JOIN (Jurnal Online Informatika)
Volume 6 No.1 | June 2021: 63-72
DOI: 10.15575/join.v6i1.704
p-ISSN: 2528-1682
e-ISSN: 2527-9165

Identification of White Blood Cells Using Machine Learning Classification Based on Feature Extraction

Anwar Siswanto[1], Abdul Fadlil[2], Anton Yudhana[3]
[1]Informatics Engineering Master Program, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[23]Departement of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Article Info

Article history:

Received February 20, 2021
Revised March 16, 2021
Accepted March 17, 2021
Published June 15, 2021

Keywords:

Sel Darah Putih
Gray Level Co-occurrence Matrix
K-nearest Neighbor
Naïve Bayes
Multilayer Perceptron

ABSTRACT

In various disease diagnoses, one of the parameters is white blood cells, consisting of eosinophils, basophils, neutrophils, lymphocytes, and monocytes. Manual identification takes a long time and tends to be subjective depending on the staff's experience, so the automatic identification of white blood cells will be faster and more accurate. White blood cells are identified by examining a colored blood smear (SADT) and examined under a digital microscope to obtain a cell image. Image identification of white blood cells is determined through HSV color space segmentation (Hue, Saturation Value) and feature extraction of the Gray Level Cooccurrence Matrix (GLCM) method using the Angular Second Moment (ASM), Contrast, Entropy, and Inverse Different Moment (IDM) features. The purpose of this study was to identify white blood cells by comparing the classification accuracy of the K-nearest neighbor (KNN), Naïve Bayes Classification (NBC), and Multilayer Perceptron (MLP) methods. The classification results of 100 training data and 50 white blood cell image testing data. Tests on the KNN, NBC, and MLP methods yielded an accuracy of 82%, 80%, and 94%, respectively. Therefore, MLP was chosen as the best classification model in the identification of white blood cells.

Corresponding Author:

Anwar Siswanto,
Informatics Engineering Master Program,
Universitas Ahmad Dahlan,
Jl. Prof. Dr. Soepomo, S.H. Janturan Yogyakarta, Indonesia
Email: anwar1907048006@webmail.uad.ac.id

1. INTRODUCTION

Human blood consists of blood and fluid components. White blood cells are part of blood components, which are responsible for the immune system and have the function of destroying objects that are considered foreign and harmful to the body. There are five types of white blood cells, namely eosinophils, basophils, neutrophils, lymphocytes, and monocytes [1]. The shape and characteristics of white blood cells vary [2].

Observation of white blood cells was carried out using a digital microscope by examining the smear stained with Giemsa, Wright, and May Grunwald. Identification of white blood cells is done manually, has shortcomings that are subjective depending on the workload and experience of staff, and requires a long time [3]. Automatic identification based on digital images [4] is expected to help identify cells quickly, precisely, and efficiently

This study uses HSV color space segmentation (Hue, Saturation, and Value) to represent actual color values [5]. The results of the segmentation were carried out with feature extraction using the GLCM method. GLCM is a texture extraction method in images and has computational advantages and better accuracy than other feature extraction methods [6]. The features used are Angular Second Moment (ASM), Contrast, Entropy, and Inverse Different Moment (IDM).

Fitri conducted a white blood cell identification study comparing the KNN [7] and SVM (Support Vector Machine) classifications based on color and shape by adding contrast stretching to improve image quality. The KNN [8] method gets the best accuracy of 94.3%.

Mahran et al. conducted research on mushroom classification based on the first-order statistical feature extraction with the Gaussian Naïve Bayes classification method [9] on 60 images of mushrooms. The test results with the Cross-Validation method and the value of K = four obtained an accuracy of 98.75%.

Liyantoko et al. Researching the classification of white blood cells and lymphoblasts using the Multilayer Perception (MLP) and backpropagation method [10] on GLCM feature extraction of geometric and color features resulted in an accuracy value of 91.43% and a precision of 50.63%.

From the three studies, it can be seen that each classification method has its advantages. Therefore, this study was to determine the comparison of the accuracy in the identification of white blood cells from the three classification models. The classification models chosen in this study are KNN, NBC, and MLP. The dataset used is from the LISC (leukocyte images from segmentation and classification) (Rezatofighi & Soltanian-Zadeh, 2011), which consists of five classes of white blood cells, namely eosinophils, basophils, neutrophils, lymphocytes, and monocytes.

2. METHOD

This research uses octave software with the GLCM method for feature extraction. The classification uses the K-nearest Neighbor (KNN) method, Naïve Bayes Classification (NBC) and Multilayer Perceptron (MLP), and Weka software. The author tested three ways to obtain high accuracy and a time-efficient method. The research stages are shown in Figure 1.

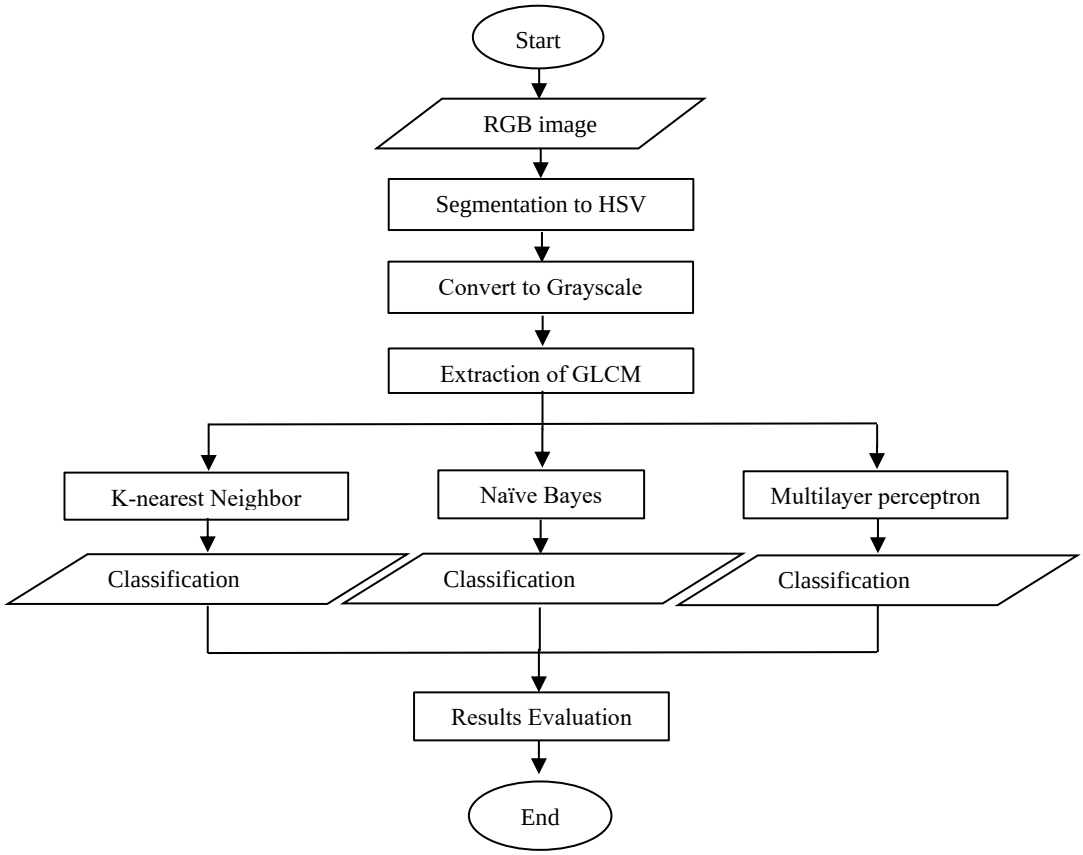
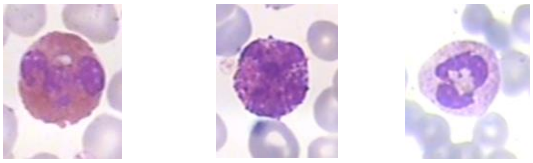


Figure 1. Feature Extraction and Classification Process

The first stage is cropping the image of white blood cells manually using a digital microscope. The second stage is uploading to the HSV color space segmentation program (Hue, Saturation, and Value) [11]. The third stage is image conversion to grayscale. The fourth stage features feature extraction of Angular Second Moment (ASM), Contrast, Entropy and Different Moment (IDM), Entropy. The fifth stage of the feature extraction results is classified using the KNN, NBC, and MLP methods. The sixth stage results in the classification of 5 classes of white blood cells. The seventh stage evaluates the results of the classification of the 3 methods.

2.1. White blood cell

White blood cells have different characteristics and shapes, as in Figure 2.

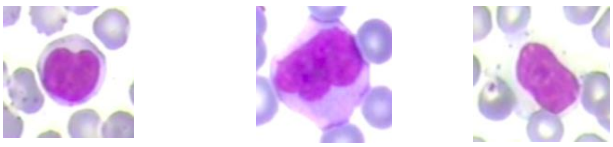


Eosinophil Cells Basophil Cells Neutrophil Cells

Lymphocyte Cells Monocyte Cells

Figure 2. Shape and characteristics of white blood cells.

Position of white blood cells about other cells [12]. Do not coincide with each other, overlap with other cells; the outline is broken or not intact as in Figure 3.



Not Crushed Coincide Broken Edges

Figure 3. Shape and characteristics of white blood cells.

The shape and characteristics of white blood cells are described in table 1

Table 1. Types, sizes, and characteristics of white blood cells

Type	Sizes	Characteristics
Eosinophil	16µ	Cytoplasm granules are coarser and orange in color.
Basophil	14µ	The granules are varied, irregular in arrangement to cover the nucleus, and are azurophilic.
Neutrophil	14µ	Granules in the form of thin, refined grains, pink and faint in color.
Limphocyte	12µ	The cell is almost covered with a dense, granular nucleus.
Monocyte	18µ	The largest size is solid core and curves like a kidney.

2.2. Hue Saturation and Value (HSV) Color Space

A.R Smith introduced the 1978 HSV color space. Hue represents the colors known to humans, such as green and red, which are produced by wavelengths. Saturation is the strength of color or color purity. Value reflects of the intensity of the object, expressed as a change in white to a dark color or known as grayscale (grey level). The value is between 0-100%. The value 0 is black, as in Figure 4.

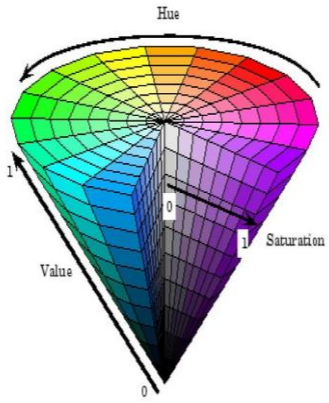


Figure 4. HSV Color Space

The h (hue) color space is like formula 1..

$$h(hue) = \begin{cases} 0, if \max = min \\ 60^0 x \left(\frac{G-B}{\max - min} mod\ 6\right), if \max = R \\ 60^0 x \left(\frac{B-R}{\max - min} + 2\right), if \max = G \\ 60^0 x \left(\frac{R-G}{\max - min} + 4\right), if \max = B \end{cases} \quad (1)$$

The h (hue) color space represents fractional values around the circle, starting at red, which has a zero shade. The color space s (saturation) is like formula 2.

$$s(saturation) = \begin{cases} 0\ , if \max = min \\ \frac{max - min}{V}, otherwise \end{cases} \quad (2)$$

Color space s (saturation) in RGB range (0,1). The RGB transformation (65, 27,234) is divided by 255 into HSV form [13]. Color space v (value) is like formula 3.

$$V\ (value) = max \quad (3)$$

Color space v (value) maximum value = 0.918.

2.3. RGB Color Space

The three essential components are Red (red), Green (green), and Blue (blue). The 3-dimensional cube is an RGB model [14] with red, green, and blue colors at the corners of the axis, as shown in figure 5.

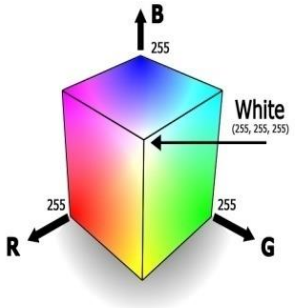


Figure 5. RGB Color Space

Each image pixel is represented by 24 bits, namely 8 bits R (red), 8 bits G (green), 8 bits B (blue). The grayscale color space [15] is displayed in white with the highest intensity (255) [16] and black with the lowest power (0)—formulas in equations such as formula 4.

$$Gray=((R*0.2989)+(G*0.5870)+(B*0.1140)) \quad (4)$$

R is the red value, G the green value, and B the blue value.

2.4. Gray Level Co-occurrence Matrix (GLCM)

Second-order texture extraction is GLCM information [17]. The GLCM matrix is a matrix presentation between neighboring pixels in various spatial distance directions (d) and orientation direction θ. The method of calculating the cooccurrence matrix is based on the angle of the neighbor direction of 2 pixels [18], namely the double-angle cooccurrence matrix and single-angle cooccurrence matrix [19]. The angle 0 °, 45 °, 90 °, 135 ° [20] is the direction of the angle used, as shown in Figure 6.

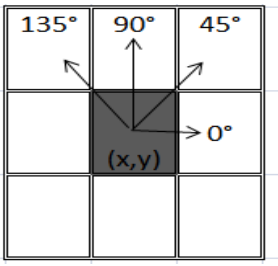


Figure 6. Directional Angle Matrix 0 °, 45 °, 90 °, 135 °

Order statistical features, namely Angular Second Moment (ASM), Correlation, Contrast, Entropy, and Inverse Different Moment (IDM) [21]. The equation formula is as follows:

1. Angular Second Moment(ASM)
   It is the inverse of Entropy and a measure of local homogeneity, as in formula 5.

$$Energy = \sum_{j=0}^{g-1}\sum_{i=0}^{g-1}(p(i,j))^2 \tag{5}$$

   Is $p(i,j)$ normalized matrix, is the value in row i and column j.
2. Contrast
   Is the degree of the greyness of the image area, the difference is measured, as in formula 6.

$$Contr = \sum_{i}\sum_{j}(i-j)^2.p(i,j) \tag{6}$$

   Is the i value in the row, and j is the column value.
3. Entropy
   It is to show the measure of texture shape irregularity, like formula 7.

$$E = -\sum_{i}\sum_{j}p(i,j).2_{\log\,[p(i,j)]} \tag{7}$$

   Where $p(i,j)$ is the value of i and j multiplied by log2 at the value of $p(i,j)$.
4. Inverse Different Moment (IDM)
   The weight value is the inverse of the contrast, measured the homogeneity level of the texture structure repetition.

$$IDM = \sum_{i,j=0}^{N-1}\frac{p(i,j)}{1+(i-j)^2} \tag{8}$$

   p (i, j) is the value of i and j, divided by the number1 and added by the weight of k

2.5. K-nearest Neighbor (KNN).

The K-Nearest Neighbor (K-NN) method is one of the oldest and most popular N-based methods. The K value used here represents the number of closest neighbors involved in determining the class label prediction on the test data. From the nearest K, the closest neighbor is selected, then a class voting is conducted from the nearest neighbor K. It is the class with the highest number of neighboring votes that are given as the class label of the predicted results on the test data.

The most commonly used distance calculation in the KNN algorithm [22] is the Euclidean distance calculation. The formula is as formula 9.

$$d = \sqrt{(x2-x1)^2+(y2-y1)^2} \tag{9}$$

Where d is the distance of the object, x is the training data, and y is the testing data.

2.6. Naïve Bayes.

The British scientist Thomas Bayes [23] proposed a classification method using statistics and probability to predict future odds based on previous data. Bayes' theorem is like the formula 10

$$P(C_i|\,X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{10}$$

Where C_i is a data hypothesis X is a specific class, x is data with a new level, $P(C_i|\,X)$ is the probability of the C_ii hypothesis based on the condition x, $P(X|C_i)$ is the probability x based on the hypothetical condition C_i. Data with numerical value has a standard or Gaussian probability by calculating the value of the mean μ and standard deviation Ϭ for each class formulated by the equation formula 11.

$$p(X_i = x_i|Y = y_i) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} exp^{\frac{-(x_i-\mu_i)^2}{2\sigma^2{}_{ij}}} \tag{12}$$

It is explained that Ϭ is the variance of one variant for the population, X_i is the midpoint of the value in one attribute, μ is the average or mean of the community, and n is the number of data.

2.5. Multilayer Perceptron (MLP).

The function and structure of the human brain are examples of models in the preparation of Artificial Neural Networks (ANN). Neurons are components of an artificial neural network. Neurons are connected to many neurons; each neuron connection has a weight. Biological neural networks, there are equal parts of ANN. The essential characteristic of ANN is learning through the weight adjustment of neuron connections.

MLP consists of several neurons connected by connecting weights. Each perceptron is connected to form layers (layers). Each MLP [24] consists of one input layer (input layer), one or more hidden layers (hidden layer), and one output layer (output layer), as Figure 7.

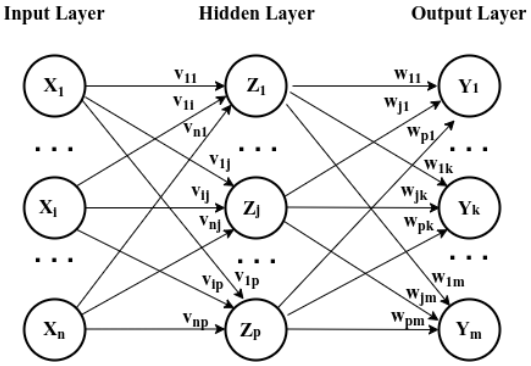


Figure 7. Artificial Neural Network Architecture.

In Figure 7, the input layer is denoted by X, the hidden layer is denoted by Z, and the output layer is denoted by Y. The weight between X and Y is denoted by v, while the weight between Z and Y is denoted by w.

The MLP learning method is backpropagation consisting of 4 stages. The first stage is initialization, the second is activation, the third is weight training, and the fourth is iteration. The initialization stage of the initial weight value and the threshold value are determined randomly within certain limits. The activation stage is given the input and output values predicted. In the weight training stage, the actual output value is compared to the expected value, and the weight adjustment is made. The second and third stages are repeated (iteration) until certain conditions.

3. RESULTS AND DISCUSSION

Image size used for this research, 200 x 200 pixels, and digital data format * PNG. The total number of training data for each class of white blood cells was 20, 100 training data, and 10 test data for each type, 50 test data, as in table 2.

Table 2. Databases Used in Research

Data	Eosinophil	Basophil	Neutrophil	Limphocyte	Monocyte	Total
Training	20	20	20	20	20	100
Test	10	10	10	10	10	50
Total						150

3.1. Image Processing

The steps taken are image acquisition and segmentation in the HSV color space, including blue, purple, magenta, and pink areas. The image is cropped manually, and the HSV color space is segmented. The results of cropping and segmentation are as in Table 3.

Table 3. Cropping of RGB Image and HSV Space Segmentation.

Cell Class	RGB Image	Cropped	Blue	Purple	Magenta	Pink
Eosinophil						
Basophil						
Neutrophil						
Limphocyte						
Monocyte						

3.2. Feature Extraction

Image of blue, purple, magenta, and red color space is segmented to grayscale. Each color space features GLCM extraction of 4 features, namely Angular Second Moment (ASM), Contrast, Entropy, and Different Moment (IDM). Each feature is tested with an Angle of 0 °, 45 °, 90 °, 135 °. Then one white blood cell class totals 64 attributes. The results of feature extraction for one of the training data are as shown in table 4.

Table 4. Extraction Results of GLCM Features

Color Space	ASM				Contrast				Entropy				IDM			
	0°	45°	90°	135°	0°	45°	90°	135°	0°	45°	90°	135°	0°	45°	90°	135°
Blue	0.999	0.999	0.999	0.999	0.003	0.004	0.004	0.004	0.002	0.002	0.002	0.002	1	1	1	1
Purple	0.859	0.853	0.860	0.854	162.8	244.2	165.9	230.0	0.346	0.361	0.342	0.360	0.945	0.939	0.948	0.940
Magenta	0.571	0.563	0.572	0.563	177.3	269.4	179.8	249.5	1.063	1.107	1.052	1.104	0.830	0.813	0.835	0.813
Pink	0.956	0.952	0.955	0.952	12.20	22.06	16.40	21.79	0.106	0.115	0.108	0.115	0.984	0.981	0.984	0.981

The total test attributes are 4 (color space) x 4 (GLCM features) x 4 (directional angle) = 64 attributes.

3.3. Classification

The results of the GLCM feature extraction for each class of white blood cells were classified against three methods as training data. The results were learning classification methods for testing the test data.

3.3.1. K-nearest Neighbor (KNN)

The results of the classification of the data model, the time needed to build the classification model is 0 seconds. The classification model data is correct 84% and wrong 16%, as in Figure 8.

```
Time taken to build model: 0 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances          84               84      %
Incorrectly Classified Instances        16               16      %
```

Figure 8. Data of the KNN Classification Model.

The test results on the 50 test data obtained results as in Table 5.

Table 5.Test Results on the KNN Classification

Cell Class	E	B	N	L	M
E = Eosinophil	10	0	0	0	0
B = Basophil	0	9	0	0	1
N = Neutrophil	0	3	7	0	0
L = Limphocyte	0	0	1	8	1
M = Monocyte	0	0	3	0	7
% Accuracy = $\frac{41}{50}$ x 100 = 82 %					

The highest identification accuracy was for ten eosinophil cells. The lowest was 7 for neutrophil cells and monocytes. The KNN classification was able to identify 41 types of white blood cells from 50 test data with or an accuracy of 82%.

3.3.2. Naïve Bayes

The results of the classification of the data model, the time needed to build the classification model is 0.02 seconds. The classification model data is correct 80% and wrong 20%, as in Figure 9.

```
Time taken to build model: 0.02 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances          80               80      %
Incorrectly Classified Instances        20               20      %
```

Figure 9. Data of the Naïve Bayes Classification Model.

The test results on the 50 test data obtained results as in Table 6.

Table 6.Test Results on the Naïve Bayes Classification.

Cell Class	E	B	N	L	M
E = Eosinophil	9	1	0	0	0
B = Basophil	0	9	0	0	1
N = Neutrophil	0	0	10	0	0
L = Limphocyte	0	1	3	6	0
M = Monocyte	0	4	0	0	6
% Accuracy = $\frac{40}{50}$ x 100 = 80 %					

The highest test results were for ten neutrophil cells; the lowest identification was 6 for lymphocytes and monocytes. The Naïve Bayes classification can identify 40 types of white blood cells from 50 test data with or an accuracy of 80%.

3.3.3. Multilayer Perceptron (MLP)

MLP testing uses 64 input layer neurons and one hidden layer with 64 neurons and five output layers. Epoch value of 500 and learning rate of 0.3. For the test results of the data model, the time needed to build the classification model is 5.31 seconds. Classification model data is accurate 96% and false 4%, as in Figure 10.

```
Time taken to build model: 2.08 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances          96               96      %
Incorrectly Classified Instances         4                4      %
```

Figure 10. Multilayer Perceptron Classification Model Data.

The results of testing on 50 test data obtained results as in Table 7.

Table 7. Test Results on the Multilayer Perceptron Classification

Cell Class	E	B	N	L	M
E = Eosinophil	10	0	0	0	0
B = Basophil	0	10	0	0	0
N = Neutrophil	0	4	6	0	0
L = Limphocyte	0	0	0	10	0
M = Monocyte	0	0	0	0	10
% Accuracy = $\frac{46}{50}$ x 100 = 92 %					

The highest identification accuracy of eosinophils, basophils, lymphocytes, and monocytes was ten cells. The lowest was 6 for neutrophil cells. The MLP classification was able to identify all types of white blood cells by 46 out of 50 test data with or an accuracy of 92%.

3.4. Evaluation Testing

After testing three classification methods and five classes of white blood cells, the results are shown in Table 8, Table 9.

Table 8. Test Results Against 3 Methods.

Cell Class	KNN (%)	Naïve Bayes (%)	MLP (%)
Eosinophil	100	90	100
Basophil	90	90	100
Neutrophil	70	100	60
Limphocyte	80	60	100
Monocyte	70	60	100
Average	82	80	92

Table 9. Classification Test Results.

Testing	KNN	Naïve Bayes	MLP
Classification corrections	84 %	80 %	96 %
Time (seconds)	0	0.02	2.08
Test Results	82 %	80 %	92 %

The test results on the KNN method, accurately 82% of the predicted data models were 84% with a test time of 0 seconds. Identify each class of white blood cells as much as 1 class (eosinophils) with a value of 100%. The Naïve Bayes method, accurately 80% of the prediction data model, is 80% with a testing time of 0.02 seconds. Identify each class of white blood cells as much as 1 class (neutrophils) with a value of 100%. The MLP method, accurately 92% of the prediction model data, is 96% with a testing time of 2.08 seconds. Identify each class of white blood cells as many as four classes (eosinophils, basophils, lymphocytes, and neutrophils) with a value of 100%.

4. CONCLUSION

The classification method test has been successfully carried out. The conclusion of the study of white blood cell identification using the MLP method resulted in the best accuracy. The research results on the performance of the NBC method have an accuracy of 80%, the KNN method has an accuracy of 82%, and the MLP method has an accuracy of 92% with the longest time of 2.08 seconds. MLP was chosen as the best classification model for identifying of white blood cells even though it requires a longer time. The high accuracy value can be due to the color absorption of the stain time properly and the cropping of white blood cells that do not overlap. The accuracy value is low because the blood cells have been damaged or broken, and the cropping of cells that is coincided causes the color of white blood cells to be mixed with the color of other cells, and the cell border is not intact.

5. REFERENCES

[1] R. Ardina and S. Rosalinda, "Morfologi Eosinofil Pada Apusan Darah Tepi Menggunakan Pewarnaan Giemsa, Wright, dan Kombinasi Wright-Giemsa," *J. Surya Med.*, vol. 3, no. 2, pp. 5–12, Feb. 2018, doi: 10.33084/jsm.v3i2.91.

[2] D. Arwie and Islawati, "Penentuan Kriteria Penilaian Kesan Jumlah Leukosit Pada Pemeriksaan Apusan Darah Tepi," *J. Kesehat. Panrita Husada*, vol. 3, no. 2, pp. 118–127, Sep. 2018, doi: 10.37362/jkph.v3i2.188.

[3] A. Setiawan, E. Suryani, and , W., "Segmentasi Citra Sel Darah Merah Berdasarkan Morfologi Sel Untuk Mendeteksi Anemia Defisiensi Besi," *J. Teknol. Inf. ITSmart*, vol. 3, no. 1, 2016, doi: 10.20961/its.v3i1.638.

[4] M. N. Khasanah, A. Harjoko, and I. Candradewi, "Klasifikasi Sel Darah Putih Berdasarkan Ciri Warna dan Bentuk dengan Metode K-Nearest Neighbor (K-NN)," *IJEIS (Indonesian J. Electron. Instrum. Syst.*, vol. 6, no. 2, p. 151, Oct. 2016, doi: 10.22146/ijeis.15254.

[5] D. S. Alham and D. Herumurti, "Segmentasi Dan Perhitungan Sel Darah Putih Menggunakan Operasi Morfologi Dan Transformasi Watershed," *INFORMAL Informatics J.*, vol. 4, no. 2, 2019, doi: 10.19184/isj.v4i2.13347.

[6] R. Widodo, A. W. Widodo, and A. Supriyanto, "Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (Citrus reticulata Blanco) untuk Klasifikasi Mutu," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, 2018.

[7] A. Siswanto, A. Fadlil, and A. Yudhana, "Ekstraksi Ciri Metode Gray Level Co-Occurrence Matrix Untuk Identifikasi Sel Darah Putih," *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 5, no. 2, 2020, doi: 10.31328/jointecs.v5i2.1334.

[8] Z. E. Fitri, L. N. Y. Syahputri, and A. M. N. Imron, "Classification of White Blood Cell Abnormalities for Early Detection of Myeloproliferative Neoplasms Syndrome Based on K-Nearest Neighbor," *Sci. J. Informatics*, vol. 7, no. 1, pp. 136–142, Jun. 2020, doi: 10.15294/sji.v7i1.24372.

[9] A. A. Mahran, R. K. Hapsari, and H. Nugroho, "Penerapan Naive Bayes Gaussian Pada Klasifikasi Jenis Jamur Berdasarkan Ciri Statistik Orde Pertama," *Netw. Eng. Res. Oper.*, vol. 5, no. 2, 2020, doi: 10.21107/nero.v5i2.165

[10] A. N. Liyantoko, I. Candradewi, and A. Harjoko, "Klasifikasi Sel Darah Putih dan Sel Limfoblas Menggunakan Metode Multilayer Perceptron Backpropagation," *IJEIS (Indonesian J. Electron. Instrum. Syst.*, vol. 9, no. 2, 2019, doi: 10.22146/ijeis.49943.
[11] A. F. Hastawan, R. Septiana, and Y. E. Windarto, "Perbaikan hasil segmentasi hsv pada citra digital menggunakan metode segmentasi rgb grayscale," *Edu Komputika J.*, vol. 6, no. 1, 2019, doi: 10.15294/EDUKOMPUTIKA.V6I1.23025.
[12] N. N. Arisa and C. Fatichah, "Perhitungan dan pemisahan sel darah putih berdasarkan Centroid dengan menggunakan metode multi pass voting dan k-means pada citra sel acute leukemia," *JUTI J. Ilm. Teknol. Inf.*, vol. 16, no. 2, p. 105, Oct. 2018, doi: 10.12962/j24068535.v16i2.a661.
[13] G. A. Pradipta and P. D. Wulaning Ayu, "Perbandingan segmentasi citra telur ayam menggunakan metode otsu berdasarkan perbedaan ruang warna rgb dan hsv," *JST (Jurnal Sains dan Teknol.*, vol. 6, no. 1, 2017, doi: 10.23887/jst-undiksha.v6i1.9329.
[14] H. Sanusi, S. H. S., and D. T. Susetianingtias, "Pembuatan aplikasi klasifikasi citra daun menggunakan ruang warna rgb dan hsv," *J. Ilm. Inform. Komput.*, vol. 24, no. 3, 2019, doi: 10.35760/ik.2019.v24i3.2323.
[15] S. Saifullah, S. Sunardi, and A. Yudhana, "Perbandingan segmentasi pada citra asli dan citra kompresi wavelet untuk identifikasi telur," *Ilk. J. Ilm.*, 2016, doi: 10.33096/ilkom.v8i3.75.190-196.
[16] S. I. Syafi'i, R. T. Wahyuningrum, and A. Muntasa, "Segmentasi obyek pada citra digital menggunakan metode otsu thresholding," *J. Inform.*, 2016, doi: 10.9744/informatika.13.1.1-8.
[17] K. Novar Setiawan and I. M. Suwija Putra, "Klasifikasi Citra Mammogram Menggunakan Metode K-Means, GLCM, dan Support Vector Machine (SVM)," *J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi)*, 2018, doi: 10.24843/jim.2018.v06.i01.p02.
[18] Y. Agussationo, I. Soesanti, and W. Najib, "Klasifikasi Citra X-Ray Diagnosis Tuberkulosis Berbasis Fitur Statistis," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 3, 2018, doi: 10.29207/resti.v2i3.523.
[19] S. Saifudin and A. Fadlil, "Sistem identifikasi citra kayu berdasarkan tekstur menggunakan gray level coocurrence matrix (glcm) dengan klasifikasi jarak euclidean," *SINERGI*, 2015, doi: 10.22441/sinergi.2015.3.003.
[20] R. A. Surya, A. Fadlil, and A. Yudhana, "Identification of Pekalongan Batik Images Using Backpropagation Method," in *Journal of Physics: Conference Series*, 2019, vol. 1373, no. 1, doi: 10.1088/1742-6596/1373/1/012049.
[21] S. Saifullah, S. Sunardi, and A. Yudhana, "Analisis Ekstraks Ciri Fertilitas Telur Ayam Kampung dengan Grey Level Cooccurrence Matrix," *J. Nas. Tek. ELEKTRO*, 2017, doi: 10.25077/jnte.v6n2.376.2017.
[22] F. Yunita, "Diabetes Mellitus Menggunakan Metode K-Nearest Neighbor (K-Nn)," *Bappeda*, vol. 2, 2016.
[23] M. A. Bianto, K. Kusrini, and S. Sudarmawan, "Perancangan Sistem Klasifikasi Penyakit Jantung Mengunakan Naïve Bayes," *Creat. Inf. Technol. J.*, vol. 6, no. 1, 2020, doi: 10.24076/citec.2019v6i1.231.
[24] M. Sofie and A. Rizal, "Klasifikasi citra rekaman sinyal elektrokardiogram menggunakan metode analisis tekstur, k-nn dan multilayer perceptron," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 7, no. 1, 2016, doi: 10.24176/simet.v7i1.509.