Deep transfer learning-based hologram classification for molecular diagnostics

Sung-Jin Kim[1], Chuangqi Wang[1], Bing Zhao[1], Hyungsoon Im[2,3], Jouha Min[2,3], Nu Ri Choi[1], Cesar M. Castro[2], Ralph Weissleder[2,3,4], Hakho Lee[2,4], Kwonmoo Lee[1]

[1]Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts

[2]Center for Systems Biology, Massachusetts General Hospital, Boston, Massachusetts

[3]Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts

[4]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts

Correspondence

Kwonmoo Lee (klee@wpi.edu), Hakho Lee (hlee@mgh.harvard.edu)

ABSTRACT

Lens-free digital in-line holography (LDIH) is a promising microscopic tool that overcomes several drawbacks (e.g., limited field of view) of traditional lens-based microcopy. However, extensive computation is required to reconstruct object images from the complex diffraction patterns produced by LDIH, which limits LDIH utility for point-of-care applications, particularly in resource limited settings. Here, we describe a deep transfer learning (DTL) based approach to process LDIH images in the context of cellular analyses. Specifically, we captured holograms of cells labeled with molecular-specific microbeads and trained neural networks to classify these holograms without reconstruction. Using raw holograms as input, the trained networks were able to classify individual cells according to the number of cell-bound microbeads. The DTL-based approach including a VGG19 pretrained network showed robust performance even with noisy experimental data. Combined with the developed DTL approach, LDIH could be realized as a low-cost, portable tool for point-of-care diagnostics.

INTRODUCTION

Lens-free digital in-line holography (LDIH) is a powerful imaging platform that overcomes many of the limitations of traditional microscopy[1-6]. LDIH records diffraction patterns produced by samples, which can later be used to computationally reconstruct original object images. This strategy enables LDIH to image a large area (~mm^2) while achieving a high spatial resolution (~μm). Furthermore, the simplistic optical design allows for compact setups, consisting of a semiconductor imager chip and a coherent light source. LDIH has been previously tested for potential point-of-care (POC) diagnoses[7]. Recently, we have advanced LDIH for the purpose of molecular diagnostics (D3, digital diffraction diagnostics)[3] wherein cancer cells were labeled with antibody-coated-microbeads, and bead-bound cells were counted for molecular profiling.

A major hurdle to translating LDIH into POC tests is the need for extensive computational power. In principle, diffraction patterns can be back-propagated to reconstruct human-friendly object images. The bottleneck lies in the recovery of phase information, lost during the imaging process. It has been shown that this information can be numerically recovered through iterative optimization[1,8-13], but the process is costly in computation time and requires high-end resources (e.g., graphical processing unit). To overcome this issue, we demonstrated that a deep neural network could be trained to recover phase information and reconstruct object images, substantially reducing the total computational time[14]. However, this method still required an input of back-propagation images obtained from the holograms. In this paper, we explored an alternative approach in which diagnostic information could be extracted from the raw hologram images without the need for hologram reconstruction. In the microbead-based assay, we reasoned that cell-bead objects could generate distinct hologram patterns, albeit imperceptible to human

eyes, recognizable by machine vision classifiers. Developing such a capacity would eliminate the need for image reconstruction, further advancing LDIH utility for POC operations.

We here report on new machine-learning (ML) based approaches for LDIH image analysis. ML has been making significant progress in extracting information from complex biomedical images and started to outperform human experts for many data sets[15-18]. In this paper, we compared three different ML schemes: the support vector machine (SVM)[19], convolutional neural networks (CNN)[20-22] and deep transfer learning (DTL)[23-28] . SVM has been known to perform well with a small dataset if an appropriate feature extraction for given dataset is provided, while CNN can outperform SVM without *a priori* feature extraction once the size of the dataset is large[25]. DTL extracts feature information from input data using the convolution part of pre-trained networks and subsequently feeds the information to classifiers. It has been known that pretrained networks can be exploited as a general-purpose feature extractor[24]. In this DTL approach, we used a VGG19[29] model that was pretrained with a large number of ordinary images (i.e., not holograms) available in the ImageNet[30], and fine-tuned the classifier to obtain high-performance classification. We applied all three schemes to classify holograms generated from cells and microbeads without a reconstruction process. Specifically, algorithms were developed to i) recognize the holograms of cells labeled with microbeads and ii) classify the cells according to the number of attached beads. We found that a DTL approach offered highly reliable performance in hologram classification. When applied to experimental holograms, the DTL algorithm achieved good accuracy (92.8%), allowing for reconstruction-free image classification.

RESULTS

System and assay setup

Figure 1A shows the schematic of LDIH system[3]. As a light source, we used a light-emitting diode (LED; λ = 420 nm). The light passes through a circular aperture (diameter, 100 μm), generating a coherent spherical wave on the sample plane. The incidence light and the scattered light from the sample interfere with each other to generate holograms which are then recorded by a CMOS imager[10,31]. The system has a unit (×1) optical magnification, resulting in a field-of-view equal to the imager size.

To enable molecular-specific cell detection, we used antibody-coated microbeads (diameter, 6 μm) for cell labeling. The number of attached beads is proportional to the expression level of a target marker, allowing for quantitative molecular profiling[3]. Diffraction patterns from unlabeled and bead-bound cells have subtle differences that are hard to detect with human eyes (**Fig. 1B**). Only after image reconstruction can beads and cells be differentiated and counted; cells have high amplitude and phase values, whereas microbeads have negligible phase values.

Reconstruction-free ML approaches

Conventional LDIH reconstruction (**Fig. 2A**) requires multiple repetitions of back-propagation, constraint application, and transformation[8]. This iterative algorithm is computationally intensive, either incurring long processing time or requiring high-end resources (e.g., a high-performance graphical processing unit server) for faster results[3]. Furthermore, human curation is occasionally needed to correct for stray reconstruction (e.g., debris, twin images). In contrast, our ML-based approach is a reconstruction-free classification method (**Fig. 2B**). As an off-line task, we first trained a network using annotated holograms of bead-bound cells. After the training was complete, the network was used for on-line classification tasks; holograms, without any image preprocessing, were entered as an input.

To develop the optimal classifier, we compared three different models (see Methods for detail). The first one (PCA-SVM; **Fig. 3A**) was a conventional machine learning model. The second model (**Fig. 3B**) was a basic convolutional neural network (CNN)[20] consisting of two convolutional layers and one fully-connected layer (**Fig. 3B**). The third and fourth models (VGG19-FC and VGG19-PCA-SVM; **Fig. 3C, D**) incorporated a pre-trained network (VGG19)[29] with a fully connected layer or PCA-SVM classifications. The VGG19-PCA-SVM used the SVM classifier with PCA preprocessing, where the optimal number of principal components was determined based on the classifier performance.

To make training sets, we prepared holograms and corresponding reconstruction images of bead-bound cells. In each image pair, individual cells were annotated according to the number of beads attached: class 1 (0 or 1 bead) and class 2 (> 2 beads). Cells with more than 2 bead attachments are considered positive for a given target biomarker[3]. For the robust performance evaluation in classifying experimental data, we performed multiple 5-fold cross-validation processes[32] followed by statistical testing (see Methods for detail).

Testing with synthetic hologram images

We first tested the feasibility of the reconstruction-free classification, using numerically generated data sets (**Fig. 4A**, see Methods for detail**).** We trained each classifying network (**Fig. 3**) with either 320 object images or 320 synthetic holograms. Additional 80 images of each type (object images or holograms) were used for performance validation. With the synthetic test data, all models achieved 100% accuracy (**Fig. 5A**) except SVM (no PCA). For the VGG19-FC model, overfitting was minimal in the case of object image training and hologram training (**Figs. 5B, C**). Notably, the VGG19 pre-trained model, which was trained only with public ImageNet datasets[29], was capable of extracting key features from holograms, making the classification

effective. Overall, the results confirmed that i) bead-bound cells can be classified directly from holograms, and ii) classification of holograms can be as accurate as with object images.

We also considered how the variation of the optical distance z (the distance between samples and the imager) affects the network performance. In real experiment settings, precisely controlling the optical distance is challenging; errors can be introduced due to the thickness variation or misalignment of sample slides. In the conventional hologram reconstruction, using incorrect optical distance leads to blurred, un-focused object images[8]. We tested whether the reconstruction-free classification is robust to such perturbations. We varied the optical distance, z according to $z = z_0 + n \cdot \sigma$, where z_0 is the nominal value (600 μm) used in the experiment, σ is the standard deviation, and n is a random number from the standard normal distribution [$N(0,1)$]. For a given σ value, we generated a new set of synthetic holograms; Due to the random noise addition, each synthetic hologram had a different z value. **Figure 6A** shows two example sets with σ = 120 and 240 μm. We then trained the VGG19-FC network for each set of synthetic holograms generating using various σ values. The accuracy and the loss performance of VGG19-FC with respect to σ are shown in **Fig. 6B.** We observed no significant performance degradation of the classification up to σ = 60 μm (10% of the nominal value z_0). Since the optical distance variation in our experimental setup[33] was recorded to be about 12 μm (2%), these results suggest that the reconstruction-free classification is robust to the optical distance variation in our experimental condition.

Classification with experimental data

We next prepared the training set from the experimental images **(Fig. 7A)** taken by LDIH system[3] (see Methods for detail). We grouped these training images into two classes: cells with

<2 beads and cells with ≥2 beads for holograms (**Fig. 7B**) and object images (**Fig. 7C**). Compared to the synthetic dataset, these images included noisy objects such as unbound beads and other surrounding cells.

We then trained each classifier separately for holograms or object images. **Figure 8A** summarizes the results. Compared to the synthetic image cases, the overall accuracy of SVM, PCA-SVM and CNN classifiers substantially decreased, presumably due to real-world variations (i.e., inherent noise) in both object and hologram datasets. In contrast, VGG19-FC, VGG19-SVM, VGG19-PCA-SVM maintained high accuracy and their accuracy was significantly larger in both datasets (**Fig. 8B-D**). Also, all VGG19 based classifiers perform similarly for the object images and hologram, confirming the feasibility of reconstruction-free classification. SVM is usually considered to be highly effective in small datasets[34]. SVM and PCA-SVM achieved only moderate accuracy in holograms (62.1% and 72.7%) using the experimental dataset, even if PCA-SVM achieved 100% accuracy using the synthetic dataset (**Fig. 5A)**. However, VGG19-SVM and VGG19-PCA-SVM achieved 92.8 and 92.5% accuracy in the holograms. The better performance of these approaches can be interpreted as follows. VGG19 extracted highly meaningful features from the holograms even though VGG19 was trained using natural images, suggesting that transfer learning from natural images to hologram was highly effective. Consistently, VGG19-FC also achieve the similar accuracy in both datasets. To validate this reasoning, we inspected outputs of VGG19 and VGG19-PCA by generating t-SNE plots[35]. In both training sets (holograms, object images), each class of bead-bound cells was visually more segregated in VGG19 and VGG19-PCA outputs (**Fig. 9**), which helps efficient downstream classification. It is conceivable that VGG19, pre-trained with a large number of natural images, is robust against noise and artifacts. Taken together, these results demonstrate the feasibility of

hologram classification without reconstruction in the experimental images, simplifying the workflow and decreasing computational cost.

DISCUSSION

We have demonstrated that ML approaches can effectively classify holograms without reconstructing their original object images. The conventional reconstruction requires high computational complexity due to iterative phase recovery steps. Our ML approach offers an appealing new direction to further advance the utility of LDIH: i) once trained, deep learning-based classification can be executed at the local device level without complex computation; ii) not relying on high resolution images, the classification network is robust to experimental noises; and iii) the network is elastic and can be continuously updated for higher accuracy. With these merits, we envision that the developed ML networks will significantly empower LDIH, realizing a truly POC diagnostic platform.

METHODS

ML architecture

PCA-SVM (**Fig. 3A**) was a conventional machine learning model. We used principal component analysis (PCA)[36] for the dimensional reduction of imaging data. Two principal components were obtained from each image, and used as inputs to support vector machine (SVM) with the radial basis function kernel[19,37]. While the optimal best number of the principal components was found to be two by searching the numbers sequentially, we additionally test SVM with all principal components instead of the best two components, to show the base-line performance of the SVM

classification. The optimal number of the principal components was found to be two. The second model (**Fig. 3B**) was a basic convolutional neural network (CNN)[20] consisting of two convolutional layers and one fully-connected layer (**Fig. 3B**). After two connective convolution processes with a 3×3 convolutional filter, a 2×2 max pooling was conducted with the dropout rate $P_d = 0.5$ [38]. Then a 128-node fully-connected layer was applied ($P_d = 0.5$), followed by a 2-node output layer that produced a classification output with softmax activation function. Convolution layers and the 128-node layer included ReLU (Rectified Linear Unit) activation function[39].

The third and fourth models (VGG19-FC and VGG19-PCA-SVM; **Fig. 3C, D**) incorporated a pre-trained network (VGG19)[29] with a fully connected layer or PCA-SVM classifications. We used the original VGG19 from Keras Python Package[40], to extract features from the hologram images. After taking feature information from the final convolutional layer of the VGG19 model, the two classification approaches (**Fig. 3C, D**) were executed. The VGG19-FC approach used a fully connected layer of 64 nodes with 70% dropout rate, and a ReLU activation function. The VGG19-PCA-SVM used the SVM classifier with PCA preprocessing, where the optimal number of principal components was determined based on the classifier performance.

Synthetic training dataset

The data generator (**Fig. 4A**) took two partial images from an experimental image, one containing a cell and the other a microbead; each image had both magnitude and phase information (**Fig. 4B**) obtained from hologram reconstruction. These inputs were then combined; the number of beads and their relative position on the cell were changed, mimicking cell labeling with microbeads. Finally, the combined images were transformed into holograms using the Fresnel diffraction integral formula[41] (**Fig. 4C**) given by

$$E(x,y,z) = \frac{1}{i\lambda}\int\int_{-\infty}^{+\infty} E(x',y',0)\frac{e^{ikr}}{r}dx'dy' \quad (1)$$

where (x, y, z) and (x, y, 0) represent geometrical positions of a hologram and an object image, respectively, E represents images, k is a wave number, λ is a wavelength. The z represents the optical distance between the object and the imaging planes and $r = \sqrt{(x-x')^2 + (y-y')^2 + z^2}$. In our experiment, λ and z are 405 nm and 600 μm, respectively. The absolute values from Eq. 1 were used for synthesizing hologram images.

Experimental training dataset

We next evaluated the networks using experimental images taken by LDIH system[3]. Samples were prepared by labeling cancer cells (SkBr3, breast carcinoma) with polystyrene beads (diameter, 6 μm) conjugated with EpCAM antibodies. Labeled cells, suspended in buffer, were loaded on a microscope slide, and their holograms were imaged (**Fig. 7A, left**). To prepare the training set for classification, we reconstructed object images from holograms (**Fig. 7A, right**), manually annotated cells with the number of bound beads, and prepared single-cell images with individual cells at the image center (270 × 270 pixels). We grouped these training images into two classes: cells with <2 beads and cells with ≥2 beads. We obtained two training sets, one for holograms (**Fig. 7B**) and the other for object images (**Fig. 7C**). Each set had 179 images.

Performance evaluation of Each ML approach

For the robust performance evaluation in classifying experimental data, we applied a 5-fold cross-validation process[32]. In the process, about 80% of data samples were iteratively used to train the network for classification. Following training, the remaining 20% of data samples were used as a testing set to measure the classification performance. Then, this cross-validation

process was repeated ten times. The differences in the accuracy of the ML approaches in this study were tested using unpaired two-tailed Wilcoxon rank sum test.

Code availability statement

The code used in the current study is available from the corresponding author upon reasonable request.

Data availability statement

The datasets used in the current study are available from the corresponding author on reasonable request.

Acknowledgments

We thank NVIDIA for providing us with TITAN X GPU cards (NVIDIA Hardware Grant Program, K.L.) and Microsoft for Azure cloud computing resources (Microsoft Azure Research Award, K.L.). This work is supported by the WPI Start-up Fund for new faculty. The authors were supported in part by the WPI Start-up Fund for new faculty (K.L); a generous gift by Boston Scientific (K.L); NIH grants R21-CA205322 (H.L.), R01-HL113156 (H.L.), R01-CA204019 (R.W.), R01-EB010011 (R.W.), R01-EB00462605A1 (R.W.), K99-CA201248-02 (H. I.); Liz Tilberis Award - Ovarian Cancer Research Fund (C.M.C); the Lustgarten Foundation (R.W.); and MGH Scholar Fund (H.L.).

Author Contributions

S.K. initiated the project, designed and trained the classifiers, coordinated the collaboration as a research scientist at WPI and non-employee research personnel at MGH, and wrote the final version of the manuscript; C.W. designed and trained the classifiers; B.Z performed training of CNNs. H.I. and J.M set up the imaging system and generated the hologram data; C.M.C and

R.W coordinated the experiments with cancer cells; N.C. prepared the training set; K.L. and H.L. coordinated the study and wrote the final version of the manuscript and supplement. All authors discussed the results of the study.

Competing Interests

The authors declare no competing interests.

Author Information

Correspondence and requests for materials, code and data should be addressed to K.L. (klee@wpi.edu) and H. L. (hlee@mgh.harvard.edu)

REFERENCES

1 Garcia-Sucerquia, J. *et al.* Digital in-line holographic microscopy. *Appl Opt* **45**, 836-850 (2006).

2 Greenbaum, A. *et al.* Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy. *Nat Methods* **9**, 889-895, doi:10.1038/nmeth.2114 (2012).

3 Im, H. *et al.* Digital diffraction analysis enables low-cost molecular diagnostics on a smartphone. *Proc Natl Acad Sci U S A* **112**, 5613-5618, doi:10.1073/pnas.1501815112 (2015).

4 Xu, W., Jericho, M. H., Meinertzhagen, I. A. & Kreuzer, H. J. Digital in-line holography for biological applications. *Proc Natl Acad Sci U S A* **98**, 11301-11305, doi:10.1073/pnas.191361398 (2001).

5 Gurkan, U. A. *et al.* Miniaturized lensless imaging systems for cell and microorganism visualization in point-of-care testing. *Biotechnol J* **6**, 138-149, doi:10.1002/biot.201000427 (2011).

6 Greenbaum, A. *et al.* Increased space-bandwidth product in pixel super-resolved lensfree on-chip microscopy. *Sci. Rep.* **3**, 1717, doi:10.1038/srep01717 (2013).

7 Zhu, H., Isikman, S. O., Mudanyali, O., Greenbaum, A. & Ozcan, A. Optical imaging techniques for point-of-care diagnostics. *Lab Chip* **13**, 51-67, doi:10.1039/c2lc40864c (2013).

8 Fienup, J. Phase retrieval algorithms: a comparison. *Appl Opt* **21**, 2758-2769, doi:10.1364/AO.21.002758 (1982).

9 Mudanyali, O., Oztoprak, C., Tseng, D., Erlinger, A. & Ozcan, A. Detection of waterborne parasites using field-portable and cost-effective lensfree microscopy. *Lab Chip* **10**, 2419-2423, doi:10.1039/c004829a (2010).

10 Mudanyali, O. *et al.* Compact, light-weight and cost-effective microscope based on lensless incoherent holography for telemedicine applications. *Lab Chip* **10**, 1417-1428, doi:10.1039/c000453g (2010).

11 Gerchberg, R. & Saxton, W. A practical algorithm for the determination of phase from image and diffraction plane pictures. *SPIE milestone series MS* **93**, 306-306 (1994).

12 Fienup, J. R. Reconstruction of an object from the modulus of its Fourier transform. *Optics letters* **3**, 27-29 (1978).

13 Latychevskaia, T. & Fink, H.-W. Solution to the twin image problem in holography. *Physical review letters* **98**, 233901 (2007).

14 Rivenson, Y., Zhang, Y., Gunaydin, H., Teng, D. & Ozcan, A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light: Science & Applications*, doi:10.1038/lsa.2017.141 (2017).

15 Shen, D., Wu, G. & Suk, H.-I. Deep learning in medical image analysis. *Annual review of biomedical engineering* **19**, 221-248 (2017).

16 Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118, doi:10.1038/nature21056 (2017).

17 Gulshan, V. *et al.* Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **316**, 2402-2410, doi:10.1001/jama.2016.17216 (2016).

18 Ehteshami Bejnordi, B. *et al.* Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **318**, 2199-2210, doi:10.1001/jama.2017.14585 (2017).

19 Cortes, C. & Vapnik, V. Support-vector networks. *Machine learning* **20**, 273-297 (1995).

20 LeCun, Y. *et al.* 396-404.

21 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444, doi:10.1038/nature14539 (2015).

22 Krizhevsky, A., Sutskever, I. & Hinton, G. E. in *Advances in neural information processing systems.* 1097-1105.

23 Pratt, L. Y. in *Advances in neural information processing systems.* 204-211.

24 Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. in *Advances in neural information processing systems.* 3320-3328.

25 Razavian, A. S., Azizpour, H., Sullivan, J. & Carlsson, S. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on.* 512-519 (IEEE).

26 Donahue, J. *et al.* in *International conference on machine learning.* 647-655.

27 Oquab, M., Bottou, L., Laptev, I. & Sivic, J. in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on.* 1717-1724 (IEEE).

28 Zeiler, M. D. & Fergus, R. in *European conference on computer vision.* 818-833 (Springer).

29 Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations* (2015).

30 Deng, J. *et al.* in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* 248-255 (IEEE).

31 Fung, J. *et al.* Measuring translational, rotational, and vibrational dynamics in colloids with digital holographic microscopy. *Opt Express* **19**, 8051-8065, doi:10.1364/OE.19.008051 (2011).

32 Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society. Series B (Methodological)*, 111-147 (1974).

33 Pathania, D. *et al.* Holographic Assessment of Lymphoma Tissue (HALT) for Global Oncology Field Applications. *Theranostics* **6**, 1603-1610, doi:10.7150/thno.15534 (2016).

34 Sakr, G. E., Mokbel, M., Darwich, A., Khneisser, M. N. & Hadi, A. in *Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International.* 207-212 (IEEE).

35 Maaten, L. v. d. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9**, 2579-2605 (2008).

36 Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**, 559-572, doi:10.1080/14786440109462720 (1901).

37 Vert, J.-P., Tsuda, K. & Schölkopf, B. A primer on kernel methods. *Kernel methods in computational biology* **47**, 35-70 (2004).

38 Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* **15**, 1929-1958 (2014).

39 Nair, V. & Hinton, G. E. in *Proceedings of the 27th international conference on machine learning (ICML-10).* 807-814.

40 Chollet, F. & others. Keras. *GitHub* (2015).

41 Oberst, H., Kouznetsov, D., Shimizu, K., Fujita, J. & Shimizu, F. Fresnel diffraction mirror for an atomic wave. *Phys Rev Lett* **94**, 013203, doi:10.1103/PhysRevLett.94.013203 (2005).

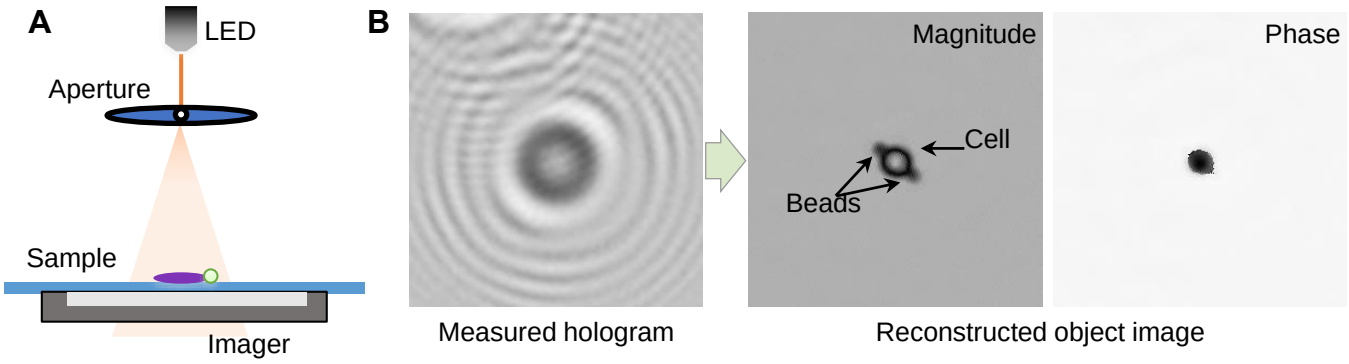


Figure 1. In-line holographic imaging. (A) A holography system includes LED, a sample glass and a sensor where a light is passed to a sample through a pinhole disk. **(B)** A hologram image and its associated reconstructed images consisting of magnitude and phase images.

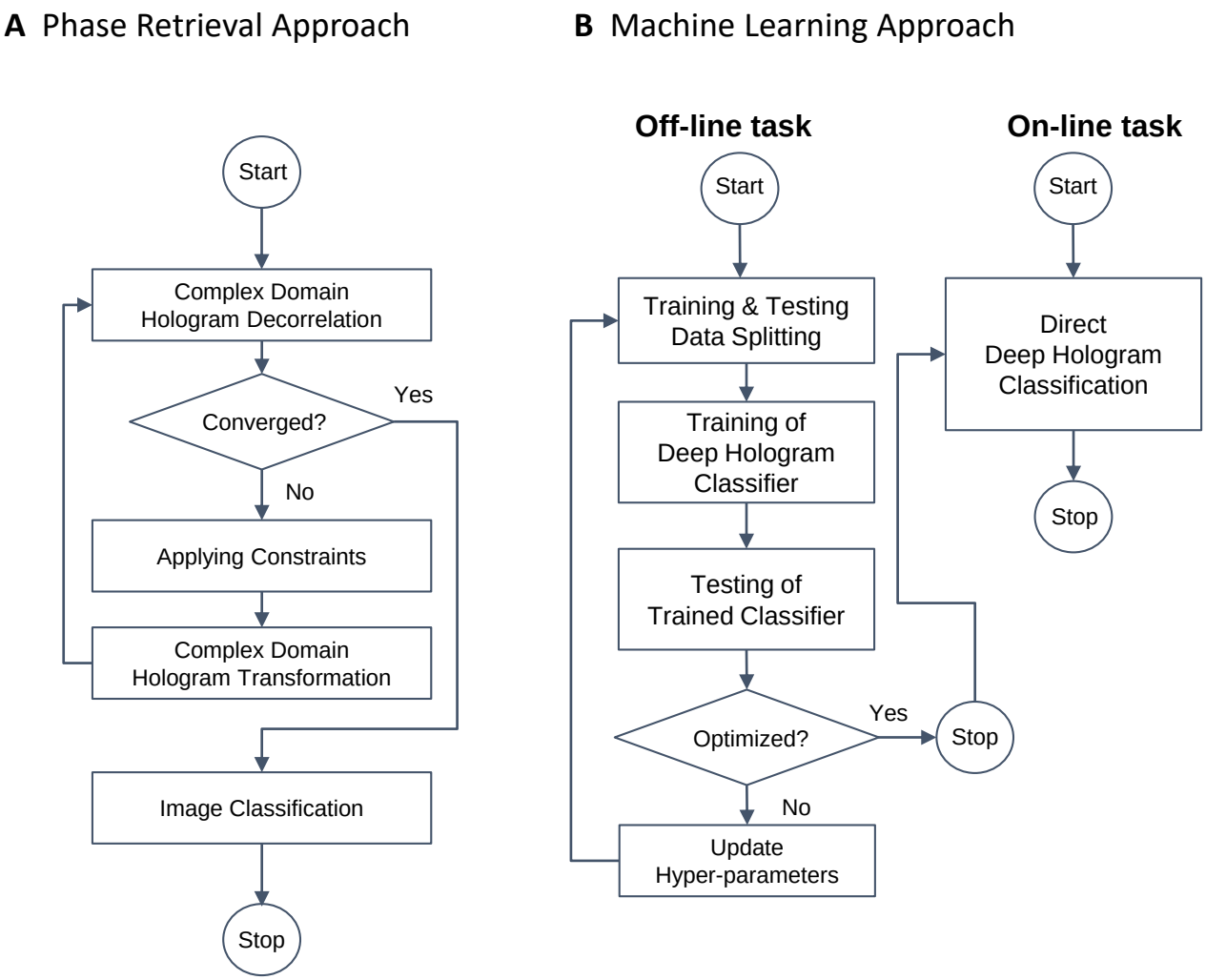


Figure 2. Flow charts of holographic classification approaches. **(A)** The conventional approach includes an iterative reconstruction process before image classification. **(B)** The machine learning based approach performs a training process before the classification stage.

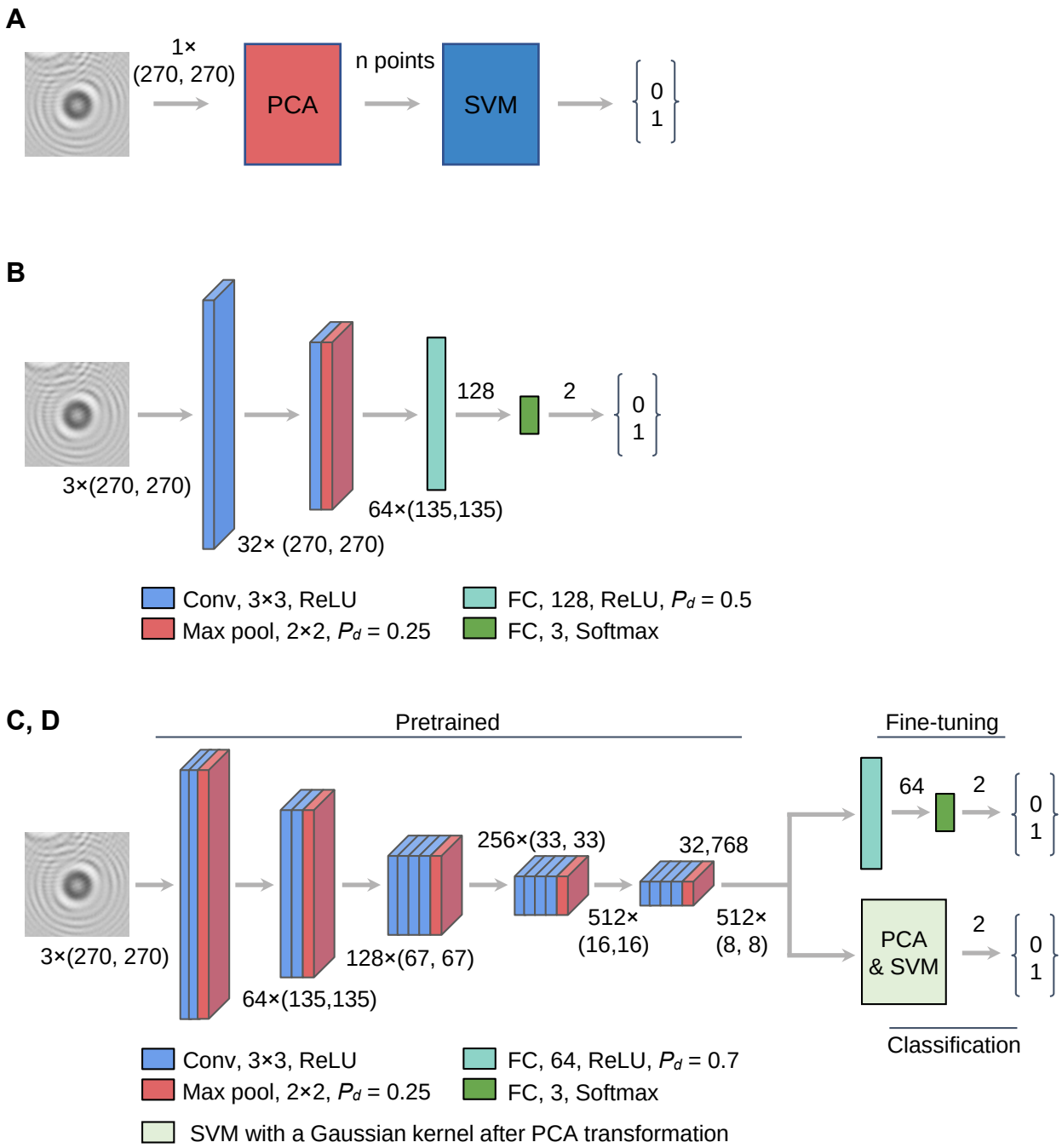


Figure 3. Machine learning approaches in this study. (**A**) The conventional machine learning approach consisting of PCA and SVM (PCA-SVM). (**B**) A basic convolutional network approach (CNN). (C-D) VGG19 pretrained model-based approaches with a fully connected layer and PCA & SVM classifier, respectively

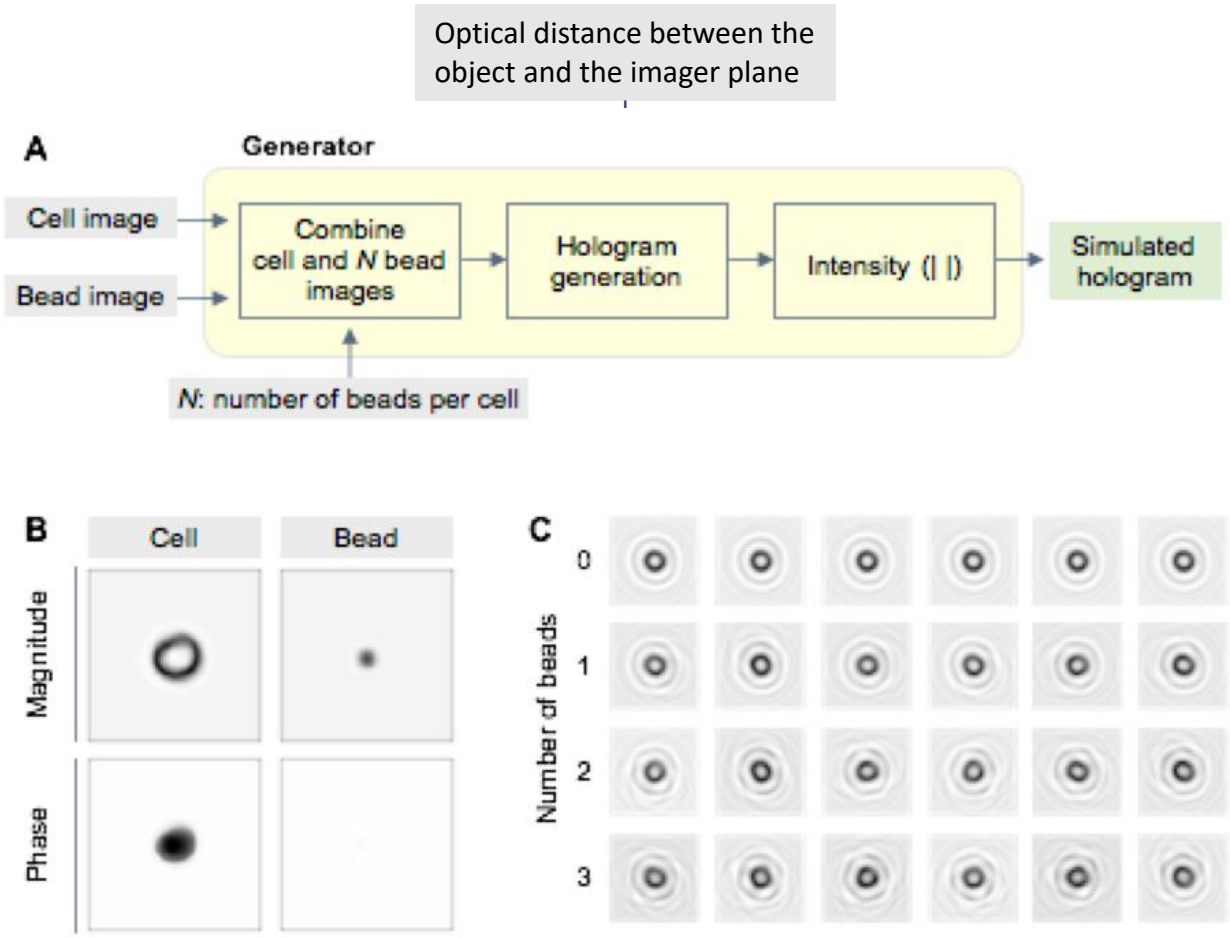
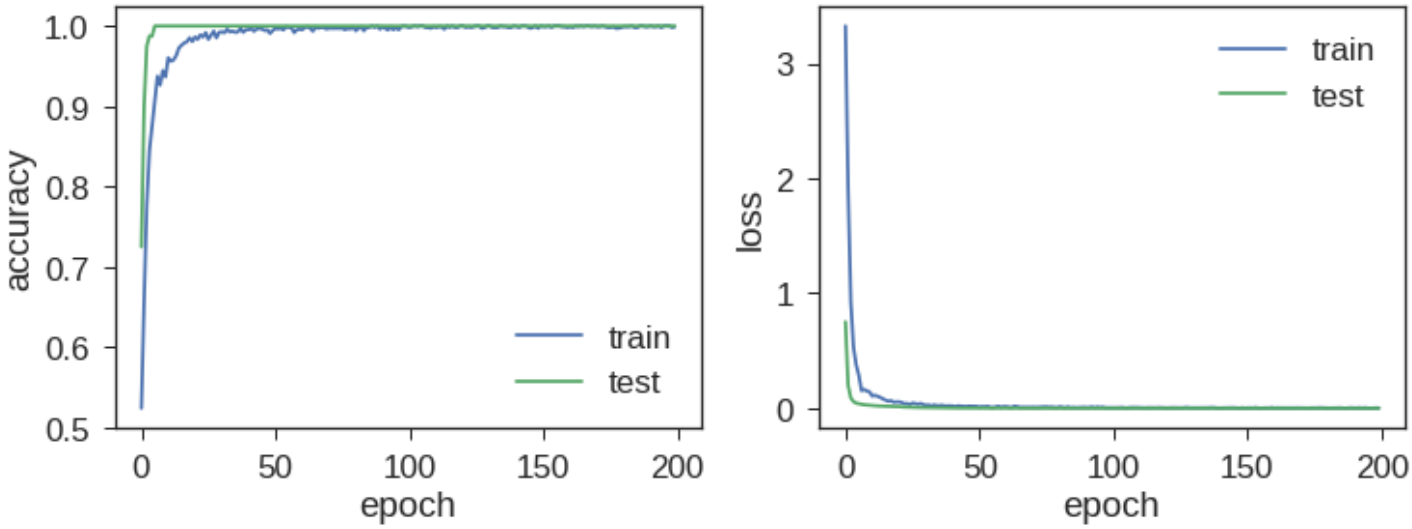


Figure 4. Synthesized holographic dataset. (A) Data synthesis workflow. **(B)** Experimental cell and bead images for synthesizing the data. **(C)** Synthesized hologram images.

A

	Object Image		Hologram	
	Parameters	Accuracy	Parameters	Accuracy
SVM	C=1, g=10^{-1}	81.3%	C=1, g=10^{-1}	84.8%
PCA-SVM	C=10, g=10^{-3}, n=2	100%	C=10, g=10^{-3}, n=2	100%
CNN	-	100%	-	100%
VGG19-FC	N=64, p=0.7	100%	N=64, p=0.7	100%

B Object images



C Hologram Images

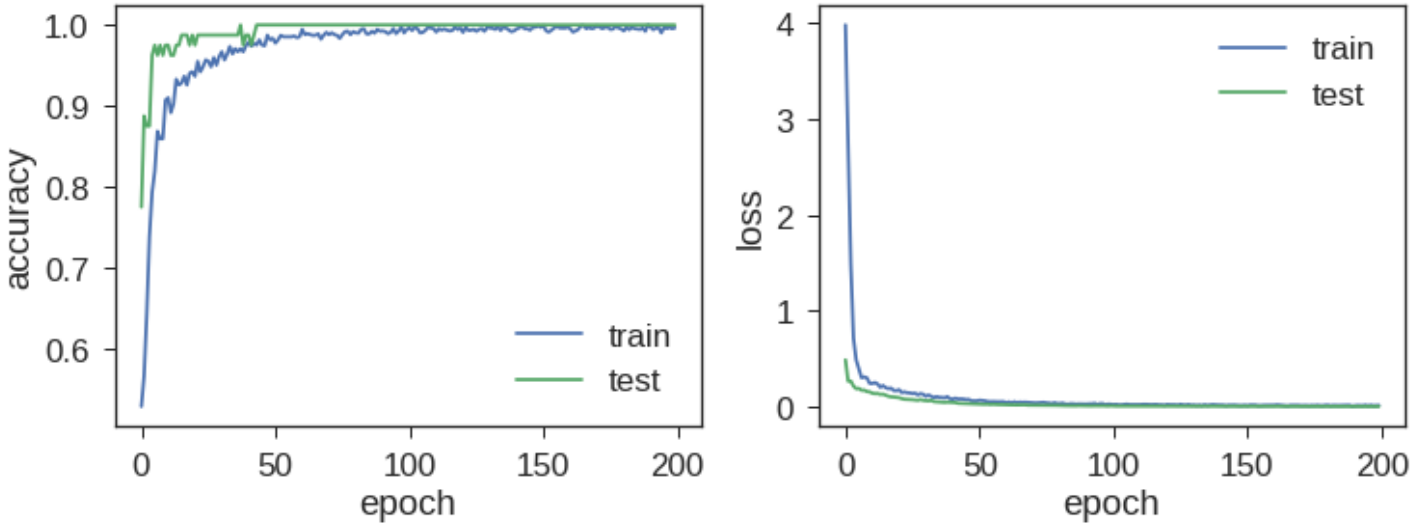


Figure 5. Classification results of the synthesized dataset. (A) Accuracy and loss results by the four machine learning approaches, where C and g are hyper-parameters for SVM, n is an order of PCA, N and p are a number of nodes and dropout rate in the fully connected layer part. **(B)** VGG-19-FC training trajectories with object images. **(C)** VGG-19-FC training trajectories with holograms.

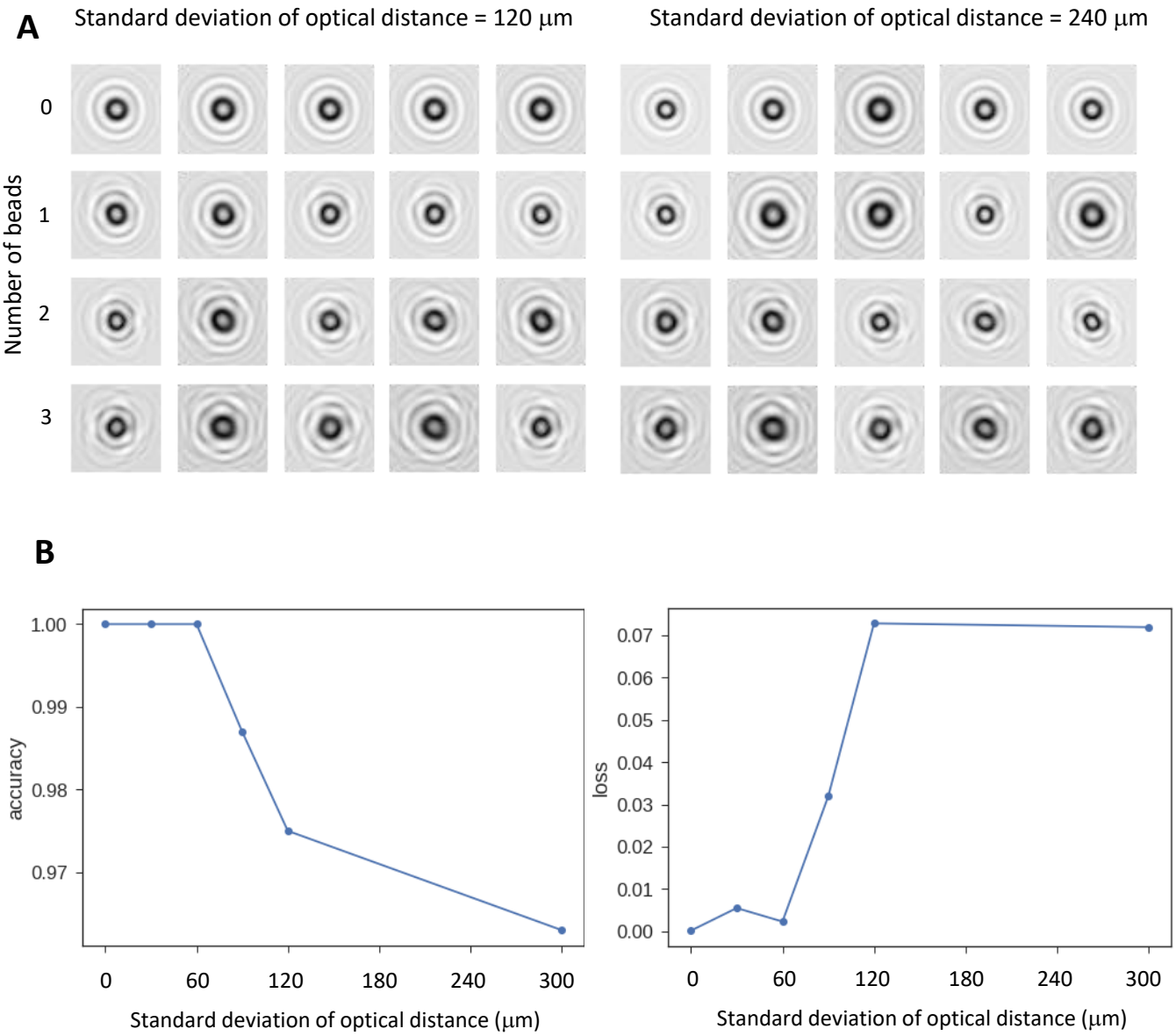


Figure 6. Classification results with respect to optical distance variations. (**A**) Holograms when the standard deviations of optical distance are given 120 and 240 μm, respectively. (**B**) Accuracy and loss performance as a function of the standard deviation of optical distance.

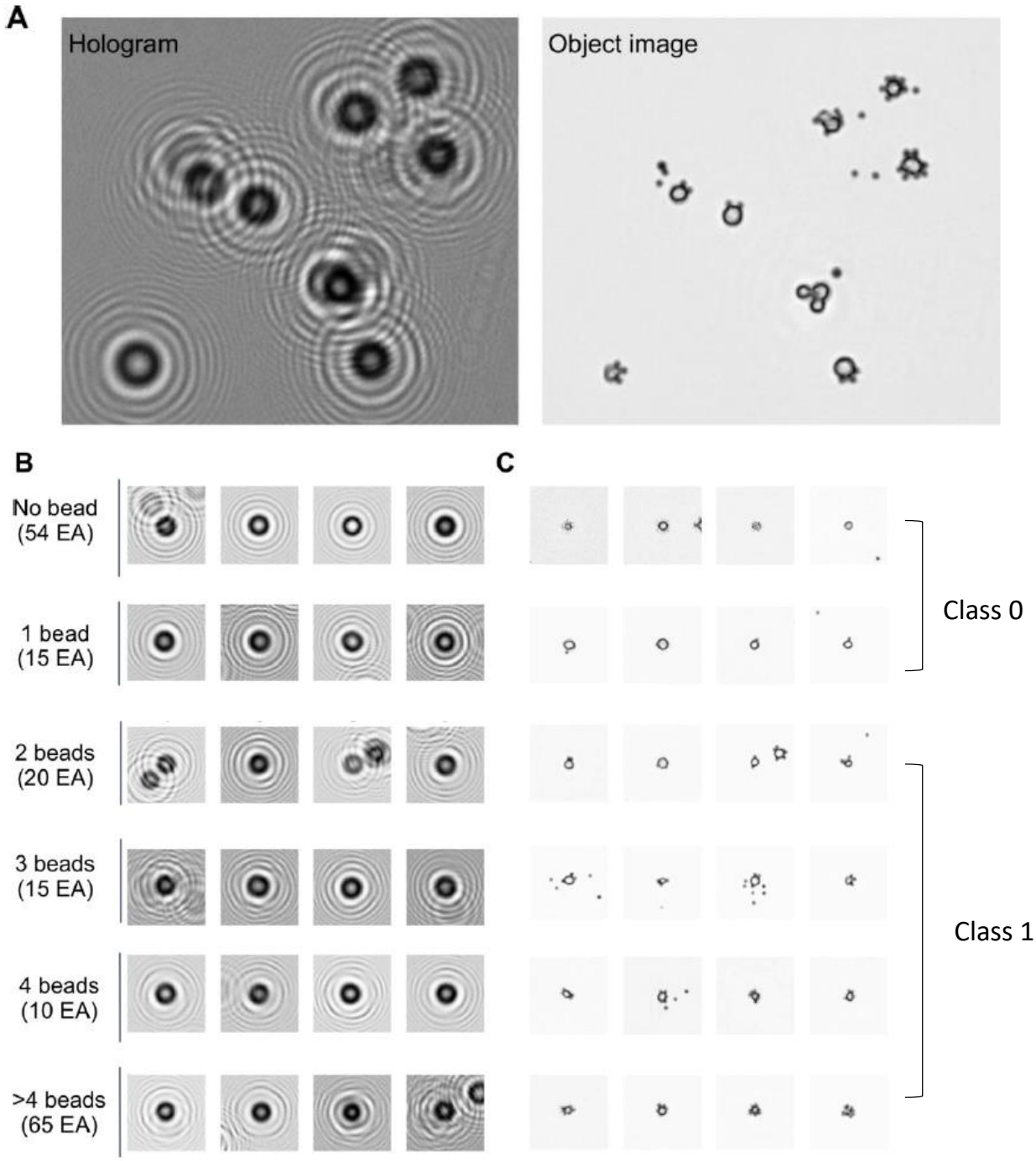
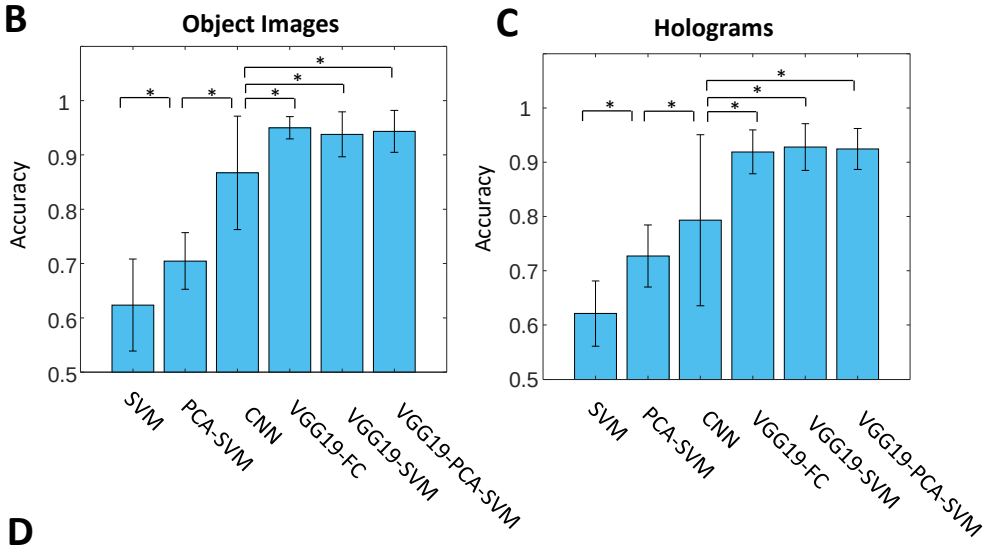


Figure 7. Experimental hologram dataset. (A) Experimental hologram and reconstructed image used for the preparation of training dataset. **(B, C)** Examples of hologram (B) and object images (C) in the training data. Each image has 270 × 270 pixels.

A

	Object Images			Holograms		
	SVM Parameters	Average Accuracy	Standard Deviation	SVM Parameters	Average Accuracy	Standard Deviation
SVM	C=1, g=10^{-1}	62.4%	8.5%	C=1, g=10^{-1}	62.1%	6.0%
PCA-SVM	C=10, g=10^{-2}, n=2	70.5%	5.2%	C=1, g=10^{-1}, n=2	72.7%	5.7%
CNN	-	86.7%	10.4%	-	79.3%	15.8%
VGG19-FC	-	95.0%	2.0%	-	91.9%	4.0%
VGG19-SVM	C=1, g=10^{-3}	93.8%	4.2%	C=1, g=10^{-3}	92.8%	4.3%
VGG19-PCA-SVM	C=1, g=10^{-2}, n=5	94.4%	3.9%	C=10, g=10^{-3}, n=31	92.5%	3.7%



D

	P-values	
	Object Images	Holograms
SVM vs PCA-SVM	1.4 X 10^{-7}	1.1 X 10^{-11}
PCA-SVM vs CNN	1.9 X 10^{-12}	1.3 X 10^{-3}
CNN vs VGG19-FC	8.2 X 10^{-11}	2.3 X 10^{-7}
CNN vs VGG19-SVM	1.8 X 10^{-5}	5.3 X 10^{-6}
CNN vs VGG19-PCA-SVM	1.6 X 10^{-6}	1.5 X 10^{-5}

Figure 8. Classification results of the experimental dataset. (A-C) Accuracy results by the six machine learning approaches. where C and g are hyper-parameters for SVM, n is an order of PCA. * indicates the statistical significance ($p < 0.05$) . The error bars are standard deviation. **(D)** p-values for the statistical analyses of the accuracies of the machine learning approaches by two-tailed Wilcoxon rank sum test.

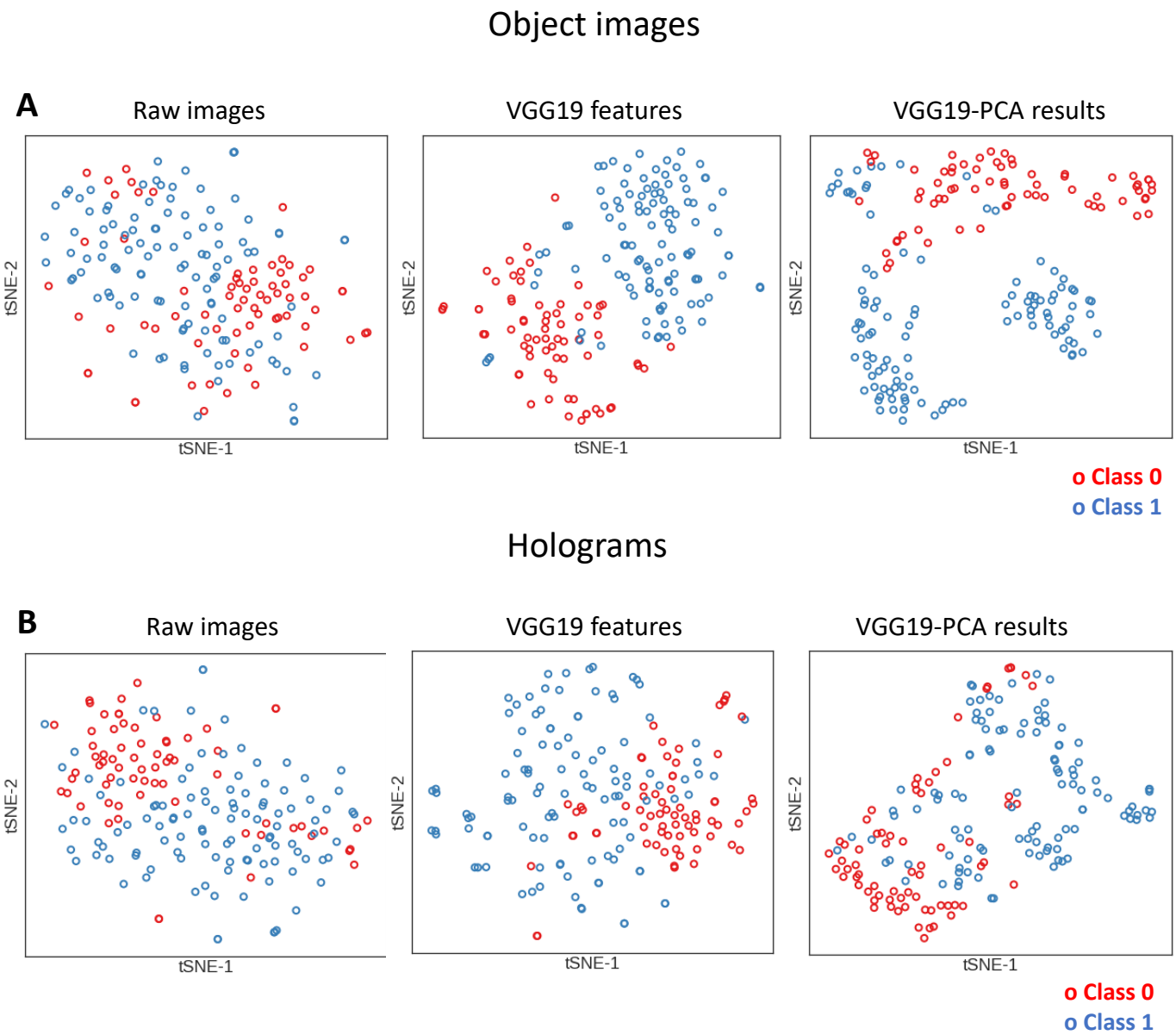


Figure 9. Visualization of the experimental training dataset. (A, B) t-SNE results of raw images, VGG19 output features and VGG19-PCA output results in the case of object images (**A**) and holograms (**B**).