Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.7 n.32, p. 12-22. Nov/Dec, 2021

DOI: https://doi.org/10.5935/jetia.v7i32.781

RESEARCH ARTICLE OPEN ACCESS

COUNT OF BACTERIA AND YEAST IN MICROBIAL BIOPRODUCT USING DIGITAL IMAGE PROCESSING

Jorge Peña Martín*[1], Yelenys Alvarado-Capó[2], Rubén Orozco Morales[3], Tatiana Pichardo[4] and Ailet Abreu López[5]

[1, 3, 5] Automatic Control Department, Central University "Marta Abreu" of Las Villas - UCLV. Santa Clara, Cuba.
[2, 4] Laboratory of Applied Microbiology, Institute of Plant Biotechnology (IBP), Central University "Marta Abreu" of Las Villas - UCLV. Santa Clara, Cuba.

[1] http://orcid.org/0000-0002-4627-8775, [2] http://orcid.org/0000-0003-1721-717X, [3] http://orcid.org/0000-0002-6240-1569, [4] http://orcid.org/0000-0001-9416-2649, [5] http://orcid.org/0000-0002-9716-7696

Email: *jopena@uclv.cu, yelenys@ibp.co.cu, rorozco@uclv.edu.cu, tatiana@ibp.co.cu, aileta@uclv.cu

ARTICLE INFO

Article History
Received: October 28th, 2021
Accepted: December 06th, 2021
Published: December 15th, 2021

Keywords:
Microorganism counting,
Bacteria detection,
Yeast detection,
Digital microbiology imaging,
Morphological operations.

ABSTRACT

The count of microorganisms in substances from different industries, like the count of bacteria and yeasts, is a necessary and important process since long time ago. Traditionally, in the industries this process is performed by experts observing the samples in the microscopes, which is time-consuming and varies depending on the degree of expertise of the experts. Currently, the use of digital images of the samples to be analyzed is a variant widely used for such count task. In that sense, several methods have been created in recent years to make this process, but none of them covers the wide range of diversity that can be found in the real microbiological world. With these ideas as premises, a new method for count bacteria and yeasts in microbial bioproducts using digital images is presented in this paper, in order to provide to experts the approximate number of those microorganism. The method involves basic operations of digital image processing like contour detection, morphological operations and statistical analysis; and it was developed in Python language using the OpenCV library. The results obtained were evaluated by microbiological experts proved to have an acceptable performance for the context of use.

I. INTRODUCTION

Due to the importance of the microorganisms in society, the detection and count of those small organisms has long been a task of vital importance in the laboratories of different industries such as chemical, pharmaceutical, agricultural, food and environmental [1], [2], as it allows to know the status and properties of microorganisms themselves, the substances and environment where they live.

Depending on the application where the sample is analyzed, is the degree of accuracy required for detection and count the microorganisms. In some scenarios, such as medical sciences, it is necessary to know the classification of the microorganisms present in a sample [3], while in others it is necessary to know the quantity of microorganisms [4]. On certain occasions, as in the detection of diseases, it is required that such information is present in a relatively short time, and with the least margin of error [5], [6] because it can be the cause of the late application of treatment in sick patients; or even if it is possible that information can be used to prevent diseases in the persons [7].

For the count of microorganisms and their colony forming units (CFUs) [8], [9] there are several methods. One of the most widely used are the observation of samples of the substances under microscopes by trained persons, usually called experts.

This counting task requires a great effort and time for part of the experts, sometimes becoming inaccurate when the volume of samples analyzed is large and eye fatigue appears [10], or the counting performed among different specialists may also vary [11], especially when the number of microorganisms involved in the counting process is high [12].

The morphological change in the structure of the microorganisms and their characteristics along their growth and

development, constitutes other factors that can cause difficulties in the identification and count task. The process of sample preparation is also other factor that could be lead a number of impurities than can cause the apparition of many false positives during this process.

Trying to eradicate these adverse situations and looking for the humanization in the counting of microorganisms, in recent years the use of digital images of the samples has been increased to perform this counting process. The use of images of the samples also provides additional possibilities, since they can be stored for future investigations, training, verification, audits and legal processes, among others.

Computer vision and digital image processing techniques have been one of the fields that have given an impulse to this variant, trying to achieve a high, consistent and fast repeatability, free from the subjectivities that inevitably arise due to fatigue or the degree of experience of experts [2].

Different software's have emerged for the analysis of bacteriological and other cell culture images, counting of microorganisms and their colonies. Some of those are proprietary or sometimes require the purchase of specialized equipment to perform an accurate count, as reported by some researchers [9], [13], making them highly expensive solutions.

As alternative to these difficulties, open source software has emerged, such as OpenCFU [13], COVASIAM [14], NICE [15], ImageJ [16] or the new project surged from that software: Fiji [17]; but their results, such as accuracy, versatility and speed, varies depending on the context where they are applied, as demonstrated in the research presented in [18] for different software with this functionality.

That is why several studies have tried to improve or extend the use of those software's, such as the research presented in [11], [19]; but until now, due to the wide range of characteristics and circumstances that can be found on the different microorganisms to be counted, such as shape, size, texture and overlap between them [20], it has been difficult to have general methods to carry out this task.

There is also another wide spectrum of changing parameters that affect the detection and count such as the physical media to take the image and the way that are acquired it, in which environment the microorganisms are placed, the rotation of the microorganisms, the scale at which the images are taken, the illumination, the angle at which the image is taken, the quality of staining, among others.

With this complex scenario, the settings parameters of those algorithms are tuned for specific experiments and scenarios; leaving the usefulness of algorithms with a restricted field of action [5], because in this count variant the steps to do are simple but the way to perform it could be diverse. Ideally, the microorganisms or their colonies should be isolated from the background [21], and then, if they are clustered, they should be separated from each other [13]. This makes the techniques employed in these processes diverse, especially when try to separate the clusters.

Generally, the techniques used to detect and count microorganisms in images can be grouped in two main groups: "classical techniques" of computer vision (CV) and digital image processing (DIP), and machine learning (ML) techniques, which differ in the way the counting is performed. While conventional techniques can achieve a detection and count that depends only on the characteristics of the image analyzed, ML techniques can drastically increase the efficiency compared to classical techniques but require large amounts of data from experiments, especially in the most advanced trend of ML, called deep learning (DL).

CV and DIP techniques can include the use of methods such as: multilevel threshold [14], morphological operations [7], edge detection [20], watershed algorithm [8], [12], [22], distance transform [8], and active contours [23] among others. Usually these algorithms are preceded by a series of steps in the preprocessing stage, which may include conversion between color spaces, image filtering and contrast adjustment.

The watershed algorithm with applied threshold, usually in conjunction with other techniques, has been one of the most used variants due to the possibility of divide and classify in different regions objects that are grouped in an image, in this case different microorganism. Examples of research using this algorithm can be seen in [24] where the H-Dome algorithm is also used whit images that are captured under different lighting conditions.

In [8] they also use the watershed algorithm to separate the colonies that form clusters. In this case, the images were acquired from samples placed in Petri dishes. For their method, the authors propose to work in two regions: one in the center of the disk and another in the periphery where they report that counting is much more difficult.

Those methods of segmentation using watershed could cause an over segmentation of the components in the image, that is the reason that in some cases that technic is combined whit technic to mark the regions to detect, as the work reported in [25].

In [20] propose a method composed of several processing techniques such as morphological operations, edge detection (Canny and Laplacian of Gaussian) and threshold, for counting microorganisms in a general way in digital images. Although the method returns 11 possible variants of microorganism detection, it requires user intervention to obtain good results. The author himself recognizes that the method may fail when the microorganisms are tightly clustered or the area of contact between them is large, as well as when the microorganisms have a color similar to that of the substance in which they are placed.

Other less commonly techniques reported in this context that have demonstrated acceptable accuracy are the circular Hough transform [10], [26], or the use of granulometry based on the morphological characteristics of the microorganisms [9], [27]. There are also reports of less general techniques such as the Chan-Vese algorithm [28], which is a variant of active model contours.

Inside the ML techniques used for this task, the convolutional neural networks (CNN), support vector machines (SVM), k-nearest neighbors (K-NN), decision trees and random forest (RF) are the most widely used techniques, demonstrating a superiority in obtaining results compared to classical techniques.

With the use of such methods, the detection, classification, count or analysis become very effective because they can learn patterns from the images. In many cases, as in [4], these more advanced techniques are combined with classical techniques to achieve greater accuracy or to save time in training the models.

Examples of the use of CNN can be seen in [3] where they use a deep convolutional neural network (DCNN) applying the transfer learning paradigm when employ a pre-trained network for the classification of bacteria, as in [29] for environmental microorganisms. CNN are also used in [4], [5].

The same technique, CNN, but for yeast segmentation is employed in the research presented in [30] where they use the watershed algorithm and the distance transform algorithm along with known CNN architectures.

The same purpose of bacterial classification is reached in [31] where three ML techniques are used: SVM, K-NN and Random Forest.

In [32] employ the RF algorithm and two variants of SVM: linear SVM (LinSVM), and Cross-Validation SVM (CVSVM) for the classification of bacteria causing tuberculosis disease.

With a need to automate the counting of microorganisms in microbial bioproduct obtained by fermentation, specialists from the Instituto de Biotecnología de las Plantas (IBP) [33] of the UCLV joined efforts with researchers from the Departmento de Automatica of the Universidad Central "Marta Abreu" de Las Villas (UCLV), to develop a method using digital images for bacteria and yeasts counting in the microbial bioproduct where microorganisms of both groups form a consortia. Until now, the detection and count of microorganism of different characteristic in the same image, as in [34], has been poorly covered, and there is not evidence that yeast and bacteria was detected and counted in the same image.

The main objective of this research was to develop a method to determine the number of bacteria and yeast in these microbial suspension, using DIP technics and digital images taken under different magnifications and conditions. This is the first step to establish, in later stages of the research and which are out of the scope of this article, relationships between the amount of that microorganisms and other properties measured by specialists such as pH and conductivity in these microbial bioproduct.

The structure of this paper after the Introduction follow as: Section II describes the proposed method for bacteria and yeast counting, Section III presents an analysis of the results obtained using the proposed method in images of different experiments and discusses circumstances that may cause the method fail, while Section IV presents the conclusions reached during the research.

II. MATERIALS AND METHODS

For the detection and count of bacteria and yeasts in this research, it were used samples of microbial bioproduct, which are under investigation in the IBP for agricultural use. These bioproduct present a composition where coexist bacteria of the coccus type and yeasts, in which bacteria are the great majority.

The process of sample preparation is out of the scope of this work, starting from the point of image acquisition. This image acquisition process was made changing the color of the illumination source, the suspension and the magnification of the microscope lens-ocular assembly, so that images were obtained with different sizes of microorganisms and under different conditions.

The images were taken with an Olympus BH series (Japan) microscope and a digital camera coupled. Under the objective for 40x magnification (PL40, 0.65 numerical aperture and 0.17 working distance) the scale factor was 2.3 px/ µm. In the case of 100x (HI plan 100/1.25 with immersion oil) it was 6.0 px/ µm.

The camera used to obtain the images was a HDCE-X camera [35] with a ½" CMOS sensor, with a resolution of 2592 pixel by 1944 pixels (5 Mp), where the pixel size is 3.2µm X 3.2µm and adjustable exposure time. This camera is coupled via a USB cable to a desktop computer with Ubuntu 18.04 LTS operating system, producing an output RGB image of 640 X 480 pixel with 24-bit depth (8 bit per channel).

The method code was written in Python 3 language, using the digital image processing library OpenCV in its version 4.0.0.

II.1 PROPOSED METHOD

The method proposed in this research to count bacteria and yeasts in digital images is composed of three fundamental stages, which were denominated: Primary Operations or Stage I, Statistic Operations or Stage II, and Classification or Stage III, as its show in Figure 1.

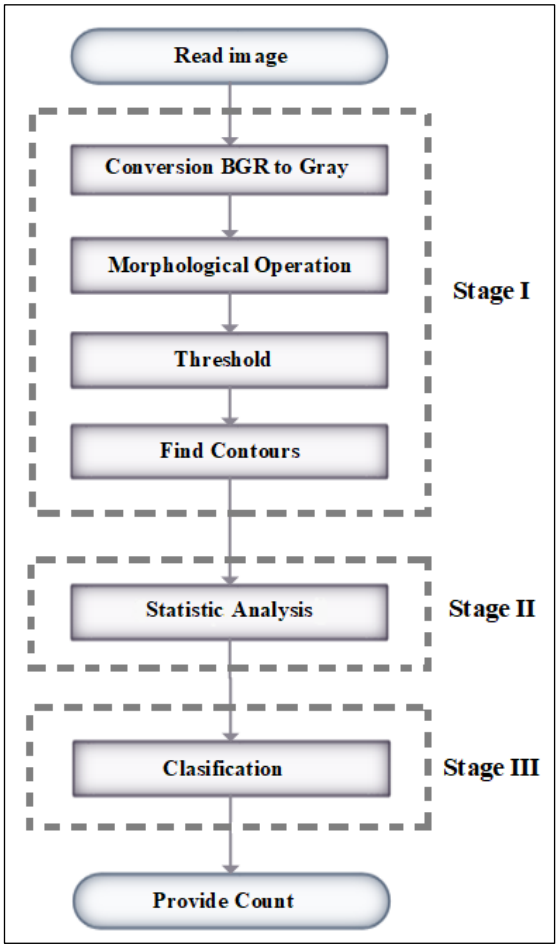

Figure 1: Flow of the proposed method.
Source: Authors, (2021).

The first step is read the image as a 3-channel image (BGR format, by default in OpenCV). Once the image is read, "Stage I" begins, which aims to delimit the possible microorganisms that appears in the image. In this stage the image is converted to grayscale, and to this new image is applied a process of morphological operations as filter, consisting of the "Black Top-Hat" method [36] whit the intention to highlight the visible contours of the microorganism.

This morphological operation consists on the difference between the process of making a morphological closing on the grayscale image and the grayscale image itself [37] as defined in Equation (1).

Black Top-hat morphological operation

$$\text{blackhat (img)} = close\ (img) - img \qquad (1)$$

Due the first part of the Black Top-Hat method is make an image "closing", the closing operation in gray scale images remove objects, relative to the size of the structuring element, that are more dark than their neighbors [38]; the effect of the Black Top-Hat method reveals areas more darker in the surrounding area of the objects of the image [36].

For this process, a square structuring element was used composed of a flat top with value of 1 in all its positions and size of nine pixel. The selection of the structuring element size was adjusted empirically, by applying the same operation in different

experiments, observing which was the most suitable size to cover the range of the experiments. The result of those steps can be seen in Figure 2.

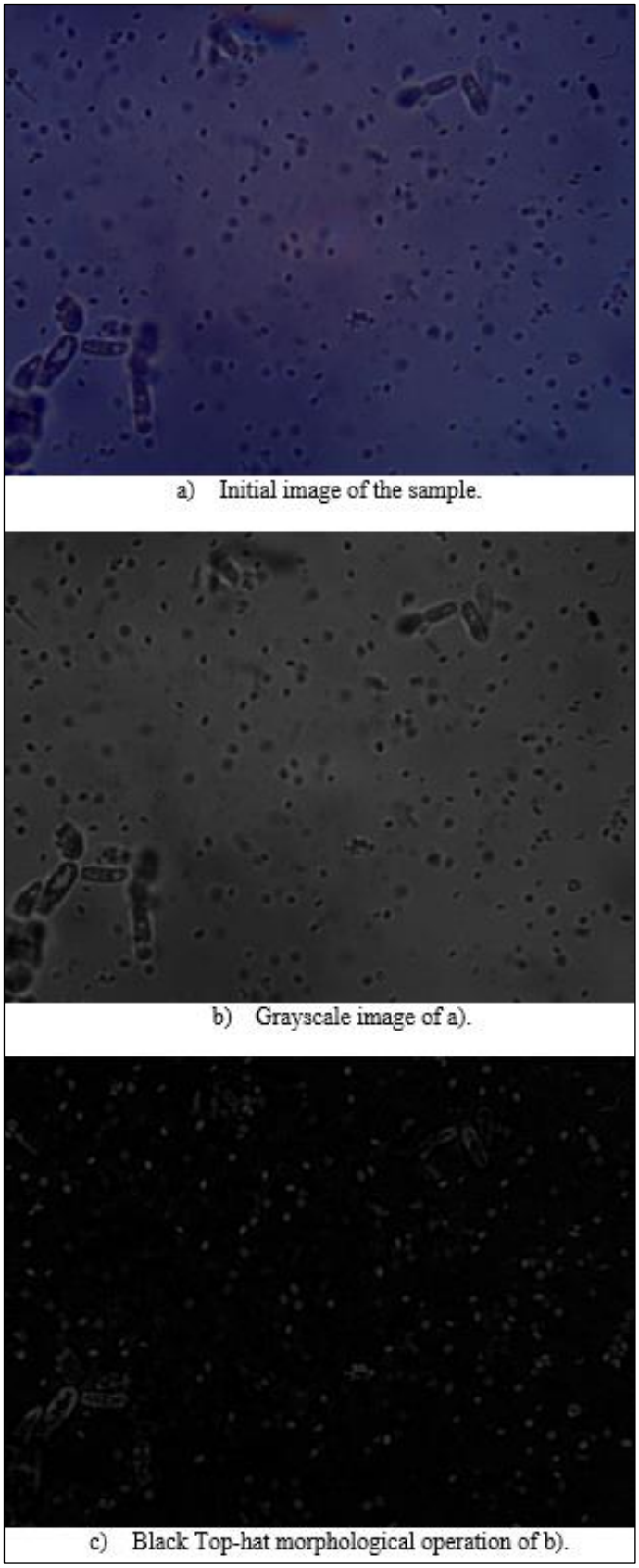

Figure 2: Example of Stage I in a microscopic image of microbial bioproduct (40X objective): a) Original image, b) Grayscale image, c) Black Top-hat morphological operation.
Source: Authors, (2021).

Once the image with the morphological operation is obtained a threshold is applied to it, selecting the Otsu method to apply a global binarization process, resulting in a binary image as shown in Figure 3.

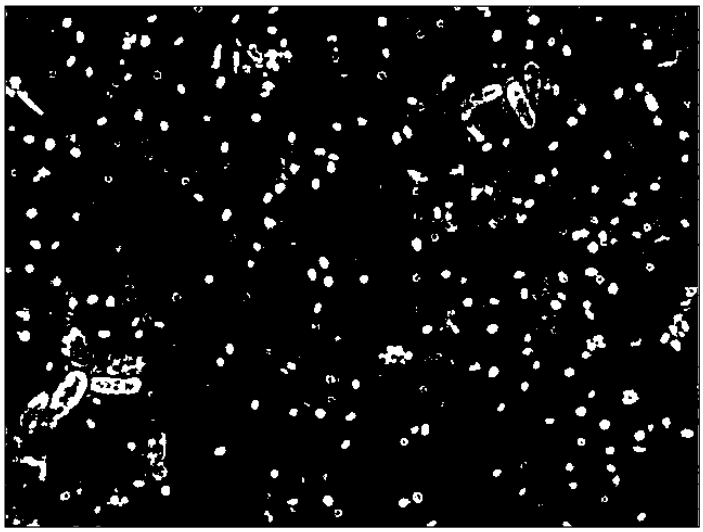

Figure 3: Image obtained after a threshold process on the image resulting of the Black Top-hat morphological operation.
Source: Authors, (2021).

After obtaining the binary image, as part of Stage I, is applied a process to obtain the contours present in the image. In this case, only the external contours that appear in the image were considered valid for the analysis, because sometimes microorganisms, especially yeasts, could project certain contour in the image that could be classified as internal, as illustrated in Figure 4. In this figure the external contours are drawn with a white line and the internal contours with black lines.

In order to avoid incorrect classifications, only the external contours are selected.

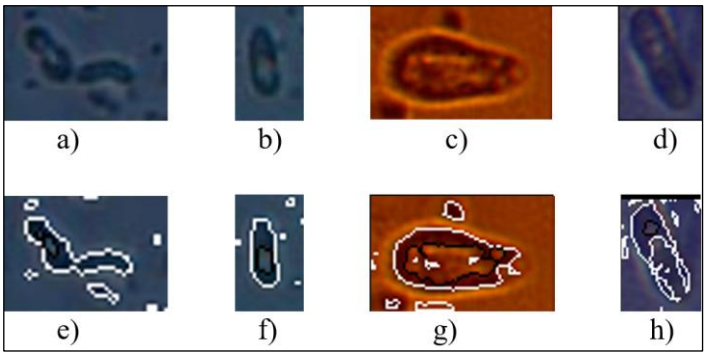

Figure 4: Example of classifications of internal and external contours in different microorganisms: a), b), c), d) image segments with different microorganisms, e), f), g), h) Internal and external contours detected in those image segments.
Source: Authors, (2021).

Once all the external contours of the microorganisms present in the image have been determined, “Stage II” begins. The main objective of this phase is determinate the mode of the size of the contours of the bacteria present in the image, as shown in Algorithm 1.

For achieve that objective, for each external contour found, the convex hull shape (cHS) of that perimeter is determined: that is

the convex polygon with the smallest area containing all the points of the contour [36].

The idea of the cHS of a contour is a figure to contain the entire area of that contour, all their points, within a polygon that is fully convex at any point on it.

Algorithm 1 Estatistic Analysis

```
for all external contour in image do
    find convex hull shape of contour
    find area of convex hull shape
    if area of hull contour > 0.01 then
        find area of contour
        calculate solidity factor
        if solidity factor is > 0.90 then
            find radius of mEC of contour
            take radius in consideration for determine mode
        end if
    end if
end for
extract mode of radius of minimal enclosing circle
```

As the scenario of this type of samples is high diverse, sometimes small contours appear after all preceded steps that can be ignored because they are too small. That is the reason to employ an adjustable parameter, using the area of the convex polygon, to discriminate or select the contours for subsequent analysis. In this case, if the area is less than 0.01 pixel it is not taken into account in the Stage II because such small areas may belong to noise in the images, or impurities in the environment where the microorganisms are placed [20]. This parameter was adjusted based on observations made during different iterations, and is applied in the Stage III too.

If the cHS has an area that must be taken into consideration, then the area of the set of internal binary pixel encapsulated by the analyzed contour is determined in order to calculate the contour solidity factor using Equation 2.

Solidity factor:

$$solidity = \frac{encapsulate\ area\ of\ contour}{area\ of\ convex\ hull\ shape} \quad (2)$$

The goal of the solidity factor in this research is to determine how regular are the pattern analyzed, which should correspond to well defined, non-clustered, bacteria or yeast, considering when more regular is the analyzed contour it solidity factor will be higher due the concordance between the areas.

Similar ideas to the solidity factor has been handled in other investigations like [12], but with a different formula.

At this point of the method, the comparison factor for the solidity of a contour was set to 0.90, using expert knowledge, thus considering for the statistic analysis to obtain contours as regular as possible.

If the analyzed contour satisfies the solidity criteria then the minimum enclosing circle (mEC) of that contour [39] is calculated; in other words, the circle of minimum area containing all the points of contour, as can be seen in Figure 5.

The integer part of the radius of the mEC are stored to take in account to determine, once all contours were analyzed, the mode of the radius size of the regular contours, otherwise it is not taken in consideration.

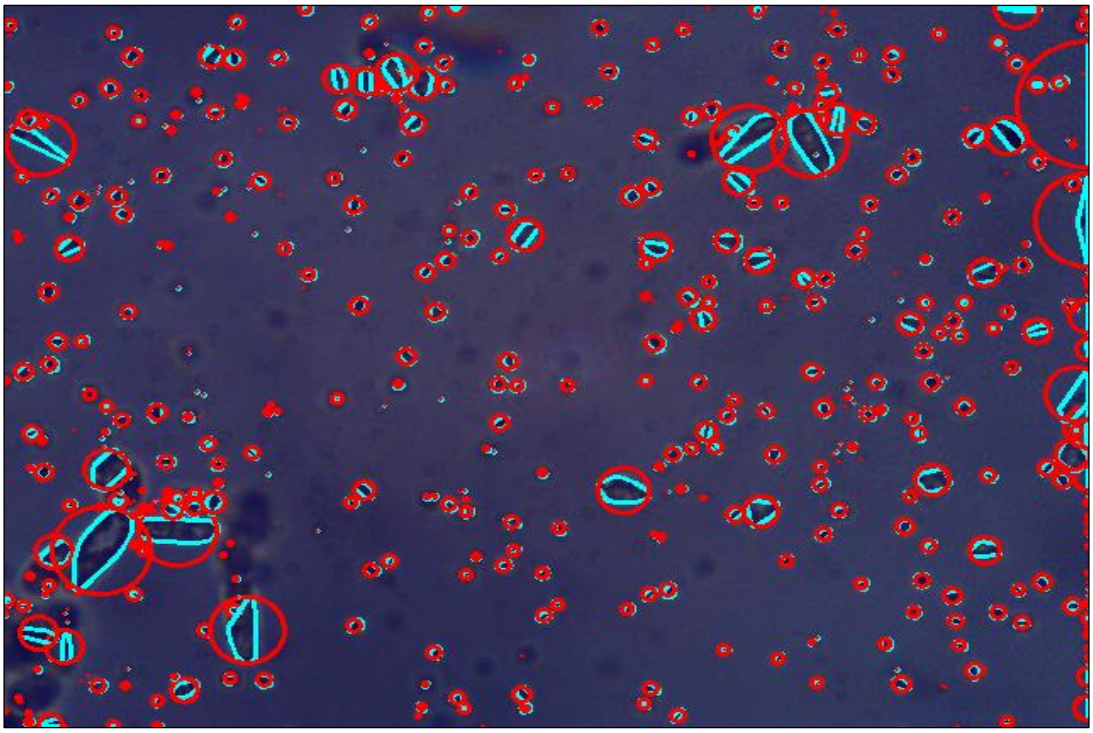

Figure 5: Determination of convex hull shape (cyan) and minimal enclosing circle (red).
Source: Authors, (2021).

After stored the integer part of the radius of mEC of contours, the first and second modes of these radius measurements are determined. One characteristic to employ these method whit this kind of samples is that always the number of bacteria have to be greater than the number of yeast. In case that property don't can established the method could achieve bad results, or even could fail. In the case that the number of bacteria is greater than yeast, was confirmed during the experiments realized, that can be

assumed that the first mode calculated, and in many times the second mode too, always going to belong to bacteria, due to the high number of them.

Once the second phase of statistic analysis is completed, the "Stage III" is carried out with the objective of classifying the microorganisms, bacteria or yeasts, that produce the contour detected and finally provide a count of them to the users.

To perform this task, all previously external contours determined are reanalyzed, as shown in Algorithm 2.

Algorithm 2 Clasification

```
for all external contour in image do
    find convex hull shape of contour
    find area of convex hull shape
    if area of hull contour > 0.01 then
        find radius of mEC of contour
        if radius of mEC> 0.5 then
            if radius is < 3 * maximun mode of radios then
                calculate solidity factor
                if solidity > 0.7 then
                    classify contour as bacteria
                else
                    classify contour as unknown
                end if
            else
                obtain coordinates of contour
                call function find_yeast for that contour
            end if
        else
            discard
        end if
    else
        discard
    end if
end for
```

For each of those external contours, the encapsulating cHS and its area are determined again, and the same criteria used in phase two are used to discriminate between the contours, being the very small contours excluded from the classification process. If the contour encapsulates a minimum established area, then the radius of the encapsulating mEC is calculated.

At this point in the method is placed a condition that allows ignore small contours for classification, those contours whit the radius of mEC is less than 0.5 pixel; otherwise the analysis of contour continue. That criteria were adjusted based on the experiments carried out using expert criteria.

When the contour satisfies that selection criteria, is used another condition that allows distinguish if the analyzed contour has a higher possibility of being a bacteria or not. This condition is based on compare the radius of the mEC of the contour with the mode of the radius. If the mEC radius is less than 3 times the mode, then it is more probably that the contour could be a bacteria; otherwise it is more probably to be a yeast or a union, cluster, of several microorganisms.

If the contour classifies to be analyzed as bacteria, its solidity factor is calculated again and this time if it is greater than 0.7 it is classified as bacteria; otherwise it is classified as unknown due to the irregularity of the shape.

Larger contours, those where the radius of the mEC is greater than three times the maximun of two modes of the radius, require another process for better classification as described in Algorithm 3. These contours can be composed of yeasts, or clusters of bacteria and yeasts, which is why another analysis is required for their classification. In that last case the contour is passed to a new function called "Find yeast", which is shown in Algorithm 3.

In this process, an empty image is created and the contour under study is drawn and filled. Two morphological operation processes are applied to this new image whit the intention to discover and separate, if it is necessary, the microorganism that are clustered.

Algorithm 3 Find yeast function

```
for specific contour in image do
    create a black image to draw contour
    draw and fill contour in black image
    apply erosion process to new image
    apply dilation process to new image
    find externals contours in new image
    for all new external contours do
        find radius of mEC of contour
        if radius of mEC > 0.5 then
            if radius of mEC < maximun mode of radios
then
                classify contour as bacteria
            else
                classify contour as yeast
            end if
        end if
    end for
end for
```

The first morphological operation consists of an erosion process who have the intention to erode the new figure seeking if this new figure is compose by different microorganism. To make that operation the size of the structuring element to be used is selected according to the mode of the radius of the mECs: if this mode is less than two, the size of the structuring element is set in 2, and if it is greater than 2 then the size of them is equal to the integer part of the mode of the radius.

The second morphological operation applied is a dilation process. That operation is applied whit the intention of increase the area of the shapes remaining from the erosion process, but preventing that the segments that were separated from each other coming back to be in touch again. In this operation the size of the structuring element to be used decreases by one unit with respect to the size of the structuring element used in the erosion process.

Once these two processes have been carried out, the new external contours that appear in this new image are determined. For each of these contours the radius of the mEC is calculated.

In this point of the method, the condition used previously is used again: if the radius of the mEC is greater than 0.5 pixel, it is classified, otherwise the contour is discarded for being too small.

If the radius of the new contour meets the requirement for classification, then it is compared to the higher value of the two most representative modes of the radius of the mECs determined in the statistic phase. If this new radius is smaller than the greater value of the two modes of the radius, then it is classified as bacteria, otherwise it is classified as yeast.

Finally, the count is provided to the specialists, highlighting bacteria in green circles, unknown forms in red circles and yeasts in magenta rectangles.

III. RESULTS AND DISCUSSIONS

For the development and test of the proposed method, several experiments were carried out in which the microbial suspension, the magnification of the camera-lens set, the focus, among other characteristics, were different; trying to capture a wide range of real characteristics and operations in this process.

As can be seen in Figure 6, the method is able to detect bacteria and yeasts in microbial bioproduct. In that image, bacteria were highlighted with a green circle and yeasts in purple rectangle. Microorganisms or microbial detritus that were classified as unknown are marked in red.

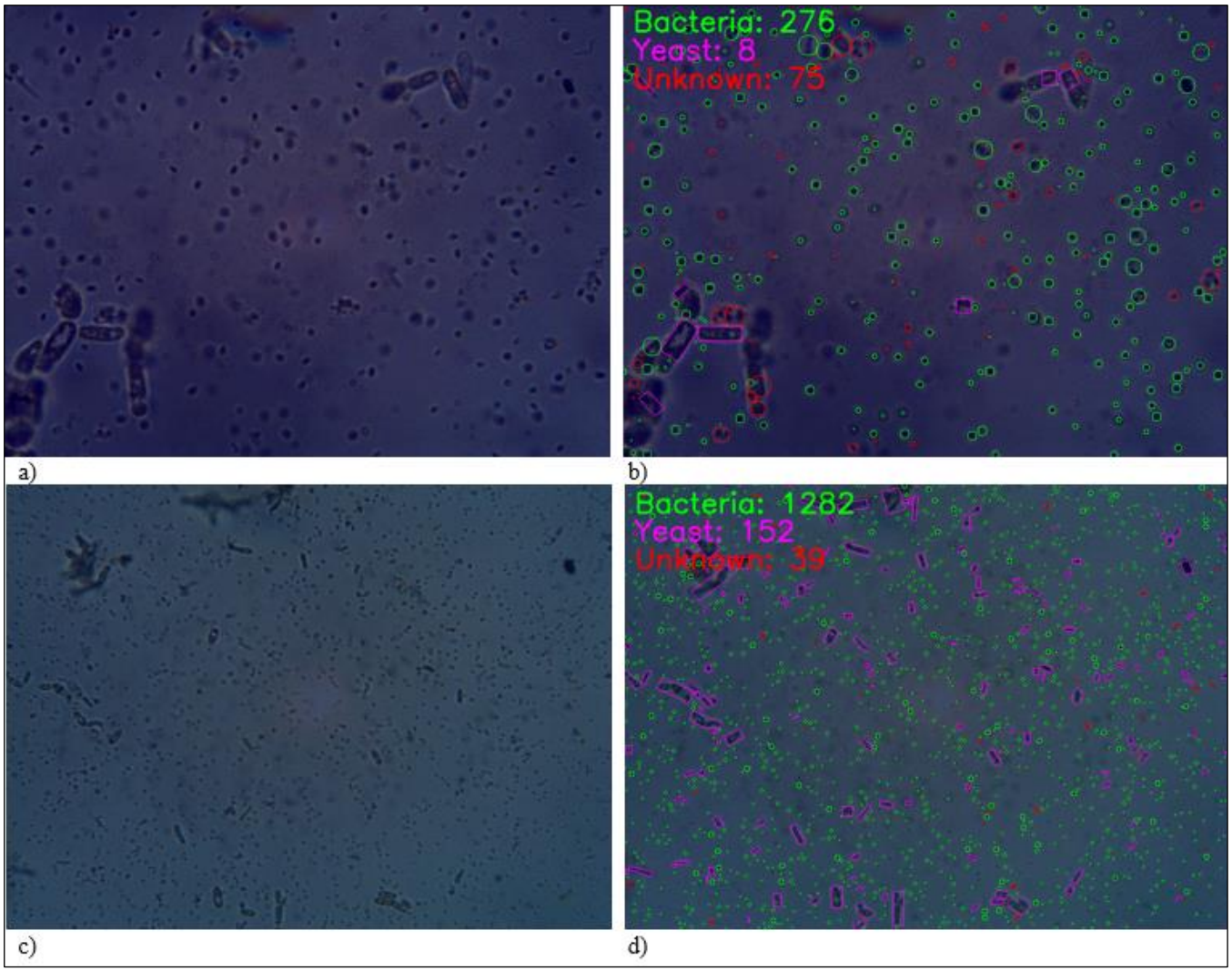

Figure 6: Count of bacteria and yeast in microbial suspension by the proposed method in experiments with different conditions: a) Experiment 1, b) count on the a) image, c) Experiment 2, d) count on the c) image.
Source: Authors, (2021).

Analyzing the results obtained through the observation of the images, it can be seen that the detection and counting of bacteria was achieved with acceptable accuracy, but not in the case of yeasts, whose precision is considerably lower. In this case the user should discern in those contours marked as unknown which is its possible classification, or if it should be discard.

The method was able to separate cluster of bacterias or yeast with a few microorganism touching between them, in large concentrations of those microorganism the accuracy is lower.

This issue is closely related to the small regular shape characteristic of bacteria, which don't present great variation in their color or in their shape, whereas the larger size of yeasts and their changes in color can result in a detection not precisely after image processing.

Other aspect that can also observed in the images are shadows that belong to microorganisms that are below the analysis surface or that are poorly focused, which can cause many false positives to be classified as bacteria because at least a part of the microorganism contour meets all the selection criteria, and in most cases this contour does not appear visibly clear.

There are also contours and shapes that do not meet the established matching criteria and were not classified.

III.1 FAIL CASES

As result of the wide range of experimental conditions, it was observed different scenarios where the method can make an incorrect classification.

One of common problems than can cause poor results have already been reported in similar investigations, such as the low contrast between the microorganisms and the background.

Other condition that can cause classification errors as it was mentioned before, is the shape itself of the figures of microorganisms in the image, especially the larger ones: yeast.

The non-homogeneous shape characteristics, especially in the color, can result in the fact that during the image processing, part of their shape is removed from the images by the different operations, resulting in a partition of the same microorganism into several sorting regions.

This two conditions make very difficult the process of detection and count as can be observed in Figure 7.

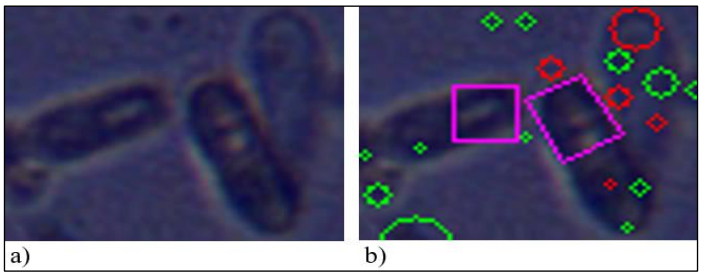

Figure 7: Example of misclassification of a yeast due to the irregularity in its shape structure in the image and poor contrast between the microorganism and the bacground: a) segment of an image of the sample and b) results obtained after processed the image in a).
Source: Authors, (2021).

In addition, sometimes the silhouette of the microorganism in the image cannot be observed completely, or there are clusters of bacteria and yeasts, in which it is very difficult to separate them, as can be seen in Figure 8 where is visible misclassification or bad count of microorganism, more accentuated in the classification of yeasts.

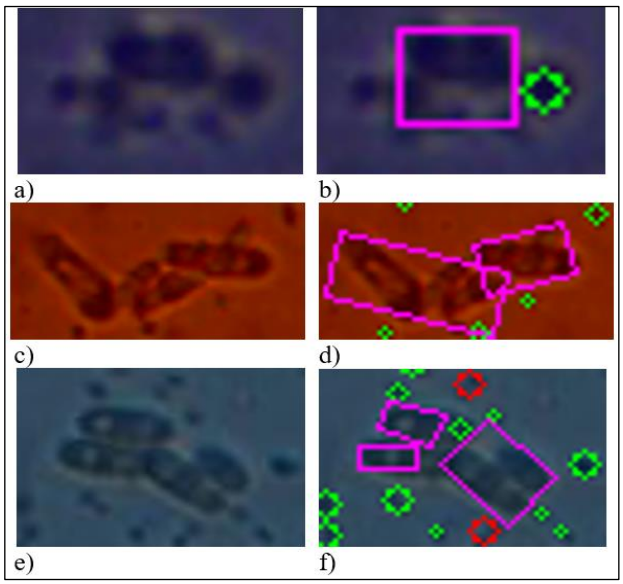

Figure 8: Clustering of several microorganisms causing misclassification: a) segment of an image of the sample and b) results obtained after processed the image in a); c) segment of an image of the sample and d) results obtained after processed the image in c); e) segment of an image of the sample and f) results obtained after processed the image in e).
Source: Authors, (2021).

Other great difficulty in obtaining good results with the method is the degree of focus with which the image is acquired, a condition that can be challenging for any computer vision algorithm. In this case the focus over the sample of the microscope-camera has a great influence on the results that can be obtained. Figure 9 shows this problem in different frames of a video taken on the same sample where the degree of focus varies.

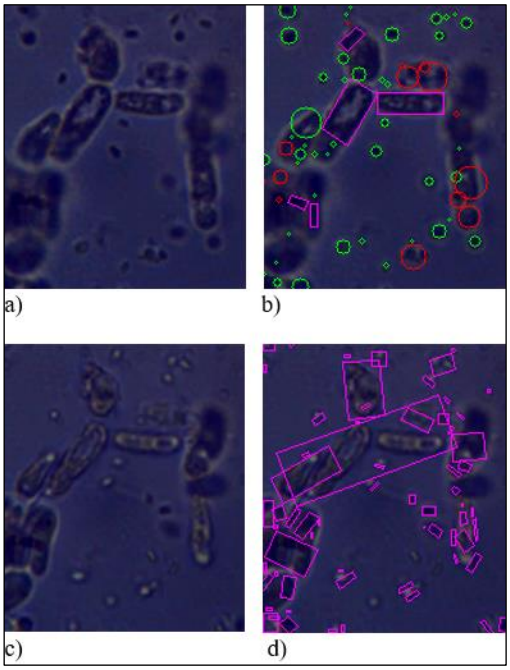

Figure 9: Example of variation of the degree of focus on the same sample: a) segment off an image of the sample and b) results obtained after processed the image in a), c) new segment image of the same microorganism by varying the focus of the camera-microscope d) results obtained after processed the image in c).
Source: Authors, (2021).

There are also problems associated with dust, fermentation residues or unusual objects that can be observed in the images and can belong to the imperfections in the camera, microscope lens or where the dissolution is prepared and are not the microorganism analyzed in the research.

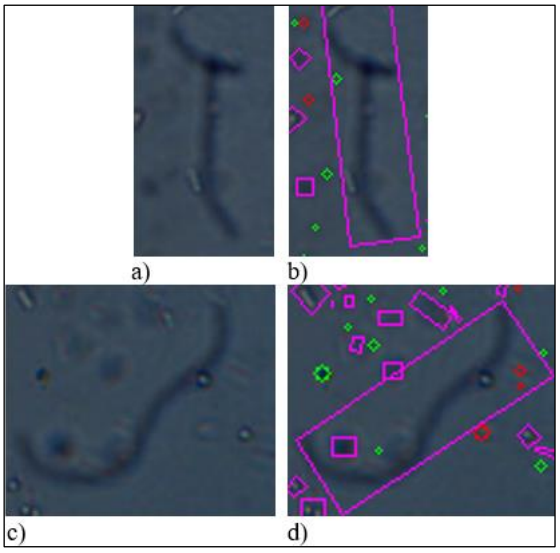

Figure 10: Example of count bias due to the appearance in the image of large objects that are not parts of the biological specimens: a) segment of an image of the sample and b) results obtained after processed the image in a); c) segment of an image of the sample and d) results obtained after processed the image in c).
Source: Authors, (2021).

Figure 10 shows how the visible silhouette of objects that are not part of the experiments, named artifacts, making the classification and counting process incorrect.

III.2 METHOD EVALUATION BY EXPERTS

After all parameters were adjusted for cover all experiments, the evaluation of the performance of the method was made by experts from the microbiologic laboratory of IBP.

The proceeding was made making a subjective evaluation [40], where the experts evaluate a set of images of microbial suspension with the count provided after processed by the method, and the original image whit out count.

Microorganisms detected as unknown were not taken in consideration because is decision of the user, when employ the method, make a visual inspection of the unknown shapes.

The experts analyzed the results taking the following assumptions:

- True Positive (TP): bacteria or yeasts that were correctly classified and counted.
- False Positive (FP): bacteria or yeasts that were counted but they are not.
- False Negative (FN): those microorganisms that were not classified or counted by the method (excluding unknowns) and which in the opinion of the experts should be counted because they were visible and well defined in the image.
- True Negatives (TN): were not counted because the intention of the method is not detect those elements that are not considered microorganisms, such as noise in the image, dirt on the microscope lens, etc.

If the research were approached from a point of view of image segmentation, where generally there is for each segmented image a ground truth label image, then it could be possible obtain values of TN; but in this investigation there is not intention follow that line.

The metrics used for the evaluation of the method were accuracy (Acc), precision (Pr), and sensitivity (Se) or recall [41], [42], as expressed in Equation of Accuracy, Equation of Precision and Equation of Recall, respectively.

Equation of Accuracy

$$Acc = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \quad (3)$$

Equation of Precision

$$Pr = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (4)$$

Equation of Sensitivity or Recall

$$Se = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (5)$$

For the validation process, nine images randomly chos n was given to the experts. This images were representative of the widely range of experiments, whit images taken under different magnifications. The annotations obtained by experts can be seen in Table 1 for bacteria, and in Table 2 for yeast.

In that tables can be appreciate the metrics results obtained after process the validation data.

Table 1: Experts evaluation and metrics in bacteria count.

Image	Method	TP	FP	FN	Acc	Pr	Se
A	280	265	15	2	0.93	0.94	0.99
B	728	726	2	32	0.95	0.99	0.95
C	1199	1190	9	54	0.94	0.99	0.95
D	1173	1153	20	46	0.94	0.98	0.96
E	77	74	3	0	0.96	0.96	1.00
F	1282	1243	39	5	0.96	0.96	0.99
G	1039	1018	21	10	0.97	0.97	0.99
H	877	866	11	7	0.97	0.98	0.99
I	269	263	6	2	0.97	0.97	0.99
Average	-	-	-	-	0.95	0.97	0.98

Source: Authors, (2021).

Table 2: Experts evaluation and metrics in yeast count.

Image	Method	TP	FP	FN	Acc	Pr	Se
A	9	6	3	1	0.6	0.66	0.85
B	166	152	14	34	0.76	0.91	0.81
C	101	85	16	18	0.71	0.84	0.82
D	99	83	16	20	0.69	0.83	0.80
E	0	0	0	0	0	0	0
F	152	149	3	4	0.95	0.98	0.97
G	123	117	6	0	0.95	0.95	1
H	124	123	1	4	0.96	0.99	0.96
I	11	8	3	2	0.61	0.72	0.8
Average	-	-	-	-	0.78	0.86	0.88

Source: Authors, (2021).

Analyzing the results obtained after processing the set of images, concluded that the method is able to detect with accepted accuracy the presence of bacteria whit a mean of 0.96, and the lowest value a 0.93. The precision and recall for those microorganism was high to, whit a 0.97 of precision and a recall of 0.98.

Lower level of count was achieve for yeasts. In this case, the accuracy obtained was around 0.69, whit the precision mean equal to 0.77 and a recall of 0.78.

Those results is mainly related to the situations explained above, which are accentuated in the case of yeasts.

In this case, the user can more easily corroborate and correct the yeast count since the number of yeasts in these solutions is much lower than the number of bacteria, enough to obtain a ratio that serves as an indicator for future research to establish the relationship between the number of microorganisms and other measurable parameters of the substances.

IV. CONCLUSIONS

General methods for the classification of microorganisms is an arduous task that has not yet been completely solved by the research reported so far. Most of the research are focused on determining and/or counting the presence of specific microorganisms, mainly bacteria due to the importance they report in different current contexts such as disease transmission, in the chemical, food and pharmaceutical industries, among others.

The method proposed in this work effectively classifies and count the presence of bacteria and yeasts in microbial bioproduct obtaining a average accuracy of 0.95 in the count of bacteria and 0.78 in the case of yeast. It is composed of three fundamental modules called: Primary Operations, Statistic Operations and Classification; and it mainly uses morphological operations contour detection operations, and statistical analysis.

As a prerequisite for the correct classification of microorganisms by means of this method, it is necessary that the presence of bacteria in the suspension be greater than that of yeast, and that the bacteria must have a regular circular shape.

The accuracy achieved is closely linked to the conditions under which the images are acquired, the degree of focus of the camera, the cleaning of the lens, and greatly influenced by the diversity of sizes and shapes of the microorganisms of interest in the research, the contrast between the microorganisms and the background in which they are found, and the degree of clustering of the microorganisms.

In the verification of the results obtained by the method by means of expert criteria, it was achieved an acceptable performance for the conditions of the experiments.

V. AUTHOR'S CONTRIBUTION

Conceptualization: Jorge Peña Martín and Yelenys Alvarado-Capó.
Methodology: Jorge Peña Martín, Yelenys Alvarado-Capó, Rubén Orozco Morales and Tatiana Pichardo.
Investigation: Jorge Peña Martín.
Discussion of results: Jorge Peña Martín, Yelenys Alvarado-Capó, Rubén Orozco Morales, Tatiana Pichardo and Ailet Abreu López.
Writing – Original Draft: Jorge Peña Martín.
Writing – Review and Editing: Jorge Peña Martín, Yelenys Alvarado-Capó, Rubén Orozco Morales and Ailet Abreu López.
Supervision: Jorge Peña Martín, Yelenys Alvarado-Capó, Rubén Orozco Morales and Ailet Abreu López.
Approval of the final text: Jorge Peña Martín, Yelenys Alvarado-Capó, Rubén Orozco Morales, Tatiana Pichardo and Ailet Abreu López.

VI. ACKNOWLEDGMENTS

Thanks to the *Laboratorio de Microbiología* (Microbiology Laboratory) of the *Instituto de Biotecnología de las Plantas* (Institute of Plant Biotechnology) and its workers for dedicating time and effort to carry out the experiments necessary for this research, especially to Tatiana.

VII. REFERENCES

[1] C. Li, K. Wang, y N. Xu, «A survey for the applications of content-based microscopic image analysis in microorganism classification domains», *Artif. Intell. Rev.*, vol. 51, n.º 4, pp. 577-646, 2019, doi: https://doi.org/10.1007/s10462-017-9572-4.

[2] F. Kulwa *et al.*, «A state-of-the-art survey for microorganism image segmentation methods and future potential», *IEEE Access*, vol. 7, pp. 100243-100269, 2019, doi: https://doi.org/10.1109/ACCESS.2019.2930111.

[3] Md. F. Wahid, T. Ahmed, y Md. A. Habib, «Classification of Microscopic Images of Bacteria Using Deep Convolutional Neural Network», en *2018 10th International Conference on Electrical and Computer Engineering (ICECE)*, 2018, pp. 217-220. doi: https://doi.org/10.1109/ICECE.2018.8636750.

[4] A. Ferrari, S. Lombardi, y A. Signoroni, «Bacterial colony counting with convolutional neural networks in digital microbiology imaging», *Pattern Recognit.*, vol. 61, pp. 629-640, 2017, doi: https://doi.org/10.1016/j.patcog.2016.07.016.

[5] A. Ferrari, S. Lombardi, y A. Signoroni, «Bacterial colony counting by convolutional neural networks», en *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 7458-7461. doi: https://doi.org/10.1109/EMBC.2015.7320116.

[6] J. H. Jung y J. E. Lee, «Real-time bacterial microcolony counting using on-chip microscopy», *Sci. Rep.*, vol. 6, n.º 1, pp. 1-8, 2016, doi: https://doi.org/10.1038/srep21473.

[7] M. Woźniak, D. Polap, L. Kośmider, y T. Clapa, «Automated fluorescence microscopy image analysis of Pseudomonas aeruginosa bacteria in alive and dead stadium», *Eng. Appl. Artif. Intell.*, vol. 67, pp. 100-110, 2018, doi: https://doi.org/10.1016/j.engappai.2017.09.003.

[8] P.-J. Chiang, M.-J. Tseng, Z.-S. He, y C.-H. Li, «Automated counting of bacterial colonies by image analysis», *J. Microbiol. Methods*, vol. 108, pp. 74-82, 2015, doi: https://doi.org/10.1016/j.mimet.2014.11.009.

[9] D. T. Boukouvalas, R. A. Prates, C. R. L. Leal, y S. A. de Araújo, «Automatic segmentation method for CFU counting in single plate-Serial dilution», *Chemom. Intell. Lab. Syst.*, vol. 195, p. 103889, 2019, doi: https://doi.org/10.1016/j.chemolab.2019.103889.

[10] M. Grossi, C. Parolin, B. Vitali, y B. Riccò, «Computer Vision Approach for the Determination of Microbial Concentration and Growth Kinetics Using a Low Cost Sensor System», *Sensors*, vol. 19, n.º 24, p. 5367, 2019, doi: https://doi.org/10.3390/s19245367.

[11] D. K. Maurya, «ColonyCountJ: A User-Friendly Image J Add-on Program for Quantification of Different Colony Parameters in Clonogenic Assay», *J ClinToxicol*, vol. 7, n.º 358, pp. 2161-0495.1000358, 2017. https://doi.org/10.4172/2161-0495.1000358.

[12] H. Ates y O. N. Gerek, «An image-processing based automated bacteria colony counter», en *2009 24th International Symposium on Computer and Information Sciences*, 2009, pp. 18-23. doi: https://doi.org/10.1109/ISCIS.2009.5291926.

[13] Q. Geissmann, «OpenCFU, a new free and open-source software to count cell colonies and other circular objects», *PloS ONE*, vol. 8, n.º 2, p. e54072, 2013, doi: https://doi.org/10.1371/journal.pone.0054072.

[14] G. Corkidi, R. Diaz-Uribe, J. L. Folch-Mallol, y J. Nieto-Sotelo, «COVASIAM: an image analysis method that allows detection of confluent microbial colonies and colonies of various sizes for automated counting», *Appl. Environ. Microbiol.*, vol. 64, n.º 4, pp. 1400-1404, 1998, doi: https://doi.org/10.1128/AEM.64.4.1400-1404.1998.

[15] M. L. Clarke, R. L. Burton, A. N. Hill, M. Litorja, M. H. Nahm, y J. Hwang, «Low-cost, high-throughput, automated counting of bacterial colonies», *Cytometry A*, vol. 77, n.º 8, pp. 790-797, 2010, doi: https://doi.org/10.1002/cyto.a.20864.

[16] «ImageJ», ago. 24, 2020. https://imagej.net/Welcome (accessed Aug. 24, 2020).

[17] «Fiji», abr. 16, 2021. https://fiji.sc/ (accessed Apr. 16, 2021).

[18] P. Choudhry, «High-throughput method for automated colony and cell counting by digital image analysis based on edge detection», *PloS One*, vol. 11, n.º 2, p. e0148469, 2016, doi: https://doi.org/10.1371/journal.pone.0148469.

[19] B. Brzozowska, M. Gałecki, A. Tartas, J. Ginter, U. Kaźmierczak, y L. Lundholm, «Freeware tool for analysing numbers and sizes of cell colonies», *Radiat. Environ. Biophys.*, vol. 58, n.º 1, pp. 109-117, 2019, doi: https://doi.org/10.1007/s00411-018-00772-z.

[20] J. G. A. Barbedo, «An algorithm for counting microorganisms in digital images», *IEEE Lat. Am. Trans.*, vol. 11, n.º 6, pp. 1353-1358, 2013, doi: https://doi.org/10.1109/TLA.2013.6710383.

[21] Z. Wang, B. Ma, y Y. Zhu, «Review of Level Set in Image Segmentation», *Arch. Comput. Methods Eng.*, pp. 1-18, 2020, doi: https://doi.org/10.1007/s11831-020-09463-9.

[22] Z. Cai, N. Chattopadhyay, W. J. Liu, C. Chan, J.-P. Pignol, y R. M. Reilly, «Optimized digital counting colonies of clonogenic assays using ImageJ software and customized macros: comparison with manual counting», *Int. J. Radiat. Biol.*, vol. 87, n.º 11, pp. 1135-1146, 2011, doi: https://doi.org/10.3109/09553002.2011.622033.

[23] P. S. Hiremath y P. Bannigidad, «Digital image analysis of cocci bacterial cells using active contour method», en *2010 International Conference on Signal and Image Processing*, 2010, pp. 163-168. https://doi.org/10.1109/ICSIP.2010.5697462.

[24] A. Liu, Z. Liu, L. Song, y D. Han, «Adaptive ideal image reconstruction for bacteria colony detection», en *Information Technology and Agricultural Engineering*, Springer, 2012, pp. 353-360. https://doi.org/10.1007/978-3-642-27537-1_44.

[25] A. González-Betancourt, P. Rodríguez-Ribalta, A. Meneses-Marcel, S. Sifontes-Rodríguez, J. V. Lorenzo-Ginori, y R. Orozco-Morales, «Automated marker identification using the Radon transform for watershed segmentation», *IET Image Process.*, vol. 11, n.º 3, Art. n.º 3, 2017, doi: https://doi.org/10.1049/iet-ipr.2016.0525.

[26] B. Kis, M. Unay, G. D. Ekimci, U. K. Ercan, y A. Akan, «Counting Bacteria Colonies Based on Image Processing Methods», en *2019 Medical Technologies Congress (TIPTEKNO)*, Izmir, Turkey, 2019, pp. 1-4. doi: https://doi.org/10.1109/TIPTEKNO.2019.8895213.

[27] D. T. Boukouvalas, P. Belan, C. R. L. Leal, R. A. Prates, y S. A. de Araújo, «Automated colony counter for single plate serial dilution spotting», en *Iberoamerican Congress on Pattern Recognition*, Madrid, Spain, 2018, pp. 410-418. doi: https://doi.org/10.1007/978-3-030-13469-3_48.

[28] P. Kalavathi y S. Naganandhini, «A Hybrid Method For Automatic Counting Of Microorganisms In Microscopic Images», *Adv. Comput. Int. J.*, vol. 7, n.º 1/2, pp. 51-60, 2016. https://doi.org/10.5121/ACIJ.2016.7206.

[29] J. Zhang *et al.*, «A Multiscale CNN-CRF Framework for Environmental Microorganism Image Segmentation», *BioMed Res. Int.*, vol. 2020, 2020. https://doi.org/10.1155/2020/4621403.

[30] N. Dietler *et al.*, «YeaZ: A convolutional neural network for highly accurate, label-free segmentation of yeast microscopy images», *bioRxiv*, 2020. https://doi.org/10.1101/2020.05.11.082594.

[31] Y. He, W. Xu, Y. Zhi, R. Tyagi, Z. Hu, y G. Cao, «Rapid bacteria identification using structured illumination microscopy and machine learning», *J. Innov. Opt. Health Sci.*, vol. 11, n.º 01, p. 1850007, 2018, doi: https://doi.org/10.1142/S1793545818500074.

[32] C. Zheng, J. Liu, y G. Qiu, «Tuberculosis bacteria detection based on Random Forest using fluorescent images», en *2016 9th international congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI)*, Datong, China, 2016, pp. 553-558. doi: https://doi.org/10.1109/CISP-BMEI.2016.7852772.

[33] «Instituto de Biotecnología de las Plantas», abr. 18, 2021. https://www.ibp.co.cu/es/ (accessed Apr. 18, 2021).

[34] J. Moinar, A. I. Szucs, C. Molnar, y P. Horvath, «Active contours for selective object segmentation», en *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1-9. https://doi.org/10.1109/WACV.2016.7477572.

[35] «HDCE-X5 Camera for Microscopic», abr. 18, 2021. http://beltec.com.pe/site/beltec1/catalog/product_info.php?products_id=44 (accessed Apr. 18, 2021).

[36] A. Kaehler y G. Bradski, *Learning OpenCV 3: computer vision in C++ with the OpenCV library*. O'Reilly Media, Inc., 2016.

[37] E. R. Davies, *Computer and machine vision: theory, algorithms, practicalities*. Academic Press, 2012.

[38] J. Beyerer, F. P. León, y C. Frese, *Machine vision: Automated visual inspection: Theory, practice and applications*. Springer, 2016.

[39] A. F. Villán, *Mastering OpenCV 4 with Python: a practical guide covering topics from image processing, augmented reality to deep learning with OpenCV 4 and Python 3.7*. Packt Publishing Ltd, 2019.

[40] Z. Wang, E. Wang, y Y. Zhu, «Image segmentation evaluation: a survey of methods», *Artif. Intell. Rev.*, vol. 53, n.º 8, Art. n.º 8, 2020, doi: https://doi.org/10.1007/s10462-020-09830-9.

[41] J. Zhang *et al.*, «A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches», *Artif. Intell. Rev.*, pp. 1-70, 2021. https://doi.org/10.1007/s10462-021-10082-4.

[42] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, y D. Terzopoulos, «Image segmentation using deep learning: A survey», *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021, doi: https://doi.org/10.1109/TPAMI.2021.3059968.