Detection and Quantification of Neuron's Cellular Bodies in C. elegans

Jorge M. Santos[1,2] and Daniel Dias[1]
[1]ISEP, School of Engineering, Polytechnic of Porto, Portugal
[2]INEB, Biomedical Engineering Institute, Porto, Portugal
Email: {jms, 1130182}@isep.ipp.pt

***Abstract*—Parkinson's disease is a neuronal degenerative disease characterized by problems on movement ability originated by the presence of mutant forms of leucine-rich repeat kinase in neurons causing its degradation and death. Studying the mechanisms involved in this process is fundamental to find a desirable therapeutical treatment. On way this is being done is to use the Caenorhabditis elegans (C. elegans), a simple nematode, to model this disease and to simulate the influence of certain proteins and dysfunctions in the way the disease evolves. Laboratory experiments are being carried out to study this influence by submitting C. elegans to different conditions and analyzing how the degradation evolves. Images of C. elegans are taken in several stages of the process and several characteristics are extracted to quantify the degradation. In this work, we implement a pipeline using Matlab and CellProfiler to automatize and speedup this process and to alleviate the burden of doing it manually. Preliminary results are very promising and motivate us to continue with this work.**

***Index Terms*—detection, cellular bodies, C. elegans, Parkinson's disease**

I. Introduction

Caenorhabditis elegans (C. elegans) is a small organism widely used to study neuronal degenerative diseases. It is a small and transparent nematode from which we can easily acquire images without the need of sampling processing techniques. In this work, we analyse C. elegans phenotype changes and its relation to Parkinson disease. C. elegans has eight dopaminergic neurons, six of them in the head region. In Fig. 1 (right), we show an example of a normal nematode (head region) where we highlight some of its cellular bodies (oval) and axons. These two components are the neuron's structures. In Fig. 1 (left), we can see an example of the neuronal degradation exposed by the small bumps along what is left from the axons and also some deformity of the cellular bodies.

Neurons integrity is a direct measure of a disease manifestation. In this kind of images, we are especially interested on detecting and counting cellular bodies, to measure its area and to evaluate axons integrity. In this work, we focus on the first two objectives. The integration of these parameters will give us a level of neuro-degeneration. Studies that are being conducted on these kinds of organisms for the same purpose are mainly based on manual and tedious observation. With the purpose of alleviating the burden of analyzing hundreds of images, to eliminate some common observational errors and to expedite the extraction of important characteristics we implemented an automatic tool based on an open source application.

This work is structured as follows: in Section 2 we make a brief introduction to C. elegans anatomy and also to Parkinson disease and its relation with the nematode; in Section 3 we introduce the data set, the CellProfiler and how to perform the automatic detection of Cellular Bodies; in Section 4 we analyze the results and in the last section we present the conclusions.

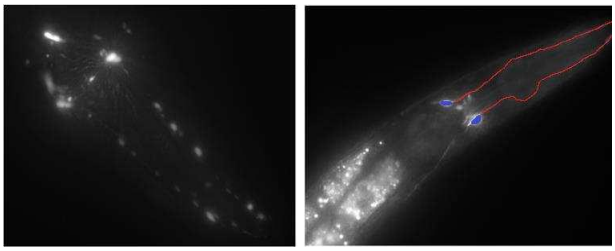

Figure 1. Two examples of C. elegans (head region). In the right side a normal one with neurons highlighted (cellular bodies and axons). On the left side an abnormal one with visible axons degeneration.

II. C. elegans and Parkinson's Disease

C. elegans (Fig. 2) is a small soil nematode (about 1mm length) used in different lab experiments for biological development, genetics, toxicology and aging studies. Its characteristics like small size, short life period a small genome and an easy cell culture make it ideal for these kinds of studies [1]-[3]. Similar to other nematodes, they have a non-segmented and cylindrical body, straightening in the extremities. Its nervous system's cell is organized in ganglion from head to tail. Most part of the neurons is located in the head around the pharynx and will be our focus of attention.

Parkinson's Disease (PD) belongs to the group of diseases characterized by problems in movement ability. These problems are not related with spasticity, motor weakness, sensorial loss or ataxia but solely with neurological degradation. They are characterized by movement backwardness, increase in tonus or abnormal involuntary movements and they occur when there is a

Manuscript received May 11, 2016; revised August 22, 2016.

doi: 10.18178/joig.4.2.99-104

disfunction on subcortical cerebral structures, especially in the basal ganglia [4].

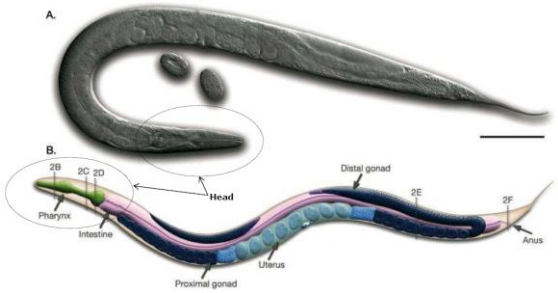

Figure 2. Anatomy of an adult nematode. A. DIC image of an adult, left lateral side. Scale bar 0.1 mm. B. Schematic drawing of anatomical structures, left lateral side (adapted from [5]).

There are four fundamental characteristics in Parkinson disease that can be grouped under the acronym TRAP: Tremor at rest, Rigidity, Akinesia and Postural instability [6]. Different rating scales are used to evaluate motor impairment and disability in patients with PD but most of them have not been fully validated. The Unified Parkinson's Disease Rating Scale is the most well established scale to assess disability and impairment. Studies using this scale to track the progression of PD suggest that its course is not linear and that the rate of deterioration is variable and more rapid in the early phase of the disease.

There are at least nine different genes known to cause Parkinson disease. Functional studies with genes related with PD implicate mitochondria disfunction, protein aggregation and the stress response in its pathophysiology. Autosomal-dominant mutations in leucine-rich repeat kinase 2 (LRRK2) are associated with PD. The G2019S mutation is one of the most commons and is present in different populations. Some studies demonstrate that G2019S mutation lead to increased LRRK2 auto-phosphorylation and kinase activity [7], [8]. Mutant forms of LRRK2 in neurons cause cell death and also decreases neurite length in rodents [9]. The protein LRRK1 is the homolog of LRRK2 in Mammals. Biochemical evidence indicates that mutations are connected to activities of aberrant kinase and that these modifications can be responsible for neural toxicity and pathogens of Parkinson's disease.

Since there is no therapeutical treatment to reduce speed or improve dopaminergic neuron degeneration in PD, studies are being conducted in organisms to simulate the influence of certain proteins and disfunctions in the way the disease evolves. Caenorhabditis elegans is an excellent organism to model neuronal degenerative diseases and is being used to study the degeneration mechanism on PD models and other neuro-degenerative diseases. The LRRK2 gene, in C. elegans, is highly preserved between species and its over-expression may become a valuable model for preclinical testing of compounds for the treatment of PD. Some observations support the hypothesis that LRRP2 is responsible for modeling mitochondrial function that is associated with hereditary PD. Transgenic expression of LRRK2 and the disease-causing mutations of LRRK2 in C. elegans results in loss of dopaminergic neurons and behavioral deficits. These transgenic models recapitulated several key features of human parkinsonism, indicating that over-expression of LRRK2 in C. elegans can provide a valuable model for preclinical testing of compounds for PD. Experiments with C. elegans can, therefore, contribute to better understand what factors most influence the evolution of PD and what kind of treatment may be produced to alleviate the symptoms of this disease.

III. Automatic Detection of Cellular Bodies of C. elegan's Neurons

Although there are no previous works on image segmentation of C. elegans images like the ones used in our work, there are some works on feature extraction using image segmentation. In [10], for example, some image processing and analysis technics are used to extract grey images from video of C. elegans in order to produce a segmentation and to identify the body of the organism. In other work [11], the maximum intensity projection on labeled training data is used to extract the dendritic tree. In this work, we use the open-source software CellProfiler, to build a pipeline for automatic detection of neuron's cellular bodies in C. elegans images collected in INEB's laboratory under protocol conditions with the goal of evaluating neurons degeneration in C. elegans life cycle.

TABLE I. The C. elegans Growing Regime and the Protocol for Image Extraction

Days	Description
0	Synchronization
1	Seeding
2	Growing at 20 ℃
3	Growing at 20 ℃
4	Adult day 1
5	Growing at 20 ℃
6	Growing at 20 ℃
7	Adult day 4 – Image Acquisition
8	Adult day 5 – Image Acquisition
9	Growing at 20 ℃
10	Growing at 20 ℃
11	Adult day 8 – Image Acquisition

A. C. elegans's Images

In the laboratory, C. elegans nematodes were cultivated following a rigorous growing method during eleven days. Images were acquired in three different days from its adult life. The growing scheme and image acquisition protocol is depicted in Table I.

For image acquisition, the nematode is placed on a microscope slide and paralyzed using Azide Sodium.

Next, the slide is placed on an Axio Observer Z1 Feiss microscope with a 640× magnification and a 16 bit image is produced. The acquired images are then preprocessed in order to select the proper ones to serve as training samples. An example of such an image with a single nematode is shown in Fig. 3.

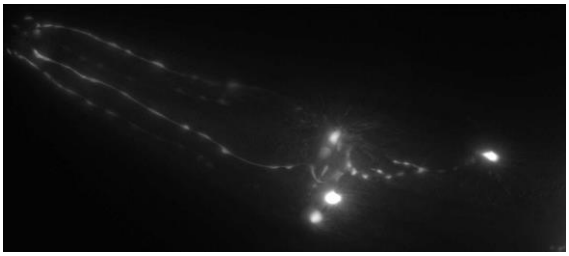

Figure 3. Example of a C. elegans image acquired with an Axio Oberver Z1 Feiss microscope.

Since the first goal of this project is to detect and count the cellular bodies of each nematode (image), the set of images used to train the system were annotated by an expert in this research area. An example of an annotated image is presented in Fig. 4 where we can see the cellular bodies highlighted by a white line (yellow line in color scale). The expert has performed this annotation on 72 images presenting nematodes in different positions and with different number of visible cellular bodies and degradation conditions.

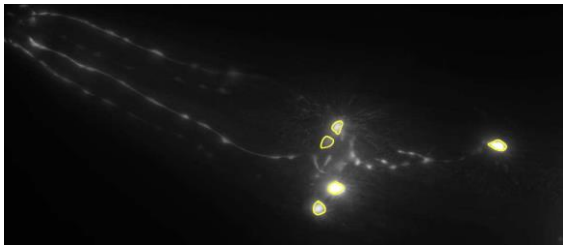

Figure 4. The annotation performed by the expert on Fig. 3 image.

Neuron detection on these images are usually manually performed by an technician and, therefore, subject to error. In the next subsection we explain how CellProfiler was used to perform an automatic analysis of the images.

B. *CellProfiler*

CellProfiller [12] is an open-source application extremely versatile used to analyse data in biological images. It was conceived to analyse images of high resolution and to measure the size, shape, intensity and texture of each cell (or other objects) in the image. Using a point-and-click interface, the user can build a pipeline by creating a sequence of modules each one performing a given image processing task like lightning correction, segmentation, image identification, measuring, etc. Users can mix and combine modules, adjusting its configurations, to measure the phenotype of interest. It can also be used with low resolution images. The pipeline used in this work is depicted in Fig. 5.

C. *Building the Pipeline*

The first stage of the pipeline it's a cropping action that allows the user to verify that in each image a single nematode is selected. This step can be automatic but the user can also, by manual inspection, refine the selected region. This cropping step introduces a normalization in the image that must be corrected so that it does not influence the following steps of the process. To achieve this we apply a threshold mask to select the region of interest. The obtained image is focus on the region of interest containing the cellular bodies that we want to detect.

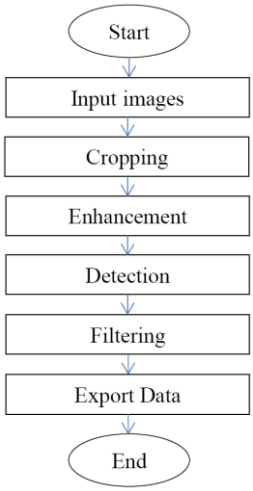

Figure 5. The CellProfiller pipeline built for this problem.

After the cropping step, the following stage of the pipeline should be the detection of the cellular bodies. We tried to detect them directly on the original images using CellProfiller's module *IdentifyPrimaryObjects* but the results were not satisfactory. In Fig. 6 one can see the extremely high number of objects detected on an original image.

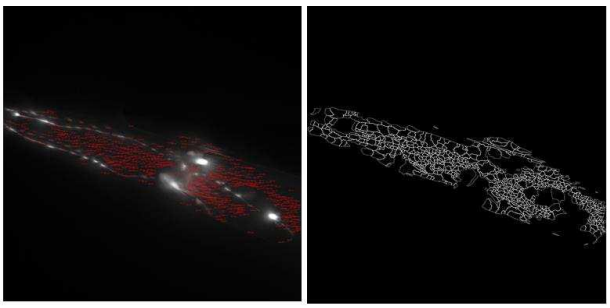

Figure 6. An example of the detection of primary objects in an original, without any treatment.

The next action was to improve the images in order to highlight the cellular bodies to facilitate its detection. Since cellular bodies are circular shaped, we tried to use the module *EnhanceOrSupressFeatures* to increment the quality of circular objects, together with the module *EnhanceEdges* to enhance the borders of the objects. The module *EnhanceOrSupressFeatures* was used with its enhance function that is responsible for enhancing the intensity of certain pixels relative to the rest of the image. This module can enhance three kinds of features: speckles, neurites and dark holes. We choose to enhance the speckles because a speckle is an area of enhanced intensity relative to its immediate neighborhood. The

module enhances the speckles using a white *tophat* filter, which is the image minus the morphological grayscale opening of the image. The module takes as tuning parameter the feature size, in this case being the diameter of the largest speckle to be enhanced, which will be used to calculate an adequate filter size. The module *EnhanceEdges* enhances the edges of a grayscale image. We can select several edge-finding methods such as the Sobel, Prewitt, Roberts, LoG or the Canny. There is a threshold parameter that one can select or choose to be automatically computed. The module *EnhanceOrSuppressFeatures* produces a grayscale image that is used as input for the *EnhanceEdges* module which output is further used to identify objects.

We performed several experiments with different edge finding methods and different values for the tuning parameters but the results were not satisfactory. An example of the resulting image of the *EnhanceEdges* module is shown in Fig. 7 (top), where we can see that several object were bounded together and some cellular bodies were not detected. In Fig. 7 (bottom), we can see the results of the detection not using the module *EnhanceEdges*. The results are better but there are some false positives and the borders of the detected objects are far apart from its original place.

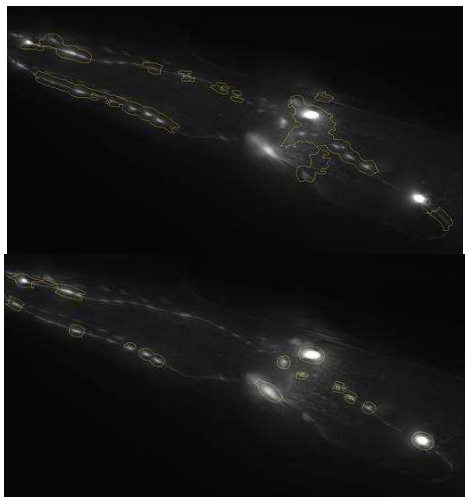

Figure 7. Objects detected using two different enhancement strategies.

To improve these results, we focus only on subject each image to an enhancement stage to improve the speckles that are present in the image. This is done because the cellular bodies are characterized by areas of higher intensity and this step will facilitate its detection. In Fig. 8 we can see an example of an original image (left) and the result of the enhancement step (right).

The module *IdentifyPrimaryObjects* was applied to these enhanced images. This module is usually used to detect cell's nucleus but, in this case, we use it to detect the neuron's cellular bodies because both structures present analogous characteristics. Although one can use the default options for this module, there is a very important parameter that one has to tune to achieve the best results on detecting the cellular bodies: *the typical diameter of objects, in pixels*. This parameter will be responsible to define the size of the globular shapes detected by the module. In preliminary experiments, using 30 pixels as diameter, there were a considerable number of cellular bodies that were not detected, making us reduce the smallest number of pixels to detect an cellular body to 10 pixels. After several experiments with the given images, the interval [10, 100] was the one that showed to present the best results. Approximately 85% of the cellular bodies were identified using this interval. However, for being so wide, this interval of values produce some false positives and, therefore, a filtering stage must be applied to reduce them. There are two options in the filtering step: automatic (default) and manual.

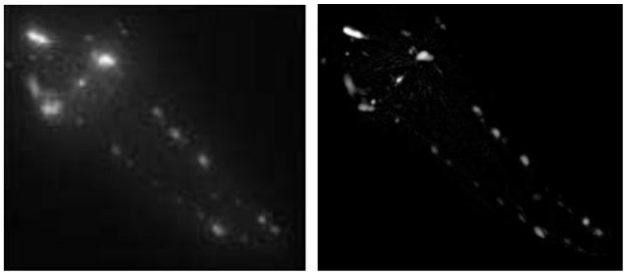

Figure 8. An example of an original image (left) and the resulting image after speckles enhancement (right).

Automatic filtering is performed using two filters: one that filters objects by area and another that filters objects by eccentricity. These two filters are applied sequentially and receive as input the result of a previous module that measures objects size and shape (*MeasureObjectSizeShape*). The decision to filter objects by area and eccentricity was taken after several experiments on the data using these and several other options like the perimeter, solidity, extent, Euler number or compactness of the detected objects. These experiments were made using the cross-validation approach. This approach consists on splitting the data in several groups, use the images in one group to tune (train) the filter's parameters for the best results (detecting all cellular bodies with a minimum number of false positives) and test the same parameter's values in the remaining images to evaluate its performance. We perform this process using as training images a group of 10 images that were chosen randomly and the training was repeated 10 times. The final values used in filtering are displayed in Table II.

TABLE II. AREA AND ECCENTRICITY VALUES USED IN THE FILTERING MODULE

	Area	Eccentricity
Min	328	0.334
Max	2370	0.926
Mean	1278.4	0.68

Manual filtering is an option for the user to manually remove false positive detections. An image with the original nematode and the detection produced by the pipeline is presented to the user and he can delete objects that do not correspond to cellular bodies. This will help to improve the results with the cost of some extra time. This

manual intervention only eliminates false positives but do not interfere with the characteristics of the true positives. Although this option is implemented in the pipeline we recommend to use it only if the number of cellular bodies is much higher than the one expected. This can occur if the images are obtained in deficient conditions.

In the final stage of the pipeline, the module *ExportToSpreadsheet* is used to export the data corresponding to the characteristics of the detected objects. These characteristics can be selected by the user among those available in this module.

To build this pipeline, we performed a considerable number of experiments with a variety of modules and parameters. All the results were evaluated according to the final objective of detecting the cellular bodies and extracting its characteristics such as size, intensity or shape. We compared the detected objects with the experts annotation. This process is described in the following section.

IV. Analysing Data

To compare the results of the detected objects against the expert annotation, we use the Dice similarity coefficient [13], [14],

$$D = \frac{2(I \cap A)}{I \cap A + I \cup A} \quad (1)$$

This coefficient is mostly used to compare segmentations of a given image and measures the spatial overlap between two binary images, I and A. It varies from 0 (no overlap) to 1 (total overlap agreement). In this work, I corresponds to an image produced by the pipeline and A to the correspondent image annotated by the expert.

To perform these comparisons we use Matlab [15] to build an algorithm that takes as input the images with the outlines of the detected objects and the images with the outlines of the cellular bodies annotated by the expert. The algorithm counts the number of objects in each image and it also fills the regions inside the outlines of both images in order to compare them and to evaluate its spatial overlap. In Fig. 9 we show an example of the outlines produced by the pipeline and by the expert in the example presented in Fig. 4. We can see that the number of objects and its location is the same with only some small differences in the shape and size between the detected objects (Fig. 9, left) and the annotated ones (Fig. 9, right). This difference is especially visible when comparing the smaller detected object with its respective expert's annotation.

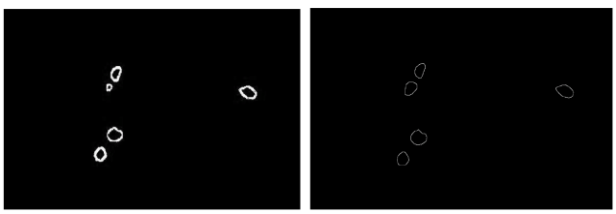

Figure 9. An example of the outlines produced by the pipeline (left) and by the expert (right) in the example presented in Fig. 3 (focus on the cellular bodies region).

The results using the automatic filtering and detection mentioned in the previous section are resumed in Table III. We present the number of images, the total number of cellular bodies annotated by the expert, the total number of objects detected by the pipeline and the mean of the Dice coefficient of all the comparisons.

We can see in the table that the number of detected objects is higher than the real number of cellular bodies, showing the difficulty of this task due to the similarity between cellular bodies and other artifacts present in the image. In this case, the manual filtering would be advised.

TABLE III. Results of the Application of the Automatic Detection and Quantification on the Nematodes Data Set

Number of Images	Total number of Cellular Bodies annotated by the expert	Total number of detected objects	DICE coefficient mean
72	277	381	0.73

V. Conclusion

The purpose of this work was to build an automatic detection of neuron's cellular bodies in C. elegans images. The usual process to quantify the characteristics of the cellular bodies is to perform, with a specific software, a manual annotation followed by some calculations like intensity, sphericity, counting, etc. This manual process is extremely slow and subject to errors. The automation is able to expedite the process and to alleviate the burden of making it manually.

The study of the degradation of neurons cellular bodies and neurites in C. elegans will contribute to the study of the phenomena related with Parkinson disease. It has been proven that there is some influence of certain proteins and disfunctions in the way the disease evolves and consequently how neurons degrade and this work is an additional contribution to help researcher on finding a therapeutical treatment to neuron degeneration in Parkinson's disease.

The limited number of images affected the performance of the implemented pipeline but we expect to tune it better when more images are available. Nevertheless, the presented results are very encouraging, motivating us to continue this work.

Acknowledgment

Authors would like to thank Conceição Pereira, from INEB, for making available the data set and for her collaboration as an expert on annotating the original images.

References

[1] S. Brenner, "The genetics of behaviour," *Br. Med. Bull.*, vol. 29, no. 3, pp. 269–271, 1973.
[2] D. Riddle, T. Blumenthal, and B. Meyer, *C. elegans II*, Cold Spring Harbor Laboratory Press, 1997.
[3] A. Diogo and A. Mota, "Caenorhabditis elegans: Modelo biológico para o século xxi," *Biologias*, vol. 4, pp. 1–8, 2001.

[4] E. Pereira and T. Aziz, "Parkinson's disease and primate research: Past, present, and future," *Postgraduate Medical Journal*, vol. 82, no. 967, pp. 293–299, 2006.
[5] Wormatlas. (2006). Introduction to c. elegans anatomy. [Online]. Available: http://www.wormatlas.org
[6] J. Jankovic, "Parkinson's disease: Clinical features and diagnosis," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 79, pp. 368–376, 2008.
[7] A. West, *et al.*, "Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity," *Proc. Natl. Acad. Sci. USA*, vol. 102, p. 1684216847, 2005.
[8] E. Greggio, *et al*., "Kinase activity is required for the toxic effects of muta nt lrrk2/dardarin," *Neurobiol. Dis.*, vol. 23, pp. 329-341, 2006.
[9] D. Macleod, J. Dowman, R. Hammond, T. Leete, K. Inoue, and A. Abeliovich, "The familial parkinsonism gene lrrk2 regulates neurite process morphology," *Neuron*, vol. 52, pp. 587-593, 2006.
[10] K. L. Moy, "C. elegans segmentation," Tech. rep., 2014.
[11] A. Greenblum, *et al.*, "Dendritic tree extraction from noisy maximum intensity projection images in C. elegans," *Biomed. Eng. Online*, vol. 13, no. 74, 2014.
[12] A. E. Carpenter, *et al.*, "Cellprofiler: Image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, pp. 1–11, 2006.
[13] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, pp. 297–302, 1945.
[14] A. Zijdenbos, B. Dawant, R. Margolin, and A. Palmer, "Morphometric analysis of white matter lesions in MR images: Method and validation," *IEEE Trans Med Imagin*g, vol. 13, pp. 716–724, 1994.
[15] *MATLAB*, version 7.10.0 (R2010a), Natick, Massachusetts: The MathWorks Inc., 2010.

Jorge M. Santos received his PhD in Engineering Sciences from the Engineering Faculty of Porto University, Portugal, in 2007. He is presently an Assistant Professor in the Department of Mathematics at the Engineering School, Polytechnic of Porto, and a member of INEB — the Biomedical Engineering Institute at Porto. His research focused on the application of entropic concepts in data classification where he has published several works, including a book. His current interests are related to pattern recognition and machine learning, and particularly with supervised classification with neural networks and deep learning, with applications to biomedical data.

Daniel Dias received is Master degree in Medical Computation and Instrumentation Engineering from the Engineering School, Polytechnic of Porto, Portugal, in 2015. He is currently enrolled in a second Master course in Computer Science from the Engineering Faculty of Porto University, Portugal.