Journal of Physics: Conference Series

PAPER • OPEN ACCESS

# Corn Leaf Disease Classification Using Local Binary Patterns (LBP) Feature Extraction

To cite this article: Aeri Rachmad *et al* 2022 *J. Phys.: Conf. Ser.* **2406** 012020

View the article online for updates and enhancements.

You may also like

- A review and experimental study on the application of classifiers and evolutionary algorithms in EEG-based brain–machine interface systems
  Farajollah Tahernezhad-Javazm, Vahid Azimirad and Maryam Shoaran
- Solar Flare Prediction Model with Three Machine-learning Algorithms using Ultraviolet Brightening and Vector Magnetograms
  N. Nishizuka, K. Sugiura, Y. Kubo et al.
- A *k*-NN METHOD TO CLASSIFY RARE ASTRONOMICAL SOURCES: PHOTOMETRIC SEARCH OF BROWN DWARFS WITH *SPITZER*/IRAC
  Massimo Marengo and Mayly C. Sanchez


This content was downloaded from IP address 140.0.123.248 on 23/12/2022 at 10:43

# Corn Leaf Disease Classification Using Local Binary Patterns (LBP) Feature Extraction

Aeri Rachmad[1*], Mohammad Syarief[2], Silfia Rifka[3], Fifin Sonata[4], Wahyudi Setiawan[5], Eka Mala Sari Rochman[6]

[1,2,5,6]Departemen of Informatics, Faculty of Engineering, University of Trunojoyo, Madura, Bangkalan, Indonesia
[3]Telecommunication Engineering, Padang State Polytechnic, Indonesia
[4]Informatics Management, STMIK Triguna Dharma, Medan, Indonesia

**Abstract**. Corn is a plant that is widely grown in developing countries such as Indonesia. To increase maize yields, researchers are always innovating on the current state of technology for classifying maize plant diseases. Three kinds of diseases attack corn leaves, namely Gray leaf Spot, Blight, and Common Rush. The amount of data that we use is 3500 data consisting of 500 Gray Leaf Spots, 1000 Blights, 1000 Common Rushes, and 1000 healthy leaves. This study aims to develop an artificial intelligence model. The artificial intelligence model that we developed uses LBP feature extraction combined with k-NN for the classifier. In addition to using the k-NN method, our tests were carried out using several classification methods such as Naïve Bayes and Adaboost. The result of our test is that the k-NN method has the highest value compared to the Naïve Bayes and Adaboost methods. The results of the performance using k-NN with k=5 resulted in a value of 81.1%, the AUC value of 94.1%, the F1-Score of 80.9%, Precision of 81.8%, and Recall of 81.1%.

**Keywords**: Corn Leaf Diseases, LBP, k-NN, Classification.

## 1. Introduction

Indonesia is one of the corn producers in the world [1]. This corn plant is found in almost all parts of Indonesia, such as on the island of Madura. This is because the island of Madura has the type of soil and weather that supports the growth of corn plants [2]. Currently, corn fields are not only used for agricultural fields but are also used as tourist destinations on Madura Island.

One of the factors that can affect corn yields is the health of corn plants which can be seen in the diseases experienced in plants. This disease in corn plants is an abnormality in plants that can interfere with the growth of corn plants. Some of the diseases that attack corn plants from leaves include blight, rust, and spot [3]. Diseases that attack corn stalks are gibberella stem rot, Pythium stem rot, and Charcoal stem rot [4]. Diseases that attack the cob are red cob rot, penicillium cob rot, diplodia cob rot. Corn plant diseases can not only be seen through the human sense of sight, it is because human traits such as fatigue, emotions, and others can affect the quality of the classification. So, we need a Digital Image Processing technique in the context of disease classification in corn plants [5].

Previous studies used the Local Binary Patterns (LBP) method for texture feature extraction and Color Histogram for color feature extraction. As for the classification of the extraction results, modeling using Random Forest will be carried out and the results will be stored for use in the testing phase. In addition, there is also research on the identification of diseases in the leaves of apple plants. The feature

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

extraction used is LBP and Color Histogram. Random forest was used as a classification. Based on the results of testing on apple plants using a dataset of 2,154 leaf images, an average accuracy of 91.41% [6].

In another study, namely food image classification using HSV Color Moment and LBP with Naïve Bayes Classifier. Testing is done by dividing the data into two parts, namely 30 training data and 20 testing data. The accuracy results obtained are 60% for the HVS method and 60% for LBP [7].

Based on the above problems, researchers will detect diseases in corn leaves using the k-NN method. The k-NN method will be combined with LBP feature extraction to produce high accuracy in identifying diseases on corn leaves.

## 2. Methodology

In the system design made in Figure 2, a disease image search system on corn leaf texture in the database with the lowest similarity and closest to the image desired by the user, there are processes carried out from the beginning of the data selected, until the Final, the data is found similarity distance by using Euclidean Distance. There are several steps in this research, the first is data. Figure 1 shows corn leaf data which has 4 classes, namely Health, Gray leaf Spot, Blight, and Common Rush. Images of corn leaves can be accessed through the Kaggle.com website.
https://www.kaggle.com/datasets/arhasnaazzahra/cornleavediseases

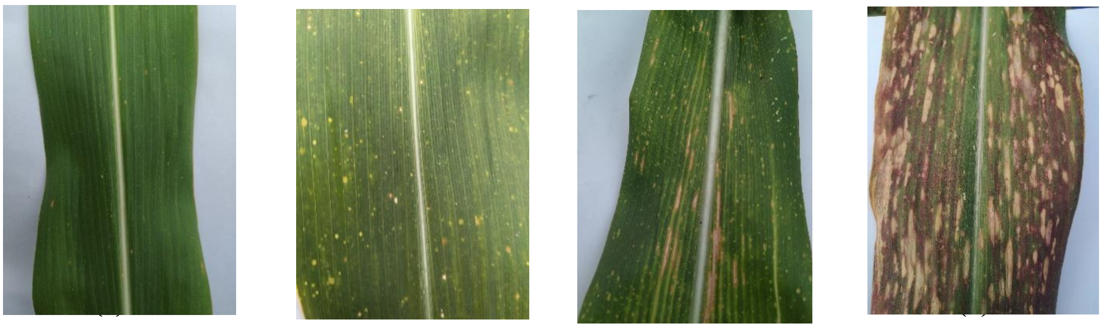


**Figure 1.** image of healthy corn leaves and leaf disease on corn with a size of 3000 x 4000 pixels (a) Healthy (b) Gray leaf Spot (c) Blight (d) Common Rush

Second, equalize the size of the 3000 x 4000 corn leaf image to 300 x 300 then change the RGB leaf image to grayscale. The third step is to perform feature extraction of corn leaves using the LBP method. In the fourth step, after getting the results of feature extraction using the LBP method, classification was carried out using k-NN using Euclidean Distance to determine the class of corn leaf disease.

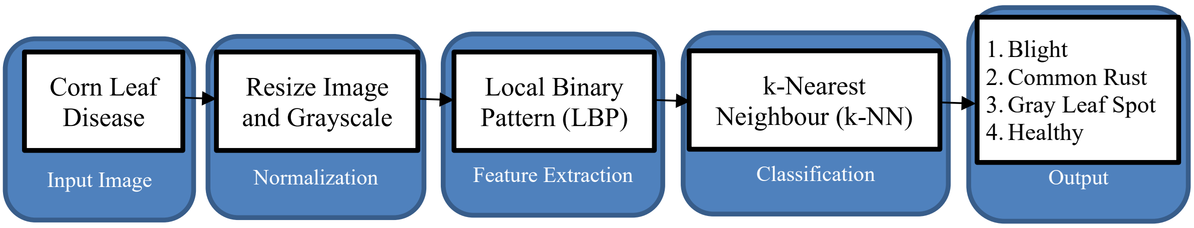



**Figure 2.** System Design

### 2.1. *Grayscale*

Grayscale is an image or image format in which each image pixel consists of 1 color composition [8]. The grayscale technique is to calculate the value of the RGB color image, namely red, green, and blue to distinguish between the shadow and the original color of the image which in this final project uses a leaf image. To calculate the grayscale image using equation (1).

$$Y = \frac{1}{3}(R + G + B) \tag{1}$$

The following is an explanation of the equation:
Y = Output image
R = Red input image
G = Green input image
B = Blue input image

### *2.2. Local Binary Patterns (LBP)*

Local Binary Patterns (LBP) is a feature extraction algorithm that extracts statistical and structural image texture features [9]. The LBP method was first introduced by Timo Ojala. The way the LBP operator works is to use a comparison of the gray values of neighbouring pixels. Local Binary Patterns is a texture feature extraction method that is rotation invariant. The LBP value itself is obtained from the thresholding process and then the value is multiplied by the binary weight [10]. The most basic form of LBP uses the value of the center pixel as a threshold at 3x3 neighbouring pixels. The threshold operation will create a binary pattern that represents the texture character. The basic LBP operator is shown in Figures 3 and 4.

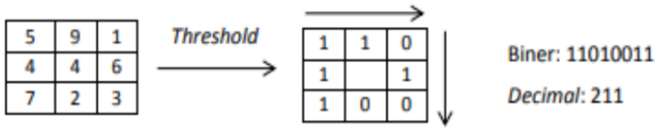



**Figure 3**. LBP Feature Extraction Basic Operator Example

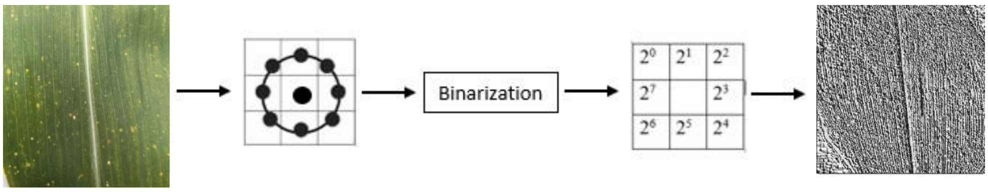



**Figure 4.** Sample pictures before and after the LBP process

The following are the steps for calculating LBP [11]:

1. Input is an image measuring 300 x 300 which has been done by the Grayscale and resizing process previously.
2. Determine the radius value of 1 and the neighbour intensity value of 8.
3. The binarization process is carried out, ie 8 neighbouring pixel values will be compared with the value at the center of the image. The 8-pixel values will be reduced by the center value and will be changed to 1 if the result is positive, if the result is negative then it will be changed to 0.
4. The results from the previous step will be multiplied by the weight value
5. decimal with the direction of rotation of the weights from the smallest to the largest clockwise.
6. The output is the image from the LBP.

### *2.3. K-Nearest Neighbour (k-NN)*

k-Nearest Neighbour is a classification method to distinguish objects based on the closest learning data to the object. Learning data is projected onto many dimensional spaces, where each dimension represents a data feature. The room is divided into several parts based on the classification of learning data. The best value of k for this algorithm depends on the data, in general, the value of k can reduce the effect of noise on the classification, but it can make the boundaries between each classification less clear. There are many ways to measure the distance between the proximity of new data (proximity) and old data (training data), including the Euclidean distance and Manhattan distance, the most commonly used is the Euclidean distance [12].

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_{ik} - y_{ik})^2} \quad (2)$$

Keterangan:
$d$ : Euclidean distance
$x_{ik}$ : the i-th value of the k-th variable from the training data
$y_{ik}$ : the i-th value of the k-th variable from the testing data
$n$ : number of variables / data dimension

The KNN algorithm basically consists of 4 stages. In the first stage, a calculation is made between the distance from the new data to all the data. In the second stage, the distances are sorted. The third stage chooses the smallest k value and the last stage determines the class [13].

### 2.4. *Confusion Matrix*

Confusion Matrix is a table used for performance analysis that will facilitate visualization as shown in Table1. The matrix can distinguish between True Positive, False Negative, False Positive, and True Negative [14] [15][16].

Table 1. *Confusion Matrix*

| Actual Class | Prediction Class | |
|---|---|---|
| | Positive | Negative |
| Positive | True Positive (TP) | False Negative (FN) |
| Negative | False Positive (FP) | True Negative (TN) |

Accuracy is the ratio of correctly classified over all available data quantities:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Precision or positive predictive value (PPV), is the relationship between true positive and all positive predictions:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

Recall or true positive value (TPV), is the relationship between true positives and all positive elements:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

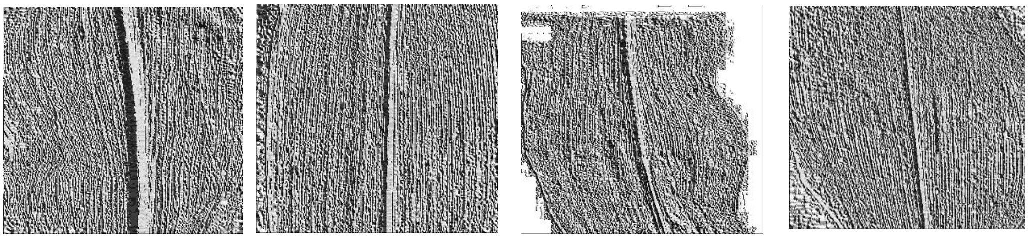
F1-Score is two times the precision value and the recall value divided by the total precision value and recall value:

$$\text{F1} - \text{Score} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall}+\text{Precision}} \quad (6)$$

## 3. Result and Analysis

Figure 1 shows the experimental data using 4 classes of the image of corn leaf disease, namely Health (b) Gray Leaf Spot (c) Blight (d) Common Rush. The amount of data used is 3500 consisting of

1000 health leaves data, 500 leaves with Gray Leaf Spot disease, 1000 leaves with Blight disease, and 1000 leaves with Common Rush disease. The data has a size of 3000 x 4000 pixels which is then resized to 300 x 300 pixels. The corn leaf image will be converted to grayscale using formula (1). After the image has changed size and changed color to grayscale, the LBP feature extraction process will be carried out. Waiting for LBP feature extraction gives the image a different texture. it is as shown in Figure 5.



(a) Healthy (b) Blight (c) Common Rush (d) Gray leaf Spott

**Figure 5**. Feature extraction results using LBP with a size of 300 x 300

The results of the LBP feature extraction are then classified using several classification methods such as Adaboost, Naïve Bayes, and k-NN. Figure 6 shows a comparison of three classifiers 5-fold, namely Adaboost, Naïve Bayes, and k-NN. The highest value obtained by the k-NN method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.9408, 0.8108, 0.8095, and 0.8108. For the lowest value, the Adaboost method is obtained with AUC values, Accuracy values, F1-Score, Precision, and Recall, namely 0.82, 0.74, 0.74, and 0.74. The values obtained in the Naïve Bayes method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.93, 0.75, 0.75, 0.76, and 0.75

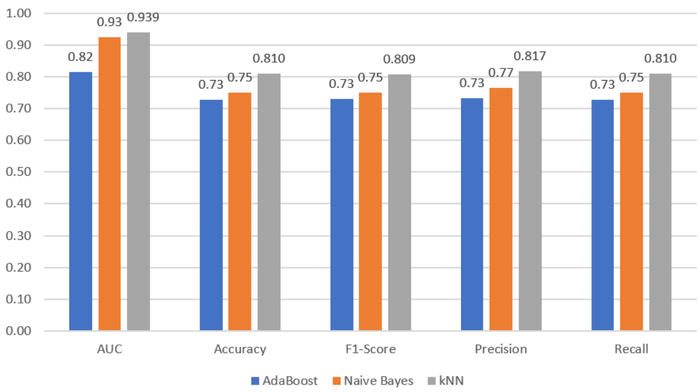



**Figure 6**. Comparison of results of three classification models on maize leaf disease with 5-fold

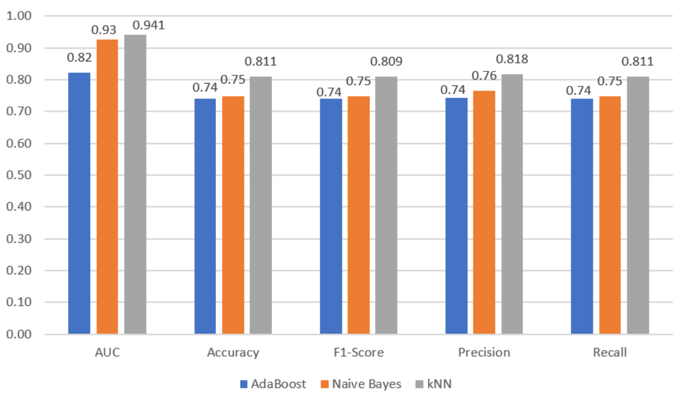



**Figure 7**. Comparison of results of three classification models on corn leaf disease with 10-fold

While at 10-fold, classification using k-NN is still the best classifier of the other two, namely Adaboost and Naïve Bayes. Figure 7 shows a comparison of the three classifiers with 10-fold trials. The highest value obtained by the k-NN method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.941, 0.811, 0.809, 0.818, and 0.811. For the lowest value, the Adaboost method is obtained with AUC values, Accuracy values, F1-Score, Precision, and Recall, namely 0.82, 0.74, 0.74, 74, and 0.74. The values obtained in the Naïve Bayes method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.93, 0.75, 0.75, 0.76, and 0.75.

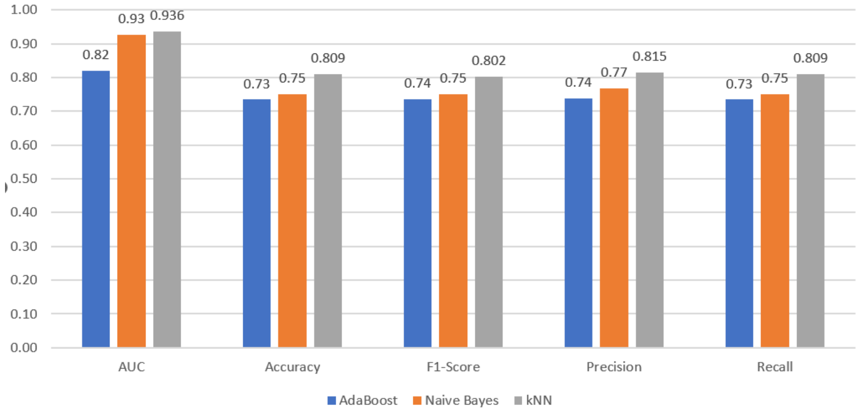



**Figure 8**. Comparison of results of three classification models on corn leaf disease with 20-fold

In the 20-fold trial, the classification using k-NN remains the best classifier of the other two, namely Adaboost and Naïve Bayes. Figure 8 shows a comparison of the three 20-fold classifier trials. The highest value obtained by the k-NN method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.936, 0.809, 0.802, 0.815, and 0.809. For the lowest value, the Adaboost method is obtained with AUC values, Accuracy values, F1-Score, Precision, and Recall, namely 0.82, 0.73, 0.74, 74, and 0.73. The values obtained in the Naïve Bayes method with AUC values, Accuracy values, F1-Score, Precision, and Recall are 0.93, 0.75, 0.75, 0.77, and 0.75. The results of the comparison of three trials using folds 5, 10, and 20 can be seen in Table 2, where the best fold results are in the 10-fold.

**Table 2**. The results of the fold comparison in the k-NN method

| K-Fold | AUC | Accuracy | F1-Score | Precision | Recall |
|---|---|---|---|---|---|
| K=5 | 0.939 | 0.810 | 0.809 | 0.817 | 0.810 |
| **K=10** | **0.941** | **0.811** | **0.809** | **0.818** | **0.811** |
| K=20 | 0.936 | 0.809 | 0.802 | 0.815 | 0.809 |

## 4. Conclusion
From the experiments that have been carried out, it can be concluded that the k-NN method with LBP feature extraction has better accuracy results than the Adaboost and Naïve Bayes methods in classifying corn leaf disease.

Corn leaf disease classification using LBP feature extraction and k-NN method obtained an AUC level of 94.1%, F1-Score of 80.9%, Precision of 81.8%, and recall of 81.1%. The results of the performance of the k-NN method using the parameter k=10 have resulted in an accuracy of 81.1%.

By conducting a trial using 3-fold, namely 5-fold, 10-fold, and 20-fold, the test did not show a significant value.

## 5. Acknowledgment
Thank you to the Institute for Research and Community Service and the Faculty of Engineering, the University of Trunojoyo Madura for allowing researchers to complete this research. We also thank the research team of the Multimedia and Network Laboratory for supporting the success of this research.

This research article is supported by UTM DIPA Fund No. 225/UN46.4.1/ PT.01.03/2022 for Independent Research.

**References**

[1] Kementrian Pertanian. Outlook Jagung 2020: Komoditas Pertanian Subsektor Tanaman Pangan. Pusat Data Dan Sistem Informasi Pertanian, 78, 2020. Retrieved from http://epublikasi.setjen.pertanian.go.id

[2] Putro, S. S., Syakur, M. A., Rochman, E. M. S., & Rachmad, A. 2022 Comparison of backpropagation and ERNN methods in predicting corn production. Commun. Math. Biol. Neurosci., 2022.

[3] Syarief, M., & Setiawan, W. 2020. Convolutional neural network for maize leaf disease image classification. Telkomnika (Telecommunication Computing Electronics and Control), 18(3), 1376–1381. https://doi.org/10.12928/TELKOMNIKA.v18i3.14840.

[4] Chauhan, M. D. 2021. Detection of maize disease using random forest classification algorithm. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(9), 715-720.

[5] Syarief, M., Mukminin, A., Prastiti, N., & Setiawan, W. (2017). Penerapan Metode Naive Bayes Classifier Untuk Deteksi Penyakit pada Tanaman Jagung. Network Engineering Research Operation, 3(1).

[6] Putri, F. D., Ramadhani, K. N., & Yunanto, P. E. 2021. Identifikasi Penyakit Pada Tanaman Apel Menggunakan Local Binary Pattern (lbp) Dan Color Histogram. eProceedings of Engineering, 8(1).

[7] Mulyawan, G., Sari, Y. A., & Furqon, M. T. (2019). Klasifikasi Citra Makanan Menggunakan HSV Color Moment dan Haralick Feature Extraction dengan Naive Bayes Classifier. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X.

[8] Chen, L., Chen, J., Hajimirsadeghi, H., & Mori, G. (2020). Adapting grad-CAM for embedding networks. Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020, 2783–2792.

[9] Mujib, K., Hidayatno, A., and Prakoso, T., 2018. Pengenalan wajah dengan Local Binari Pattern (LBP) dan support vector machine (SVM), Transient, vol. 7, no. 1, pp. 123–130.

[10] Indrawanti, A. S. and Mandyartha, E. P. 2018, Deteksi Limfoblas pada Citra Sel Darah Menggunakan Fitur Geometri dan Local Binary Pattern, JNTETI, vol. 7, no. 4.

[11] Duanita Putri Firlia, Nur Ramadhani Kurniawan,Eko Yunanto Prasti, 2021, "Identifikasi Penyakit pada Daun Apel Menggunakan Local Binary Pattern (LBP) dan Color Histogram", e-Proceeding of Engineering (ISSN:2355-9365) Vol 8 No 1.

[12] Ummah, K., Nafi'iyah, N., & Masruroh, M. (2019). IDENTIFIKASI PENYAKIT DAUN JAGUNG BERDASARKAN TEKSTUR DENGAN K-NN. Musamus Journal of Technology & Information, 2(1), 1-4.

[13] Siregar, M. R., Hidayat, N., & Dewi, R. K. Implementasi Metode Modified K-Nearest Neighbor (MK-NN) untuk Diagnosis Penyakit Tanaman Kentang. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN, 2548, 964X.

[14] Rachmad, A., Chamidah, N., & Rulaningtyas, R. 2020. Mycobacterium tuberculosis identification based on colour feature extraction using expert system. Ann. Biol 36, 196-202.

[15] Rachmad, A., Chamidah, N., & Rulaningtyas, R. 2020. Mycobacterium tuberculosis images classification based on combining of convolutional neural network and support vector machine. Commun. Math. Biol. Neurosci. 2020.

[16] Rachmad, A., Chamidah, N., & Rulaningtyas, R. 2021. Classification of mycobacterium tuberculosis based on color feature extraction using adaptive boosting method. In AIP Conference Proceedings (Vol. 2329, No. 1, p. 050005). AIP Publishing LLC.