

Designing task-oriented cameras by machine learning

Hajime Nagahara¹

¹D3 Center, Osaka University

2-8, Yamadaoka, Suita, Osaka, 565-0871, Japan

E-mail:nagahara@ids.osaka-u.ac.jp

Abstract Deep neural network (DNN) is a powerful tool for solving image processing and computer vision tasks such as image and video reconstructions, object recognition, and scene understanding, etc. However, DNN have been used for only digital domain in the imaging pipeline, such as the feature extractor and classifier models after an image is captured and digitized. In this research, we propose a new framework called “deep sensing.” The proposed framework also models the analog layer to the neural network model and jointly optimizes the parameters in optics and sensor designs of a camera, as well as reconstruction and classification models by the same training strategy.

Keywords: Deep neural network, Computational photography, Compressive sensing

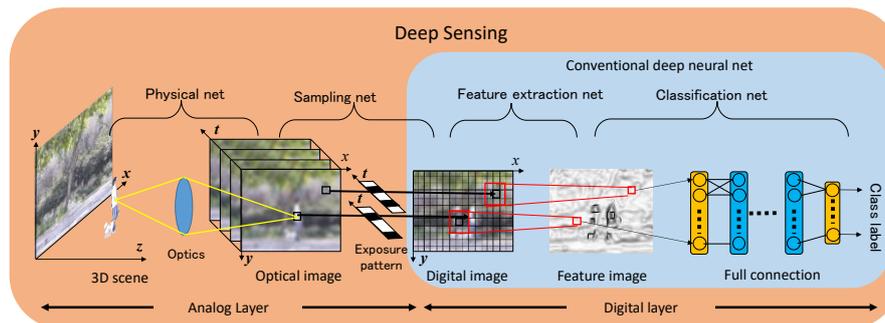


Figure 1: Conceptual figure of Deep sensing

1. Overview

Deep neural network (DNN) is a powerful tool for solving image processing and computer vision tasks such as image and video reconstructions, object recognition, and scene understanding, etc. It realizes to drastically improve the quality of reconstruction and the accuracy of recognition to the classical methods since feature extractor and classifier models are designed by training based on the target data. However, DNN have been used for only the digital domain in the imaging pipeline, such as the feature extractor and classifier models after an image is captured and digitized, as shown in the blue part of figure 1. On the other hand, optics and sensors in the analog layer still have been designed by hand based on theoretical or empirical analysis. It is not always grantee that the designs and hardware setting parameters are optimal to the applications and target tasks. In this research, we propose a new framework called “deep sensing,” as shown in figure 1. The proposed framework also models the analog layer to the neural network model and jointly optimizes the parameters in optics and sensor designs of a camera as well as reconstruction and classification models by the same training strategy. In this talk, I introduce the concept of deep sensing and show our work; compressive light field sensing [1,4,5,8], compressive video sensing [2], compressive ToF [9], action recognition [3], and privacy-preserving imaging. The details are in the corresponding papers.

References

- [1] Y. Inagaki, Y. Kobayashi, K. Takahashi, T. Fujii, and H. Nagahara, “Learning to capture light fields through a coded aperture camera,” European Conference on Computer Vision, pp. 418–434, 2018,.
- [2] M. Yoshida, A. Torii, M. Okutomi, K. Endo, Y. Sugiyama, R. Taniguchi, and H. Nagahara, “Joint optimization for compressive video sensing and reconstruction under hardware constraints,” European Conference on Computer Vision, pp. 634–649, 2018.
- [3] T. Okawara, M. Yoshida, H. Nagahara, and Y. Yagi, “Action Recognition from a Single Coded Image,” in Proceedings of IEEE International Conference on Computational Photography, 2020.
- [4] K. Sakai, K. Takahashi, T. Fujii, H. Nagahara: “Acquiring Dynamic Light Fields through Coded Aperture Camera”, European Conference on Computer Vision, 2020.
- [5] R. Mizuno, K. Takahashi, M. Yoshida, C. Tsutake, T. Fujii, H. Nagahara, “Acquiring a Dynamic Light Field through a Single-Shot Coded Image”, IEEE Conference on Computer Vision and Pattern Recognition, June, 2022.
- [6] S. Kumawat, T. Okawara, M. Yoshida, H. Nagahara, Yasushi Yagi, “Action Recognition From a Single Coded Imaging”, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1-14, Aug., 2022.
- [7] S. Kumawat, H. Nagahara, “Privacy-Preserving Action Recognition via Motion Difference Quantization”, European Conference on Computer Vision, Oct, 2022.
- [8] Shuji Habuchi, Keita Takahashi, Chihiro Tsutake, Toshiaki Fujii, Hajime Nagahara, “Time-Efficient Light-Field Acquisition Using Coded Aperture and Events”, IEEE Conference on Computer Vision and Pattern Recognition, June, 2024.
- [9] Michitaka Yoshida, Daisuke Hayashi, Lioe De Xing, Keita Yasutomi, Shoji Kawahito, Keiichiro Kagawa, Hajime Nagahara, “Single-shot efficient depth imaging based on time-compressive CMOS image sensor”, IEEE Sensors, Oct., 2024.