Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

Graph-Based Unsupervised Segmentation Algorithm for Cultured Neuronal Networks' Structure Characterization and Modeling

Daniel de Santos-Sierra,[1,2]* Irene Sendiña-Nadal,[1,3] Inmaculada Leyva,[1,3] Juan A. Almendral,[1,3] Amir Ayali,[4] Sarit Anava,[4,5] Carmen Sánchez-Ávila,[2] Stefano Boccaletti[6,7]

[1]Centre for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain

[2]Group of Biometrics, Biosignals and Security, Universidad Politécnica de Madrid, Madrid, Spain

[3]Complex Systems Group, Universidad Rey Juan Carlos, Madrid, Spain

[4]Department of Zoology, Tel Aviv University, Tel Aviv, Israel

[5]Department of Neurobiology, Wise Faculty of Life Sciences & Sagol School, Tel Aviv University, Tel Aviv, Israel

[6]Embassy of Italy in Tel Aviv, Tel Aviv, Israel

[7]CNR–Istituto dei Sistemi Complessi, Florence, Italy

Received 16 June 2014; Revised 27 September 2014; Accepted 28 October 2014

Grant sponsor: Spanish Ministerio de Economía y Competitividad, Grant numbers: FIS2012–38949-C03-01 and FIS2013–41057-P

Grant sponsor: Comunidad de Madrid through the R+D Activity Program NEUROTEC-CM (D.S.S.), Grant number: 2010/BMD-2460

Additional Supporting Information may be found in the online version of this article.

Correspondence to: Daniel de Santos-Sierra, Centre for Biomedical Technology, Universidad Politécnica de Madrid, Campus de Montegancedo, Pozuelo de Alarcón 28223, Madrid, Spain. E-mail: daniel.desantos@ctb.upm.es

ISAC
International Society for Advancement of Cytometry

Abstract

Large scale phase-contrast images taken at high resolution through the life of a cultured neuronal network are analyzed by a graph-based unsupervised segmentation algorithm with a very low computational cost, scaling linearly with the image size. The processing automatically retrieves the whole network structure, an object whose mathematical representation is a matrix in which nodes are identified neurons or neurons' clusters, and links are the reconstructed connections between them. The algorithm is also able to extract any other relevant morphological information characterizing neurons and neurites. More importantly, and at variance with other segmentation methods that require fluorescence imaging from immunocytochemistry techniques, our non invasive measures entitle us to perform a longitudinal analysis during the maturation of a single culture. Such an analysis furnishes the way of individuating the main physical processes underlying the self-organization of the neurons' ensemble into a complex network, and drives the formulation of a phenomenological model yet able to describe qualitatively the overall scenario observed during the culture growth.

Key terms

neuron image segmentation; cultured neuronal network; neurite tracing; complex networks; network topology analysis; automated tracing; high throughput; neuronal morphology; light microscopy; connectome reconstruction

ALONG the past decades, cultured neuronal networks (CNNs) have constituted a fundamental tool for scientists, as one of the benchmark models for the study of the central nervous system. They, indeed, allow performing very well controlled laboratory experiments, thus providing a systematic way to approach fundamental questions, as for example, unveiling the principles and mechanisms underlying memory, connectivity, and even information processing of their *in vivo* counterparts (1–5).

CNNs have also important practical applications, when computer-connected to a real or a simulated robot (to create what is called a hybrot (6–8) or an animat (9,10), respectively), in that scientists are then endowed with the possibility of studying some basic neuronal processes in realistic contexts, such as learning and plasticity. Possibly, the most relevant advantage of CNNs is the unique option they offer of following the footprints of the self (or induced) organization of the network's functionality and dynamics (usually by means of a multi-electrode array [MEA] or calcium fluorescence, recording the CNN electrophysiological data, or inducing electrical stimulations in given spatial positions) together with the monitoring and tracking of the structural organization of the neurons' connectivity along the entire course of the culture's growth (11–13).

Although culturing neurons on top of a MEA equipped chamber implicates, in general, only mild constraints, following the development of the culture's structure is

Published online 00 Month 2014 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/cyto.a.22591

a far more delicate issue. Indeed, image-based systems biology essentially requires to gather sequential imaging of the culture and its processing to seek the evolution of the main network's indicators and measures along the CNN's maturation (14–16).

Now, it turns out that the accuracy of the existing image processing tools [as e.g., NeuronJ (17), NeuriteTracer (18), NeuronMetrics (19), NeurphologyJ (20), MorphoNeuroNet (21), HCA-Vision (22), and for a recent review on the subject consult Ref. (23)] in segmenting neurons and neurons' connections depends crucially on furnishing, as inputs, pictures with a high level of contrast. This traditionally led experimentalists to rely on immunocytochemistry techniques, which however implies cell fixation and death.

As a result, while comparing different CNNs at different stages of their maturation is nowadays a common practice, tracking variations of the network's structure on a single CNN requires a completely novel approach, wherein image processing could analyze non-invasively image acquisitions.

In this article, we describe a novel graph-based segmentation algorithm which operates on large scale images acquired by phase-contrast microscopy, and therefore by a fully non invasive technique, that is, without the need of adding chemicals to the culture. The algorithm accurately identifies the relevant network's units, and reconstructs the wiring of network connectivity with an overall computational cost (in terms of time and memory) which scales linearly with the image size. We show that we are able to fully track the main parameters characterizing the morphology and network's topology of a single culture during its maturation, and to identify the basic mechanisms that take place at different stages of the culture development. Finally, we show how such an information can be used for the construction of a data driven model of CNN's growth which is able, on its turn, to qualitatively reproduce the whole scenario observed in the experiments.

Materials and methods

Neuronal Cultures and Image Acquisition

Primary neuronal cultures were prepared from the frontal ganglion of adult locusts *Schistocerca gregaria*. Frontal ganglia dissection and dissociation procedures are thoroughly described in Refs. (14) and (24). In brief, neuronal cells are isolated (and removed from their original neurites) by enzymatic and mechanical dissociation. Cells are then resuspended in Leibovitz medium (L-15) with L-glutamine (Sigma–Aldrich, L4386), supplemented with 0.01% penicillin–streptomycin (Biological Industries, Israel), and seeded on a Concanavalin A (Sigma–Aldrich, C0412) pre-coated circular area ($r \sim 2.5$ mm) of a 35 mm Petri dish at a density of about 1,500 cells per cm^2, and left eventually for 2 h to allow adherence. Neurons are then incubated with 2 ml conditioned medium L-15 enriched with 5% locust hemolymph, and cultured in darkness for 18 days in vitro (DIV) under controlled temperature (29°C) and humidity (70%).

High-resolution and large scale phase-contrast images are acquired daily with an inverted microscope (Eclipse Ti-S, Nikon Instruments) equipped with a motorized XYZ stage (H117 ProScan, Prior Scientific), and using a charge coupled device camera (DS-Fi1, Nikon Instruments) with a 10× air (Achromat, ADL, NA 0.25) objective. The automated control of the motorized XYZ stage and camera is performed using the NIS-Elements software (Nikon Instruments Software, Nikon). Mosaic images with a pixel size of 1.34 μm are captured with the "large image" method implemented in NIS-Elements, which does automatic blended stitching with an overlap of 25%. Therefore, at each day of measurement, the result is a large high-resolution jp2 image file consisting of an ensemble of images acquired with 10× magnification in mosaic (usually 8 × 9 images). An example of a typical acquisition of the whole culture after 7 DIV is shown in Supporting Information Figure S1, while in the zoomed areas neurons, ranging from 10 to 50 μm in size, grow neuronal processes (neurites) trying to target neighboring cells.

Image Segmentation and Analysis

The proposed image-processing pipeline for the detection and analysis of CNNs comprises eight steps whose workflow is outlined in Figure 1. In the first step, the red RGB-layer of a phase-contrast large image corresponding to a time point is loaded (Fig. 1A). In the second and third steps (Fig. 1B), the image is segmented and thresholded, respectively, to separate background (noise, death cells, and cell debris) and foreground (neurons and neurites). Subsequently, in the fourth and fifth steps neurons and neurites are detected (Fig. 1C), after which the different components of the CNN are connected (Fig. 1D) and coded in the adjacency matrix (step 6), which is eventually analyzed to extract and store the morphological and topological properties of the CNN (steps 7 and 8). Background segmentation (second and third steps) is written in C++ while the rest of the algorithm is written in MATLAB. Both the MATLAB pipeline and the binary C++ MEX-file are available as Supporting Information.

Graph-based image segmentation and thresholding (steps 2 and 3). A first action to provide an accurate and fast segmentation of a CNN from phase-contrast illumination images is to aggregate areas with similar features according to a specific function along different scales. The implemented algorithm is graph-based (25,26), and it considers an input image I as an undirected weighted graph $G = (V, \varepsilon, w)$, in which every pixel of the image is a node in V, and each pair of pixels

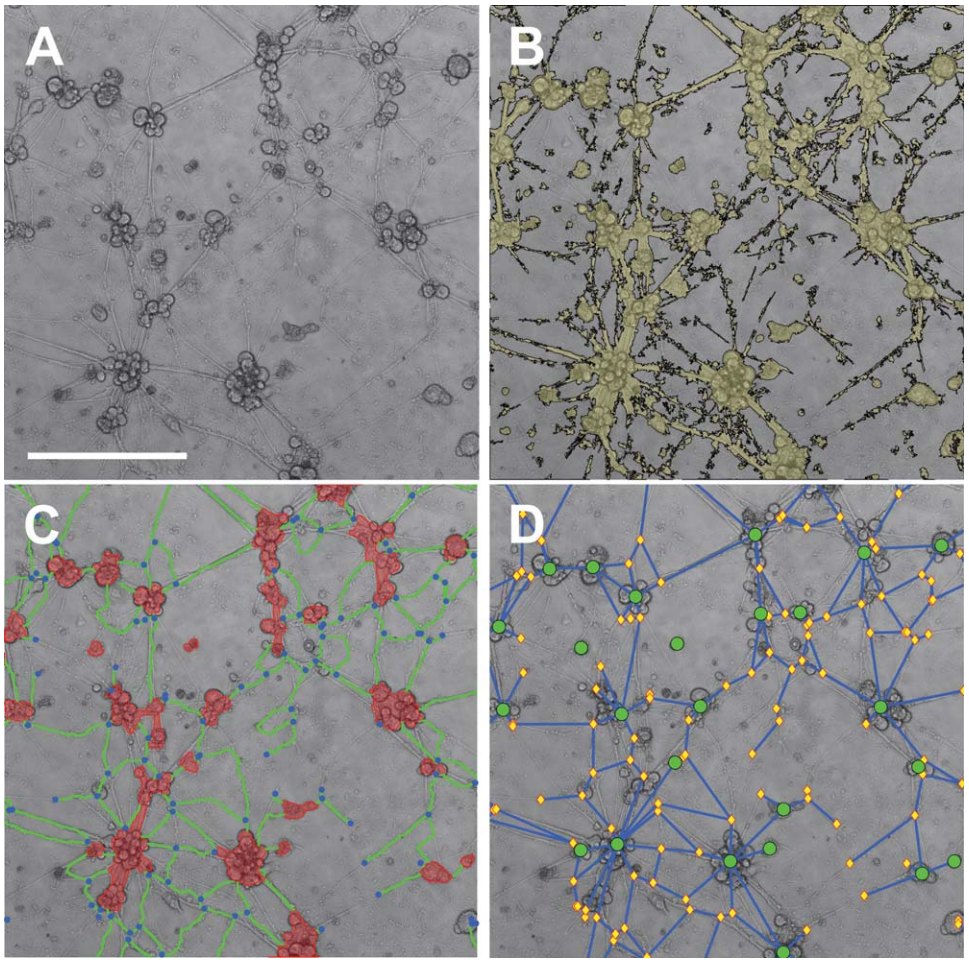

Figure 1. Image processing steps. (**A**) Red RGB-layer of an enlarged area of a culture 11 DIV old. (**B**) Output of the segmentation algorithm. The region of interest, where neurons and processes are located, is highlighted in yellow (boundaries are in black). (**C**) Output of the neuron and neurite searching algorithm. Identified clusters of neurons are marked in red, neurites in green, and processes' forks in blue. (**D**) Graph representing the network structure where green circles (clusters of neurons) and yellow diamonds (forks and processes' endings) are the nodes and blue straight lines (neurites) are the links between nodes whenever there is a process connecting them. Scale bar 500 μm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

is connected by an edge in ε. Given two nodes i and j, the existence of an edge between them depends on their spatial location and its weight w_{ij} depends on a similarity function of the pixel intensities. More precisely, for each node i we consider its pixel (px) location (x_i, y_i) in the spatial grid. We then consider the neighborhood $N(i)$ of node i to be formed by its eight first neighbors in the grid (local neighborhood), together with the four nodes located at positions $(x_i \pm 100$ px, $y_i)$ and $(x_i,\ y_i \pm 100$ px) (non-local neighborhood), and weight each link according to the absolute value of the difference of the pixel intensities. Local neighborhood attempts to link those pixels belonging to the same region, while non-local neighborhood aims for linking those regions with similar features. In our application, the purpose of the non-local neighborhood is to allow the segmentation task the merging of isolated background regions. Notice that, as the graph is undirected, the calculation of the weights only involves the half of them. The resulting matrix is then thresholded (adopting a suitable threshold value, which is then a parameter of the algorithm) to produce a binary graph, and all its connected components are determined, each one of them being labeled with the corresponding average intensity. In this way, all the pixels belonging to the background appear in a histogram grouped in a delta function, while the rest of pixels are uniformly distributed. Subsequently, a straightforward segmentation is performed with threshold-based segmentation obtaining two binary masks: one corresponding to those pixels belonging to background and the other with those pixels forming the neuronal network mask (Fig. 1B).

Neuronal cluster detection (step 4). Given the mask considered as no-background, to segment both isolated and clustered neurons we first distinguish between ROIs (regions of interest) larger and smaller than 10^5 pixels, to subsequently perform selective morphological operations on each class of ROI, being all of then standard techniques in image processing. On the large ROIs, we performed two consecutive erosions using rectangles as structural elements. After first

eroding with a 10 × 30 rectangle, regions smaller than 500 pixels are eliminated, and a second erosion is performed using a rectangle of 30 × 10 pixels. Finally, a dilation is applied using a disk with radius 10 pixels to smoothen the contours. The erosion in the smallest ROIs is performed by performing eight successive linear erosions at 45° from each other. After that, regions smaller than 100 pixels are eliminated and contours are eventually smoothed by means of a dilation with a 5 pixels disk. The combination of the two resulting ROIs yields the final neuronal cluster mask.

Neurite detection (step 5). After segmenting the neuronal clusters, the corresponding mask is subtracted from the neuronal network mask to proceed with the neurite detection. To that end, we apply a further morphological operation consisting in linearly dilating at (45°, −45°, 30°, −30°) to connect those areas which were separated due to small errors in the initial background segmentation, especially relevant in the case of thinner neurites. The segmentation is then finalized by performing a dilation with a 5 pixels circular mask, and those holes of size smaller than 500 pixels are filled in afterwards. The neurite mask is then generated from the skeletonized image, after which the neuronal cluster and neurite masks are combined to produce the CNN mask as shown in Figure 1C, were branching points and ends of neurites are marked as blue pixels, neurites as green lines, and neuronal clusters as red areas.

Adjacency matrix reconstruction and analysis (steps 6–8). The obtained CNN mask is used in a twofold way: to extract morphological parameters characterizing both neuronal clusters (number, size, centroid, roundness, etc.) and neurites (length, orientation, etc.), and to extract the actual adjacency matrix encoding the topology of the neuronal network, as shown in Figure 1D. The adjacency matrix is constructed as a binary and undirected graph, whose nodes are either branching points or cluster centroids, and two nodes are linked if there is a neurite process connecting them. Treating all links as identical, that is, ignoring edge length and edge directionality, such a graph can be described in terms of a symmetric matrix A whose elements a_{ij} are equal to 1 if nodes i and j are linked, and 0 otherwise.

Our graph is bipartite, as it is made of two types of nodes. However, we focus on the network statistical properties at the level of the neuronal clusters, ignoring the dynamics of both neurite connections and branching points. Therefore, we ended up with a subgraph defining the connectivity among the neuronal clusters in such a way that two of them are linked either directly or through a connected path of branching points. Such an object allows us to calculate all classical parameters characterizing the topology of a complex network (degree distribution, shortest path, node clustering, etc.) (27).

Algorithm Evaluation

To the best of our knowledge, there is no automated software dealing with segmentation of large size phase-contrast images of CNNs to compare with. Therefore, to quantitatively asses the capability of our proposed algorithm to reconstruct the CNN, we design a manual and an automated evaluation procedure. In the former case, we compare the topological similarity between the manual and algorithm's graph solutions while in the latter, segmentation results will be compared to ground-truth synthetic images whose network structure is known. To test the robustness of our segmentation algorithm, different levels of blurring were added to the synthetic images by applying a Gaussian filter of the same size of the images (blur kernel from MATLAB) and using the blur metric introduced in Ref. (28).

Manual annotation. Manual evaluation is performed, by eye inspection, labeling clusters of neurons and neurite's bifurcation points, and connecting those labels to create the graph representing the network structure. This manual annotation is done by the same subject by clicking and dragging using "gaimc" (Graph Algorithms In MATLAB Code) (29), a MATLAB tool modified to show the annotated image as a background. Because of the tool limitations, the image being annotated cannot be zoomed in or out causing some subjective decisions like resolving whether two neurons are in the same cluster or whether two neurons are connected through a very thin neurite. Supporting Information Figure S2 illustrates some of those situations where the user could make a subjective call. Conversely, if the image size (in pixels) is smaller than the screen resolution, it can happen that a neurite is present but not its origin nor its end and, consequently, that link is lost in the manual classification. Therefore, we consider images of size 1,000 × 1,000 pixels and compare the connectivity matrices resulting from the manual annotation and from applying our segmentation algorithm (see Supporting Information Fig. S3 for a visual inspection of the manual and algorithm's outputs). Finally, to proceed with the comparison of the two adjacency matrices, the following rules are taken into account: (i) a node in the manual solution is identified as the same node in the algorithm's one if the Euclidean distance between their positions is below a given threshold and vice versa; (ii) those nodes not having a reciprocal mapping are removed from both graphs; (iii) those nodes which are connected through a deleted node become connected through a direct link. This reduction process avoids the subjective judgments issue as each node has to be detected by the two labeling methods and the evaluation will reflect the ability of the algorithm to find an existing path between two detected clusters.

Synthetic images database. In the automated evaluation instead, we create a synthetic database of images emulating the observed CNN's. The image database is built using somas and neurites extracted from real images, and the background under the same optical conditions as in the experiments when only the culture medium is present. Eleven different neuron sizes and shapes (see Supporting Information Fig. S4 for an illustration) were used to create an artificial neuronal network as follows. Neurons are arbitrarily chosen from this set and

randomly placed in a 9,000 × 9,000 pixels square grid, allowing for the formation of larger clusters if neurons intersect. Next, cluster centroids are computed and a Delaunay triangulation is performed to obtain a connectivity matrix where some of those connections are randomly removed. Every pair of linked cluster centroids are physically connected in the image by placing a neurite whose thickness is randomly scaled and whose length and orientation are given by the Euclidean distance and relative angle of the straight line between them. Finally, the background is added into the image by blurring the borders between the artificial network and background. Supporting Information Figure S5 depicts the process to create one synthetic neuronal network.

Accuracy analysis. Let $A=(a_{ij})$ be the binary matrix representing the manual/synthetic graph connectivity and $B=(b_{ij})$ the matrix representing the algorithm's solution. We define true positive (TP) as the number of pixels in B such that $a_{ij}=b_{ij}=1$ and true negative (TN) if $a_{ij}=b_{ij}=0$ (that is, in both cases there is a true segmentation). A false positive (FP) is defined as those pixels which satisfy $a_{ij}=0$ and $b_{ij}=1$, while a false negative (FN) corresponds to $a_{ij}=1$ and $b_{ij}=0$. The proposed evaluation method is based on the F-measure (30) defined as $F=2RP/R+P$, where $P=\mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$ stands for the Precision or Confidence in retrieving true segmentations, while $R=\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$ represents the Recall or Sensitivity of the algorithm to return most of the links in A. F-measure is within the [0, 1] interval, so that 0 states for a bad segmentation, while on the contrary 1 represents the best segmentation result. Together with these measures, we also computed the % of coincidence E between the two matrices.

Results

Neuronal Network Analysis

To automatically analyze the neuronal morphology and network topology, we have developed a graph-based segmentation algorithm which combines C++ and MATLAB. The pipeline, thoroughly described in Materials and Methods section and illustrated in Figure 1, processes RGB phase-contrast images of neuronal cultures as large as 20 mm^2 at different stages of the network development. First, it takes the red layer and clusters nearby pixels with similar features allowing to separate the background from the foreground containing the pixels forming the neuronal mask (neurons and neurites). Second, the algorithm operates in the foreground applying several morphological operations to segment the neuronal cluster and neurite masks. During this process, a skeletonization of the neurite mask finds the bifurcation points. Finally, the pipeline analyses and exports multiple morphological and topological characteristics of the neuronal network. Figure 2 shows an example of the kind of information it can be extracted. There, the pipeline is applied to images captured at three different ages of the same culture, 3, 6, and 11 DIV. Panels A–C just show a small area of each image while panels D–F show, superimposed to the original images, the corresponding adjacency matrices extracted by the program and capturing the observed network structure. As already introduced in Materials and Methods section, we just focused on the network at the level of the neuronal clusters such that two of them are linked either directly or through a connected path of branching points. This is illustrated in the inset of Figure 2G, sketching how a small subgraph with neurons and bifurcation points, like the one encircled in Figure 2D, it is reduced to a subgraph whose nodes are only neuronal clusters. Among the classical measures characterizing the connectivity of a graph is the degree distribution, $P(k)$ that is the probability of having a node connected to k neighbors, being k the degree of a node. A related quantity is the cumulative degree distribution which refers to the probability of having a node with a degree greater or equal to k. In panels G–I, we plot the cumulative degree distributions of the whole networks developed in days 3, 6, and 11, respectively. We observe that during the culture maturation, degree distributions had a fast decay with a non monotonous increase in the average connectivity, with most of the nodes having a similar number of connections and only a few ones with degrees deviating significantly from such a number.

Performance

To evaluate the performance of the proposed algorithm in the analysis of real CNNs, we consider aspects concerning the computational cost and accuracy.

Computational cost. Figure 3A shows the overall processing time for each one of the images of the dataset. The height of the bars accounts for the total time consumed in both the segmentation and analysis tasks (time consumed by specific tasks are color coded), and is computed assuming that the graph-based algorithm is implemented in C++ and the rest of the pipeline uses MATLAB running in a PC computer (Intel Core i5, 6GB RAM, 3.00 GHz). As it can be observed from the inset of Figure 3A, background segmentation and adjacency matrix extraction are the most demanding tasks, consuming, respectively, 30% and 45% of the total computational time. To further quantify the computational efficiency of the algorithm, we plot in Figure 3B the image processing time as a function of the image size (in pixels). The algorithm is able to manage images as large as 125 Mpixels in less than 6 min (while the segmentation task only takes 1.6 min, see inset of Fig. 3B) without RAM saturation. For larger image sizes, the RAM becomes saturated and the performance worsens, as the pipeline needs the hard drive to caching data. Moreover, a detailed examination of the code shows that the algorithm performs linearly with the number of pixels. As shown in the inset, performing time also scales linearly when only the time consumed during just the segmentation task is taken into account. This is also supported by the linear regressions in both RAM and hard drive memory usage regions.

Accuracy. As for the algorithm accuracy in reconstructing the CNN, that is, understanding accuracy as the capability to properly detect a link, we compare the adjacency matrices coming from manual annotation or from synthetic images

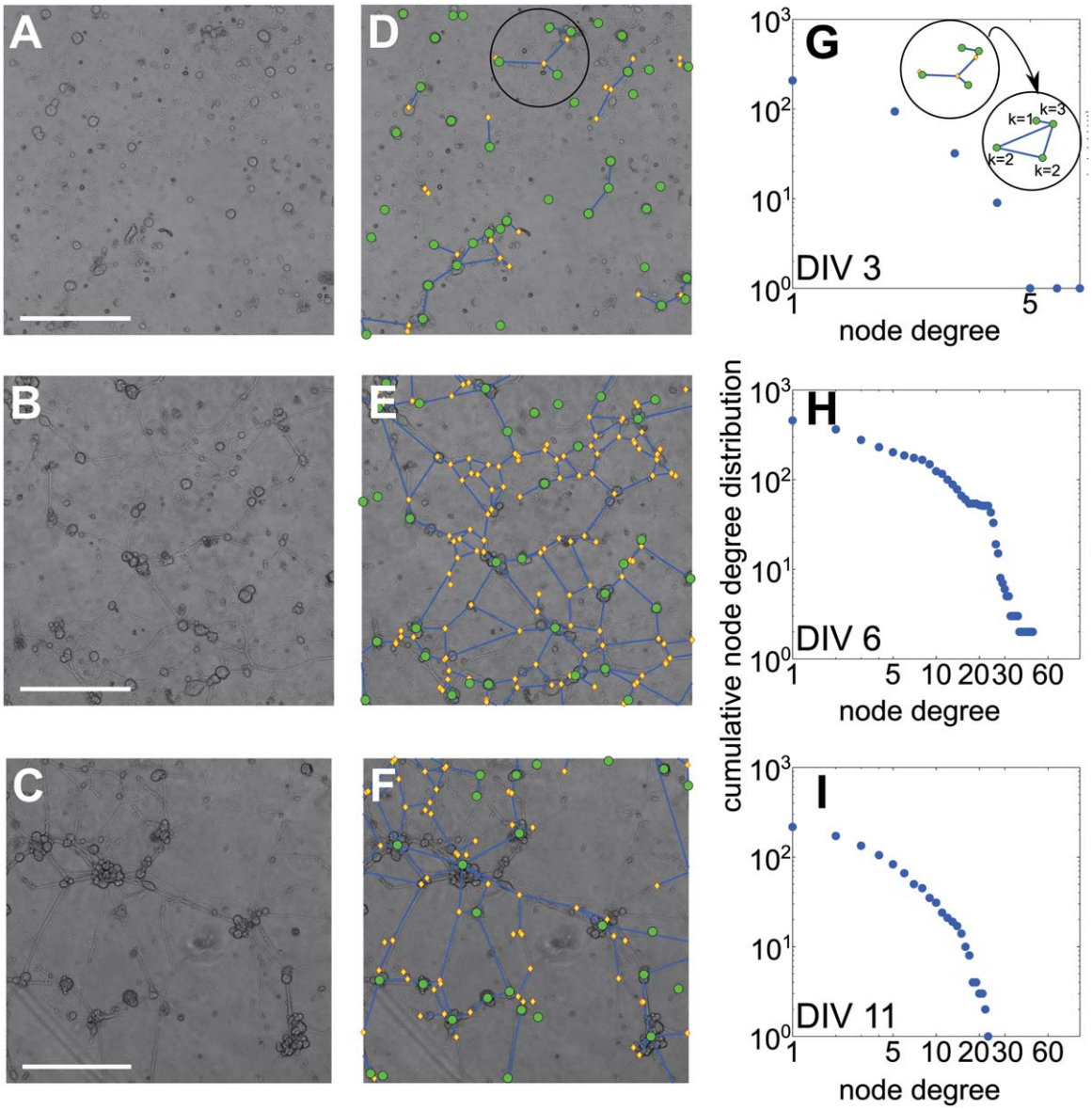

Figure 2. Neuronal network analysis. Panels (**A**)–(**C**), phase-contrast images at different locations of a self-organizing neuronal culture from frontal ganglia of locusts after 3, 6, and 11 days of plating, respectively. Panels (**D**)–(**F**), the corresponding graph representations of the neuronal networks extracted using the image segmentation algorithm. In this representation, green circles (clusters of neurons) and yellow diamonds (forks and processes' endings) are the nodes of the network, and blue straight lines (neurites) are the links between nodes whenever there is a process connecting them. Panels (**G**)–(**I**) show the cumulative degree distributions of the adjacency matrices encoding the graphs in panels (D)–(F) when only the network of paths connecting neuronal clusters (green circles) is considered. Inset in panel (G) shows how that network of neuronal clusters is built. Scale bar 500 μm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

whose structure is known (see Algorithm Evaluation in Materials and Methods section for more details) with the adjacency matrix resulting from the algorithm's segmentation.

To this purpose, we applied the algorithm to small size images (1,000 × 1,000 pixels) of a 6 DIV neuronal culture, and manually annotated each image as previously described. An example of such a comparison is shown in the top row of Figure 4 after an original image (Fig. 4A1) is being manually (Fig. 4A2) and automatically (Fig. 4A3) annotated, where neurons (green circles), bifurcations (yellow diamonds), and neurites (blue straight lines) are labeled. The resulting F-score obtained for 32 images is $73 \pm 3\%$ ($R = 65 \pm 4\%$, $P = 92 \pm 3\%$) ($\pm$ refer to the standard error of the mean) and the level of coincidence $E = 80 \pm 2\%$, indicating that the segmentation algorithm is able to statistically detect, with high precision, most of the existing links in the manually annotated graph. However, there are several causes affecting the algorithm's performance like the presence of debris or image artifacts (blurriness, illumination problems, . . .) which makes difficult the segmentation of thin neurites like the ones marked with black arrows in

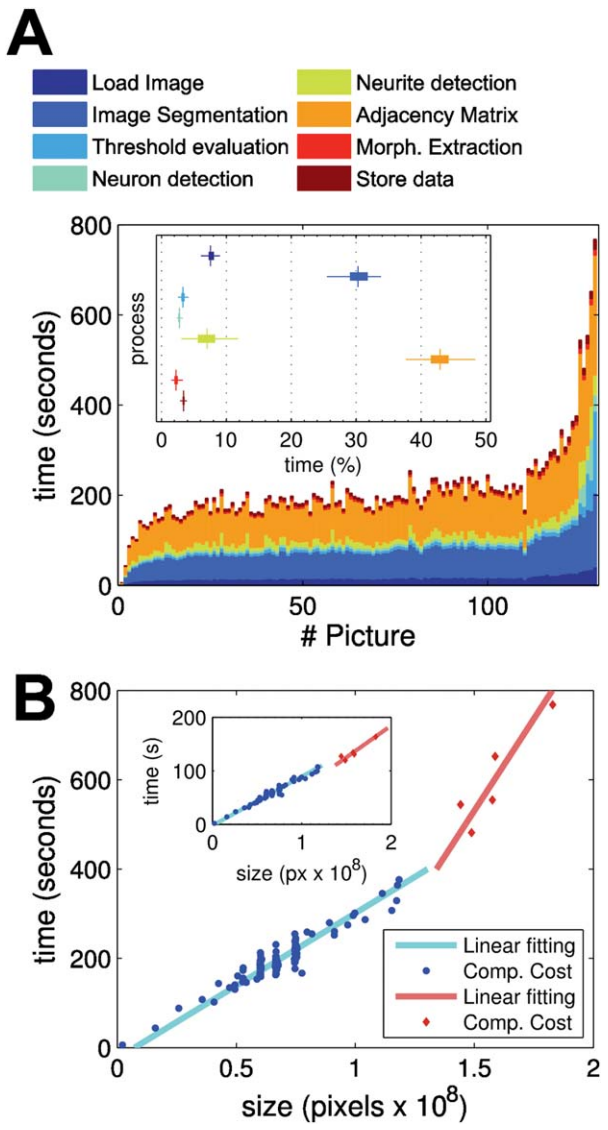

Figure 3. Computational cost. (**A**) Image processing time for each one of the whole set of images ordered according to their sizes (from left to right). Each bar is proportionally divided in colored segments (see legend) according to the time each specific task consumes. (Inset) Task time percentage in relation to the whole process. (**B**) Image processing time as a function of the image size (in pixels). Blue dots account for the processing time when the RAM is not saturated while red diamonds account for the processing time when hard drive is used to caching data (poor performance). Linear regressions show that the algorithm computational cost scales linearly with the image size, that is $O(n)$ where n is the image size. (Inset) The same as is in the main plot but time corresponds only to the time consumed during segmentation. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figures 4A2 and 4A3 (a more detailed revision of the problems affecting the algorithm's performance can be found in Supporting Information Fig. S6).

In addition to the problems associated with the user bias, discussed in Materials and Methods section, the manual evaluation of the performance restricted to small areas of the culture, hampers the possibility of evaluating the ability of the algorithm to detect long-range connections, and, therefore, of working with a real ground-truth topology to compare with. To overcome this issue, we devised an automated method of evaluation based on synthetic images of neuronal networks whose topology is known. The bottom row of Figure 4 shows a detail of one artificial network designed as explained in Materials and Methods section. The underlying network structure used to create the artificial network image is shown in Figure 4B2 while Figure 4B3 depicts the graph solution rendered by the algorithm. We use 16 images (9,000 × 9,000 pixels) to evaluate the performance with different network configurations and backgrounds, obtaining an accuracy of about 80.4 ± 0.7% (± means standard error).

Automatic evaluation gives us the possibility of comparing the neuronal network mask obtained just after the image segmentation (see Image Segmentation and Analysis in Materials and Methods section) with the exact location of neurons and neurites used in the synthetic image creation. The synthetic image can be mapped into a 9,000 × 9,000 matrix, encoding the existence of a neuron or a neurite pixel in the image if the matrix element equals 1 and 0 otherwise. Taking this matrix as ground-truth, we evaluated the algorithm's capacity for recovering the neuronal network mask (or, equivalently, for distinguishing the background from the foreground) yielding a 96.6 ± 0.3% of *F*-score. Taking advantage of this framework, we measured the robustness of the algorithm's segmentation by blurring the synthetic images with a Gaussian filter. The blur level was measured as in Ref. (28) and it is within the interval [0, 1] where 0 means sharp and 1 means blurred. Figure 5 shows the effect on the accuracy of the segmentation process as a function of the blur level. Left panels in Figures 5A and 5B show the synthetic blurred images for two values of blurring 0.22 and 0.41, respectively, while the right panels show the differences between the ground-truth (in red) and the segmented mask (in yellow). The more yellow is the picture the better is the accuracy (98% in *A* and 73% in *B*). In a more systematic way, Figure 5C quantifies the *F*-score in the location of neurons and neurites' pixels as a function of the blur level, showing a relatively slow decay up to blur levels of about 0.35, after which the accuracy no longer sustains. As a matter of comparison, the graph also shows the usual blurriness value of a real culture image (0.23 ± 0.02), which falls within the region where the algorithm exhibits a more stable behavior.

Network Development and Topology

Typically, when the evolution of a CNN from an initial state of dissociated neurons is monitored, one observes that neurons create a network of synaptic connections from scratch, and that the growth of the network is more pronounced in the early stages of the culture, until it reaches a sort of asymptotic state (14–16,31). The tracking of the full process with our image processing tool allows us to identify (and, more importantly, to quantify) the different network stages through which neurons arrange their connectivity until they reach a functional mature state. In particular, the analysis of the images suggests that CNNs tend to develop from

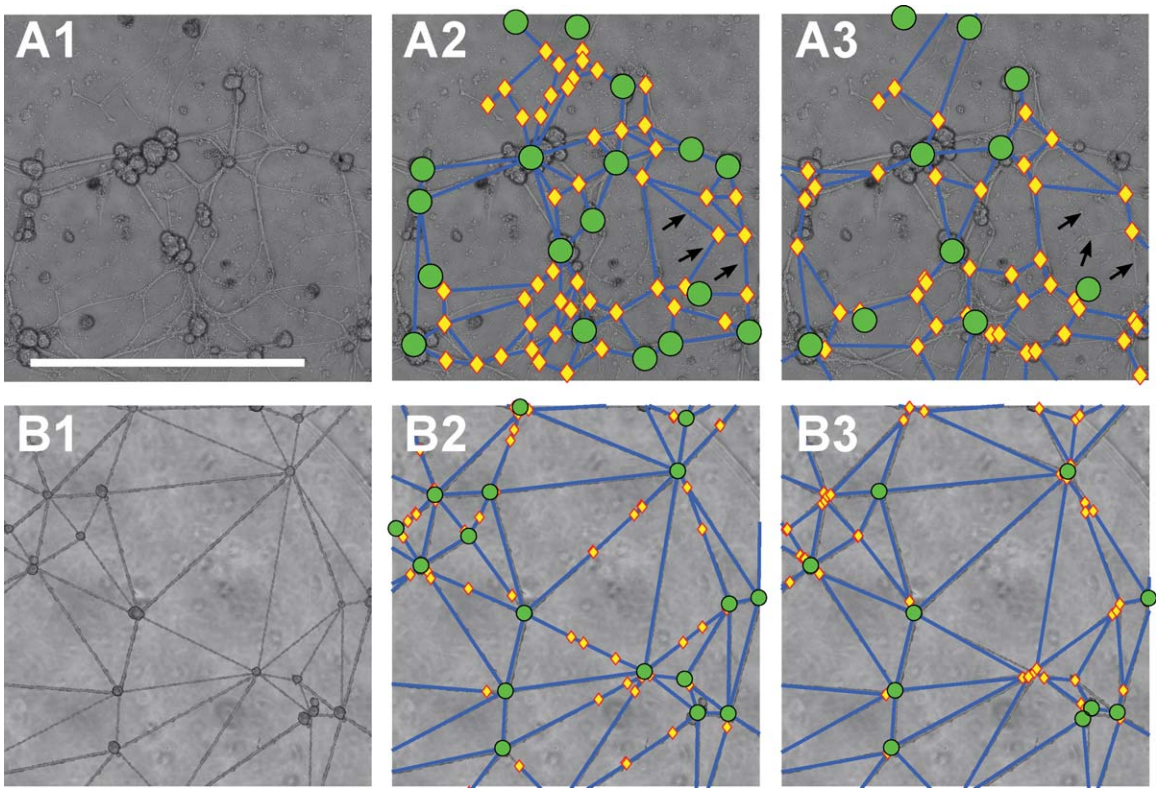

Figure 4. Evaluation of the algorithm's accuracy using manual annotation (top row) and synthetic images (bottom row). **(A1–A3)** Phase-contrast images (1,000 × 1,000 pixels) from a CNN after 6 DIV used to compute the algorithm's performance (A1), manual labeling of the image (A2), and the resulting graph proposed by the algorithm (A3). Scale bar 1 mm. **(B1–B3)** Image segment (1,000 × 1,000 pixels) of an artificial neuronal network (B1), network structure used to create the synthetic image (B2), and the algorithm's solution when applied to the image on the left (B3). In panels (A2-A3) and (B2-B3), green circles (clusters of neurons) and yellow diamonds (forks and processes' endings) represent the nodes of the network, and blue straight lines (neurites) are the links between nodes whenever there is a process connecting them. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

random toward small-world (SW) networks (13,14). This is evidenced in Figure 6A, where we report the simultaneous increase of the clustering coefficient and decrease in the mean path length of the network, a clear fingerprint of the onset of a SW network configuration (32). SW features prominently manifest at DIV 6, and the resulting graph remains relatively stable in topology through the rest of the life of the culture. Furthermore in Ref. (14), we already demonstrated that the algorithm is also able to extract the most relevant network statistics at the micro-scale, like degree distribution, degree–degree correlations, and so forth.

Even more importantly, the unique possibility of tracing the development of a specific, single, CNN endows us also with the option of measuring topological changes over different days of the culture maturation, and therefore of individuating the main physical processes underlying the culture's morphological transformations. On its turn, such a latter information can be conveyed to engineer a data-driven growth model, in view of qualitatively reproducing the overall scenario of experimental evidences.

Data-Driven Model

Namely, we now move to show how a relatively simple growth model can in fact encompass an initial phase of random growth that lasts up to a particular critical point, where optimization processes, such as neuronal clusterization and neuronal cable minimization driven by tension forces (33), start to be the dominating mechanisms of the evolution. The interplay and alternation of these two mechanisms (that were already taken into account, but separately, in previous studies (34,35)) are indeed the core of information that is given to us by the comparative analysis allowed by the algorithm.

We start from a zero-model, previously reported in Ref. (14) that considers a set of N cells randomly distributed with a uniform cell density ρ. Each cell is associated with an interaction disk of radius $r_i(t)$ with an initial radius a. Whenever two cell's disk overlap, a link between them is established, whose length equals the distance between the centers of the two cell's disks. If this distance is less than $2a$, then the two cells are merged into a single cluster. The time evolution of the growth of each disk depends on the actual connectivity $k_i(t)$ of the cell. At each discrete time step t (here denoting the sequence of the DIV), the radius $r_i \geq a$ increases by a quantity $\delta r_i(t)$ which decays as: $\delta r_i(t) = V/t[1-(1/K_i)k_i(t-1)]$ where V is the neurite growth velocity (the same for all cells), K_i a random number in the interval $[1, N]$, and k_i the degree of the node (cell) at the time step t. The term k_i/K_i introduces heterogeneity in the cell population, and represents the fraction of

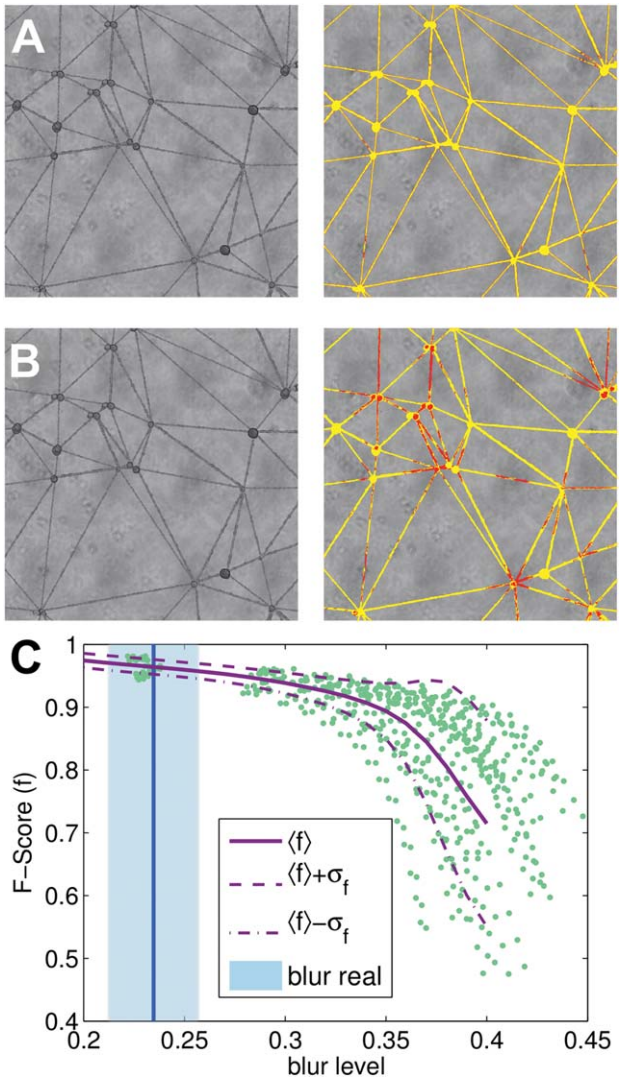

Figure 5. Robustness of the algorithm's segmentation to blur. (**A**, **B**) Left panels show synthetic images with different blur levels (A) 0.22, and 0.41 (B), while the right panels represent the comparison between the image segmentation outcome (in yellow) and the ground truth mask (in red). (**C**) Algorithm's accuracy processing synthetic images as a function of the blur level. Green dots are the *F*-scores of all the synthetic images analyzed for a given blur level. Purple solid (dashed) line represents the average (standard deviations) value. Blue area stands for the usual blurriness values observed in real culture images (0.23 ± 0.02). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

links acquired by the cell in the previous steps from the initial randomly assigned endowment K_i. A very large K_i indicates that, potentially, a cell is very active and could connect to many other cells. The wiring process is iterated up to a given time step T_s, at which the formation of new connections is stopped.

The only variation that we are here considering with respect to the original model of Ref. (14) is that now the neurite growth velocity V is no longer isotropic, and rather each cell disk i has an associated velocity $V_{ij} = V\,\xi_{ij}$ with respect to cell j, where ξ_{ij} is a random number drawn from a uniform distribution between 0 and 1. This leads to the fundamental consequence that two disks initially located very close to each other do not necessarily establish a connection, but they have instead an associated probability of linking, as observed in the experiment.

Furthermore, and following Ref. (14), whenever a new link is formed between cells i and j, a tension force $T_{ij} = 0.1$ is created along the line connecting them, and it is incremented in 0.1 force units at each time step. The total force acting on a

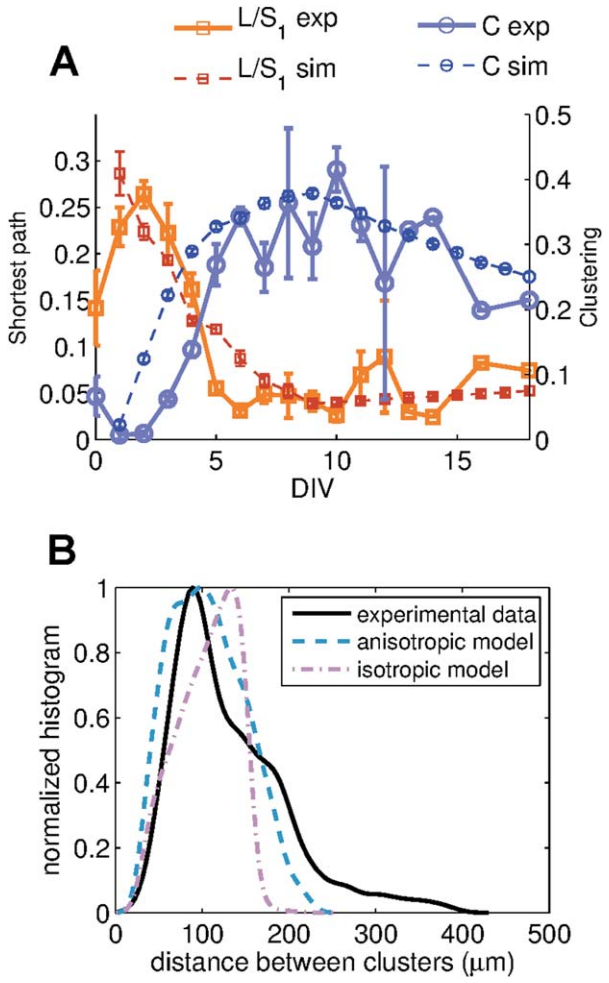

Figure 6. Comparison between model and experiment. (**A**) Shortest path L (left vertical axis, orange lines and symbols), normalized by the size of the giant connect component, and clustering coefficient C (right vertical axis, blue lines and symbols) as a function of the culture age (DIV) for both experiments (solid curves) and simulations (dashed lines). (**B**) Normalized histograms of the distance between connected clusters in DIV 5 obtained in the experiments (solid line), and in the simulations using the anisotropic (dashed line) and isotropic (dot-dashed line) versions of the model. All experimental quantities are averaged for the set of six cultures at each day of measure (DIV). The simulation parameters are $N = 700$, $T_s = 9$, and $V = 65$, 40 for the anisotropic and isotropic models, respectively, and each point is the ensemble average over 50 independent runs of the growth algorithm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

cell or cluster i is given by the vectorial sum of all the tension forces acting on it. Finally, each cell is "anchored" to the substrate by a force $F_a = 10$ force units, and the ith cell can only be detached if it experiments a net force F_i larger than the adhesion to the substrate. In the case of a cluster of cells, the adhesion force to the substrate is considered to be the sum of the individual adhesions of the cells composing the cluster. Therefore, cells and clusters move in a certain direction in all circumstances in which the net force acting on them overcomes the adhesion force, and an equilibrium point is reached at a new position in which the new net force balances (or is smaller) than the adhesion to the substrate.

To validate our model, we ran a large number of simulations for different values of the model parameters N, V, and T_s. Remarkably, when comparing the statistical topological features of the simulated networks to those measured from the experiments, we found that high correlation values exist only in a very narrow window of V and T_s. For instance, the parameter values which better fit the experimental observations for $N = 700$ neurons, are $V = 65 \pm 5$ and $T_s = 9 \pm 1$.

Precisely, Figure 6A reports the comparison of the time evolution of the clustering coefficient C and of the shortest path L (normalized by the size of the giant connected component of the network) for both experiments (continuous lines) and simulations (dashed lines), and the agreement is good, indicating that the model is able to capture the main mechanisms leading the culture to organize into a small-world structure. The new version of the model also outperforms the previous one reported in Ref. (14) regarding the evolution of the morphological properties of the culture, as it is the case of the neurite length distribution. Figure 6B shows the comparison between the normalized histograms of the average distance between connected clusters typically observed in the DIV 5 of the culture both in experiments and simulations. There, it is clear that the anisotropic growth of the disks makes the model to better fit the trend measured in the experiments, as this modification allows the connection of more distant neurons, giving rise to a not so sharp fall in the tail of the distribution as observed in the isotropic version of the model.

Discussion

Despite the existence of some functional studies of CNNs (36,37) and also of a few theoretical–numerical attempts to model the relationships between CNNs' structure and function (38), no experimental verifications are available so far of such a connection, basically due to the lack of tools allowing a simultaneous tracking of the culture's dynamical activity and morphological/topological changes.

To cater for this need, we developed an algorithm and a set of tools to enable automatic location of neurons and tracing of neurites in non-invasive phase-contrast images, which can be acquired simultaneously with electrophysiological measures (with e.g., multi-electrode arrays), hence potentially allowing for a combined study of network structure and dynamics.

The relevance and value of our work is then to make a first step in the direction of unveiling and uncovering structure/function relationships during the evolution of a CNN.

In particular, in this report we operated a fully longitudinal inspection of a single culture's life and maturation process, in which some aspects of the topological and functional organization can be compared (39). The accuracy of the proposed algorithm (in comparison to the manual and automated evaluations) is very good, and the computational demand is rather low.

It is worth saying that one way to improve the validation of our algorithm would have been the use of green fluorescent protein (GFP) transfection or simply anti-horseradish peroxidase (HRP) staining. These tools label the connections and would therefore provide precise information on the physiological, ground-truth topology of the network. Although these labelling techniques are difficult to combine (the former) or incompatible (the latter) with a continuous monitoring of the culture maturation, their availability at a particular age of the culture would have allow the access to other neuronal tracing automated software to compare with and to produce ground truths with most of the inter-neuronal connectivity. Apart from quality image issues, a potential bottleneck of our algorithm is working with high dense neuronal networks as the extraction of the network structure could be compromised mainly due to neurite crossing or high interconnectedness.

More importantly, the possibility of pursuing the footprints of the different mechanisms at the basis of the culture's morphological changes allowed us to set up and assemble a fully data-driven growth model, whereby the observed phenomenology can be qualitatively reproduced to a large extent.

As a perspective work, the algorithm can be easily customized (and its parameters conveniently tuned) to study the spontaneously emerging morphology and organization in cultures of dissociated neurons of other animals, and in particular of vertebrates (rats and mice) given the current level of attention and their vast use in a series of other experiments with CNNs.

Literature Cited

1. Eckmann J, Feinerman O, Gruendlinger L, Moses E, Soriano J, Tlusty T. The physics of living neural networks. Phys Rep 2007;449:54–76.
2. Baruchi I, Ben-Jacob E. Towards neuro-memory-chip: Imprinting multiple memories in cultured neural networks. Phys Rev E 2007;75:1–4.
3. Fuchs E, Ayali A, Robinson A, Hulata E, Ben-Jacob E. Coemergence of regularity and complexity during neural network development. Dev Neurobiol 2007;67:1802–1814.
4. Marom S, Shahaf G. Development, learning and memory in large random networks of cortical neurons: Lessons beyond anatomy. Q Rev Biophys 2002;35:63–87.
5. Van Pelt J, Vajda I, Wolters PS, Corner MA, Ramakers GJA. Dynamics and plasticity in developing neuronal networks in vitro. Prog Brain Res 2005;147:173–188.
6. De Santos D, Lorente V, de la Paz F, Cuadra JM, Álvarez-Sánchez JR, Fernández E, Ferrández JM. A client–server architecture for remotely controlling a robot using a closed-loop system with a biological neuroprocessor. Rob Auton Syst 2010;58:1223–1230.
7. Bakkum D, Shkolnik A, Ben-Ary G, Gamblen P, DeMarse TB, Potter SM. Removing some "A" from AI: Embodied cultured networks. In: Iida F, Pfeifer R, Steels L, Kuniyoshi Y, editors. Embodied Artificial Intelligence. Berlin: Springer; 2004. pp 130–145.
8. Bakkum DJ, Chao ZC, Gamblen P, Ben-Ary G, Shkolnik AG, DeMarse TB, Potter SM. Embodying cultured networks with a robotic drawing arm. Conf Proc IEEE Eng Med Biol Soc 2007;2007:2996–2999.
9. Wilson S. Knowledge growth in an artificial animal. In: Proc Int Conf Genet Algorithms their Appl Hillsdale, NJ, USA: L. Erlbaum Associates Inc.; 1985. pp 16–23.
10. Demarse TB, Wagenaar DA, Blau AW, Potter SM. The neurally controlled animat: Biological brains acting with simulated bodies. Auton Robots 2001;11:305–310.
11. Bologna L, Nieus T, Tedesco M, Chiappalone M, Benfenati F, Martinoia S. Low-frequency stimulation enhances burst activity in cortical cultures during development. Neuroscience 2010;165:692–704.
12. Orlandi JG, Soriano J, Alvarez-Lacalle E, Teller S, Casademunt J. Noise focusing and the emergence of coherent activity in neuronal cultures. Nat Phys 2013;7:1–9.
13. Downes JH, Hammond MW, Xydas D, Spencer MC, Becerra VM, Warwick K, Whalley BJ, Nasuto SJ. Emergence of a small-world functional network in cultured neurons. PLoS Comput Biol 2012;8:e1002522.

14. De Santos-Sierra D, Sendiña-Nadal I, Leyva I, Almendral JA, Anava S, Ayali A, Papo D, Boccaletti S. Emergence of small-world anatomical networks in self-organizing clustered neuronal cultures. PLoS One 2014;9:e85828.
15. Woiterski L, Claudepierre T, Luxenhofer R, Jordan R, Käs JA. Stages of neuronal network formation. New J Phys 2013;15:025029.
16. Shefi O, Golding I, Segev R, Ben-Jacob E, Ayali A. Morphological characterization of in vitro neuronal networks. Phys Rev E 2002;66:1–5.
17. Meijering E, Jacob M, Sarria J-CF, Steiner P, Hirling H, Unser M. Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. Cytometry A 2004;58A:167–176.
18. Pool M, Thiemann J, Bar-Or A, Fournier AE. NeuriteTracer: A novel ImageJ plugin for automated quantification of neurite outgrowth. J Neurosci Methods 2008;168: 134–139.
19. Narro ML, Yang F, Kraft R, Wenk C, Efrat A, Restifo LL. NeuronMetrics: Software for semi-automated processing of cultured neuron images. Brain Res 2007;1138:57–75.
20. Ho S-Y, Chao C-Y, Huang H-L, Chiu T-W, Charoenkwan P, Hwang E. NeurphologyJ: An automatic neuronal morphology quantification method and its application in pharmacological discovery. BMC Bioinformatics 2011;12:230.
21. Pani G, De Vos WH, Samari N, de Saint-Georges L, Baatout S, Van Oostveldt P, Benotmane MA. MorphoNeuroNet: An automated method for dense neurite network analysis. Cytometry A 2014;85A:188–199.
22. Vallotton P, Lagerstrom R, Sun C, Buckley M, Wang D, De Silva M, Tan S-S, Gunnersen JM. Automated analysis of neurite branching in cultured cortical neurons using HCA-Vision. Cytometry A 2007;71A:889–895.
23. Meijering E. Neuron tracing in perspective. Cytometry A 2010;77A:693–704.
24. Anava S, Saad Y, Ayali A. The role of gap junction proteins in the development of neural network functional topology. Insect Mol Biol 2013;22:457–472.
25. De Santos Sierra A, Avila CS, Casanova JG, del Pozo GB. Gaussian multiscale aggregation applied to segmentation in hand biometrics. Sensors 2011;11:11141–11156.
26. Shi J, Malik J. Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 2000;22:888–905.
27. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. Complex networks: Structure and dynamics. Phys Rep 2006;424:175–308.
28. Crete F, Dolmiere T, Ladret P, Nicolas M. The blur effect: Perception and estimation with a new no-reference perceptual blur metric. In: Proc SPIE 6492, Hum Vis Electron Imaging XII, Vol. 6492; 2007. p 64920I. Available at: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1298489.
29. Gleich D. Gaimc: Graph algorithms in MATLAB code. MATLAB 2009. MATLAB 7.5 (R2007b): File ID #24134.
30. Alpert S, Galun M, Brandt A, Basri R. Image segmentation by probabilistic bottom-up aggregation and cue integration. IEEE Trans Pattern Anal Mach Intell 2012;34: 315–326.
31. Lai P-Y, Jia L, Chan C. Growth of cortical neuronal network in vitro: Modeling and analysis. Phys Rev E 2006;73:051906.
32. Watts DJ, Strogatz SH. Collective dynamics of "small-world" networks. Nature 1998; 393:440–442.
33. Anava S, Greenbaum A, Jacob E, Hanein Y, Ayali A. The regulative role of neurite mechanical tension in network development. Biophys J 2009;96:1661–1670.
34. Shefi O, Ben-Jacob E, Ayali A. Growth morphology of two-dimensional insect neural networks. Neurocomputing 2002;44–46:635–643.
35. Segev R, Benveniste M, Shapira Y, Ben-Jacob E. Formation of electrically active clusterized neural networks. Phys Rev Lett 2003;90:168101.
36. Teller S, Granell C, De Domenico M, Soriano J, Gomez S, Arenas A. Emergence of assortative mixing between clusters of cultured neurons. PloS Comput Biol 2014;10: e1003796.
37. Downes JH, Hammond MW, Xydas D, Spencer MC, Becerra VM, Warwick K, Whalley BJ, Nasuto SJ. Emergence of a small-world functional network in cultured neurons. PLoS Comput Biol 2012;8:e1002522.
38. Fuchs E, Ayali A, Ben-Jacob E, Boccaletti S. The formation of synchronization cliques during the development of modular neural networks. Phys Biol 2009;6:036018.
39. Feldt S, Bonifazi P, Cossart R. Dissecting functional connectivity of neuronal microcircuits: Experimental and theoretical insights. Trends Neurosci 2011;34:225–236.