

PENGARUH REDUKSI DIMENSI PADA *CLUSTERING* CITRA DAUN TOMAT MENGGUNAKAN *DARKNET19* DAN *K-MEANS*

EFFECT OF DIMENSIONALITY REDUCTION ON TOMATO LEAF IMAGE CLUSTERING USING DARKNET19 AND K-MEANS

Aeri Rachmad¹⁾, Muhammad Holilur Rohman²⁾, Wahyudi Setiawan³⁾

^{1), 2), 3)} Program Studi Sistem Informasi, Fakultas Teknik, Universitas Trunojoyo
Jl Raya Telang, PO. BOX 2, Kamal, Bangkalan
Email : aery_r@trunojoyo.ac.id¹⁾

Abstrak

Image Clustering adalah pengelompokan citra dari kelas tanpa label sebelumnya. Pada penelitian ini menggunakan citra *clustering* dengan data daun tomat leaf panthogen sebanyak 900 citra yang terdiri dari 3 cluster yaitu *Bacterial Spot*, *Yellow Leaf Curl Virus* dan *Healthy*. Pada setiap cluster memiliki 300 citra. Langkah awal yang dilakukan adalah ekstraksi fitur menggunakan *DarkNet19*. *DarkNet19* menerapkan beberapa parameter seperti epoch sebanyak 160 dengan menggunakan *Stochastic Gradient Descent*, *Learning Rate* dimulai dari 0.01, *weight decay* 0.0005, dan augmentasi data termasuk *random crops*, *rotations*, *hue*, *saturation*, and *exposure shifts*. Selain ekstraksi fitur, pada penelitian ini juga melakukan pengurangan dimensi menggunakan *Principal Component Analysis (PCA)*. Selanjutnya, proses *clustering* menggunakan *K-Means Clustering*. Penelitian ini bertujuan untuk mengetahui tingkat akurasi dalam proses *Clustering* penyakit daun tomat menggunakan metode *CNN DarkNet19*, *PCA* dan *K-Means*. Hasil klastering yang terbaik menggunakan *K-Means Clustering* dengan *PCA 20* yang menghasilkan *accuracy* 97.7%, *precision* 97.73%, dan *recall* 97.67% dengan waktu komputasi 1 menit 16 detik.

Kata Kunci: *Image Clustering*, *Penyakit Daun Tomat*, *DarkNet19*, *Principal Component Analysis*, *K-Means Clustering*.

Abstract

Image Clustering is a grouping of images from classes without previous labels. In this study using image clustering with tomato leaf panthogen data as many as 900 images consisting of 3 clusters namely *Bacterial Spot*, *Yellow Leaf Curl Virus* and *Healthy*. Each cluster has 300 images. The initial step taken is feature extraction using *DarkNet19*. *DarkNet19* applies several parameters such as 160 epochs using *Stochastic Gradient Descent*, *Learning Rate* starting from 0.01, *weight decay* 0.0005, and data augmentation including *random crops*, *rotations*, *hue*, *saturation*, and *exposure shifts*. In addition to feature extraction, this research also reduces dimensions using *Principal Component Analysis (PCA)*. Next, the clustering process uses *K-Means Clustering*. This study aims to determine the level of accuracy in the *Clustering* process of tomato leaf disease using *CNN DarkNet19*, *PCA* and *K-Means* methods. The best clustering results use *K-Means Clustering* with *PCA 20* which produces 97.7% accuracy, 97.73% precision, and 97.67% recall with a computation time of 1 minute 16 seconds.

Keywords: *Image Clustering*, *Tomato Leaf Disease*, *DarkNet19*, *Principal Component Analysis*, *K-Means Clustering*.

1. PENDAHULUAN

Secara geologis, Indonesia merupakan wilayah dengan tingkat kekayaan daratan terbesar di dunia, hal ini dikarenakan Indonesia memiliki keseimbangan musim yang cukup baik, terutama di sekitar daerah khatulistiwa yang memiliki musim berangin dan musim kemarau. Dengan berada di daerah pusat, Indonesia memiliki curah hujan yang tinggi dan musim kemarau yang hampir menyesuaikan, sehingga menciptakan lahan subur yang dapat dimanfaatkan untuk pertanian dan non-pertanian serta aset normal yang melimpah baik dari segi organik maupun awal makhluk hidup. Hal ini juga diperparah dengan kuatnya lapangan pekerjaan individu di Indonesia yang sebagian besar hampir 70% menangani lahan pedesaan [1].

Salah satu produk pertanian yaitu tomat. Yang merupakan salah satu jenis tanaman budidaya di Indonesia. Buah tomat merupakan kebutuhan manusia yang memiliki sumber vitamin A dan juga memiliki vitamin C yang begitu baik [2]. Selain itu dapat dikonsumsi dan diolah dalam

bentuk sayuran segar, juga bisa dijadikan sebagai bahan-bahan pangan lain seperti: selai, sambal, saus dan buah kaleng [3].

Dari data Kementerian Pertanian tentang produksi tomat pada tahun 2015 produksi tomat sebesar 877.792 Ton, lalu pada tahun 2016 sebesar 883.233 Ton, selanjutnya pada tahun 2017 sebesar 962.845 Ton, sedangkan pada tahun 2018 sebesar 976.772 Ton, dan pada tahun 2019 sebesar 1.020.333 Ton, dengan persentase peningkatan produksi tomat sebesar 4,46 % [4].

Seiring dengan meningkatnya populasi dan meningkatnya perhatian publik terhadap pentingnya kesejahteraan, minat terhadap tomat diperkirakan akan terus meningkat. Selama ini pengembangan produksi tomat masih dipengaruhi oleh faktor infeksi tanaman [2].

Gejala penyakit ini dapat dilihat dari perubahan daunnya. Adapun penyakit pada daun tomat diantaranya *a. bacterial spot*, *b. yellow leaf curl virus* dan *c. healthy* seperti yang di tampilkan pada Gambar 1.



Gambar 1. Citra Daun Tomat

Oleh karena itu, pada penelitian ini dirancang sebuah sistem Clustering penyakit daun tomat menggunakan metode *Convolution Neural Network (CNN)* yaitu DarkNet19 untuk ekstraksi fitur, PCA untuk reduksi dimensi, serta K-Means Clustering untuk menghasilkan clustering yang tepat.

2. DASAR TEORI

2.1 *Convolution Neural Network (CNN)*

Convolution Neural Network (CNN) adalah salah satu algoritma deep learning yang dikembangkan oleh *Multilayer Perceptron* dan dirancang untuk mengolah data dalam dua dimensi seperti image atau audio. CNN digunakan untuk mengklasifikasikan data menggunakan metode pembelajaran terawasi. *Supervised learning* bekerja sedemikian rupa sehingga ada data yang dilatih dan ada variabel yang menjadi sasaran sehingga tujuan dari metode ini adalah mengelompokkan data dengan data yang ada. CNN dapat memiliki puluhan atau ratusan lapisan, masing-masing mengenali image

Tahapan pada CNN adalah *convolution layer*, *Padding*, *aktivasi ReLu (Rectifier Linear Unit)*, dan *Pooling*[5].

2.2 DarkNet19

DarkNet19 adalah arsitektur di CNN yang memiliki total 64 layer. Layer tersebut terdiri dari 19 *Convolution layer*, 18 *Batch Normalization (BN)*, 18 *Leaky ReLU*, 5 *max-pooling*, 1 *global-average-pooling*, dan 1 *softmax*, serta *input dan output layer*.

Metode DarkNet19 menerapkan beberapa parameter seperti epoch sebanyak 160 dengan menggunakan *stochastic gradient descent*, *learning rate dimulai dari 0.01*, *weight decay 0.0005*, dan augmentasi data termasuk *random crops*, *rotations*, *hue*, *saturation*, dan *exposure shifts*[6].

2.3 PCA (Principal Component Analysis)

PCA adalah pendekatan statistik untuk menemukan fitur utama dari dataset terdistribusi berdasarkan total varian[7]. Hasil dari fitur DarkNet19 memiliki dimensi yang tinggi, maka di perlukan untuk mengurangi dimensi. Pengurangan dimensi menggunakan PCA juga memunculkan fitur ekstraksi baru yang digunakan untuk proses selanjutnya[8]. Langkah-langkah pengurangan dimensi menggunakan PCA adalah[9]:

a. Memasukan image dalam sebuah matriks

Langkah pertama proses pengenalan pada citra daun tomat adalah dengan input data image. Misal ada sebuah image dengan ukuran 100 x 100 piksel = 10.000. Kemudian matriks tersebut diubah menjadi 10.000 x m.

$$u = \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,m} \\ u_{2,1} & u_{2,2} & \dots & u_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ u_{10000,1} & u_{10000,2} & \dots & u_{10000,m} \end{bmatrix} \tag{1}$$

b. Mencari rata-rata *image*

Langkah berikutnya setelah pengolahan citra awal adalah mengambil citra rata-rata dari seluruh citra yang ada di database. Rumus untuk mencari rata-rata adalah sebagai berikut:

$$\bar{u} = \frac{1}{M} \sum_{n=1}^M u_{i,k} \tag{2}$$

Image rata-rata adalah jumlah $u_{i,k}$ dibagi dengan M merupakan jumlah image dalam database. Selain itu, data di atas memberikan matriks dengan ukuran 10.000 x 1.

c. Mencari kovarian matriks PCA

Matriks kovarian terdiri dari semua kemungkinan variasi yang diperoleh dari sepasang vektor kolom. Mencari matriks kovarian dilakukan dengan mengalikan u dengan transposisi. Seperti pada Rumus persamaan 2 sebagai berikut:

$$C = u x \bar{u} \tag{3}$$

d. Kemudian dilakukan pencarian nilai eigen menggunakan rumus (4):

$$C x v = \lambda x v \tag{4}$$

λ adalah matriks eigen value dengan v adalah matriks eigen vector. λ dan v merupakan matriks berdimensi n x n dimana n merupakan jumlah piksel image. Setelah Eigen value diperoleh maka nilai tersebut diurutkan dari yang terbesar ke yang terkecil. Hasil dari operasi matriks menghasilkan matriks V berukuran 10.000 x 10.000.

e. Mencari fitur PCA

Fitur PCA adalah informasi terpenting dalam image yang sebelumnya dicari nilai eigen vektornya. Fitur PCA dapat ditemukan dengan mentransformasikan citra asli ke ruang citra daun tomat menggunakan persamaan 4 sebagai berikut :

$$f = \sum_{i=1}^m (I_i - \bar{u})^T x V \tag{5}$$

I adalah data tiap piksel dari setiap training I, m merupakan jumlah image training, V adalah matriks eigen vector dan f merupakan matriks berukuran m x 10.000.

2.4 K-Means Clustering

Pengelompokan data adalah salah satu metode *data mining* yang *unsupervised*. Dua jenis pengelompokan data yang umum digunakan dalam pengelompokan data, yaitu pengelompokan data secara hierarkis dan pengelompokan data non-hierarkis[10]. K-Means adalah metode pengelompokan data non-hierarkis yang mencoba mengklasifikasikan data yang ada ke dalam satu atau beberapa cluster/grup. Metode ini mengelompokkan data ke dalam klaster sehingga data dengan karakteristik yang sama dikelompokkan dalam klaster yang sama dan data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok yang berbeda.

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut[11]:

- a) Tentukan jumlah cluster yang akan di Set
- b) Alokasikan data ke dalam *cluster* secara *random*
- c) Hitung *centroid* dari data yang ada pada tiap *cluster*
- d) Alokasikan masing-masing data ke *centroid* terdekat
- e) Ulangi ke langkah-3, Jika masih ada data yang dipindahkan ke *cluster*, atau jika nilai *centroid* berubah, berarti ada sesuatu yang berada di atas ambang batas yang ditentukan, atau jika nilai fungsi tujuan yang digunakan berubah di atas ambang batas yang ditentukan.

2.5 Performance Measure

Performance Measure merupakan suatu metode untuk melakukan perhitungan akurasi pada Image Clustering [12]. Untuk perhitungan akurasinya menggunakan *Multiclass Confusion*[13]. *Matrix*Tabel untuk contoh *Multiclass Confusion Matrix* dapat dilihat pada Tabel 1.

Tabel 1. *Confusion Matrix*

<i>Predicted Value</i>	<i>Actual Value</i>	
	<i>True</i>	<i>False</i>
<i>True</i>	<i>TP (True Positive)</i>	<i>FP (False Positive)</i>
<i>False</i>	<i>FN (False Negative)</i>	<i>TN (True Negative)</i>

Ada beberapa metode untuk mengukur performance pada K-Means Clustering, yaitu dengan mencari nilai *Accuracy*, *Precision* dan *Recall* [14][15].

Rumus untuk menghitung akurasi dapat dilihat pada persamaan 6.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

Rumus untuk menghitung precission dapat dilihat pada persamaan 7.

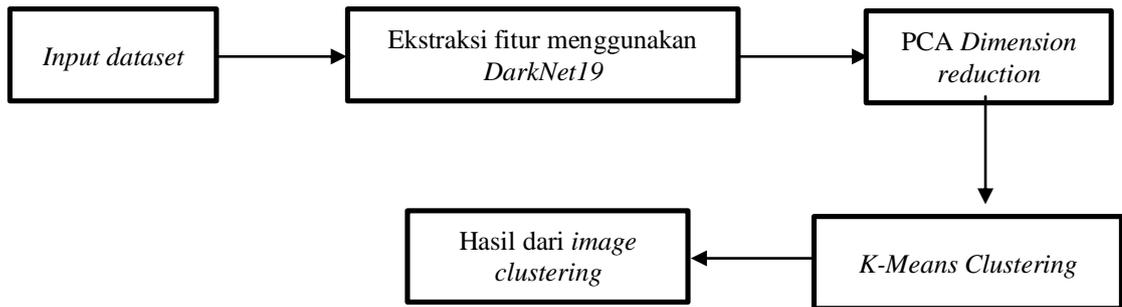
$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Rumus untuk menghitung recall dapat dilihat pada persamaan 8.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

3. METODOLOGI PENELITIAN

Pada penelitian ini sistem melakukan proses ekstraksi fitur, reduksi dimensi, dan pengelompokan. Metode yang digunakan termasuk DarkNet19 untuk ekstraksi fitur, PCA untuk dimension reduction, dan K-Means untuk pengelompokan. Gambar 2 menunjukkan diagram blok sistem sedang dibangun.



Gambar 2. Diagram blok sistem *Clustering* citra daun tomat

4. PENGUJIAN DAN PEMBAHASAN

Secara umum pelaksanaan penelitian ini menggunakan perangkat keras berupa laptop dengan sistem operasi Windows 10 Home Single Language. Detail spesifikasi laptop ini antara lain: Prosesor Intel(R) Core (TM) i5 6200U CPU @2.30GHz (4 CPUs), ~2.4 GHz, RAM 8 GB. Adapun perangkat lunak yang digunakan dalam penelitian ini adalah MATLAB R2020a.

Uji coba pada penelitian ini menggunakan 900 data citra sampel yang terdiri dari 3 kelas dengan ukuran 256 x 256 piksel. Untuk kelas pertama merupakan kelas Bacterial Spot yang terdiri dari 300 citra. Kelas kedua merupakan kelas Yellow Leaf Curl Virus yang terdiri dari 300 citra. Dan untuk kelas yang ketiga merupakan kelas Healthy yang terdiri dari 300 citra. Contoh dataset dapat dilihat pada Gambar 3.



a. Bacterial spot b. Yellow Leaf Curl Virus c. Healthy

Gambar 3. Contoh Citra Daun Tomat

Dalam proses uji coba dilakukan 4 skenario. Skenario ke-1 tanpa menggunakan PCA, scenario ke-2 menggunakan PCA 10, Skenario ke-3 menggunakan PCA 20, Skenario ke-4 menggunakan PCA 30.

Confusion Matrix

Output Class	Bacterial _s pot	280 31.1%	0 0.0%	0 0.0%	100% 0.0%
	Healthy	14 1.6%	300 33.3%	1 0.1%	95.2% 4.8%
	Yellow _L eaf _C url _V irus	6 0.7%	0 0.0%	299 33.2%	98.0% 2.0%
		93.3% 6.7%	100% 0.0%	99.7% 0.3%	97.7% 2.3%
		Bacterial _s pot	Healthy	Yellow _L eaf _C url _V irus	
		Target Class			

Gambar 4. Hasil Skenario ke-1

Pada Skenario ke-1 confusion matrix menjelaskan presentase accuracy 97.7%, precision untuk kelas Bacterial Spot 100%, Healthy 95.2%, Yellow Leaf Curl Virus 98.0%, dan recall untuk kelas Bacterial Spot 93.3%, Healthy 100%, Yellow Leaf Curl Virus 99.7%.

Confusion Matrix

Output Class	Bacterial _s pot	278 30.9%	0 0.0%	0 0.0%	100% 0.0%
	Healthy	16 1.8%	300 33.3%	3 0.3%	94.0% 6.0%
	Yellow _L eaf _C url _V irus	6 0.7%	0 0.0%	297 33.0%	98.0% 2.0%
		92.7% 7.3%	100% 0.0%	99.0% 1.0%	97.2% 2.8%
		Bacterial _s pot	Healthy	Yellow _L eaf _C url _V irus	
		Target Class			

Gambar 5. Hasil Skenario ke-2

Pada skenario ke-2 confusion matrix menjelaskan presentase accuracy 97.2%, precision untuk kelas Bacterial Spot 100%, Healthy 94.0%, Yellow Leaf Curl Virus 98.0%, dan recall untuk kelas Bacterial Spot 92.7%, Healthy 100%, Yellow Leaf Curl Virus 99.0%.

Confusion Matrix

Output Class	Bacterial _s pot	280 31.1%	0 0.0%	0 0.0%	100% 0.0%
	Healthy	14 1.6%	300 33.3%	1 0.1%	95.2% 4.8%
	Yellow _L eaf _C url _V irus	6 0.7%	0 0.0%	299 33.2%	98.0% 2.0%
		93.3% 6.7%	100% 0.0%	99.7% 0.3%	97.7% 2.3%
		Bacterial _s pot	Healthy	Yellow _L eaf _C url _V irus	
		Target Class			

Gambar 6. Hasil Skenario ke-3

Pada skenario ke-3 confusion matrix menjelaskan presentase accuracy 97.2%, precision untuk kelas Bacterial Spot 100%, Healthy 94.0%, Yellow Leaf Curl Virus 98.0%, dan recall untuk kelas Bacterial Spot 92.7%, Healthy 100%, Yellow Leaf Curl Virus 99.0%.

Confusion Matrix

Output Class	Bacterial _S pot	280 31.1%	0 0.0%	0 0.0%	100% 0.0%
	Healthy	14 1.6%	300 33.3%	1 0.1%	95.2% 4.8%
	Yellow _L eaf _C url _V irus	6 0.7%	0 0.0%	299 33.2%	98.0% 2.0%
		93.3% 6.7%	100% 0.0%	99.7% 0.3%	97.7% 2.3%
		Bacterial_Spot	Healthy	Yellow_Leaf_Curl_Virus	
		Target Class			

Gambar 7. Hasil Skenario ke-4

Pada skenario ke-4 confusion matrix menjelaskan presentase accuracy 97.7%, precision untuk kelas Bacterial Spot 100%, Healthy 95.2%, Yellow Leaf Curl Virus 98.0%, dan recall untuk kelas Bacterial Spot 93.3%, Healthy 100%, Yellow Leaf Curl Virus 99.7%.

Tabel 2. Hasil Ujicoba Sistem

Uji Coba	Uji Coba Clustering			
	Accuracy (%)	Precision (%)	Recall (%)	Waktu
Tanpa menggunakan PCA	97.7	97.73	97.67	1 menit 33 detik
Menggunakan PCA 10	97.2	97.33	97.23	1 menit 14 detik
Menggunakan PCA 20	97.7	97.73	97.67	1 menit 16 detik
Menggunakan PCA 30	97.7	97.73	97.67	1 menit 21 detik

Seperti terlihat pada Tabel 2, hasil pengujian pada *scenario* ke-1 yaitu tanpa menggunakan PCA memiliki nilai *Accuracy* sebesar 97.7%, *Precision* 97.73%, *Recall* 97.67% dan waktu komputasi 1 menit 33 detik. hasil pengujian pada *scenario* ke-2 yaitu menggunakan PCA 10 memiliki nilai *Accuracy* sebesar 97.2%, *Precision* 97.33%, *Recall* 97.23% dan waktu komputasi 1 menit 14 detik. Hasil pengujian pada *scenario* ke-3 yaitu menggunakan PCA 20

memiliki nilai *Accuracy* sebesar 97.7%, *Precision* 97.73%, *Recall* 97.67% dan waktu komputasi 1 menit 16 detik. Hasil pengujian pada *scenario* ke-4 yaitu menggunakan PCA 30 memiliki nilai *Accuracy* sebesar 97.7%, *Precision* 97.73%, *Recall* 97.67% dan waktu komputasi 1 menit 21 detik.

5. KESIMPULAN

Hasil dari 4 skenario diatas menunjukkan bahwa *scenario* terbaik terdapat pada *scenario* yang menggunakan PCA 20. Dimana hasil *Accuracy* sebesar 97.7%, *Precision* 97.73%, *Recall* 97.67% dan waktu komputasi 1 menit 16 detik. penggunaan PCA pada metode DarkNet19 bisa mempengaruhi tingkat akurasi dan waktu komputasi pada saat proses K-Means *Clustering*.

DAFTAR PUSTAKA

- [1] S. Maulana, “Septian Maulana, 2018 Zonasi kesesuaian agroklimat untuk menentukan wilayah potensial pengembangan tanaman tomat (*LYCOPERSICUM ESCULENTUM* MILL) di Kabupaten Garut Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu,” no. 1, pp. 43–59, 2018.
- [2] M. Astiningrum, P. P. Arhandi, and N. A. Ariditya, “Identifikasi Penyakit Pada Daun Tomat Berdasarkan Fitur Warna Dan Tekstur,” *J. Inform. Polinema*, vol. 6, no. 2, pp. 47–50, 2020, doi: 10.33795/jip.v6i2.320.
- [3] W. O. Sahera, L. Sabaruddin, and L. O. Safuan, “Pertumbuhan dan produksi tomat (*Lycopersicum esculentuma* Mill) pada berbagai dosis bokhasi kotoran sapi dan jarak tanam,” *J. Berk. Penelit. Agron.*, vol. 1, no. 2, pp. 102–106, 2012.
- [4] Kementerian Pertanian Indonesia, “Produksi Tomat menurut Provinsi, 2015 - 2019,” vol. 2019, p. 1, 2020, [Online]. Available: <https://www.pertanian.go.id/home/?show=page&act=view&id=61>
- [5] A. Rachmad, M. Fuad, & E. M. S. Rochman, (2023). Convolutional Neural Network-Based Classification Model of Corn Leaf Disease. *Mathematical Modelling of Engineering Problems*, 10(2).
- [6] J. Redmon and A. Farhadi, “YOLO9000: Better, faster, stronger,” *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 6517–6525, 2017, doi: 10.1109/CVPR.2017.690.
- [7] S. C. Ng, “Principal component analysis to reduce dimension on digital image,” *Procedia Comput. Sci.*, vol. 111, no. 2015, pp. 113–119, 2017, doi: 10.1016/j.procs.2017.06.017.
- [8] A. Rachmad, R. K. Hapsari, W. Setiawan, T. Indriyani, E. M. S. Rochman, & B. D. Satoto. (2023, May). Classification of Tobacco Leaf Quality Using Feature Extraction of Gray Level Co-occurrence Matrix (GLCM) and K-Nearest Neighbor (K-NN). In *1st International Conference on Neural Networks and Machine Learning 2022 (ICONNSMAL 2022)* (pp. 30-38). Atlantis Press.
- [9] A. Suryadi, “Sistem Pengenalan Wajah Menggunakan Metode Principal Component Analysis (PCA) Dengan Algoritma Fuzzy C-Means (FCM),” *J. Pendidik. Mat.*, vol. 4, no. 2, pp. 58–65, 2015, [Online]. Available: http://e-mosharafa.org/index.php/mosharafa/article/view/mv4n2_2/194
- [10] Rochman, E. M. S., Herawati, S., Khozaimi, A., Indriyani, E., Khatimah, B. K., & Rachmad, A. (2023, January). Clustering tourism places in Madura based facilities using fuzzy C-means. In *AIP Conference Proceedings* (Vol. 2679, No. 1). AIP Publishing.
- [11] Y. Agusta, “K-Means – Penerapan, Permasalahan dan Metode Terkait,” *J. Sist. dan Inform.*, vol. 3, no. Februari, pp. 47–60, 2007.
- [12] A. Rachmad, M. Fuad, & E. M. S. Rochman, (2023). Convolutional Neural Network-Based Classification Model of Corn Leaf Disease. *Mathematical*

Modelling of Engineering Problems, Vol 10, No. 2. doi: <https://doi.org/10.18280/mmep.100220>

- [13] A. Rachmad, M. Syarief, S. Rifka, F. Sonata, W. Setiawan, E.M.S. Rochman, Corn leaf disease classification using local binary patterns (LBP) feature extraction. In *Journal of Physics: Conference Series*, 2406(1): 012020. 2022, <https://doi.org/10.1088/1742-6596/2406/1/01202>
- [14] M. Syarief, A. Mukminin, N. Prastiti, & W. Setiawan. (2017). Penerapan Metode Naive Bayes Classifier Untuk Deteksi Penyakit pada Tanaman Jagung. *Network Engineering Research Operation*, 3(1). <http://dx.doi.org/10.21107/nero.v3i1.75>
- [15] E. M. S. Rochman, A. Khozaimi, I. O. Suzanti, R. Jannah, Khotimah, , & Rachmad, A. (2022). A combination of algorithm agglomerative hierarchical cluster (AHC) and K-means for clustering tourism in Madura-Indonesia. *J. Math. Comput. Sci.*, 12.