Colour segmentation of Gram-Negative bacteria using graph Quadratic Form and Random Walker

B D Satoto¹, I Utoyo² and R Rulaningtyas³

¹S3 Student Applied Mathematics, Sains and Technology, Airlangga university of Surabaya
²Applied Mathematics, Sains and Technology, Airlangga university of Surabaya
³Physics, Sains and Technology, Airlangga university of Surabaya

E-mail: m.i.utoyo@fst.unair.ac.id, budids@gmail.com

Abstract. Gram-negative bacteria are one of the bacteria that are often pathogenic to the human body. This bacterium causes resistance due to nosocomial with other Gram-negative bacteria. In the medical stage, the bacteria that cause nosocomial traits removed first before antibiotic therapy carried out on the main bacteria. To identify these bacteria, the clinical laboratory needs to make manual observations under a microscope. The approach taken in this research is using the image processing technique. There are four stages: pre-processing, segmentation, feature extraction, and identification. Segmentation is a stage to emphasize the object sought in an image. In this research, the approach used to capture objects is one of them using the Graph Quadratic Form algorithm. This algorithm chose because it can determine the shortest distance of the object from the nearest node so that the process of convergence of the object search becomes faster. The result is that this algorithm is better than the morphology-based algorithm and the contour-based algorithm, while the number of samples taken from 50 patients affected by Gram-negative bacteria. The image under research has a size of 512x512 pixels, a resolution of 72 dpi with a bit depth of 24. The segmentation process is carried out on Gram-negative bacterial images using two classes producing an average accuracy of 89% to Ground truth.

1. Introduction
Gram-negative bacteria are one of the microorganisms that caused respiratory tract diseases in Indonesia. It is related to research conducted that many Gram-negative bacteria cause diseases in the Asian region. In several cases of examination, it found that Gram-negative bacteria have a nosocomial infection [1]. If there are two bacteria found in the sputum sample, the second bacterium must be cleaned first before conducting therapy on the main bacteria. It is related to the handling of antibiotic treatment in patients. This research needs to differentiate it, so the manual observation of identifying bacteria still does in the laboratory [2].

The development of visual observations in laboratories began to use image processing and machine learning in the learning process. There are four main steps in the image processing process, including initial processing, division of the image area, determination of features, and recognition. At the stage of segmentation, the image will be clustered based on the value of its closeness. There are several studies of bacteria related to image processing. Including those conducted by Hiremath (2013) about the characteristics of bacterial cell image by Gram staining using RGB color space to YCbCr, this...
research is doing Image enhancement using Histogram entropy for equalization, improving global thresholding and improving morphological image operations [3]. From one hundred samples taken, 58 patients identified as Gram-positive. Thirty-four patients were Gram-negative, and eight patients were misidentified so that the accuracy was 92% [4].

Amir Reza Miandarhoie conducted image processing research using graphs (2014) said that One of the most versatile of the Voronoi algorithms is being able to segment the optimal model [5]. The weakness of the iterative algorithm is the long turnaround time caused by this algorithm requires more iteration to produce the best segmentation. With the Voronoi diagram estimation, the approach to get the object distance value in the segmentation process did with a shorter time compared to the iterative algorithm [6].

This research conducted to get the benefits of graph algorithms in the segmentation of color images of bacteria. The gap with previous research is to get the results of segmentation that has been done by many researchers is image enhancement, morphologically based approaches, and active contours. The uniqueness of this research is the use of Graphs to help the segmentation process by finding the closest path to the object. It is not only intended to obtain convergence but also to reduce the need for ground truth by finding a way taken by a computer.

2. Literature Review
The literature review chapter discusses the research done previously, namely, research on bacterial identification, image processing, and the use of Graph in image processing.

2.1. Gram-negative bacteria
This bacterium has two layers in the cell wall. First, the outer layer consists of proteins and polysaccharides. Second, the inner layer contains thinner peptidoglycan. This cell wall absorbs the red color during Gram’s coloring [7]. Pathogenic Gram-negative bacteria are more dangerous because the outer layer of the cell wall can protect bacteria. It could be a host defense system by blocking the entry of antibiotic drugs. Lipopolysaccharide compounds on the outer membrane are toxic to the place of residence[8].

2.2. Image processing
Image processing is one of the methods used to obtain information from an image. There are four main processes in image processing, including initial processing with repairs, division of the image object area or segmentation, determination of features that used, and identification of objects [9]. Pre-processing prepares data so that the data is ready to be processed in the segmentation process. At the scene of segmentation, the image is divided into several colors so that it become identified. The feature selection phase will select the features needed in the identification process[10].

The stages of image processing shown in Figure 1.

![Figure 1. Image processing.](image)

2.3. Segmentation
The division of object areas is an essential part of the image analysis process [11]. At this stage, the process will partition the image into sections based on the value of the closeness of the color, shape, texture, and pixel membership [12].

2.4. Graph Quadratic Form
If there are input images with a resolution of n x n pixels, It can liken it to having an n x n matrix. For every n x n matrix A = (\(a_{ij}\)) with non-negative entries on the diagonal [13].

\[
\max_{x_r \in [-1, 1]} \sum_{r,s} a_{rs} x_r x_s = \max_{x_r \in [-1, +1]} \sum_{r,s} a_{rs} x_r x_s ;
\]  

(1)
\[
\max_{\|x\|_2 \leq 1} \sum_{r,s} a_{rs} \langle x_r, x_s \rangle = \max_{\|x\|_2 \leq 1} \sum_{r,s} a_{rs} \langle z_r, z_s \rangle.
\]

(2)

Proof. Since the diagonal entries are non-negative, the quadratic form is convex in each variable, and thus its maximum is attained at the boundary [14].

An immediate consequence of this observation is that if the diagonal entries of \( A \) are non-positive, then the integrity gap \( \text{GAP}_A \) is equal to

\[
\text{GAP}_A = \max_{\|x\|_2 \leq 1} \sum_{r,s} a_{rs} \langle x_r, x_s \rangle
\]

(3)

It will be often more convenient for us to search for an approximate solution among \( x_r \in [-1, 1] \). Then it can sequentially round each \( x_r \) to -1 or 1 without decreasing the value of the objective function [15]. There were consider a slightly modified SDP relaxation if the diagonal entries of \( A \) are non-negative:

\[
\max_{\|x\|_2 \leq 1} \sum_{r,s} a_{rs} \langle z_r, z_s \rangle
\]

(4)

2.5. Generalized Voronoi Diagrams (GVD)

Voronoi is the division of a flat surface into regions or groups based on the distance to the points contained in a subset in the flat surface region. A collection of points includes seeds or scattered points, locations, or generator points that were determined initially, and for each seed in the appropriate area will contain the same color because it is closer to the initial seed than other grain. The region formed is called the Voronoi cell.

Let \( Z = \{z_1, \ldots, z_n\} \) be a set of points of \( \mathbb{R}^d \). To each \( z_i \), there was associate its Voronoi region \( V(z_i) \) with

\[
V(z_i) = \{ x \in \mathbb{R}^d : \|x - z_i\| \leq \|x - z_j\|, \forall j \leq n \}.
\]

(5)

Region \( V(z_i) \) is an intersection of \( n - 1 \) with a size of half-space. In this region, each half of the space contains \( z_i \) enclosed by the bisector of \( z_i \) and several other \( Z \) points. The bisector of this process is hyperplanes, \( V(z_i) \) is a convex polyhedron, so it may not be bound as shown in figures 2 and 3 [16].

![Figure 2](image2.png)

**Figure 2.** The Voronoi diagram contains 9 points.

![Figure 3](image3.png)

**Figure 3.** The Delaunay triangulation of a set of 9 points.

2.6. Sub Markov Random Walker Algorithm

It gave a weighted graph \( G \), where one set of nodes has the label \( V_M \), and another set of nodes not labeled \( V_U \), where \( V_U \cup V_M = V \), the division of multilabel image object areas can be formulated as labelling [17]. Each node \( v_i \in V \) assigned by adding a label from the set. Problem-solving is done by comparing the probability of \( r_{ik}^l \) on each node that has the label \( l_i \) in the algorithm [18]. Before the system calculates the probability value, then the subMarkov transition probability value \( p \) on \( V \) is defined as follows:

\[
\sum_{j \sim i} p(i,j) \leq 1
\]

(6)
SubMarkov transition probabilities in which there is a property value that is by adding an auxiliary node \( \Delta \), then the subMarkov transition probability value \( p \) on \( G \) can be made into a transition probability value (Markov) on \( V \cup \{ \Delta \} \) by specifying \( p(\Delta,\Delta) = 1 \) and \( p(i,\Delta) = 1 - \sum_{j \neq i} p(i,j) \) [19]. Random walkers will leave graph \( G \) randomly by looking at the Probability Value \( p(i,\Delta) \) [20].

3. Research Method

The method used in this research is shown in Figure 4 and 5.

![Figure 4. Graph Quadratic form.](image1)

![Figure 5. Random Walker.](image2)

This research began with reading image data. The image used here has a size of 512x512 pixels, 72 dpi, with a bit depth of 24. Then the generalized Voronoi diagram (GVD) is calculated. The results used to represent the graph so that it forms figures from GVD pixels. The final step is to apply the quadratic form to the diagram to arrange the segmentation. Then the segmentation is done using a sub-Markov random walker, as shown in figure 5. The results used as a comparison of the results of the graph-based segmentation process with ground truth made both manually and with Voronoi diagrams.

4. Result and discussion

The output of this research is the division of the area or segmentation of an object image, which done using a graph — shown in Figure 6,7,8. Figure 6 shows the original vision of the bacterial image, Figure 7 shows the regional division carried out by the Voronoi Graph Diagram (GVD), and Figure 7 shows a graph formed and surrounding the bacteria as a barrier and calculating the distance to the center of the object. Figure 8 shows the Quadratic Shape used to help get the closest position to the object sought. It is equipped with labeling in the form of numbers to find out the number of nodes connected to bacterial objects in the image so that the segmentation process can prioritize particular objects.
After determining the nearest path using Voronoi Graph Diagram to get the results of segmentation, a random walker algorithm used to approach the object. It showed in figures 9, 10, and 11. The next stage used a random walker, a reference to ground truth obtained from medical analysts. Ground truth is guidance to get the object sought and non-object. Figure 9 shows an object image that is given ground truth by giving a blue color to the picture instead of the object and giving the purpose the color red. This step helps to search for objects with the color approach of the image object. Figure 10 shows the results where the image object turns red, and the non-object turns blue and is ready to be segmented. Figure 11 shows the results of segmentation done on the object image while still taking the red color on the image object and removing blue or non-bacterial objects. The next step is segmenting the area of the bacterial object sought with the help of the distance approach.

The segmentation process, as a minimization of the quadratic energy function. In the formed graph requires a unique minimizer, using an iterative scheme. The Random walker assumes an image consisting of \( k \) objects. Each point labelled \( v_m \) belongs to one of these \( k \) objects. It is considering the weight of the edge where the endpoint or the final form of segmentation is \( v_i \) and \( v_j \), where \( v_i \) is the original purpose and \( v_j \) is the target object, for example, \( e = (v_i, v_j) \). The weight of the edge \( \omega(e) \), which is in the range \((0,1)\), the Random walker runs from one point to the next using a probability value \( p \). If the \( v_i \) and \( v_j \) objects are similar in color or intensity, then the shape changes from \( v_i \) to \( v_j \). The shape changes and does not change when it is very different. For each vertex \( v_i \), according to the \( k \) object segmentation if \( p^k_k > p_{k'}^{k'} \) with all \( k' \neq k \). Low transition probability values determined through the edge in the image, where the position is opposite to the side in the graph. Rapid changes in color or intensity depend on this probability value.

Table 1 shows the calculation of segmentation accuracy. Accuracy results obtained by comparing the outputs of the segmentation process with ground truth. An object that has been validated by a laboratory assistant. Fifty 512x512 pixel image samples, 72 dpi resolution and 24-bit depth, trained and compared with four image samples that marked as a reference or ground truth, and the results show the average accuracy of segmentation results is 89%. 

Table 1
Table 1. Calculation of accuracy

|                          | TP (True positive) | TN (True negative) | FP (False positive) | FN (False negative) | Accuracy | Sensitivity | Specificity | Precision Rate |
|--------------------------|-------------------|--------------------|---------------------|--------------------|----------|-------------|-------------|----------------|
| Klebsiella pneumonia     | 18                | 2                  | 1                   | 1                  | 0.91     | 0.95        | 0.67        | 0.95           |
| Acinetobacter Baumanii   | 18                | 2                  | 2                   | 1                  | 0.87     | 0.95        | 0.50        | 0.90           |
| Pseudomonas aerugenusa   | 26                | 5                  | 2                   | 2                  | 0.89     | 0.93        | 0.71        | 0.93           |
| Eschericia Coli          | 18                | 2                  | 1                   | 1                  | 0.91     | 0.95        | 0.67        | 0.95           |

Average: 0.89 0.94 0.64 0.93

A comparison of the advantages and disadvantages of the graph-based method approach compared with research on previous segmentation. The reference used is the range of accuracy of the segmentation produced from each method, although during the last case, it did not use bacterial image objects. The aim is to get GAP and research features to develop. This comparison saw in table 2.

Table 2. Comparison Between Method

| No | Segmentation          | Technique | Method Advantages |
|----|-----------------------|-----------|-------------------|
| 1  | Histogram thresholding| Each color object in the image has a peak, and each represents a specific region | Works with a low complexity value, does not require prior information about the object image |
|    |                       |           | Does not consider the spatial details of an image and Not too good for images that do not have peaks and unclear image details |
|    |                       |           | 87.5%-92%         | Image Segmentation by using Histogram Thresholding. P Daniel Ratna,[21] January 2012 |
| 2  | K-means clustering    | Work-based on the assumption. Its region has specific clusters and corresponding to segments. | Easy implementation and Supervised learning |
|    |                       |           | Problems in determining the number of groups and Segmented areas are not always areas of concern. |
|    |                       |           | 90.83%-93.1%  | Implementation of the K-Means Algorithm as Image Segmenting Method in Identifying the Citrus Leaves Disease. F G, 2018[22] |
| 3  | Edge detection        | Based on discontinuities in the image, and regional segments based on sudden changes in pixel values. | Works well for images that have good contrast. |
|    |                       |           | It does not provide efficient results in images with too many edges, and it is not easy to produce closed boundaries and curves. |
|    |                       |           | 84.72%-96%     | A Combinational Clustering Based Method for cDNA Microarray Image Segmentation Guifang Shao, 2015[23] |
| 4  | Watershed             | It segmented the region based on contour separating different intensity regions. | Simple and fast, Gives better result in |
|    |                       |           | Leads to over segmentation, Sensitive to noise. |
|    |                       |           | 85.25%-90.25% | Detection of Oral Tumor based on Marker Controlled Watershed Algorithm, 2012[24], Anuradha |
Table 2 of the method comparison shows that a graph-based approach considered for the segmentation process. In general, the accuracy value generated also depends on the pre-processing process carried out on the observed object. But sometimes the amount of accuracy is also not a consideration when the processing time needed is the speed in terms of service. Object identification becomes more critical with acceptable accuracy results.

5. Conclusion
The main goal of the current research was to determine that area-based segmentation done using the Graph approach. Bacterial colonies counted and labeled based on the closest distance from each region. The result is that bacterial objects can be searched based on the closest average distance from the node in the Voronoi diagram. The significance of the results obtained that the addition of a random walker can help complete the segmentation process carried out by the Voronoi layout. The results of segmentation show that the suitability of objects compared to ground truth is an average of 89%. The open problem faced is the form of a shifted object or translated image, which may be solved by an invariant moment or Scale Feature Invariant Feature Transform approach.

Acknowledgment
Thank you to the Dr. Soetomo Hospital microbiology lab team in Surabaya for supporting primary data and promoters and co-promoters who have allowed this publication.

References
[1] Sharika R, Subbaiah P, and Balamurugan K 2018 Studies on reproductive stress caused by candidate Gram positive and Gram negative bacteria using model organism Caenorhabditis elegans *Gene* 649 113-126
[2] Farzaneh Dehghan N Z, Vahid Boostani A S, Eghbal T 2016 Resistance of gram negative bacteria in hospital acquired pneumonia: a prospective study *The Brazilian Journal of Infectious Diseases* 20 113-114
[3] Mamou G, Ganesh B, Mali M, et al 2016 Early developmental program shapes colony morphology in bacteria *Cell Reports* 14 1850-1857
[4] Hiremath P S 2013 An Improved Automated Method for Identification of Bacterial Cell Morphological Characteristics *International Journal of Advanced Trends in Computer Science and Engineering* 2(1) 11-16
[5] Aurenhammer F, Su B, et al 2014 A note on visibility-constrained Voronoi diagrams *Discrete Applied Mathematics* 174 52-56
[6] Miandarhoie A R and Khalili K 2015 iterative mesh segmentation using approximated voronoi diagram *Procedia Technology* 19 106-111
[7] Miller S I and Salama N R 2018 The gram-negative bacterial periplasm: Size matters PLoS biology 16 e2004935-e2004935
[8] Maria-Neto S, de Almeida K C, et al 2015 Understanding bacterial resistance to antimicrobial peptides: From the surface to deep inside Biochimica et Biophysica Acta (BBA) - Biomembranes 1848 3078-3088
[9] Wamiq M, Ahmed B B, Arun K B, Daniel Hirlme Dan, Paul Robinson J, and Bartek Rajwa 2013 Classification of Bacterial Contamination Using Image Processing and Distributed Computing IEEE Journal of Biomedical and Health Informatics
[10] Goswami M, Maheshwari S, and Poonia A 2017 Combinational feature approach: Performance improvement for image processing based leaf disease classification 4th International Conference on Signal Processing, Computing and Control (ISPCC) 568-574
[11] Dongdong Y, Rong F, et al 2015 SAR image segmentation by selected principal components and kernel graph cuts ensembles 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) 1976-1980
[12] Wang D, Li H, et al 2017 An efficient iterative thresholding method for image segmentation Journal of Computational Physics 350 657-667
[13] Gerstein L J 2020 Integral quadratic forms and graphs Linear Algebra and its Applications 585 60-70
[14] Singaraju D, Grady L, and Vidal R 2008 Interactive image segmentation via minimization of quadratic energies on directed graphs IEEE Conference on Computer Vision and Pattern Recognition 1-8
[15] Albers C J, Critchley F, and Gower J C 2011 Quadratic minimisation problems in statistics Journal of Multivariate Analysis 102 698-713
[16] Chadha A and Andreopoulos Y 2017 Voronoi-based compact image descriptors: efficient region of interest retrieval with vlad and deep learning based descriptors IEEE Transactions on Multimedia 19 1596-1608
[17] Jha S K, Bannerjee P, and Banik S 2013 Random walks based image segmentation using color space graphs Procedia Technology 10 271-278
[18] Kanas V G, Zacharaki E I, et al 2015 A low cost approach for brain tumor segmentation based on intensity modeling and 3D Random Walker Biomedical Signal Processing and Control 22 19-30
[19] Dong X, Shen J, et al 2016 sub-Markov random walk for image segmentation IEEE Transactions on Image Processing 25 516-527
[20] Faragallah O S, Abdel-Aziz G, and Kelash H M 2017 Efficient cardiac segmentation using random walk with pre-computation and intensity prior model Applied Soft Computing 61 427-446
[21] Daniel P, Raju R, and Neelima G 2012 Image segmentation by using histogram thresholding Int. J. Comp. Sci. Eng. Technol. 2
[22] Febrinanto F G, Dewi C, and Triwidarto A 2019 The implementation of k-means algorithm as image segmenting method in identifying the citrus leaves disease IOP Conference Series: Earth and Environmental Science 243 012024
[23] Shao G, Li T, et al 2015 A combinational clustering based method for CDNA microarray image segmentation PLOS ONE 10 e0133025
[24] Anuradha and Sankaranarayanan 2012 Detection of oral tumor based on marker controlled watershed algorithm International Journal of Computer Applications 52 15-18
[25] Chekkoury A, Khurd P, et al 2012 Automated malignancy detection in breast histopathological images Conference: SPIE Medical Imaging 8315