Article

Role 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, Andhra Pradesh 522240, India.
[2] LiFE Laboratory, Department of Electronics and Communication Engineering, Alliance College of Engineering and Design, Alliance University, Bengaluru, Karnataka 562106, India.
* Correspondence: inbarasan.muniraj@alliance.edu.in; inbarasan.m@gmail.com

Abstract: Imaging based problem-solving approaches have shown an illustrative way of handling problems for various scientific applications. With an increased demand for automation, artificial intelligence techniques have shown an exponential growth in the recent years. In this context, deep learning-based "learned" solutions are widely opted for many applications thus slowly becoming an inevitable alternative tool. It is known that in contrast to the conventional "physics-based" approach, deep learning models are known to be a "data-driven" approach where the outcomes are based on data analysis and interpretation. Thus, the deep learning approaches have been applied for 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 voxels removal.

Keywords: Optical 3D Imaging; unsupervised denoising; off-focus removal; Integral Imaging.

Citation: Vineela Chandra, Dodda.; Inbarasan, Muniraj.; Role of deep learning in Optical Imaging . *Journal Not Specified* **2023**, *1*, 0.

Received:
Revised:
Accepted:
Published:

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 a wide attention as the applications of II spans over several research problems in optical engineering research areas [2–4]. For instance, biomedicine, security, autonomous vehicles, and remote sensing, to name a few [5].

Efficacy of the advanced machine learning (ML) and deep learning (DL) algorithms were shown to be producing the superior results in computer vision based applications. Thereafter, such approaches have also extended to solve several problems in various other scientific research areas. In particular, DL framework is proven to be an important tool to make an automatic decision as it solves numerous image-based problems without much human intervention. Convolution Neural Network (CNN) is a widely used DL algorithm for several problems such as image classification [6], and autonomous driving [7], etc. Furthermore, a CNN framework for 3D face recognition and classification in the photons starved environment is also 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). To note, 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 the 3D sectional images [10]. Consequently, the objects or 3D points which are located at the corresponding depth position in an imaging plane is properly overlapped and looks in focus, while the other points from different depth location does not overlap properly hence appears off-focused or defocused. The defocused points in the 3D sectional image does not convey any valuable information and are therefore redundant.

Recently, we have demonstrated a way to manually identify and remove the off-focused points from a 3D sectional image [11]. Furthermore, under some special imaging scenarios (e.g. biomedical imaging and night vision) low-light level or photons-starved illumination conditions may be encountered. In such cases, since the image capturing happens at a completely darker condition the recorded image looks degraded due to the presence of noises [8,10]. Nevertheless, this system shown to be providing a better 3D reconstruction in terms of 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, transform-based methods, rank reduction methods, and dictionary learning methods, to name a few. In addition to these, DL algorithms have also been applied for image-denoising problem [12]. In this regard, there are two ways 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 supervised manner (we use clean data as labels). We note, 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 are employed with a learning rate of 0.001.

Figure 1. Denoised results for supervised learning.

Figure. 1(c) shows the denoised 3D sectional image. We analysed the performance of the proposed method quantitatively in terms of peak signal-to-noise ratio (PSNR). For instance, the PSNR value given in Fig. 1(c) is an estimation between Fig. 1(a) and Fig. 1(c). It is evident from the Fig. 1(c) that the proposed denoising method has a better performance, in terms of PSNR. Second, we also tested the unsupervised learning for 3D image denoising. For this study, we opted a U-Net architecture [8]. This is an end-to-end fully unsupervised denoising approach where the noisy photons counted 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) are added to the skip connection strategy in U-Net architecture to avoid vanishing gradients problem. In the training process, 3D input image is given in the form of patches to the network. The patched input image is converted to 1D vector and fed as an input to the network. After removing the noises, we unpatch the 1D vector and convert back to the size of input data. In our experiments, to test the performance of the proposed method, we used two 3D objects: one is a tri-colored ball known as Object 1 in Fig. 2 (a) and second is a toy bird referred as Object 2 in Fig. 2(a). Fig. 2(a) depicts the two 3D objects used in our experiments and Fig. 2(b) and 2(c) show the clean sectional images i.e., reconstructed 3D depth images via computational approach as described in [9] without using the photon counting technique. To note, we used 20% of the PCSI patches for validation and 60% of patches are allotted for training purpose. In this work, 15 epochs were used with a learning rate of 0.001 to train the network. The PSNR values are given in Fig. 2.

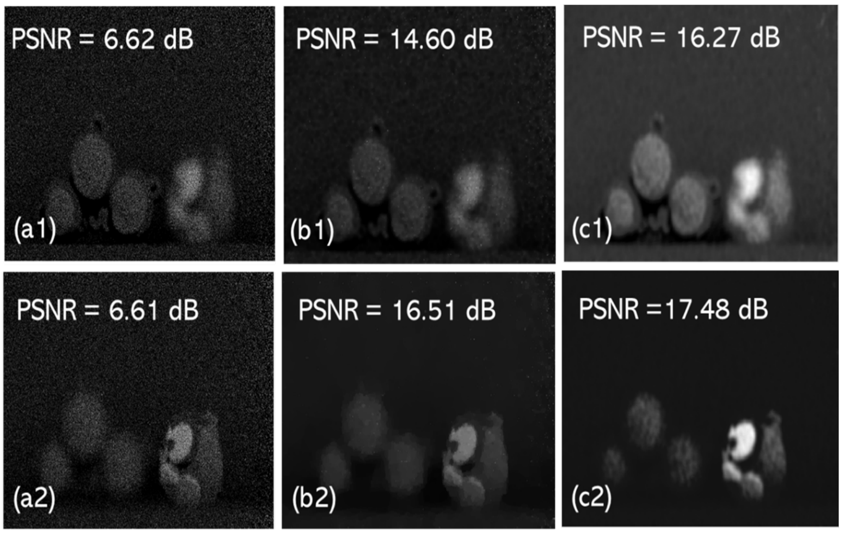

Figure 2. Denoised results: a1, b1, c1 represents noisy Photon counted 3D sectional image, TV denoised image and result of our proposed denoising method when object 1 is in focus, respectively and a2, b2, c2 represents the noisy Photon counted 3D sectional image, TV denoised image and result of our proposed denoising method when object 2 is in focus, respectively.

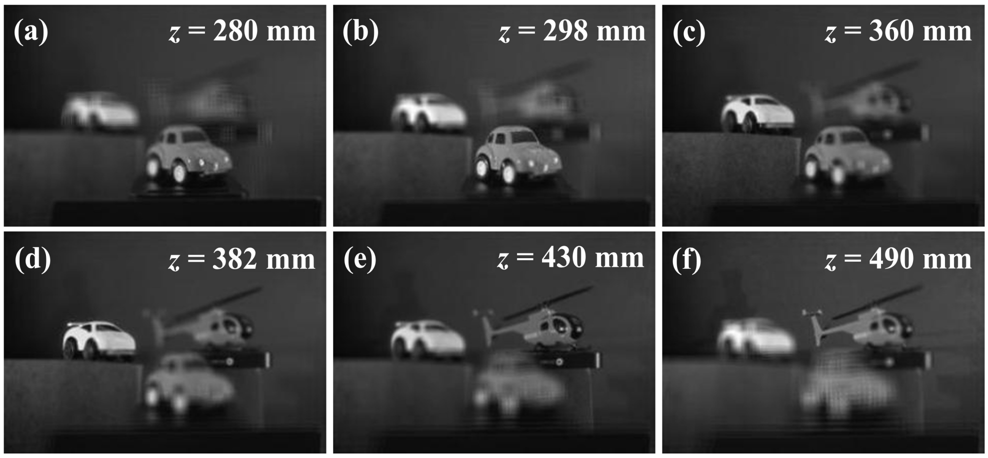

Figure 3. Reconstructed 3D CII sectional images at various depth locations.

2.2. *Off-focus removal*

Several studies have been conducted to demonstrate the feasibility of combining photons detection imaging or photons counting imaging (PCI) techqniue with the conventional 3D integral imaging systems, known as photon counted integral imaging (PCII) [2,9,10,17,18]. In such systems, it is known that the reconstructed depth images contain both the focused and off-focused (or out-of-focus) voxels, simultaneously. Off-focused pixels often look blurred and therefore do not convey acceptable information about the scene. Several approaches have been proposed to efficiently remove the off-focused points from the reconstructed 3D images[4,18]. 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.

Here, we propose a new ensembled Dense Neural Network (DNN) model, that comprises six different DNN models each trained with its own set of training dataset, for removing the off-focused points from the 3D sectional images. It is known that data pre-processing enhances the accuracy of the network, therefore, we used Otsu-thresholding algorithm [19] to remove the unwanted (and obvious) background from the 3D sectional

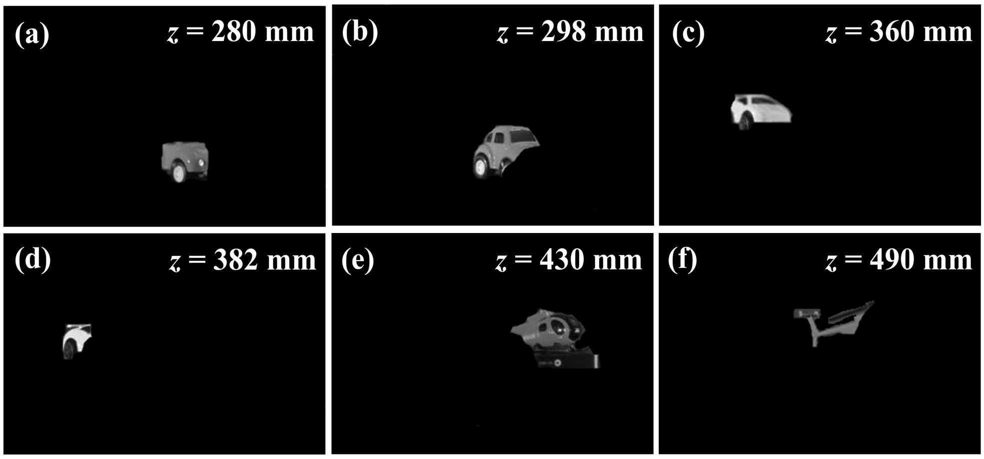

Figure 4. Reconstructed focused-only CII sectional images by using the proposed DL network.

images. In this work, we employed an ADAM optimizer to update the weights and bias [13], and a standard Mean Squared Error (MSE) was used as the cost function in our training process. To note, the proposed ensembled deep neural network is trained (supervised way) using the conventional 3D sectional images from various depth locations and the corresponding focused images (labels). We tested on the 3D scene which contains two toy cars and one toy helicopter [13]. We used Intel® Xeon® Silver 4216 CPU @2.10 GHz (2 processors) with 256 GB RAM, 64-bit operating system to simulate the all the results.

Conclusion

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

Data availability

Data for this paper are not publicly available but shall be provided upon reasonable request to the corresponding author.

Acknowledgement

VC acknowledge the support of SRM University AP research fund. IM acknowledges the financial support from Science and Engineering Research Board (SERB) under SRG/2021/001464 scheme, Department of Science and Technology, Government of India. Authors thank Mr. Suchit Patel of Poornima College of Engineering, India in lending support in the simulations and we sincerely thank Prof Bahram Javidi of University of Connecticut and Prof Moon Inkyu of DGIST, Korea for providing the dataset.

Author contributions statement

V.C., I.M. contributed equally to the manuscript preparation.

Corresponding author

Correspondence and requests should be addressed to Inbarasan Muniraj (inbarasan.m@gmail.com).

Competing interests

The authors declare no competing interests.

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 Continuum* **2018**, *1*, 373–383.
3. Yeom, S.; Javidi, B.; Watson, E. Three-dimensional distortion-tolerant object recognition using photon-counting integral imaging. *Optics Express* **2007**, *15*, 1513–1533.
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. *Applied optics* **2014**, *53*, 2777–2786.
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. *Optics Express* **2020**, *28*, 32266–32293.
6. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters* **2016**, *23*, 1499–1503.
7. Chen, C.; Seff, A.; Kornhauser, A.; Xiao, J. Deepdriving: Learning affordance for direct perception in autonomous driving. In Proceedings of the Proceedings of the IEEE international conference on computer vision, 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. *Scientific Reports* **2023**, *13*, 1365.
9. Moon, I.; Muniraj, I.; Javidi, B. 3D visualization at low light levels using multispectral photon counting integral imaging. *Journal of Display Technology* **2013**, *9*, 51–55.
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. *Optics Express* **2015**, *23*, 15907–15920.
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. *Optics Communications* **2017**, *387*, 196–201.
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. *Optics Express* **2019**, *27*, 4927–4943.
13. Dodda, V.C.; Kuruguntla, L.; Elumalai, K.; Muniraj, I.; Chinnadurai, S. An undercomplete autoencoder for denoising computational 3D sectional images. In Proceedings of the Computational Optical Sensing and Imaging. Optica Publishing Group, 2022, pp. 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. IEEE, 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 Transactions on Geoscience and Remote Sensing* **2021**.
16. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
17. Tavakoli, B.; Javidi, B.; Watson, E. Three dimensional visualization by photon counting computational integral imaging. *Optics Express* **2008**, *16*, 4426–4436.
18. 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. *Optics Communications* **2017**, *387*, 196–201.
19. Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* **1979**, *9*, 62–66.