Multi-Class Categorization of Three-Dimensional (3-D) Objects for Digital Holographic Information Using Deep Learning

Uma Mahesh R N[1] and Yogesh N[2]
[1]*Dept of CSE (AI&ML), ATME College of Engineering, Mysore, Karnataka, India*
[2]*Dept of CSD (Computer Science and Design), ATME College of Engineering, Mysore, Karnataka, India*

Keywords: Deep Learning, Receiver Operating Characteristic (ROC), 3-D Objects Categorization, Phase-Shifting Digital Holography (PSDH).

Abstract: In this paper, n-class (n=3) categorization of three-dimensional (3-D) objects using digital holographic data has been achieved with a deep learning network. For n=3 categories, the 3-D object “triangle-square” is assigned to category 1, the 3-D object “circle-square” to category 2, and the 3-D objects “triangle-circle” and “square-triangle” are grouped into category 3. The dataset, comprising phase-only images derived from digital holographic data, was generated using the phase-shifting digital holography (PSDH) technique. It includes 2880 images created through the application of a rotation invariance method. The deep learning network was trained on the dataset to generate the output. The results, including the n-class (n=3) error matrix, receiver operating characteristic (ROC), and positive predictive value (PPV)–true positive rate (TPR) characteristic are presented to validate the work.

1 INTRODUCTION

Digital holography is a three-dimensional (3D) imaging technique that captures digital holograms of 3D objects using charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) sensors. The recorded digital hologram can be numerically processed using the phase-shifting digital holography (PSDH) technique to obtain a complex-valued image containing both intensity and phase information. The phase-only digital holographic information derived from PSDH was subsequently utilized for deep learning-based applications, including categorization and prediction tasks. Deep learning, a branch of artificial intelligence, encompasses various deep neural networks, such as multi-layer perceptron (MLP), convolutional neural network (CNN), long short-term memory (LSTM) model, Alex Net, and generative adversarial networks (GANs). These networks have been applied to numerous deep learning-based digital holographic applications, including single-pixel imaging (Mizutani, Kataoka, et al. , 2024), quantitative phase imaging (Butola, Hellberg, et al. , 2024), fast particle characterization (Schneider, Dambre, et al., 2015), hologram generation (Kang, Park, et al. , 2021), and categorization and prediction of 3D objects(Basavaraju, 2024), (Reddy, Mahesh, et al. , 2022), (Mahesh, Reddy, et al. , 2022), (U. M. R N, and, K. B, 2024), (Mahesh, R.N.U., et al. , 2022), (Mahesh, R.N.U., et al. , 2023). A CNN is a deep neural network comprising multiple stages of convolutional and pooling layers for feature extraction, followed by dense and output layers in the classification stage. The feature extraction layer process the input data, and their output is passed to the classification layer to perform the n-class (n=3) categorization task. In the classification stage, the dense layer receives input from the final pooling layer and generates an intermediate output, which is then passed to the output layer to produce the final result. Categorization, a supervised machine learning technique, determines the decision boundary between the input features and the target labels. The categorization output provides discrete labels as the final result. Lam et al. (Lam, H.H., et al. , 2019) performed hologram categorization of deformable objects using a deep CNN, while Kim et al. (Kim, Wang, et al. , 2018) conducted hologram categorization of microbeads employing deep learning technique. Additionally, Pitkäaho et al. (Pitkäaho, Manninen, et al. , 2018) categorized

Mahesh R N, U. and N, Y.
Multi-Class Categorization of Three-Dimensional (3-D) Objects for Digital Holographic Information Using Deep Learning.
DOI: 10.5220/0013592800004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 2*, pages 384-388
ISBN: 978-989-758-763-4

phase-only cancer cell images using a deep learning technique. In this paper, n-class (n=3) categorization of digital holographic data for 3-D objects is performed using a deep learning network. For n=3 categories, the 3-D object "triangle-square" is assigned to category 1, "circle-square" to category 2, and "triangle-circle" and "square-triangle" are grouped into category 3. The primary distinction of this work from previous studies lies in its focus on n-class (n=3) categorization of 3-D objects using a deep learning network. The dataset, comprising phase-only images derived from digital holographic data, was generated using the PSDH technique. This dataset includes 2880 images produced through a rotation invariance method. The deep learning network was trained on this dataset to generate the output. Results, including the n-class (n=3) error matrix, receiver operating characteristic (ROC), and positive predictive value (PPV)–true positive rate (TPR) characteristic are presented to validate the effectiveness of the proposed approach.

2 DESIGN AND PRINCIPLE OF OPERATION

2.1 METHODOLOGY

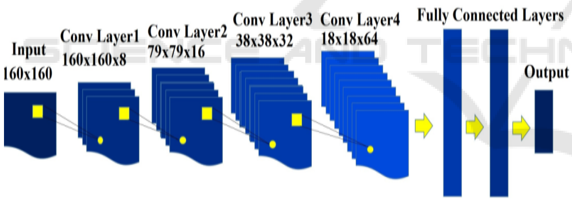


Figure 1. Architecture of Deep CNN for three-class categorization

Figure 1 shows the architecture of deep CNN used for the n-class (n=3) categorization task. The CNN takes the input as digital holographic image of size 160×160. The feature extraction layer has four consecutive convolutional, pooling layers. The convolutional layer expression is given by

$$Z_{pq}^{(n)} = f(\sum_{a=0}^{s}\sum_{b=0}^{s} h_{ab}^{(n)} X_{pq} + B_{pq}) \ldots\ldots(1)$$

In the above eqn. (1), $Z_{pq}^{(n)}$ represents the output and X_{pq} represents the input. $h_{ab}^{(n)}$ represents kernel coefficients, n represents the number of kernels, s represents the kernel size, f represents the activation function, and B_{pq} represents bias(Mahesh, Nelleri, et al. , 2023). The value of n is varied $n = 8,16,32,64$. The value of s is $s = 3 \times 3$. The activation function f represents the Rectified Linear Unit (ReLU) activation function, which is used in both convolutional and dense layers. Next, the pooling technique is employed consisting of Max-Pooling2D. The expression for pooling layer is given by

$$Z_{pq} = Y_{pq} \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

In the above eqn. (2), Z_{pq} represents the output and Y_{pq} represents the input (Mahesh, Nelleri, et al. , 2023). The output of the final pooling layer is flattened and passed to the dense layer. The expression for the dense layer is then given by

$$Z_p = f(\sum_{p=1}^{q} W_{mn} X_p + B_p) \ldots \quad \ldots(3)$$

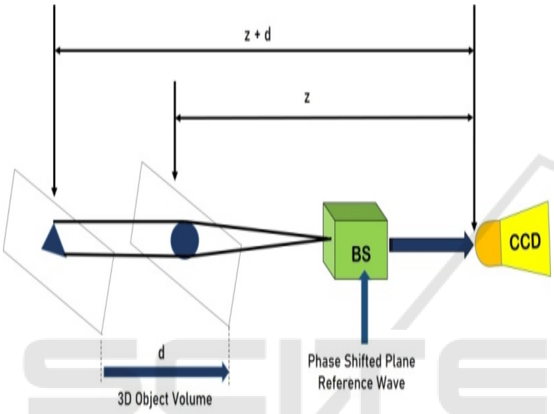
In the above eqn. (3), Z_p represents the output and B_p represents the bias, f represents the ReLU activation function, W_{mn} represents weight values, X_p represents the one dimensional (1-D) data obtained through the flatten layer, and q represents the number of neurons. The output of the final pooling layer is $8 \times 8 \times 64$. The value of q is $q = 16$. The output of the dense layer is fed into the output layer. For n-class (n=3) categorization, the output layer comprises three neurons along with a softmax activation function to produce the output. The equation for the softmax activation function is given by

$$Z_k = \frac{\exp(Y_k)}{\sum_{q=1}^{N} \exp(Y_q)} \ldots\ldots\ldots\ldots\ldots(4)$$

In the above eqn. (4), where Z_k represents the output, Y_k represents the input, and N represents the number of neurons.

3 DATASET PREPARATION WITH SIMULATION RESULTS AND DISCUSSION

For n-class (n=3) categorization, the 3D object "triangle-square" is assigned to category 1, "circle-square" to category 2, and both "triangle-circle" and "square-triangle" are grouped into category 3. The 3D object "circle-square" is designed such that the circle feature is positioned in the front plane, while the square feature is located in the back plane. Each plane is separated by various distances d_1, and d_2 respectively. The remaining three 3D objects were constructed in a similar manner, with different features positioned in the front and back planes, respectively. Four phase-shifted holograms of all four 3-D objects were formed at 0°, 90°, 180°, and 270° at the camera plane and these holograms were post-processed to obtain complex-valued image containing intensity and phase information using a four-step PSDH technique. The holograms and reconstructed intensity/phase images of all four 3-D objects are of size 1024 × 1024. The reconstructed intensity and phase images were generated at both distances. Figure 2 illustrates the schematic of the 3D object "triangle-circle" which belongs to category 3. Additionally, Figure 2 presents the geometry for digital hologram recording using four-step phase-shifted plane reference waves.

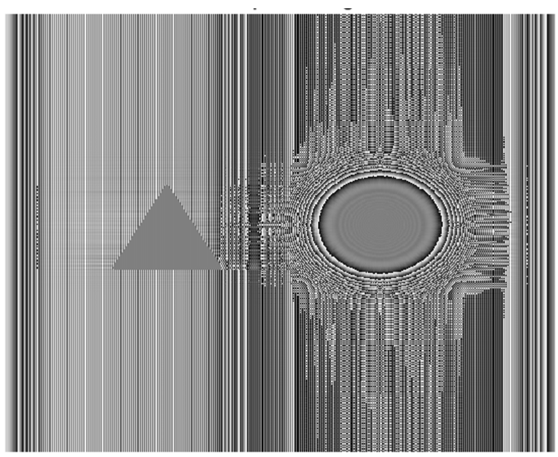


(a)

Figure 2. schematic of the geometry for the recording of the digital hologram of 3-D object volume with different features in the front and back planes and separating distances z=5 cm and d=1 cm. (a) triangle-circle. BS : beam splitter CCD : charge coupled device.

Digital holograms and the reconstructed intensity and phase images of four different 3D objects were rotated incrementally in steps of 0.5° resulting in a dataset of 2,880 images for each type. The dataset was prepared in MATLAB. For the n-class (n=3) categorization of 3D objects, only phase information was utilized. The dataset of 2,880 phase images was divided into training, validation, and test sets comprising 2,160 images (75%), 432 images (15%), and 288 images (10%) respectively. For the training of the deep learning network, the size of the phase image considered was 160 × 160 from 1024 × 1024. The deep learning network was implemented in a TensorFlow environment using python programming. A sample of a reconstructed phase image of a 3D object, specifically the "triangle-circle" belonging to category 3 is shown in Figure 3.

The deep learning network was tested on a batch of 24 images from the test set. The n-class (n=3) error matrix generated by the deep learning network is presented in Figure 4.



(a)

Figure 3. reconstructed phase-only image of 3-D object (a) triangle-circle.

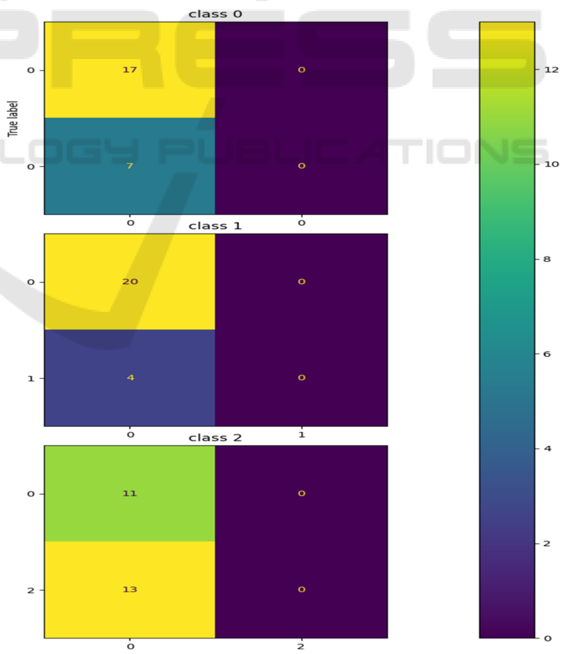
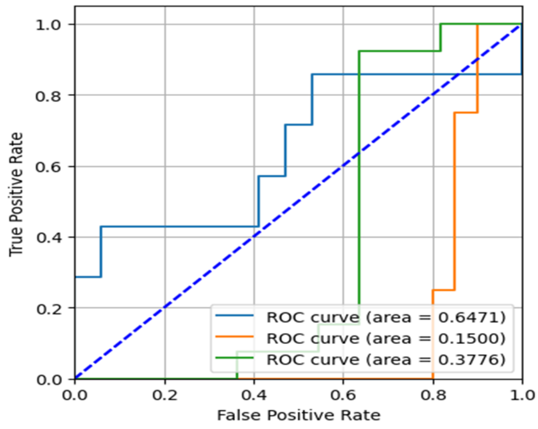
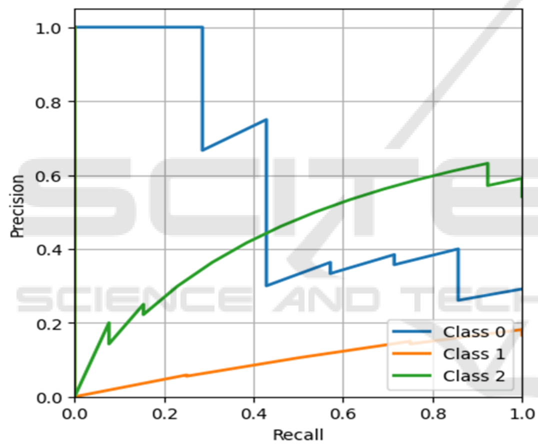


Figure 4. Three-class confusion matrix from phase-only image dataset.

From Fig. 4, it is evident that the error matrix represents the categorization results for n=3 categories. Additionally, the ROC and the PPV-TPR characteristic derived from the deep learning network are displayed in Figure 5.



(a)



(b)

Figure 5. a) receiver operating characteristic (ROC). b) positive predictive value (PPV)-true positive rate (TPR) characteristic.

From Figure 5 (a), it can be said that the deep learning network has a higher area under curve (AUC) value for category 1 compared to other categories. Similarly, from Figure 5 (b), it can be said that the deep learning network has lower PPV as the TPR approaches higher for categories 1, and 2 whereas, for category 3, the deep learning network has higher PPV compared to the other two categories.

4 CONCLUSIONS

This paper presents the n-class (n=3) categorization of three-dimensional (3D) objects using phase-only digital holographic data with a deep learning network. For the three categories, the 3D object “triangle-square” is assigned to category 1, “circle-square” to category 2, and “triangle-circle” and “square-triangle” are grouped into category 3. The dataset, consisting of phase-only images obtained from digital holographic data, was generated using the PSDH technique. It comprises 2,880 images created through a rotation invariance method. The deep learning network was trained on this dataset to produce categorization results. The results, including the n-class (n=3) error matrix, ROC, and PPV–TPR characteristic validate the approach. The error matrix reveals a higher number of images categorized as FALSE compared to TRUE for categories 1, and 2 compared to category 3. For category 3, the error matrix has higher number of images for TRUE compared to FALSE. Additionally, the ROC analysis indicates that the AUC is highest for category 1 compared to the other two categories. These findings demonstrate that deep learning network is a suitable method for n-class (n=3) categorization of 3D objects using phase-only digital holographic data.

REFERENCES

Mizutani, Y., Kataoka, S., Uenohara, T., Takaya, Y. and Matoba, O., 2024, July. “Machine learning assisted single pixel imaging for weak light detection”, In 3D Image Acquisition and Display: Technology, Perception and Applications (pp. DW3H-4). Optica Publishing Group. ttps://doi.org/10.1364/3D.2024.DW3H.4

Butola, A., Hellberg, S., Nystad, M. and Agarwal, K., 2024, July. “Quantitative phase imaging and machine learning for spermatozoa analysis”, In Imaging Systems and Applications (pp. JM4A-6). Optica Publishing Group. https://doi.org/10.1364/3D.2024.JM4A.6

Schneider, B., Dambre, J. and Bienstman, P., 2015. “Fast particle characterization using digital holography and neural networks”, Applied optics, 55(1), pp.133-139. https://doi.org/10.1364/AO.55.000133

Kang, J.W., Park, B.S., Kim, J.K., Kim, D.W. and Seo, Y.H., 2021. “Deep-learning-based hologram generation using a generative model”, Applied Optics, 60(24), pp.7391-7399. https://doi.org/10.1364/AO.427262

Mahesh R N, U.; Nelleri, A. “Multi-Class Classification and Multi-Output Regression of Three-Dimensional Objects Using Artificial Intelligence Applied to Digital

Holographic Information", Sensors 2023, 23, 1095. https://doi.org/10.3390/s23031095

RN UM, Basavaraju L. Deep Learning-based Multi-class Three-dimensional (3-D) Object Classification using Phase-only Digital Holographic Information. IgMin Res. Jul 09, 2024; 2(7): 550-557. IgMin ID: igmin216; DOI:10.61927/igmin216; Available at: igmin.link/p216

Reddy, B.L., Uma Mahesh, R.N. and Nelleri, A., 2022. "Deep convolutional neural network for three-dimensional objects classification using off-axis digital Fresnel holography", Journal of Modern Optics, 69(13), pp.705-717. https://doi.org/10.1080/09500340.2022.2081371

Uma Mahesh, R.N., Lokesh Reddy, B., Nelleri, A. (2022). Deep Learning-Based Multi-class 3D Objects Classification Using Digital Holographic Complex Images. In: Sivasubramanian, A., Shastry, P.N., Hong, P.C. (eds) Futuristic Communication and Network Technologies. VICFCNT 2020. Lecture Notes in Electrical Engineering, vol 792. Springer, Singapore. https://doi.org/10.1007/978-981-16-4625-6_43

U. M. R N and K. B, "Three-dimensional (3-D) objects classification by means of phase-only digital holographic information using Alex Network", 2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 2024, pp. 1-5, doi: 10.1109/IConSCEPT61884.2024.10627906.

Mahesh, R.N.U., Nelleri, A. "Deep convolutional neural network for binary regression of three-dimensional objects using information retrieved from digital Fresnel holograms", Appl. Phys. B 128, 157 (2022). https://doi.org/10.1007/s00340-022-07877-w

Mahesh, R.N.U., Nelleri, A. (2023). "Machine Learning-Based Binary Regression Task of 3D Objects in Digital Holography", In: Subhashini, N., Ezra, M.A.G., Liaw, SK. (eds) Futuristic Communication and Network Technologies. VICFCNT 2021. Lecture Notes in Electrical Engineering, vol 995. Springer, Singapore. https://doi.org/10.1007/978-981-19-9748-8_34

Lam, H.H., Tsang, P.W.M. and Poon, T.C., 2019. "Ensemble convolutional neural network for classifying holograms of deformable objects", Optics Express, 27(23), pp.34050-34055. https://doi.org/10.1364/OE.27.034050

Kim, S.J., Wang, C., Zhao, B., Im, H., Min, J., Choi, H.J., Tadros, J., Choi, N.R., Castro, C.M., Weissleder, R. and Lee, H., 2018. "Deep transfer learning-based hologram classification for molecular diagnostics", Scientific reports, 8(1), p.17003. doi: 10.1038/s41598-018-35274-x.

Pitkäaho, T., Manninen, A. and Naughton, T.J., 2018, June. "Classification of digital holograms with deep learning and hand-crafted features", In Digital Holography and Three-Dimensional Imaging (pp. DW2F-3). Optica Publishing Group. https://doi.org/10.1364/DH.2018.DW2F.3

Trieu, Q. and Nehmetallah, G., 2024. "Deep learning based coherence holography reconstruction of 3D objects", Applied Optics, 63(7), pp.B1-B15. https://doi.org/10.1364/AO.503034

Tahara, T., 2024. "Incoherent digital holography with two polarization-sensitive phase-only spatial light modulators and reduced number of exposures", Applied Optics, 63(7), pp.B24-B31. https://doi.org/10.1364/AO.505624

Störk, T., Seyler, T., Fratz, M., Bertz, A., Hensel, S. and Carl, D., 2024. "Detecting vibrations in digital holographic multiwavelength measurements using deep learning", Applied Optics, 63(7), pp.B32-B41. https://doi.org/10.1364/AO.507303