Discrimination Colonies of *Staphylococcus aureus* and *Salmonella enterica* by Using Machine Learning

Manao Bunkum and SarinpornVisitsattapongse
Department of Biomedical Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
Email: {61601007, sarinporn.vi}@kmitl.ac.th

***Abstract*—Discrimination of bacteria is necessary because bacteria are remarkably crucial as they can contaminate in many possible ways such as food, water, air or environment, and lead to pathogenesis. In recent times, the technique of discrimination and counting bacteria are not only have a large expenditure of time and funds, but the discrimination method also required an expert in field. According to above, this research introduces an algorithm for solve those problems by using *Staphylococcus aureus* and *Salmonella enterica* as a pilot study. The stab technique, to inoculated bacteria in Luria-Bertani agar (LB agar) plates, was used for providing a good detection. In addition, the image segmentation technique was invited to train the separate colony of each bacterium for machine learning algorithm. It took approximately 800 images of bacteria for training process. The algorithm had resulted in 98.75 % of accuracy in bacteria's colony counted and 98.12% in discriminated between *Staphylococcus aureus* and *Salmonella enterica* from stab technique. The algorithm had resulted in 33.61 % of accuracy in bacteria's colony counted and 41.73% in discriminated between *Staphylococcus aureus* and *Salmonella enterica* from spread technique.**

***Index Terms*—bacteria, *Staphylococcus aureus*, *Salmonella enterica*, webcam camera, machine learning algorithm**

I. Introduction

These days, Bacteria are wildly found in environment and some of them have a role in human's daily life. The bacteria can contaminate in food, water, air or soil. Some types of bacteria are pathogenic bacteria that cause many illnesses or diseases. The secretion or blood is used to find the bacterial or the cause of the diseases [1]. This research is focused on 2 bacteria including *Staphylococcus aureus (S. aureus*) and *Salmonella enterica* (*S. enterica*).

S. aureus is gram-positive cocci in cluster. The colonies of *S. aureus* are round, convex and rather golden in some colonies. *S. aureus* is bacteria in human body that usually found on the skin (skin flora) and nasal carrier. *S. aureus* has cell wall protein A, which has function to protect cell from white blood cell destruction. *S. aureus* has fibronectin-binding protein for cell fixation. *Staphylococcus sp.* groups can secrete many toxins such as staphylococcal enterotoxin, toxic shock syndrome toxin1 (TSST-1), exfoliative toxin, cytolytic toxin and other enzymes. These toxins are the causes of food poisoning, watery diarrhea, feverish, and red blood cell, white blood cell or epidermal layer destruction [2].

S. enterica is a rod-shaped and gram-negative. The colonies of *S. enterica* are round, convex and rather white. *S. enterica* can be found in gastrointestinal tract of mammals, poultry and reptiles. This bacterium is the lead of causes Salmonella gastroenteritis, enteric fever and salmonella septicemia [3].

The common methods for bacteria characterization are selective media, differential media and biochemical test, but these techniques take quite a long time and cost a lot. Moreover, these technics must be done by specialist in the step of bacteria characterization and bacteria counting.

Machine learnings are divided in two sub-topics, which are Supervised machine learning and Unsupervised machine learning. Unsupervised learning algorithm has not a required to train the data input and output by human. This algorithm will analyze data by itself and automatically find structure in data by extracting useful features and analyzing its structure. On the other hand, Supervised learning algorithm has required the train data (input and output) by human. The supervised algorithm will learn the relationship between input and output. The training will have completed when the algorithm has a desire accuracy rate [4]. Supervised learning algorithm has two main types; Classification and Regression. The Regression algorithm estimates function from the input to numerical or continuous output, while the Classification algorithm estimates the function from the input to discrete or categorical output. In the Fig. 1 shows the difference between classification that category output and regression that has continuous output [5].

In this research, the samples of *S. aureus* and *S. enterica* on enriched media (Luria-Bertani: LB agar) plate were studied by using image segmentation and Classification algorithm of Machine Learning. The image segmentation is selected for detecting and counting colony of bacteria and Classification algorithm is chose

Manuscript received August 11, 2020; revised October 23, 2020.

doi: 10.18178/ijpmbs.10.1.30-34

for discrimination the bacteria colony on enriched media (Luria-Bertani: LB agar). For these reasons, this technic will decrease time spending on determination and bacteria counting.

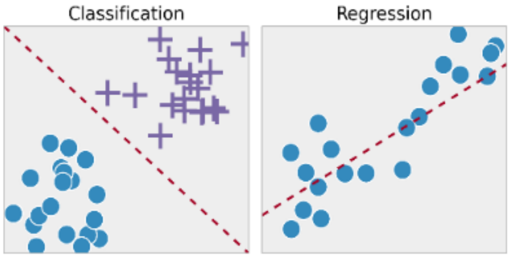


Figure 1. Show the difference between classification and Regression.

II. METHODOLOGY

A. Bacteria

S. aureus and *S. enterica* were selected for the study in this research field. The colonies of both bacteria are round, convex and several color. The samples were inoculated on LB agar plate by using the stab technique to separate the single colony for clearly detection. Ten-fold serial-dilution methodology was used for determining the proper concentration of bacteria. The concentration between 10^{-6} to 10^{-10} plaque forming unit (pfu) were counted by using image processing [5]. In this experiment, the concentration 10^{-7} were applied to spreading on LB agar plate because this concentration shows 30-50 bacterial colony on agar plate, by using visual inspection. All of the culture plates were cultivated at 37° C for 16 hours in incubator chamber. After that, the colonies result of *S. aureus* and *S. enterica* will appear on LB agar plate as show in Fig. 2.

(a) (b)

Figure 2. (a). Colonies of *S. aureus* concentration 10^{-7}; (b). Colonies of *S. enterica* concentration 10^{-7}.

All of the culture plate samples were captured by using webcam camera, which has high quality glass lens, resolution 1920×1080 pixels and auto focus function. The distance between camera and plate must be fixed for the same size of image's colonies. The dark background was used for contradistinction with colonies of bacteria on culture plate.

B. Agar Plate Preparation

This research used the Luria-Bertani agar (LB agar) for culture bacteria. LB agar consists of 10 g. of peptone, 5 g. of yeast, 10 g. of sodium chloride, and 15 g. of agar powder for solution 1 Liter. The Erlenmeyer flask of mixed ingredient was set into autoclave for sterilization process. The plastic petri dish was prepared by aseptic technique in biosafety cabinet class II. The agar solution was poured in petri dish approximately 20 milliliters per plate then place the lid on the plate and left them to cooling for 30-60 minutes.

C. Image Segmentation

The images of sample were detected by using edge detection feature. The feature edge detection was applied for finding the edge and frame of the colony. The feature edge detection consists of pixels with derivative values that exceed a certain threshold and measure the discontinuity of the grayscale [7], [8].

The image segmentation was used with images of sample to simplify image analysis. The segment can represent the objects or part of objects, which contain a number of pixels. The image segmentation was sorted the pixels into larger components and cut-off the process of consider for each pixel as units of observation [4].

D. Training Data

The images from the image segmentation part were trained in Classification algorithm of Machine Learning. The Classification algorithm was used with Logistic Regression Analysis (Binary Regression Analysis) for separate two groups of data. The trained data in algorithms called the training set. The training set is the material that the computer uses to learn how to process image or data. Machine learning algorithms is the same as the abilities of human brain to take many inputs and weigh them [9].

This research used the Classification for counting and discriminating bacteria. All sample of bacteria were trained in algorithm and defined the type of bacteria, then the model was exported from training.

III. EXPERIMENTAL AND RESULT

A. Sample of Bacterial

First of all, stab *S. aureus* and *S. enterica* on LB agar 40 plates, each plate stab 15-20 times these are sample of controlling. Second, spread *S. aureus* and *S. enterica* concentration 10^{-7} on LB agar 20 plates, these are sample of testing. Next, all plates were incubated at 37° C for 16 hours. After 16 hours the colonies will appear on LB agar.

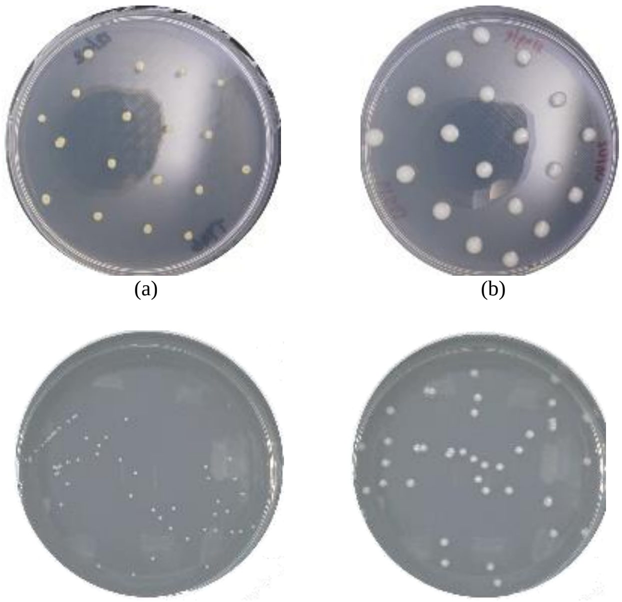
The result of stab technique shows the colonies of *S. aureus* are round, convex, golden, in the middle of colony, and dimeter about 3-4 millimeter as show in Fig. 3(a). The colonies of *S. enterica* are round, convex, white and dimeter about 6-10 millimeter as show in Fig. 3(b).

The result of spread technique, the colonies of *S. aureus* are round, convex, rather light or dark yellow tone and 2 millimeter of dimeter as show in Fig. 3(c). The colonies of *S. enterica* are round, convex, white and dimeter about 5 millimeter as show in Fig. 3(d).

B. Image Segmentation

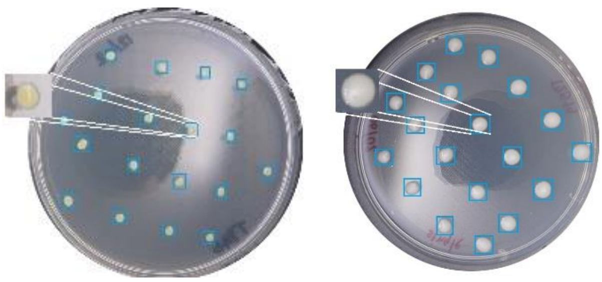
After the image of bacteria has done in edge detection feature and image segmentation. The edge detection feature can detect edge of colonies and use image segmentation with the sample image of controlling.

The colonies of the stab technique sample *S. aureus* and *S. enterica* were detected by edge detection feature and frame by the image segmentation. The feature can detect all 18 colonies of *S. aureus* as show in Fig. 4(a), and the feature can detect all 19 colonies of *S. enterica* as show in Fig. 4(b).



(a) (b) (c) (d)

Figure 3. (a). Show result of *S. aureus* controlling; (b). Show result of *S. enterica* controlling; (c). Show result of *S. aureus* testing; (d). Show result of *S. enterica* testing.



(a) (b)

Figure 4. (a). The result when use image segmentation with *S. aureus*; (b). The result when use image segmentation with *S. enterica*.

C. Training Data

The sample of controlling image of *S. aureus* and *S. enterica* from image segmentation was trained in machine learning algorithms. In this research used the keyword "T746" for *S. aureus* and "T2519" for *S. enterica*.

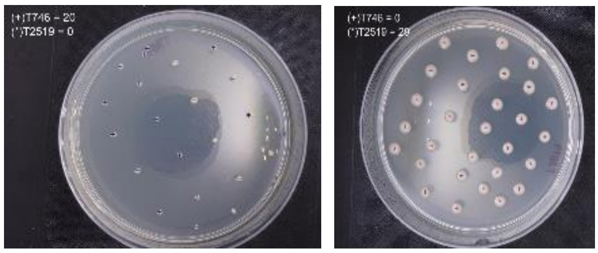
This research used the training set; 784 samples of *S. aureus* and 810 samples of *S. enterica* for training in algorithm. The example of the training set show in the Fig. 5.

D. Evaluation

The untrained stab technique sample controlling image of *S. aureus* and *S. enterica* was captured for evaluate the accuracy of algorithm. The Fig. 6(a) is the result of controlling image of *S. aureus* (T746). The algorithm can define type of bacteria and count the colonies of bacteria correctly. As the result, *S. aureus* (T746) sample, the algorithm-counted is 20 colonies while the eye-counted is 21 colonies, which can assume that the algorithm has a high accuracy performance.

The Fig. 6(b) is the result of controlling image of the stab technique sample *S. enterica* (T2519) sample, there are 29 colonies by algorithm-counted while the eye-counted exist 29 colonies.

Figure 5. Show the training set of *S. aureus* and *S. enterica* that train in machine learning algorithm.



(a) (b)

Figure 6. The result when test controlling image of *S. aureus* or T746 (a) and *S. enterica* or T2519 (b).

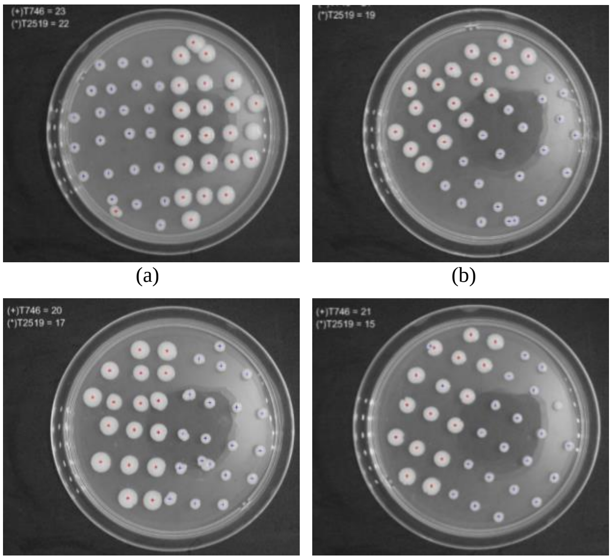
The algorithm can define type and count bacteria correctly when compare with human eye. The Figure 7 is the result of controlling image of the stab technique sample both of *S. aureus* and *S. enterica* in the same plate.

The Fig. 7(a) there are 23 colonies of *S. aureus* and 22 colonies of *S. enterica* by using this algorithm compare with 24 colonies of *S. aureus* and 22 colonies of *S. enterica* by using visibility of human eye.

The Fig. 7(b) there are 21 colonies of *S. aureus* and 19 colonies of *S. enterica* by using this algorithm compare with 21 colonies of *S. aureus* and 19 colonies of *S. enterica* by using visibility of human eye.

The Fig. 7(c) there are 20 colonies of *S. aureus* and 17 colonies of *S. enterica* by using this algorithm compare with 21 colonies of *S. aureus* and 17 colonies of *S. enterica* by using visibility of human eye.

The Fig. 7(d) there are 21 colonies of *S. aureus* and 15 colonies of *S. enterica* by using this algorithm compare with 20 colonies of *S. aureus* and 16 colonies of *S. enterica* by using visibility of human eye.



(a) (b) (c) (d)

Figure 7. (a), (b), (c) and (d) are the results when test controlling image of *S. aureus* or T746 and *S. enterica* or T2519 in the same plate.

The evidence indicated that the accuracy of counting colony of bacteria is 98.75% and the types discrimination of bacteria is approximately 98.12%.

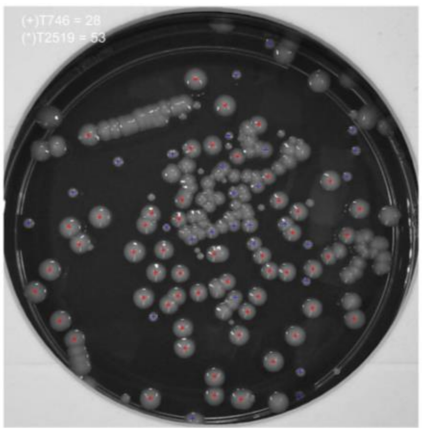


Figure 8. The result when test testing image of *S. aureus* or T746 and *S. enterica* or T2519 in the same plate.

Alternatively, in practical, the captured of the spread technique sample images of *S. aureus* and *S. enterica* in the same plate for algorithm testing as shows in the Fig. 8. It illustrated the discriminates of two bacteria precisely. Nevertheless, in the counting process, there are 28 colonies of *S. aureus* and 53 colonies of *S. enterica*, whereas there are 33 colonies of S. aureus and 127 colonies of *S. enterica* by using visibility of human eye. Thus, the accuracy of algorithm in counted is 33.61 % and the discrimination accuracy is 41.73%

IV. Conclusion and Discussion

The main aim of this research is to create algorithm that can discriminate the sample of bacteria (*S. aureus* and *S. enterica*) with a high accuracy. The outcome of this research can reduce time, for discrimination type of bacteria, and budget, for hiring the experts to analyses the bacteria result. The result in part of controlling image counted, the colonies of bacteria and has an accuracy at 98.75 %; also, the part of discriminating by using controlling image give an accuracy at 98.12 %. However, in testing, the testing image has low accuracy than the control. The accuracy of testing image for the Classification algorithm in counted is 33.61 % and discriminated is 41.73%. As a consequence, the error occurred because there has a lot of connecting colonies, low quality of image, for each picture has difference the concentration of light. For this reason, the improve of the aggregated and connecting colonies by dilute the solution concentration, or select an appropriate method to spread plate for more clearly single colony of bacteria are planning for future work. Furthermore, the camera specification would be change for good quality of image, and select more accurate algorithm to apply in part segmentation image for the good detection colonies. Also, this research can be a prototype for another bacteria discrimination.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Sarinporn Visitsattapongse is responsible for research design, research summary, recommendation and contributes for data analysis. Manao Bunkum's contribution is data analysis and research experiment.

Acknowledgment

We were acknowledged, and overwhelming supported by our respected professor and Mr. Nuttanan Wanluk, staff from Department of Biomedical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL). This work was supported by King Mongkut's Institute of Technology Ladkrabang Research Fund [grant number: KREF016012]

References

[1] P. Phonchalermpong, *et al.* Sources of microorganisms contaminated in food. [Online]. Available: http://www.foodnetworksolution.com/wiki/word/3222

[2] Staphylococcus aureus. (2019). [Online]. Available: https://en.wikipedia.org/wiki/Staphylococcus_aureus#:~:text=Staphylococcus%20aureus%20is%20a%20Gram,tract%20and%20on%20the%20skin

[3] Salmonella enterica. (2019). [Online]. Available: https://en.wikipedia.org/wiki/Salmonella_enterica

[4] I. Salian. (July 2019). Supervize me: What's the difference between supervised, unsupervised, semi supervised and reinforcement learning. [Online]. Available: https://blogs.nvidia.com/blog/2018/08/02/supervised-unsupervised-learning/

[5] M. J. Garbade. (July 2019). Regression versus classification machine learning: What's the difference? [Online]. Available: https://medium.com/quick-code/regression-versus-classification-machine-learning-whats-the-difference-345c56dd15f7

[6] N. Vongmanee, *et al.*, "An automated colony counter for serial-dilution culture method," in *Proc. 10th International Conference on Bioinformatics and Biomedical Technology*, New York, 2018, pp. 42-45.
[7] J. G. A. Barbedo, "A review on methods for automatic counting of objects in digital images," *IEEE Latin America Transaction*, vol. 10, pp. 2112-2124, 2012.
[8] Image segmentation in deep learning: Methods and application. (July 2019). [Online]. Available: https://missinglink.ai/guides/neural-network-concepts/image-segmentation-deep-learning-methods-applications/
[9] Training data. (July 2019). [Online]. Available: https://www.techopedia.com/definition/33181/training-data

Sarinporn Visitsattapongse is a lecturer in Department of Biomedical Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang. She received a bachelor's degree of science in major of biology, master's degree of science in major of molecular genetics and genetic engineering and philosophy of doctoral degree in medical technology. She used to work in medical technology field in image processing, and work in molecular biology of proteins. Now her current focus is about medical that may be healthcare device, home healthcare or healthcare technology.