

Supplementary information

CNN-based cell analysis: from image to quantitative representation

Cédric Allier¹, Lionel Hervé¹, Chiara Paviolo¹, Ondrej Mandula¹, Olivier Cioni¹, William Pierré¹, Francesca Andriani², Kiran Padmanabhan², Sophie Morales¹

¹ Univ. Grenoble Alpes, CEA, LETI, DTBS, 38000, Grenoble, France

² Institut de Génomique Fonctionnelle de Lyon, CNRS UMR 5242, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, 69007 Lyon, France

*** Correspondence:**

Cédric Allier
cedric.allier@cea.fr

Supplementary figure captions

FIGURE S1. (A) OPD reconstructed images are 32-bit, values being given in micrometer. OPD images are not transformed for CNN training. This therefore implies that the obtained CNNs will work best with OPD images in the range of 0 to 1 μm for which they have been trained. In the considered dataset cell OPD values are well distributed into the considered range of (0, 1 μm) as depicted in (B), being the OPD value distribution over image (A).

FIGURE S2 | (A) OPD image reconstructed with reference algorithm (200 s). The cell concentration varies between 180 and 550 cells/ mm^2 . All images are details (500x500 μm^2) of a $\sim 30 \text{ mm}^2$ reconstructed OPD image. (B) OPD image reconstructed with accelerated algorithm (3 s). SSIM calculated between (A) and (B) are given. (C) Difference between OPD images reconstructed with reference algorithm (200 s) and accelerated algorithm (3 s). Root means square deviation of the OPD images difference is given.

FIGURE S3 | (A) Test OPD images featuring 5381 fibroblast cells (183 cells/ mm^2). (B) Test OPD images featuring 7490 cells (255 cells/ mm^2). (C) Test OPD images featuring 10714 cells (364 cells/ mm^2). (D) Test OPD images featuring 15500 cells (527 cells/ mm^2). All OPD images are reconstructed with the reference reconstruction algorithm (200 s computation time).

FIGURE S4 | OPD image degradations. (A) 30 nm maximum amplitude non-uniform background used for a CNN robustness evaluation tasks. Field of view is that of the lens-free acquisition ($\sim 30 \text{ mm}^2$). (B) OPD profile along the yellow line in (A). (C) Detail of a test OPD image of fibroblast cells. For a CNN robustness evaluation task image (C) has been degraded with (D) noise (20 nm spatial variation S.D.), (E) Non-uniform background (30 nm max. amplitude, see (A, B)) and (F) a combination of noise and non-uniform background. (G, H, I, J) OPD profiles along diagonal of (C, D, E, F).

CNN-based cell culture analysis: from phase image to quantitative representation

	N cells	Cell conc.	Mean	S.D.	SSIM	Prec.	Recall	Slope	Intercept	R ²	Error MAD
Cell area	5381	183	584.8	157.9	0.93	0.99	0.99	0.76	120.4	0.858	44.1
[μm^2]	7490	255	641.3	175.5	0.91	0.99	0.99	0.76	125.8	0.860	50.6
	10714	364	673.7	175.9	0.86	0.99	0.99	0.72	149.3	0.820	61.0
	15500	527	655.0	166.4	0.80	0.98	0.99	0.66	184.1	0.734	82.6
Cell dry mass	5381	183	184.9	44.7	0.94	0.99	0.99	0.74	33.5	0.912	13.2
[pg]	7490	255	194.9	49.7	0.92	0.99	0.99	0.73	36.0	0.903	14.3
	10714	364	209.0	55.3	0.88	0.99	0.99	0.72	40.0	0.895	16.1
	15500	527	203.4	55.8	0.82	0.99	0.98	0.72	39.0	0.860	17.5

TABLE S1 | Results of a CNN conducting two tasks in parallel, namely determination of cell area and cell dry mass. OPD test images are reconstructed with the reference reconstruction algorithm (200 s computation time). The results of a standard image processing pipeline are the reference. We used four test images taken at different time point of the cell culture. They feature respectively 5,381, 7,490, 10,714 and 15,500 fibroblast cells. Corresponding cell concentrations are given in cells/mm². The cell measurements obtained with the standard image processing follow normal distributions of which means and standard deviations are listed. SSIM is calculated between the CNN predictions and ground truth quantitative representations. SSIM values below 0.9 are overlaid in orange. The detection ability of the CNN solution is assessed with precision and recall value. To estimate the discrepancies between the cell measurements obtained with the CNN prediction and the reference values, we performed linear regression fits to obtain the coefficient of determination R², the slope and the intercept. Slope and R² values below 0.85 are overlaid in orange. Error MAD refers to the absolute deviation from the median of the error (Eq. 6).

CNN-based cell culture analysis: from phase image to quantitative representation

	N cells	Cell conc.	Mean	S.D.	SSIM	Prec.	Recall	Slope	Intercept	R ²	Error MAD
Cell area [μm^2]	5381	183	584.8	157.9	0.93	0.99	0.99	0.96	129.1	0.858	39.4
	7490	255	641.3	175.5	0.90	0.98	0.99	0.95	131.9	0.846	45.6
	10714	364	673.7	175.9	0.86	0.98	0.98	0.90	166.9	0.800	52.5
	15500	527	655.0	166.4	0.79	0.98	0.98	0.81	226.3	0.709	63.7
Cell dry mass [pg]	5381	183	184.9	44.7	0.94	0.99	0.99	1.08	33.0	0.853	16.3
	7490	255	194.9	49.7	0.92	0.99	0.99	1.03	42.0	0.848	16.6
	10714	364	209.0	55.3	0.88	0.99	0.99	1.04	39.2	0.839	19.2
	15500	527	203.4	55.8	0.83	0.98	0.99	1.02	41.2	0.797	22.1
Cell max. OPD [nm]	5381	183	79.4	20.8	0.93	0.99	0.99	0.99	12.90	0.921	5.9
	7490	255	75.5	17.0	0.90	0.99	1.00	1.00	11.16	0.938	5.8
	10714	364	76.8	16.5	0.86	0.99	0.99	0.99	10.86	0.926	5.9
	15500	527	74.7	15.9	0.80	0.98	0.99	0.99	9.84	0.923	6.2
Cell major axis length [μm]	5381	183	40.4	11.2	0.92	0.98	1.00	0.94	9.9	0.798	3.3
	7490	255	42.3	11.5	0.89	0.99	0.99	0.93	10.6	0.802	3.6
	10714	364	42.9	11.2	0.84	0.98	0.99	0.87	12.8	0.740	4.1
	15500	527	40.0	9.4	0.77	0.98	0.99	0.80	14.8	0.638	4.2

TABLE S2 | Results of CNN-based quantifications obtained on OPD test images reconstructed with the accelerated reconstruction algorithm (3 s computation time). The results of a standard image processing pipeline are the reference. We used four test images taken at different time point of the cell culture. They feature respectively 5,381, 7,490, 10,714 and 15,500 fibroblast cells. Corresponding cell concentrations are given in cells/mm². The cell measurements obtained with the standard image processing follow normal distributions of which means and standard deviations are listed. SSIM is calculated between the CNN predictions and ground truth quantitative representations. SSIM values below 0.9 are overlaid in orange. The detection ability of the CNN solution is assessed with precision and recall value. To estimate the discrepancies between the cell measurements obtained with the CNN prediction and the reference values, we performed linear regression fits to obtain the coefficient of determination R², the slope and the intercept. Slope and R² values below 0.85 are overlaid in orange. Error MAD refers to the absolute deviation from the median of the error (Eq. 6).

CNN-based cell culture analysis: from phase image to quantitative representation

CNN-A trained with experimental images	SSIM	Prec.	Recall	Slope	Intercept [pg]	R ²	Error MAD [pg]
Raw image	0.92	0.99	0.99	0.79	64	0.872	11.0
+Noise 10 nm S.D.	0.92	0.99	0.99	0.80	54	0.855	15.0
+Noise 20 nm S.D.	0.91	0.99	0.99	0.77	55	0.779	23.2
+BG 15 nm max. amplitude	0.91	0.99	0.99	0.77	107	0.755	19.4
+BG 30 nm max. amplitude	0.88	0.99	0.99	0.70	164	0.457	36.0
+Noise 10 nm +BG 15 nm	0.91	0.99	1.00	0.79	89	0.799	18.0
+Noise 20 nm +BG 30 nm	0.90	0.99	0.99	0.75	124	0.622	28.7
CNN-B trained with experimental images numerically degraded	SSIM	Prec.	Recall	Slope	Intercept [pg]	R ²	Error MAD [pg]
Clean image	0.88	0.99	0.98	0.97	74	0.813	24.9
Raw image	0.91	0.99	0.99	0.98	49	0.852	19.6
+Noise 20 nm S.D.	0.92	0.99	0.99	0.98	23	0.866	17.5
+BG 15 nm max. amplitude	0.89	0.99	0.99	1.02	45	0.882	16.2
+BG 30 nm max. amplitude	0.91	0.99	0.99	1.01	38	0.884	15.4
+Noise 10 nm +BG 15 nm	0.92	0.99	0.99	1.00	43	0.880	15.2
+Noise 20 nm +BG 30 nm	0.92	0.99	1.00	1.00	29	0.874	15.1

TABLE S3 | Results of two CNNs (CNN-A and CNN-B) performing cell dry mass quantification tested on degraded OPD images. CNN-A was trained with raw OPD images (spatial noise of 1.2 nm S.D. in background area of 130x130 μm^2). CNN-B was trained with numerically degraded OPD images (spatial noise of 20 nm S.D. and non-uniform background of 30 nm maximum amplitude). The results of a standard image processing pipeline applied to the raw images are the reference. We used the test image featuring 7,490 cells (255 cells/ mm^2). The cell dry mass values obtained with the standard image processing follow a normal distributions with mean of 195 pg and 50pg S.D. SSIM is calculated between the CNN predictions and ground truth quantitative representations. SSIM values below 0.9 are overlaid in orange. The detection ability of the CNN solution is assessed with precision and recall value. To estimate the discrepancies between the cell measurements obtained with the CNN prediction and the reference values, we performed linear regression fits to obtain the coefficient of determination R², the slope and the intercept. Slope and R² values below 0.85 are overlaid in orange. Intercept values larger than 50 pg are overlaid in orange. Error MAD refers to the absolute deviation from the median of the error (Eq. 6).