Received: 1 July 2019 | Revised: 1 December 2019 | Accepted: 4 December 2019

DOI: 10.1002/jbio.201900242

JOURNAL OF BIOPHOTONICS

FULL ARTICLE

Deep learning of diffraction image patterns for accurate classification of five cell types

Jiahong Jin[1,2,3] | Jun Q. Lu[2,1] | Yuhua Wen[1,3] | Peng Tian[1,3] | Xin-Hua Hu[2,1*]

[1]Institute for Advanced Optics, Hunan Institute of Science and Technology, Yueyang, Hunan, China

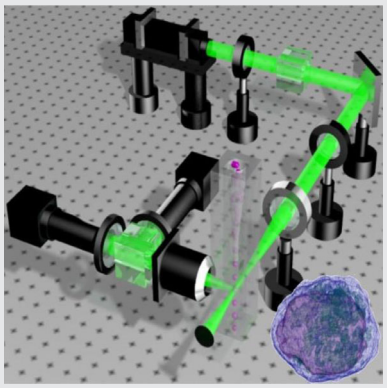
[2]Department of Physics, East Carolina University, Greenville, North Carolina

[3]School of Physics & Electronic Science, Hunan Institute of Science and Technology, Yueyang, Hunan, China

*Correspondence
Xin-Hua Hu, Department of Physics, East Carolina University, Greenville, NC.
Email: hux@ecu.edu

Funding information
Golfers Against Cancer, Grant/Award Number: 2012–13-GAC; East Carolina University, Grant/Award Number: IRA-2016

Abstract



Development of label-free methods for accurate classification of cells with high throughput can yield powerful tools for biological research and clinical applications. We have developed a deep neural network of DINet for extracting features from cross-polarized diffraction image (p-DI) pairs on multiple pixel scales to accurately classify cells in five types. A total of 6185 cells were measured by a polarization diffraction imaging flow cytometry (p-DIFC) method followed by cell classification with DINet on p-DI data. The averaged value and SD of classification accuracy were found to be 98.9% ± 1.00% on test data sets for 5-fold training and test. The invariance of DINet to image translation, rotation, and blurring has been verified with an expanded p-DI data set. To study feature-based classification by DINet, two sets of correctly and incorrectly classified cells were selected and compared for each of two prostate cell types. It has been found that the signature features of large dissimilarities between p-DI data of correctly and incorrectly classified cell sets increase markedly from convolutional layers 1 and 2 to layers 3 and 4. These results clearly demonstrate the importance of high-order correlations extracted at the deep layers for accurate cell classification.

KEYWORDS

cell assay, deep neural network, diffraction imaging, light scattering

1 | INTRODUCTION

Rapid classification of cells according to their phenotypes and/or pathologic conditions is a significant and long-standing problem [1]. While genomic analysis offers the gold standard [2], accurate and cost-effective methods at the single-cell level are highly valuable and attract intense research efforts [3]. Conventional methods of microscopy and flow cytometry are current tools of choice, but the former requires time-consuming manual analysis and the latter yields very limited morphology information. Furthermore, both need cell staining by often multiple fluorescent reagents [4]. Besides the burdens of labor intensive preparation and cost, these reagents can disturb cellular functions to be investigated to the extent of cell toxicity and exhibit inconsistency in molecular marking due to inherent variations in affinity and brightness for quantifying molecules of interest [5, 6].

Unlike fluorescence, light scattered elastically by a cell under laser illumination is coherent and forms diffraction patterns in far field given by the spatial distribution of intensity I(**k**) with **k** as the wave vector. Since I(**k**) arises from the heterogeneous distribution of refractive index (RI) n(**r**) at an intracellular position **r**, imaging of I(**k**) thus opens up the possibilities for morphology-based and label-free cell assay. Determination of intracellular phase distribution or reconstruction of n(**r**) can be achieved with images of I(**k**) acquired over a sufficiently large angular range of **k** with or without time-stretch technique [7–11]. It should be noted that I(**k**) needs to be oversampled relative to the Nyquist limit, with image number proportional inversely to the sub-cellular organelle sizes, to avoid pitfalls such as non-convergence and non-uniqueness in phase unwrapping and subsequent determination of n(**r**). The requirements of multi-shot and oversampled image acquisition adds to the high computation costs of reconstruction from the measured data and make it unsuitable for rapid classification of cells in large numbers. On the other end of speed range, time-stretch ultrashort laser pulses of wide bandwidth are expanded spatially into lines by grating. With multiple pules at a high repetition rate projected on a moving cell, the interferograms obtained with these probing pulses are picked up by a fast photodetector to form a phase image of the cell. While the time-stretch technique affords fast imaging of moving cells, the phase images are of low resolution because of the limitation of pulse bandwidth for spectral expansion and repetition rate relative to the cell moving speed [12]. Furthermore, in either case of n(**r**) reconstruction and phase imaging, the acquired data are similar to those microscopy images formed by noncoherent light that require complex segmentation of intracellular organelles to extract morphological features. Despite rapid advances in machine learning, segmentation presents significant challenges for full automation.

We have developed a single-shot approach to image I(**k**) over a limited angular range along the side scattering directions. It is clear that the acquired images are insufficient for reconstruction of n(**r**). However, they yield truly big data for rapid assay of single cells and led to a polarization diffraction imaging flow cytometry (p-DIFC) method [13–19]. By flowing single cells through the focus of a linearly polarized laser beam, one pair of cross-polarized diffraction images (p-DI) is acquired per cell by an imaging unit with a microscope objective. Since the diffraction patterns embedded in the p-DI data correlate strongly with n(**r**) of an imaged cell, the p-DIFC method offers a rapid and cost-effective approach to pursue morphology based single-cell assay, which needs no fluorescent staining and image segmentation. Indeed, we have shown that texture parameters extracted from p-DIs as dark-field images can be used to accurately distinguish two cell lines or types at a time, which include Jurkat versus Ramos cells derived from human T and B cancer cells and other cell types [17–20].

Our previous methods for cell classifications with p-DI data are based on parameter analysis that consist of two steps. The input images were first processed by an algorithm for extraction of texture parameters followed by a classifier operating in the parameter space. Different algorithms have been explored for extracting image texture parameters, which include gray level co-occurrence matrix (GLCM), short-time Fourier transform, contourlet and Gabor transforms [18, 20–24]. Despite the variations in effectiveness for classification of cells in two types, the parameter based approach requires labor intensive assessment and validation. More importantly, the extracted p-DI parameters show mixed sensitivity to texture types in classification of multiple cell types that lead to fluctuations in accuracy due to inability to characterize textures on multiple pixel scales [19, 24]. To take the advantages of the big data nature of p-DIs, it is necessary to develop algorithms that can learn features of different pixel scales from the input image data for improvement of classification accuracy. Achieving this goal should markedly enhance the robustness of cell assay with p-DI data and reduce cost of algorithm development. In this study, we have investigated various designs of convolutional neural networks (CNNs) for classifying cultured cells in five types derived from human white blood and epithelial cells. The process of feature extraction by an optimized architect, termed as DINet, has been quantitatively analyzed to gain insight on the feature transfer through the layers. Our results show that the DINet architect can quantify high-order correlations in the p-DI data for accurate cell classification.

2 | METHODS

2.1 | Cell sample preparation and confocal imaging

Five types of cultured human cells were purchased from ATCC (Manassas, VA) that include four cancer cell lines of Jurkat, Ramos, PC3, MCF-7 and a normal prostate epithelial cell type of PCS (PCS440010, ATCC). The Jurkat and Ramos cell lines are suspension cells derived from T and B lymphocytes while the PC3, MCF-7, and PCS are adherent epithelial cells derived respectively from patients of prostate cancer, breast cancer, and from normal human prostate tissues. The cell lines were maintained in RPMI-1640 (Gibco, ThermoFisher) supplemented with 10% fetal calf serum maintained while

PCS cells were in the prostate epithelial cell basal medium (PCS440030, ATCC) supplemented with the prostate epithelial cell growth kit (PCS440040, ATCC). All suspended cell samples were prepared during the logarithmic phases of cell growth and adherent cells were detached with the trypsin–EDTA solution. Cell viability of each sample was assessed by trypan blue exclusion tests with percentages of viable cells found to range from 95% to 98%. Cell concentrations were adjusted to values between 1×10^6 and 2×10^6 cells/mL for p-DIFC measurement and all samples were kept on ice before confocal and diffraction imaging. For each cell type, small portions of the cell suspensions were doubly stained with Syto-61 (S11343, ThermoFisher) for nucleus and MitoTracker Orange (M-7510, ThermoFisher) for mitochondria. Confocal image stacks were acquired using a laser scanning confocal microscope (LSM 510 or 700, Zeiss) followed by 3D reconstruction and quantitative measurement of morphologic parameters [17, 18, 25, 26].

2.2 | Diffraction imaging flow cytometry

Optical and fluidic designs of the p-DIFC system have been described in details elsewhere [15, 17, 18]. Briefly, cell suspension was injected into the square channel of a flow chamber from a nozzle as the core fluid and hydrodynamically focused by a sheath fluid. A continuous-wave solid state laser (MGL-III-532 nm-100 mW, CNI Optoelectronics Tech. Co.) was used to produce an incident beam of 532 nm in wavelength. As shown in Figure 1, the linearly polarized laser beam was focused on the core fluid with a spot size of about 30 μm at about 1 mm from the nozzle tip. The power and polarization direction of the incident beam were adjusted with two half-wave plates and a polarizing beam splitter. An imaging unit collects the coherent light I(**k**) scattered by a cell with an infinity-corrected 50x objective of 0.55 in NA (378-805-3, Mitutoyo). The angular cone for measurement of I(**k**) has its center axis along the side scattering direction of 90° from the incident beam direction $\mathbf{k}_0$. The scattered light collected by the objective was divided into two beams of s- and p-polarizations and focused by two tube lenses on cameras for acquisition of one p-DI pair. The imaging unit consisting of the objective, an interference filter centered at 532 nm with width of 10 nm, a polarizing beam splitter, two tube lenses and cameras was translated as a whole toward the flow chamber by an off-focus distance of Δx from the position conjugate to the imaged cell. It has been shown that nonconjugate imaging of the coherent light with $\Delta x > 0$ for moving toward the chamber allows adjusting the cone angle of scattered light collected by the imaging unit to optimize image contrast [14, 27]. We set $\Delta x = 100$ μm for p-DI measurement from smaller white blood cells of Jurkat and Ramos derived from human lymphocytes and $\Delta x = 150$ μm for p-DI from larger epithelial cells of PC3, PCS, and MCF-7. The change of Δx from 100 to 150 μm corresponds to variation of the half-cone angle for scattered light collection by the imaging unit from 23.76° to 23.27° [27]. The flow speed of cell through the incident beam was set at about 4 mm/s and exposure time of the CCD cameras at 1 ms.

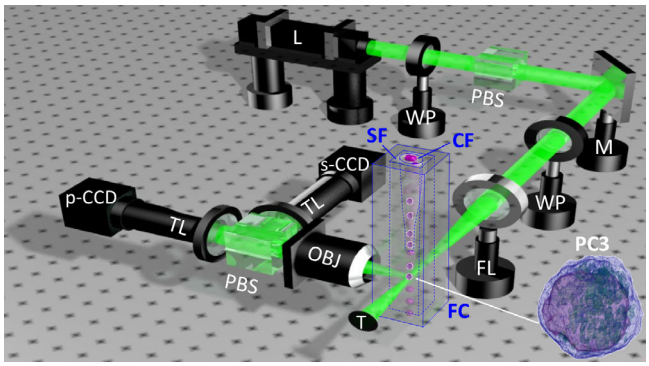


FIGURE 1 Schematic of a p-DIFC system: L, laser; WP, half-wave plate; PBS, polarizing beam splitter; M, mirror; FL, focusing lens; FC, flow chamber; CF, core fluid; SF, sheath fluid; T, trap; OBJ, objective; TL, tube lens. The magnified view of one reconstructed PC3 cell has its nucleus colored in purple and mitochondria in green

2.3 | Data preparation and CNN training

In p-DIFC measurement, light scattering occurs only during the time of a cell moving across the focus of the incident laser beam. Thus, the acquired p-DI pairs are dark-field images of single cells with 12-bit pixel depth and little background. After data acquisition, the overexposed and underexposed images were removed automatically followed by manual prescreening, which eliminated the p-DIs due to cell debris and aggregated homogeneous particles based on their significantly different image patterns from those of intact cells [16]. Each 12-bit image in a p-DI pair of single cells was normalized by its minimum and maximum pixel intensities, and the two cross-polarized images were saved into one false-color image of 8-bit pixels in two channels as a combined DI for reducing data sizes. The measured set of combined DIs were mixed among the five cell types and divided into five groups as listed in Table 1 to form training and test data sets. All CNN classifiers were investigated with a 5-fold cross-validation scheme in which network training and test were

TABLE 1 Groups of combined DIs

Cell type	N_{DI}[a]	Group 1	Group 2	Group 3	Group 4	Group 5
Jurkat	1504	291	304	301	312	296
Ramos	1179	234	239	233	239	234
PC3	1071	211	224	218	201	217
PCS	1138	235	207	236	223	237
MCF-7	1293	266	263	249	262	253
Total	6185	1237	1237	1237	1237	1237

[a]N_{DI} = number of combined DIs or cells in the measured set.

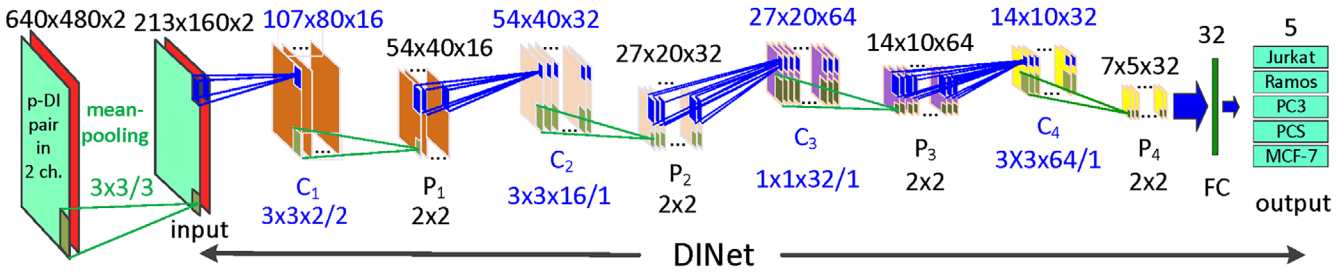


FIGURE 2 The dataflow graph of DINet with eight layers of convolution and pooling before the FC layer. Each layer has its element numbers of each dimension shown on top as N_{mi}, N_{mj}, N_{md}, and receptive field size/stride at bottom. The blue numbers are for convolutional layers and black for other layers

repeated five times by rotating one group of p-DI data as the held-out test data set and the rest as the training-validation data set. Each round started with the training-validation data set and ended after 250 epochs to classify the test data set. The data sets of training and validation were formed by a 9:1 ratio and a batch size of 256 combined DIs was used for training. In addition, we have obtained an expanded set of combined DIs to test invariance of an investigated network architect to pixel translation, rotation, and blurring. Each combined DI listed in Table 1 underwent three consecutive transforms of horizontal translation by 64 pixels, vertical flipping and pixel smoothing by a 5×5 window to form the expanded set. Thus, the expanded set was made of the same five groups with 8-fold increases in image numbers.

TensorFlow was employed as the platform with the python library of TFlearn for construction and optimization of CNNs [28]. Training and test of all investigated networks were performed on one GPU card (Tesla K40, Nvidia) in a high-performance computing cluster node at the Institute for Advanced Optics. Different configurations of convolutional and fully connected (FC) layers, connected by pooling layers, have been evaluated with the combined DIs listed in Table 1 to optimize performance. An optimized CNN classifier has been obtained as DINet that produced excellent classification performance with training times of 250-epoch at about 25 minutes. Once trained, DINet took only about 5 seconds on one CPU (Xeon E5-2630, Intel) to complete classification of 1237 combined DIs in group 1. Figure 2 presents the architecture of DINet.

Cascading feature extraction in DINet is handled by convolutional layers of C_m with m = 1, ..., 4. Each layer outputs a 3D array of elements $y_m(i, j, d)$ from its input array of elements $x_m(i, j, d)$ by

$$y_m(i,j,d) = f(x_m(i,j,d)) = f(\mathbf{w}_m(d) \otimes \mathbf{p}_{m-1} + b_m(d)) \quad (1)$$

where $1 \le i, j, d \le N_{mi}, N_{mj}, N_{md}$, respectively, with N_{mi}, N_{mj}, and N_{md} as the element numbers of each dimension for arrays of $\mathbf{x}_m$ and $\mathbf{y}_m$, f is the activation function given by rectified linear unit (ReLU) [29], $\mathbf{p}_{m-1}$ denotes the 3D input array in the previous pooling layer as seen by $y_m(i, j, d)$, $\mathbf{w}_m(d)$ is the 3D convolution feature map by weight parameters and $b_m(d)$ is the bias. The convolutional stride was set to two for C_1 and one for C_2 to C_4. With shared elements for $\mathbf{w}_m(d)$, DINet needs only to optimize 2.5×10^4 weight parameters for marked decrease in training time with robust performance.

For each round of CNN classifier training and test, we adopted the following definition of the classification accuracy A_i for cell type i with i = 1, ..., 5 [30].

$$A_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (2)$$

where TP_i is the number of combined DIs labeled as *i*th cell type and are correctly classified as *i*th type, TN_i is the number labeled as one of other types and are correctly

classified as that type, FP_i is the number labeled as one of other types and are incorrectly classified as *i*th type and FN_i is the number labeled as *i*th type and are incorrectly classified as other types. The averaged and standard deviation values of A_i over the five cell types were obtained as A_{av} and A_{std}, respectively. After training, the classifier was tested on the held-out group to obtain A_{av} and A_{std} and their mean values were determined as μ_A and σ_A respectively to assess the classifier's performance on five test data sets.

3 | RESULTS

3.1 | Imaging measurement of five cell types

Both confocal fluorescence and diffraction imaging were carried out on suspended cell samples. The former was performed with double fluorescence staining of nucleus and mitochondria, which typically took up to 2 minutes to complete one stack of confocal images with one to five cells in a field of view of about 35 μm. Figure 3A presents confocal images of two cells per type selected near the middle of respective image stacks of about 40 to 70 slices per stack. A software has been developed for reconstruction of 3D cell structures and determination of morphology parameters for cells of each type [25, 26]. The p-DI data were acquired from single cells carried by the core fluid and excited by a vertically polarized laser beam of 532 nm in wavelength. The throughput rate was about one cell/s that was limited by the speeds of cameras and data transmission of the experimental system.

Each acquired p-DI pair consists of two cross-polarized images of 640 × 480 pixels. For vertically polarized incident beam, the s-polarized scattered light is much stronger than the p-polarized component as shown in Figure 3B by the average pixel intensity values. It should be noted that the linear depolarization ratio δ_L is given by the averaged 12-bit pixel value ratio of p- and s-polarized diffraction images. We have shown that δ_L can be used for cell classification for its dependence on the types and distributions of molecular dipoles [23]. For this study, however, we import only the combined DIs into a CNN classifier without δ_L. This allows us to examine the effectiveness of diffraction patterns alone for cell classification. Figure 4 presents the combined DIs acquired from single cells for each type. While pattern differences among the five cell types are visible, it is important to establish an objective approach of p-DI analysis for cell classification by machine learning.

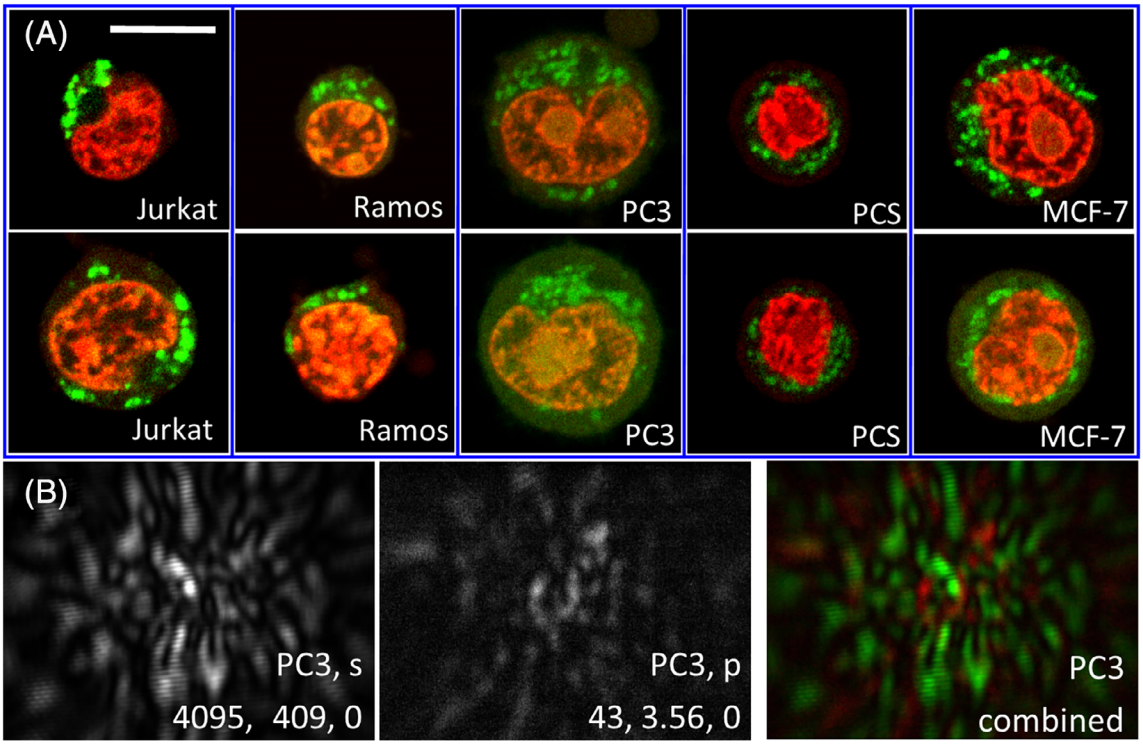


FIGURE 3 A, Fluorescence confocal image slices of two cells per type with nuclear fluorescence saved in red and mitochondrial fluorescence saved in green channels; B, one normalized p-DI pair and combined DI in false colors from one PC3 cell. The labels indicate cell type, scattered light polarization, maximum, average, and minimum 12-bit pixel values

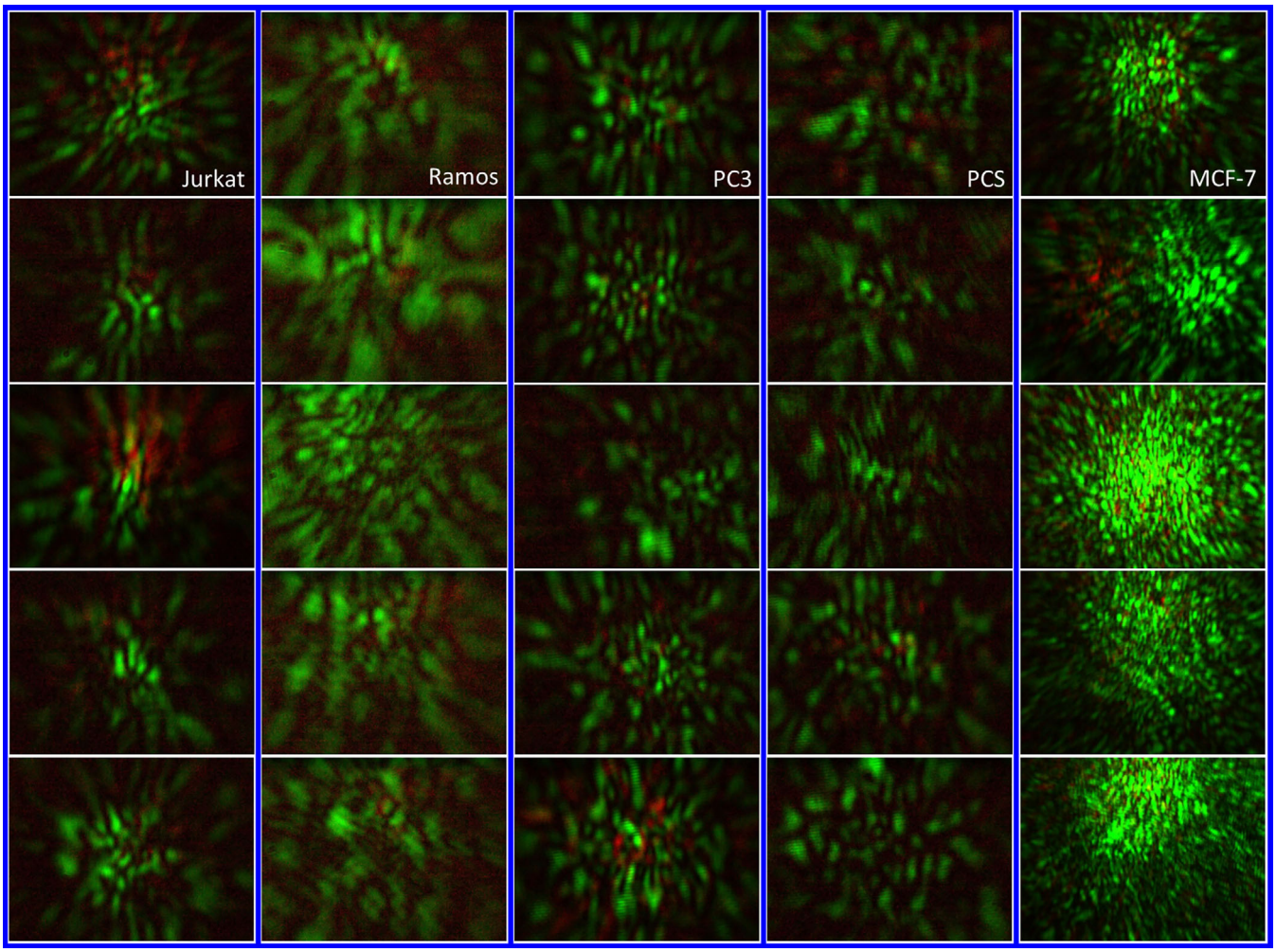


FIGURE 4 False-color combined DIs from five cells for each type with green and red colors representing, respectively the normalized intensity of s- and p-polarized scattered light

3.2 | DINet Training and classification results

Our investigation of CNN architects was inspired by the widely known structure of AlexNet that has five convolutional layers and three fully connected (FC) layers with 6.2×10^7 network parameters for optimization [29]. While high classification accuracy can be achieved, AlexNet takes long time to train due to the large number of network parameters. More importantly, its depth and high dimensionality in weight parameters makes it challenging to quantitatively analyze the cascading process of feature extraction. To accelerate training and improve understanding of the feature extraction process, we have developed DINet with 100-fold less weight parameters for improved robustness and reduced computation cost.

Training of CNNs with different architects and learning rate η was performed in batch mode. The best performance was achieved with DINet and η set at 0.01. We have also examined the possibility of overfitting by varying value of dropout r from 100% to 30% over the 250-epoch training processes. Typical results of the mean cross-entropy loss L and accuracy A_{av} obtained with the validation data are shown in Figure S1 file. The L and A_{av} curves of different r values in Figure S1 exhibit little changes after epoch 80 and indicate negligible overfitting.

The classification accuracy of five cell types by DINet in terms of $A_{av} \pm A_{std}$ were found to range from 98.7% ± 1.09% for group 5 to 99.2% ± 0.89% for group 1. The average training time for the five-round rotation is about 25 minutes. Figure 5A shows three confusion matrices of classification results by the DINet. Training and test were repeated with the expanded set of combined DIs to test the invariance of DINet to pixel translation, rotation, and blurring. The confusion matrices and classification accuracies of selected test data sets are presented in Figure 5B. Confusion matrices of all five groups in Table 1 used as test data sets and those of the expanded set are provided in

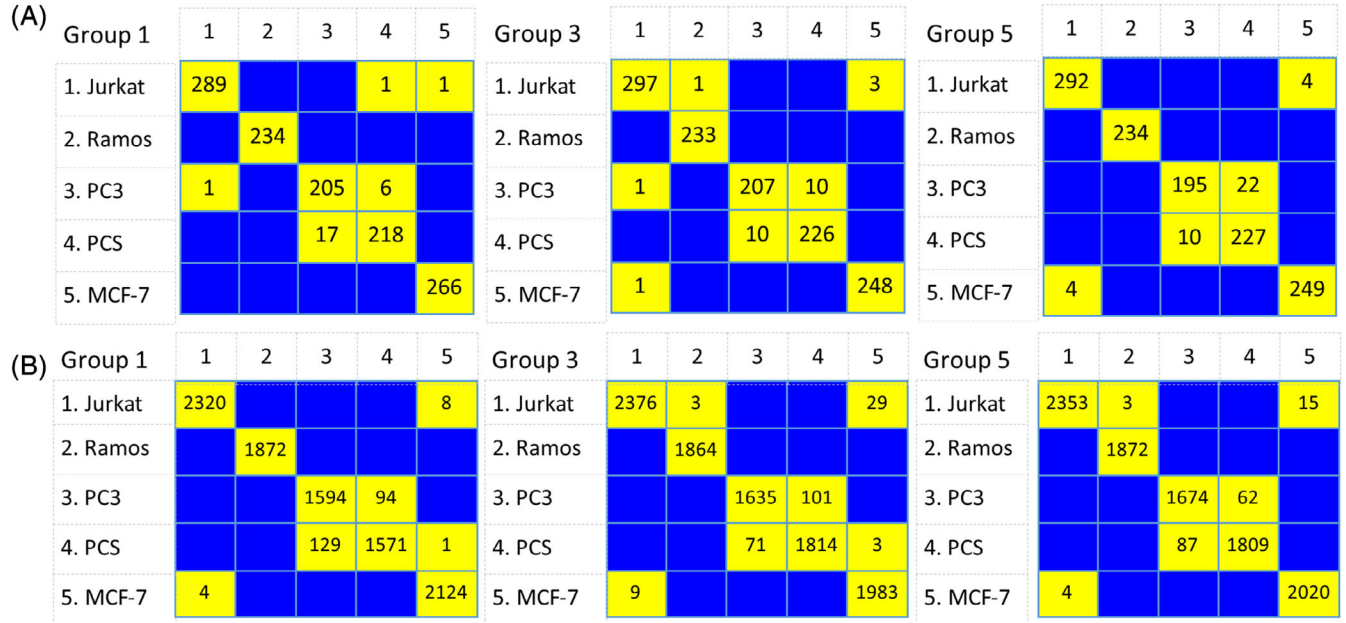


FIGURE 5 A, Confusion matrices obtained with test data set of groups 1, 3, and 5 of the measured set of combined DIs with rows as ground truth types and columns as predicted types. Blue squares indicate zero elements. The values of $A_{av} \pm A_{std}$ were given by 99.2% ± 0.89% for group 1, 99.2% ± 0.68% for group 3 and 98.7% ± 1.09% for Group 5. B, Similar results obtained with the eight times expanded set of combined DIs and corresponding values of $A_{av} \pm A_{std}$ given by 99.0% ± 1.07%, 99.1% ± 0.74% and 99.3% ± 0.68%

Figure S2 file. It is clear from these results that the pixel-based CNN classification yielded a very robust method of cell classification by combined DIs.

3.3 | Hierarchical representation of DI features

To gain insight on discriminative representation of input images across multiple pixel scales, we analyzed the outputs of each layer C_m in the optimized DINet. For each cell type, two sets of combined DIs of were selected from a test data set as correctly and incorrectly classified DIs. The 3D output array $\mathbf{y}_m$ of C_m layer as defined in Equation (1) was exported for cell c of the selected DI as a 2D feature image stack $\{Y_{mdc}: d = 1, ..., N_{md}\}$ in which $Y_{mdc}(i, j)$ was obtained from $y_m(i, j, d; c)$. Six histogram parameters $\Phi_{md}(c, p)$ were calculated from $Y_{mdc}(i, j)$ with p (=1, ..., 6) as the parameter index. Then $\Phi_{md,av}(p)$ was obtained from $\Phi_{md}(c, p)$ as the mean values of all cells in a selected DI set. Finally, the relative difference between $\Phi_{md,av}(p)$ of the two sets was determined as $\Delta_{md}(p)$ to quantify the averaged difference in parameter p for layer C_m between the correctly and incorrectly classified cells. For this study, we focus only on the parameters with significant difference or $|\Delta_{md}(p)| \geq 60\%$ as the "signature features" to reduce the effect of image noise. Five of six histogram parameters were found to satisfy this criterion, which are the mean gray-level μ, variance σ^2, entropy s, skewness γ_1, and kurtosis γ_2 with γ_2 exhibiting the highest values of $|\Delta_{md}(p)|$. The histogram parameter lacking significant difference is energy ε. Parameter definitions and details of calculating signature features from $\mathbf{y}_m$ are given in the Supporting Information file.

For quantifying variation of signature features among different layers, we calculated $n_m(p)$ as the number of 2D feature images having p with $|\Delta_{md}(p)| \geq 60\%$ among the N_{md} images of layer C_m and $|\Delta_{m,av}(p)|$ as the averaged values over the $n_m(p)$ images. The results for PC3 and PCS cells are presented in Figure 6 since they have sufficiently large numbers of incorrectly classified cells in the measured DIs among the five cell types. Study of the differences of histogram parameters between the 2D feature images of the two selected cell sets allows us to understand the hierarchical representation and selection of diffraction pattern features through the four convolutional layers in DINet's prediction for cell type.

Figure 6 shows that the values of $n_m(p)$ and $|\Delta_{md}(p)|$ for γ_2 increase markedly from C_2 to C_3 whereas the increases from C_1 to C_2 or from C_3 to C_4 are much less for both types of PC3 and PCS. These results indicate an important role in cell classification by γ_2 which characterizes the "peak-tail distribution" of histogram in the feature images of $Y_{mdc}(i, j)$. Furthermore, the ability of C_3 to extract large number of feature images with high values of $|\Delta_{m,av}(\gamma_2)|$ suggests that three convolution

layers may be sufficient for accurate classification by these signature features. To test this hypothesis, we modified the DINet by elimination of C_4 and P_4 layers and connecting the P_3 layer directly to the FC layer. The reduced DINet was then optimized into two different networks: DINet-A with only the network parameters after

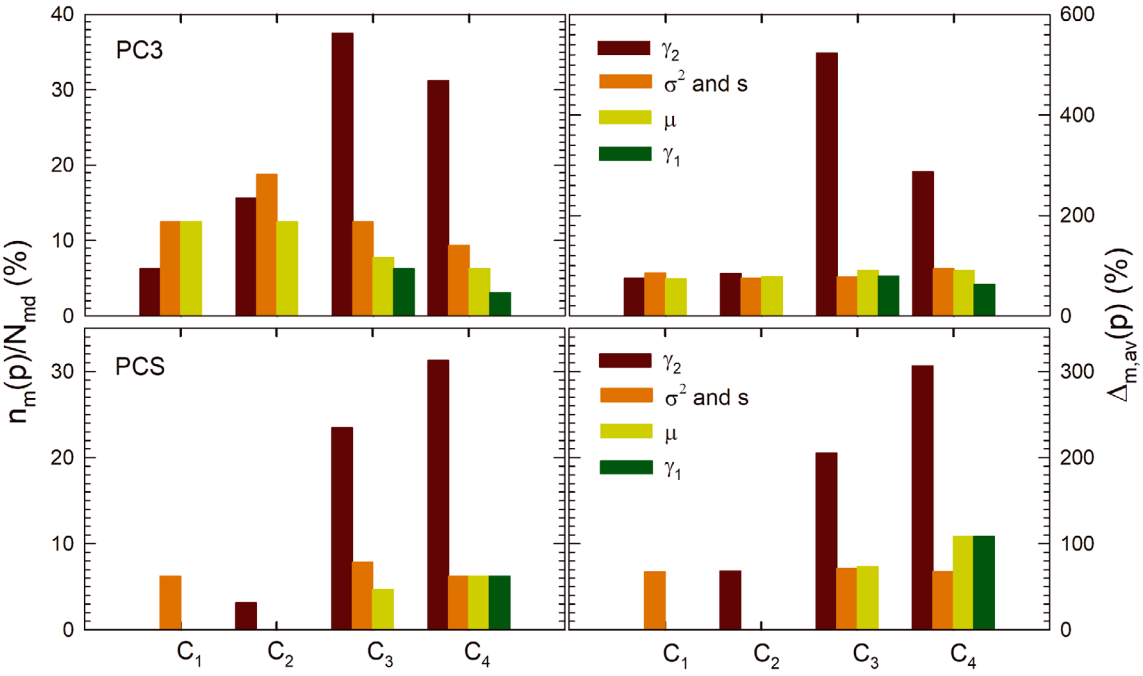


FIGURE 6 Number ratio $n_m(p)/N_{md}$ and averaged difference $|\Delta_{m,av}(p)|$ of 2D feature images with signature features by histogram parameter p determined from $y_m(i, j, d; c)$ versus layer C_m for PC3 (upper) and PCS (lower) cells. Histogram parameter of μ, σ^2, s, γ_1, and γ_2 for different values of p are defined in Appendix S1

(A)

Group 1	1	2	3	4	5
1. Jurkat	288		1	1	1
2. Ramos		234			
3. PC3	1		196	13	2
4. PCS			23	212	
5. MCF-7	3				263

Group 3	1	2	3	4	5
1. Jurkat	297	1	1		2
2. Ramos		233			
3. PC3	1		210	7	
4. PCS			15	220	1
5. MCF-7	3				246

Group 5	1	2	3	4	5
1. Jurkat	294				2
2. Ramos		234			
3. PC3			197	20	
4. PCS			13	234	
5. MCF-7	2		1	3	247

(B)

Group 1	1	2	3	4	5
1. Jurkat	288	1	1	1	
2. Ramos		234			
3. PC3			194	16	1
4. PCS			17	217	1
5. MCF-7	2			1	263

Group 3	1	2	3	4	5
1. Jurkat	297	1		1	2
2. Ramos		233			
3. PC3	1		206	11	
4. PCS			10	225	1
5. MCF-7	1				248

Group 5	1	2	3	4	5
1. Jurkat	292				4
2. Ramos		234			
3. PC3			209	8	
4. PCS			16	221	
5. MCF-7	2				251

FIGURE 7 Confusion matrices obtained with, A, DINet-A and, B, DINet-B similar to those in Figure 5. The values of $A_{av} \pm A_{std}$ of groups 1, 3, and 5 were given, respectively as (A) 98.6% ± 1.36%, 99.0% ± 0.76%, and 98.7% ± 1.24%; (B) 98.7% ± 1.27%, 99.1% ± 0.76%, and 99.0% ± 0.81%

the P_3 layer re-trained and DINet-B with all network parameters retrained. Figure 7 shows the confusion matrices of these two modified networks, which demonstrate clearly the little impact of layer C_4 on classification accuracy. However, the training time of DINet-B was found to be about the same as the DINet. For increased robustness, we keep the DINet as our choice of CNN for this study.

4 | DISCUSSION

The process of translating genotype to phenotype in cell differentiation is complex and pathway dependent [31]. Cells of the same phenotype can exhibit substantial heterogeneity in their morphology and one usually prefers immunofluorescent measurement over analysis of morphology alone for cell classification. It thus remains to be answered whether single-shot images of scattered light intensity I(**k**) under coherent excitation can be applied for rapid and accurate classification of cells in multiple types. Here we have demonstrated the feasibility of the p-DIFC approach to achieve such a goal by cross-polarized recording of I(**k**) over a limited angular range along the side scattering directions. The results presented in this report show the key role of deep networks for learning and extracting signature features from the input p-DI data by its multilayer architect. For comparison, we have classified the same set of p-DI data used for DINet development by a clustering algorithm in the GLCM parameter space. A Gaussian mixture model was combined with the hierarchical clustering algorithm to obtain stable clustering results [24]. Thirty GLCM parameters were extracted from each p-DI pair to quantify the second order correlations among the image pixels [18]. Different combination of the GLCM parameters were investigated for classification of the cells in five types and the classification accuracies were mostly below 90% with the best value of $A_{av} \pm A_{std}$ found to be 91.3% ± 7.62%, which are significantly lower than those by DINet due to much increased cross-over errors between Ramos and PCS and misclassifying MCF-7 into PC3.

Convolution of two functions as defined in Equation (1) is equivalent to their cross-correlation by coordinate reversal. Consequently, the multilayer structure of CNN classifiers provides a powerful means to quantify different orders of statistical relations among the pixel intensities of an input image. We note that $\mathbf{w}_m(d)$ as feature maps serve as "adaptive sieves" to screen useful local patterns of the input image for object classification. After optimization, these maps can be regarded as an adapted convolutional operator on the output of previous layer and generate the feature images in $\mathbf{y}_m$. For combined DIs concerned in this report, the diffraction patterns exhibit the angular distribution of the coherent light intensity I(**k**) scattered by a cell. One can argue that convolutional layers present features closely related to the correlations of I(**k**) with the optimized $\mathbf{w}_m(d)$ in succeeding orders as m increases from 1 to 4. DINet can thus yield much improved performance due to its ability to extract high-order correlation features from the combined DIs, which is entirely different from the conventional image analysis algorithms like the GLCM method. Indeed, the results in Figure 6 present strong evidences that the high-order correlations revealed by γ_2 of $\{Y_{mdc}: d = 1, ..., N_{md}\}$ in the deep layers of C_3 and C_4 are essential for the excellent performance of DINet. In comparison, the lower-order correlations revealed by the shallow layers of C_1 and C_2 do not provide sufficient signature features for accurate classification. The analysis of layers' output arrays presented here thus provides a practical approach to explore quantitatively the effect of network depth on pattern recognition [32].

5 | CONCLUSION

We have developed a compact deep neural network for classifying five cultured cells types by extracting diffraction pattern features from p-DI data. It has been shown that the unique ability of deep network to learn and extract high order correlation among the input image pixels is critical for accurate cell classification. These results demonstrate the strong potentials of accurate and rapid cell classification by the p-DIFC method with single-shot diffraction images combined with deep neural networks.

ACKNOWLEDGMENTS

XHH acknowledge research grant supports from East Carolina University (IRA-2016) and from Golfers Against Cancer (2012–13-GAC).

ORCID

Xin-Hua Hu https://orcid.org/0000-0002-4353-9028

REFERENCES

[1] M. Sadelain, I. Rivière, S. Riddell, *Nature* **2017**, *545*, 423.
[2] C. Klijn, S. Durinck, E. W. Stawiski, P. M. Haverty, Z. Jiang, H. Liu, J. Degenhardt, O. Mayba, F. Gnad, J. Liu, G. Pau, J. Reeder, Y. Cao, K. Mukhyala, S. K. Selvaraj, M. Yu, G. J. Zynda, M. J. Brauer, T. D. Wu, R. C. Gentleman, G. Manning, R. L. Yauch, R. Bourgon, D. Stokoe, Z. Modrusan, R. M. Neve, F. J. de Sauvage, J. Settleman, S. Seshagiri, Z. Zhang, *Nat. Biotechnol.* **2015**, *33*, 306.
[3] E. M. Christiansen, S. J. Yang, D. M. Ando, A. Javaherian, G. Skibinski, S. Lipnick, E. Mount, A. O'Neil, K. Shah, A. K. Lee, P. Goyal, W. Fedus, R. Poplin, A. Esteva,

M. Berndl, L. L. Rubin, P. Nelson, S. Finkbeiner, *Cell* **2018**, *173*, 792.
[4] N. C. Pégard, H.-Y. Liu, N. Antipa, M. Gerlock, H. Adesnik, L. Waller, *Optica* **2016**, *3*, 517.
[5] J. C. Waters, *J. Cell Biol.* **2009**, *185*, 1135.
[6] U. Schnell, F. Dijk, K. A. Sjollema, B. N. G. Giepmans, *Nat. Methods* **2012**, *9*, 152.
[7] W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, M. S. Feld, *Opt. Lett.* **2008**, *33*, 171.
[8] T. Kim, R. Zhou, M. Mir, S. D. Babacan, P. S. Carney, L. L. Goddard, G. Popescu, *Nat. Photonics* **2014**, *8*, 256.
[9] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, D. Psaltis, *Optica* **2015**, *2*, 517.
[10] A. Mahjoubfar, D. V. Churkin, S. Barland, N. Broderick, S. K. Turitsyn, B. Jalali, *Nat. Photonics* **2017**, *11*, 341.
[11] Z. A. Steelman, W. J. Eldridge, J. B. Weintraub, A. Wax, *J. Biophotonics* **2017**, *10*, 1714.
[12] C. L. Chen, A. Mahjoubfar, L. C. Tai, I. K. Blaby, A. Huang, K. R. Niazi, B. Jalali, *Sci. Rep.* **2016**, *6*, 21471.
[13] K. M. Jacobs, J. Q. Lu, X. H. Hu, *Opt. Lett.* **2009**, *34*, 2985.
[14] K. M. Jacobs, L. V. Yang, J. Ding, A. E. Ekpenyong, R. Castellone, J. Q. Lu, X. H. Hu, *J. Biophotonics* **2009**, *2*, 521.
[15] Y. Sa, Y. Feng, K. M. Jacobs, J. Yang, R. Pan, I. Gkigkitzis, J. Q. Lu, X. H. Hu, *Cytometry A* **2013**, *83*, 1027.
[16] J. Zhang, Y. Feng, M. S. Moran, J. Q. Lu, L. V. Yang, Y. Sa, N. Zhang, L. Dong, X. H. Hu, *Opt. Express* **2013**, *21*, 24819.
[17] Y. Feng, N. Zhang, K. M. Jacobs, W. Jiang, L. V. Yang, Z. Li, J. Zhang, J. Q. Lu, X. H. Hu, *Cytometry A* **2014**, *85*, 817.
[18] W. Jiang, J. Q. Lu, L. V. Yang, Y. Sa, Y. Feng, J. Ding, X. H. Hu, *J. Biomed. Opt.* **2016**, *21*, 071102.
[19] H. Wang, Y. Feng, Y. Sa, J. Q. Lu, J. Ding, J. Zhang, X.-H. Hu, *Pattern Recogn.* **2016**, *61*, 234.
[20] K. Dong, Y. Feng, K. M. Jacobs, J. Q. Lu, R. S. Brock, L. V. Yang, F. E. Bertrand, M. A. Farwell, X. H. Hu, *Biomed. Opt. Express* **2011**, *2*, 1717.
[21] Y. Sa, J. Zhang, M. S. Moran, J. Q. Lu, Y. Feng, X. H. Hu, *Opt. Express* **2012**, *20*, 22245.
[22] J. Zhang, Y. Feng, W. Jiang, J. Q. Lu, Y. Sa, J. Ding, X. H. Hu, *Opt. Express* **2016**, *24*, 366.
[23] S. Wang, J. Liu, J. Q. Lu, W. Wang, S. A. Al-Qaysi, Y. Xu, W. Jiang, X. H. Hu, *J. Biophotonics* **2019**, *12*, e201800287.
[24] W. Wang, Y. Wen, J. Q. Lu, L. Zhao, S. A. Al-Qaysi, X.-H. Hu, *J. Quant. Spectrosc. Radiat. Transf.* **2019**, *224*, 453.
[25] Y. Zhang, Y. Feng, C. R. Justus, W. Jiang, Z. Li, J. Q. Lu, R. S. Brock, M. K. McPeek, D. A. Weidner, L. V. Yang, X. H. Hu, *Integr. Biol.* **2012**, *4*, 1428.
[26] Y. Wen, Z. Chen, J. Lu, E. Ables, J. L. Scemama, L. V. Yang, J. Q. Lu, X. H. Hu, *PLoS ONE* **2017**, *12*, e0184726.
[27] R. Pan, Y. Feng, Y. Sa, J. Q. Lu, K. M. Jacobs, X. H. Hu, *Opt. Express* **2014**, *22*, 31568.
[28] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: a system for large-scale machine learning. *Proc 12th USENIX Conf Oper Syst Des Implement*, Savannah, GA, USA2016.
[29] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks. in *Advances in neural information processing systems* (Eds: F. Pereira, C. Burges, L. Bottou, K. Weinberger), Curran Associates Inc., Red Hook, NY **2012**, pp. 1097.
[30] M. Sokolova, G. Lapalme, *Inf. Process. Manag.* **2009**, *45*, 427.
[31] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, M. W. Covert, *Cell* **2012**, *150*, 389.
[32] K. Kawaguchi, Y. Bengio, *Neural Netw.* **2019**, *118*, 167.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Jin J, Lu JQ, Wen Y, Tian P, Hu X-H. Deep learning of diffraction image patterns for accurate classification of five cell types. *J. Biophotonics*. 2019;e201900242. https://doi.org/10.1002/jbio.201900242