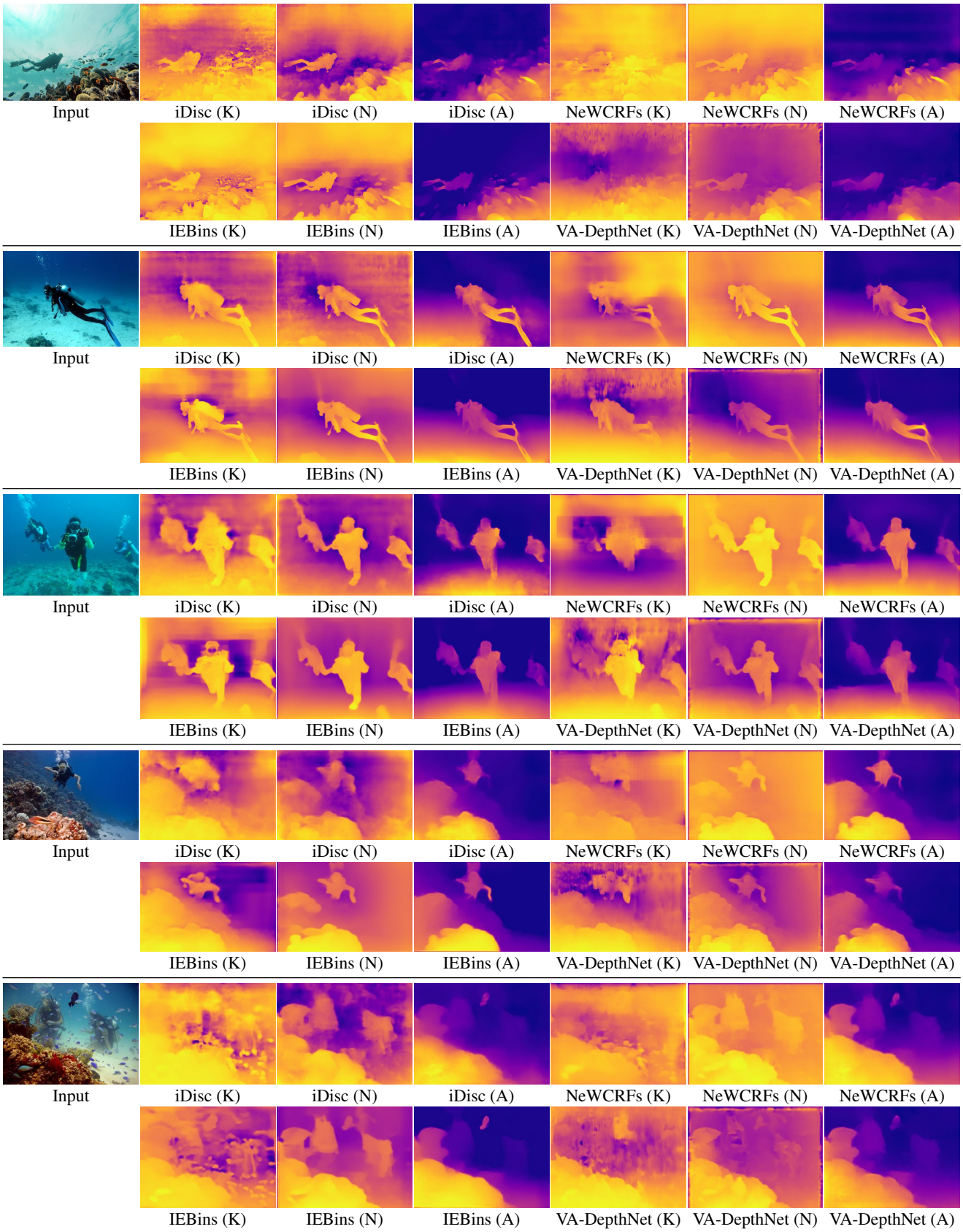

Figure 7. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

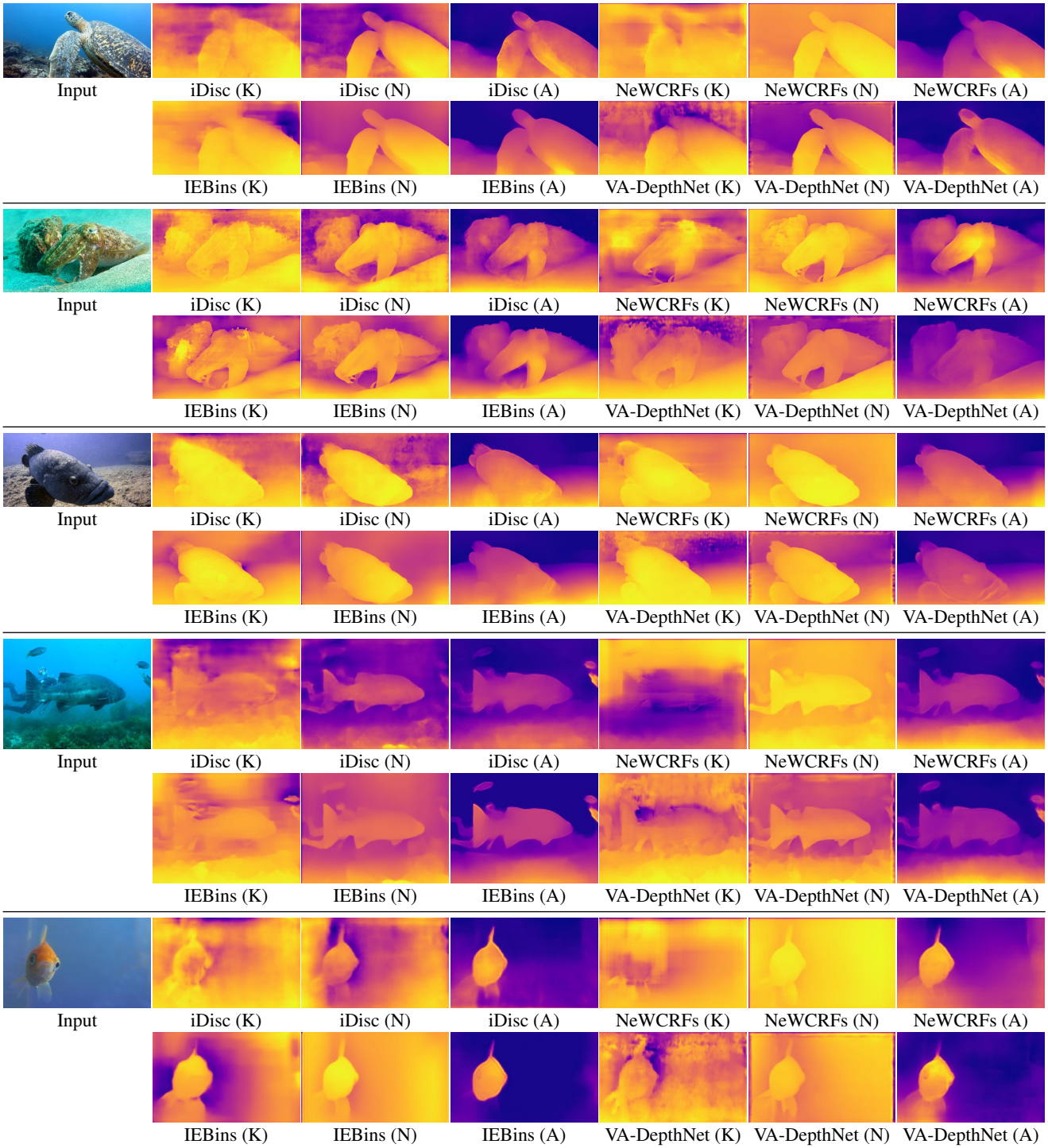

Figure 8. Qualitative results on test set of UIEB dataset [5]. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

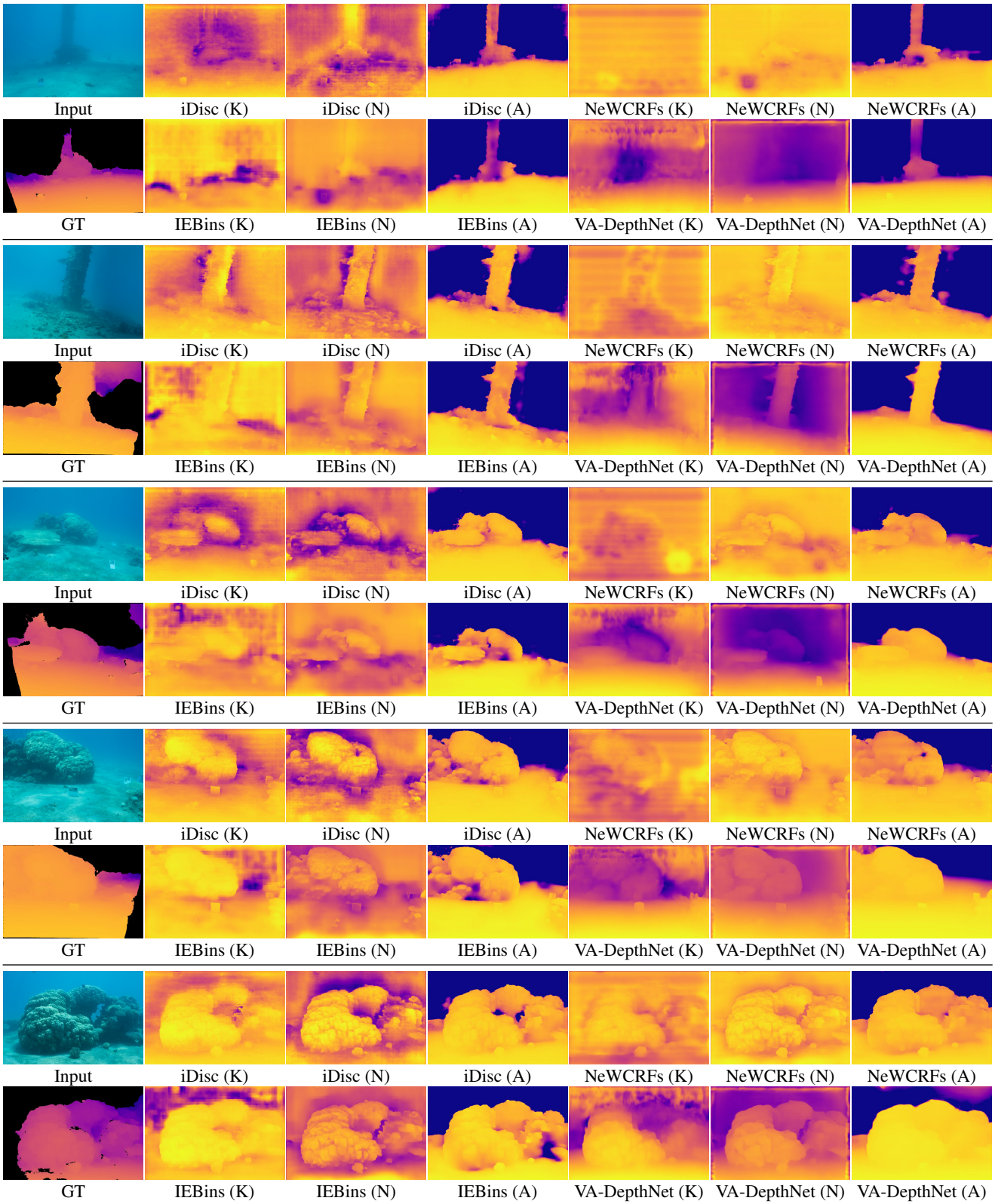

Figure 9. Qualitative results on SQUID [2] dataset. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

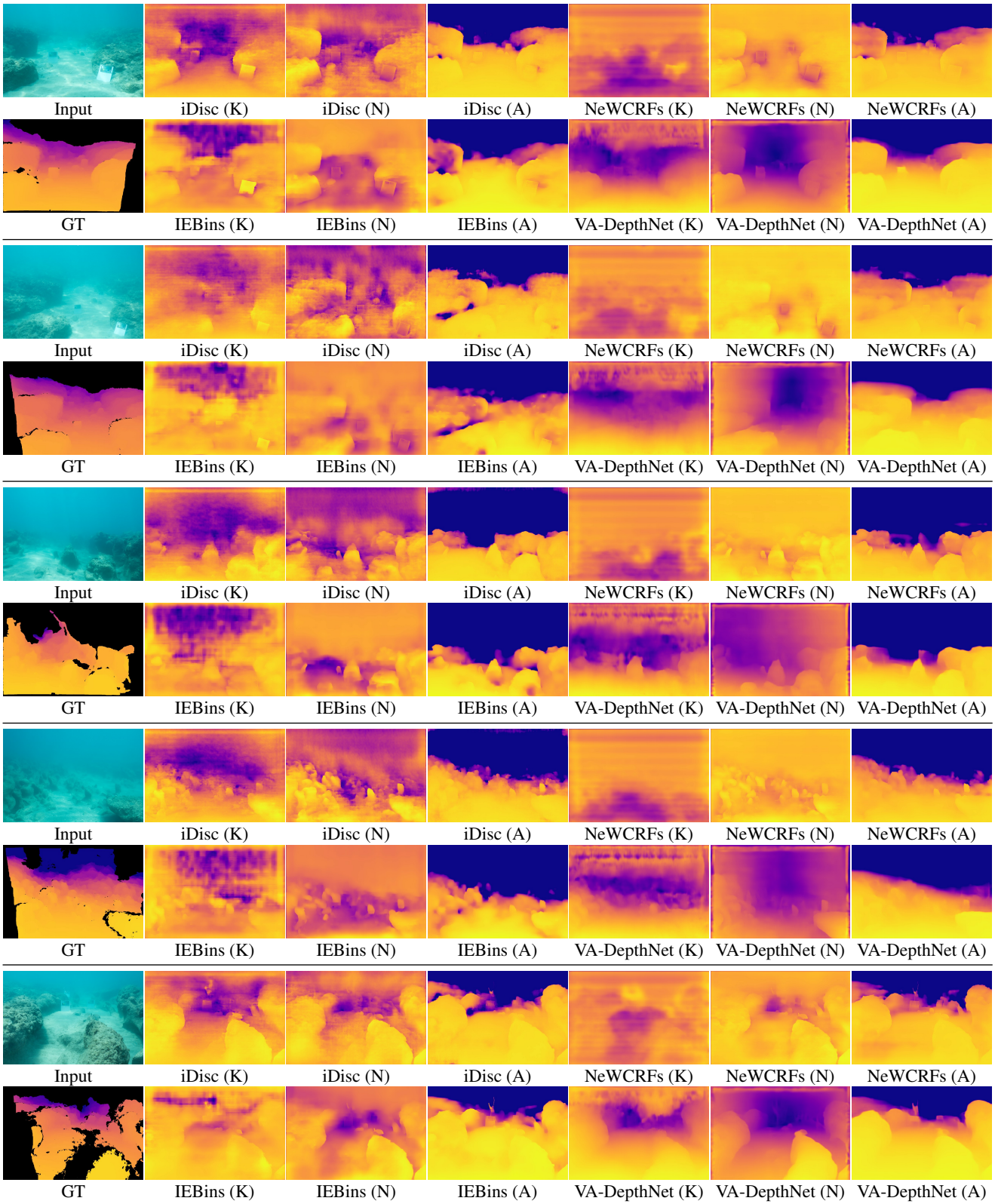

Figure 10. Qualitative results on SQUID [2] dataset. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

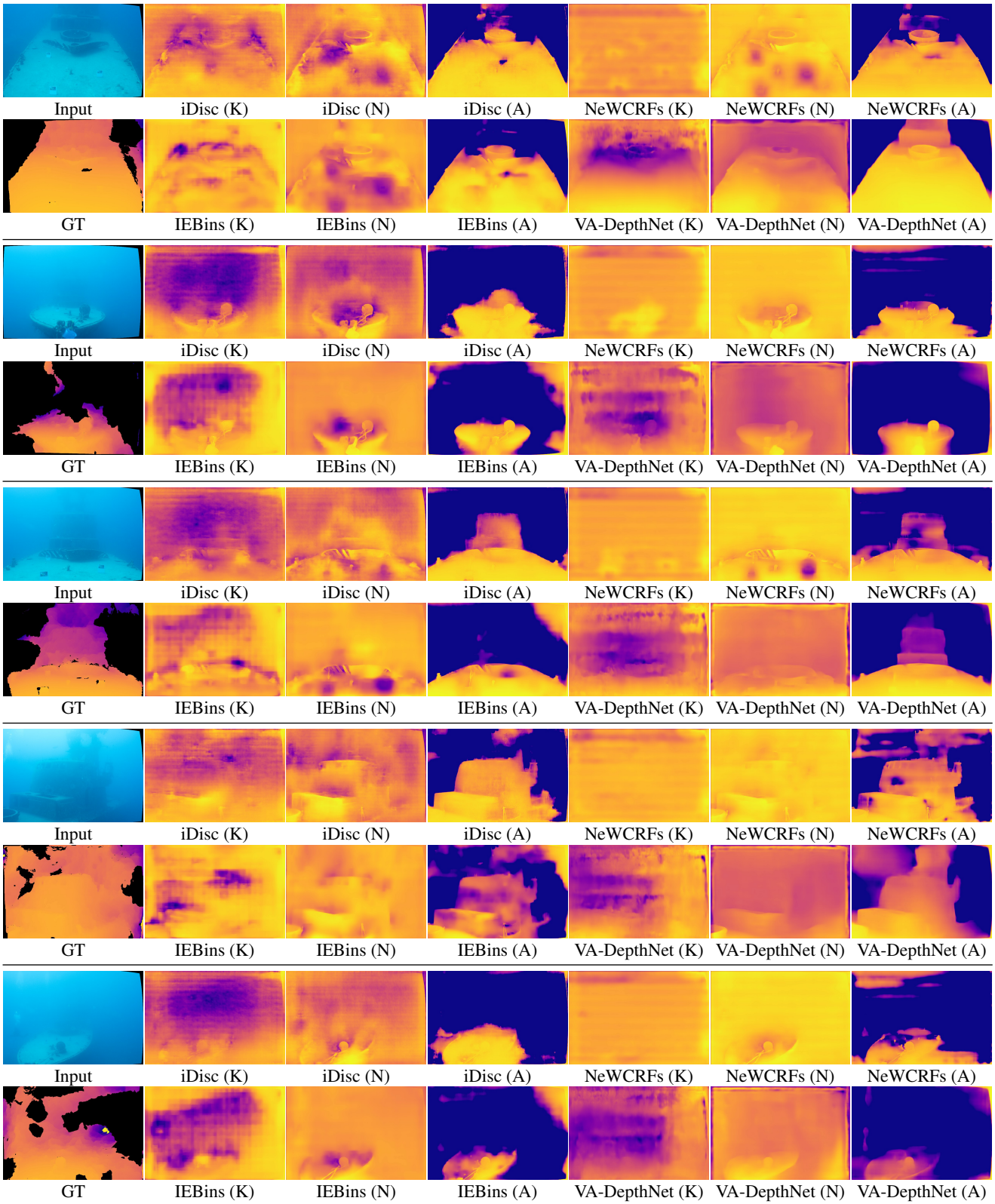

Figure 11. Qualitative results on SQUID [2] dataset. K, N denote pretrained models of KITTI [3], NYU Depthv2 [9] datasets and A represents the models trained on *Atlantis*. Depth results are evidently improved after training on *Atlantis*.

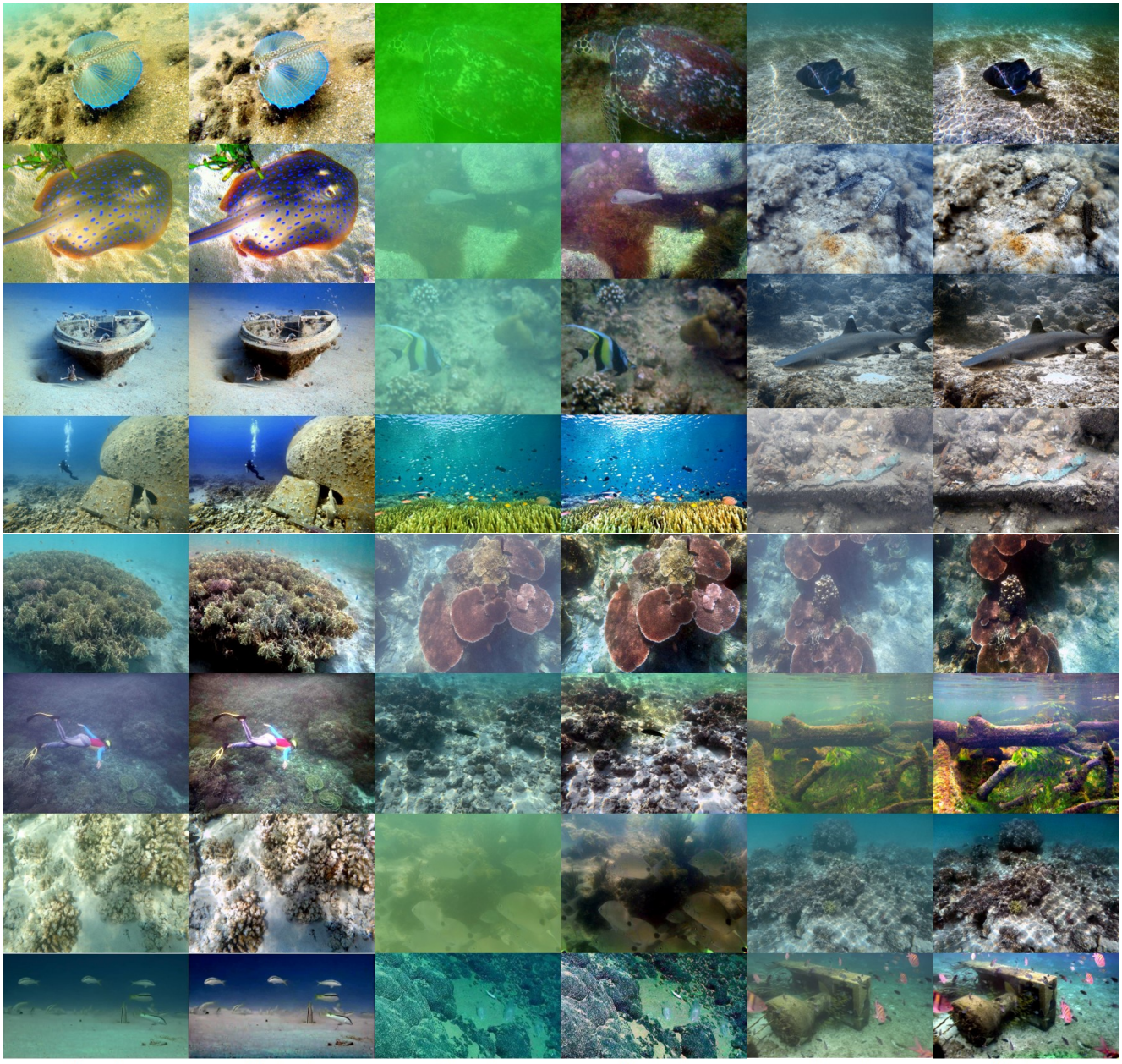

Figure 12. More underwater image enhancement results using the improved depth on images from UIEB dataset [5]. Enhancement outputs well show the practical utility of *Atlantis* on training depth models for reliable underwater depth estimation.