Supplementary Information

Deep-learning-assisted biophysical imaging cytometry at massive throughput delineates cell population heterogeneity

Dickson M.D. Siu,[a] Kelvin C.M. Lee,[a] Michelle C.K. Lo,[a] Shobana V. Stassen,[a] Maolin Wang,[a] Iris Z.Q. Zhang,[a] Hayden K.H. So,[a] Godfrey C.F. Chan,[b] Kathryn S.E. Cheah,[c] Kenneth K.Y. Wong,[a] Michael K.Y. Hsin,[d] James C.M. Ho,[e] Kevin K. Tsia*[a]

[a]Department of Electrical and Electronic Engineering, Faculty of Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong
[b]Department of Pediatrics and Adolescent Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam Road, Hong Kong
[c]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam Road, Hong Kong
[d]Department of Cardiothoracic Surgery, Queen Mary Hospital, Pokfulam Road, Hong Kong
[e]Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam Road, Hong Kong

Table of content

Supplementary Methods

Deep neural network and transfer learning. To test the label-free specificity of optophysical phenotyping in multi-ATOM, a neural-network-based model was built to conduct lung cancer subtype classification and H2170 detection in spike-in test.

Lung cancer subtype classification - The neural network model was employed to classify 3 lung cancer subtypes (adenocarcinoma, squamous cell carcinoma, small cell carcinoma) from 7 cell lines (H358, H1975, HCC827, H520, H2170, H526, H69) based on the single-cell images captured from the multi-ATOM system. Three replicates (batches) of cell image datasets, each consisting of ~120,000 cells per cell line, were obtained on different dates. 69 optophysical features were extracted as the input of the neural network. A dataset with 14,000 cells, having 1,000 cells per cell line from 2 of the batches, was used to train the classification model, whereas the remaining batch was used as the test set. The model was first trained with the training set for 25 epochs at a learning rate of 0.005 with a mini-batch size of 128.

Here, transfer learning was applied to the neural network models for adapting the batch-to-batch variations. With such, the batch effects could be practically overcome by using a small set of labeled data from the new batch each time, streamlining the workflow for time-critical laboratory routine. The general approach of transfer learning starts from training the neural network model with the source datasets (Batch A and B in our case). Then, the model was further trained using another (small) dataset for another task (i.e. classification of batch C). Upon successful transfer of knowledge, the resultant training curve will have a few advantages, including a higher intercept (Improved starting position), a higher initial slope (Fewer epochs or data required) and a higher resultant asymptote (i.e. higher classification accuracy). [1] It has been demonstrated as a promising model in single-cell analysis [2-4], as in this study, similar concepts of minimizing batch effect have yet been commonly adopted in image-based cellular analysis. To overcome the batch effects among the training and test data, 100 cells per cell line from the test data were isolated to form a transfer learning dataset. Such dataset was further used to train the neural network for 120 epochs, at a learning rate of 0.005 with a mini-batch size of 128. Python was employed to develop the classification model. With the developed classification model, a test set of 105,000 cells (with 15,000 cells per cell line in the test batch) was used to assess the classification accuracy. To facilitate a thorough examination, the three batches were interchanged between the source datasets and test dataset, which is also the transfer learning dataset. Through integrating the performance of each models, the prediction accuracies were then reported within a range. By increasing the size of the transfer learning dataset, the accuracy improved and levelled off at around 95% with 500 cells per cell line (**Fig. S8**). The prediction accuracies of the training set and test set were monitored continuously to prevent overfitting, while the leveling off of the accuracy in the test set helps in avoiding underfitting.

The overall classification procedures can be divided into 3 steps, training, transfer learning retraining and testing. For all of these computations, we used a consumer grade computer with a CPU (6 cores, 12 threads; 4.00 GHz), a 64 GB RAM. We only rely on the CPU for the computations. The training time takes the longest, of around 0.346 s per epoch. Then, the transfer learning could be accomplished with a speed of 10 times faster due to the reduction of sample size, which takes around 0.028 s per epoch. With the training accomplished, the testing was done at a speed of 1,133,624 cells/s at last.

Spike-in test - In the spike-in test, the neural-network-based model was used to spot the H2170 from peripheral blood mononuclear cell (PBMC). In order to prove the significance of all the optophysical phenotypes, 3 classification models were trained: one was trained and tested only based on 2 phenotypes (cell area and volume), one was trained and tested based on easily measurable bulk phenotypes (volume, attenuation density and dry mass density), whereas the other are based on 81 optophysical phenotypes. To train the first two models, a training set with 1,000 H2170 cells and 5,000 PBMCs was used. For the model with 81 phenotypes, the training set was composed of 1,000 H2170 cells and 6,000 PBMCs. They

were trained for 200 epochs at a learning rate of 0.0001 with a mini-batch size of 128 respectively. As in the lung cancer subtype classification, Python was employed to develop the classification model in a consumer grade computer with a CPU (6 cores, 12 threads; 4.00 GHz), a 64 GB RAM. Following that, 3 test sets with different spike ratios of 1 in 1,000 (H2170: 234 cells, PBMC: 26,618 cells), 1 in 10,000 (H2170: 26 cells, PBMC: 118,762 cells) and 1 in 100,000 (H2170: 4 cells, PBMC: 144,121 cells) were employed to evaluate their classification performance.

Supplementary Note S1: Full list of optophysical phenotypes and their hierarchical categories

Bulk		Area	Local	Optical Density	BF STD Var
		Volume			BF STD Skewness
		Circularity			BF STD Kurtosis
		Eccentricity			BF STD Range
		Aspect Ratio			BF STD Peak
		Orientation			BF STD Min
		Dry Mass			BF STD Centroid Displacement
Global	Optical Density	Attenuation Density			BF STD Radial Distribution
		Amplitude Var			BF Fiber Texture Centroid Displacement
		Amplitude Skewness			BF Fiber Texture Radial Distribution
		Amplitude Kurtosis			BF Fiber Texture Pixel >Upper Percentile
		Peak Amplitude			BF Fiber Texture Pixel >Median
		Peak Absorption			BF Fiber Mean
		Amplitude Range			BF Fiber Variance
	Mass Density	Dry Mass Density			BF Fiber Skewness
		Dry Mass Var			BF Fiber Kurtosis
		Dry Mass Skewness		Mass Density	Phase STD Mean
		Dry Mass Radial Distribution			Phase STD Var
		Dry Mass Centroid Displacement			Phase STD Skewness
		Peak Phase			Phase STD Kurtosis
		Phase Var			Phase STD Centroid Displacement
		Phase Skewness			Phase STD Radial Distribution
		Phase Kurtosis			Fit Texture Mean
		Phase Range			Fit Texture Variance
		Phase Min			Fit Texture Skewness
		Phase Radial Distribution			Fit Texture Kurtosis
		Phase Centroid Displacement			Fit Texture Centroid Displacement
		Mean Phase Arrangement			Fit Texture Radial Distribution
		Phase Arrangement Var			Phase Entropy Mean
		Phase Arrangement Skewness			Phase Entropy Var
		Phase Orientation Var			Phase Entropy Skewness
		Phase Orientation Kurtosis			Phase Entropy Kurtosis
Local	Optical Density	BF Entropy Mean			Phase Entropy Centroid Displacement
		BF Entropy Var			Phase Entropy Radial Distribution
		BF Entropy Skewness			Phase Fiber Centroid Displacement
		BF Entropy Kurtosis			Phase Fiber Radial Distribution
		BF Entropy Range			Phase Fiber Pixel >Upper Percentile
		BF Entropy Peak			Phase Fiber Pixel >Median
		BF Entropy Min			Phase Fiber Mean
		BF Entropy Centroid Displacement			Phase Fiber Var
		BF Entropy Radial Distribution			Phase Fiber Skewness
		BF STD Mean			Phase Fiber Kurtosis

All the optophysical phenotypes were extracted with a custom MATLAB code. Optical density phenotypes were based on the brightfield (BF) image of the cell whereas mass density phenotypes were extracted dry mass density map which is converted from the quantitative phase image (ϕ) using the well-

known linear relationship between refractive index and mass density of most intracellular biomolecules. The slope of this relationship, d*n*/d*c*, is called the specific refractive increment. Specific refractive increments for most biomolecules, (especially those for proteins and nucleic acids) fall within a very narrow range (0.19 ml/g) [5] and thus permits valid evaluation of cell mass inferred from the quantitative phase (ϕ).

Bulk features

Bulk features were extracted according to the mask of the cell in quantitative phase image, using basic thresholding. They describe the cell size, cell mass, and the cell shape (i.e. Circularity, Eccentricity, Aspect Ratio, Orientation).

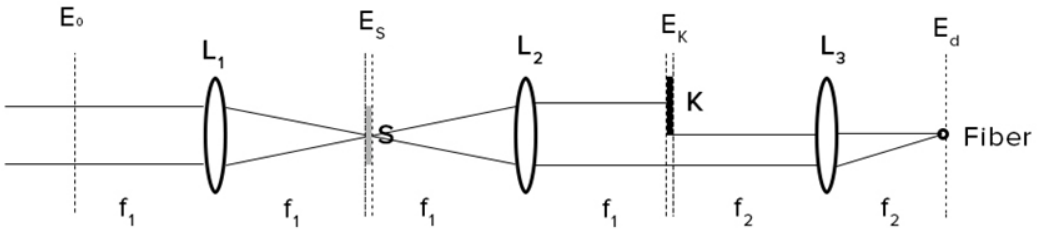
Global texture features

In each BF and ϕ images, the global texture phenotypes were extracted based on the statistical distribution of the gray-scale values in the images. They include four basic statistical moments of the global distribution (i.e. mean, variance, skewness and kurtosis), and the peak, minimum values and the range of the distribution. Also included is the dry mass density, which is extracted based on the assumption that the cell in suspension is in spherical shape. The phase arrangement phenotypes characterize the phase distribution along the radial directions, i.e. distribution of phase times its corresponding radial position. The phase orientation phenotypes on the other hand describe the relationship of phase values, its angular position and angular "repetitiveness". The phase values were first represented in the angular coordinates. Then the distribution was Fourier transformed to obtain a distribution of phase in the angular frequency domain. The statistical moments of this distribution were used as the phenotypes. We also quantified the centroid displacement and radial distribution of the mass density phenotypes. Centroid displacement measures the displacement of the weighted centroid of the mass/phase from the unweighted centroid obtained from the mask alone. Radial distribution characterizes the tendency of distribution going closer to the edge or to the centre of the cell.

Local texture features

To extract the local texture phenotypes, various local filters were used. They include entropy filter with a kernel size of 2μm, standard deviation filter with kernel size of 1 μm, and Hessian-based multiscale filter. The features extracted with these filters have "Entropy", "STD" for optical density features and mass density features, and "Fiber" in their feature names. Hence, Phase STD, Phase Entropy and Phase Fiber phenotypes are equivalent to BF STD, BF Entropy and BF Fiber phenotypes, but performed in the quantitative phase map of the cell. We also quantified the centroid displacement and radial distribution of these local texture features. Finally, the Fit Texture phenotypes were obtained by characterizing the statistical moments of the profile that emphasizes the high spatial frequency of the phase. It was obtained by subtracting the phase profile of the cell with a smoothed phase profile of the cell (computed from a fitted polynomial surface, along the x and y directions up to the 5th degree). The subtracted profile thus contains the high spatial frequency details of the cell.

Supplementary Note S2: Theory of phase retrieval in multi-ATOM



Multi-ATOM simplified schematic.

Here we describe a theoretical model using Fourier optics analysis to validate the phase-gradient and thus the phase quantified by multi-ATOM. In this analysis, the aperture size of the optics is assumed to be significantly larger in diameter than the extent of light field in the x-y plane. The overall schematic is shown in **simplified schematic**. In multi-ATOM, when a focused spectrally-encoded light illuminates on a biological cells on the sample plane S (focused by the objective lens L1), the local phase gradient of the cell would create a wavefront tilt in each *minimally resolvable* spectrally-encoded beam along the spectral shower (i.e. along y-direction). For clarity without loss of generality, here we consider only one minimally resolvable spectrally-encoded beam, denoted as $E_0(x_i, y_i)$ as the input illuminating field (as shown in **simplified schematic**). (x_s, y_s) are the coordinates on the front focal plane of L1. We also omit the diffraction grating between the L2 and knife-edge as it does not alter beam propagation characteristic along the x-direction. Based on the above configuration, the field right after the sample plane, which has the amplitude profile of $A(x_s, y_s)$ and phase profile of $\phi(x_s, y_s)$, is expressed as:

$$E_s(x_s, y_s) = A(x_s, y_s)e^{j\phi(x_s, y_s)} \cdot \frac{1}{j\lambda f_1}\mathcal{F}(E_0(x_i, y_i)) \quad (1)$$

where $\mathcal{F}$ stands for 2D Fourier transform operation (the notation follows the convention adopted in ref. [6]). Note that the phase profile of $\phi(x_s, y_s)$ is referred to the relative phase to the background phase of the medium within which the cell is suspended; λ is the wavelength of the light; and f_1 is the focal length of the lens L1. (x_s, y_s) is the coordinate on the sample plane. Subsequently, a knife-edge K, located on the Fourier plane of the sample plane S, is used to half-block the light field along the x-direction – an essential procedure to decode the quantitative phase. The field is thus expressed as:

$$E_k(x_k, y_k) = \frac{1 + sgn(x_k)}{2} \cdot \frac{1}{j\lambda f_1}\mathcal{F}(E_s(x_s, y_s)) \quad (2)$$

where $sgn(x_k)$ is the sign function along the x-direction; and (x_k, y_k) is the coordinate on the knife-edge plane. The lens L2 has a focal length of f_1. Afterwards, the light is coupled to a single-mode fiber positioned on the Fourier plane of the knife-edge K (i.e. conjugate plane of sample plane S) through the relay lens L3. As a result, the detected light field is written as:

$$E_d(x_d, y_d) = \frac{1}{j\lambda f_2}\mathcal{F}(E_k(x_k, y_k)) \quad (3)$$

where f_2 is the focal length of the relay lens L3; (x_d, y_d) is the coordinate on the detection plane. We further employ smooth phase approximation [7], $e^{j\phi(x_s, y_s)} \approx 1 + j\phi(x_s, y_s) \approx 1 + j\frac{\partial\phi}{\partial x_s}$, which is generally valid for biological cells. Hence, the detection field can then be written as,

$$E_d = \frac{A}{2\lambda^3 {f_1}^2 f_2}(C_1 + C_2 \frac{\partial \phi}{\partial x_s}) \tag{4}$$

where $C_1 = \left[-\mathcal{F}(E_0) - \frac{1}{x_d} * \mathcal{F}(E_0)\right] + j\left[\frac{1}{x_d} * \mathcal{F}(E_0) - \mathcal{F}(E_0)\right]$ and $C_2 = \mathcal{F}(E_0) \cdot x_d - j\{\frac{1}{x_d} * [\mathcal{F}(E_0) \cdot x_d]$. Note that the coordinate arguments are dropped for the sake of simplicity. Hence, the integrated intensity received at the fiber is

$$I_x{}^+ = \iint |E_d|^2 \, dx_d dy_d = B \cdot A + A \cdot C_3 \frac{\partial \phi}{\partial x_s} \tag{5}$$

where $B = \frac{2}{6\lambda^6 {f_1}^4 {f_2}^2} \iint \left\{[\mathcal{F}(E_0)]^2 + \left[\frac{1}{x_d} * \mathcal{F}(E_0)\right]^2\right\} dx_d dy_d$ and $C_3 = \frac{2}{6\lambda^6 {f_1}^4 {f_2}^2} \iint \mathcal{F}(E_0) \left\{\frac{1}{x_d} * [\mathcal{F}(E_0) \cdot x_d] - \frac{1}{x_d} * \mathcal{F}(E_0)\right\} dx_d dy_d$. When the knife edge is applied from the opposite direction, the intensity is

$$I_x{}^- = \iint |E_d|^2 \, dx_d dy_d = B \cdot A - A \cdot C_3 \frac{\partial \phi}{\partial x_s} \tag{6}$$

Thus, the phase gradient can be extracted from the subtraction of the intensities obtained from opposite knife edges (i.e. Eqs (5-6)), i.e.

$$\frac{I_x{}^+ - I_x{}^-}{I_x{}^+ + I_x{}^-} = \frac{C_3}{B} \frac{\partial \phi}{\partial x_s} \tag{7}$$

And the same derivation can be applied to the knife edges applied in the y-direction. It proves that the intensity-only measurement in multi-ATOM yields phase-gradient information along both the x- and y-directions. By applying this interrogation across the sample plane (as multi-ATOM is essentially a scanning technology), phase gradient images ($\nabla\phi(x,y) = \frac{\partial \phi}{\partial x_s}(x,y) + i\frac{\partial \phi}{\partial y_s}(x,y)$) can be obtained. The QPI can then be obtained by applying complex Fourier integration on the phase gradient images,

$$\phi(x,y) = CF \cdot Im\left\{\mathcal{F}^{-1}\{NF \cdot \mathcal{F}[\nabla\phi(x,y)]\}\right\} \tag{8}$$

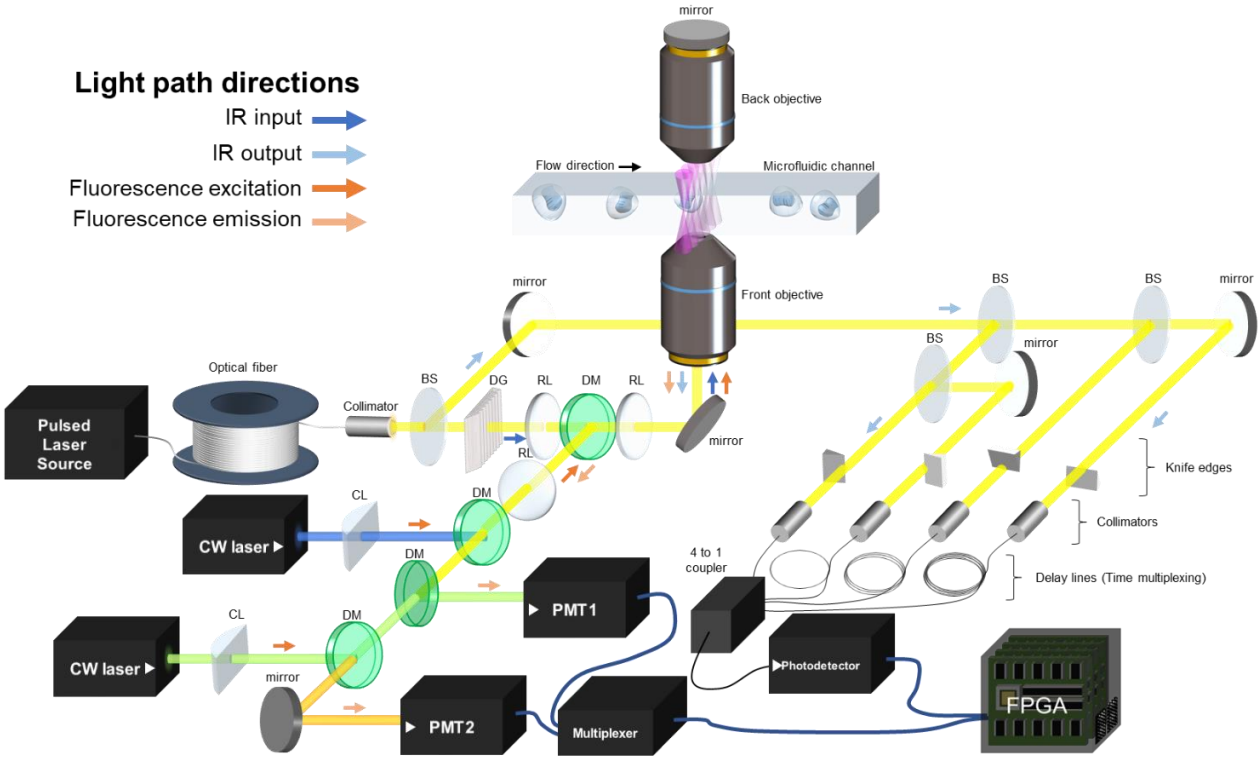
$$where\ NF = \begin{cases} FOV/[2\pi j \cdot k(x,y)] & , k(x,y) \neq 0 \\ 0 & , k(x,y) = 0 \end{cases}$$

where *Im* is the imaginary part of a complex number; and $\mathcal{F}^{-1}$ is inverse Fourier transform operator; NF is a normalization factor for quantifying the phase and avoiding singularity in the integration operation; $k(x,y)$ is the 2D wavenumber; FOV is the 2D field-of-view; CF is the calibration factor for correcting the systematic phase deviation arise from non-ideal system setting. On the other hand, the bright-field, i.e. amplitude, image of the cell ($BF(x,y)$) is the sum of two images obtained from opposite knife edges normalized by the background (i.e. B, regions without samples).

$$BF(x,y) = \frac{1}{B}(I_x{}^+ + I_x{}^-) = \frac{1}{B}(I_y{}^+ + I_y{}^-) \tag{9}$$

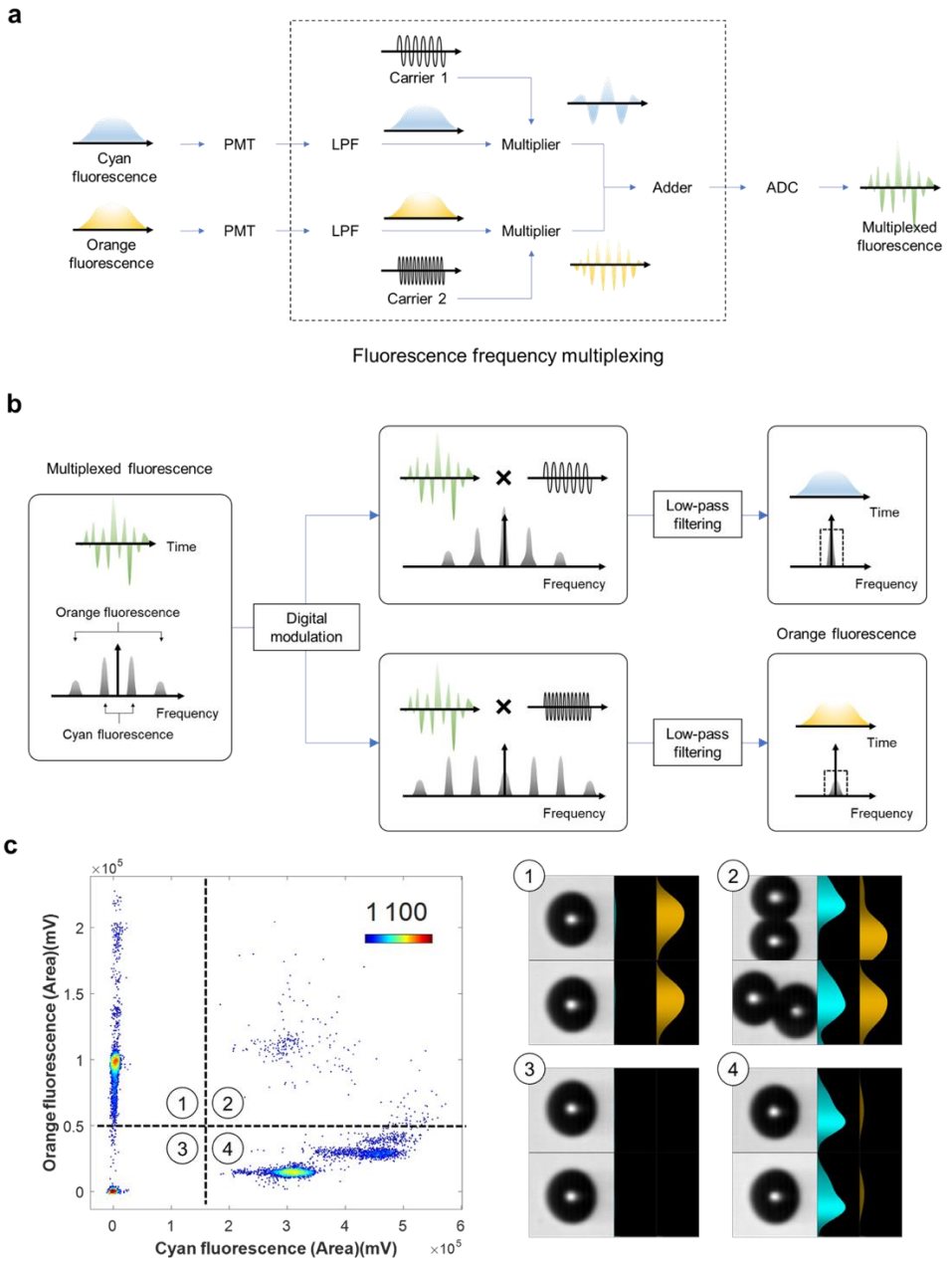
We validate the complex field information (i.e. amplitude and quantitative phase) extracted from multi-ATOM by Fourier light-scattering analysis, as shown in **Fig. S3**, compared with both Mie-scattering theory as well as the experimentally measured results.

Figure S1: Imaging setup



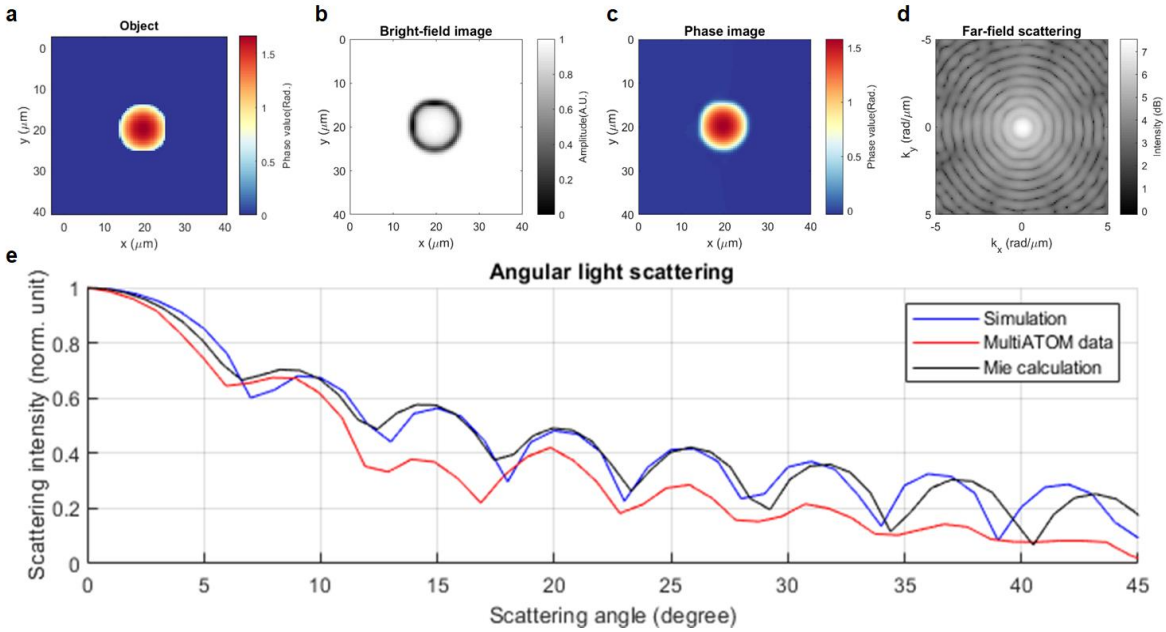
A general schematic of the multi-ATOM system integrated with the fluorescence detection module. BS: Beam splitter. DG: Diffraction grating. RL: Relay lens. DM: Dichroic mirror. CL: Cylindrical lens. The optical system consists of two integrated parts: the multi-ATOM and fluorescence detection module. Without using dedicated interferometry, multi-ATOM acquires the complex-field image information at high speed by optical time-stretch combined with multiplexed differential phase-gradient contrast encoding. For each cell, the 2D complex-field information (i.e., bright-field and QPI) were retrieved from the four different phase-gradient contrasts based on an algorithm using complex Fourier integration.[8] In the fluorescence detection module, two continuous wave (CW) lasers (wavelength: 488nm and 532nm) were employed to generate line-shaped fluorescence excitation, that were spatially overlapped with the multi-ATOM illumination (**Fig. 1**). The two epi-fluorescence signals were detected by two photomultiplier tubes (PMT) separately. In the analog electronics backend, we multiplexed the PMT-detected signals by frequency modulation (11.8 MHz and 35.4 MHz respectively, using a multichannel direct digital synthesizer). The multiplexed signals were then separated by digital demodulation and low-pass filtering (**Fig. S2**). The same FPGA was configured to synchronously obtain the signal from multi-ATOM and fluorescence detection from each single cell at high-speed.

Figure S2: Fluorescence detection scheme



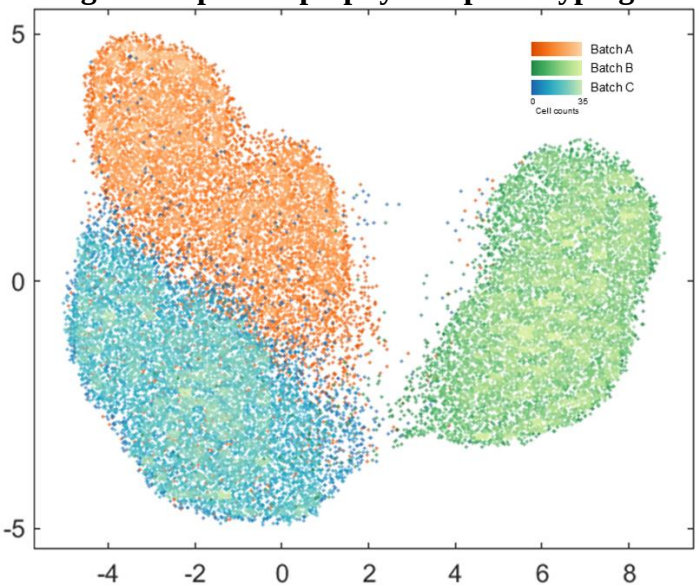
(a) Working principle of the fluorescence signal multiplexing scheme. Two channels of fluorescence signal were detected by two separate photomultiplier tubes (PMTs). Then the signals were modulated by two carrier signals at the frequencies much higher than the fluorescence signal bandwidth (100 kHz). The modulated signals were then combined together and digitized by an analog-to-digital conversion (ADC) circuit. (b) Working principle of the fluorescence signal recovery (demodulation). The multiplexed signal was modulated with their corresponding carrier frequencies respectively. A low-pass filter was applied to retrieve the selected fluorescence channel. (c) Experimental verification of 2-color fluorescence multiplexing scheme in multi-ATOM using a mixed population of microspheres, consisting of label-free and 2-color fluorescently-labelled microspheres. (Left) A 2D scatter plot shows four distinct regions, each of which corresponds (1) orange fluorescent beads and (2) doublets with orange and cyan fluorescence (3) label-free beads and (4) cyan fluorescent beads. The cyan and orange fluorescence signals were modulated at the frequencies of 11.8 MHz, 35.4 MHz. (Right) Examples of the (left) bead images (middle) cyan fluorescence signal, and (right) orange fluorescence signal that correspond to the 4 regions of the scatter plot respectively.

Figure S3: Fourier light scattering analysis and validation of multi-ATOM



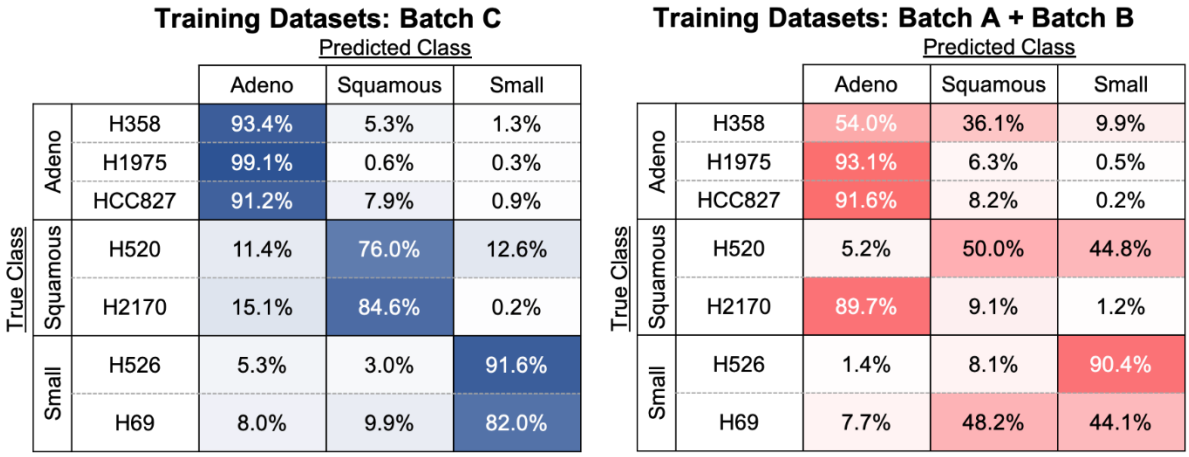
(a) A phase profile of the simulated sphere (diameter of 11.8 μm) used for analysis. (b-c) The retrieved BF and quantitative phase (ϕ) images based on our Fourier optics model. (d) The far-field light scattering pattern computed from the complex field obtained from (b) and (c). (e) The angular light scattering (ALS) profile computed from (d) (blue). This is consistent with the ALS of the bead with the same size, obtained experimentally from the multi-ATOM system (red). They also in good agreement with the ALS of the sphere evaluated based on the classical Mie light scattering theory (black).

Figure S4: Batch effect in single-cell spatial optophysical phenotyping



An example of the batch effect observed in the optophysical phenotypic profile of the H358 cells. The cells are imaged by multi-ATOM on three different days using the same setup and system configuration. We used UMAP to visualize the high-dimensional data on the three different datasets. The batch effect is manifested by the three separate clusters in the UMAP plot.

Figure S5: The influence of batch effect on neural network classification - comparing between the single-batch-trained model and the multi-batch-trained model



Training Datasets: Batch C

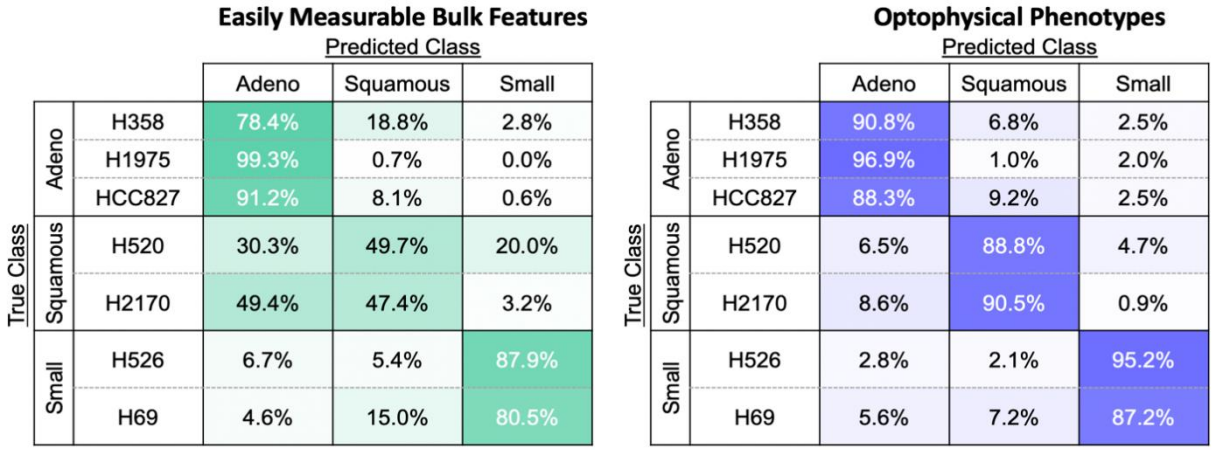
True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	93.4%	5.3%	1.3%
	H1975	99.1%	0.6%	0.3%
	HCC827	91.2%	7.9%	0.9%
Squamous	H520	11.4%	76.0%	12.6%
	H2170	15.1%	84.6%	0.2%
Small	H526	5.3%	3.0%	91.6%
	H69	8.0%	9.9%	82.0%

Training Datasets: Batch A + Batch B

True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	54.0%	36.1%	9.9%
	H1975	93.1%	6.3%	0.5%
	HCC827	91.6%	8.2%	0.2%
Squamous	H520	5.2%	50.0%	44.8%
	H2170	89.7%	9.1%	1.2%
Small	H526	1.4%	8.1%	90.4%
	H69	7.7%	48.2%	44.1%

Confusion matrices (left) of single-batch-trained model (training datasets: Batch C) and (right) of multi-batch-trained model (training datasets: Batch A + B). For both of the model, they were trained with 14,000 cells from the training datasets and tested with 105,000 cells from Batch C. The prediction results are reported above. The superior performance of the single-batch-trained model when compared with the others indicates the presence of batch effect. Color gradient of the grids are set proportional to the value within it. Darker colors mean higher values. The highest value in each row is marked with white texts.

Figure S6: The significance of the high-dimensional optophysical phenotypes in improving the classification accuracy in comparing with the common biophysical features



Easily Measurable Bulk Features

True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	78.4%	18.8%	2.8%
	H1975	99.3%	0.7%	0.0%
	HCC827	91.2%	8.1%	0.6%
Squamous	H520	30.3%	49.7%	20.0%
	H2170	49.4%	47.4%	3.2%
Small	H526	6.7%	5.4%	87.9%
	H69	4.6%	15.0%	80.5%

Optophysical Phenotypes

True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	90.8%	6.8%	2.5%
	H1975	96.9%	1.0%	2.0%
	HCC827	88.3%	9.2%	2.5%
Squamous	H520	6.5%	88.8%	4.7%
	H2170	8.6%	90.5%	0.9%
Small	H526	2.8%	2.1%	95.2%
	H69	5.6%	7.2%	87.2%

Confusion matrices (left) with the common biophysical features (i.e. cell size, the averaged mass/optical density) and (right) with the high-dimensional optophysical phenotypic profile. The transfer-learning-assisted results are reported above. Color gradient of the grids are set proportional to the value within it. Darker colors mean higher values. The highest value in the row is marked with white texts.

Figure S7: The significance of transfer-learning-assisted neural network in improving the classification accuracy using label-free optophysical phenotypes

Before Transfer Learning

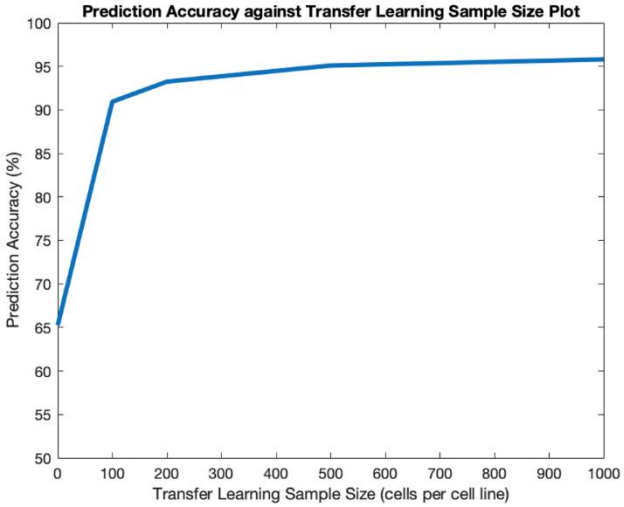
True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	54.0%	36.1%	9.9%
	H1975	93.1%	6.3%	0.5%
	HCC827	91.6%	8.2%	0.2%
Squamous	H520	5.2%	50.0%	44.8%
	H2170	89.7%	9.1%	1.2%
Small	H526	1.4%	8.1%	90.4%
	H69	7.7%	48.2%	44.1%

After Transfer Learning

True Class		Predicted Class: Adeno	Squamous	Small
Adeno	H358	90.8%	6.8%	2.5%
	H1975	96.9%	1.0%	2.0%
	HCC827	88.3%	9.2%	2.5%
Squamous	H520	6.5%	88.8%	4.7%
	H2170	8.6%	90.5%	0.9%
Small	H526	2.8%	2.1%	95.2%
	H69	5.6%	7.2%	87.2%

Confusion matrices (left) before and (right) after transfer learning. Color gradient of the grids are set proportional to the value within it. Darker colors mean higher values. The highest value in the row is marked with white texts.

Figure S8 : Classification accuracy with different transfer learning sample size



The prediction accuracy curves were acquired by training the transfer-learning-assisted neural network with different size of the transfer learning dataset, ranging from an empty dataset to 1,000 cells per cell line. With more transfer learning data, the classification accuracy increases and levels off at around 95%.

Figure S9 (same as Fig. 2e): Mean phenotypic heatmap and correlations with the feature labels

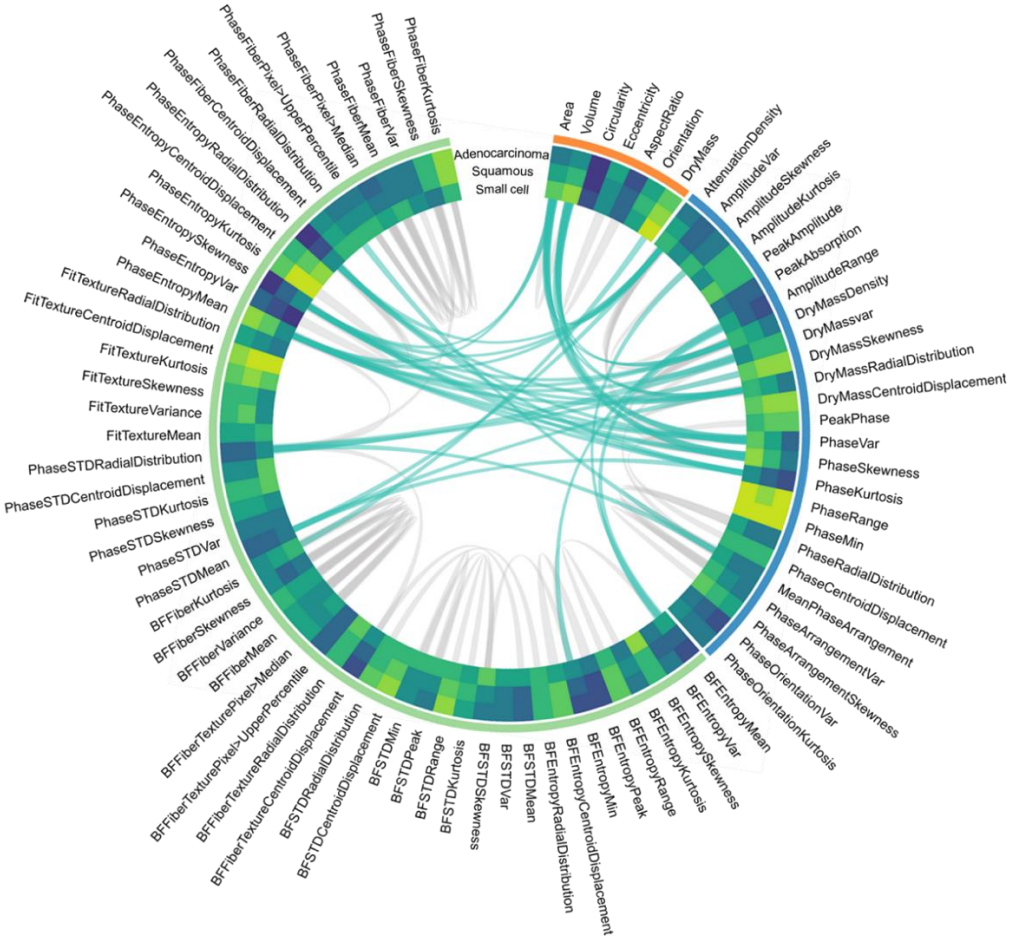
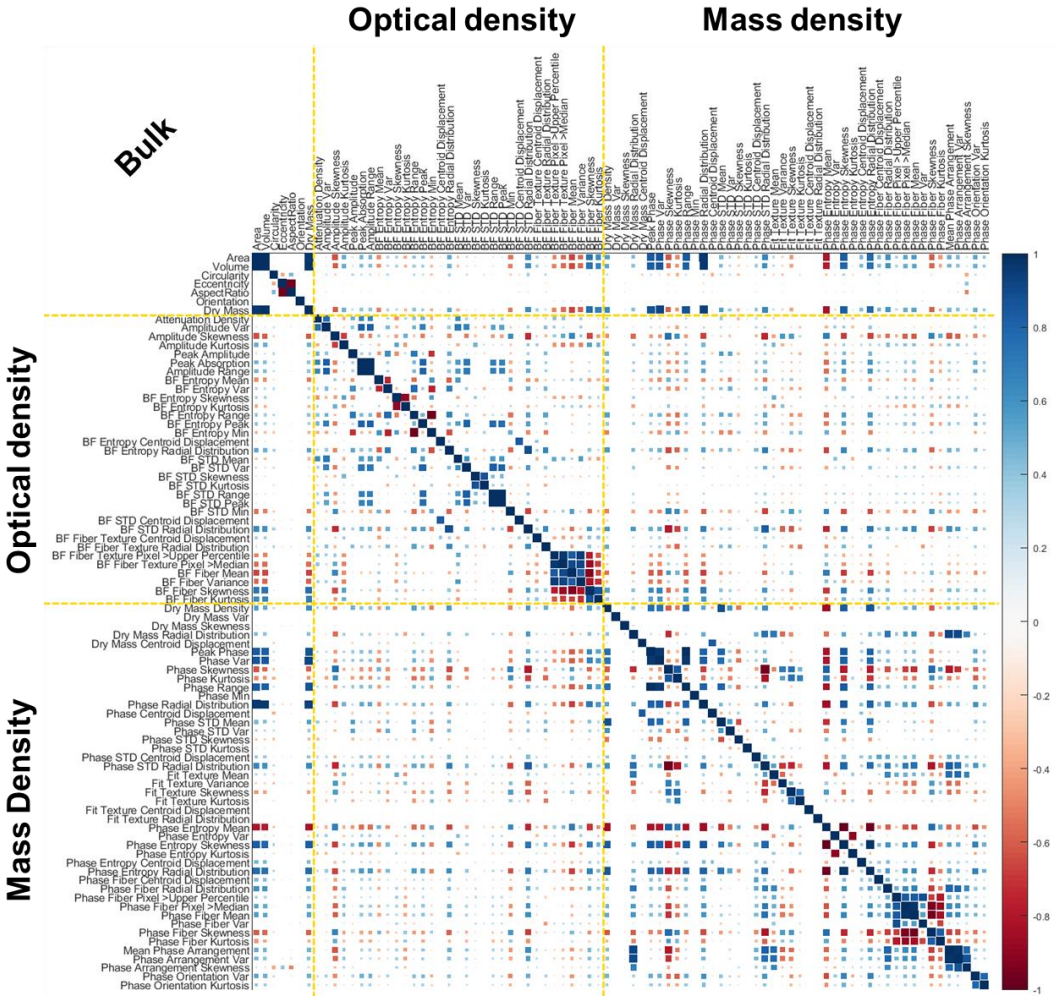
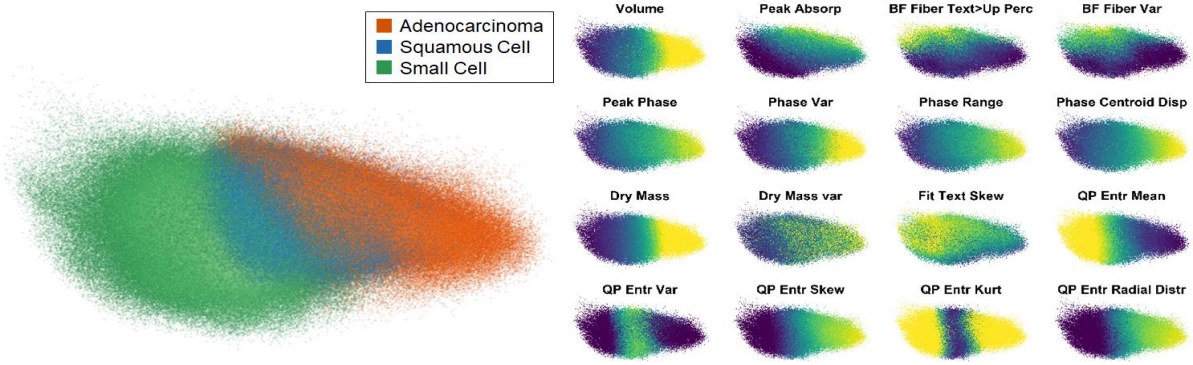


Figure S10: Correlation matrix of hierarchical spatial optophysical phenotypes



Blue colour denotes positive correlation and red colour denotes negative correlation (Pearson's correlation coefficient). The magnitude of correlation is represented by the size of the grid. The yellow dotted lines divide the phenotypes into their hierarchical categories (i.e. bulk, optical density and mass density).

Figure S11: Color-encoded spatial optophysical phenotype map



(Left) 3D PCA plot of the three lung-cancer subtypes based on the spatial optophysical phenotypes (same as **Fig. 2f**). (Right) the optophysical phenotypes that show distinguishing differences among the three subtypes.

Figure S12 (same as Fig. 3c): Ranking of phenotypes (with the feature labels) in classifying H2170 and PBMC

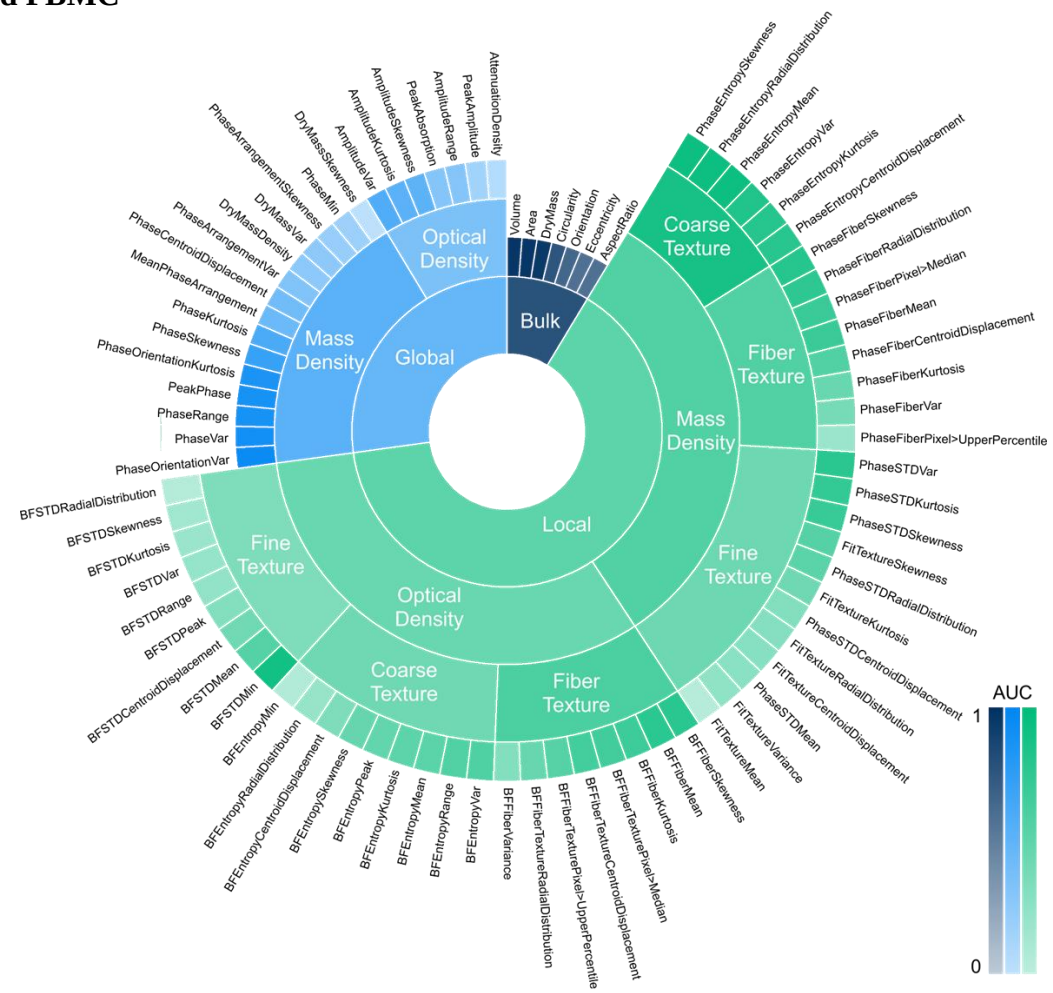
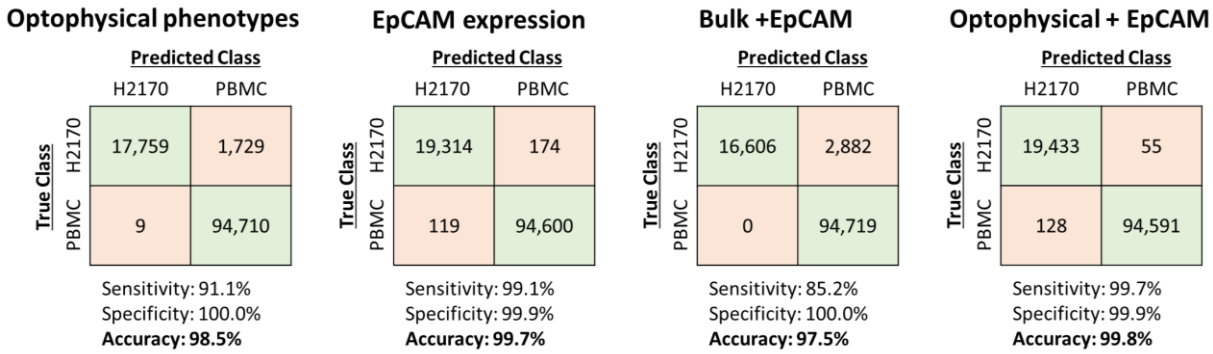
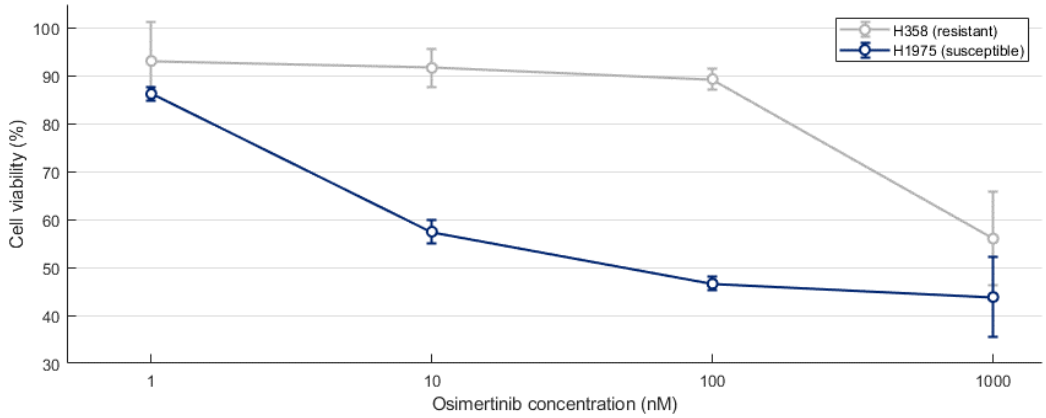


Figure S13: The roles of label-free optophysical phenotypes and EpCAM expression for classification of H2170 cells and PBMC.



Confusion matrices comparing the performances of neural network model using (Left) optophysical phenotypes only, (middle left) EpCAM expression only, (middle right) easily measurable bulk features with EpCAM expression and (right) optophysical phenotypes with EpCAM expression for detecting lung cancer cells in PBMCs. Using label-free optophysical phenotypes allows high accuracy (98.5%) and sensitivity (91.1%). Despite the lower accuracy than that using EpCAM expression, using both optophysical phenotypes and EpCAM expression can further augment both sensitivity (99.7%) and accuracy (99.8%), demonstrating the essential role of optophysical phenotypes in cancer cell detection in PBMCs. Compared to the use of EpCAM expression alone, using both optophysical phenotypes and EpCAM expression indeed increases the number of true positives detected by 119, which improved both the sensitivity and accuracy despite a slight compromise of specificity (<0.1%). The superior sensitivity (99.7%) of using both optophysical phenotypes and EpCAM expression in comparison with that of using easily measurable bulk features with EpCAM expression also justifies the importance of the high-dimensional optophysical phenotypic profile. However, when comparing the EpCAM-expression-only model and the model of bulk features with EpCAM expression, the accuracy is decreased (from 99.7% to 97.5%) after the addition of bulk features. Further investigation is required to explain this change.

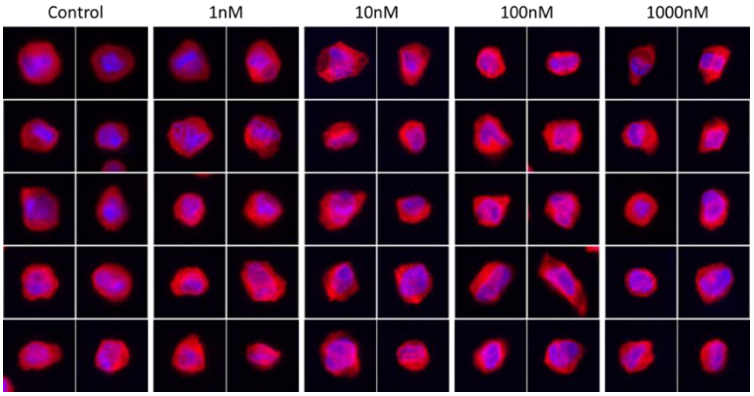
Figure S14: Cell viability of H1975 and H358 cells treated with Osimertinib



The dose-response curves (across 5 orders of magnitude of osimertinib concentration) were acquired by the MTT assays after 72 hours of drug treatment. The MTT assay was done on a 96-well plate. The error bar is constructed from the standard deviation of 2 replicates.

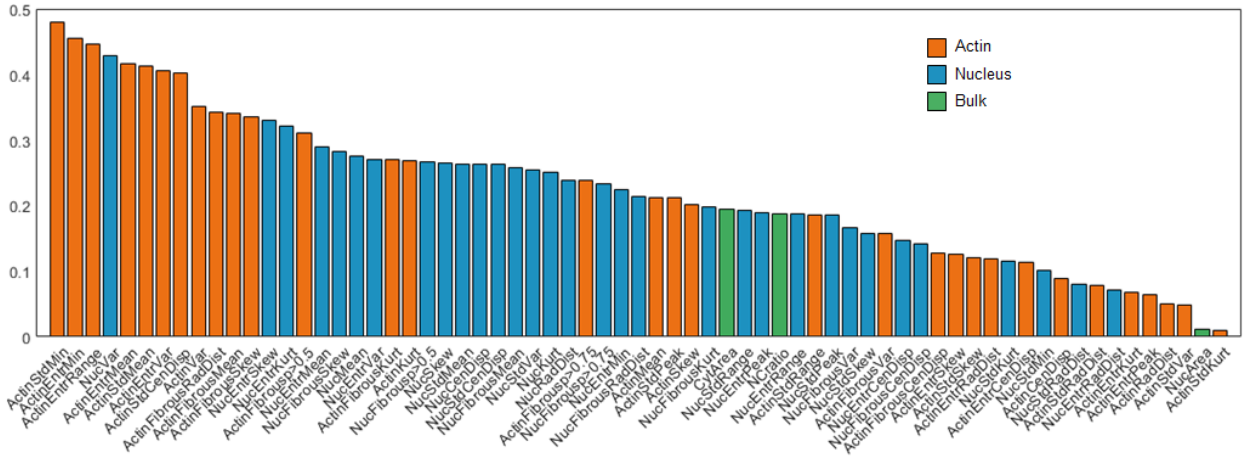
Figure S15 (same as Fig. 4e): Ranking of phenotypes (with the feature labels) in showing the response to the Osimertinib treatment

Figure S16: Fluorescence images of H1975 in response to the osimertinib treatments



Fluorescence images of osimertinib-treated H1975 (fixed with 4% paraformaldehyde). The red channel represents actin (stained by Alexa Fluor 546 phalloidin) whereas the blue channel represents nucleus (stained by DAPI). Each image is 40µm x 40µm.

Figure S17: Overall ranking of the spatial features obtained from fluorescence images of cells treated with Osimertinib



The ranking is arranged in a descending order of Spearman's correlation coefficients of individual features from left to right.

Table S1: Feature equation table (List of variables used can be found in Table S2)

Feature name	*Abbreviation*	*Equation*
Area	A	${L_{pix}}^2 \cdot N_{pix}$
Volume	V	$\frac{4}{3}\pi \cdot (\frac{L_{minor}}{2})^2 \cdot (\frac{L_{major}}{2})$
Circularity		$4\pi A/P$
Eccentricity		$\frac{L_{ellip}}{L_{major}}$
Aspect Ratio		$\frac{L_{minor}}{L_{major}}$
Orientation		θ_{major}
Dry Mass	MD_{total}	$\frac{\lambda}{2\pi\,\alpha}\iint_A MD(x,y)\,dxdy$
Attenuation Density		$\frac{\iint_A (1-OD(x,y))\,dxdy}{N_{pix}}$
Amplitude Variance	${\sigma_{OD}}^2$	$\iint_A (OD(x,y)-\overline{OD})^2\,dxdy/(N_{pix}-1)$
Amplitude Skewness		$\frac{\iint_A (OD(x,y)-\overline{OD})^3\,dxdy/N_{pix}}{{\sigma_{OD}}^3}$
Amplitude Kurtosis		$\frac{\iint_A (OD(x,y)-\overline{OD})^4\,dxdy/N_{pix}}{{\sigma_{OD}}^4}$
Peak Amplitude		$\max\{OD(x,y)\}$
Peak Absorption		$\min\{OD(x,y)\}$
Amplitude Range		$\max\{OD(x,y)\}-\min\{OD(x,y)\}$
Dry Mass Density	$\overline{DMD}$	$\frac{\iint_A DMD(x,y)\,dxdy}{N_{pix}}$
Dry Mass Variance	${\sigma_{DMD}}^2$	$\iint_A (DMD(x,y)-\overline{DMD})^2\,dxdy/(N_{pix}-1)$
Dry Mass Skewness		$\frac{\iint_A (DMD(x,y)-\overline{DMD})^3\,dxdy/N_{pix}}{{\sigma_{DMD}}^3}$
Dry Mass Radial Distribution		$\frac{\iint_A DMD(r,\theta)\,r\,drd\theta}{\iint_A DMD(r,\theta)\;drd\theta}$
Dry Mass Centroid Displacement		$\sqrt{(x_{DMD,cen}-x_{cen})^2+(y_{DMD,cen}-y_{cen})^2}\cdot L_{pix}$
Peak Phase		$\max\{MD(x,y)\}$
Phase Minimum		$\min\{MD(x,y)\}$
Phase Radial Distribution		$\frac{\iint_A r\cdot MD(r,\theta)\,drd\theta}{\iint_A MD(r,\theta)\;drd\theta}$
Phase Centroid Displacement		$\sqrt{(x_{DMD,cen}-x_{cen})^2+(y_{DMD,cen}-y_{cen})^2}\cdot L_{pix}$
Mean Phase Arrangement		$\frac{\iint_A MD(r,\theta)\,r\,drd\theta}{\iint_A MD(r,\theta)\;drd\theta}$

Phase Arrangement Variance	$\sigma_{MDarr}{}^2$	$\frac{\iint_A (MD(r,\theta)\, r)^2\, drd\theta}{\iint_A MD(r,\theta)\; drd\theta}$
Phase Arrangement Skewness		$\frac{\iint_A (MD(r,\theta)\cdot r)^3\, drd\theta}{\sigma_{MDarr}{}^2 \cdot \iint_A MD(r,\theta)\; drd\theta}$
Phase Orientation Variance	$\sigma_{MDang}{}^2$	$\frac{\int_0^\infty (\widetilde{MD}(\omega)\cdot\ \omega)^2\, d\omega}{\int_0^\infty \widetilde{MD}(\omega)\, d\omega}$
Phase Orientation Kurtosis		$\frac{\int_0^\infty (\widetilde{MD}(\omega)\cdot\ \omega)^4\, d\omega}{\sigma_{MDang}{}^2 \cdot \int_0^\infty \widetilde{MD}(\omega)\, d\omega}$
BF Entropy Mean	$\overline{OD_{ent}}$	$\frac{\iint_A OD_{ent}(x,y)\, dxdy}{N_{pix}}$
BF Entropy Variance	$\sigma_{ODent}{}^2$	$\frac{\iint_A (OD_{ent}(x,y) - \overline{OD_{ent}})^2 dxdy}{N_{pix}-1}$
BF Entropy Skewness		$\frac{\iint_A (OD_{ent}(x,y) - \overline{OD_{ent}})^3 dxdy/N_{pix}}{\sigma_{ODent}{}^3}$
BF Entropy Kurtosis		$\frac{\iint_A (OD_{ent}(x,y) - \overline{OD_{ent}})^4 dxdy/N_{pix}}{\sigma_{ODent}{}^4}$
BF Entropy Range		$\max\{OD_{ent}(x,y)\} - \min\{OD_{ent}(x,y)\}$
BF Entropy Peak		$\max\{OD_{ent}(x,y)\}$
BF Entropy Min		$\min\{OD_{ent}(x,y)\}$
BF Entropy Centroid Displacement		$\sqrt{(x_{ODent,cen} - x_{cen})^2 + (y_{ODent,cen} - y_{cen})^2} \cdot L_{pix}$
BF Entropy Radial Distribution		$\frac{\iint_A r \cdot OD_{ent}(r,\theta)\, drd\theta}{\iint_A OD_{ent}(r,\theta)\; drd\theta}$
BF STD Mean	$\overline{OD_{STD}}$	$\frac{\iint_A OD_{STD}(x,y)\, dxdy}{N_{pix}}$
BF STD Variance	$\sigma_{ODstd}{}^2$	$\frac{\iint_A (OD_{STD}(x,y) - \overline{OD_{STD}})^2 dxdy}{N_{pix}-1}$
BF STD Skewness		$\frac{\iint_A (OD_{STD}(x,y) - \overline{OD_{STD}})^3 dxdy/N_{pix}}{\sigma_{ODstd}{}^3}$
BF STD Kurtosis		$\frac{\iint_A (OD_{STD}(x,y) - \overline{OD_{STD}})^4 dxdy/N_{pix}}{\sigma_{ODstd}{}^4}$
BF STD Range		$\max\{OD_{STD}(x,y)\} - \min\{OD_{STD}(x,y)\}$
BF STD Peak		$\max\{OD_{STD}(x,y)\}$
BF STD Min		$\min\{OD_{STD}(x,y)\}$
BF STD Centroid Displacement		$\sqrt{(x_{ODSTD,cen} - x_{cen})^2 + (y_{ODSTD,cen} - y_{cen})^2} \cdot L_{pix}$
BF STD Radial Distribution		$\frac{\iint_A r \cdot OD_{STD}(r,\theta)\, drd\theta}{\iint_A OD_{STD}(r,\theta)\; drd\theta}$
BF Fiber Texture Centroid Displacement		$\sqrt{(x_{ODfiber,cen} - x_{cen})^2 + (y_{ODfiber,cen} - y_{cen})^2} \cdot L_{pix}$

BF Fiber Texture Radial Distribution		$\frac{\iint_A r \cdot OD_{fiber}(r,\theta)\, drd\theta}{\iint_A OD_{fiber}(r,\theta)\ drd\theta}$
BF Fiber Texture Pixel>Upper Percentile		$\frac{Number\ of\ pixels\ in\ OD_{fiber}(x,y) > 75th\ percentile}{N_{pix}}$
BF Fiber Texture Pixel>Median		$\frac{Number\ of\ pixels\ in\ OD_{fiber}(x,y) > median}{N_{pix}}$
BF Fiber Mean	$\overline{OD_{fiber}}$	$\frac{\iint_A OD_{fiber}(x,y)\, dxdy}{N_{pix}}$
BF Fiber Variance	${\sigma_{ODfiber}}^2$	$\frac{\iint_A (OD_{fiber}(x,y) - \overline{OD_{fiber}})^2\, dxdy}{N_{pix} - 1}$
BF Fiber Skewness		$\frac{\iint_A (OD_{fiber}(x,y) - \overline{OD_{fiber}})^3 dxdy/N_{pix}}{{\sigma_{ODfiber}}^3}$
BF Fiber Kurtosis		$\frac{\iint_A (OD_{fiber}(x,y) - \overline{OD_{fiber}})^4 dxdy/N_{pix}}{{\sigma_{ODfiber}}^4}$
Phase STD Mean	$\overline{MD_{STD}}$	$\frac{\iint_A MD_{STD}(x,y)\, dxdy}{N_{pix}}$
Phase STD Var	${\sigma_{MDstd}}^2$	$\frac{\iint_A (MD_{STD}(x,y) - \overline{MD_{STD}})^2 dxdy}{N_{pix} - 1}$
Phase STD Skewness		$\frac{\iint_A (MD_{STD}(x,y) - \overline{MD_{STD}})^3 dxdy/N_{pix}}{{\sigma_{MDstd}}^3}$
Phase STD Kurtosis		$\frac{\iint_A (MD_{STD}(x,y) - \overline{MD_{STD}})^4 dxdy/N_{pix}}{{\sigma_{MDstd}}^4}$
Phase STD Centroid Displacement		$\sqrt{(x_{MDSTD,cen} - x_{cen})^2 + (y_{MDSTD,cen} - y_{cen})^2} \cdot L_{pix}$
Fit Texture Mean	$\overline{MD_{fit}}$	$\frac{\iint_A MD_{fit}(x,y)\, dxdy}{N_{pix}}$
Fit Texture Variance	${\sigma_{MDfit}}^2$	$\frac{\iint_A (MD_{fit}(x,y) - \overline{MD_{fit}})^2 dxdy}{N_{pix} - 1}$
Fit Texture Skewness		$\frac{\iint_A (MD_{fit}(x,y) - \overline{MD_{fit}})^3 dxdy/N_{pix}}{{\sigma_{MDfit}}^3}$
Fit Texture Kurtosis		$\frac{\iint_A (MD_{fit}(x,y) - \overline{MD_{fit}})^4 dxdy/N_{pix}}{{\sigma_{MDfit}}^4}$
Fit Texture Centroid Displacement		$\sqrt{(x_{MDfit,cen} - x_{cen})^2 + (y_{MDfit,cen} - y_{cen})^2} \cdot L_{pix}$
Fit Texture Radial Distribution		$\frac{\iint_A r \cdot MD_{fit}(r,\theta)\, drd\theta}{\iint_A MD_{fit}(r,\theta)\ drd\theta}$
Phase Entropy Mean	$\overline{MD_{ent}}$	$\frac{\iint_A MD_{ent}(x,y)\, dxdy}{N_{pix}}$
Phase Entropy Var	${\sigma_{MDent}}^2$	$\frac{\iint_A (MD_{ent}(x,y) - \overline{MD_{ent}})^2 dxdy}{N_{pix} - 1}$

Phase Entropy Skewness		$\frac{\iint_A (MD_{ent}(x,y) - \overline{MD_{ent}})^3 dxdy / N_{pix}}{\sigma_{MDent}{}^3}$
Phase Entropy Kurtosis		$\frac{\iint_A (MD_{ent}(x,y) - \overline{MD_{ent}})^4 dxdy / N_{pix}}{\sigma_{MDent}{}^4}$
Phase Entropy Centroid Displacement		$\sqrt{(x_{MDent,cen} - x_{cen})^2 + (y_{MDent,cen} - y_{cen})^2} \cdot L_{pix}$
Phase Entropy Radial Distribution		$\frac{\iint_A r \cdot MD_{ent}(r,\theta)\, drd\theta}{\iint_A MD_{ent}(r,\theta)\ drd\theta}$
Phase Fiber Centroid Displacement		$\sqrt{(x_{MDfiber,cen} - x_{cen})^2 + (y_{MDfiber,cen} - y_{cen})^2} \cdot L_{pix}$
Phase Fiber Radial Distribution		$\frac{\iint_A r \cdot MD_{fiber}(r,\theta)\, drd\theta}{\iint_A MD_{fiber}(r,\theta)\ drd\theta}$
Phase Fiber Pixel>Upper Percentile		$\frac{Number\ of\ pixels\ in\ MD_{fiber}(x,y) > 75th\ percentile}{N_{pix}}$
Phase Fiber Pixel>Median		$\frac{Number\ of\ pixels\ in\ MD_{fiber}(x,y) > median}{N_{pix}}$
Phase Fiber Mean	$\overline{MD_{fiber}}$	$\frac{\iint_A MD_{fiber}(x,y)\, dxdy}{N_{pix}}$
Phase Fiber Var	$\sigma_{MDfiber}{}^2$	$\frac{\iint_A (MD_{fiber}(x,y) - \overline{MD_{fiber}})^2\, dxdy}{N_{pix} - 1}$
Phase Fiber Skewness		$\frac{\iint_A (MD_{fiber}(x,y) - \overline{MD_{fiber}})^3 dxdy / N_{pix}}{\sigma_{MDfiber}{}^3}$
Phase Fiber Kurtosis		$\frac{\iint_A (MD_{fiber}(x,y) - \overline{MD_{fiber}})^4 dxdy / N_{pix}}{\sigma_{MDfiber}{}^4}$

Table S2: List of variables and abbreviations

Variable	*Description*	*Equation/Remarks*
C	Contour of binary mask	
CM	Cell mask function	$CM(x,y) = \begin{cases} 1 & if\ inside\ cell \\ 0 & otherwise \end{cases}$
DMD	Dry mas density map	$DMD(x,y) = \frac{\lambda \cdot MD(x,y)}{2\pi\alpha \cdot h(x,y)}$
h	Cell height map	$h(x,y) = \sqrt{(\frac{L_{minor}+L_{major}}{2})^2 - ((x-x_{cen})^2 + (y-y_{cen})^2)}$
L_{ellip}	Distance between foci of ellipse	
L_{major}	Major axis length	
L_{minor}	Minor axis length	
L_{pix}	Physical length of one pixel	
MD	Mass density map	$MD(x,y)$
$MD(\theta)$	Mass density projected to polar angle	
$\widetilde{MD}(\omega)$	Mass density in angular frequency domain	$\widetilde{MD}(\omega) = \mathcal{F}(MD(\theta))$
$\overline{MD_{STD,ker}}(x,y)$	Mean value of QPI within STD filter kernel	$\frac{\int_{x-w_{STD}/2}^{x+w_{STD}/2}\int_{y-w_{STD}/2}^{y+w_{STD}/2} MD(u,v)\ dvdu}{{w_{STD}}^2}$
$MD_{STD}(x,y)$	QPI STD map	$\int_{x-w_{STD}/2}^{x+w_{STD}/2}\int_{y-w_{STD}/2}^{y+w_{STD}/2}\sqrt{\frac{(MD(u,v)-\overline{MD_{STD,ker}}(x,y))^2}{{w_{STD}}^2}}\ dvdu$
$MD_{cubic}(x,y)$	Cubic polynomial surface fit of mass density map	
$MD_{fit}(x,y)$	Fit texture map of mass density map	$MD(x,y) - MD_{cubic}(x,y)$
$MD_{ent}(x,y)$	Entropy filtered mass density map	$\sum_{k=0}^{255} p_{MD,k} \cdot \log_2 p_{MD,k}$
$MD_{fiber}(x,y)$	Fiber texture enhanced mass density map	$FF(MD(x,y))$, ref. [9]
N_{pix}	Pixel number in cell mask	$\iint CM(x,y)\ dA$
OD	Optical density map	$OD(x,y)$
$\overline{OD}$	Amplitude mean	$\iint_A OD(x,y)\ dxdy/N_{pix}$
$OD_{ent}(x,y)$	Entropy filtered optical density map	$\sum_{k=0}^{255} p_{OD,k} \cdot \log_2 p_{OD,k}$
$\overline{OD_{STD,ker}}(x,y)$	Mean value of BF within STD filter kernel	$\frac{\int_{x-w_{STD}/2}^{x+w_{STD}/2}\int_{y-w_{STD}/2}^{y+w_{STD}/2} OD(u,v)\ dvdu}{{w_{STD}}^2}$

$OD_{STD}(x,y)$	BF STD map	$\int_{x-w_{STD}/2}^{x+w_{STD}/2}\int_{y-w_{STD}/2}^{y+w_{STD}/2}\sqrt{\frac{(OD(u,v)-\overline{OD_{STD,ker}}(x,y))^2}{{w_{STD}}^2}}\,dvdu$
$OD_{fiber}(x,y)$	Fiber texture enhanced optical density map	$FF(OD(x,y))$, ref. [9]
P	Perimeter	$\oint_C\sqrt{\left(\left(\frac{dx}{d\theta}\right)^2+\left(\frac{dy}{d\theta}\right)^2\right)}\,d\theta$
$p_{MD,k}(x,y)$	Normalized histogram counts within kernel of mass density map	$p_{MD,k}(x,y)=\frac{number\ of\ pixels\ in\ kernel\ (w_{ent})\ with\ MD=k}{Total\ number\ of\ pixels\ in\ kernel}$, where k = 0 to 255
$p_{OD,k}(x,y)$	Normalized histogram counts within kernel of optical density map	$p_{OD,k}(x,y)=\frac{number\ of\ pixels\ in\ kernel\ (w_{ent})\ with\ OD=k}{Total\ number\ of\ pixels\ in\ kernel}$, where k = 0 to 255
r,θ	Polar coordinates centered at cell centroid	
w_{ent}	Kernel size of entropy filter	
w_{STD}	Kernel size of STD filter	
x,y	Cartesian coordinates	
x_{cen} y_{cen}	Coordinates of cell centroid	$x_{cen}=\frac{\iint_A x\cdot CM(x,y)\,dxdy}{N_{pix}}, y_{cen}=\frac{\iint_A y\cdot CM(x,y)\,dxdy}{N_{pix}}$
$x_{MD,cen}$ $y_{MD,cen}$	Coordinates of mass density weighted cell centroid	$x_{MD,cen}=\frac{\iint_A x\cdot MD(x,y)\,dxdy}{N_{pix}}$ $y_{MD,cen}=\frac{\iint_A y\cdot MD(x,y)\,dxdy}{N_{pix}}$
$x_{DMD,cen}$ $y_{DMD,cen}$	Coordinates of dry mass density weighted cell centroid	$x_{DMDcen}=\frac{\iint_A x\cdot DMD(x,y)\,dxdy}{N_{pix}}$ $y_{DMDcen}=\frac{\iint_A y\cdot DMD(x,y)\,dxdy}{N_{pix}}$
$x_{MDent,cen}$ $y_{MDent,cen}$	Coordinates of entropy filtered MD weighted cell centroid	$x_{MDent,cen}=\frac{\iint_A x\cdot MD_{ent}(x,y)\,dxdy}{N_{pix}}$ $y_{MDent,cen}=\frac{\iint_A y\cdot MD_{ent}(x,y)\,dxdy}{N_{pix}}$
$x_{MDfiber,cen}$ $y_{MDfiber,cen}$	Coordinates of fiber enhanced MD weighted cell centroid	$x_{MDfiber,cen}=\frac{\iint_A x\cdot MD_{fiber}(x,y)\,dxdy}{N_{pix}}$ $y_{MDfiber,cen}=\frac{\iint_A y\cdot MD_{fiber}(x,y)\,dxdy}{N_{pix}}$
$x_{MDfit,cen}$ $y_{MDfit,cen}$	Coordinates of MD fit texture weighted cell centroid	$x_{MDfit,cen}=\frac{\iint_A x\cdot MD_{fit}(x,y)\,dxdy}{N_{pix}}$ $y_{MDfit,cen}=\frac{\iint_A y\cdot MD_{fit}(x,y)\,dxdy}{N_{pix}}$
$x_{MDSTD,cen}$ $y_{MDSTD,cen}$	Coordinates of STD filtered MD weighted cell centroid	$x_{MDSTD,cen}=\frac{\iint_A x\cdot MD_{STD}(x,y)\,dxdy}{N_{pix}}$ $y_{MDSTD,cen}=\frac{\iint_A y\cdot MD_{STD}(x,y)\,dxdy}{N_{pix}}$

$x_{ODent,cen}$ $y_{ODent,cen}$	Coordinates of entropy filtered OD weighted cell centroid	$x_{ODent,cen} = \frac{\iint_A x \cdot OD_{ent}(x,y)\, dxdy}{N_{pix}}$ $y_{ODent,cen} = \frac{\iint_A y \cdot OD_{ent}(x,y)\, dxdy}{N_{pix}}$
$x_{ODfiber,cen}$ $y_{ODfiber,cen}$	Coordinates of fiber enhanced OD weighted cell centroid	$x_{ODfiber,cen} = \frac{\iint_A x \cdot OD_{fiber}(x,y)\, dxdy}{N_{pix}}$ $y_{ODfiber,cen} = \frac{\iint_A y \cdot OD_{fiber}(x,y)\, dxdy}{N_{pix}}$
$x_{ODSTD,cen}$ $y_{ODSTD,cen}$	Coordinates of STD filtered OD weighted cell centroid	$x_{ODSTD,cen} = \frac{\iint_A x \cdot OD_{STD}(x,y)\, dxdy}{N_{pix}}$ $y_{ODSTD,cen} = \frac{\iint_A y \cdot OD_{STD}(x,y)\, dxdy}{N_{pix}}$
α	Specific refractive increment	$0.19\ ml/g$ (ref. [5])
θ_{major}	Angle between major axis and x-axis	
$\mathcal{F}$	Fourier transform	

References

1. Olivas, E. S.; Guerrero, J. D. M.; Sober, M. M.; Benedito, J. R. M.; Lopez, A. J. S., *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques - 2 Volumes*. Information Science Reference - Imprint of: IGI Publishing: 2009.
2. Ogoe, H. A.; Visweswaran, S.; Lu, X.; Gopalakrishnan, V., Knowledge transfer via classification rules using functional mapping for integrative modeling of gene expression data. *BMC Bioinformatics* **2015,** *16* (1), 226.
3. Kim, S.-J.; Wang, C.; Zhao, B.; Im, H.; Min, J.; Choi, H. J.; Tadros, J.; Choi, N. R.; Castro, C. M.; Weissleder, R., Deep transfer learning-based hologram classification for molecular diagnostics. *Sci. Rep.* **2018,** *8* (1), 1-12.
4. Niazi, M. K. K.; Tavolara, T. E.; Arole, V.; Hartman, D. J.; Pantanowitz, L.; Gurcan, M. N., Identifying tumor in pancreatic neuroendocrine neoplasms from Ki67 images using transfer learning. *Quantifying biomass changes of single CD8+ T cells during antigen specific cytotoxicity* **2018,** *13* (4).
5. Zhao, H.; Brown, P. H.; Schuck, P., On the distribution of protein refractive index increments. *Biophys. J.* **2011,** *100* (9), 2309-2317.
6. Goodman, J. W., *Introduction to Fourier optics*. Roberts and Company Publishers: 2005.
7. Hamilton, D.; Sheppard, C.; Wilson, T., Improved imaging of phase gradients in scanning optical microscopy. *Optical coherence tomography* **1984,** *135* (3), 275-286.
8. Lee, K. C.; Lau, A. K.; Tang, A. H.; Wang, M.; Mok, A. T.; Chung, B. M.; Yan, W.; Shum, H. C.; Cheah, K. S.; Chan, G. C., Multi-ATOM: Ultrahigh-throughput single-cell quantitative phase imaging with subcellular resolution. *J. Biophotonics* **2019,** *12* (7), e201800479.
9. Frangi, A. F.; Niessen, W. J.; Vincken, K. L.; Viergever, M. A. In *Multiscale vessel enhancement filtering*, International conference on medical image computing and computer-assisted intervention, Springer: 1998; pp 130-137.