Label-free classification of cells based on supervised machine learning of subcellular structures

Label-free classification of cells based on supervised machine learning of subcellular structures

Yusuke Ozaki[1☯], **Hidenao Yamada**[2☯]*, **Hirotoshi Kikuchi**[1], **Amane Hirotsu**[1],
Tomohiro Murakami[1], **Tomohiro Matsumoto**[1], **Toshiki Kawabata**[1], **Yoshihiro Hiramatsu**[1],
Kinji Kamiya[1], **Toyohiko Yamauchi**[2], **Kentaro Goto**[2], **Yukio Ueda**[2], **Shigetoshi Okazaki**[3],
Masatoshi Kitagawa[4,5], **Hiroya Takeuchi**[1], **Hiroyuki Konno**[6]

1 Second Department of Surgery, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka, Japan,
2 Central Research Laboratory, Hamamatsu Photonics K.K., Hamamatsu, Shizuoka, Japan, **3** Department
of Medical Spectroscopy, Institute for Medical Photonics Research, Preeminent Medical Photonics Education
and Research Center, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka, Japan,
4 Department of Molecular Biology, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka,
Japan, **5** Laboratory Animal Facilities and Services, Preeminent Medical Photonics Education and Research
Center, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka, Japan, **6** Hamamatsu University
School of Medicine, Hamamatsu, Shizuoka, Japan

☯ These authors contributed equally to this work.
* iwai@crl.hpk.co.jp

Abstract

It is demonstrated that cells can be classified by pattern recognition of the subcellular structure of non-stained live cells, and the pattern recognition was performed by machine learning. Human white blood cells and five types of cancer cell lines were imaged by quantitative phase microscopy, which provides morphological information without staining quantitatively in terms of optical thickness of cells. Subcellular features were then extracted from the obtained images as training data sets for the machine learning. The built classifier successfully classified WBCs from cell lines (area under ROC curve = 0.996). This label-free, non-cytotoxic cell classification based on the subcellular structure of QPM images has the potential to serve as an automated diagnosis of single cells.

Introduction

Morphological classification of cells and tissue on a subcellular scale under a microscope has a long history in pathology, including cytology and histology. The subcellular organelles cause subcellular features such as increased nuclear-to-cytoplasmic ratio, granular cytoplasm, and a large round nucleus with prominent nucleolus [1]. In identifying and classifying diseases, physicians recognize and analyze the pattern (features) in the microscopic image and interpret its meaning from their past training (experience). In cell biology, cytology, and pathology, the features of cells or tissue to be recognized and analyzed can be enhanced in two ways: one is staining with dyes or labeling the molecules to be observed with fluorescence light; the other is optical filtering by dark- or bright-field microscopy, including label-free optical imaging such

as phase-contrast and differential-interference-contrast imaging. The former describes subcellular features as a distribution map of specific proteins or molecules. The later describes the features as a refractive index map of various proteins or molecules. In this paper, we refer to the refractive index map inside a cell as a subcellular structure.

For the last decade, as one task in computer vision, pattern recognition has been the most-topical area in fields such as autonomous cars and security. In these fields, patterns on target objects are described in certain manners, such as Haar-like features, local binary patterns (LBPs), and histogram of oriented gradients (HOG) [2][3], to enhance features of imagery or suppress artifacts such as illumination because general video or still cameras offer images as they are (i.e., no staining, labelling, or optical filtering). The combination of pattern recognition and machine learning is opening up new fields in not only industry but also biomedical and medical imaging.

Recently, as artificial intelligence becomes more advanced, automated diagnosis [4] of tissue and cells is gaining popularity. In such diagnosis, the texture of labeled or stained tissue slices and cells are automatically recognized and classified as normal or abnormal by computer vision trained by machine learning. On the other hand, label-free automated detection [5–8] and classification [9–13] of single cells (not in tissue) have been developed over the last decade or so. For instance, cells have been classified via imaging-flow cytometry in a manner of bright-field microcopy except quantitative-phase microscopy (QPM), dark-field microscopy, and machine learning of subcellular morphology [14]. Also, label-free drug assessment of cells is performed in the same manner as imaging-flow cytometry [15]. These two applications of imaging-flow cytometry utilize optical filtering to enhance the subcellular features of single cells.

Among the types of bright-field microscopy, QPM complements [16] the existing conventional label-free interferometric imaging, namely, phase-contrast microscopy, from the viewpoint of giving quantitative information about optical-path length of specimens and suppressing artifacts such as the "halo" [17–20] and "shade off" that hinder image segmentation [20]. Since QPM provides a non-optically-filtered image about optical path-length (OPL), a QPM image consists of all spatial-frequency components. The drawback of QPM images is that the subcellular structure is "buried" in the exiting spatial low-frequency components; consequently, instead of subcellular structure, physical parameters (physical features such as cellular volume, surface area, sphericity, dry mass, and dry mass density) in a QPM image have been be utilized for cell classification [16,21–23]. Here we refer to the physical parameters as cellular outlines to be contrasted with subcellular structures. A set of statistics of subcellular structures (standard deviations, variances, skewness, kurtosis, and so on) of a QPM image have been also utilized for cell classification by means of machine learning [24,25].

In consideration of pathological diagnosis, namely, a doctor diagnoses dysplasia of cells through a microscope on the basis of not only such cellular outline but also the subcellular structure of the tissue slices or cells stained, the subcellular structure of a single cell should also be extracted (recognized) and on the basis of which, cells are classified when computer vision also diagnoses. Among the various algorithms for extracting patterns from an image, histograms of oriented gradients [26] (HOG) is a de-facto standard for detecting humans by computer vision. HOG is generally variant to image size and rotation. To the best of our knowledge, HOG-based feature extraction has not been applied to QPM images of label-free single-cells, although modified HOG has been used to extract features from an image of stained cells [27]. Alternative method for extracting patterns from an image and classifying them is deep-learning [28].

In this paper, we describe a shallow-learning [29] approach for classifying rule-based auto-segmented images of healthy white blood cells (WBCs) and cancer cell-lines (CLs) on the basis of the subcellular structure in QPM images. After the sizes of the cells in the QPM images were

longitudinally and laterally normalized, the subcellular structure was extracted by the HOG-based feature-extraction algorithm. A support vector machine [30–32] (SVM) classifier was trained on the subcellular features by HOG descriptor, and the classification performance was plotted as a detection-error trade-off curve [33] (DET). For comparison, another SVM classifier was also trained by using a set of statistics of subcellular structures (statistical subcellular-structures).

Materials and methods

Quantitative phase microscopy

Among the various label-free imaging techniques, actively stabilized phase-shifting reflection-type QPM [34–37]—one kind of bright-field microcopy—was used in this study (Fig 1(A)). It provides quantitative morphological information about live cells without need for cytotoxic methods such as photo-bleaching and photo-toxicity [17], which are commonly used in fluorescence-labeled microscopy. As shown schematically in Fig 1(B), the difference between the refractive index (RI) of a cell (n_1) and that of its surrounding medium (n_0) and the physical path length (PL) cause phase delay of the incident light (wavelength; λ) to a sample according to Eq (1).

$$\Delta\phi(x, y) = \frac{2\pi}{\lambda}(n_1 - n_0) \cdot 2PL(x, y) = \frac{4\pi}{\lambda}OPL(x, y) \tag{1}$$
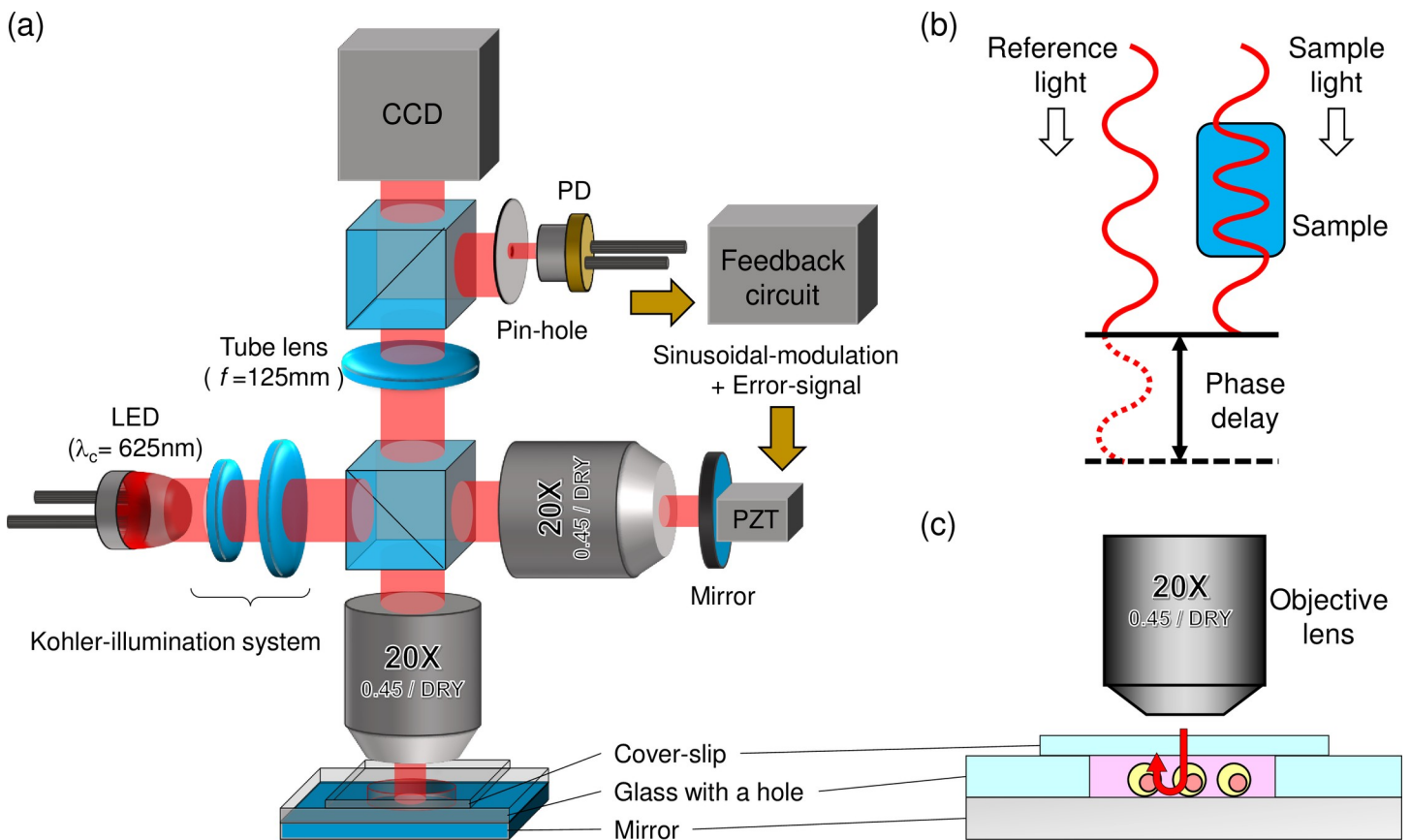

Fig 1. Schematic of imaging system and principle of phase measurement. (a) Schematic of reflection-type quantitative phase microscope. (b) Principle of phase measurement of the sample in QPM. (c) Configuration of custom-made slide chamber for cells in reflection-type QPM.
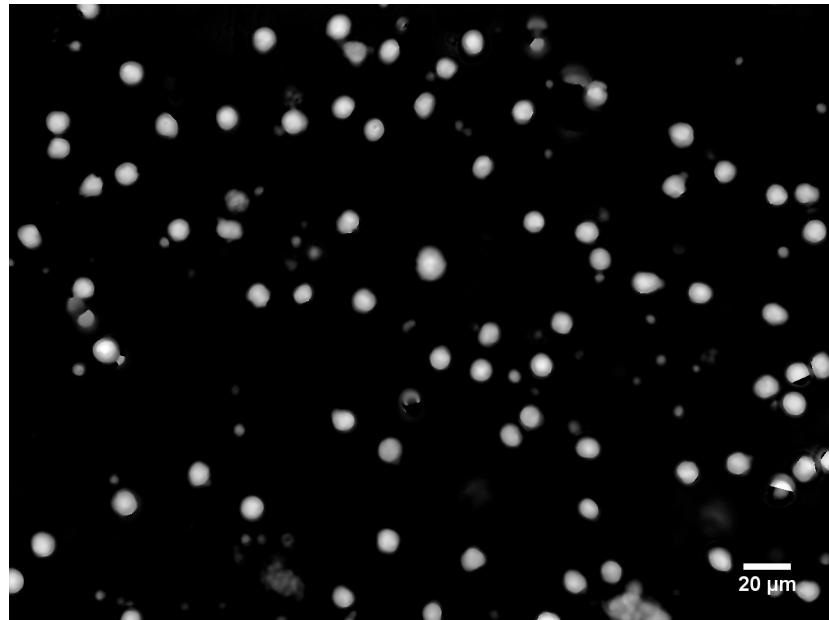
Fig 2. Full-field of view of a QPM image of WBCs after the removal of curvature in phase. White spots are QPM images of single WBCs, and the black region is background. Image size is 350×260 μm (1292×964 pixels). In the sample preparation, only WBCs were extracted from whole blood and then suspended in PBS.

https://doi.org/10.1371/journal.pone.0211347.g002

The phase delay (Δϕ) of the incident light is referred as the "optical path length" (OPL) of an object. RI acts as an intrinsic contrast agent [38] that enhances contrast of transparent samples. In the reflection-type configuration of a QPM shown in Fig 1(C), light incident on the sample passes through the sample twice; therefore, according to Eq (1), the physical path length (PL) is doubled.

In our QPM shown in Fig 1(A), the imaging light emitted from an LED (center wavelength: 625 nm) is split into two lights: one is incident on the sample, and the other is modulated (in phase) by a mirror mounted on the PZT transducer. These lights generate an interference pattern called "interferograms." The phase delay is retrieved from the interferograms by applying a temporal-phase-shifting algorithm. Magnification of our QPM is 13.89×, and the numerical aperture of the objective lens (Olympus 20×) is 0.45. The theoretical depth of focus is 1.5 μm. Tube lens with focal length of 125 mm is used for the objective lens (designed for a tube length of 180 mm). Interferograms are acquired by a two-dimension-arrayed sensor (CCD) with a pixel size of 3.75 μm and 12-bit depth (Guppy Pro F-125B, Allied Vision Technologies GmbH). The histogram of the background noise in a phase image was shown in S1 Fig. The standard deviation of the background noise was about 5 nm (50 milli-radians). The cells in the phosphate buffered-saline medium are configured in a custom-made slide chamber, as shown in Fig 1(C), where the slide underneath the cells is reflection coated to make the light passing through the cells reflect back to the objective lens. The reflection-coated glass slide is covered by a coverslip with a gap of around 1 mm so that the cells are not compressed and maintain their shape.

Rule-based auto-segmentation of cells

Our QPM provides a field of view of 350×260 μm (Fig 2). The background of the acquired QPM image typically has curvature in terms of phase because of the mismatch of the curvatures of the wavefronts of the sample- and reference lights. After the curvature is removed (the
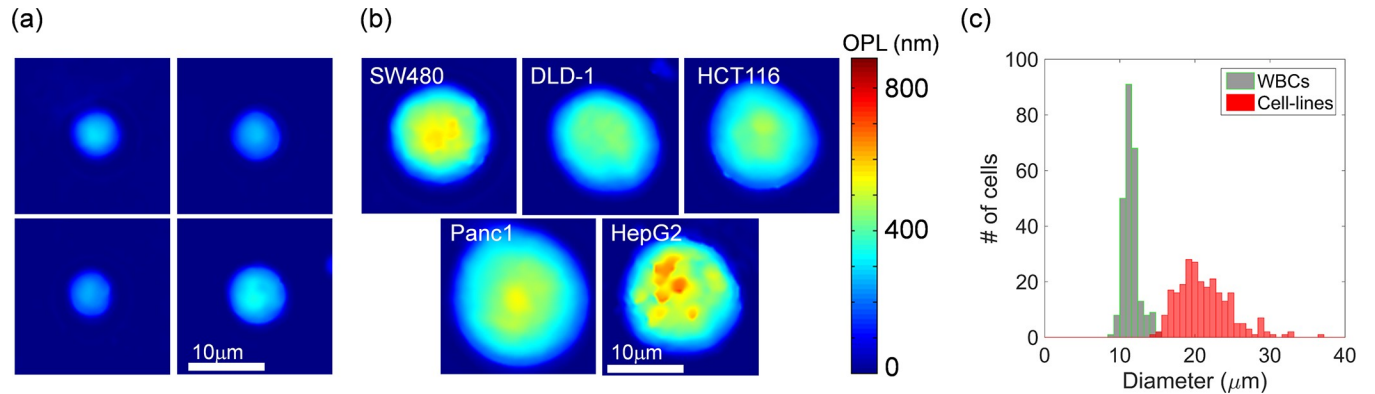
Fig 3. Representative segmented QPM images of single cells and their size distribution. (a) QPM images of WBC of a healthy donor. (b) QPM images of five types of cancer cell lines, namely, SW480, DLD-1, HCT116, Panc-1, and HepG2. Images were captured immediately after trypsinization to imitate CTCs floating in circulation. Pseudo-color represents OPL in nanometers. (c) Size distributions of diameters of 250 WBCs (green bars) and 250 cell lines (red bars).

procedure was shown in S2 Fig), the backgrounds of the image are uniform, so it is easy to segment each cell on the basis of a triangle algorithm [39], which is one of the auto thresholding algorithms implemented in the open-source image-processing software ImageJ. Representative segmented QPM images of WBCs of a healthy donor and five types of cancer cell lines (SW480, DLD-1, HCT116, Panc-1, and HepG2) are shown in pseudo color in Fig 3(A) and 3 (B), respectively. Since the difference between the diameters of the WBCs and the CLs is obvious in Fig 3(C), and it is easy to classify them, the sizes of cells were normalized to emphasize the classification based on subcellular structure by the method described in following section.

OPL normalizations for size-invariant feature extraction

To eliminate the influence of size on the classification, longitudinal (OPL) and lateral (size) directions of QPM images should be considered, as shown in Fig 4. Statistical features such as mean and standard deviation are already normalized so they are not influenced by the size of
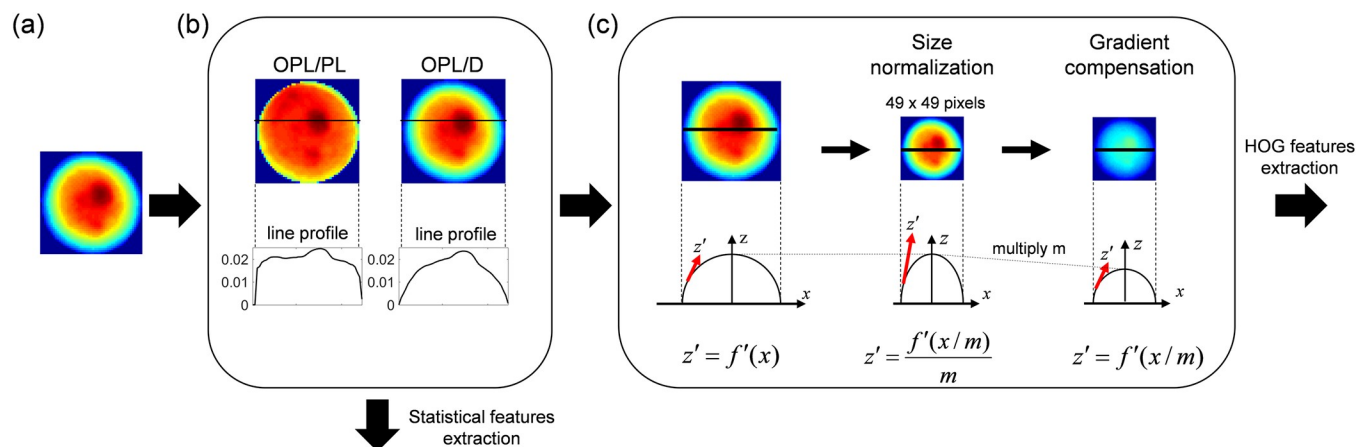
Fig 4. Two-branch flowchart of preprocessing for size-invariant feature extraction for statistical- and HOG-based subcellular classification. (a) Segmented image of a cell. (b) OPL normalizations by path length (PL) or diameter (D). Line profiles along the black lines in the pseudo color images of OPL/ PL or OPL/D are shown for comparison of two normalizations. The details of subcellular structure are unclear in the profile of OPL/D because of spatial low frequency components (the hemisphere), whereas the details of intracellular structure are recognized in the flat region of the profile of OPL/PL. (c) Size normalization and gradient compensation. Function f represents the profile of OPL/D or OPL/PL. Red arrows represent the spatial gradients of f (i.e., f'). The gradients were divided by resize factor m for compensations.

cells. In the longitudinal direction, as the size of cells increases, OPLs in the QPM images get thicker, because according to Eq (1), the OPLs are proportional to the thickness and RI. One longitudinal normalization is done by dividing OPLs by diameter (D) as representative thickness. The diameter can be predicted from a two-dimensional image under the assumption that that a floating cell is a sphere. Another longitudinal normalization is also proposed hereafter. Under the same assumption that a floating cell is almost a sphere with radius D/2, its path length along which light penetrates can be expressed as follows:

$$PL(x, y) = 2\sqrt{(D/2)^2 - x^2 - y^2}. \tag{2}$$

OPLs are divided by path length (PL) as real (not representative) thickness. A QPM image divided by its diameter or path length and line profiles along the solid lines in the images are shown in Fig 4(B). The ratio of OPL to PL or the ratio of OPL to D is a representative parameter as a dimensionless number.

Statistical subcellular-structures of single cells

A single cell contains organelles, which have different refractive indices [40,41]. As a result, a QPM can image subcellular structure in OPLs [42,43] in the lateral direction due to the distribution of refractive indices. Besides, it is reported that the structure might be related to malignancy [44–47]. As shown in Fig 5(A) and 5(B), OPL distribution of a CL cell is heterogeneous, whereas that of a WBC is smooth. These appearances can be clues for classification of cells.

As statistical subcellular-structures, five statistical parameters of the OPL/PL or OPL/D of single cells, namely, mean, standard deviation, maximum, skewness, and kurtosis, were utilized for the training data sets. Ensemble means of the five statistical parameters of OPL/PL and OPL/D over all 250 WBCs or 250 CLs are listed in Tables 1 and 2. After the above normalizations, the SVM classifier was trained on the five statistical parameters. The cost parameter (C) was searched for the maximum AUC. AUCs for each C, which varied from 2^{-1} to 2^8, were calculated by means of five-fold cross validation.

Subcellular features of single cells by HOG descriptor

We propose extracting subcellular feature from QPM images of single cells by a HOG descriptor, and training SVM classifier on the features. The procedure of the feature extraction was shown in Fig 4. After the longitudinal and lateral normalization, HOG feature was extracted and the gradient was compensated for the lateral normalization.

A HOG descriptor (see S3 Fig) enhances the feature of an image because the HOG descriptor counts occurrences of certain gradient orientations in localized portions of an image. Actually, the HOG descriptor extracted such gradients in Fig 5(C) and 5(D). In these figures show the visualized features extracted by the HOG descriptor (simply referred to as "HOG features") [48] of (a) the CL and (b) the WBC, respectively. An image of a single cell is divided into seven by seven compartments. The angles and lengths of the yellow lines in each compartment represent orientation and strength of spatial gradient in OPL of an image, respectively. In the preprocessing of the HOG feature extraction, images were normalized not only in the longitudinal but also in the lateral direction. After the longitudinal normalization in the same way as described in the previous section, Fig 4(B), the segmented images were resized to the size of 49×49-pixel images in Fig 4(C) because the HOG descriptor accepts only same-size images (i.e., the HOG descriptor is size-variant). The resizing of an image enhances or degrades the gradients. For instance, the gradient depicted by a red arrow is enhanced by the
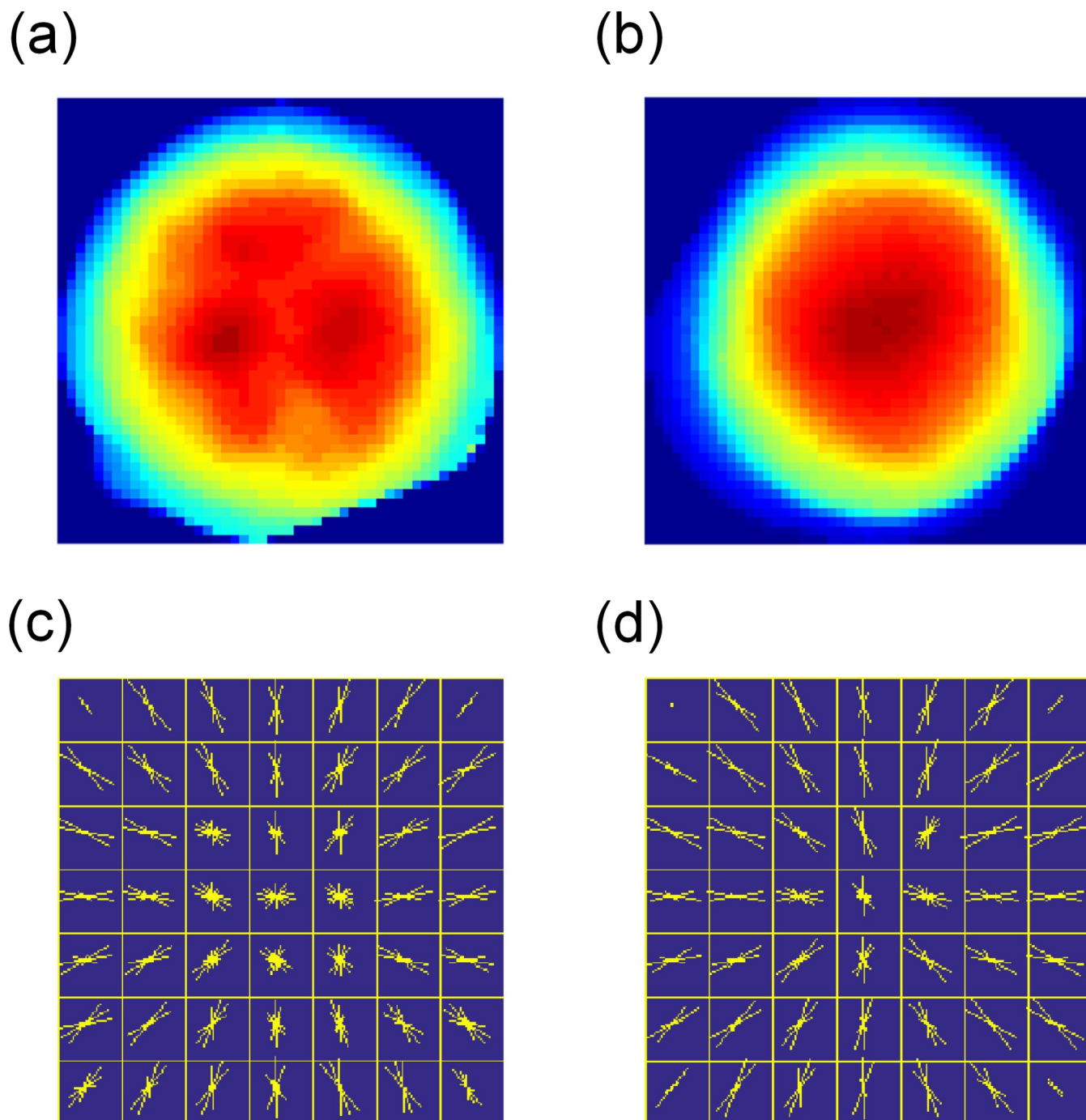
Fig 5. Representative QPM images and their visualized HOG features. (a) QPM image of cell line. (b) QPM image of WBC. In (a) and (b), pseudo color represents OPL. OPL is not normalized by diameter or path length. (c) Visualized HOG feature of (a). (d) Visualized HOG feature of (b). In (c) and (d), angles and lengths of the yellow lines in each compartment respectively represent orientation and strength of spatial gradient in OPL of the image.

resizing shown in Fig 4(C). To compensate the enhanced gradients, in the right panel in Fig 4 (C), OPL was divided by size-normalization factor m.

Training-data sets consisting of HOGs of 250 WBCs and HOGs of 250 CLs of QPM images normalized by path length or diameter were used. Maximum AUC was then found for by

varying the cost of SVM parameters from 2^{-1} to 2^8. AUCs for each parameter C were calculated by five-fold cross validation.

Image acquisition and data processing

The proposed QPM is a kind of interference microscopy, so the interferograms were captured by a CCD device and send to a computer, they were processed on MATLAB (Mathworks, Inc.) software to retrieve phase delay caused by OPL of a cell, and the phase was unwrapped by Goldstein's algorithm [49]. The QPM image of the cell is segmented so that it fit the size (width w and height h) in ImageJ. The segmented image (w by h) is divided by its path length or its diameter, and resized (size-normalization) to 49-by-49 pixels by factors $m_x = 49/w$ and $m_y = 49/h$. Factor m is averaged value of m_x and m_y. Enhancement of the spatial gradients of OPL due to the size-normalization was compensated by multiplying m to OPL in the resized image shown in Fig 4(C). After that compensation, HOG features of the resized-and-compensated image were extracted by utilizing the HOGDescriptor function in the open-source computer-vision library OpenCV (version 2.4.10). This function can setup the parameters of HOG, such as image size (49 by 49), block size (14 by 14), stride size (7 by 7), cell size (7 by 7), and number of orientation bins (9). These parameters of HOG were set to the values in parentheses. The part of the source code for utilization of OpenCV is shown in S1 Text. The HOG descriptor translated a two-dimensional image (49×49 = 2401 pixels) to vector $\boldsymbol{x}$ consisting of 1296 elements.

For the training data set used in machine learning, 250 HOG features (vectors) of WBCs (positive) and CLs (negative) images were collected. Since the sample size (250 pairs) was less than the number of elements (1296), a SVM with a linear kernel was used as the machine-learning algorithm. The SVM plots given vectors (x_i) of two classes in high-dimensional space and finds the optimal hyper plane that maximizes the margin between the two classes. It then builds a SVM classifier trained on the training images. The decision (discriminant) function [30,50] is given as

$$d_i = \frac{\mathbf{w}^T \mathbf{x}_i + b}{\|\mathbf{w}\|}, \tag{3}$$

where $\boldsymbol{x}_i$ is a vector of HOG, $\boldsymbol{w}$ is a weight vector, and b is a bias. If decision value d is positive, the cell having HOG $\boldsymbol{x}_i$ is classified as a positive image (i.e., a WBC). An SVM library, LIBSVM [51] version 3.20, on MATLAB software was utilized for that purpose. In the option of function svmtrain, svm_type was set to C-SVC (-s 0), kernel_type was set to linear function (-t 0), and other options were set to default values. Five-fold cross validation was taken to find the maximum area under the ROC curve (AUC) by varying cost (-c xx) parameters of the SVM. Cost "C" is the parameter for the soft-margin cost function, which controls the influence of each individual support vector [50]. Function svmpredict predicts a decision value of a given HOG feature. The script of MATLAB is shown in S1 Text.

Collection and processing of blood samples

Blood samples were obtained from a healthy donor, and ethical approval for the study was obtained from the Institutional Review Board of Hamamatsu University School of Medicine (No.16-101), and the methods were carried out in accordance with the approved guidelines. Informed consent was obtained in a written form from each donor before sample collection. The subjects consented to cooperate after they were informed that they would not incur any disadvantage, that they could resign from the study, that the researchers were obliged to protect their privileged information, and that their identities would not be revealed. They were

Table 1. Ensemble mean values of statistical parameters of OPL/PL over all WBCs and CLs.

	Mean	Std. dev.	Max.	Skewness	Kurtosis
WBCs	0.0185	0.00877	0.0378	1.039	0.0124
Cell lines	0.0208	0.00659	0.0397	-2.073	4.95

collected in tubes containing an anticoagulant (EDTA) and diluted with an equal volume of PBS. After carefully layering 4 ml of the diluted blood over 3 ml of Lymphoprep in a 15-ml centrifuge tube, the cells were centrifuged at 15,000 rpm at 4˚C for 30 minutes. Mononuclear cells were collected from the sample/medium interface using a Pasteur pipette. The cells were diluted with PBS to reduce the density of the solution, pelleted at 1500×g for 10 minutes, and then incubated in VersaLyse Lysing Solution (1 ml/5×10^5 leucocytes) for 10 minutes. They were pelleted again at 1500×g for 10 minutes to remove contaminating red blood cells and platelets. The obtained cells were suspended in PBS at 1.5×10^5 cells/ml. All the QPM images taken of both WBCs and CLs are shown in Fig 6.

Cells and cell culture

Human-colorectal-tumor cell lines HCT116, SW480, and DLD-1, human-pancreatic-tumor cell line Panc-1, and human-hepatocellular-liver-tumor cell line HepG2 were purchased from the European Collection of Authenticated Cell Cultures (ECACC, England, UK) and maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37˚C in a humidified atmosphere containing 5% CO_2.

Results

OPL normalizations for size-invariant feature extraction

As for size-invariant feature extraction, OPL was normalized by path length or diameter. Histograms of mean OPL/PL or OPL/D over all 250 WBCs or 250 CLs of the training data sets are shown in Fig 7(A) and 7(B). Although the diameters of the two classes differ significantly (Fig 3(C)), the histograms of OPL normalized by PL or D in Fig 7 almost overlap.

Training on statistical subcellular structures

The SVM classifier were built with the cost parameters C that gave the maximum AUC. Weight vectors [50] (**w**) of the classifiers are given in Fig 8. Both graphs indicate that the standard deviation is the main component for the classification in the case of both training data sets (OPL/PL and OPL/D). This result agrees with the fact that the standard deviations for the two classes (WBCs and CLs) significantly differ in Tables 1, 2, and S4 Fig. To evaluate the SVM classifier, histograms of decision values (i.e., outputs of the decision function given as x_i in Eq 3) of WBCs and CLs of the training-data sets are shown in Fig 9. In Fig 9(A) and 9(B), most cells are properly classified, but some cells still overlap. The plot of sensitivity versus 1 – specificity is a called "receiver operating characteristic (ROC)" curve, while the plot of 1 –sensitivity versus 1 –specificity is called a "detection-error trade-off (DET)" curve [33]. DET curves

Table 2. Ensemble mean values of statistical parameters of OPL/D over all WBCs and CLs.

	Mean	Std. dev.	Max.	Skewness	Kurtosis
WBCs	0.0146	0.00751	0.0267	0.334	1.235
Cell lines	0.0154	0.00609	0.0265	0.317	-0.9312
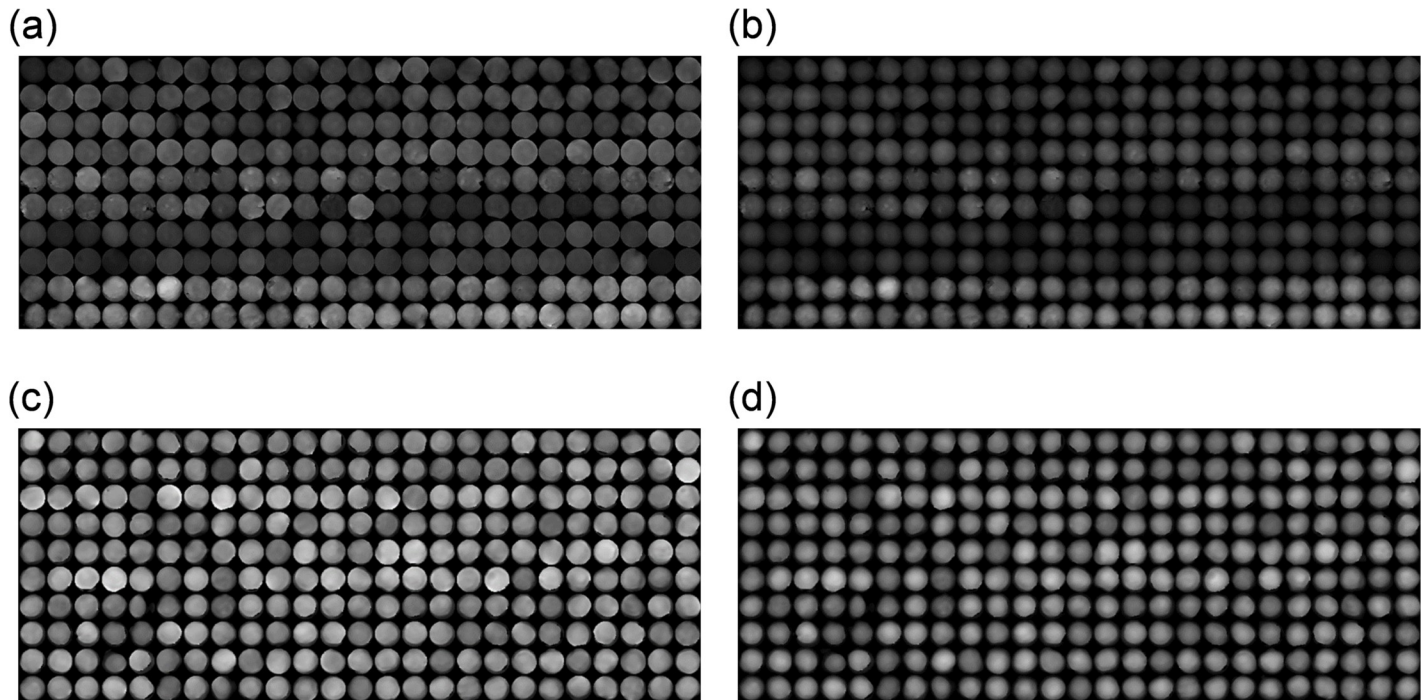
Fig 6. All QPM images (OPL) normalized by path-length or diameters. (a) OPL/PL: QPM images normalized by path lengths of cell lines; (b) OPL/D: QPM images normalized by diameters of cell lines; (c) OPL/PL: QPM images normalized by path lengths of WBCs; and (d) OPL/D: QPM images normalized by diameters of WBCs.

https://doi.org/10.1371/journal.pone.0211347.g006

Fig 7. Statistical characteristics of two classes of training-data sets after two OPL normalizations. (a) Histograms of mean OPL normalized by path length (OPL/PL) over all WBCs and CLs and (b) histograms of mean OPL normalized by diameter (OPL/D) over all WBCs and CLs.

https://doi.org/10.1371/journal.pone.0211347.g007

(a)

(b)

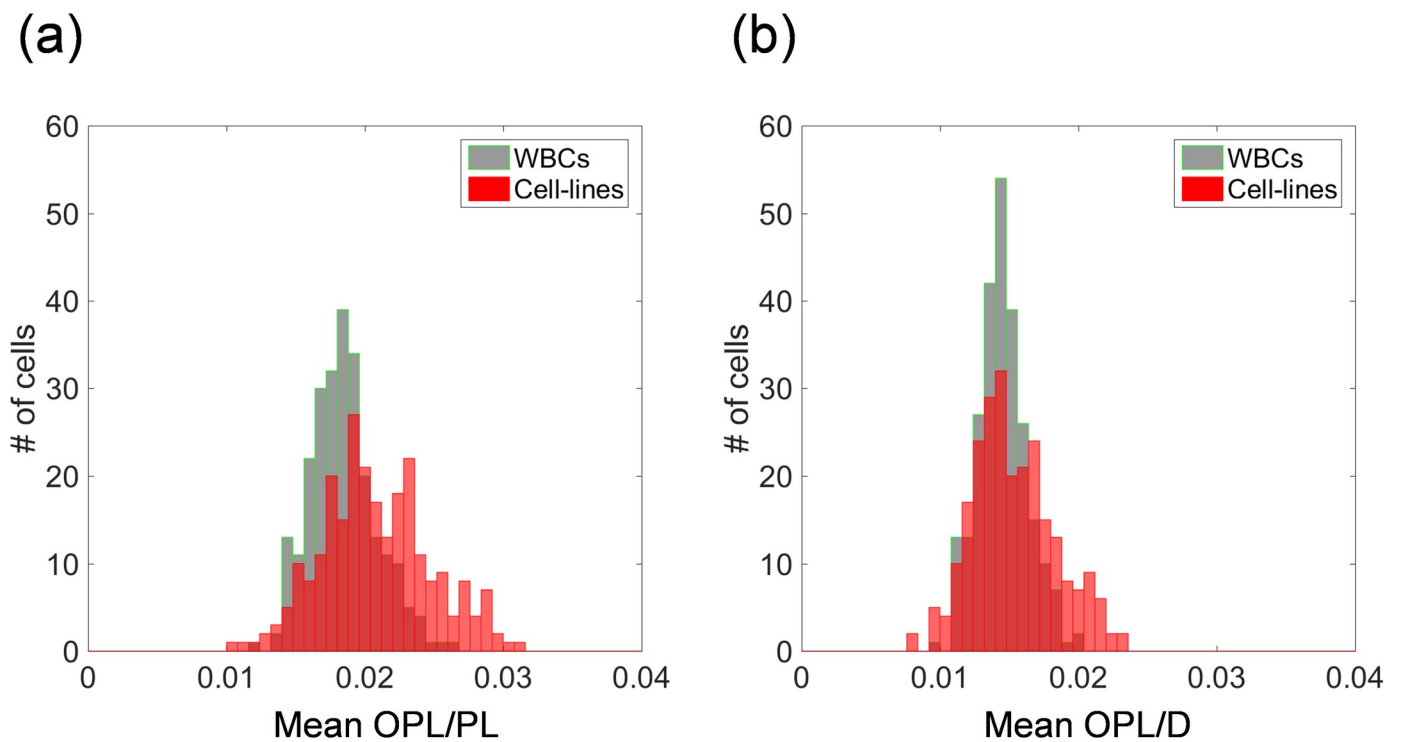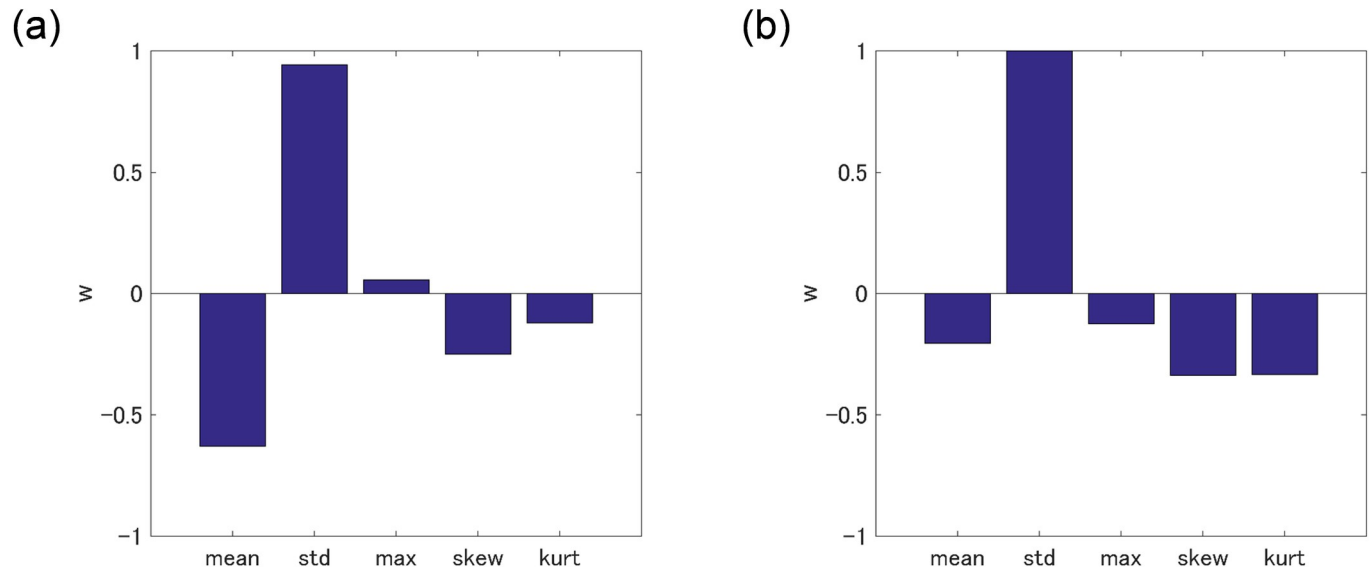**Fig 8. Weight vectors (*w*) of SMV classifiers visualized in five statistical parameters.** (a) Weight vector for statistical parameters of OPL/PL and (b) weight vector for statistical parameters of OPL/D.

for SVM classifiers trained by various training-data sets are shown in Fig 10. AUCs of ROC curves corresponding to the DET curves are shown in the legend of Fig 10.

Training on subcellular structure by HOG descriptor

SVM Classifiers were built with the C parameters that gave the maximum AUC. Weight vectors (*w*) of the models are visualized in Fig 11, where the angles and lengths of the yellow lines in each cell represent orientation and strength of spatial gradient of OPL/PL or OPL/D in the images. Most arrows of the WBC in the center region are short, implying that the OPL/PL or OPL/D is uniform in that region. In contrast, the magnitude and orientation of the arrows of the HOG feature of CLs vary, especially in the center of the cell. These observations agree with the diagnostic criteria of cancer cells or abnormal cells.

To evaluate the SVM classifier histograms of decision values (output of decision function x_i given in Eq 3) of WBCs and CLs of training-data sets are shown in Fig 9(C) and 9(D), respectively. The overlapping area of the histograms trained on HOGs of OPL/PL in Fig 9(C) is the smallest among all the panels in Fig 9. This result agrees with the facts that the corresponding DET curve is the lowest (red line) among the four DET curves and that the corresponding AUC (0.996) is the maximum listed in the legend in Fig 10.

Interpretation of mechanism of classifications of cells

To understand what the SVM classifier "sees" in the images, phantoms with bumps on a flat surface were tested. Those bumps imitate organelle inside a cell. The five type of phantoms used are shown in Fig 12. Forty-nine (seven by seven) bumps are set on the flat surface. The heights of the bumps are variable from 0% to 14% with respect to the height of the flat surface. The visualized HOG features of three phantoms (7.1%, 11%, and 14%) shown in Fig 12(A) are similar to those of CLs shown in Fig 5(C). In Fig 12(A), the decision values of each phantom decease gradually from 4.7 to -2.6 with respect to the height of the bump on the top-hat phantom. The phantom with bump height of more than 3.5% is classified as CLs. These simulations suggest that cancer cells can be distinguished according to their intracellular heterogeneity
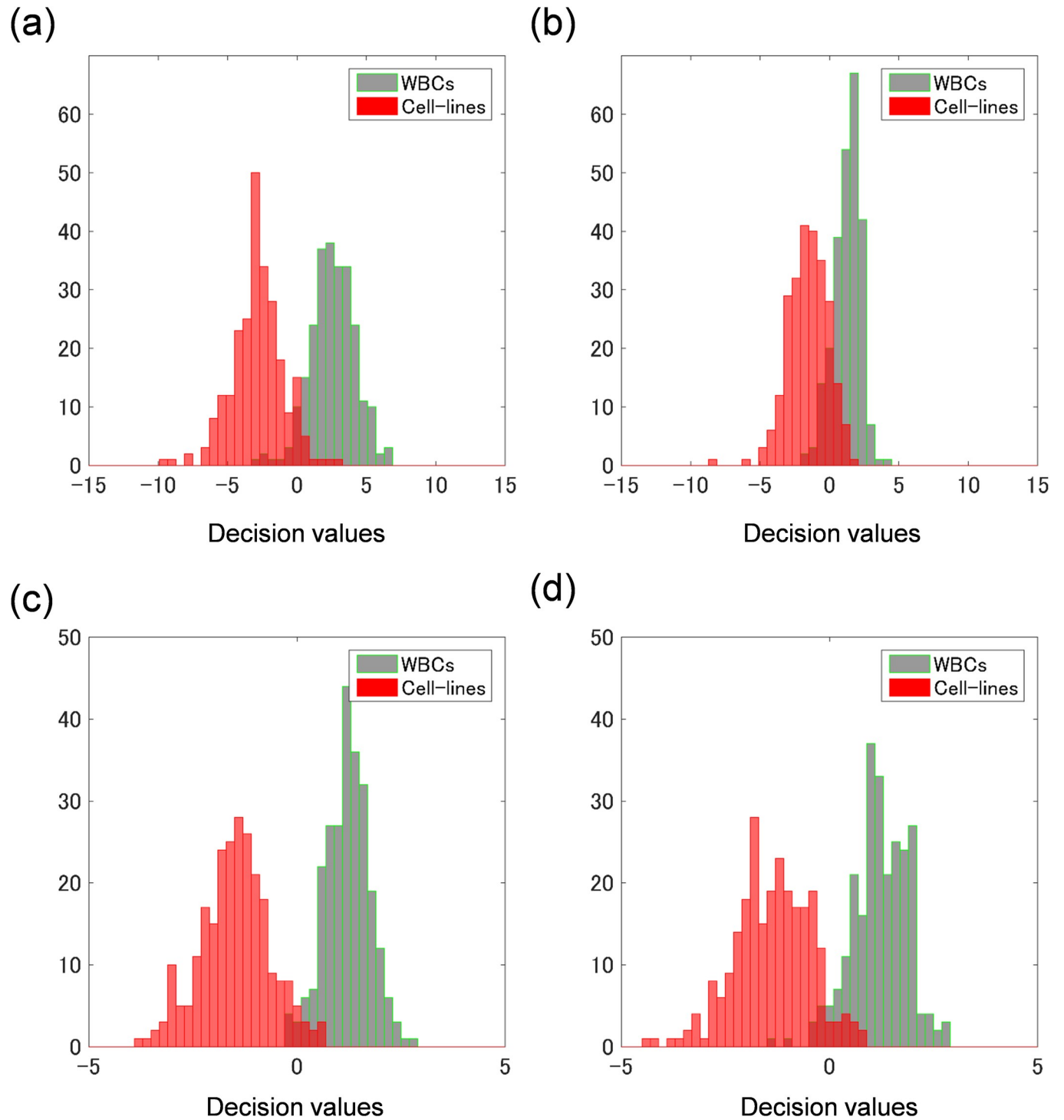
Fig 9. Histograms of decision values of training-data sets. Positive decision values for each cell are classified as a WBC, whereas negative decision values are classified as a cell line. Most WBCs (green bars) are classified as a positive decision value, and most cell lines (red bars) are classified as a negative one. The training data set consists of (a) statistical features of OPL/PL, (b) statistical features of OPL/D, (c) HOG features of OPL/PL, and (d) HOG features of OPL/D.

from WBCs. These simulations suggest that the proposed classifier can recognize subcellular heterogeneity, especially in the center of the cell.
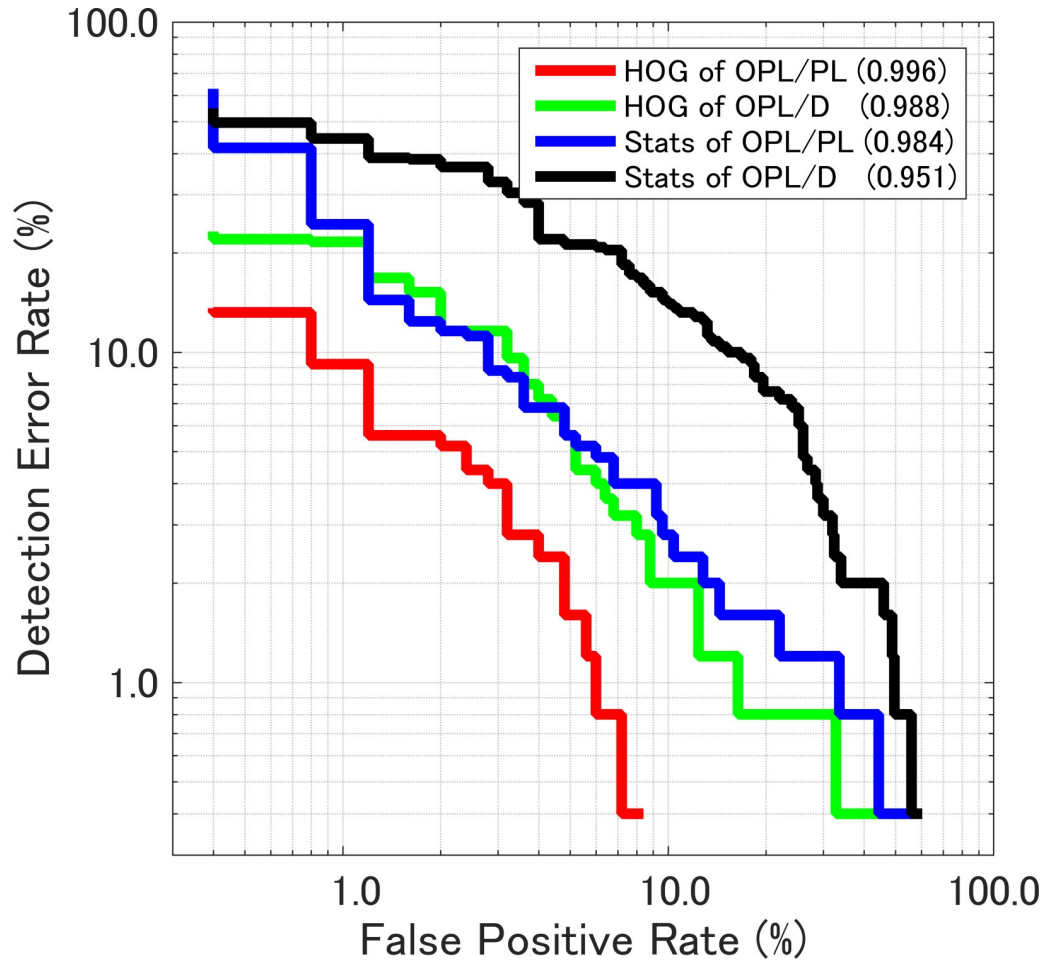
Fig 10. Detection-error trade-off (DET) curves for various training-data sets. Values in parentheses in the legend represent AUCs of their ROC curves.
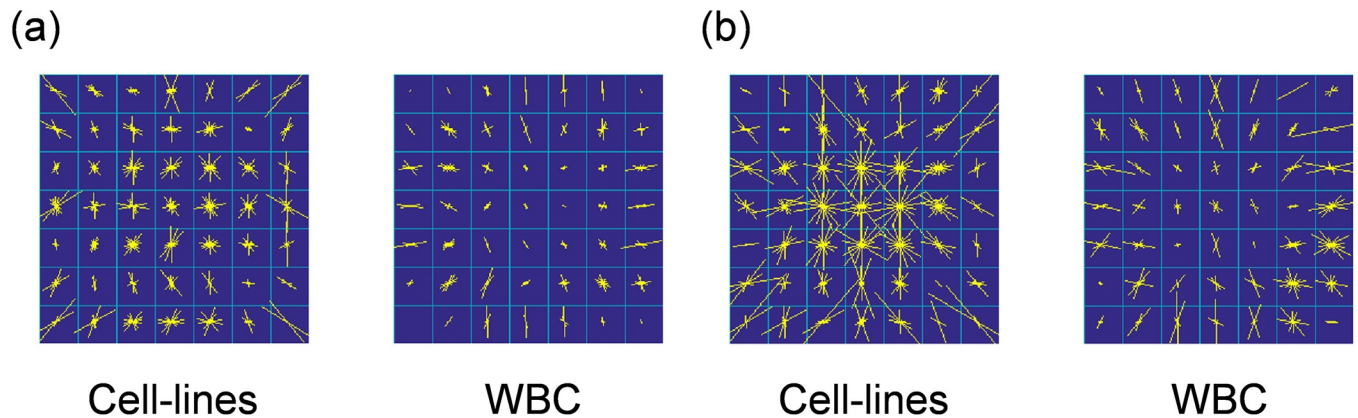
Fig 11. Weight vectors of SVM classifier found by parameter search visualized in seven-by-seven compartments. The angles and lengths of the yellow lines in each cell represent direction and strength of decline in an QPM image. (a) Weight vectors of QPM images (OPL) normalized by path length. Left side represents visualized weight vectors of HOG feature of cell lines. Right side represents vectors of WBCs. (b) Weight vectors of QPM images (OPL) normalized by diameter. In each panel, the left sides represent weight vectors of HOG feature of cell lines. The right sides represent the vectors of WBCs.
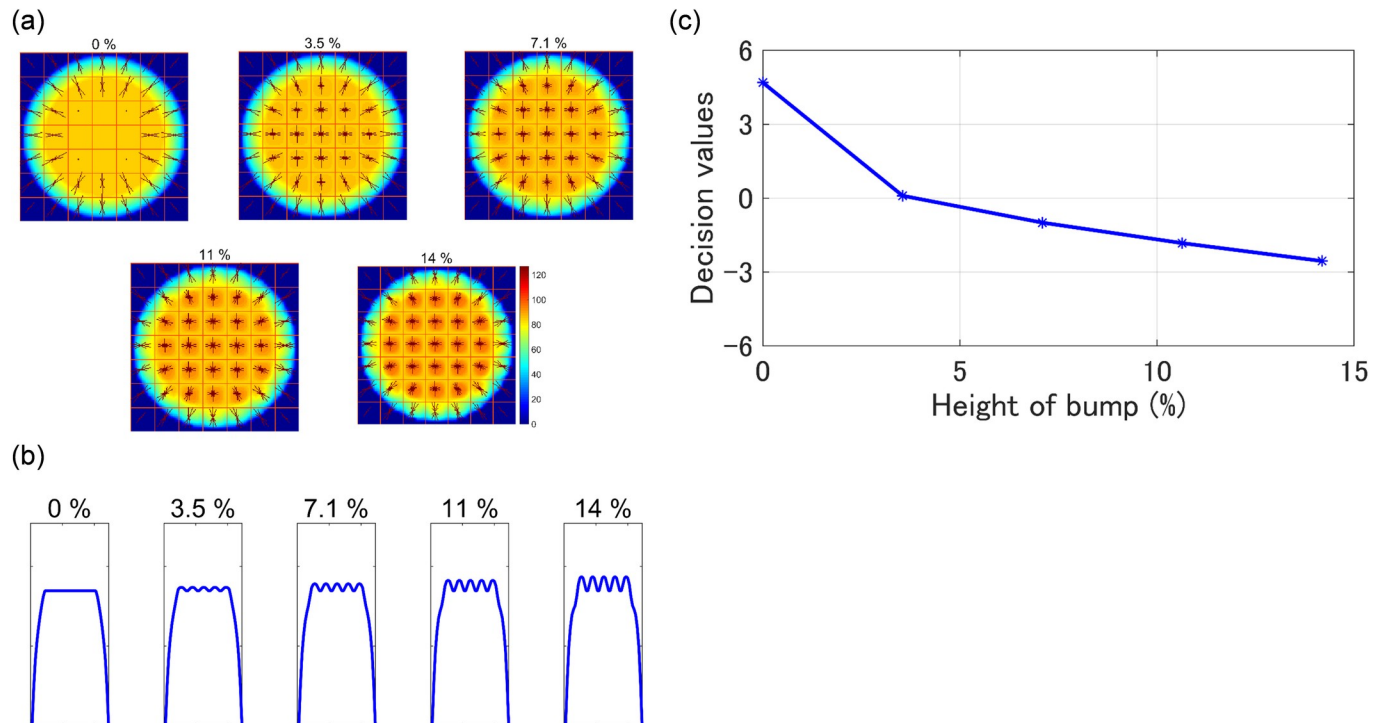
Fig 12. Classifications of heterogeneous phantoms. (a) Visualized HOG features superimposed on OPL/PL represented in pseudo color. The numbers above the figures are bump heights (%). (b) Cross sections of the phantoms. (c) Decision values of simulated phantoms with respect to bump heights are determined by the SVM classifier trained on HOG features of OPL/PL.

https://doi.org/10.1371/journal.pone.0211347.g012

Discussion

A method for label-free cell classification by computer-vision technologies for pattern recognition based on subcellular structures of QPM images of cells was demonstrated. This cell-recognition method differs from conventional methods in several respects. First, because QPM does not require a contrast agent nor staining to observe live cells, image acquisition by QPM is much less harmful to cells. This aspect is a great advantage for cell sorting based on features in QPM images. For example, the sorted cells can be cultivated to perform biological investigations. Second, cell classification is based on the heterogeneity of subcellular structures rather than the cellular outline used in conventional methods. The difference between weight vectors for WBCs and cancer cell lines in Fig 11 reflects the heterogeneity of subcellular components because compared with benign cells, malignant cells express a heterogeneous distribution of subcellular components [52]. As far as cancer CLs are concerned, the proposed classifier could find the difference between them.

In the sample preparation, Lymphoprep was used to separate mononuclear cells and platelets from red blood cells and granulocytes. Consequently, the target cells are lymphocytes, monocytes, and platelets. Platelets are smaller than the other cells. In morphology, nuclei of lymphocytes are round, and nuclei of monocytes are segmented. S3 Fig shows the projection of the RI along the optical axis (i.e., OPL) of a monocyte and lymphocyte. In contrast to these WBCs, nuclei of cultured cancer cells are expanded and bumpy. According to Figs 11 and 12, the proposed classifier appears to recognize the bumps which emulate cancer cell-lines' bumps. Although a number of different types of WBCs exist, the SVM classifier successfully classified those different types of WBCs as one class distinguishable from another class of CLs (i.e., two-class classification).

In this paper, all the phase images were not refocused [53] after segmentations. As for all the experiments on imaging six kinds of cells (DLD-1, HCT116, HepG2, Panc-1, SW480, and WBC), only the one experimenter carefully focused cells after they are settled at the bottom of the sample chamber, by adjusting the position of the optical system including objective lenses so as to focus entire cells in the field of view. This adjustment gives rise to defocus for a larger or smaller cell than the cells having average diameters even though only one kind of cell is in the field of view. The distributions of predicted diameters from images of WBCs in Fig 3(C), and five types of CLs are shown individually in S5 Fig. The distributions of CLs are broader than those of WBCs, and their images were more blurred (smoother, not bumpy). This fact implies that defocusing mostly occurs with CLs, and increases their decision values (i.e. shift to positive). Consequently, the defocusing adversely affected the classification performances (especially, specificity and AUC).

A HOG descriptor is generally rotation-variant. It is used for detecting pedestrians on roads because they can be supposed to be stand-up (no rotation). In contract, a single cell has no orientation (i.e., no left and right and no upside down). We confirmed that the built SMV classifier is robust in relation to rotation of a cell in S6 Fig. We also confirmed that the sample size is sufficient by drawing a learning curve (see S7 Fig). These facts indicate that a HOG descriptor is applicable to feature extraction of a single cell.

Among the label-free, subcellular classifications, the most-reliable classifier is trained on subcellular structures (HOG features) of OPL/PL [Fig 6(A) and 6(C) and Fig 9(A) and 9(C)]. As shown in Fig 10, AUC of this model is 0.996. This value might result from the OPL normalization shown in Fig 4(B). OPL divided by PL [expressed by Eq (2)] enhances bumps on the line profile of OPL/PL in Figs 4(B) and 11. The randomly oriented yellow lines representing the visualized weight vector of cell lines in Fig 11(A) exit all compartments, whereas the randomly oriented yellow lines of cell lines in Fig 11(B) exit only the center compartments in three-by-three compartments.

In the statistical subcellular-structures classification, according to the box plot in S4 Fig, it is predicted that adding statistical subcellular parameters (standard deviation, skewness, and kurtosis) to the cellular outline (mean and maximum) improved classification (Fig 7, Fig 10) up to the level of HOG-based classification.

The sensitivity and specificity of the classification on the basis of cell size are the best among the other parameters in Table 3 because the sizes of the two populations we prepared in this research seldom overlap in Fig 3(C). In future applications of the proposed classification, we are focusing on CTC [54] detection in an enormous number of nucleated cells such as WBCs in blood. In this application, it would be difficult to classify CTCs on the basis of cell size because it has been reported that there is considerable overlap between the two cell populations [55]. It is reported that the morphology of some CTCs has been reported to resemble that of primary cancer cells [55]. This fact imply that we can use training images of cancer CLs substitute for that of CTCs.

For future applications of this technology, we need to develop another technologies for acquiring QPM images of flowing cells, and sorting cells based on the label-free, subcellular classification This advanced technology will be a useful, non-cytotoxic, marker-free isolation

Table 3. Performance of the classifications on the basis of various parameters.

	Stats of OPL/PL	Stats of OPL/D	HOG of OPL/PL	HOG of OPL/D	Diameter
Sensitivity [%]	95.2	87.6	98.4	95.2	100.0
Specificity [%]	93.2	86.8	95.2	94.8	99.2
Accuracy [%]	94.2	87.2	96.8	95.0	99.6
AUC	0.984	0.951	0.996	0.988	0.999

https://doi.org/10.1371/journal.pone.0211347.t003

method for CTCs. Since cell surface markers are dependent on tumor types and the tumor progresses, conventional surface-marker-dependent methods for detecting CTCs miss some populations of CTCs [56]. A cell-sorting system based on the label-free, subcellular classification can also be applied to classification of WBC examination. According to recent studies, tumor cells interfere with T-cells through co-inhibitory factors such as programmed cell death 1 and cytotoxic T-lymphocyte antigen 4, and they regulate the immune response, promoting tumor progression [57]. Antibodies blocking these immune checkpoint proteins have demonstrated clinical activity in the treatment of advanced malignant melanoma and non-small-cell lung cancer. In addition, inhibitory ligands and receptors that regulate T-cell effector functions in tissues are commonly overexpressed on tumor cells and non-transformed cells in the tumor microenvironment. These findings imply that tumor cells affect lymphocyte functions, which may result in a change of the morphological features of lymphocytes. Close examination of cancer patient-derived WBCs obtained by the proposed label-free, subcellular classification method may clarify the morphological features of tumor-educated WBCs, and it will be useful for estimating the effect of immune checkpoint inhibitors.

In summary, we developed a method for image classification of live cells. The classification was executed by computer-vision approach for pattern recognition of the subcellular structure of images. These images are label-freely acquired by QPM. This method, label-free subcellular classification, successfully differentiated human WBCs from cancer cell-line cells. We revealed that what features SVM sees in subcellular structure of a single cell and it provides a clue for classifying them label-freely. This method is expected to apply to detecting CTCs in a non-cytotoxic manner, thus providing good opportunities for studying intact CTCs.

Supporting information

S1 Fig. Spatial noise distribution of a background image. The standard deviation of the spatial background noise in a phase image was 5 nm (50 milli-radians) in single optical-path-length, while the mean maximum OPL of WBCs was about 300 nm. The noise distribution in S5 Fig has offset from the center. This may be due to not sufficient masking cells.
(TIF)

S2 Fig. Images in the procedure for making a mask image, fitting a background image to a paraboloid, and subtracting the background image of a phase image from the original phase image. (a) original phase image with background curvature. (b) Paraboloid fitted to the phase image without masking cells. (c) phase image corrected by subtracting the paraboloid (b). (d) mask image to cover the cells on the phase image (a). (e) background image without cells (cells are masked by mask image (d). (f) paraboloid fitted to (a). (g) phase image corrected by subtracting the paraboloid (f); We compensated the difference in wave-fronts of the sample and reference light by fitting a background image to a paraboloid and subtracting it. In step one, a mask image (d) is extracted by fitting a paraboloid (b) to an original phase image (a) and setting a threshold (c) for distinguishing the background from objects. In step two, the original phase image is masked (e) by the mask image made in step one in order to obtain a background image without cells. Then, it was fitted to a paraboloid (f). Finally, a phase image corrected y subtracting the background image is obtained (g).
(TIF)

S3 Fig. Projection images of cells in terms of OPLs and their gradients. Projection images of a cell in terms of optical path length (OPL) are shown in S1 Fig. OPL is proportional to refractive index (RI) or physical path length. HOG describes spatial gradients of OPL corresponding to the inclination of OPL in S1 Fig. The directions of the red arrows represent the

directions of spatial gradients of OPL, and their lengths represent the magnitude of the spatial gradients. In practice, a captured QPM image is sectioned into 7×7 compartments (To avoid confusion, a cell, that is properly named in the field of computer vision, is referred to as a compartment), and the spatial gradient of OPL is visualized in each compartment. (a) schematic of a WBC, its profile of OPL, and visualized HOG feature (red arrows); and (b) schematic of a cancer cell, its profile of OPL, and visualized HOG feature (red arrows).
(TIF)

S4 Fig. Characteristics of five statistical subcellular structures. Five statistical parameters are plotted in Box and whisker plots. The first quartile (Q_1) and 3rd quartile (Q_3) are boxed. Interquartile range is referred to as IQR. The upper whisker is Q3+1.5IQR, and the lower whisker is Q1-1.5IQR. Outliers are plotted as red crosses. Mean values are expressed as circles. The red boxes represent CLs, and the green boxes represent WBCs. (a) Five statistical parameters of OPL/PL and (b) five statistical parameters of OPL/D.
(TIF)

S5 Fig. Distributions of predicted diameter of various types of cell-lines. Five types of cell-lines (DLD-1, HCT116, HepG2, Panc-1, and SW480) were imaged separately. We predicted the diameters of the segmented cells by averaging the width and the height of boundary box of a cell. No refocusing was done before segmentation of the cell in an image.
(TIF)

S6 Fig. Robustness of HOG to rotation of cell images. The robustness of the SVM classifier trained on OPL/PL shown in Fig 9(C) against rotation of images was tested as follows. Two representative QPM images of phantoms were chosen: a heterogeneous hemi-ellipsoid phantom with a bump height of 11% for CLs (a), and a homogeneous hemi-ellipsoid with a top-hat phantom for WBCs (b). Two phantom models are shown in panel (a) and (b) respectively as maps of OPL/PL and their cross-sections. These phantoms were rotated from 0 to 350˚ in 10˚ steps and classified by the built classifier. In panel (c), the WBC phantom (green line) showed almost no change in the decision value with respect to rotational angles, and the CL phantom (red line) showed a slight fluctuation in the decision value (which remained in the minus range). These results suggest that the effects of rotation of an image or cell are relatively small and do not affect the classification.
(TIF)

S7 Fig. Learning curve for sample sizes of HOG features of QPM images. It was confirmed that sample size is sufficient for a SVM by drawing the learning curve in S4 Fig. A SVM was trained on 250 images pairs (positive and negative image pairs). The images to be extracted HOG features are normalized by path length (OPL/PL). SVM parameter (C) is fixed at 16.
(TIF)

S1 Text. Source codes for extracting HOG features, training and predicting them.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Yusuke Ozaki, Hidenao Yamada, Hirotoshi Kikuchi.

Data curation: Hidenao Yamada, Toyohiko Yamauchi.

Funding acquisition: Hirotoshi Kikuchi, Toshiki Kawabata, Hiroyuki Konno.

Investigation: Yusuke Ozaki, Amane Hirotsu.

Methodology: Toyohiko Yamauchi.

Project administration: Hiroyuki Konno.

Resources: Tomohiro Murakami, Tomohiro Matsumoto, Yoshihiro Hiramatsu, Kinji Kamiya, Kentaro Goto.

Supervision: Yukio Ueda, Shigetoshi Okazaki, Masatoshi Kitagawa, Hiroya Takeuchi.

Visualization: Hidenao Yamada.

Writing – original draft: Yusuke Ozaki, Hidenao Yamada.

Writing – review & editing: Hirotoshi Kikuchi.

References

1. Cibas ES, Ducatman BS. Cytology: Diagnostic Principles and Clinical Correlates. third. Elsevier Health Sciences; 2009.

2. Neumann D, Langner T, Ulbrich F, Spitta D, Goehring D. Online vehicle detection using Haar-like, LBP and HOG feature based image classifiers with stereo vision preselection. 2017 IEEE Intelligent Vehicles Symposium (IV). 2017. pp. 773–778. https://doi.org/10.1109/IVS.2017.7995810

3. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In CVPR. 2005. pp. 886–893.

4. Shamir L, Delaney JD, Orlov N, Eckley DM, Goldberg IG. Pattern recognition software and techniques for biological image analysis. PLoS Comput Biol. 2010; 6. https://doi.org/10.1371/journal.pcbi.1000974 PMID: 21124870

5. Pan J, Kanade T, Chen M. Learning to detect different types of cells under phase contrast microscopy. Proc Microscopic Image Analysis with Applications in Biology. 2009;

6. Mualla F, Scholl S, Sommerfeldt B, Maier A, Hornegger J. Automatic cell detection in bright-field micro-scope images using SIFT, random forests, and hierarchical clustering. IEEE Trans Med Imaging. 2013; 32: 2274–2286. https://doi.org/10.1109/TMI.2013.2280380 PMID: 24001988

7. Essa E, Xie X, Errington RJ, White N. A multi-stage random forest classifier for phase contrast cell seg-mentation. Conf Proc IEEE Eng Med Biol Soc. 2015; 2015: 3865–3868. https://doi.org/10.1109/EMBC.2015.7319237 PMID: 26737137

8. Essa E, Xie X. Phase contrast cell detection using multilevel classification. Int J Numer Method Biomed Eng. 2017; https://doi.org/10.1002/cnm.2916 PMID: 28755437

9. Theriault D, L. Walker M, Wong J, Betke M. Cell morphology classification and clutter mitigation in phase-contrast microscopy images using machine learning. Machine Vision and Applications—MVA. 2011;23. https://doi.org/10.1007/s00138-011-0345-9

10. Bayramoglu N, Kaakinen M, Eklund L, Åkerfelt M, Nees M, Kannala J, et al. Detection of tumor cell spheroids from co-cultures using phase contrast images and machine learning approach. 2014 22nd International Conference on Pattern Recognition. 2014. pp. 3345–3350. https://doi.org/10.1109/ICPR.2014.576

11. Xu M, Papageorgiou DP, Abidi SZ, Dao M, Zhao H, Karniadakis GE. A deep convolutional neural net-work for classification of red blood cells in sickle cell anemia. PLoS Comput Biol. 2017; 13: e1005746. https://doi.org/10.1371/journal.pcbi.1005746 PMID: 29049291

12. Jiang Y, Lei C, Yasumoto A, Kobayashi H, Aisaka Y, Ito T, et al. Label-free detection of aggregated platelets in blood by machine-learning-aided optofluidic time-stretch microscopy. Lab Chip. 2017; 17: 2426–2434. https://doi.org/10.1039/c7lc00396j PMID: 28627575

13. Bhaskar D. Morphology based cell classification: unsupervised machine learning approach [Internet]. University of British Columbia. 2017. https://doi.org/10.14288/1.0345604

14. Blasi T, Hennig H, Summers HD, Theis FJ, Cerveira J, Patterson JO, et al. Label-free cell cycle analysis for high-throughput imaging flow cytometry. Nature Communications. 2016; 7: ncomms10256. https://doi.org/10.1038/ncomms10256 PMID: 26739115

15. Kobayashi H, Lei C, Wu Y, Mao A, Jiang Y, Guo B, et al. Label-free detection of cellular drug responses by high-throughput bright-field imaging and machine learning. Scientific Reports. 2017; 7: 12454. https://doi.org/10.1038/s41598-017-12378-4 PMID: 28963483

16. Kasprowicz R, Suman R, O'Toole P. Characterising live cell behaviour: Traditional label-free and quantitative phase imaging approaches. The International Journal of Biochemistry & Cell Biology. 2017; 84: 89–95. https://doi.org/10.1016/j.biocel.2017.01.004 PMID: 28111333

17. Kang M-S, Song S-M, Lee H, Kim M-H. Cell morphology classification in phase contrast microscopy image reducing halo artifact. Progress in Biomedical Optics and Imaging—Proceedings of SPIE. 2012. p. 34. https://doi.org/10.1117/12.908070

18. Nguyen TH, Kandel M, Shakir HM, Best-Popescu C, Arikkath J, Do MN, et al. Halo-free phase contrast microscopy. Sci Rep. 2017; 7. https://doi.org/10.1038/srep44034 PMID: 28338086

19. Kandel ME, Fanous M, Best-Popescu C, Popescu G. Real-time halo correction in phase contrast imaging. bioRxiv. 2017; 227025. https://doi.org/10.1101/227025

20. Yin Z, Kanade T, Chen M. Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation. Med Image Anal. 2012; 16: 1047–1062. https://doi.org/10.1016/j.media.2011.12.006 PMID: 22386070

21. Yoon J, Kim K, Park H, Choi C, Jang S, Park Y. Label-free characterization of white blood cells by measuring 3D refractive index maps. Biomed Opt Express. 2015; 6: 3865–3875. https://doi.org/10.1364/BOE.6.003865 PMID: 26504637

22. Mahjoubfar A, Chen C, Niazi KR, Rabizadeh S, Jalali B. Label-free high-throughput cell screening in flow. Biomed Opt Express, BOE. 2013; 4: 1618–1625. https://doi.org/10.1364/BOE.4.001618 PMID: 24049682

23. Chen CL, Mahjoubfar A, Tai L-C, Blaby IK, Huang A, Niazi KR, et al. Deep learning in label-free cell classification. Scientific Reports. 2016; 6: 21471. https://doi.org/10.1038/srep21471 PMID: 26975219

24. Roitshtain D, Wolbromsky L, Bal E, Greenspan H, Satterwhite LL, Shaked NT. Quantitative phase microscopy spatial signatures of cancer cells. Cytometry Part A. 2017; 91: 482–493. https://doi.org/10.1002/cyto.a.23100 PMID: 28426133

25. Park HS, Rinehart MT, Walzer KA, Chi J-TA, Wax A. Automated Detection of P. falciparum Using Machine Learning Algorithms with Quantitative Phase Images of Unstained Cells. Sullivan DJ, editor. PLOS ONE. 2016; 11: e0163045. https://doi.org/10.1371/journal.pone.0163045 PMID: 27636719

26. Dalal N, Triggs B. Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005. pp. 886–893 vol. 1. https://doi.org/10.1109/CVPR.2005.177

27. Ryota O, Hideto O, Ryoko U, Yukio K, Noriji K, Mikiko K, et al. Automatic detection of nucleated red blood cells from microscope images using Cell-Hog feature. J Jpn Soc Prec Eng. 2013; 79: 1074–1077. https://doi.org/10.2493/jjspe.79.1074

28. Jo Y, Park S, Jung J, Yoon J, Joo H, Kim M, et al. Holographic deep learning for rapid optical screening of anthrax spores. Science Advances. 2017; 3: e1700606. https://doi.org/10.1126/sciadv.1700606 PMID: 28798957

29. Kim GB, Jung K-H, Lee Y, Kim H-J, Kim N, Jun S, et al. Comparison of Shallow and Deep Learning Methods on Classifying the Regional Pattern of Diffuse Lung Disease. J Digit Imaging. 2018; 31: 415–424. https://doi.org/10.1007/s10278-017-0028-9 PMID: 29043528

30. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory. ACM Press; 1992. pp. 144–152.

31. Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995; 20: 273–297.

32. Schölkopf B, Burges C, Vapnik V. Extracting Support Data for a Given Task. Proceedings, First International Conference on Knowledge Discovery & Data Mining, Menlo Park. AAAI Press; 1995. pp. 252–257.

33. Martin AF, Doddington GR, Kamm T, Ordowski M, Przybocki MA. The DET curve in assessment of detection task performance. Proc Eurospeech, Rhodes, Greece. 1997; 1895–1898.

34. Iwai H, Fang-Yen C, Popescu G, Wax A, Badizadegan K, Dasari RR, et al. Quantitative phase imaging using actively stabilized phase-shifting low-coherence interferometry. Optics Letters. 2004; 29: 2399–2401. https://doi.org/10.1364/OL.29.002399 PMID: 15532280

35. Yamauchi T, Iwai H, Miwa M, Yamashita Y. Low-coherent quantitative phase microscope for nanometer-scale measurement of living cells morphology. Optics Express. 2008; 16: 12227–12238. https://doi.org/10.1364/OE.16.012227 PMID: 18679500

36. Yamauchi T, Iwai H, Yamashita Y. Label-free imaging of intracellular motility by low-coherent quantitative phase microscopy. Optics Express. 2011; 19: 5536. https://doi.org/10.1364/OE.19.005536 PMID: 21445192

37. Yamauchi T, Yamada H, Goto K, Matsui H, Yasuhiko O, Ueda Y. Transportable and vibration-free full-field low-coherent quantitative phase microscope. Quantitative Phase Imaging IV. International Society for Optics and Photonics; 2018. p. 105031U. https://doi.org/10.1117/12.2289473

38. Curl CL, Bellair CJ, Harris T, Allman BE, Harris PJ, Stewart AG, et al. Refractive index measurement in viable cells using quantitative phase-amplitude microscopy and confocal microscopy. Cytometry. 2005; 65A: 88–92. https://doi.org/10.1002/cyto.a.20134 PMID: 15800856

39. Zack GW, Rogers WE, Latt SA. Automatic measurement of sister chromatid exchange frequency. J Histochem Cytochem. 1977; 25: 741–753. https://doi.org/10.1177/25.7.70454 PMID: 70454

40. Dunn AK, Smithpeter CL, Welch AJ, Richards-Kortum RR. Finite-difference time-domain simulation of light scattering from single cells. J Biomed Opt. 1997; 2: 262–266. https://doi.org/10.1117/12.275219 PMID: 23014912

41. Schürmann M, Scholze J, Müller P, Guck J, Chan CJ. Cell nuclei have lower refractive index and mass density than cytoplasm. J Biophotonics. 2016; 9: 1068–1076. https://doi.org/10.1002/jbio.201500273 PMID: 27010098

42. Mir M, Tangella K, Popescu G. Blood testing at the single cell level using quantitative phase and amplitude microscopy. Biomed Opt Express. 2011; 2: 3259–3266. https://doi.org/10.1364/BOE.2.003259 PMID: 22162816

43. Iwai H, Yamauchi T, Miwa M, Yamashita Y. Doppler-spectrally encoded imaging of translational objects. Optics Communications. 2014; 319: 159–169. https://doi.org/10.1016/j.optcom.2014.01.013

44. Drezek R, Guillaud M, Collier T, Boiko I, Malpica A, Macaulay C, et al. Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, DNA content, and chromatin texture. J Biomed Opt. 2003; 8: 7–16. https://doi.org/10.1117/1.1528950 PMID: 12542374

45. Chalut KJ, Ostrander JH, Giacomelli MG, Wax A. Light Scattering Measurements of Subcellular Structure Provide Noninvasive Early Detection of Chemotherapy-induced Apoptosis. Cancer Res. 2009; 69: 1199–1204. https://doi.org/10.1158/0008-5472.CAN-08-3079 PMID: 19141640

46. Gschwendtner A, Hoffmann-Weltin Y, Mikuz G, Mairinger T. Quantitative assessment of bladder cancer by nuclear texture analysis using automated high resolution image cytometry. Mod Pathol. 1999; 12: 806–813. PMID: 10463483

47. Diaz G, Cappai C, Setzu MD, Diana A. Nuclear pattern recognition by two-parameter texture analysis. Comput Methods Programs Biomed. 1996; 49: 1–9. PMID: 8646832

48. Brauer J. HOG descriptor computation and visualization [Internet]. 3 May 2016 [cited 9 Jan 2018]. Available: http://www.juergenwiki.de/old_wiki/doku.php?id=public:hog_descriptor_computation_and_visualization

49. Ghiglia DC, Pritt MD. Two-dimensional phase unwrapping: Theory, Algorithms, and Software. New York etc.: Wiley-Interscience; 1998.

50. Ben-Hur A, Weston J. A user's guide to support vector machines. SpringerLink. 2010; 223–239. https://doi.org/10.1007/978-1-60327-241-4_13

51. Chang C-C, Lin C-J. LIBSVM: A Library for Support Vector Machines. ACM Trans Intell Syst Technol. 2011; 2: 27:1–27:27. https://doi.org/10.1145/1961189.1961199

52. Gerö E, Hageman P, Hilkens J, Sugar J, Hilgers J. Diversity of Antigenic Expression on Normal and Malignant Human Mammary Cells Recognized by a New Panel of Monoclonal Antibodies. Immunological Approaches to the Diagnosis and Therapy of Breast Cancer. Springer, Boston, MA; 1987. pp. 95–118.

53. Langehanenberg P, Kemper B, Dirksen D, von Bally G. Autofocusing in digital holographic phase contrast microscopy on pure phase objects for live cell imaging. Applied Optics. 2008; 47: D176. https://doi.org/10.1364/AO.47.00D176 PMID: 18594573

54. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004; 351: 781–791. https://doi.org/10.1056/NEJMoa040766 PMID: 15317891

55. Ozkumur E, Shah AM, Ciciliano JC, Emmink BL, Miyamoto DT, Brachtel E, et al. Inertial Focusing for Tumor Antigen–Dependent and–Independent Sorting of Rare Circulating Tumor Cells. Sci Transl Med. 2013; 5: 179ra47. https://doi.org/10.1126/scitranslmed.3005616 PMID: 23552373

56. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009; 119: 1420–1428. https://doi.org/10.1172/JCI39104 PMID: 19487818

57. McDermott DF, Atkins MB. PD-1 as a potential target in cancer therapy. Cancer Med. 2013; 2: 662–673. https://doi.org/10.1002/cam4.106 PMID: 24403232