CABaNe, an automated, high content ImageJ macro for cell and neurite analysis

Nathan Thibieroz[1], Fabrice Cordelières[2], Paul Machillot[1], Akshita Singh[1], Catherine Picart[1],Elisa Migliorini[1]

[1] Univ Grenoble Alpes , CEA, INSERM U1292 Biosanté , CNRS EMR BRM 5000, Grenoble, France

[2] CNRS UMS 3420, Bordeaux Imaging Center , Bordeaux, France

Abstract

Measuring neurite length is crucial in neurobiology because it provides valuable insights into the growth, development, and function of neurons. In particular, neurite length is fundamental to study neuronal development and differentiation, neurons responses to drugs, neurodegenerative diseases and neuronal plasticity. Surprisingly, there is currently a lack of tools for high content neurite analysis. In this article, we present CABaNe, as an open source, high content, rule based Image J macro for cell analysis, including their neurite length. This macro possesses a graphical interface, metadata production, as well as verification means before and after analysis. Rule based and machine learning based programming have been tested for cell identification. After testing, we had better precision and adaptability using rule based cell identification. We challenged CABaNe with currently used techniques, which are manual or assisted. When tested on a small sample, CABaNe demonstrated a massive speed increase in capacity to treat dataset while maintaining or increasing precision when compared to manual measurement. When tested on a large data set, comparing different conditions, we successfully highlighted differences between conditions, in a fully automated manner. Therefore, CABaNe is viable as a high content option for cell analysis, for neurite length and other parameters. It is a base of code that can be used for other analysis or to train deep learning models. In the future, we expect this tool to be widely used in both basic and applied neurobiology research.

Significance statement

When studying neuronal cell differentiation, an important morphological parameter is neurite length. This parameter requires measuring the protrusions length of analysed cells. However, this analysis done manually can be long, as each individual cell must be measured independently. Currently, efficient single cell tools exist to assist the measurement, such as NeuronJ. However, there

is currently no available automated tool for this analysis, and manual techniques suffer operator bias. In this paper, we present a macro to fully automatize neurite length and other parameters measurement, for each cell, in each image, in each condition.

1 Introduction

Measuring neurite length is crucial in neurobiology because it provides valuable insights into the growth, development, and function of neurons. Growth of neurites is essential for the establishment of neuronal networks. Neurite length is a measure of neuronal maturation, polarity and functional potential. It can help assess how different signaling pathways, growth factors, or drugs influence neuronal development (Wiatrak et al., 2020; Rumpansuwon et al., 2022; Sánchez-Sarasúa et al., 2022), or in the field of regenerative medicine, where the aim is to differentiate cells in neurons to repair injuries (Gao et al., 2022). In addition, analyzing neurite length changes can also help the understanding of how neural circuits adapt to experiences and environmental stimuli (Crutcher and Collins, 1982; Gilbert and Man, 2017). In diseases like Alzheimer's, Parkinson's, and Huntington's, neurite shortening or atrophy can serve as biomarkers for disease progression (Schwab and Ebert, 2015; Barnat et al., 2017; Ghatak et al., 2019).

Finally, neurite length is important to study cancers such as glioblastoma and neuroblastoma (NB) where more differentiated cancers are less aggressive than their more immature counterpart (Mohlin et al., 2011; Sabelström et al., 2019). In particular, NB in particular represents 7-10% of diagnosed paediatric cancers, and is responsible for 15% of childhood cancer fatalities (Gurney et al., 1997; Davidoff, 2012; Gatta et al., 2014).

Mature, differentiated NB cells usually form neurites, therefore measuring their length is often used as parameter to assess differentiation. Antibodies for fluorescent staining (e.g. anti-beta III tubulin for neurite and anti-MAP2 for dendrites), facilitate quantification or morphological analysis (dos Santos et al., 2023). We reviewed some of the work in literature on NB and cell differentiation to evaluate how the measurement of neuronal differentiation is generally performed. The methods used, including for measuring neurite length, are outlined in Table 1. We noticed that the analysis is mainly manual.

While there exist automated analysis solutions, these are either linked to a specific equipment (Ramm et al., 2003), expensive and running on outdated versions of Windows (Beier et al., 2017),or free but not adapted to simple, cell culture, 2D images (Wang et al., 2010).

Table 1: Summary of read out in NB cell lines differentiation studies. For neurite length, when technique is not mentioned, it is assumed to be manual. Manual refers to measurement of neurite length using ImageJ or equivalent. When the analysis is assisted with code, but the operation is still manual, it is considered as assisted manual.

Article	Cell line	Neurite length measurement method/magnification	Neurite length	Fluorescence (intensity, signal)	qPCR	Western blot
(Thomson et al., 2022)	SH	**Manual**, **assisted** 20x	(**80 x8 images x8 conditions**) Longest Neurite length, Number of branches, using NeuronJ.	BETA-III, MAP2	X	MAP2
(Lauzon et al., 2017)	SH	**Manual** 20x	Euclidian distance between cell body and end point. **300 neurite** per conditions.	Smad 1/5, Beta III, MAP2, NeuN, NSE, ChAT, VAchT, GSK3	X	MAP2
(Mundhara and Panda, 2023)	SH	**Manual**, 10x	from periphery of cell to terminal end. (**60/conditions**)	Tubulin	X	Tubulin
(Ozgun et al., 2021)	SH	**Manual**	no information on how the process took place	TUJ1 (Beta III tubulin)	X	TUJ1 (Beta III tubulin)
(Labour et al., 2012)	SH	**Manual**	All neurite of cells, and the longest of each. (**80/conditions**). Cell considered differentiated if neurite length > 2 diameter of cell body.	X	X	X
(Kumar and Katyal, 2018)	N2A	**Manual, assisted** 20x	Neurite length, only accounting neurite longer than 1 µm. Branching patterning was obtained using Sholl analysis. (**800/day**)	MAP2	MAP2	X
(Shimizu et al., 2002)	N2A	X	X	X	X	Foxa1 and 2
(Namsi et al., 2018)	N2A	**Manual**	assessing presence of dendrites and/or axons (**20 images**)	X	NGF, BDN	X
(El Merhie et al., 2019)	N2A	**Manual**	number of cell with neurite, and length	MAP2	X	X

To test the role of different molecules, drugs, cellular responses to external stimuli and have a statistical relevance a large number of conditions have to be tested in parallel and multiple images per conditions have to be acquired. As illustrated in Figure 1, this could produce a large volume of data to process. With current methods, that are mainly manual (Table 1), analysing neurite length of this dataset of images would take hundreds of hours. Manual analysis suffer from operator bias, variability in measurement and

protocols, and lack of quality control. Long session can also induce mental and physical strain on the operator. Therefore, there is a need to develop automated means of measuring neurite length.

This paper presents and explains an open, fully automated, and free Image J macro able to measure neurite length of differentiating neurons in batches: Cell Analyser in Batch for Neurites (CABaNe).

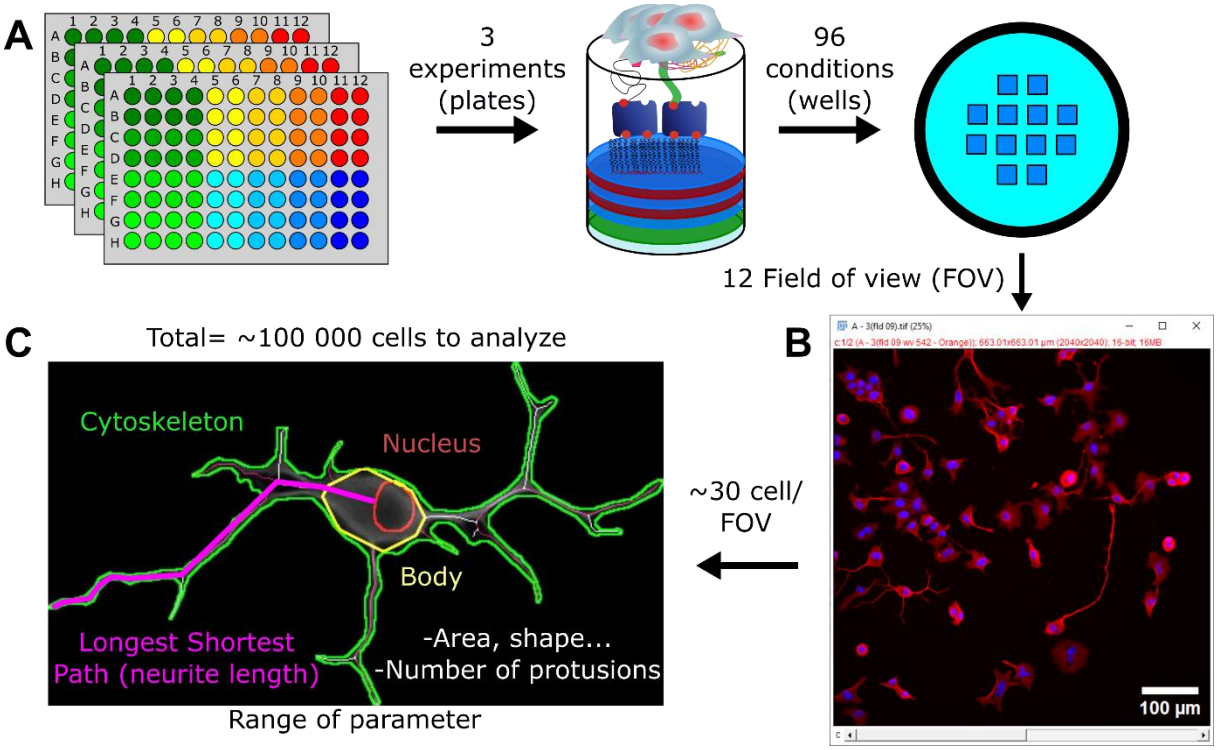

Figure 1: High content analysis of NB differentiation, via neurite length measurement, for multiple experimental conditions studied in parallel, with approximation of the number of cell to analyse, and representation of the parameters of interest. A: visual representation of a high content experiment, with platemap and fields of view. B: example image, composed by the merging of cell channel (Beta tubulin immunostaining, red), and a nuclei channel (DAPI, blue). C: visual representation of main parameters extracted from each cells in each images by CABaNe, as well as an example of verification image produced during the analysis.

2 Material and methods

2.1 What is the solution we bring?

CABaNe automatically analyses fluorescent images of cells, extracting the length of the longest neurite for each cell on a field of view, among other parameters. The software facilitates the study of neuronal cell differentiation in high content settings, using low resolution images, but can also analyse single images at higher magnification/resolution and the protrusions of other cell types.

This macro allows analysing large batches of images, without any operator inputs apart at the initialization step. The macro has been designed to analyse multiple conditions in succession, and to save results for each condition and for each single cell. Most importantly, past initial parameter set up, it is entirely automated therefore limiting human bias, and it can be run on a computer and left for analysis.

While we implemented an option to analyse the nuclei intensity signal from a third channel, it will not be developed here: it is functional, works in single cell to link neurite length and intensity of signal, but we decided on focusing on the morphological analysis.

Overall, three main steps can be identified. The initialisation, during which metadata are created and the images prepared. The processing, in which images are going through filters and operations to isolate the shapes of interest, namely the nucleus, the body, the whole cell, and the skeleton. Finally, the analysis, in which the data is extracted and stored, in .csv tables. Verification images and metadata of parameters used are automatically provided, for each condition.

The readout provided includes, but is not limited to, single cell neurite length, total cell, nuclei, and body area, shape indicators, and number of branches (Figure 1C). Results are produced in .csv, one for each single cell, and one per condition. CABaNe analyses a combination of the nucleus and cell images of a given condition. A graphical interface, presented in Figure 2 was created, to ease the use of CABaNe, but also to give to the user insights on the parameters used.

When developing this macro, we had to choose between filtering using machine learning or traditional rule based programing. We selected Trainable Weka Segmentation tool (Arganda-Carreras et al., 2017; Wasswa et al., 2019) for its well established in ImageJ bibliography. However, machine learning in our application faced limitation: our limited computer power, problems with cells and background heterogeneity, which lead to a lack of flexibility, the lack of transparency in machine learning parameters and overall the intricacy of machine learning for novice users. All these aspects led us to revert to what we define here as Rule-Based Programming (RBP), which can be described as a set of filters from ImageJ, using neither deep or machine learning.

The user has wide control over the parameters; however, it should be noted that the default values have been calibrated to work with our own images, such as the one displayed in Figure 1B.

The main menu is illustrated in Figure 2. The parameters have been divided into 3 categories: Initialisation (Figure 2A) selection of the operating parameters; Processing (Figure 2B) to set the filters that will be applied during the analysis; and Analysis (Figure 2C) to select the type of measurements that will be done.

A user guide is available in extended data Figure 2-1, as well as basic trouble shooting in extended data Figure 2-2. The code itself is in extended data 1.

2.1.1 Initialisation

In Initialisation, the operator can select the folder where the images to be tested are located, and the folder to save the analysed images. This must contain one folder per condition, we advise following the architecture displayed in extended data Figure 2-3. To identify which channel correspond to the nuclei or the cytoskeleton of the cells, a character string only present in the channel of interest has to be inputted. The macro will search it in the image names when they get opened. Then, the type of operation to perform can be defined under the “chose operation” drop-down” menu. The three types of operation are “Analyse”, “Sort”, and “Test_filter”. A fourth option, “Sort_then_analyse”, is the combination of “Sort”, followed by “Analyse”.

Neurite analysis is started if the “Analyze” option in “chose operation” is selected. “Sort” allows to pre-sort images using defined parameters, to remove images without analysable cells and store them in a verification folder, and to store valid images in a specific condition folder if the image name format is compatible, or in a general folder if not. To be compatible with automated sorting, images must be named following template: “X - Y(info).tif”, were X is the first condition identifier and Y the second, for example A-3(green).tif. “test_filter” allows to dynamically test filter values, in case the default parameters do not work for a set of images, e.g. images with a strong background or artefacts. This tool provides visual feedback on the filter parameters. The operator can select the number of images to test per condition. “Sort_then_analyse” is to start “sort” and “analyse” in sequence, if the operator wants to process large amounts of data without human intervention.

To guarantee optimal cell detection in defined conditions we suggest to test images using ”test_filter” . Then, using the ”sort” option if some images potentially do not wield cells to analyze, then “test_filter” options to finely tune the filtering parameter and use the selected values for the analysis.

When relevant, the user may restrict the analysis to a range of continuous conditions within the list, stating the start and end conditions to be analysed.

2.1.2 Processing

In Processing, the operator can adjust the filter parameters for analysis. While “test_filter” is running, the macro will loop back to the menu and save the parameter used, allowing to fine tune the macro. Unsharp mask is used to reduce noise, but can sometime produce artefacts in empty images;

weigth of this mask can be tuned, with values closer to 0 reducing the strength. The radius is set per analysis. It is recommended to separate touching cells using the nuclei, but separation using body identification can be used if the nuclei channel is missing. In this situation, the operator can use any channel with a good body imaging to replace the nuclei channel. Thresholds are used to detect pixel values, to differentiate “objects pixels” from the “background pixels”. “Minimum” is the most discriminating threshold, reducing noise and potentially giving the most accurate shape contours, but also the least robust. In contrast, “Default” threshold is more robust but has poor discrimination, and may reduce detail or separation between cells. “Triangle” threshold is in between. Size of the different element can be slightly adjusted if during testing the operator notices the segmentation differ from the true size of the bodies, nuclei, or whole cells. The operator can also select parameters to discriminate the size and shape of cell and nuclei, mainly to remove small artefacts.

2.1.3 Analysis

Finally, in Analysis, the operator can chose whether the macro would measure neurites from the body of the cell, or from the nucleus.

Additionally, we provide a Python script (in Extended Data 2), using Panda, a data analysis library, to summarise and analyse the tables generated as results during the analysis phase.

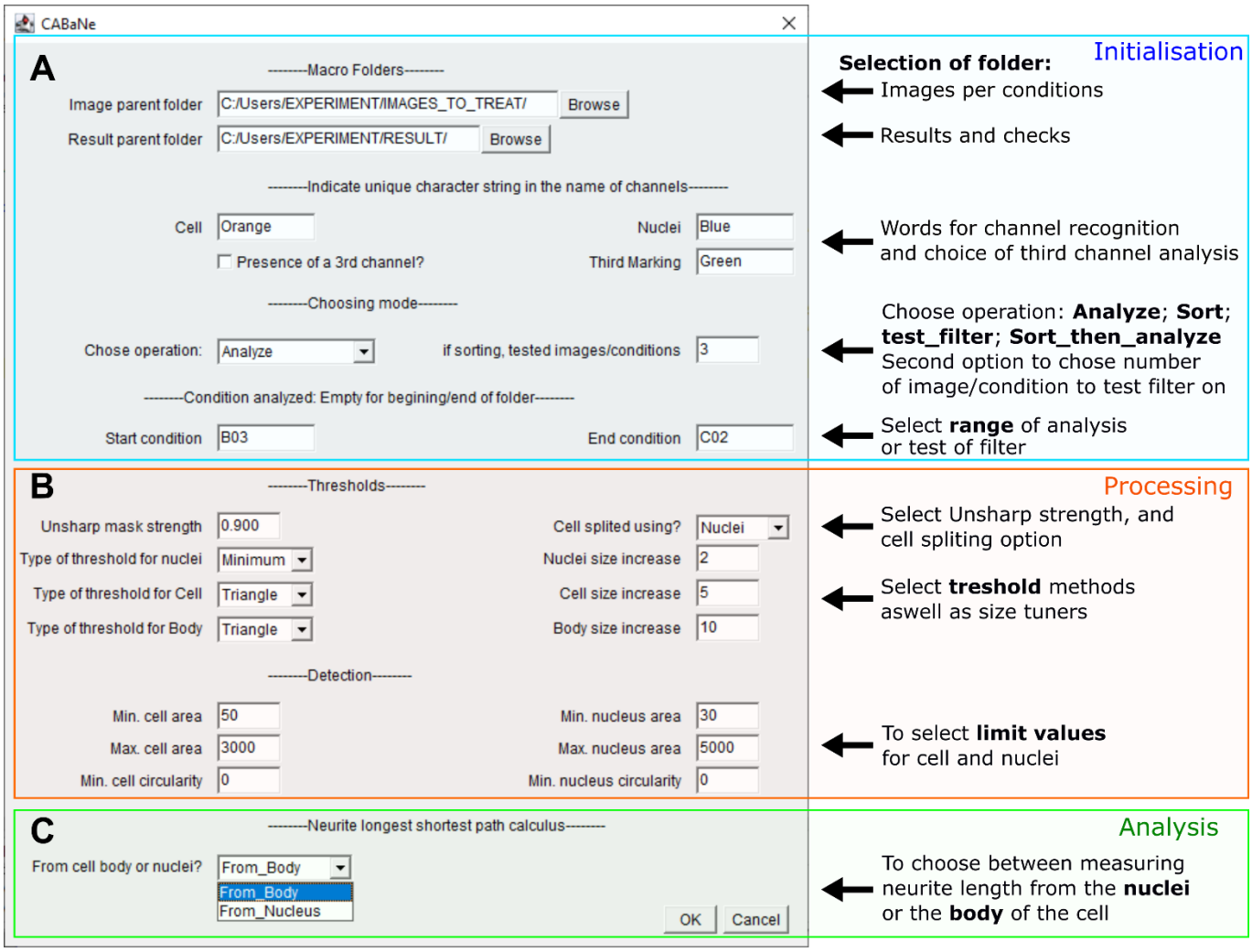

Figure 2: CABaNe graphical interface with explanation of parameters of interest. Values are either default or examples. Interface is divided in 3 parts. A: initialisation, to define the folders of interest, the type of operation, and the experimental conditions used. B: Processing, to change the parameter related to the image processing. C: Analysis, to choose the calculus mean for neurite length. A user guide is available in extended data Figure 2-1, a trouble shouting guide in extended data Figure 2-2, and an overview of the architecture in extended data Figure 2-3.

Requirement and recommendation for running CABaNe are provided in Table 2. During acquisition of cells, it is recommended to acquire two channels: one for the cytoskeleton and one for the nuclei. They have to be separated files, not stacked. In the image folder images for each field of view must be following each other alphabetically, and have a distinctive string of characters for each channel.

Imaging is done with immunofluorescent labelling for better cell segmentation. Background noise is ideally low to permit a good analysis of the neurites, even if filters are applied to reduce it. The best results were obtained at 20x magnification, numerical aperture 0.75, but we could also run the macro at 10x. At 4x the resolution may be too low and at magnifications greater than 20x the cells are likely to touch the edges. Regarding cell contact, even if the macro separates touching cells, better results are obtained when cells do not touch. Otherwise, neurites may be shared between two or more cells. For cells that are already well differentiated and forming networks, performance is reduced for the same reason.

There is no limit to the number of images that can be analysed at once. We were able to do hundreds of images at a time. Only time is a limit. In average, with the computer we used, which hardware is described in Table 2, we accounted for approximately 30 seconds per image. This value depends on the number of cells on the image and the hardware used, with time values halved when we used a better computer.

Table 2: Pre-requisite for CABaNe optimal functioning. The tables recapitulate, for a range of parameters, the specifications for CABaNe to work in an optimal manner, as well as the importance of following the specifications. When a parameter is required, failure to follow the specification will result in failure of the macro. When highly recommended, other ways could work, but were either not tested or tested with low success. Recommended parameters are advices to get quality analysis. Low are for informative purposes.

Parameter	**Specification**	**Importance**
Channels	2, one for cytoskeleton and one for Nuclei.	Required
Format	.tif	Required
Order of images in files (alphabetically)	1st nuclei first, 1st cytoskeleton, then 2nd nuclei, 2nd cytoskeleton, etc…	Required
Imaging	Fluorescent imaging, on glass slide or plates, or low auto-fluorescence plastics support.	Highly recommended
Objective	20x. Higher works but is more susceptible to neurite touching the edges of the FOV. (tested with 0.75 numerical aperture)	Highly recommended, 10x is possible
Type of experiment	Differentiating cells, not yet forming complex network	Highly recommended
Number of cells/images	No upper limit, larger number will take additional computation time. <50 cells is optimal, with cells of comparable size as the example.	recommended
Cell contact	Non-contacting cells	recommended
Number of images	Unlimited	Low
Hardware	SSD hard drive should make running faster. Tested with a i5-8265U 1.60 GHz, 8 Go RAM, SSD. Macro is 33 Ko.	Low
Software	FiJi 2.14 Plugin: MorpholibJ 1.6.3 Tested on Windows 10	Required plugin, and version are the one we've tested

2.2 Description of the macro

Workflow of the analysis loop of the macro is presented in Figure 3A. We will present the different steps occurring after running the code.

2.2.1 Initialisation

First, all other windows are closed, results and ROI manager are reset, to insure nothing interferes with CABaNe. Then, the parent image folder, the folder containing all the image folders, is scanned to find the first condition to analyse, as inputted in “Start condition”. Once the condition is found, a loop is started to analyse all the conditions up to the end condition. Two images of the folder are opened (Figure 3B and C), and each image will be identified and renamed using the character string unique for each channel. The folder architecture for each condition is created for the given image (see extended data Figure 2-3), and if it is the first one of the conditions, a metadata file is generated. The metadata contains the values of the parameters used and the names of the folders.

2.2.2 Processing

The cell image is processed, first with “auto-contrast” to normalize the images, then with an “unsharp” mask, to reduce background noise, and to make cell shape more contrasting to the background. It is then binarised using a threshold, and further processings, closing and dilatation, are applied to remove artefacts and smooth the cells (Figure 3D). The global structure of the cells is then segmented using “Analyze Particles” (Figure 3E). A similar background cleaning process is applied to the nuclei images. Nucleus are then segmented following operator inputs values, and the ROI of their position is saved in an external file (Figure 3F). Nucleus touching the edges of the images won’t be analysed. Nucleus will be used as references to identify cells position. An unmodified version of the cell image is then treated with minimum and median filters, using user’s parameters, to segment cell bodies. A “Watershed” step is applied, using the nucleus as a seed, and applied to extract the body of each identified cell (Figure 3G). The body ROIs are saved in a separate file. The segmented cell image is also treated with a watershed using nucleus as markers, and user’s parameters on cell size and shape limits. Cells touching the edges of the image are excluded (Figure 3H). Cells’ ROIs are also saved in a separate file.

Cell ROI are then skeletonized (Figure 3I), and either bodies or nucleus are removed from their structure, depending on chosen parameters.

2.2.3 Analysis

An image is generated from the base image, cell, nuclei, body ROI and the skeleton, both for each Field Of View (FOV) (Figure 3J) and for each cell (Figure 3K), as a verification mean. Then, for each cell, data, including area, shape and others, are extracted from each ROI and from the skeleton, from which the longest shortest path, number of branches and other parameters will be saved into the summary .csv file, while the values for each branch is saved for the single cell analysis .csv files.

Once this field of view analysis is complete, the images are closed, and the next pair of images is opened. The summary .csv is kept and amended for each image, until the condition is over. Finally, all output (images, tables, RoiManager) are closed or reset, and the analysis for the next condition begins, until all conditions have been completed or the condition specified by the operator has been reached.

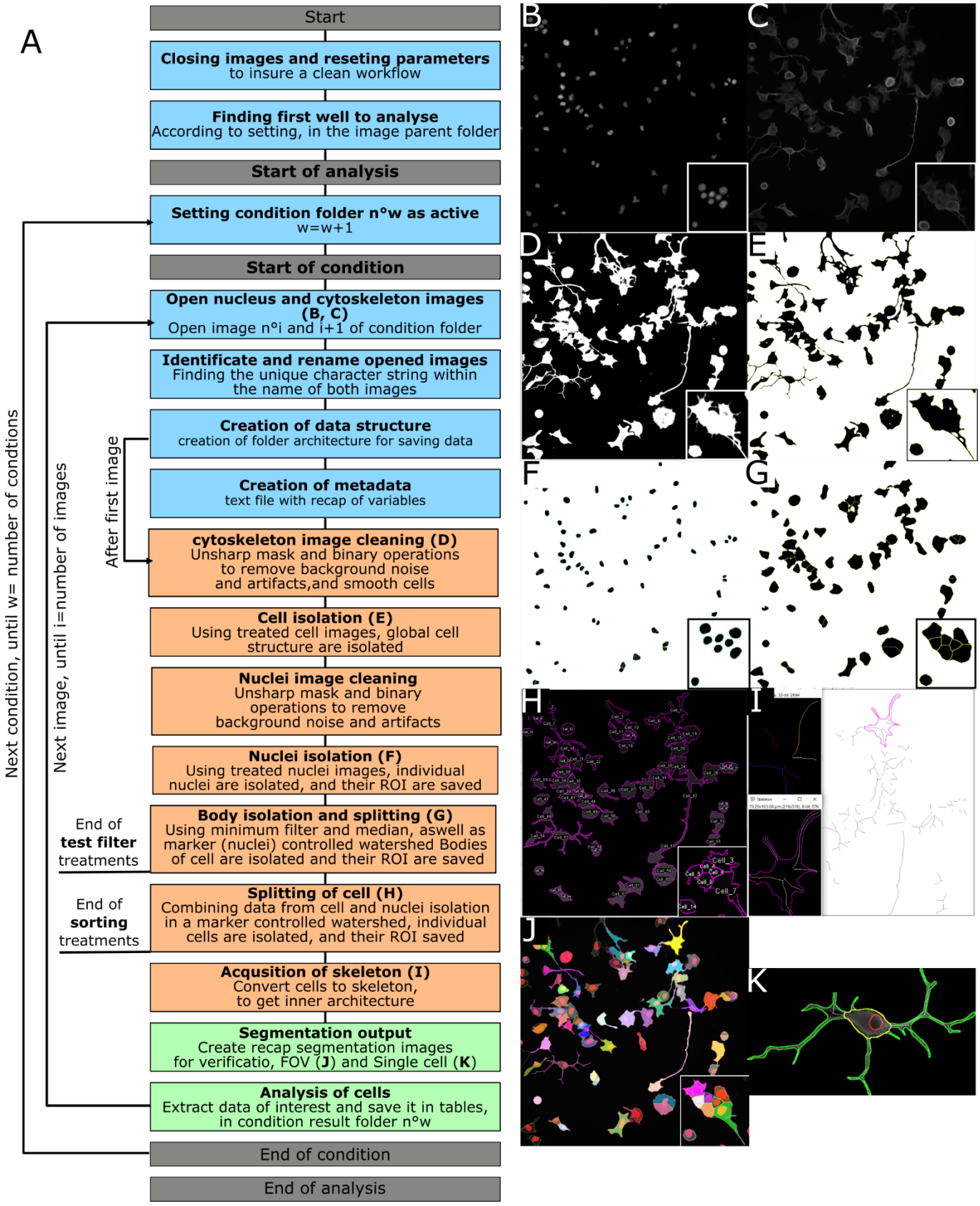

Figure 3: Overview of "Analyze" operation workflow, with example images. Blue is for initialisation steps, orange for image processing, and green for analysis. A: Process diagram. Other letters are in the workflow. E, F, G, H, J images have labels in imageJ, but we only kept them in H for visibility sake. A zoomed insert of analysed cells is displayed on the bottom right corner of images. Images have a size of 663x 663 µm.

2.2.4 “Test_filter” description

If selected in the main menu and started, CABaNe will carry on with the “test_filter” operation. It is used to test chosen parameters on cell, nucleus, and body detection. It follows the same step as the analysis seen in Figure 3, up to step G, but without saving the ROIs. The workflow of this operation is presented in Figure 4A. After splitting the global cytoskeleton shape, nucleus, and body, an image is created by merging and colouring the base image (black and white), the nucleus (blue), as well as the detected contour of cytoskeleton (red), body (yellow) and nucleus (green). All images, with the merging, is then displayed, as in the example presented in Figure 4B. The operator can check the selected parameter and click “OK” in the explanatory window, presented in Figure 4C, at which point the macro will revert to batch mode and close all images. As many images per condition will be tested as selected, until the end of the conditions. The macro will then loop back to the main menu, for the user to change parameters if necessary, or to start the main analysis.

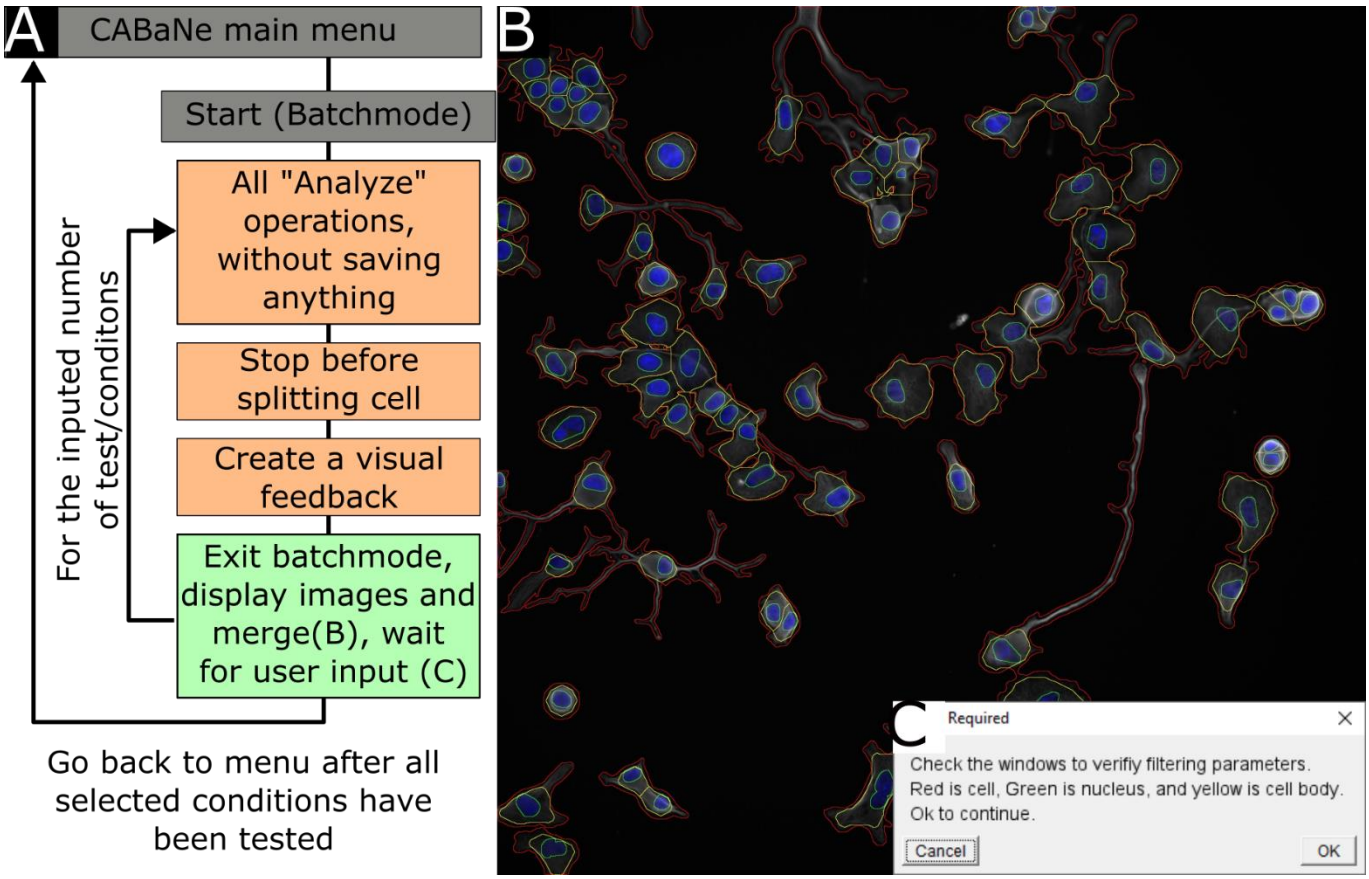

Figure 4: Workflow of “Test_Filter” operation (A), with example of visual feedback provided (B) and waiting window (C)

2.2.5 Code accessibility

The code is fully accessible in the Extended Data 1 and 2, or on Github, at this address:

https://github.com/NathanThibieroz/Cell_Analyser_in_Batch_for_Neurites_-CABaNe-

/blob/c83accf8b62e0f930abca69ec0b3409a160b30ba/CABaNe_V31.ijm . Anyone is free to use or modify the code, as long as it is not for commercial use.

2.3 Macro output

For each condition, the "_Skeleton_analysis_summary.csv" from the last field of view will be a summary of all single cells analyses of the conditions. Each row is a cell. There are a variety of parameters extracted from the analysis, presented in Table 3.

The main parameters of interest are: the longest shortest path, which is related to neurite length; the cell, nucleus and body area; and the number of segments, which represents the number of neurites. Note not length-based discrimination is being carried out.

Table 3: Result variable of CABaNe found in the result .csv, with name of the parameter, unit and a description. Bold parameters are the one used here as means of comparisons between the various techniques.

Parameter	**Unit**	**Description**
Skeleton_analysis_summary		
Label		Name of cell analysed (number in image)
Cell_area	μm^2	Area of cell derived from cytoskeleton
Nucleus_area	μm^2	Area of Nucleus
Cytoplasm_area	μm^2	Cell_area – Nucleus_area
Body_area	μm^2	Area of body
nSegments		Number of paths starting from nucleus or body, going to an end point, an end of protrusion.
nBranches		Connection between endpoints, junctions, or both
nJunctions		Number of junction between branches
nEndPoints		Number of terminal points of protrusions
Longest shortest path	µm	Length of the shortest path to the most far away end point
Major	µm	Distance between the two furthest points of the cell
Minor	µm	Distance between the two closest point of the cell
Angle	°	Angle between major and x axis
Feret	µm	Longest distance between two points of the selection, within its perimeter.
Circ.		4π*area/perimeter^2. A value of 1.0 indicates a perfect circle. As the value approaches 0.0, it indicates an increasingly elongated shape
AR (Aspect Ratio)		Major/minor
Round		4*area/(π*major_axis^2)
Solidity		area/convex area

2.4 WEKA-based Machine learning

For our first test, in machine learning, we used Trainable Weka segmentation (Arganda-Carreras et al., 2017; Wasswa et al., 2019). A classifier was created by manually identifying the cytoskeleton, nucleus, and background, using a reference image for training, different from the analysed images. The classifier was then saved and used in the macro, replacing the RBP filters used on the cells and nuclei channels, replacing the cleaning of nuclei and cell images, as well as their isolation. Once overall cell shape and

nuclei are detected, the next steps are the same for both machine learning and RBP macro, including the "watershed" to split the cells.

2.5 Cell culture and fixation and immuno-staining

In order to test CABaNe on cell images, we cultured N2A cells (NB) in 75 cm^2 in DMEM (1X) + GlutaMAX (1 g/L + Pyruvate) (Dulbecco's Modified Eagle Medium) supplemented with 10% FBS medium, 1% non-essential amino acids, and 1% Antibiotic-Antimycotic. Cells were kept in low passage (<10). After passing and sufficient growth, cells were starved (replacement of medium with a medium without FBS) for 4h. For the single image analysis, cells were then detached and seeded in a 10-well slide from Greiner, ref 543079, in 2% medium. Cell density was 250 cell/well. Cells were cultivated for 7 days, and new medium was added each 48h. For the batch analysis, cells were seeded in 3 different occasion, in similar passages, in 96 well plate from Greiner, ref 655986. Cell were seeded at 125, 250, and 500 cell/well, in a 10% medium. After 72h, medium was replaced with 0 or 10% serum medium. Medium was added once at day 5. Both experiments were then fixed in 4% paraformaldehyde, at day 7, and marked with mouse anti beta-tubulin, from Sigma, coupled with goat anti mouse Alexa Fluor plus 555, from Invitrogen, for cytoskeleton, and DAPI for nucleus.

Image acquisition was carried out in high content using an INCELL 2500 HS, with 12 FOV per well, ordered around the center of each wells. Used objective was 20x, of 0.75 numerical aperture, resulting in 663x663 µm images. Channels acquired were the blue, excitation/emission wavelength of 390/432.5 nm, and the orange, excitation/emission wavelength of 542/587nm, with an exposition of 20 and 50 ms.

Experiment was repeated 3 times, on different days.

2.6 Statistical analysis

Statistical analysis was carried out using Origin. Sample were determined not to have a normal distribution. Datasets are independent. Thus, we decided on using the Mann-Whitney-Wilcoxon test, with $p=0.05$.

2.7 Machine learning, RBP, assisted manual comparison

We decided to create a classifier to compare classical image processing filters, the machine learning approach, both using the macro, and analysis done manually using NeuronJ, which we will here call the

manual method, to be considered as the reference method to analyse neurite length. We decided to compare selected, specific cell parameters, and the analysis time, as it is important for high content analysis. Time is manually measured, from the moment the macro, CABaNe or NeuronJ is started, to when the results are saved. We also decided to focus our comparison analysis on four reference images to compare a set of parameters, all from the same condition.

Manual analysis of parameters other than neurite length was carried out using basic ImageJ tools.

Cell limit parameter, here min/max area of cell and nucleus, were set to be the same for the machine learning and RBP. RBP was set to default values, as they are set in Figure 2, with the exception of the cell threshold that was set to “minimum” due to the good quality of the image in the 10 well slide. The 96 well plate, with plastic bottoms, required the following threshold: “Default” for Nuclei, “Mean” for Cell, and “Triangle” for Body. Cell touching the edges are excluded.

During first tests, accepted cell size was 200 to 3000 μm^2 and nucleus size from 50 to 5000 μm^2, but minimum nuclei size was reduced to 30 μm^2 because the machine learning macro severely under estimate nucleus outlines, and the limit of 50 μm^2 led to recognition of very few cells. Therefore, for machine learning, we had to adapt the parameters for this image to get optimal results. All images were analysed using the same parameters, and by the same operator.

3 Result

3.1 Manual, RBP and machine learning analysis comparison

The same image has been analysed using the WEKA-based machine learning (**Erreur ! Source du renvoi introuvable.**A), RBP (**Erreur ! Source du renvoi introuvable.**B), and a manual analysis using NeuronJ (**Erreur ! Source du renvoi introuvable.**C), to compare these three methods. For machine learning and the RBP processing, each colour corresponds to a cell, and black cells are not analysed. Nucleus are in red but sometime are not visible due to colour incompatibilities. In the machine learning example, even if cells are globally accounted, and detached from the background, splitting between cells is inaccurate, creating straight lines between cells. Too many nuclei are accounted, creating an overabundance of cells. In addition, nuclei size and positions are incorrect, which can bear an influence on calculated neurite length. Different cells outlines are often merged together. For RBP, overall, all cells

of interest are accounted. Elongated neurite end up split halfway, as highlighted by the red arrows in **Erreur ! Source du renvoi introuvable.**B. Apart from this specific situation, cells seem properly segmented, even when touching. Lonely cells are perfectly segmented. The manual measurement highlights the quality of neurite tracing in every cells, even in overlapping cells. However, the monochrome image, with no colour or nucleus marking, increase the difficulty of neurite placement. All cell not touching edges are accounted.

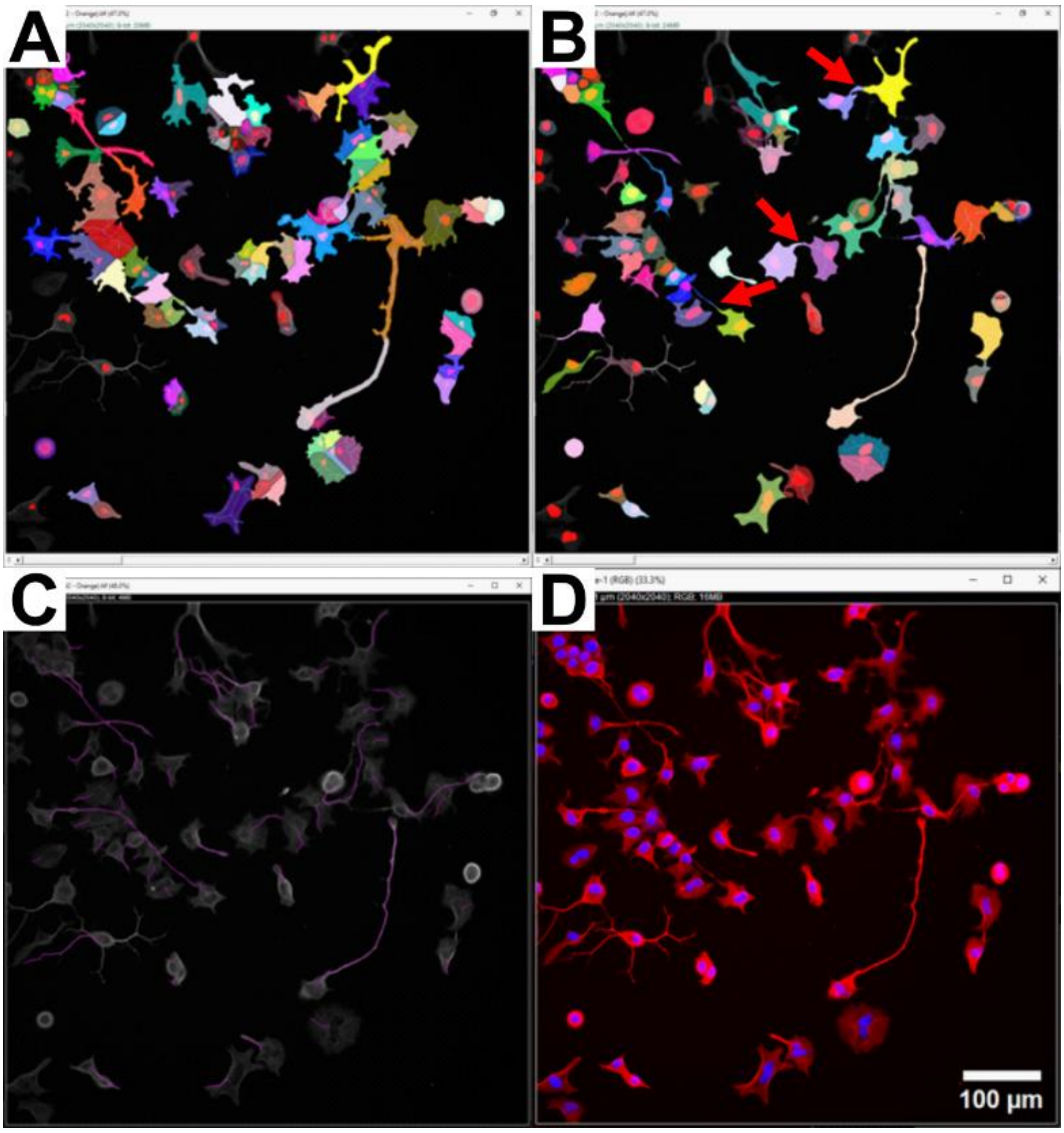

Figure 5: 3 different analysis of the same image. A and B are done using our macro, and C using NeuronJ. A is the result of a machine learning processing using WEKA, and B of a RBP. More detail on cell and nucleus fitting is available in extended data Figure 5-1.

Table 4 contains the selected measured parameters needed to compare the three different methods. Neurite length was slightly higher in the RBP analysis, yet not statistically meaningful, with an average increase of 6.4% compared to manual measurement. The machine learning filtering produced higher neurite length values, with an average increase of 21.5% of the manual measurement. In terms of the number of cells analysed, machine learning filtering accounted for 255 out of the 223 cells present in the image, or an increase of 14,3% of real value, due to nuclei fragmentation. RBP and manual measurement accounted for 233 and 210, or a variation of 4.5% and -5.8% of the cells. RBP in macro was the fastest, roughly taking 40% less time than the machine learning-based method, and being more than 6 times faster than manual measurements, if only considering neurite length measurement, and 27 time faster for the whole analysis.

Both automated analysis provided more analysis data in addition to the average neurite lengths. We consider the RBP technique to be exact on nuclei area; it was the technique that wielded the best fitting result (see extended data Figure 5-1). We selected some parameters of interest: Cell area, body area, nucleus area, circularity, and number of segments per cell. Body area could not be obtained by machine learning, as there is no distinction between the body and the cell: detection would require additional filtering and a different classifier. Overall, cell and nuclei area were under estimated in machine learning when compared to both RBP and manual measurement. This is especially true for the nuclei area, which is reduced by 54.4% of the exact value. Manual measurement was lower than the RBP measurement. Mean cell area stays close to manual measurements for both RBP and machine learning, with respective difference in values of -3% and 0.4%. Circularity is similar in all conditions. The number of segment in RBP is higher than in the machine learning approach.

Figure 6 provides the results obtained by the three different methods for several parameters: the neurite length analysis, nuclei area, and cell area. All techniques produce somewhat similar results, except from the nuclei area with machine learning. However, machine learning results in neurite length are statistically different from the manual measurement and RBP, while RBP is not. All values are statistically different when it comes to nuclei area.

Table 4: Result of different analysis on the same image. Machine learning based algorithm, RBP based algorithm, and manual measurement were compared, based on total time required for analysis, from the starting of the macros, to saving of the results, as well as a range of parameters of interests. Manual did not permit having the average segment number per cell using NeuronJ.

Analysis	Machine learning	RBP	Manual
Total analysis time	10 min 09	6 min 02	162 min 55 (36 min 37 for neurites)
Per image	2 min 32	1 min 30	41 min 29 (9 min 9 for neurites)
Total number of cells (hand: 223)	255 ± 21,19	233 ± 8,22	210 ± 15,13
Average neurite length (µm)	41,9 ± 28,8	36,7 ± 32,6	34,5 ± 37,2
Average Cell area (µm²)	808 ± 512	787 ± 467	811 ± 603
Average Body area (µm²)	N/A	672 ± 515	553 ± 408
Average Nucleus area (µm²)	73 ± 40	160 ± 70	140 ± 77
Circularity	0,45 ± 0,18	0,44 ± 0,21	0,47 ± 0,27
Average segment number per cell	2,58 ± 1,09	2,77 ± 1,18	N/A

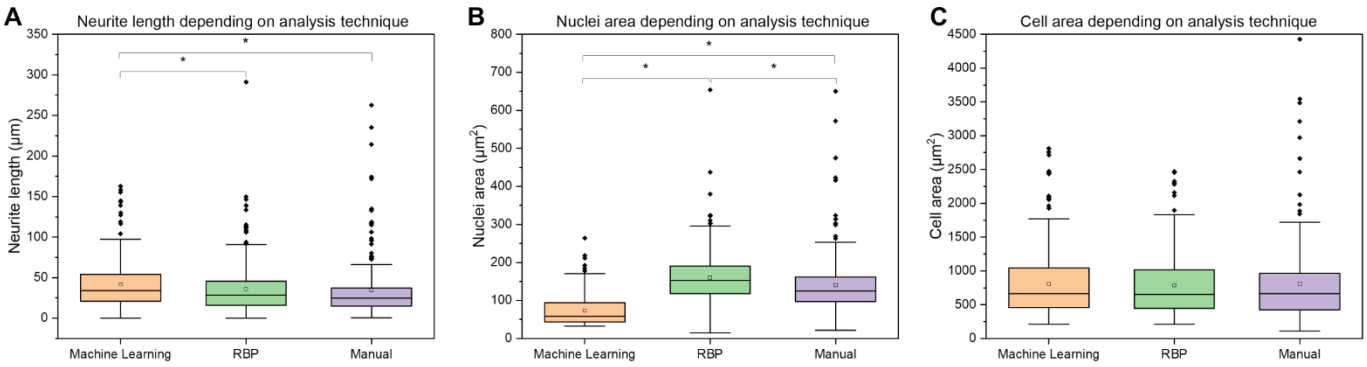

*Figure 6: Result of the analysis using three different methods, the machine learning, the Rule Based Programing, and manual anlaysis, using neuronJ and basic imageJ tools, on image presented in figure 5, and three other of the same condition. P=0.05. Normality rejected. Independent sample. Probality test runned: Mann-Whitley. A: Neurite length (in µm), measured from nucleus. B: Nuclei area. C: Cell area. *p<0.05.*

3.2 Batch analysis

We experimentally tested the macro on N2A cells exposed to starving and different percentage of serum in medium. Neurite length was measured on every single cell we imaged. Experiment is composed of 12 wells, grouping 6 conditions, with 12 FOV per well, for a total of 144 FOV. In Figure 7, we added a representative image for each condition. From the image, we observe higher average cell area and neurite length than in the 0% conditions.

We then applied the macro to an entire experiment. We sorted the images using CABaNe, and analysed all the data at once. Results are summarized in Figure 8. Number of cells, Figure 8A, increases with initial cell density, but surprisingly only with the highest starting value. Neurite length, calculated with the longest shortest path coming from the body of each cell, Figure 8B, are overall higher in the 0% conditions values than their 10% counterparts. Relative body area/cell area ratio is used as a secondary mean to assess the quantity of prolongations, in single cell. The lower is the ratio, the more the cell have elements outside of

the body, such as neurites. Following this analysis, with data in Figure 8C, all 0% conditions have lower values and a wider interval (toward low values) for ratio. Some point are higher than 1, caused by artefacts in the analysis. This data supports the results of neurite length from nuclei. Cell area, Figure 8D, decreases in relation with the overall number of cells. Nuclei area is superior in the 10% condition in regards to their 0% counterpart (Figure 8E). While body area is statistically higher in the 10% conditions for the 100 and 500 cell/well conditions, the difference is minimal.

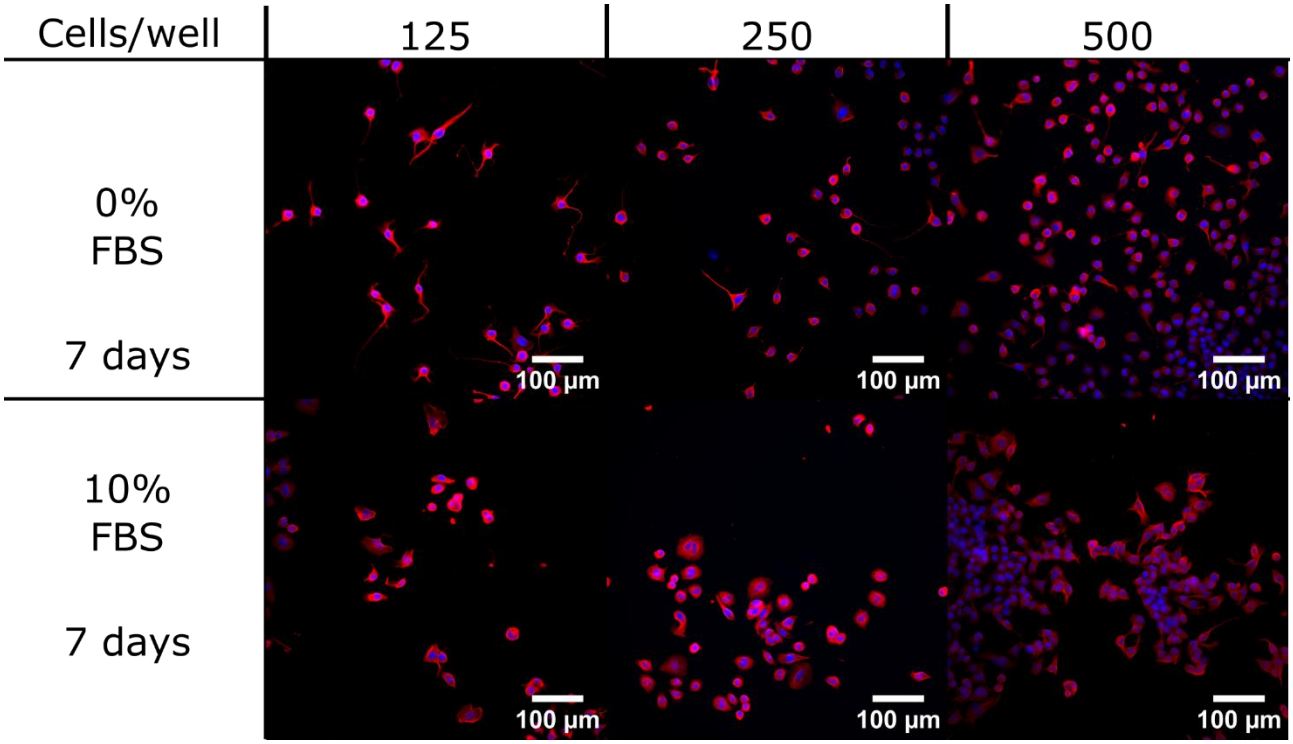

Figure 7: Merged DAPI and Beta tubulin channel, one image per condition.

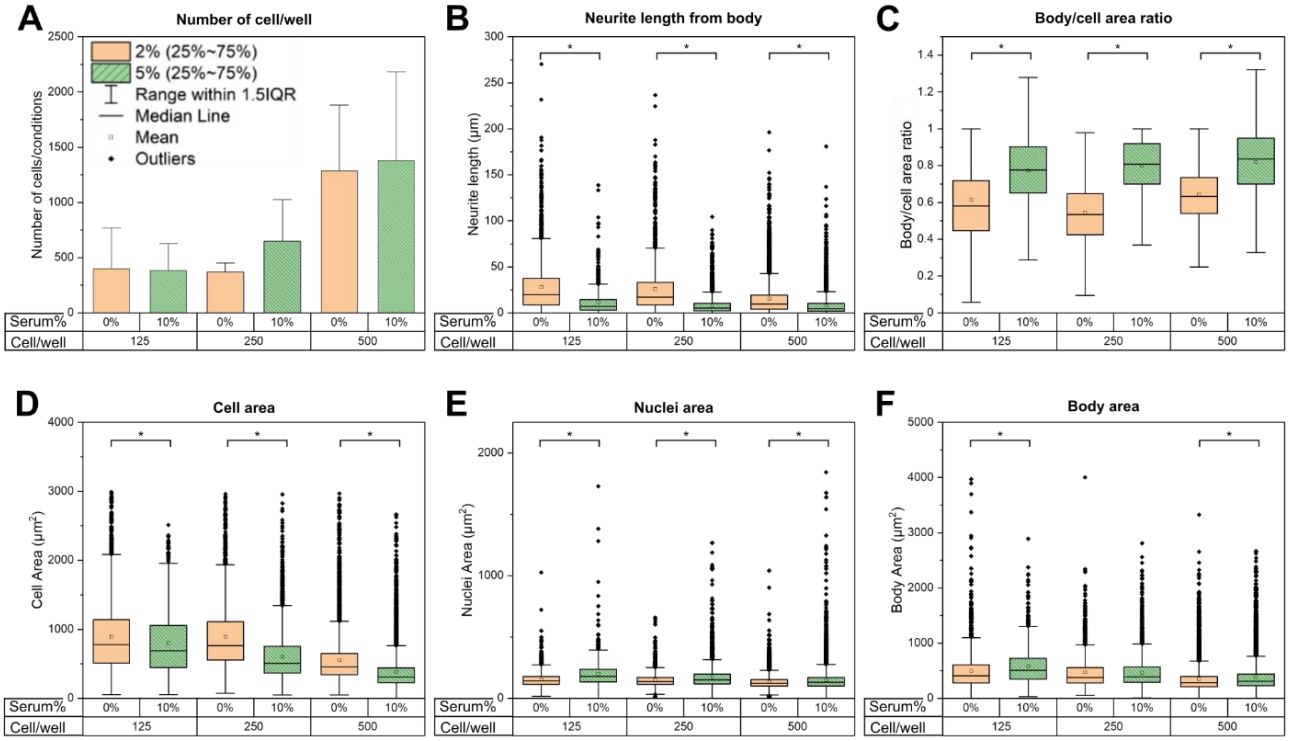

*Figure 8: Data extracted from data set using CABaNe. A: average number of cell per well, per condition. B: Single cell box plot of nuclei area per condition. C: single cell body area D: Single cell whole area E: Single cell body area/cell area ratio per condition. F: Single cell neurite length from nuclei, per condition. G: Single cell neurite length from body, per condition. N=3 experiments, realised on different days. *p<0.05*

4 Discussion

In this paper we have developed a novel open source macro based on RBP to analyse cellular morphological parameters, including neurites length. Our macro named CABaNe, can be used on a wide range of images, when used in accordance to the limitation and eventually adapting the parameters. This technique is based on RBP which is based on traditional filtering, cell segmentation and particle analysis proved to have quantitative advantages with respect to the reference manual measurement of neurite length and to the machine learning techniques.

When compared with other technique, our RBP macro measured similar values with respect to the manual measurement, while the cellular morphological characteristics measured by machine learning were statistically different. The differences between these methodologies could be attributed to the difficulty of automated techniques to properly segment overlapping cells, especially in 2D. We have restricted our exploration to WEKA segmentation, which relies on the Random Tree Forest Approach. We wanted to test an easy to use procedure that could be fully integrated into an ImageJ macro. We lacked the infrastructure and knowledge to make much more advanced machine or deep learning models, that could have improved the analysis. Obviously, use of external python-based tools could reveal successful for the segmentation step. Its training and integration would however require additional efforts that are not worth, knowing that traditional methods allowed us to reach a decent result in regard to the manual approach. Difference in appreciation of nucleus in ambiguous cases, such as when a cell nucleus is dividing, can further change neurite length, cell counting and other parameters, since each nucleus represent a cell to analyse. Indeed, while an operator could take the decision that a cell with two distinct nuclei, but not yet split, is either a single or two cells for analysis, automated means will always consider, in CABaNe, that there are two cells to analyse. Machine learning also had the disadvantage of being more computationally heavy, and to required re-training when either the background or the cells changed, or to possess a really large training data set accounting for this variability. The training process was also highly manipulator dependant. Manual neurite length measurement was tricky as the nucleus channel cannot be included in the images in the NeuronJ plug-in, making measurement from nucleus, as well as cell identification, harder and less precise. In addition, CABaNe provided more parameters of interest, not all being used here, but than can be of interest for other studies, and which can increase the analytical data obtained from one experiment. The rising energy usage of deep learning is to be considered: RBP, when

applicable, could be a more energy and material efficient way to run analysis, especially for the lab that do not possess the capacity to utilise advanced models.

One parameter not taken into account is the bias of manual measurement, which causes physical and mental strain when analysing large datasets, or the variability operator to operator. Already on four images, nuclei area was statistically different in manual measurement when compared with RBP, and analysis of verification files confirmed a better quality of the RBP segmentation. This error is likely to be found in the other manual analysis as well. We are confident that the precision of the CABaNe, on top of representing a significant time and data gain, would be advantageous compared to the one manual measurement in high content experiments. In this paper we do not aim to reduce the impact of machine learning tools but here, the Weka script we used, based on random tree forest, while being accessible, was not advanced enough given our technical knowledge. In the future, CABaNe could be used to train more advanced machine and deep learning macros, classifying the large amount of data required.

The Python script, while basic, further optimises the analysis, summarizing the results.

From the proof-of-concept that we provided in this article, the discrimination between conditions was successful and the amount of data allows to demonstrate the advantages and the limits of CABaNe. Indeed, while we obtained a good analysis and statistically different results, we observed some limitations when cells became too dense in images, but the results still seemed coherent with observed morphology.

5 Conclusion

In conclusion, we have succeeded in providing an open, free, fully automated, batch-designed ImageJ macro for neurite length analysis of differentiating cells. As far as we know, it is the first tool of this type widely available. The user-friendly interface and automated sorting make it possible for anyone, even without knowledge of ImageJ or neuronal studies, to use it efficiently and without inter-operator variability. Although designed for neurite analysis, the number of parameters extracted by this tool, as presented in Table 3, is also an alternative for other types of analysis as it is not limited to neurite length, and can be used as a FAIR alternative to paid software. Notably, we intend on testing the marker intensity analysis in nuclei, still in a single cell manner.

Furthermore, it can be used as a base for anyone who wants to create a similar algorithm, adapt it to their needs, or just learn it, as it is fully annotated, especially in the case of protrusions in other cancer cells types. We encourage sharing with the community in these cases.

Acknowledgment

Thanks to AGASSE Fabienne (Saudou team, GIN, U1216) and EGLES Christophe (BIOMAT team, PBS, UMR 6270) for their suggestion regarding the macro functionalities. This project has been financed by Ligue contre le cancer Isére. NT is a recipient of Pfizer Innovation fellowship. EM is part of GDR 2000 BIOMIM and GDR B2i. EM acknowledges GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). This work was co-funded by the ANR (GlyCON, grant number [ANR-19-CE13-0031-01 PRCI] and by the “Investissements d’avenir” program Glyco@Alps, grant number [ANR-15-IDEX-02].

Credit

NT, FC, CP, EM Designed research; NT, FC, LM, AS, PM performed research; NT, AS analysed data; NT,FC, CP, EM wrote the paper

BIBLIOGRAHIC REFERENCES

Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, Sebastian Seung H (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 33:2424–2426.

Barnat M, Le Friec J, Benstaali C, Humbert S (2017) Huntingtin-Mediated Multipolar-Bipolar Transition of Newborn Cortical Neurons Is Critical for Their Postnatal Neuronal Morphology. Neuron 93:99–114.

Beier T, Pape C, Rahaman N, Prange T, Berg S, Bock DD, Cardona A, Knott GW, Plaza SM, Scheffer LK, Koethe U, Kreshuk A, Hamprecht FA (2017) Multicut brings automated neurite segmentation closer to human performance. Nat Methods 14:101–102.

Crutcher KA, Collins F (1982) In Vitro Evidence for Two Distinct Hippocampal Growth Factors: Basis of Neuronal Plasticity? Science 217:67–68.

Davidoff AM (2012) Neuroblastoma. Semin Pediatr Surg 21:2–14.

dos Santos MG, Gomes JR, Costa MDM (2023) Methods used to achieve different levels of the neuronal differentiation process in SH-SY5Y and Neuro2a cell lines: An integrative review. Cell Biol Int 47:1883–1894.

El Merhie A, Salerno M, Toccafondi C, Dante S (2019) Neuronal-like response of N2a living cells to nanoporous patterns of thin supported anodic alumina. Colloids Surf B Biointerfaces 178:32–37.

Gao X, Han Z, Huang C, Lei H, Li G, Chen L, Feng D, Zhou Z, Shi Q, Cheng L, Zhou X (2022) An anti-inflammatory and neuroprotective biomimetic nanoplatform for repairing spinal cord injury. Bioact Mater 18:569.

Gatta G et al. (2014) Childhood cancer survival in Europe 1999–2007: results of EUROCARE-5—a population-based study. Lancet Oncol 15:35–47.

Ghatak S, Dolatabadi N, Trudler D, Zhang X, Wu Y, Mohata M, Ambasudhan R, Talantova M, Lipton SA (2019) Mechanisms of hyperexcitability in Alzheimer’s disease hiPSC-derived neurons and cerebral organoids vs isogenic controls. eLife 8:e50333.

Gilbert J, Man H-Y (2017) Fundamental Elements in Autism: From Neurogenesis and Neurite Growth to Synaptic Plasticity. Front Cell Neurosci 11.

Gurney JG, Ross JA, Wall DA, Bleyer WA, Severson RK, Robison LL (1997) Infant Cancer in the U.S.: Histology-Specific Incidence and Trends, 1973 to 1992. J Pediatr Hematol Oncol 19:428.

Kumar M, Katyal A (2018) Data on retinoic acid and reduced serum concentration induced differentiation of Neuro-2a neuroblastoma cells. Data Brief 21:2435–2440.

Labour M-N, Banc A, Tourrette A, Cunin F, Verdier J-M, Devoisselle J-M, Marcilhac A, Belamie E (2012) Thick collagen-based 3D matrices including growth factors to induce neurite outgrowth. Acta Biomater 8:3302–12.

Lauzon M-A, Drevelle O, Faucheux N (2017) Peptides derived from the knuckle epitope of BMP-9 induce the cholinergic differentiation and inactivate GSk3beta in human SH-SY5Y neuroblastoma cells. Sci Rep 7:4695.

Mohlin SA, Wigerup C, Påhlman S (2011) Neuroblastoma aggressiveness in relation to sympathetic neuronal differentiation stage. Semin Cancer Biol 21:276–282.

Mundhara N, Panda D (2023) β-II tubulin isotype directs stiffness and differentiation of neuroblastoma SH-SY5Y cells. Mol Cell Biochem 478:1961–1971.

Namsi A, Nury T, Hamdouni H, Yammine A, Vejux A, Vervandier-Fasseur D, Latruffe N, Masmoudi-Kouki O, Lizard G (2018) Induction of Neuronal Differentiation of Murine N2a Cells by Two Polyphenols Present in the Mediterranean Diet Mimicking Neurotrophins Activities: Resveratrol and Apigenin. Diseases 6:67.

Ozgun A, Erkoc-Biradlı FZ, Bulut O, Garipcan B (2021) Substrate stiffness effects on SH-SY5Y: The dichotomy of morphology and neuronal behavior. J Biomed Mater Res B Appl Biomater 109:92–101.

Ramm P, Alexandrov Y, Cholewinski A, Cybuch Y, Nadon R, Soltys BJ (2003) Automated screening of neurite outgrowth. J Biomol Screen 8:7–18.

Rumpansuwon K, Prommahom A, Dharmasaroja P (2022) eEF1A2 knockdown impairs neuronal proliferation and inhibits neurite outgrowth of differentiating neurons. NeuroReport 33:336.

Sabelström H, Petri R, Shchors K, Jandial R, Schmidt C, Sacheva R, Masic S, Yuan E, Fenster T, Martinez M, Saxena S, Nicolaides TP, Ilkhanizadeh S, Berger MS, Snyder EY, Weiss WA, Jakobsson J, Persson AI (2019) Driving Neuronal Differentiation through Reversal of an ERK1/2-miR-124-SOX9 Axis Abrogates Glioblastoma Aggressiveness. Cell Rep 28:2064-2079.e11.

Sánchez-Sarasúa S, Meseguer-Beltrán M, García-Díaz C, Beltrán-Bretones MT, ElMlili N, Sánchez-Pérez AM (2022) IRS1 expression in hippocampus is age-dependent and is required for mature spine maintenance and neuritogenesis. Mol Cell Neurosci 118:103693.
Schwab AJ, Ebert AD (2015) Neurite Aggregation and Calcium Dysfunction in iPSC-Derived Sensory Neurons with Parkinson's Disease-Related LRRK2 G2019S Mutation. Stem Cell Rep 5:1039–1052.
Shimizu S, Kondo M, Miyamoto Y, Hayashi M (2002) Foxa (HNF3) Up-regulates Vitronectin Expression during Retinoic Acid-induced Differentiation in Mouse Neuroblastoma Neuro2a Cells. Cell Struct Funct 27:181–8.
Thomson AC, Schuhmann T, de Graaf TA, Sack AT, Rutten BPF, Kenis G (2022) The Effects of Serum Removal on Gene Expression and Morphological Plasticity Markers in Differentiated SH-SY5Y Cells. Cell Mol Neurobiol 42:1829–1839.
Wang D, Lagerstrom R, Sun C, Bishof L, Valotton P, Götte M (2010) HCA-Vision: Automated Neurite Outgrowth Analysis. J Biomol Screen 15:1165–1170.
Wasswa W, Obungoloch J, Basaza-Ejiri AH, Ware A (2019) Automated Segmentation of Nucleus, Cytoplasm and Background of Cervical Cells from Pap-smear Images using a Trainable Pixel Level Classifier In: 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) , Presented at the 2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) pp1–9.
Wiatrak B, Kubis-Kubiak A, Piwowar A, Barg E (2020) PC12 Cell Line: Cell Types, Coating of Culture Vessels, Differentiation and Other Culture Conditions. Cells 9:958.

Legend list

Figures

Figure 1: High content analysis of NB differentiation, via neurite length measurement, for multiple experimental conditions studied in parallel, with approximation of the number of cell to analyse, and representation of the parameters of interest. A: visual representation of a high content experiment, with platemap and fields of view. B: example image, composed by the merging of cell channel (Beta tubulin immunostaining, red), and a nuclei channel (DAPI, blue). C: visual representation of main parameters extracted from each cells in each images by CABaNe, as well as an example of verification image produced during the analysis.

Figure 2: CABaNe graphical interface with explanation of parameters of interest. Values are either default or examples. Interface is divided in 3 parts. A: initialisation, to define the folders of interest, the type of operation, and the experimental conditions used. B: Processing, to change the parameter related to the image processing. C: Analysis, to choose the calculus mean for neurite length. A user guide is available in extended data Figure 2-1, a trouble shouting guide in extended data Figure 2-2, and an overview of the architecture in extended data Figure 2-3.

Figure 3: Overview of “Analyze” operation workflow, with example images. Blue is for initialisation steps, orange for image processing, and green for analysis. A: Process diagram. Other letters are in the workflow. E, F, G, H, J images have labels in imageJ, but we only kept them in H for visibility sake. A zoomed insert of analysed cells is displayed on the bottom right corner of images. Images have a size of 663x 663 µm.

Figure 4: Workflow of “Test_Filter” operation (A), with example of visual feedback provided (B) and waiting window (C)

Figure 5: 3 different analysis of the same image. A and B are done using our macro, and C using NeuronJ. A is the result of a machine learning processing using WEKA, and B of a RBP. More detail on cell and nucleus fitting is available in extended data Figure 5-1.

Figure 6: Result of the analysis using three different methods, the machine learning, the Rule Based Programing, and manual anlaysis, using neuronJ and basic imageJ tools, on image presented in figure 5, and three other of the same condition. P=0.05. Normality rejected. Independent sample. Probality test

runned: Mann-Whitley. A: Neurite length (in µm), measured from nucleus. B: Nuclei area. C: Cell area. *p<0.05.

Figure 7: Merged DAPI and Beta tubulin channel, one image per condition.

Figure 8: Data extracted from data set using CABaNe. A: average number of cell per well, per condition. B: Single cell box plot of nuclei area per condition. C: single cell body area D: Single cell whole area E: Single cell body area/cell area ratio per condition. F: Single cell neurite length from nuclei, per condition. G: Single cell neurite length from body, per condition. N=3 experiments, realised on different days. *p<0.05

Tables

Table 1: Summary of read out in NB cell lines differentiation studies. For neurite length, when technique is not mentioned, it is assumed to be manual. Manual refers to measurement of neurite length using ImageJ or equivalent. When the analysis is assisted with code, but the operation is still manual, it is considered as assisted manual.

Table 2: Pre-requisite for CABaNe optimal functioning. The tables recapitulate, for a range of parameters, the specifications for CABaNe to work in an optimal manner, as well as the importance of following the specifications. When a parameter is required, failure to follow the specification will result in failure of the macro. When highly recommended, other ways could work, but were either not tested or tested with low success. Recommended parameters are advices to get quality analysis. Low are for informative purposes.

Table 3: Result variable of CABaNe found in the result .csv, with name of the parameter, unit and a description. Bold parameters are the one used here as means of comparisons between the various techniques.

Table 4: Result of different analysis on the same image. Machine learning based algorithm, RBP based algorithm, and manual measurement were compared, based on total time required for analysis, from the starting of the macros, to saving of the results, as well as a range of parameters of interests. Manual did not permit having the average segment number per cell using NeuronJ.

Extended data

Extended data 1: CABaNe ImageJ code, to be used as a macro.

Extended data 2: Python code to help with treatment of the data produced by CABaNe.

Figure 2-1: User guide for the use of CABaNe, guiding in more detail through its use, and giving advice and its usage, as well as giving information on data analysis and verification.

Figure 2-2: Trouble shooting guide covering some of the basic possible errors that the user can encounter.

Figure 2-3: Data architecture in macro workflow. Yellow is for folders, green for tables, blue for images and white for texts. The macro will create every folder but the Experiment, Image, and Results, and the conditions if the sorting is not used. This assures a normalisation of the architecture as well as concordance with quality standards.

Figure 5-1: Result of segmentation of nuclei and cells, on the cell image, using the different techniques analysed in the article. RBP stand for rule base programming. Machine learning is WEKA-based and is trained in similar images, and is trained to detect nuclei and cells on the same image. Manual is done using ImageJ, segmentation being done by an operator. All images are 663x663 µm in size.