Integrated Dielectric Scatterers for Speeding up Classification of Cell Diffraction Patterns

Alessio Lugnan, Joni Dambre*, and Peter Bienstman
Photonics Research Group[1], UGent – imec, Technologiepark 15, 9052 Gent, Belgium
Center for Nano- and Biophotonics (NB-Photonics), Ghent University,
Technologiepark 15, 9052 Ghent, Belgium
**IDLab, UGent - imec, Technologiepark 15, 9052 Gent, Belgium*
Tel: +32 9 264 3450, Fax: +32 9 264 3593, e-mail: alessio.lugnan@ugent.be

ABSTRACT
The computational power required to classify diffraction pattern images of biological cells is a major limit to the implementation of fast label-free cell sorting based on digital holographic microscopy. This work aims to demonstrate that the performance of a computationally cheap linear classifier can be significantly improved by the scattering action of a simple collection of integrated dielectric pillars. FDTD simulations show that the error of cell classification based on nucleus size can be decreased by a factor 5 when dielectric scatterers are employed and cells are made to flow in a suitable optical cavity.
Keywords: integrated optics, flow cytometry, cell classification, machine learning, extreme learning machine (ELM).

1. INTRODUCTION

The sorting of biological cells is of key importance in several biomedical applications, like diagnostics, therapeutics and cell biology. However, an accurate classification and separation of different cell types is usually expensive, time consuming and often requires alterations of the samples due to the use of labels, e.g. fluorescent tags, that may hinder subsequent analyses [1].

For these reasons, the development of label-free, high-speed, automated and integrated cell sorting solutions is of particular interest. Among several options, the employment of digital holographic microscopy in microfluidic flow cytometry is a promising candidate. In this technique, the classification is carried out through the analysis of the interference pattern (hologram) projected by the cell when illuminated by monochromatic light. The hologram is acquired by an image sensor and contains information on the 3D refractive index structure of the cells [2].The large amount of information contained in a cell hologram enables nontrivial analysis and classifications. On the other hand, the computational cost of elaborating such a complex source of information by reconstructing the image from the hologram is a major hindrance to an increase in the cell sorter throughput, e.g. by parallelization of the process.

An important reduction in the required computing power was achieved by bypassing the reconstruction of the cell image and directly processing the acquired hologram with a machine learning algorithm that carries out the classification task [2][3]. However, a further simplification of the processing in the electric domain is required in order to fully exploit the potential of this implementation.

2. DIELECTRIC SCATTERERS FOR IMPROVING CELL HOLOGRAMS CLASSIFICATION

We provided a proof of concept, based on finite-difference time-domain (FDTD) simulations, of an integrated photonics system for fast and label-free classification of biological cells [4]. In this application, a passive optical stage comprising a collection of pillar silica scatterers embedded in a silicon nitride cladding is used to process the light forward-scattered by a cell when illuminated via a green monochromatic source (Fig. 1).

The photonic stage containing the scatterers is intended to exploit the nonlinearity of the transfer function that relates the phase shift accumulated by the light through the cell to the corresponding interference pattern measured by an image sensor. Basically, the system can be seen as an hardware implementation of an extreme learning machine (ELM), i.e. a feedforward neural network whose hidden nodes are randomly generated and kept fixed [5] (see *Results and discussion* in [4] for a further argumentation). In the past decade, this machine learning paradigm has been the object of an increasing interest in various research field due to its remarkable efficiency, simplicity, and generalization performance [6]. Generally, the main advantages of ELMs with respect to other machine learning techniques are that only a linear readout (in this case a linear classifier) needs to be trained and that it is easily implemented in hardware. In this case, the pillar scatterer stage determines the ELM hidden node structure by projecting onto the far-field intensity a very intricate nonlinear mapping (based on sinusoidal functions of the phase information). This parallel processing is carried out nearly instantaneously with respect to both the cell movement and the operating speed of an electronic computer, providing an important

[1] This work was funded by: Research Foundation Flanders (FWO)(Grant No. G024715N)

advantage over other machine learning solutions in the electric domain. It should be stressed that the phase-to-intensity nonlinearity is already expressed by the interference pattern projected by the cell alone, without scatterers. However, the complexity of such a nonlinear mapping can be enhanced and controlled by the use of scatterers in order to increase the performance of a subsequent linear classification.

In order to properly train the readout classifier and to test its performance once trained, a sufficient number (thousands) of diffraction pattern samples had to be computed and provided to the training algorithm. Randomized cell models were employed to create reasonable variability in the diffraction pattern acquisition. The employed models are the result of a trade-off between computational cost and closeness to reality. Such a trade-off is legitimated by the fact that the goal of the work is not to provide absolute references for real applications but, instead, to investigate a relative difference between the classification performance with and without using scatterers.

In [4], two different classification tasks were considered. The first is based on average nucleus size and aims to distinguish between "normal" cells (small nucleus) and "cancer" cells (bigger nucleus). The names in quotation marks were chosen because of the common tendency of cancer cells to show evident irregularities in nucleus size [7]. The second task is based on nucleus shape and aims to distinguish between ``lymphocytes" (big quasi-spherical nucleus) and "neutrophils" (nucleus divided in 3 lobes). The names in quotation marks refer to two among the most common white blood cells that are present in human blood.

Even if significant improvements (~ 50%, Fig. 2a) due to the introduction of pillar scatterers were achieved in the classification based on nucleus size, a simple quantitative reasoning on how nucleus size changes affect the input of the readout classifier suggested that better results could be in principle obtained. In particular, the difference in phase shift accumulated by the light through the 2 types of nucleus corresponding to the 2 cell classes was too small and therefore was expressed in an almost linear way by the acquired intensity pattern, while nonlinearity is required by the ELM approach (see [4] under *Nonlinear phase sensitivity* in *Results and discussion* section). The employment of an UV laser source (λ = 337.1 nm), that implies an increase in the optical path through the given nuclei, was considered as a solution and a much better classification performance was obtained (Fig. 2b). However, when real applications are concerned, this option is highly impractical as UV lasers are usually quite expensive and would probably damage or even kill the illuminated cells.

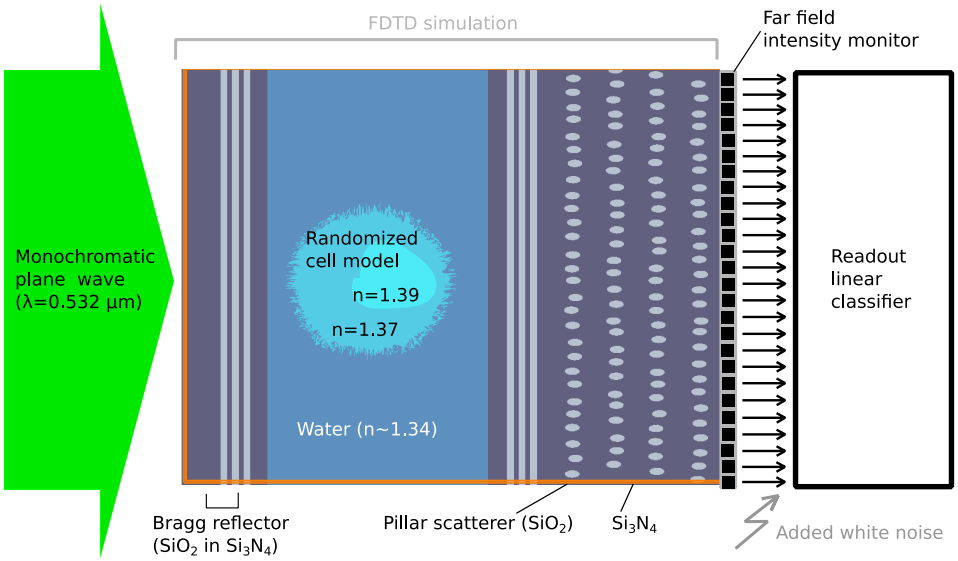

Figure 1. Schematic of the classification process. A monochromatic plane wave impinges on a Fabry-Pérot optical cavity composed by Bragg reflectors and containing a microfluidic channel with a cell in water (n_{H_2O} ~ 1.34), which has a low refractive index contrast ($n_{cytoplasm}$=1.37, $n_{nucleus}$=1.39); the forward scattered light passes through a collection of silica scatterers (n_{SiO_2} ~ 1.461) embedded in silicon nitride ($n_{Si_3N_4}$ ~ 2.027) and organized in layers; the radiation intensity is then collected by a far-field monitor, which is divided into bins (pixels); each pixel value is fed into a trained linear classifier (logistic regression) that consists of weighted sums (one per class) of the pixel values. The weights are trained so that the sum exceeds a certain threshold value only if the corresponding input class is recognized.

3. INTEGRATED OPTICAL CAVITY FOR ENHANCED PERFORMANCES

A more feasible solution with respect to the employment of UV light consists of increasing the effective optical path length through the cell by inserting it in an optical cavity. Intuitively, this makes the impinging light pass, on average, more than once through the cell. In practice, in the FDTD simulation design 2 Bragg reflectors were placed at the 2 external sides of the microfluidic channel, orthogonally to the light beam direction, creating

a Fabry-Pérot cavity (Fig. 1). The employed Bragg reflectors were each composed of 3 layers of SiO_2 with a width of (455 ± 10) nm in a Si_3N_4 cladding. The error in the layer width was implemented by adding a random value sampled from an uniform probability distribution between -10 and 10 nm. It approximately accounts for fabrication errors. The distance D = 21.02 μm between the reflectors was chosen so that the portion of light passing through and near the nucleus of the cell was resonant. This was done by monitoring the light intensity inside the cavity for different values of D. Note that such a tuning was relatively easy to perform because the cell acts as a weak converging lens, providing an additional light confinement along the microfluidic channel direction.

The reflectivity R of the reflectors is also a crucial parameter, since it controls the cavity Q-factor and therefore controls both the sensitivity of the resonance to intracavity optical path lengths and how long the light stays, on average, inside the cavity. In this case the sensitivity of the acquired intensity pattern to the nucleus size needed to be improved by increasing the average time that the resonant light passing through the nucleus stays in the cavity. On the other hand, if the cavity Q-factor was too high, the resonance strength might be strongly influenced by uninteresting small details of the cell structure or by fabrication errors. Generally, the cavity should be designed so that the average light phase shift differences due to the optical feature of interest corresponding to the considered classes are roughly between π/2 and 2π. Indeed, a phase shift difference significantly lower than π/2 would induce a nearly linear difference on the acquired intensity pattern, undermining the nonlinearity requirement of the ELM approach. On the other hand, a phase shift difference significantly bigger than 2π would make the acquired intensity pattern too sensitive to small changes in the cell optical features, undermining the stability requirement of the ELM approach. In particular, the reflectors employed in the simulations (composed of 3 layers) have a satisfying reflectivity of ~ 56%, while it turned out that similar reflectors with 4 and 5 layers have a too high reflectivity, respectively of ~ 73% and ~ 85%. All the simulation and machine learning aspects concerning the investigation presented in this paper are as in [4], apart from the introduction of the described optical cavity, that also required longer FDTD simulation times (1.2 ps).

Figure 2 compares the classification errors obtained considering respectively green (λ = 532 nm) and UV (λ = 337.1 nm) laser light [4] and green laser light employing the described optical cavity. A substantial improvement is observed when the cavity is used with respect to the green laser counterpart (Fig. 2c). In particular, the classification improvement due to the use of scatterers is increased by factor 5 by the cavity, for sufficiently low but still plausible noise levels (< 10%). At these noise levels, the results are similar to what was obtained with an UV light source, without the drawback of possible cell damage. For higher noise levels an increased sensitivity to noise pushes the classification error rate to significantly higher values. Finally, it should be stressed that an additional advantage arising from the use of an optical cavity is that it can be designed to increase the intensity pattern sensitivity towards specific optical path lengths, making the optical features of interest more evident to the readout classifier with respect to other competing ones.

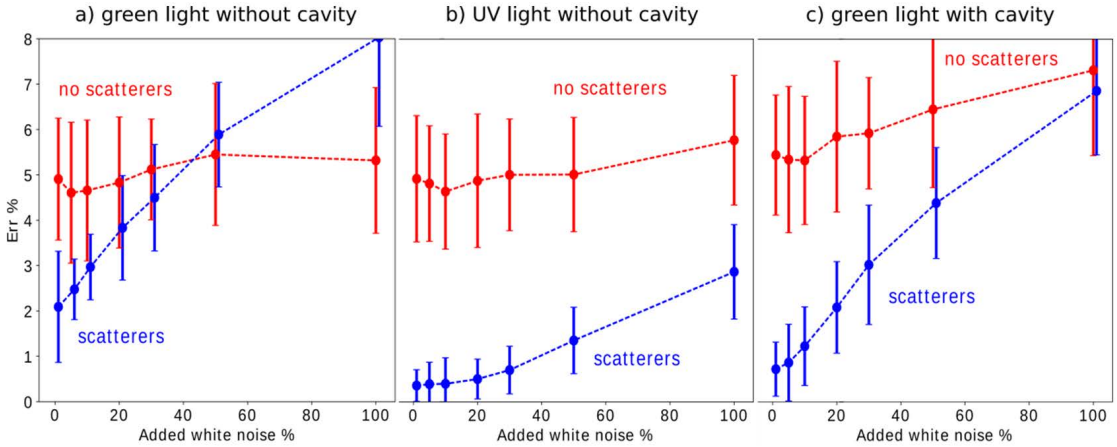

*Figure 2. Error rates of cell classification based on nucleus size as a function of the white noise percentage added to the acquired intensity patterns, corresponding to the absence (in red) and the presence (in blue) of scatterers. The displayed error rate values are averages on the values obtained using 250, 260,270,280,290, 300 number of pixels in the simulation far field monitor. The graphed error bars represent a confidence interval of 2 standard deviations and corresponds to the error obtained through validation, by randomly shuffling the simulated intensity patterns 20 times between training and test sets. The other simulation and machine learning details are as described in [4]. **a)** A green laser source (λ = 532 nm) was employed, without optical cavity. **b)** An UV laser source (λ = 337.1 nm) was employed, without optical cavity. **c)** A green laser source (λ = 532 nm) and Fabry-Pérot cavity were employed, as in Fig. 1.*

4. CONCLUSIONS

In this work, a simple passive integrated photonic structure is proposed to improve the performance of a linear classifier employed for the classification of biological cells based on nucleus size. The presented approach enables for significant reduction of the computational cost required for both training and application of the classifier, with respect to an equivalent application in the electric domain.

The proposed integrated structure consists of a collection of silica pillars embedded in silicon nitride that scatters the monochromatic light diffracted by a cell. When a green laser source is employed, the scatterers presence leads to a significant improvement in performance (up to 50%) of the classification of simulated cells with different nucleus size, provided that a sufficient number of pixels and a low enough noise level are considered. When an UV laser source (λ = 337.1 nm) is employed, the scatterers presence leads to a much greater error rate reduction (up to an order of magnitude) for all the considered pixel numbers and noise levels.

The much more feasible implementation with green light, though, is improved to a performance (a factor ~ 5 improvement) similar to the one obtained with UV light, if an integrated Fabry-Pérot optical cavity is properly designed and included in the simulation.

ACKNOWLEDGEMENTS

We thank Bendix Schneider for sharing his technical and theoretical insight regarding the topics treated in this work. We thank Geert Vanmeerbeeck and Yuqian Li from Imec for providing valuable updates on the challenges of cell classification via digital holographic microscopy.

REFERENCES

[1] D. R. Gossett, W. M. Weaver, A. J. Mach, S. C. Hur, H. T. K. Tse, W. Lee, H. Amini, and D. D. Carlo: Label-free cell separation and sorting in microfluidic systems, *Anal. Bioanal. Chem.*,vol. 397, no. 8, pp. 3249-3267, 2010.

[2] L. Lagae, D. Vercruysse, A. Dusa, C. Liu, K. de Wijs, R. Stahl, G. Vanmeerbeeck, B. Majeed, Y. Li, and P. Peumans: High throughput cell sorter based on lensfree imaging of cells, in *Proc. IEEE International Electron Devices Meeting*, pp. 333-336, 2015.

[3] B. Schneider, G. Vanmeerbeeck, R. Stahl, L. Lagae, J. Dambre, and P. Bienstman: Neural network for blood cell classification in a holographic microscopy system, in *Proc. ICTON*, Budapest, 2015.

[4] A. Lugnan, J. Dambre, and P. Bienstman: Integrated pillar scatterers for speeding up classification of cell holograms, *Opt. Express*, vol. 25, no. 24, pp. 30526-30538, 2017.

[5] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew: Extreme learning machine: Theory and applications, *Neurocomputing*, vol. 70, no. 1, pp. 489-501, 2006.

[6] G. Huang, G.-B. Huang, S. Song, and K. You: Trends in extreme learning machines: A review, *Neural Networks*, vol.61, pp. 32-48, 2015.

[7] D. Zink, A. H. Fischer, and J. A Nickerson: Nuclear structure in cancer cells, *Nat. Rev. Cancer*, vol. 4, no. 9, pp. 677-687, 2004.