

# 3D Side Scattering Imaging Flow Cytometer and Convolutional Neural Network for Label-Free Cell Analysis

Rui Tang<sup>1</sup>, Zunming Zhang<sup>1</sup>, Xinyu Chen<sup>1</sup>, Lauren Waller<sup>2</sup>, Alex Ce Zhang<sup>1</sup>, Jiajie Chen<sup>3</sup>, Yuanyuan Han<sup>1</sup>, Cheolhong An<sup>1</sup>, Sung Hwan Cho<sup>4</sup>, and Yu-Hwa Lo<sup>1,a)</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, California 92093, USA

<sup>2</sup> Department of Bioengineering, University of California, San Diego, La Jolla, California 92093, USA

<sup>3</sup> College of Physics and Optoelectronics Engineering, Shenzhen University, Shenzhen 518060, China

<sup>4</sup> NanoCollect Biomedical Inc., San Diego, California 92121, USA

<sup>a)</sup>Author to whom correspondence should be addressed: [ylo@ucsd.edu](mailto:ylo@ucsd.edu)

## ABSTRACT

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 not only produce 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. Thereby we have opened up a new avenue for cell phenotype analysis in biomedical and clinical research.

## I. 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<sup>1-3</sup>. 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 the 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<sup>4</sup>. 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<sup>5</sup>. 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 the high-content single-modal or multi-modal cell images<sup>6-8</sup>.

Several recent works have applied imaging techniques with the IFC systems and the 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<sup>9</sup>. 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<sup>10</sup>. 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%<sup>11</sup>. 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 predicts the same subject in the validation step<sup>12</sup>. 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%<sup>13</sup>. 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<sup>14</sup>.

All label-free IFC technologies are developed based on the understanding that light scattering properties are related to cell morphology since the 1970's<sup>15,16</sup>. Until the past two decades, very few studies examined the correlation between the intracellular structure and scattered light pattern. 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<sup>17</sup>. A recent experimental study also demonstrated that while cellular organelles contribute to the side-scattering pattern,

the nucleus has the largest contribution<sup>18</sup>. 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<sup>19</sup>.

In this paper, 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<sup>19</sup>. This offers a new modality to reveal the 3D internal structures of cells through the 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 of internal cellular structure than the 2D SSC image.

## II. METHODS AND MATERIALS

### A. Intelligent 3D-IFC-based cell type analysis workflow

The three-stage workflow of intelligent 3D-IFC-based cell type analysis is illustrated in Figure 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 applying the temporal-spatial

transformation method<sup>19,20</sup>. 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 Supplementary Figure 1 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 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.

## B. Cell preparation and image acquisition

### 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. 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  $20 \times 20 \times 20 \mu\text{m}$  space. The 3D image stack was  $80 \times 80 \times 80$  pixels. In the meantime, a 2D projection image was generated by collapsing the depth dimension of the 3D image stack. Figure 2 shows the 3D SSC image stacks of human cancer cells in comparison with the 2D SSC projection and bright-field microscopic images.

Figure 2. 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:  $5 \mu\text{m}$ .

## C. 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 images.

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<sup>21</sup>. 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.

## D. 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 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<sup>22</sup>. 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

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

where  $\mathbf{x}$  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

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \mathbf{y}_i(\mathbf{x}) \cdot \log(\hat{\mathbf{y}}_i(\mathbf{x})) \quad (2)$$

where  $\mathbf{y}_i$  is the ground truth class vector,  $\hat{\mathbf{y}}_i$  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<sup>23</sup>. UNet requires very few labeled images and has reasonable training time. In the meantime, recent research demonstrated that using autoencoder and the latent space to conduct could improve classification performance<sup>24,25</sup>. 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

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

Where  $\mathbf{x}$  and  $\hat{\mathbf{x}}$  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 3.

Figure 3. 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

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, and one representative confusion matrix was shown in Figure 4 and more information on the CV training curves can be found in the supplementary materials. Apart from the confusion matrix, the balanced classification accuracy was reported. The balanced accuracy  $\bar{\sigma}$  is the arithmetic mean of class-specific accuracies and is calculated as

$$\bar{\sigma} = \frac{1}{C} \sum_{i=1}^C \sigma_i \quad (4)$$

where  $\sigma_i$  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<sup>26</sup>.

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 image 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<sup>27</sup>. 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<sup>28</sup>

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

### III. 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 feedbacks 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.

#### A. Classifying human cancer cells

This experiment aims to evaluate the classification performance on the SSC image for cells with different sizes and internal structures. The confusion matrices and corresponding t-SNE visualizations for all classification experiments are presented in Figure 4. 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 the side-scattering images.

Figure 4. Confusion matrices and t-SNE visualizations from cross-validation experiments on the human cancer cell dataset

#### B. 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<sup>29</sup>. Conventional methods require fluorescent antibodies



tagging that are 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 5 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 dimension, 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 5. Confusion matrices and t-SNE visualization from cross-validation experiment on the human white blood cell dataset

#### IV. CONCLUSION AND 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.

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 1  $\mu\text{m}$ , 2  $\mu\text{m}$ , and 2  $\mu\text{m}$  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 subpopulation 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 to improve the supervised machine-learning models and provide a better understanding and explanation of the morphology-based cell type characterization and classification studies.

#### **DATA AND CODE AVAILABILITY**

The human cancer cell and leukocyte datasets 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>

#### **SUPPLEMENTARY MATERIAL**

See the supplementary material for the processing pipeline of image reconstruction, learning curves of CNN models, and protocols for sample preparation.

#### **ACKNOWLEDGEMENTS**

The research reported in this publication was supported by the National Institutes of Health under award number 2R44DA045460-02. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was performed in part at the San Diego Nanotechnology Infrastructure (SDNI) of UCSD, a member of the National Nanotechnology Coordinated Infrastructure (NNCI), which is

supported by the National Science Foundation (Grant ECCS-1542148). Yu-Hwa Lo has an equity interest in NanoCollect Biomedical, Inc. as a co-founder, share holder, and a member of the company's Scientific Advisory Board. NanoCollect may potentially benefit from the results of this research.

### LIST OF FIGURE LEGENDS

Figure 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. 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  $\mu\text{m}$ .

Figure 3. 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 4. Confusion matrices and t-SNE visualizations from cross-validation experiments on the human cancer cell dataset

Figure 5. Confusion matrices and t-SNE visualization from cross-validation experiment on the human white blood cell dataset

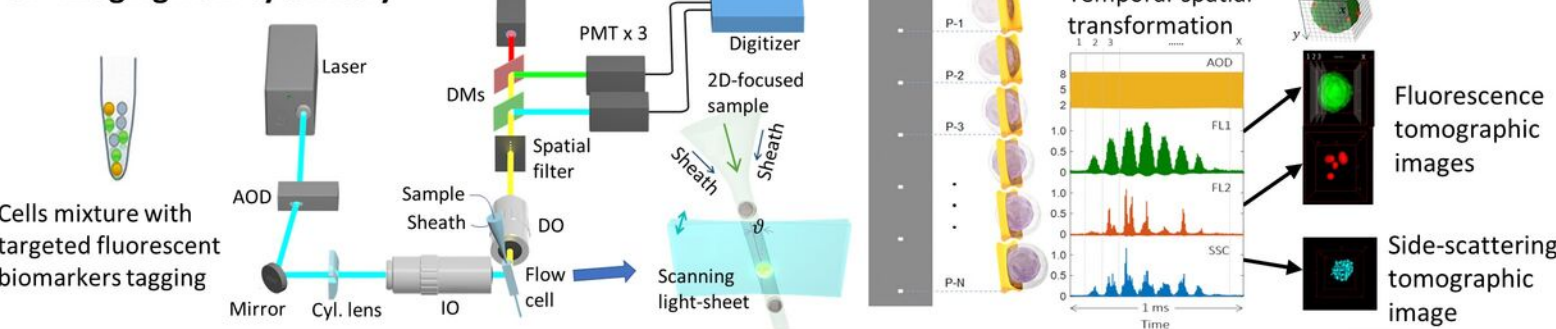
### REFERENCES

- [1] Altschuler, S. J. and Wu, L. F., "Cellular Heterogeneity: Do Differences Make a Difference?," *Cell* **141**(4), 559–563 (2010).
- [2] Snijder, B. and Pelkmans, L., "Origins of regulated cell-to-cell variability," *Nat. Rev. Mol. Cell Biol.* **12**(2), 119–125 (2011).

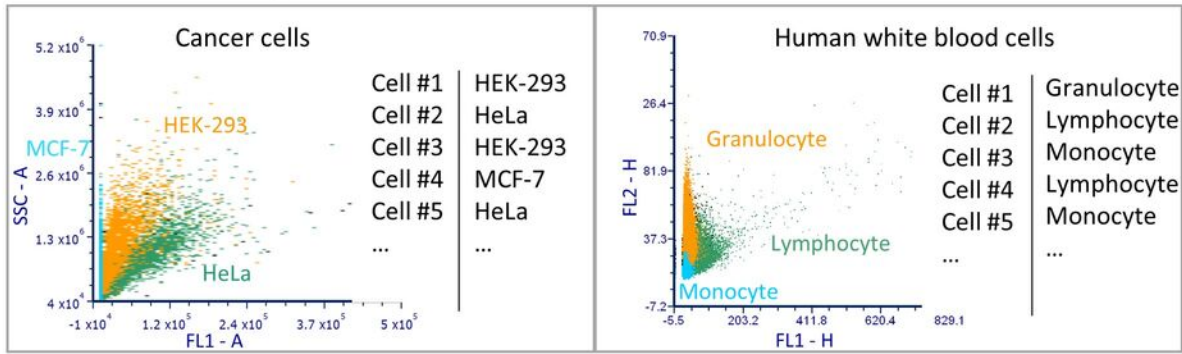
- [3] Pelkmans, L., “Using Cell-to-Cell Variability--A New Era in Molecular Biology,” *Science* (80-. ). **336**(6080), 425–426 (2012).
- [4] Han, Y., Gu, Y., Zhang, A. C. and Lo, Y.-H., “Review: imaging technologies for flow cytometry,” *Lab Chip* **16**(24), 4639–4647 (2016).
- [5] Boddington, S. E., Sutton, E. J., Henning, T. D., Nedopil, A. J., Sennino, B., Kim, A. and Daldrup-Link, H. E., “Labeling human mesenchymal stem cells with fluorescent contrast agents: The biological impact,” *Mol. Imaging Biol.* **13**(1), 3–9 (2011).
- [6] Doan, M., Vorobjev, I., Rees, P., Filby, A., Wolkenhauer, O., Goldfeld, A. E., Lieberman, J., Barteneva, N., Carpenter, A. E. and Hennig, H., “Diagnostic Potential of Imaging Flow Cytometry,” *Trends Biotechnol.* **36**(7), 649–652 (2018).
- [7] Li, Y., Mahjoubfar, A., Chen, C. L., Niazi, K. R., Pei, L. and Jalali, B., “Deep Cytometry: Deep learning with Real-time Inference in Cell Sorting and Flow Cytometry,” *Sci. Rep.* **9**(1), 11088 (2019).
- [8] Isozaki, A., Harmon, J., Zhou, Y., Li, S., Nakagawa, Y., Hayashi, M., Mikami, H., Lei, C. and Goda, K., “AI on a chip,” *Lab Chip* **20**(17) (2020).
- [9] Chen, C. L., Mahjoubfar, A., Tai, L. C., Blaby, I. K., Huang, A., Niazi, K. R. and Jalali, B., “Deep Learning in Label-free Cell Classification,” *Sci. Rep.* **6**(March), 1–16 (2016).
- [10] Wu, Y., Zhou, Y., Huang, C.-J., Kobayashi, H., Yan, S., Ozeki, Y., Wu, Y., Sun, C.-W., Yasumoto, A., Yatomi, Y., Lei, C. and Goda, K., “Intelligent frequency-shifted optofluidic time-stretch quantitative phase imaging,” *Opt. Express* **28**(1), 519 (2020).
- [11] Li, Y., Cornelis, B., Dusa, A., Vanmeerbeeck, G., Vercruysse, D., Sohn, E., Blaszkiewicz, K., Prodanov, K., Schelkens, P. and Lagae, L., “Accurate label-free 3-part leukocyte recognition with single cell lens-free imaging flow cytometry,” *Comput. Biol. Med.* **96**, 147–156 (2018).
- [12] Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. and Kording, K., “Voodoo Machine Learning for Clinical Predictions,” *bioRxiv*, 059774 (2016).
- [13] Nassar, M., Doan, M., Filby, A., Wolkenhauer, O., Fogg, D. K., Piasecka, J., Thornton, C. A., Carpenter, A. E., Summers, H. D., Rees, P. and Hennig, H., “Label-Free Identification of White Blood Cells Using Machine Learning,” *Cytom. Part A* **95**(8), 836–842 (2019).
- [14] Lippeveld, M., Knill, C., Ladlow, E., Fuller, A., Michaelis, L. J., Saeys, Y., Filby, A. and Peralta, D., “Classification of human white blood cells using machine learning for stain-free imaging flow cytometry,” *Cytom. Part A* **97**(3), 308–319 (2020).
- [15] Cram, L. S. and Brunsting, A., “Fluorescence and light-scattering measurements on hog cholear-infected PK-15 cells,” *Exp. Cell Res.* **78**(1), 209–213 (1973).
- [16] Salzman, G. C., Crowell, J. M., Trujillo, T. T., Romero, A., Mullaney, P. F. and Labauve, P. M., “Cell Classification by Laser Light Scattering: Identification and Separation of Unstained Leukocytes,” *Acta Cytol* **19**(4), 374–377 (1975).
- [17] Liu, C., Capjack, C. and Rozmus, W., “3-D simulation of light scattering from biological cells and cell differentiation,” *J. Biomed. Opt.* **10**(1), 014007 (2005).
- [18] Marina, O. C., Sanders, C. K. and Mourant, J. R., “Correlating light scattering with internal cellular structures,” *Biomed. Opt. Express* **3**(2), 296 (2012).
- [19] Han, Y., Tang, R., Gu, Y., Zhang, A. C., Cai, W., Castor, V., Cho, S. H., Alaynick, W. and Lo, Y.-H., “Cameraless high-throughput three-dimensional imaging flow

- cytometry,” *Optica* **6**(10), 1297 (2019).
- [20] Han, Y. and Lo, Y.-H., “Imaging Cells in Flow Cytometer Using Spatial-Temporal Transformation,” *Sci. Rep.* **5**, 13267 (2015).
- [21] Japkowicz, N. and Stephen, S., “The class imbalance problem: A systematic study,” *Intell. Data Anal.* **6**, 429–449 (2002).
- [22] He, K., Zhang, X., Ren, S. and Sun, J., “Deep Residual Learning for Image recognition,” *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, 770–778 (2016).
- [23] Ronneberger, O., Fischer, P. and Brox, T., “U-net: Convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **9351**, 234–241 (2015).
- [24] Yeh, C. K., Wu, W. C., Ko, W. J. and Wang, Y. C. F., “Learning deep latent spaces for multi-label classification,” *31st AAAI Conf. Artif. Intell. AAAI 2017*, 2838–2844 (2017).
- [25] Aytekin, C., Ni, X., Cricri, F. and Aksu, E., “Clustering and Unsupervised Anomaly Detection with 12 Normalized Deep Auto-Encoder Representations,” *Proc. Int. Jt. Conf. Neural Networks* **2018-July** (2018).
- [26] Brodersen, K. H., Ong, C. S., Stephan, K. E. and Buhmann, J. M., “The Balanced Accuracy and Its Posterior Distribution,” *2010 20th Int. Conf. Pattern Recognit.*, 3121–3124 (2010).
- [27] Kingma, D. P. and Ba, J. L., “Adam: A method for stochastic optimization,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, 1–15 (2015).
- [28] van der Maaten, L. J. J. and Hinton, G., “Visualizing Data using t-SNE,” *J. Mach. Learn. Res.* **1**, 1–48 (2008).
- [29] Walker, H. K., Hall, W. D. and Hurst, J. W., [Peripheral Blood Smear--Clinical Methods: The History, Physical, and Laboratory Examinations], Butterworths (1990).

### 3D imaging flow cytometry



### Ground truth labeling



Intensity-based features

|      |
|------|
| 0.11 |
| 0.95 |
| 0.34 |
| 0.62 |
| 1.54 |
| 9.98 |

15 March 2023 21:48:14

Obtaining ground truth cell class labels

### Deep learning

**Input**

Ground truth cell class label + side-scattering tomographic image

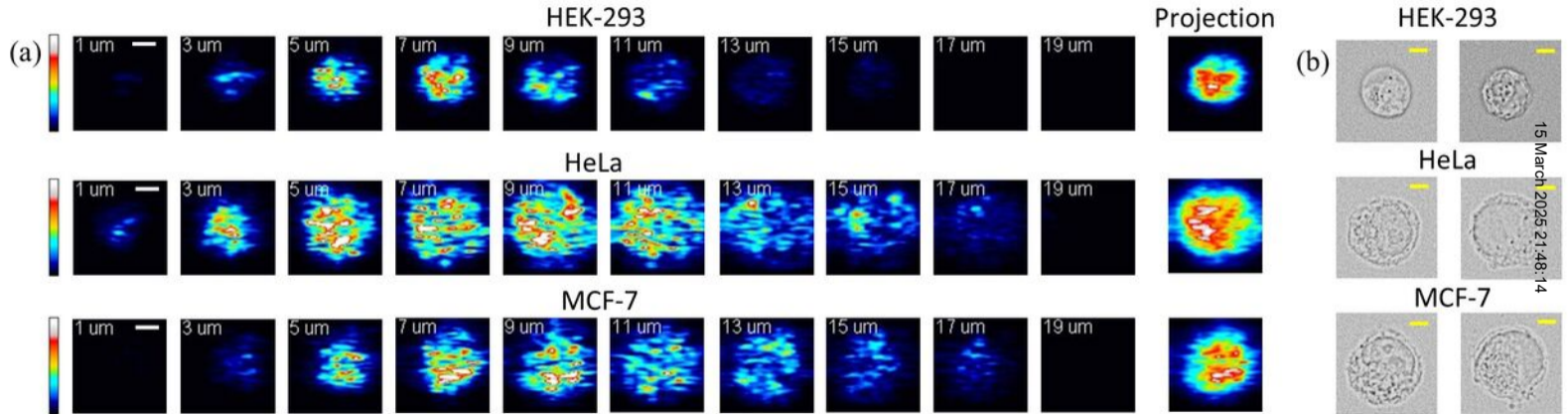
**Cell type classification**

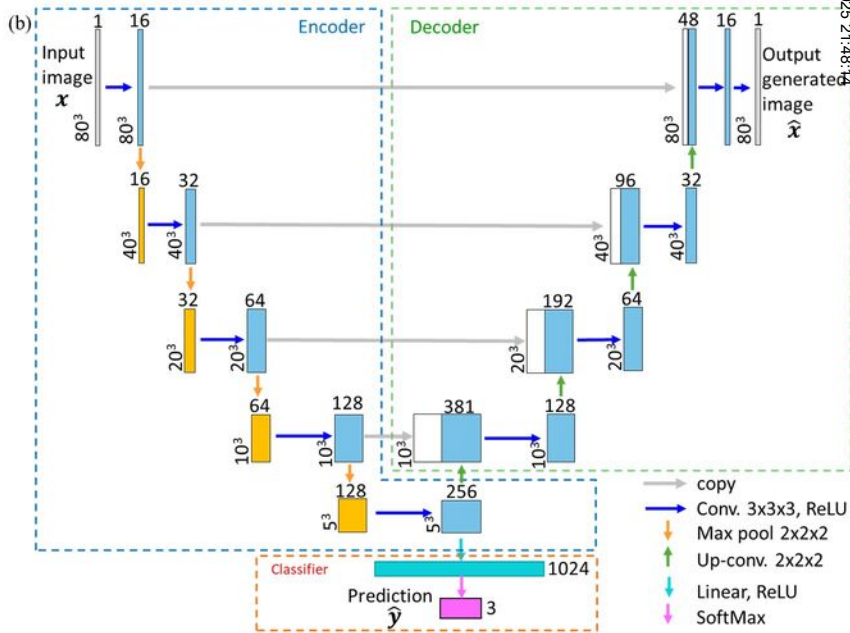
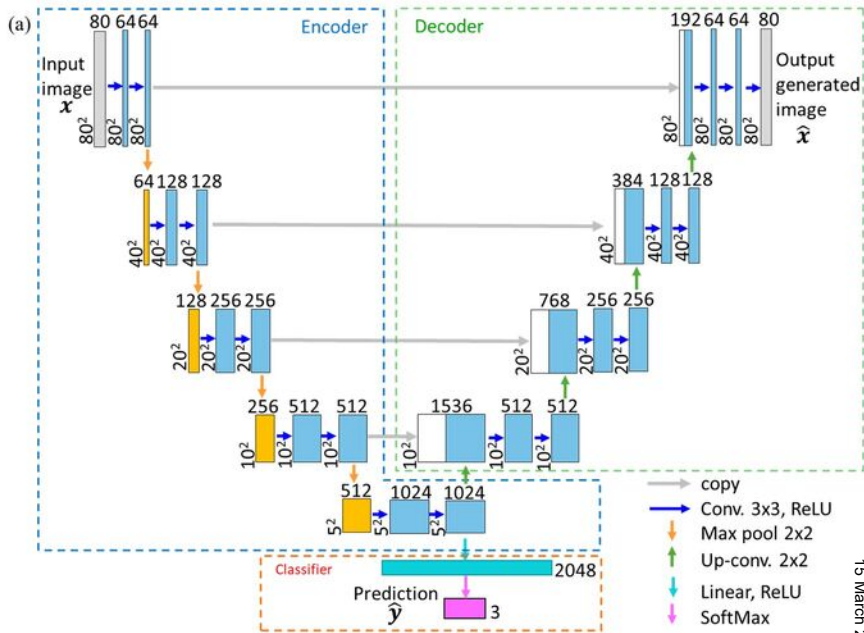
Convolutional neural networks

**Evaluation**

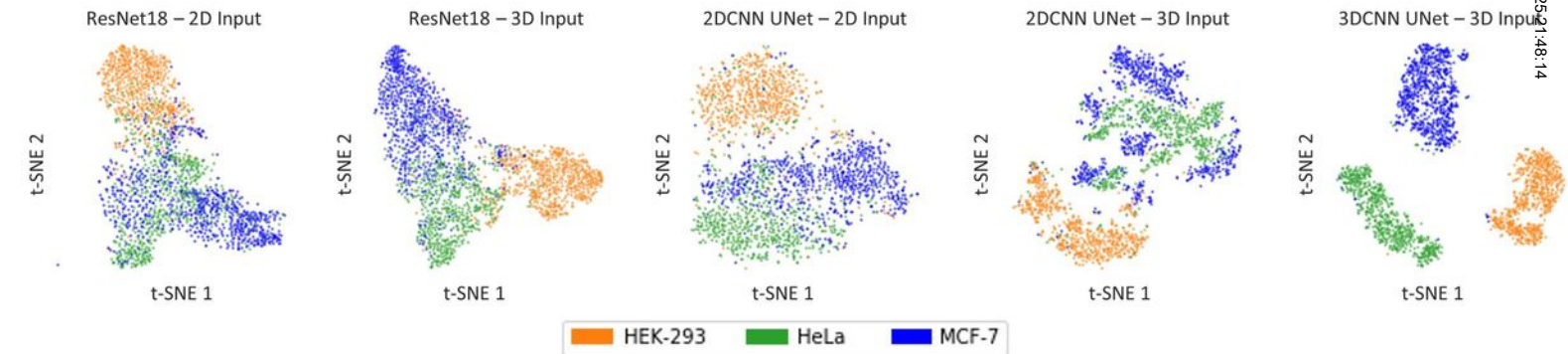
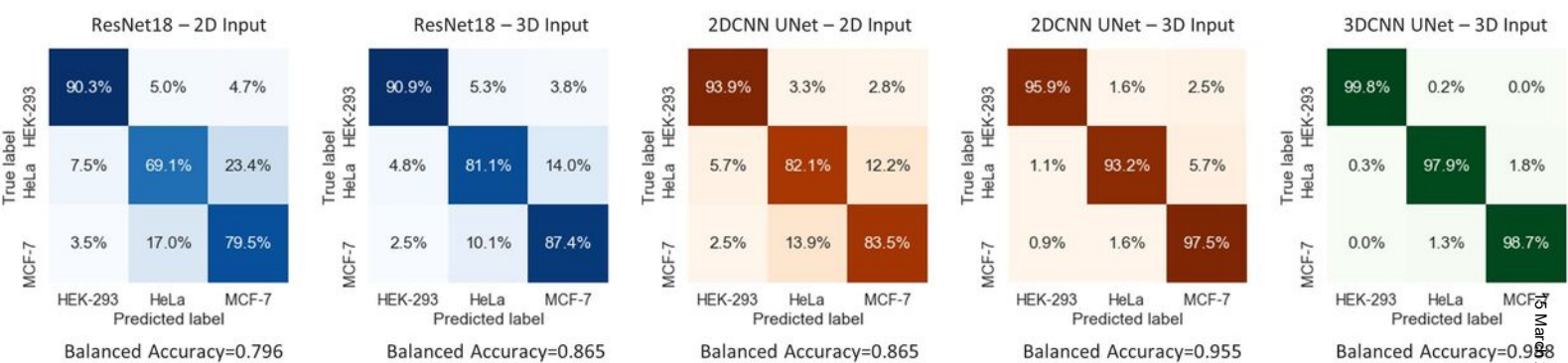
Confusion matrices + Balanced accuracy

|       |       |       |       |       |       |
|-------|-------|-------|-------|-------|-------|
| 94.9% | 1.1%  | 4.0%  | 97.2% | 0.8%  | 2.0%  |
| 3.0%  | 83.0% | 14.0% | 1.4%  | 87.5% | 11.0% |
| 3.6%  | 9.3%  | 87.1% | 2.1%  | 7.4%  | 90.5% |







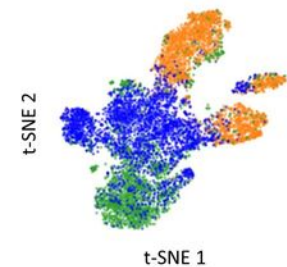


ResNet18 – 2D Input

| True label | Predicted label |         |       |
|------------|-----------------|---------|-------|
|            | Granulo.        | Lympho. | Mono. |
| Granulo.   | 93.0%           | 1.6%    | 5.4%  |
| Lympho.    | 5.7%            | 77.7%   | 16.6% |
| Mono.      | 4.8%            | 15.0%   | 80.2% |

Balanced Accuracy=0.836

ResNet18 – 2D Input

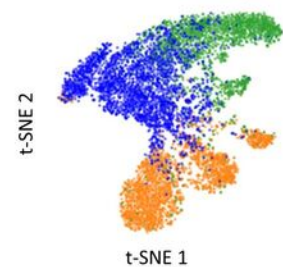


ResNet18 – 3D Input

| True label | Predicted label |         |       |
|------------|-----------------|---------|-------|
|            | Granulo.        | Lympho. | Mono. |
| Granulo.   | 94.9%           | 1.1%    | 4.0%  |
| Lympho.    | 3.0%            | 83.0%   | 14.0% |
| Mono.      | 3.6%            | 9.3%    | 87.1% |

Balanced Accuracy=0.883

ResNet18 – 3D Input

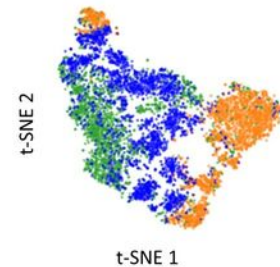


2DCNN UNet – 2D Input

| True label | Predicted label |         |       |
|------------|-----------------|---------|-------|
|            | Granulo.        | Lympho. | Mono. |
| Granulo.   | 94.4%           | 0.6%    | 5.0%  |
| Lympho.    | 2.8%            | 85.1%   | 12.1% |
| Mono.      | 3.2%            | 11.4%   | 85.4% |

Balanced Accuracy=0.883

2DCNN UNet – 2D Input

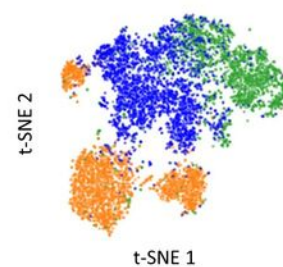


2DCNN UNet – 3D Input

| True label | Predicted label |         |       |
|------------|-----------------|---------|-------|
|            | Granulo.        | Lympho. | Mono. |
| Granulo.   | 97.2%           | 0.8%    | 2.0%  |
| Lympho.    | 1.4%            | 87.5%   | 11.0% |
| Mono.      | 2.1%            | 7.4%    | 90.5% |

Balanced Accuracy=0.918

2DCNN UNet – 3D Input

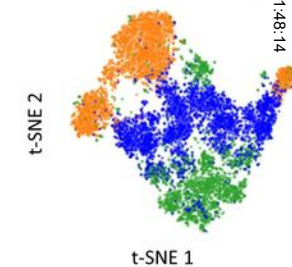


3DCNN UNet – 3D Input

| True label | Predicted label |         |       |
|------------|-----------------|---------|-------|
|            | Granulo.        | Lympho. | Mono. |
| Granulo.   | 97.8%           | 0.4%    | 1.8%  |
| Lympho.    | 1.8%            | 90.0%   | 8.2%  |
| Mono.      | 3.5%            | 7.2%    | 89.2% |

Balanced Accuracy=0.922

3DCNN UNet – 3D Input



Granulocyte Lymphocyte Monocyte