

## Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network



Ristohadi S<sup>1</sup>, Andi Sunyoto<sup>2</sup>, Emha Taufiq Luthfi<sup>3</sup>

<sup>1,2,3</sup> Master of Informatics, Universitas Amikom Yogyakarta

**ABSTRACT:** This study aims to identify the condition of corn plants based on imagery leaf using the gray level co-occurrence matrix (GLCM) method and artificial neural network (ANN) backpropagation. The GLCM method is used for extracting features from image leaf corn, whereas ANN backpropagation is used for classification condition plant corn based on features. The classification was done using a dataset of corn leaves with four conditions: healthy, leaf spot, blight, and leaf rust. Next, the leaf features are extracted using method GLCM and training on model ANN backpropagation to classify conditions of corn plants. After training on the model, the next step is model evaluation using the confusion matrix method. The research results show that the technique can produce accuracy, which is tall enough to identify condition corn plants, with an accuracy of 95%. This indicates that the use of GLCM and ANN backpropagation can be a good alternative in determining the condition of corn plants. This research provides benefits in facilitating the identification of the state of corn plants quickly and accurately.

**KEYWORDS:** Leaf Corn, GLCM, ANN, Backpropagation

### I. INTRODUCTION

Corn (*Zea mays ssp. mays*) is one of the carbohydrate-producing food plants the world's most important, besides wheat and rice, seeds corn is food tree for resident America Middle and South, as well as for part resident Africa and part region Indonesia, besides That corn Also become component important feed livestock. Production corn in Indonesia year on 2020 as big 29.02 million tons, provinces with corn production the largest national is East Java with the number production of 23.16 percent of national corn production on year 2020. Opportunities to increase corn production for meet domestic and export needs Still Enough big, program repair corn national level to increase productivity and expand area will held on environment or agroecosystem Which different, starting from a high productivity environment (land optimal) to low productivity environments (land marginal/dry) by Because That, different maize cultivation techniques are required And Specific in a manner ecological. Wrong One constraint moment production corn that is in land dry Which caused by lack water Because very depend on condition bulk rain, besides that, the level of pest attack and disease is factor Which very influence production corn dry, reason main low results corn in Indonesia is use of local varieties, soil fertility bad, fertilization which no adequate, as well as attack pest and disease. For can know plant corn caught disease Wrong only one that is with see condition leaf on plant corn, plant corn Which indicated disease usually No will bear fruit. Still Lots farmer Which need time long to know the condition of the corn plant, this can make plant corn future no grow in accordance hope para farmers.

There are many ways you can do it to get it help farmers to know the conditions plant corn with condition Healthy or caught disease with fast that is Wrong only one with use Machine Learning Which previously proven capable finish topic like this. As example do Classification of leaves that have properties as drug, study This done for can give information to people related to what kind of leaves have properties medicinal and has no medicinal properties, research This use method machine learning. There is a number of studies previously with studies case Which The same like Which done by (Iswantoro & Handayani UN, 2022), this study carried out a classification corn plant diseases using the method convolutional neural network (CNN) results from this research is the CNN algorithm is sufficient Good in do classification, and results testing produce level accuracy in Classification of diseases of maize plants as big 94%. Next study automatic Fuzzy Logic-Based Maize common Rust disease Severity Predictions with Thresholding and Deep Learning done by (Sibiya & Sumbwanyambe, 2021), identification process study This do approach with network VGG-16 produce accuracy testing 89%. Furthermore, study Which conducted by (Kshyanaprava et al., 2020) the

# Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

study conducted maize leaf disease detection and classification using machine learning algorithms results from study This is algorithm random forest (RF) with accuracy highest 79.23% compared to the classification technique other. Study final related with classification pest and disease plant corn use method fuzzy random forest based on resampling repeated k-fold cross validation conducted by (Nearadiaty, 2022), the method used in this study is fuzzy random forest. Results from this research Namely classification using fuzzy method random forest that has been done to obtain model accuracy rate reaches 92%. Study previously with this research, using which The same that is dataset leaf corn.

Study previously average No do normalization data and alteration background rear, it can make the model inside do classification No get results optimal. The aim of this research is improving the results of the accuracy of the identification condition plant corn with propose method gray levels co-occurrence matrix (GLCM), artificial neural network (ANN) backpropagation, and enhance with cover lack Which there is on previous research to be able to improve accuracy results are better than that accuracy generated previously.

## II. RESEARCH METHODS

Draft study This that is do classification of conditions based on corn plants Leaf images use the GLCM method for extraction feature on image the leaves and JST backpropagation for do classification. Following explanation of the steps study This.

### A. Taking Datasets

In this process, the authors search for datasets that according to the research topic, therefore writer fetch public dataset coming from GitHub the image to be converted later become feature numeric.

### B. Pre-processing Data

Process This done for inspect and repair error Which found in the dataset, so it can next on step next. Technique Pre-processing data Which used in study This that is change the image background. This matter done so that the extraction process will focus on the main object only. There is a number of cases in previous research, where the image was not changed beforehand the background behind, so that own brightness or noisy on image Which extracted, thus reducing contrast object and causes the model to not be optimal when committing extraction features.

### C. Extraction Feature GLCM

Process This will extract characteristic to get the feature value in pixels on every image. With use GLCM get feature texture dissimilarity, correlation, homogeneity, contrast, ASM, And energy. Feature Which must There is on method GLCM namely (1) contrast which serves to measure pixel intensity, (2) working correlation measure how much correlate pixels and other angles, (3) functioning energy for measure mark diversity intensity, and (4) homogeneity function to measure the value of the similarity of variations image intensity. Some angles are used in do extraction features in the image are 0, 45, 90, and 135 represent direction Where pixel pairs are counted in the matrix GLCM. Each corner has different meanings and related implications analysis texture image.

### D. Normalization data

Normalization data done for group attribute from something relation so that form relation Which Good. Technique normalization This use equality 7.

$$\text{new\_data} = \frac{\text{data}}{10^7}$$

### E. Splits data

This stage is the division stage dataset become data training and data testing. Data sharing on \_ study This is 80/20 data training and data testing. Amount distribution the aim is to look at the model in predict testing data with a total of 770. In general, the model deep learning gets accuracy results Which Good If own amount data testing Which A little. So, in study this testing data is increased in number and then tested is model get results which is better?

### F. Classification backpropagation & Tuning Parameters \_

Next do the classification with method backpropagation. At stage in this case, identify the condition of the plant corn based on the leaf image.

### G. Evaluation Method

Process this done for evaluate performance of the model that has been designed previously for do identification condition corn plant. When the accuracy results generated does not exceed the result of previous research then back to process classification backpropagation and parameter set up produce accuracy Which Better. But when, accuracy results have exceeded the results from study previously so proceed to the next process viz taking conclusion.

# Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

In this study the evaluation used the confusion method matrix. he is table Which containing calculations based on evaluations model classification based on amount studies case, and predicted Correct or wrong. More clearly can see on table 1.

**Table 1. Confusion Matrix**

<i>Classification</i>	<i>predicted Positive</i>	<i>predicted Negative</i>
<i>actual Positive</i>	<i>true Positive (TP)</i>	<i>False Negative (FN)</i>
<i>actual Negative</i>	<i>False Positive (FP)</i>	<i>true Negative (TN)</i>

In measurement performance using the existing confusion matrix four part for identify something prediction, including:

1. True Positive (TP) is amount data with positive actual values and mark prediction positive.
2. True Negative (TN) is amount data with mark actual positive and mark prediction negative.
3. False Positive (FP) is sum of data with actual value negative and mark prediction positive.
4. False Negatives (FN) is amount of data with actual value negative and mark prediction negative.

Accuracy is mark evaluation Which often worn on classification binary. Can see based on mark confusion matrix: accuracy (acc) is effectiveness of the results obtained in the process classification; accuracy done use equality (8) following.

$$Accuracy (\%) = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

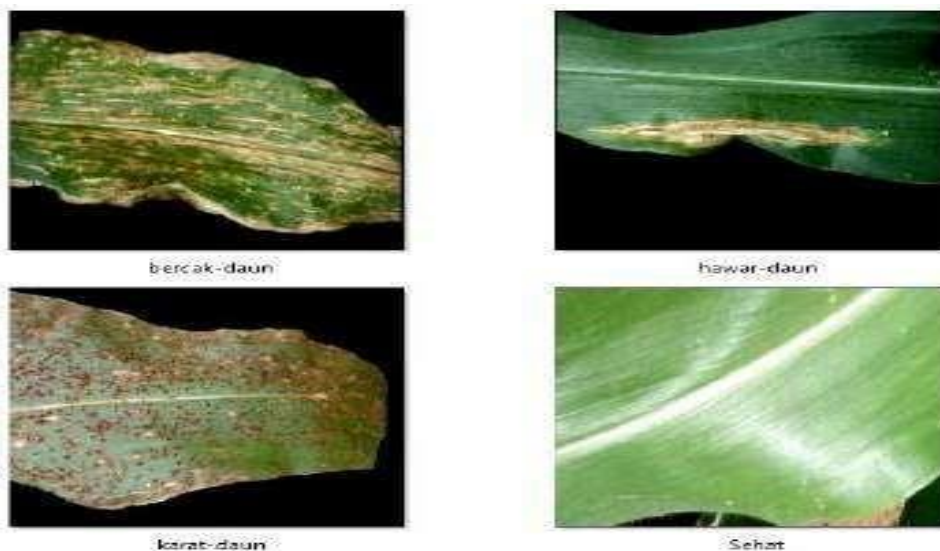
## H. Withdrawal Conclusion

Stage This is stage gift conclusion based on design model applied to previous studies and results testing Which has conducted in this study.

## III. RESULTS AND DISCUSSION

### A. Taking dataset

Study This use dataset leaf corn, dataset This taken from github <https://github.com/ibnujakaria/dataset-leaf-corn>, This dataset is used to train and test model moment do classification specifically on identification condition type leaf. following This is a viewable view of the dataset on figure 1.



**Figure 1. Datasets**

In figure 1 of this dataset there are 4 types picture that is leafspot, blight, rust- leaf, And Healthy. Amount \_ from each respectively picture in dataset they can see on table 1.

**Table 2. Number of classes dataset**

Type Picture	Number of Images
leaf spot	508
blight	985
leaf-rust	1192
Healthy	1162



# Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

Table 2 is the result of the changes leaf type label through one hot encoding process where each type of leaf has a number value different binary. On this type of image example that is leafspot, Blight, Rust-Leaves, and Healthy. Label Encoder results on class dataset using the python program can be seen on figure 3.

```
1 enc = LabelEncoder()

1 Y = glm_df['Label'] = enc.fit_transform(glm_df['Label'].values)
```

Figure 3. Variable categorical

Picture 3 is appearance from attribute Which Still *categorical* the data, those attributes will in change become form *numeric*, can see on figure 4.

```
1 le = LabelEncoder()
2 le.fit(glm_df["Label"].values)
3 print(" categorical label : \n", le.classes_)
4 Y = le.transform(glm_df['Label'].values)
5 Y = to_categorical(Y)
6 print("\n\n one hot encoding for sample 0 : \n", Y[0])

categorical label :
['bercak-daun' 'hawar-daun' 'karat-daun' 'sehat']

one hot encoding for sample 0 :
[1. 0. 0. 0.]
```

Figure 4. Results label encoding

Figure 4 is the result that has been done alteration data *categorical* become *numerical* in *the class* dataset. Example is shown in the image *class* with mark 'Leafspot' so changed become '1.0.0.0'.

### C. Feature Extraction using GLCM

The results of the GLCM process for the image Corn leaves, saved in .csv format previously Already executed use program python. Following This results whole from the results of image feature extraction dataset can see on figure 5.

dissimilarity_0	dissimilarity_45	dissimilarity_90	dissimilarity_135	correlation_0	correlation_45	correlation_90	correlation_135	homogeneity_0	...	contrast_135	ASM_0
24.016524	24.111059	15.746410	22.781114	0.637701	0.668183	0.648255	0.677658	0.673211	...	1067.228136	0.000515
21.273662	28.642381	29.821731	39.239803	0.458419	0.146192	0.085471	0.056667	0.065214	...	1429.982998	0.000478
15.418217	11.117596	15.802449	20.052812	0.515600	0.296119	0.219700	0.204345	0.067116	...	689.550296	0.000554
13.983251	39.397245	39.881721	37.796844	0.807023	-0.206875	-0.228531	-0.172087	0.130513	...	3498.339383	0.000889
13.486328	87.953400	15.488304	17.203613	0.381588	0.072738	0.138328	0.076175	0.104838	...	879.479347	0.000870
19.279699	30.914930	34.618115	11.723208	0.769294	0.448144	0.624452	0.917488	0.988113	...	284.788191	0.000817
0.158731	15.055297	21.025704	68.258245	0.972122	0.619518	0.615125	0.675254	0.153288	...	748.678816	0.000880
13.054847	25.794142	31.326183	31.895260	0.849846	0.466836	0.199707	0.218882	0.082488	...	1573.610765	0.000690
15.142256	12.625717	23.438964	25.123465	0.648584	0.756227	0.195222	0.624072	0.851914	...	1533.365227	0.001248
9.504264	34.983553	39.124847	31.428893	0.922296	0.448797	0.324536	0.553855	0.718025	...	2148.898284	0.000657

Figure 5. The results of the gray level co-occurrence matrix (GLCM)

Figure 5 is results from extraction feature use GLCM, each image in datasets have number of features as much 24 features. because feature texture Which worn namely (1) contrast, (2) correlation, (3) ASM, (4) energy, (5) homogeneity, and (6) dissimilarity. Medium corner Which used in extraction This there are four ie 0°, 45°, 90°, and 135°.

### D. Normalization Data

Results from extraction feature use GLCM has a value with a very large scale big or very different, these things can make model machine \_ learning No optimal in do classification. Because attribute on datasets This own scale which is different. Scale Which different can see on figure 1.

## Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

On features contrast\_135 and homogeneity\_0, therefore necessary done standardization in order to have the same scale when doing process classification. Technique normalization data This share each column on feature dataset with numbers raised to the power of 10. Normalization technique this use equality following. Where are many digits mark largest in each column? Here are the results standardization use technique normalization on dataset.

### E. Splits Data

Stage This is stages share dataset become data training and data testing, distribution on study This shared data training and data testing to 80/20. Amount distribution the aim see model in predict a total of 770 data testing. kindly general model deep learning gets good accuracy results when amount data testing Which A little. On study in this case, the data testing is increased in number, then tested to get better results. Table 3. describes the division of the data in do in this research.

**Table 3. Splits Data**

Information	Data Training	Data testing	Amount
Proportion	80%	20%	100%
Amount	3077	770	3847

On Table 3 explained that distribution data training and data testing is done is to divide data become 80/20, which is 80% For data training which amounted to 3077 data and 20% for total testing data 770 data, meaning the whole data from dataset amount 3847.

### F. Classification Backpropagation

Testing This use method backpropagation Which done for now ability something model in identifying plant conditions corn based on the leaf image. In previous research, the power of the system in identifying types of herbal leaves depend on the process training for produce accuracy in determine type leaf Which used on stage testing parameter backpropagation. A number of Tests carried out previously on the image of herbal plant leaves can be seen on table 4.

**Table 4. Testing backpropagation**

No	Neurons Inputs	Hidden Neurons	Neuron's output	Epoch	Batches size	Acc.
1	24	34, 150, 300, 200	4	1000	22	90%
2	24	34, 150, 300, 200	4	2000	22	89%
3	24	42, 150, 300, 200	4	1000	22	87%
4	24	24, 150, 300, 200	4	1000	22	88%
5	24	24, 150, 300, 200,250	4	1000	22	95%
6	24	24, 150, 300, 200.50	4	1000	22	93%

On table 4 parameter Which used is *inputs neurons layers, hidden layers neurons, output layers neurons, epoch, And batch size*. Test This consists of six classification trials, every test done with set different parameters, starting with enhancement mark *hidden neurons* on *layer* and decrease the value of *hidden neurons* on *layers* each parameter. Change *epoch* and *batch sizes* Also tested to get the right combination that produce high accuracy values in the classification process backpropagation. Experiment fifth own accuracy highest, that is 95%, with architecture *neurons inputs 24, concealed neurons 24, 150, 300, 200,250, neurons output, epoch 1000 And batch size 22*. Mark in euros *i input 24* Because according to the number of features dataset processing. *input*

# Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

*neurons* This will bring data enter inside \_ system for later processing on the layer furthermore. There are 5 *hidden layers*, where n is the value *hidden layer* is the randomization value of each test Which done. Randomization \_ increase neurons and decline neurons for sent to epoch as Wrong One stage for get an evaluation value. *Output layers* consists of 4 neurons to be equal to dimensions 4 *categorical columns* on *y\_train* and *y\_test*. Made to equalize the number of classes in the dataset i.e amount 4 class. Furthermore *epoch*, is amount *complete passed* Which must done on dataset *training*, the *epoch* value is set to 1000 because on previous trials *epoch* not enough from 1000, or try to raise but not produce accuracy Which Better. So, *batch sizes* is a number sample Which processed before the model is *updated*. Results accuracy of all experiment's classification is done, can be seen on picture 6.

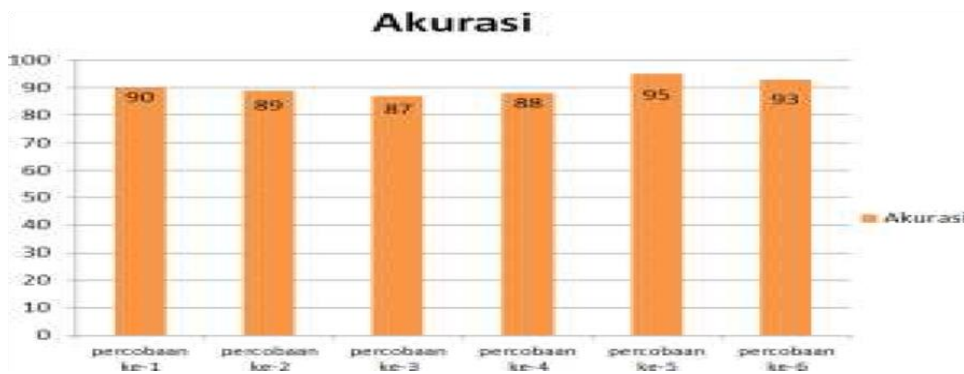


Figure 6. Results accuracy Test

## G. Evaluation Confusion Matrix

Study This do 6 testing with use method classification JST backpropagation. The experiment that produces the best accuracy is the fifth try. Fifth experiment use inputs layers 24, own 5 hidden layers with each neuron, i.e., 24, 150, 300, 200,250, output layer 4, epoch 1000, And batch sizes 22. Confusion matrix from test fifth showed on Picture 7.

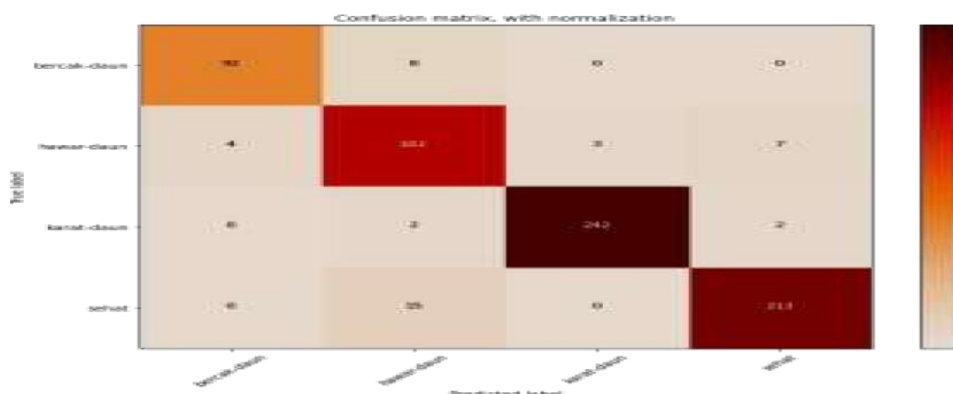


Figure 7. Results confusion matrix

Figure 7 is the result of the evaluation of all classes use the confusion matrix. As much eight data Which should category leaf spot identified Wrong namely as leaf blight. Then four, three, and seven that should be categories blight- leaf but Wrong identification that is four identified as leafspot, three as rust-leaf, and seven as healthy. Also, there is error identification which should be rust-leaf but as much as 2 data identified as blight leaf, And 2 as Healthy. last\_ happened 15 Wrong identification Which should Healthy but identified as leaf blight.

Application of ANN backpropagation with some parameters on the corn leaf dataset get results best with mark accuracy of 95% on the fifth try. The results can be seen in figure 7. accuracy This is ratio prediction Which appropriate in identify condition plant corn in a manner whole in dataset. Amount data Which predicted Wrong is 41 data, and total data Which expected correct is 729 data of the total whole data test Which amount 770 data.

## V. CONCLUSIONS

Based on the analysis performed with use dataset leaf corn Which own 4 class leaf, can concluded that:

- In this study there were six tests. Results the best is in the 5th test with accuracy 95%. A curation best This happen enhancement 1% from results accuracy Which generated in previous studies that is 94%.

## Identification of Condition of Corn Plant Based on Leaf Image Features Using Gray Level Co-Occurrence Matrix and Backpropagation Neural Network

- There is a number of stages process study Which help find condition model Which appropriate so that produce accuracy Which Good moment process classification done namely: (1) pre-processing data, (2) extraction features, (3) normalization data, (4) split data, (5) classification, (6) evaluation use confusion matrix.

### REFERENCES

- 1) K. Agriculture, "Analysis Performance Trading Corn," Pussy. Data and System. inf. Question. Secretary Ministry General. Question. 2021, pp. 5–24, 2021.
- 2) MA Suparlan, Nurali, Edi Wati, "CONTROL INTEGRATED PEST MAIN PLANT CORN (Zea Mays, L) IN LAND DRY," 2022.
- 3) Koesrini, "Technology Cultivation Corn in Land Swamp", 2016.
- 4) F. Afandi, "Disease Corn and Method get over it, 2019.
- 5) Agriculture, "Cultivation of Corn," 2022.
- 6) P Semitera, "Abstract Seminar National Technology Applied (SEMITERA) 2021," Monday. Technol. therapy. 2021.
- 7) K. Saputra S and M. I. Wind-War, "Classification Plant Drug Based on Extraction Feature Morphology Leaves Using Artificial Neural Networks," J. Inform, vol. 5, no. 2, pp. 169–174, 2018, doi: 10.31311/ji. v5i2.3770.
- 8) D. Iswantoro and D. Handayani UN, "Classification Disease Plant Corn Use Method Convolutional neural network (CNN)," J. Science. Univ. Batanghari Jambi, vol. 22, no. 2, p.s. 900, 2022.
- 9) M. Sibya and M. Sumbwanyambe, "Automatic fuzzy logic-based cornstarch common rust disease severity predictions with thresholding and deep learning," Pathogens, vol. 10, no. 2, pp. 1–17, 2021, doi: 10.3390/pathogens10020131.
- 10) A K. S. & S. C. M. Kshyanaprava Panda Panigrahi, Himansu Das, "Maize Leaf disease detection and Classification using Machine Learning Algorithms," Prog. Comput. Anal. netw. , vol. volume 111, 2020.
- 11) A NEARDIATY, "Classification Pest and Disease Corn Plant Using Fuzzy Random Method Forest Based on Resampling Repeated K-Fold Cross validation," 2022.
- 12) R. Widodo, A W. Widodo, and A Supriyanto, "Utilization Characteristic Gray Levels Co-Occurrence Matrix (GLCM) Image Fruit Orange tangerine (Citrus reticulata Blanco) for Quality Classification," J. Developer. Technol. inf. and Computing Science. , vol. 2, no. 11, p. 5769–5776, 2018.
- 13) H. Wibawanto, A. Susanto, TS Widodo, and BC Tjokronegoro, "Identification of Cystic Mass Image Based Feature GrayLevel Co-Occurrence Matrix," Monday. Nas. app. Technol. inf. , vol. 2008, no. Snati, p.s. 55281, 2008.
- 14) WI Praseptiyana, AW Widodo, and MA Rahman, "Utilization Characteristic Gray Levels Co-occurrence Matrix (GLCM) for Melasma Detection in Facial Images," J. Development Technol. inf. and Computing Science. , vol. 3, no. 11, pp. 10402–10409, 2019.
- 15) Nunik Rahmawati K, "Algorithm of Artificial Neural Networks with Method Learning Vector Quantization Nunik Rahmawati K STUDY PROGRAM TECHNIQUE INFORMATICS," 2007.
- 16) EPC Edi Ismanto, "Algorithmic Neural Networks Backpropagation In predict Availability Riau Province Food Commodities," Rabbit J. Technol. And Sist. inf. Univrab , vol. 2, no. 2, pp. 196–209, 2017, doi: 10.36341/rabit.v2i2.152.
- 17) F. Marpaung et al., COMPUTER VISION AND DIGITAL IMAGE PROCESSING. 2022.
- 18) adi ultimate, "Textures Analysis – Gray-Level Co- Occurrence Matrix (GLCM) – GUI matlab," 2017.
- 19) L. Veronica, I. H. Al Amen, B. Hartono, and Q Kristianto, "Extraction Feature Texture Use GLCM Matrix on Images with Variations in Object Direction," Pros. SENDI\_U , pp. 978–979, 2019.
- 20) S. Goddess, "Comparison 5 Method Algorithm Classification Data Mining," J. Techno Nusa Mandiri , vol. XIII, no. 1, pp. 60–66, 2016.
- 21) IW Saputro and BW Sari, "Algorithmic Performance Test Naïve Bayes For predictions Period Studies Student," Creat. inf. Technol. J. , vol. 6, no. 1, p.s. 1, 2020.
- 22) Kusuma wahannggar, "Classification Image Cancer Skin Melanoma Use Method support Vector Machine ( Svm ) Image Classification of Skin Cancer Melanoma using support Vector Machine ( Svm ) Method," 2016.



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.