

# Automatic Growth Detection of Cell Cultures through Outlier Techniques using 2D Images

P.A. Gagniuc, C. Ionescu-Tîrgoviște, C. H. Rădulescu

## Paul Aurelian Gagniuc

Department of Genetics, University of Bucharest, Romania  
E-mail: paulgagniuc@yahoo.com

## Constantin Ionescu-Tîrgoviște

National Institute of Diabetes,  
Nutrition and Metabolic Diseases "N.C. Paulescu", Romania  
E-mail: cit@paulescu.ro

## Clara Hortensia Rădulescu

National R&D Institute for Textile and Leather, Romania  
E-mail: clarabios@yahoo.com

### Abstract:

Using conventional statistics, we have developed a new method for cell culture analysis through outlier detection techniques. Statistical methods enable researchers in microbiology to identify experimental parameters that are critical for colony growth and inhibition. This paper reports a method for analysing 2D images of cell cultures in Petri dishes, such as fungi, bacteria or yeast. The aim of this study was to obtain a sensitive and robust method for detection of growth rate, surface coverage and the approximate number of cells in the colony. For testing we have implemented a software application called MoldATRIX. This software generates useful statistics and displays critical information about the cell colony area. Our results were obtained by analyzing a series of digital images of *Aspergillus niger* cultures at different time intervals. Moreover, our results show the behavior of *Aspergillus niger* on leather.

**Keywords:** outliers, cell cultures, biodegradation, growth detection.

## 1 Introduction

New textile materials are conceived in order to increase their biodegradability properties [1]. For research, small quantities of bacteria or fungal cells are usually grown in a petri dish on a solid support that contains embedded nutrients (agar gel with a particular mix of nutrients), and new textile samples. Molds extract energy from organic matter in which they live and some of them play an important role in biodegradation [3, 4], food production, antibiotics [2] or enzyme synthesis. These microorganisms can be found in divisions such as Ascomycota, Deuteromycota or Zygomycota. Usually, molds secrete hydrolytic enzymes [4–6] which degrade complex biopolymers [8] such as cellulose [2, 3] or starch into simpler substances. Therefore, molds play a major role in decomposition of organic material. There are many known species of molds (eg. pathogens, aquatic species, thermophiles), but their impact is not known for new products (ie. new textile materials). Accordingly, their behavior should be measured quickly by different laboratories. The criteria for evaluation of the behavior of some leather samples [8] toward the fungi [1, 6] is represented by the percentage of the surface coverage with mycelium. Nevertheless, the EN 14 119 standard provides a table of correspondence for this percentage. Yet, there is no indication in this standard of a technique for measuring the coverage of the mycelium surface. The immediate option of the laboratory operator is a subjective visual approximation of the mycelium percentage from the total surface of the sample. Yet, this remains an empirical evaluation. Biodegradation process is currently studied for textile industry [7, 14]. Many leather

materials are tested with different mold cultures in order to characterize the type of fabric and the impact it has on the environment or the behavior in special conditions [16] (eg. extreme temperatures, long-term resistance). Parameters such as rate of growth, surface coverage or the number of cells are very difficult to quantify. We have developed an algorithm based on outlying observations [17, 18] which can extract data from pictures (photos) of the mold culture surface [19], made at different time intervals. An outlying observation is defined as an anomaly that appears to deviate from other members of the sample. Outliers may indicate data points that belong to a different population than the rest of the sample set. Identifying an observation as an outlier depends on the underlying distribution of the data. For testing we have implemented a software application called MoldATRIX (Supplementary material 1), which can work with bitmap or JPEG (Joint Photographic Experts Group) files taken from a digital camera. The aim of this paper is to provide a method of measurement for biodegradation, in order to certify the integration of new materials. Continuing the work presented in [20] with additional results and several refinements, we show a novel method for detection of colony area, growth rate and the approximate number of cells in the colony.

## 2 Materials and Methods

The European standard EN 14 119:2003 "Testing of textiles - Evaluation of the action of microfungi" was used in order to test the behavior of some leather samples towards the action of the filamentous fungus *Aspergillus niger* IMI 045551 [21]. Agar plates are frequently used in microbiology. Agar medium with mineral salts and a carbon source (glucose) was used as a culture medium. The medium was sterilized in an autoclave at  $115^{\circ}\text{C}$  for 30 minutes. In order to obtain the inoculum (cell suspension), a fresh culture tube was used. The surface of the culture was gently scraped, and then the spores were washed and centrifuged twice in a EBA 21 centrifuge. Next, the medium was held at  $50^{\circ}\text{C}$  in a waterbath. The inoculum was mixed with the medium and then was poured in Petri dishes (of 90 mm diameter, with a depth of maximum 5 mm) for cooling at room temperature for 20 minutes. The leather samples were cut in 25 mm square form and placed in the center of the Petri dish. Next, all Petri dishes were incubated at a temperature of  $28 \pm 2^{\circ}\text{C}$  for 14 days. Throughout this period, the cell cultures were photographed at 3, 7 and 14 days.

### 2.1 Implementation

In order to detect the surface coverage on leather materials we designed an algorithm consisting of five main loops. First, we consider a digital image (of the mold sample - 255x255 pixels) as a square matrix  $A$  (Figure 1A). Each pixel color can be represented by a 24-bit or 32-bit RGB value. Therefore, each pixel can be considered an element with a value between 0 and 32 million. We then consider a void square matrix  $B$  (255x255 pixels) that stores the outlier values (Figure 1B). Initially, the algorithm requires a traversal through every pixel of the image (matrix  $A$ ) in order to store the pixel values inside variable  $e$ . The mean is obtained by dividing variable  $e$  to the total number of elements (pixels) from matrix  $A$ . Next, the algorithm calculates the variance, standard deviation and the Z-score for all pixels of matrix  $A$ . Outlier values are listed in matrix  $B$  depending on the Z-score value for each pixel of matrix  $A$ . If the Z-score for a pixel of matrix  $A$  exceeds a threshold value (user specified parameter), the program copies the pixel position and value into matrix  $B$ . Next, the algorithm calculates the density of pixels from matrix  $B$  and displays the results. The growth rate of the cell culture is calculated based on photographs taken at various time intervals. The program can store the results of tests carried out for other pictures of the same cell culture. Accordingly, surface coverage can be calculated by:

$$C = \left( \frac{100}{E} \right) \times OL \quad (1)$$

where  $C$  represents the surface coverage,  $E$  represents the total number of elements from matrix  $A$  and  $OL$  represents the number of outliers found. For our experiment, we used three intervals, namely  $C_1=3$  days,  $C_2=7$  days and  $C_3=14$  days.

In our test, a total of four pictures were analyzed. The first picture represents a control sample used to verify the average pixel color before the cell culture growth. The other three images are taken from  $C_1$ ,  $C_2$  and  $C_3$  intervals. In order to detect the culture growth rate, we used the difference between two images of the mold surface taken at different time intervals ( $G = C_{n-1} - C_n$ ). Nevertheless, the number of photos and the intervals between

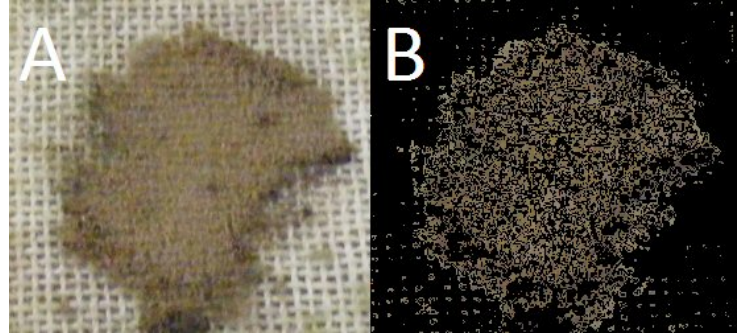


Figure 1: Textile material analysis. (A) a textile material image containing a mold culture (B) outlier pixels.

photo shots can be chosen by the user. The approximative number of cells is estimated by converting the number of pixels into millimeters. Thus, the result is multiplied by the number of outliers found and by cell density/millimeter (eg. yeast size can vary depending on the species, typically measuring  $3\mu m$  up to  $40\mu m$ ),

$$NC = \left( OL \times \left( \frac{P}{M} \right) \right) \times CM \quad (2)$$

where  $NC$  represents the total number of cells,  $OL$  represents the number of outliers found,  $P$  stores the number of pixels on  $y$ -axis or  $x$ -axis (number of rows or columns on the square matrix  $A$ ),  $M$  represents the size of the sample (millimeters) and  $CM$  represents the number of cells/millimeter. Depending on the chosen yeast or fungal species, the density of cells/millimeter is a parameter left to the user appreciation. Yeasts are unicellular, although some species may become multicellular through the formation of pseudohyphae or false hyphae (ie. molds). These unicellular and pseudohyphae yeasts lead to different densities of cells/millimeter, due to a different spatial positioning inside the colony (ie. formation of 3D structures that can not be neglected in some cases). For these special cases, a calculation of the colony surface is not an adequate solution.

## 2.2 Algorithm implementation

We aimed at assessing the effectiveness of the proposed algorithm by implementing a GUI application (Figure 3). Below we show the source code implementation of the algorithm, syntactically compatible with VBA, VBScript, Visual Basic 4,5,6, Visual Basic .NET and 2005.

```

1 Dim Mean, Variance, SD, Background_Mean, pixel As Double
2
3 Dim X, Y, X1, Y1, e, p1 As Double
4 Matrix_B.Cls
5
6 X1 = Matrix_A.ScaleWidth - 1
7 Y1 = Matrix_A.ScaleHeight - 1
8 e = 0

```

```

9
10 For x = 0 To X1
11     For Y = 0 To Y1
12         e = e + 1
13         Mean = Mean + Matrix_A.Point(X, Y)
14     Next
15 Next
16
17 Mean = Mean / e
18
19 e = 0
20 For x = 0 To X1
21     For Y = 0 To Y1
22         e = e + 1
23         pixel = control_pic.Point(X, Y)
24         Background_Mean = Background_Mean + pixel
25     Next
26 Next
27
28 Background_Mean = Background_Mean / e
29
30 e = 0
31 For x = 0 To X1
32     For Y = 0 To Y1
33         e = e + 1
34         pixel = Matrix_A.Point(X, Y)
35         Variance = Variance + (pixel - Mean) ^ 2
36     Next
37 Next
38
39 Variance = Variance / e
40 SD = Sqr(Variance)
41
42 For x = 0 To X1
43     For Y = 0 To Y1
44         p1 = Matrix_A.Point(X, Y)
45         zP = zP + p1
46         z = (p1 - Background_Mean) / SD
47
48         If (Abs(z) >= (1/100) * Int(Lim.LowerValue) And Abs(z) <= (1/100) * _
49             Int(Lim.UpperValue)) Then
50             Matrix_B.PSet (X, Y), p1
51         End If
52     Next
53 DoEvents
54 Next
55
56 zMare = (zP - Mean) / SD
57
58 For x = 0 To X1
59     For Y = 0 To Y1
60         q1 = q1 + 1
61         If Matrix_B.Point(X, Y) <> 0 Then q2 = q2 + 1
62     Next
63 Next
64
65 total_elements = (100 / q1)
66 procentage = total_elements * q2
67 CM = val(CM_txt.Text)
68 M = val(M_txt.Text)
69 P = Matrix_A.ScaleHeight
70 pm = P / M
71 NC = (q2 * pm) * CM

```

Variable `Background_Mean` represents the average pixel color of the control sample and `Mean` variable represents the average pixel color of the matrix *A*. The input objects, such as `CM_txt.Text` or `M_txt.Text`, are responsible for taking user input parameters. `CM_txt.Text` object stores the side length of a square (fabric size from the petri dish) and `M_txt.Text` object stores the amount of cells/millimeter. For the `Background_Mean` value, the user can also click with the mouse on the matrix *A* object to choose the cell culture approximative color value.

### 2.3 Equipment

We used a digital camera connected to a tripod. The distance between the petri dish and the digital camera was set to 25 centimeters (9.84 inch). Initially, for calibration reasons, the petri dish contained a ruler (the measuring unit was the centimeter) next to the sample. The ruler is important for further image processing on MoldATRIX software. The leather sample was cut from the picture at 25 mm by using the ruler from the image. Next, the sample area (of 25x25 mm) is stretched to 255x255 pixels and is processed by MoldATRIX program. However, the ruler is not a mandatory criteria. Another method consists of cutting a square-shaped sample (25x25 mm), thus avoiding the use of the ruler from the petri dish. An alternative method is to find the size of the sample directly from the photo by performing a pixel to millimeter transformation using dots or pixels per inch (dpi, 1 inch = 25.4 mm),

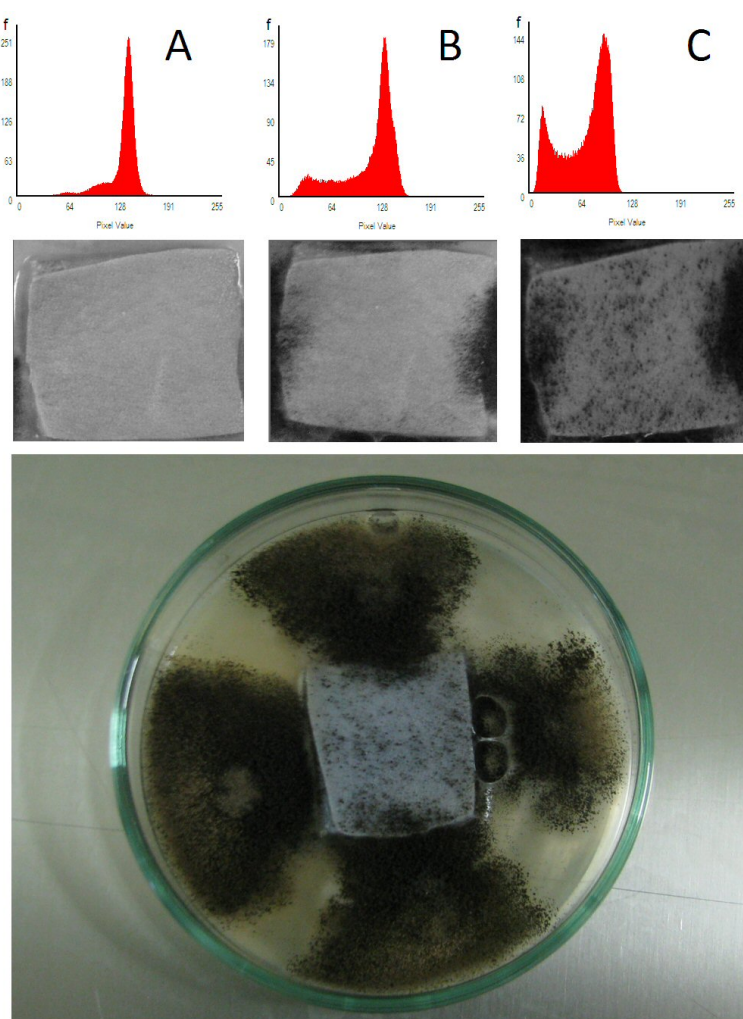


Figure 2: Results from the analysis of leather materials against *Aspergillus niger* cultures. The alignment of cell culture images following (A)  $C_1$  phase, (B)  $C_2$  phase and (C)  $C_3$  phase. Their corresponding histograms are presented at the top of the figure. The petri dish is shown at the bottom of the figure.

$$mm = \frac{(pixel \times 25.4)}{dpi} \quad (3)$$

MoldATRIX runs on all Windows operating systems, no installation required and the complete package has 9.44Mb (Supplementary material 1). MoldATRIX memory requirements are between 1.9Mb and 5Mb, depending on Windows OS version. MoldATRIX was tested on a computer equipped with an Intel P4 - 2800 MHz and 512 MB DDR dual channel.

### 3 Results and Discussion

An outlier is an observation that is numerically distant from the rest of the data. A common method for outlier detection is to look for observations that deviate more than three times ( $3\sum$  edit rule) the standard deviation from the mean. However, this is not the only criterion that could be used. It is also common practice to use Z-scores to identify possible outliers [22]. The proposed algorithm uses Z-scores and is tested on various images with different characteristics.

We also made additional tests to verify the correctness of the algorithm. Perhaps the most useful tool available in digital photography is the histogram [23, 24]. We generate histograms from a gray image sample of the cell culture by plotting the number of pixels for each color value. As shown in Figure 2, once the cell culture increases, the second peak appears in the histogram (Figure 2(A,B,C)). On the  $x$ -axis, the histogram shows color variations of the sample, while on the  $y$ -axis shows the number of pixels of that particular color. The left side of the histogram represents the black areas of the square sample, whereas the middle represents medium grey and the right side represents the pure white areas. For tests we used leather against *Aspergillus niger*. In  $C_1$  phase the

mold coverage accounted for approximately 2%. In  $C_2$  phase the mold coverage increased to 7% and in the last phase,  $C_3$ , the mold coverage has reached 22% (Figure 4). Growth rate is calculated according to each pre-final phase ( $G = C_{n-1} - C_n$ ). Table 1 shows the results of  $C_1 \dots C_3$  phases, where  $MC$  represents the mold coverage and  $MG$  represents the mold growth rate.  $MaxPV$  stands for the maximum pixel value, whereas  $MinPV$  represents the minimum pixel value (Supplementary material 2).

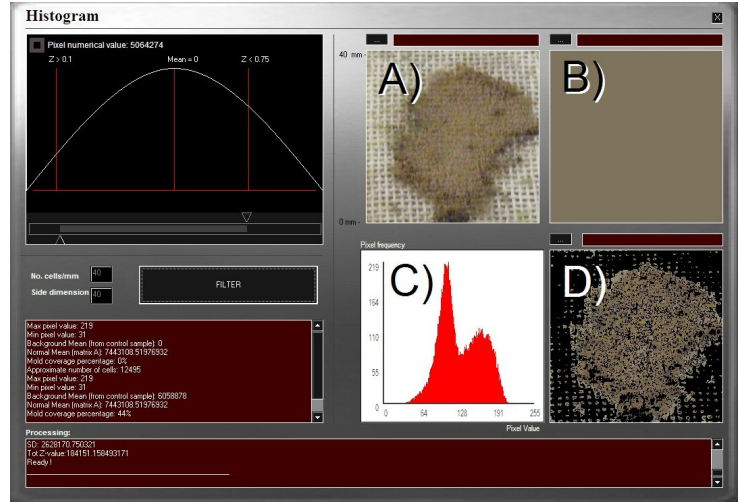


Figure 3: MoldATRIX program. On the right we show four matrices, (A) matrix A that contains the picture of the cell culture at a certain stage of growth ( $C_1$ ,  $C_2$  or  $C_3$ ), (B) represents the control sample, (C) represents the histogram of matrix A, (D) represents matrix B. On the left panel are the control parameters, results window and the status window.

$C_n$	$MC$	$MG$	$MinPV$	$MaxPV$
$C_1$	2%	2%	32	251
$C_2$	8%	6%	11	179
$C_3$	22%	14%	3	144

In MoldATRIX, users can define specific parameters, such as color differentiation (outlier pixels), sample size or number of cells/millimeter. For a better implementation of an experiment, users can create particular camera settings allowing different environments, such as illumination or color intensity. Nevertheless, in order to obtain clear and reliable results, the same initial settings must be preserved for each experiment.

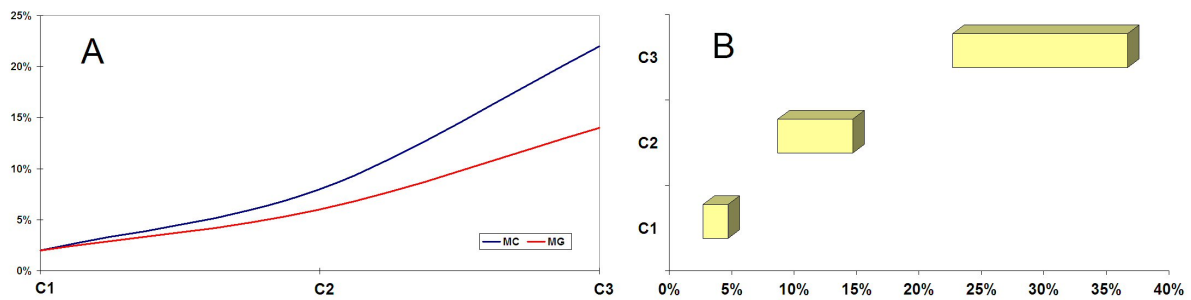


Figure 4: *Aspergillus niger* growth rate on leather. Two visualization methods are used: (A) blue line represents the mold coverage (MC) and the red line represents the mold growth rate (MG), (B) light yellow bars start from the MC value and show the mold growth for each phase.

Different methods, algorithms and automated systems have been proposed for cell culture studies. The outlier data detection may also be a potential solution for numerous studies performed on diabetes-related issues [25–28], such as electron microscopy images of b-cells or adipocyte cells. Some other possible applications are in the area of genomic signal processing [29–31], where outlier data usually exhibits biological significance.

**Notice:** Part of this paper was presented in a preliminary version as [20].

## 4 Conclusion and Future Works

Mold species within the genus *Aspergillus* have a large chemical range. Among scientists working on molds, there is an endless fascination with their biodeterioration potential. In this work, a new method was proposed in order to improve the quality of results for cell culture (ie. fungi, bacteria or yeast) studies on textile materials. Our immediate step is to use this method in cooperation with a neural network in order to detect different cellular structures on electron microscopy images. Future applications may include other areas, such as antibiotic research, hygiene research, contamination control or bioinformatics.

## 5 Acknowledgments

This research was partially funded by *Ministry of Education, Youth and Sports* through the PN2IDPCE2011-30429 project.

## Bibliography

- [1] J.W. Raymond, T.N. Rogers, D.R. Shonnard, A.A. Kline, A review of structure-based biodegradation estimation methods, *J Hazard Mater*, 84(2-3):189-215, 2001.
- [2] K.N. Bharath, R.P. Swamy, G.C. Mohan Kumar, Experimental studies on biodegradable and swelling characteristics of natural fibers composites, *International Journal of Agriculture Sciences*, Vol.2, No.1, pp. 01-04, 2010.
- [3] K.B. Raper, C. Thom, *A Manual of the Penicillium*, The Williams & Wilkins Co., Baltimore, 1949.

- 
- [4] S. Moularat, E. Robine, O. Ramalho, M.A. Oturan, Detection of fungal development in closed spaces through the determination of specific chemical targets, *Chemosphere*, 72, pp. 224-232, 2008.
- [5] S. Lehrer and A. Nowotny, Isolation and Purification of Endotoxin by Hydrolytic Enzymes, *Infection and immunity*, Vol. 6, No. 6, pp. 928-933, 1972.
- [6] W.E.S. Carr et al., The role of degradative enzymes in chemosensory processes, *Chem. Senses*, 15(2):181-190, 1990.
- [7] S.P. Vinogradova and S.N. Kushnir, Biosynthesis of Hydrolytic Enzymes during Cocultivation of Macro- and Micromycetes, *Applied Biochemistry and Microbiology*, Vol. 39, No. 6, pp. 573-575, 2003.
- [8] N. Lucas et al., Polymer biodegradation: Mechanisms and estimation techniques, *Chemosphere*, 73, pp. 429-442, 2008.
- [9] D. Montegut, N. Indictor, R.J. Koestler, Fungal Deterioration of Cellulosic Textiles: a Review, *Intern. Biodet. Bull.*, 28, pp. 209-226, 1991.
- [10] C. Aranyank, Microscopical Study of Fungal Growth on Paper and Textiles, *Proceedings of the 3rd International Conference on Biodeterioration of Cultural Property*, 4-7, Bangkok, Thailand, pp. 83-102, 1995.
- [11] C.H. Xue et al., Preparation of superhydrophobic surfaces on cotton textiles, *Sci. Technol. Adv. Mater*, Vol. 9, No.3, pp. 1-7, 2008.
- [12] K.H. Domsch et al., *Compendium of soil fungi*, Vol. 1 and 2, Academic Press, London, 1980.
- [13] J.C. Gilman, *A Manual of Soil Fungi*, Second Edition, U.S.A., 1975.
- [14] J. Szostak-Kotowa, Biodeterioration of Textiles, *Int. Biodeter. Biodegrad.*, 53, pp. 165-170, 2004.
- [15] D. K. Setua, G. D. Pandey, R. Indusekhar, G. N. Mathur, Biodeterioration of coated nylon fabric, *Journal of Applied Polymer Science*, 75(5):685-691, 2000.
- [16] E.D. Weil and S.V. Levchik, Flame Retardants in Commercial Use or Development for Textiles, *Journal of Fire Sciences*, 26(3):243-281, 2008.
- [17] M. Agyemang, K. Barker, and R. Alhajj, A comprehensive survey of numeric and symbolic outlier mining techniques, *Intell. Data Anal.*, 10:521-538, 2006.
- [18] V. Barnett and T. Lewis, *Outliers in Statistical Data*, John Wiley & Sons., 3rd edition, 1994.
- [19] D. J. Watts and J. M. Ashworth, Growth of Myxamoebae of the Cellular Slime Mould *Dictyostelium discoideum* in Axenic Culture, *Biochem. J.*, Vol. 119, pp. 171-174, 1970.
- [20] P. Gagniuc et al., Automatic growth detection of cell cultures through outlier techniques, GSP 2011, *2nd International Workshop on Genomic Signal Processing*, pp. 143-147, 2011.
- [21] K.A. Powell, A. Renwick, and J.F. Peberdy, *The Genus Aspergillus from taxonomy and genetics to industrial application*, Plenum Press in New York, 1994.



- [22] Richard J. Larsen and Morris L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, Third Edition, Prentice Hall, Upper Saddle River, USA, 2000.
- [23] H.O. Lancaster, *An Introduction to Medical Statistics*, John Wiley and Sons, 1974.
- [24] D.W. Scott, On optimal and data-based histograms, *Biometrika*, 66(3):605-610, 1979.
- [25] C. Ionescu-Tîrgoviște and F. Despa, Biophysical alteration of the secretory track in b-cells due to molecular overcrowding: the relevance for diabetes, *Integrative Biology*, 3(3):173-179, 2010.
- [26] C. Ionescu-Tîrgoviște, A short personal view on the pathogenesis of diabetes mellitus, *Proc. Rom. Acad., Series B*, 12(3):219-224, 2010.
- [27] A. Constantin & G. Costache, The emerging role of adipose tissue-derived leptin in inflammatory and immune responses in obesity: an update, *Proc. Rom. Acad., Series B*, (1):3-12, 2010.
- [28] N. Milici, A short history of the metabolic syndrome definitions, *Proc. Rom. Acad., Series B*, (1):13-20, 2010.
- [29] P.D. Cristea, R. Tuduce, *Use of Nucleotide Genomic Signals in the Analysis of Variability and Inserts in Prokaryote Genomes*, BIOCOMP 2008, pp. 241-247, 2008.
- [30] P.D. Cristea, Large scale features in DNA genomic signals, *Signal Processing*, 83(4):871-888, 2003.
- [31] P. Gagniuc et al., A sensitive method for detecting dinucleotide islands and clusters through depth analysis, *RJDNMD*, 18(2):165-170, 2011.