

Roles of Deep Learning in Optical Imaging †

Vineela Chandra Dodda ¹ and Inbarasan Muniraj ^{1,2,*}

¹ Department of Electronics and Communication Engineering, School of Engineering and Applied Sciences, SRM University AP, Amaravathi 522240, India; vineelachandra_dodda@srmmap.edu.in

² LiFE Laboratory, Department of Electronics and Communication Engineering, Alliance College of Engineering and Design, Alliance University, Bengaluru 562106, India

* Correspondence: inbarasan.muniraj@alliance.edu.in

† Presented at the International Conference on “Holography Meets Advanced Manufacturing”, Online, 20–22 February 2023.

Abstract: Imaging-based problem-solving approaches are an exemplary way of handling problems in various scientific applications. With an increased demand for automation, artificial intelligence techniques have shown exponential growth in recent years. In this context, deep-learning-based “learned” solutions have been widely adopted in many applications and are thus slowly becoming an inevitable alternative tool. It is known that in contrast to the conventional “physics-based” approach, deep learning models are a “data-driven” approach, where the outcomes are based on data analysis and interpretation. Thus, deep learning approaches have been applied in several (optical and computational) imaging-based scientific problems such as denoising, phase retrieval, hologram reconstruction, and histopathology, to name a few. In this work, we present two deep-learning networks for 3D image denoising and off-focus voxel removal.

Keywords: optical 3D imaging; unsupervised denoising; off-focus removal; integral imaging

1. Introduction

Integral imaging (II) is one of the passive three-dimensional (3D) imaging techniques invented by Gabriel Lippmann in 1908 [1] and has received wide attention, as the applications of II span several research problems in optical engineering research areas [2–4]. For instance, these include biomedicine, security, autonomous vehicles, and remote sensing, to name a few [5].

Advanced machine learning (ML) and deep learning (DL) algorithms have been shown to produce superior results in computer-vision-based applications. Thereafter, such approaches have also been extended to solve several problems in various other scientific research areas. In particular, the DL framework has been proven as an important tool to make automatic decisions, as it solves numerous image-based problems without much human intervention. Convolution Neural Networks (CNN) are a widely used DL algorithm for several problems such as image classification [6], autonomous driving [7], etc. Furthermore, a CNN framework for 3D face recognition and classification in a photon-starved environment has also been demonstrated [2,8].

2. Integral Imaging

Integral imaging (II) captures a 3D scene in the form of two-dimensional (2D) elemental images (EIs) in addition to the directional information (i.e., angle of propagation). Notably, 3D scene reconstruction can be achieved in two ways: (i) optical methods and (ii) computational methods [9]. In computational integral imaging (CII), a geometric ray back-propagation method is employed which magnifies and superimposes the EIs onto each other to reconstruct 3D sectional images [10]. Consequently, the objects or 3D points which are located at the corresponding depth position in an imaging plane are properly



Citation: Dodda, V.C.; Muniraj, I. Roles of Deep Learning in Optical Imaging. *Eng. Proc.* **2023**, *34*, 6. <https://doi.org/10.3390/HMAM2-14123>

Academic Editor: Vijayakumar Anand

Published: 6 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

overlapped and in focus, while the other points at different depth locations do not overlap properly and hence appear off-focus or defocused. The defocused points in the 3D sectional image do not convey any valuable information and are therefore redundant. Recently, we have demonstrated a way to manually identify and remove the off-focus points from a 3D sectional image [11]. Furthermore, under some special imaging scenarios (e.g., biomedical imaging and night vision), low light levels or photon-starved illumination conditions may be encountered. In such cases, since image capturing happens in much darker conditions, the recorded image looks degraded due to the presence of noise [8,10]. Nevertheless, this system has been shown to provide a better 3D reconstruction in terms of the PSNR even with fewer photons, e.g., 100 photons [10].

2.1. Denoising

For image denoising, various methods have been proposed in the literature such as prediction filtering, transformation-based methods, rank reduction methods, and dictionary learning methods, to name a few. In addition to these, DL algorithms have also been applied to the image denoising problem [12]. In this regard, there are two methods that are commonly followed to train the DL network: (i) supervised and (ii) unsupervised. First, we discuss supervised learning, where an under-complete autoencoder is used to denoise the noisy 3D integral (sectional) images with a patch-based approach. In this process, the noisy input 3D sectional image is divided into multiple patches, which are then used to train the neural network in a supervised manner (we use clean data as labels). We note that by using the patch-based approach, the time required to prepare the labeled training data is greatly reduced. Then, after denoising, the acquired denoised patches can be combined via an unpatching process. Figure 1 depicts the supervised denoising technique used on our dataset [13]. To train the network, 20 epochs were employed with a learning rate of 0.001.



Figure 1. Denoised results for supervised learning.

Figure 1c shows the denoised 3D sectional image. We analyzed the performance of the proposed method quantitatively in terms of the peak signal-to-noise ratio (PSNR). For instance, the PSNR value given in Figure 1c is an estimation from Figure 1a,c. It is evident from Figure 1c that the proposed denoising method has a better performance in terms of the PSNR. Second, we proposed an unsupervised learning method for 3D image denoising. In this study, we opted for a U-Net architecture [8]. This is an end-to-end, fully unsupervised denoising approach where the noisy photons in the 3D sectional image are fed as an input to the network. The major components in the U-Net are encoder and decoder blocks with skip connection layers [14–16]. In addition to this, skip blocks (SB) were added to the skip connection strategy in the U-Net architecture to avoid the vanishing gradients problem. In the training process, the 3D input image is given in the form of patches to the network. The patched input image is converted to a 1D vector and fed as an input to the network. After removing the noise, we unpatch the 1D vector and convert it back to the size of the input data. In our experiments, to test the performance of the proposed method, we used two 3D objects: a tri-colored ball known as Object 1 in Figure 2a and a toy bird referred to as Object 2 in Figure 2(a1,a2). Figure 2(b1,b2) are obtained after the TV denoising method. The proposed method results are given in Figure 2(c1,c2). Notably, we used 20% of the PCSI patches for validation and 60% of the patches for training purposes. In this work, 15 epochs were used with a learning rate of 0.001 to train the network. The PSNR values are shown in Figure 2.

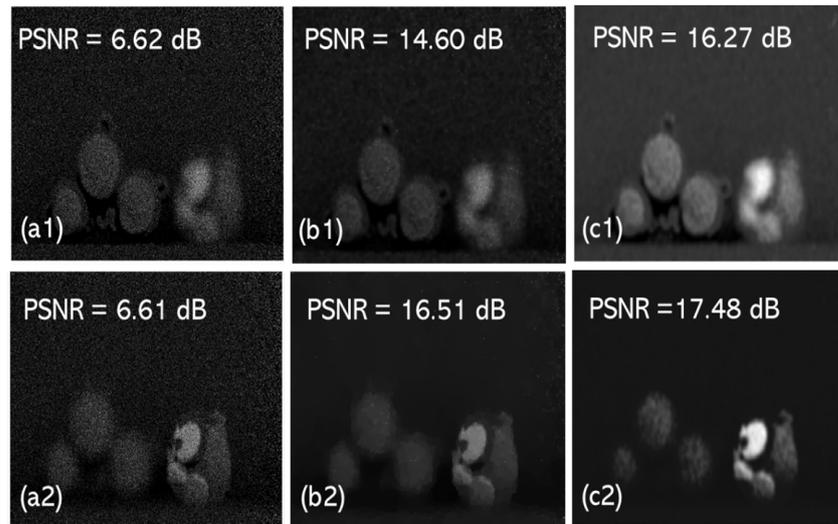


Figure 2. Denoising results: (a1,b1,c1) represent the noisy photon-counted 3D sectional image, the TV denoised image, and the result of our proposed denoising method when object 1 is in focus, respectively, and (a2,b2,c2) represent the noisy photon-counted 3D sectional image, the TV denoised image, and the result of our proposed denoising method when object 2 is in focus, respectively.

2.2. Off-Focus Removal

Several studies have been conducted to demonstrate the feasibility of combining photon detection imaging or photon counting imaging (PCI) techniques with conventional 3D integral imaging systems, known as photon counted integral imaging (PCII) [2,9–11,17]. In such systems, it is known that the reconstructed depth images contain both the focused and off-focus (or out-of-focus) voxels simultaneously (see for instance Figure 3). Off-focus pixels often look blurred and therefore do not convey clear information about the scene. Several approaches have been proposed to efficiently remove the off-focus points from the reconstructed 3D images [4,11]. We note that the existing approaches are subjective as they involve manual calculation of algorithm parameters such as variance, threshold, etc., which is time consuming.

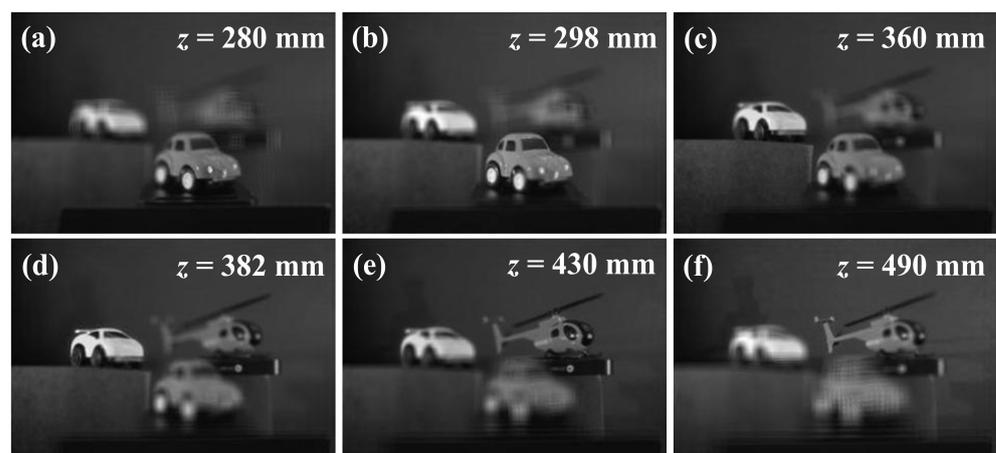


Figure 3. Reconstructed 3D CII sectional images at various depth locations.

Here, we propose a new ensemble Dense Neural Network (DNN) model that is composed of six different DNN models, each trained with its own set of training datasets for removing off-focus points from 3D sectional images. It is known that data pre-processing enhances the accuracy of the network; therefore, we used the Otsu thresholding algorithm [18] to remove the unwanted (and obvious) background from the 3D sectional images. In this work, we employed an ADAM optimizer to update the weights and bias [13], and the

standard mean squared error (MSE) was used as the cost function in our training process. Notably, the proposed ensemble deep neural network was trained (supervised way) using the conventional 3D sectional images from various depth locations and the corresponding focused images (labels). We tested the method on a 3D scene that contains two toy cars and one toy helicopter (see Figure 4) [13]. We used an Intel® Xeon® Silver 4216 CPU @2.10 GHz (two processors) with 256 GB RAM and a 64-bit operating system to simulate all the scenarios.

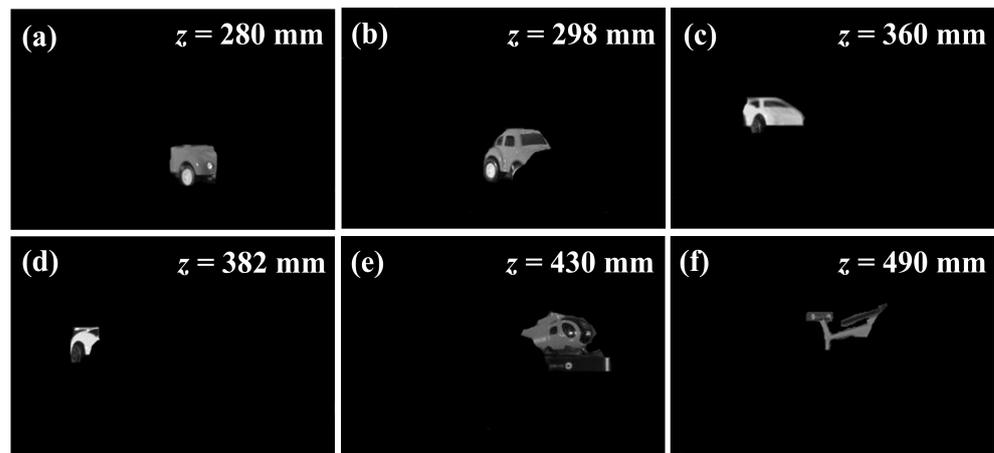


Figure 4. Reconstructed focus only CII sectional images by using the proposed DL network.

3. Conclusions

In summary, we demonstrated that it is possible to use deep learning networks to solve some of the inherent problems of 3D optical imaging systems. For instance, we have tackled two important problems that exist in 3D integral imaging systems, i.e., denoising and off-focus removal, using two different datasets. For our study, it is evident that DL can be used to solve problems that are too complex to carry out manually. It is therefore expected that we will further expand our analysis to various other imaging modalities such as holography, microscopy, etc.

Author Contributions: V.C.D. and I.M. contributed equally to manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Department of Science and Technology (DST) under the Science and Engineering Research Board (SERB) grant number SRG/2021/001464.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data for this paper is not publicly available but shall be provided upon reasonable request to the corresponding author.

Acknowledgments: Authors thank Suchit Patel of Poornima College of Engineering, India, for lending his support in the simulations and we sincerely thank Bahram Javidi of the University of Connecticut and Moon Inkyu of DGIST, Korea, for providing the dataset.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lippmann, G. La photographie integrale. *Comptes-Rendus* **1908**, *146*, 446–451.
2. Markman, A.; Javidi, B. Learning in the dark: 3D integral imaging object recognition in very low illumination conditions using convolutional neural networks. *OSA Contin.* **2018**, *1*, 373–383. [[CrossRef](#)]
3. Yeom, S.; Javidi, B.; Watson, E. Three-dimensional distortion-tolerant object recognition using photon-counting integral imaging. *Opt. Express* **2007**, *15*, 1513–1533. [[CrossRef](#)] [[PubMed](#)]

4. Yi, F.; Lee, J.; Moon, I. Simultaneous reconstruction of multiple depth images without off-focus points in integral imaging using a graphics processing unit. *Appl. Opt.* **2014**, *53*, 2777–2786. [[CrossRef](#)] [[PubMed](#)]
5. Javidi, B.; Carnicer, A.; Arai, J.; Fujii, T.; Hua, H.; Liao, H.; Martínez-Corral, M.; Pla, F.; Stern, A.; Waller, L.; et al. Roadmap on 3D integral imaging: Sensing, processing, and display. *Opt. Express* **2020**, *28*, 32266–32293. [[CrossRef](#)] [[PubMed](#)]
6. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* **2016**, *23*, 1499–1503. [[CrossRef](#)]
7. Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. Deepdriving: Learning affordance for direct perception in autonomous driving. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 2722–2730.
8. Dodda, V.C.; Kuruguntla, L.; Elumalai, K.; Chinnadurai, S.; Sheridan, J.T.; Muniraj, I. A denoising framework for 3D and 2D imaging techniques based on photon detection statistics. *Sci. Rep.* **2023**, *13*, 1365. [[CrossRef](#)] [[PubMed](#)]
9. Moon, I.; Muniraj, I.; Javidi, B. 3D visualization at low light levels using multispectral photon counting integral imaging. *J. Disp. Technol.* **2013**, *9*, 51–55. [[CrossRef](#)]
10. Muniraj, I.; Guo, C.; Lee, B.G.; Sheridan, J.T. Interferometry based multispectral photon-limited 2D and 3D integral image encryption employing the Hartley transform. *Opt. Express* **2015**, *23*, 15907–15920. [[CrossRef](#)] [[PubMed](#)]
11. Muniraj, I.; Guo, C.; Malallah, R.; Maraka, H.V.R.; Ryle, J.P.; Sheridan, J.T. Subpixel based defocused points removal in photon-limited volumetric dataset. *Opt. Commun.* **2017**, *387*, 196–201. [[CrossRef](#)]
12. Choi, G.; Ryu, D.; Jo, Y.; Kim, Y.S.; Park, W.; Min, H.S.; Park, Y. Cycle-consistent deep learning approach to coherent noise reduction in optical diffraction tomography. *Opt. Express* **2019**, *27*, 4927–4943. [[CrossRef](#)] [[PubMed](#)]
13. Dodda, V.C.; Kuruguntla, L.; Elumalai, K.; Muniraj, I.; Chinnadurai, S. An undercomplete autoencoder for denoising computational 3D sectional images. In *Computational Optical Sensing and Imaging*; Optica Publishing Group: Washington, DC, USA, 2022; p. JW2A-19.
14. Lempitsky, V.; Vedaldi, A.; Ulyanov, D. Deep image prior. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 9446–9454.
15. Yang, L.; Wang, S.; Chen, X.; Saad, O.M.; Chen, W.; Oboué, Y.A.S.I.; Chen, Y. Unsupervised 3-D Random Noise Attenuation Using Deep Skip Autoencoder. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5905416. [[CrossRef](#)]
16. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
17. Tavakoli, B.; Javidi, B.; Watson, E. Three dimensional visualization by photon counting computational integral imaging. *Opt. Express* **2008**, *16*, 4426–4436. [[CrossRef](#)] [[PubMed](#)]
18. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.