MICROSCOPY

Microscopy, 2020, 1–8
doi: 10.1093/jmicro/dfaa005

Invited Special Issue

Implementing machine learning methods for imaging flow cytometry

Sadao Ota[1,2,3,*], Issei Sato[3,4,5] and Ryoichi Horisaki[2,6]

[1]Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan, [2]JST, PRESTO, 4-1-8 Honcho, Kawaguchi-shi 332-0012, Saitama, Japan, [3]Thinkcyte Inc., 7-3-1 Hongo, Bunkyo-ku 113-8654, Tokyo, Japan, [4]Department of Computer Science, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku 113-8654, Tokyo, Japan, [5]RIKEN AIP, Nihonbashi 1-chome Mitsui Building, 15th floor, 1-4-1 Nihonbashi, Chuo-ku, 103-0027, Tokyo, Japan and [6]Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

*To whom correspondence should be addressed. E-mail: sadaota@solab.rcast.u-tokyo.ac.jp

Received 14 October 2019; Revised 19 January 2020; Editorial Decision 4 February 2020

Abstract

In this review, we focus on the applications of machine learning methods for analyzing image data acquired in imaging flow cytometry technologies. We propose that the analysis approaches can be categorized into two groups based on the type of data, raw imaging signals or features explicitly extracted from images, being analyzed by a trained model. We hope that this categorization is helpful for understanding uniqueness, differences and opportunities when the machine learning-based analysis is implemented in recently developed 'imaging' cell sorters.

Key words: optical imaging, machine learning, image analysis, flow cytometry, cell sorting

Introduction

The aim of machine learning is to achieve human-like behavior in a machine by extracting the hidden properties of observed examples through the use of empirical data and past human experiences. Machine learning is related to many fields, including probability theory and statistics, computational neuroscience, computer science and statistical physics, and has a range of applications, such as in natural language processing, computer vision, recommendation systems, speech recognition, bioinformatics and medical image analysis. In the analysis of microscopic cell images, remarkable advances have already been reported in the application of powerful machine learning methods [1].

Imaging flow cytometry (IFC) is an experimental technique combining the spatial resolution of optical microscopy with the fast sampling capability of flow cytometry [2,3]. After the first experimental demonstration in 1979 [4,5], a commercial IFC system was introduced by Amnis Corporation (acquired by EMD Millipore in 2011 and currently part of Luminex Corporation). This commercial machine boosted the research area of the application of IFC to a variety of biological and clinical studies worldwide. Recent biological studies [6,7,8], for example, have utilized IFC to effectively analyze the cell cycles, disease states, and cell divisions and characterize localized fluorescence molecules to study DNA damage and repair processes, and in fluorescence *in situ* hybridization (FISH). Recent studies in clinical fields have also shown that IFC can be a useful tool for the detection of disease cells, including disease and cancerous cells and infectious pathogens, based on the high content of information mapped in the images [9,10]. Cell image data generated by IFC systems in many laboratories are now shared across biological and technological groups in different institutions to promote and accelerate fruitful collaborations. Meanwhile, in recent decades, the throughputs and modalities of other IFC systems have continually increased beyond those of Amnis' IFC systems and have diversified according with the advances made in measurement and analysis technologies [11–13].

A range of software tools that employ machine learning methods in microscopic cell image analysis have already been applied to the image data of IFC. But why do we need IFC? Compared with the conventional high-content analysis of images taken using non-flow microscopy methods, IFC cannot capture dynamic cell information or

cell–cell interactions in complex cellular structures, such as in tissues or organs, and is limited in terms of its spatial resolution. However, despite these disadvantages, IFC is capable of high-throughput imaging of many flowing single cells with their positions controlled in space. The simplicity of both the measurement and analysis of IFC allows for the generation of a large data set of images with smaller risks of artifacts, to which machine learning-based analysis is readily applicable. Specifically, the fact that each IFC image captures a distinct single cell alleviates the need for cell segmentation in image processing. These characteristics of IFC are the most advantageous for cells in suspension, such as in blood, saliva and urine samples.

Another significant advantage of IFC is its capability to selectively sort cells based on the analysis of their image information. The availability of sorted cells is expected to allow for the use of functional cell products in therapeutic applications, as well as the further analysis of single cells with a variety of molecular methods, including genomics, transcriptomics and proteomics. However, increasing its throughput is not straightforward due to the challenges associated with speeding up the real-time analysis of high content image information. Pioneering efforts for the sorting of cells based on the analysis of CCD camera images started in the early 2000s, to our knowledge, as shown in Fig. 1a [14–16]. More recently, researchers have reported developments in cell sorting based on image information analysis, with significant increases in throughputs by integrating machine learning methods with the real-time analysis of temporal signals recorded by single pixel imagers, as shown in Fig. 1b and c [13,17,18].

In this review, after summarizing the advances made in machine learning-based cellular image analysis over the years, we will discuss how machine learning methods can be implemented in IFC platforms.

Machine learning methods

We humans typically decide how to behave based on knowledge learned from past experiences. In real life, the set of all possible solutions for all possible problems is too large to be covered by a set of observed examples. Moreover, a large amount of information makes it difficult for us to learn useful knowledge and give accurate solutions to problems. Therefore, we need to generalize the given examples to make useful decisions for future observations. This ability humans possess is called 'generalization'. Machine learning is a subtopic of artificial intelligence that aims to achieve the ability of generalization, more concretely, developing systems that automatically learn underlying mechanisms from a large number of observed data and that make predictions on phenomena yet to be observed.

Machine learning tasks are categorized into two basic tasks according to the type of observed data: supervised and unsupervised learning. A supervised learning task acquires the input-output relationship, $y = f(x)$, given the training data where x is the input, y is the output and f is a function indicating the relationship. A major supervised learning task is a classification that constructs a function to map inputs into labels. An unsupervised learning task deals with only the inputs and acquires the decomposition of input data into a function common to all input data and data-specific representations, i.e. $x = f(z)$, where x is the input, f is a function common to all input data and z is a data-specific representation for input x, called a latent variable.

Kernel methods are a class of well-known machine learning algorithms in which support vector machine [19,20], kernel principal component analysis [21] and kernel k-means method [22] are typical algorithms in supervised and unsupervised learning. For kernel methods, the features of the observations are provided intrinsically on the basis of a positive-semi-definite symmetric kernel function, which can be interpreted as a similarity measure in a Hilbert space.

That is, formulating the similarity measure between data designs the feature space of the data. The kernel function can be learned from observations [23]. Learning algorithms based on kernel methods are statistically well-defined and stable in terms of optimization, and their generalization abilities are solidly analyzed using statistical learning theory.

Boosting is a family of machine learning algorithms that convert weak learners into strong ones [24,25]. Gradient boosting [26] is one of the most widely used algorithms in machine learning competitions such as kaggle. The term weak/strong learners are defined in probability approximately correct (PAC) learning [27] where the strong learner is a PAC learnable learner such as support vector machine, and the weak learner is not. The 'PAC learnable' means that, for any approximate error and confidence, there must be a polynomial-time algorithm that outputs a hypothesis that satisfies a concentration inequality for the misclassification rate. On the other hand, the weak learner satisfies a concentration inequality for the misclassification rate for a 'fixed' approximate error and confidence, but not for 'any'. Most boosting algorithms iteratively learn weak classifiers and combine them into a strong classifier such as weighted voting of classifiers' outputs. Boosting algorithms are also statistically well-defined and easy to implement, and their overfitting properties are well analyzed.

Deep learning is a class of machine learning methods that uses multi-layer neural networks [28]. The name 'deep' comes from a structure consisting of many, often hundreds of, layers of artificial neurons stacked over each other. Briefly, deep learning is who to acquire synthetic functions with a large number of simple nonlinear functions, $y = f\mathrm{L}(f\mathrm{L}{-}1(\ldots(f3(f2(f1(x))))))$ from a large number of training examples $\{(xi,yi)\}$, where L is the number of layers and simple nonlinear functions $\{fi\ ()\}$ are often composed of affine transformations and an activation function such as a rectified linear function [29]. Multiple abstractions with multi-layer stacking are considered to give effective representations of the data.

Deep learning is not a specific method but a set of various techniques for learning complex neural networks such as network architecture, activation functions, optimization methods, loss functions, evaluation metrics, preprocessing, data augmentations and many heuristics. The basic model is the convolutional neural network (CNN), which is composed of a convolutional layer and pooling layer [30]. Although it has been studied for several decades, it was the focus in the 2012 ImageNet Large Scale Visual Recognition Challenge. Many algorithms have been proposed since that event: Alex Net (8 layers) [31], VGGNet (19 layers) [32], GoogLeNet (22 layers) [33] and ResNet (152 layers) [34].

The selection of machine learning algorithms has to be designed to solve problems specifically. In an example of cell classification in IFC, a deep learning algorithm was previously found to outperform a boosting algorithm in terms of accuracy; however, the boosting algorithm seemed to be better than the deep learning algorithm when evaluated with F1 macro [7]. Since we currently have no theoretical reason that DNN outperforms the boosting algorithms or vice versa, it is therefore essential to know what kind of algorithm works better for each type of problem.

Machine learning in microscopy cell image analysis

Here, we briefly review four use cases of machine learning in microscopy image analysis: image classification, image segmentation,

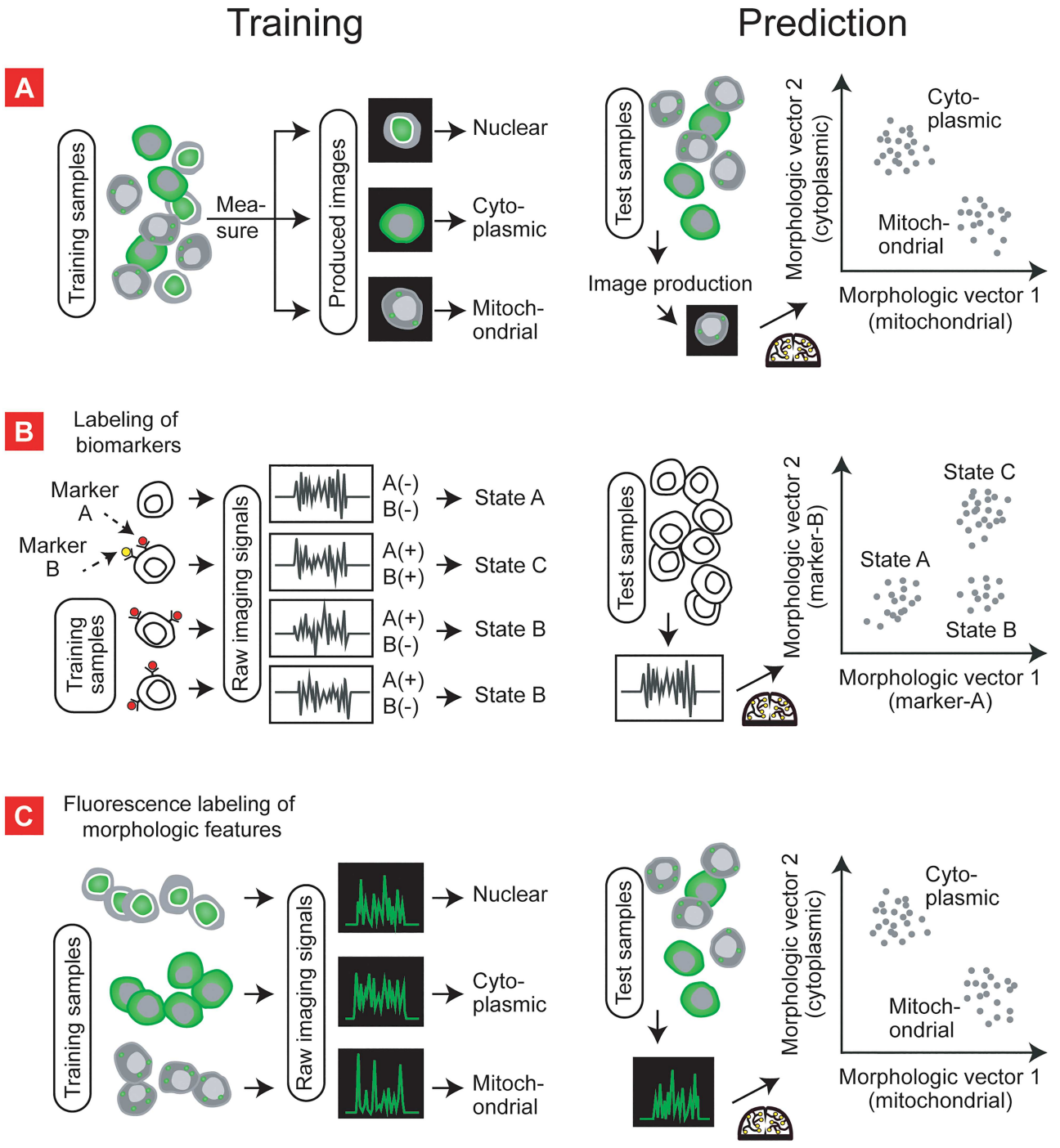

Fig. 1. Implementation of supervised machine learning methods to imaging flow cytometry. (a) Cell classification based on the analysis of explainable features extracted from the produced images. Left: workflow of constructing a training data set. After measuring the samples and producing their image library, the library is analyzed to explicitly extract and evaluate the morphological features. Right: the trained model is used to predict a label reflecting the morphological features from the produced image. (b) Application of a model directly to raw imaging signals, which are signal pixel intensities in an arrayed pixel detector or their equivalents (depicted here as waveforms from a single pixel detector as an example) for the prediction of cell information such as cell type, state and differentiation. Left: workflow involved in the construction of a training data set. After labeling the cells with fluorescently labeled biomarkers, raw imaging signals are measured with the label information. The imaging signals are paired with the classification results based on analysis of the label information. Right: directly from the raw imaging information, the trained model is used to predict a label reflecting the morphological information relevant to the biomarker labeling. (c) Application of a model directly to the raw imaging signals which are the signal pixel intensities or their equivalents (depicted here as waveforms again) for the prediction of explainable feature information of cells. Left: workflow of constructing a training data set. After fluorescently labeling the features of the cells in advance, raw imaging signals are measured with the labeled feature information. Right: the trained model is used to predict a label reflecting the morphological features directly from the raw imaging information.

object tracking and image generation. Please refer to the articles provided by Xing *et al*. [35] and Moen *et al*. [1] for a detailed survey of these. Although not all of the cases are not applicable to the image data obtained in IFC, we believe that overview of this advancing research field is helpful to foresee the future combination of machine learning and IFC.

Image classification involves predicting a label to an image. One example would be identifying whether or not a stem cell has been differentiated. A fluorescent marker of differentiation is often used as a ground truth, i.e. a true label, where a classifier is learned to identify and prospectively predict differentiation of cells from bright-field images [1,36]. The automatic detection of circulating tumor cells in blood has been developed using the microscopic image of unstained cells [37]. Supervised classification was recently used to classify spatial patterns in fluorescence images, such as protein localization [38]. In another study, a machine learning algorithm outperformed experts in identifying a specific protein localization in high-throughput fluorescence microscopy imaging [39].

Image segmentation involves partitioning an image into several meaningful parts. One example would be identifying single cells in microscopy images. U-Net [40, 41] and DeepCell [42] are representative methods that formulate single-cell segmentation as a pixel-level classification, where pixel-level predictions of the cell interiors, cell edges and background are obtained. The performance of image segmentation is important for object tracking.

Object tracking involves following objects through live-cell time-lapse imaging. One biological application of this is single-cell tracking in live-cell imaging movies. The supervised machine learning method can robustly segment the fluorescent images of cell nuclei as well as phase images of the cytoplasm of individual bacterial and mammalian cells from phase-contrast images without the need for a fluorescent cytoplasmic marker, allowing for simultaneous segmentation and identification of different mammalian cell types grown in co-culture [43,44]. DeepLabCut is an object-tracking package developed to quantify cell behavior during neuroscience experiments [45]. It provides an efficient system using markerless pose estimation based on transfer learning; thus, it can be applied in several kinds of neuroscience experiments if minimal training data are prepared.

Image generation involves generating synthetic images representing the characteristics of training data. It is typically based on a generative model that applies unsupervised learning. When we obtain the decomposition of an image x, $x = f(z)$, where f is a function common to training images and z is a specific low-dimensional representation, called a latent variable, for image x, we obtain a synthetic image similar to x by manipulating z in the latent space. Generative models have a wide variety of applications and models, e.g. auto-encoding models [46,47], generative adversarial networks [48–50] and flow-based models [51–53]. The targets of generation models are not limited to image data. One biological example is the generation of protein structures for applications in fast *de novo* protein design [54].

Generative modeling for fluorescence images has been studied [55,56]. These studies aim at learning to generate cell and nuclear morphology and at predicting the location of subcellular structures from fluorescence images. Image-to-image translation methods of medical blood smear data have been developed to generate new samples for data augmentation aiming at meaningfully increasing small datasets [57]. These methods generate photorealistic images of blood cells given the segmentation mask of the microscopy image. The generated images are used alongside real data during training for segmentation and object detection tasks.

Implementing machine learning-based analysis in IFC

These fast-evolving machine learning methods have also been applied to image data acquired using IFC technologies. Under the assumption that the effective application of supervised machine learning methods requires three key steps, i.e. (i) constructing a training dataset; (ii) training models on that dataset; and (iii) deploying the trained model for analyzing new data; the difference between IFC and non-flow microscopes appears to lie in Step (iii) when one performs cell sorting by deploying the trained model for the real-time analysis of imaging information.

For analyzers, the training model can be applied offline to a new image that is downloaded from a measurement instrument to a local computer. In this case, abundant computational resources, such as commercial central processing units (CPUs) and graphics processing units (GPUs), can be applied with abundant calculation time to complete heavy analysis. On the other hand, to allow for cell sorting based on the real-time analysis of image information, the calculation has to be completed within a limited time scale: the calculation starts after the measurement of a cell and has to be finished before sorting the same cell inside the same fluidic device. For example, it is often the case that a cell travels from the measurement site to the sorting site within 1 ms in commercial high-throughput cell sorters. Considering possible variations in the calculation time and the flowing speed, the calculation time in average is desired to be much shorter than that. To satisfy this requirement, in recently developed systems, trained models were implemented onto a limited choices of information processing units, such as field-programmable gate arrays (FPGAs) or specialized calculation units [13,17,18].

While novel architectures are being theoretically proposed (or currently under intensive developments), processing hardware that is able to perform intensive calculations for high content analysis at a high speed is not yet available. Consequently, the speed and depth of the real-time analysis and the size of the image information are at the trade-off. It is thus an open question how to implement any machine learning model used offline, such as very deep neural networks, constructed with more than 100 layers with a very high throughput, such as in commercial fluorescence-activated cell sorters. We believe that further technological evolutions will bridge the gap between the models that can be used offline and online (real-time).

In the following sections, we will discuss how machine learning methods have been implemented in the flow-based imaging analysis and sorting. First, we apply a new categorization based on the modality being analyzed to approaches for supervised image classification, and divide them into two groups, a group depicted in Fig. 1a, and another group in Fig. 1b and c. The first and traditional approach has a step involving the extraction of the morphological features of images and their evaluation to assign labels to the image information. The second and relatively new approach that has gained in accuracy with the power of machine learning is the application of a model directly to raw imaging information, such as pixel values from pixel-arrayed cameras and temporal waveforms from single pixel imagers. Next as shown in Fig. 2, we introduce two types of cell sorters based on the real-time analysis of image information, either with or without image reconstruction, which will be referred to as 'imaging' cell sorters in this review.

The discussions presented here focus on the strategies used for image classification, the task of assigning a meaningful label to each image information, and for supervised methods which have been the most successful for maximizing the performance on labeled data sets. While it is out of scope in this review, we can foresee that augmented microscopy will also be realized using IFC platforms

soon. We would also like to note that image segmentation is still a very important task in IFC, especially when cell–cell communications are of interest, as well as when a specific type of organelles are being segmented. Moreover, we also note that there is significant potential in unsupervised learning approaches, which may become more powerful in upcoming technological realization of index sorting of many single cells followed by single cell analysis.

Analyzing explicit features extracted from images produced in IFC

Figure 1a depicts a simplified example workflow of the first category of machine learning-based image classification, based on the extraction of explicit features. Here, one would be interested in discriminating whether a fluorescently labeled protein is localized within the nucleus or the cytoplasm. The workflow starts with the construction of a training dataset comprising pairs of an image and information derived from extracted features (Fig. 1a left), then proceeds to the training of models on that dataset, and ends up with the deployment of the trained model onto the new imaging data for the prediction of feature vectors reflecting morphological features (Fig. 1a right). The features to be extracted include intensity, size, shape, texture, correlations and subcellular components, as well as their combinations for more complex phenotyping that manifests in multiple features at a time.

The features can be annotated and extracted manually using our eyes and/or automatically using image analysis algorithms and software. The visual assessment of the cell images is the most widespread approach used in both biological research and clinical studies to date. However, considering the large scale of datasets obtained with IFC technologies, which can easily exceed thousands of images, employing trained and expensive human experts to analyze images efficiently and with the least number of discrepancies is challenging. It is therefore reasonable to synergistically combine the power of human-based and software-based recognitions. In classical image processing implemented in analysis software, a researcher designs an algorithm to recognize each cell and draw the borders between the cells and the relevant subcellular compartments, such as nuclei in the images, whereby various measurements over the identified regions can be performed at a time.

For the purpose of extracting image features, fluorescence imaging methods are likely to be the best suited, as many of fluorescence labeling methods allow for the extraction of information on specific biological features with a very high specificity and signal-to-noise ratios. Fluorescent proteins or antibody-based labeling are good examples of revolutionary tools used to determine the existence of molecules of interest either genetically or immunochemically. Other immunohistochemical methods have also been also commonly used for the purpose of enhancing the visibility of specific molecules and for cell phenotyping.

Machine learning-based direct analysis of imaging information in IFC

Figure 1b and c depict simplified example workflows of the second category for predicting labels by directly applying trained models for the analysis of raw images (pixel intensities) or their equivalent, such as compressed imaging waveforms. In other words, this direct analysis approach performs image analysis without explicitly extracting morphological features from the raw image information.

Direct analysis is powerful particularly when we do not know or cannot easily tell which feature to recognize in the images to efficiently identify the state of the cells, their phenotypes or their functions, as depicted in an example workflow of label-free (non-fluorescence) morphological analysis in Fig. 1b. In a recent article [58], Doan *et al.* described a workflow of analyzing IFC images with deep learning methods: after training a model with example images that are known to have a particular phenotype, the model takes the pixels of images as its input and predict an answer for new images. The direct analysis approach was also adopted in a recent study using conventional microscopy, wherein a fluorescent marker of differentiation was used as a ground truth for training a classifier that prospectively predicted cell differentiation from their bright-field images [36]. In these studies, biologically meaningful labels were predicted directly from label-free images whose morphological features were not necessarily easy to recognize. This machine learning-based direct analysis will be expected to allow for biological interpretations of image modalities taken by so called label-free microscopy methods, including not only bright field and dark field images but also phase images, phase contrast images and Raman spectroscopic images.

It is also worth noting that, in this direct analysis approach, cell classification based on the information of features implicitly embedded in raw imaging information can be realized without extracting features explicitly from the images produced. As depicted in Fig. 1c, for example, one can obtain a data set comprising pairs of raw image information and a morphological feature label by carefully preparing and measuring cells with their specific molecules or subcellular structures fluorescently stained. A model trained using this training data can predict a label from new image information, instead of performing analysis based on morphological feature extraction from a produced image. A wide range of fluorescence-based techniques, including the staining of different organelles and the localization of fluorescent proteins, are suited for this purpose, owing to their ability to enhance signals specific to targeted features. Since this workflow does not rely on the extraction or annotation of morphological features that users can recognize in two- or three-dimensional (2D or 3D) images, this analysis can be performed on representations in different and/or compressive coordinates. Importantly, the fact that image inferring becomes free from image reconstruction mitigates the bottleneck associated with fast real-time morphological analysis, as described in the sorting section [13,59,60].

Imaging cell sorter with image reconstruction

The first approach follows the tradition of image production and image analysis, followed by activated sorting action. Based on this approach, to our knowledge, the pioneering developments were performed using an arrayed pixel camera by the Yasuda group in the early 2000s, as depicted in Fig. 2a [14–16]. More recently, based on this approach and using single pixel detectors, high-speed cell sorters were developed, as depicted in Fig. 2b. In 2018, Nitta *et al.* reported the development of an intelligent image-activated cell sorter by integrating a single pixel imager, real-time analysis utilizing six or eight layers of CNNs and a microfluidic cell sorting device. Using this technology, cell isolations based on protein localizations and cell-cell interactions were demonstrated [17]. In 2019, using a different type of a single pixel imaging technique, Gu *et al.* reported another development of an image-guided cell sorting and classification system where cell sorting was realized based on the localization analysis of targeted molecules and particles [18].

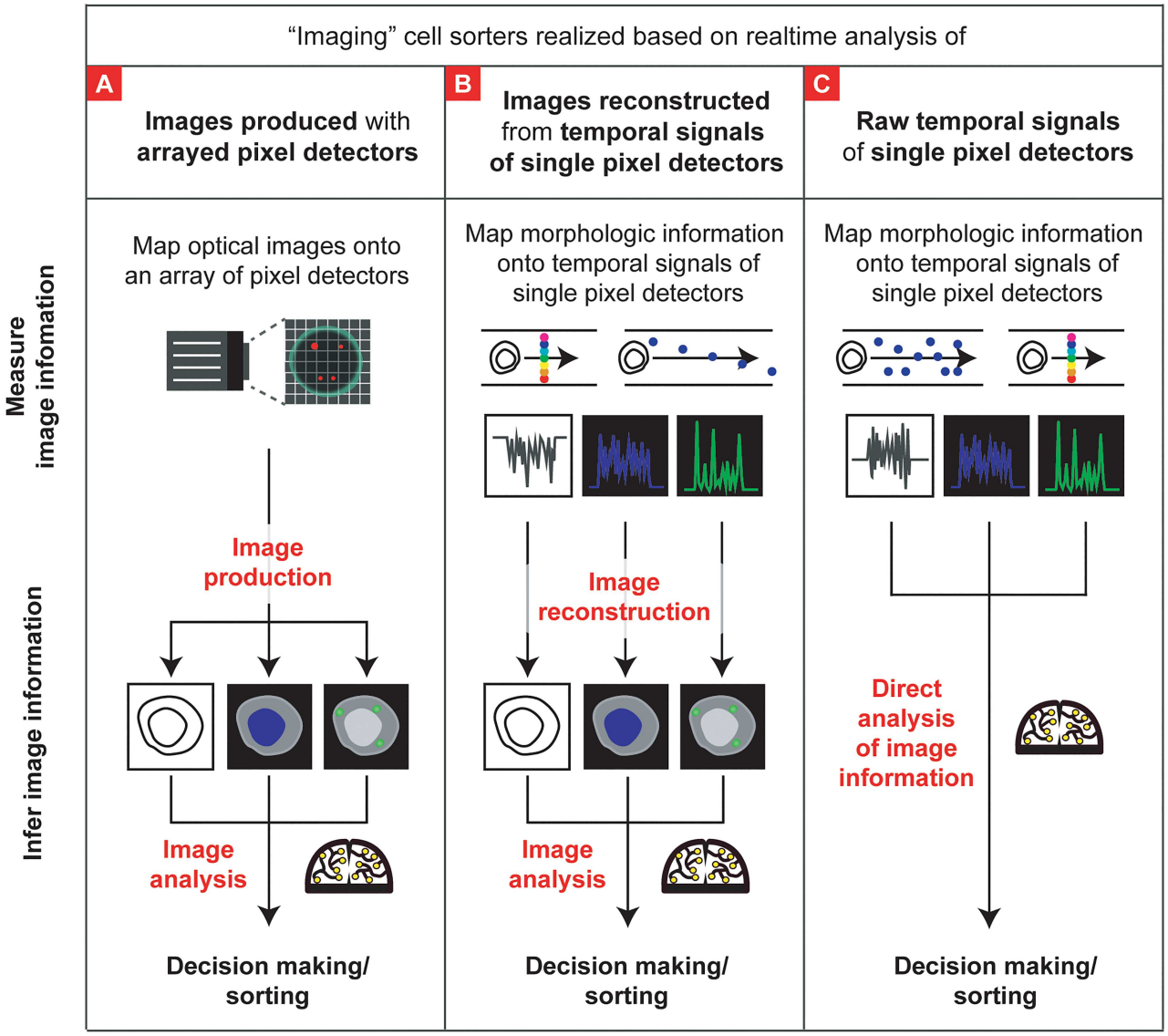

Fig. 2. 'Imaging' cell sorters categorized according to the type of data analyzed in real-time for decision-making on sorting. (a) Cell classification based on the analysis of images captured using an array of pixel detectors. (b) Cell classification based on the analysis of images reconstructed from temporal signals of single pixel detectors. Different imaging modalities can be recorded independently by multiple single pixel detectors in parallel. (c) Cell classification based on the direct analysis of temporal imaging signals of single pixel detectors. Different modalities can be recorded independently by multiple single pixel detectors in parallel.

The advantage of the machine learning-based analysis of reconstructed images is that users can make arbitrary annotations on the images produced. This is especially useful when one can tell the morphological features to be extracted from the images for analysis, as summarized in the section 'Analyzing explicit features extracted from images produced in IFC'. A wide range of software and algorithms developed for application in microscopic cellular images could also be applied for the annotation or feature extraction from the reconstructed images, as well as for their quantification. Moreover, it is convenient to confirm the analysis and sorting results by analyzing the reconstructed images one by one. In addition, image reconstruction may improve classification due to the potentially good fit of existing neural network models in 2D or 3D image analysis. On the other hand, considering the reported calculation time of longer than milliseconds required for inferring the cell classification [17,18], the throughput of the machine learning-based analysis of reconstructed images currently seems still limited. This is mainly due to the heavy calculations required for image reconstruction, which can be mitigated by further developments in the hardware units and algorithms.

Imaging cell sorter without image reconstruction

Another rather newly developed approach takes advantage of machine learning and directly analyzes raw image information without image reconstruction, followed by activated sorting action, as depicted in Fig. 2c. By directly applying a trained model to one-dimensional compressive imaging signals detected by a single pixel detector, we introduced this approach for the first time as image-free 'imaging' cytometry in 2018, which we named 'ghost

cytometry'. Ghost cytometry carries out cell classification based on morphological information [13,59,60] and was integrated into a microfluidic device to demonstrate the high-speed sorting of cells exhibiting various image patterns. The process of analyzing image information in these demonstrations is summarized in the section 'Machine learning-based direct analysis of imaging information in IFC'. In 2019, Li *et al.* further demonstrated the deployment of a deep CNN to directly process the one-dimensional time-series signals of a single pixel imager without image reconstruction for low-latency inference [61].

A significant technological advantage of this approach is the significant reduction in inferring time for cell morphological analysis, thanks to the removal of time-consuming steps, such as image formation and feature extraction. In the case of ghost cytometry [13], the inference time was shorter than 10 μs, an improvement of orders of magnitude compared with the approach performing image reconstruction, such that it will not limit the throughput of the entire IFC system. Meanwhile, this approach may mitigate possible errors associated with the process of image reconstruction. Furthermore, one may reduce the sampling points to further shorten the acquisition time to increase the throughput at the cost of image recovery while maintaining a sufficiently high classification accuracy.

As shown in Fig. 1b, this approach is effective when the morphological features are not easily recognizable in the images. In a recent work [60], after training a model using partial label-free compressive waveforms and biological labels obtained via fluorescence staining, the model was able to predict the biological labels directly from the waveforms without seeing the fluorescence labels. This demonstrates the ability of this technology to perform high-speed label-free classification of a variety kinds and states of cells, including healthiness and differentiation, among others. These findings demonstrate the high potential of this method in applications wherein fluorescence labeling of the final product is undesirable, such as in cell manufacturing and regenerative medicine.

Another study demonstrated cell classification based on implicit feature information embedded in raw imaging information [59]. A model able to recognize morphological features was trained using cell samples with their targeted subcellular features stained fluorescently at high specificity. Here, the cell samples consisted of a population of cells whose cytoplasm was stained and another population of the same cells whose mitochondria were stained in the same color. The trained model was then able to distinguish the two populations directly from the waveform with a high accuracy and without image reconstruction. Currently, the quantitative evaluation of morphological features, such as the nuclear-cytoplasmic ratio and spatial correlations, remains a challenge and will require the training of a regression model in future research.

Future perspectives

The combination of IFC technologies and machine learning methods will allow for a wide range of applications in pharmaceutical and medical industry as well as in biological research. Fluorescence-based methods including specific labeling of subcellular organelles, membranes and biological functions as well as the use of fluorescent proteins are powerful tools to enhance and extract image features of cells with high sensitivity. The cell classification based on such fluorescence imaging biomarkers and the selective sorting holds potential in the cell-based medical diagnosis and phenotypic screening. On the other hand, label-free imaging cytometry and augmented microscopy are good examples that enable inferring cell classification based on the cell morphology data which is not easy to explain manually by humans. These label-free, high-throughput cell analyzers and sorters are finding applications in the fields of cell therapy and regenerative medicine where molecular labeling of the cell products is undesirable. Despite the remarkable advances, several fundamental questions remain in this field. What is the best ground truth from which to evaluate the machine learning-based classification results? How can the content of measurement, the depth of calculation and the throughput of overall IFC system be increased simultaneously? Further research and technological innovations are expected to answer these questions and improve on the current applications of cytometers and microscopes.

Conflict of interest

S.O., I.S. and R.H. are the founders and shareholders of ThinkCyte Inc., a company engaged in the development of the ultrafast imaging cell sorter. S.O. is a board member, and I.S. is a scientific advisor of ThinkCyte Inc. S.O., I.S. and R.H. filed patent applications related to ghost cytometry technologies.

Acknowledgements

S.O., I.S. and R.H. greatly appreciate fruitful discussions with friends and collaborators in academic and industrial communities.

References

1. Moen E *et al.* (2019) Deep learning for cellular image analysis. *Nat. Methods* . https://doi.org/10.1038/s41592-019-0403-1.
2. Basiji D A, Ortyn W E, Liang L, Venkatachalam V, and Morrissey P (2007) Cellular image analysis and imaging by flow cytometry. *Clin. Lab. Med.* 27: 653–670.
3. Filby A (2019) Imaging flow cytometry: A brief overview. *iBiology*. https://www.ibiology.org/techniques/imaging-flow-cytometry/.
4. Kay D B, Cambier J L, and Wheeless L L J (1979) Imaging in flow. *J. Histochem. Cytochem.* 27: 329–334.
5. Kachel V, Benker G, Lichtnau K, Valet G, and Glossner E (1979) Fast imaging in flow: A means of combining flow cytometry and image analysis. *J. Histochem. Cytochem.* 27: 335–341.
6. Blasi T *et al.* (2016) Label-free cell cycle analysis for high-throughput imaging flow cytometry. *Nat. Commun.* 7: 10256.
7. Eulenberg P *et al.* (2017) Reconstructing cell cycle and disease progression using deep learning. *Nat. Commun.* 8: 463.
8. Thaunat O *et al.* (2012) Asymmetric segregation of polarized antigen on B cell division shapes presentation capacity. *Science* 335: 475–479.
9. Doan M *et al.* (2018) Diagnostic potential of imaging flow cytometry. *Trends Biotech.* 36: 649–652.
10. Haridas V, Ranjbar S, Vorobjev I A, Goldfeld A E, and Barteneva N S (2017) Imaging flow cytometry analysis of intracellular pathogens. *Methods* 112: 91–104.
11. Goda K, *et al.* (2012) High-throughput single-microparticle imaging flow analyzer. *Proc. Natl. Acad. Sci. U S A* 109 (29): 11630–11635.
12. Rane AS, Rutkauskaite J, deMello A., and Stavrakis S. (2017) High-throughput multi-parametric imaging flow cytometry. *Chem.* 3 (4): 588–602.
13. Ota S *et al.* (2018) Ghost cytometry. *Science* 360: 1246–1251.
14. Takahashi K, Hattori A, Suzuki I, Ichiki T, and Yasuda K. (2004) Non-destructive on-chip cell sorting system with real-time microscopic image processing. *J. Nanobiotech.* 2 (1): 5.
15. Yasuda K, *et al.* (2013) Non-destructive on-chip imaging flow cell-sorting system for on-chip cellomics. *Microfluid. Nanofluidics* 14 (6): 907–931.

16. Girault M *et al.* (2017) An on-chip imaging droplet-sorting system: A real-time shape recognition method to screen target cells in droplets with single cell resolution. *Sci. Rep.* 7: 40072.
17. Nitta N, *et al.* (2018) Intelligent image-activated cell sorting. *Cell* 175: 266–276.e13.
18. Gu Y, *et al.* (2019) Machine learning based real-time image-guided cell sorting and classification. *Cytometry A,* 95 (5): 499–509.
19. Vapnik V, and Learner A (1963) Pattern recognition using generalized portrait method. *Automat. Rem. Control* 24: 774–780.
20. Boser B E, Guyon I M, and Vapnik V N (1992) A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* 144–152.
21. Schölkopf B, Smola A, and Müller K R (1998) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* 10: 1299–1319.
22. Dhillon I S, Guan Y, and Kulis B (2004) Kernel k-means: Spectral clustering and normalized cuts. *Proceedings of the Tenth ACM SIGKDD* 551–556.
23. Gonen M, and Alpaydın E (2011) Multiple kernel learning algorithms. *J. Mach. Learn. Res.* 12: 2211–2268.
24. Freund Y, and Schapire R E (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55: 119–139.
25. Schapire R E, and Singer Y (1999) Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.* 37: 297–336.
26. Friedman J H (2001) Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29: 1189–1232.
27. Valiant L G (1984) A theory of the learnable. *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing* 436–445.
28. LeCun Y, Bengio Y, and Hinton G (2015) Deep learning. *Nature* 521: 436–444.
29. Nair V, and Hinton G E (2010) Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learn* 807–814.
30. LeCun Y, Haffner P, Bottou L, and Bengio Y (1999) Object recognition with gradient-based learning. *Shape, Contour and Grouping in Computer Vision* 319–345.
31. Krizhevsky A, Sutskever I, and Hinton G E (2012) ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst* 25: 1097–1105.
32. Simonyan K, and Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. *Proceedings of the 3rd International Conference on Learning Representations*.
33. Szegedy C *et al.* (2015) Going deeper with convolutions. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition* 1–9.
34. He K, Zhang X, Ren S, and Sun J (2016) Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition* 770–778.
35. Xing F, Xie Y, Su H, Liu F, and Yang L (2018) Deep learning in microscopy image analysis: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* 29: 4550–4568.
36. Felix B *et al.* (2017) Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* 14: 403–406.
37. Ciurte A, Selicean C, Soriţău O, and Buiga R (2018) Automatic detection of circulating tumor cells in darkfield microscopic images of unstained blood using boosting techniques. *PloS One* 13: e0208385-e0208385.
38. Pärnamaa T and Parts L (2017) Accurate classification of protein subcellular localization from high throughput microscopy images using deep learning. *G3: Genes, Genomes, Genetics* 7: 5, 1385–1392.
39. Rumetshofer E, Hofmarcher M, Röhrl C, Hochreiter S, and Klambauer G (2019) Human-level protein localization with convolutional neural networks. In: *International Conference on Learning Representations*.
40. Ronneberger O, Fischer P, and Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. *Med. Image Comput. Assist. Interv.* 234–241.
41. Falk T *et al.* (2019) U-net: Deep learning for cell counting, detection, and morphometry. *Nat. Methods* 16: 67–70.
42. Valen V *et al.* (2016) Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Comput. Biol.* 11: 12.
43. Kraus O, Ba J, and Frey B (2016) Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics* i52–i59.
44. Kraus O *et al.* (2017) Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.* 13: 924.
45. Mathis A *et al.* (2018) DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21: 1281–1289.
46. Kingma D P, and Welling M (2014) Auto-encoding variational Bayes. In: *International Conference on Learning Representations*.
47. Kingma D P, Mohamed S, Rezende D J, and Welling M (2014) Semi-supervised learning with deep generative models. *Adv. Neural Inf. Process. Syst.* 27: 3581–3589.
48. Goodfellow I *et al.* (2014) Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* 27: 2672–2680.
49. Radford A, Metz L, and Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. In: *International Conference on Learning Representations*.
50. Arjovsky M, Chintala S, and Bottou L (2017) Wasserstein generative adversarial networks. *Proceedings of the 34th International Conference on Machine Learning* 214–223.
51. Dinh L, Sohl-Dickstein J, and Bengio S (2017) Density estimation using real NVP. In: *International Conference on Learning Representations*.
52. Kingma D P, and Dhariwal P (2018) Glow: Generative flow with invertible 1x1 convolutions. *Adv. Neural Inf. Process. Syst.* 31: 10215–10224.
53. Behrmann J, Grathwohl W, Chen R T Q, Duvenaud D, and Jacobsen J H (2019) Invertible residual networks. *Proceedings of the 36th International Conference on Machine Learning* 573–582.
54. Anand N, and Huang P (2018) Generative modeling for protein structures. *Adv. Neural Inf. Process. Syst.* 31: 7494–7505.
55. Goldsborough P, Pawlowski N, Caicedo J C, Singh S, and Carpenter A E (2017) CytoGAN: Generative modeling of cell images. In: *Workshop On Machine Learning In Computational Biology, Neural Information Processing Systems*.
56. Johnson G, Donovan-Maiye R, and Maleckar M (2017) Generative modeling with conditional autoencoders: Building an integrated cell. *arXiv* 1705: 00092.
57. Bailo O, Ham D S, and Shin Y M (2019) Red blood cell image generation for data augmentation using conditional generative adversarial networks. *arXiv* 1901: 06219.
58. Doan M and Carpenter AE (2019). Leveraging machine vision in cell-based diagnostics to do more with less. *Nat. Mater.* 18 (5): 414–418.
59. Adachi H *et al.* (2019) Use of ghost cytometry to differentiate cells with similar gross morphologic characteristics. *arXiv* 1903: 09538.
60. Ugawa *et al.* (2018) Real-time Optofluidic diffractive "imaging" cell analyzer. *The 22nd International Conference on Miniaturized Systems for Chemistry and Life Sciences* MicroTAS 2018.
61. Li Y *et al.* (2019) Deep cytometry: Deep learning with real-time inference in cell sorting and flow cytometry. *Sci. Rep.* 9: 11088.