



Classification of unlabeled cells using lensless digital holographic images and deep neural networks

Duofang Chen¹, Zhaohui Wang¹, Kai Chen¹, Qi Zeng¹, Lin Wang², Xinyi Xu¹, Jimin Liang¹, Xueli Chen¹

¹Engineering Research Center of Molecular and Neuro Imaging, Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, China; ²School of Computer Science, Xi'an Polytechnic University, Xi'an, China

Contributions: (I) Conception and design: D Chen, X Chen; (II) Administrative support: J Liang, X Chen; (III) Provision of study materials or patients: Q Zeng, X Xu; (IV) Collection and assembly of data: K Chen, L Wang; (V) Data analysis and interpretation: D Chen, Z Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Xueli Chen. Engineering Research Center of Molecular and Neuro Imaging, Ministry of Education, China and School of Life Science and Technology, Xidian University, Xi'an 710071, China. Email: xlchen@xidian.edu.cn.

Background: Image-based cell analytic methodologies offer a relatively simple and economical way to analyze and understand cell heterogeneities and developments. Owing to developments in high-resolution image sensors and high-performance computation processors, the emerging lensless digital holography technique enables a simple and cost-effective approach to obtain label-free cell images with a large field of view and microscopic spatial resolution.

Methods: The holograms of three types of cells, including MCF-10A, EC-109, and MDA-MB-231 cells, were recorded using a lensless digital holography system composed of a laser diode, a sample stage, an image sensor, and a laptop computer. The amplitude images were reconstructed using the angular spectrum method, and the sample to sensor distance was determined using the autofocus criteria based on the sparsity of image edges and corner points. Four convolutional neural networks (CNNs) were used to classify the cell types based on the recovered holographic images.

Results: Classification of two cell types and three cell types achieved an accuracy of higher than 91% by all the networks used. The ResNet and the DenseNet models had similar classification accuracy of 95% or greater, outperforming the GoogLeNet and the CNN-5 models.

Conclusions: These experiments demonstrated that the CNNs were effective at classifying two or three types of tumor cells. The lensless holography combined with machine learning holds great promise in the application of stainless cell imaging and classification, such as in cancer diagnosis and cancer biology research, where distinguishing normal cells from cancer cells and recognizing different cancer cell types will be greatly beneficial.

Keywords: Lensless digital holography; label-free imaging; convolutional neural networks (CNN); cell classification

Submitted Jan 06, 2021. Accepted for publication May 08, 2021.

doi: 10.21037/qims-21-16

View this article at: <https://dx.doi.org/10.21037/qims-21-16>

Introduction

Cancer is a major cause of death in developed countries and increasingly, also in developing countries. Accurate cancer cell identification and efficient therapy are extremely desirable but challenging in the clinical setting (1). In

addition, distinguishing tumor cells from normal cells holds the key to precise diagnosis and effective intervention of tumors (2). By analyzing the images of cells, image-based cell analytic methodologies offer a relatively simple and economical way to understand cell heterogeneities

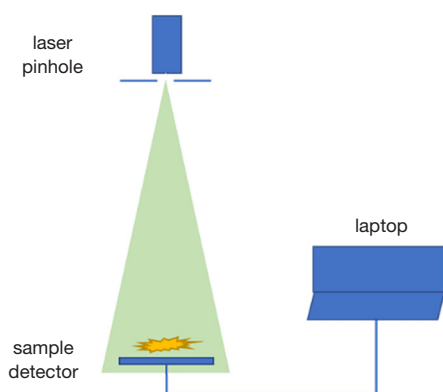


Figure 1 Schematic diagram of the lensless digital holography system.

and developments. Usually, the cell images are acquired through microscopy-based assays, which provide ample visual information that allows the investigation of cellular phenotypes induced by genetic or chemical treatments (3). Label-free cell imaging and analysis avoid the adverse effects of staining reagents on cellular viability and cell signaling. Thus, it is essential for personalized genomics, drug development, and cancer diagnostics (4,5). Digital holographic microscopy (DHM) is a well-known imaging technique that allows the recovery of the complex field information of label-free microscopic samples (6,7). The lensless digital holography is an emerging technology that has the typical configuration of DHM. It works without any objective lens or another intermediate optical component, and thus, does not carry the same limitations as traditional DHM in space-bandwidth product or device size (8). By placing the samples as close as possible to the imaging sensors, the lensless digital holography has unique advantages of a large effective numerical aperture (NA) approaching 1 across the native field of view (FOV) of the imaging sensor (tens of mm^2) (9,10). The system can be built in a miniaturized format, providing a potential solution to reducing health care costs for point-of-care diagnostics in resource-limited environments (11). It enables a simple and cost-effective approach to obtain label-free cell images with large fields of view and microscopic spatial resolution due to high-resolution image sensors and high-performance computation processors.

Due to the escalation of computing power and the availability of massive datasets, the past few years have seen a dramatic surge of interest in deep learning (DL), which is a subfield of machine learning. DL allows computational

models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction (12). DL's unique "representation learning" capability enables direct training from raw images instead of manually extracted features. DL can discover intricate structures in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer (12). Striking breakthroughs have occurred in the field of object recognition, detection, and classification. In recent years, methods based on DL have been introduced in the study of lensless digital holography, including wave field reconstruction (13-15), autofocusing (16,17), noise suppression (18,19), particle classification (20), and molecular diagnostics (21).

In this current study, cell imaging and classification were performed based on recovered holographic images acquired by the lensless digital holography technique. Three cell lines, including the MCF-10A human mammary gland epithelial cell line, the MDA-MB-231 breast cancer cell line, and the EC-109 esophageal cancer cell line, were imaged. The lensless digital holographic imaging system is composed of a laser diode, a sample stage, an image sensor, and a laptop computer. The holograms of the three cell lines were recorded. The images were reconstructed using the angular spectrum method, and the sample to sensor distance was determined using the autofocusing criteria based on the sparsity of image edges and corner points. Based on the reconstructed images, three networks, including ResNet, DenseNet, and GoogLeNet, were used to classify these cells. In addition, a simple CNN was also configured to perform the cell classification task.

Methods

Ethical approval was not required as there were no human experiments, animal experiments, or case reports included in this work. Only cell lines were used for stainless cell imaging.

Reconstruction of cell images from lensless holograms

The schematic diagram illustrated in *Figure 1* is a typical lensless digital holography system, which is composed of a laser diode, a spatial filter, a sample stage, an image sensor, and a laptop computer. The light emitted by the laser diode passes through the pinhole and illuminates the sample. In

this configuration, all the optical components lie along a common axis with the sample. Part of the illumination wave reaching the sample is scattered and propagates to the image detector. The remaining waves are transmitted through the sample and pass to the sensor chip. The scattered wave, also known as the object wave, carries the information about the sample. The non-scattered transmitted wave can be used as the reference wave. A hologram results from the interference of the object wave and the reference wave. Usually, the transmitted wave is known, and the structural information about the sample can be recovered from the interference between the transmitted and the scattered waves.

Mathematically, the interference pattern recorded in the hologram plane can be described using the following formula:

$$I(x,y) = |O(x,y) + R(x,y)|^2 = |O(x,y)|^2 + |R(x,y)|^2 + O(x,y)R^*(x,y) + O^*(x,y)R(x,y) \quad [1]$$

Here, $O(x,y)$ is the object wave, $R(x,y)$ is the reference wave, and (x,y) is the coordinate on the sensor. $|R(x,y)|^2$ is a constant term, $|O(x,y)|^2$ is usually smaller than the other terms and can be neglected, and the sum of two terms $O(x,y)R^*(x,y) + O^*(x,y)R(x,y)$ describes the interference pattern. The first two terms are called the zero-order or dc terms, and the last two terms are the holographic virtual and real images, also called the first-order terms.

In digital holography, numerical reconstruction approaches are used to recover the object wavefront $O(x,y)$ from the recorded hologram $I(x,y)$. The reconstruction algorithms are usually based on a Fresnel–Kirchhoff integral and can be implemented using a Fresnel approximation or an angular-spectrum method. As there is no minimum reconstruction distance constraint imposed before the algorithm breaks down and no pixel scaling between the hologram and its reconstruction, the angular-spectrum method (22) is used in this work. The recovered complex amplitude of the object is expressed as follows:

$$U(x,y;z) = F^{-1} \{ F [I(x,y)] H(u,v;z) \} \quad [2]$$

where F and F^{-1} represent Fourier transform and inverse Fourier transform, respectively, and z is the recording distance from the sample to the sensor. $H(u,v;z)$ is defined as the transfer function, which can be expressed as follows:

$$H(u,v;z) = \begin{cases} e^{jkz\sqrt{1-(\lambda u)^2-(\lambda v)^2}}, & \sqrt{u^2+v^2} < 1/\lambda \\ 0, & \text{otherwise} \end{cases} \quad [3]$$

where $k=2\pi/\lambda$ is the wavenumber, λ is the wavelength, and u and v are the spatial frequencies along the x-axis and y-axis, respectively.

Autofocusing

As shown in Eq. [2], the accurate distance from the sensor to the sample z should be known to recover the object wavefront numerically from the recorded hologram by the angular spectrum method. Usually, a focused image is considered to have the sharpest and sparsest edges. The sharpness or the sparsity of the image edges can be used as an autofocusing criterion. The tensor structure, also referred to as the second-moment matrix, is a widely used tool for corner detection, texture orientation assessment, and sharpness evaluation. Here, the tensor-based sharpness is used as the focus metric. If we let $U(x,y;z)$ be the reconstructed intensity image at distance z , the 2D structure tensor at a pixel (x,y) can be written as follows (23):

$$S(x,y) = G(x,y) * [\nabla U(x,y) \nabla U(x,y)^T] = \begin{bmatrix} G(x,y) * U_m^2(x,y) & G(x,y) * U_m(x,y) U_n(x,y) \\ G(x,y) * U_m(x,y) U_n(x,y) & G(x,y) * U_n^2(x,y) \end{bmatrix} \quad [4]$$

where $\nabla U(x,y) = [U_x(x,y), U_y(x,y)]^T$ is the 2D spatial gradient, and G is a nonnegative convolution kernel which is normally selected as a 2D Gaussian function. Let Δ_1 and Δ_2 be the larger and smaller eigenvalues of the matrix $S(x,y)$. As $S(x,y)$ is a symmetric and semi-positive-definite matrix, we have $\Delta_1 \geq \Delta_2 \geq 0$. The structure tensor measures the geometry of the image structures in the neighborhood of each pixel. The tensor-based sharpness metric is defined as follows:

$$TF(U) = \sum_{x=1}^M \sum_{y=1}^N [|\Delta_1(x,y)|^p + |\Delta_2(x,y)|^p]^{1/p} \quad [5]$$

The larger the TF value, the sharper the image. Autofocusing is used to find the z that leads to the maximum TF value.

Cell classification based on convolutional neural networks

In the field of deep learning, convolutional neural networks (CNNs) are one of the most commonly used types of artificial neural networks. Usually, a CNN is composed of three basic layers: the convolutional layer, the pooling layer, and the fully-connected layer with a rectified linear activation function. The convolution layer extracts image features by convolving the input image with a convolution

kernel. The pooling layer is a down-sampling layer, which reduces the dimensionality of the feature map but retains important information. The fully connected layer is similar to the classical neural networks. Each node of the fully connected layer is connected to all pixels of all input feature maps. In this work, several deep CNNs were utilized for cell classification. The workflow is illustrated in *Figure 2*, which shows the following four main steps: (I) the cell hologram images are obtained; (II) the object images are reconstructed from the holograms; (III) the individual cells are segmented from the recovered images; and (IV) the individual cells are classified.

Three popular deep CNNs were used, namely, ResNet (24), DenseNet (25), and GoogLeNet (26). In addition, a self-configured CNN, named CNN-5, was also

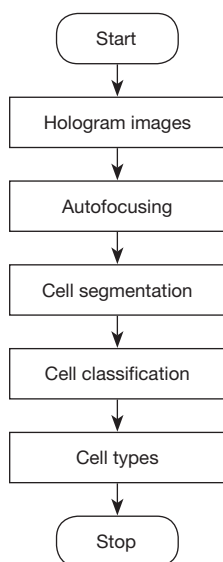


Figure 2 Flow chart of the cell classification process.

used. ResNet uses a residual structure that allows a few intermediate layers to be directly connected to auxiliary classifiers to address vanishing/exploding gradients. DenseNet is very similar to ResNet, except that it uses a dense connection instead of the addition operation that is used in ResNet. This allows the features in the front of the network to be reused in the network's back. GoogLeNet introduces the inception architecture, which integrates convolution and pooling layers into the same layer in parallel, increasing the adaptability to multi-scale feature processing. As shown in *Figure 3*, the CNN-5 has five convolutional layers, with each layer following a batch normalization. The batch normalization acts as a regularizer that can accelerate the deep network training, resulting in the network being less sensitive to the initialization (27). The structure of the self-configured CNN-5 is listed in *Table 1*, where $n=2$ is for the two cell type classification and $n=3$ is for the three cell type classification.

Cell culture and imaging

A lens-free digital in-line holography system was built to acquire holograms of different cells. A 532 nm diode laser beam passes through a 20 μm pinhole and illuminates the sample holder. The scattered light from the sample interferes with the non-scattered light. A complementary metal-oxide-semiconductor (CMOS) sensor with 1.67 μm pixel size was used to record the hologram. The sensor was placed close to the sample holder.

The MDA-MB-231, EC-109, and MCF-10A cell lines were generously supplied by the Xi'an Medical College (Xi'an, China). Cells were incubated for 24 hours in Dulbecco's modified essential medium (DMEM, Sigma-Aldrich) containing 10% fetal bovine serum (FBS, Corning),

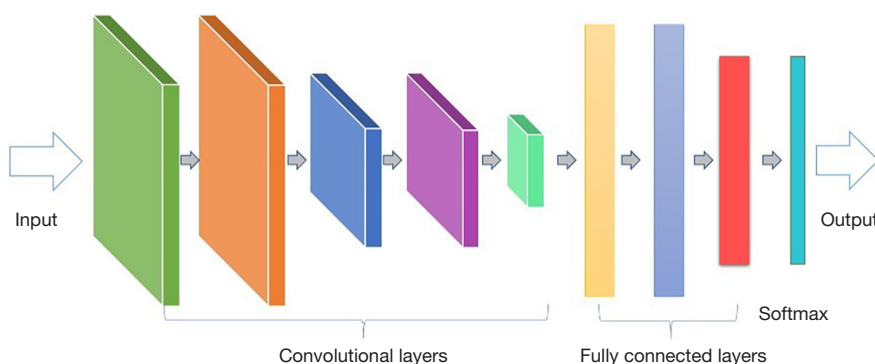


Figure 3 CNN-5 architecture.

Table 1 Detailed configuration of the CNN-5

Layer	Type	Filter dimension	Stride	Pad	Date dimension
0	Input	–	–	–	20×20×1
1	Conv	3×3×15	1	1	20×20×15
2	BNorm	–	–	–	20×20×15
3	ReLU	–	–	–	20×20×15
4	Conv	3×3×15	1	1	20×20×15
5	BNorm	–	–	–	20×20×15
6	ReLU	–	–	–	20×20×15
7	MaxPool	2×2	2	–	10×10×15
8	Conv	3×3×20	1	1	10×10×20
9	BNorm	–	–	–	10×10×20
10	ReLU	–	–	–	10×10×20
11	Conv	3×3×20	1	1	10×10×20
12	BNorm	–	–	–	10×10×20
13	ReLU	–	–	–	10×10×25
14	MaxPool	2×2	2	–	5×5×25
15	Conv	3×3×25	1	1	5×5×25
16	BNorm	–	–	–	5×5×25
17	ReLU	–	–	–	5×5×25
18	FConn	360×625	–	–	1×1×360
19	BNorm	–	–	–	1×1×360
20	ReLU	–	–	–	1×1×360
21	FConn	60×360	–	–	1×1×60
22	BNorm	–	–	–	1×1×60
23	ReLU	–	–	–	1×1×60
24	FConn	n×60	–	–	1×1×n
25	softmax	–	–	–	1×1×n

100 units/mL penicillin, and 100 µg/mL streptomycin (Corning) in standard tissue culture conditions of 37 °C, 5% CO₂, and 100% humidity. Approximately 1 µL suspension containing several thousand cells was dropped onto a glass slide and imaged by the lens-free holography system.

The precision, recall, and accuracy were used to evaluate the performance of the classification networks. The assumption was a binary classification problem where some samples required identification, and the samples belong to two classes, namely, the positive class and the negative class. The three metrics were defined as follows:

$$precision = \frac{TP}{TP + FP} \quad [6]$$

$$recall = \frac{TP}{TP + FN} \quad [7]$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad [8]$$

where TP, TN, FP, and FN represent the true positives, true negatives, false positives, and false negatives, respectively. Each was defined as the number of cases where the model

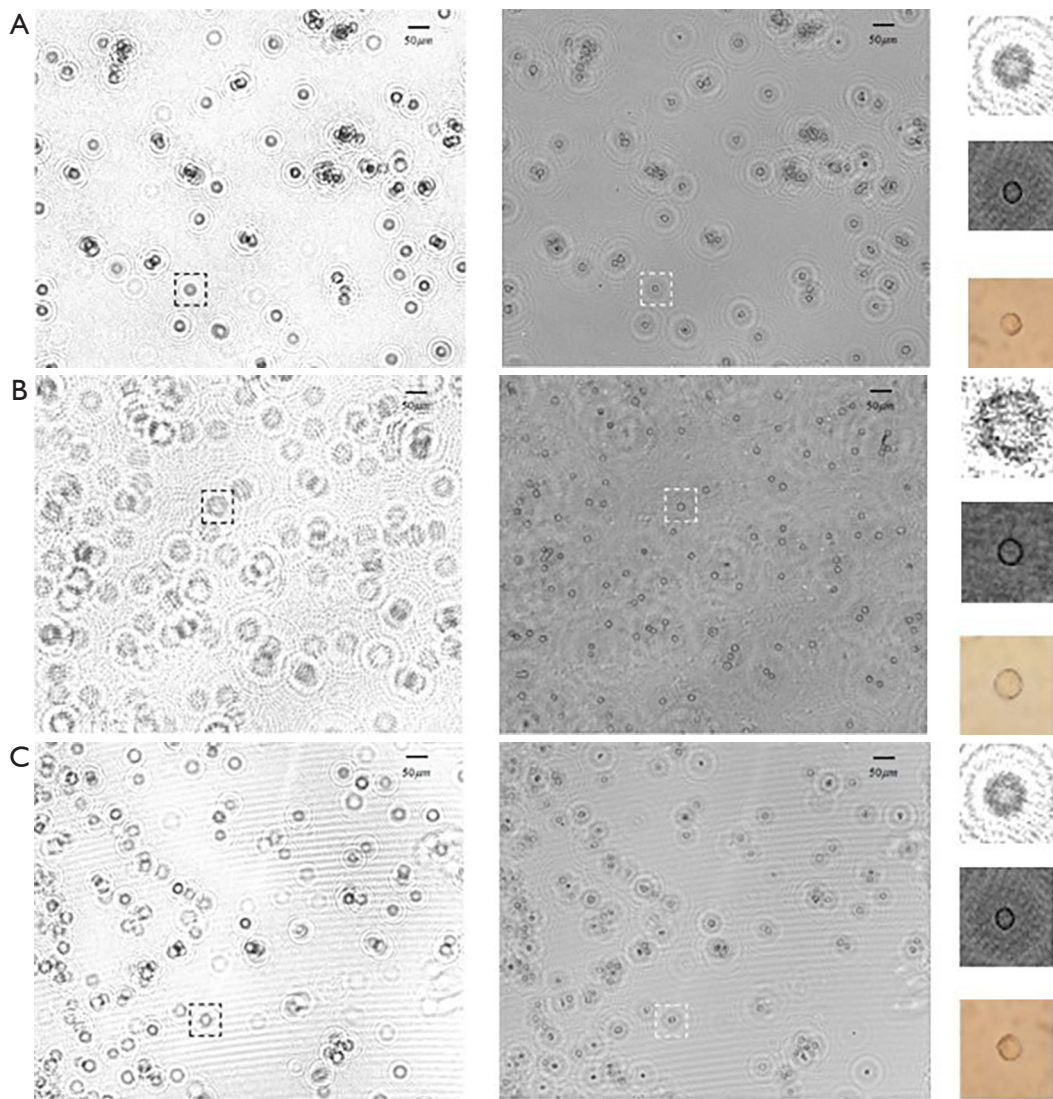


Figure 4 The recorded holograms and reconstructed images for (A) MDA-MB-231, (B) EC-109, and (C) MCF-10A cells. The left, middle, and right columns show the holograms, reconstructed images at the focal plane, and the typical single cell. The images from top to bottom represent the recorded, reconstructed, and bright-field microscopic single cell in the dashed boxes in the right column, respectively. The scale bars represent 50 μm .

correctly predicts the positive class, correctly predicts the negative class, incorrectly predicts the positive class, and incorrectly predicts the negative class, respectively.

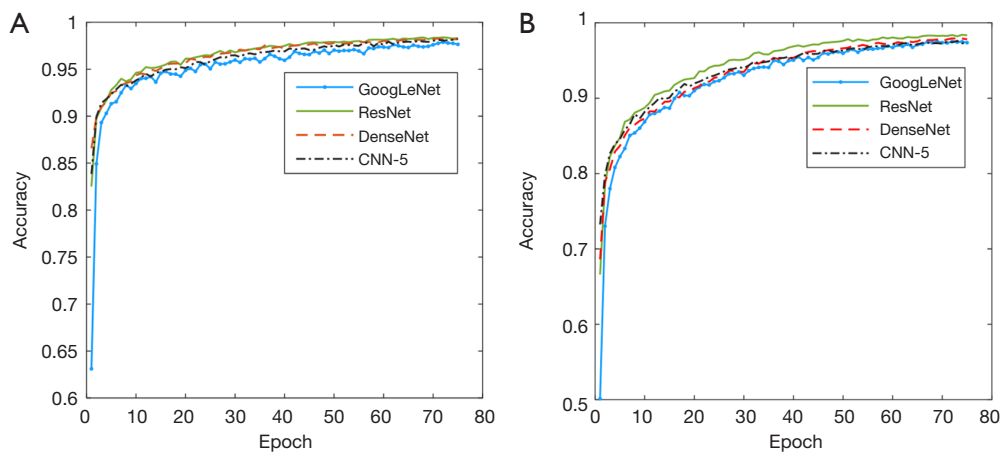
Results

The recorded holograms of the MDA-MB-231, EC-109, and MCF-10A cells are shown in the left column of *Figure 4*, and the reconstructed cell images are illustrated in the middle column. The single enlarged cell in the dashed

boxes is illustrated in the right column in detail, where the recorded, reconstructed, and bright-field microscopic images are listed from top to bottom. The reconstructed images were cropped by a 20×20 pixel box, each of which is centered on a cell. To obtain the position of each cell, the pixels presenting a cell edge were determined based on the fact that the edge of an object has the lowest brightness along different depths of the 3D reconstruction space. These points were then projected onto the same plane, and smoothing and threshold segmentation were

Table 2 Computation cost to train the networks for two cell type classification

	ResNet	DenseNet	GoogLeNet	CNN-5
Floating-point operations	2.288G	4.643G	3.431G	0.803G
Number of parameters	11.171M	6.951M	5.596M	2.912M
Time (s)	1,102.5	3,287.2	1,815.6	680.9

**Figure 5** Validation accuracy vs. epoch during training. (A) Validation accuracy for EC-109 and MDA-MB-231 classification; (B) validation accuracy for EC-109, MDA-MB-231, and MCF-10A classification.

performed. Finally, all cells were extracted from the focused reconstruction image.

The number of segmented cells was 8,160, 6,623, and 6,145 for MDA-MB-231, EC-109, and MCF-10A. The cells were divided into two groups with a ratio of 8:2 for the training and test sets. Training a network is essentially optimizing a nonlinear function concerning weights and biases. The Adam optimizer (28) was utilized to minimize the categorical cross-entropy, which computes the dissimilarity of the approximated output distribution. The learning rate is 0.01 and learns for 75 epochs during training. All the networks were implemented using Python3.7 and computed using a computer with an Intel(R) Core(TM) i5-8400 CPU at 2.80 GHz and an NVIDIA GeForce GTX 1070 8GB card. To investigate the computation costs of ResNet, DenseNet, GoogLeNet, and CNN-5 training, the number of floating point operations, the number of parameters, and the time for MDA-MB-231 and EC-109 classification were recorded. The results in *Table 2* show that CNN-5 had the fewest parameters and the fastest learning speed. DenseNet had the most operations and the longest time needed for training.

During network training, the 5-fold validation strategy was used. *Figure 5* shows the variation of the validation accuracy during the training process, where (I) plots the training results for EC-109 and MDA-MB-231 classification, and (II) plots the training result for the three cell type classification. The results demonstrated that for both two and three cell type classifications, the ResNet and DenseNet models obtained higher accuracy rates than the GoogLeNet and CNN-5 networks.

After network training, the weights of the different networks were restored for testing. The precision rates and recall rates of testing for the two cell type classification are listed in *Tables 3-5*. Classification was performed between MCF-10A and MDA-MB-231, MCF-10A and EC-109, and EC-109 and MDA-MB-231. In these tables, the largest values are illustrated in bold. For the two cell type classification tasks, the best mean precision rates and recall rates were obtained by either ResNet or DenseNet, and GoogLeNet outperformed CNN-5.

The performance of the networks on classifying the three different cell types, MCF-10A, EC-109, and MDA-MB-231, is shown in *Table 6*. DenseNet demonstrated

Table 3 Performance of networks on MCF-10A and MDA-MB-231 classification

Measurement	Cell type	Networks			
		ResNet	DenseNet	GoogLeNet	CNN-5
Precision	MCF-10A	0.967	0.970	0.954	0.964
	MDA-MB-231	0.981	0.968	0.977	0.965
	Mean	0.974	0.969	0.966	0.965
Recall	MCF-10A	0.976	0.958	0.958	0.954
	MDA-MB-231	0.974	0.978	0.978	0.973
	Mean	0.975	0.968	0.968	0.964

Table 4 Performance of networks on MCF-10A and EC-109 classification

Measurement	Cell type	Networks			
		ResNet	DenseNet	GoogLeNet	CNN-5
Precision	MCF-10A	0.940	0.934	0.945	0.918
	EC-109	0.941	0.944	0.929	0.931
	Mean	0.941	0.939	0.937	0.925
Recall	MCF-10A	0.938	0.934	0.924	0.928
	EC-109	0.943	0.943	0.948	0.921
	Mean	0.941	0.939	0.936	0.925

Table 5 Performance of networks on EC-109 and MDA-MB-231 classification

Measurement	Cell type	Networks			
		ResNet	DenseNet	GoogLeNet	CNN-5
Precision	EC-109	0.971	0.962	0.957	0.960
	MDA-MB-231	0.968	0.980	0.978	0.956
	Mean	0.970	0.971	0.968	0.958
Recall	EC-109	0.960	0.975	0.974	0.945
	MDA-MB-231	0.977	0.969	0.964	0.968
	Mean	0.969	0.972	0.969	0.957

superior performance, with mean precision and mean recall rates above 92%. The mean precision and recall rates obtained by the CNN-5 network were both 90.9%, which was the lowest among all the networks.

The accuracy rates of the networks were calculated and illustrated in *Figure 6*, where EC, MDA, and MCF represent EC-109, MDA-MB-231, and MCF-10A respectively. All the networks achieved an accuracy of higher than 91%.

The ResNet and DenseNet networks showed similar classification accuracy that was higher than 95%, and they both outperformed GoogLeNet and CNN-5. Although the number of weight parameters in CNN-5 was much less than that in GoogLeNet, the two networks obtain similar accuracy rates. In addition, the accuracy of the networks for three cell type classification was lower than that for two cell type classifications, except with GoogLeNet.

Table 6 Performance of the networks on three cell type classification

Measurement	Cell type	Networks			
		ResNet	DenseNet	GoogLeNet	CNN-5
Precision	MCF-10A	0.885	0.911	0.908	0.875
	MDA-MB-231	0.959	0.948	0.932	0.935
	EC-109	0.934	0.927	0.939	0.917
	Mean	0.926	0.929	0.926	0.909
Recall	MC-F10A	0.953	0.929	0.915	0.908
	MDA-MB-231	0.934	0.948	0.965	0.950
	EC-109	0.898	0.912	0.894	0.869
	Mean	0.928	0.930	0.925	0.909

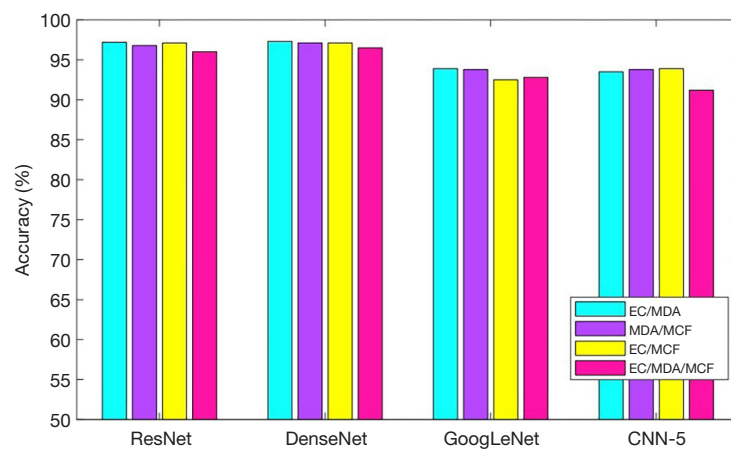


Figure 6 Accuracy of the networks for two cell type and three cell type classification. EC/MDA, MDA/MCF, and EC/MCF denote the two cell type classification. EC/MDA/MCF denotes the three cell type classification. EC, MDA, and MCF represent EC-109, MDA-MB-231, and MCF-10A respectively.

As shown in *Figure 4*, the cells in the reconstructed images were a little blurred, and there were interference fringes around the focused cells. We analyzed 58 misclassified images in the ResNet-based MDA-MB-231 and EC-109 classification experiments to investigate the effects of the twin images on classification accuracy. Among these 58 images, 29 were MDA-MB-231 cells, and the other 29 were EC-109 cells. The interference fringes were removed manually, and the clear cell images were input to the trained ResNet network. Principal component analysis (PCA) was performed on the features of the last convolutional layer output, and the results are shown in *Figure 7A*, where MDA/EC indicates the MDA-MB-231 cells that were misclassified into the EC-109 cell type. A

total of 5 MDA-MB-231 images and 8 EC-109 images were incorrectly predicted after the twin images were eliminated. In addition, the cell images were manually removed, and the left interference fringe images were input to the trained CNNs. The PCA results in *Figure 7B* show the 23 MDA-MB-231 cells and 15 EC-109 cells that were wrongly predicted. These results suggested that the interference fringes mainly caused the misclassifications in the images.

Discussion

In this work, label-free imaging was performed using the digital holography system without a lens. Based on the

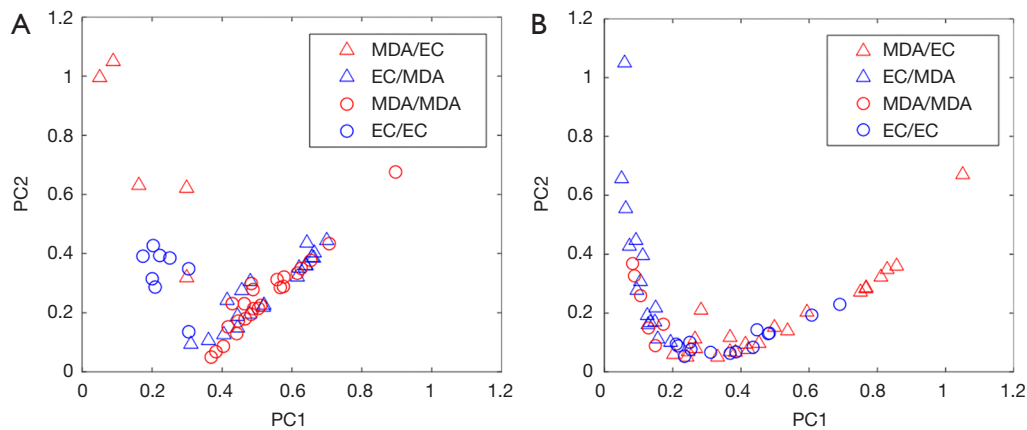


Figure 7 Principal component analysis results of the extracted features when the trained ResNet is input with (A) cell images without interference fringes and (B) interference fringe images. The triangle and circle markers show the cells that were incorrectly and correctly predicted by ResNet, respectively. MDA/EC represents the MDA-MB-231 cell that was misclassified into the EC-109 cell type.

reconstructed holographic images, the cells were classified using four deep neural networks. To recover the wave amplitude of the cells, the angular spectrum method was used. For digital autofocusing, the sample to sensor distance was determined based on the sparsity of image edges and corner points. After reconstruction, each cell was detected and segmented. The images were cropped into thousands of sub-images with a size of 20×20 pixels, and each image was centered on a single cell. For cell classification, the popular ResNet, DenseNet, and GoogLeNet networks were used. In addition, a CNN containing five convolutional layers, named CNN-5, was configured and utilized.

For MDA-MB-231 and EC-109 cell classification, the number of floating point operations during ResNet, DenseNet, GoogLeNet, and CNN-5 training was 2.3G, 4.6G, 3.4G, and 0.8G, respectively. The number of network parameters was 11.2M, 7.0M, 5.6M, and 2.9M, respectively. The training time for 75 epochs ranged from 11 minutes to 55 minutes. As the DenseNet model had the most floating point operations, it had the longest training time. During testing, the precision rates, recall rates, and accuracy rates were calculated to evaluate the performance of the networks. For two cell type classifications, ResNet and DenseNet showed higher mean precision and mean recall rates than the other two networks. The accuracy of three cell type classification was lower than that of two cell type classifications as the number of parameters increases, and the task becomes more difficult. For all the classification tasks, the CNN-5 network performed the fastest while achieving similar accuracy to GoogLeNet.

The CNN-5 network was formulated directly based on the classical CNN. GoogLeNet introduces the inception module to the traditional CNNs, which allows information to be processed at various scales (26). The ResNet model uses shortcut connections to perform identity mapping (24). The DenseNet model connects each layer to every other layer in a feed-forward fashion (25). Thus, GoogLeNet, ResNet, and DenseNet can all alleviate the vanishing gradient problem in the deep networks, and the last two models encourage feature reuse. Although it is not easy to find the optimal CNNs for cell classification tasks, our experiments suggested that ResNet and DenseNet performed better than the other two networks.

This investigation demonstrated that all four networks were effective at classifying two and three cell types. To further improve the performance of the networks and obtain better results, the size of the training data may be increased by imaging more cells and using data augmentation strategies. Due to the effects of the twin images, the cells in the reconstructed images were slightly blurred, and there were interference fringes around the focused cells. To further investigate this effect, 58 misclassified cell images in the ResNet-based MDA-MB-231 and EC-109 classification experiments were analyzed. After manual removal of the fringes, more than half of the cell images were correctly predicted by the pre-trained ResNet, suggesting that the interference fringes mainly caused the misclassification in the images. Therefore, the quality of the reconstructed images should be improved by removing the twin image, thereby enhancing classification performance.

Autofocusing was performed before classification as there were overlapping interference patterns among cells in the recorded holograms to extract individual cells. The focusing operation may be unnecessary if the cells are sparsely located as the cell information has been encoded in the holograms. Future studies should investigate whether cell types can be directly distinguished from the recorded holograms.

Conclusions

In recent years, lensless digital holography has advanced as a label-free modality. This has been facilitated by developments in inexpensive digital image sensors with small pixel size and high pixel counts and improvements in computing power and reconstruction algorithms used to process the captured diffraction patterns. This imaging technique can provide a spatial resolution of several hundred nanometers to several micrometers without any focusing lens, offering the advantages of a large field of view, high resolution, cost-effectiveness, and portability. In this study, three types of unlabeled cells, specifically, EC-109, MDA-MB-231, and MCF-10 cells, were imaged using the lensless digital holography technique, and four deep learning networks including ResNet, DenseNet, GoogLeNet, and CNN-5 were used for cell classification. The results demonstrated that lensless digital holography combined with deep CNNs could provide a powerful imaging modality for large field label-free cell classification. This combined technique can distinguish normal cells from cancer cells and recognize different types of cancer cells. This will afford enormous benefits to the field of cancer diagnosis and cancer biology research.

Acknowledgments

Funding: This work was supported in part by the National Key R&D Program of China (No. 2018YFC0910600), the National Natural Science Foundation of China (No. 81627807, 81871397, 62007026), the National Young Top-notch Talent of “Ten Thousand Talents Program”, the Shaanxi Science Fund for Distinguished Young Scholars (No. 2020JC-27), the Fok Ying-Tung Education Foundation of China (No. 161104), the Shaanxi Young Top-notch Talent of “Special Support Program”, the China Scholarship Council (No.201906965029), and the Fundamental Research Funds for Central Universities (No. JB211211).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/qims-21-16>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Ethical approval was not obtained because no human experiments, animal experiments, or case reports were included in this work.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Su J, Wu F, Xia H, Wu Y, Liu S. Accurate cancer cell identification and microRNA silencing induced therapy using tailored DNA tetrahedron nanostructures. *Chem Sci* 2019;11:80-6.
2. Liu Z, Zhao J, Zhang R, Han G, Zhang C, Liu B, Zhang Z, Han M, Gao X. Cross-Platform Cancer Cell Identification Using Telomerase-Specific Spherical Nucleic Acids. *ACS Nano* 2018;12:3629-37.
3. Meng N, Lam EY, Tsia KK, So KH. Large-scale multi-class image-based cell classification with deep learning. *IEEE J Biomed Health Inform* 2019;23:2091-8.
4. Chen CL, Mahjoubfar A, Tai LC, Blaby LK, Huang A, Niazi KR, Jalali B. Deep Learning in Label-free Cell Classification. *Sci Rep* 2016;6:21471.
5. Park JH, Go T, Lee SJ. Label-free sensing and classification of old stored blood. *Ann Biomed Eng* 2017;45:2563-73.
6. Go T, Byeon H, Lee SJ. Label-free sensor for automatic identification of erythrocytes using digital in-line holographic microscopy and machine learning. *Biosens Bioelectron* 2018;103:12-8.
7. Cucho E, Bevilacqua F, Depeursinge C. Digital holography

- for quantitative phase-contrast imaging. *Opt Lett* 1999;24:291-3.
8. Greenbaum A, Luo W, Su TW, Gorocs Z, Xue L, Isikman SO, Coskun AF, Mudanyali O, Ozcan A. Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy. *Nat Methods* 2012;9:889-95.
 9. Garcia-Sucerquia J, Xu W, Jericho M, Kreuzer HJ. Immersion digital in-line holographic microscopy. *Opt Lett* 2006;31:1211-3.
 10. Ozcan A, McLeod E. Lensless imaging and sensing. *Annu Rev Biomed Eng* 2016;18:77-102.
 11. Zhang J, Sun J, Chen Q, Li J, Zuo C. Adaptive pixel-super-resolved lensfree in-line digital holography for wide-field on-chip microscopy. *Sci Rep* 2017;7:11777.
 12. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44.
 13. Rivenson Y, Zhang Y, Günaydin H, Teng D, Ozcan A. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light Sci Appl* 2018;7:17141.
 14. Sinha A, Lee J, Li S, Barbastathis G. Lensless computational imaging through deep learning. *Optica* 2017;4:1117-25.
 15. Ren Z, Xu Z, Lam EY. End-to-end deep learning framework for digital holographic reconstruction. *Advanced Photonics* 2019;1:016004.
 16. Ren Z, Xu Z, Lam E Y. Learning-based nonparametric autofocusing for digital holography. *Optica* 2018;5:337-44.
 17. Jaferzadeh K, Hwang SH, Moon I, Javidi B. No-search focus prediction at the single cell level in digital holographic imaging with deep convolutional neural network. *Biomed Opt Express* 2019;10:4276-89.
 18. Zeng T, So HKH, Lam EY. Computational image speckle suppression using block matching and machine learning. *Appl Opt* 2019;58:B39-45.
 19. Chen L, Chen X, Cui H, Long Y, Wu J. Image enhancement in lensless inline holographic microscope by inter-modality learning with denoising convolutional neural network. *Opt Commun* 2020; 126682.
 20. Wu Y, Ray A, Wei Q, Feizi A, Tong X, Chen E, Ozcan A. Deep learning enables high-throughput analysis of particle-aggregation-based biosensors imaged using holography. *ACS Photonics* 2018;6:294-301.
 21. Kim S J, Wang C, Zhao B, Im H, Min J, Choi HJ, Tadros J, Choi NR, Castro CM, Weissleder R, Lee H, Lee K. Deep transfer learning-based hologram classification for molecular diagnostics. *Sci Rep* 2018;8:17003.
 22. Poon TC, Liu J. *Introduction to Modern Digital Holography*. Cambridge: Cambridge University Press, 2014; 95-115.
 23. Ren Z, Chen N, Lam EY. Automatic focusing for multisectional objects in digital holography using the structure tensor. *Opt Lett* 2017;42:1720-3.
 24. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc Comput Vision Pattern Recognit* 2016; 770-778.
 25. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proc Comput Vision Pattern Recognit* 2017; 4700-4708.
 26. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. *Proc Comput Vision Pattern Recognit* 2015; 1-9.
 27. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv* 2015; 1502: 03167. <https://arxiv.org/pdf/1502.03167.pdf>.
 28. Kingma DP, Ba J. Adam: A method for stochastic optimization. *Computer Science arXiv preprint arXiv* (2014);1412.6980. Available online: <https://arxiv.org/pdf/1412.6980v8.pdf>

Cite this article as: Chen D, Wang Z, Chen K, Zeng Q, Wang L, Xu X, Liang J, Chen X. Classification of unlabeled cells using lensless digital holographic images and deep neural networks. *Quant Imaging Med Surg* 2021;11(9):4137-4148. doi: 10.21037/qims-21-16