Title

Cameraless Image Flow Cytometry and Image-Activated Cell Sorting Using Artificial Intelligence

Permalink

https://escholarship.org/uc/item/8df3b31x

Author

Tang, Rui

Publication Date

2022

UNIVERSITY OF CALIFORNIA SAN DIEGO

Cameraless Image Flow Cytometry and Image-Activated Cell Sorting Using Artificial Intelligence

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Medical Devices & Systems)

by

Rui Tang

Committee in charge:

Professor Yu-Hwa Lo, Chair
Professor Drew Hall, Co-Chair
Professor Manmohan Chandraker
Professor Duygu Kuzum
Professor Kun Zhang

2022

The dissertation of Rui Tang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

To my fiancée Qian and our parents.

EPIGRAPH

Look deep into nature, and then you will understand everything better.

-Albert Einstein

It's not what you look at that matters, it's what you see.

-Henry David Thoreau

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere thanks to my advisor, Prof. Yu-Hwa Lo, for his immeasurable support and guidance during my Ph.D. study over the past five years. He introduced me to an exciting research area of flow cytometers and cell sorting instruments. During the past five years, his patience, vast and deep knowledge of numerous fields, and attention to fundamentals have always inspired me. His training and continual encouragement have allowed me to work towards the success of research projects. His enthusiasm, persistence, and optimism have significantly influenced my character both in my professional and personal life.

I would also like to thank my committee members, Prof. Drew Hall, Prof. Duygu Kuzum, Prof. Kun Zhang, and Prof. Manmohan Chandraker, for their invaluable time and feedback. I am grateful that they shared their knowledge and experience with me and kept me on the right track during the pursuit of my Ph.D. degree. I want to express my special thanks to Prof. Cheolhong An for his insightful input in my Ph.D. research.

I also wish to thank the impressive former members of Prof. Lo's research group, from whom I have learned a lot. Specifically, Dr. Tony Yen gave me a great first impression of working on an interdisciplinary research project. Dr. Yuanyuan Han had a significant influence on bringing me to the IFC and IACS research fields. Dr. Sung Hwan Cho inspired the technology transfer from academia to industry and guided me in technology development. Also, I am very thankful to Dr. Yu-Jui (Roger) Chiu, Dr. Wei Cai, Violet Castor, Dr. Brian Lewis, Dr. Alex Ce Zhang, Dr. Yi Gu, and Dr. Chi-Yang Tseng for the help that you offered after I became a Lo research group member. I appreciate the fantastic students Prof. Lo has gathered in his group. Working daily with such people has made my time in graduate school such a memorable and worthwhile experience. I wish to thank Zunming Zhang, Xinyu Chen, Ramkumar Subramanian, Edward Wang, Lauren Waller, and Zhilin Guo.

Furthermore, I would like to express my gratitude and appreciation for the support and help of the colleges, Dr. Jose Morachis, Dr. Willian Alynick, Dr. Mike Benchimol, Chris Neary, Lin Xia, Bien Gutierrez, Ivan Gagne, Adonary Munoz, Dr. Nicole Jagnandan, and Korina Eribez at NanoCellect Biomedical, Inc. I have learned how to be a professional engineer in collaboration with NanoCellect.

Next, I would like to thank my friends for sharing their happiness with me.

Most important of all, I owe thanks to my fiancée, Qian Liao, for her unrequited love and support. You helped me to keep things in perspective. I greatly value her encouragement and sincerely appreciate her belief in me.

Last but not least, I am supremely grateful to my family for their unswerving love and support. You have been the constant thread throughout my entire life, especially to my mom, Hong Deng, and my dad, Hong Tang. I salute you all for showing faith in me and giving me the liberty to choose what I desire. I am honored to be your child in my life. Thank you for always being there.

Chapter 1, in part, is a reprint of the material as it appears in Optica 2019, Y. Han, R. Tang, Y. Gu, A. Zhang, W. Cai, V. Castor, S. Cho, W. Alaynick, and Y.-H. Lo. "Cameraless high-throughput 3D imaging flow cytometry". The dissertation author was the second author of this paper.

Chapter 1, in part, is a reprint of the material as it appears in High-Speed Biomedical Imaging and Spectroscopy V, Proceeding of SPIE 2020, R. Tang, X. Chen, Z. Zhang, L. Waller,

J. Chen, Y. Gu, Y. Han, C. Lee, I. Gagne, A. Zhang, S. Cho and Y.-H. Lo. "2D image-guided cell sorter and 3D imaging flow cytometer". The dissertation author was the first author of this paper.

Chapter 1, in part, is a reprint of the material as it appears in Proceedings of the National Academy of Sciences 2022, Z. Zhang, R. Tang, X. Chen, L. Waller, A. Kau, A. Fung, B. Gutierrez, C. An, S. Cho, L. Shi, and Y.-H. Lo. “A high-throughput technique to map cell images to cell positions using a 3D imaging flow cytometer”. The dissertation author was the second author of this paper.

Chapter 2, in part, is a reprint of the material as it appears in APL Photonics 2020, R. Tang, Z. Zhang, X. Chen, L. Waller, A. Zhang, J. Chen, Y. Han, C. An, S. Cho, and Y.-H. Lo. "3D side-scattering imaging flow cytometer and convolutional neural network for label-free cell analysis". The dissertation author was the first author of this paper.

Chapter 2, in part, is a reprint of the material as it is under peer review in Scientific Reports 2022, R. Subramanian, R. Tang (co-first), Z. Zhang, V. Joshi, J. Miner, Y.-H. Lo. “Multimodal NASH prognosis using 3D imaging flow cytometry and artificial intelligence to characterize liver cells”. The dissertation author was the co-first author of this paper.

Chapter 3 has been submitted for publication of the material, R. Tang, L. Xia, B. Gutierrez, I. Gagne, A. Munoz, K. Eribez, N. Jagnandan, X. Chen, Z. Zhang, L. Waller, W. Alaynick, S. Cho, C. An, Y.-H. Lo. “Low-latency label-free image-activated cell sorting using fast deep learning and AI inferencing”. The dissertation author is the first author of this paper.

VITA

2016	Bachelor of Engineering in Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong, China
2016	Bachelor of Business Administration in Management, The Hong Kong Polytechnic University, Hong Kong, China
2018	Master of Science in Electrical Engineering (Medical Devices & Systems), University of California San Diego, San Diego, USA
2022	Doctor of Philosophy in Electrical Engineering (Medical Devices & Systems), University of California San Diego, San Diego, USA

PUBLICATIONS

- **Tang, R.**, Lin, X., Gutierrez, B., Gagne, I., Munoz, A., Eribez, K., Jagnandan, N., Chen, X., Zhang, Z., Waller, L., Alaynick, W., Cho, S., An, C., & Lo, Y.-H., (submitted). Low-latency label-free image-activated cell sorting using fast deep learning and AI inferencing. *Nature Communications*.
- Subramanian, R., **Tang, R.(co-first)**, Zhang, Z., Josh, V., Miner, J., & Lo, Y.-H. (peer review). Multimodal NASH prognosis using 3D imaging flow cytometry and artificial intelligence to characterize liver cells. *Scientific Reports.*
- Zhang, Z., **Tang, R.**, Chen, X., Waller, L., Kau, A., Fung, A. A., Gutierrez, B., An, C., Cho, S., Shi, L., & Lo, Y.-H. (2022). A high-throughput technique to map cell images to cell positions using a 3D imaging flow cytometer. *Proceedings of the National Academy of Sciences*, 119(8), e2118068119.

- Waller, L., Guo, Z., **Tang, R.**, Zhang, Z., Wang, E., Yasuhara-Bell, J., Laurent, L., & Lo, Y.-H. (2022). High Sensitivity, Rapid Detection of Virus in High Traffic Environments. *Frontiers in bioengineering and biotechnology*, *10*.
- Wang, E., Guo, Z., **Tang, R.**, & Lo, Y.-H. (2021). Using airflow-driven, evaporative gradients to improve sensitivity and fluid control in colorimetric paper-based assays. *Lab on a Chip*, *21*(21), 4249-4261.
- Chen, X., Waller, L., Chen, J., **Tang, R.**, Zhang, Z., Gagne, I., Gutierrez, B., Cho, S., Tseng, C.-Y., Lian, I. Y., & Lo, Y.-H. (2021). Label-free image-encoded microfluidic cell sorter with a scanning Bessel beam. *APL Photonics*, 6(7), 076101.
- **Tang, R.**, Zunming, Z., Chen, X., Waller, L., Zhang, A., Chen, J., Han, Y., An, C., Cho, S., & Lo, Y.-H. (2020). Using 3D side scattering imaging flow cytometer and convolutional neural network for label-free cell analysis, *APL Photonics*, 5(12), 126105.
- **Tang, R.**, Chen, X., Zhang, Z., Waller, L., Chen, J., Gu, Y., Han, Y., Lee, C.H., Gagne, I., Zhang, A., Cho, S., & Lo, Y.-H. (2020). 2D Image-guided cell sorter and 3D imaging flow cytometer, SPIE Photonics West, 11250-15
- Chen, X., Gu, Y., Chen, J., Lee, C., Gagne, I., **Tang, R.**, Waller, L., Zhang, Z., Zhang, A., Han, Y., Wang, W., Cho, S., & Lo, Y.-H. (2020). Image-guided cell sorting using ultrafast scanning lasers, *APL Photonics*, 5(4), 040801.
- Han, Y., **Tang, R.**, Gu, Y., Zhang, A., Cai, W., Castor, V., Cho, S., Alaynick, W., & Lo, Y.-H. (2019). Cameraless high-throughput three-dimensional imaging flow cytometry, *Optica*, 6, 1297-1304

ABSTRACT OF THE DISSERTATION

Cameraless Image Flow Cytometry and Image-Activated Cell Sorting Using Artificial Intelligence

by

Rui Tang

Doctor of Philosophy in Electrical Engineering (Medical Devices & Systems)

University of California San Diego, 2022

Professor Yu-Hwa Lo, Chair

Professor Drew Hall, Co-Chair

High-throughput cellular image processing and analysis based on imaging flow cytometry (IFC) technology can bring significant insight to biology and medicine. The ability to classify, map and isolate cells based on a high-content cellular image provides a powerful tool for biological and biomedical researchers and doctors to understand the connection between the phototype and genotype among the heterogeneous cell populations. This dissertation details the approach to conducting high-throughput cellular image analysis and low-latency real-time image processing using the IFC and artificial intelligence. As a result, we demonstrated the workflow for conducting high-throughput label-free cell study on IFC systems. In addition, we developed a low-latency image-activated activated sorting (IACS) system using artificial intelligence and machine vision,

opening a new venue for high-throughput cellular analysis and cell sorting based on machine vision and artificial intelligence.

Chapter 1 Introduction to Cameraless Imaging Flow Cytometry and Image-Activated Cell Sorting

Flow cytometer and fluorescence-activated cell sorter (FACS) are widely used biological instrument for a various biomedical and clinical study that reveals the cell-related phenotype and genotype information. To meet the increasing demand of understating the heterogeneity of cellular phenotype and relating phenotype with the genotype information, the development of cameraless imaging flow cytometry (IFC) and image-activated cell sorter (IACS) has become a competitive research field. This chapter introduces the concept of cameraless IFC and IACS. In particular, a 2D IACS and 3D IFC system will be reviewed, and the potential need to analyze high-content image information will be discussed. The scope of this dissertation will be presented at the end of this chapter.

1.1 Introduction

A central challenge of biology is to correlate the phenotype of heterogeneous individuals in a population to their genotype in order to understand whether they conform to the observed population behavior or stand out as exceptions that drive disease or become threats to health [1]–[4]. While optical microscopy is a cornerstone method for studying biological specimens' morphology and molecular composition, flow cytometry is a gold standard for quantitative high-throughput single-cell characterization in numerous biomedical applications [5], [6]. Recognizing the need to merge these two powerful platforms, several groups have proposed techniques for imaging flow cytometry (IFC) [7]. IFC simultaneously produces ensemble-averaged measurements and high-content spatial metrics from individual cells in a large population of cells without perturbation due to experiment condition change. In addition to the ability to obtain high-content cell images under a high-throughput approach, classification and isolation of different cell types among cell populations based on phenotype features can bring significant insight to biology and medicine [8], [9]. Image-activated cell sorter (IACS) combines the IFC system with cell-sorting capabilities and enables isolating cells based on high-content image features.

To facilitate high photon sensitivity and fast response of the photodetector, the so-called cameraless IFC and IACS system use photomultiplier tubes (PMTs) to replace CMOS or CCD cameras to form the image based on the temporal signal from PMTs. The cameraless design makes the system more compatible with conventional FACS and user-friendly. The cell image features can be directly extracted from the temporal waveforms or the reconstructed cell images via temporal-spatial transformation [8]. A cameraless 2D IACS system and a 3D IFC system will be introduced in the following sections.

1.2 Cameraless 2D image-activated cell sorter

A cameraless 2D image-activated cell sorter (IACS) system utilizes a high-speed linear laser scanner, coupled with cell flow, to form a 2D scanning system that surveys every single cell passing the interrogation area. The 2D scanning process encodes the cell image into a temporal signal, which can be transformed mathematically to reconstruct the 2D cell image. The system contains a dual-wavelength excitation laser and four photomultiplier tubes (PMTs) to detect fluorescent and transmission signals for each cell. The signal from each PMT detector can be constructed into a single modality 2D image, which can be overlaid to produce multi-color, multi-parameter cell images similar to the images generated by a fluorescent microscope. To perform image-guided cell sorting with single-cell resolution, we integrate a piezoelectric on-chip actuator with the microfluidic channel and adopt a hybrid signal/image processing architecture where FPGA is used for real-time image processing and CPU (or GPU) is used for image feature extraction in parallel. High sorting purity (>96%), high enrichment factor (~1200), and a decent throughput of >350 cells/s, limited by the response of the on-chip piezoelectric cell sorting actuator, have been demonstrated

The overall system architecture of the 2D image-activated cell sorter is shown in Fig. 1.2.1 (a). Each traveling cell in the microfluidic channel is illuminated by the scanning laser excitation beams, which generate temporal fluorescent and transmission signals. Photomultiplier tubes (PMTs) convert the optical signals to electronic waveforms, which are sent to a field-programmable-gate-array (FPGA) for real-time cell image reconstruction. The image features are then extracted and compared against the cell sorting criteria. If the cell falls within the gated region, the FPGA will send a voltage signal to activate the on-chip piezoelectric actuator (PZT), which bends upward or downward mechanically. The cytometry cell sorter provides two sorting modes. In the population sorting mode, the PZT deflects the flow to the designated channel and sorts selected cells to a specific collection tube. In the single-cell sorting mode, each deflected cell exits the designated channel in a cartridge and enters a programmable cell-placement unit that places each cell into one well of a 386-well plate.

Fig. 1.2.1 (b) shows the design of the imaging system. Suspended cells are hydro-dynamically focused to the center of the microfluidic channel by a sheath flow 6. At the optical interrogation zone, each cell is illuminated simultaneously by a scanning dual-wavelength 488/561 nm laser for imaging, and a non-scanning 455 nm LED for cell speed measurement. The 488/561 nm wavelength laser is coupled to an Acousto Optic Deflector (OAD948, Isomet) to create a scanning beam along the y-direction (width of the microfluidic channel) with a spot size of 1μm in diameter. Two 10X (NA=0.28) objective lenses (378-803-3, Mitutoyo) on opposite sides of the microfluidic channel are used, one for focusing the scanning laser illumination beam and the other to collect the fluorescent or transmitted light. A series of dichroic mirrors with different reflection bands separate the LED, transmission, and fluorescent signals to different channels, which are detected by PMTs (H10721-20, Hamamatsu). A spatial mask, with two slits separated in the cell flow direction, is placed at the image plane of the optical system to generate speed information. The field of view, dependent on the scanning range of AOD and the signal recording time period, is chosen to be 30μm by 30μm, which covers the size of most biological cells.

An FPGA-CPU hybrid design is adopted to meet the requirement of real-time processing and computation of cell image features with high throughput. The FPGA reconstructs the cell images from the acquired signals, and the reconstructed cell images are transferred to the CPU for imaging feature extraction, including area, perimeter, circularity, aspect ratio (major axis length/minor axis length), integrated intensity, mean intensity (intensity divided by area), standard variation of intensity over space, granularity, spot count, etc. Features directly related to the cell sorting criteria are then transferred back to the FPGA. The image reconstruction and feature extraction are completed in less than 1ms in most cases

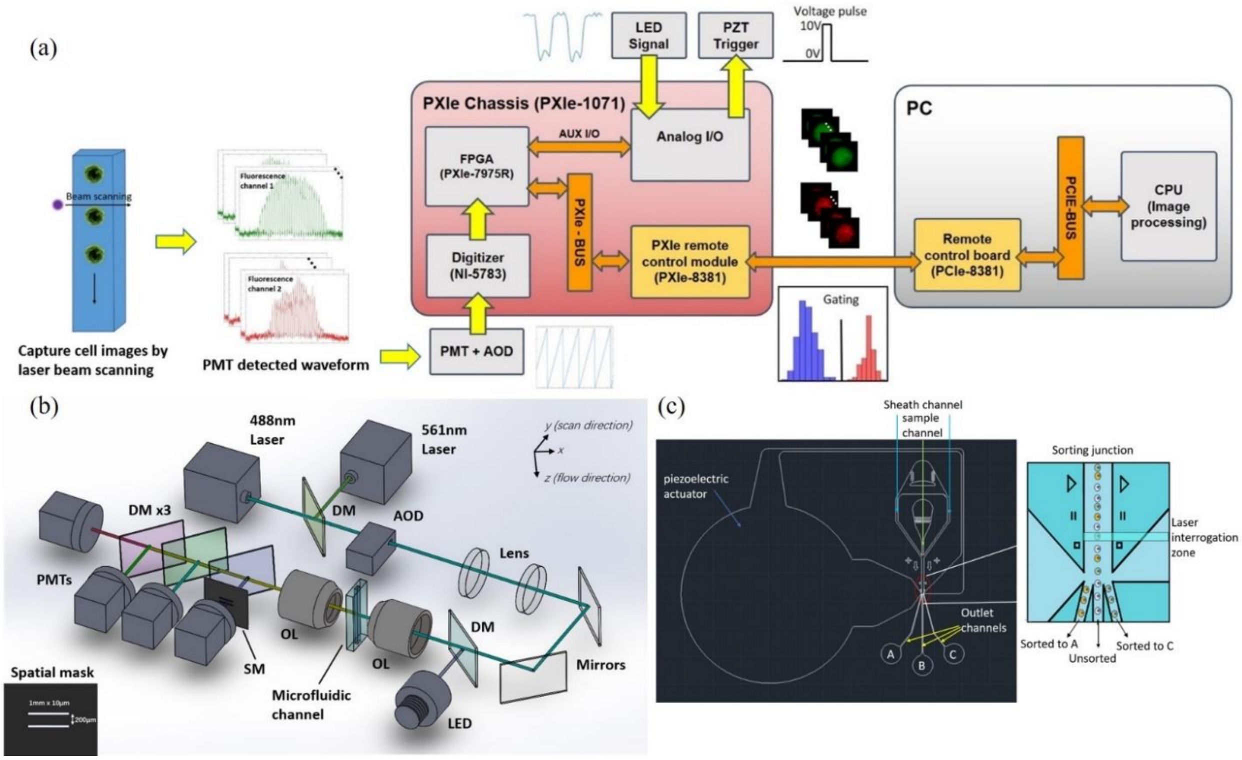


Figure 1.2.1. 2D IACS system overview. (a). Overall system architecture. The scanning laser beam and the cell travel produce an equivalence of a 2D raster scanning system. The bright field and fluorescent signals of the cell are detected by PMTs and the temporal signals are reconstructed to form cell images. Meanwhile, the features of each cell image are extracted by a PC. According to the sorting criteria (gating) based on user-selected image features, the on-chip piezoelectric (PZT) actuator is triggered to sort out cells that have the target features. (b). Design of imaging system. AOD, acousto-optic deflector; DM, dichroic mirror; OL, 10X/0.28 objective lens; PMTs, photomultiplier tubes; SM, the double-slit mask for cell speed detection with its design shown on the left. (c). Microfluidic chip design. Suspended cells are focused to the center of the microfluidic channel by a sheath flow. The on-chip piezoelectric actuator bends upward or downward mechanically, deflecting the flow and the target cell within the flow into the designated channel

1.3 Cameraless 3D imaging flow cytometer

A significant limitation of 2D IFC systems is that, regardless of the optical detection method and computation algorithm is used, only 2D cell images can be obtained [10]–[12]. The absence of 3D tomography results in the occlusion of objects, blurring by focal depth, loss of z-axis spatial resolution, and artifacts due to the projection of a 3D cell into a 2D image. For a range of applications, such as internalization measurements, probe co-localization, and spot counting, relative to 2D imaging that is dependent on the cellular orientation to the imaging plane, 3D images provide complete and accurate phenotyping of cell and organelle morphology, as well as nucleic acid and protein localization to support biological insights [13].

Here, a cameraless high-throughput three-dimensional imaging flow cytometry (3D-IFC) based on optical sectioning microscopy [14] is introduced. The combination of light-sheet scanning illumination technique and spatial-temporal transformation detection technique enables fluorescent and label-free 3D cell image reconstruction from single-element photodetector readout without a camera [8], [15]. Building upon the speed and sensitivity benefits of the photomultiplier tube (PMT), the 3D-IFC uses multiple scanning techniques to add spatial information in a conventional flow cytometry architecture. 3D imaging is achieved by laser scanning across the first (z-) axis, the cell translating by flow across the second (y-) axis, and the use of multiple pinholes arranged along the third (x-) axis to produce fluorescent and label-free information from 6,000 voxels per scanning volume. By precisely mapping time to space, the photodetector readout at one timepoint corresponds to one voxel in a 3D space. Here we demonstrate 3D-IFC of fluorescence and 90-degree label-free side-scattering (SSC) imaging of single cells in the microfluidic channel at a velocity of 0.2 m/s, corresponding to a throughput of approximately 500 cells per second.

A schematic of the 3D-IFC system is shown in Fig. 1.3.1 (a). In the 3D-IFC system, suspended cells form a 2D hydrodynamically focused single file in a quartz flow cell with a square cross-section [16]. Laser excitation is via a light sheet (x-y plane) with a diffraction-limited beam waist and a height of 200 to 400 µm, scanning in the z-direction at 200 kHz. When a cell flows through the whole optical interrogation at 0.2 m/s, a pixelated field of view is represented by a 3D space with X by Y by Z voxels, as shown in Fig. 1.3.1 (c). A pinhole array on the spatial filter is aligned at a tilting angle, ϑ, to the flow stream, so the pinhole array also steps along the x-direction. In this manner, each pinhole allows light from voxels with a distinct x-index to reach the PMT detector (see Fig. 1.3.1 (g) for details of the pinhole mask design). The imaging process begins when a flowing cell appears at the first pinhole of the spatial filter. During the first light-sheet scanning period (5 µs), the light intensity of voxels z_{1-Z} with $x_1 y_1$ index is collected. As the cell flows downstream in the y-direction to the next position, $x_1 y_2$, the corresponding z_{1-Z} voxels are produced. In this manner, when a cell completely passes pinhole 1, the entire 2D yz-slice at x_1 is imaged. As the cell travels farther downstream the y-direction and passes through all the following pinholes, yz-slices of at x_2 to x_X are recorded.

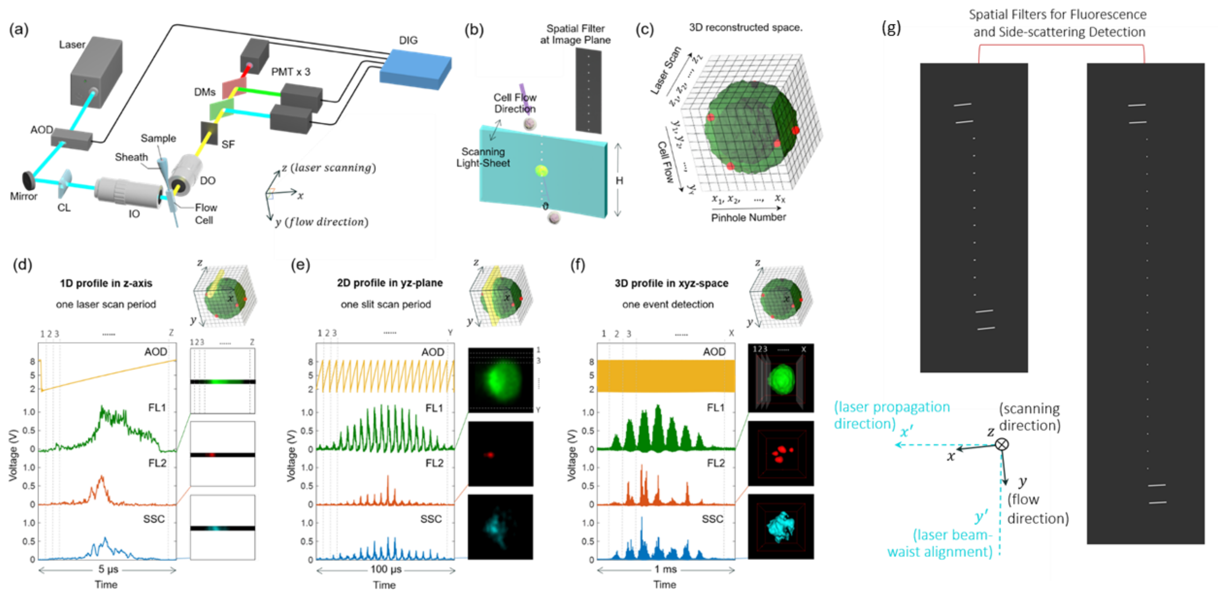


Figure 1.3.1. Implementation of the 3D imaging flow cytometer (3D-IFC). (a) Schematic diagram of the 3D-IFC system. AOD, acousto-optic deflector; CL, cylindrical lens; IO, 50X/0.55 illumination objective lens (Mitutoyo, Plan Apo Infinity Corrected Long WD Objective); DO, 10X/0.28 detection objective lens (Mitutoyo, Plan Apo Infinity Corrected Long WD Objective); SF, spatial filter; DMs, dichroic mirrors; PMT, photomultiplier tube; DIG, 125 MS s-1 digitizer. The AOD and CL produce a scanning light sheet. The sample is 2D hydrodynamically focused by a sheath flow before entering the square cross-section quartz flow cell. (b) Optical interrogation area. H, the height of the light sheet; ϑ, tilt angle between flow (y-axis) and vertical line. Illumination light sheet propagates horizontally and scans in the z-axis, the sample flows in the y-axis, x is the orthogonal axis. The spatial filter at the image plane uses pinholes to produce line scans across the x-axis. (c) 3D reconstructed space. The resolution on the X-axis is determined by the number of pinholes (pixelated field of view in the x-direction); resolution of Y by the distance between two slits (pixelated field of view in the y-direction); and resolution of Z by the light-sheet scanning range (pixelated field of view in the z-direction). (d) One light sheet scan period produces a 1D light intensity profile in the z-axis. The PMT voltage readout of one sample point corresponds to the light intensity of one voxel in the z-axis. (e) While the object travels along the y-axis, multiple scans produce a 2D profile in the yz-plane within one pinhole scan period. Each section—separated by dotted lines—corresponds to the light intensity of one row in the 2D image stack. (f) When an object completely passes through the spatial filter covering the area, the time-domain signal contains the complete information of the 3D profile in the xyz-space. Each section corresponds to one 2D image slice. AOD, tuning voltage of the AOD driver; FL1, PMT readout of fluorescence detection channel 1; FL2, PMT readout of fluorescence detection channel 2; SSC, PMT readout of side-scattering light detection channel. (g) Spatial Filter Design. Two examples of spatial filters placed at the image plane. The top two and bottom two long slits with dimensions of 10 µm by 200 µm are for speed detection. The other pinholes on the spatial filter are 10 µm by 20 µm (left) and 10 µm by 10 µm (right), for 3D image capturing with a pixel size of 2 µm and 1 µm in the x-direction, respectively.

1.4 Single-cell position and image mapping techniques using 3D-IFC system

To relate single-cell position with its images from the 3D IFC system, we develop a high throughput technique to relate positions of individual cells to their 3D imaging features with single-cell resolution. The technique is particularly suitable for non-adherent cells where existing spatial biology methodologies relating cell properties to their positions in a solid tissue do not apply. Our design consists of two parts: recording 3D cell images at high throughput (500 to 1000 cells/s) using a custom 3D imaging flow cytometer (3D-IFC) and dispensing cells in a first-in-first-out (FIFO) manner using a robotic cell placement platform (CPP). To prevent errors due to violations of the FIFO principle, we have invented a method that uses marker beads and DNA sequencing software to detect errors. Experiments with human cancer cell lines demonstrate the feasibility of mapping 3D side scattering and fluorescent images, as well as 2D transmission images of cells to their locations on the membrane filter for around 100,000 cells in less than 10 minutes. While the current work uses our specially designed 3D imaging flow cytometer to produce 3D cell images, our methodology can support other imaging modalities. The technology and method form a bridge between single-cell image analysis and single-cell molecular analysis.

Figure 1.4.1 shows the overall design and workflow of our approach. The system consists of two interconnected hardware modules, a 3D imaging flow cytometer and a robotic cell dispenser. Cells and beads were pre-mixed and examined using the 3D-IFC system. The 3D hydrodynamically focused sample flow establishes a single-cell stream with a sample concentration of ~500 samples/uL. When a cell or bead passes through the laser interrogation area, it is illuminated by a scanning light sheet at a 200kHz scanning rate. The spatial filter placed at the image plane contains a series of spatially positioned pinholes aligned with the cell flow direction by a predetermined separation. The emitted light from a specific portion of a cell is detected by photomultiplier tubes (PMTs). The spatial-temporal transformation is applied to reconstruct the 3D tomographic images. The forward spatial filter contains a long slit aligned with the laser scanning range. The transmitted light is collected by a PMT and the signal can produce a 2D transmission image. In this cameraless design with a scanning light-sheet and spatial masks, the 3D-IFC system can produce 3D side scattering and fluorescent images plus a 2D transmission image of traveling cells at a rate of 1000 cells/s. The details of the 3D-IFC can be found in our earlier publications[29,30].

The cell sample was premixed with three non-fluorescent beads of different sizes: 10um beads which we represented as nucleobase A, 20um beads which we represented as nucleobase T, and 30um beads which we represented as nucleobase C. Hence the sequence consists of these 3 types of marker beads and cells. By matching the marker bead sequences between the 3D-IFC signals and cell placement platform, we were able to align the two sequences, which subsequently enabled us to map the cells between marker beads. To keep the average number of cells between marker beads to be a relatively small number ($\bar{n} = 2$) and minimize the chance of error, we kept the ratio between cells and the total number of marker beads to be 2:1.

When the cells and beads passed the laser interrogation area of the 3D-IFC and exited the flow cell, they were dispensed in a first-in-first-out (FIFO) manner on a template consisting of a 12 μm thick transparent porous film (Sterlitech, SKU. 1300026) on a holder with an array of groves. The liquid out of the 3D-IFC (at around 300 μL/min) was immediately absorbed by the porous membrane filter through the capillary effect and drained by vacuuming the groves under the membrane. Only the beads and cells were left on the wetted porous membrane. The moving speed of the template was programmed according to the cell density in the sample to achieve an average cell-to-cell spacing of 250 μm along the line of travel and a spacing of 500 μm between two adjacent lines of cells. In this design, a filter plate of the same size of a 384-well plate can house around $6x10^4$ cells. For a 30-minute run of the 3D-IFC at a throughput of 300 cells/s, we can record around 500,000 3D images of single cells deposited on 10 cell plates with full knowledge of the position of every single cell and its 3D image.

1.5 High-throughput signal and image processing needs

Compared with Conventional FACS only a few intensity features can be extracted from the examined cell, IFC systems need to analyze high-content image feature that comprises thousands to hundreds of thousands of pixel intensity values and the spatial distribution of pixels. Such complex image-based features lead to a tremendous burden for back-end data transmission and image processing. For example, the 3D-IFC system generates at least 1GB of data per second at a throughput of 500 cells per second, each cell with more than 6,000 voxels. Such a large volume is challenging for traditional data and image processing workflow. In addition to the computation requirements for processing the large volume of image data, the IACS system also faces the real-time constraints that all data processing and image features extraction must be completed during the milliseconds time window when the cell travels from the optical interrogation region to the sorting junction. Tradition intensity-based processing algorithms adopted by the FACS system and conventional image-feature extraction approaches cannot meet the above requirements. With the recent advances in real-time processing hardware like field-programmable gate arrays (FPGA) and graphical processing units (GPU) and evolving development of deep learning algorithms, analyzing the big data generated from IFC and IACS systems could be possible and become artificially intelligent.

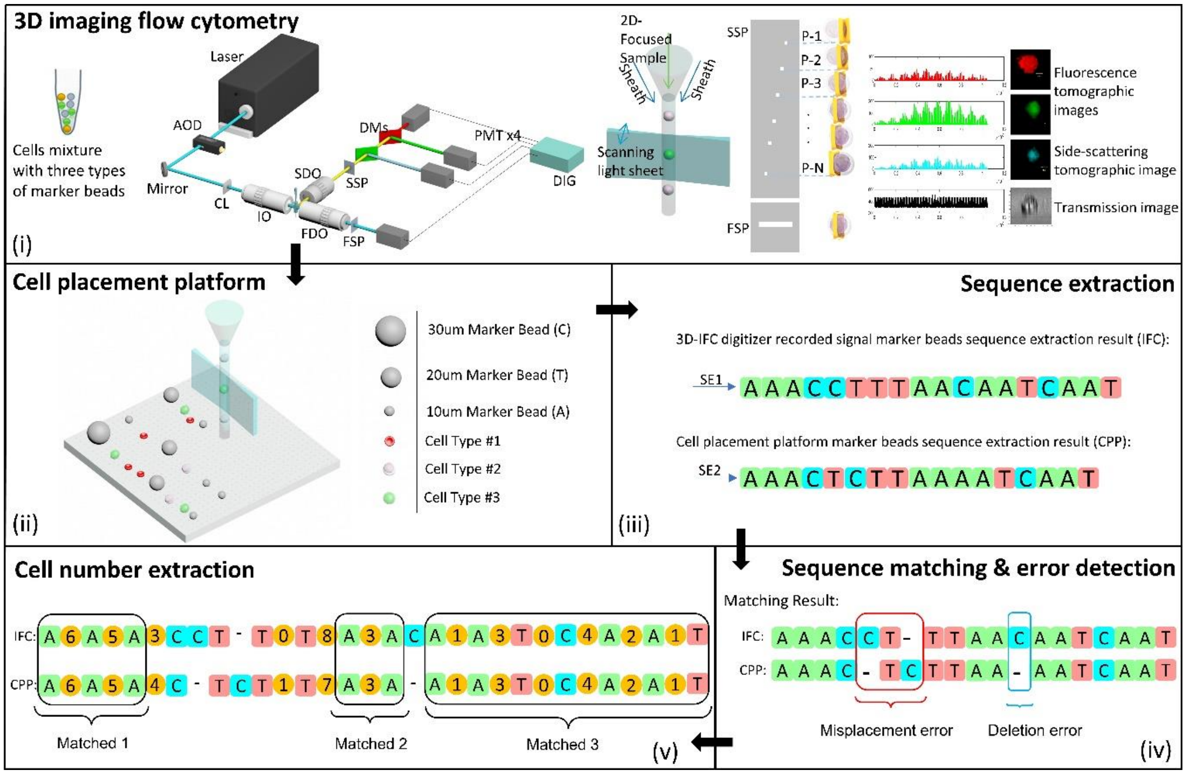


Figure 1.5.1. Single-cell analysis and isolation workflow based on 3D imaging flow cytometer (3D-IFC) and robotic cell dispensing and pickup. Workflow: (i) Three types of marker beads are mixed with cell mixture to help map cell images from the image stream to cell locations. Run the sample through a 3D-IFC that can capture 3D fluorescent and side-scattering (SSC) images as well as 2D cell transmission images. (ii) After cells exit the 3D-IFC, they are dispensed by a robotic dispenser in a first-in-first-out (FIFO) manner on a template. (iii) Marker beads sequences from the 3D-IFC signal and the cell placement platform (CPP) are extracted using developed sequence extraction pipelines described in supplementary Fig. S1 and Fig. S2. (iv) Two sequences are compared and matched using a sequencing bioinformatics tool. (v) Marker beads sequence errors (including deletion and misplacement errors) are detected, and for the correctly registered marker beads, the number of cells between two marker beads is identified. In this manner, one can relate hundreds of thousands of individual cells to their respective 3D fluorescent and scattering images in high accuracy. If certain cells with specific image features are of interest, one can locate and pick up those cells individually for downstream analyses. Legends: AOD, acousto-optic deflector; CL, cylindrical lens; IO, 20X/0.42 illumination objective; SDO, 10X/0.28 side detection objective; SSP, side spatial filter; DMs, dichroic mirrors; FDO, forward detection objective; FSP, forward spatial filter; PMT, photomultiplier tube; DIG,125 MSs−1 digitizer; SE1, Sequence Extraction from 3D-IFC, SE2: Sequence Extraction from cell placement platform (CPP).

1.6 Scope of Dissertation

This dissertation focuses on the development of cameraless imaging flow cytometry and image-activated cell sorting technology using artificial intelligence (AI). The development of deep learning algorithms and the real-time implementation of AI models enables high-throughput data analysis image processing in the IFC system and supports AI-guided image-activated cell sorting in the IACS system.

Earlier in Chapter 1, we introduced the cameraless IACS and 3D-IFC systems and indicated the current limitation on data and image processing for IFC and IACS technologies.

Chapter 2 presents the development of label-free cell analysis using the 3D imaging flow cytometer (3D-IFC) and convolutional neural network (CNN) and demonstrates the clinical application of multimodal NASH prognosis using 3D-IFC and CNN. This chapter also describes the capability of conducting label-free intracellular structure segmentation using the IFC system.

Chapter 3 extends the implementation of the AI model to meet the sorting processing time constraint and presents the development of morphology-based artificial intelligence (AI)-guided image-activated cell sorting system.

Chapter 4 summarizes the dissertation and briefly discusses the future direction of research.

Chapter 1, in part, is a reprint of the material as it appears in Optica 2019, Y. Han, R. Tang, Y. Gu, A. Zhang, W. Cai, V. Castor, S. Cho, W. Alaynick and Y.-H. Lo. "Cameraless high-throughput 3D imaging flow cytometry". The dissertation author was the second author of this paper.

Chapter 1, in part, is a reprint of the material as it appears in High-Speed Biomedical Imaging and Spectroscopy V, Proceeding of SPIE 2020, R. Tang, X. Chen, Z. Zhang, L. Waller, J. Chen, Y. Gu, Y. Han, C. Lee, I. Gagne, A. Zhang, S. Cho and Y.-H. Lo. "2D image-guided cell sorter and 3D imaging flow cytometer". The dissertation author was the first author of this paper.

Chapter 1, in part, is a reprint of the material as it appears in Proceedings of the National Academy of Sciences 2022, Z. Zhang, R. Tang, X. Chen, L. Waller, A. Kau, A. Fung, B. Gutierrez,

C. An, S. Cho, L. Shi and Y.-H. Lo. “A high-throughput technique to map cell images to cell positions using a 3D imaging flow cytometer”. The dissertation author was the second author of this paper.

Chapter 2 Label-Free Cell Analysis Using Imaging Flow Cytometer and Convolutional Neural Networks

Compared with conventional fluorescence biomarker labeling, the classification of cell types based on their stain-free morphological characteristics enables the discovery of new biological insight and simplifies the traditional cell analysis workflow. Most artificial intelligence (AI) aided image-based cell analysis methods primarily use transmitted bright-field images or holographic images. Here, we present the first study of the convolutional neural network (CNN) analysis on three-dimensional side scattering cell images out of a unique three-dimensional (3D) imaging flow cytometer study. Human cancer cell lines and leukocyte classifications were performed to investigate the information carried by the spatial distribution of side scattering imaging of single cells. We achieved a balanced accuracy of 98.8% for cancer cell line classification and 92.3% for leukocyte classification, respectively. The results demonstrate that the side scattering signals can produce not only general information about cell granularity following the common belief but also carry rich information about the properties and functions of cells, which can be uncovered by the availability of side-scattering imaging flow cytometer and the application of CNN. In addition, we extended the analysis to the cell classification of key regulated cell types involved in the development of non-alcoholic steatohepatitis (NASH) and demonstrated the capability of intracellular structure segmentation of CHO cells. Thereby we have opened a new avenue for cell phenotype analysis in biomedical and clinical research.

2.1 Introduction

Characterization and classification of different morphologies and phenotypes in a heterogeneous cell population generate biomedical applications and significant insight in biological research to correlate cell phenotype and genotype information [2], [4], [17]. Although cell classification is beneficial for understanding cell heterogeneity, it usually requires the revelation of spatial information of the intercellular structures of single cells and the analysis of a large amount of data. In recent years, significant advances have been made in both the hardware of producing the cell imaging data from single cells and the data processing algorithms based on deep learning. The rapid development and proliferation of imaging flow cytometry (IFC) facilities data-driven cell analysis as IFC can generate a large amount of cell imaging data at a very high rate [12]. Most IFC systems (e.g., the Amnis® ImageStream IFC system) can generate label-free images by detecting the transmitted light from cells without fluorescent biomarker labeling, which complicates the workflow and might disturb the cell morphology and viability during the staining process [18]. On the other hand, the advances in the fields of deep learning and artificial intelligence have transformed the traditional cell image processing by greatly enhancing our ability of discerning and classifying cell features unattainable by traditional cell features, such as size, shape, etc. Recently, in combination with the advances in artificial intelligence, IFC has driven the exploitation of cell classification pipelines for biomedical practice using high-content single-modal or multi-modal cell images [19]–[21].

Several recent works have applied imaging techniques with the IFC systems and data-driven machine learning algorithms to demonstrate label-free cell classifications. Chen *et al*. reported a time-stretch quantitative phase imaging (TS-QPI) system and the artificial neural network model to identify white blood cells from colon cancer cells with a 96.4% balanced accuracy[22]. Wu *et al.* demonstrated an intelligent frequency-shifted optofluidic time-stretch quantitative phase imaging (OTS-QPI) system that applies a convolutional neural network (CNN) autoencoder to extract image features from the captured image intelligently and achieved high accuracy of over 96% to classify leukemia cells among healthy white blood cells [23]. Though OTS-QPI methods facilitate label-free binary and multi-class classification, they only capture cell images with a single modality, lacking companion fluorescent images for verification and ground truth determination. Li *et al.* presented an IFC system that implements digital holographic microscopy (DHM) imaging for 3-part leukocyte recognition using machine learning algorithms with a high balanced classification accuracy of 99% [24]. However, the adopted machine learning method is prone to over-fitting because of the random shuffle and split of the dataset when evaluating the prediction model. In the reported methodology, the classifier might have already seen a particular subject during the training step and then predicted the same subject in the validation step [25]. The latest approaches with improved reliability and repeatability include adopting commercially available IFC systems to conduct label-free multi-class human white blood cells (WBC) classification. Nassar *et al*. established a label-free WBC approach using the Amnis® ImageStream IFC system to capture the two-dimensional (2D) transmission (bright-field) and side-scattered (dark-field) images of human WBC to achieve four-class WBC classification with an average F1-score of 97% [26]. Lippeveld *et al.* also used the bright-field and dark-field 2D images captured by the Amnis IFC system to compare the human WBC classification performance using conventional machine learning with that using the deep learning approach. They have reported an eight-class classification accuracy of 77.8% and 70.3% for classical machine learning and deep learning approaches, respectively [27].

All label-free IFC technologies are developed based on the understanding that light scattering properties are related to cell morphology since the 1970's [28], [29]. Until the past two decades, very few studies examined the correlation between the intracellular structure and scattered light patterns. A theoretical study on three-dimensional (3D) simulation of light scattering patterns from biological cells shows that cells with a slightly different intracellular structure can be differentiated by measuring the side-scattered light at optimal angles[30]. A recent experimental study also demonstrated that while cellular organelles contribute to the side-scattering pattern, the nucleus has the largest contribution [31]. However, cell classification based on only the side-scattering pattern has not been demonstrated yet, mainly because the side-scattering images are hard to obtain from most custom-designed systems, and the SSC dark-field images obtained from the commercial system appear to be much darker and less informative than the bright-field images. A more general and significant limitation of the existing IFC technologies is that the captured 2D images from 3D cells are suffering from projection problems, and the collapsed 2D images cannot fully reveal the 3D spatial information. Utilizing the label-free 3D cell tomographic images for cell classification had not been demonstrated, due to a lack of IFC technology to generate a label-free 3D tomographic cell. In order to utilize label-free 3D cell tomographic images for cell classification, our group produced a flow cytometer design capable of capturing 3D 90-degree side scattering and fluorescent images [32].

In this chapter, we demonstrate an intelligent label-free cell type analysis workflow that utilizes 3D cell tomography of the side-scattered light captured by a recently developed camera-less high-throughput 3D imaging flow cytometry (3D-IFC) system [32]. This offers a new modality to reveal the 3D internal structures of cells through tomographic imaging from the side-scattered light. The 3D side-scattering (SSC) tomographic image is captured along with the 3D fluorescent tomographic images from the assayed cell sample. Then, the SSC image and ground-truth label extracted from corresponding fluorescence images are used as the input for the deep learning process. Experiments were conducted to evaluate our workflow performance for multi-class classification of human cancer cell lines and white blood cells, respectively. Our results demonstrate a three-part classification balanced accuracy of 98.8% and 92.3% for cancer cell lines and human WBC type classification using customized 3D CNN models, respectively. We also found that models with 3D SSC images as inputs outperform the ones using the projected 2D SSC images, confirming that the 3D SSC tomographic image encodes a greater depth of information on internal cellular structure than the 2D SSC image.

2.2 Methods and materials

2.2.1 Intelligent 3D-IFC-based cell type analysis workflow

The three-stage workflow of intelligent 3D-IFC-based cell type analysis is illustrated in Figure 2.2.1. First, a mixture of cells is labeled with fluorescent biomarkers in the sample preparation process, targeting each cell type with corresponding biomarkers. The cell mixture is examined using the 3D-IFC system. The cell flow stream is 3D hydrodynamically focused by a 2D sheath flow confinement in a flow cell cuvette (Hamamatsu, Cat. J11020-000-004). Such flow confinement establishes a single-cell stream with a cell concentration of approximately 1000 cells per microliter in the sample flow stream. When a cell is flowing through the optical interrogation area at 0.2m/s, it is illuminated with a scanning light sheet (at 200 kHz). With a series of spatially positioned (10 to 20) pinholes aligned with the cell flow direction by a predetermined angle, the emitted fluorescent and scattered light from a specific spatial position of a cell is detected by photomultiplier tubes (PMTs) chronologically with a series of dichroic mirrors and spectral filters to separate the detection light spectra. The single-cell stream is examined with a throughput of approximately 500 cells per second. 3D tomographic images are reconstructed based on the detected temporal signal intensity by applying the temporal-spatial transformation method [8], [32]. The temporal-spatial transformation was implemented using MATLAB, and the acquired raw waveforms were processed offline. The reconstructed 3D images were inspected through image gradient analysis. The truncation check step eliminates the truncated events. (see Figure 2.2.2 for image reconstruction pipeline). The intensity features are then extracted from the images and are analyzed by manual gating to identify the ground truth label for each cell in a sample mixture, analogous to phenotyping by conventional flow cytometry. At the last stage, the preprocessed 3D SSC images, together with the ground truth labels, are used as the input of customized deep CNN models for the training and evaluation process compared with the state-of-the-art architecture as a benchmark. The classification performance is evaluated by the balanced prediction accuracy and confusion matrices of the test dataset.

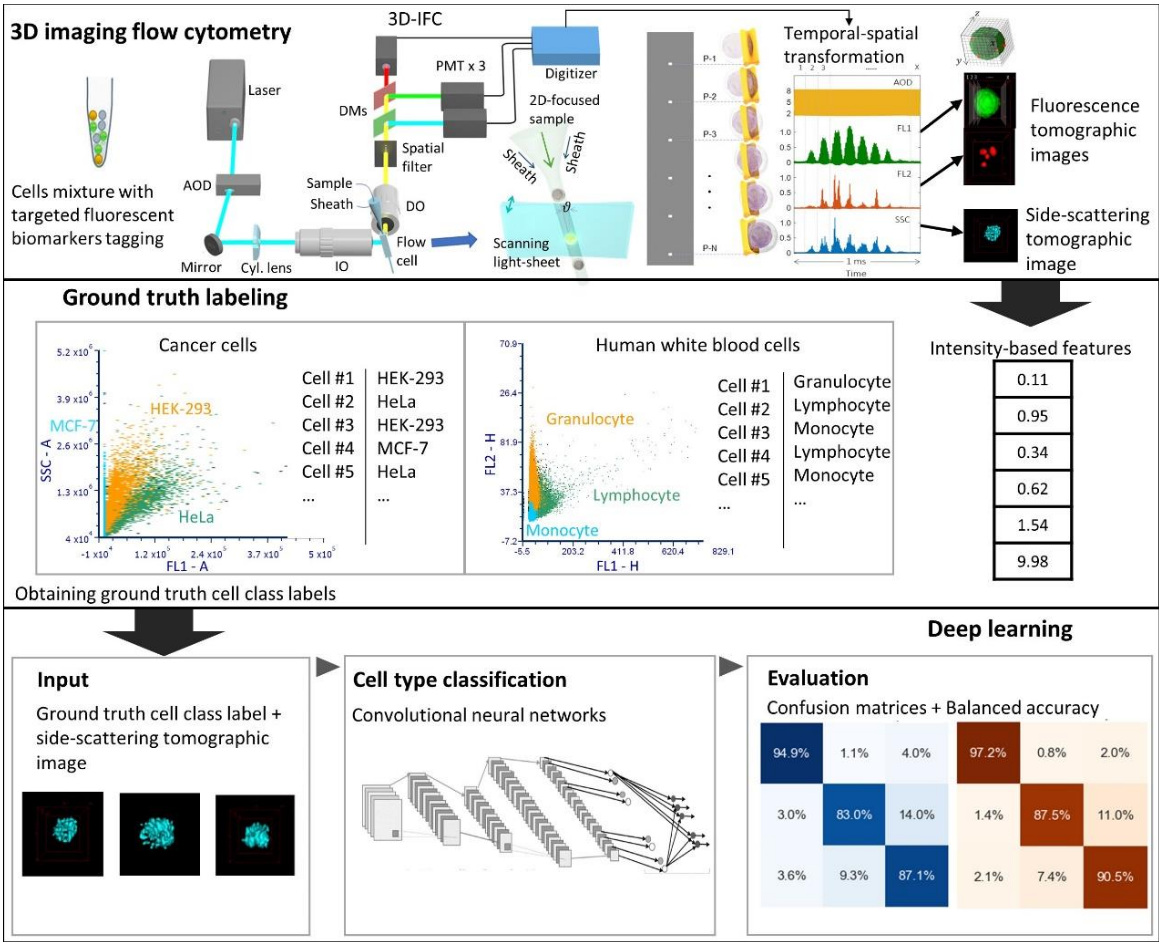


Figure 2.2.1. Intelligent three-dimensional imaging flow cytometry (3D-IFC) based cell type analysis workflow. 3D-IFC system images targeted biomarkers labeled cell mixture, it generates fluorescent and side-scattering (SSC) 3D tomographic images through temporal-spatial transformation. In order to find the ground truth labels, cells are fluorescently labeled, while the label-free 3D SSC images are used for training and classification. Convolutional neural network (CNN) models use SSC images to make predictions of the cell type. Classification performance is evaluated through confusion matrices and balanced prediction accuracy. AOD,acousto-optic deflector; Cyl. lens, cylindrical lens; IO, illumination objective lens; DO, detection objective lens; DMs, dichroic mirrors; PMT, photomultiplier tube; FL1, fluorescent channel 1; FL2, fluorescent channel 2; SSC, side-scattering channel.

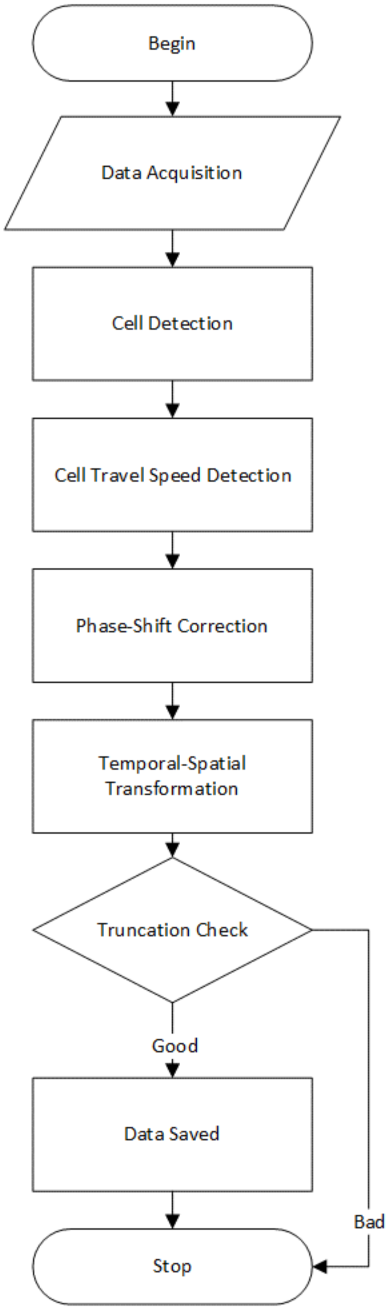


Figure 2.2.2. Batch processing pipeline for automatic image reconstruction using temporal-spatial transformation

2.2.2 Cell preparation and image acquisition

2.2.2.1 Human cancer cell type classification experiment

The human embryonic kidney 293 (HEK-293) cells, the Michigan Cancer Foundation-7 (MCF-7) cells, and the cervical cancer cells (HeLa) were used in the human cancer cell type classification experiment. Each cell line was cultured and harvested separately and then separated into three batches for targeted fluorescence staining. Two batches were fluorescently stained with the CFSE Cell Proliferation Kit (Ex/Em 492/517nm, Cat. 34554, Thermo Fisher) and the CellTrace Yellow Proliferation Kit (Ex/Em 546/579nm, Cat. 34567, Thermo Fisher), respectively. Both stained and unstained cells were fixed by 4% paraformaldehyde. Three cell lines were evenly mixed with one unstained cell line and two fluorescently stained cell lines with separate emission bands and then were analyzed by the 3D-IFC system in multiple image acquisition experiments. For each image acquisition experiment, the 3D SSC images from the unstained cell type were isolated based on the detected fluorescence intensity and were used for deep learning. A total of 10,270 images were acquired from the unstained cancer cells for the cancer cell dataset. The dataset contains 3191 HEK-293 cells, 3315 HeLa cells, and 3764 MCF-7 cells.

2.2.2.2 Human white blood cell type classification experiment

This experiment used Veri-Cells™ Leukocyte Kit (Cat. 426003, BioLegend), prepared from lyophilized human peripheral blood leukocytes. Before the image acquisition experiment, the WBC sample was immuno-stained with the following antibodies cocktail for phenotyping: CD3, CD14, CD19, CD66b. Three WBC types were phenotyped using the antibody cocktail: lymphocytes, granulocytes, and monocytes. Detailed staining protocols are described in the supplementary materials. The 3D-IFC system analyzed the phenotyped cell sample in multiple image acquisition experiments. In each image acquisition experiment, the 3D SSC images from the three WBC cell types were isolated based on the detected fluorescence intensity and were used for deep learning. A total of 24,230 images were acquired from the unstained cancer cells. For the human WBC dataset, the dataset contains 13,573 granulocytes, 3,061 lymphocytes, and 7,596 monocytes.

For both cell type classification experiments, the captured 3D SSC image was stored as a 3D image stack with a field of view of 20 × 20 × 20 µm space. The 3D image stack was 80 × 80 × 80 pixels. In the meantime, a 2D projection image was generated by collapsing the depth dimension of the 3D image stack. Figure 2.2.3 shows the 3D SSC image stacks of human cancer cells in comparison with the 2D SSC projection and bright-field microscopic images.

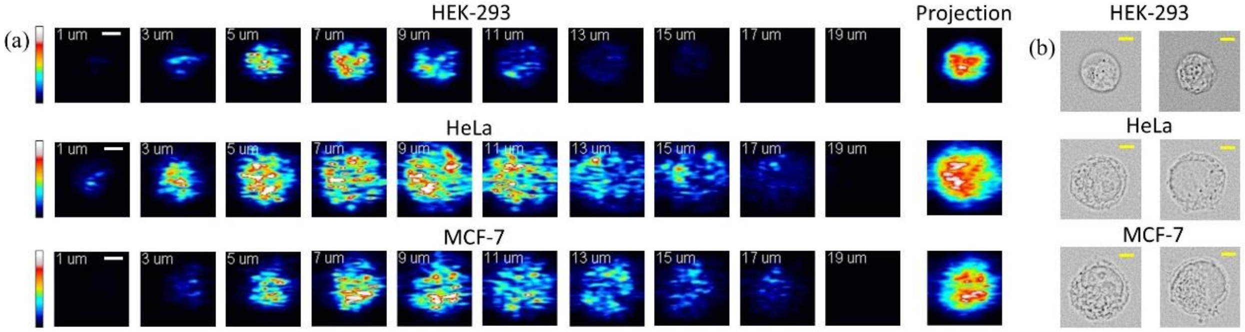


Figure 2.2.3. SSC images of cancer cells by 3D-IFC. (a) Example 3D SSC image stacks of HEK-293, HeLa and MCF-7 cells and corresponding 2D projections. (b) Example microscope images of HEK-293, HeLa and MCF-7 cells. Scale bar: 5 µm.

2.2.2.3 NASH disease characterization experiment

Cell samples were harvested and prepared from two healthy donors and two NASH patients. For the image acquisition experiment, each batch of cell samples was independently run by the 3D-IFC system. Hepatic stellate cell (HSC) and liver endothelial cell (LEC) samples were analyzed separately for every subject. Each cell produced an image pair of a 2D transmission and a 3D side-scattering image. The captured 3D side-scattering image was stored as a 3D image stack of 40 × 40 × 40 voxels with a field of view of 40 × 40 × 40 μm^3. The acquired 2D transmission was stored to 80 × 80 pixels with a field of view of 40 × 40 μm^2. The label-free dual-modality 3D-IFC system illustration is shown in Figure 2.2.4. The representative LEC and HSC cells acquired from the 3D-IFC system are presented in Figure 2.2.5.

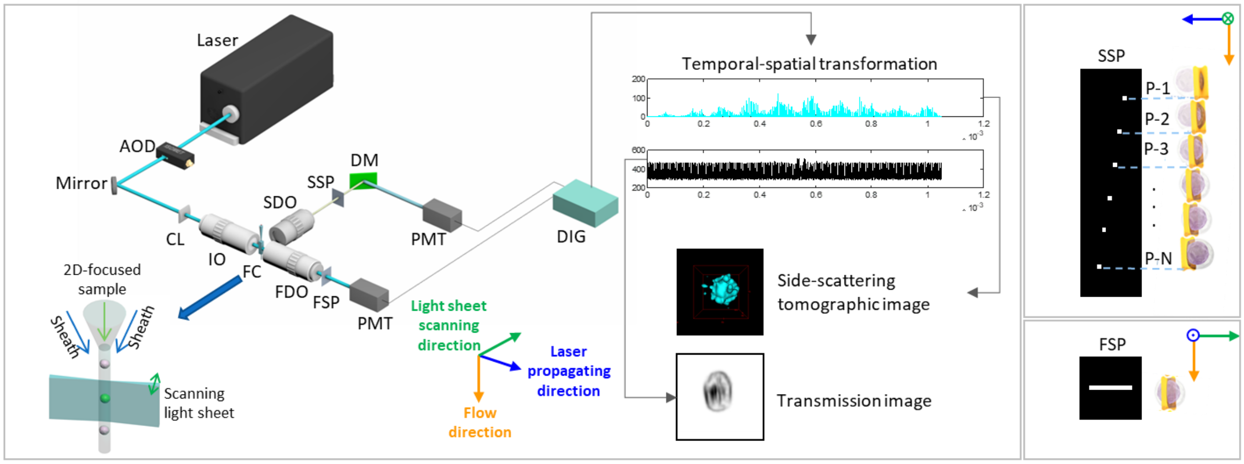


Figure 2.2.4. 3D imaging flow cytometer (3D-IFC) dual-modality imaging system. AOD, acousto-optic deflector; CL, cylindrical lens; IO, 20X/0.42NA illumination objective; SDO, 10X/0.28 NA side detection objective; SSP, side spatial filter; DM, dichroic mirror; FC, flow cell; FDO, 50X/0.55 NA forward detection objective; FSP, forward spatial filter; PMT, photomultiplier tube; DIG, 25MSps digitizer

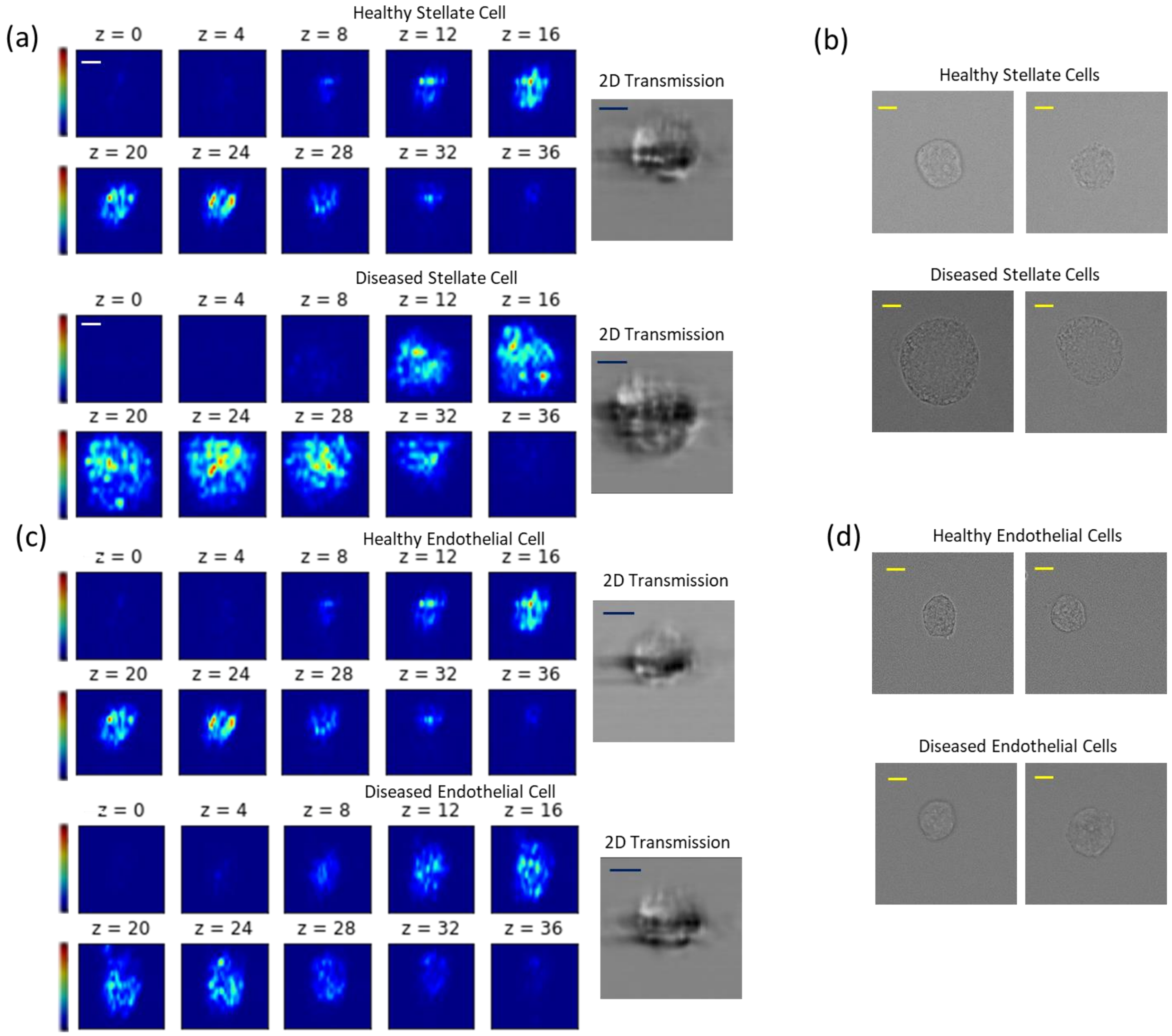


Figure 2.2.5. 2D transmission and 3D SSC images of liver cells by 3D-IFC. (a) Example 3D SSC image stacks of healthy (top stack) and diseased HSC cells (bottom stack) and corresponding 2D transmission images. (b) Example microscope images of healthy and diseased HSC cells. (c) Example 3D SSC image stacks of healthy (top stack) and diseased LEC cells (bottom stack) and corresponding 2D transmission images. (d) Example microscope images of healthy and diseased LEC cells. Scale bar: 10 µm.

2.2.2.4 Preparation of CHO-K1 cells for cell body and nuclear segmentation

CHO-K1 (ATCC CCL-61) was used for nuclear staining. The cells were harvested when they were at a confluency of around 80%. The cells were centrifuged at 350 x g for 5 minutes to remove the supernatant to perform a wash. After washing the cells with PBS (Genesee Scientific, CA, USA), the cells were centrifuged at 350 x g for 5 minutes. The supernatant was removed and 100 µL of 4% formaldehyde, methanol-free (Cell Signaling Technology, Massachusetts, USA) was added for every million cells to be fixed. The cells were incubated for 15 minutes at 37ºC. After fixation, the cells were washed again and resuspended in PBS with 0.5% BSA (Thermo Scientific) to a concentration of $1.0\text{x}10^6$ cells/mL. Lastly, the cells were stained with 0.5 µM Vybrant DyeCycle Green Stain (Invitrogen) for 30 minutes and filtered with a 35 µm strainer cap (Genesee Scientific, CA, USA) before analysis.

2.2.3 Ground truth labeling and data preparation

The detected fluorescence and SSC intensity levels were extracted from the 3D images after image acquisition. A manual gating process was implemented to identify the ground truth cell type based on the detected fluorescence intensity. For each experiment, an unstained cell type and two stained cell types were assayed with different fluorescence emission bands. Therefore, different cell types could be gated manually based on the detected fluorescence intensity level, and the ground truth cell type labels were assigned to each 3D SSC image.

In the data preparation process, we first removed the background noise, leaving only the pixels of interest in the 3D SSC images through intensity thresholding. Then, the pixel intensity of 3D SSC images was normalized globally to the highest intensity level of the dataset. As many deep learning classifiers are sensitive to class imbalance, we augmented the imbalanced WBC dataset before training to balance the class occurrence frequencies [33]. We applied a rotational augmentation to the minority classes for the training dataset. For each 3D SSC image, a 2D SSC projection image was generated by collapsing the depth dimension of the corresponding 3D SSC image.

2.2.4 Deep learning models and validation

Three CNN models were used to take the SSC images as input and make the cell type prediction. ResNet is a state-of-the-art architecture in image classification. It eases the weight parameters optimization by reformulating the convolutional layers as learning residual functions, with the layer inputs as a reference. ResNet architecture converges faster and contains fewer parameters compared with other neural network architectures like VGG or InceptionNet, while ResNet could maintain a lower error rate compared with others [34]. In this work, a variant of 18 layers (ResNet18) was used as a benchmark architecture. We conducted a transfer learning with the ResNet18 model with the modification of the fully connected and Softmax layer. The output of the Softmax layer can be written as

$$\hat{\boldsymbol{y}}_{\boldsymbol{i}} = \frac{e^{x_i}}{\sum_{j=1}^{C} x_j}, i = 1,2, \ldots, C \quad (1)$$

where $\boldsymbol{x}$ is the input vectors, C is the number of classes.

In our work, the ResNet18 model optimizes the averaged cross-entropy loss between the predicted class and ground-truth class through mini-batch gradient descent. The averaged cross-entropy loss L_{CE} can be expressed as

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}_{\boldsymbol{i}}(\boldsymbol{x}) \cdot \log\left(\hat{\boldsymbol{y}}_{\boldsymbol{i}}(\boldsymbol{x})\right) \quad (2)$$

where $\boldsymbol{y}_{\boldsymbol{i}}$ is the ground truth class vector, $\hat{\boldsymbol{y}}_{\boldsymbol{i}}$ is the predicted class vector, and N is the data size in the mini-batch.

In comparison with the contracting architecture of ResNet, a "fully-connected" convolutional neural network, the so-called UNet architecture was applied for biomedical image classification and segmentation. UNet supplements the usual contracting network by successive up-sampling and convolution layers, and the high-resolution features from the contracting path are combined with the upsampling output. A successive convolution layer can then learn to assemble a more precise output based on the passed information [35]. UNet requires very few labeled images and has reasonable training time. In the meantime, recent research demonstrated that using an autoencoder and the latent space to conduct could improve classification performance [36], [37]. However, none has attempted to use the latent activation of a dense UNet architecture for both classification and image regeneration. In this work, we have developed two customized autoencoder models based on UNet architecture: (i) 2DCNN UNet that contains 2D convolution layers from single-channel 3D images and multi-channel 2D images; (ii) 3DCNN UNet that uses 3D convolution layers from multi-channel 3D images. For both UNet architectures, the contracting path takes the image as input. Image features are extracted by the convolution layers and encoded to subsequent layers through max pooling. A fully connected layer and Softmax layer are connected to the latent space to make a classification decision. The upsampling path takes the features from the latent space in combination with the high-resolution features that are passed from the convolution layers to generate an output image with the same dimension as the input image. This path is trained such that the generated images look similar to the input image while suppressing the noise. In both UNet architectures, the contracting path works as an encoder, while the upsampling path works as a decoder. We use a weighted loss that incorporates the mini-batch averaged cross-entropy loss between the predicted class and ground-truth class (Eqs. (2)) and the mean-square error loss between the input and generated output image pixel values. The mini-batch averaged mean-square error loss L_{MSE} can be expressed as

$$L_{MSE}(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{M}(x_{i,j} - \hat{x}_{i,j})^2 \quad (3)$$

Where $\boldsymbol{x}$ and $\widehat{\boldsymbol{x}}$ are the input image and generated image vectors, respectively, M is the flattened image vector dimension, and N is the data size in the mini-batch.

The weighted total loss L is defined as

$$L = w \cdot L_{CE} + (1 - w) \cdot L_{MSE} \quad (4)$$

where w is the weight coefficient to balance the loss function.

The overall structures of 2DCNN and 3DCNN UNet are illustrated in Figure 2.2.6.

A stratified 5-fold cross-validation (CV) approach is used to train the deep learning models and evaluate their performance. For each fold, training data were augmented by rotating the 3D image matrix to balance the class occurrence frequencies and used to train a model. The model was then validated using the instance from the validation set. The predictions made on the validation set were summarized in a confusion matrix per fold. more information on the CV training curves can be found in Figure 2.2.7-2.2.18. Apart from the confusion matrix, balanced classification accuracy was reported. The balanced accuracy $\overline{\sigma}$ is the arithmetic mean of class-specific accuracies and is calculated as

$$\overline{\sigma} = \frac{1}{C}\sum_{i=1}^{C} \sigma_i \quad (5)$$

where σ_i is the class-specific accuracy, and C is the number of classes. The balanced accuracy does not favor a classifier that exploits class imbalance by biasing toward the majority class [38].

For each classification experiment, ResNet18 and 2DCNN UNet use 2D SSC projections and 3D SSC images as the input to train the model separately, while 3DCNN UNet uses only 3D SSC images as the input. For 2DCNN UNet, the 3D SSC images were sliced along the depth direction and the 2D slices were treated as the multi-channel input of the 2DCNN UNet model. The deep learning models were trained for 100 epochs with the Adam optimization algorithm [39]. 2DCNN and 3DCNN UNet used a learning rate of 1e-4 and 5e-5 in the initial training epochs for human cancer cell line and leukocyte classification, respectively. ResNet18 adopted an initial training learning rate of 5e-6 for the first five epochs to avoid getting stuck into the local minimum and then increased to 5e-5. The exponential decay parameters for Adam optimizer were set as $\beta_1 = 0.9$, $\beta_2 = 0.999$. In order to allow the optimization to converge, the learning rate was reduced by half if the validation metric stopped improving for five epochs. The parameters of the trained deep learning models were stored and validated using the validation dataset to generate the confusion matrix of the classification results, and balanced classification accuracy was calculated based on the confusion matrix. In addition, the activation output of the fully connected layer was projected to a 2D plane as the t-distributed stochastic neighbor embedding (t-SNE) plot [40]

All deep learning models were implemented using the PyTorch framework and trained on a 12-core machine with an Intel® Core™ i9-10920X processor running at 3.50GHz. The deep learning experiments were run on an NVIDIA Titan RTX GPU with 24GB of VRAM.

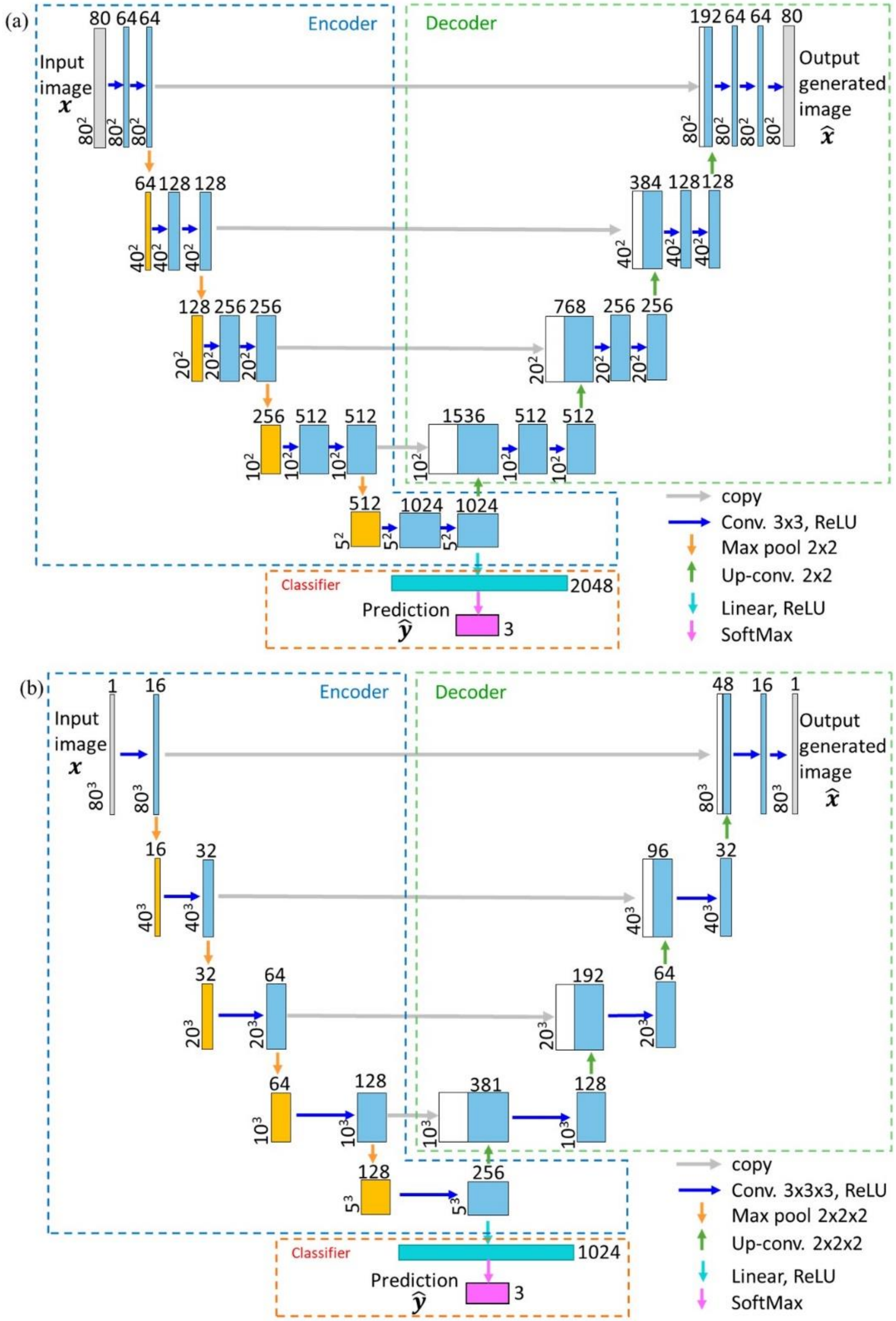


Figure 2.2.6. 2DCNN and 3DCNN UNet structures. (a) 2DCNN UNet structure with multi-channel 2D input. (b) 3DCNN UNet structure with single-channel 3D input. Each box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The feature map size is shown at the lower-left edge of the box. The arrows denote the different operations

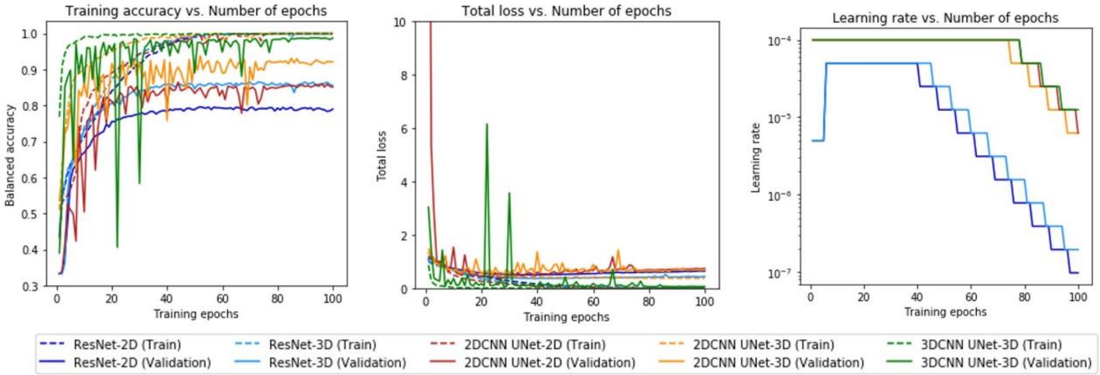


Figure 2.2.7. Learning curves and learning rate schedules during training for human cancer cell classification

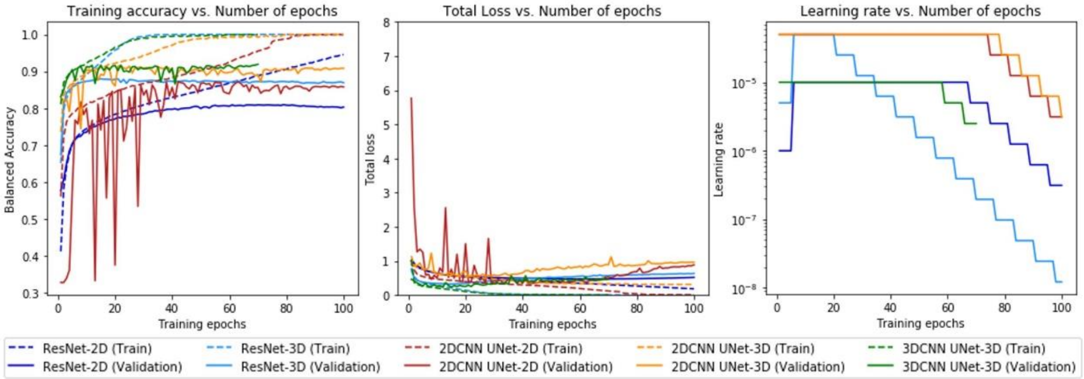


Figure 2.2.8. Learning curves and learning rate schedules during training for human leukocyte classification

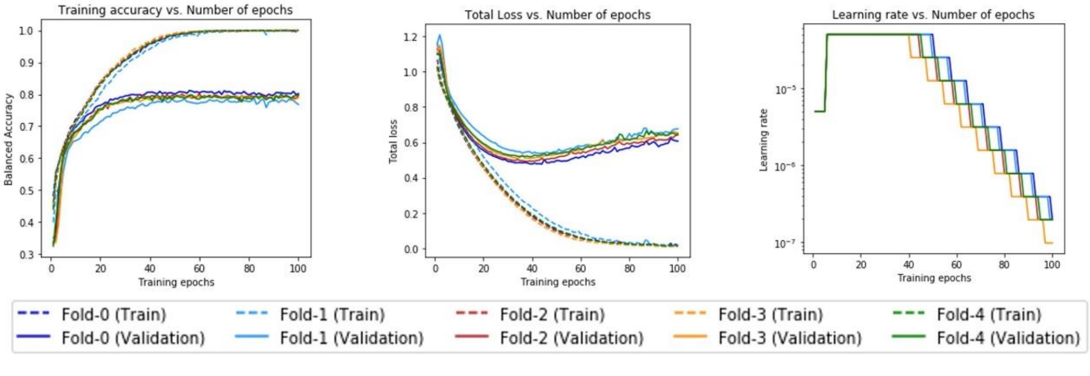


Figure 2.2.9. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human cancer cell classification using ResNet model with 2D input

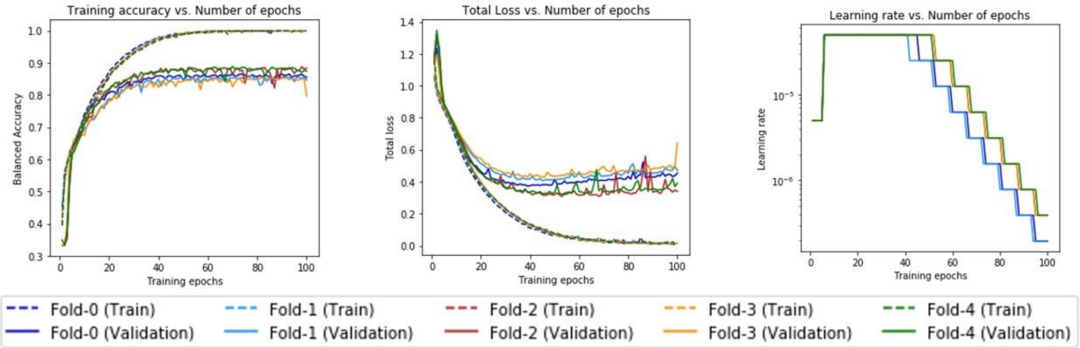


Figure 2.2.10. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human cancer cell classification using ResNet model with 3D input

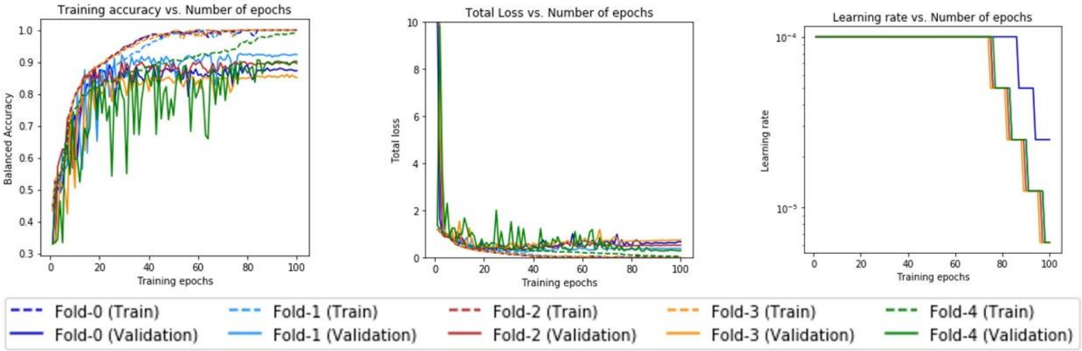


Figure 2.2.11. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human cancer cell classification using UNet2D model with 2D input

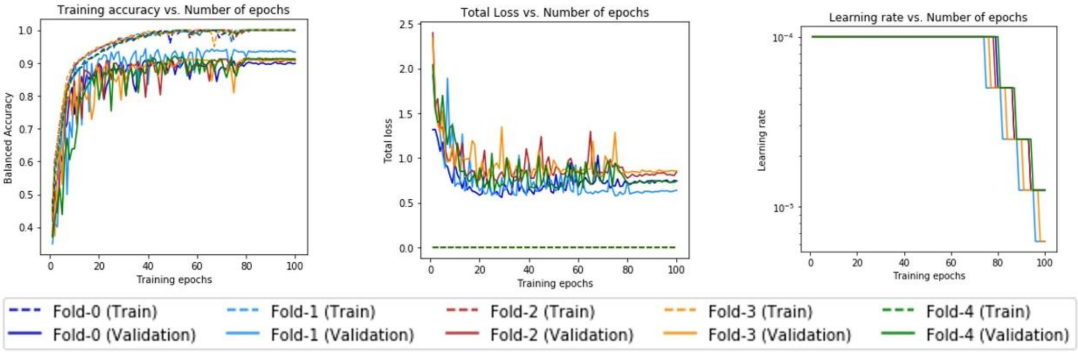


Figure 2.2.12. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human cancer cell classification using UNet2D model with 3D input

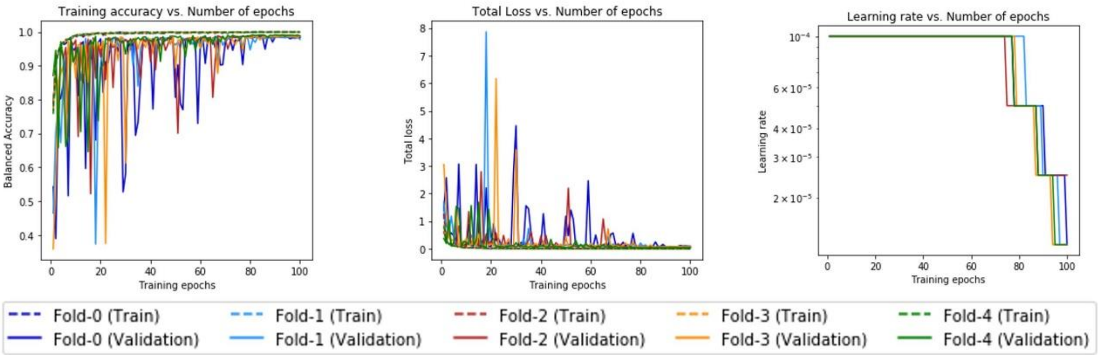


Figure 2.2.13. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human cancer cell classification using UNet3D model with 3D input

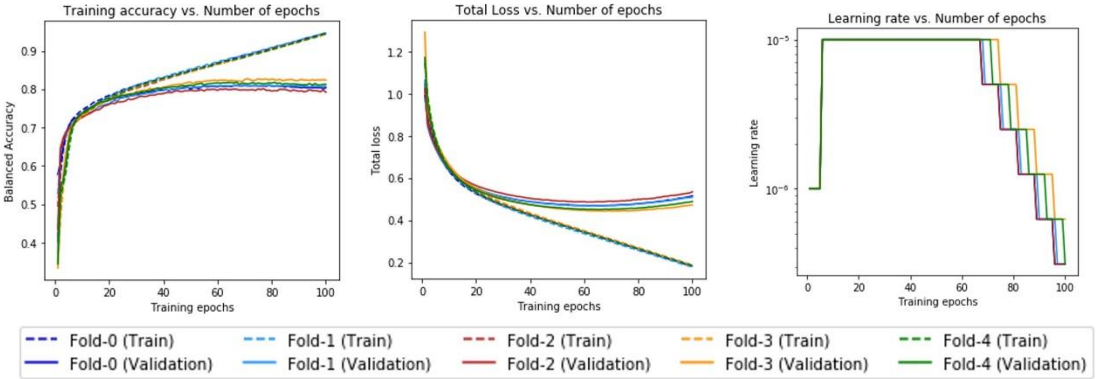


Figure 2.2.14. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human leukocyte classification using ResNet model with 2D input

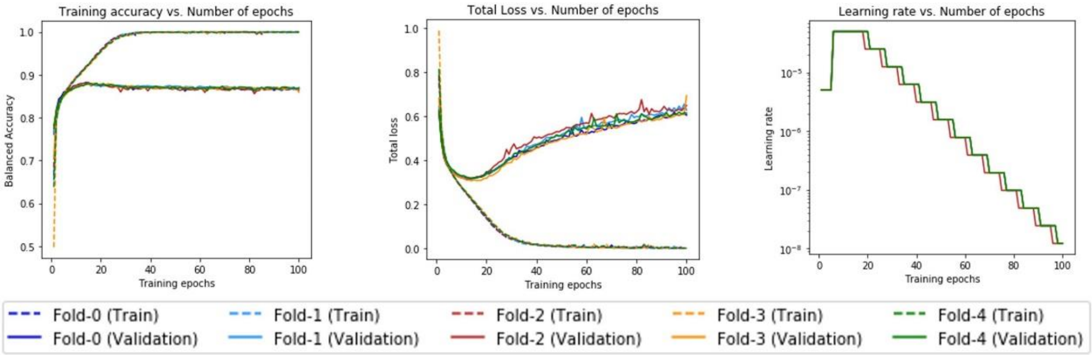


Figure 2.2.15. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human leukocyte classification using ResNet model with 3D input

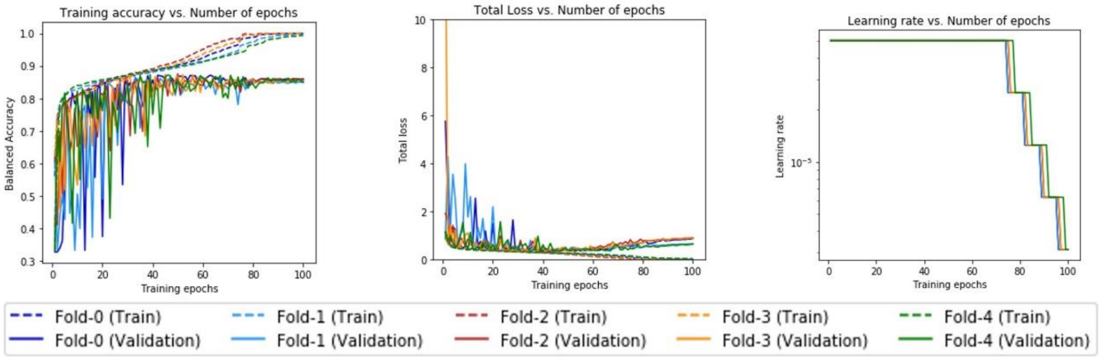


Figure 2.2.16. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human leukocyte classification using UNet2D model with 2D input

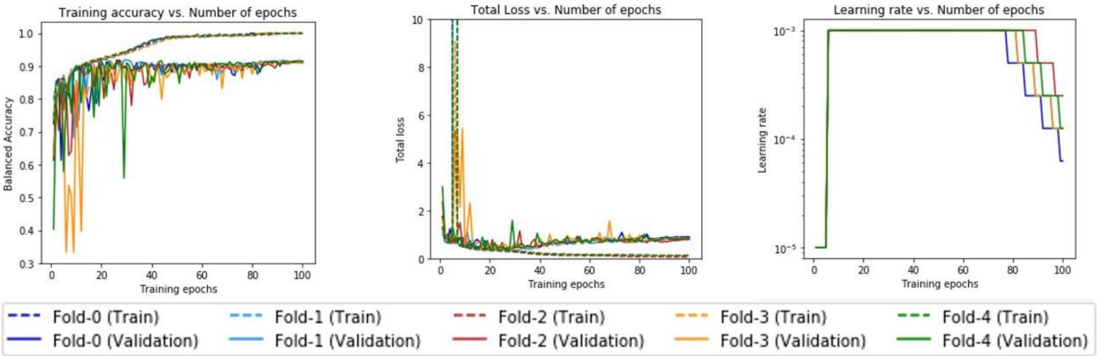


Figure 2.2.17. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human leukocyte classification using UNet2D model with 3D input

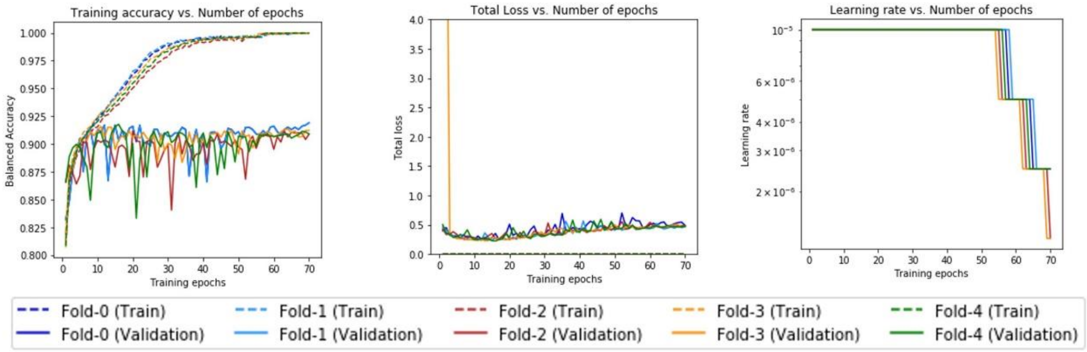


Figure 2.2.18. Learning curves and learning rate schedules for all folds in cross-validation acquired during training for human leukocyte classification using UNet3D model with 3D input

2.2.5 Fused UNet CNN for NASH cell characterization

To exploit the high information content in the dual-modality cell images from the 3D-IFC system, a customized autoencoder model, Fused UNet, was developed based on UNet architectures introduced in the previous section. The previous section examined the utilization of side-scattering images for cell classification. However, the side-scattering image only reflects the 90-degree scattering light profile of the cell, while the light loss information caused by the occlusion of the cell when it passes through the optical interrogation area can only be captured by the 2D transmission image. Therefore, combining the two imaging modalities could provide a complete light profile information of the cell, which leads to our 'Fused UNet' architecture. Fused UNet architecture combines two UNet structures (2D UNet and 3D UNet), each taking a single image modality as the input. Image features are extracted by the convolutional layers and encoded to subsequent layers by the max-pooling kernels during the contracting paths. The latent spaces from the two UNets are concatenated together with a Softmax layer to make a classification decision for the classifier. The images generated by the upsampling paths are also used to optimize the loss function during the model training process. The Fused UNet structure is illustrated in Figure 2.2.19.

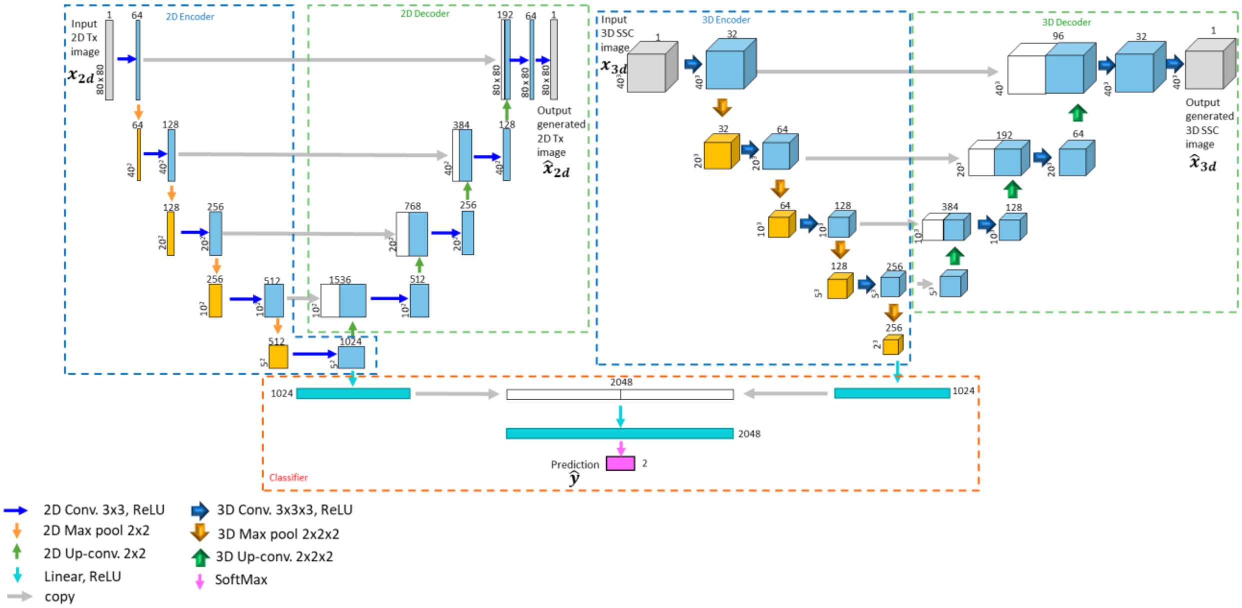


Figure 2.2.19. Fused UNet autoencoder model architecture

The output of the Softmax layer can be written as

$$\hat{\boldsymbol{y}}_i = \frac{e^{x_i}}{\sum_{j=1}^{C} e^{x_j}}, i = 1,2, \dots, C \quad (2)$$

where $\boldsymbol{x}$ is the input image pair vectors, C is the number of classes.

The Fused UNet uses a weighted loss that consists of the mini-batch averaged cross-entropy loss (Eqs. (3)) between the ground truth class and the predicted class and the weighted mean-square error loss (Eqs. (4)) between the input and generated output images pixel values. The averaged cross-entropy loss L_{CE} can be expressed as

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}_i(\boldsymbol{x}) \cdot \log\left(\hat{\boldsymbol{y}}_i(\boldsymbol{x})\right) \quad (3)$$

where $\boldsymbol{y}_i$ is the ground truth class vector, $\hat{\boldsymbol{y}}_i$ is the predicted class vector, and N is the data size in the mini-batch.

The mini-batch averaged mean-square error loss L_{MSE} can be expressed as

$$L_{MSE}(\boldsymbol{x}, \hat{\boldsymbol{x}}) = w_1 L_{MSE,2d} + (1 - w_1) L_{MSE,3d}$$

$$= \frac{1}{N}\sum_{j=1}^{N}\left(\frac{w_1}{M_{2d}}\sum_{i=1}^{M_{2d}}\left(x_{2d,i,j} - \hat{x}_{2d,i,j}\right)^2 + \frac{1-w_1}{M_{3d}}\sum_{i=1}^{M_{3d}}\left(x_{3d,i,j} - \hat{x}_{3d,i,j}\right)^2\right) \quad (4)$$

Where $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ are the input and generated image pair vectors, respectively, M_{2d} is the flattened transmission image vector dimension, M_{3d} is the flattened side-scattering image vector dimension, $L_{MSE,2d}$ is the mean-square error loss between the input and generated transmission image, $L_{MSE,3d}$ is the mean-square error loss between the input and generated 3D side-scattering image, N is the data size in the mini-batch, w_1 is the mean-square error loss weight assigned to the two image modalities.

The weighted total loss L is defined as

$$L = w_2 \cdot L_{CE} + (1 - w_2) \cdot L_{MSE} \quad (5)$$

where w_2 is the weight coefficient to balance the loss between the averaged cross-entropy loss and the mean-square error loss.

The same stratified 4-fold cross-validation (CV) approach used in the conventional machine learning approach is also applied to prepare the image pair dataset for the training of the Fused UNet autoencoder model and the performance evaluation. The predictions made on the validation set were summarized in a confusion matrix per fold. More information on the CV training curves can be found in Figure 2.2.20-2.2.21. In addition to the confusion matrix, the balanced classification precision, recall, and F_1 score are reported for cell morphology characterization. The balanced accuracy $\overline{\sigma}$ is the arithmetic mean of class-specific accuracies and is calculated as

$$\overline{\sigma} = \frac{1}{C}\sum_{i=1}^{C}\sigma_i \quad (6)$$

where σ_i is the class-specific accuracy, and C is the number of classes.

For the cell characterization experiment, the Fused UNet model was trained for 50 epochs with the Adam optimization algorithm [39]. The initial learning was set to 1×10^{-5} for the first five epochs and then increased to 5 × 10^{-5} to avoid getting stuck into the local minimum. The exponential decay parameters for the Adam optimizer were set as β1 = 0.99 and β2 = 0.9999. To allow the optimization to converge, the learning rate was reduced by half if the validation metrics stopped improving for five epochs.

The Fused UNet model was implemented using the PyTorch framework and was trained on a 12-core machine with an Intel® Core™ i9-10920X processor and an NVIDIA Titan RTX GPU with 24GB of VRAM.

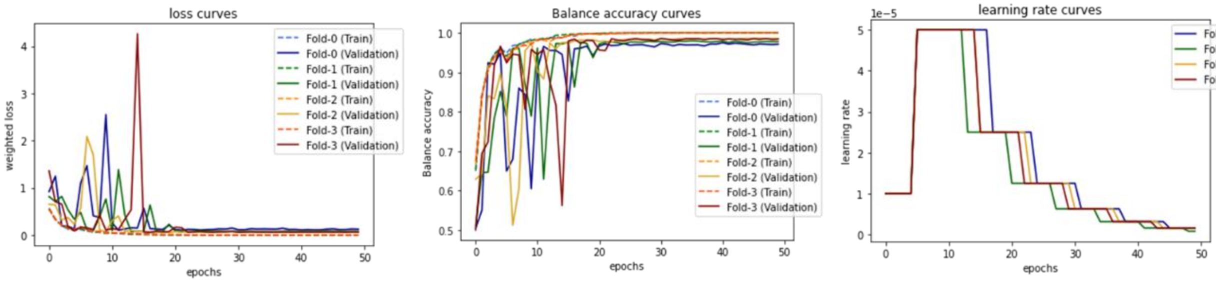


Figure 2.2.20. Fused UNet learning curves and learning rate schedules during training for stellate cells characterization

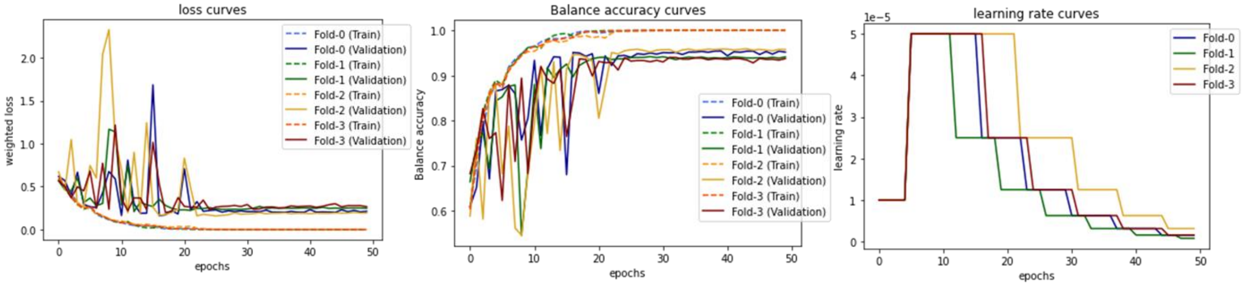


Figure 2.2.21. Fused UNet learning curves and learning rate schedules during training for endothelial cells characterization

2.2.6 Image segmentation UNet for intracellular structure segmentation

In addition to the cell type classification application, we developed an image segmentation UNet model to conduct pixel-level classification and delineate the intracellular structure based on the image captured from the 2D IFC system. The image segmentation UNet takes 2D transmission images and the binary segmentation maps generated from the fluorescence detection channels. Image features are extracted by the convolutional layers and encoded to subsequent layers by the max-pooling kernels during the contracting paths. The upsampling path takes the features from the latent space in combination with the high-resolution features that are passed from the convolution layers to generate a predicted segmentation mask

The image segmentation UNet uses a weighted loss that consists of the mini-batch averaged binary cross-entropy loss (Eqs. (1)) between the ground truth class and the predicted class and the Dice loss (Eqs. (2)) between the input and predicted segmentation mask pixel values. The averaged cross-entropy loss L_{CE} can be expressed as

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}\frac{1}{C}\sum_{c=1}^{C}\boldsymbol{x_{c,i}} \cdot log\ (\boldsymbol{\hat{y}_{c,i}}) \quad (1)$$

where $\boldsymbol{x_{c,i}}$ is the ground truth class vector, $\boldsymbol{\hat{y}_{c,i}}$ is the predicted class vector, and N is the data size in the mini-batch, C is the class number in the mini-bath.

The Dice loss L_{Dice} can be expressed as

$$L_{Dice} = 1 - \frac{1}{C}\sum_{c=1}^{C}\frac{2\sum_{i=1}^{N}x_{c,i}\hat{y}_{c,i}}{\sum_{i=1}^{N}x_{c,i}^{2}+\sum_{i=1}^{N}\hat{y}_{c,i}^{2}} \quad (2)$$

Where, $x_{c,i}$ and $\hat{y}_{c,i}$ are the input and predicted pixel values. N is the data size in the mini-batch, C is the class number in the mini-bath.

The weighted total loss L is defined as

$$L = w \cdot L_{BCE} + (1-w) \cdot L_{Dice} \quad (5)$$

where w is the weight coefficient to balance the loss between the averaged binary cross-entropy loss and the Dice loss.

The same stratified 5-fold cross-validation (CV) approach used in the conventional machine learning approach is also applied to prepare the image pair dataset for the training of the Fused UNet autoencoder model and the performance evaluation. The predictions made on the validation set were summarized using the intersection-over-union (IoU) score and F_1 score for each CV fold.

The image segmentation UNet structure is illustrated in Figure 2.2.22.

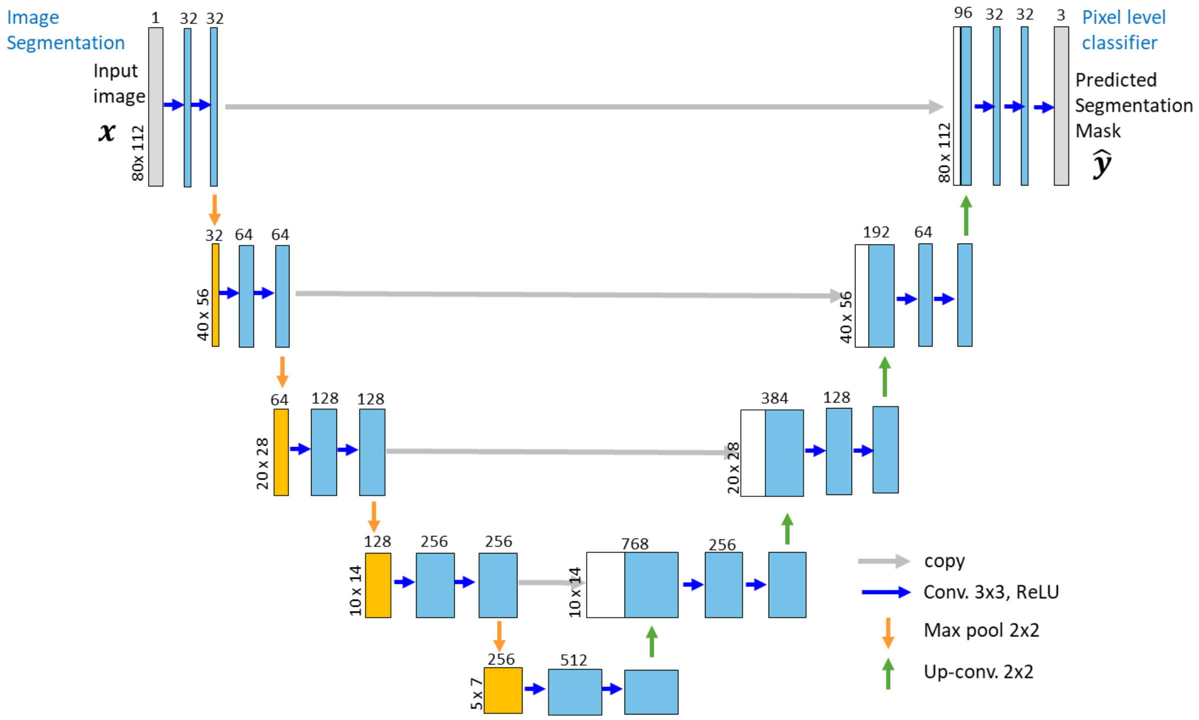


Figure 2.2.22. Image segmentation UNet architecture

2.3 Results

We started with the ResNet18 model for the classification experiments by setting a benchmark performance on the human cancer cell lines and white blood cell datasets, trained on 2D SSC projection images and 3D SSC tomographic images, respectively. We then used our customized 2D CNN and 3D CNN UNet models to conduct the same classification and found that they achieved better performance for both datasets than the benchmark ResNet18 model with the feedback of the upsampling path in the model architectures. We also found that 3D SSC tomographic images as input outperform 2D SSC projection images substantially for the same deep learning model. For NASH disease cell type characterization, we found that the Fused UNet returns a high classification F_1 score compared with the conventional machine learning results based on morphological statistics. In addition, our image segmentation UNet returns high IoU and F_1 scores for cell body and nuclear area segmentation for CHO-K1 cells.

2.3.1 Classifying human cancer cells

This experiment aims to evaluate the classification performance of the SSC image for cells with different sizes and internal structures. The confusion matrics and corresponding t-SNE visualizations for all classification experiments are presented in Figure 2.3.1. It is observed that all deep learning models were able to differentiate three cancer cell types using the SSC images, while the performance varies with the model architecture and the input dimension. Regarding the model performance, both UNet models outperform the benchmark ResNet18 by a large margin in terms of 2D and 3D input classification. For 2D input, 2DCNN UNet shows better classification accuracies for all cell types (2DCNN UNet: 0.865, versus ResNet18: 0.796, balanced accuracy). For 3D input, 3D CNN UNet achieved a very high performance among the three architectures (3DCNN UNet: 0.988, 2DCNN UNet: 0.955, versus ResNet18: 0.865, balanced accuracy). The balanced accuracy improvement of using the 3D input instead of 2D input for ResNet18 and 2DCNN UNet is 8.67% and 10.40%, respectively. In addition, we observed a more apparent separation in the t-SNE plots with 3D input compared with 2D input. The improvement of classification performance using 3D input implies that 3D SSC tomographic images contain more spatial information, especially along with the depth dimension, than 2D SSC projection images. Such information is beneficial for cell type detection based on side-scattering images.

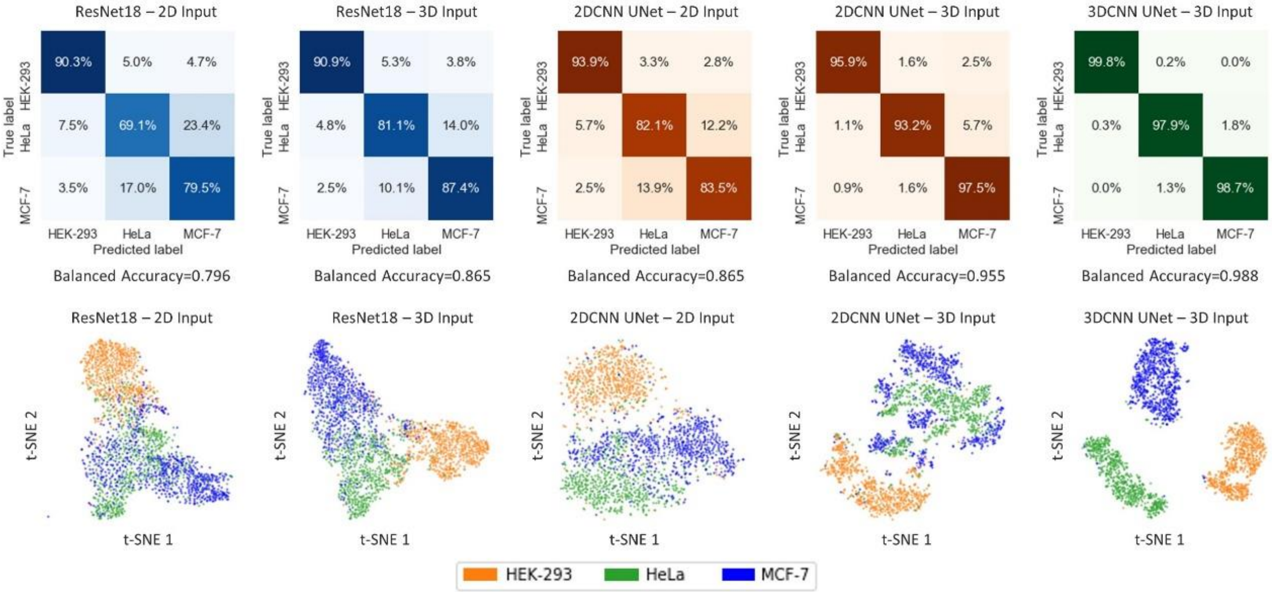


Figure 2.3.1. Confusion matrices and t-SNE visualizations from cross-validation experiments on the human cancer cell dataset

2.3.2 Classifying human white blood cells

WBC differential count is a widely adopted clinical test that measures the percentage number and percentage of the WBC types [41]. Conventional methods require fluorescent antibody tagging that is known to label the WBC population differentially. However, this method provides low content cell phenotype information and can only be applied to the cell type with the known antibody tagging. Cells in the WBC population have different sizes and different cell types vary in the intracellular structure. Distinguishing WBCs using the stain-free morphological information could avoid the biomarker labeling process and enable the discovery of new subtypes that cannot be labeled by fluorescent biomarkers. In this experiment, we demonstrate a three-part WBC cell type classification to identify granulocytes, lymphocytes, and monocytes based on the SSC images.

We conducted the same classification experiments in the cancer cell line dataset using the benchmark and customized models with both 2D and 3D SSC inputs. Figure 2.3.2 shows the confusion matrices and corresponding t-SNE visualizations for all classification experiments. From the results, all deep learning models were able to classify the WBC, though the performance varies with model architectures and input dimensions, showing similar trends compared with the cancer cell line dataset. In terms of model architectures, 3DCNN UNet achieved the highest performance compared with 2DCNN UNet and the benchmark ResNet18 with 3D input (3DCNN UNet: 0.923, 2DCNN UNet: 0.918, versus ResNet18: 0.883, balanced accuracy). For the experiments on the input dimension, models with 3D input (0.923, balanced accuracy) outperform the ones with 2D input for both ResNet18 (0.836, balanced accuracy) and 2DCNN UNet models (0.883, balanced accuracy). Apart from the accuracy performance, we also found that the t-SNE visualization has better separation among each cell type for the models with 3D input. Overall, we found that the classification performance for the WBC dataset is not as high as the cancer cell line dataset. One of the contributing factors is that there is a more significant variation of the cell morphology for the WBC dataset compared with the cancer cell line dataset. The primary WBC types identified in this experiment could be further divided into several subtypes. For example, granulocytes can be further separated into neutrophils, eosinophils, basophils. Nevertheless, we observed multiple well-separated clusters within the single cell type, especially in the t-SNE visualization of 2DCNN and 3DCNN UNet with 3D input.

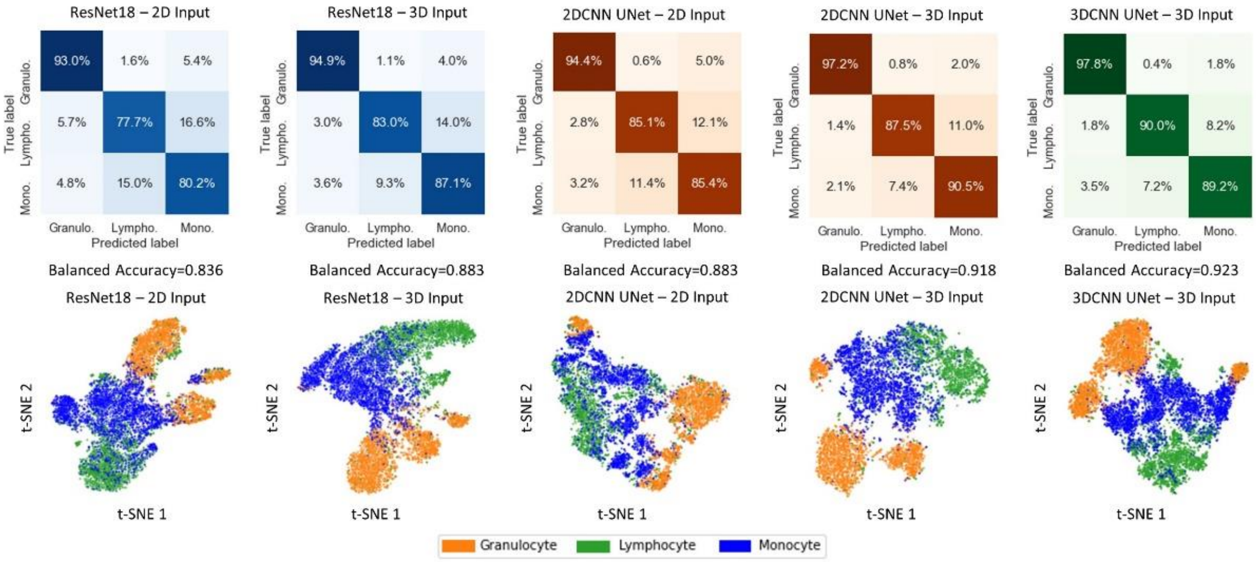


Figure 2.3.2. Confusion matrices and t-SNE visualization from cross-validation experiment on the human white blood cell dataset

2.3.3 Fused UNet CNN NASH cell characterization based on the dual-modality image input

We conducted the cell characterization on both HSC and LEC cells datasets using the Fused UNet CNN model. The confusion matrices for HSC and LEC cells characterization are presented in Figure 2.3.3 and Figure 2.3.4, respectively. The results show that our Fused UNet model was able to characterize both HSC and LEC cells with high accuracy between NASH and healthy donors with the dual-modality image inputs acquired by the 3D IFC system. For HSC characterization, Fused UNets return a high balanced characterization F_1 score (0.975-0.987, mean F_1 score: 0.982) for all CV folds. The characterization of LEC cells generally returns a balanced characterization F_1 score (0.939-0.959, mean F_1 score: 0.949) for all CV folds. Our result shows that we can separate the cells based on the physical structure, which suggests that there are fundamental changes that occur during the disease process that alter the structure and likely the function of these cells during disease progression.

For comparisons between the Fused CNN UNet and conventional machine learning, we also conducted the cell characterization based on the morphological features of HSC and LEC cells datasets using conventional machine learning techniques. Our experiment shows that conventional machine learning classification algorithms generally return relatively high F_1 scores for the HSC cells dataset but yielded lower F_1 scores for the LEC dataset. The best conventional classifier model based on morphological features is the Gaussian process classifier model. It returns a mean F_1 score of 0.905 for the HSC cells dataset and a mean F_1 score of 0.81 for the LEC cells dataset. Figure 2.3.5 shows the comparison between the Fused CNN UNets and different conventional machine learning methods. The Fused CNN UNet outperforms all conventional machine learning algorithms by a significant margin for the detection of early-stage NASH.

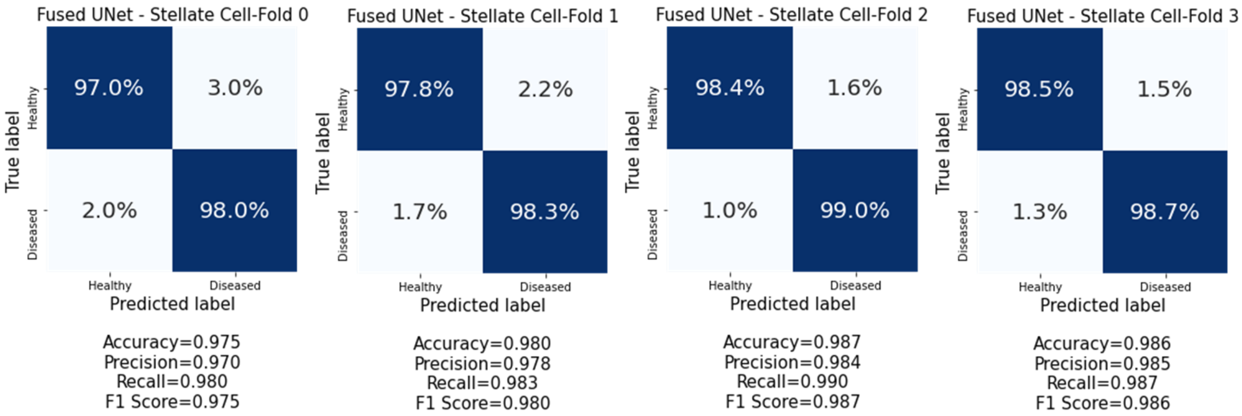


Figure 2.3.3. Confusion matrices from cross-validation experiments on the HSC characterization dataset.

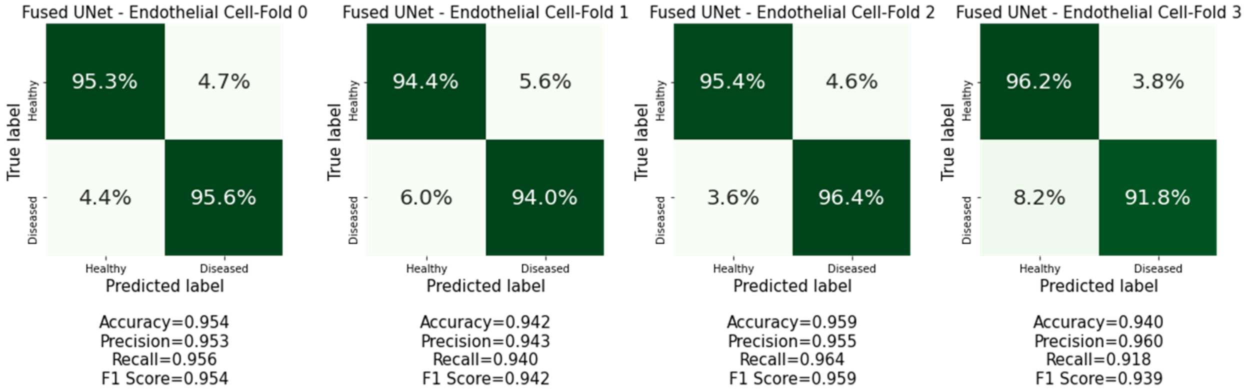


Figure 2.3.4. Confusion matrices from cross-validation experiments on the LEC cells characterization dataset.

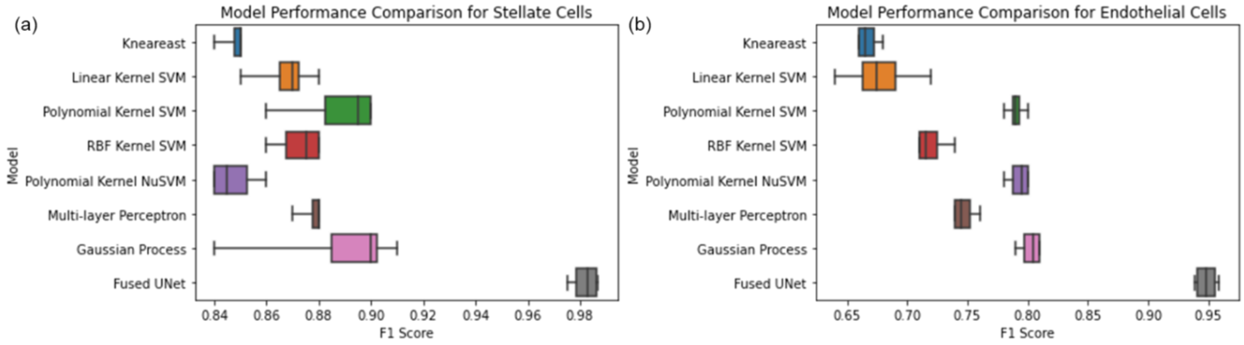


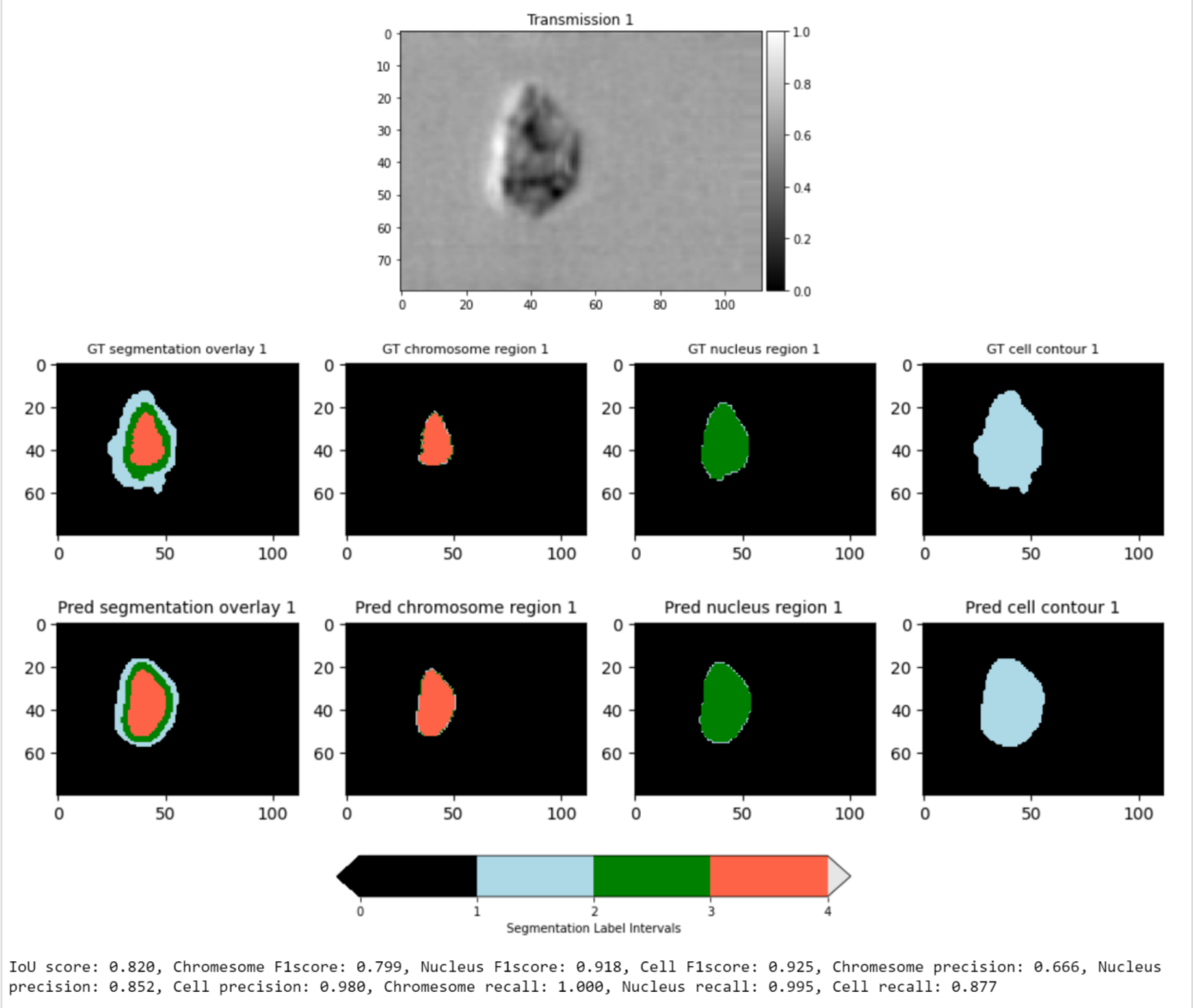
Figure 2.3.5. Model performance comparison. (a) Model performance comparison for HSC cells characterization; (b) Model performance comparison for LEC cells characterization.

2.3.4 Image segmentation UNet CHO-K1 cell body and nuclear area segmentation

We conducted pixel-level cell body and nuclear area segmentation based on the 2D transmission images of CHO-K1 cells captured by the 2D-IFC system. We use the image segmentation UNet to predict the cell body area, nucleus, and chromosome area using a data set of 10,000 images. An example visualization of image segmentation UNet prediction visualization is presented in Figure 2.3.6. The predicted segmentation masks from the 2D transmission images are very similar to the segmentation masks generated from the ground truth fluorescence channel. The IoU and F_1 score metrics of all CV folds are included in Table 2.3.1. It is observed that the image segmentation UNet can accurately predict the cell area and nucleus area of CHO-K1 cells with a high F_1 score (mean cell body area F_1 score: 0.914, mean nucleus area F_1 score: 0.923). The mean IoU score for the prediction mask is 0.826.

Table 2.3.1 CHO-K1 cell body and nuclear segmentation performance metrics

Performance metrics	**Fold-0**	**Fold-1**	**Fold-2**	**Fold-3**	**Fold-4**
IoU score	0.822	0.831	0.828	0.823	0.825
Chromosome area F_1 score	0.817	0.826	0.822	0.820	0.821
Nucleus area F_1 score	0.921	0.926	0.924	0.922	0.922
Cell body area F_1 score	0.912	0.916	0.915	0.912	0.913



```
IoU score: 0.820, Chromesome F1score: 0.799, Nucleus F1score: 0.918, Cell F1score: 0.925, Chromesome precision: 0.666, Nucleus
precision: 0.852, Cell precision: 0.980, Chromesome recall: 1.000, Nucleus recall: 0.995, Cell recall: 0.877
```

Figure 2.3.6. Example CHO-K1 cell segmentation output visualization compared with the ground truth segmentation map

2.4 Discussion

Leveraging the unique 3D tomographic side-scattering imaging capability of the 3D-IFC system, we present an intelligent cell type analysis workflow enabled by customized 2DCNN and 3DCNN UNet deep learning neural networks, demonstrating for the first time that 3D tomographic side-scattering patterns can be applied to distinguish different cell types based on the morphologically separable features. Two multi-class cell type classification experiments have been demonstrated with high performance using only the SSC images from biological cells. For the human cancer cell type experiment, 3DCNN UNet achieves an accuracy of 98.8% for classifying HEK-293, HeLa, and MCF-7 cells. In the human white blood cell experiment, 3DCNN UNet can differentiate granulocyte, lymphocyte, and monocyte from the WBC population with an accuracy of 92.3% and the potential to further identify subtypes. In the NASH cell characterization experiment, Fused UNet can differentiate the healthy and diseased HSE and LSC with an average F_1 score of 0.982 and 0.949, respectively. In the CHO-K1 cell body and nuclear segmentation experiment, the image segmentation UNet can accurately delineate the cell body and nucleus area of CHO-K1 cells with high average F_1 scores (cell body: 0.914, nucleus: 0.923).

The current 3D imaging IFC has some constraints that limit its performance for label-free cell classification and cell type discovery.

A significant constraint is the image resolution for the SSC images. The current optical resolution for the 3D-IFC system is 1µm, 2 µm, and 2 µm for the horizontal, vertical, and depth directions. An improvement in the optical resolution of the 3D-IFC system can improve the resolving power of the subtle morphological differences among different cell types, which might be beneficial for identifying cell types that are challenging for the current IFC system to identify morphologically (e.g., helper T cells versus killer T cells)

In terms of the machine learning aspect, a natural next step for future work is to apply deep learning methods with the aim of increasing the performance of the model. The current deep learning model uses simple data preprocessing approaches. The impacts on different data preprocessing approaches to deep learning performance have not been quantified yet. In addition to the data preprocessing methods, one could examine different deep learning approaches to improve model performance. Semi-supervised, unsupervised learning approaches and other methods of dimensionality reduction and visualization could be used for cell type analysis. Such approaches could potentially define new subpopulations and provide insights when combined with quantitative and qualitative clustering techniques. Furthermore, semi-supervised and unsupervised machine learning approaches can be applied to further investigate the label-free classification potential of new cell type discovery and improve the supervised machine-learning models and provide a better understanding and explanation of the morphology-based cell type characterization and classification studies.

Chapter 2, in part, is a reprint of the material as it appears in APL Photonics 2020, R. Tang, Z. Zhang, X. Chen, L. Waller, A. Zhang , J. Chen, Y. Han, C. An, S. Cho and Y.-H. Lo. "3D side-scattering imaging flow cytometer and convolutional neural network for label-free cell analysis". The dissertation author was the first author of this paper.

Chapter 2, in part, is a reprint of the material as it is under peer review in Scientific Reports 2022, R. Subramanian, R. Tang (co-first), Z. Zhang, V. Joshi, J. Miner, Y.-H. Lo. “Multimodal NASH prognosis using 3D imaging flow cytometry and artificial intelligence to characterize liver cells”. The dissertation author was the co-first author of this paper.

Chapter 3 Low-latency label-free image-activated cell sorting using fast deep learning and AI inferencing

Classification and sorting of cells using image-activated cell sorting (IACS) systems can bring significant insight into biomedical sciences. Incorporating deep learning algorithms into IACS enables cell classification and isolation based on complex and human-vision uninterpretable morphological features within a heterogeneous cell population. However, current realizations of deep learning–assisted IACS systems either have insufficient capabilities or complicated implementation to make the utilization of such systems unfeasible for wide adoption in biomedical research. Here, we present image-activated cell sorting by applying fast deep learning algorithms to conduct cell sorting without labeling. The overall sorting latency, including signal processing and AI inferencing, is less than three milliseconds, and the training time for the deep learning model is less than 30 minutes with a training dataset of 20,000 images. The utility of the system was demonstrated in a 3-part white blood cell sorting, showing purities of 92.0%, 89.05%, and 98.4%, for lymphocyte, monocyte, and granulocyte, respectively, using the label-free, brightfield cell images captured by the system. The system provides a compact, sterile, low-cost, label-free, and low-latency cell sorting solution based on real-time AI inferencing and fast training of the deep learning model.

3.1 Introduction

Image-based detection, classification, and separation of target cells among the total cell population can bring phenomenal insight to biomedical research and application [9], [19], [42]–[46]. Conventional fluorescent-activated cell sorting (FACS) technology optically interrogates each cell in a single-cell flow stream and isolates cells based on scattering and fluorescence intensity features of the interrogated cells [47], [48]. In comparison, image-activated cell sorting (IACS) systems can classify and sort cells based on spatial features obtained from cell images, which offer much greater information content than conventional FACS which is limited to a single value per parameter. By extracting the spatial and morphological features carried by light transmission, scattering, and fluorescent properties of cells, IACS can classify and isolate the targeted cell types from a large heterogeneous cell population using image-feature based gating (e.g., cellular size and shape, nuclear size, and shape, nucleus-to-cytoplasm ratio, DNA and RNA localization, cellular organelle localization, cellular aggregation, as well as non-intuitive features) [49]–[52]. The emergence of the image-activated cell sorting (IACS) technique provides a powerful biomedical research tool for cell cycle analysis, cell-cell interaction, protein localization, DNA and RNA localization, and exploiting the relationship between cellular phenotype to genotype [45], [46], [53]–[57].

The IACS system requires real-time data processing and sorting actuation to process high-content image data at a high data transfer rate and extract many image-related features based on which sorting decisions are made [53], [58]. The computing power of the processor limits the number of cell image features that can be extracted in real-time as many image-related features require heavy computation [43], [58], [59]. Perhaps the most important concern is that in some cases, cell phenotypical and morphological features can be complex and convoluted, not resolvable or correctly identifiable by human vision or some subjective criteria, partly because humans can only process a very small set of images out of a very large sample size. As a result, mathematical representations of image features driven by human-vision-based gating can have deficiencies and miss important biological insight [9], [22], [23]. In this paper, we try to address this limitation by artificial intelligence (AI) derived gating and sorting with real-time AI inferencing.

With the recent advances in deep learning algorithms and artificial intelligence (AI) computing hardware, convolutional neural networks (CNN) can solve complex image-driven pattern recognition problems. Conducting a real-time AI model inference computation with millisecond latency has become feasible with the state-of-art graphical processing unit (GPU) hardware [12], [49], [60]–[65]. The recent advent of AI image-activated cell sorting (AI-IACS) achieved cell sorting based on CNN predicted sorting decision at a speed of 82.8 events per second (eps) with a sorting latency of 32 ms using a field-programmable gate array (FPGA) processor, one personal computer (PC) for distributed data processing, and a two-PC cluster deep learning module[53]. An updated version based on the same design achieved a cell sorting throughput of 1133 eps using a nine-PC cluster deep learning processing module for parallel computing. Increasing the PC cluster size to achieve a higher throughput has limited scalability due to cost and complexity. An alternative hardware design of the IACS system used a high-speed CMOS camera and a fast deep learning algorithm to reduce the CNN model inference time to less than 3 ms and the overall sorting latency to 5.5 ms [57]. However, the camera-based optics design suffers from sensitivity and motion blur issues, and the system does not support fluorescence detection due to its optics design. Apart from hardware limitations, none of the reported systems disclosed the training time for their CNN model in the learning process even though the CNN model training turnover time is a critical factor in the overall workflow.

In this chapter, we present a low-latency image-activated cell sorting platform based on a fast, deep-learning approach and low-latency AI inference. The system enables label-free cell classification and sorting by AI gating, while it also offers fluorescent imaging for labeled applications. Our system integrates a scanning microscopy technique (200-350 kHz scanning rate) with a disposable microfluidic chip containing an on-chip piezoelectric (PZT) cell sorting actuator. To leverage the photocounting sensitivity of the photomultiplier tubes (PMTs), the transmitted and fluorescent light was detected by the PMTs as a temporal readout. The detected cell images were reconstructed from the temporal signals by a field-programmable-gate-array (FPGA) and then sent to a PC with a dedicated GPU for conducting real-time AI model inference based on an optimized UNet CNN autoencoder model [35], [66]. Instead of running conventional image feature algorithms which varies in computing power need, the image features were extracted by a deep learning model in less than 1ms to enable real-time cell sorting with a deterministic computing power request. The entire sorting latency, including the data processing time and piezoelectric actuation time, was less than 3 ms for a sorting throughput of greater than 200 eps. Unlike most IACS systems that require extraction of image features in real time to generate image features and scatter plots for “sort gate” determination, our system replaces the conventional image feature extraction unit with an AI module and uses machine vision to automatically extract image features in combination with user inputs in the offline analysis. The AI module can learn the image features through a fast-training process and use the trained model to make sorting decisions. The entire learning and sorting gate definition process takes less than 30 minutes, which is the shortest time among any AI learning for IACS to the best of our knowledge.

To demonstrate the system's capability, we used the deep learning framework to sort polystyrene beads and three classes of human white blood cells without labeling (cells and beads were fluorescently labeled only for post-sorting evaluation of the performance of the AI inference model). Using brightfield images of cells without accessing any fluorescent signals, our system achieved 96.6% sorting purity for 7 µm and 15 µm beads, and sorted granulocytes, lymphocytes, and monocytes in human leukocyte samples with 98.4%, 92.0% and 89.05% purity, respectively. With an optimized CNN model and a small dataset of around 20,000 images, the optimized three-part human white blood cell classification CNN model training can be completed and ready for deployment in less than 40 minutes using a single PC, showing a marked improvement in training and inferencing time and computing resources compared to any AI-based IACS demonstrated to date.

3.2 Methods

3.2.1 Design of the imaging system

Suspended cells are hydrodynamically focused to the center of the microfluidic channel by a sheath flow in the microfluidic chip [67]. An acousto-optical deflector (AOD) (OAD948, Isomet) controlled scanning 488nm laser (W488-25FS-025, 25mW, PIC) optically interrogates each cell in the single-cell core flow stream. The beam size is adjusted by a series of beam-shaping lenses. Two 10x (0.28NA Plan Apo, Mitutoyo) objective lenses are positioned on the opposite sides of the microfluidic chip. The illumination laser forms a Gaussian illumination laser beam with a focal depth of approximately 25 µm and a full-width-half-maximum (FWHM) circular spot size of 1.6 µm on the object plane where the cells are being examined. A series of dichroic mirrors and band-pass optical filters separate the transmitted laser light and laser-excited fluorescent light into different PMT detection channels. The laser scanning range and scanning speed can be adjusted to accommodate samples of different cell sizes for a suitable image field of view. The maximum field of view of the system is 60 x 60 µm and the maximum laser scanning speed is 350kHz. The optical schematics and CAD layout is illustrated in Figure 3.2.1. All the image acquisition and sorting experiments demonstrated in this paper were conducted at a laser scanning frequency of 200kHz and an image field of view of 35 x 35 µm.

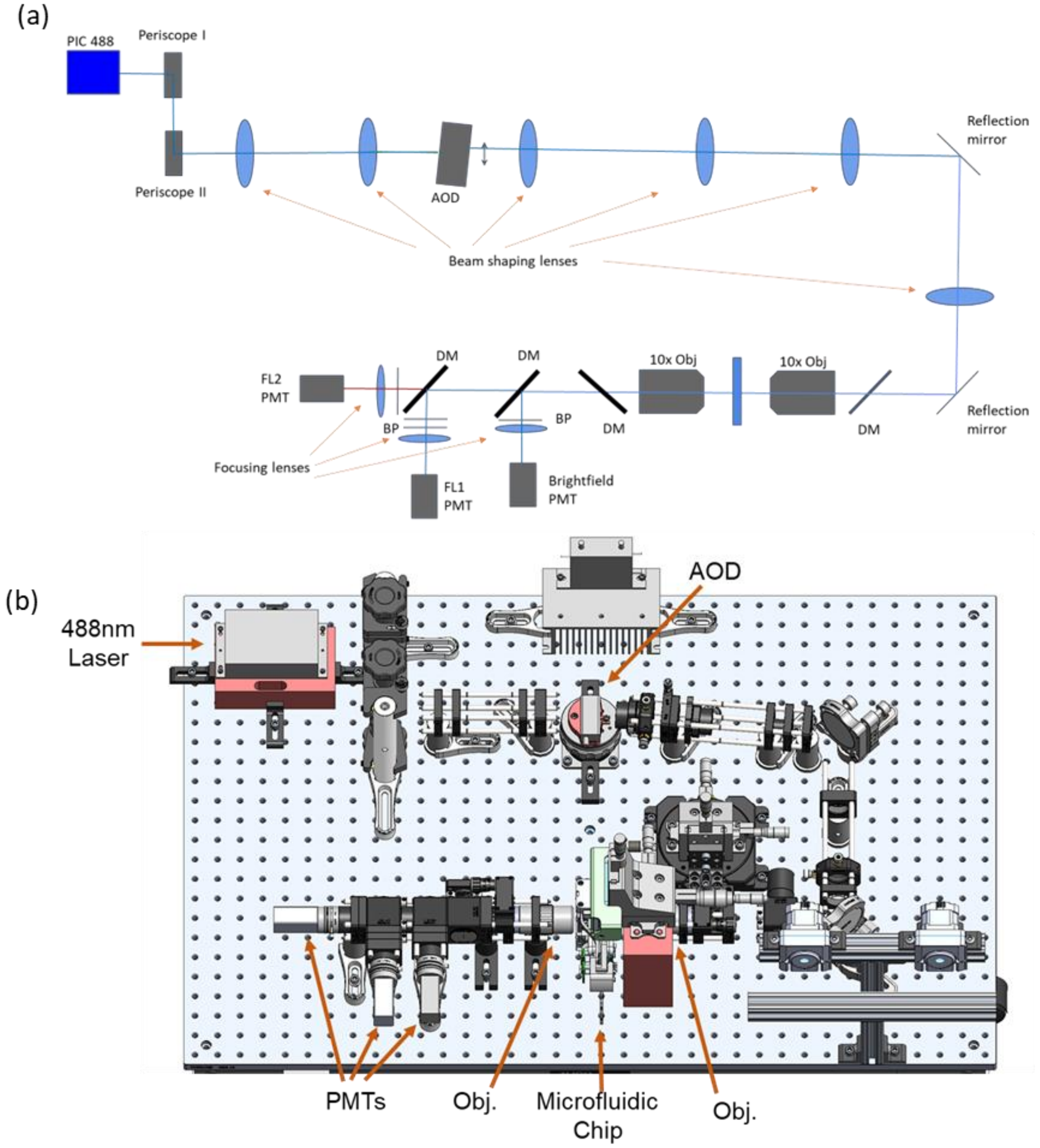


Figure 3.2.1. Major optical components of the low-latency IACS platform. (a) Optics schematics of the ultrafast raster scanning microscope. (b) CAD layout of the major optical components in low-latency IACS platform

We conducted optical calibration experiments to measure the illumination spot size, and depth of focus with a CMOS camera (DCC1645C, Thorlabs). The detection optics resolution limit was measured using a high-resolution optical test target (HIGHRES-1, Newport). The measured spot size, illumination beam depth of focus, and detection optics resolution limit are presented in

Supplementary Figure 3.2.2–3.2.3. The detection optics was diffraction-limited by the objective lens numerical aperture (NA) and excitation laser light wavelength.

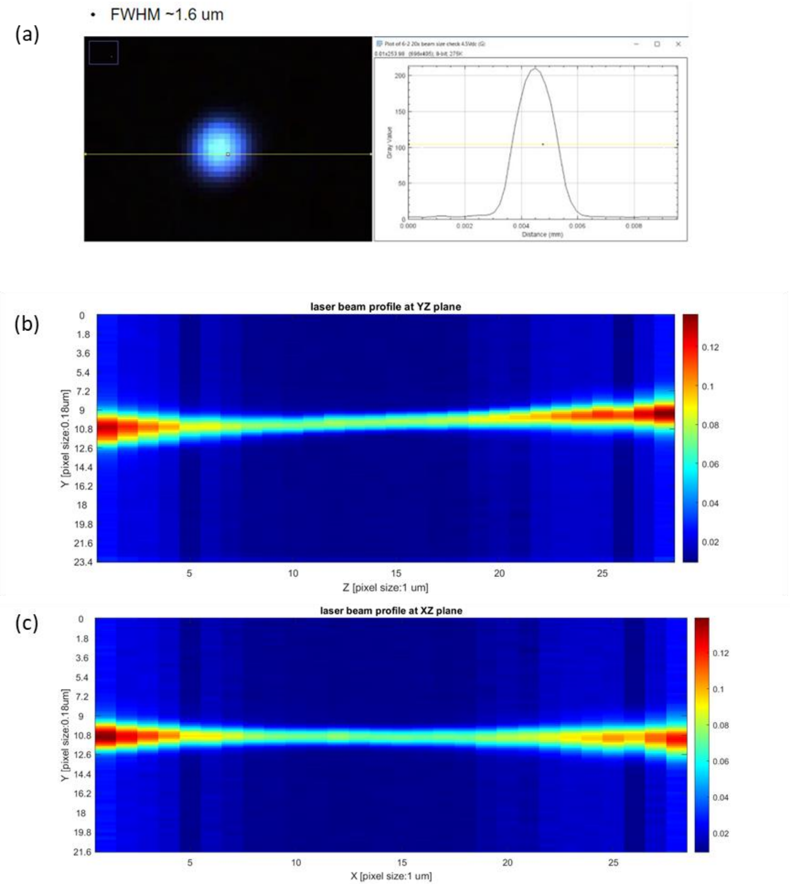
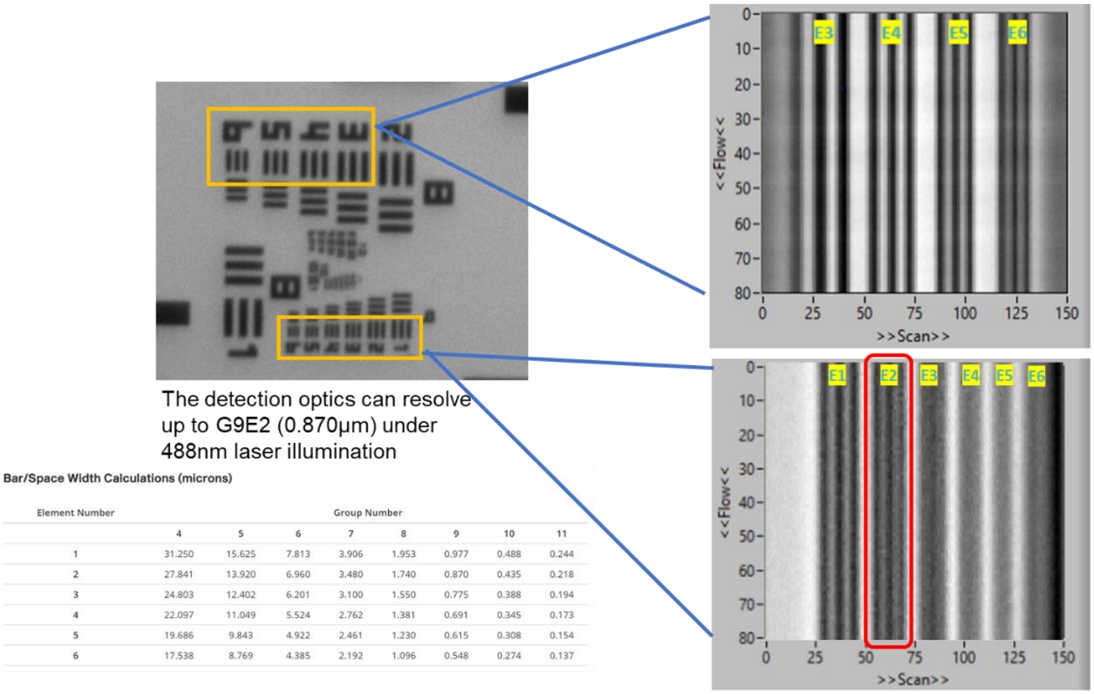


Figure 3.2.2. Optical performance measurement of low-latency IACS system. (a) Illumination spot size measurement. (b) Measured illumination light depth of focus measurement at YZ plane. (c) Measured illumination light depth of focus measurement at XZ plane. X: laser scanning direction, Y: cell traveling direction, Z: laser propagation direction



The detection optics can resolve up to G9E2 (0.870μm) under 488nm laser illumination

Bar/Space Width Calculations (microns)

Element Number	Group Number							
	4	5	6	7	8	9	10	11
1	31.250	15.625	7.813	3.906	1.953	0.977	0.488	0.244
2	27.841	13.920	6.960	3.480	1.740	0.870	0.435	0.218
3	24.803	12.402	6.201	3.100	1.550	0.775	0.388	0.194
4	22.097	11.049	5.524	2.762	1.381	0.691	0.345	0.173
5	19.686	9.843	4.922	2.461	1.230	0.615	0.308	0.154
6	17.538	8.769	4.385	2.192	1.096	0.548	0.274	0.137

Figure 3.2.3. Detection optics resolution limit measurement with resolution target under scanning laser illumination

3.2.2 Real-time data processing

To achieve low data processing latency with AI inference, an FPGA-CPU-GPU hybrid design is adopted to meet such a requirement. The architecture of the real-time data processing system is illustrated in Figure 3.2.4. The digitizer (NI-5783) samples and converts the voltage waveforms at a sampling speed of 25MSps. The digitizer continuously streams the waveforms to FPGA (PXIe-7975R) which segments the streaming waveforms by a moving sum cell detection algorithm and then undergoes image reconstruction using a temporal-spatial transformation algorithm [8]. FPGA corrects the image phase shift due to the electronic delay between the AOD control signal and detected PMT readout waveforms. The reconstructed images are then transferred through the PCIe bus to the dynamic random-access memory (DRAM) in the AI module consisting of a stand-alone multi-core PC workstation with a dedicated Nvidia GPU module. The AI module offers a graphical user interface (GUI) that displays the reconstructed images. Two operation modes are provided to users. In analysis mode, users can save the image data to an internal or external solid-state storage disk for offline image processing and AI model training. In sorting mode, users can set the sorting criteria (i.e., cell class, confidence level) and the pre-trained AI model conducts real-time inference to automatically determine the cell class along with a prediction confidence level. After generating the sorting decision, AI modules send the sort trigger to the sorting module. The sorting module subsequently triggers the on-chip PZT actuator to deflect the cell to user-defined downstream channels. An optical sorting verification detector detects the sorting outcome and sends the feedback signal to the AI module to show users the sorting yield on the GUI. The real-time data processing software is developed based on LabVIEW with a customized Python Node calling Python code that conducts real-time AI inference in sorting mode. The real-time data processing pipeline is presented in Figure 3.2.5.

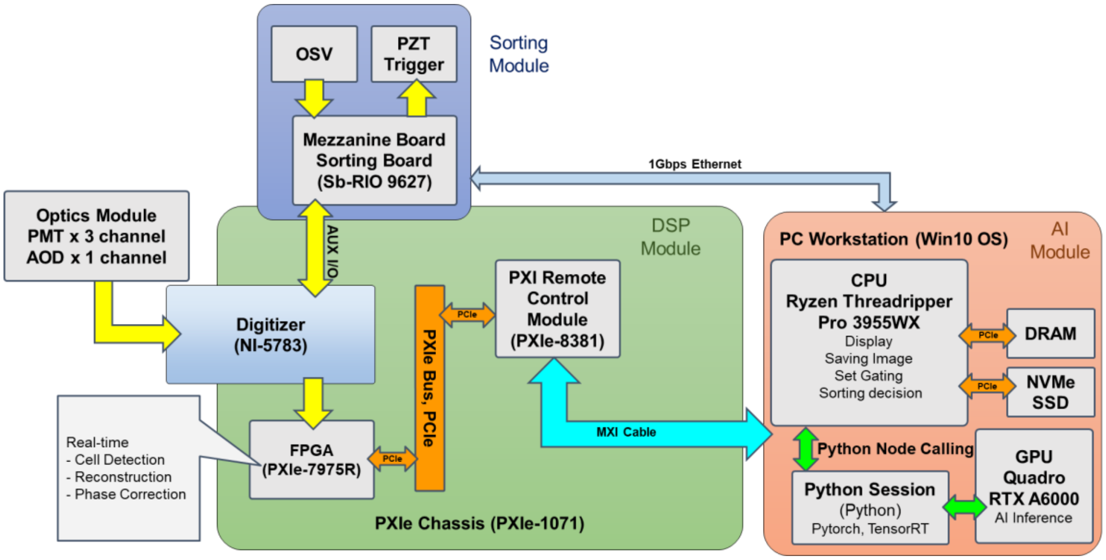


Figure 3.2.4. Real-time data processing system architecture. OSV, optical sorting verification detector; PZT, piezoelectric transducer; PMT, photomultiplier tube; AOD, acoustic optical deflector; FPGA, field-programmable gate array; CPU, central processing unit; GPU, graphics processing unit; DRAM, dynamic random-access memory; NVMe SSD, non-volatile memory express solid-state disk.

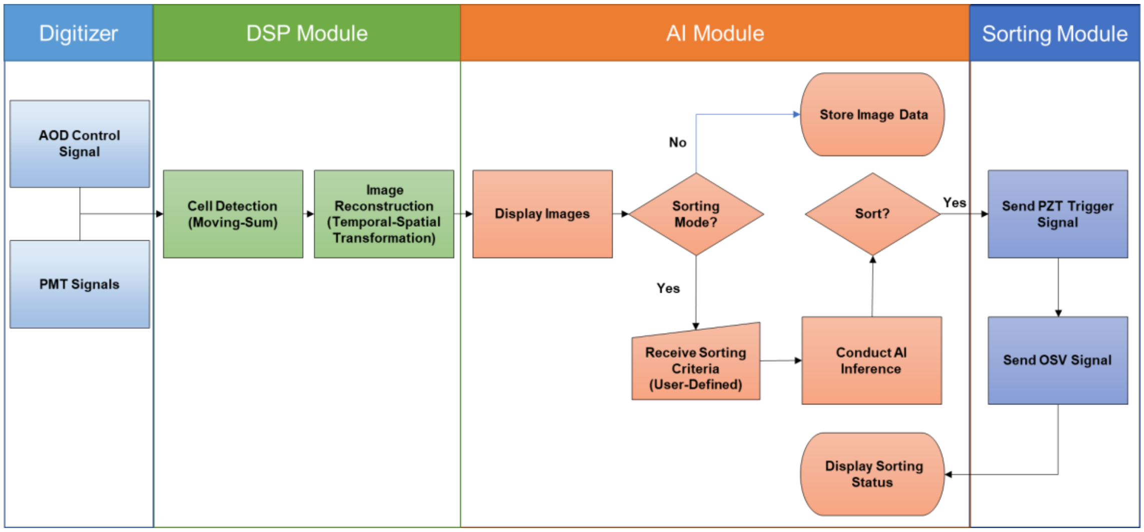


Figure 3.2.5. Real-time data processing pipeline

3.2.3 Image preprocessing for CNN model training

In the CNN model training process, a custom MATLAB image preprocessing code is used to conduct the conventional image feature extraction steps to produce human interpretable image features and generate the preprocessed image dataset. The list of extracted features is imported to the FCS Express software, where the user can define the gating to select targeted image data for the CNN model training. FCS Express exports the selected image indices, and the selected images are prepared for CNN model training using a custom MATLAB code. The image features list extracted by the MATLAB program is presented in Table 3.2.1. The typical processing time for a 20,000-image data set is approximately 5-10 minutes using the current approach.

Table 3.2.1. Human-vision image features extracted in the image preprocessing step

Feature Category	**Image feature Name**
Intensity	Raw Integral Intensity, Raw Maximum Intensity, Raw Average Intensity, Masked Integral Intensity, Masked Maximum Intensity, Masked Average Intensity
Size	Major Axis Length, Minor Axis Length, Area
Shape	Aspect Ratio
Texture	Root-Mean-Square Gradient Magnitude

3.2.4 CNN model training and optimization for real-time inference

To achieve fast CNN model training and low-latency model inference for sorting, we developed a customized 2D UNet by significant improvement of our previously reported work [66]. UNet requires very few labeled images and has reasonable training time. In the meantime, recent research demonstrated that using an autoencoder and the latent space to conduct could improve classification performance [36], [37]. Our previous work also demonstrated that the customized UNet model had better performance than the state-of-art ResNet-18 CNN architecture, which has a faster convergence rate and fewer model parameters than other CNN architectures like VGG, or InceptionNet [34], [66].

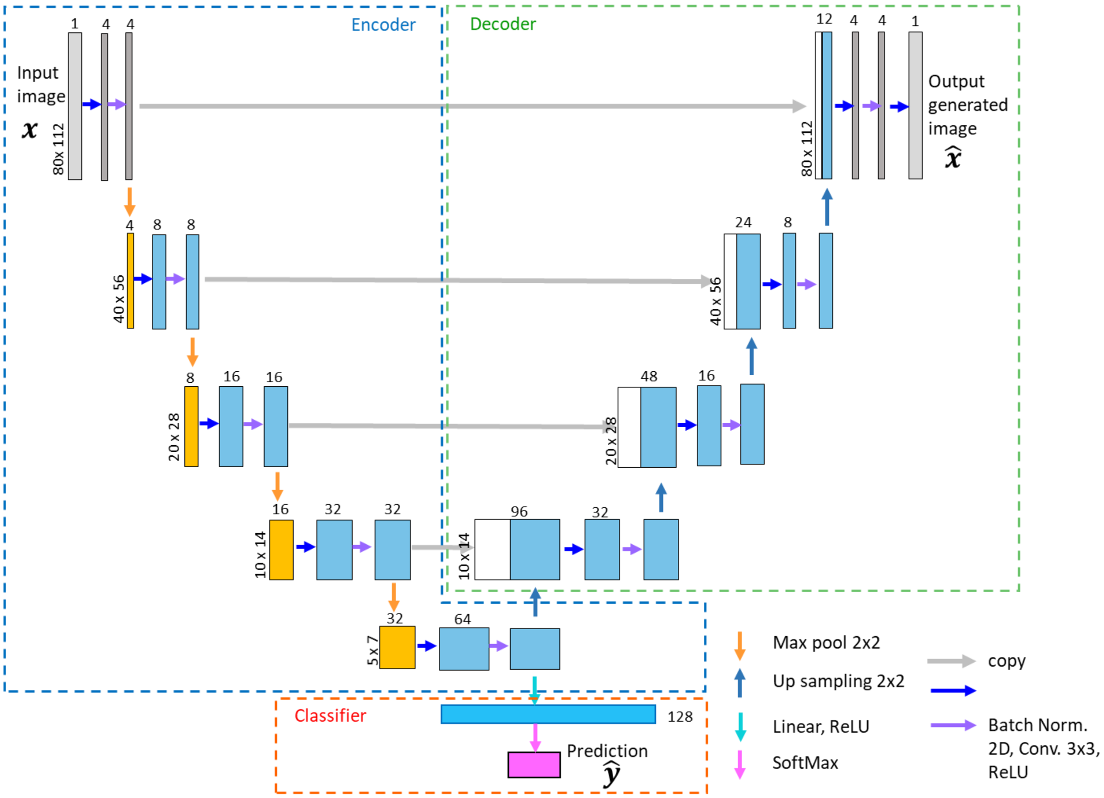


Figure 3.2.6. 2D CNN UNet Architecture. The number of channels is denoted on top of the box. The feature map size is shown at the lower-left edge of the box. The arrows denote the different operations.

The architecture of the 2D UNet is illustrated in Figure 3.2.6. Our UNet model contains a contracting path to encode the image features and takes the input images. Image features are extracted by the convolution layers and encoded to subsequent layers through max pooling. A fully connected layer and Softmax layer are connected to the latent space to make a classification decision. The upsampling path takes the features from the latent space in combination with the higher resolution features passing the convolution layers to generate an output image of the same dimension as the input image. This path acts as a confirmation to ensure the UNet can extract image features effectively. The contracting path works as an encoder, while the upsampling path works as a decoder. The fully connected layer and the Softmax layer act as a classifier. The output of the Softmax layer can be written as

$$\widehat{\boldsymbol{y}}_{\boldsymbol{i}} = \frac{e^{x_i}}{\sum_{j=1}^{C} e^{x_j}}, i = 1,2, \dots, C \quad (1)$$

where $\boldsymbol{x}$ is the input vectors, C is the number of classes.

During the model training process, we use a weighted loss that incorporates the mini-batch averaged cross-entropy loss between the predicted class and ground-truth class (Eqs. (2)) and the mean-square error loss between the input and generated output image pixel values. The averaged cross-entropy loss L_{CE} can be expressed as

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}_{\boldsymbol{i}}(\boldsymbol{x}) \cdot \log\left(\widehat{\boldsymbol{y}}_{\boldsymbol{i}}(\boldsymbol{x})\right) \quad (2)$$

where $\boldsymbol{y}_{\boldsymbol{i}}$ is the ground truth class vector, $\widehat{\boldsymbol{y}}_{\boldsymbol{i}}$ is the predicted class vector, and N is the data size in the mini-batch.

The mini-batch averaged mean-square error loss L_{MSE} can be expressed as

$$L_{MSE}(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{M}(x_{i,j} - \hat{x}_{i,j})^2 \quad (3)$$

Where $\boldsymbol{x}$ and $\widehat{\boldsymbol{x}}$ are the input image and generated image vectors, respectively, M is the flattened image vector dimension, and N is the data size in the mini-batch.

The weighted total loss L is defined as

$$L = w \cdot L_{CE} + (1 - w) \cdot L_{MSE} \quad (4)$$

where w is the weight coefficient to balance the loss function.

To optimize the UNet model architecture, a CNN model architecture search was performed to reduce the initial convolutional kernel number in the UNet model. We examined the impact on model size, model parameters, training time, inference time, and classification accuracy. In the model optimization process, A stratified 5-fold cross-validation (CV) approach was used to train the UNet models and evaluate their performance. For each fold, training data were augmented by conducting random horizontal and vertical flips on the image data. The model was then validated using the instance from the validation set. The model performance was evaluated using balanced classification accuracy. The balanced accuracy $\overline{\sigma}$ is the arithmetic mean of class-specific accuracies and is calculated as

$$\overline{\sigma} = \frac{1}{C}\sum_{i=1}^{C}\sigma_i \quad (5)$$

where σ_i is the class-specific accuracy, and C is the number of classes. The balanced accuracy does not favor a classifier that exploits class imbalance by biasing toward the majority class. We used a 15,000-image dataset acquired in the white blood cells imaging experiment to conduct the model architecture search. 5,000 cell images were used for each cell type to balance the data occurrence. To investigate the impact on the inference time utilizing different GPU acceleration frameworks, a comparison of using Pytorch and TensoRT frameworks was performed during the UNet architecture search.

All deep learning model training and model performance tests were conducted on the same PC workstation in the AI module of the low-latency IACS platform. The deep learning development was conducted under Python 3.6.8, Pytorch 1.10.2, and TensorRT 8.2.2.1 frameworks.

3.2.5 Sorting experiment

Two sorting experiments were performed to demonstrate the system's capability. In the bead sorting experiment, we sorted beads with the targeted size out of a mixture of 7 µm and 15 µm polystyrene (PS) fluorescent microsphere beads sample. To demonstrate the cell sorting capabilities, we performed 3-part human white blood cells (WBCs) sorting experiments to sort out the targeted WBC type from leukocyte samples (Veri-Cells, BioLegend).

In the beads sorting experiment, we prepared 7 µm and 15 µm PS beads samples separately and ran them through our system under the analysis mode to obtain the training dataset. A total of 4,000 images were obtained, with 2,000 images for 7 µm PS beads and 2,000 images for 15 µm PS beads. A two-part image classification training was performed to train the UNet model with an 80/20 train/validation split of the dataset. The pre-trained model performance was evaluated by generating the confusion matrix of the classification results using the validation dataset. In addition, the activation output of the fully connected layer was projected to a 2D plane as the t-distributed stochastic neighbor embedding (t-SNE) plot [40]. We then deployed the pre-trained CNN model to sort the beads when the prediction confidence level of the AI model was higher than 99%. We sorted 15 µm PS beads out of the mixture and analyzed the collected outputs from the target and waste channels using a commercial flow cytometer (Accuri C6 plus, BD Biosciences). We also used centrifugation to enrich the output samples from the sorting channel and the waste channel and examined them under a fluorescence microscope.

In the WBC sorting experiments, the WBC samples were immunostained with antibodies panels (See Table 3.2.2) to provide a ground truth label for each cell type. Lymphocytes, granulocytes, and monocytes were phenotyped using an antibody panel. We processed the immunostained WBC samples to obtain the training dataset. A total of 17,876 cell images were used, with 6,788 lymphocytes images, 5,001 granulocytes images, and 6,087 Monocytes images. A three-part image classification training was performed to train the UNet model with a stratified 80/20 train/validation split of the dataset. We then deployed the pre-trained CNN model to sort cells when the AI model prediction confidence level was higher than 99%. Three sorting experiments were performed. For each sorting experiment, we fluorescently labeled the target cell type with one color and other cell types with another color using the antibodies panel for the sole purpose of post-sorting performance evaluation (See Table 3.2.3) while the AI inference and the sorting decision were entirely based on the label-free, transmission images. To verify the system performance, we ran the sorted sample and the waste sample through a commercial flow cytometer (Accuri C6 plus, BD Biosciences) and used the fluorescent signals of each WBC cell type to evaluate the sorting purity.

The sorting purity measurement was evaluated by counting the ratio between the sorted target particle number and the total sorted particle number described by Eqn. (6):

$$Sorting\ Purity = \frac{N_{target}}{N_{target}+N_{non-target}} \quad (6)$$

Where N_{target} is the sorted target particle number and $N_{non-target}$ is the sorted non-target particle number.

During the sorting experiment, we also observed and recorded the event processing time to evaluate the sorting latency.

Table 3.2.2. Antibody panel design for human white blood cell training data collection ground truth labeling

	Ground Truth Labelling 1	**Ground Truth Labelling 2**	**Unlabeled**
Panel 1	Monocytes (PE anti-human CD14 Antibody)	Lymphocytes (BB515 Mouse Anti-Human CD3 and CD19)	Granulocytes and others

Panel 2	Granulocytes (PE anti-human CD66b Antibody)	Lymphocytes (BB515 Mouse Anti-Human CD3 and CD19)	Monocytes and others
Panel 3	Lymphocytes (PE Mouse Anti-Human CD3 and CD19 antibodies)	Granulocytes (BB515 anti-human CD66b Antibody)	Monocytes and others

Table 3.2.3. Antibody panel design for human white blood cell sorting ground truth labeling

	Ground Truth Labelling 1	**Ground Truth Labelling 2**	**Unlabeled**
Panel 4 (Sorting target: Monocytes)	Monocytes (BB515 CD14 Antibody)	Lymphocytes (PE Mouse Anti-Human CD3 and CD19) Granulocytes (PE Mouse Anti-Human anti-human CD66b)	others
Panel 5 (Sorting target: Granulocytes)	Granulocytes (BB515 Mouse Anti-Human anti-human CD66b Antibody)	Lymphocytes (PE Anti-Human CD3 and CD19) Monocytes (PE CD14 Antibody)	others
Panel 6 (Sorting target: Lymphocytes)	Lymphocytes (BB515 Mouse Anti-Human CD3 and CD19)	Granulocytes (PE anti-human CD66b Antibody) Monocytes (PE CD14 Antibody)	others

3.2.6 Image acquisition and sorting experiment sample preparation

3.2.6.1 Preparation of fluorescent polystyrene particles

The evaluate the imaging and sorting performance of the low-latency IACS, fluorescent PS beads were used. A 1:6 mixture of 15 µm PS particles (Fluorescent microspheres, Dragon Green, Cat. No. FSDG009, Bangs Laboratories, Inc.) and 7 µm PS particles (Fluorescent microspheres, Dragon Green, Cat. No. FSDG007, Bangs Laboratories, Inc.) was fed from the sample inlet of the microfluidic chip. The concentration of the particles was adjusted to be 500 particles µL^{-1}.

3.2.6.2 Preparation of CHO-ES cells and DNA staining

CHO-K1 (ATCC CCL-61) were used for DNA staining. The cells were harvested when they were at a confluency of around 80%. The cells were centrifuged at 350 x g for 5 minutes to remove the supernatant to perform a wach. After washing the cells with PBS (Genesee Scientific, CA, USA), the cells were centrifuged at 350 x g for 5 minutes. The supernatant was removed and 100 µL of 4% formaldehyde, methanol-free (Cell Signaling Technology, Massachusetts, USA) was added for every million cells to be fixed. The cells were incubated for 15 minutes at 37ºC. After fixation, the cells were washed again and resuspended in PBS with 0.5% BSA (Thermo Scientific) to a concentration of 1.0x10^{6} cells/mL. Lastly, the cells were stained with 0.5 µM Vybrant DyeCycle Green Stain (Invitrogen) for 30 minutes and filtered with a 35 µm strainer cap (Genesee Scientific, CA, USA) before analysis.

3.2.6.3 Preparation of MCF7 cells and mitochondrial staining

To image the mitochondria of MCF7 cells (ATCC HTB-22), the cells were harvested when they were at a confluency of 70%. The cells were diluted to a concentration of 1.0x10^{6} cells/mL using a buffer composed of PBS, 0.5% BSA, 12.5 mM HEPES (Gibco) and 5 mM EDTA (Invitrogen). The diluted cells were stained with 100 mM of MitoView Green (Biotium, San Francisco, USA) and incubated for 15 minutes at 37ºC. After incubation, the cells were filtered with a 35 µm stainer cap and analyzed.

3.2.6.4 Preparation of human iPSC cells and viability staining

Human iPSCs reprogrammed from fibroblasts were cultured in DMEM/F-12 50/50 1X (Corning™, #10-092-CM) supplemented with HiDef B8 500X (Defined Bioscience, #LSS-201). Non-TC treated 6-well plates (CELLTREAT, #229506) were treated with vitronectin (Gibco™, #A14700), a recombinant human protein that provides a defined surface for feeder-free culture. Samples were maintained with a visual assessment of <30% differentiation per well. Cells were passaged in aggregates ranging from 50-100 µm, using the enzyme-free Gentle Cell Dissociation Reagent (STEMCELL Technologies, #100-0485). Healthy iPSC colonies were determined by examining the morphology under phase microscopy for colony compactness, defined borders, well-outlined edges, and a large nucleus-to-cytoplasm ratio. A single-cell suspension was obtained using Accutase® (Innovative Cell Technologies, Inc. #AT104), centrifuged at 200 x g for 3 minutes, and resuspended in sheath buffer (basal media + 10% Accutase) at a concentration of 3.0 x 10^5 cells/mL. Live calcein AM (Invitrogen™, # C3099) stained iPSCs were imaged by capturing conversion of the green, fluorescent calcein (Ex/Em: 494/517 nm).

3.2.6.5 Preparation of human white blood cells and immune staining

The Veri-Cells™ Leukocyte Kit, prepared from lyophilized human peripheral blood leukocytes were sused (BioLegend Cat. 426003). These cells have been verified to work with commonly tested cell surface markers such as CD3, CD14, CD19, and CD66b. CD66b is a glycosylphosphatidylinositol (GPI) linked protein with a molecular weight of 100 kDa expressed on granulocytes. CD3 and CD19 are expressed on T-cell and B-cell, respectively. CD14 is expressed at high levels on monocytes. In this experiment, we used different combinations of the following antibodies: BB515 Mouse Anti-Human CD66b (Ex/Em 490/515nm, Clone G10F5, Cat. 564679, BD Biosciences), PE anti-human CD66b Antibody (Ex/Em 496/578nm, Clone 6/40c, Cat.

392903, BioLegend Biomedical Inc), Alexa Fluor 488 anti-human CD14 Antibody (Ex/Em 498/520nm, Clone 63D3, Cat. 367129, BioLegend Biomedical Inc), PE anti-human CD14 Antibody (Ex/Em 496/578nm, Clone 63D3, Cat. 367103, BioLegend Biomedical Inc), BB515 Mouse Anti-Human CD3 (Ex/Em 490/515nm, Clone UCHT1, Cat. 564466, BD Biosciences), PE anti-human CD3 Antibody (Ex/Em 496/578nm, Clone UCHT1, Cat. 300407, BioLegend Biomedical Inc.), BB515 Mouse Anti-Human CD19 (Ex/Em 490/515nm, Clone HIB19, Cat. 564457, BD Biosciences), and PE anti-human CD19 Antibody (Ex/Em 496/578nm, Clone HIB19, Cat. 302207, BioLegend Biomedical Inc.), Table 3.2.5.1-3.2.5.2 show the panel design used for leukocyte phenotyping. The concentration of the particles was adjusted to be 500-1000 particles μL^{-1} to achieve an event rate of approximately 100-200 eps.

3.3 Results

3.3.1 Low-latency IACS system overview and functionality

The low-latency IACS platform uses a scanning laser beam to interrogate the cell that is traveling in a microfluidic channel and PMTs to detect the signal, which is then transformed into cell images via a temporal-spatial transformation (Figure 3.3.1). The image characteristics are analyzed by an AI inference model to determine whether the cells meet the sorting criteria defined by the trained deep learning model. If so, an actuation signal is sent to the on-chip piezoelectric (PZT) actuator to deflect the cells into the sorting channel(s) while the unselected cells leave the microfluidic device via the waste channel. The details of the optical and microfluidic hardware and the image reconstruction algorithms of the IACS were implemented as previously described [58].

The major enhancement of the system presented here is the method of deep learning and AI-guided cell sorting without image feature extraction. We will elucidate the hardware and software design that supports this approach in detail here. The temporal waveforms are sampled by an analog-to-digital converter (NI-5783, National Instruments) and then are sent to a field-programmable-gate-array (FPGA, PXIe-7975R, National Instruments) for real-time cell image reconstruction using a temporal-spatial transformation [68]. The reconstructed cell image data are transferred to a stand-alone PC workstation (AI module) through a wide-band PCIe bus for conducting real-time AI model inference on a dedicated GPU (Quadro RTX A6000, Nvidia). The AI module predicts the cell types with a prediction confidence level on the AI inference prediction. The sorting decisions are made by comparing the AI inference prediction with the user-specified cell classes and the classification confidence level. A clock is used to register the latency for the process. If the overall processing time is within the set value, then the sorting action can be triggered and the sorting decision is sent through the PCIe bus to the FPGA that controls the action of the on-chip piezoelectric actuator. In our system, the entire data processing including AI model inference and PZT actuation for 99% of cells is shorter than 3 ms to enable sorting.

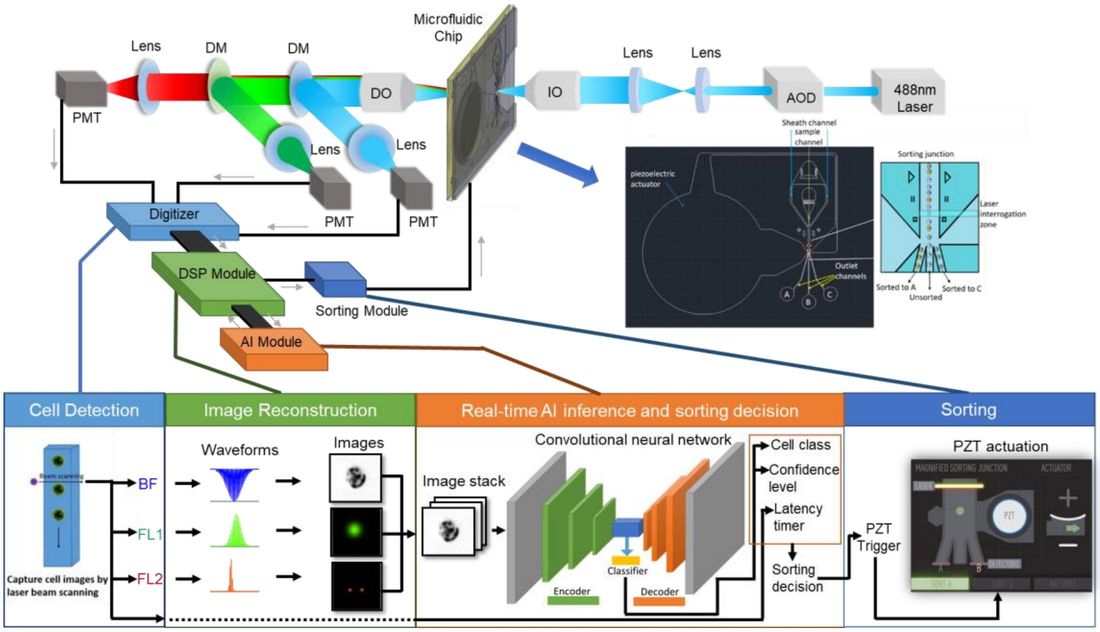


Figure 3.3.1. Schematics and functionality of low-latency IACS system. The scanning laser beam and the cell travel produce the equivalence of a 2D raster scanning system. The bright field and fluorescent signals of the cell are detected by PMTs and the temporal signals are reconstructed to form cell images via real-time processing by an FPGA. Meanwhile, each cell image is fed to a convolutional neural network (CNN) to conduct real-time AI inference. According to the CNN prediction output, the on-chip PZT actuator is triggered to sort out cells. AOD: acousto-optic deflector; DM: dichroic mirror; IO: 10X/0.28NA illumination objective lens; DO: 10X/0.28NA detection objective lens; PMT: photomultiplier tube.

3.3.2 Imaging beads and cells using the low-latency IACS system

Our system can produce both transmission and fluorescent images of objects traveling in the microfluidic channel at approximately 20 cm/s, as shown in Figure 3.3.2. We have used our system to capture single and multiple 7 µm and 15 µm fluorescent bead images (Figure 3.3.2 a-b), resolved intracellular DNA and mitochondrial localization distribution of fluorescently labeled CHO-ES and MCF7 cells (Figure 3.3.2 c-d), and imaged human iPSC (Figure 3.3.2 e) of which the bright-field images showed intracellular structures. In the experiment with white blood cells, we used the system to capture the surface antibody distribution of immunostained cells (Figure 3.3.2 f-h). To summarize, our system can capture particles of different sizes ranging from 1 to 40 µm. Our system also clearly shows images of doublets.

3.3.3 CNN model optimization

Through a CNN model architecture search, we have optimized an UNet CNN model to conduct 2-part or 3-part particle classification. During the CNN model architecture search, we found that the initial convolutional kernel number has a significant impact on the model size, parameter number, training time, and inference time; while it has a relatively low impact on model prediction accuracy for our system. By reducing the initial convolutional kernel number from 64 to 4, we reduced the model parameter and model size by approximately 100-fold (Figure 3.3.3 a). The reduced model parameter and model size decreases the total training time and the model inference time. We achieved a 4-fold reduction in model training time from 7200 seconds to less than 1800 seconds in a CNN training process with a dataset of 15,000 images while maintaining a high model prediction accuracy of 0.977. In addition, by leveraging the optimized GPU acceleration framework under the TensorRT framework [69], [70], we achieved approximately 5-fold reduction in the model inference time compared to the Pytorch framework. By shrinking the initial convolutional kernel number from 64 to 4, we also improved the model inference time from 1.735 ms to 0.518 ms under the TensorRT framework, further reducing the sorting latency for real-time CNN inference during the sorting experiment (Figure 3.3.3 c). Additional details about CNN model training and inference time optimization are found in Figure 3.3.4-3.3.13.

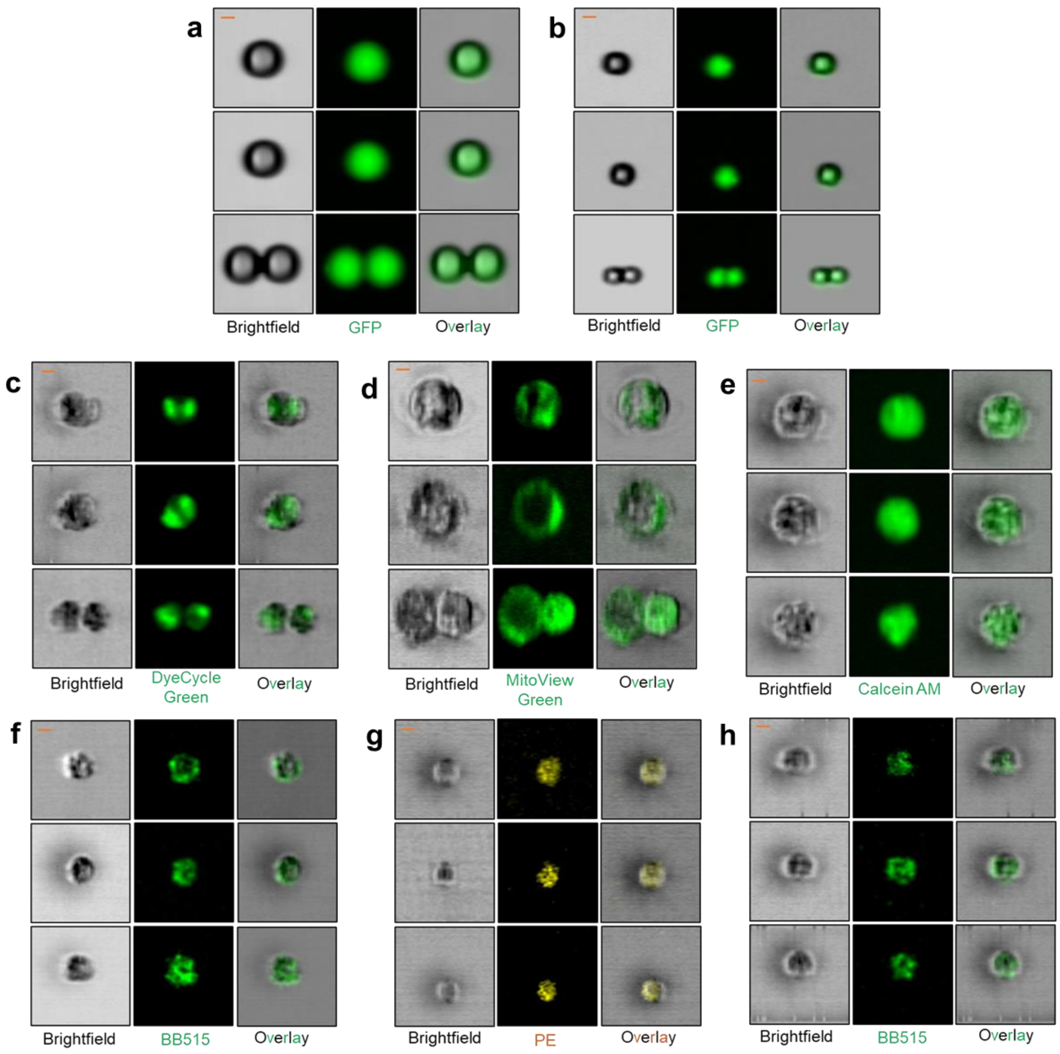


Figure 3.3.2. Image gallery of example beads and cell images captured by the low-latency IACS platform. A. 15 µm fluorescent beads; B. 7 µm fluorescent beads; C. CHO-ES cells with DNA staining (Vybrant DyeCycle Green); D. MCF7 cells with mitochondrial staining (MitoView Green); E. Human iPSC with vitality dye staining (Calcein AM; F. Human granulocytes with anti-CD66b BB515 immunostaining; G. Human lymphocytes with anti-CD3 and anti-CD19 PE immunostaining; H. Human monocytes with anti-CD14 BB515 immunostaining. Scale bar: 5 µm.

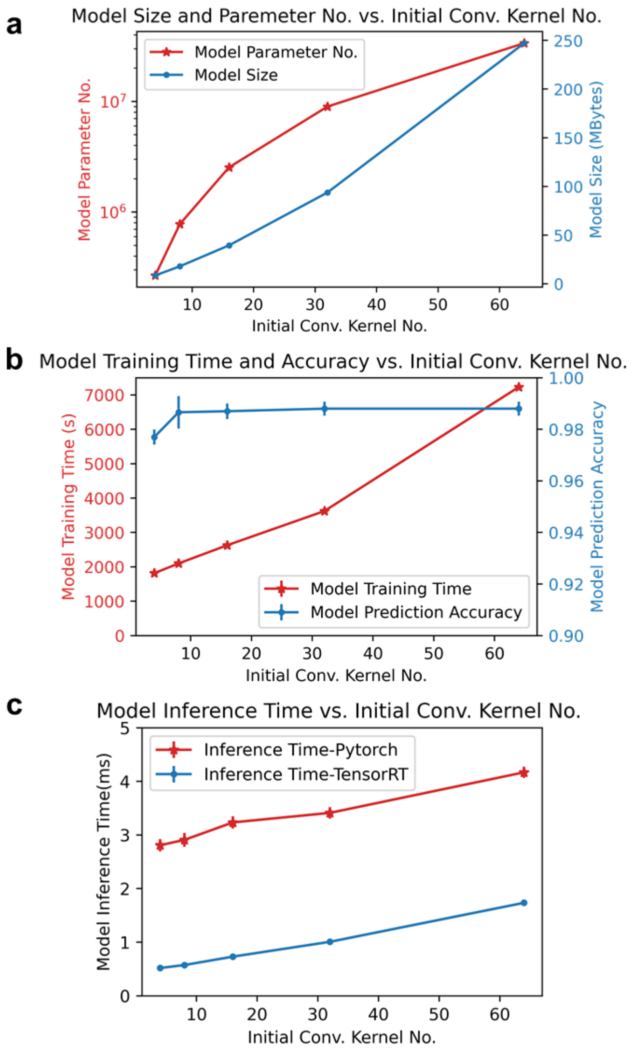


Figure 3.3.3. Custom UNet model optimization on model size, training time, and inference time. A. Model size and parameter number vs. Initial convolutional kernel number. B. Model training time vs. Initial convolutional kernel number. C. Model inference time under Pytorch and TensorRT frameworks vs. Initial convolutional kernel number. Error bars ± SD.

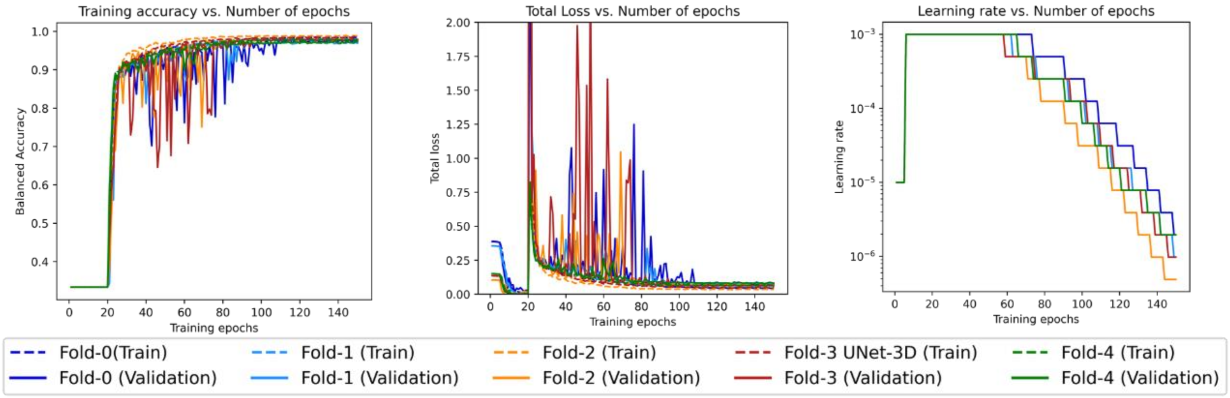


Figure 3.3.4. UNet training curves with initial convolutional kernel size = 4

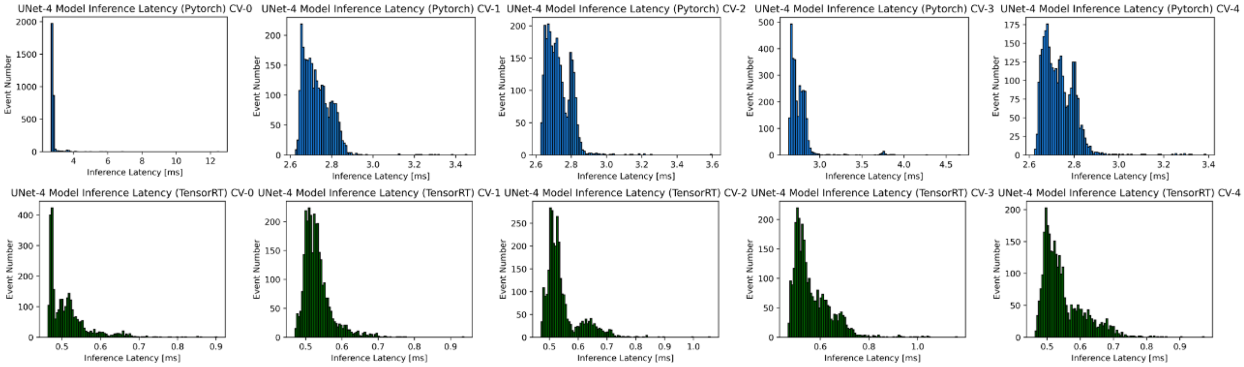


Figure 3.3.5. UNet inference time with initial convolutional kernel size = 4

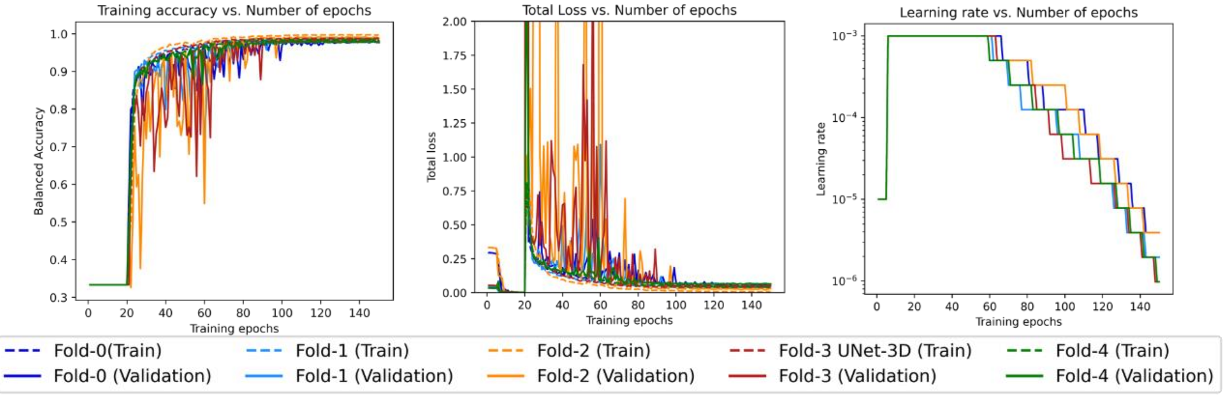


Figure 3.3.6. UNet training curves with initial convolutional kernel size = 8

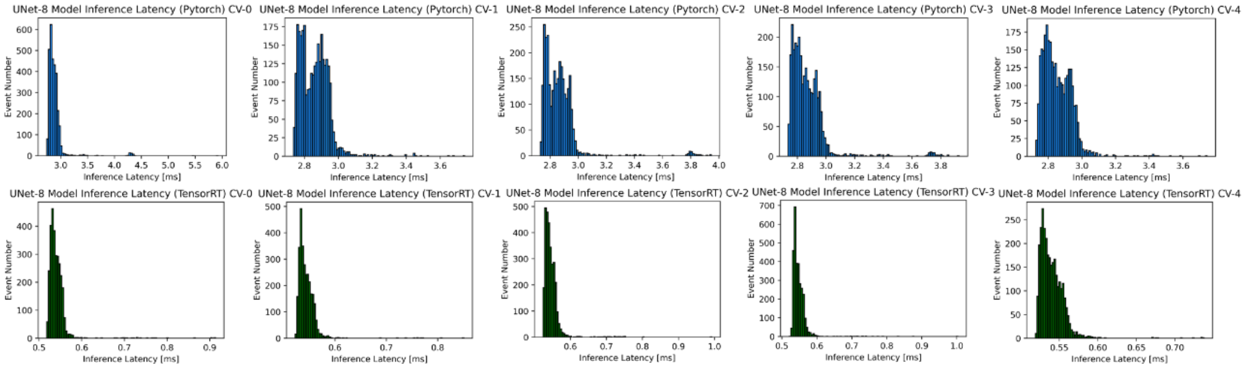


Figure 3.3.7. UNet inference time with initial convolutional kernel size = 8

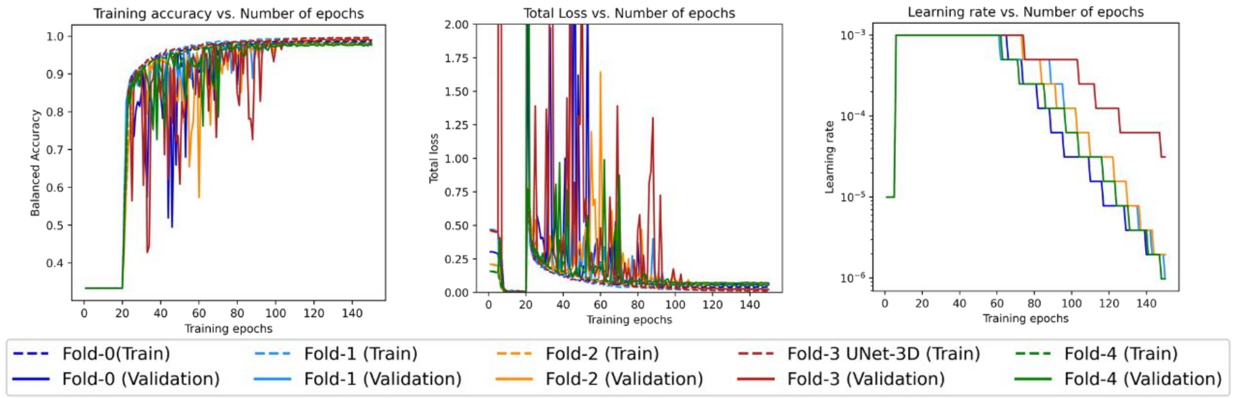


Figure 3.3.8. UNet training curves with initial convolutional kernel size = 16

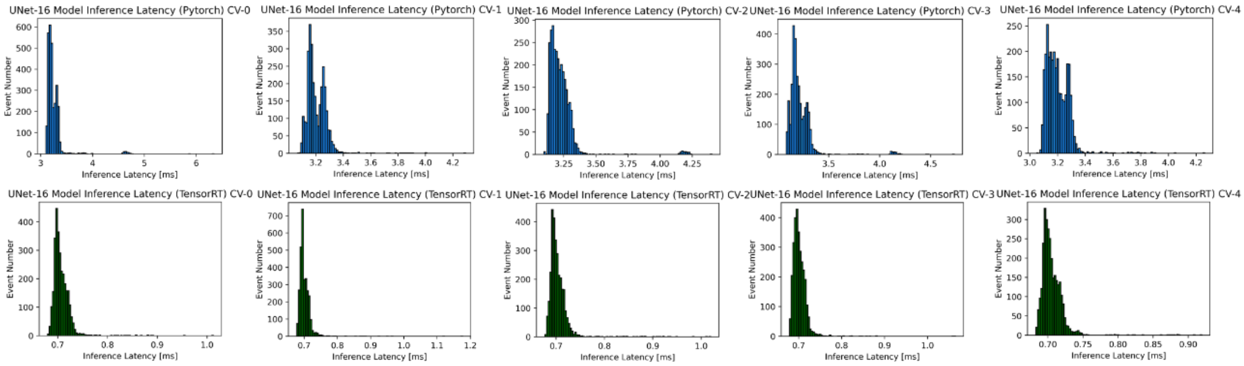


Figure 3.3.9. UNet inference time with initial convolutional kernel size = 16

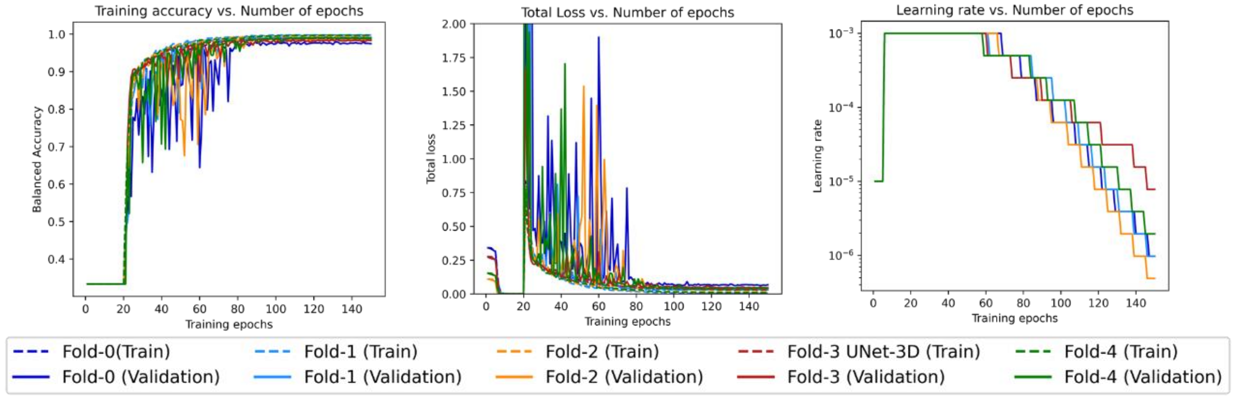


Figure 3.3.10. UNet training curves with initial convolutional kernel size = 32

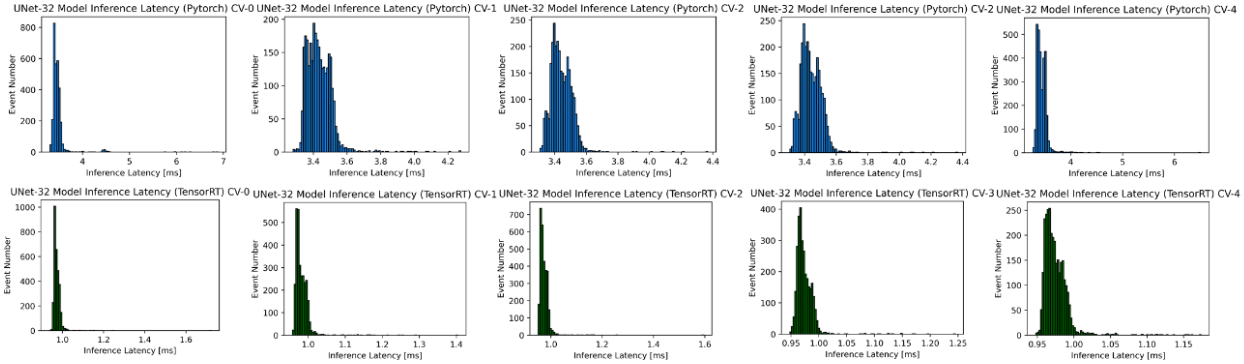


Figure 3.3.11. UNet inference time with initial convolutional kernel size = 32

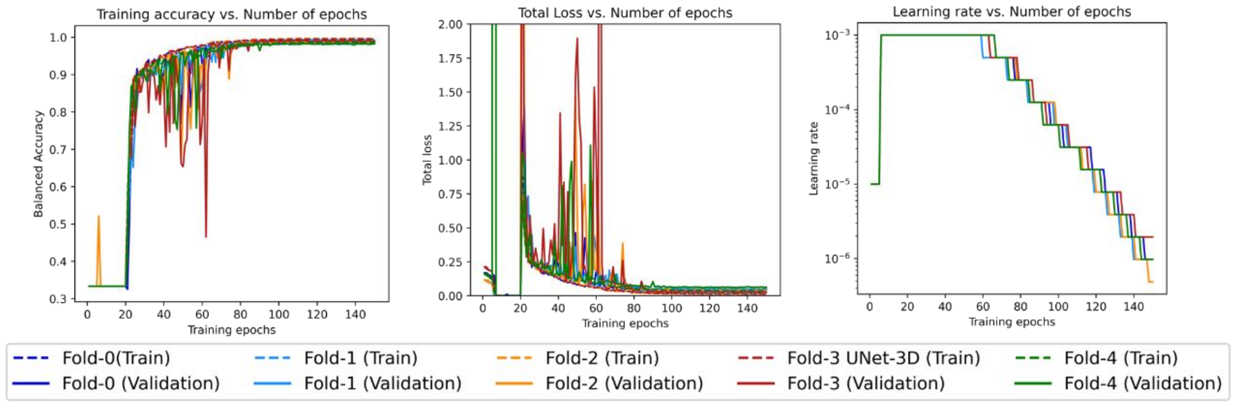


Figure 3.3.12. UNet training curves with initial convolutional kernel size = 64

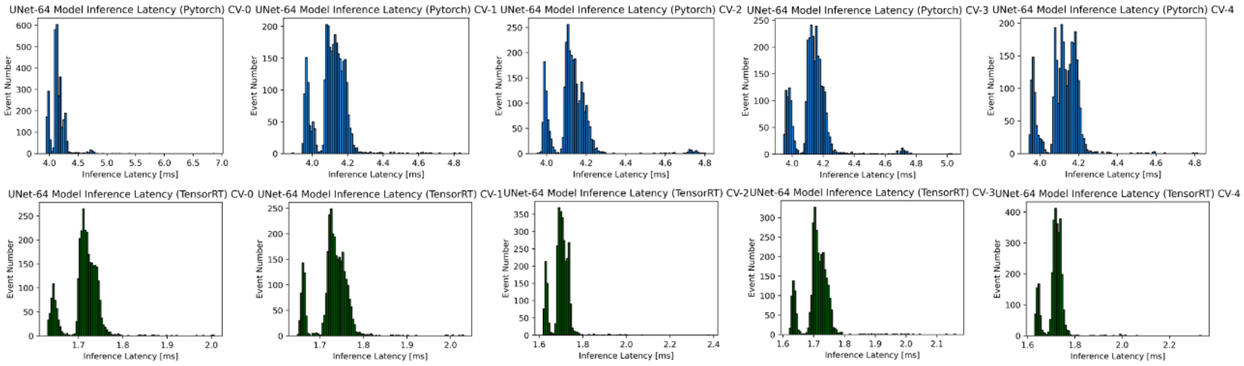


Figure 3.3.13. UNet inference time with initial convolutional kernel size = 64

3.3.4 7 µm and 15 µm bead sorting

In the bead sorting experiment, we evaluated the training progress and the classification performance of the pre-trained CNN model. We completed the model training process within 540 seconds using a dataset of 4,000 images to conduct a 2-part classification. The pre-trained CNN model returns a balanced prediction accuracy of 100% and the tSNE visualization shows two distinctly separated clusters of 7 µm and 15 µm beads events (Figure 3.3.14 a-b). The model training detail for beads sorting is included in Figure 3.3.15.

During the bead sorting experiment, we monitored the data processing time of the deployed pre-trained CNN model. The processing time of 1,118 sorting events was recorded and the rank-ordered processing time distribution is presented in Fig. 3.3.14 c. It was observed that 98.4% of the events can be processed within 2,287 µs and the average data processing time was 1,802 µs. To measure the purity of the AI-inferred bead sorting, we confirmed particle composition analysis on the collected sample post sorting by a commercial cytometer (Accuri C6, BD Biosciences) and achieved 96.6% purity (See Figure 3.3.16-3.3.17). The microscopic images of pre-sort, post-sort, and waste samples are presented in Figure 3.3.18, and the results are consistent with confirmatory flow cytometer results.

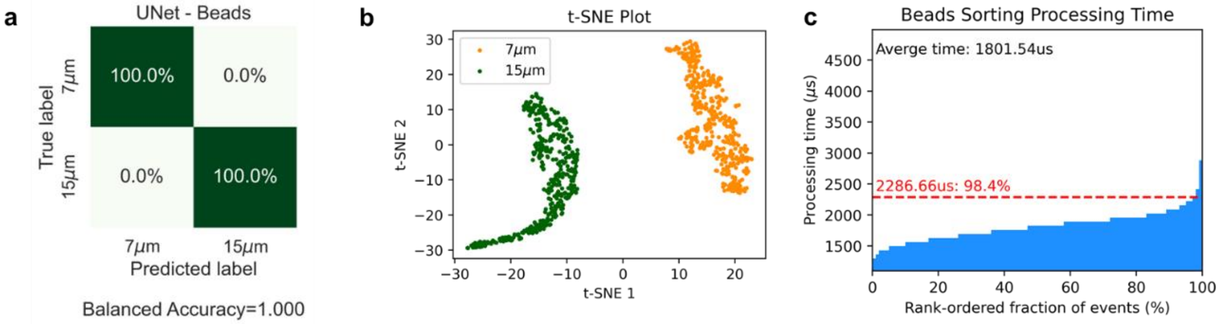


Figure 3.3.14. 7 µm and 15 µm bead sorting results. A. Confusion matrix for bead sorting experiment. B. t-SNE visualization on the beads classification. C. Beads sorting experiment data processing time.

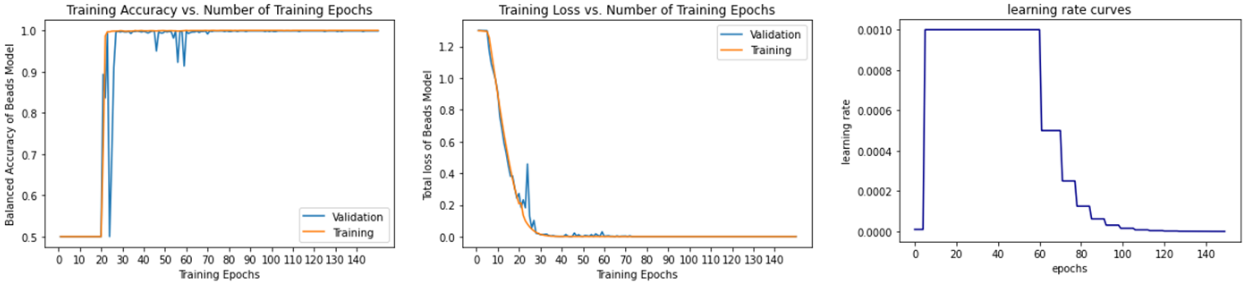
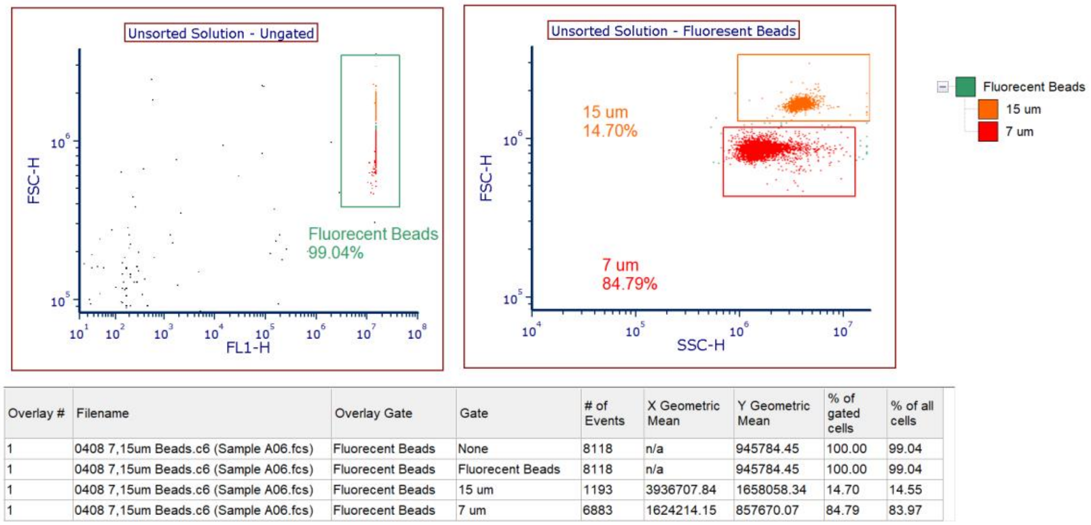
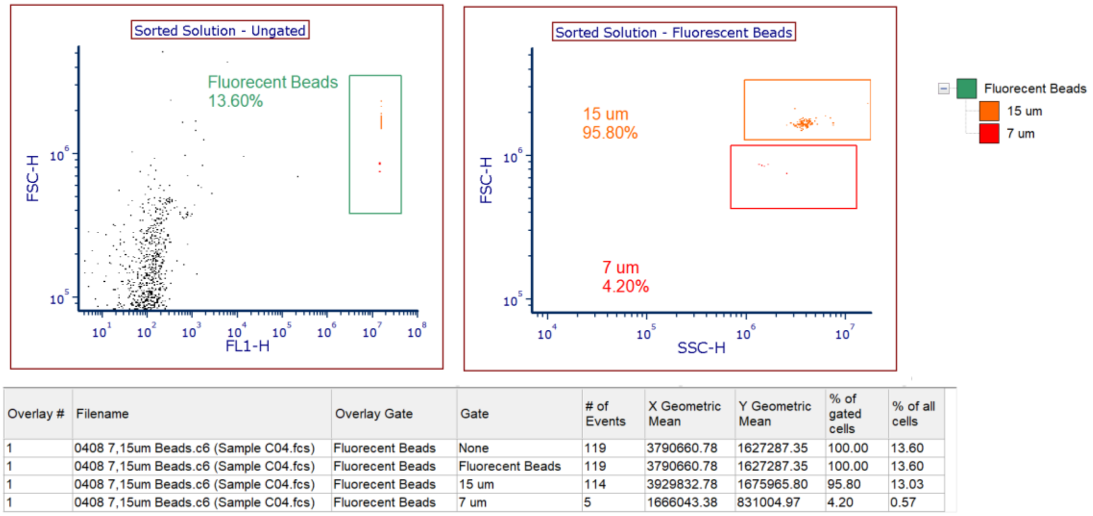


Figure 3.3.15. UNet training curves for beads sorting experiment



Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0408 7,15um Beads.c6 (Sample A06.fcs)	Fluorecent Beads	None	8118	n/a	945784.45	100.00	99.04
1	0408 7,15um Beads.c6 (Sample A06.fcs)	Fluorecent Beads	Fluorecent Beads	8118	n/a	945784.45	100.00	99.04
1	0408 7,15um Beads.c6 (Sample A06.fcs)	Fluorecent Beads	15 um	1193	3936707.84	1658058.34	14.70	14.55
1	0408 7,15um Beads.c6 (Sample A06.fcs)	Fluorecent Beads	7 um	6883	1624214.15	857670.07	84.79	83.97

Figure 3.3.16. Beads sorting experiment pre-sorting Accuri particle composition analysis



Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0408 7,15um Beads.c6 (Sample C04.fcs)	Fluorecent Beads	None	119	3790660.78	1627287.35	100.00	13.60
1	0408 7,15um Beads.c6 (Sample C04.fcs)	Fluorecent Beads	Fluorecent Beads	119	3790660.78	1627287.35	100.00	13.60
1	0408 7,15um Beads.c6 (Sample C04.fcs)	Fluorecent Beads	15 um	114	3929832.78	1675965.80	95.80	13.03
1	0408 7,15um Beads.c6 (Sample C04.fcs)	Fluorecent Beads	7 um	5	1666043.38	831004.97	4.20	0.57

Figure 3.3.177. Beads sorting experiment post-sorting Accuri particle composition analysis

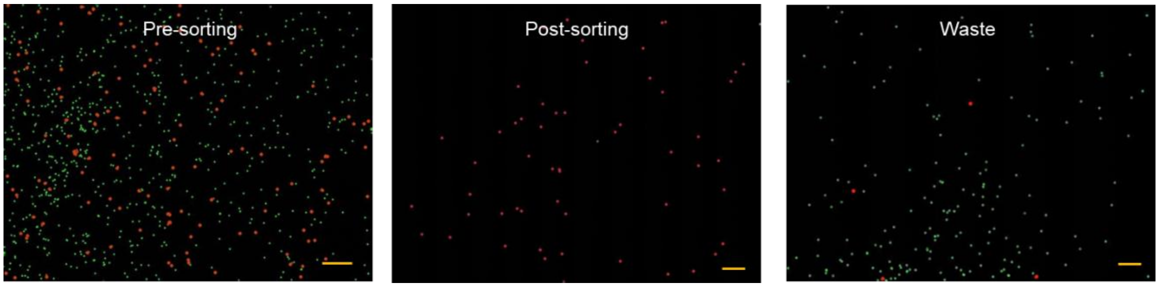


Figure 3.3.18. fluorescence microscopy images of the pre-sorting, post-sorting, and waste beads mixture, 15 µm (target, red) and 7 µm (non-target, light green) beads. Scale bar: 200 µm

3.3.5 Three-part white blood cell sorting

The white blood cell (WBC) differential is a broadly used in clinical test and usually requires fluorescence antibody labeling[71]. Conducting label-free WBC classification and sorting can avoid the biomarker labeling process and minimize cell degradation or morphological changes during labeling. Although CNN-based label-free white blood cell type classification has been demonstrated using both research and commercial imaging flow cytometer systems [24], [26], [27],

[66], none of these systems supported AI-based cell sorting with a real-time AI inferencing process. This AI-enabled cell sorting capability not only enables broad biomedical research applications, but also allows us to evaluate the effectiveness of the deep learning model and discover new biology such as identifying new cell types and their variation under health and disease.

Using our customized UNet CNN model, we achieved a high-accuracy and feasible 3-part WBC type classification. The CNN model training was completed within 40 minutes with a training set of approximately 18,000 images. This pre-trained CNN model yielded a balanced classification accuracy of 99.5% for a three-part WBC type classification and the tSNE visualization using the validation dataset that was reflected in well-separated clusters of the cell groups (Figure 3.3.19 a-b). The model training detail for WBC sorting is included in Figure 3.3.20.

In the AI-inferred WBC sorting experiments, we sorted relatively small numbers of cells: 2,082, 2,840, and 10,124 for lymphocyte, monocyte, and granulocyte, respectively. More than 99% of sorting events were processed within 2.312 ms and the average data processing time ranged from 1.687 to 1.834 ms (Figure 3.3.19 d-f). Three replicates of the sorted samples were collected for sorting purity analysis. The average sorting purity for lymphocytes, monocytes, and granulocytes were 92.0%, 89.05%, and 98.4%, respectively (Figure 3.3.19 c). Because of the sorting mechanism and Poisson factors, sorting purity variation was influenced by the cell compositions of the initial samples which had the percentage for lymphocytes, monocytes, and granulocytes to be 13.12%, 6.51%, and 79.48%, respectively. The post-sorting cell composition analysis details are illustrated in Figure 3.3.20-3.3.32.

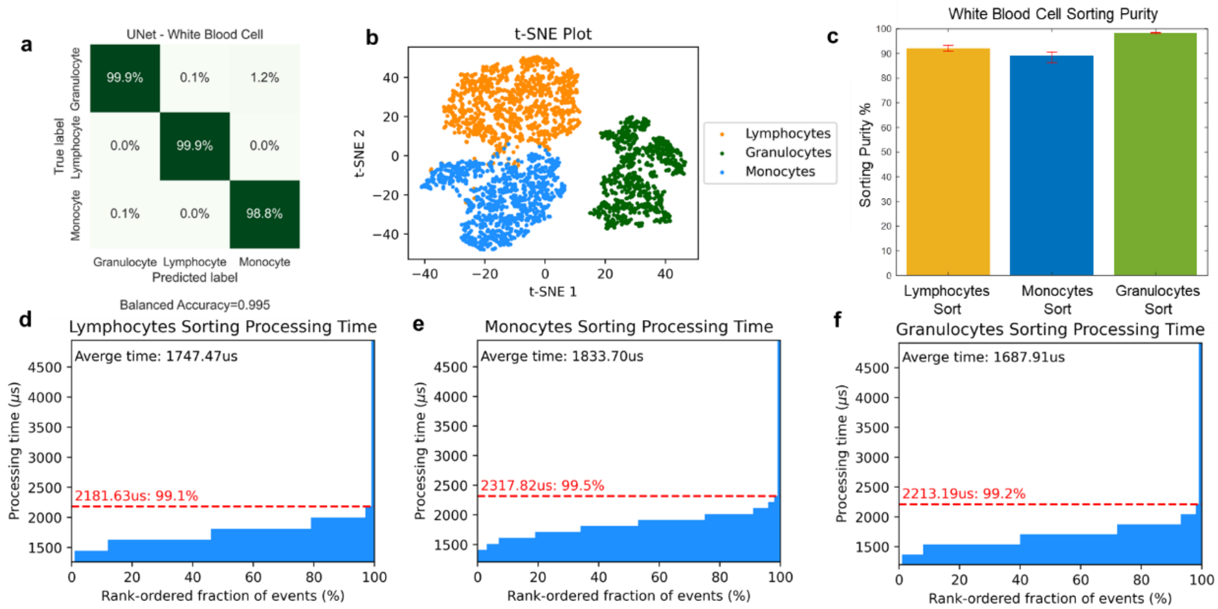


Figure 3.3.19. Human white blood cell sorting results. A. Confusion matrix for white blood cell sorting experiment. B. t-SNE visualization on the white blood cell classification C. White blood cell sorting purity plot; D. Lymphocytes sorting experiment data processing time; E. Monocytes sorting experiment data processing time; F. Granulocytes sorting experiment data processing time

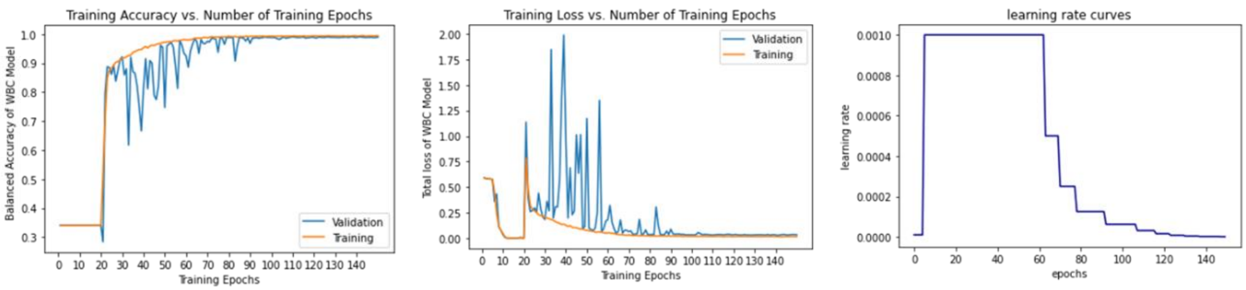
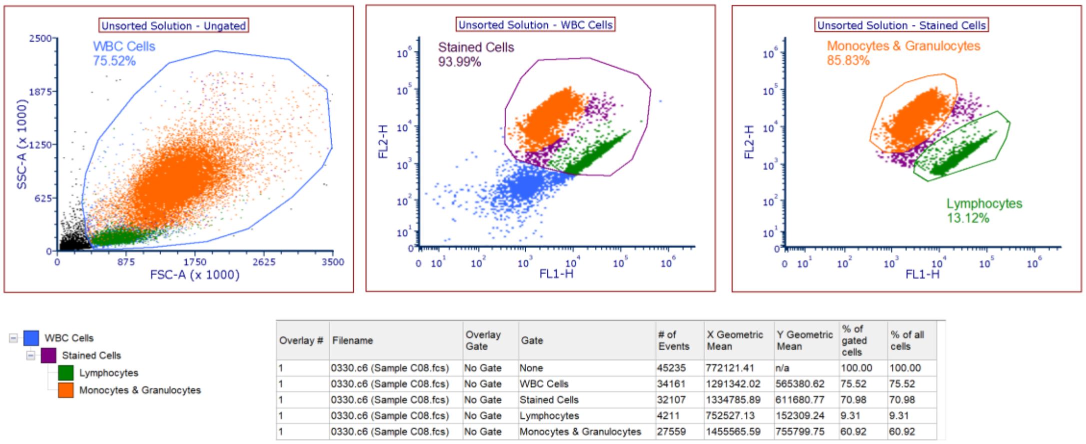
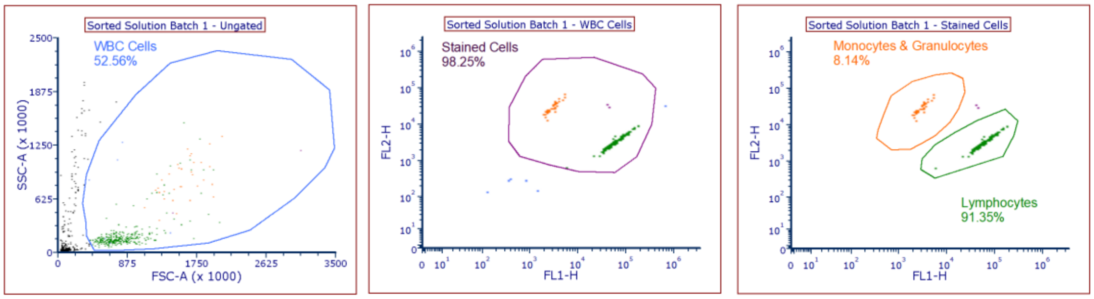


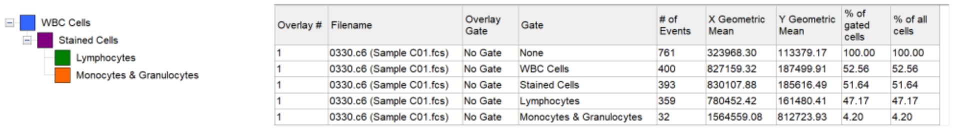
Figure 3.3.20. UNet training curves for human white blood cell sorting experiment



Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample C08.fcs)	No Gate	None	45235	772121.41	n/a	100.00	100.00
1	0330.c6 (Sample C08.fcs)	No Gate	WBC Cells	34161	1291342.02	565380.62	75.52	75.52
1	0330.c6 (Sample C08.fcs)	No Gate	Stained Cells	32107	1334785.89	611680.77	70.98	70.98
1	0330.c6 (Sample C08.fcs)	No Gate	Lymphocytes	4211	752527.13	152309.24	9.31	9.31
1	0330.c6 (Sample C08.fcs)	No Gate	Monocytes & Granulocytes	27559	1455565.59	755799.75	60.92	60.92

Figure 3.3.21. Lymphocyte sorting experiment pre-sorting Accuri particle composition analysis

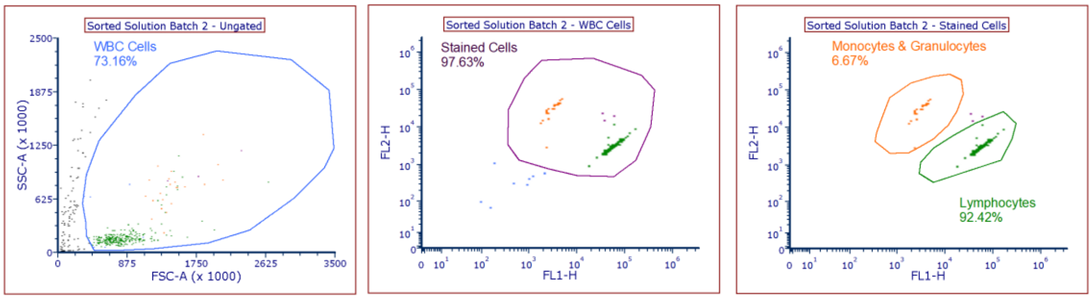


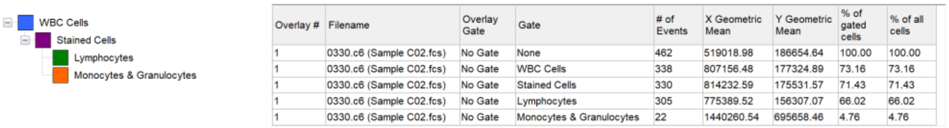


🟦 WBC Cells
  🟪 Stained Cells
    🟩 Lymphocytes
    🟧 Monocytes & Granulocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample C01.fcs)	No Gate	None	761	323968.30	113379.17	100.00	100.00
1	0330.c6 (Sample C01.fcs)	No Gate	WBC Cells	400	827159.32	187499.91	52.56	52.56
1	0330.c6 (Sample C01.fcs)	No Gate	Stained Cells	393	830107.88	185616.49	51.64	51.64
1	0330.c6 (Sample C01.fcs)	No Gate	Lymphocytes	359	780452.42	161480.41	47.17	47.17
1	0330.c6 (Sample C01.fcs)	No Gate	Monocytes & Granulocytes	32	1564559.08	812723.93	4.20	4.20

Figure 3.3.22. Lymphocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 1 solution

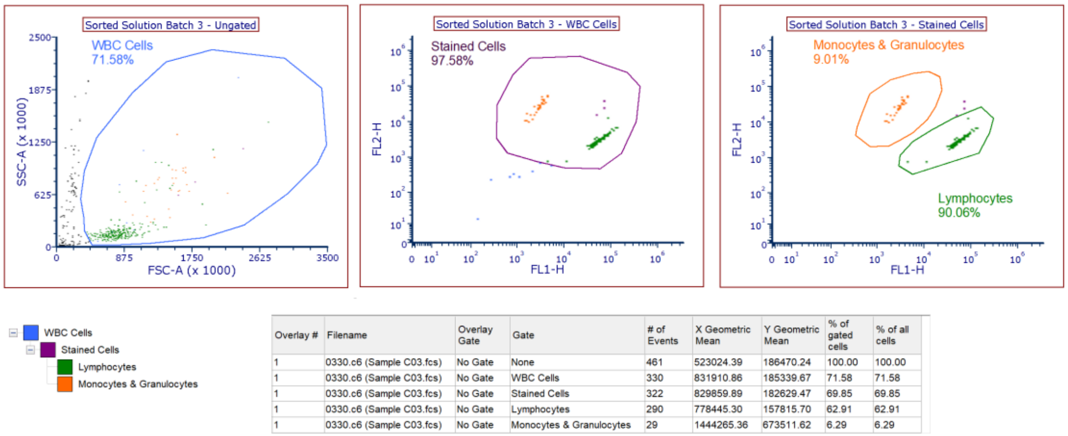




🟦 WBC Cells
  🟪 Stained Cells
    🟩 Lymphocytes
    🟧 Monocytes & Granulocytes

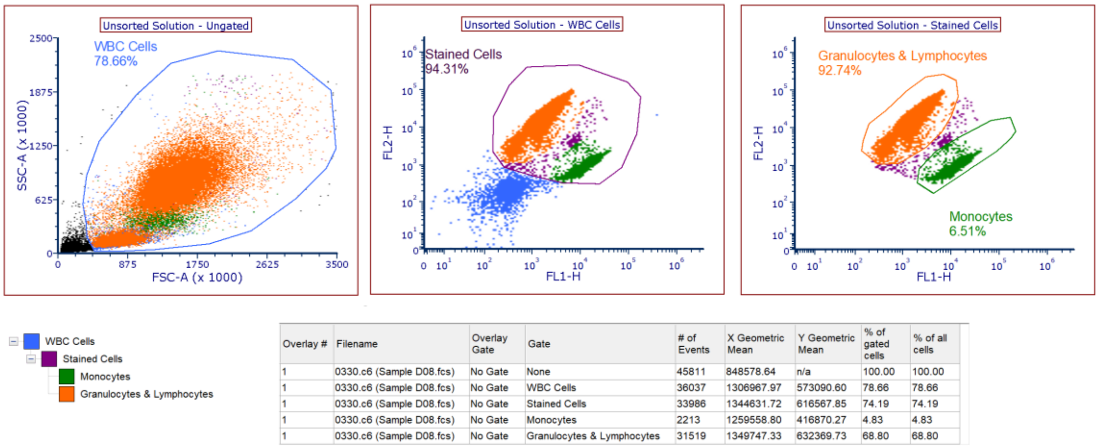
Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample C02.fcs)	No Gate	None	462	519018.98	186654.64	100.00	100.00
1	0330.c6 (Sample C02.fcs)	No Gate	WBC Cells	338	807156.48	177324.89	73.16	73.16
1	0330.c6 (Sample C02.fcs)	No Gate	Stained Cells	330	814232.59	175531.57	71.43	71.43
1	0330.c6 (Sample C02.fcs)	No Gate	Lymphocytes	305	775389.52	156307.07	66.02	66.02
1	0330.c6 (Sample C02.fcs)	No Gate	Monocytes & Granulocytes	22	1440260.54	695658.46	4.76	4.76

Figure 3.3.23. Lymphocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 2 solution



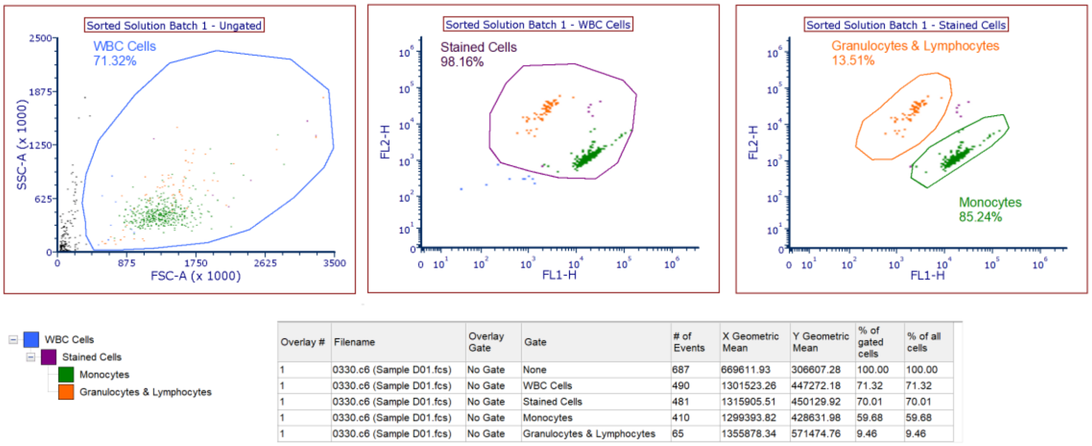
Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample C03.fcs)	No Gate	None	461	523024.39	186470.24	100.00	100.00
1	0330.c6 (Sample C03.fcs)	No Gate	WBC Cells	330	831910.86	185339.67	71.58	71.58
1	0330.c6 (Sample C03.fcs)	No Gate	Stained Cells	322	829859.89	182629.47	69.85	69.85
1	0330.c6 (Sample C03.fcs)	No Gate	Lymphocytes	290	778445.30	157815.70	62.91	62.91
1	0330.c6 (Sample C03.fcs)	No Gate	Monocytes & Granulocytes	29	1444265.36	673511.62	6.29	6.29

Figure 3.3.24. Lymphocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 3 solution



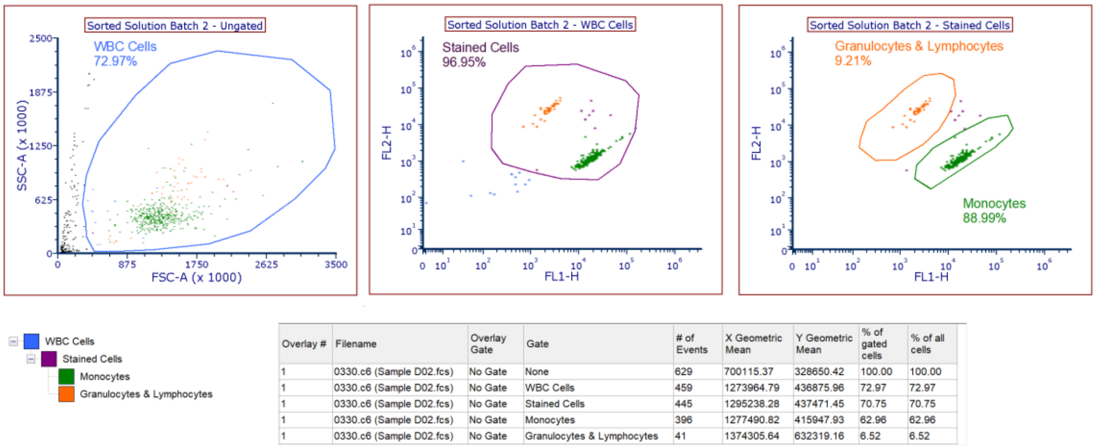
Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample D08.fcs)	No Gate	None	45811	848578.64	n/a	100.00	100.00
1	0330.c6 (Sample D08.fcs)	No Gate	WBC Cells	36037	1306967.97	573090.60	78.66	78.66
1	0330.c6 (Sample D08.fcs)	No Gate	Stained Cells	33986	1344631.72	616567.85	74.19	74.19
1	0330.c6 (Sample D08.fcs)	No Gate	Monocytes	2213	1259558.80	416870.27	4.83	4.83
1	0330.c6 (Sample D08.fcs)	No Gate	Granulocytes & Lymphocytes	31519	1349747.33	632369.73	68.80	68.80

Figure 3.3.25. Monocyte sorting experiment pre-sorting Accuri particle composition analysis



Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample D01.fcs)	No Gate	None	687	669611.93	306607.28	100.00	100.00
1	0330.c6 (Sample D01.fcs)	No Gate	WBC Cells	490	1301523.26	447272.18	71.32	71.32
1	0330.c6 (Sample D01.fcs)	No Gate	Stained Cells	481	1315905.51	450129.92	70.01	70.01
1	0330.c6 (Sample D01.fcs)	No Gate	Monocytes	410	1299393.82	428631.98	59.68	59.68
1	0330.c6 (Sample D01.fcs)	No Gate	Granulocytes & Lymphocytes	65	1355878.34	571474.76	9.46	9.46

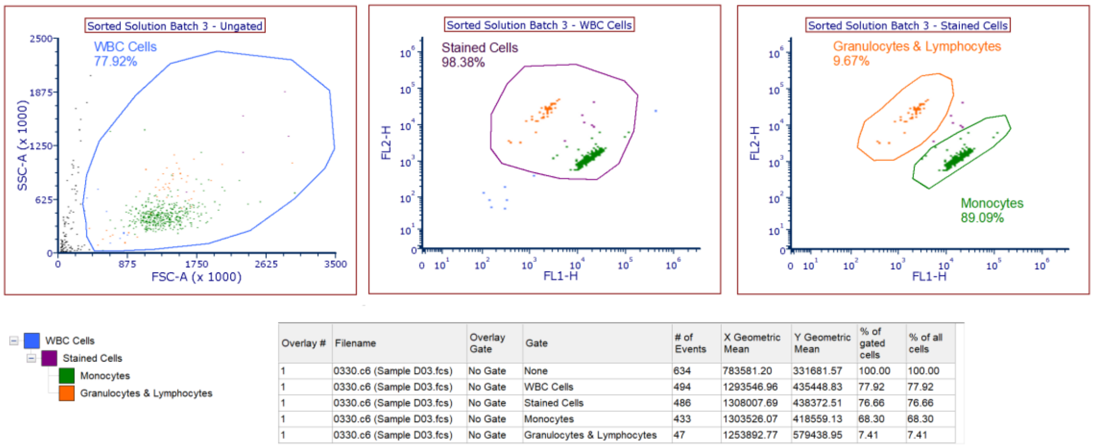
Figure 3.3.26. Monocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 1 solution



- 🟦 WBC Cells
  - 🟪 Stained Cells
    - 🟩 Monocytes
    - 🟧 Granulocytes & Lymphocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample D02.fcs)	No Gate	None	629	700115.37	328650.42	100.00	100.00
1	0330.c6 (Sample D02.fcs)	No Gate	WBC Cells	459	1273964.79	436875.96	72.97	72.97
1	0330.c6 (Sample D02.fcs)	No Gate	Stained Cells	445	1295238.28	437471.45	70.75	70.75
1	0330.c6 (Sample D02.fcs)	No Gate	Monocytes	396	1277490.82	415947.93	62.96	62.96
1	0330.c6 (Sample D02.fcs)	No Gate	Granulocytes & Lymphocytes	41	1374305.64	632319.16	6.52	6.52

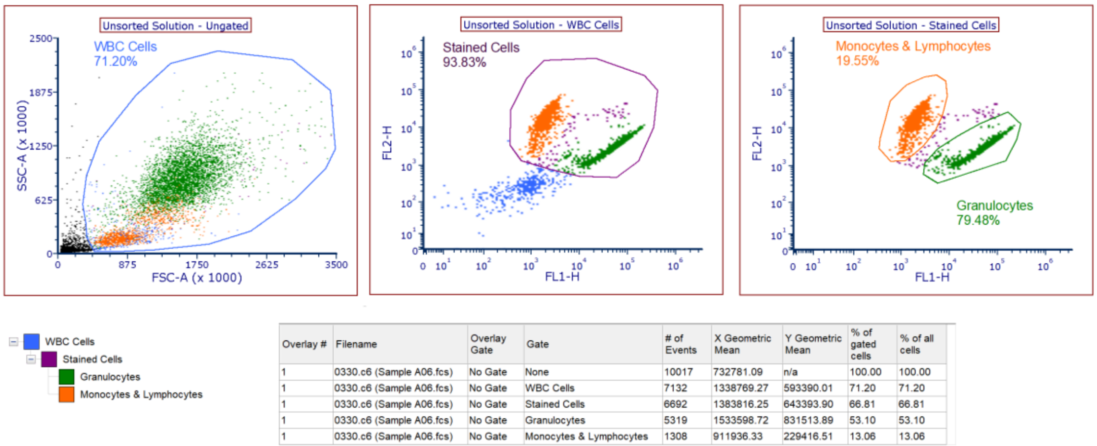
Figure 3.3.27. Monocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 2 solution



- 🟦 WBC Cells
  - 🟪 Stained Cells
    - 🟩 Monocytes
    - 🟧 Granulocytes & Lymphocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample D03.fcs)	No Gate	None	634	783581.20	331681.57	100.00	100.00
1	0330.c6 (Sample D03.fcs)	No Gate	WBC Cells	494	1293546.96	435448.83	77.92	77.92
1	0330.c6 (Sample D03.fcs)	No Gate	Stained Cells	486	1308007.69	438372.51	76.66	76.66
1	0330.c6 (Sample D03.fcs)	No Gate	Monocytes	433	1303526.07	418559.13	68.30	68.30
1	0330.c6 (Sample D03.fcs)	No Gate	Granulocytes & Lymphocytes	47	1253892.77	579438.95	7.41	7.41

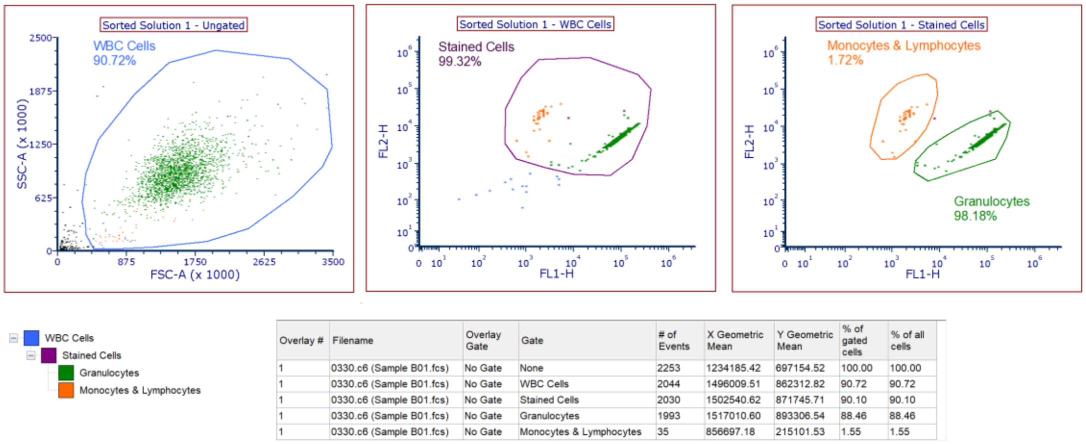
Figure 3.3.28. Monocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 3 solution



🟦 WBC Cells
  🟪 Stained Cells
    🟩 Granulocytes
    🟧 Monocytes & Lymphocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample A06.fcs)	No Gate	None	10017	732781.09	n/a	100.00	100.00
1	0330.c6 (Sample A06.fcs)	No Gate	WBC Cells	7132	1338769.27	593390.01	71.20	71.20
1	0330.c6 (Sample A06.fcs)	No Gate	Stained Cells	6692	1383816.25	643393.90	66.81	66.81
1	0330.c6 (Sample A06.fcs)	No Gate	Granulocytes	5319	1533598.72	831513.89	53.10	53.10
1	0330.c6 (Sample A06.fcs)	No Gate	Monocytes & Lymphocytes	1308	911936.33	229416.51	13.06	13.06

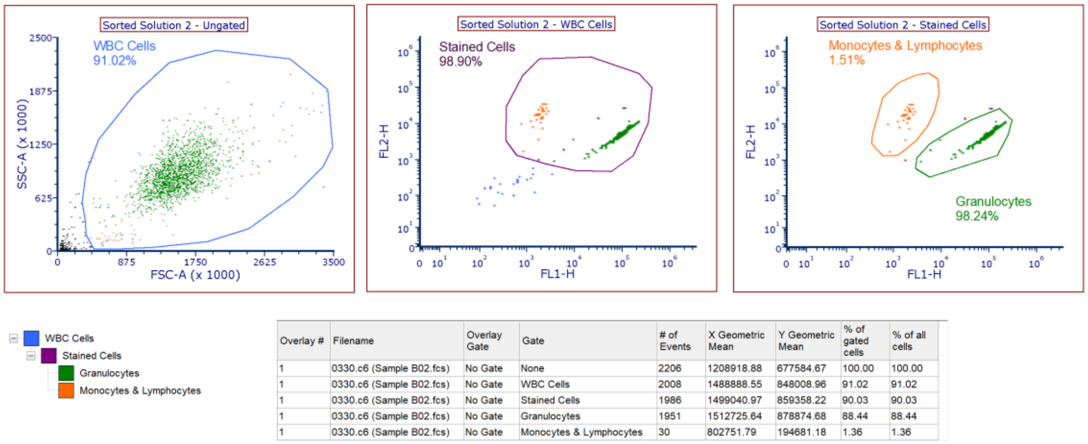
Figure 3.3.29. Granulocyte sorting experiment pre-sorting Accuri particle composition analysis



🟦 WBC Cells
  🟪 Stained Cells
    🟩 Granulocytes
    🟧 Monocytes & Lymphocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample B01.fcs)	No Gate	None	2253	1234185.42	697154.52	100.00	100.00
1	0330.c6 (Sample B01.fcs)	No Gate	WBC Cells	2044	1496009.51	862312.82	90.72	90.72
1	0330.c6 (Sample B01.fcs)	No Gate	Stained Cells	2030	1502540.62	871745.71	90.10	90.10
1	0330.c6 (Sample B01.fcs)	No Gate	Granulocytes	1993	1517010.60	893306.54	88.46	88.46
1	0330.c6 (Sample B01.fcs)	No Gate	Monocytes & Lymphocytes	35	856697.18	215101.53	1.55	1.55

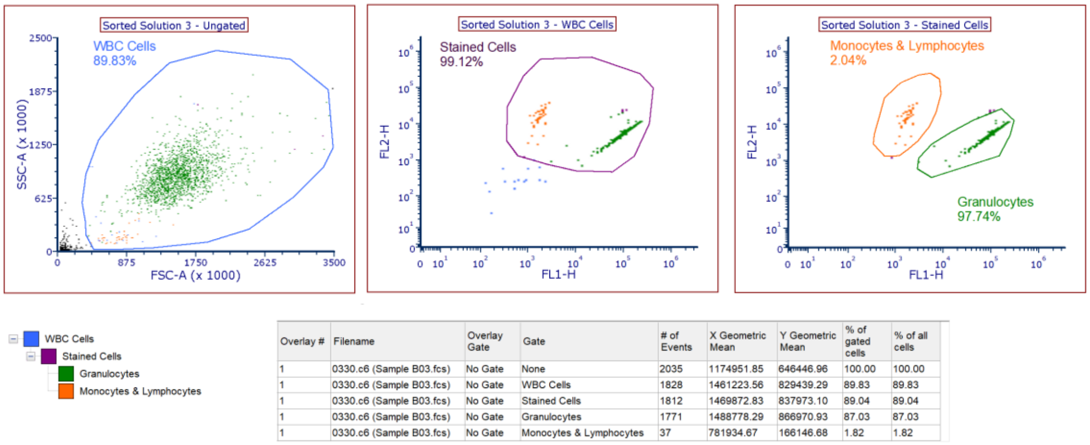
Figure 3.3.30. Granulocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 1 solution



🟦 WBC Cells
  🟪 Stained Cells
    🟩 Granulocytes
    🟧 Monocytes & Lymphocytes

Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample B02.fcs)	No Gate	None	2206	1208918.88	677584.67	100.00	100.00
1	0330.c6 (Sample B02.fcs)	No Gate	WBC Cells	2008	1488888.55	848008.96	91.02	91.02
1	0330.c6 (Sample B02.fcs)	No Gate	Stained Cells	1986	1499040.97	859358.22	90.03	90.03
1	0330.c6 (Sample B02.fcs)	No Gate	Granulocytes	1951	1512725.64	878874.68	88.44	88.44
1	0330.c6 (Sample B02.fcs)	No Gate	Monocytes & Lymphocytes	30	802751.79	194681.18	1.36	1.36

Figure 3.3.31. Granulocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 2 solution



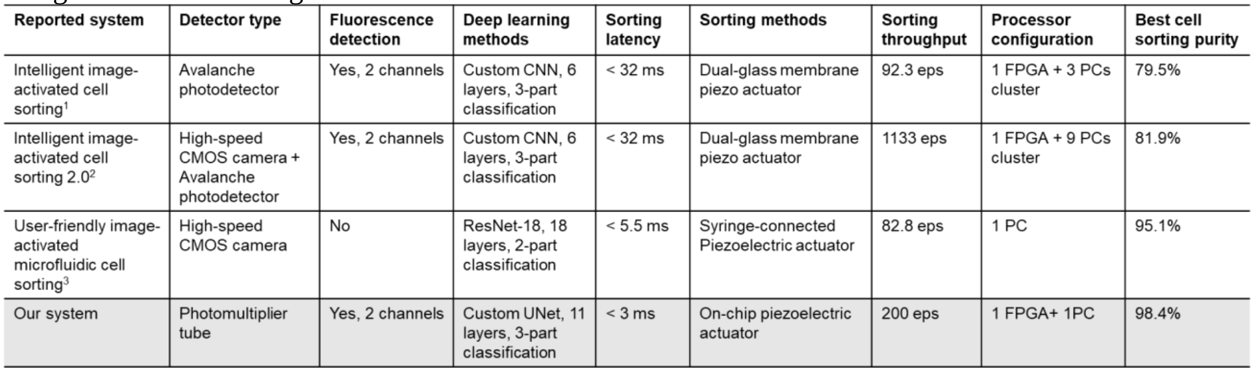
Overlay #	Filename	Overlay Gate	Gate	# of Events	X Geometric Mean	Y Geometric Mean	% of gated cells	% of all cells
1	0330.c6 (Sample B03.fcs)	No Gate	None	2035	1174951.85	646446.96	100.00	100.00
1	0330.c6 (Sample B03.fcs)	No Gate	WBC Cells	1828	1461223.56	829439.29	89.83	89.83
1	0330.c6 (Sample B03.fcs)	No Gate	Stained Cells	1812	1469872.83	837973.10	89.04	89.04
1	0330.c6 (Sample B03.fcs)	No Gate	Granulocytes	1771	1488778.29	866970.93	87.03	87.03
1	0330.c6 (Sample B03.fcs)	No Gate	Monocytes & Lymphocytes	37	781934.67	166146.68	1.82	1.82

Figure 3.3.32. Granulocyte sorting experiment post-sorting Accuri particle composition analysis for post-sorting batch 3 solution

3.4 Discussion

In this study, we demonstrated a low-latency AI-inferencing microfluidic image-activated cell sorting system using a fast-learned customized UNet CNN model and a single FPGA-PC processor. Our system achieved less than 3 ms sorting latency from image acquisition to sorting, including data processing with AI. We used the system to achieve label-free sorting with a throughput greater than 200 cells/s and sorting purity of 96.6% for beads and 92.0%, 89.05%, and 98.4% lymphocyte, monocyte, and granulocyte in 3-part human white blood cell sorting experiments, all based on the CNN model inference prediction from the brightfield transmission images. Our system provides a compact, sterile, low-latency, and low-cost IACS platform for biomedical research with superior performance compared with other reported CNN-assisted IACS systems (See Table 3.4.1).

Table 3.4.1. Comparison of imaging-activated cell sorters using CNN. Sorting purities were compared using the best cell sorting results



Reported system	Detector type	Fluorescence detection	Deep learning methods	Sorting latency	Sorting methods	Sorting throughput	Processor configuration	Best cell sorting purity
Intelligent image-activated cell sorting[1]	Avalanche photodetector	Yes, 2 channels	Custom CNN, 6 layers, 3-part classification	< 32 ms	Dual-glass membrane piezo actuator	92.3 eps	1 FPGA + 3 PCs cluster	79.5%
Intelligent image-activated cell sorting 2.0[2]	High-speed CMOS camera + Avalanche photodetector	Yes, 2 channels	Custom CNN, 6 layers, 3-part classification	< 32 ms	Dual-glass membrane piezo actuator	1133 eps	1 FPGA + 9 PCs cluster	81.9%
User-friendly image-activated microfluidic cell sorting[3]	High-speed CMOS camera	No	ResNet-18, 18 layers, 2-part classification	< 5.5 ms	Syringe-connected Piezoelectric actuator	82.8 eps	1 PC	95.1%
Our system	Photomultiplier tube	Yes, 2 channels	Custom UNet, 11 layers, 3-part classification	< 3 ms	On-chip piezoelectric actuator	200 eps	1 FPGA+ 1PC	98.4%

One constraint of the current system is that the AI-assisted IACS system generates only 2D cell images, which cannot reveal the tomographic information of cells as 3D bodies and may experience occlusions and dependence on viewing perspective. A natural net is to incorporate the deep learning and AI-inference model, as well as the soring capability into the 3D imaging flow cytometer design reported in our earlier publications [32], [66], [72]. Second, the current data processing system is constructed based on a generic FPGA-PC architecture for cost advantages, but the design is not optimized for specific data processing algorithms of the image-activated cell sorting. We observed relatively large event-to-event processing time variations and large data transfer latency (>1.2 ms) from the FPGA module to the PC-GPU system using a generic data transfer framework (National Instruments API). The data transfer overhead can be reduced by developing a customized System-on-Chip (SOC) board and transferring the current data processing software to hardware description languages such as Verilog. Regarding the deep learning aspect, different deep learning algorithms can be used in addition to supervised learning methods. Semi-supervised, unsupervised learning approaches and other methods could be used for cell analysis and sorting. Such approaches could potentially define new subpopulations and provide insights to researchers when combined with cell sorting techniques.

Despite the areas for further improvements, our system has substantially advanced state of the art for image-activated cell sorting (IACS) that is viewed as a paradigm-shifting technology for biomedical research by: 1) incorporating deep learning and AI-inferencing for the sorting gate, a fundamental change from conventional approaches of sorting by image processing and pattern recognition; 2) setting a record for the shortest AI inferencing latency by using a single FPGA-PC for a low-cost solution; 3) achieving short training time of ~30 minutes for high performance deep learning models via network architecture optimization. These accomplishments make the AI-assisted IACS system a unique and attractive tool for biomedical experiments that involve the detection and isolation of cell types or cell properties from large cell images with or without presumptive image features for cell classification in a labeled or label-free manner.

Chapter 3 has been submitted for publication of the material, R. Tang, L. Xia, B. Gutierrez, I. Gagne, A. Munoz, K. Eribez, N. Jagnandan, X. Chen, Z. Zhang, L. Waller, W. Alaynick, S. Cho, C. An, Y.-H. Lo. "Low-latency label-free image-activated cell sorting using fast deep learning and AI inferencing'. The dissertation author is the first author of this paper.

Chapter 4 Summary and Outlook

4.1 Summary

This dissertation enabled the high-throughput label-free cellular image analysis based on the IFC technology and demonstrated the low-latency image-activated cell sorting using the deep learning CNN models. We have developed a label-free cellular analysis workflow based on the 3D-IFC system and artificial intelligence to support high-throughput image processing and analysis, explored the clinical and biological application of this cellular analysis workflow, and constructed a low-latency IACS system using the fast deep learning and AI inference. In contrast of conventional cellular image analysis workflow that is limited by the number of morphological features defined by human vision, the AI solution driven by machine vision provides a highly versatile and scalable solution to address the high-throughput and low-latency real-time image requirements in the IFC and IACS technologies, opening a new venue for relating cell phenotype and genotype in biology and biomedical fields where cell heterogeneity is recognized.

4.2 Outlook: Predictive cellular analysis and high-throughput single-cell manipulation

One of the major challenges of cellular biology is to conduct predictive cellular analysis from existing data. The IFC technology enables high-throughput data acquisition and analysis for discrete observation intervals. However, the utilization of the IFC technology to conduct the time-lapse study is very useful for evaluating drug response, developing cancer therapy, and isolating target cell groups. In addition, the approach to conducting data-driven predictive cellular analysis has not been explored due to a lack of enabling technology or tools. Fortunately, this dissertation provides a path to bridge the gap between high-throughput image acquisition and data analysis using artificial intelligence approaches, yet we only explore the possibility of using supervised learning approaches for high-throughput real-time image analysis. One can extend the scope of artificial intelligence methods to explore the possibility of using semi-supervised and unsupervised learning techniques for high-throughput predictive cellular analysis using the IFC and IACS platforms. Such techniques also reduce or eliminate the requirement of ground truth label generation during the training process. The ground truth label is usually technically challenging to obtain during the predictive cellular analysis or time-lapse study.

Compared with the cell population sorting techniques in the majority of the IACS system, high-throughput single-cell manipulation based on high-throughput image processing can provide an unprecedented bridge between single-cell molecular analysis and single-cell image analysis to connect phenotype and genotype analysis with single-cell resolution. One can explore the possibility of integrating single-cell isolation and manipulation technology with the IFC platform based on high-throughput image processing and a real-time cellular image analysis pipeline. Possible approaches to couple closed-loop high-precision cell isolation and manipulation robotics platform with the IFC system to achieve single-cell level manipulation and isolation.

Appendix

The human cancer cell and leukocyte datasets used in Chapter 2 are available on request. The PyTorch implementation of all deep learning models and training code is publicly available on Github: https://github.com/rut011/3D-IFC-ML.git

The CHO-K1 dataset used in Chapter 2 is available on request. The PyTorch implementation of the deep learning model and training code is publicly available on Github: https://github.com/rut011/TIGSImageSegmentation

The liver cells datasets used in Chapter 2 are available on request. The PyTorch implementation of the deep learning model and training code is publicly available on Github: https://github.com/rut011/LiverCellML

The white blood cells dataset used in Chapter 3 are available on request. The MATLAB implementation of image preprocessing code and the Pytorch implementation of the deep learning model and training code is publicly available on Github: https://github.com/rut011/CNN-assisted-IACS

Bibliography

[1] S. J. Altschuler and L. F. Wu, “Cellular heterogeneity: do differences make a difference?,” *Cell,* vol. 141, no. 4, pp. 559–63, May 2010, doi: 10.1016/j.cell.2010.04.033.

[2] B. Snijder and L. Pelkmans, “Origins of regulated cell-to-cell variability,” *Nat. Rev. Mol. Cell Biol.*, vol. 12, no. 2, pp. 119–125, Feb. 2011, doi: 10.1038/nrm3044.

[3] V. Almendro, A. Marusyk, and K. Polyak, “Cellular Heterogeneity and Molecular Evolution in Cancer,” *Annu. Rev. Pathol. Mech. Dis.*, vol. 8, no. 1, pp. 277–302, Jan. 2013, doi: 10.1146/annurev-pathol-020712-163923.

[4] L. Pelkmans, “Using Cell-to-Cell Variability--A New Era in Molecular Biology,” *Science (80-.).*, vol. 336, no. 6080, pp. 425–426, Apr. 2012, doi: 10.1126/science.1222161.

[5] O. D. Laerum and T. Farsund, “Clinical application of flow cytometry: A review,” *Cytometry,* vol. 2, no. 1, pp. 1–13, Jul. 1981, doi: 10.1002/cyto.990020102.

[6] W. G. Telford, T. Hawley, F. Subach, V. Verkhusha, and R. G. Hawley, “Flow cytometry of fluorescent proteins,” *Methods,* vol. 57, no. 3, pp. 318–330, Jul. 2012, doi: 10.1016/J.YMETH.2012.01.003.

[7] N. S. Barteneva, E. Fasler-Kan, and I. A. Vorobjev, “Imaging flow cytometry: coping with heterogeneity in biological systems.,” *J. Histochem. Cytochem.*, vol. 60, no. 10, pp. 723–33, Oct. 2012, doi: 10.1369/0022155412453052.

[8] Y. Han and Y.-H. Lo, “Imaging Cells in Flow Cytometer Using Spatial-Temporal Transformation.,” *Sci. Rep.*, vol. 5, p. 13267, 2015, doi: 10.1038/srep13267.

[9] T. Blasi H. Hennig, H.D. Summers, F.j. Theis, J.O. Patterson, D. Davies, A. Filby, A.E. Carpenter, and P. Rees, “Label-free cell cycle analysis for high-throughput imaging flow cytometry,” *Nat. Commun.*, vol. 7, 2016, doi: 10.1038/ncomms10256.

[10] D. A. Basiji, W. E. Ortyn, L. Liang, V. Venkatachalam, and P. Morrissey, “Cellular image analysis and imaging by flow cytometry.,” *Clin. Lab. Med.*, vol. 27, no. 3, pp. 653–70, viii, Sep. 2007, doi: 10.1016/j.cll.2007.05.008.

[11] Keisuke Godaa, A. Ayazia, D. R. Gossett, J. Sadasivama, C. K. Lonappana, E. Sollier, A. M. Farda, S. C. Hurb, J. Adama, C. Murrayc, C. Wanga, N. Brackbilla, D. D. Carlo, and B. Jalalia, “High-throughput single-microparticle imaging flow analyzer,” *Proc. Natl. Acad. Sci.*, vol. 109, no. 29, pp. 11630–11635, 2012, doi: 10.1073/pnas.1204718109.

[12] Y. Han, Y. Gu, A. C. Zhang, and Y.-H. Lo, “Review: imaging technologies for flow cytometry,” *Lab Chip,* vol. 16, no. 24, pp. 4639–4647, Nov. 2016, doi: 10.1039/C6LC01063F.

[13] F. Huang, G. Sirinakis, E. S. Allgeyer, L. K. Schroeder, W. C. Duim, E. B. Kromann, T. Phan, F. E. Rivera-Molina, J. R. Myers, I. Irnov, and M. Lessard, “Ultra-High Resolution 3D Imaging of Whole Cells.,” *Cell*, vol. 166, no. 4, pp. 1028–1040, Aug. 2016, doi: 10.1016/j.cell.2016.06.016.

[14] J.-A. Conchello and J. W. Lichtman, “Optical sectioning microscopy,” *Nat. Methods*, vol. 2, no. 12, pp. 920–931, Dec. 2005, doi: 10.1038/nmeth815.

[15] P. G. Pitrone, J. Schindelin, L. Stuyvenberg, S. Preibisch, M. Weber, K. W. Eliceiri, J. Huisken, and P. Tomancak, “OpenSPIM: an open-access light-sheet microscopy platform,” *Nat. Methods*, vol. 10, no. 7, pp. 598–599, Jul. 2013, doi: 10.1038/nmeth.2507.

[16] O. O. Dada, B. J. Huge, and N. J. Dovichi, “Simplified sheath flow cuvette design for ultrasensitive laser induced fluorescence detection in capillary electrophoresis.,” *Analyst*, vol. 137, no. 13, pp. 3099–101, Jul. 2012, doi: 10.1039/c2an35321k.

[17] S. J. Altschuler and L. F. Wu, “Cellular Heterogeneity: Do Differences Make a Difference?,” *Cell*, vol. 141, no. 4, pp. 559–563, 2010, doi: 10.1016/j.cell.2010.04.033.

[18] S. E. Boddington, E. J. Sutton, T. D. Henning, A. J. Nedopil, B. Sennino, A. Kim, and H. E. Daldrup-Link, “Labeling human mesenchymal stem cells with fluorescent contrast agents: The biological impact,” *Mol. Imaging Biol.*, vol. 13, no. 1, pp. 3–9, 2011, doi: 10.1007/s11307-010-0322-0.

[19] M. Doan, I. Vorobjev, P. Rees, A. Filby, O. Wolkenhauer, A.E. Goldfeld, J. Lieberman, N. Barteneva, A.E. Carpenter, and H. Hennig, “Diagnostic Potential of Imaging Flow Cytometry,” *Trends Biotechnol.*, vol. 36, no. 7, pp. 649–652, 2018, doi: 10.1016/j.tibtech.2017.12.008.

[20] Y. Li, A. Mahjoubfar, C. L. Chen, K. R. Niazi, L. Pei, and B. Jalali, “Deep Cytometry: Deep learning with Real-time Inference in Cell Sorting and Flow Cytometry,” *Sci. Rep.*, vol. 9, no. 1, p. 11088, 2019, doi: 10.1038/s41598-019-47193-6.

[21] A. Isozaki, J. Harmon,Y. Zhou,S. Li,Y. Nakagawa, M. Hayashi,H. Mikami,C. Lei, and K. Goda, “AI on a chip,” *Lab Chip*, vol. 20, no. 17, 2020, doi: 10.1039/D0LC00521E.

[22] C.L. Chen, A. Mahjoubfar, L.-C. Tai, I.K. Blaby, A. Huang, K.R. Niazi, and B. Jalali, “Deep Learning in Label-free Cell Classification,” *Sci. Rep.*, vol. 6, no. March, pp. 1–16, 2016, doi: 10.1038/srep21471.

[23] Y. Wu, Y. Zhou, C.-J. Huang, H. Kobayashi, S. Yan, Y. Ozeki, Y. Wu, C.-W. Sun, A. Yasumoto, Y. Yatomi, C. Lei, and K. Goda “Intelligent frequency-shifted optofluidic time-stretch quantitative phase imaging,” *Opt. Express*, vol. 28, no. 1, p. 519, 2020, doi: 10.1364/oe.380679.

[24] Y. Li, B. Cornelis, A. Dusa, G. Vanmeerbeeck, D. Vercruyssem R. Sohn, K. Blaszkiewicz, D. Prodanov, P. Schelkens, and L. Lagae, “Accurate label-free 3-part leukocyte recognition with single cell lens-free imaging flow cytometry,” *Comput. Biol.*

Med., vol. 96, pp. 147–156, 2018.

[25] S. Saeb, L. Lonini, A. Jayaraman, D. Mohr, and K. Kording, “Voodoo Machine Learning for Clinical Predictions,” *bioRxiv*, p. 059774, 2016, doi: 10.1101/059774.

[26] M. Nassar, M. Doan, A. Filby, O. Wolkenhauer, D. K. Fogg, J. Piasecka, C. A. Thornton, A. E. Carpenter, H. D. Summers, P. Rees, and H. Hennig, “Label-Free Identification of White Blood Cells Using Machine Learning,” *Cytom. Part A*, vol. 95, no. 8, pp. 836–842, 2019, doi: 10.1002/cyto.a.23794.

[27] M. Lippeveld, C. Knill, E. Ladlow, A. Fuller, L. J. Michaelis, Y. Saeys, A. Filby, and D. Peralta, “Classification of human white blood cells using machine learning for stain-free imaging flow cytometry,” *Cytom. Part A*, vol. 97, no. 3, pp. 308–319, 2020, [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/cyto.a.23920.

[28] L. S. Cram and A. Brunsting, “Fluorescence and light-scattering measurements on hog cholear-infected PK-15 cells,” *Exp. Cell Res.*, vol. 78, no. 1, pp. 209–213, 1973, [Online]. Available: https://www.sciencedirect.com/science/article/pii/0014482773900566?via%3Dihub.

[29] G. C. Salzman, J. M. Crowell, T. T. Truijillo, A. Romero, P. F. Mullaney, and P. M. Labauve, “Cell Classification by Laser Light Scattering: Identification and Separation of Unstained Leukocytes,” *Acta Cytol*, vol. 19, no. 4, pp. 374–7, 1975, doi: 808927.

[30] C. Liu, C. Capjack, and W. Rozmus, “3-D simulation of light scattering from biological cells and cell differentiation,” *J. Biomed. Opt.*, vol. 10, no. 1, p. 014007, 2005, doi: 10.1117/1.1854681.

[31] O. C. Marina, C. K. Sanders, and J. R. Mourant, “Correlating light scattering with internal cellular structures,” *Biomed. Opt. Express*, vol. 3, no. 2, p. 296, 2012, doi: 10.1364/boe.3.000296.

[32] Y. Han, R. Tang, Y. Gu, A. Zhang, W. Cai, V. Castor, S. Cho, W. Alynick, and Y.-H. Lo, “Cameraless high-throughput three-dimensional imaging flow cytometry,” *Optica*, vol. 6, no. 10, p. 1297, 2019, doi: 10.1364/optica.6.001297.

[33] N. Japkowicz and S. Stephen, “The class imbalance problem: A systematic study,” *Intell. Data Anal.*, vol. 6, pp. 429–449, 2002, doi: 10.3233/IDA-2002-6504.

[34] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[35] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.

[36] C. K. Yeh, W. C. Wu, W. J. Ko, and Y. C. F. Wang, “Learning deep latent spaces for multi-label classification,” *31st AAAI Conf. Artif. Intell. AAAI 2017*, pp. 2838–2844, 2017.

[37] C. Aytekin, X. Ni, F. Cricri, and E. Aksu, “Clustering and Unsupervised Anomaly Detection with l2 Normalized Deep Auto-Encoder Representations,” *Proc. Int. Jt. Conf. Neural Networks*, vol. 2018-July, 2018, doi: 10.1109/IJCNN.2018.8489068.

[38] K. H. Brodersen, C. S. Ong, K. E. Stephan, and J. M. Buhmann, “The Balanced Accuracy and Its Posterior Distribution,” in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 3121–3124, doi: 10.1109/ICPR.2010.764.

[39] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[40] L. V. U. Nl and G. Hinton, “Visualizing Data using t-SNE Laurens van der Maaten,” *J. Mach. Learn. Res.*, vol. 1, pp. 1–48, 2008, doi: 10.1007/s10479-011-0841-3.

[41] H. K. Walker, W. D. Hall, and J. W. Hurst, *Peripheral Blood Smear--Clinical Methods: The History, Physical, and Laboratory Examinations*. Butterworths, 1990.

[42] C. A. LaBelle, A. Massaro, B. Cortés-Llanos, C. E. Sims, and N. L. Allbritton, “Image-Based Live Cell Sorting,” *Trends Biotechnol.*, vol. 39, no. 6, pp. 613–623, 2021, doi: 10.1016/j.tibtech.2020.10.006.

[43] D. Schraivogel, T. M. Kuhn, B. Rauscher, M. Rodríguez-Martínez,M. Paulsen, K. Owsley, A. Middlebrook, C. Tischer, B. Ramasz,D. Ordoñez-Rueda, M. Dees, S. Cuylen-Haering, E. Diebold, L. M. Steinmetz, “High-speed fluorescence image – enabled cell sorting,” *Science (80-.).*, vol. 320, no. January, pp. 315–320, 2022.

[44] J. P. Roberts and J. Tesdorpf, “Phenotypic Drug Discovery with High Content Screening,” 2015.

[45] Y. Gu, A. C. Zhang, Y. Han, J. Li, C. Chen, and Y.-H. Lo, “(2019) Machine Learning Based Real-Time Image-Guided Cell Sorting and Classification.pdf,” *Cytometery Part A*, vol. 95.5, pp. 499–509, 2019.

[46] C. Brasko, K. Smith, C. Molnar, N. Farago, L. Hegedus, A. Balind, T. Balassa, A. Szkalisity, F. Sukosd, K. Kocsis, B, Balint, L, Paavolainen, M, Z. Enyedi, I. Nagy, L. G. Puskas, L. Haracska, G. Tamas, and P. Horvath, “Intelligent image-based in situ single-cell isolation,” *Nat. Commun.*, vol. 9, no. 1, pp. 1–7, 2018, doi: 10.1038/s41467-017-02628-4.

[47] W. A. Bonner, H. R. Hulett, R. G. Sweet, and L. A. Herzenberg, “Fluorescence activated cell sorting,” *Rev. Sci. Instrum.*, vol. 43, no. 3, pp. 404–409, 1972, [Online]. Available: https://herzenberglab.stanford.edu/sites/g/files/sbiybj10046/f/publications/lah063.pdf.

[48] L. A. Herzenberg, R. G. Sweet, L. A. Herzenberg, L. A. Herzenberg, R. G. Sweet, and L. A. Herzenberg, “Fluorescence-activated Cell Sorting,” *Sci. Am.*, vol. 234, no. 3, pp. 108–

118, 1976, [Online]. Available: https://www.jstor.org/stable/24950310.

[49] Y. Shen, Y. Yalikun, and Y. Tanaka, “Recent advances in microfluidic cell sorting systems,” *Sensors Actuators, B Chem.*, vol. 282, no. November, pp. 268–281, 2019, doi: 10.1016/j.snb.2018.11.025.

[50] D. Gao, F. Jin, M. Zhou, and Y. Jiang, “Recent advances in single cell manipulation and biochemical analysis on microfluidics,” *Analyst*, vol. 144, no. 3, pp. 766–781, 2019, doi: 10.1039/c8an01186a.

[51] N. Nitta, T. Iino, A. Isozaki, M. Yamagishi, Y. Kitahama, S. Sakuma, Y. Suzuki, H. Tezuka, M. Oikawa, F. Arai, T. Asai, D. Deng, H. Fukuzawa, M. Hase, T. Hasunuma, T. Hayakawa , K. Hiraki, K. Hiramatsu, Y. Hoshino, M. Inaba, Y. Inoue, T. Ito, M. Kajikawa, H. Karakawa, Y. Kasai, Y. Kato, H. Kobayashi, C. Lei, S. Matsusaka, H. Mikami, A. Nakagawa, K. Numata, T. Ota,T. Sekiya, K. Shiba, Y. Shirasaki, N. Suzuki, S. Tanaka, S. Ueno, Hiroshi Watarai, T. Yamano, M. Yazawa, Y. Yonamine, D. D. Carlo, Y. Hosokawa, S. Uemura, T. Sugimura, Y. Ozeki. and K. Goda, “Raman image-activated cell sorting,” *Nat. Commun.*, vol. 11, no. 1, pp. 1–16, 2020, doi: 10.1038/s41467-020-17285-3.

[52] Broad Institue, “CellProfiler,” 2021. https://cellprofiler-manual.s3.amazonaws.com/CellProfiler-4.2.1/modules/measurement.html.

[53] N. Nitta, T. Sugimura, A, Isozaki, H. Mikami, K. Hiraki, S. Sakuma, T. Iino, F. Arai, T. Endo, Y. Fujiwaki, H. Fukuzawa, M. Hase, T. Hayakawa, K. Hiramatsu, Y. Hoshino, M. Inaba, T. Ito, H. Karakawa, Y. Kasai, K. Koizumi. S. W. Lee, C. Lei, M. Li, T. Maeno, S. Matsusaka, D. Murakami, A. Nakagawa, Y. Oguchi, M. Oikawa, T. Ota, K. Shiba, H. Shintaku, Y. Shirasaki, K. Suga, Y. Suzuki, N. Suzuki, Y. Tanaka,H. Tezuka, C. Toyokawa, Y. Yalikun,M. Yamada, M. Yamagishi, T. Yamano,A. Yasumoto, Y. Yatomi, M. Yazawa, D. D. Carlo, Y.Hosokawa, S.Uemura, Y. Ozeki, and K. Goda, “Intelligent Image-Activated Cell Sorting.,” *Cell*, vol. 175, no. 1, pp. 266-276.e13, Sep. 2018, doi: 10.1016/j.cell.2018.08.028.

[54] A. Isozaki, H. Mikami, H. Tezuka, H. Matsumura, K. Huang, M. Akamine, K. Hiramatsu, T. Iino, T. Ito, H. Karakawa, Y. Kasai, Y. Li, Y. Nakagawa, S. Ohnuki, T. Ota, Y. Qian, S. Sakuma, T. Sekiya, Y. Shirasaki, N. Suzuki, E. Tayyabi, T. Wakamiya, M. Xu, M. Yamagishi, H. Yan, Q. Yu, S. Yan, D. Yuan, W. Zhang, Y. Zhao, F. Arai, R. E. Campbell, C. Danelon, D. D. Carlo, K. Hiraki, Y. Hoshino, Y. Hosokawa, M. Inaba, A. Nakagawa, Y. Ohya, M. Oikawa, S. Uemura, Y. Ozeki, T. Sugimura, N. Nitta, and K. Goda. “Intelligent image-activated cell sorting 2.0,” *Lab Chip*, 2020, doi: 10.1039/d0lc00080a.

[55] A. A. Nawaz, M. Urbanska, M. Herbig, M. Nötzel, M. Kräter, P. Rosendahl, C. Herold, N.e Toepfner, M. Kubánková, R. Goswami, S. Abuhattum, F. Reichel, P. Müller, A. Taubenberger, S. Girardo, A. Jacobi, and J. Guck, “Intelligent image-based deformation-assisted cell sorting with molecular specificity,” *Nat. Methods*, vol. 17, no. 6, pp. 595–599, 2020, doi: 10.1038/s41592-020-0831-y.

[56] X. Chen, L. Waller, J. Chen, R. Tang, Z. Zhang, I. Gagne, B. Gutierrez, S. Cho, C.-Y.

Tseng, I. Y. Lian, and Y.-H. Lo, “Label-free image-encoded microfluidic cell sorter with a scanning Bessel beam,” *APL Photonics*, vol. 6, no. 7, 2021, doi: 10.1063/5.0051354.

[57] K. Lee, S. E. Kim, J. Doh, K. Kim, and W. K. Chung, “User-friendly image-activated microfluidic cell sorting technique using an optimized, fast deep learning algorithm,” *Lab Chip*, vol. 21, no. 9, pp. 1798–1810, 2021, doi: 10.1039/d0lc00747a.

[58] X. Chen, Y. Gu, J. Chen, C.-H. Lee, I. Gagne, R. Tang, L. Waller, Z. Zhang, A. Zhang, Y. Han, W. Wang, I. Y. Lian, S. Cho, and Y.-H. Lo, “Image-guided cell sorting using fast scanning lasers,” *APL Photonics*, vol. 5, no. 4, 2020, doi: 10.1063/1.5144117.

[59] N. S. Barteneva and I. A. Vorobjev Editors, *Imaging Flow Cytometry Methods and Protocols Methods in Molecular Biology*. 2016.

[60] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” pp. 1–11, 2015, [Online]. Available: http://arxiv.org/abs/1511.08458.

[61] J. Koushik, “Understanding Convolutional Neural Networks,” *NIPS Proc.*, no. 3, pp. 1–6, 2016, [Online]. Available: http://arxiv.org/abs/1605.09081.

[62] J. Wu, “Introduction to Convolutional Neural Networks,” 2017. [Online]. Available: https://web.archive.org/web/20180928011532/https://cs.nju.edu.cn/wujx/teaching/15_CNN.pdf.

[63] W. Dai and D. Berleant, “Benchmarking contemporary deep learning hardware and frameworks: A survey of qualitative metrics,” *Proc. - 2019 IEEE 1st Int. Conf. Cogn. Mach. Intell. CogMI 2019*, pp. 148–155, 2019, doi: 10.1109/CogMI48466.2019.00029.

[64] A. Isozaki, J. Harmon, Y. Zhou, S. Li, Y. Nakagawa, M. Hayashi, H. Mikami, C. Lei, and K. Goda, “AI on a chip,” *Lab Chip*, vol. 20, no. 17, pp. 3074–3090, 2020, doi: 10.1039/d0lc00521e.

[65] K. Huang, H. Matsumura, Y. Zhao, M. Herbig, D. Yuan, Y. Mineharu, J. Harmon, J. Findinier, M. Yamagishi, S. Ohnuki, N. Nitta, A. R. Grossman, Y. Ohya, H. Mikami, A. Isozaki, and K. Goda, “Deep imaging flow cytometry,” *Lab Chip*, vol. 22, no. 5, pp. 876–889, 2022, doi: 10.1039/d1lc01043c.

[66] R. Tang, Z. Zhang, X. Chen, L. Waller, A. Zhang, J. Chen, Y. Han, C. An, S. Cho, and Y.-H. Lo, “3D side-scattering imaging flow cytometer and convolutional neural network for label-free cell analysis,” *APL Photonics*, vol. 5, no. 12, 2020, doi: 10.1063/5.0024151.

[67] Y. J. Chiu, S. H. Cho, Z. Mei, V. Lien, T. F. Wu, and Y.-H. Lo, “Universally applicable three-dimensional hydrodynamic microfluidic flow focusing,” *Lab Chip*, vol. 13, no. 9, pp. 1803–1809, 2013, doi: 10.1039/c3lc41202d.

[68] Y. Han and Y. H. Lo, “Imaging Cells in Flow Cytometer Using Spatial-Temporal Transformation,” *Sci. Rep.*, vol. 5, pp. 1–10, 2015, doi: 10.1038/srep13267.

[69] H. Vanholder, "Efficient inference with tensorrt," *GPU Technol. Conf.*, vol. 1, p. 2, 2016, [Online]. Available: https://on-demand.gputechconf.com/gtc-eu/2017/presentation/23425-han-vanholder-efficient-inference-with-tensorrt.pdf.

[70] B. Ulker, S. Stuijk, H. Corporaal, and R. Wijnhoven, "Reviewing inference performance of state-of-the-art deep learning frameworks," *Proc. 23rd Int. Work. Softw. Compil. Embed. Syst. SCOPES 2020*, pp. 48–53, 2020, doi: 10.1145/3378678.3391882.

[71] H. K. Walker, W. D. Hall, and J. W. Hurst, "Clinical Methods: The History," *Phys. Lab. Exam.*, vol. 3, no. 60, 1990.

[72] Z. Zhang, R. Tang, X. Chen, L. Waller, A. Kau, A. A. Fung, B. Gutierrez, C. An, S. Cho, L. Shi, and Y.-H. Lo, "A high-throughput technique to map cell images to cell positions using a 3D imaging flow cytometer," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 119, no. 8, pp. 1–8, 2022, doi: 10.1073/pnas.2118068119.