This is the pre-peer reviewed version of the article "Speeding up reconstruction of 3D tomograms in holographic flow cytometry via deep learning" which has been published in the final form in ***Lab on a Chip, DOI: 10.1039/D1LC01087E****.*

Speeding up Reconstruction of 3D Tomograms in Holographic Flow Cytometry via Deep Learning

Daniele Pirone,[a,b] Daniele Sirico,[a] Lisa Miccio,[a] Vittorio Bianco,[a] Martina Mugnano, [a] Pietro Ferraro [†,a] and Pasquale Memmolo[††,a]

[a.] *CNR-ISASI, Institute of Applied Sciences and Intelligent Systems "E. Caianiello", Via Campi Flegrei 34, 80078 Pozzuoli, Napoli, Italy.*

[b.] *DIETI, Department of Electrical Engineering and Information Technologies, University of Naples "Federico II", via Claudio 21, 80125 Napoli, Italy.*

† pietro.ferraro@isasi.cnr.it

†† pasquale.memmolo@isasi.cnr.it

Abstract

Tomographic flow cytometry by Digital Holography is an emerging imaging modality capable of collecting multiple views of moving and rotating cells with the aim of recovering their refractive index distribution in 3D. Although this modality allows to access to high-resolution imaging with high-throughput, the huge amount of time-lapse holographic images to be processed (hundreds of digital holograms per cell) constitutes the actual bottleneck. This prevents the system to be suitable for lab-on-a-chip platforms in real-world applications, where fast analysis of measured data is mandatory. Here we demonstrate a significant speeding-up reconstruction of phase-contrast tomograms by introducing in the processing pipeline a multi-scale fully-convolutional context aggregation network. Although it was originally developed in the context of semantic image analysis, we demonstrate for the first time it can be successufly adapted to holographic lab-on-chip platform for achieving 3D tomograms through a faster computational process. We trained the network with input-output image pairs to reproduce the end-to-end holographic reconstruction process, i.e. recovering quantitative phase maps (QPMs) of single cells from their digital holograms. Then, the sequence of QPMs of the same rotating cell is used to perform the tomographic reconstruction. The proposed approach significantly reduces the computation time for retrieving tomograms, thus making them available in few seconds instead of tens of minutes, while essentially preserving the high-content information of tomographic data. Moreover, we have accomplished a compact deep convolutional neural network parameterization that can fit into on-chip SRAM and a small memory footprint, thus demonstrating its possible exploitation to provide onboard computations for lab-on-chip devices with low processing hardware resources.

This is the pre-peer reviewed version of the article "Speeding up reconstruction of 3D tomograms in holographic flow cytometry via deep learning" which has been published in the final form in ***Lab on a Chip, DOI: 10.1039/D1LC01087E****.*

Introduction

Digital Holography (DH) in microscopy is a label-free computational imaging technique able to provide a posteriori multiple refocusing capability and quantitative phase-contrast imaging[1-3]. Thanks to these features, DH has been successfully employed in a variety of biomedicine applications[4], including cancer cells identification and characterization[5-7], diagnostics of blood diseases[8-11], inflammations[12] and infectious diseases[13-15], study of cells motility and migration[16], and marker-free detection of lipid droplets[17]. The possibility to probe biological samples from different directions opens to the full 3D label-free imaging achieved by holographic tomography technology[18,19], which represents the leading edge of biological inspection at single-cell level. The combination of compact holographic microscopes and flow cytometry allows the high-throughput screening of cells flowing in microfluidic channels, thus permitting to study biological specimens in their natural environment for point-of-care diagnostics at lab-on-a-chip scale[20-22]. Recently, the possibility to perform the 3D tomographic reconstruction of rotating cells in flow cytometry conditions has been investigated[23-31] for a variety of applications, such as blood anemia detection[23], cancer cells identification[25], label-free liquid biopsy[29], intracellular nanoparticles inspection[30], and plant cells biology[31]. However, when a very large number of digital holograms have to be recorded by a flow cytometry assay, the numerical holographic reconstruction process becomes the bottleneck that prevents reliable and exploitable applications at lab-on-a-chip scale due to the demanding algorithms and the huge computational time. In fact, it is well known that, in DH microscopy, the complete processing pipeline to retrieve the Quantitative Phase Map (QPM) reconstruction from a digital hologram consists of several steps[2,3], namely suppression of zero order and twin-image, numerical refocusing, aberration correction and, finally, phase unwrapping. Actually, in off-axis DH, the suppression of undesired diffraction orders is essentially made by real time methods based on Fourier spectrum filtering[32], while iterative phase retrieval algorithms need to be employed for in-line recording schemes[33]. The numerical refocusing is usually solved by first reconstructing a stack of images, obtained by varying the reconstruction distance, then a suitable image sharpness metric is computed on each reconstruction, which maximum or minimum value corresponds to the in-focus distance depending on the nature of the object. To date, a large variety of metrics has been proposed to perform the in-focus reconstruction automatically[16,34]. Of course the channel height might influence the computational burden of this step, since it sets the volume to be digitally scanned to look for the sample best focus plane. The phase aberration compensation step can be achieved with fitting based processing[35] or by acquiring a reference hologram (i.e. without the sample in the imaged field of view (FOV)) to be subtracted[36] to the aberrated phase image. Finally, the phase unwrapping is performed to correct phase jumps caused by the imaged objects that introduce an optical path difference larger than 2π. Usually, accurate and robust methods are based on global image unwrapping algorithms[37] that can be very time consuming, especially depending on the level of speckle noise on the recorded digital hologram. Therefore, to speed up the phase unwrapping, it is advisable to use a preliminary denoising algorithm[38]. Definitely,

depending on the hologram size (i.e., the number of pixels of the sensor camera) and the imaging configuration (i.e. in-line or off-axis) the entire processing pipeline to reconstruct one single QPM from the recorded digital hologram can take minutes on a basic desktop computer.

To cope with these computational limitations, recently, deep convolutional neural networks (DCNNs) have been employed to speed up the holographic processing pipeline[39,40]. In particular, the numerical refocusing step is addressed as a DCNN-based in-focus distance regression problem[41,42], or through DCNNs for classification[43]. The phase aberration compensation has been solved by using a simplified version of the U-Net model[44] for the background detection and subtraction[45]. Instead, DCNNs inspired by the ResNet model[46] have been employed for both the automatic phase aberration compensation[47] and phase unwrapping process[48]. Recently, by suitably adapting the encoder-decoder models, it has been demonstrated that the entire holographic reconstruction process can be skipped, thus enabling the direct reconstruction from raw holograms without any prior knowledge about the imaging parameters[49-53]. Despite the remarkable results achieved in these works in terms of the accuracy about the recovering QPMs from digital holograms, such DCNN architectures usually employ tens of millions learnable parameters and need a remarkable amount of memory to store them. Here we propose for the first time in DH, the multi-scale context aggregation network (CAN)[54], that was originally developed in the context of semantic image analysis[55]. The network operates on input-output image pairs to approximate an operator's action, hence we train such fully-convolutional DCNN to reproduce the end-to-end holographic reconstruction process. CANs applied to DH has shown to guarantee accurateness in approximating image processing operators thus achieving interactive rates on high-resolution images and a constant runtime. This latter feature points out that such networks are potentially deployable within the constraint of compact devices. In our work we have investigated the trade-off between input image sizes, network deepness and runtime to achieve a compact DCNN parameterization that can fit into on-chip SRAM and a small memory footprint[56], thus making possible to provide onboard computations for lab-on-chip devices with low processing hardware resources. We demonstrate that the proposed approach allows a single QPM reconstruction in 168 ms at constant runtime, thus making tomographic reconstructions in few seconds instead of tens of minutes, while preserving the high-resolution content of tomographic data, i.e. over 98.5% accuracy with respect to the corresponding ground-truth.

Materials and Methods

Holographic recording and numerical processing

In order to perform the holographic recording of flowing cells, we used a DH microscope in off-axis configuration employing a Mach-Zehnder interferometric configuration, as sketched in Fig. 1A. In particular, we coupled a 532 nm laser source (Laser Quantum Torus 532) with a 750 mW output power. A polarizing beam splitter (PBS) separates the laser beam in an object and a reference wave. The reference beam is

transmitted, while the object beam is reflected. In addition, to balance the ratio between the intensity of the object and reference beam without changing their polarization, two half-wave plates (HWPs) are placed in front of and behind the PBS. The object beam illuminates the cells while flowing within a microfluidic channel (Microfluidic ChipShop 10000107), and the scattered radiation is collected by a microscope objective (MO_1) (Zeiss Plan-Apochromat 40× - NA=1.3 - Oil immersion) and sent to a tube lens (TL_1). The reference beam passes through a beam expander shaped by a microscope objective and a second tube lens (MO_2 and TL_2). Then, both beams are recombined by a beam splitter cube (BS) with a non-zero angle between them because of the off-axis configuration, and the resulting interference fringe pattern is recorded by a CMOS camera (Genie Nano-CXP Cameras - 5120×5120 pixels, 4.5 µm pixel) at 30 fps. An example of recorded digital hologram is shown in Fig. 1B. Within the microfluidic channel, a laminar flow is generated by an automatic syringe pump (CITONI Syringe Pump neMESYS 290N), which is a low-pressure system that allows a high-precision and pulsation-free dosing of liquids in a nanolitre scale, thus ensuring a very homogeneous flow. Cells not flowing at the centre of the microchannel undergo a velocity gradient due to the parabolic velocity profile. Therefore, while flowing, cells experience rotation thanks to the hydrodynamic forces of the laminar stream[25].

Each recorded hologram is pre-processed with the aim to detect and track each flowing cell within the FOV. This is made by a semi-automatic algorithm that performs cells tagging by a threshold-based detection method when a cell enters in the FOV. Then a sliding patch of sizes 384×384 pixels, centered on the cell, shifts along the flow direction with a velocity provided by the nominal flow rate (see red box in Fig. 1B). In summary, the video sequence of 5120×5120 pixels holographic images turns into several sequences of 384×384 pixels sub-holograms, each of them containing the cell during its rotation. Finally, each sub-hologram is numerically reconstructed by implementing the processing pipeline summarized in Fig. 1C. In particular, the suppression of undesired diffraction orders is made by Fourier spectrum filtering to select and centre the real diffraction order (i.e., hologram demodulation in Fig. 1C). Afterwards, the numerical refocusing is obtained by searching the in-focus distance through the minimization of the Tamura Coefficient (TC)[16]. The propagation of the demodulated hologram at such distance provides the in-focus complex amplitude from which the phase-contrast image is obtained by calculating its argument. The residual optical aberrations, superposed to the retrieved phase-contrast image, are subtracted by using a reference hologram, acquired without the sample in the imaged FOV[35]. Finally, we employ the two-dimensional windowed Fourier transform filtering as the denoising method[57] and the PUMA algorithm for the unwrapping[58]. The entire processing pipeline takes about 8 seconds to run for one single position of a single cell. The processing is then repeated for all the different positions occupied by the cell while it flows along the FOV. After having computed the QPMs of the same cell, the corresponding rolling angles are estimated from its transversal positions[27]. Finally, the tomographic reconstruction is performed by using the Filtered Back Projection algorithm[23].

*This is the pre-peer reviewed version of the article "Speeding up reconstruction of 3D tomograms in holographic flow cytometry via deep learning" which has been published in the final form in **Lab on a Chip, DOI: 10.1039/D1LC01087E**.*

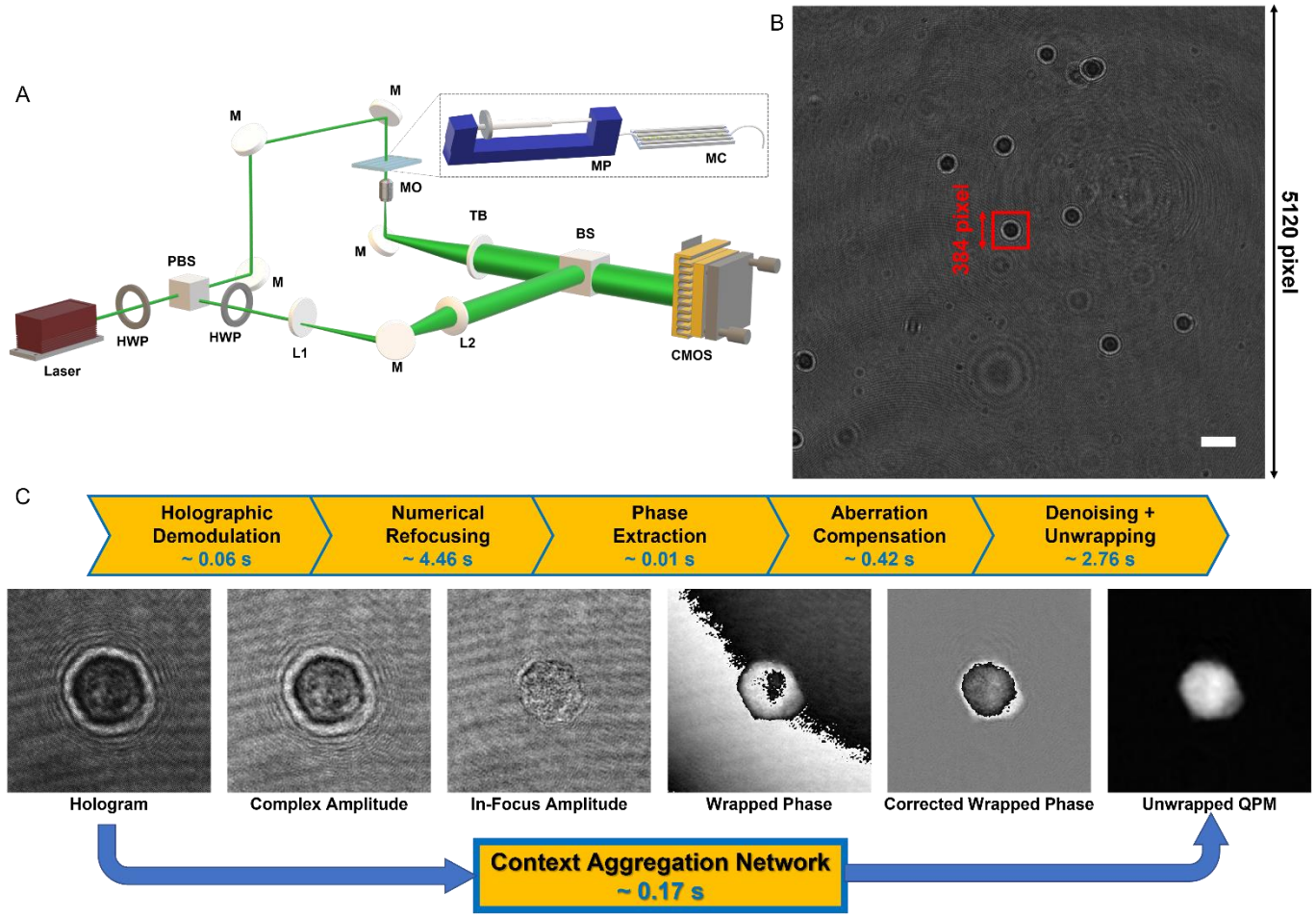

Figure 1. Holographic recording and QPM retrieval. A DH microscope in off-axis configuration. HWP – Half-Wave Plate; PBS – Polarizing Beam Splitter; L1, L2 – Lens; M – Mirror; MO – Microscope Objective; MP – Microfluidic Pump; MC – Microfluidic Channel; TL – Tube Lens; BS – Beam Splitter; CMOS – Camera. **B** Digital hologram recorded by the DH setup in A. The scale bar is 50 µm. **C** Holographic processing pipeline to compute the QPM of a cell from the corresponding holographic ROI selected in B (red box). The standard operations are reported at the top with the corresponding computational times.

Sample preparation

Mouse embryonic fibroblasts NIH 3T3 were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 4.5 g/L D-glucose and supplemented with 10% fetal bovine serum (FBS) (Life Technologies, Carlsbad, CA, USA), 100 units/mL penicillin and 100 µg/mL streptomycin (Sigma, St. Louis, MO). Then they were harvested from the tissue culture flasks by incubation with a 0.05% trypsin–EDTA solution (Sigma, St. Louis, MO) for 5 min. After centrifugation and resuspension in Phosphate buffered saline (PBS), cells were injected into the microfluidic channel at final concentration of 2×10^5 cells/mL. In order to assure the right condition for the cell culture medium during the manipulation outside of a CO_2 incubator, the addition of 20 mM HEPES (Sigma-Aldrich) is made to provide extra buffering capacity.

CAN architecture

The CAN architecture sketched in Fig. 2 has been trained to reconstruct the QPM from the recorded digital hologram[54]. The CAN is a fully convolutional network as the resolution of the input (i.e., $M \times M$) is not changed throughout the layers of the network up to the output layer. The network has a depth $d = 8$ and all the convolutional layers CL_S, with $1 \leq S \leq d$, have a width w. In particular, the layers CL_S, with $1 \leq S \leq d-1$, are based on w 3×3 kernels and are followed by a Leaky ReLU nonlinearity[59], while the last layer CL_d employs w 1×1 kernels with no nonlinearity. Let x_S be the output of a convolutional layer CL_S, with $1 \leq S \leq d-1$. The input y_S of the successive Leaky ReLU layer is computed through an adaptive normalization process as follows

$$y_S = \lambda_S x_S + \mu_S BN(x_S) \tag{1}$$

where BN is the batch normalization operator[60] and $\lambda_S, \mu_S \in \mathcal{R}$ are learnable scalar weights. Thanks to the adaptive normalization step, the batch normalization layer is strengthened and the overall model can better approximate the phase retrieval operator. More important, the fully-resolution intermediate layers allow aggregating the global context of the input image through a multi-scale analysis due to the several dilation factors r_S of the convolutional layers CL_S. In particular, the dilation factor r_S increases as 2^{S-1} for $= 1, \dots, d-2$, while the layer CL_{d-1} has no dilation. Hence, the receptive field expands exponentially with the network's depth, thus including the multi-scale global context despite the compactness of the CAN architecture. In fact, the number of learnable parameters of the model with the described configuration is very low with respect to the tens of millions of parameters of the classical encoder-decoder networks like the U-Net.

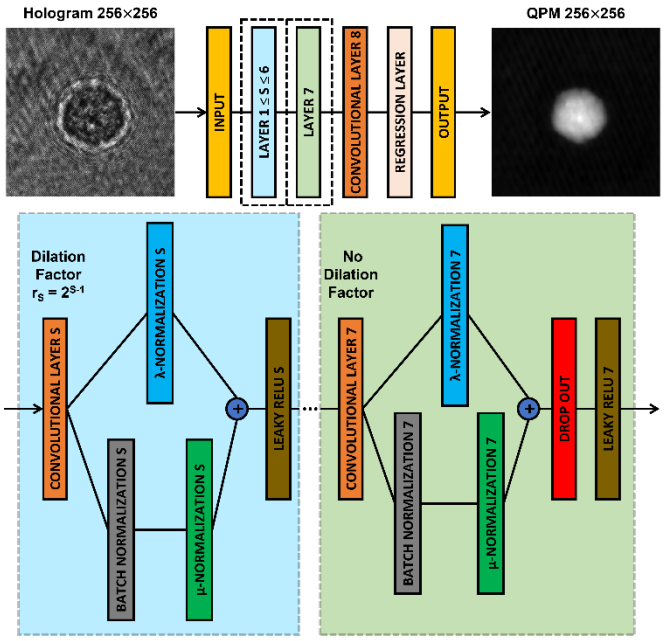

Figure 2. CAN model. Sketch of the CAN architecture for the end-to-end prediction of the QPM from the recorded digital hologram. The Layers 1-7 are zoomed in the dotted boxes.

Furthermore, the network requests a small memory during the forward step because there are no skip connections across non-consecutive layers[61]. For these reasons, the CAN model is expected to be accurate (due to aggregation of the multi-scale global context), fast (due to the compactness of the architecture), and particularly suited for on board computing (due to the small memory requested)[54]. Of course, these three properties depend on the setting of the width w and the input size M. In Table 1, the comparison among different network configurations have been reported to identify a suitable trade-off. In fact, the memory occupation increases with the width w, while the prediction time increases with both the width w and the input size M. Obviously, the smallest and fastest configuration is $w = 32$ and $M = 64$, which allows reaching a video-rate QPM prediction with a 214 kB memory occupation. However, as shown in Fig. 1B, the whole cell information is contained in a 384×384 region of interest (ROI), which means that a down sampling up to 64×64 leads to an excessive loss of resolution. On the other hand, the most accurate configuration is $w = 128$ and $M = 512$, which however requires megabytes for the memory occupation and seconds for the QPM prediction. Therefore, an intermediate configuration is more suitable in respect to our DH recording system. We choose $w = 64$ and $M = 256$, since it allows computing a QPM in 168 ms by occupying only 818 kB of memory (due to just 223183 parameters) with a negligible resolution loss with respect to the original ROI size.

Table 1. Comparison among different CAN configurations. Prediction times (in milliseconds) obtained by varying the width w (and then the memory occupation) and the input size M. The selected configuration is highlighted in green.

	w = 32 214 kB	**w = 64 818 kB**	**w = 128 3231 kB**
M = 64	36	43	57
M = 128	49	69	36
M = 256	94	168	333
M = 512	258	547	1207

The dataset for training the neural network has been created by considering the NIH-3T3 cell line. For multiple flowing cells, hundreds of 5120×5120 holograms have been recorded. For each cell within the FOV, a 384×384 ROI has been cropped from the recorded hologram around the cell to preserve all the diffraction information useful for the autofocusing operator. Then, the corresponding QPM has been computed through the conventional holographic processing. Then, the hologram and the corresponding QPM have been resized to 256×256 thus becoming the input and the target of the network, respectively. In particular, the training set and the validation set have been created by randomly selecting respectively 4000 and 1000 images from 100 flowing and rotating cells. To train the network, some hyperparameters have been tuned. In particular, the Leaky ReLU coefficient has been set to 0.2, and a Dropout operation has been added to the Layer 7 with 0.5 factor to improve the generalization property of the network. Moreover, a

mini-batch with 100 observations has been used. Finally, the ADAM optimizer[62] has been employed to learn the parameters by minimizing the Mean Absolut Error (MAE)[63] computed as follows by the final Regression Layer

$$MAE = \frac{1}{K}\sum_k \frac{1}{P}\sum_p \left| z_{k,p} - \hat{z}_{k,p} \right| \quad (2)$$

where K is the number of observations in the mini-batch, P is the number of pixels p in an image, z is an output image, and $\hat{z}$ is the corresponding target image.

Results

The CAN model has been trained for 100 epochs by using an Intel® Core™ i9-9900K CPU with a 64Gb RAM through the Matlab® 2021a environment. The training step has required about 92h. However, as shown in Fig. 3A, the loss function curves of both the training and validation sets drop quickly after a few epochs, and they saturate around the 50th epoch. Therefore, the training time can be safely halved. The correctness in the QPM restoration has been evaluated by computing the Structural Similarity Index Measure (SSIM)[64] between the target QPM and the predicted QPM reported at the original $\mathbf{384 \times 384}$ size. In particular, the trained CAN model allows reaching a $\mathbf{0.962 \pm 0.013}$ accuracy measured over the 4000 images of the training set and a $\mathbf{0.961 \pm 0.015}$ accuracy measured for the 1000 images of the validation set. Moreover, a test set has been created by randomly selecting other 2000 images (not used to feed the network during the training step), achieving $\mathbf{0.961 \pm 0.013}$ of accuracy, thus indicating the substantial capability of generalization of this network. In Figs. 3B-D we show an observation belonging to the test set, made of the input (i.e., the holographic ROI), the target (i.e., the QPM obtained by the standard processing), and the output (i.e., the QPM computed by the network), respectively, in which an average SSIM of $\mathbf{0.961}$ is obtained. Beyond the numerical assessment of the regression performance, a comparison between phase profiles is also reported in Fig. 3E. This clearly shows the proficiency of the CAN in the realm of DH to preserve not only the mere cell morphology, but also the quantitative content of its 2D phase-contrast map. It is worth to remark that such process is also very fast. In fact, the trained model takes about 0.17 seconds to get from the hologram the unwrapped QPM. On the contrary, by the conventional holographic process, the same operation takes about 7.71 seconds, that is notably longer (i.e. about 45 times). This property is crucial in flow cytometry systems, since it allows analysing a much larger number of cells in the same time-period, thus enabling statistically relevant studies about specific cell populations. In fact, from the QPMs, 2D label-free features can be measured for diagnostic purposes[5-17]. As a consequence, a further way to validate the architecture we propose consists in checking whether the QPM outputs lead to the same features that would be measured from the corresponding QPM targets. To this aim, the 2000 cells belonging to the test set have been segmented from the background within the QPMs. The average phase and the average area have been computed for each cell in both the target and the output cases, which

corresponding histograms are compared in Figs. 3F,G, respectively. To quantify the matching between the histograms, the percentage error has been computed as follows

$$Err = 100\frac{1}{N}\sum_{i=1}^{N}\left|\frac{f_i - \hat{f}_i}{\hat{f}_i}\right| \quad (3)$$

where N is the number of observation, f is a feature computed from the output image, and $\hat{f}_i$ is the same feature computed from the corresponding target image. In the case of the average cell phase we obtained a 3.90% error, while in the case of the average cell area we obtained a 2.35% error.

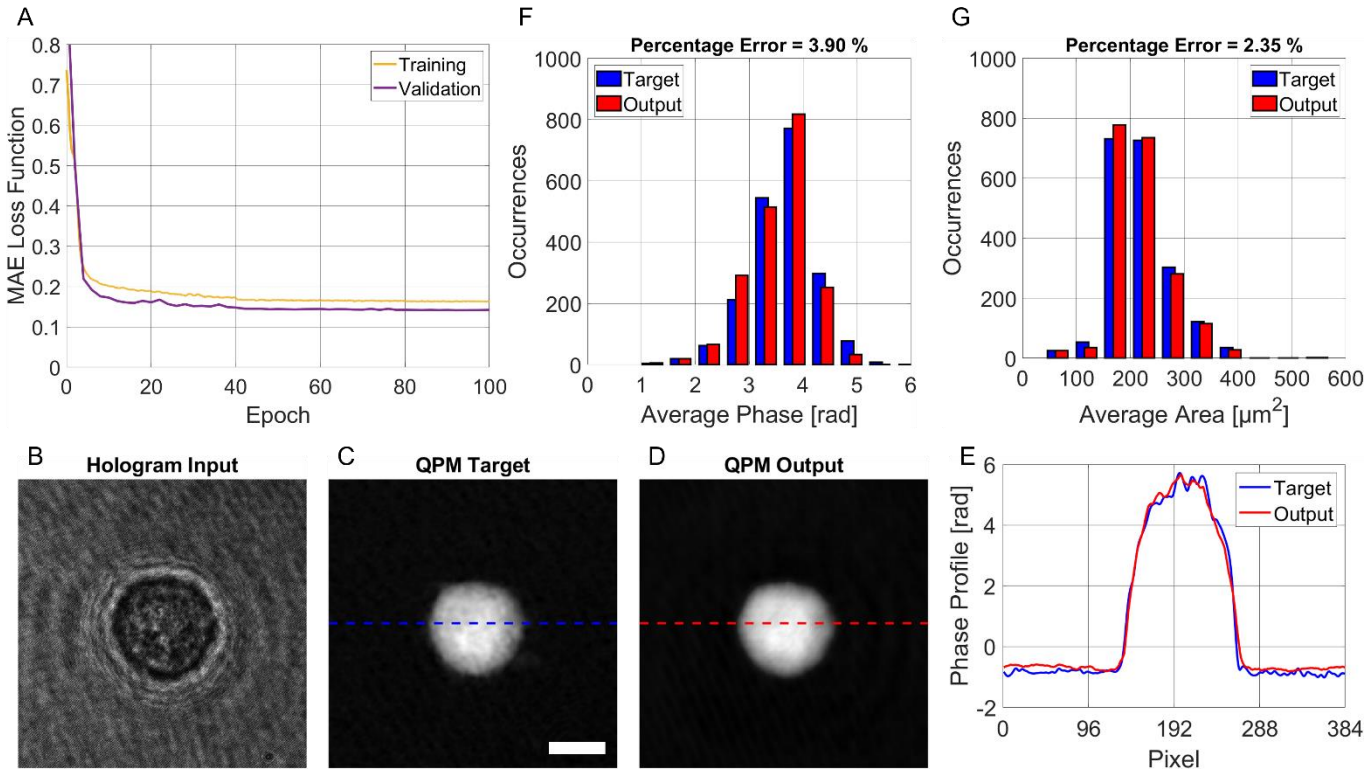

Figure 3. Assessment of the QPM reconstruction by deep learning. A MAE loss function computed at different epochs from the training set (yellow) and the validation set (violet). **B-D** Input, target, and output, respectively, of the trained CAN model containing a test cell. The SSIM between the target and the output is 0.961. The scale bar is 10 μm. **E** Phase profiles of the QPM target (blue) and QPM output (red) selected from the lines highlighted in C,D, respectively. **F,G** Comparison between the histograms of the average phase and the average area, respectively, computed from the QPM targets (blue) and the QPMs outputs (red) of each cell belonging to the test set. The percentage error between the measured features is reported at the top.

Performances of the trained network in non-conventional cases

In order to further assess the performances of the CAN architecture in correctly reconstructing the QPM, here we show some particular cases in which the network works surprisingly well, thus supporting its generalization property. In fact, the model has been trained by using single live spherical-like cells centred in their ROIs, as displayed in Figs. 3B-D. An example is reported in Fig. 4A, in which the phase-contrast maps of a dead cell are reported after reconstruction with the conventional method and the proposed one. A healthy cell has a convex phase profile, while a dead

cell loses its content thus resulting in an internal phase concavity[65]. The phase profiles plotted in Fig. 4A highlight that the network prediction is accurate in reproducing also the phase concavity, even though the network model had never received this kind of image as an example during the training step. Moreover, as the cells are suspended in the microfluidic flow, they usually have a spherical shape. However, when the cell suffers a stress condition, it could exhibit a distorted shape, and the reconstruction architecture should be able to reproduce it in order to be reliable for diagnostics. Remarkably, also in this case the network correctly predicts the QPM, as shown in Fig. 4B. It is worth remarking that, besides being not spheroid-like shaped, the cell in Fig. 4B is not centred in its ROI, unlike all the cells used in the training set. Finally, in the case reported in Fig. 4C, the network is even able to exceed the performances of the standard processing.

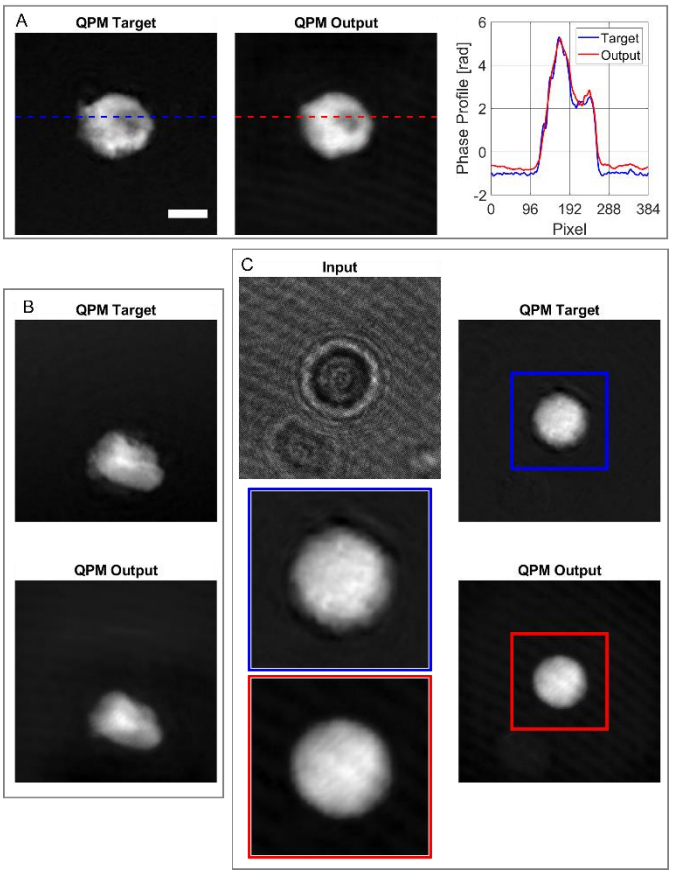

Figure 4. Assessment of the CAN performances in some particular cases. A QPM target and QPM output of a dead cell, with the phase profile corresponding to the highlighted lines. **B** QPM target and QPM output of a deformed cell, not centred in its ROI. **C** Input, QPM target, and QPM output of a cell not refocused well by the standard holographic processing (blue box), but rightly refocused by the CAN network (red box). The scale bar is 10 µm.

In fact, in this case the presence of a severely out-of-focus object near the analysed cell and overlapping with its ROI perturbs the minimization of the TC in the autofocusing process, thus resulting in a wrong estimation of the focal distance. This is clearly visible in Fig. 4C, where the zoomed QPM target shows residual diffraction rings near the cell contour, i.e. the cell is incorrectly

returned by the conventional method out-of-focus. Instead, as displayed in the zoomed QPM output, the proposed network is able to accurately refocus the cell. The highlighted special cases, in addition to the results reported in the previous section, underline that the network has learned the right mathematical operator that converts a hologram into the corresponding QPM.

3D tomographic Imaging

It is well known that phase-contrast values can be interpreted as the integral of the refractive index (RI) contrast along the optical axis. Therefore, in a QPM, the information about the 3D RI spatial distribution and the 3D morphology are coupled in the same 2D image[4]. Starting from the different QPMs recorded at multiple viewing angles around the flowing and rolling cell, this two information can be decoupled, thus permitting the reconstruction of the corresponding 3D RI tomogram[23]. In tomographic flow cytometry (TFC), the sample rotation is exploited while keeping the light probe and the optical system fixed. Thus, one of the main advantages of TFC with respect to the more conventional holographic tomography techniques is its potentially higher throughput. However, so far, this property has not been exploited in full, because the heavy computational burden of holographic processing turns up to significantly slow down the tomographic reconstruction, thus leading to very few cells as output of each TFC experiment. In fact, to reconstruct the 3D tomogram of each cell, on average more than 200 holographic ROIs must be processed to obtain the corresponding QPMs. To test the potentiality of the proposed network in the TFC framework, we have reconstructed 65 tomograms by using both the standard and the DCNN based approaches to recover the QPMs. In particular, 65 cells correspond to 14462 recorded holograms, whose QPM retrieval takes about 31 hours by using the standard processing and only 41 minutes by exploiting the CAN model, i.e. the CAN inference can do the same task using only the 2% of the time required to the conventional method. In Fig. 5A we display the central slices of the 3D RI tomograms of the same cell, respectively reconstructed from the QPMs obtained in the standard processing modality and through the proposed network. A great similarity has been reached, as also underlined by the good agreement between the RI profiles reported in Fig. 5B. A further proof is the high symmetry of the corresponding RI violin histogram in Fig. 5C, which is a visual representation of the 0.997 SSIM computed between the two tomograms. As well as in the 2D case, also in the 3D case it is important to preserve the truthfulness of the statistical measurements, especially their adherence to the quantitative ground-truth. To this aim, in Figs. 5D,E we report the histograms of the average RI and the equivalent radius (i.e., the radius of a sphere having the same volume of the analysed cell) calculated for the 65 reconstructed tomograms, where a 0.07% and a 0.70% percentage errors are obtained, respectively. An important quantitative feature that can be inferred from the 3D RI tomogram is the dry mass[66]. It is defined as the mass of the cell in the absence of water content, i.e.

$$m = \frac{(\bar{n} - n_0)V}{\alpha} \tag{4}$$

where $\bar{n}$ is the cell average RI, n_0 is the RI of the surrounding medium, V is the cell volume, and α is the RI increment, which is 0.2 mL/g for a nucleated cell[67]. The dry mass can be considered a bioindicator of the health state of the cell as it is related to its biophysical properties. Therefore, the fast and accurate quantification of the dry mass for a large number of cells could encourage further developments of TFC-based diagnostic applications in biomedicine[68]. For this reason, the low percentage error of 3.77% obtained in the case of the dry mass reported in Fig. 5F by using the CAN architecture acquires even more importance. As a counterweight to the abovementioned advantages of the reported results, a limitation can be recognized in the partial loss of internal RI contrast. This effect is clearly visible in the central slice comparison in Fig. 5A, and is quantified as a percentage error of 11.92% about the RI standard deviation in Fig. 5G, where we notice a shift to lower values of the histogram obtained from DCNN processing with respect to the standard one. Therefore, the proposed CAN architecture allows for an accurate characterization of the global cellular features. However, in order to perform an intracellular analysis, we infer that a contrast enhancement could be easily obtained by increasing the width of the network, which leads to a larger number of parameters.

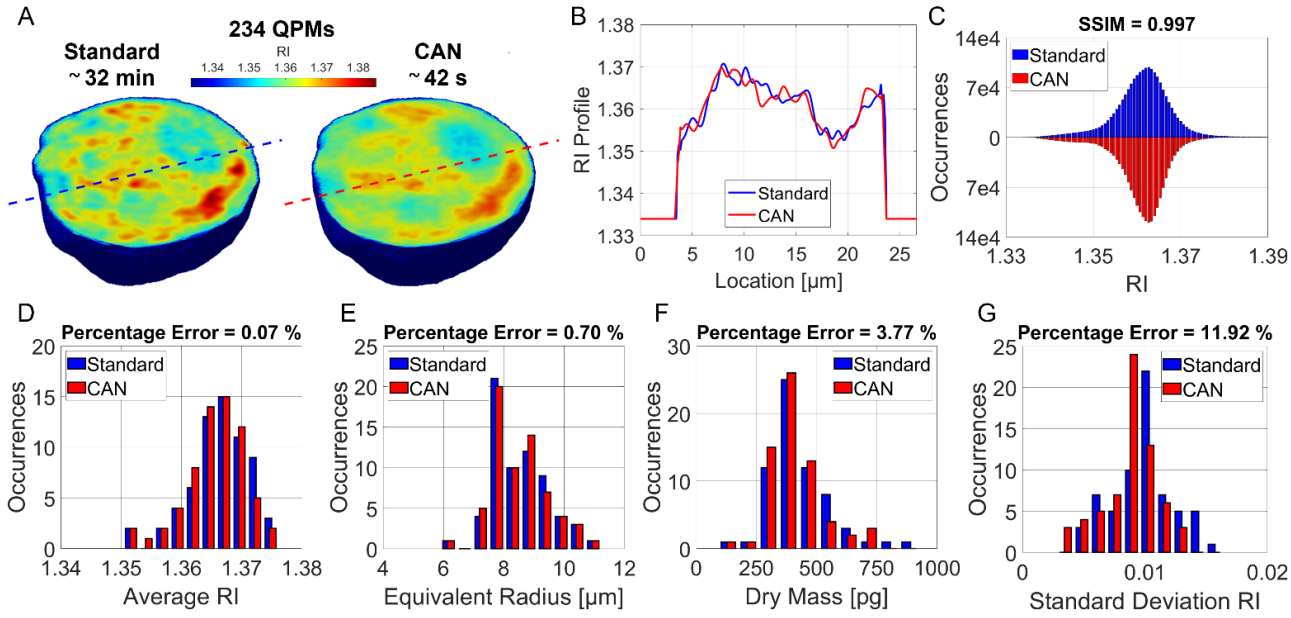

Figure 5. Assessment of the tomographic reconstruction by deep learning. A Central slice of the 3D RI tomogram reconstructed from 234 QPMs retrieved (on the left) in the standard way in about 32 min and (on the right) by the CAN model in about 42 s. **B** RI profile corresponding to the lines highlighted in A from the standard tomogram (blue) and the CAN tomogram (red). **C** Violin histogram of the 3D RI distribution about the standard tomogram (blue) and the CAN tomogram (red) in A, which SSIM is reported at the top. **D-G** Comparison between the histograms of the average RI, the equivalent radius, the dry mass, and the standard deviation RI, respectively, computed from 65 standard tomograms (blue) and CAN tomograms (red). The percentage error between the measured features is reported at the top.

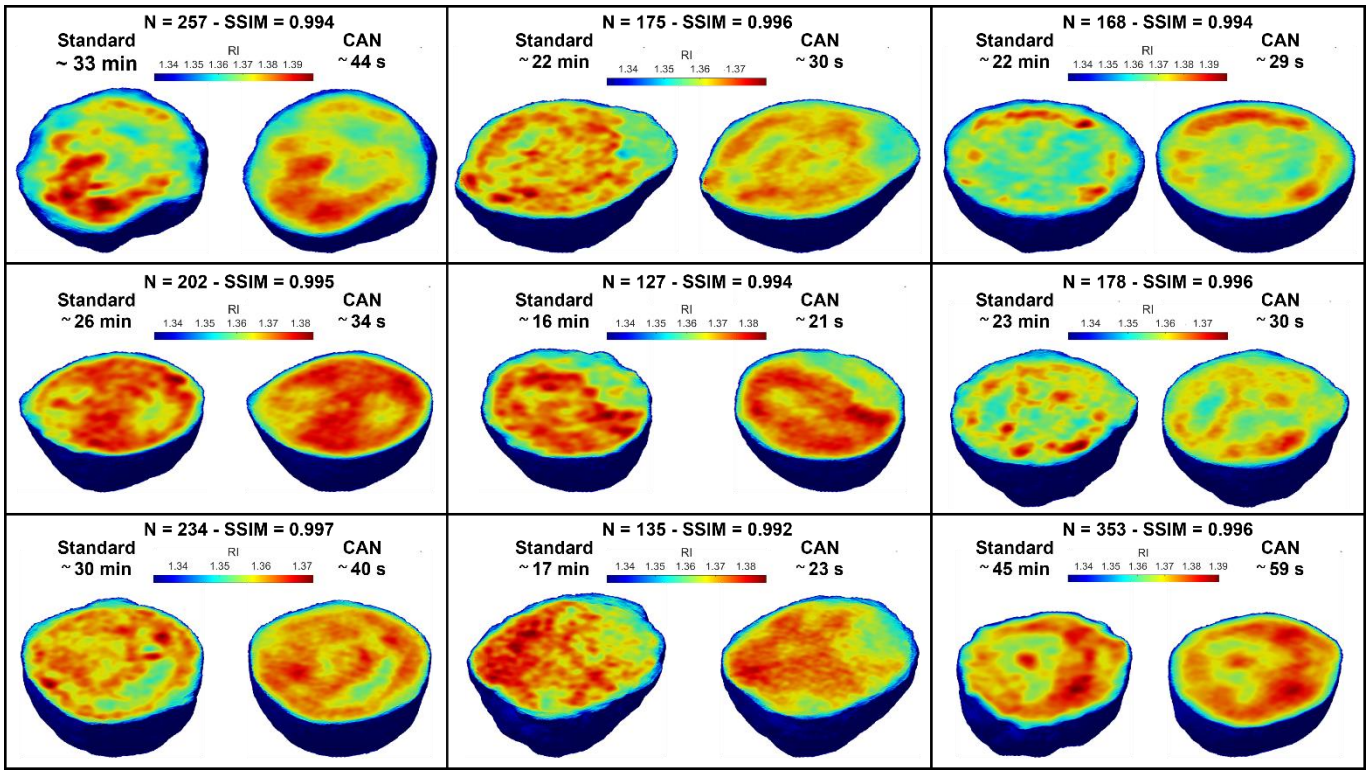

Figure 6. 3D tomographic reconstructions. Comparison between the central slices of 9 3D RI tomograms reconstructed by using the QPMs obtained through the standard holographic processing (on the left) and the CAN model (on the right). For each cell, the number of QPMs, the SSIM between the tomograms, and the computational time of the standard phase retrieval and the CAN-based phase retrieval are reported at the top.

Conclusions

In this paper we have proposed a DCNN to reach very fast processing in obtaining tomographic reconstruction of flowing and rotating cells in a holographic flow cytometry system. In particular, we have implemented, for the first time in holographic imaging, a suitable DCNN architecture, namely CAN, able to guarantee high prediction accuracy, constant runtime and very low memory usage. In fact, by investigating the trade-off among the aforementioned properties, we have allowed the computing of a single QPM in 168 ms, with a negligible details' loss and with a trained model that occupies only 818 kB of memory. By employing the proposed DCCN based reconstruction method on a sequence of QPMs of the same rotating cell, we have demonstrated the possibility to recover its tomogram in few seconds instead of tens of minutes, while essentially preserving the high-content information of tomographic data. Reported results show a very good agreement between data measured using the conventional DH processing and the quantitative parameters measured from the network output, substantially showing their equivalence for diagnostic purposes. Besides, the network has been demonstrated able to outperform the conventional DH processing in all the cases where the presence of multiple objects within the same ROI can determine a failure of the autofocusing algorithms. The proposed DCNN model provides enough compactness and computing velocity to be fit into

on-chip SRAM, opening to the possibility of performing onboard computations, which is highly demanded property for lab-on-chip devices with low processing hardware resources.

Author Contributions

D.S. L.M. and V.B. set the holographic flow cytometry system and were responsible of the holographic acquisitions; M.M. prepared the biological samples and contributed to the experiments; D.P. were in charge for the data analysis and numerical results. All the authors contributed to critical discussion of the results and contributed to write the manuscript. P.M. and P.F. conceptualized and supervised the research.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was funded by project PRIN 2017, Morphological Biomarkers for early diagnosis in Oncology (MORFEO) Prot. 2017N7R2CJ.

References

1 P. Ferraro, A. Wax and Z. Zalevsky, *Coherent Light Microscopy*, Springer Series in Surface Sciences (Springer, 2011), Vol. 46.
2 M. K. Kim, *SPIE Rev.*, 2010, **1**, 018005.
3 Y. Cotte, F. Toy, P. Jourdain, N. Pavillon, D. Boss, P. Magistretti, P. Marquet and C. Depeursinge, *Nat. Photonics*, 2013, **7**, 113–117.
4 Y.K. Park, C. Depeursinge and G. Popescu, *Nat. Photonics*, 2018, **12**, 578–589.
5 B. Kemper, D. D. Carl, J. Schnekenburger, I. Bredebusch, M. Schäfer, W. Domschke and G. von Bally, *J. Biomed. Opt.*, 2006, **11**, 034005–034008.
6 Z. El-Schich, A. L. Mölder, and A. Gjörloff Wingren, *Appl. Sci.*, 2018, **8(7)**, 1027.
7 A. V. Belashov, A. A. Zhikhoreva, T. N. Belyaeva, A. V. Salova, E. S. Kornilova, I. V. Semenova and O. S. Vasyutinskii, *Cells*, 2021, **10**, 2587.
8 S. Cho, S. Kim, Y. Kim and Y. K. Park, *Trends Biotechnol.*, 2011, **30(2)**, 71-79.
9 F. Yi, I. Moon and B. Javidi, *Biomed. Opt. Express*, 2016, **7**, 2385-2399.
10 L. Miccio, P. Memmolo, F. Merola, P.A. Netti and P. Ferraro, *Nat. Comm.*, 2015, **6**, 6502.
11 M. Mugnano, P. Memmolo, L. Miccio, F. Merola, V. Bianco, A. Bramanti, A. Gambale, R. Russo, I. Andolfo, A. Iolascon and P. Ferraro, *Anal. Chem.*, 2018, **90**, 7495−7501.
12 P. Lenz, D. Bettenworth, P. Krausewitz, M. Brückner, S. Ketelhut, G. von Bally, D. Domagk and B. Kemper, *Integr. Biol.*, 2013, **5**, 624–630.
13 Y. Park, M. Diez-Silva, G. Popescu, G. Lykotrafitis, W. Choi, M. S. Feld and S. Suresh, *Proc. Natl Acad. Sci. USA*, 2008, **105**, 13730–13735.
14 T. O'Connor, J.-B. Shen, B. T. Liang and B. Javidi, *Opt. Lett.*, 2021, **46**, 2344-2347.
15 N. Goswami, Y. R. He, Y.-H. Deng, C. Oh, N. Sobh, E. Valera, R. Bashir, N. Ismail, H. Kong, T. H. Nguyen, C. Best-Popescu and G. Popescu, *Light: Science & Applications*, 2021, **10**, 176.
16 P. Memmolo, L. Miccio, M. Paturzo, G. Di Caprio, G. Coppola, P.A. Netti and P. Ferraro, *Adv. Opt. Photon.*, 2015, **7**, 713-755.
17 L. ShenemanI, G. Stephanopoulos and A. E. Vasdekis, *PLoS ONE*, 2021, **16(4)**, e0249196.
18 V. Balasubramani, A. Kuś, H.-Y. Tu, C.-J. Cheng, M. Baczewska, W. Krauze and M. Kujawińska, *Appl. Opt.*, 2021, **60**, B65-B80.
19 D. Jin, R. Zhou, Z. Yaqoob and P. T. C. So, *J. Opt. Soc. Am. B*, 2017, **34**, B64-B77.
20 H. Zhu, S.O. Isikman, O. Mudanyali, A. Greenbauma and A. Ozcan, *Lab Chip*, 2013, **13**, 51-67.
21 F. Merola, P. Memmolo, L. Miccio, V. Bianco, M. Paturzo and P. Ferraro, *Proc. IEEE*, 2015, **103(2)**, 192-204.

22 V. Bianco, B. Mandracchia, V. Marchesano, V. Pagliarulo, F. Olivieri, S. Coppola, M. Paturzo and P. Ferraro, *Light Sci. Appl.*, 2017, **6**, e17055.
23 F. Merola, P. Memmolo, L. Miccio, R. Savoia, M. Mugnano, A. Fontana, G. D'Ippolito, A. Sardo, A. Iolascon, A. Gambale and P. Ferraro, *Light Sci. Appl.*, 2017, **6**, e16241.
24 F. Merola, P. Memmolo, L. Miccio, M. Mugnano and P. Ferraro, *Methods*, 2018, **136**, 108-115.
25 M.M. Villone, P. Memmolo, F. Merola, M. Mugnano, L. Miccio, P.L. Maffettone and P. Ferraro, *Lab Chip*, 2018, **18**, 126-131.
26 N.C. Pégard, M.L. Toth, M. Driscoll and J.W. Fleischer, *Lab Chip*, 2014, **14**, 4447-4450.
27 D. Pirone, P. Memmolo, F. Merola, L. Miccio, M. Mugnano, A. Capozzoli, C. Curcio, A. Liseno and P. Ferraro, *Appl. Opt.*, 2021, **60**, A277-A284.
28 A. Kleiber, D. Kraus, T. Henkel, and W. Fritzsche, Lab Chip, 2021, **21**, 3655-3666.
29 L. Miccio, F. Cimmino, I. Kurelac, M.M. Villone, V. Bianco, P. Memmolo, F. Merola, M. Mugnano, M. Capasso, A. Iolascon, P.L. Maffettone and P. Ferraro, View, 2020, **1**, 20200034.
30 D. Pirone, M. Mugnano, P. Memmolo, F. Merola, G. C. Lama, R. Castaldo, L. Miccio, V. Bianco, S. Grilli and P. Ferraro, *Nano Lett.*, 2021, **21(14)**, 5958–5966.
31 Z. Wang, V. Bianco, D. Pirone, P. Memmolo, M. M. Villone, P. L. Maffettone and P. Ferraro, *Light Sci. Appl.*, 2021, **10**, 187.
32 P. Memmolo, V. Renò, E. Stella and P. Ferraro, *Appl. Opt.*, 2019, **58**, G155-G161.
33 T. Latychevskaia, *J. Opt. Soc. Am. A*, 2019, **36**, D31-D40.
34 S. K. Mohammed, L. Bouamama, D. Bahloul and P. Picart, *Appl. Opt.*, 2017, **56**, F158-F166.
35 C. Trujillo, R. Castañeda, P. Piedrahita-Quintero and J. Garcia-Sucerquia, *Appl. Opt.*, 2016, **55**, 10299-10306.
36 W. Zhou, Y. Yu and A. Asundi, *Opt. Lasers Eng.*, 2009, **47(2)**, 264-270.
37 J. Bioucas-Dias, V. Katkovnik, J. Astola and K. Egiazarian, *Appl. Opt.*, 2008, **47**, 5358–5369.
38 S. Montrésor, P. Memmolo, V. Bianco, P. Ferraro and P. Picart, *J. Opt. Soc. Am. A*, 2019, **36**, A59-A66.
39 G. Barbastathis, A. Ozcan and G. Situ, *Optica*, 2019, **6**, 921-943.
40 Y. Rivenson, Y. Wu and A. Ozcan, *Light Sci. Appl.*, 2019, **8**, 85.
41 Z. Ren, Z. Xu and E.Y. Lam, *Optica*, 2018, **5**, 337-344.
42 K. Jaferzadeh, S.-H. Hwang, I. Moon and B. Javidi, *Biomed. Opt. Express*, 2019, **10**, 4276-4289.
43 T. Pitkäaho, A. Manninen and T.J. Naughton, *Appl. Opt.*, 2019, **58**, A202-A208.
44 O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, 234-241.
45 T. Nguyen, V. Bui, V. Lam, C.B. Raub, L.-C. Chang, and G. Nehmetallah, *Opt. Express*, 2017, **25**, 15043-15057.
46 K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, 770–778.
47 W. Xiao, L. Xin, R. Cao, X. Wu, R. Tian, L. Che, L. Sun, P. Ferraro and F. Pan, *Lab Chip*, 2021, **21(7)**, 1385–1394.
48 G. Dardikman-Yoffe, D. Roitshtain, S.K. Mirsky, N.A. Turko, M. Habaza and N.T. Shaked, *Biomed. Opt. Express*, 2020, **11**, 1107-1121.
49 Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng and A. Ozcan, *Light Sci. Appl.*, 2018, **7**, 17141.
50 H. Wang, M. Lyu and G. Situ, *Opt. Express*, 2018, **26**, 22603-22614.
51 Z. Ren, Z. Xu, and E.Y. Lam, *Adv. Photonics*, 2019, **1(1)**, 016004.
52 K. Wang, J. Dou, Q. Kemao, J. Di and J. Zhao, *Opt. Lett.*, 2019, **44**, 4765-4768.
53 T. Zeng, H. K.-H. So and E. Y. Lam, *Opt. Express*, 2020, **28**, 4876-4887.
54 Q. Chen, J. Xu, and V. Koltun. "Fast Image Processing with Fully-Convolutional Networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision*, 2017, 2516-2525.
55 F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in *International Conference on Learning Representations*, 2016.
56 S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," in *International Conference on Learning Representations*, 2016.
57 Q. Kemao, *Appl. Opt.*, 2004, **43**, 2695–2702.
58 J. Bioucas-Dias and G. Valadão, *IEEE Trans. Image Process.*, 2007, **16**, 698–709.
59 A. L. Maas, A. Y. Hannun and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proceedings of the 30th International Conference on Machine Learning*, 2013.
60 S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 30th International Conference on Machine Learning*, 2015.
61 X. Mao, C. Shen and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, 2802–2810.
62 D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," in *International Conference on Learning Representations*, 2015.
63 H. Zhao, O. Gallo, I. Frosio and J. Kautz, *IEEE Trans. Comput. Imaging*, 2017, **3**, 47–57.
64 W. Zhou, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, *IEEE Trans. Image Process*, 2004, **13**, 600–612.
65 M. Mugnano, P. Memmolo, L. Miccio, S. Grilli, F. Merola, A. Calabuig, A. Bramanti, E. Mazzon and P. Ferraro, *J. Biophotonics*, 2018, **11**, e201800099.
66 T. Zangle and M. Teitell, *Nature Methods*, 2014, **11(12)**, 1221–1228.
67 G. Popescu, Y. Park, N. Lue, C. Best-Popescu, L. Deflores, R. R. Dasari, M. S. Feld and K. Badizadegan, *American Journal of Physiology – Cell Physiology*, 2008, **295(2)**, C538–C544.
68 Y. Liu, P. Y. Liu, L. K. Chin, W. Ser, H. F. Chen, C.-M. Hsieh, C.-H. Lee, K.-B. Sung, T. C. Ayi, P. H. Yap, B. Liedberg, K. Wang, T. Bourouina and Y. Leprince-Wang, Lab Chip, 2016, 16, 634–644.