Non-Proprietary Bacterial Colony Enumeration

1st Kayle Ransby
Department of Computer Science
and Software Engineering
University of Canterbury
Christchurch, New Zealand
krr39@uclive.ac.nz

2nd Richard Green
Department of Computer Science
and Software Engineering
University of Canterbury
Christchurch, New Zealand
richard.green@canterbury.ac.nz

***Abstract*—This paper proposes methods to enumerate bacterial colonies present on agar plates without the use of proprietary hardware or software. This is a very common activity in microbiology labs, however it is usually performed by hand as current dedicated equipment for this task is slow, expensive and not user-friendly. The proposed methods utilize the Hough Circle and Watershed Transforms to detect circular features in images of agar plates. A control group of 19 high quality images with optimal contrast was compared with a test group of 19 low quality smartphone images with high image noise retrieved from the University of Canterbury Microbiology Labs. These images provided a sufficient data set to develop the methods. The methods are able to achieve a control enumeration accuracy of up to 98% and 80% for the Hough Circle and Watershed Methods respectively, matching the accuracy of past research. The test enumeration accuracy of 5.5% and 8% for the same two methods compares unfavourably to prior research's enumeration accuracy but required less algorithm complexity and required no proprietary hardware or software.**

I. Introduction

One of the most common tasks performed in microbiology labs is enumerating (counting) bacterial colonies grown on agar plates. This is commonly achieved by using a pen to mark where you've been and a clicker to keep count, shown in Figure 1. This method is cheap and simple, however it is also subject to human error, as depending on the plate dilution, there could be upwards of 1000 colonies present. This also puts strain on the microbiologist performing the counting, as it is a time consuming and tedious task.

To reduce some of the strain on microbiologists, slightly more advanced counting devices utilise a magnifying glass, grid and light (typically below the plate) shown in Figure 2. Following the grid, each colony is tapped with either a pen or a probe connected to the apparatus. Each tap registers as a colony being counted, which is kept track of by a screen on the device. While this reduces eye strain for the microbiologists, it only slightly reduces the human error and the amount of time to count the plate. Additionally, a basic counter like this could cost around $650.00 USD [2], making the cost heavily outweigh the benefits.

The next step is fully or semi-automated colony counters. These machines typically consist of an enclosed chamber, HD camera, various lights and some form of partner software running either directly on the machine or on a separate computer. They capture an image of the plate using their integrated cameras, and perform different image filtering techniques, such as sharpening and edge detection, to process the image and count the number of colonies present. These machines are always extremely expensive and run proprietary software, which can't be customized beyond what the software supports [2]. An example of such a machine is shown in Figure 3.

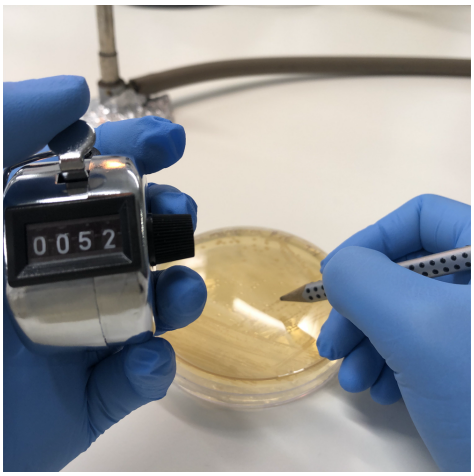

Fig. 1. Using a pen and clicker to manually counting bacterial colonies [1].

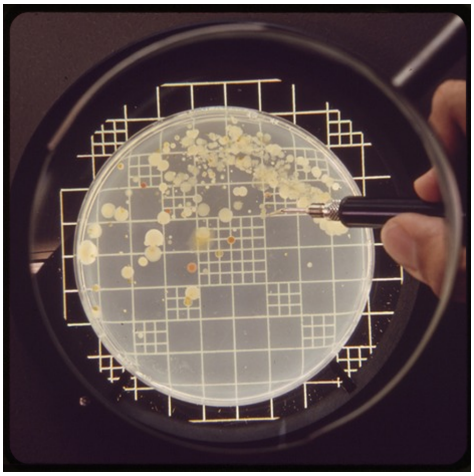

Fig. 2. View of a more advanced manual colony counting technique [3].

Clearly, all of the methods outlined above have their draw-

Fig. 3. Automated colony counter, with integrated interface. Captures image of plate, analyzes and filters image, and returns the number of colonies as well as an image of what the device determined to be colonies [4].

backs. Whether it's taking time, paying excessive money or fiddling with proprietary software, none of these solutions are ideal. It is the third, automated method that this paper aims to implement using a simple, accurate and non-proprietary solution.

II. Background

A. Review of Literature

Current solutions to automated and assisted bacterial colony counting in the form of machines pictured in Figures 2 and 3 are exorbitantly expensive. A basic manual counter costs approximately $650.00 USD and an advanced automated counter can cost anywhere between $4,000.00 to $100,000.00 USD for the machine and software [2]. Because of this, access to these machines is extremely limited and inconvenient for the vast majority of microbiologists. Current open-source solutions attempt to solve this problem by building their own automated machines for around $100.00 USD. These typically consist of an HD camera, back-light with diffuser and a jig or enclosure to hold everything in the correct place [5]–[7]. While this is certainly more affordable than purchasing a typical colony counting machine, it does not eliminate the inconvenience of using a proprietary design.

Previous research shows that automated colony counters encounter certain obstacles that can be reduced through pre-processing. Visual noise such as shadows, reflections, bacterial colony overlap, the surface the plate is sitting on and imperfections in or on the agar plate itself can introduce potentially problematic pixels that could interfere with the counting and introduce false positives. Sharpening, greyscale, noise-reducing and binarization filters can be applied to input images to reduce the effects of this unwanted noise [8].

Even with these preprocessing techniques, some methods require additional steps such as dying the bacteria [8], lack appropriate noise filtering resulting in false positives [9] or rely on third-party proprietary software to work [10]. Further research identified that simply masking the area of interest in the image (i.e. removing all pixels except the agar and colonies) eliminates the chance of any false positives outside the plate area [11].

Past research has produced solutions that align with the proposed outcome of this paper [7], [12]–[14]. In some cases, these solutions sacrifice simplicity and customizability for accuracy and speed through the use of neural networks [15]. Depending on the use-case, the ability to customize the methods is important, giving the added requirement of simple inputs and basic algorithms.

A candidate method for identifying bacterial colonies in an image is the Hough Transform, which Paul Hough developed in 1960 as a "method and means for recognizing complex patterns " [16]. Hough's goal was to optimize the analysis of bubble chamber pictures, since many photographs needed analysis. When done by hand, each photograph took several hours to analyze [16]. Essentially, Hough created a method for solving the same problem for the study of atomic particles as we would like to solve for the enumeration of bacterial colonies.

A second candidate method is the automated Watershed Transform, which works by viewing a greyscale image as a topographic surface (like a map) [17]. This method is a popular choice amongst previous researchers due it its ability to automatically segment different features and label them individually.

B. Aims and Objectives

The aim of this paper is to identify a simple method for accurately enumerating bacterial colonies using pictures taken by a smartphone. This eliminates the need for any proprietary hardware, such as a dedicated colony counting machine or HD camera, needing to be purchased or constructed, as almost everyone already has a powerful computer with an integrated HD camera with them at all times in the form of a smartphone.

The methods will be integrated using Python and use some OpenCV library functions, however the methods outlined here could be translated into the form of a mobile app for use directly on smartphones. Implementations of the Hough Circle Transform and the Watershed Transform will be compared with hand counting for speed, accuracy and error rates.

Through assistance from microbiologists at the University of Canterbury [1], the test group of images consists of two different types of bacteria: *Escherichia coli* (Gram negative) and *Serratia marcescens* (Gram negative). In addition to this, four different types and colors of agar will be used: Luria-Bertani (LB), Luria-Bertani + rifampicin (LB+R), Reasoner's 2A agar (R2A), and Tryptone Soya Agar (TSA). All of which will be prepared in 85mm agar plates.

III. Method

To prepare the agar plates, the following procedure was used: A single colony was taken from a serially diluted stock plate (LB+R) and used to inoculate a liquid culture of LB. This was incubated overnight at 37°C. The inoculum was then

serially diluted in LB by a factor of 10000. 100 µl of diluted inoculum was then plated onto the various agar types and spread with a glass spreader. Plates were incubated overnight at 37°C [1]. The results are pictured in Figure 4

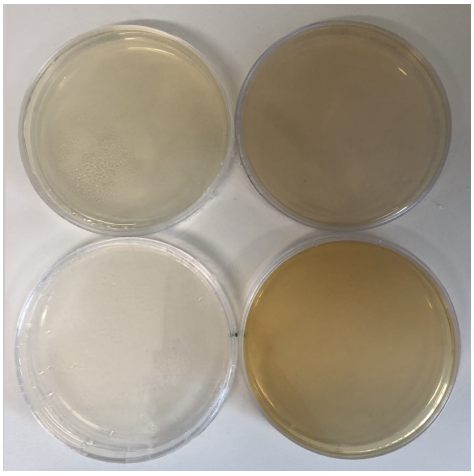

Fig. 4. The different colours and transparencies of agar to be used, from left to right: (top) LB, LB+R, (bottom) R2A, and TSA [1].

This produced a good sized sample group of images for initial testing, ensuring that the methods can cope with a range of bacteria and agar types.

Image preprocessing techniques and the two main feature detection methods were implemented as individual Python functions. The two feature detection methods used were the Hough Circle and Watershed transform. The code pipeline[1] for these methods is shown in Figure 5.

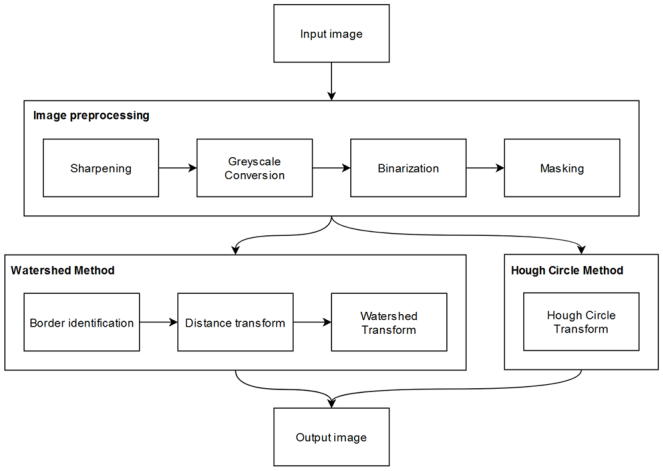

Fig. 5. The analysis pipeline.

A. Image Preprocessing

To prepare an image for enumeration with the two different methods, it needs to be preprocessed. As shown in Figure 5, the following filters are applied to the input image:

[1]See [18] for the source code.

1. Laplacian sharpening (make edges more noticeable),
2. Greyscale conversion (black-and-white, 0 - 255),
3. Binarization (two-tone, 0 - 1),
4. Masking (remove noise).

The first three steps are automatic and require no input from the user. The image could be enumerated after performing steps 1 - 3, but as mentioned in [8], there may still be problematic noise present. To prevent this, step four uses the Hough Circle Transform to identify the agar plate in the image with assistance from the user. Once an appropriate circle is identified, a mask is created using the circle as a guide. This eliminates the noise outside the plate area.

Given the vast number of contrasts for both the agar types and bacterial colonies, it is naive to assume that this method is fool-proof and will always produce a binary image containing a black background with white colonies. This can be mitigated by allowing the user to see the binary image before it is processed and decide whether to invert it or add some extra preprocessing steps if the colonies are not defined enough [12]. The extra preprocessing consists of two rounds of dilation with erosion in the middle all of which are a single pass, producing acceptable results during testing.

The implementations for both Method 1 and Method 2 rely on an image with white features on a black background, allowing for the same preprocessing pipeline to be utilised by both methods. Figure 6 shows the preprocessing pipeline applied to an input image [19].

B. Method 1: Hough Circle Transform

Our version of the Hough Circle Transform implements a detection method more complex than the standard Hough Transform called The Hough gradient method, which consists of two steps [20]. The first step performs edge detection by analyzing high gradient pixels to find possible circle centers. The preprocessing steps applied to the input images makes this a simple task, as the image only has two colours. The second step finds the best radius for each candidate center, requiring the user to enter a range over a minimum and maximum radius. These identified circles are recorded as vectors of the form (x, y, r), where x is the X position of the circle center, y is the Y position of the circle center and r is the best radius found in the specified range. Using these vectors, the circles can be drawn onto the image to provide visual feedback to the user about their choice of radii. They are also the metric used for counting the number of colonies present.

The Hough Circle Transform is a good option to use due to its robustness in the face of image noise and its ability to detect circles that are only partially visible, which is important for the detecting colonies right on the edge of the plate or that are overlapping.

C. Method 2: Watershed Transform

There are two key steps to the Watershed method, in addition to the Watershed transform itself, as seen in Figure 5. The first of these steps is to generate a border for the components in the preprocessed image. This is achieved by

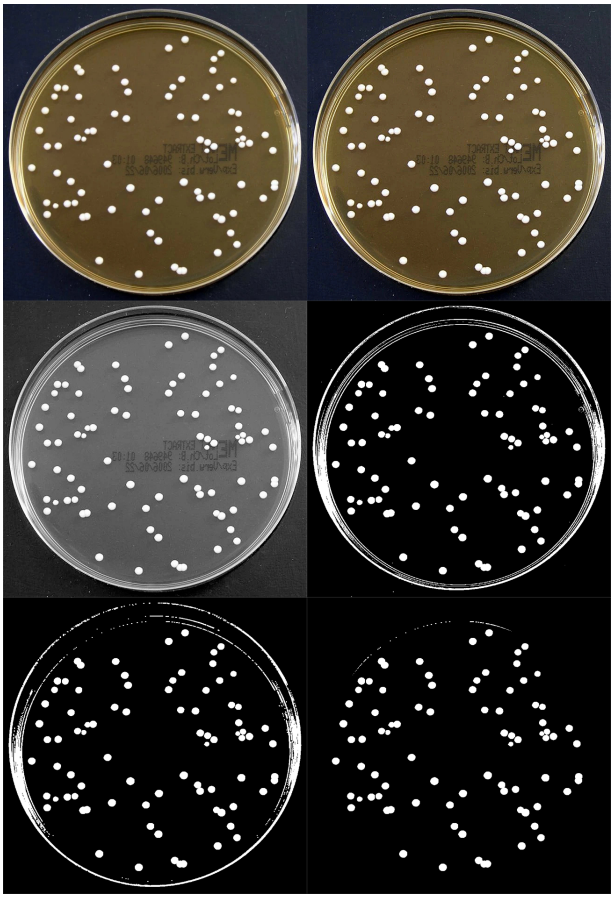

Fig. 6. Pre-processing pipeline, from left-to-right, top-to-bottom: original image, sharpening, greyscale conversion, binarization, noise reduction, masking.

subtracting an eroded image from a dilated image, resulting in a thin border surrounding all components in the preprocessed image. Once the border is generated, a distance transform is used to convert the binary preprocessed image into a topological image, with peaks being white (255) and troughs being black (0). This topological image is then binary thresholded between the values (heights) 150 - 255, this ensures only the peaks are visible in the final image. Once the binarized distance transform image is ready, the watershed transform can be applied. The Watershed Transform returns the number of individual components it has detected as well as a mask containing their positions. Once this mask has been retrieved, it is coloured red and overlaid onto the original image, giving visual feedback to the user about the number of colonies identified.

D. Testing Conditions

All tests will be conducted using the following hardware and software:

- **OS:** Xubuntu 20.04
- **Processor:** AMD Ryzen 7 1700; Eight-Cores, Sixteen-Threads; 3.00 GHz.
- **IDE:** Visual Studio Code
- **Language:** Python 3.8
- **Device:** Desktop computer
- **Camera:** IPhone 10; 12MP
- **OpenCV Version:** 4.5.1

IV. Results

A. Control Group

The control group is a set of 19 high quality images of *Staphilococcus aureus* colonies on LB agar plates with a bright background [21]. These images, which are under ideal conditions, allow for a control group to compare the lower-quality smartphone images to. Results for the control group testing are pictured in Figure 7.

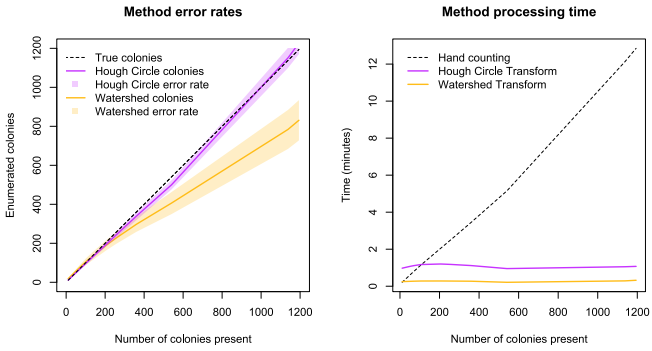

Fig. 7. Graph showing the total number of colonies enumerated by each method with their respective error rates (left) and the time taken to enumerate the colonies using each method (right)

In order to achieve comparable results, the number of colonies enumerated by hand is assumed to be the ground truth. With this in mind, the error rate graph shows the number of colonies enumerated by each method as the purple and orange lines, and their respective error rates are shown as the light purple and light orange areas surrounding these lines. These error rates are calculated as the number of colonies enumerated by a particular method plus or minus the sum of the false positive rate (FPR) and the false negative rate (FNR). Written as an equation, this becomes:

$$error(x) = C(x) \pm (FPR(x) + FNR(x)) \tag{1}$$

Where $C(x)$ is the number of colonies enumerated by the method at expected value x, and $(FPR(x) + FNR(x))$ is the sum of the FPR and FNR at expected value x. The equation returns two values for a given x value of the form (upper-bound, lower-bound). This upper and lower-bound are what is used to define the error rate of the respective methods.

For better accuracy, the light-coloured areas should be small and the slope of the line should be as close to the "True colonies" (black dashed) line as possible. The control group's results show that both methods are viable for appropriate plate densities. The peak appropriate plate density for the Watershed Method is $\approx 300 \pm 53$. The peak appropriate plate density for the Hough Circle Method of $\approx 1200 \pm 60$ is far greater as its

deviation from the true count is minimal. If we simplify these ranges to 0 - 1200 for the Hough Circle Method and 0 - 300 for the Watershed Method, the former is $\approx 75\%$ more accurate than the latter. Despite this, both methods are essentially 100% accurate with plate densities at or below ≈ 200.

The mean time taken to enumerate the colonies using each method was identified as (M:SS) $\approx$ 2:29, 1:04 and 0:16 for the Hand counting, Hough Circle and Watershed Methods respectively. The mean overall detection accuracy (calculated as the number of enumerated colonies divided by the number of expected colonies) was identified as 98% for the Hough Circle Transform and 80% for the Watershed Transform.

The important comparators for the different methods are shown in Table I, allowing conclusions to be drawn from the data.

TABLE I
CONTROL TEST RESULTS

Test Result	Comparators			
	Time	*Accuracy*	*FPR*	*FNR*
Mean Hough	01:04.69	98%	7.26	5.53
Min Hough	00:28.76	82%	0	0
Max Hough	02:13.29	100%	60	22
Mean Water	00:16.11	80%	6.05	19.47
Min Water	00:08.55	62%	0	0
Max Water	00:26.69	100%	53	100

The same original image [19], shown in Figure 8, was enumerated using Method 1 and 2. The results from each method are shown below in their respective subsection.

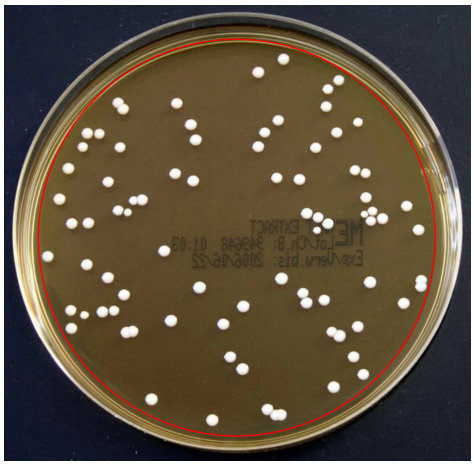

Fig. 8. Original image with plate identified.

The identified plate (red) circle can be seen in Figure 8. This circle is the once transformed into a mask during the preprocessing stage. The fully preprocessed image is shown in Figure 9

B. Method 1: Hough Circle Transform

After performing the Hough Circle Transform on the image shown in Figure 9, the output image shown in Figure 10 is generated. This image consists of the original input image

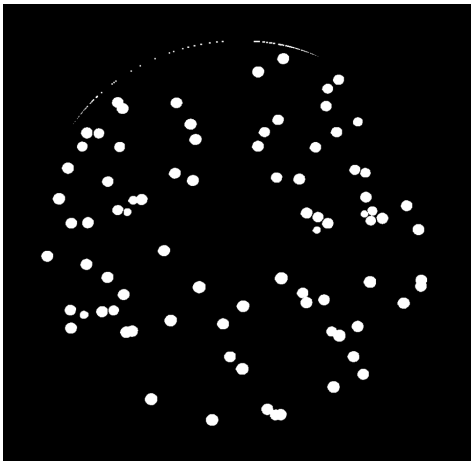

Fig. 9. Pre-processed image ready to be enumerated.

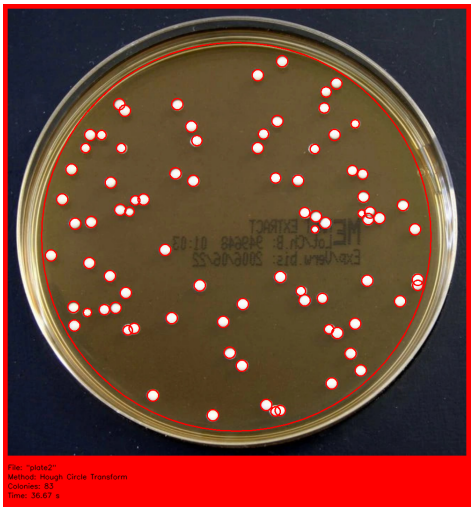

Fig. 10. Output image after performing the Hough Circle Transform.

with the plate mask and identified colonies highlighted by red circles overlaid on top. A border containing relevant information about the test is also generated, containing the filename, method used, number of colonies identified and the time taken to achieve the result. In this example, the number of colonies identified is 83, and the number of expected colonies is 83, giving an accuracy of 100.00%. As mentioned earlier, the Hough Circle Transform is robust in the face of image noise and partially obscured circles, meaning it is fully capable of detecting colonies that overlap if the original image is of sufficient quality.

The test group is a set of 19 low quality images of the plates described in section III. These images are not ideal but provide a valuable contrast to the control group and represent the image quality expected from smartphone cameras in a lab. Results for the test group after performing the Hough Circle Method are pictured in Figure 11.

From Figure 11 alone, it is clear to see that the test group did not perform as well as the control group. The average

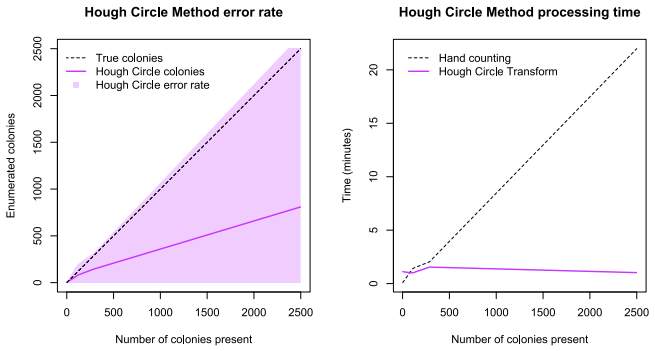

Fig. 11. Graph showing the test results for the Hough Circle Method

accuracy for the test group is 0.36% compared to the control groups 98%. The important comparators for the Hough Circle Method are shown in Table II.

TABLE II
HOUGH CIRCLE METHOD TEST RESULTS

Hough Circle Method	Comparators			
	Time	*Accuracy*	*FPR*	*FNR*
Mean	01:05.91	0.36%	38.16	416.89
Min	00:31.06	0.08%	0	0
Max	01:38.50	5.50%	153	2178

C. Method 2: Watershed Transform

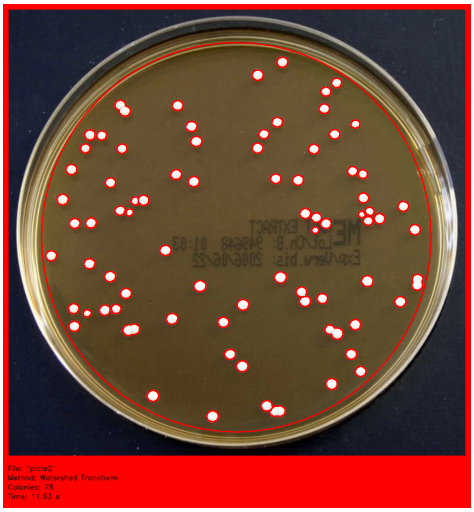

Fig. 12. Output image after performing the Watershed Transform.

After performing the Watershed Transform on the image shown in Figure 9, the output image shown in Figure 12 is generated. This image has the same composition as the one generated by Method 1. In this example, the number of colonies identified is 78, and the number of expected colonies is 83, giving an accuracy of 93.98%. Examining the output, it is clear that the watershed method is not as robust as the Hough Circle method when it comes to detecting overlapping colonies. Although, the overlapping colonies are still detected, just not separated. In general, the watershed method requires higher quality images and more distinguished colonies for the required separation and best accuracy.

Results for the same test group described in section IV-B after performing the Watershed Method are pictured in Figure 13.

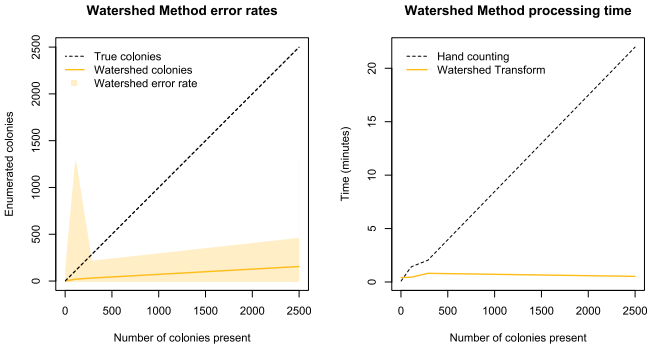

Fig. 13. Graph showing the test results for the Watershed Method

Given that the Watershed Method requires higher quality images for acceptable accuracy, it is no surprise that the low quality test group images resulted in an average accuracy of 0.07%, compared to 80% for the control group images. The important comparators for the Watershed Method are shown in Table III.

TABLE III
WATERSHED METHOD TEST RESULTS

Watershed Method	Comparators			
	Time	*Accuracy*	*FPR*	*FNR*
Mean	00:29.73	0.07%	42.68	601.16
Min	00:16.20	0.00%	0	2
Max	01:04.41	8.00%	310	2500

V. CONCLUSION

This paper proposed methods to enumerate bacterial colonies present on agar plates without the use of proprietary software or hardware. These methods performed well with high quality images, but lacked the ability to enumerate lower quality images as expected, particularly for the Watershed Method.

The performance of past research on high quality images that align with the control group were able to achieve accuracy's of approximately 98% [10], [11], matching the accuracy of the Hough Circle Method presented in this paper. As stated earlier, the Watershed Method is only able to achieve an accuracy of approximately 80%, however this is due to insufficient segmentation within the images, not lack of colony detection. Past research's performance when using low quality images that align with the test group were able to achieve

accuracy's of approximately 92% [7], [12]. Unfortunately, the methods presented here were only able to achieve a maximum accuracy of 5.5% and 8.0% for the Hough Circle and Watershed Methods respectively.

A. Limitations

The test data set consisting of 19 pictures of agar plates taken by a smartphone, was relatively limited in terms of variation, only really providing minor contrast differences in the pigment of the agar and bacteria. They were also low resolution and subject to noise from blur and glare.

In order to have consistent test metrics for the data collected during analysis, the number of colonies counted by hand was used as the ground truth. This is okay when comparing the methods to each other, but introduces incorrect scaling when comparing the data to past research. In reality, there is always some level of human error involved with hand counting.

B. Future Research

Future improvements to the methods outlined here could include automation of plate identification, as currently this is the majority of the time spent when using the Watershed method. This improvement could make the watershed method essentially instant.

Testing on a larger data-set containing different pigmentation and opacity levels for both the agar and bacteria would allow for further optimization of the preprocessing techniques and perhaps creation of "presets" for different contrast levels.

Additionally, removal and optimization of the OpenCV debug windows through the use of graphics libraries such as Tkinter or implementation into a mobile application for use and testing on a smartphones would make the algorithms more user-friendly.

The Watershed Method is very accurate for quickly and automatically identifying where the colonies are on the plates, but not so accurate for enumerating them. Given this knowledge, the Watershed method could be used as a "seed" for the Hough Circle Method, perhaps by extrapolating a radii range from the maximum and minimum area of the contours identified by the Watershed Method.

References

[1] T. Hiscox, Private communication, 2021.

[2] A. Dance, "Down for the count," *The Scientist*, May 2013. [Online]. Available: https://www.the-scientist.com/lab-tools/down-for-the-count-39400

[3] L. Learning. Counting bacteria — boundless microbiology. [Accessed: 028-08-2021]. [Online]. Available: https://courses.lumenlearning.com/boundless-microbiology/chapter/counting-bacteria/

[4] Synbiosis, A Division of the Synoptics Group. Protocol 3 - automatic colony counting and zone measuring equiptment. [Accessed: 028-08-2021]. [Online]. Available: https://www.synbiosis.com/product/automated-colony-counting-zone-measurement-protocol-3/

[5] K. Claytor, "Development and implementation of an efficient automated cell colony and plaque counter," *University of Illinois Internal Physics Publication*, 2008.

[6] J. Marlin and J. Starr, "A matlab approach to automated bacterial colony counting," Washington University in St. Louis, 1 Brookings Drive St. Louis, MO 63130USA, Tech. Rep., 2017.

[7] S. Kumar, R. V. R. Choudhury, and S. Laxman, "Colonizer: Anandroid os based automated microbial colony counter," *PeerJ Preprints*, vol. 5, no. e2792v1, 2017. [Online]. Available: https://doi.org/10.7287/peerj.preprints.2792v1

[8] G. Kaur and P. Sethi, "Article: A novel methodology for automatic bacterial colony counter," *International Journal of Computer Applications*, vol. 49, no. 15, pp. 21–26, July 2012, full text available.

[9] P. Choudhry, "High-throughput method for automated colony and cell counting by digital image analysis based on edge detection," *PLOS ONE*, vol. 11, no. 2, pp. 1–23, 02 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0148469

[10] Q. Geissmann, "Opencfu, a new free and open-source software to count cell colonies and other circular objects," *PLOS ONE*, vol. 8, no. 2, pp. 1–10, 02 2013. [Online]. Available: https://doi.org/10.1371/journal.pone.0054072

[11] A. Khan, A. Torelli, I. Wolf *et al.*, "Autocellseg: robust automatic colony forming unit (cfu)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques." *Sci Rep 8*, no. 7302, 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-24916-9

[12] W. chun foong, J. Y. Yeo, and S. K.-E. Gan, "Apd colony counter app: Using watershed algorithm for improved colony counting," *Nature Methods*, 08 2016. [Online]. Available: https://dx.doi.org/10.1038/an9774

[13] W.-B. Chen and C. Zhang, "An automated bacterial colony counting and classification system," *Information Systems Frontiers*, vol. 11, no. 4, pp. 349–368, 09 2009. [Online]. Available: https://doi.org/10.1007/s10796-009-9149-0

[14] 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, pp. 1–6, 03 2012. [Online]. Available: https://doi.org/10.1371/journal.pone.0033695

[15] Medixgraph Inc. Cfu.ai- cfu counting with ai. [Accessed: 028-08-2021]. [Online]. Available: https://www.cfu.ai/

[16] P. V. C. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3 069 654, Dec. 18, 1962.

[17] OpenCV. Opencv - image segmentation with watershed algorithm. [Accessed: 028-08-2021]. [Online]. Available: https://docs.opencv.org/master/d3/db4/tutorial_py_watershed.html

[18] K. Ransby. A python script to count bacterial colonies on agar plates. [Accessed: 028-08-2021]. [Online]. Available: https://github.com/krransby/colony-counter

[19] Thermo Fisher Scientific Inc. Malt extract agar. [Accessed: 028-08-2021]. [Online]. Available: https://www.fishersci.se/shop/products/malt-extract-agar-4/10168882

[20] OpenCV. Opencv - hough circle transform. [Accessed: 028-08-2021]. [Online]. Available: https://docs.opencv.org/3.4/d4/d70/tutorial_hough_circle.html

[21] Q. Geissmann. Opencfu - count your colonies. [Accessed: 028-08-2021]. [Online]. Available: http://opencfu.sourceforge.net/samples.php