Electronic Supplementary Material (ESI) for Lab on a Chip.

Real-time fluorescence imaging flow cytometry enabled by motion deblurring and deep learning algorithm

Yiming Wang,‡[ab] Ziwei Huang,‡[ab] Xiaojie Wang,[ab] Fengrui Yang,[c] Xuebiao Yao,[c] Tingrui Pan,[ad] Baoqing Li*[ab] and Jiaru Chu[ab]

a. Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, Hefei 230027, China.

E-mail: bqli@ustc.edu.cn

b. Key Laboratory of Precision Scientific Instrumentation of Anhui Higher Education Institutes, University of Science and Technology of China, Hefei 230027, China.

c. MOE Key Laboratory for Membraneless Organelles and Cellular Dynamics, Hefei National Center for Cross-disciplinary Sciences, University of Science and Technology of China School of Life Sciences, Hefei 230026, China.

d. Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou 215123, China.

‡These authors contributed equally to this work.

1. The purpose of sheath flow

The sheath flow is used to focus the cells at the middle of the streamline, and it can prevent cells from approaching the wall of the microchannel. To explain the necessity of sheath flow, we presented an analysis of the flow conditions with and without sheath flow in our microfluidic system, as shown in Fig. S1. In this study, the Reynolds number is below 1. Therefore, the flow of the cell sample is laminar. Under this condition, the streamwise average velocity profiles across the microchannel width follow a parabola[1], as shown in Fig. S1a. Here, the aspect ratio ε is the ratio of the height (h = 30 μm) to the width (w = 250 μm) of the microchannel. U is the normalized speed. y is the distance from the center of the microchannel.

Without the sheath flow, the flow state of cells is shown in Fig. S1b. When cells approach the walls, the velocity near the wall approaches zero due to laminar flow, while the velocity of cells on the side far from the wall is relatively high. This creates shear forces that cause the cells to rotate, which can adversely affect image quality.

However, with the sheath flow, the cells are focused in the middle of the streamline, as shown in Fig. S1c. Therefore, the cells should be focused on the microchannel to avoid lateral displacement. In this study, ε is equal to 0.12. As shown in Figure R3a, at y/w = ±0.324, the flow speed is 0.99 times the maximum speed, allowing the cells to flow forward in a straight line. The maximum width of the cell sample is calculated as 157 μm (0.628 × 250), ensuring proper cell focusing and detection.

Meanwhile, the length of a single HeLa cell is about 20 μm and the length of a HeLa cell cluster is about 40 μm. Therefore, the width of the cell sample should be larger than 40 μm to enable cells to flow into the outlet. We can conclude that:

$$40\ \mu m \leq W_{sample} \leq 157\ \mu m \tag{S1}$$

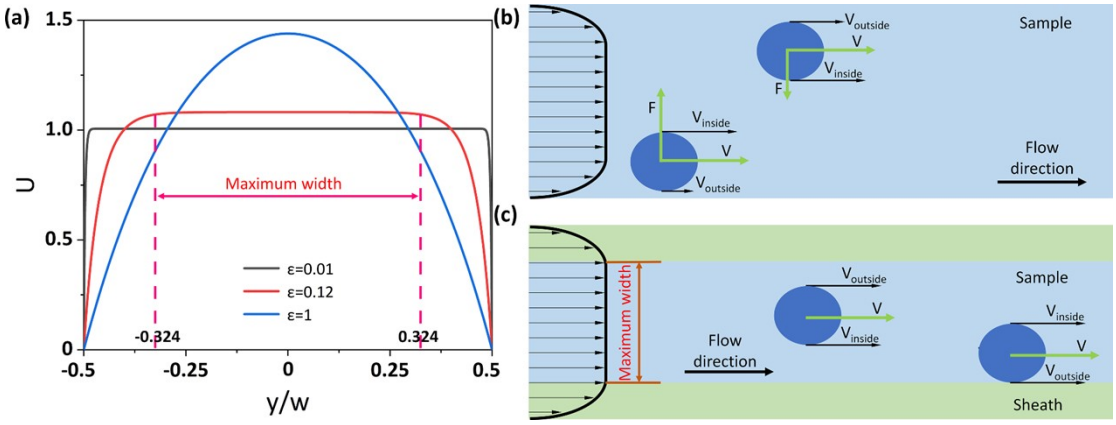

Fig. S1. The purpose of sheath flow. (a) Velocity distribution across the microchannel width with different aspect ratios. ε is equal to 0.12 in this study. (b) The flow condition of cells without sheath flow. Cells located near the wall experience shear forces and undergo rotation. (c) The flow condition of cells with sheath flow. Sheath flow induces rigid translational motion of the cells without any significant displacement in the transverse direction.

2. Fabrication and design of the microfluidic chip

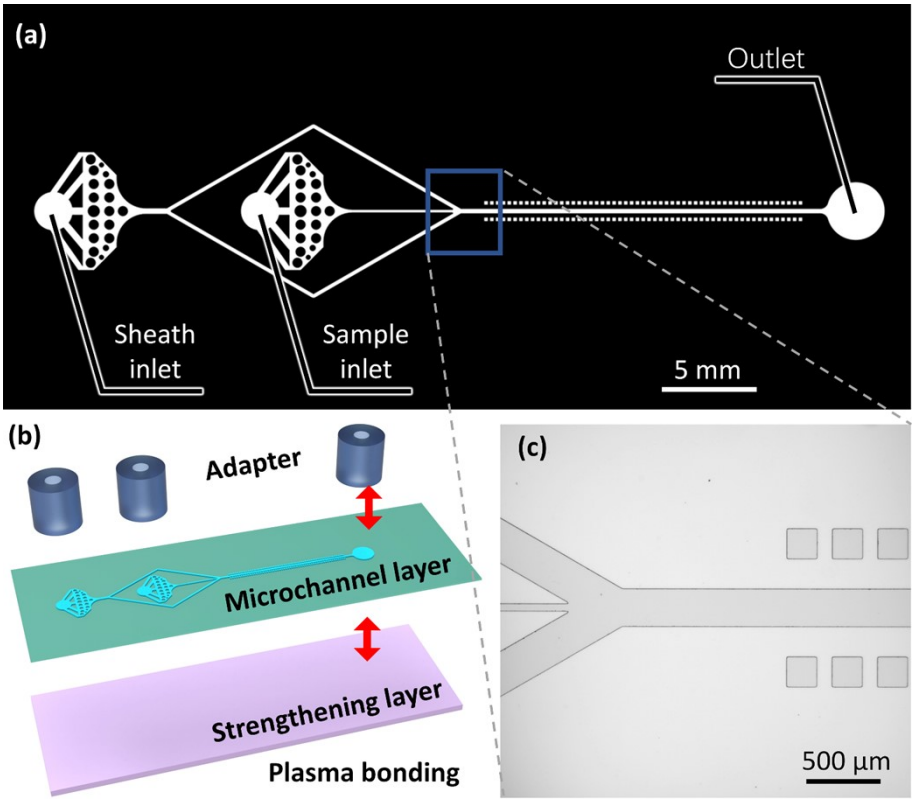

Fig. S2. Fabrication and design of the microfluidic chip. (a) CAD design of the microchannel of the microfluidic chip. (b) Schematic illustration showing the structure of the microfluidic chip. (c) A magnified view of the flow focusing and detection region.

3. An edge detection algorithm to remove incomplete cell images

It is possible for an image to not contain a complete cell due to cells entering and exiting the detection region during the exposure process, as shown in Fig. S3a. To address this issue, we developed an algorithm to detect whether cells are complete. First, the binary image is generated using threshold segmentation, as shown in Fig. S3b. We then search for the leftmost pixel in the binary image and determine if it is equal to 1. If any pixels in the leftmost column have a value of 1, we consider the cell to be incomplete and abandon this image. The same algorithm is applied to detect incomplete cells on the right side of the images. Meanwhile, we would like to clarify that we also set the size of the detection region to ensure complete cell detection. The length of detection region L should meet equation (S2).

$$L = 2 \cdot V \cdot T_{FPS} \tag{S2}$$

Where V is the flow speed of cells, T_{FPS} is the time interval between two continuous frames.

This ensures that each cell can be completely detected at least once. In this condition, even if the cell cannot be fully imaged in a previous frame, as shown in Fig. S3d, it can be fully imaged in the next frame, as shown in Fig. S3e.

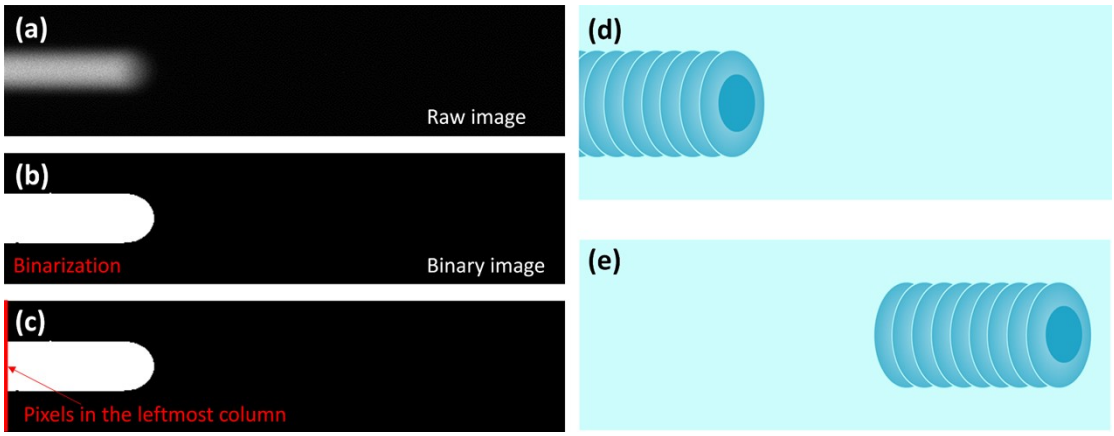

Fig. S3. Approach to obtain the complete images of cells. (a) Raw image does not contain the complete cell. (b) A binarization process is applied to the raw image. (c) The presence of a pixel value equal to 1 in the leftmost column is used to determine whether the cell is fully present in the image. (d) The cell cannot be fully imaged in the previous frame. (e) The cell can be fully imaged in the next frame.

4. Rigid translational motion of cells

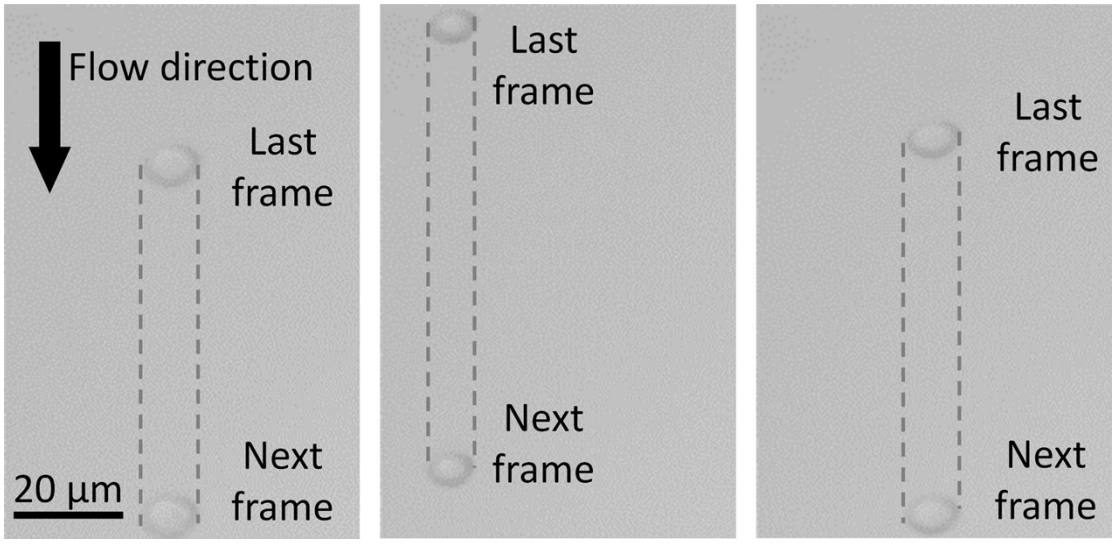

Fig. S4. Bright-field images depict the flowing of single cells in two consecutive frames, which clearly demonstrate that the cells undergo rigid translational motion without any movement in the transverse direction. The frames were captured using sCMOS camera, and each image is a result of merging two consecutive frames.

5. Feature extraction algorithm

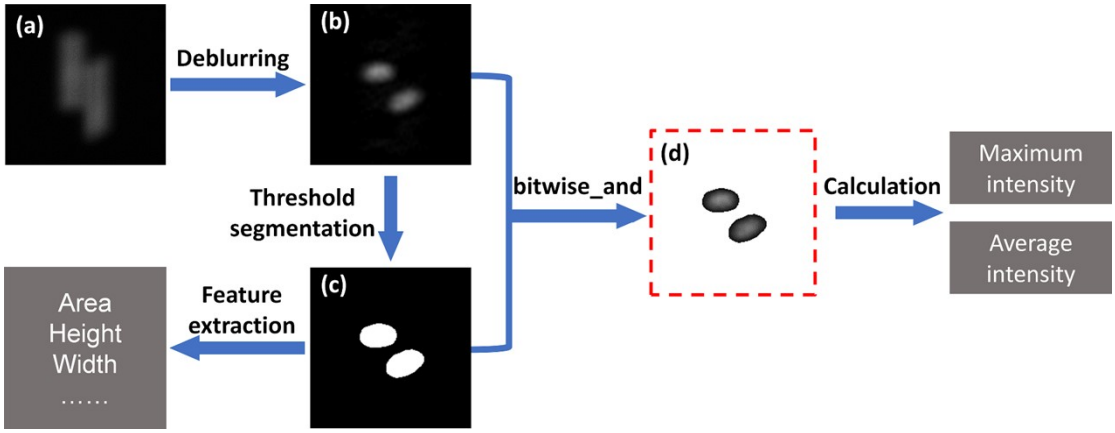

Fig. S5. Feature extraction algorithm. (a) A raw blurred image of the HeLa cell nucleus during the mitosis stage is used as an example. (b) The optimal deblurred image is reconstructed from the motion deblurring algorithm. (c) Binary image is calculated by threshold segmentation for calculating the size and morphology of cells. (d) For calculating intensity of cells, the ROI (region of interest) image is obtained by performing a bitwise_and operation on deblurred image and binary image. Bitwise_and is a standard bitwise function in OpcnCV. The maximum intensity and average intensity are calculated from the ROI image. Specifically, maximum intensity is the maximum grayscale value of pixels within the ROI image, and average intensity is the average grayscale value of pixels within the ROI image.

6. Selection of sample concentration

In this study, we used a lower concentration (5×10^5 cells/ml) for the convenience of the experiment. We choose this concentration to ensure that the cells flow through in order and do not overlap. With an average interval of 100 μm between consecutive cells, the maximum concentration should be calculated as:

$$n = \frac{1}{L_{space} * w * h} \tag{S3}$$

Where w is the width of cell sample. h is the height of microchannel and is equal to 30 μm. L_{space} is the average length of cell spacing. When the w is equal to 157 μm and L_{space} is equal to 100 μm, the maximum concentration is equal to 2.1×10^6 cells/ml.

Meanwhile, our system has the advantage of accommodating a large detection region, enabling the detection of multiple cells simultaneously. As discussed earlier, the detection region's width is optimized to 157 μm, allowing the observation of over 10 cells simultaneously in one frame. To demonstrate our system's imaging ability at high concentrations, we conducted an experiment using a cell concentration of 1×10^7 cells/ml, as shown in Fig. S6. As the cell concentration increases, cells tend to flow alongside each other. Nevertheless, each cell was successfully extracted from the image and reconstructed using the deblurring algorithm. Hence, we are confident that our system's ability is not limited by the concentration parameter.

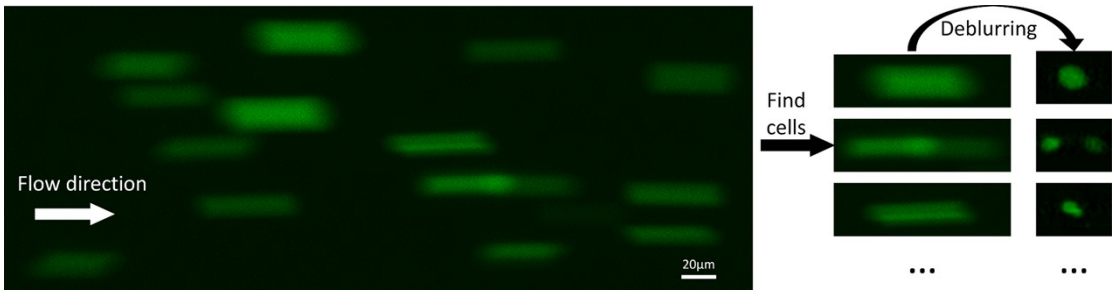

Fig. S6. Parallel detection of cells at high concentration.

7. The applicability of diverse particle and cell types

To demonstrate the versatility of our method, we applied it to different types of particles or cells, as shown in Fig. S7. 15 μm PMMA particles, Jurkat, and 231 cells were imaged by the imaging flow cytometry, and the deblurred images were reconstructed from the raw images. We are confident that our system is suitable for various biological samples that can be detected by a fluorescence microscope.

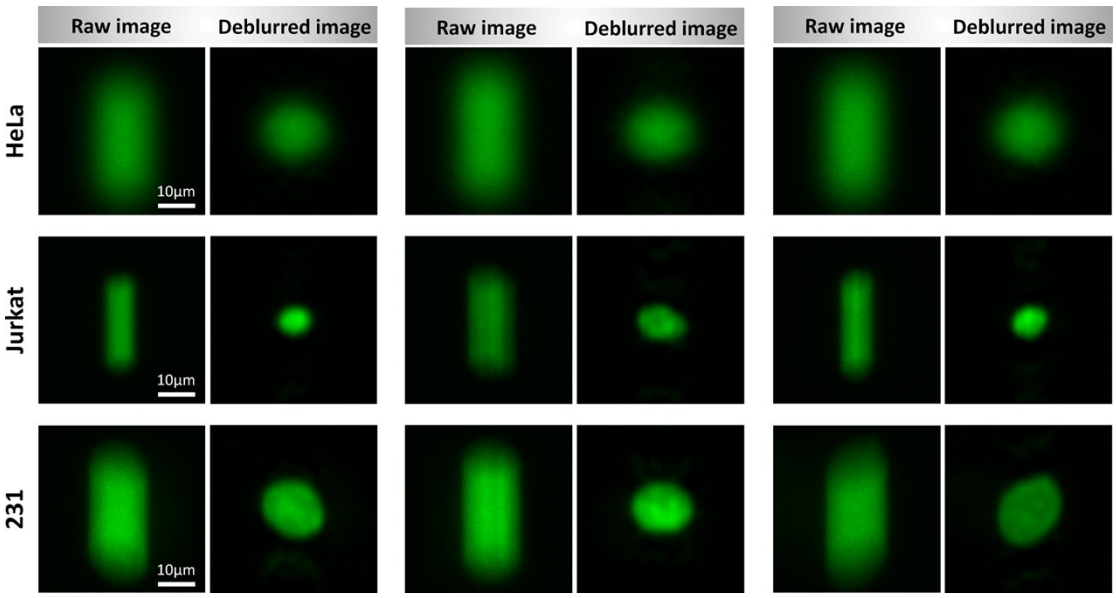

Fig. S7. Fluorescence images of diverse particle and cell types obtained by VIFFI flow cytometry. (a) Fluorescence images of 15 μm PMMA particles. (b) Fluorescence images of Jurkat cells and (c) 231 cells. The nucleus was stained by SYTO16.

To further demonstrate the imaging capability of the system, we performed imaging of a 1 μm fluorescent PS particle. For this experiment, the 100× objective was applied to improve resolution. Of course, higher magnification will sacrifice throughput because it can lead to a decrease in the flow speed of cells according to equation S6 (20 should be changed to 100 in the equation). As shown in Fig. S8, 1 μm fluorescent PS particles flowed in the microchannel at 4.6 mm s^{-1} and were imaged. The deblurred image was reconstructed from the raw images and has high roundness.

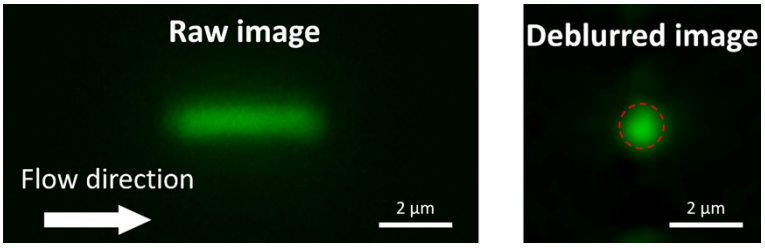

Fig. S8. Imaging of the 1 μm polystyrene particle. The motion-blur image was acquired by microscope with a 100× objective.

8. Evaluation of image quality

The image quality was evaluated using the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Due to the same size and morphology of fluorescent particles, a stationary particle image was used as the ground truth image, as shown in Fig. S9a. Therefore, we compared the deblurred particle images to the stationary particle images. The PSNR and SSIM were calculated by a C++ programmer with the image library OpenCV to qualitatively evaluate the image quality. The PSNR and SSIM of our method are 30.04 and 0.732, respectively. Furthermore, we compared our method with state-of-the-art deblurring techniques, including inverse filter, Wiener filter[2], and CLSF[3] (constrained least squares filter). Representative images generated by each algorithm are shown in Fig. S9a. The PSRN and SSIM values of the four methods are shown in Fig. S9b and Fig. S9c, respectively. These results indicate that our method is superior to other methods. However, the deblurred image generated by our method still has some problems due to the out-of-focus problem of the particles. The image quality of our method can be improved by using a three-dimension sheath flow, which is able to focus the particles into the focal plane of the microscope.

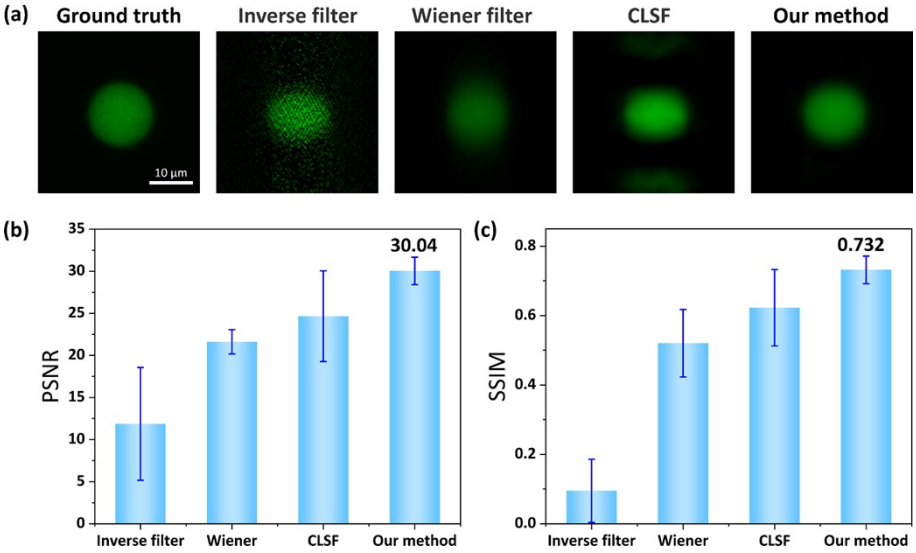

Fig. S9. Evaluation of image quality. (a) A stationary particle image was used as the ground truth image for the assessment of image quality. Deblurred images were reconstructed from the blurred images using different deblurring techniques, including inverse filter/Wiener filter/CLSF (constrained least squares filter) and our proposed method. (b) The deblurred images obtained from different deblurring techniques were compared with the ground-truth image using the Peak signal-to-noise ratio (PSNR). (C) The structural similarity index (SSIM) metric was also used to compare the deblurred images obtained from different deblurring techniques with the ground-truth image. The evaluation was performed on a dataset of 500 images. The result of each approach is calculated by 500 images.

9. Classification of different cells types

We conducted a classification study involving HeLa cells, 231 cells, and Jurkat cells. Three types of cells were individually loaded into the microfluidic chip and were imaged by the imaging flow cytometry, as shown in Fig. S10a. 10000 images of each cell type were acquired and used for training a three-classification ResNet model. The pre-trained ResNet model was called by C++ for cell type classification. The scatterplot of t-SNE visualization demonstrates that the three cell types can be clearly distinguished, as shown in Fig. S10b. Meanwhile, the confusion matrix of classification accuracy is shown in Fig. S10c. The classification accuracy was 99.6% for 231, 99.9% for HeLa, and 99.1% for Jurkat.

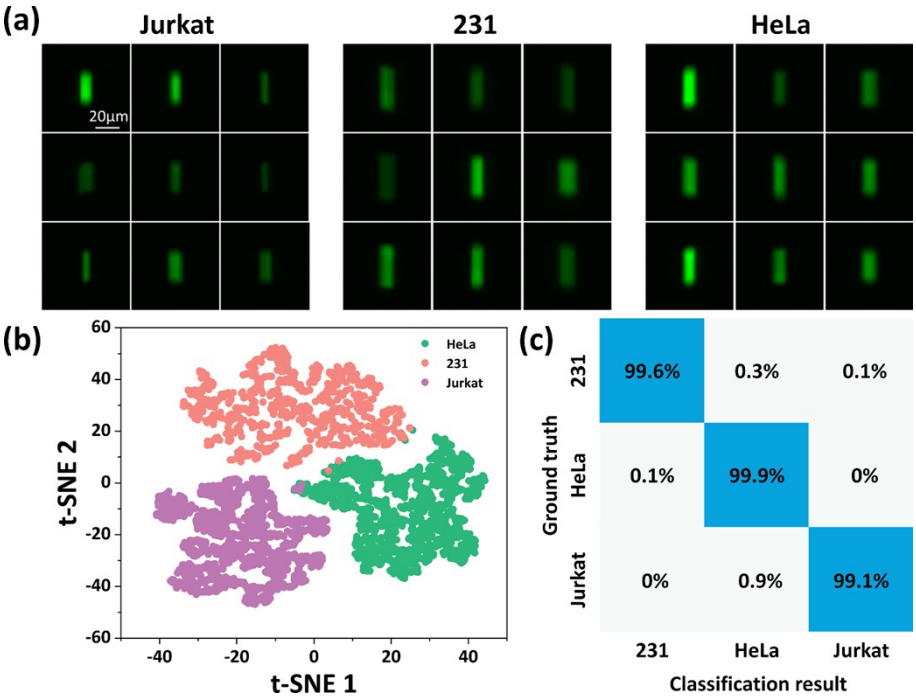

Fig. S10. Classification of Jurkat cells, 231 cells, and HeLa cells using imaging flow cytometry and the ResNet algorithm. (a) Raw blurred images of three types of cells (n = 10000 for each group) were obtained using the imaging flow cytometry, with the nucleus stained by SYTO16. (b) t-SNE visualization of the classification of the three cell types using the ResNet algorithm (n = 2000 for each group), which demonstrates that the cell types can be clearly distinguished. (c) Confusion matrix showing the classification accuracy of the three cell types.

10. Feature description of HeLa cells

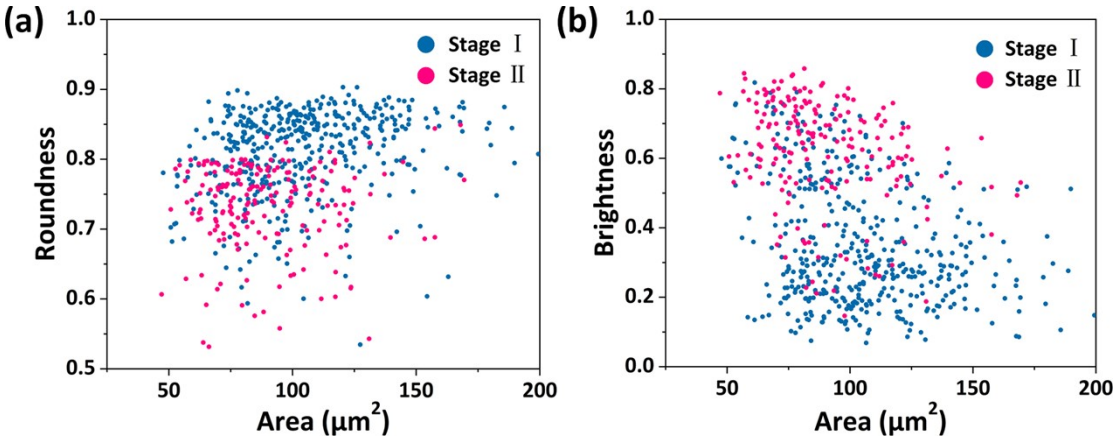

Fig. S11. Feature description of Hela cells in stage I and stage II. (a) Distinguish stage I and stage II according to area and roundness. (b) Distinguish stage I and stage II according to area and brightness.

11. Reasons for cell classification errors

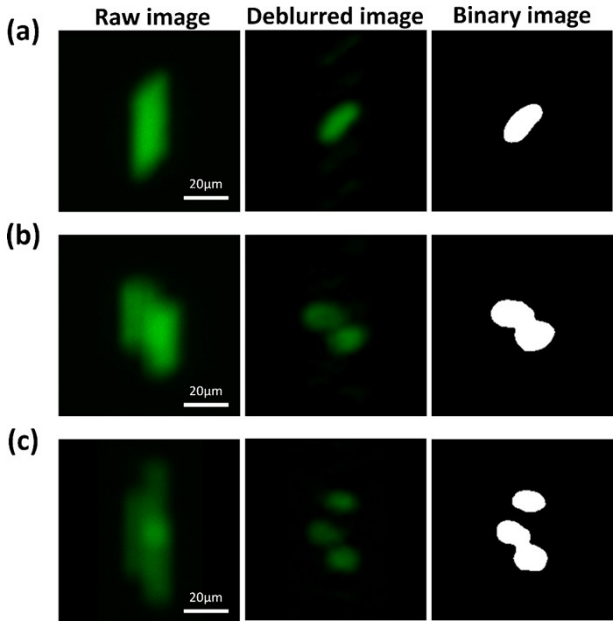

Fig. S12. Reasons for cell classification errors. (a) 14.6% of cells in stage II were misclassified as stage I due to lower brightness. (b) 15.2% of cells in stage III were misclassified as stage I due to two nuclei being too close, resulting in a binary image with a single connected domain. (c) 26.2% of cell clusters were misclassified as stage III due to two nuclei being too close, resulting in a binary image with two connected domains. Each situation is illustrated with one example.

12. Preparation of training dataset

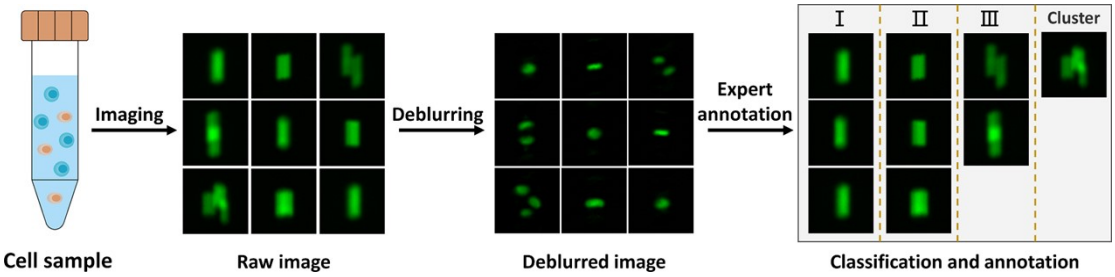

Fig. S13. The preparation of training dataset of the residual network. It contains three main steps: firstly, nuclear of Hela cells are stained by the SYTO16 dye and fluorescent images of single cells are captured by the sCMOS camera while flowing at the microfluidic chip. Secondly, the deblurred images are reconstructed from the raw blurred image by the deblurring algorithm and cell types can be easily distinguished. Thirdly, human experts classify the raw blurred images according to the deblurred image and the dataset is established. Lastly, the blurred images of cells with class labels are used to train the residual network.

13. The necessity of imaging deblurring

The raw images were labeled by human experts without the deblurring step. Then, the blurred images of cells with annotation were used to train the ResNet model. The pre-trained ResNet model was called by C++ in real-time for cell classification, and the result of classification accuracy is shown in Fig. S14a. Compared with the original method employed in main text, the accuracy has significantly decreased. There is a lot of overlap in the scatterplot of t-SNE visualization, as shown in Fig. S14b. This result indicates that the process of image deblurring is necessary for training deep learning models.

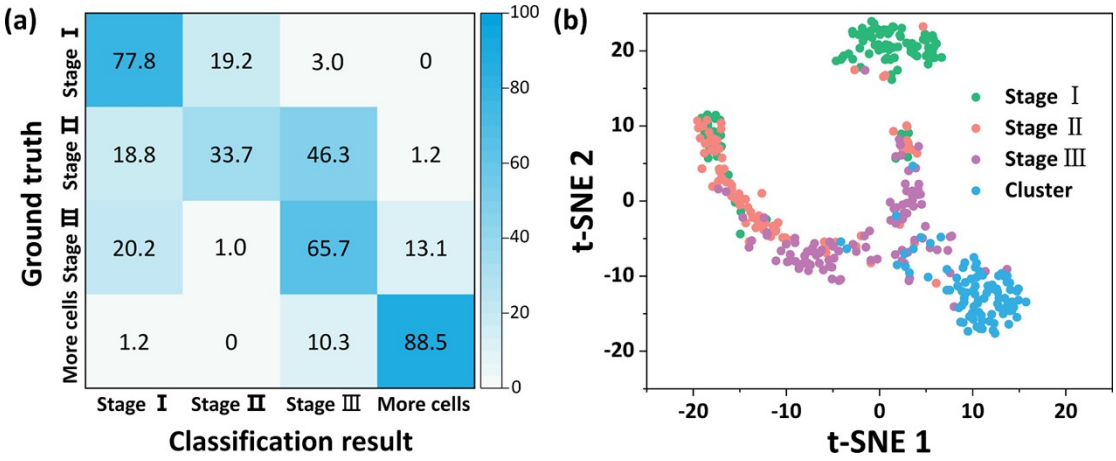

Fig. S14. Classification of HeLa cells using the ResNet algorithm trained by the dataset without deblurring step. (a) Confusion matrix showing the classification accuracy of HeLa cells at different cell cycle stages. (b) t-SNE visualization of HeLa cells at different cell cycle stages.

13. Flow speed and throughput

The throughput of imaging flow cytometry is determined by flow speed. Based on this criterion, we can estimate the flow speed using the formula:

$$V = \frac{P_{blur} * L_P}{20T} \tag{S4}$$

Where V represents the flow speed of cells. P_{blur} is the pixel number of the optimal motion blur. In order to obtain completely blurred images of cells, the length of the blurred image must be smaller than the size of the camera sensor. L_P is the size of each pixel, and this value is 6.5 μm. The magnification of the objective lens is 20 times. T is the exposure time of the camera. Specifically, long exposure times will reduce the noise but increase motion blur, and short exposure times minimize motion blur at the expense of increased noise and reduced photon collection. We choose the appropriate exposure time based on the fluorescent brightness and flow rate of the samples to achieve the optimal image quality.

In order to determine the upper limit of flow speed for this system, we conducted calibration experiments using standard 15 μm particles. A fluorescence particle was imaged at 96.9 mm s^{-1} with a 1 ms exposure time, as shown in Fig. S15. As the flow speed increases, the blur distance also increases. However, the search-based motion deblurring algorithm was able to effectively remove the motion blur, resulting in high roundness of the deblurred image. It should be noted that the maximum flow speed is limited by the camera in our system. Specifically, the available size of the camera sensor in our system is only 400 pixels, which limits the theoretical maximum flow speed of cells to 0.13 m s^{-1}. At an average cell spacing of 100 μm, this corresponds to a cell throughput of 1300 cells s^{-1}.

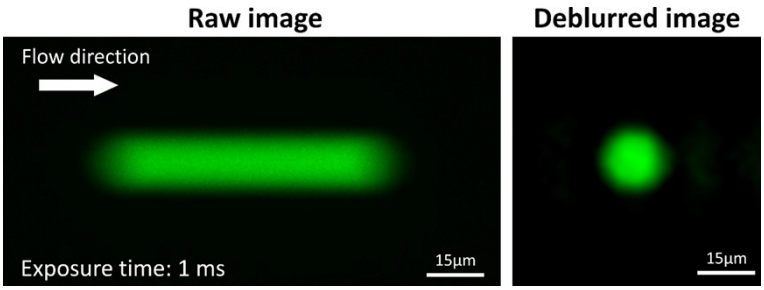

Fig. S15. Imaging of the 15 μm particle at 96.9 mm s^{-1} with a 1 ms exposure time. The motion-blur image is shown on the left and the deblurred image is shown on the right.

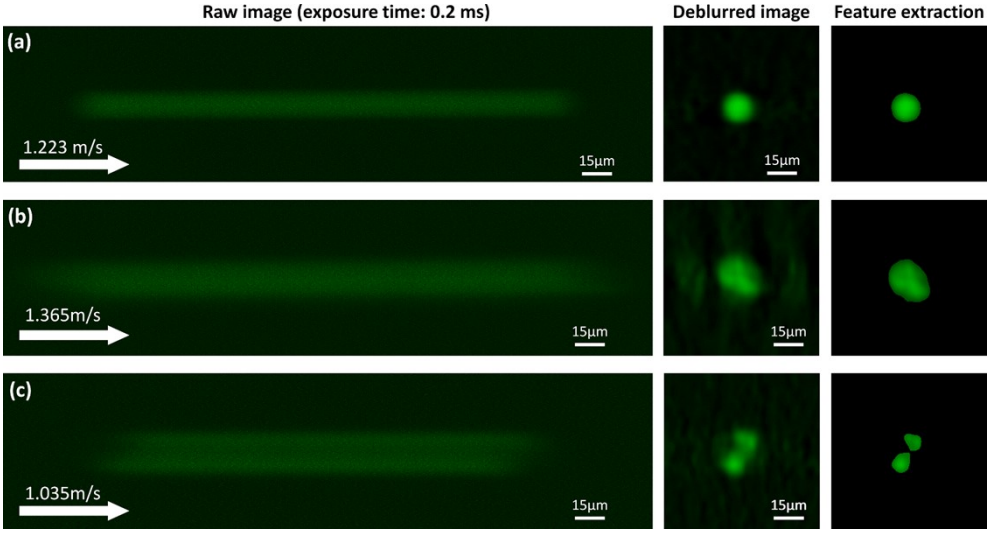

Fig. S16. High throughput imaging of particles and cells using a sCMOS camera (Teledyne Kinetix) with 3200 pixels in length. (a) Imaging of the 15 μm particle at 1.223 mm s^{-1} with a 0.2 ms exposure time. (b) Imaging of a 231 cell before the mitosis stage at 1.365 m s^{-1} with a 0.2 ms exposure time. (c) Imaging of a 231 cell under mitosis stage at 1.035 m s^{-1} with a 0.2 ms exposure time. The motion-blur image is shown on the left, the deblurred image is shown in the middle, and the feature extract result is shown on the right.

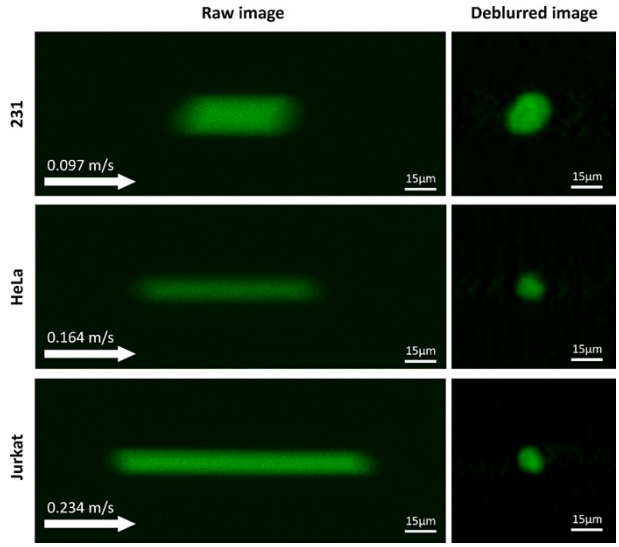

Fig. S17. Raw and deblurred images of 231, HeLa cells, and Jurkat cells were obtained at high flow speed with a 0.5 ms exposure time, demonstrating the versatility of our method.

14. The imaging sensitivity

The sensitivity is evaluated by the signal-to-noise ratio (SNR). The Signal-to-Noise Ratio is calculated using the following equation:

$$SNR = \frac{FS}{\sqrt{FS + \sigma^2}} \tag{S5}$$

Here, σ represents the readout noise of the image sensor. The fluorescence signal (FS) is the number of electrons and is defined by the equation:

$$FS = PS^{-1}TNA\varphi\eta_{IFC}q_s \tag{S6}$$

Where P is the power of the excitation beam, S is the cross-section of the excitation beam, T is the time a molecule spends in the excitation beam, N is the number of fluorescent molecules in a single-pixel area, A is the absorption cross-section of the fluorescent molecule, φ is the fluorescence quantum yield of the molecule, η_{IFC} is the photon collection efficiency of the imaging platform, q_S is the quantum efficiency of the camera sensor.

Therefore, the sensitivity of our system is influenced by two key parameters: optical power (optical power density) and exposure time. The remaining parameters are determined by the characteristics of the fluorescent sample and camera.

In our system, we utilize a mercury lamp as the light source, coupled with a wide-field microscope. There is no denying that the power and power density of the excitation beam of the mercury lamp are lower compared to a laser light source. The signal intensity is one order of magnitude lower than when using light sheet illumination.

Another important parameter that affects sensitivity is the exposure time. Specifically, long exposure times will reduce the noise but increase motion blur, and short exposure times minimize motion blur at the expense of increased noise and reduced photon collection. In our system, **we apply a long exposure time approach to improve the sensitivity and have developed a search-based motion deblurring algorithm to remove motion blur.** The maximum exposure time of our system can be calculated using the formula:

$$T = \frac{P_{max} * L_P}{20V} = \frac{3200 \times 6.5 \times 10^{-6}}{20 * V} = \frac{1.04 \times 10^{-3}}{V} \tag{S7}$$

Where V represents the flow speed of cells. P_{max} is the maximum pixel number in length of the camera sensor, which is equal to 3200 pixels for the Teledyne Kinetix sCMOS camera. L_P is the size of each pixel, and this value is 6.5 μm. The magnification of the objective lens is 20 times.

The TDI-based IFC suffers from the large noise of CCD, which limits its detection sensitivity according to equation (S5). PMT-based IFC achieves both high throughput and high spatial resolution but comes at the expense of sensitivity due to low exposure time. sCMOS-based IFC is a good way to realize single-cell imaging with high-sensitivity, which has low readout noise and high quantum efficiency. The temporally coded excitation and our approaches use long exposure time to realize high-sensitivity. The stroboscopic illumination-based IFC uses light sheet illumination to improve sensitivity by concentrating photons to a small volume, which can realize sub-cellular localization detection of features close to the diffraction limit.

15. Constrained least squares filter algorithm and Tenengrad gradient function

The constrained least squares filter algorithm is a popular method used in image reconstruction to reduce noise and enhance image quality. The algorithm is described as follows:

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma|P(u,v)|^2}\right] G(u,v) \tag{S8}$$

Where $\hat{F}$(u,v) represents the Fourier transform of the deblurred image. H(u,v) represents the Fourier transform of the motion blur kernel. H*(u, v) is the complex conjugate of H (u, v). G(u,v) represents the Fourier transform of the blurred image. P(u,v) is the Fourier transform of the following functions:

$$p(u,v) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{S9}$$

γ is a constant, which is calculated by the mean value and variance of background images.

The Tenengrad gradient function is a widely used image definition evaluation function that is based on the gradient of the image. It is calculated by taking the sum of the squares of the gradient values in both the horizontal and vertical directions. The function uses the Sobel operator to extract the gradient values, which is a widely used edge detection filter in image processing. The Tenengrad gradient function is defined as:

$$T(I) = \begin{cases} \sum_x \sum_y [S(x,y)]^2 & G(x,y) > T \\ 0 & G(x,y) < T \end{cases} \tag{S10}$$

where T is a given threshold for edge detection, and S (x, y) represents the gradient of image at point (x, y). To be more specific, S (x, y) is obtained as follows:

$$S(x,y) = \sqrt{G_x * I(x,y) + G_y * I(x,y)} \tag{S11}$$

Gx and Gy is Sobel operator in the horizontal direction and vertical direction respectively, it is described as follow:

$$G_x = \frac{1}{4}\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad G_y = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{S12}$$

References:

1. G.-B. Lee, C.-C. Chang, S.-B. Huang and R.-J. Yang, *Journal of Micromechanics and Microengineering*, 2006, **16**, 1024-1032.

2. P. D. Sankhe, M. Patil and M. Margaret, presented in part at the Proceedings of the International Conference & Workshop on Emerging Trends in Technology, Mumbai, Maharashtra, India, 2011.

3. N. Liu, X. Zheng, H. Sun and X. Tan, Pattern Recognition Letters, 2013, 34, 124-130.