

The Classification of Acute Respiratory Infection (ARI) Bacteria Based on K-Nearest Neighbor

by Zilvanhisna Emka Fitri

Submission date: 08-Jul-2022 06:49PM (UTC+0700)

Submission ID: 1868065776

File name: Lontar_Komputer_2021.pdf (637.89K)

Word count: 4883

Character count: 25686

The Classification of Acute Respiratory Infection (ARI) Bacteria Based on K-Nearest Neighbor

Zilvanhisna Emka Fitri^{a1}, Lalitya Nindita Sahenda^{a2}, Pramuditha Shinta Dewi Puspitasari^{a3}, Prawidya Destarianto^{a4}, Dyah Laksito Rukmi^{b5}, Arizal Mujibtamala Nanda Imron^{c6}

^aDepartment of Information Technology, Politeknik Negeri Jember

^bDepartment of Animal Science, Politeknik Negeri Jember

Jl. Mastrip PO.BOX 164 Jember, 68121, Indonesia

¹zilvanhisnaef@polije.ac.id (Corresponding author)

²lalitya.ns@polije.ac.id

³pramuditha@polije.ac.id

⁴prawidya@polije.ac.id

⁵dyah.laksito@polije.ac.id

^cDepartement of Electrical Engineering, Universitas Jember

Jl. Kalimantan No. 37, Kampus Tegalboto, Jember, 68121, Indonesia

⁶arizal.tamala@unej.ac.id

Abstract

Acute Respiratory Infection (ARI) is an infectious disease. One of the performance indicators of infectious disease control and handling programs is disease discovery. However, the problem that often occurs is the limited number of medical analysts, the number of patients, and the experience of medical analysts in identifying bacterial processes so that the examination is relatively longer. Based on these problems, an automatic and accurate classification system of bacteria that causes Acute Respiratory Infection (ARI) was created. The research process is preprocessing images (color conversion and contrast stretching), segmentation, feature extraction, and KNN classification. The parameters used are bacterial count, area, perimeter, and shape factor. The best training data and test data comparison is 90%: 10% of 480 data. The KNN classification method is very good for classifying bacteria. The highest level of accuracy is 91.67%, precision is 92.4%, and recall is 91.7% with three variations of K values, namely K = 3, K = 5, and K = 7.

Keywords: Bacteria, Acute Respiratory Infection, Image Processing, KNN

1. Introduction

Acute Respiratory Infections (ARI) are included in the list of the top ten infectious diseases whose incidence of infectious diseases (disease prevalence) and mortality (a measure of the number of deaths in a population) are quite high in the world [1]. ARI is divided into two, namely upper respiratory tract infections (URTIs) and lower respiratory tract infections (LRTIs). The upper respiratory tract consists of the ears, nose, and throat, while the lower respiratory tract consists of the trachea, bronchi, bronchioles, and lungs [2]. Some examples of ARI diseases caused by bacteria are pneumonia, tuberculosis (TB), diphtheria, and pharyngitis [3].

Pneumonia is an infectious disease caused by an infection that causes the lungs to become inflamed. The causative pathogens (bacteria) are *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Haemophilus influenza*, *Mycoplasma pneumonia*, *Chlamydomphila pneumonia*, and *Legionella pneumophila* [4]. Tuberculosis (TB) is one of the serious health problems in Indonesia. TB is an infection caused by *Mycobacterium tuberculosis* in the lower respiratory tract. Diphtheria is an acute infectious disease caused by *Corynebacterium diphtheriae* which attacks the upper respiratory tract [2]. From year to year in East Java, the number of diphtheria sufferers is reported to continue to increase until, in 2019, there were 358 cases [5]. In addition, *Neisseria gonorrhoeae* is a bacterial pathogen that causes pharyngitis [4], which usually occurs in sexually transmitted diseases (STD) without symptoms (asymptomatic) [3].

Achievement of performance indicators of infectious disease control and handling programs, namely discovery, treatment, and success of treatment [5]. Generally, the discovery process is carried out by examining specimens or sputum from the patient, which is then carried out by a microscopic examination process. However, the problems that often occur are the limited number of medical analysts, a large number of patients, differences in perceptions and experiences of medical analysts in identifying bacteria in sputum/throat sputum samples, and the time required for the examination process is relatively longer. Based on the description of the problem above, the researchers created an automatic and accurate bacterial classification system for the early detection of acute respiratory infections (ARI).

Several references are used as references by researchers regarding the identification of bacteria that cause pneumonia and tuberculosis. In 2016, a *Streptococcus pneumoniae* detection system was created from digital microscope images with an accuracy rate of 80% [6]. Then the bacterial segmentation was developed using the Channel Area Thresholding (CAT) segmentation method so that the system was able to identify bacilli with an accuracy of 97.58% on the sputum image dataset [8]. Meanwhile, the identification of Mycobacterium tuberculosis bacteria has also been carried out using image segmentation and the K-Means clustering method in 2015 [7]. The following research compares two classification methods: backpropagation and K-Nearest Neighbor (KNN), to obtain an accuracy rate of 93.22% for backpropagation and 94.92% for KNN [9].

Based on the references above, the researcher uses the K-Nearest Neighbor (KNN) method. The KNN method is a general and straightforward classification method used, but this research is an early stage of research on ARI bacterial classification, so we focus on selecting the right features to classify ARI bacteria. There is a difference with previous research, namely the type of bacteria studied. In this research, researchers added *Staphylococcus aureus* and *Streptococcus pneumoniae* as bacteria for pneumonia disease, *Corynebacterium diphtheriae* as bacteria for diphtheria disease, and *Neisseria gonorrhoeae* as pathogens for pharyngitis disease.

2. Research Methods

This study uses the personal data of the researcher, namely the bacterial image dataset from throat sputum. Several stages were carried out in this research, namely bacterial image, image preprocessing, image segmentation, feature extraction, and bacterial classification using the KNN method, as shown in Figure 1.

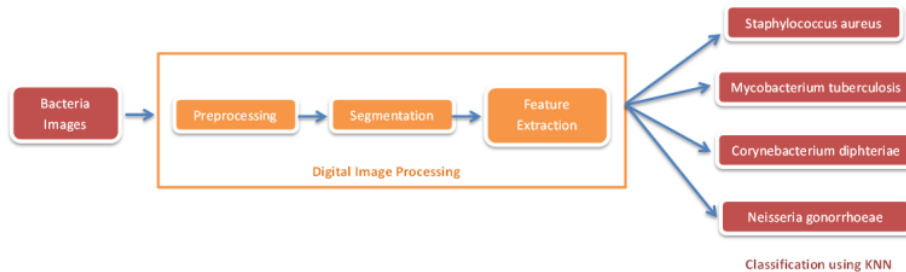


Figure 1. Block Diagram of The Bacterial Classification System Proposed

2.1. Bacteria Images

Generally, the size of bacteria is 0.4 to 2 μ m, consisting of three general forms, namely cocci, bacilli, and spirochetes [4]. The three forms have other specific forms such as *Staphylococcus aureus*, which is included in the cocci in a cluster group, *Streptococcus pneumoniae* is included in the cocci in chains group, *Corynebacterium diphtheriae* is included in the clubshaped and pleomorphic rods group, and *Neisseria gonorrhoeae* is included in the diplococci group [10].

Meanwhile, *Mycobacterium tuberculosis* belongs to the aerobic acid-fast rods group [11], as shown in Figure 2, and based on the literature. The research image data is shown in Table 1.

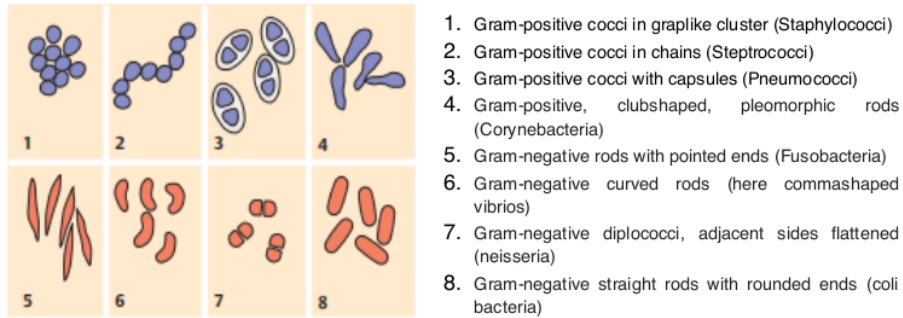


Figure 2. Bacterial Morphology [10].

Table 1. Variation of Acute Respiratory Infection Bacterial Image

Bacterial Name	Disease	Bacterial Images	
<i>Staphylococcus aureus</i>	Pneumonia		
<i>Streptococcus pneumoniae</i>	Pneumonia		
<i>Corynebacterium diphtheriae</i>	Diphtheria		
<i>Neisseria gonorrhoeae</i>	Pharyngitis		
<i>Mycobacterium tuberculosis</i>	TB		

Table 1 shows that the research data consisted of 5 classes, namely *Staphylococcus aureus* and *Streptococcus pneumoniae* as pneumonia disease bacteria, *Corynebacterium diphtheriae* as diphtheria disease bacteria, *Neisseria gonorrhoeae* as asymptomatic pharyngitis bacteria, and *Mycobacterium tuberculosis* as tuberculosis (TB) bacteria.

2.2. Preprocessing Images

The data normalization process is carried out at this research stage, such as uniformity of image size and uniformity of color space used before the image segmentation process. Initially, the size of the bacterial image varied from 1920x1080 pixels, but the size was very large, and it was necessary to cut the image to 151x151 pixels, as shown in Figure 3. The result of the cropping process is part of the normalization of data that represents the shape of the ARI bacteria. In addition, the cropping process aims to reduce the computational load [12].

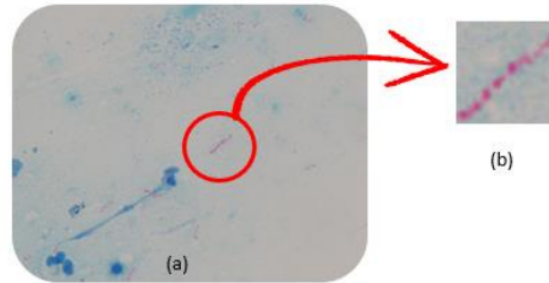


Figure 3. Image Size (a) 1920x1080 pixels to (b) 151x151 pixels

The cropped image is an RGB color space image where the color space consists of 3 color components, namely red components, green components, and blue components. RGB color space has a large size, so it isn't easy to segment, so it needs to be converted to another color space [13], for example, HSV color space. The HSV color space is a color space that also consists of 3 color components, namely the Hue color component, the Saturation color component, and the Value color component. The process of converting color from RGB color space to HSV color space with the formula equation [14] :

$$hue = \tan \left(\frac{3x(G - B)}{(R - G) + (R - B)} \right) \quad (1)$$

$$saturation = 1 - \frac{\min(R, G, B)}{V} \quad (2)$$

$$Value = \frac{R + G + B}{3} \quad (3)$$

Next is the process of adding contrast (contrast stretching). Its function is to even out the distribution of light and dark intensity over the entire intensity scale so that the image has a high contrast value.

2.3. Segmentation

At this stage, the aim is to separate the research object from the background. This stage uses a threshold process where we have to find the threshold value with formula equation [15] :

$$segmentation(x, y) = \begin{cases} 1, & \text{if } grayscale(x, y) \leq T \\ 0, & \text{if } grayscale(x, y) > T \end{cases} \quad (4)$$

To find the threshold value (T), we have to look at the histogram of the grayscale image to find out the gray-level value of the research object and the background. In addition to using the thresholding technique, the segmentation process is also carried out using the chain-code technique. This method uses a labeling system for each binary object. It then calculates the proximity of the pixel values based on the direction of 4 or 8 surrounding neighbors, as shown in Figure 4.

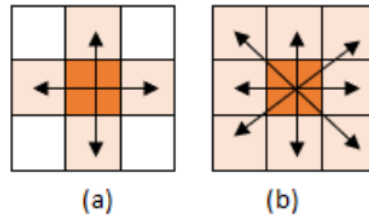


Figure 4. (a) 4-connected and (b) 8-connected

2.4. Feature Extraction

At this stage, the aim is to find characteristic values that can distinguish the first class from other classes. Feature extraction used in this research is morphological or shape features such as Bacterial count, area, perimeter, and form factor. Determination of the area and perimeter using a chain code, where area (A) represents the area of the bacteria, the perimeter or circumference (P) represents the edge length, and the shape factor (S) represents the shape of the bacteria. The three parameters are expressed by the equation formula [16] :

$$A = \text{Number of pixels in row} - 1 + \text{row to} - 2 + \dots + \text{row to} - 8 \quad (5)$$

$$P = \sum \text{Even code} + \sqrt{2} \sum \text{odd code} \quad (6)$$

$$S = \frac{P^2}{A} \quad (7)$$

2.5. K-Nearest Neighbor Classification

K-Nearest Neighbor (KNN) classification is one of the classification methods with supervised learning methods. In supervised learning, the classification target is known. The KNN method uses the closest distance to the object to classify data, so that the method is often known as lazy learning. The basic principle of KNN is to find the value of K where the value of K is the closest amount of data that will determine the classification results and to calculate the closest distance using Euclidean distance (ED) calculations with the equation formula [16]–[18]:

$$ED(x_i, x_j) = \sqrt{\sum_{r=1}^n (x_{ir} - x_{jr})^2} \quad (8)$$

Where X_{ir} is the testing data and X_{ij} is the training data

The total number of data is 481 images, consisting of 94 images of *Corynebacterium diphtheriae* bacteria, 91 images of *Mycobacterium tuberculosis*, and 95 images of *Neisseria gonorrhoeae* 92 images of *Staphylococcus aureus*, and 108 images of *Streptococcus pneumoniae* bacteria. In this research, the classification process is to find the highest level of accuracy from the KNN method in comparing training data and testing data. The comparison of the data carried out is 50% : 50%, 60% : 40%, 70% : 30%, 80% : 20% and 90% : 10%.

3. Result and Discussion

In the research of bacterial images, which were originally in the RGB color space, they were converted into HSV color spaces using the equations (1), (2), and (3) so that the HSV color space channel that best represented the shape of the bacteria was shown in Figure 5. The figure shows that the Hue component image best represents the shape of the bacteria *Streptococcus pneumoniae*, *Corynebacterium diphtheriae*, and *Mycobacterium tuberculosis*. Meanwhile, *Staphylococcus aureus* and *Neisseria gonorrhoeae* bacteria can be represented well on the

image of the saturation component. To clarify the shape of the bacteria, the following process is contrast stretching which causes the image to have a high contrast value so that it also affects the histogram of the image. In addition, there is a change in the image before and after the contrast stretching process, as shown in Figure 6.

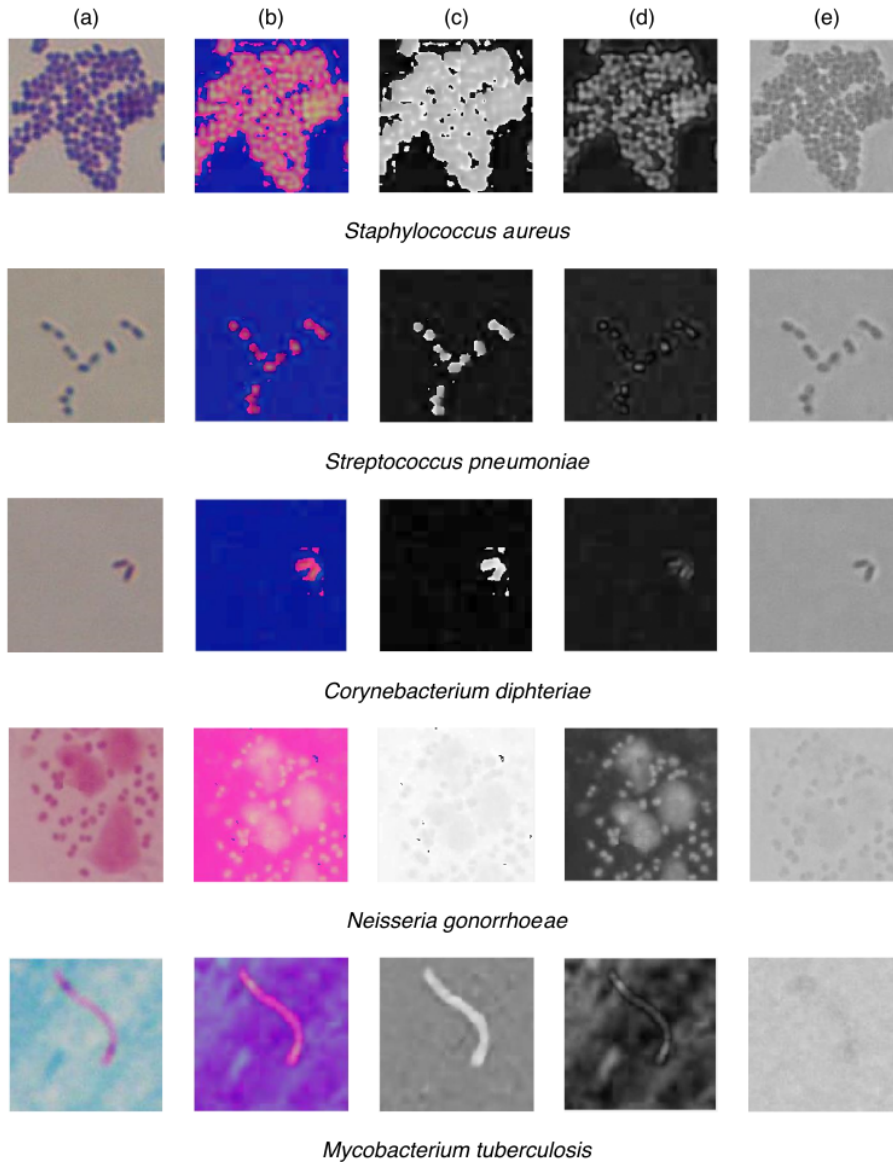


Figure 5. Image of (a) RGB, (b) HSV, (c) Hue, (d) Saturation and (e) Value on various types of bacteria

Figure 6 shows a difference between the Hue image histogram before and after the contrast stretching process. The range of gray values of the HSV image is 0 - 1. This is certainly different

from the range of gray values of the grayscale image, which is 0 - 255. In the image before contrast stretching, there are two peaks in the histogram, namely 0.58 and 0.78, while after doing the contrast stretching, there are two peaks in the histogram contrast stretching, the distribution of light and dark intensity throughout the intensity scale so that the image histogram looks bigger than before. In addition to changes in the histogram, Figure 6 also shows changes in the hue image before and after contrast stretching. The contrast stretching process helps the process of segmentation because the image of (a) the value of gray level is similar between the object and the background, while the image (b) occur significant color difference between the object and the background that will ease the process of segmentation using a threshold.

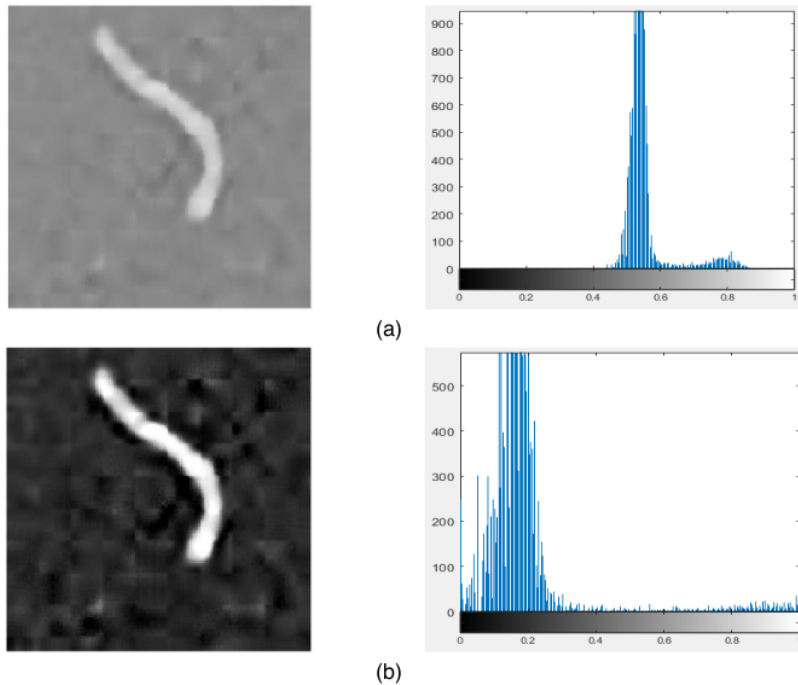


Figure 6. (a) Hue image and histogram before contrast stretching while (b) Hue image and histogram after contrast stretching on Mycobacterium tuberculosis bacteria image

After the contrast stretching process, the segmentation process is carried out based on the threshold value with the equation (4). Because this study used hue images and saturation images, the threshold value used has a range of 0.4 to 0.7. It depends on the results of the contrast stretching the image, whether it is dark or light. The thresholding process results are a binary image, an image with two values, namely 0 (black) and 1 (white), as shown in Figure 8.

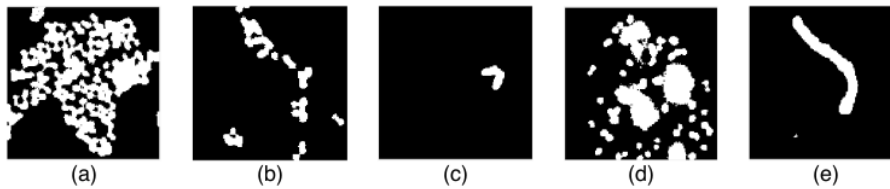


Figure 7. Thresholding image of (a) Staphylococcus aureus, (b) Streptococcus pneumoniae, (c) Corynebacterium diphtheriae, (d) Neisseria gonorrhoeae and (e) Mycobacterium tuberculosis

Figure 7 shows that the threshold image can represent most forms of bacteria, but there are some bacteria such as *Neisseria gonorrhoeae* and *Mycobacterium tuberculosis* that need to be re-segmented. This is because there is still noise in the segmentation image based on the threshold value. Noise is meant objects that are not parts of the bacterial body, such as paint residues and other objects like Polymorphonuclear (PMN) cells. PMN itself is one of the white blood cells that will appear if there is an infection in the body. In the image of *Neisseria gonorrhoeae* bacteria, the shape of polymorphonuclear cells (PMN) is also segmented, so it is necessary to do segmentation based on area. To perform the segmentation, the process is continued by labeling the object and finding the area value using a chain code with the proximity of 8 neighboring pixels. This process is known as the Channel Area Thresholding (CAT) segmentation technique [19].

In this research, all bacterial images were segmented based on the Channel Area Threshold (CAT) value, but the range of the threshold values differed depending on each bacterium's area. The threshold value for the CAT segmentation technique is denoted [S.Area]. The value of the threshold area [S. Area] used varies depending on the area of each bacterium. Determination of S.Area is to use two threshold values, namely $[S. Area] \geq 5$ & $[S. Area] \leq 100$ was used to remove other objects such as PMN in *Neisseria gonorrhoeae*, while *Streptococcus pneumoniae*, *Corynebacterium diphtheriae*, and *Mycobacterium tuberculosis*, two threshold values were $[S. Area] \geq 50$ & $[S. Area] \leq 7000$. Then in *Staphylococcus aureus*, the difference in the upper threshold value $[S. Area] \leq 10000$. The results of the CAT segmentation are shown in Figure 8.

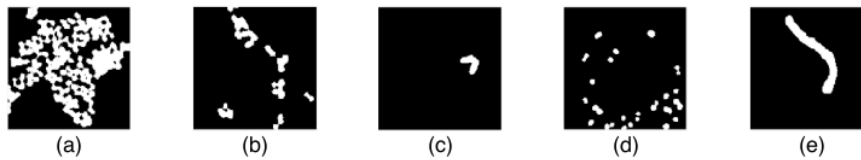


Figure 8. CAT segmentation images (a) *Staphylococcus aureus*, (b) *Streptococcus pneumoniae*, (c) *Corynebacterium diphtheriae*, (d) *Neisseria gonorrhoeae* and (e) *Mycobacterium tuberculosis*

Figure 8 shows the results of the segmentation image where there are only bacterial objects, without any noise such as staining or other cells (PMN) in the bacterial image (d) *Neisseria gonorrhoeae*. The following process is feature extraction based on the shape (morphology) of bacteria. Morphological features are used to classify bacteria that cause ARI because the characteristics of the shape of the bacteria are in accordance with Figure 2, so the features used are bacterial count, area, perimeter, and Shape factor. The results of bacterial feature extraction are shown in Table 2.

Table 2. Feature extraction on each type of bacteria

Feature		<i>Staphylococcus aureus</i>	<i>Streptococcus pneumoniae</i>	<i>Corynebacterium diphtheriae</i>	<i>Neisseria gonorrhoeae</i>	<i>Mycobacterium tuberculosis</i>
Bacterial count	Minimum	1	1	1	3	1
	Maximum	8	12	5	29	1
	Average	3	6	2	12	1
Area	Minimum	1555	832	208	414	613
	Maximum	9984	4679	1580	3949	2465
	Average	4639	2438	728	1502	1332
Perimeter	Minimum	586	384	101	263	257
	Maximum	4009	1993	696	2495	877
	Average	1774	1088	292	1021	487
Shape factor	Minimum	220,833	154,793	25,584	167,075	81,061
	Maximum	1697,013	958,275	311,889	1750,571	351,729
	Average	687,847	491,619	123,410	700,310	181,752

Table 2 shows that the area, perimeter, and shape factor of the bacteria with the largest value is *Staphylococcus aureus*, while the smallest is *Corynebacterium diphtheriae*. The highest number of bacteria is *Neisseria gonorrhoeae*, as many as 29 bacteria in one image. In comparison, the least number of bacteria is *Mycobacterium tuberculosis*, as many as one bacteria in one image.

These features will be the input of the K-Nearest Neighbor (KNN) classification method. The basic principle of KNN is to find the value of K where the value of K is the closest amount of data that will determine the classification results and calculate the closest distance using the Euclidean distance calculation using the equation (8). The learning process of the KNN method is supervised learning, where the target is known beforehand. When testing the test data (unknown class label), the KNN algorithm looks for the training data closest to the test data. The test data is classified according to the class from the training data with the closest Euclidean distance. This study uses 480 data which is divided into training data and testing data with the provisions of 50% : 50%, 60% : 40%, 70% : 30%, 80% : 20% and 90% : 10% with variations in K values, the results accuracy, precision, recall are shown in Table 3.

Table 3. The results of accuracy, precision, and recall on variations in the value of K

Training Data : Testing Data	K. value	Accuracy (%)	Precision (%)	Recall (%)
50% : 50%	1	87,5	87,9	87,5
	3	85	85,6	85
	5	87,5	87,8	87,5
	7	86,67	87	86,7
	9	86,25	86,5	86,3
60% : 40%	11	85,42	85,7	85,4
	1	86,46	87,1	86,5
	3	88,02	88,6	88
	5	88,54	88,8	88,5
	7	85,94	86,3	85,9
70% : 30%	9	85,94	86,1	85,9
	11	85,94	86,3	85,9
	1	86,81	87,5	86,8
	3	88,19	88,7	88,2
	5	90,28	90,5	90,3
80% : 20%	7	88,19	88,6	88,2
	9	86,81	86,9	86,7
	11	86,81	86,9	86,8
	1	84,38	85,7	84,4
	3	85,42	86,2	85,4
90% : 10%	5	88,54	89,4	88,5
	7	88,54	89	88,5
	9	90,63	91,6	90,6
	11	89,58	90,3	89,6
	1	87,5	89,8	87,5
	3	91,67	92,4	91,7
	5	91,67	92,4	91,7
	7	91,67	92,4	91,7
	9	89,58	90,1	89,6
	11	89,58	90,1	89,6

Table 3 contains the comparison of training data and test data used with variations in the value of K to produce the best level of accuracy, precision, and recall. In comparing data 50%: 50%, the best accuracy rate is 87.5%, with a K = 1. Comparison of data 60%: 40%, the best accuracy rate is 88.54% with a K = 5. Comparison of data 70%: 30 % the best accuracy rate is 90.28% with a Value of K = 5. Comparison of data 80%: 20% the best accuracy rate is 90.63% with a Value of K = 9. This is different in the comparison of training data and test data 90%: 10 %, the best level of accuracy is 91.67%, precision is 92.4%, and recall is 91.7% with three variations in the value of K, namely K = 3, K = 5 and K = 7. To find out the results of the KNN classification, a confusion matrix was made, as shown in Table 4.

Table 4. Confusion Matrix with a data ratio of 90%: 10% at the value of $K = 7$

Output					Target
a	b	c	d	e	
10	0	0	0	0	a = <i>Corynebacterium diphtheriae</i>
1	8	0	0	0	b = <i>Mycobacterium tuberculosis</i>
0	0	4	0	1	c = <i>Neisseria gonorrhoeae</i>
0	1	0	12	0	d = <i>Staphylococcus aureus</i>
1	0	0	0	10	e = <i>Streptococcus pneumoniae</i>

Table 4 shows that as many as 10 data were correctly classified as *Corynebacterium diphtheriae*, while for *Mycobacterium tuberculosis*, 8 data were correctly classified, and 1 data was misclassified as *Corynebacterium diphtheriae*. 4 data were correctly classified as *Neisseria gonorrhoeae* and 1 data was misclassified as *Streptococcus pneumoniae*. The *Staphylococcus aureus* was classified correctly as many as 12 data, and 1 data was misclassified into *Mycobacterium tuberculosis*. *Streptococcus pneumoniae* were classified correctly as many as 10 data, and 1 data was misclassified into *Corynebacterium diphtheriae*. These results can occur due to the closeness of the values between the KNN input parameters (number of bacteria, area, perimeter, and shape factor) for each bacterium, as shown in Table 2. An example of the average perimeter feature is shown in Table 5 below.

Table 5. Perimeter feature values for each type of bacteria

Feature	<i>Staphylococcus aureus</i>	<i>Streptococcus pneumoniae</i>	<i>Corynebacterium diphtheriae</i>	<i>Neisseria gonorrhoeae</i>	<i>Mycobacterium tuberculosis</i>
Minimum	586	384	101	263	257
Maximum	4009	1993	696	2495	877
Average	1774	1088	292	1021	487

Table 5 shows a closeness of the average value of perimeter features between *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Neisseria gonorrhoeae* bacteria, namely 1774, 1088, and 1021. Of course, this proximity affects the classification results using the KNN method, causing misclassification between bacteria so that a confusion matrix is created and is shown in Table 4.

Suppose we compare with previous research where the accuracy of KNN is 94.92% while the accuracy of KNN in this research is 91.67%. This difference occurs because the previous research only classified one bacterium, namely *Mycobacterium tuberculosis*. Still, in this research, we added four other bacteria, namely *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Corynebacterium diphtheriae*, and *Neisseria gonorrhoeae*.

4. Conclusion

This research is one of the computer vision studies that aims to classify acute respiratory tract infection (ARI) bacteria using the K-Nearest Neighbor (KNN) method. The parameters used in this study are shape parameters, namely Bacterial count, area, perimeter, and form factor. The data used are 480 data with the best comparison of training data and test data, namely 90%: 10%. The KNN classification method can classify these bacteria with the highest level of accuracy, namely 91.67%, precision 92.4%, and recall 91.7% with 3 variations in the value of K , namely $K = 3$, $K = 5$ and $K = 7$. In this study, it is necessary to add other features and compare them with other classification methods to get the best classification method to classify bacteria that cause acute respiratory infections (ARI).

References

- [1] E. Setyowati and S. Mariani, "Penerapan Jaringan Syaraf Tiruan dengan Metode Learning Vector Quantization (LVQ) untuk Klasifikasi Penyakit Infeksi Saluran Pernapasan Akut (ISPA)," vol. 4, p. 10, 2021.
- [2] S. J. Pitt, *Clinical Microbiology for Diagnostic Laboratory Scientists*. Chichester, UK: John Wiley & Sons, Ltd, 2017. doi: 10.1002/9781118745847.
- [3] K. Struthers, *Clinical Microbiology*, Second Edi. New York: CRC Press, 2017.

- [4] C. R. Mahon and D. C. Lehman, *Textbook of Diagnostic Microbiology*, Sixth Edit. St. Louis, Missouri: Elsevier, 2019. doi: 10.1309/u0mb-0p7r-rrwf-4bth.
- [5] Dinas Kesehatan Provinsi Jawa Timur, *Profil Kesehatan Provinsi Jawa Timur 2019*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2020.
- [6] R. Yuliwardana, "Deteksi Bakteri Streptococcus pneumoniae Berbasis Jaringan Syaraf Tiruan dari Citra Mikroskop Digital," Universitas Airlangga, Surabaya, 2016.
- [7] R. Rulaningtyas, Andriyan Bayu Suksmono, T. Mengko, and P. Saptawati, "Multi patch approach in K-means clustering method for color image segmentation in pulmonary tuberculosis identification," in *2015 4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)*, Bandung, Indonesia, Nov. 2015, pp. 75–78. doi: 10.1109/ICICI-BME.2015.7401338.
- [8] K. S. Mithra and W. R. S. Emmanuel, "Segmentation of Mycobacterium Tuberculosis Bacterium From ZN Stained Microscopic Sputum Images," in *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, Dec. 2018, pp. 150–154. doi: 10.1109/ICSSIT.2018.8748294.
- [9] L. N. Sahenda, M. H. Purnomo, I. K. E. Purnama, and I. D. G. H. Wisana, "Comparison of Tuberculosis Bacteria Classification from Digital Image of Sputum Smears," in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM)*, Surabaya, Indonesia, Nov. 2018, pp. 20–24. doi: 10.1109/CENIM.2018.8711386.
- [10] F. H. Kayser, Ed., *Medical microbiology*. Stuttgart; New York, NY: Georg Thieme Verlag, 2005.
- [11] P. R. Murray, *Basic Medical Microbiology*. Philadelphia: Elsevier, 2018.
- [12] Z. E. Fitri, A. Baskara, M. Silvia, A. Madjid, and A. M. N. Imron, "Application of backpropagation method for quality sorting classification system on white dragon fruit (*Hylocereus undatus*)," *IOP Conference Series : Earth Environmental Science*, vol. 672, no. 1, p. 012085, Mar. 2021, doi: 10.1088/1755-1315/672/1/012085.
- [13] A. M. Nanda Imron and Z. E. Fitri, "A Classification of Platelets in Peripheral Blood Smear Image as an Early Detection of Myeloproliferative Syndrome Using Gray Level Co-Occurrence Matrix," *Journal of Physics: Conference Series*, vol. 1201, p. 012049, May 2019, doi: 10.1088/1742-6596/1201/1/012049.
- [14] Z. E. Fitri, U. Nuhanatika, A. Madjid, and A. M. N. Imron, "Penentuan Tingkat Kematangan Cabe Rawit (*Capsicum frutescens* L.) Berdasarkan Gray Level Co-Occurrence Matrix," *Jurnal Teknologi Informasi Dan Terapan (JTIT)*, vol. 7, no. 1, pp. 1–5, Jun. 2020, doi: 10.25047/jtit.v7i1.121.
- [15] Z. E. Fitri, R. Rizkiyah, A. Madjid, and A. M. N. Imron, "Penerapan Neural Network untuk Klasifikasi Kerusakan Mutu Tomat," *Jurnal Rekayasa ElektriKa*, vol. 16, no. 1, May 2020, doi: 10.17529/jre.v16i1.15535.
- [16] Z. E. Fitri, L. N. Y. Syahputri, and M. N. Imron, "Classification of White Blood Cell Abnormalities for Early Detection of Myeloproliferative Neoplasms Syndrome Based on K-Nearest Neighbor," *Scientific Journal of Informatics*, vol. 7, no. 1, p. 7, 2020.
- [17] I. M. A. S. Widiatmika, I. N. Piarsa, and A. F. Syafiandini, "Recognition of The Baby Footprint Characteristics Using Wavelet Method and K-Nearest Neighbor (K-NN)," *Lontar Komputer Jurnal Ilmiah Teknologi Informasi*, vol. 12, no. 1, p. 41, Mar. 2021, doi: 10.24843/LKJITI.2021.v12.i01.p05.
- [18] R. J. Al Kautsar, F. Utaminingrum, and A. S. Budi, "Helmet Monitoring System using Hough Circle and HOG based on KNN," *Lontar Komputer Jurnal Ilmiah Teknologi Informasi*, vol. 12, no. 1, p. 13, Mar. 2021, doi: 10.24843/LKJITI.2021.v12.i01.p02.

The Classification of Acute Respiratory Infection (ARI) Bacteria Based on K-Nearest Neighbor

ORIGINALITY REPORT

11 %
SIMILARITY INDEX

6 %
INTERNET SOURCES

0 %
PUBLICATIONS

5 %
STUDENT PAPERS

PRIMARY SOURCES

1 download.atlantispress.com **6** %
Internet Source

2 Submitted to Forum Perpustakaan Perguruan Tinggi Indonesia Jawa Timur **5** %
Student Paper

Exclude quotes On

Exclude matches < 5%

Exclude bibliography On

The Classification of Acute Respiratory Infection (ARI) Bacteria Based on K-Nearest Neighbor

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11
