Measuring the size of pores by the segmentation of images from scanning electron microscopy

MEASURING THE SIZE OF PORES BY THE SEGMENTATION OF IMAGES FROM SCANNING ELECTRON MICROSCOPY

A.C. Sparavigna[1], R. Pisano[1], A.A. Barresi[1]

[1] *Department of Applied Science and Technology, Politecnico di Torino, Torino, Italy*
*E-mail of the corresponding author: **antonello.barresi@polito.it***

Abstract: Segmentation is an image processing method used for partitioning an image into multiple sets of pixels, which are defined as its "super-pixels". Here, we are proposing a method based on segmentation, for determining the super-pixels corresponding to the pores resulting from freeze-drying of a pharmaceutical solution. The sizes of these pores, evidenced by the Scanning Electron Microscopy (SEM), are estimated through the areas of the super-pixels of the segmented image. These sizes can be used to estimate the resistance to mass transfer and hence optimise the cycle of production.

Keywords: freeze-drying, pore sizing, image segmentation, image processing, pharmaceutical solutions

Introduction

The freeze-drying process is currently used as a common industrial processing and preservation technique, applied to a wide variety of products, the large part of them being food, bio-products and pharmaceuticals. Consequently, freeze-drying is generating technological processes which have a high impact on industry and society.

Because of this fundamental role, the researches on freeze-drying processes are attracting the efforts of several scientific disciplines, rendering it the subject of a highly interdisciplinary area of studies. Let us consider, for instance, the scaffolding of a dried solution created by such processes. This structure is requiring models which can help us to understand the reasons of its formation as a consequence of concentrations and heat fluxes applied to the sample. However, models are based on observations from microscopy and tomography and were demonstrated to be useful for the determination of the resistance to vapor flow and hence for cycle optimisation. For this reason, the modelling method can receive a large benefit from the processing of the data coming from such analyses, and therefore from the discipline of the processing of the images obtained from the abovementioned resources.

Here we consider the observation of the pores in the lyophilized cakes, evidenced by the Scanning Electron Microscopy (SEM). The lyophilized samples are obtained from aqueous solutions of sucrose and dextran. The size of the pores can be estimated through the areas of the super-pixels obtained by a segmentation of the SEM images. The size of the pores can be used to estimate the resistance to mass transfer. Data on such a feature are relevant for optimising the cycle of the freeze-drying process.

Materials and Methods

To illustrate the approach, let us start using a lyophilized sample produced from an aqueous solution of sucrose, having 5% w/w as solid content. The solution was freeze-dried as discussed in [1]. In the Figure 1 (left), we can see a SEM image of the sample, which is

showing the presence of the typical pores created by the process. To estimate the size of the pores, the approach we follow is that of segmenting the image.

Before discussing the detail of the method, let us sketch it shortly. To obtain the segmentation, the SEM image, such as that given in the left panel of the Figure 1, must be pre-processed to enhance brightness and contrast. After, some further filtering is necessary: in the case of this image, it was smoothed by a Gaussian filter. Then, the image is mapped into a binary black and white map (Fig. 1, middle), which is segmented through a thresholding method [2]. The result of the segmentation is the partitioning of the SEM image into sets of super-pixels, which are represented in the Fig.1 (right), by the colored domains. Each super-pixel is characterized by a label, by means of which we can easily evaluate the area (in pixels) covered by the domain.

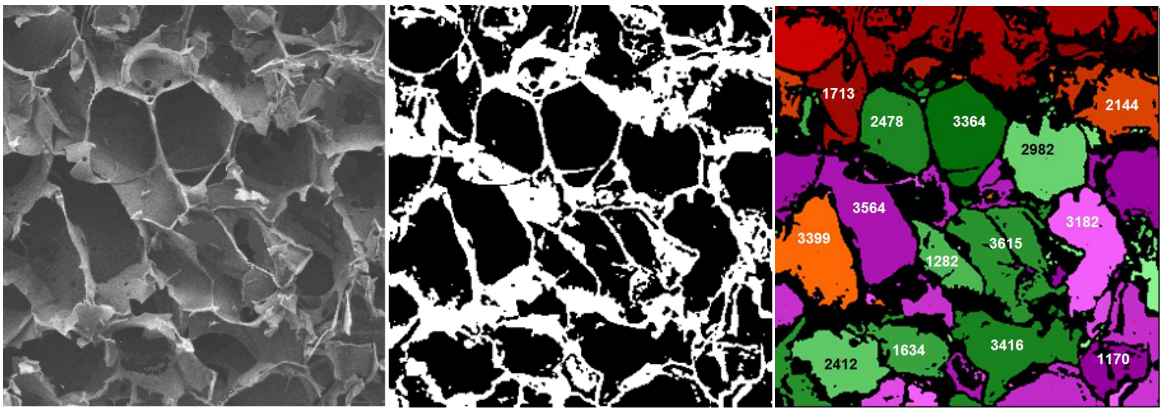

Fig. 1. A SEM image (left) is pre-processed to have the binary image (middle) used for segmentation. The segmented image (right) shows super-pixels (here rendered by different colour tones) and their area (in pixels). The image is 300 x 315 pixels (540 x 568 μm).

The area of the super-pixel is a measure of the observed cross-section of the considered pore. Then, the segmentation gives us a set of data reporting the areas of the several cross-sections given by the image. As in the example given above, the cross-section of pores is 2600 ± 400 (in pixels), where we have estimated the uncertainty by means of the standard deviation of the chosen sample. That is, we obtain a cross-section of (8000 ± 1200) μm^2, corresponding to a radius of (50 ± 8) μm.

In the following section, we will discuss the algorithm we used for segmentation. Then, we will apply it to a quite larger section of the sucrose cake, from which the Figure 1 has been obtained as a detail. When we have to analyse a large sample, the method used for the Figure 1, that is, that of determining the cross-section from the areas of some chosen pores is unreasonable, being it time-consuming and strongly dependent on the observer who is preforming the choice. Therefore, the results will be proposed as deduced by the counting of all the super-pixels.

Image segmentation

As previously told, in image processing, a segmentation is a process of partitioning an image into multiple sets of pixels, defined as super-pixels, in order to have a representation which is simpler than the original one and more useful to the following desired analyses [3]. For this reason, the segmentation of images is often used in many applications of image

processing, in particular in the medical image processing, where the aim is that of determining the presence of pathologies, and for the stacking of the maps coming from tomography to have the 3D reconstructions [4-5].
The typical use of the image segmentation is that of locate objects, or domains, and boundaries among them. Specifically, the segmentation is a process of assigning a label to every pixel in an image, such that the pixels having the same label share certain characteristics. As a consequence, the result of the segmentation is a set of "segments", or "super-pixels", that are covering the whole image. Several methods exist for segmentation, mainly based on the use of binary (black and white) images (see for instance, [6]).
In general, we have a colour RGB source image of $N_x \times N_y$ pixels, represented by the three-channel brightness function $b_c : I \to B$, where $I = [1, N_x] \times [1, N_y] \subset \mathbb{N}^2$ and $B = [0,255]^3 \subset \mathbb{N}^3$. From this function, a grey-tone map can be obtained by giving:

$$\tilde{\beta}(i, j) = \frac{1}{3} \sum_{c=1}^{3} b_c(i, j) \tag{1}$$

The index *c* corresponds to the three RGB channels. The integer indices *i* and *j* are ranging in the *x* and *y* directions of the Cartesian frame corresponding to the image frame. In this manner, we obtain a brightness map of grey-tone pixels. From this map, we can obtain some statistical parameters of the whole image.
The images that we are here considering are coming from a SEM instrument and therefore have already grey-tone pixels. Our aim is that of measuring the size of the pores, which are, in the SEM images, represented by dark or black pixels. In the segmentation of the image then, we move on with a method based on the thresholding of the brightness map. The thresholding is using a clip-level (the threshold value τ), according to which a grey-scale image is turned into a binary image *T*.

$$\begin{aligned} &\tilde{\beta}(i, j) \le \tau \to T(i, j) = 0 \quad \textit{black} \\ &\tilde{\beta}(i, j) > \tau \to T(i, j) = 255 \; \textit{white} \end{aligned} \tag{2}$$

The clip-level has a crucial role in the rendering of the binary image; often it is determined by means of the entropy, evaluated on the histogram of grey-tones of the pixels [7-9].
In the case of the SEM images we prefer the use of software, such as GIMP (GNU Image Manipulation Program, for X Windows systems) and make a visual choice of thresholding. Usually, it is necessary to equalize the histogram of the grey-tones of the SEM image, followed by an adjustment of the levels, before converting the image into a binary one. Sometimes, other tools of GIMP are necessary to have a further adjustment of the binary image, as we will discuss in the following given examples. Once we have obtained the binary image, we have at our disposal a matrix of pixels containing black and white domains. Starting from the left/upper corner of this matrix, we move following rows and columns of the matrix. We focus on black pixels and characterize each of them by a sequential integer number *k*, which is acting as a label of the single pixel. Some of these labels will be the labels identifying the domains, or super-pixels, to which the pixels belong (see Fig.2).

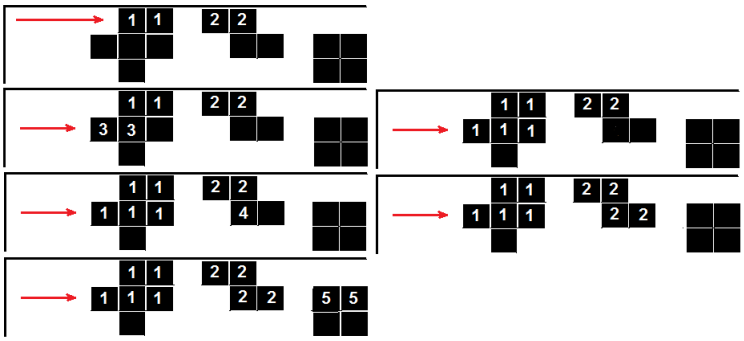

Fig. 2: The white pixels have a label k=0. Each black pixel has a label different from zero, with a value which is increasing as we move from left to right on the rows. However, the final values is fixed according to the labels of the black pixels above and on the left of the considered pixel**.**

The label of a black pixel is determined according to the labels k of the nearest black pixels above (A) and on the left (L) of the considered pixel. If the labels k_A and k_L are the same, their value is the label of the pixel under examination. If these labels are different, the pixel assumes as its label the lower value among them. Then, all the pixels having the label with the larger value change their labels into the value of the smaller one (see Figure 2). This approach can be easily obtained by logic instructions in any programming language (here we use Fortran 77). It gives to each of the super-pixels a different label. After this procedure, we have a matrix of labels:

$$\begin{aligned} &T(i,j)=255 \rightarrow K(i,j)=0 \\ &T(i,j)=0 \quad \rightarrow \quad K(i,j)=k,\ k\neq 0, k\subset \mathbb{N} \end{aligned} \tag{3}$$

Then, a new map can be proposed, where a colour tone is associated to such labels and therefore to each super-pixel, such as in the example given in the Figure 1. Since each black pixel of the original grey-tone image has the label of the super-pixel to which it belongs, we can easily do some calculations. First of all, we have the number of the super-pixels. Then, for each of them, we can give the number of pixels, that is, the area (in pixels) the super-pixel is covering, as given in the example of Figure 1.

Results on a SEM image of a sucrose cake

Let us consider the section of the sucrose cake (Figure 3), from which the detail given in the Figure 1 has been deduced. Again, we determine the thresholding by means of GIMP, with a visual inspection of the SEM image, to obtain the binary image from the grey-tone original one. A smoothing is required, to avoid the presence of a multitude of domains containing a few of black pixels. The smoothing, that in the case of the Figure 1 was obtained through a Gaussian filter, can be made by other filters too. For instance, quite useful is the GIMP filter which is simulating the brushstrokes of oil painting.

Sometimes, the filters applied to remove the small domains are also removing the thin walls between pores. To restore them, we can use the binary watershed tool of ImageJ, a public domain Java-based image program. From the SEM image of the sucrose cake (Figure 3), using the abovementioned tools of GIMP and ImageJ, the binary image given in the right panel of the Figure 3 is obtained.

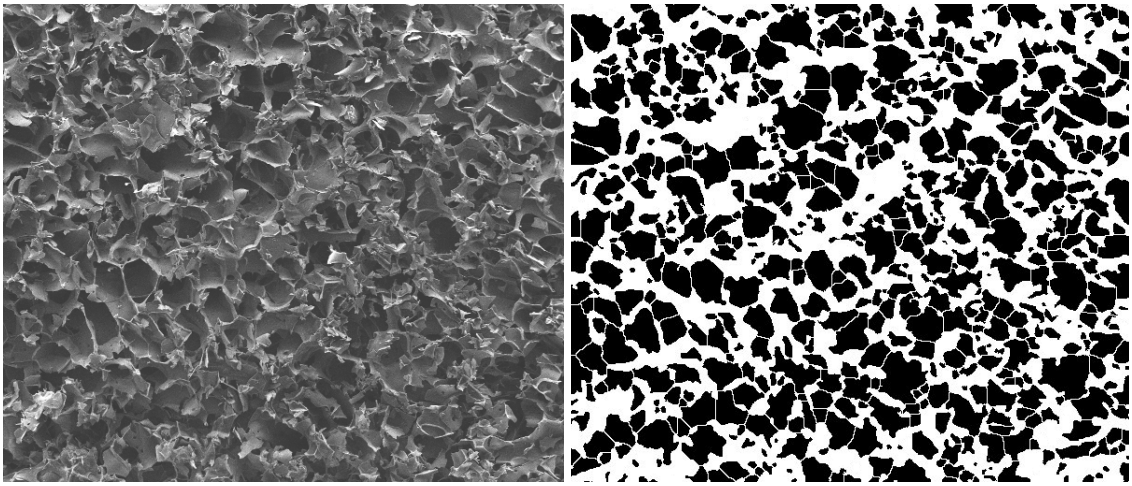

Fig. 3: SEM image (left) of a cross-section of the sucrose cake. On the right, the binary image filtered by the tool of GIMP which is simulating the oil painting, with the thin walls between pores restored by the Watershed tool of Image J. The image is 600 x 511 pixels (1785 x 1520 μm).

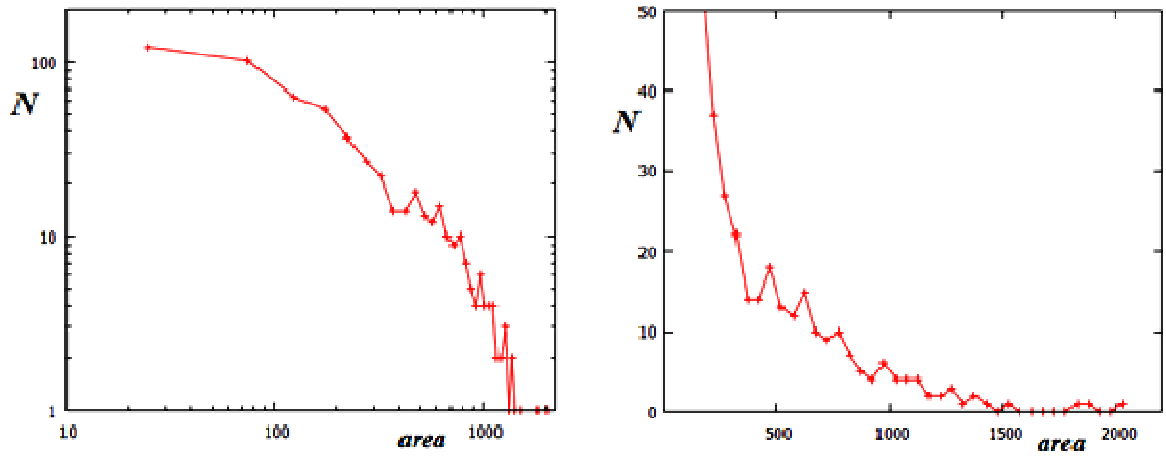

Fig.4 : Distribution of the super-pixels. by counting them according to their area (in pixels) within intervals spaced of 50 pixels.

As we have segmented the image, it is possible to deduce the area of each pore, because it is the area of the corresponding super-pixel. Then, we can plot the distribution of the super-pixels, and consequently the occurrences *N* of them, for given areas of the pores within intervals spaced of 50 pixels (Figure 4). It is clear than many small pores could be an artifact of the segmentation; for this reason, it is necessary to determine a threshold for the results we would like to consider. To help us in the choice of the threshold on occurrences, they are given in log scale too. From the plots, we can see that the pores we observe in the image have areas comprised between 500 and 1500 pixels. Using the scale provided by the scan from SEM instrument, we have areas comprised between 4400 μm^2 and 13000 μm^2.

Results from other cakes

Of course, the method we have proposed in the previous sections can be applied to other samples. Here we give two examples of image segmentation. The first concerns the section of a 5% dextran cake. ….. In the Figure 5 we can see, on the left, the original SEM image and, on the right, the binary image, processed by means on GIMP and ImageJ. In the Figure 6, the results of calculations are given.

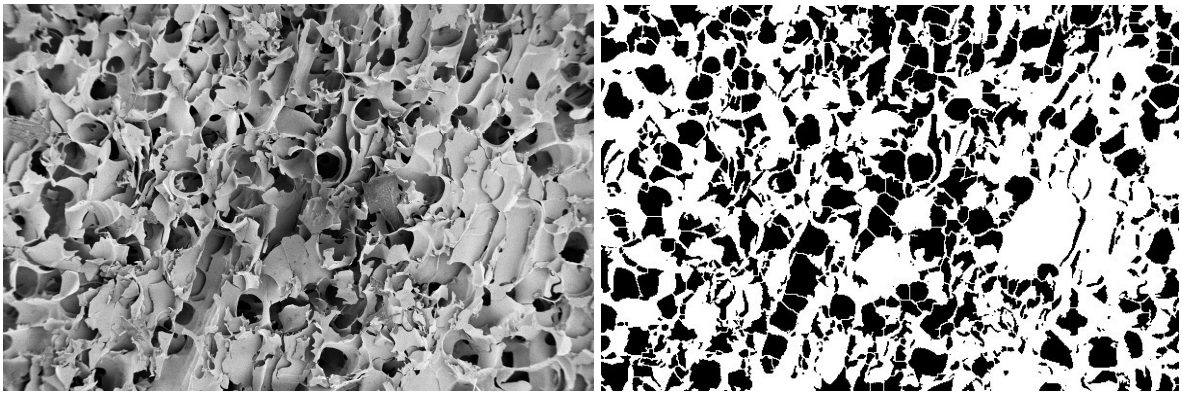

Fig. 5: SEM image (left) of a cross-section of a (A8_sect-b1-2013-11-27) 5% dextran cake. On the right, the binary image filtered by the tool of GIMP which is simulating the oil painting, with the thin walls between pores restored by the watershed tool of ImageJ. The image is 600 x 397 pixels (1875 x 1240 μm).

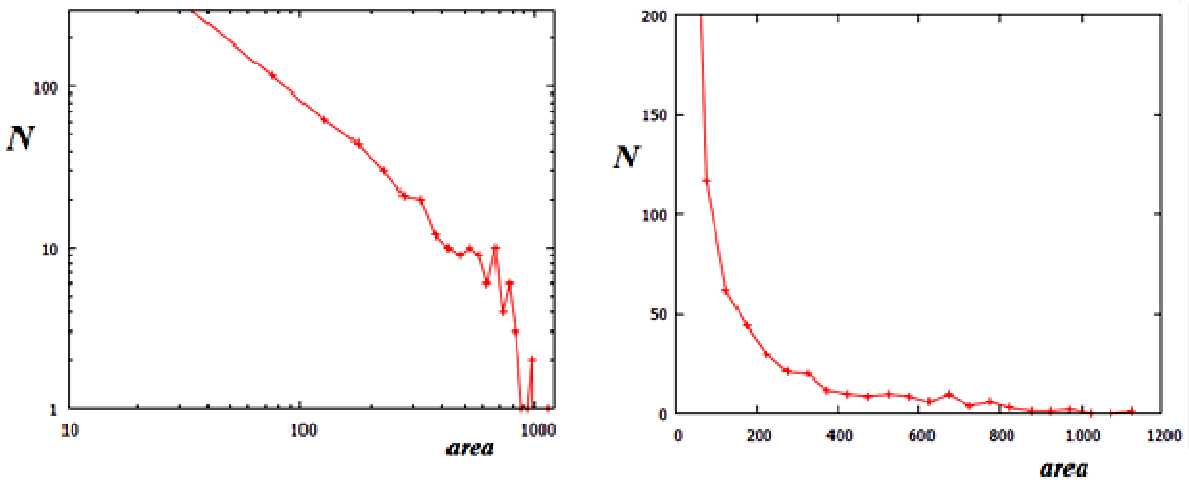

Fig.6 : Distribution of the super-pixels, by counting them according to their area (in pixels) within intervals spaced of 50 pixels.

As we have segmented the image, we deduce the area of the pores. We plot the distribution of the super-pixels, and the occurrences of them, for given area of the pores within intervals spaced of 50 pixels (Figure 6). From the plots, we can see that the cross-sections of the pores we observe in the image have an area comprised between 200 and 800

pixels. Again, using the scale provided by the scan from SEM instrument, the areas are between 1950 μm^2 and 7800 μm^2. It corresponds to radii from about 25 μm to 50 μm.

Here another example. It is the section of a cake of 1% mannitol + 4% dextran ….

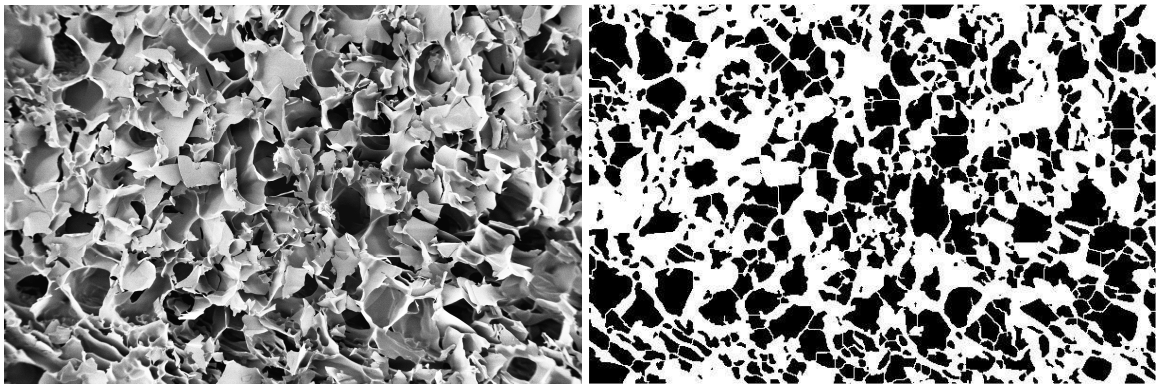

Fig. 7: SEM image (left) of a cross-section of a (D7_sect-c1-2013-11-19) 1% mannitol + 4% dextran cake. On the right, the binary image filtered by the tool of GIMP which is simulating the oil painting, with the thin walls between pores restored by the watershed tool of ImageJ. The image is 600 x 400 pixels (1875 x 1250 μm).

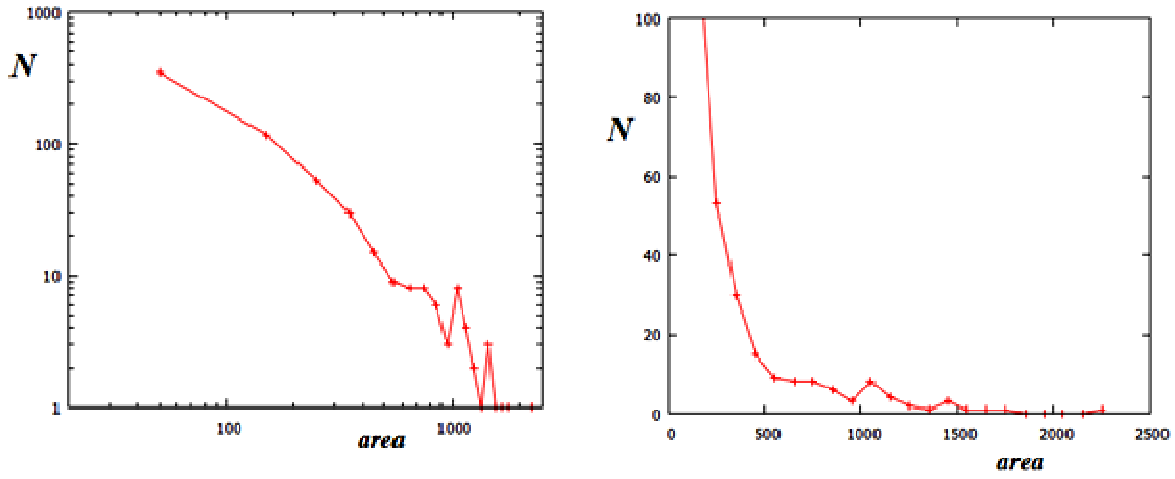

Fig.8 : Distribution of the super-pixels, by counting them according to their area (in pixels) within intervals spaced of 50 pixels.

The segmentation of the image gives us the area of the cross-section of the pores. We plot the distribution of the super-pixels, and the occurrences of them, for given area of the pores within intervals spaced of 50 pixels (Figure 8). Let us note in this case, a peak about 1200 pixels. This corresponds to a cross-section of 11700 μm^2, that is, to a radius of about 60 μm.

Conclusions

As we have illustrated, the study of the pores of lyophilized samples can be improved by methods of image processing. Here, for instance, we have determined the segmentation of the super-pixels which are containing the pores, in order to estimate their size. Let us add that the

approach we have used, the segmentation, is the first step for any further 3D reconstruction of the sample from images given by tomography. Actually, a 3D analysis is requiring the registration of such images, an operation usually performed through the alignment of super-pixels. Knowledge of the average pore size and its distribution is essential for the determination of the resistance to mass transfer and hence for the use of the most recent model-based tools for the development and scale-up of freeze-drying cycles.

References

[1] Pisano R., Barresi, A.A., Capozzi, L.C., Novajra, G., Oddone, I., Vitale-Brovarone, C. (2016). Characterization of the mass transfer of lyophilized products based on x-rays micro-computed tomography images. Drying Technology, DOI: 10.1080/07373937.2016.1222540.

[2] Zhang, Y. J. (1996). A survey on evaluation methods for image segmentation. Pattern recognition, 29(8), 1335-1346.

[3] Shapiro, L. G., & Stockman, G. C. (2001). Computer vision, New Jersey, Prentice-Hall, ISBN 0-13-030796-3

[4] Pham, D. L., Xu, Chenyang, & Prince, J. L. (2000). Current methods in medical image segmentation. Annual Review of Biomedical Engineering. 2: 315–337. DOI: 10.1146/annurev.bioeng.2.1.315. PMID 11701515.

[5] Forghani, M., Forouzanfar, M., & Teshnehlab, M. (2010). Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation. Engineering Applications of Artificial Intelligence, 23 (2), 160–168. DOI: 10.1016/j.engappai.2009.10.002

[6] Schladitz, K. (2011). Quantitative micro-CT. Journal of Microscopy, 243 (2), 111–117. DOI: 10.1111/j.1365-2818.2011.03513.x

[7] Gull, S.F., & Skilling, J. (1984). Maximum entropy method in image processing. Communications, Radar and Signal Processing, IEE Proceedings F, 131(6), 646-659.

[8] Sahoo, P.K., & Arora, G. (2006). Image thresholding using two-dimensional Tsallis – Havrda – Charvát. Pattern Recognition Letters, 27, 520-528. DOI: 10.1016/j.patrec.2005.09.017

[9] Sparavigna, A. C. (2015). Tsallis entropy in bi-level and multi-level image thresholding. International Journal of Sciences, 4(1), 40-49. DOI: 10.18483/ijsci.613