Label-free scattering snapshot classification for living cell identification

David Dannhauser[1,*], *Paolo Antonio* Netti[1,2], and *Filippo* Causa[1]

[1]Interdisciplinary Research Centre on Biomaterials (CRIB) and Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale, University of Naples "Federico II", 80125 Naples, Italy
[2]Center for Advanced Biomaterials for Healthcare@CRIB, Istituto Italiano di Tecnologia, 80125 Naples, Italy

Abstract. A scattering snapshot hold an enormous potential for cell class and state classification, allowing to avoid costly fluorescence labelling. Beside convolutional neural networks show outstanding image classification performance compared to other state-of-the-art methods, regarding accuracy and speed. Therefore, we combined the two techniques (Light Scattering and Deep Learning) to identify living cells with high precision. Neural Networks show high prediction performance for known classes but struggles when unknown classes need to be identified. In such a scenario no prior knowledge of the unknown cell class can be used for the model training, which inevitably results in a misclassification. To overcome the hurdle, of identifying unknown cell classes, we must first define an in-distribution of known snapshots to afterwards detect out of distribution snapshots as unknowns. Ones, such a new cell class is identified, we can retrain our cell classifier with the obtained knowledge, so we dynamically update the cell class database. We applied this measurement approach to scattering pattern snapshots of different classes of living cells. Our outcome shows a precise cell classification, which can be applied to a wide range of single cell classification approaches.

1 Introduction

The human body consists of various cell types, each with unique biophysical properties like size, structure, and function, useful to distinguish one cell from another.[1,2] The identification of cells like human peripheral blood mononuclear cells (PBMCs) is crucial for liquid biopsies. Knowledge of PBMC class appearance and cell states over time aids clinicians in their decisions. Nowadays costly cytometric analysis of surface receptor expressions are used for patient diagnosis, which require specialized operators and bulky instruments.[3]
Label-free cell identification methods, which work operator-free and most importantly preserve the native cell functions at low measurement costs are needed. Here, biophysical cell investigations show promise, especially for sparsely present cells, using indicators like relative cell count and anomalous cell shapes.
For instance, scattering patterns provide valuable biophysical cell properties for label-free identification. Forward scattering (<2°) reveals cell size and refractive index (inner cell structure) information, while side scattering (5°-30°) is more useful for nucleus and nucleus over cell ratio information. Small internal organelles contribute more at larger scattering angles (~90°), while backward scattering (160°-180°) can be useful for membrane roughness investigations.[4]
However, cells are heterogeneous, posing challenges in their label-free classification, which only high-performance classification methods can fulfil. Neural networks, commonly used for image identification show promising classification performance. In the case of label-free cell detection Neural Networks should detect unknown cells to avoid misclassifications. Though, Neural Networks are typically based on a closed-set assumption (training cell class are equal to testing classes), limiting their applicability in dynamically changing cell scenarios. For instance, thresholding classification scores for unknown cell classes is generally working impractical. Instead, an out-of-distribution recognition classifier should detect new cell classes by projecting known cell class inputs into compact regions of the feature space, which in the end identify cells.
In this work we show the potential of pure light scattering patterns for cell identification using a neural network for snapshot classification. A microfluidic device combined with optical single-cell investigation recorded data of different cell classes was used for data recording. Deep Learning used scattering information to predict cell types, achieving ~90% classification accuracy for single cell types and demonstrating the ability to predict out-of-distribution cells, expanding its application field.

2 MATERIALS AND METHODS

2.1 Sample

We recovered PBMC cells from healthy donors after obtaining informed consent.[5,6] Cells were diluted in

* Corresponding author: david.dannhauser@unina.it

viscoelastic alignment medium (~1x10^{5}cells/mL) and directly measured with the light scattering apparatus.

2.2 Setup

Scattering pattern snapshots were recorded with a small angle light scattering apparatus combined with a microfluidic device (see Fig. 1), composed of an alignment and observation section.[5,6] During travel through the observation channel, cell hit one after another the collimated incident beam, that generates a scattering pattern, which is recorded in a continuous angular range from 3°-33° with an angular resolution of ~0.1°.[5-8]

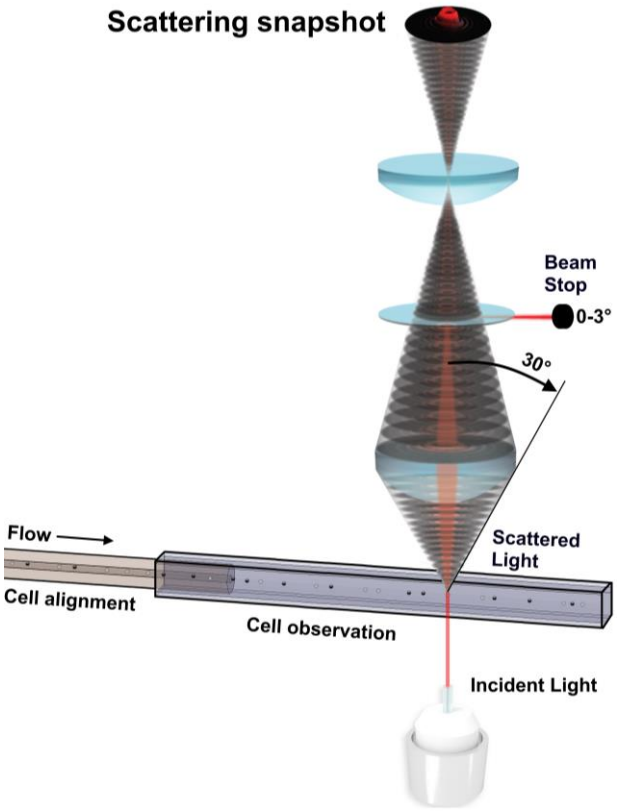

Fig. 1. Caption of the Figure 1. Below the figure.

2.3 Calculations

For Deep Learning calculations we centred and resized snapshots and passed them to a convolutional neural network (see Fig. 2).

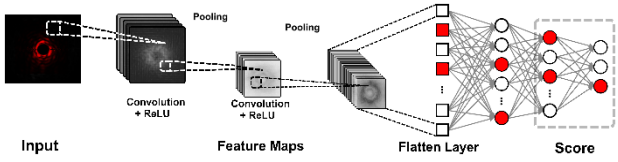

Fig. 2. Caption of the Figure 1. Below the figure.

3 Results and discussion

We optimized the network by tunning the hyperparameters, selecting the best number of epochs as a trade-off between computational efficiency and accuracy. We also defined a label space of known cell classes, while the remaining space was allocated as unknown class. Therefore, we trained the network to classify known cell classes. Then, we used the last layer before the SoftMax function as the encoded feature vector to train the novel open-set network. Next, we initialized the auxiliary domain of the network architecture by randomly generating samples in the encoded feature space and estimated the weights between the new and old encoder. The higher such an estimated weight, the more likely a scattering snapshot belongs to known classes. The main tuning parameter was β, that helps to define the auxiliary domain distribution and therefore tune the feature space, which results in a correct classification of unknown from known cell classes.

We trained the classifier model using more than 300 cells for each investigated PBMC class (4 classes). We added acute monocytic leukemia cells, which are used as proof-of-concept as unknown cell class, which were not seen by the classifier during the training phase. Results show that higher epoch numbers show a significant increase of the model performance, albeit a high epoch number does not automatically ensure a good unknown cell detection performance.

We optimized the snapshot classifier for PBMC, which resulted in an inter-cell as well as unknown from known cell class prediction of over 90% in real-life measurement scenario. Such outcomes pave the way for label-free tumor cell detection in liquid biopsy, where most state-of-the-art classifiers struggle.

In conclusion, scattering snapshot classification must perform robustly, which requires trading off between maximizing the recognition rate and minimizing the inclusion of novel data. Therefore, the goal is to minimize the risk of capturing unknown cells as known. In fact, the high amount of speckling information in scattering snapshots allowed label-free cell class identification, which can also be applied for cell state identifications.

References

1. A. Merino, L. Puigví, L. Boldú, S. Alférez, J. Rodellar, *Clin. Lab. Hematol*. **40** (2018)
2. N. Tatsumi, R.V. Pierre, *Clinics in laboratory medicine* **22** (2002)
3. H.T. Maecker, J.P. McCoy Jr, *Nat. Immunol.* **11** (2010)
4. D. Watson, N. Hagen, J. Diver, P. Marchand, M. Chachisvilis, *Biophys. J.* **87** (2004)
5. D. Dannhauser, D. Rossi, A.T. Palatucci, V. Rubino, F. Carriero, G. Ruggiero, M. Ripaldi, M. Toriello, G. Maisto, P.A. Netti, G. Terrazzano, F. Causa, *Lab Chip* **21** (2021)
6. D. Dannhauser, D. Rossi, P. Memmolo, A. Finizio, P. Ferraro, P.A. Netti, F. Causa, *Biomed. Opt. Express* **9** (2018)
7. D. Dannhauser, D. Rossi, V. De Gregorio, P.A. Netti, G. Terrazzano, F. Causa, *R. Soc. Open Sci.* **9** (2022)
8. D. Dannhauser, F. Causa, E. Battista, A. Cusano, D. Rossi, P.A. Netti, *Biomicrofluidics* **10** (2016)