Biomedical Microdevices (2022) 24: 26
https://doi.org/10.1007/s10544-022-00627-x

Biosensors and machine learning for enhanced detection, stratification, and classification of cells: a review

Hassan Raji[1] · Muhammad Tayyab[1] · Jianye Sui[1] · Seyed Reza Mahmoodi[1] · Mehdi Javanmard[1]

Accepted: 22 June 2022 / Published online: 12 August 2022

Abstract
Biological cells, by definition, are the basic units which contain the fundamental molecules of life of which all living things are composed. Understanding how they function and differentiating cells from one another, therefore, is of paramount importance for disease diagnostics as well as therapeutics. Sensors focusing on the detection and stratification of cells have gained popularity as technological advancements have allowed for the miniaturization of various components inching us closer to Point-of-Care (POC) solutions with each passing day. Furthermore, Machine Learning has allowed for enhancement in the analytical capabilities of these various biosensing modalities, especially the challenging task of classification of cells into various categories using a data-driven approach rather than physics-driven. In this review, we provide an account of how Machine Learning has been applied explicitly to sensors that detect and classify cells. We also provide a comparison of how different sensing modalities and algorithms affect the classifier accuracy and the dataset size required.

Keywords Biosensors · Machine Learning (ML) · Neural Networks · Deep learning · Microfluidics · Support Vector Machine

1 Introduction

A biosensor is typically composed of a biorecognition element and a signal transduction element (Bora 2013). It is used to selectively quantify an analyte or a biomarker. A biomarker is a biological element such as a cell, protein, or DNA which can be a sign of a normal or diseased state (Califf 2018). In a biosensor, once an analyte of interest is detected by a biorecognition element, the presence of the analyte is confirmed by a transducer quantitatively or semi-quantitatively. Then, the generated signal due to the recognition event is converted to an output signal. Biosensors are employed in a broad range of applications including but not limited to disease diagnostics, prognosis, and drug discovery (Thevenot et al. 1999).

Biosensor measurements in microchannels have attracted a lot of attention considering the small volume of the fluid required (Bamshad et al. 2017). Furthermore, sensors that detect cells have garnered special interest in the last few decades with the advent of technologies which automate and miniaturize its different components. In particular, they are important for the diagnosis and detection of various diseases which include but are not limited to: Sickle cell Disease (Lizarralde Iragorri et al. 2018), Acute Myeloid Leukemia (Jackson et al. 2016), and metastatic cancers (Kim et al. 2019) through detection of Circulating Tumor Cells (CTCs). An important stage when using these sensors in a clinical setting is converting the data obtained from these biosensors into useful information by classifying the cells into different categories. For example, Circulating Tumor Cells need to be identified and separated from Red Blood Cells. There are a number of qualities which make a biosensor that detects cells more popular including rapid performance and response (Xue et al. 2020), high specificity (Ugawa et al. 2015), high sensitivity (Hsieh et al. 2014). Also, other beneficial qualities include continuous measurement of analyte without involving experienced personnel (Lee et al. 2011), range (Zhu et al. 2016), response time (King et al. 2007), stability (Mani et al. 2017), low cost (Bardin and Lee 2014), and accuracy (Carminati et al. 2017). Processing of the

✉ Mehdi Javanmard
mehdij@alumni.stanford.edu

Hassan Raji
hassan.raji@rutgers.edu

[1] Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854, USA

generated data from biosensors can be considered as an important stage that effectively influences the improvement of the above-mentioned qualities.

Machine learning, a subset of artificial intelligence, is a framework allowing algorithms to learn automatically from data. Many techniques based on machine learning (ML) have been shown to solve significantly difficult tasks in the real world. They are especially applicable to tasks that require learning a variety of patterns obtained from data. Traditionally, data has been analyzed by people with specific-domain knowledge. However, with the recent advancements in AI and ML, we now have models that can be trained to perform these tasks with sufficient accuracy to significantly reduce or even eliminate the need for human expert intervention. The reason these models can work with such a high degree of precision is that the given problem is defined in a precise mathematical framework. That framework uses large amounts of either labeled or unlabeled data, and then some general probabilistic algorithm is applied to find patterns in the dataset. Evidently, this can have numerous advantages as well as several drawbacks. These advantages include the fact that in many applications since some general model is used, there is no further need for hand-engineered expert knowledge which can be quite expensive or even ambiguous. For medical applications in particular, it has been shown that such methods can not only significantly outperform human experts, but they are also able to discover new knowledge (Najafi et al. 2019). Another advantage is that sometimes these methods discover patterns that could not have been discovered independently and might have seemed irrelevant at first. This overall makes them much more scalable compared to human intervened knowledge discovery. However, these approaches have some drawbacks as well. For example, in many applications, they are very heavy computations that takes several weeks for some models to train. More importantly, they require costly predefined labels in some supervised scenarios. Additionally, some applications are extremely sensitive to the choice of architectures or the hyperparameters are chosen. These drawbacks are being actively improved. As an example, in many classical classification problems, quite simple methods such as logistic regression or Support Vector Machine (SVM) have been shown to perform extremely well. For more complex tasks, more complex neural net-based architectures can be required.

In some biosensors, a large amount of data is generated quickly at the output, and the analysis of this data requires further processing by an experienced user that can lead to errors. Processing by a person can take time to analyze data, which can result in significant delays and add to the sample-to-answer time attributed to a biosensor. On the other hand, ML can identify features and trends, and can also provide understandable output. A quick web search shows that the application of Machine Learning in biosensors have seen an exponential rise in the last decade.

Other review papers have reviewed deep learning applied on microfluidics and image cytometry, but no paper specifically discusses the application of ML on biosensors detecting cells (Gupta et al. 2019; Riordon et al. 2019). In this paper, a review of ML publications on biosensors detecting cells is discussed whilst some pieces of useful information will be provided for biosensor engineers and scientists who want to use ML in their research. Therefore, we present an overview of the main ML concepts to facilitate the reader in understanding fundamental differences between the main ML techniques employed in the subsequent sections. The papers in this review are divided into four main categories based upon the detection mechanism used. These include: Electrical Detection and Optical Detection. Optical Detection further subdivided into Image-based, Flow Cytometry, and Smartphone-based Detection.

2 Overview of main concepts in machine learning

Machine Learning is a very wide field of study which has seen advancement at a very accelerated rate in the past few years. In this section, we aim to provide a brief overview of the fundamental techniques in Machine Learning which are used extensively throughout the review. The introductions presented here are by no means extensive and the reader is encouraged to consult Machine Learning textbooks and literature for a more detailed description of the concepts presented here.

2.1 Unsupervised ML

Unsupervised learning is a branch of Machine Learning in which the data is unlabeled and it is up to the machine to draw meaningful and useful interpretations from the data (Ghahramani 2004). These interpretations can be used to make decisions, finding meaningful connections within the dataset, or even recognizing patterns which can be helpful in another downstream task, such as a classification or regression problem (Ghahramani 2004). These algorithms can take different approaches towards finding such meaningful representations such as by probabilistic density estimation, clustering, and latent variable modeling to name a few (Ghahramani 2004).

2.2 Supervised ML

Supervised Machine Learning is another branch of Machine Learning which encompasses a lot of the ML models currently available (Cunningham et al. 2008). In the supervised

approach, a set of pre-defined labels accompanies the raw dataset. These labels are used to assign variables to the dataset and draw conclusions. Although this approach has yielded significant results for certain applications, labeled data is expensive and time consuming to gather (Xiao et al. 2015). Labeling data presents various challenges and is often perceived as a laborious and boring task in the ML framework although it is critical to the performance of most ML algorithms (Alonso 2015). Some labeled datasets have been made publicly available such as ImageNet (Deng et al. 2010) and have proved to be instrumental in advancing the field of image classification and computer vision. Supervised ML finds applications in a wide range of tasks such as cancerous Circulating Tumor Cells (CTC) detection (Guo et al. 2014), facial recognition (Dixit and Silakari 2015), and weather forecasting (Rodrigues et al. 2018). Biosensors that detect and classify cells have specific labels associated with them and have a well-defined objective such as detecting cancerous cells. Thus, supervised ML models are more widely used than their unsupervised counterparts for biosensors. Another reason for the widespread use of supervised learning in this domain is due to the relative simplicity compared to unsupervised learning techniques (Saravanan and Sujatha 2018). We now proceed to give a brief account of the supervised techniques widely used throughout the review.

2.2.1 Support vector machine

Support vector machines (SVMs) are among the most widely used methods for most supervised classification tasks (Vaidya et al. 2008). SVM can also be used for regression, but since this review is focused mainly on classification of cells, we present SVMs in the context of classification. The SVM algorithm seeks to identify an optimal hyperplane in an n-dimensional space which maximizes the margins between the data points where n is the number of features that distinguishes data points from each other based on their true labels (Akaho 2002). In case of non-linear problems where a linear hyperplane may be insufficient, Kernel SVM is used. This algorithm uses so-called kernel functions to map the data into higher dimensional spaces with the objective that in this higher-dimensional space, the data can be separated easily (Patle and Chouhan 2013). Hyperplanes are decision boundaries that facilitate classification of the data points, and their size depends on the number of features in the dataset. A major advantage of the SVM is that it not only finds hyperplanes but maximizes margins among datapoints, which gives better generalization error (Wang and Miao 2012). The hyper-plane in 2D space, for instance, would be a line divides the space into two sub-regions, each of which represents a different class. In this respect, hyperplane position and orientation can be affected by support vectors that are datapoints closer to the hyperplane.

2.2.2 Artificial neural networks

One branch in machine learning which has recently gotten significant attention is called Artificial Neural Network (ANN). These methods are loosely inspired by the inner functioning of the human brain and aim to mathematically model problems in a way that mimic the human brain. The basic component of a neural network is the "neuron" (Maass 1997). These neurons are interconnected to make a structure which can perform classification tasks with varying degrees of complexity and nonlinearity. These nonlinear dynamics allow them to extract much more complex and useful features from raw data, thus leading to more useful representations and significantly better performance on variety of tasks such as facial recognition (Baron 1981) and classification of cells in blood (Tabrizi et al. 2010). The process of learning within such ANN models is in fact the finding of optimal parameters for synaptic weights of the neurons in order to gain a reasonable accuracy (Livni et al. 2014). Also, it is necessary to mention that in most ANN architectures, more than one layer of neural operations are cascaded to make them solve more complex tasks, thus giving them the name "Deep Learning models" (Lecun et al. 2015).

Convolutional neural networks A specific form of ANNs which is widely used in the field of medical image analysis (Li et al. 2014) and therefore deserves special mention in this review is called Convolutional Neural Network (CNN). These architectures are specifically designed for image-based tasks such as image/video classification (Li et al. 2014) and object detection (Zhiqiang and Jun 2017), although they have been applied to other problems as well. CNNs are a kind of Feed Forward Neural Network that considers spatial dependency by using convolutional kernels. More specifically, the learnable weights of the network are the parameters of a set of convolutional kernels which are convolved with the input images or the outputs of the preceding layer. Such architectures were initially designed for problems focusing on images, mainly because they take advantage of the effect of spatial-invariance in the images as well as the importance of locally-neighboring features. A result, they convolve the same shared parameters across the whole image.

3 Machine learning in different biosensing techniques

3.1 Optical detection

Optical detection of cells implies the use of optical techniques and instruments for the detection, classification, and stratification of cells. These can be divided into 3

main categories depending on the sensing modality used, mainly image-based detection, optical flow cytometry, and smartphone-based detection. Each of these sensing modalities have their unique characteristics and come with their own set of challenges. Image-based detection applies ML techniques to images to extract features and get relevant information. Smartphone-based detection, for example may be considered an offshoot of image-based detection since it also applies the ML techniques on images. However, these differ from their image-based techniques mainly due to the fact that data collection is carried out using a smartphone camera. Although offering several advantages such as being easily portable and simple, the images acquired using smartphones may have lower resolution than those acquired using expensive optical microscopes and specialized cameras and present a challenge for detection of cells when compared to traditional high resolution imaging system. In optical flow cytometry, cells flow through a microfluidic channel while being illuminated by a light source (typically a laser). The scattering of light in different directions (e.g. forward scatter and back scatter) is measured for each cell passing through the light source using optical filters. Such systems have high throughput and need faster algorithms for real-time detection and classification. Recently, some researchers have augmented data collection using this technique with images to enhance the detection and classification accuracy (Li et al. 2019; Chen et al. 2016). In the subsequent sections, we present a review of the different ML techniques applied to each of these categories and summarize our findings in Table 1.

3.1.1 Image-based detection

Image-based detection implies the use of images or videos of cells. These images or videos need to be processed to identify and quantify cells. ML has tremendous power in the analysis of image data by making accurate predictions on large sample datasets. ML algorithms eliminate much of the manual steps required to process data, thereby reducing the processing time, and eliminating human error. In the next following paragraphs, we present sensing approaches based on ML algorithms on data obtained using this detection method.

Neural networks are often utilized in image analysis making it a powerful tool for classifying cells imaged using biosensors utilizing image-based detection. To demonstrate, Koohababni et al. utilized Mixture Density Networks (MDNs) using a Gaussian Mixture Model (GMM) to identify cell nuclei (Koohababni et al. 2018). MDNs overcome the limitations of the conventional neural networks utilizing the least squares approach and are more suitable candidates for mapping single inputs to multi outputs. As such, MDNs were used by the authors for the detection of several seeds in a single image patch. The researchers compared their proposed method with the existing well established NN frameworks and demonstrated a higher F1 score for identifying cell nuclei in colorectal histology images. Classifying cells using low resolution images remains a challenge in the world of biosensing. Huang et al. (2016a) employed a CNN-based super resolution (SR) in order to extract high resolution (HR) images from low resolution (LR) images. This has been previously demonstrated using an Extreme Learning Machine (ELM) but with limited accuracy for large datasets. The authors illustrated that they were able to obtain HR images of red blood cells (RBCs), white blood cells (WBCs) and platelets from their lensless system. The application of ML is not merely limited to 2D images. Mayerich et al. (2011) have presented a method for cell soma detection based on volumetric data. The researchers obtained these volumetric images through their indigenously developed technique which they term as Knife-Edge Scanning Microscopy (KESM). This technique generates huge amount of data in a short span of time. This group has illustrated the use of commercially available GPUs in conjunction with a multi-layer feed forward NN to accurately locate neuron positions in rat brain tissue from a large dataset amounting to 200 Gigabytes (GB). A comparison highlighting the superior performance of their algorithm over standard feature detection algorithms was also presented. In another interesting application of ML applied to biosensors, authors used NNs to simulate the movement and behaviour of red blood cells in blood plasma (Bachratý et al. 2020). This was mainly used to optimize the microfluidic channel geometry. In this study, the NN was taking a numerical simulation as an input. Alternatively, the input could also be a video recorded from an actual biological experiment. Their results indicated that for uncomplicated box channels, there was no advantage of using this method instead of fluid streamlines. However, in a more complicated geometry, the NN performance showed a significant improvement. Cell gating has been traditionally used to stratify various types of blood cells. ML provides an efficient and favourable alternative to this problem. Researchers used principal component analysis (PCA) in conjunction with NNs to recognize five types of WBCs in peripheral blood (Tabrizi et al. 2010). For this purpose, nucleus and cytoplasm were segmented using the Gram-Schmidt method and snake algorithm. Moreover, three kinds of features (morphological, textural, and color) were extracted from the segmented areas. Next, the best features were selected using Principal Component Analysis (PCA). Finally, five types of white blood cells were classified using Learning Vector Quantization neural network (LVQNN). When applying Machine Learning to a specific application, there exists a wide variety of software packages and interfaces to choose from. Each of these interfaces and packages has a learning curve. Falk et al. (2019) proposed plugin in a software package for cell detection and cell segmentation

Table 1 Comparison of ML efficacy of different Biosensors for Cellular Analysis

Author & Year of Publication	Training Algorithm	Detection Technique	Sample Type	Cell type	Approximate Classifier Accuracy (Maximum)*	Dataset Size	Ratio
Dataset size = number of cells							
Toedling et al. (2006)	Multivariate Classification, SVM	Optical Flow Cytometry	Yeast Culture	Yeast Cells	82%	120 Cells	6.83E-01
Tabrizi et al. (2010)	Neural Networks	Image-based	Biological Tissue	Neuron Cell Soma	92.80%	2158 Cells	4.30E-02
Chen et al. (2011)	Neural Networks	Electrical	Cell co-culture	RBC and HepG2 Tumor Cells	99%	11909 Cells	8.31E-03
Yu et al. (2011)	GED, SVM and KNN	Image-based	Blood	Red Blood Cells	84.80%	166326 Cells	5.07E-04
Mayerich et al. (2011)	Neural Networks	Image-based	Stem Cell Solution	Mouse Embryonic Carcinoma Cells (P 19)	95%	98 Signals (Cells and Beads)	9.69E-01
Valen et al. (2016)	Deep learning	Image-based	Cell Culture	CRL-5803 cells and CCL-185 Cells	74.40%	976 Cells	7.62E-02
Zheng et al. (2012)	Neural Networks	Electrical	Cell Suspension	ML-2 and HL-60 Cells	93%	6647 Cells	1.40E-02
Song et al. (2013)	SVM	Electrical	PBS	HepG2 and RBC Cells	92.00%	3698 Cells	2.49E-02
Zhao et al. (2013)	Neural Networks	Electrical	Blood and Bone Marrow Smears	Acute lymphoblastic leukemia Cells	97%	958 Cells	1.01E-01
Zheng et al. (2013)	Neural Networks	Electrical	Blood	Monocytes, Granulocytes, and Lymphocytes	89%	7500 Cells	1.19E-02
Huang et al. (2014)	Neural Networks (ELM-SR)	Optical Flow Cytometry	Blood	Leukemia Cells	99.46%	10000 Cells	9.94E-03
Amin et al. (2015)	SVM	Image-based	Cell Culture	Fluo-N2DL-HeLa, PhC-HeLa and Hist-BM Cells	96.90%	34060 Cells	2.84E-03
Schneider et al. (2015)	Neural Networks	Optical Flow Cytometry	Lysis of RBCs	MDA-MB-231 and MCF7 Cells	False Positive Rate of at Most 0.001%	100000 Cell Training Sets	N/A
Ni et al. (2016)	SVM	Optical Flow Cytometry	Cell lines	Leukemia cell lines(K562, MOLT, and HL60)	97.60%	618 Cells	1.58E-01
Akram et al. (2016)	Deep Learning	Image-based	Cell lines	leukemia cell lines HL60, MOLT, and K562	99.52%	618 Leukemia cells	1.61E-01
Singh et al. (2017)	Decision Tree	Image-based	Buffer	Lung cancer cell lines ofH1299 and A549	90.90%	100,000 Cells	9.09E-04
Gopakumar et al. (2017)	Deep Learning	Optical flow cytometry	Online Dataset	Colorectal Adenocarcinoma Cells	Precision = 0.788	29756 Nuclei	N/A
Kalmady et al. (2017)	Neural Networks	Optical Flow Cytometry	Blood	Erythrocytes and Platelets	Minimum Error = 0.003	> 40 k Cells	N/A

Table 1 (continued)

Author & Year of Publication	Training Algorithm	Detection Technique	Sample Type	Cell type	Approximate Classifier Accuracy (Maximum)*	Dataset Size	Ratio
Dataset size = number of cells							
Zhao et al. (2018)	Neural Networks	Electrical	Buffer	Normal Human Fibroblasts (NHFs) and Senescent Human Fibroblasts (SHFs)	88%	480 Cells (240 Normal Human Fibroblasts (NHFs) and 240 Senescent Human Fibroblasts)	1.83E-01
Koohababni et al. (2018)	Deep Learning	Image-based	Buffer	T47D Cancer Cells	95.90%	> 1000 Cells	9.59E-02
Lin et al. (2018)	SVM	Optical Flow Cytometry	Blood	RBCs and Yeast Cells	RMSE = 1.2 um for particle size	17000 Single Particle Signals (Beads and Cells)	N/A
Ahuja et al. (2019)	SVM	Electrical	Yeast Culture	Yeast Cells	82%	120 Cells	6.83E-01
Fu et al. (2019)	SVM	Optical Flow Cytometry	Biological Tissue	Neuron Cell Soma	92.80%	2158 Cells	4.30E-02
Honrado et al. (2020)	Neural Networks	Electrical	Cell co-culture	RBC and HepG2 Tumor Cells	99%	11909 Cells	8.31E-03
Dataset size = Number of Images							
Long et al. (2006)	SVM	Image-based	Cell Culture	B-cell Lymphoma Cells	94%	59 Images	1.59
Su et al. (2014)	Neural Networks	Image-based	Blood	White Blood Cells	99.11%	450 Images	2.20E-01
Mao et al. (2015)	SVM, CNN	Image-based	Blood	CTCs (Breast Cancer)	91.20%	45 Images	2.03
Heo et al. (2017)	CNN	Optical Flow Cytometry	Microparticles in Buffer	K562 and RBCs	93.30%	6000 Images	1.55E-02
Xu et al. (2017)	CNN	Image-based	Blood	Sickle Cells (RBCs)	89.28%	> 7000 single RBC images	1.28E-02
Go et al. (2018)	Decision Tree, SVM, kNN, Linear Discriminant Classification	Image-based	Blood	Erythrocytes: Discocytes, Echinocytes, Spherocytes	97.37%	630 Images	1.55E-01
Turan et al. (2018)	SVM	Image-based	Blood	T and B Cells	94%	420 Scan Images	2.24E-01
Soldati et al. (2018)	CNN	Optical Flow Cytometry	Blood	CTCs and CD45 Cells	90.20%	500 Images	1.80E-01
Xia et al. (2019)	Deep Learning	Image-based	Buffer	White Blood Cells	98.40%	364 Images	2.70E-01
Sun et al. (2019)	Deep Learning	Optical flow cytometry	Buffer	T Cells and B Cells (Acute Luekemia)	93.20%	2400 Images	3.88E-02
Uslu et al. (2019)	SVM	Image-based	Buffer	B Lymphoblast	87.40%	100000 Sub-images	0.000874

Table 1 (continued)

Dataset size = Other							
Park et al. (2008)	SVM, ELS-ELM, RLS-ELM, ELM	Electrical	Blood	Red Blood Cells	RMSE = 0.74	199 Blood Samples	N/A
Ko et al. (2018)	SVM	Optical Flow Cytometry	Blood	Acute Myeloid Leukemia (AML) and Myelodysplastic Syndrome (MDS) Cells	92.40%	5333 MFC Data Points	N/A
Li et al. (2019)	Deep learning	Optical Flow Cytometry	Buffer	OT-II White Blood Cells and SW-480 Epithelial Cancer Cells	95.74%	6700 Data Points	N/A
Haan et al. (2020)	Deep Learning	Smartphone-based	Whole blood	Sickle Cells and Healthy RBCs	98%	96 Patients' Blood Smear	N/A
Zhang et al. (2019)	Deep Learning	Image-based	Whole Blood	MCF-7 Cancer Cells	78%	17,447 Videos	N/A

based on deep learning allowing users to employ this plugin without having knowledge of ML. Unlike previous similar software packages, their plugin, U-Net, had the capability to be trained and adapted to new sets of data and tasks by ImageJ® software interface.

Live cell imaging is a valuable tool for studying living organisms. However, cell segmentation is a considerable challenge in live cell imaging since there are segmentation methods that require several hours curation that should be performed manually and these methods are dependent on approaches which are difficult to share between different labs. Van Valen et al. (2016) developed a framework utilizing deep learning to successfully overcome the image segmentation problem. The authors demonstrated that deep CNN is able to successfully segment and classify different mammalian cells. The researchers have reimagined the problem of cell segmentation as cell classification using a deep learning framework. Similarly, Akram et al. (2016) presented a CNN-based method providing cell segmentation proposals. These proposals initially represented bounding boxes utilizing a fully CNN (FCN) and then predicted segmentation masks for bounding boxes using another CNN. They compared their proposed techniques with other conventional cell detection and segmentation methods and concluded that their method has a better performance in terms of common evaluation parameters. In another study, Xia et al. (2019) developed a deep learning-based object detection method, Faster Region-based CNN which is a modified version of Region-based CNN. In this method, a Region Proposal Network (RPN) was used along with a transfer learning process to detect WBCs in microscopic images. By conducting analysis on 364 images, 50 images for training and 314 images for testing, they reported a miss rate of 1.3% and a detection accuracy of 98.4%. Likewise, Faster Region-based CNN was applied for cell detection by segmentation and classification to detect cells (Yang et al. 2017). Their experiments showed that cells can be detected in microscopic images using Faster R-CNN. Furthermore, this technique improved cell detection performance, saved time, and was easily implemented. Detecting rare cells in blood is of particular interest in diagnosing disease particularly cancer such as the case of CTCs. Zhang et al. (2019) have recently demonstrated a novel cell detection and cytometry technique by incorporating magnetically modulated lensless imaging. A deep learning-based classifier was employed to enhance the specificity of their cytometer which also allowed to detect MCF-7 circulating tumor cells based on their spatio-temporal features under a controlled magnetic force. This technique enabled authors to detect 10 cells per milliliter of whole blood. Spatio-temporal features of bacteria can be utilized for classification of bacteria and characterizing bacterial growth similar to how these were used for cell classification. Wang et al. (2020) presented a live microscopy detection system for detecting

3 different kinds of bacteria. Their imaging platform comprises of a lensless image sensor which scans an agar plate every 30 min to acquire holographic images. A differential image analysis is then applied to detect objects. These objects include surface impurities in addition to the bacteria. A Deep Neural Network (DNN) is used to separate the bacteria from these impurities based on the bacterial growth; the size of the bacterial colonies will grow over time whereas the size of impurities will remain relatively constant. A second DNN is then utilized to classify the bacterial colonies into their subtypes. For this study, they used *Escherichia coli, Klebsiella aerogenes and Klebsiella pneumoniae subsp. Pneumoniae* respectively. Their platform decreases the time required for detection by > 12 h and eliminates the need for expertise required for identification of the bacteria due to use of ML techniques.

Applications of ML in cell segmentation have been extended to cell shape and morphology. Cell shape and structure says a lot about its health and is an important parameter in diagnosing various conditions. An example of this is Sickle Cell Anemia where the RBC shape deteriorates usually due to the interaction of abnormal hemoglobin with the RBCs. Researchers have developed a high-throughput and automated RBC shape classification framework utilizing CNNs on patient-specific microscopy images for aiding in diagnosis of sickle cell anemia (Xu et al. 2017). Their high-throughput classification assay consists of four main steps: 1) Hierarchical RBC patch extraction, 2) Size-invariant RBC patch normalization, 3) RBC pattern classification based on deep CNN, and 4) Automated RBC shape factor calculation. Their work differs from traditional methods used such that they have removed most of the labor-intensive tasks associated with classification which usually require specific domain knowledge. For example, instead of scanning the whole image using same-size patch method (Han et al. 2016) or manually selecting the image patches with RBCs, they automate this process through hierarchical RBC patch extraction. In this method, an entropy function is calculated to differentiate the cells from the background and to identify a Region of Interest (ROI). A patch normalization technique is then used to eliminate the variations in the data and this is fed into a deep CNN to classify the cells into as much as 8 different categories including Sickle cells. Similarly, the use of ML algorithms for differentiating cells on the basis of shape has been demonstrated for classifying T-cells and B-cells in a pillar-based microfluidic cell counting system by applying a SVM classifier (Turan et al. 2018). In object detection, a descriptor is a simplified representation of the image that contains only the most important features of the image. In this work, authors used a commonly used descriptor which focuses on the structure or the shape of an object namely as histogram of oriented gradients (HOG) along with color features to differentiate B-cells from T-cells (Fig. 1). First, a linear-kernel SVM was trained to detect cells from a background in dual dyed images. Subsequently, the cells in a single dye image were identified by the first SVM based on HOG features found in the image using a sliding window method. At last, a Radial Basis Function (RBF)-kernel SVM was trained with the color information of found cells to differentiate T-cells from B-cells.

SVMs are an important ML technique that are more robust against the otherwise notorious problem of overfitting associated with ANNs and deep learning. It's use for classifying cells has been explored extensively. Long et al. (2006) demonstrated that they can differentiate between unstained viable and non-viable cells using SVMs combined with a novel image selection technique. The problem associated with differentiating viable and non-viable cells is that the dataset is extremely unbalanced i.e. the non-viable cells detected greatly outnumber the viable cells. This problem was overcome by a method they named "Compensatory Iterative Sample Selection". The way this method works is by iteratively selecting images that are most representative of the data in the non-viable class and re-training the SVM until a performance plateau is reached, thus minimizing the misclassification cost and increasing classifier accuracy. Similarly, Uslu et al. (2019) developed an automated SVM based method of quantifying leukemia cells captured and separated using immunomagnetic beads. The immunomagnetic beads were coated with antibodies which then attached to the cells. The beads that contained cells fell into 2 categories: 1) They were either single isolated cells in which case the cell boundary was partially obstructed by the bead or 2) The beads formed cell-bead clusters, which could be distinguished easily. The authors considered both of these possibilities and applied it to an SVM based classifier. The classifier had 2 classes as the output i.e. cell and non-cell. There was an imbalance between the data as the non-cells outnumbered the cells and therefore the non-cell data was downsampled to eliminate this problem. The researchers used a radial basis function as the kernel for their proposed SVM and demonstrated successful stratification reporting a classification accuracy of 87.4%.

Geometric and statistical features can be used to separate cancerous cells from non-cancerous cells as demonstrated by Amin et al. (2015). Geometric features include area, perimeter, solidity, eccentricity, the extent of the nucleus from the binary image of the nucleus, whereas statistical features are features extracted from the gray-scale intensity values of the pixels such as mean, standard deviation, energy, entropy, skewness, kurtosis. This specific study used Acute lymphoblastic leukemia cells for demonstration. K-means algorithm was employed to segment cell nuclei after pre-processing of the images. The means of SVM classifier were used the two types of cells. These cells were further classified into their

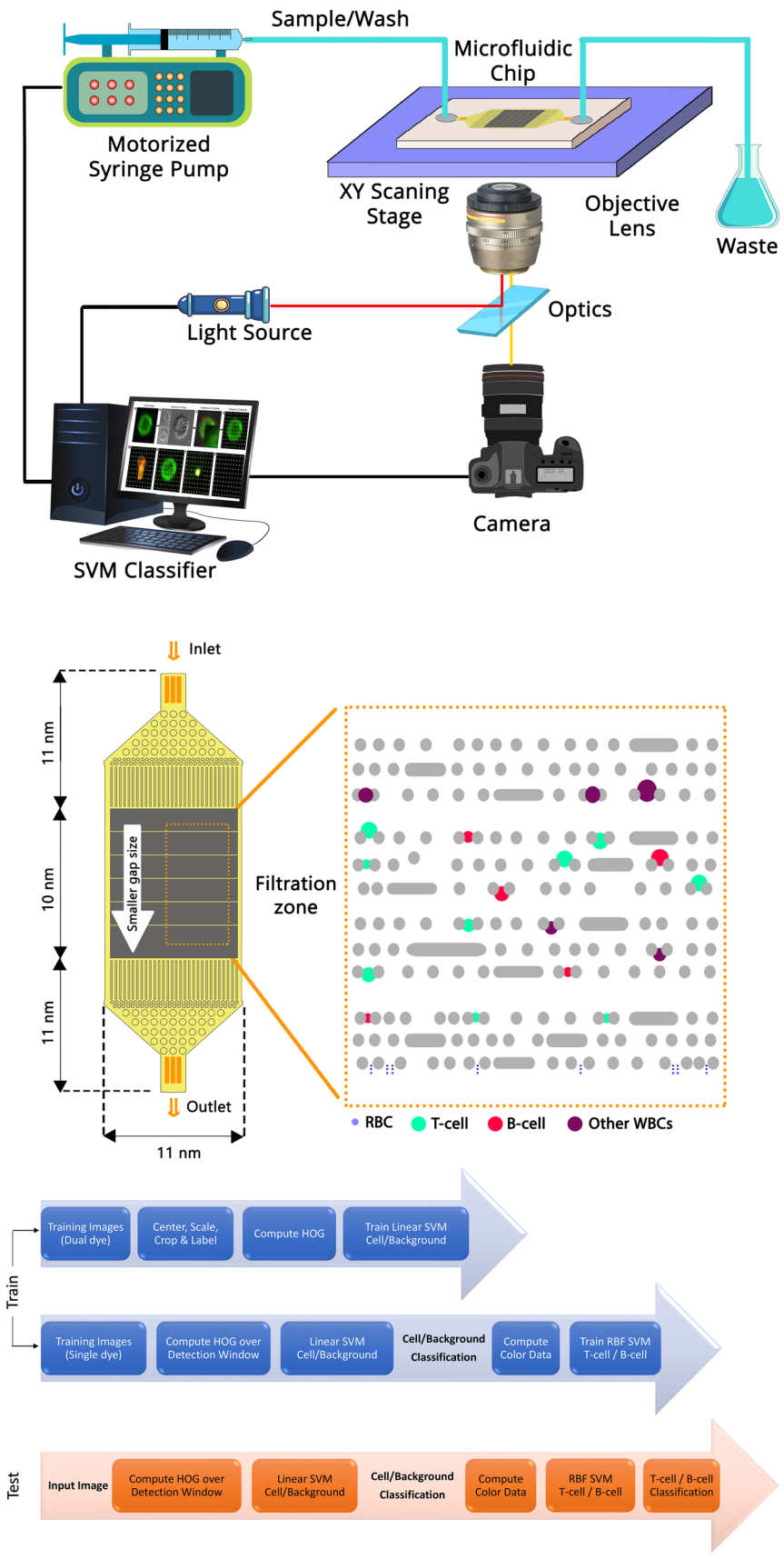

Fig. 1 Outline of the Image-based system proposed by Turan et al. (2018). Whole blood was injected into the device through inlet while leukocytes were trapped in different zones based on the deformability and size difference. The proposed experiment setup and block diagram of cell detection using ML is also shown in this figure. In the block diagram of the cell detection framework, it can be inferred that using Support Vector Machine (SVM), training images are centered, cropped, and labeled. The Histogram of Gradients (HOG) and color data was computed using the processed images for the classification of cells. Adapted from Turan et al. (2018)

different morphological subtypes by a multi-SVM classifier. The accuracies of both these classifiers were above 95%. Microalgal cells show promise as a biofuel and as an effective solution towards global warming as they perform photosynthesis and can be grown in relatively harsh conditions such as wastewater. These cells produce lipids when subjected to stress usually through nutrient deficiency, particularly nitrogen. Therefore, it is important to characterize them into nitrogen deficient and nitrogen sufficient compounds. Guo et al. (2017) proposed a high-throughput label-free single-cell method for screening lipid-producing microalgal cells by optofluidic time-stretch quantitative phase microscopy. Optofluidic time-stretch Quantitative Phase Microscopy provides a method for detecting these cells at a high rate of 10,000 cells/s (Guo et al. 2018). 188 features extracted from the images were used in the classification of nitrogen-sufficient and nitrogen-deficient E. gracilis cells. A SVM that was trained using a sequential minimal optimization algorithm was applied to analyze the intensity and phase images acquired by the optofluidic time-stretch quantitative phase microscopy. It achieved an 2.15% error rate in cell classification.

Observing cell cycle progression is an important field of study with numerous applications. One such example is the identification of cancerous cells (Srivastava et al. 2008). Yeast cells are frequently used as a model for studying cell cycle regulation. Yu et al. (2011) employed image processing algorithms to classify yeast cells in a microfluidic channel. The authors used a simple threshold algorithm based on the Mahalanobis distance (Wang et al. 2007) for image enhancement to reduce background noise, before carrying out image segmentation. The cells were differentiated into 3 classes pertaining to different stages of cell cycle development. The main difference between the cells in their cell cycles is that when they are in the Synthesis (S) stage, they have a daughter cell or a "bud" with them and have a characteristic peanut-like shape as opposed to a circular shape. The researchers extracted a combination of 3 features based on the shape and size of the cells in relation to the size of the buds which they used for training 3 different classifiers. They compared the performance of these 3 classifiers namely linear support vector machine (LSVM), distance-based classification (GED), and k-nearest-neighbor (KNN). The performance of all 3 classifiers did not vary considerably. The main drawback of the algorithm the researchers used is that their method was only applicable to single isolated cells and did not take into consideration if two cells were touching or overlapping one another and was only applicable to microfluidic devices. This is not an inherent drawback of any of the ML algorithms used but it is due to their initial assumptions of classifying cells of "peanut-like" shape from the circular shaped cells and the pre-processing image enhancement algorithms used.

A novel segmentation algorithm for the classification of five types of white blood cells by Su et al. (2014). Their segmentation algorithm was based on finding a discriminating region of white blood cell tones in the HSI color space. In their study, three different NN-based classifiers of Multilayer perceptrons (MLP), SVM and hyperrectangular composite neural networks (HRCNN) were adopted for classifying white blood cells. It was shown that the proposed system incorporated with a trained MLP can reach the highest performance. Morphological properties of erythrocytes can be an indicative of various hematological diseases (Tomaiuolo 2014). In a label-free approach, Go et al. (2018) used digital in-line holographic microscopy (DIHM) paired with ML models to identify and classify different types of erythrocytes: discocytes, echinocytes, and spherocytes. Four different models were used to determine the best algorithm: SVM, Decision Trees, Linear Discrimination Classification (LDC), and k-nearest neighbor (KNN) classification. The decision trees exhibited the best identification performance for the training sets (n = 440, 98.18%) and test sets (n = 140, 97.37%). The detection of CTCs in blood is a challenge since the CTCs are generally outnumbered by Peripheral Blood Mononuclear Cells (PBMCs). Therefore, any ML algorithm to be used for reliable detection of CTCs needs to have a very low False Positive Rate (FPR). Singh et al. employed ML-based gating criteria to differentiate MCF-7 and MDA-MB-231 cell lines from PBMCs when flowing through a microchannel (Fig. 2; Singh et al. 2017).

Initially, they used binary discriminants to differentiate these cells from the PBMCs. The authors used 3 features for differentiation: cell diameter, maximum pixel intensity, and mean pixel intensity. They used a Classification and Regression Tree (CART) algorithm (Dension 1998) for optimizing their classifiers. The authors claimed an FPR of only 0.001% and were able to identify CTCs at a concentration as low as 10 CTCs per ml. They noted that all the 3 features differed significantly between the classes. The authors also observed that SVM and Linear Discriminant classifiers provided similar improved accuracy over the binary discriminants when incorporating all 3 features. In a similar study for differentiating CTCs from blood cells, Mao et al. (2015) investigated the use of 2 classifiers on a microscopic image-based CTC detection platform. The 2 classifiers used were SVMs and CNNs. The SVM used Histogram of Gradients (HoG) as features and the CNN used 6 convolutional filters on the image patches. The authors used the MCF-7 cell lines for validating their algorithms and reported that both of these methods had a maximum F-score of 91.2%.

Manual counting of cells is time consuming, low throughput, and laboriously extensive. On the other hand, commercial flow cytometers have their own limitations such as being bulky, expensive, and require specific domain knowledge. An alternative is to use lensless microfluidic image

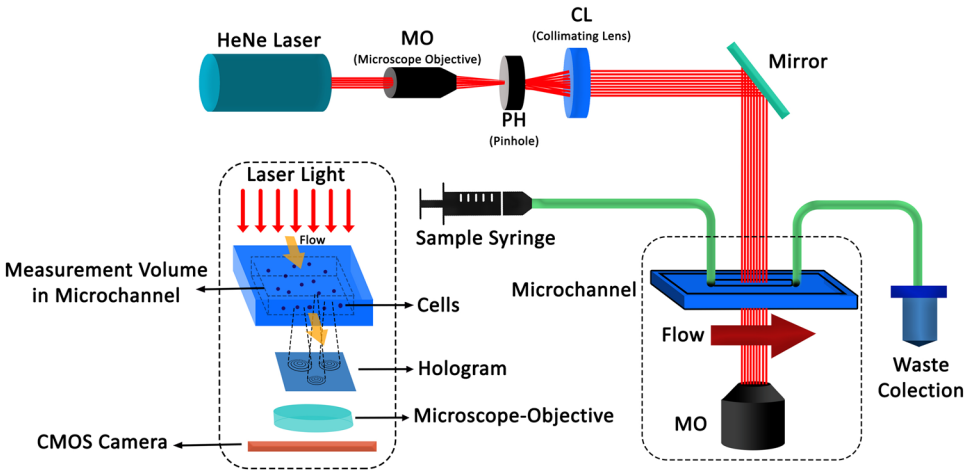

Fig. 2 Inline digital holography microscopy (DHM) utilized by Singh et al. (2017) for characterizing cells in flow. As shown, in this figure, experimental arrangement of inline-DHM is shown which enabled recording holograms of cells in bulk flow along with multiple experimental parameters. The output data was used in a classifier enabling detection of tumor cells. Adapted from Singh et al. (2017)

detection. This method produces low resolution images that present an obstacle for high-throughput analysis of samples. This is where ML assisted counting comes in. As an example 2 approaches which utilize ML, namely Extreme Machine Learning Super Resolution (ELMSR) and Convolutional Neural Network Super Resolution (CNNSR), were explored by researchers to solve the low-resolution problem in a lensless microfluidic imaging using CMOS image sensors for blood cell counting (Huang et al. 2016b). Low-resolution lensless cell images were the input and an improved high-resolution cell image was the output. At the end, cell resolution was improved 400% while the cell counting results were in line with commercial flow cytometers.

3.1.2 Optical flow cytometry

A microfluidic flow cytometer is an integrated system which consists of microchannels for flow and optical sensors for detection. Typically, the cell is detected using scattered light from laser beams illuminating the cells flowing through the detection chamber in a microchannel. Ideally, the biosensor would be portable, easy to operate, and suitable for use as a point-of-care diagnostic device. In this section, we analyze the works by researchers who have demonstrated the use of Machine Learning techniques on data gathered explicitly by Optical Flow Cytometry.

Various research groups have applied deep learning with their microfluidic flow cytometers to analyze the single-cell images for cell classification. Heo et al. (2017) developed a custom algorithm for simultaneously tracking and classifying cells in real-time by using multiple thread processing i.e. less than 2 ms. They used a CNN with fully-connected layer for supervised classification to differentiate between microparticles of different sizes, RBCs, and K562 cells. Currently, their algorithm classifies cells based on size. Although the authors did not demonstrate their algorithm's efficiency for classifying cells based on morphology, they claim that their algorithm will be useful for classifying cells with varying features and classes by making small modifications to the CNN such as using a fancy network architecture. The authors reported a mAP of 93.3%. Ben et al. (2016) proposed a method for detecting CTCs among WBCs by detecting change in pH. The detection method exploited the anomalous extra cellular acidification rate (ECAR) of CTCs. For this purpose, the CTCs and WBCs were enclosed in picolitre sized droplets so that each droplet contained at most a single cell. A pH dependent dye was used, and images were also acquired for each droplet that was detected via an electronically controlled trigger. These images then had to be inspected individually via a manual process to separate the CTCs from the empty droplets and debris. Recently, the authors demonstrated the utility of neural networks for solving this problem (Soldati et al. 2018). They tested a variety of neural network architectures, and performed various image augmentation techniques such as flipping the images across horizontal and vertical axes, modifying contrast, blur, and rotation. After testing various architectures, they concluded that the best architecture for their application was a combination of MobileNet and Inception-v2, which yielded an overall accuracy of 96%. Optical flow cytometry usually involves the use of dyes to identify micro particles of interest. Time stretch Quantitative Phase Imaging (TS-QPI) provides an alternative to these methods providing high sensitivity and producing high amounts of data in a relatively short period of time (Goda et al. 2009, 2012). Conventional methods rely on converting these data into images and performing deep learning on the images. This has two disadvantages: longer time required for conversion of images and possible loss of information which may occur during conversion. Recently, Li et al. (2019) presented a deep learning architecture which could classify cells using these raw waveforms. The authors used waveform files collected by the ultrafast ADC directly without conversion into images as input to their convolutional network (Fig. 3). Each

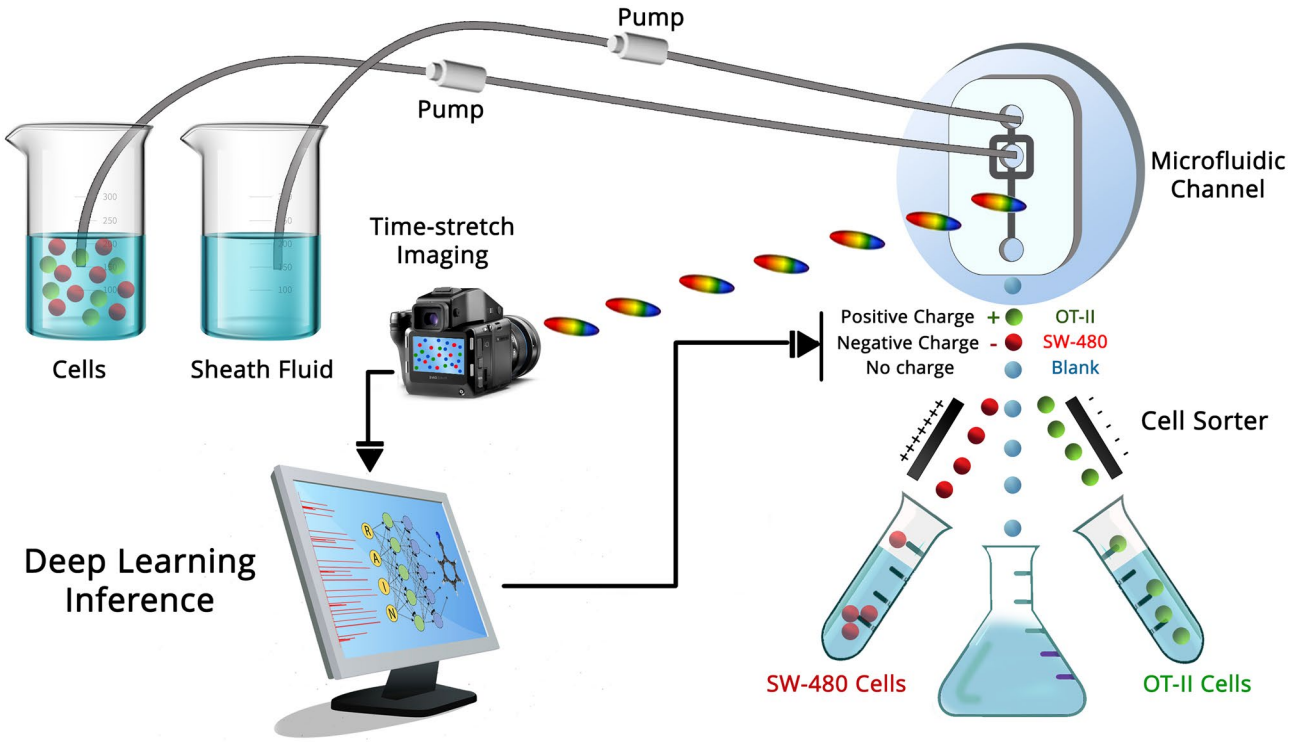

Fig. 3 Overview of the application of deep learning in flow cytometry presented by Li et al. (2019). In this research, hydrodynamic focusing mechanism was employed in a microfluidic channel to align the cells in the centerline of the main channel. Waveform pulses generated in the channel were captured by time-stretch imaging system. Consequently, without further signal processing, these waveforms of time-stretch imaging were outputted to a deep NN where cell classification was carried out rapidly with high accuracy. The 2 types of cells were then categorized before being separated into their respective collection tubes based on their charge. Adapted from Li et al. (2019)

waveform was divided into 100 elements with 50 percent overlap to increase the redundancy and augmenting their dataset thus resulting in an increased training stability. The one-dimensional time series waveform elements were reshaped into 2D data corresponding to the pulses. This digital data was further reduced by factor of 40 to achieve an acceptable trade-off between the processing time and the accuracy of the deep learning model. The learning model consisted of 16 convolutional layers to learn and extract the features of the input data, 3 max pooling layers to reduce the number of parameters and computations, and 3 fully-connected network with dropout regularization which finally concludes with a Softmax layer. The authors demonstrated the utility of their network to classify between 3 categories with 95% accuracy: SW-480 colorectal cancer cells, OT-II hybridoma white blood cells, and the blank examples i.e. running buffer. Neural networks were also used in an inline holography microscopy for label-free, high-speed, cell sorting (Schneider et al. 2015). They showed that this label-free imaging technique can be applied for ultrafast, cell sorting with classification accuracy of 89%.

In a microfluidic-based imaging flow cytometry (IFC) technique, an accurate classification framework was presented for the first time. It was based on deep learning for IFC data extracted from three unstained, unlabeled, leukemia cell lines (K562, MOLT, and HL60) (Gopakumar et al. 2017). They demonstrated that instead of using conventional fine segmentation and explicit feature extraction, deep learning algorithms could successfully classify the cell lines up to a precision of 78.8%. Sun et al. (2019) used a deep CNN to learn the biological characteristics of 2D light scattering patterns in the azimuthal and polar angle from a microfluidic cytometer and ultimately identified label-free lymphocytic leukemia cells. Their deep learning network accurately detected Jurkat and BALL-1 cells with an accuracy of 93.2%, and the sensitivity and specificity were 92% and 94.453%. Chen et al. (2016) integrated feature extraction and deep learning with high-throughput quantitative imaging enabled by photonic time stretch, achieving high classification accuracy (95.5%). Their system captures quantitative optical phase and intensity and extracts multiple biophysical features of individual cells. These biophysical measurements thus form a hyperdimensional feature space in which supervised ML is performed for cell classification.

SVMs offer an alternative to NNs for use as a classification framework and have been demonstrated as a useful tool

for a wide variety of classification problems as evident from the research reviewed in the following paragraphs. Researchers applied an SVM algorithm to analyze MFC dataset in order to automatically detect minimal residue disease in acute myeloid leukemia and myelodysplastic syndrome patients (Ko et al. 2018). The original raw data were encoded using a multivariate Gaussian mixture model and then fed into the SVM classifier. They validated this with a large-scale clinical data and clinical outcome. Another research group developed an *in vivo* Photoacoustic Flow Cytometry (PAFC) system to achieve *in vivo* melanoma inspection (Fu et al. 2019). They implemented a support vector machine algorithm to discriminate signals and noises based on the continuity, amplitude, and photoacoustic waveform pulse width extracted from photoacoustic waves. A model accuracy of 92% was accomplished. Lin et al. (2018) developed a label-free light-sheet microfluidic cytometer for single cell analysis by two-dimensional (2D) light scattering measurements. Incorporating the cytometer with SVM algorithms, a high accuracy was achieved in automatic classification of senescent and normal human fibroblasts. Four parameters (contrast, correlation, energy and homogeneity) were calculated for light scattering patterns and used as features in the SVM classifier. A linear kernel function was adopted with fivefold cross validation. Toedling et al. (2006) used an SVM for automatic detection of leukemic cells from patients' bone marrow and peripheral blood samples in flow cytometry readouts. Manually gated leukemic cells were recovered by SVM with 98.87% specificity and 99.78% sensitivity which showed the potential of a well-established multivariate-analysis technique.

Huang et al. (2014) demonstrated the use of a technique based on Extreme Learning Machines (ELM) for single-frame super-resolution processing applied on a microfluidic contact imaging cytometer platform. Compared with the commercial flow cytometer, less than 8% error was observed for the absolute number of microbeads. They demonstrated in another paper that by mixed flowing of HepG2 and Huh7 cells as the inputs, the developed scheme achieved 23% better recognition accuracy compared to the one without error recovery. Whereas, it also achieved an average of 98.5% resource-saving compared to the previous multi-frame super-resolution processing (Huang et al. 2015). Autoencoders are considered an unsupervised learning technique since they don't need explicit labels to train on. However, to be more precise they are self-supervised because they generate their own labels from the training data. A microfluidics-based platform for single-cell imaging in-flow and subsequent image analysis using Variational Autoencoders (VAE) for unsupervised characterization of cellular mixtures was demonstrated in Constantinou et al.'s paper (2019). Heterogeneous mixtures of yeast species were classified with 88% accuracy. Microfluidic Imaging Flow Cytometry (MIFC) is an emerging method of microscopic imaging, which aims to reduce the complexity of the tasks involved in cytometry by combining flow cytometry with digital microscopy (Kalmady et al. 2017). This technique promised significantly higher throughput and was easy to set up with minimal expenses in Kalmady et al. study (2017). This group employed MIFC for obtaining images instead of image cytometry. They proposed a transfer learning and ensemble learning-based approach for the automation of cytopathological analysis of Leukemia cell-line images. Compared to earlier works, the use of fine-tuned features from a modified deep NN for transfer learning provided a substantial improvement in performance.

3.1.3 Smartphone-based detection

Smartphone-based sensors are closely related to the Image-based sensors since they replace the microscope with the smartphone cameras, which are often supplemented by an attachment and have the same output i.e. image. They are becoming increasingly popular because of their small footprint and widespread availability of smartphones. Furthermore, they eliminate the need for specialized optical equipment like microscopes and spectroscopes by substituting it with relatively inexpensive and portable attachments. Smartphone-based biosensors utilizing ML, therefore, have tremendous promise for being used as point-of-care diagnostic devices with minimal training and knowledge required for operation.

A cost-effective method proposed by de Haan et al. (2020) was capable of automatic screening of sickle cells (SC) in a deep learning framework. The framework included two complementary deep NN (Fig. 4). The first one standardized and enhanced blood smear images from a smartphone microscope while the second one acted on the output of the first image and performed the semantic segmentation between SC and healthy cells in a blood smear. Furthermore, the segmented images were utilized for the diagnosis of SC disease and achieved an accuracy of 98%.

A comparison of different ML algorithms was carried out for waterborne pathogen (Giardi) detection using a smartphone- based setup (Koydemir et al. 2017). The accuracy and the Area Under the ROC curve (AUROC) of different ML models were compared including, but not limited to SVM, nearest neighbors and ensemble methods. All the models had a classification accuracy above 81%, while the AUROC values were greater than 0.7. The best predictive performance was obtained using bagged trees (Ntree = 400). Fine and cubic kNN classifiers provided fast fitting speeds, but their predictive accuracy was relatively poor. On the other hand, SVM and bagged ensemble classifiers were promising at their prediction accuracy, while their training speeds were slower.

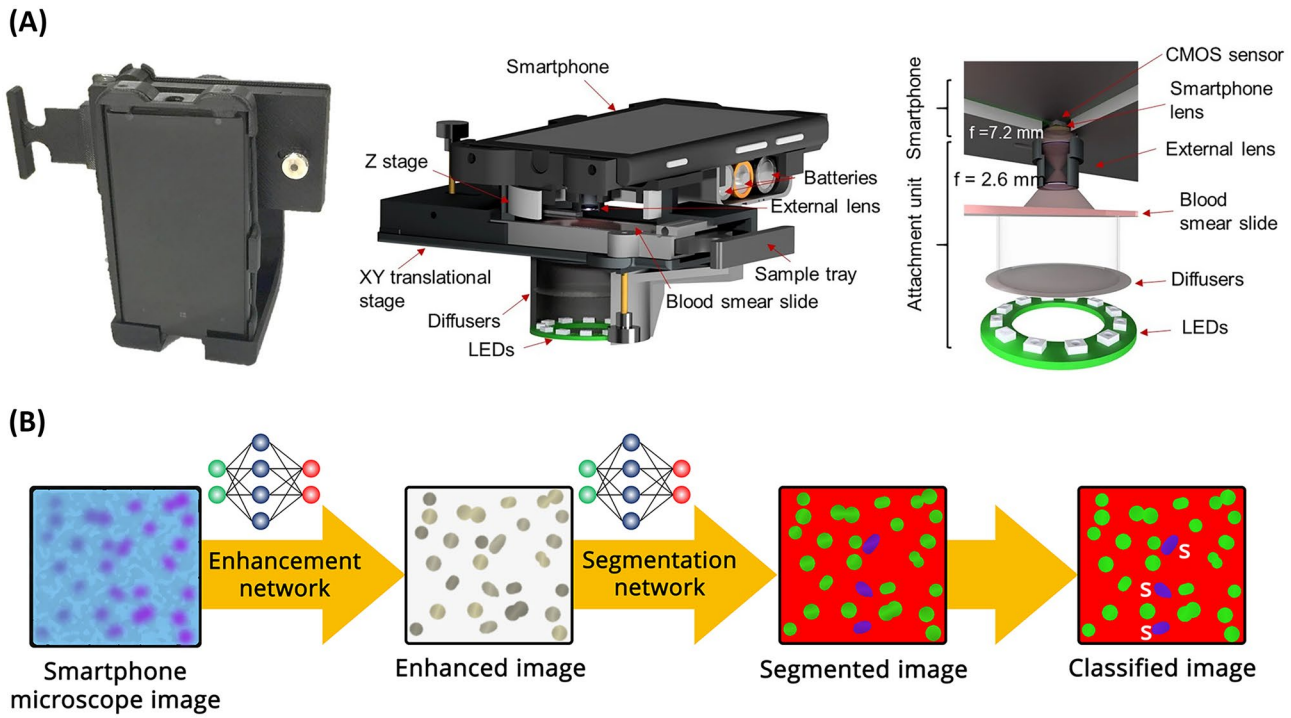

Fig. 4 Overview of smartphone-based biosensor employed by de Haan et al. (2020) **A** Photograph of the smartphone-based system, the overall design, and the light path is shown from left to right. Reprinted from de Haan et al. (2020) **B** workflow of deep learning process is presented. This learning algorithm has been used sickle cell analysis to enhance blood smear images and carry out semantic segmentation between SC and healthy cells. Adapted from de Haan et al. (2020)

3.2 Electrical detection

Electrical detection refers to the use of electrical circuits to obtain data in the form of electrical signals. These signals can be impedance, voltage, current or any other electrical signal. Impedance is generally the most commonly used parameter to identify and quantify cells. When a cell passes through the electrodes in a microfluidic channel, a change in impedance occurs. The output signal is determined by the cell's properties such as cell size, conductivity, and permittivity. Electrical Detection of cells has many advantages over traditional optical detection. Since there is no need for bulky optical equipment, electrical detection devices usually have a small footprint and are less expensive. In the following paragraphs, we present biosensors utilizing ML techniques for electrical detection of various biomarkers. A schematic diagram of an electrical impedance cytometer with SVM for data analysis is shown in Fig 5.

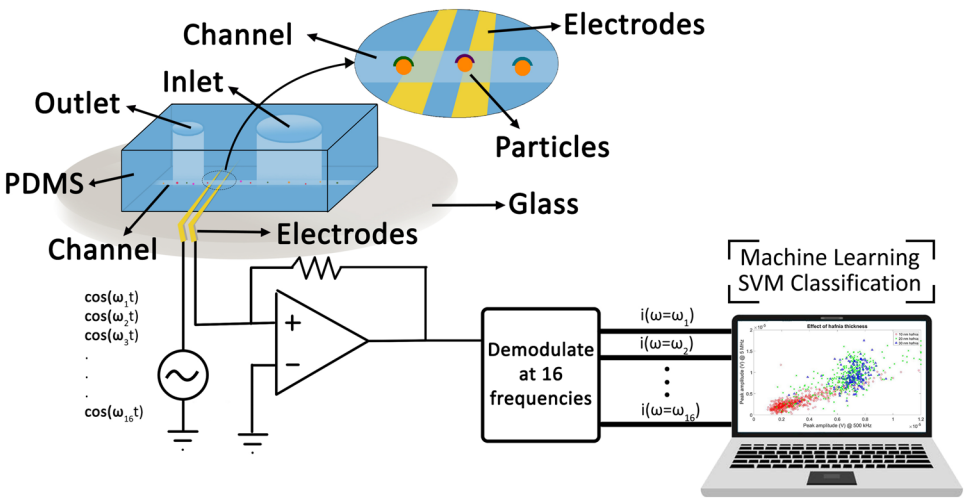

Fig. 5 Schematic diagram of an electrical impedance cytometer. As cells flow from the inlet to the outlet in these biosensors, the change in impedance is measured by a lock-in amplifier. The lock-in amplifier can apply signal in different frequencies at a time. The data are then recorded and analyzed using SVM. Adapted from Sui et al. (2020)

Integrated with ML algorithms, a microfluidic impedance cytometry for real-time, label-free multiparametric characterization of biological cells (Honrado et al. 2020). In this study, a recurrent NN was designed to predict cell diameter, velocity, and position from electric current signals, measured by a microfluidic impedance chip. The trained network was able to characterize geometric and electrical properties of beads, red blood cells, and yeasts with a good accuracy and a unitary prediction time of 0.4 ms. Zhao et al. (2018) designed a new microfluidic impedance cytometry with crossing constriction microchannels, which allows quantifying the cellular electrical markers. Using an equivalent circuit model, they translated the measured impedance values to specific membrane capacitance and cytoplasm conductivity. A NN-based pattern recognition was used to classify tumor cell lines and tumor cells with epithelial-mesenchymal transitions. Precise measurement of mechanical and/or electrical properties of cells or cell components yields useful information on the physiological and pathological state of cells and is critical for cell classification. Yang et al. (2019) extracted deformability, electrical impedance and relaxation index of single cells from impedance spectroscopy measurements with self-aligned 3D electrodes. They demonstrated the ability of their system to detect and classify cells using a back propagation NN completely based on the biophysical properties of the cells. In another study, a microfluidic constriction channel was designed to measure single-cell electrical properties (Zheng et al. 2013). A back propagation NN was used for cell classification based on three parameters of diameter, specific membrane capacitance, and cytoplasm conductivity. Finally, they showed that cell classification success rate significantly improved when information additional to cell size was included.

In Zhao et al.'s work (2013), osteoblasts and osteocytes were classified using a two-layer back propagation NN70. The input data had three groups of parameters measured on cells, namely, transit time, impedance amplitude ratio, and cell elongation length. Their results suggested that biomechanical and bioelectrical parameters, when used in combination, provided a higher cell classification success rate than using alone. In another study, a microfluidic system was presented for cell type classification based on size-independent electrical properties, specific membrane capacitance and cytoplasm conductivity. Two lung tumor cell lines were classified using a two-layer back propagation NN. The NN-based classification resulted in a fairly acceptable classification success rate of 65.4% (CSpecific Membrane), 71.4% (σcytoplasm), and 74.4% (CSpecific Membrane combined with σcytoplasm). A microfluidic system proposed by Zheng et al. (2012) with a constriction channel. The channel was marginally smaller than the RBC's diameters which was used to classify adult and neonatal RBCs using a back propagation NN through their biophysical properties (mechanical and electrical). Electrical measurements were performed to characterize these properties. The input data had three group of parameters (transit time, amplitude ratio and phase increase). The results showed that when these parameters were used in combination, yielded a relatively higher classification accuracy (84.8%) than the time each parameter was used alone. Recently a study published where the authors used Quadratic Discriminant Analysis (QDA). This is a type of supervised ML algorithm that helped them extract six features from Red Blood Cells (RBCs) and yeast cells using Impedance micro-cytometry. They achieved the maximum test accuracy (99%) by using four features on RBCs. They also demonstrated the efficacy of their platform by classifying different cancer subtypes. The accuracy decreased when more than four features was used. It was because of overfitting of the model to the training data (Joshi et al. 2020).

A study conducted cancer drug efficacy analysis using multifrequency impedance cytometry, measuring the impedance of a single cell at several discrete frequencies (Ahuja et al. 2019). Support vector machine algorithm was implemented to help differentiate alive cells from dead cells. Song et al. (2013) employed a support vector machine algorithm to help identify differentiation states of stem cells based on impedance signals collected by the microfluidic electrical impedance flow cytometer at 50 kHz, 250 kHz, 500 kHz and 1 MHz. Another research group discriminated strains of E. coli K-12, E. coli O157: H7, and Salmonella Thompson using a multichannel immunosensor incorporated with multiclass support vector machines (Zuo et al. 2006). Gini-SVM framework was adopted to design multiclass SVMs. To evaluate the performance, a 100-fold cross-validation procedure was implemented.

Detection and enumeration of circulating tumor cells from red blood cells were performed in research using a micropore-based microfluidic impedance cytometer (Guo et al. 2014). The peak amplitude and the pulse bandwidth of signal pulses were analyzed by SVM to differentiate cancer cells from red blood cells. Radial basis function (RBF) was appointed as the kernel function. The results of the proposed microfluidic sensor combined with SVM showed a good agreement with the results of a commercial flow cytometer. Wang et al. (2017) proposed a sensitive multiplex self-referencing SERS pathogen detection scheme. A linear kernel-based SVM in conjunction of PCA was performed for rapid discrimination and classification of target bacteria with a detection accuracy above 95%. An approach for hematocrit estimation from the transduced anodic current curves introduced in a study. The curves were obtained by glucose-oxidase reaction in the strip-type electrochemical biosensors (Park et al. 2008). The support vector machine was implemented for regression with the target value of accurate hematocrit values measured by a hospital analysis system.

4 Conclusions

ML has become a useful tool in analyzing and classifying the data obtained from biosensors for cellular analysis. Based on the papers we discussed here, we see that both SVM and ANN are the prevalent techniques which are effective in automating the classification of various cell types except for Image-based biosensors. This may be due to the fact that these are the most widely known techniques for researchers working in fields related to biosensors. In the case of Image-based biosensors, various NN architectures are preferred over SVM and other methods for classification of different cell types.

Table 1 compares important characteristics for several papers that are discussed in this review. It is divided into 2 main categories for easily comparing each method's efficacy. In the first part, we list the papers in which the number of cells are specified as the dataset size. The second part of the table contains the papers in which the number of images is considered as the dataset size. Other than these 2 main categories, we have listed a few papers in which the dataset neither corresponds to the number of images nor the number of cells. The last column of Table 1 is a ratio of the classifier accuracy (in percentage) to the dataset size. A higher number, therefore, indicates that the accuracy achieved corresponds to a relatively small dataset.

We also noted that Deep Learning and ANN have grown more popular recently with the majority of the newer publications using these methods. Another interesting observation is that biosensors utilizing electrical detection methods rarely employ deep learning as the analytical tool for classifying cells. This may be due to the fact that deep learning is data hungry and databases for electrical biosensing data are not yet established. On the other hand, there is a plethora of datasets easily available that may be used as training samples for image and optical detection of various biomarkers.

To summarize, the use of ML algorithms in biosensors have huge benefits that automate the cumbersome and complicated process of extracting, processing and analyzing data that is generated by the biosensors. Such an automation eliminates the need for an experienced professional to make sense of the data and moves us closer to providing Point-of-care health solutions in environments that have low resources. Although ML algorithms have been around for a while now and have huge benefits, the techniques discussed here mostly utilize code and require certain Integrated Development Environments (IDE's) e.g. Python, MATLAB etc. for their use. Researchers should consider packaging of the softwares into a GUI which will make these relatively simple to interact with and less formidable.

Funding This research was sponsored by National Science Foundation Awards 1556253 (IDBR), 1846740 (CAREER), and 1711165 (ECCS-CCSS) and also the Defense Advanced Research Projects Agency (DARPA) Biological Technologies Office (BTO) Electrical Prescriptions (ElectRx) program managed by Dr. E. V. Giesen through the DARPA Contracts Management Office Grant/Contract Number HR0011-16–2-0026 and Contract Number N660011824018. The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the National Science Foundation, the Department of Defense, or the U.S. Government.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflicts of interest Disclosure of potential conflict of interest: M. Javanmard and J. Sui has a pending patent for "Use of Multi-Frequency Impedance Cytometry in Conjunction with Machine Learning for Classification of Biological Particles" and M. Javanmard has equity in Rizlab Health Inc., a company dedicated to commercialization of a microfluidic hematology analyzer.

References

K. Ahuja, G.M. Rather, Z. Lin, J. Sui, P. Xie, T. Le, J.R. Bertino, M. Javanmard, Toward point-of-care assessment of patient response: a portable tool for rapidly assessing cancer drug efficacy using multifrequency impedance cytometry and supervised machine learning. Microsyst. Nanoeng. **5**, 1–11 (2019). https://doi.org/10.1038/s41378-019-0073-2

S. Akaho, SVM maximizing margin in the input space, in *ICONIP 2002 - Proc. 9th Int. Conf. Neural Inf. Process. Comput. Intell. E-Age*. (Institute of Electrical and Electronics Engineers Inc., 2002), pp. 1069–1073. https://doi.org/10.1109/ICONIP.2002.1198224

S.U. Akram, J. Kannala, L. Eklund, J. Heikkilä, Cell segmentation proposal network for microscopy image analysis, in *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. (Springer Verlag, 2016), pp. 21–29. https://doi.org/10.1007/978-3-319-46976-8_3

O. Alonso, Challenges with label quality for supervised learning. ACM. J. Data. Inf. Qual. **2**, 1–3 (2015). https://doi.org/10.1145/2724721

M.M. Amin, S. Kermani, A. Talebi, M.G. Oghli, Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier. J. Med. Signals. Sens. **5**, 49–58 (2015). https://doi.org/10.4103/2228-7477.150428

H. Bachratý, K. Bachratá, M. Chovanec, I. Jančigová, M. Smiešková, K. Kovalčíková, Applications of machine learning for simulations of red blood cells in microfluidic devices. BMC. Bioinf. **21**, 1–15 (2020). https://doi.org/10.1186/s12859-020-3357-5

A. Bamshad, A. Nikfarjam, M.H. Sabour, H. Raji, Theoretical and Numerical Investigation of Liquid-Gas Interface Location of Capillary Driven Flow during the Time Throughout Circular Microchannels, in *5th RSI Int. Conf. Robot. Mechatronics, IcRoM 2017*. (2018), pp. 432–438. https://doi.org/10.1109/ICRoM.2017.8466144

D. Bardin, A.P. Lee, Low-cost experimentation for the study of droplet microfluidics. Lab. Chip. **14**, 3978–3986 (2014). https://doi.org/10.1039/c4lc00424h

R.J. Baron, Mechanisms of human facial recognition. Int. J. Man. Mach. Stud. **15**, 137–178 (1981). https://doi.org/10.1016/S0020-7373(81)80001-6

U. Bora, Nucleic Acid Based Biosensors for Clinical Applications. Biosens. J. **2**, 1–8 (2013). https://doi.org/10.4172/2090-4967.1000104

R.M. Califf, Biomarker definitions and their applications. Exp. Biol. Med. **243**, 213–221 (2018). https://doi.org/10.1177/1535370217750088

M. Carminati, G. Ferrari, M.D. Vahey, J. Voldman, M. Sampietro, Miniaturized Impedance Flow Cytometer: Design Rules and Integrated Readout. IEEE. Trans. Biomed. Circuits. Syst. **11**, 1438–1449 (2017). https://doi.org/10.1109/TBCAS.2017.2748158

C.L. Chen, A. Mahjoubfar, L.C. Tai, I.K. Blaby, A. Huang, K.R. Niazi, B. Jalali, Deep Learning in Label-free Cell Classification. Sci. Rep. **6**, 1–16 (2016). https://doi.org/10.1038/srep21471

J. Chen, Y. Zheng, Q. Tan, E. Shojaei-Baghini, Y.L. Zhang, J. Li, P. Prasad, L. You, X.Y. Wu, Y. Sun, Classification of cell types using a microfluidic device for mechanical and electrical measurement on single cells. Lab. Chip. **11**, 3174–3181 (2011). https://doi.org/10.1039/c1lc20473d

I. Constantinou, M. Jendrusch, T. Aspert, F. Görlitz, A. Schulze, G. Charvin, M. Knop, Self-Learning Microfluidic Platform for Single-Cell Imaging and Classification in Flow. Micromachines. **10**, 311 (2019). https://doi.org/10.3390/mi10050311

P. Cunningham, M. Cord, S.J. Delany, Supervised learning, in *Machine Learning Techniques for Multimedia*. (Springer Verlag, 2008), pp. 21–49. https://doi.org/10.1007/978-3-540-75171-7_2

K. de Haan, H.C. Koydemir, Y. Rivenson, D. Tseng, E. Van Dyne, L. Bakic, D. Karinca, K. Liang, M. Ilango, E. Gumustekin, A. Ozcan, Automated screening of sickle cells using a smartphone-based microscope and deep learning. NPJ. Digit. Med. **3**, 1–9 (2020). https://doi.org/10.1038/s41746-020-0282-y

F. Del Ben, M. Turetta, G. Celetti, A. Piruska, M. Bulfoni, D. Cesselli, W.T.S. Huck, G. Scoles, A Method for Detecting Circulating Tumor Cells Based on the Measurement of Single-Cell Metabolism in Droplet-Based Microfluidics, Angew. Chemie. **128**, 8723–8726 (2016). https://doi.org/10.1002/ange.201602328

J. Deng, W. Dong, R. Socher, L.-J. Li, L Kai, F.-F. Li, ImageNet: A large-scale hierarchical image database, in *Institute of Electrical and Electronics Engineers (IEEE)*.(2010), pp. 248–255. https://doi.org/10.1109/cvpr.2009.5206848

D. Dension, A Bayesian CART algorithm. Biometrika **85**, 363–377 (1998). https://doi.org/10.1093/biomet/85.2.363

M. Dixit, S. Silakari, Face recognition using approximated bezier curve and supervised learning approach. Int. J. Multimed. Ubiquitous. Eng. **10**, 311–324 (2015). https://doi.org/10.14257/ijmue.2015.10.4.30

T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T.L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, O. Ronneberger, U-Net: deep learning for cell counting, detection, and morphometry. Nat. Methods. **16**, 67–70 (2019). https://doi.org/10.1038/s41592-018-0261-2

Y. Fu, Q. Zhou, Q. Liu, L. Tao, X. Wei, Photoacoustic signal classification for *in vivo* photoacoustic flow cytometry based on support vector machine, in *Opt. Heal. Care Biomed. Opt. IX*. eds. by Q. Luo, X. Li, Y. Tang, Y. Gu, D. Zhu (SPIE, 2019) p. 89. https://doi.org/10.1117/12.2537481

Z. Ghahramani, Unsupervised Learning, in *Advanced Lectures on Machine Learning. ML 2003. Lecture Notes in Computer Science*, vol 3176. eds. by O. Bousquet, U. von Luxburg, G. Rätsch (Springer, Berlin, Heidelberg, 2004), pp 72–112. https://doi.org/10.1007/978-3-540-28650-9_5

T. Go, H. Byeon, S.J. Lee, Label-free sensor for automatic identification of erythrocytes using digital in-line holographic microscopy and machine learning. Biosens. Bioelectron. **103**, 12–18 (2018). https://doi.org/10.1016/j.bios.2017.12.020

K. Goda, A. Mahjoubfar, C. Wang, A. Fard, J. Adam, D.R. Gossett, A. Ayazi, E. Sollier, O. Malik, E. Chen, Y. Liu, R. Brown, N. Sarkhosh, D. Di Carlo, B. Jalali, Hybrid dispersion laser scanner. Sci. Rep. **2**, 1–8 (2012). https://doi.org/10.1038/srep00445

K. Goda, K.K. Tsia, B. Jalali, Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena. Nature **458**, 1145–1149 (2009). https://doi.org/10.1038/nature07980

G. Gopakumar, K.H. Babu, D. Mishra, S.S. Gorthi, G.R.K.S. Subrahmanyam, Cytopathological image analysis using deep-learning networks in microfluidic microscopy. J. Opt. Soc. Am. A. **34**, 111 (2017). https://doi.org/10.1364/josaa.34.000111

B. Guo, C. Lei, H. Kobayashi, T. Ito, Y. Yalikun, Y. Jiang, Y. Tanaka, Y. Ozeki, K. Goda, High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy. Cytom. a. **91**, 494–502 (2017). https://doi.org/10.1002/cyto.a.23084

B. Guo, C. Lei, Y. Wu, H. Kobayashi, T. Ito, Y. Yalikun, S. Lee, A. Isozaki, M. Li, Y. Jiang, A. Yasumoto, D. Di Carlo, Y. Tanaka, Y. Yatomi, Y. Ozeki, K. Goda, Optofluidic time-stretch quantitative phase microscopy. Methods **136**, 116–125 (2018). https://doi.org/10.1016/j.ymeth.2017.10.004

J. Guo, Z. Chen, Y.-L. Ban, Y. Kang, Precise Enumeration of Circulating Tumor Cells Using Support Vector Machine Algorithm on a Microfluidic Sensor. IEEE. Trans. Emerg. Top. Comput. **5**, 518–525 (2014). https://doi.org/10.1109/tetc.2014.2335539

A. Gupta, P.J. Harrison, H. Wieslander, N. Pielawski, K. Kartasalo, G. Partel, L. Solorzano, A. Suveer, A.H. Klemm, O. Spjuth, I.M. Sintorn, C. Wählby, Deep Learning in Image Cytometry: A Review. Cytom. a. **95**, 366–380 (2019). https://doi.org/10.1002/cyto.a.23701

X.H. Han, J. Lei, Y.W. Chen, HEp-2 cell classification using K-support spatial pooling in deep CNNs, in *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. (Springer Verlag, 2016) pp. 3–11. https://doi.org/10.1007/978-3-319-46976-8_1

Y.J. Heo, D. Lee, J. Kang, K. Lee, W.K. Chung, Real-time Image Processing for Microscopy-based Label-free Imaging Flow Cytometry in a Microfluidic Chip. Sci. Rep. **7**, 790–784 (2017). https://doi.org/10.1038/s41598-017-11534-0

C. Honrado, J.S. McGrath, R. Reale, P. Bisegna, N.S. Swami, F. Caselli, A neural network approach for real-time particle/cell characterization in microfluidic impedance cytometry. Anal. Bioanal. Chem. **412**, 3835–3845 (2020). https://doi.org/10.1007/s00216-020-02497-9

A.T.H. Hsieh, P.J. Pan, A.P. Lee, A real-time characterization method to rapidly optimize molecular beacon signal for sensitive nucleic acids analysis. Anal. Bioanal. Chem. **406**, 3059–3067 (2014). https://doi.org/10.1007/s00216-014-7721-z

X. Huang, J. Guo, X. Wang, M. Yan, Y. Kang, H. Yu, A Contact-Imaging Based Microfluidic Cytometer with Machine-Learning for Single-Frame Super-Resolution Processing. PLoS ONE **9**, e104539 (2014). https://doi.org/10.1371/journal.pone.0104539

X. Huang, X. Wang, M. Yan, H. Yu, A robust recognition error recovery for micro-flow cytometer by machine-learning enhanced single-frame super-resolution processing. Integration. **51**, 208–218 (2015). https://doi.org/10.1016/j.vlsi.2014.07.004

X. Huang, Y. Jiang, H. Xu, X. Liu, H.W. Hou, M. Yan, H. Yu, A convolutional neural network based single-frame super-resolution for lensless blood cell counting, in *Proc. – 2016 IEEE Biomed. Circuits Syst. Conf. BioCAS 2016*. (Institute of Electrical and Electronics Engineers Inc., 2016a) pp. 168–171. https://doi.org/10.1109/BioCAS.2016.7833758

X. Huang, Y. Jiang, X. Liu, H. Xu, Z. Han, H. Rong, H. Yang, M. Yan, H. Yu, Machine Learning Based Single-Frame Super-Resolution

Processing for Lensless Blood Cell Counting. Sensors. **16**, 1836 (2016b). https://doi.org/10.3390/s16111836

M.A.L. Iragorri, S. El Hoss, V. Brousse, S.D. Lefevre, M. Dussiot, T. Xu, A.R. Ferreira, Y. Lamarre, A.C.S. Pinto, S. Kashima, C. Lapouméroulie, D.T. Covas, C. Le Van Kim, Y. Colin, J. Elion, O. Français, B. Le Pioufle, W. El Nemer, A microfluidic approach to study the effect of mechanical stress on erythrocytes in sickle cell disease. Lab. Chip. **18**, 2975–2984 (2018). https://doi.org/10.1039/c8lc00637g

J.M. Jackson, J.B. Taylor, M.A. Witek, S.A. Hunsucker, J.P. Waugh, Y. Fedoriw, T.C. Shea, S.A. Soper, P.M. Armistead, Microfluidics for the detection of minimal residual disease in acute myeloid leukemia patients using circulating leukemic cells selected from blood. Analyst. **141**, 640–651 (2016). https://doi.org/10.1039/c5an01836f

K. Joshi, A. Javani, J. Park, V. Velasco, B. Xu, O. Razorenova, R. Esfandyarpour, A Machine Learning-Assisted Nanoparticle-Printed Biochip for Real-Time Single Cancer Cell Analysis. Adv. Biosyst. **4**, 2000160 (2020). https://doi.org/10.1002/adbi.202000160

K.S. Kalmady, A.S. Kamath, G. Gopakumar, G.R.K.S. Subrahmanyam, S.S. Gorthi, Improved Transfer Learning through Shallow Network Embedding for Classification of Leukemia Cells, in *2017 9th Int. Conf. Adv. Pattern Recognition, ICAPR 2017*. (Institute of Electrical and Electronics Engineers Inc., 2018), pp. 127–132. https://doi.org/10.1109/ICAPR.2017.8593186

T.H. Kim, H.J. Yoon, S. Fouladdel, Y. Wang, M. Kozminsky, M.L. Burness, C. Paoletti, L. Zhao, E. Azizi, M.S. Wicha, S. Nagrath, Characterizing Circulating Tumor Cells Isolated from Metastatic Breast Cancer Patients Using Graphene Oxide Based Microfluidic Assay. Adv. Biosyst. **3**, 1800278 (2019). https://doi.org/10.1002/adbi.201800278

K.R. King, S. Wang, D. Irimia, A. Jayaraman, M. Toner, M.L. Yarmush, A high-throughput microfluidic real-time gene expression living cell array. Lab. Chip. **7**, 77–85 (2007). https://doi.org/10.1039/b612516f

B.S. Ko, Y.F. Wang, J.L. Li, C.C. Li, P.F. Weng, S.C. Hsu, H.A. Hou, H.H. Huang, M. Yao, C.T. Lin, J.H. Liu, C.H. Tsai, T.C. Huang, S.J. Wu, S.Y. Huang, W.C. Chou, H.F. Tien, C.C. Lee, J.L. Tang, Clinically validated machine learning algorithm for detecting residual diseases with multicolor flow cytometry analysis in acute myeloid leukemia and myelodysplastic syndrome. EBioMedicine **37**, 91–100 (2018). https://doi.org/10.1016/j.ebiom.2018.10.042

N.A. Koohababni, M. Jahanifar, A. Gooya, N. Rajpoot, Nuclei detection using mixture density networks, in *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. (2018). https://doi.org/10.1007/978-3-030-00919-9_28

H.C. Koydemir, S. Feng, K. Liang, R. Nadkarni, P. Benien, A. Ozcan, Comparison of supervised machine learning algorithms for waterborne pathogen detection using mobile phone fluorescence microscopy. Nanophotonics. **6**, 731–741 (2017). https://doi.org/10.1515/nanoph-2017-0001

Y. Lecun, Y. Bengio, G. Hinton, Deep learning. Nature **521**, 436–444 (2015). https://doi.org/10.1038/nature14539

B.S. Lee, Y.U. Lee, H.S. Kim, T.H. Kim, J. Park, J.G. Lee, J. Kim, H. Kim, W.G. Lee, Y.K. Cho, Fully integrated lab-on-a-disc for simultaneous analysis of biochemistry and immunoassay from whole blood. Lab. Chip. **11**, 70–78 (2011). https://doi.org/10.1039/c0lc00205d

Q. Li, W. Cai, X. Wang, Y. Zhou, D.D. Feng, M. Chen, Medical image classification with convolutional neural network, in *2014 13th Int. Conf. Control Autom. Robot. Vision, ICARCV 2014*. (Institute of Electrical and Electronics Engineers Inc., 2014), pp. 844–848. https://doi.org/10.1109/ICARCV.2014.7064414

Y. Li, A. Mahjoubfar, C.L. Chen, K.R. Niazi, L. Pei, B. Jalali, Deep Cytometry: Deep learning with Real-time Inference in Cell Sorting and Flow Cytometry. Sci. Rep. **9**, 11088 (2019). https://doi.org/10.1038/s41598-019-47193-6

M. Lin, Q. Liu, C. Liu, X. Qiao, C. Shao, X. Su, Label-free light-sheet microfluidic cytometry for the automatic identification of senescent cells. Biomed. Opt. Express. **9**, 1692 (2018). https://doi.org/10.1364/boe.9.001692

R. Livni, S. Shalev-Shwartz, O. Shamir, On the computational efficiency of training neural networks. Adv. Neural. Inf. Process. Syst. **27** (2014)

X. Long, L. Cleveland, Y.L. Yao, Automatic detection of unstained viable cells in bright field images using a support vector machine with an improved training procedure. Comput. Biol. Med. **36**, 339–362 (2006). https://doi.org/10.1016/j.compbiomed.2004.12.002

W. Maass, Networks of spiking neurons: The third generation of neural network models. Neural. Netw. **10**, 1659–1671 (1997). https://doi.org/10.1016/S0893-6080(97)00011-7

G.K. Mani, M. Morohoshi, Y. Yasoda, S. Yokoyama, H. Kimura, K. Tsuchiya, ZnO-Based Microfluidic pH Sensor: A Versatile Approach for Quick Recognition of Circulating Tumor Cells in Blood. ACS. Appl. Mater. Interfaces. **9**, 5193–5203 (2017). https://doi.org/10.1021/acsami.6b16261

Y. Mao, Z. Yin, J.M. Schober, Iteratively training classifiers for circulating tumor cell detection, in *Proc. - Int. Symp. Biomed. Imaging*. (IEEE Computer Society, 2015), pp. 190–194. https://doi.org/10.1109/ISBI.2015.7163847

D. Mayerich, J. Kwon, A. Panchal, J. Keyser, Y. Choe, Fast cell detection in high-throughput imagery using GPU-accelerated machine learning, in *Proc. Int. Symp. Biomed. Imaging*. (2011), pp. 719–723. https://doi.org/10.1109/ISBI.2011.5872507

A. Najafi, S. Janghorbani, S.A. Motahari, E. Fatemizadeh, Statistical association mapping of population-structured genetic data. IEEE. ACM. Trans. Comput. Biol. Bioinf. **16**, 636–649 (2019). https://doi.org/10.1109/TCBB.2017.2786239

W. Ni, B. Hu, C. Zheng, Y. Tong, L. Wang, Q. Li, X. Tong, Y. Han, Automated analysis of acute myeloid leukemia minimal residual disease using a support vector machine. Oncotarget. **7**, 71915–71921 (2016). https://doi.org/10.18632/oncotarget.12430

J.S. Park, H.T. Huynh, Y. Won, Support vector machine for hematocrit density estimation based on changing patterns of transduced anodic current, in *Proc. - 3rd Int. Conf. Converg. Hybrid Inf. Technol. ICCIT 2008*. (2008), pp. 456–460. https://doi.org/10.1109/ICCIT.2008.213

A. Patle, D.S. Chouhan, SVM kernel functions for classification, in *2013 Int. Conf. Adv. Technol. Eng. ICATE 2013*. (2013). https://doi.org/10.1109/ICAdTE.2013.6524743

J. Riordon, D. Sovilj, S. Sanner, D. Sinton, E.W.K. Young, Deep Learning with Microfluidics for Biotechnology. Trends. Biotechnol. **37**, 310–324 (2019). https://doi.org/10.1016/j.tibtech.2018.08.005

E.R. Rodrigues, I. Oliveira, R. Cunha, M. Netto, DeepDownscale: A deep learning strategy for high-resolution weather forecast, in *Proc. - IEEE 14th Int. Conf. EScience, e-Science 2018*. (Institute of Electrical and Electronics Engineers Inc., 2018) pp. 415–422. https://doi.org/10.1109/eScience.2018.00130

R. Saravanan, P. Sujatha, A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification, in *Proc. 2nd Int. Conf. Intell. Comput. Control Syst. ICICCS 2018*. (Institute of Electrical and Electronics Engineers Inc., 2019), pp. 945–949. https://doi.org/10.1109/ICCONS.2018.8663155

B. Schneider, G. Vanmeerbeeck, R. Stahl, L. Lagae, J. Dambre, P. Bienstman, Neural network for blood cell classification in a holographic microscopy system, in *Int. Conf. Transparent Opt. Networks*. (IEEE Computer Society, 2015). https://doi.org/10.1109/ICTON.2015.7193315

D.K. Singh, C.C. Ahrens, W. Li, S.A. Vanapalli, Label-free, high-throughput holographic screening and enumeration of tumor

cells in blood. Lab. Chip. **17**, 2920–2932 (2017). https://doi.org/10.1039/c7lc00149e

G. Soldati, F. Del Ben, G. Brisotto, E. Biscontin, M. Bulfoni, A. Piruska, A. Steffan, M. Turetta, V. Della Mea, Microfluidic droplets content classification and analysis through convolutional neural networks in a liquid biopsy workflow. Am. J. Transl. Res. **10**, 4004–4016 (2018). www.ajtr.org. Accessed 30 Oct 2020

H. Song, Y. Wang, J.M. Rosano, B. Prabhakarpandian, C. Garson, K. Pant, E. Lai, A microfluidic impedance flow cytometer for identification of differentiation state of stem cells. Lab. Chip. **13**, 2300–2310 (2013). https://doi.org/10.1039/c3lc41321g

S. Srivastava, J.J. Rodríguez, A.R. Rouse, M.A. Brewer, A.F. Gmitro, Computer-aided identification of ovarian cancer in confocal microendoscope images. J. Biomed. Opt. **13**, 024021 (2008). https://doi.org/10.1117/1.2907167

M.C. Su, C.Y. Cheng, P.C. Wang, A neural-network-based approach to white blood cell classification. Sci. World. J. **2014**, 796371 (2014). https://doi.org/10.1155/2014/796371

J. Sui, P. Xie, Z. Lin, M. Javanmard, Electronic classification of barcoded particles for multiplexed detection using supervised machine learning analysis. Talanta **215**, 120791 (2020). https://doi.org/10.1016/j.talanta.2020.120791

J. Sun, L. Wang, Q. Liu, X. Su, Automatic detection of leukemia cells by 2D light scattering microfluidic cytometry and deep learning, in *Opt. Heal. Care Biomed. Opt. IX*. eds. by Q. Luo, X. Li, Y. Tang, Y. Gu, D. Zhu (SPIE-Intl Soc Optical Eng, 2019) p. 65. https://doi.org/10.1117/12.2537094

P.R. Tabrizi, S.H. Rezatofighi, M.J. Yazdanpanah, Using PCA and LVQ neural network for automatic recognition of five types of white blood cells, in *2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC'10*. (2010), pp. 5593–5596. https://doi.org/10.1109/IEMBS.2010.5626788

D.R. Thevenot, K. Toth, R.A. Durst, G.S. Wilson, Electrochemical biosensors: recommended definitions and classification. Pure. Appl. Chem. **71**(12), 2333–2348 (1999). http://iupac.org/publications/pac/71/12/2333/. Accessed 25 Apr 2022

J. Toedling, P. Rhein, R. Ratel, L. Karawajew, R. Spang, Automated in-silico detection of cell populations in flow cytometry readouts and its application to leukemia disease monitoring. BMC. Bioinf. **7**, 1–11 (2006). https://doi.org/10.1186/1471-2105-7-282

G. Tomaiuolo, Biomechanical properties of red blood cells in health and disease towards microfluidics. Biomicrofluidics **8**, 051501 (2014). https://doi.org/10.1063/1.4895755

B. Turan, T. Masuda, W. Lei, A.M. Noor, K. Horio, T.I. Saito, Y. Miyata, F. Arai, A pillar-based microfluidic chip for T-cells and B-cells isolation and detection with machine learning algorithm. Robomech. J. **5**, 1–9 (2018). https://doi.org/10.1186/s40648-018-0124-8

M. Ugawa, C. Lei, T. Nozawa, T. Ideguchi, D. Di Carlo, S. Ota, Y. Ozeki, K. Goda, High-throughput optofluidic particle profiling with morphological and chemical specificity. Opt. Lett. **40**, 4803 (2015). https://doi.org/10.1364/ol.40.004803

F. Uslu, K. Icoz, K. Tasdemir, B. Yilmaz, Automated quantification of immunomagnetic beads and leukemia cells from optical microscope images, Biomed. Signal Process. Control. **49**, 473–482 (2019). https://doi.org/10.1016/j.bspc.2019.01.002

J. Vaidya, H. Yu, X. Jiang, Privacy-preserving SVM classification. Knowl. Inf. Syst. **14**, 161–178 (2008). https://doi.org/10.1007/s10115-007-0073-7

D.A. Van Valen, T. Kudo, K.M. Lane, D.N. Macklin, N.T. Quach, M.M. DeFelice, I. Maayan, Y. Tanouchi, E.A. Ashley, M.W. Covert, Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. PLOS. Comput. Biol. **12**, e1005177 (2016). https://doi.org/10.1371/journal.pcbi.1005177

C. Wang, F. Madiyar, C. Yu, J. Li, Detection of extremely low concentration waterborne pathogen using a multiplexing self-referencing SERS microfluidic biosensor. J. Biol. Eng. **11**, 9 (2017). https://doi.org/10.1186/s13036-017-0051-x

H. Wang, H.C. Koydemir, Y. Qiu, B. Bai, Y. Zhang, Y. Jin, S. Tok, E.C. Yilmaz, E. Gumustekin, Y. Rivenson, A. Ozcan, Early-detection and classification of live bacteria using time-lapse coherent imaging and deep learning. Light Sci. Appl. **9**, 118 (2020). https://doi.org/10.1038/s41377-020-00358-9

Q. Wang, F. Miao, The support vector machine and its application to hydrocarbon discriminant in oil and gas exploration, in *2012 Int. Conf. Syst. Informatics, ICSAI 2012*. (2012) pp. 2337–2341. https://doi.org/10.1109/ICSAI.2012.6223523

X. Wang, W. He, D. Metaxas, R. Mathew, E. White, Cell segmentation and tracking using texture-adaptive snakes, in *2007 4th IEEE Int. Symp. Biomed. Imaging From Nano to Macro - Proc*. (2007), pp. 101–104. https://doi.org/10.1109/ISBI.2007.356798

T. Xia, R. Jiang, Y.Q. Fu, N. Jin, Automated Blood Cell Detection and Counting via Deep Learning for Microfluidic Point-of-Care Medical Devices. IOP. Conf. Ser. Mater. Sci. Eng. **2019**, 012048. https://doi.org/10.1088/1757-899X/646/1/012048

T. Xiao, T. Xia, Y. Yang, C. Huang, X. Wang, Learning from massive noisy labeled data for image classification, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015) pp. 2691–2699

M. Xu, D.P. Papageorgiou, S.Z. Abidi, M. Dao, H. Zhao, G.E. Karniadakis, A deep convolutional neural network for classification of red blood cells in sickle cell anemia. PLOS. Comput. Biol. **13**, e1005746 (2017). https://doi.org/10.1371/journal.pcbi.1005746

W. Xue, X. Tan, M.K. Khaing Oo, G. Kulkarni, M.A. Ilgen, X. Fan, Rapid and sensitive detection of drugs of abuse in sweat by multiplexed capillary based immuno-biosensors. Analyst. **145**, 1346–1354 (2020). https://doi.org/10.1039/c9an02498k

D. Yang, Y. Zhou, Y. Zhou, J. Han, Y. Ai, Biophysical phenotyping of single cells using a differential multiconstriction microfluidic device with self-aligned 3D electrodes. Biosens. Bioelectron. **133**, 16–23 (2019). https://doi.org/10.1016/j.bios.2019.03.002

S. Yang, B. Fang, W. Tang, X. Wu, J. Qian, W. Yang, Faster R-CNN based microscopic cell detection, in *2017 Int. Conf. Secur. Pattern Anal. Cybern. SPAC 2017*. (Institute of Electrical and Electronics Engineers Inc., 2018), pp. 345–350. https://doi.org/10.1109/SPAC.2017.8304302

B.Y. Yu, C. Elbuken, C.L. Ren, J.P. Huissoon, Image processing and classification algorithm for yeast cell morphology in a microfluidic chip. J. Biomed. Opt. **16**(6), 066008 (2011). https://doi.org/10.1117/1.3589100

Y. Zhang, M. Ouyang, A. Ray, T. Liu, J. Kong, B. Bai, D. Kim, A. Guziak, Y. Luo, A. Feizi, K. Tsai, Z. Duan, X. Liu, D. Kim, C. Cheung, S. Yalcin, H.C. Koydemir, O.B. Garner, D. Di Carlo, A. Ozcan, Computational cytometer based on magnetically modulated coherent imaging and deep learning. Light. Sci. Appl. **8**, 2047–7538 (2019). https://doi.org/10.1038/s41377-019-0203-5

Y. Zhao, D. Chen, Y. Luo, H. Li, B. Deng, S.B. Huang, T.K. Chiu, M.H. Wu, R. Long, H. Hu, X. Zhao, W. Yue, J. Wang, J. Chen, A microfluidic system for cell type classification based on cellular size-independent electrical properties. Lab. Chip. **13**, 2272–2277 (2013). https://doi.org/10.1039/c3lc41361f

Y. Zhao, K. Wang, D. Chen, B. Fan, Y. Xu, Y. Ye, J. Wang, J. Chen, C. Huang, Development of microfluidic impedance cytometry enabling the quantification of specific membrane capacitance and cytoplasm conductivity from 100,000 single cells. Biosens. Bioelectron. **111**, 138–143 (2018). https://doi.org/10.1016/j.bios.2018.04.015

Y. Zheng, E. Shojaei-Baghini, A. Azad, C. Wang, Y. Sun, High-throughput biophysical measurement of human red blood cells. Lab. Chip. **12**, 2560–2567 (2012). https://doi.org/10.1039/c2lc21210b

Y. Zheng, E. Shojaei-Baghini, C. Wang, Y. Sun, Microfluidic characterization of specific membrane capacitance and cytoplasm conductivity of single cells. Biosens. Bioelectron. **42**, 496–502 (2013). https://doi.org/10.1016/j.bios.2012.10.081

W. Zhiqiang, L. Jun, A review of object detection based on convolutional neural network, in *Chinese Control Conf. CCC*. (IEEE Computer Society, 2017), pp. 11104–11109. https://doi.org/10.23919/ChiCC.2017.8029130

H. Zhu, M. Zhou, J. Lee, R. Nidetz, K. Kurabayashi, X. Fan, Low-Power Miniaturized Helium Dielectric Barrier Discharge Photoionization Detectors for Highly Sensitive Vapor Detection. Anal. Chem. **88**, 8780–8786 (2016). https://doi.org/10.1021/acs.analchem.6b02180

Y. Zuo, S. Chakrabartty, Z. Muhammad-Tahir, S. Pal, E.C. Alocilja, Spatio-temporal processing for multichannel biosensors using support vector machines. IEEE. Sens. J. **6**, 1644–1650 (2006). https://doi.org/10.1109/JSEN.2006.884445