A Deep Learning Based Approach for Synthesizing Realistic Depth Maps

| Conference paper | First Online: 05 September 2023

| pp 369–380 | Cite this conference paper

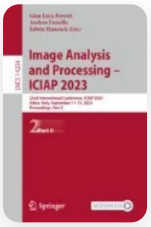

Image Analysis and Processing – ICIAP 2023

(ICIAP 2023)

Patricia L. Suárez, Dario Carpio & Angel Sappa

Part of the book series: Lecture Notes in Computer Science ((LNCS,volume 14234))

Included in the following conference series:
International Conference on Image Analysis and Processing

Abstract

This paper presents a novel cycle generative adversarial network (CycleGAN) architecture for synthesizing high-quality depth maps from a given monocular image. The proposed architecture uses multiple loss functions, including cycle consistency, contrastive, identity, and least square losses, to enable the generation of realistic and high-fidelity

depth maps. The proposed approach addresses this challenge by synthesizing depth maps from RGB images without requiring paired training data. Comparisons with several state-of-the-art approaches are provided showing the proposed approach overcome other approaches both in terms of quantitative metrics and visual quality.

Keywords

depth maps-like transfer domain cross-spectral super-resolution

References

1. Andonian, A., Park, T., Russell, B., Isola, P., Zhu, J.Y., Zhang, R.: Contrastive feature loss for image prediction. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1934–1943 (2021)

Google Scholar

2. Chen, Q., Koltun, V.: Photographic image synthesis with cascaded refinement networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1511–1520 (2017)

Google Scholar

3. Han, J., Shoeiby, M., Petersson, L., Armin, M.A.: Dual contrastive learning for unsupervised image-to-image translation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (2021)

Google Scholar

4. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134 (2017)

Google Scholar

5. Jung, C., Kwon, G., Ye, J.C.: Exploring patch-wise semantic relation for contrastive learning in image-to-image translation tasks. arXiv preprint arXiv:2203.01532 (2022)

6. Khan, M.F.F., Troncoso Aldas, N.D., Kumar, A., Advani, S., Narayanan, V.: Sparse to dense depth completion using a generative adversarial network with intelligent sampling strategies. In: Proceedings of the 29th ACM International Conference on Multimedia, pp. 5528–5536 (2021)

Google Scholar

7. Lee, S., Lee, J., Kim, D., Kim, J.: Deep architecture with cross guidance between single image and sparse lidar data for depth completion. IEEE Access **8**, 79801–79810 (2020)

Article Google Scholar

8. Liu, J., et al.: Identity preserving generative adversarial network for cross-domain person re-identification. IEEE Access **7**, 114021–114032 (2019)

Article Google Scholar

9. Mondal, T.G., Jahanshahi, M.R.: Fusion of color and hallucinated depth features for enhanced multimodal deep learning-based damage segmentation. Earthq. Eng. Eng. Vib. **22**, 55–68 (2023). https://doi.org/10.1007/s11803-023-2155-2

Article Google Scholar

10. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: European Conference on Computer Vision (2020)

Google Scholar

11. Ranasinghe, N., et al.: Season traveller: multisensory narration for enhancing the virtual reality experience. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2018)

Google Scholar

12. Schulter, S., Zhai, M., Jacobs, N., Chandraker, M.: Learning to look around objects for top-view representations of outdoor scenes. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 787–802 (2018)

Google Scholar

13. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from RGBD images. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7576, pp. 746–760. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33715-4_54

Chapter Google Scholar

14. Suárez, P.L., Sappa, A.D.: Toward a thermal image-like representation. In: Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (2023)

Google Scholar

15. Tang, H., Liu, H., Sebe, N.: Unified generative adversarial networks for controllable image-to-image translation. IEEE Trans. Image Process. **29**, 8916–8929 (2020)

Article MATH Google Scholar

16. Tian, Z., et al.: Adversarial self-attention network for depth estimation from RGB-d data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020)

Google Scholar

17. Valencia, A.J., Idrovo, R.M., Sappa, A.D., Guingla, D.P., Ochoa, D.: A 3D vision based approach for optimal grasp of vacuum grippers. In: Proceedings of the IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (2017)

Google Scholar

18. Wei, W., Qi, R., Zhang, L.: Effects of virtual reality on theme park visitors' experience and behaviors: a presence perspective. Tour. Manage. **71**, 282–293 (2019)

Article Google Scholar

19. Zhan, H., Garg, R., Weerasekera, C.S., Li, K., Agarwal, H., Reid, I.: Unsupervised learning of monocular depth estimation and visual odometry with deep feature reconstruction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 340–349 (2018)

Google Scholar

20. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)

Google Scholar

Acknowledgements

This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-22-1-0261; and partially supported by the Grant PID2021-128945NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe"; the "CERCA Programme/Generalitat de Catalunya"; and the ESPOL project CIDIS-12-2022.

Author information

Authors and Affiliations

ESPOL Polytechnic University, Guayaquil, Ecuador

Patricia L. Suárez, Dario Carpio & Angel Sappa

Computer Vision Center, Barcelona, Spain

Angel Sappa

Corresponding author

Correspondence to Patricia L. Suárez .

Editor information

Editors and Affiliations

University of Udine, Udine, Italy
Gian Luca Foresti

University of Udine, Udine, Italy
Andrea Fusiello

University of York, York, UK
Edwin Hancock

Rights and permissions

Copyright information

About this paper

Cite this paper

Suárez, P.L., Carpio, D., Sappa, A. (2023). A Deep Learning Based Approach for Synthesizing Realistic Depth Maps. In: Foresti, G.L., Fusiello, A., Hancock, E. (eds) Image Analysis and Processing – ICIAP 2023. ICIAP 2023. Lecture Notes in Computer Science, vol 14234. Springer, Cham. https://doi.org/10.1007/978-3-031-43153-1_31

DOI
https://doi.org/10.1007/978-3-031-43153-1_31

Published
05 September 2023

Publisher Name
Springer, Cham

Print ISBN
978-3-031-43152-4

Online ISBN
978-3-031-43153-1

eBook Packages
Computer Science
Computer Science (R0)

Publish with us

Policies and ethics