Semi-automatic model to colony forming units counting

Jesus Emilio Pinto Lopera[1,2], Diana Carolina Meneses Cabezas[1], Yuliana Zapata Serna[1], Yeison Alberto Garces-Gomez[1]
[1]Academic Training Unit in Natural Sciences and Mathematics, Catholic University of Manizales, Manizales, Colombia
[2]Systems Engineering, Faculty of Engineering, University of the Amazon, Florencia, Colombia

Article Info

ABSTRACT

Colony forming units counting is a conventional process carry out in bacteriological laboratories, and it is used to follow the behavior of bacteria in different conditions. Currently exist different systems, automatic or semi-automatic, to counting colony forming units exits, but, in general, many laboratories continue using manual counting, which consumes considerable time and effort from researchers and laboratory employees. This paper presents a mathematical model carry out to segment the colony forming units and, in this way, counting them from a digital image of the sample. The method uses the color space information of some points in the image and shows good behavior for images with many or few colony forming units in the sample, according to manual counting. The results show efficiencies close to 98% with MacConkey agar.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Yeison Alberto Garcés-Gómez
Academic Training Unit in Natural Sciences and Mathematics, Catholic University of Manizales
Calle 23 No. 60-63, Manizales, Colombia
Email: ygarces@ucm.edu.co

1. INTRODUCTION

The study of microorganisms is one of the most extensive fields of scientific research around the world. Bacteria are a particular case of microorganisms and can be used, for example, in the pharmaceutical and food industries, in medical diagnostics, and in quality control of drinking water [1]–[5]. A recurrent tool in the study of bacteria is the direct counting of colony forming units (CFU) in petri dishes. This process is employed in tracking bacterial growth under specific conditions and estimates the number of viable bacteria in the sample, which is a bacterial suspension on a culture medium in a petri dish. On the other hand, many laboratories around the world carry out the count of bacterial colonies process manually yet, which is a problem because consumes considerable time and effort from researchers and laboratory employees and, as exposed by Hallas and Monis [6], it can introduce an inter-operator variability and allow the data entry errors. Then, a particular interest is the automation of CFU counting using digital image techniques and computational vision systems.

Currently, commercial and automated counter systems of CFU can be found, for example, ProtoCOL 3 [7], BioSpot® Colony Counters [8], or PhenoBooth [9]. In general, these systems employ specialist image acquisition hardware, which is advantageous because they have total control of the position and illumination of the scene in the image acquisition processes, providing desirable features for the counting software. The expectancy is the systems working with excellent performance, but, in this case, the problem is they are very costly for small laboratories and institutions. Brugger *et al.* [10] and Chiang *et al.* [11] also presented prototypes with integration between specialist hardware and software, looking for the control of position and lighting of the sample. In the first case, a segmentation step is carried on for the whole image using a dynamic threshold based on Otsu´s method [12], and then an adaptive segmentation is going on to

find an appropriate threshold for each pixel, next, they use ratio aspects and Bayes classifier to separate concatenated colonies. In the second case, they use bottom-hat transform for correcting problems with the non-uniform illumination in the image and also Otsu´s method for segmentation, a watershed transform [13], [14] is employed to divide the overlapping colonies. Zhu *et al.* [15], for example, present a system based on images captured at near-infrared light and also use thresholding technique and watershed transformation to the segmentation process.

In general, arrangements between software and hardware looking for illumination control can be called active, by contrast, systems that do not use external lighting are called passive. These last systems are attractive because, without the hardware components, they are low-cost, even some may be free. In this way, different authors have indicated distinct methodologies to automate or semi-automate the process based on the passive system [16]–[21], and others, such as Bray *et al.* [22], adapt general software of biological image analysis to count CFU. In general, the first problem in the CFU automatic counter systems is to separate the bacterial colonies from the petri dish background. The second problem is to ensure the complete separation of colonies among them. However, the thresholding technique and Watershed transform are recurrent techniques in these systems, the first, generally, is to separate the petri dish from the image background and as an initial CFU segmentation, the second is used to divide the remnant clusters of colonies. Other techniques included Hough transform [23] to detect the center of the dish or circular form of colonies; color similarity metric in hue-saturation-value (HSV) color space [24] to segmentation using chromatic images; mathematical morphology, like dilation, erosion, and combination of they both [25] to segment binary images. Moreover, different development environments, and programming languages are used too, for example, MATLAB, MIT App Inventor 2, C++, or Java.

Then, there are different systems and alternatives to automate the counting of colony forming units, however, new contributions are significant because the problems of segmentation and identification in the programmed systems are more complex with a high number of colony forming units in the petri dishes. Even, there are more difficulties for correct counting when laboratories use different culture methods or kinds of agars, and when they use different types of lights in the image acquisition process, which can it mean variables are unattended in the programmed algorithms. Then, the main aim of this work is to present an alternative method to estimate the number of CFU from a digital image of a petri dish with the sample. The present method uses a predefined mathematical model and some color information about the CFU present in the image to count. This is an assisted method, and it can be seen as semi-automated, but it shows good behavior with few steps.

2. METHOD

Figure 1 presents an example of a typical petri dish with a sample of CFU; in this case, a blood agar plate and Escherichia coli ATCC 25922. In general, CFU from the same type have similar color characteristics in the image. Fact is important because it is information to be taken into account in the segment process. In general, this is similar to information used on other programs and methodologies in the segmentation processes. In the case of this work, the color information is used to adjust a quadratic function, as shown in Figure 2. It is understood that in this model, the principal components of the same object have the same color characteristics, and these change from the center to the ends in a smooth form. The model is carried out for conventional cameras in the three-color space red, green, and blue (RGB). As a reference, Figure 2 presents an example taking x1 as any space color, red, green, or blue; x2 as another one, and the remaining space color as a constant value.

Figure 1. Image of CFU sample on an agar plate. Escherichia coli ATCC 25922 in agar blood
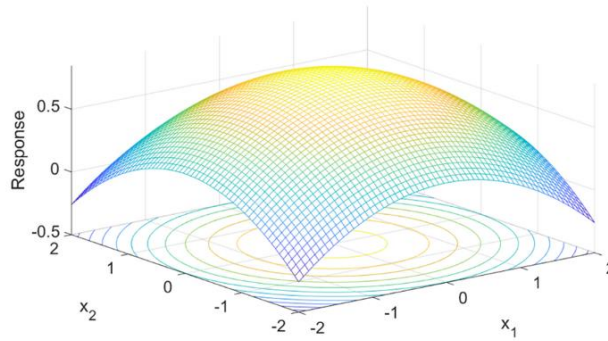
Figure 2. Quadratic model taking as a reference in the segmentation process

Figure 2 presents codified values between -2 and 2 for the input values. At the point (0,0), the function gives its maximum response, which has been encoded so that it does not take values greater than 1. This model has performed from a predefined model, which is shown in Figure 3 for one unique variable and, in that case, it is according to the relation:

$$y = 1 - 0.25x^2 \tag{1}$$

where *x* represents the input variable, *y* is the response, and produces the value zero in *y* for value 2 in *x*. It was decided to select this model to have not an abrupt change in the response and to use a coded value of 1 in the input variable as the standard deviation calculated for a sample of CFU taken from the image in the translation for real values. The codified value 0 for any input variable (red, green, blue) corresponds with the mean value calculated for the same taken sample.
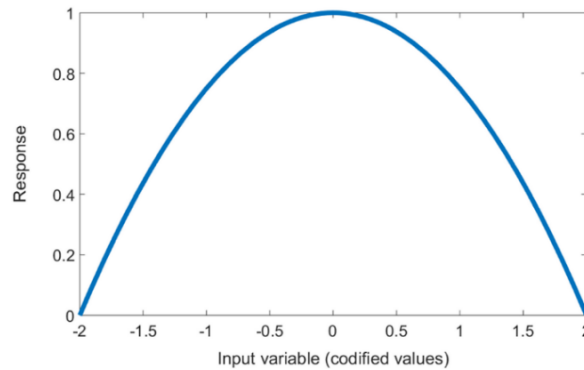

Figure 3. Quadratic model for one input variable

The model in (1) is translated for a central composite design for three input variables as shown in Figure 4 to get a second-order regression model as:

$$y_2 = b_0 + b_1 X + b_2 X + b_3 X + b_{11} X_{12} + b_{22} X_{22} + b_{33} X_{32} + b_{12} X_1 X_2 + b_{13} X_1 X_3 + b_{23} X_2 X_3 \tag{2}$$

and uses the following expression to find the regression coefficients:

$$b = (M^T M)^{-1} M^T y^* \tag{3}$$

where y^* is the vector response according to (1) for each input variable X1, X2, and X3 at the central compositive design; M is a matrix formed by the corresponding values of X1, X2, and X3 represented in Table 1, and according to with the regression coefficients of (2); M^T is the transpose matrix of M. Figure 2 is formed with the regression coefficients, (3), and the model, (2), keeping a constant input variable.
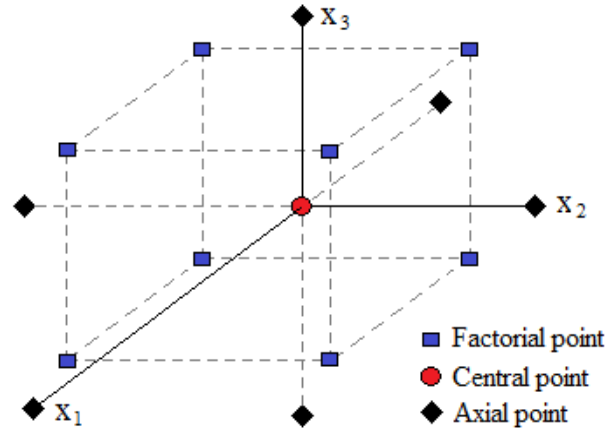
Figure 4. Central composite design for three input variables [26]

Table 1. Value coded and response according to (1)

X1	X2	X3	y*
-1	-1	1	0.75
1	-1	1	0.75
-1	1	1	0.75
1	1	1	0.75
-1	-1	-1	0.75
1	-1	-1	0.75
-1	1	-1	0.75
1	1	-1	0.75
-1.682	0	0	0.34
1.682	0	0	0.34
0	-1.682	0	0.34
0	1.682	0	0.34
0	0	-1.682	0.34
0	0	1.682	0.34
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1

As commented before, this method needs a sample of RGB values from the imagen. Then, in the methodology, one must choose manually six random points that correspond to CFU in the image, as shown in Figure 5. Six points are enough estimating ten digital levels of deviation, a confidence level of 90%, and an error of 7 digital levels. The number of digital levels in the image depends on its radiometric resolution; for a conventional 8 bits resolution, the number is 256.

Figure 5. Illustration of a chosen random point corresponding to a CFU and the point is marked with a blue cross

Values ±1 in Table 1 for X1, X2, and X3 correspond with factorial points in CCD as shown in Figure 4, value 0 to central point, and values ±1.682 to axial points. For any input red, green, and blue space color, the value is codified with the expression:

$$X_i^* = (Real\_X_i - Mean\_X_i)/Std\_X_i \tag{4}$$

where, $Real\_X_i$ is the real value for a color space input (red, green, or blue) of a pixel in the image; $Mean\_X_i$ is the mean value computed for the six choose points in this input; and Std_X_i is the standard deviation for these points in the input.

The segmentation process is carried out from (2) for all pixels in the image, values greater than zero in the response of the model are label as CFU, the others are discarded. At this point, a binarization process in the image is executed, and a morphological opening function is carried out to help separate connected CFUs. Here, the process can differentiate between a CFU pixel and a pixel of another object in the image. Then, a new labeling process is carried out to connect pixels of the same object (CFU). This process is made using connected component labeling with the eight closest pixels. Thereby, each object is identified and told apart from the others. Besides this, the method takes a mean size from all objects labeled as CFU, and all objects with a size less than 15% of the mean value size are not taken into account. Finally, counting is performed from the object labeling in the image. This method was carried out in C++ programing language, and used the OpenCV library. Different culture mediums and kinds of bacteria were performed for agar plate samples.

3.   RESULTS AND DISCUSSION

The model was achieved as:

$$y = 0.9854 - 0.1382X^2 - 0.1382X^2 - 0.1382X^2 \tag{5}$$

which is coherent with the predefined model because it is a second-order polynomial function, with the same behavior for the three independent input variables, and no linear component are found in it.

Mean values and standard deviation used to codify values change for each test because these depend of the chosen six-point, which for the same image can be different every time. However, it is expected that for the same type of colony in the same image, they will be similar. For example, Figures 6 indicate the mean value and standard deviation for different point chosen in the same image. In this case, Staphylococcus aureus in trypticase soy agar. The red cross in each image indicates the chosen point. The Figure 6(a) shows means for R, G, and B of 249, 238, and 222 respectively while in the Figure 6(b) the values are 247, 237 and 219 respectively which are very close values.
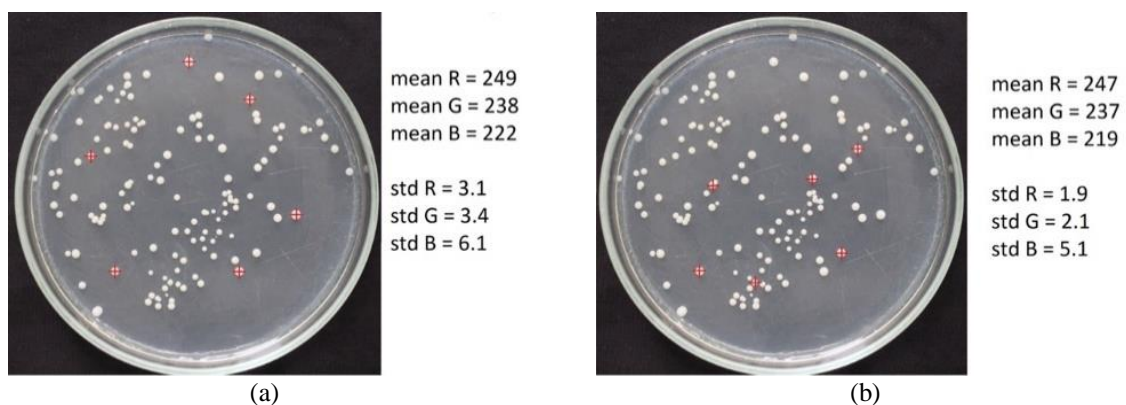

Figure 6. Comparison of mean value and corresponding standard deviation (std) for the three input color space variables (RGB) chosen for different six points in different tests for the same image: (a) first choice of random points for training and (b) second choice of random points for training on the same image

Figure 7 presents the result of segmentation for this image for the six points chosen in Figure 6(b), where, for illustration, the labeling pixels as CFU by process are highlighted in different colors; in this case the system computed 127 CFU, and the manual counting register 131, which is agree at a 96.9%. Figure 8

presents response for Escherichia coli ATCC 25922 in agar blood sample as shown in Figure 1, in this case 993 CFU was computed with 947 manually counting, agreement 95.36%

Figure 8 shows a good behavior of the systems, despite having a large number of colonies in the image. Tests with other organisms have shown high efficacies of the model, specifically for a samples of Staphylococcus aureus in blood agar, computing 41 CFU by 43 manually counting (95.34%), for Escherichia coli in MacConkey agar, computing 89 CFU with 96 manually counting (92.7%). Other example for MacConkey agar with 983 CFU computing by the system, 997 manually counted (98.59%) shows good behavior with a large number of CFU presents in the image.
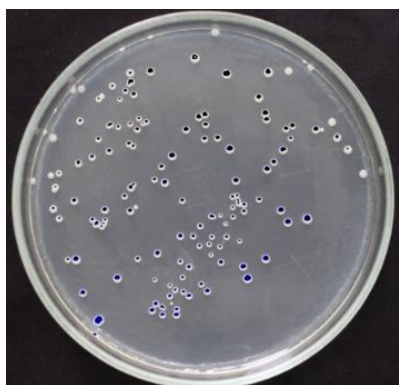
Figure 7. Sample with CFU labeling, coincident 96.9%
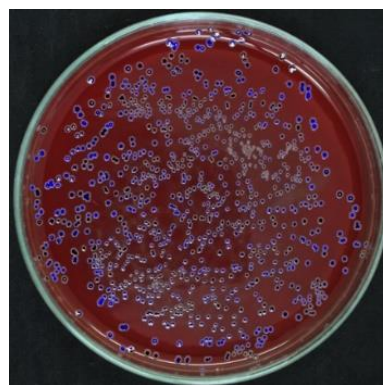
Figure 8. Sample of Figure 1 with 994 CFU labeling, coincident 95.36%

As seen from Figures 7 and 8, with an appropriate choice of six CFU points the method exhibits good behavior for images with many or few CFUs contained in the sample. However, it is clear that the segmentation process depends of the chosen points, and then, an adequate response is achieved when the chosen point corresponds to CFU in the sample. Thus, this method can be used to count CFU from the digital image and does not have problems in implementation, but it must be correctly assisted with the selection of the six points in practice, then it is acceptable to say that it can act like a semi-automatic system

4.   CONCLUSION

This paper presents a method for counting colony forming Units in agar plates from a digital image of petri dishes with the sample. The process uses a predefined second-order mathematical model, and it is based on the space color agreement between different pixels of the same CFU sample. The method is assisted and presents good behavior for images with many or few CFUs in the sample, with an agreement higher than 90% compared to manual counting. The segmentation process depends on the information of RGB color spaces of six points corresponding to CFU in the sample. The user chooses these points, and, for a correct function, the points must be chosen in an accurate CFU position in the image. In the future, the standard deviation of the chosen points must be used to segment UFC with different space colors in the same samples, and the sizes of labels in the image, to not eliminate small objects that correspond to CFU.

REFERENCES

[1]   M. Hindiyeh and A. Ali, "Investigating the efficiency of solar energy system for drinking water disinfection," *Desalination*, vol. 259, no. 1–3, pp. 208–215, 2010, doi: 10.1016/j.desal.2010.04.004.
[2]   M. Nemati, A. Hamidi, S. Maleki Dizaj, V. Javaherzadeh, and F. Lotfipour, "An overview on novel microbial determination methods in pharmaceutical and food quality control," *Advanced Pharmaceutical Bulletin*, vol. 6, no. 3, pp. 301–308, Sep. 2016, doi: 10.15171/apb.2016.042.
[3]   D. A. Burger and R. Schall, "Robust fit of Bayesian mixed effects regression models with application to colony forming unit count in tuberculosis research," *Statistics in Medicine*, vol. 37, no. 4, pp. 544–556, Feb. 2018, doi: 10.1002/sim.7529.
[4]   M. Umar *et al.*, "Bacteriological quality assessment and antibiogram profile of bacteria associated with sachet drinking water sold at Zaria, Northern Nigeria," *International Journal of Pathogen Research*, vol. 2, no. 2, pp. 1–13, 2019, doi: 10.9734/ijpr/2019/v2i230067.
[5]   P. R. Zhou, H. Hua, and X. S. Liu, "Quantity of candida colonies in saliva: a diagnostic evaluation for Oral Candidiasis," *The Chinese Journal of Dental Research : The Official Journal of The Scientific Section of The Chinese Stomatological Association (CSA)*, vol. 20, no. 1, pp. 27–32, 2017, doi: 10.3290/j.cjdr.a37739.

[6]   G. Hallas and P. Monis, "Evaluation of heterotrophic plate and chromogenic agar colony counting in water quality laboratories," *MethodsX*, vol. 2, pp. 415–422, 2015, doi: 10.1016/j.mex.2015.10.003.

[7]   "ProtoCOL 3 - automatic colony counting and zone measuring," *Synbiosis*. Accessed: Oct. 26, 2022. [Online]. Available: https://www.synbiosis.com/product/automated-colony-counting-zone-measurement-protocol-3/

[8]   CTL, "Immunospot BioSpot colony counting." *Immunospot*.com, Accessed: Oct. 26, 2022. [Online]. Available: https://immunospot.com/plaque-colony-counting

[9]   "Singer instruments phenobooth," *Singer Instrument*. https://www.singerinstruments.com/solution/phenobooth/ (accessed Oct. 26, 2022).

[10] S. D. Brugger, C. Baumberger, M. Jost, W. Jenni, U. Brugger, and K. Mühlemann, "Automated counting of bacterial colony forming units on agar plates," *PLoS ONE*, vol. 7, no. 3, Mar. 2012, doi: 10.1371/journal.pone.0033695.

[11] P.-J. Chiang, M.-J. Tseng, Z.-S. He, and C.-H. Li, "Automated counting of bacterial colonies by image analysis," *Journal of Microbiological Methods*, vol. 108, pp. 74–82, Jan. 2015, doi: 10.1016/j.mimet.2014.11.009.

[12] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[13] S. Beucher, "The watershed transformation applied to image segmentation," in *Proceedings of the 10th Pfefferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis*, 1992, vol. 1992, no. 6, pp. 299–314.

[14] J. B. T. M. Roerdink and A. Meijster, "The watershed transform: definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, pp. 187–228, 2000, doi: 10.3233/FI-2000-411207.

[15] G. Zhu, B. Yan, M. Xing, and C. Tian, "Automated counting of bacterial colonies on agar plates based on images captured at near-infrared light," *Journal of Microbiological Methods*, vol. 153, pp. 66–73, Oct. 2018, doi: 10.1016/j.mimet.2018.09.004.

[16] W.-B. Chen and C. Zhang, "An automated bacterial colony counting and classification system," *Information Systems Frontiers*, vol. 11, no. 4, pp. 349–368, Sep. 2009, doi: 10.1007/s10796-009-9149-0.

[17] S. Gupta, P. Kamboj, and S. Kaushik, "Methodology for automatic bacterial colony counter," in *Advances in Intelligent and Soft Computing*, vol. 166, no. 1, Springer, 2012, pp. 559–565.

[18] Q. Geissmann, "OpenCFU, a new free and open-source software to count cell colonies and other circular objects," *PLoS ONE*, vol. 8, no. 2, 2013, doi: 10.1371/journal.pone.0054072.

[19] J. Austerjost *et al.*, "A smart device application for the automated determination of E. coli colonies on agar plates," *Engineering in Life Sciences*, vol. 17, no. 8, pp. 959–966, Aug. 2017, doi: 10.1002/elsc.201700056.

[20] A. U. M. Khan, A. Torelli, I. Wolf, and N. Gretz, "AutoCellSeg: robust automatic colony forming unit (CFU)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques," *Scientific Reports*, vol. 8, no. 1, Dec. 2018, doi: 10.1038/s41598-018-24916-9.

[21] J. Shi *et al.*, "Noise-free microbial colony counting method based on hyperspectral features of agar plates," *Food Chemistry*, vol. 274, pp. 925–932, Feb. 2019, doi: 10.1016/j.foodchem.2018.09.058.

[22] M. Bray, M. S. Vokes, and A. E. Carpenter, "Using cellprofiler for automatic identification and measurement of biological objects in images," *Current Protocols in Molecular Biology*, vol. 109, no. 1, Jan. 2015, doi: 10.1002/0471142727.mb1417s109.

[23] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, Jan. 1981, doi: 10.1016/0031-3203(81)90009-1.

[24] G. H. Joblove and D. Greenberg, "Color spaces for computer graphics," in *Proceedings of the 5th Annual Conference on Computer Graphics and Interactive Techniques*, 1978, pp. 20–25, doi: 10.1145/800248.807362.

[25] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 532–550, Jul. 1987, doi: 10.1109/TPAMI.1987.4767941.

[26] E. Margarita *et al.*, "Parameter optimization in GMAW process with solid and metal-cored wires," in *International Congress Section I - Modelling, Control and Identification*, 2013, vol. 6, pp. 5256–5266.

BIOGRAPHIES OF AUTHORS

Jesus Emilio Pinto Lopera 🆔 📇 SC 🔗 is a Physical Engineer, Master and Ph.D. in Mechatronic Systems. Professor of the Faculty of Engineering of the Universidad de la Amazonia, Florencia-Caqueta, Colombia. He belongs to the Research Group on Informatics, Innovation and Technology of the Universidad de la Amazonia (GITUA) and to the Research Group on Technological and Environmental Developments (GIDTA) of the Universidad Católica de Manizales. His research interests are associated with computer vision and artificial intelligence. He can be contacted at email: jes.pinto@udla.edu.co.

Diana Carolina Meneses Cabezas 🆔 📇 SC 🔗 is a Biologist, Specialist in Industrial Microbiology, Magister in Agroindustrial Microbiology. Professor of the Academic Unit of Training in Natural Sciences and Mathematics of the Catholic University of Manizales, Manizales - Caldas, Colombia. Her research interests include industrial and agroindustrial microbiology, didactics of basic and applied sciences. She is a member of the Research Group on Infectious Diseases GINEI and Education and Educator Training EFE of the Catholic University of Manizales. She can be contacted at email: dmeneses@ucm.edu.co.

Yuliana Zapata Serna 🆔 📇 SC ◗ Bacteriologist. Master in Biomedical Sciences with emphasis in Microbiology. Research Group in Infectious Diseases (GINEI) of the Catholic University of Manizales. Her research interests include topics such as resistance in Gram-negative bacteria. She can be contacted at email: yzapata@ucm.edu.co.

Yeison Alberto Garcés-Gómez 🆔 📇 SC ◗ received bachelor's degree in Electronic Engineering, and master and Ph.D. degrees in Engineering from Electrical, Electronic and Computer Engineering Department, Universidad Nacional de Colombia, Manizales, Colombia, in 2009, 2011 and 2015, respectively. He is Full Professor at the Academic Unit for Training in Natural Sciences and Mathematics, Universidad Católica de Manizales, and teaches several courses such as experimental design, statistics, and physics. His main research focus is on applied technologies, embedded system, power electronics, power quality, but also many other areas of electronics, signal processing and didactics. He published more than 30 scientific and research publications, among them more than 10 journal papers. He worked as principal researcher on commercial projects and projects by the Ministry of Science, Tech and Innovation, Republic of Colombia. He can be contacted at email: ygarces@ucm.edu.co.