

Klasifikasi Jenis Penyakit Tomat Berdasarkan Daun Menggunakan Convolutional Neural Network dan Transfer Learning

Muhammad Alif Xandri¹, Murinto²✉
^{1,2}Universitas Ahmad Dahlan, Indonesia

✉Corresponding Author: murintokusno@tif.uad.ac.id

ABSTRAK

Penyakit pada tanaman tomat (*Solanum lycopersicum*) menjadi salah satu faktor utama penyebab penurunan kualitas dan kuantitas produksi. Deteksi dini penyakit secara manual memerlukan waktu yang cukup lama dan berpotensi menghasilkan kesalahan identifikasi. Penelitian ini bertujuan mengembangkan model klasifikasi penyakit daun tomat berbasis citra digital menggunakan arsitektur Convolutional Neural Network (CNN) VGG-16 dengan pendekatan transfer learning. Dataset yang digunakan terdiri dari 3.803 citra daun tomat yang mencakup sembilan kelas kondisi, diperoleh melalui platform Kaggle. Tahapan penelitian meliputi preprocessing (pembagian data, augmentasi, normalisasi, dan one-hot encoding), pelatihan model melalui pengujian 27 kombinasi hyperparameter (3 optimizer \times 3 learning rate \times 3 epoch), serta evaluasi menggunakan confusion matrix dan classification report. Hasil terbaik diperoleh pada konfigurasi optimizer SGD, learning rate 0,1, dan epoch 50, dengan nilai akurasi 81%, F1-Score 82%, recall 81%, dan precision 83%. Hasil tersebut menunjukkan bahwa CNN VGG-16 berbasis transfer learning merupakan pendekatan yang efektif dalam mendukung deteksi dini penyakit pada tanaman tomat.

Kata Kunci : penyakit daun tomat, convolutional neural network, vgg-1, transfer learning, klasifikasi citra.

A. Pendahuluan

Tomat (*Solanum lycopersicum*) merupakan komoditas hortikultura yang memiliki peranan penting dan dibudidayakan secara luas di Indonesia. Data Badan Pusat Statistik menunjukkan bahwa produksi tomat Indonesia pada tahun 2022 mencapai 1.168.744 ton, meningkat 0,21% dibandingkan tahun sebelumnya sebesar 1.114.399 ton [1]. Buah tomat memiliki kandungan gizi yang tinggi meliputi vitamin A, B, C, karbohidrat, protein, dan mineral, sehingga bernilai ekonomis besar [2].

Dalam proses budidayanya, serangan hama dan penyakit dapat mengakibatkan penurunan kualitas maupun kuantitas produksi secara signifikan. Penyakit tanaman umumnya disebabkan oleh interaksi antara *mikroorganisme patogen*, tanaman inang, dan lingkungan. Daun sering dijadikan indikator utama dalam proses deteksi penyakit karena gejala awal infeksi umumnya terlihat dari perubahan tekstur, warna, dan pola pada permukaan daun [3]. Berbagai penyakit yang sering mengganggu daun tomat meliputi *early blight*, *late blight*, *leaf mold*, *bacterial spot*, *tomato mosaic virus*, dan *tomato yellow leaf curl virus* [4].

Perkembangan teknologi pengolahan citra digital telah mendorong penggunaan pendekatan *Convolutional Neural Network* (CNN) dalam mendeteksi dan mengklasifikasikan penyakit tanaman. CNN adalah bagian dari *deep learning* yang dirancang untuk mengenali pola visual dari citra dan terbukti mampu mencapai akurasi tinggi dalam klasifikasi penyakit tanaman berdasarkan citra daun karena kemampuannya mengekstraksi fitur secara otomatis

dan efisien [5]. Model CNN populer seperti *VGG-16*, *MobileNet*, *ResNet-50*, dan *AlexNet* telah menunjukkan hasil yang menjanjikan, dengan *VGG-16* mencapai akurasi 97,78% [6], *MobileNet* 97,3% [7], dan *AlexNet* sebesar 92,18% [8].

Transfer learning merupakan teknik yang menggunakan model pra-latih untuk mengatasi keterbatasan data dan mempercepat proses pelatihan [9], [10]. Metode ini terbukti meningkatkan efisiensi serta mengoptimalkan proses klasifikasi pada dataset terbatas [11], [12], [13]. Penelitian ini mengembangkan model klasifikasi penyakit daun tomat berbasis citra digital menggunakan arsitektur CNN *VGG-16* yang dioptimalkan melalui pendekatan *transfer learning*, guna mendukung deteksi penyakit tanaman secara cepat, akurat, dan efisien.

Makalah ini disusun sebagai berikut. Bagian 1 menjelaskan tentang latar belakang permasalahan yang dihadapi dalam penelitian ini. Bagian 2 menjelaskan tentang metode yang digunakan. Bagian 3 menyajikan hasil dan pembahasan penelitian klasifikasi jenis penyakit tomat berdasarkan daun menggunakan menggunakan algoritma *Convolutional Neural Network* (CNN) dan *transfer learning*, serta Bagian 4 menjelaskan kesimpulan dan rencana kedepannya dari hasil penelitian yang dilakukan ini.

B. Metode

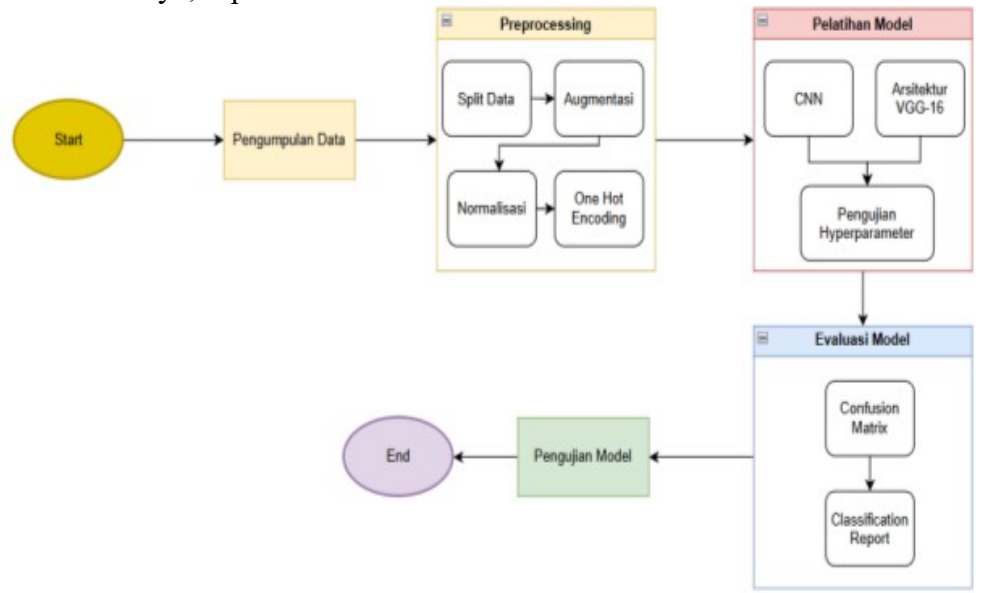
Penelitian ini menggunakan data sekunder berupa citra daun tomat yang diperoleh dari platform *Kaggle*. Dataset terdiri dari gambar daun tomat yang telah dilabeli berdasarkan kondisi sehat maupun jenis penyakit yang menyerang tanaman tomat. Seluruh citra tersedia dalam format JPG dan tersusun dalam folder sesuai kelas masing-masing sehingga memudahkan proses pemrosesan data. Dataset yang digunakan mencakup sembilan kelas utama yaitu *Early Blight*, *Late Blight*, *Leaf Mold*, *Bacterial Spot*, *Tomato Mosaic Virus*, *Septoria Leaf Spot*, *Target Spot*, *Tomato Yellow Leaf Curl Virus*, dan *Healthy*. Data yang digunakan pada penelitian ini terdiri dari 9.900 citra sebagai data pelatihan (training) dan 3.803 citra sebagai data pengujian (testing). Distribusi jumlah data antar kelas dibuat seimbang untuk menghindari bias selama proses pelatihan model sehingga model dapat mempelajari karakteristik setiap kelas secara optimal, seperti yang diperlihatkan dalam Tabel 1.

Tabel 1. Dataset Penelitian

Kelas Penyakit	Jumlah Data
<i>Early Blight</i> (Pembusukan Dini)	343
<i>Late Blight</i> (Busuk Daun)	305
<i>Leaf Mold</i> (Bercak Daun)	280
<i>Bacterial Spot</i> (Bercak Bakteri)	276
<i>Tomato Mosaic Virus</i> (Virus Mozaik Tomat)	527
<i>Tomato Septoria Leaf Spot</i> (Bercak Daun Septoria)	383
<i>Tomato Target Spot</i> (Bintik Bercak)	321
<i>Tomato Yellow Leaf Curl Virus</i> (Virus Keriting Kuning Daun Tomat)	382
<i>Tomato Healthy</i> (Tomat Sehat)	986
Total	3.803

Tahapan penelitian dilakukan secara sistematis mulai dari pengumpulan data hingga evaluasi model klasifikasi. Proses penelitian diawali dengan pengumpulan dataset citra daun

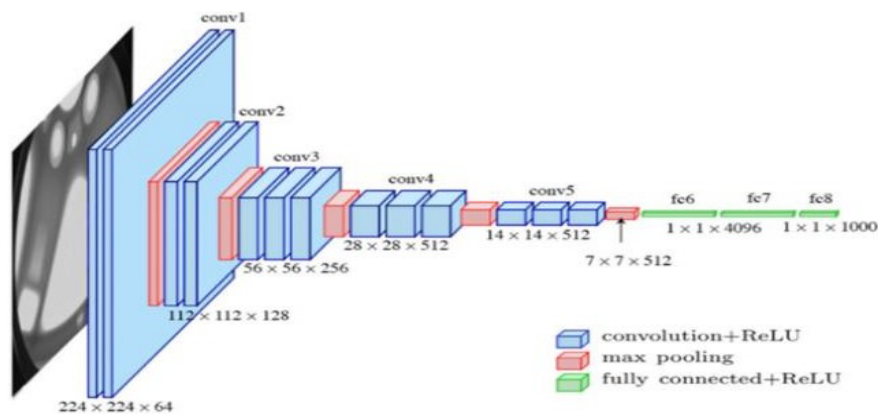
tomat, kemudian dilakukan *preprocessing* data sebelum digunakan dalam pelatihan model. Tahap selanjutnya adalah perancangan arsitektur model CNN berbasis *VGG-16* dengan pendekatan *transfer learning*. Model kemudian dilatih menggunakan beberapa kombinasi *hyperparameter* untuk memperoleh konfigurasi terbaik. Setelah proses pelatihan selesai, dilakukan evaluasi performa model menggunakan data pengujian yang belum pernah digunakan sebelumnya, seperti terlihat dalam Gambar 1.



Gambar 1. Rancangan Penelitian

Tahap preprocessing dilakukan untuk meningkatkan kualitas data sebelum proses pelatihan model. Proses *preprocessing* meliputi normalisasi nilai piksel menggunakan *rescale 1/255* sehingga rentang nilai piksel berubah dari 0–255 menjadi 0–1 untuk mempercepat proses *konvergensi* model. Selain itu dilakukan augmentasi data pada data pelatihan untuk meningkatkan variasi citra melalui rotasi, pergeseran, *zoom*, dan *flipping*. Teknik augmentasi ini bertujuan untuk meningkatkan kemampuan generalisasi model serta mengurangi risiko *overfitting*. Label kelas pada dataset diubah ke dalam format *one-hot encoding* agar dapat digunakan dalam proses klasifikasi multi-kelas pada model CNN. Dataset kemudian dipisahkan menjadi data pelatihan dan data pengujian sehingga proses evaluasi dapat dilakukan secara objektif terhadap data yang belum pernah digunakan selama pelatihan.

Model klasifikasi pada penelitian ini menggunakan *Convolutional Neural Network* dengan arsitektur *VGG-16* berbasis transfer learning. Model pra-latih digunakan sebagai *feature extractor* untuk mengekstraksi fitur visual dari citra daun tomat, kemudian ditambahkan beberapa layer klasifikasi pada bagian akhir untuk menyesuaikan jumlah kelas output. Seperti diperlihatkan dalam Gambar 2, sebagian besar layer pada arsitektur *VGG-16* dibekukan selama proses pelatihan untuk meningkatkan efisiensi komputasi dan mengurangi risiko *overfitting* pada dataset yang terbatas. Ukuran input citra yang digunakan adalah 224×224 piksel sesuai dengan spesifikasi arsitektur *VGG-16*.



Gambar 2. Arsitektur Model CNN

Pelatihan model dilakukan menggunakan beberapa kombinasi *hyperparameter* untuk memperoleh konfigurasi terbaik. Parameter yang diuji meliputi optimizer (Adam, SGD, dan Adamax), *learning rate* (0.001, 0.01, dan 0.1), jumlah *epoch* (20, 25, dan 50), serta *batch size* sebesar 32. Proses pelatihan dilakukan menggunakan GPU untuk mempercepat komputasi selama training. Selain itu, digunakan *callback Early Stopping* dengan nilai *patience* 10 untuk menghentikan pelatihan secara otomatis apabila *validation loss* tidak mengalami perbaikan. *Callback ReduceLRonPlateau* juga digunakan untuk menyesuaikan *learning rate* secara otomatis ketika performa model stagnan sehingga proses pelatihan menjadi lebih optimal.

Evaluasi performa model dilakukan menggunakan data pengujian yang tidak digunakan selama proses pelatihan. Metrik evaluasi yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan model dalam mengklasifikasikan setiap kelas penyakit daun tomat. Selain itu, *confusion matrix* digunakan untuk menganalisis pola kesalahan klasifikasi antar kelas sehingga dapat diketahui kemampuan model dalam membedakan setiap jenis penyakit secara lebih detail.

C. Hasil dan Pembahasan

Preprocessing dilakukan melalui tahap augmentasi citra dan normalisasi piksel sebelum data digunakan pada proses pelatihan model. Augmentasi diterapkan menggunakan rotasi hingga 20° , pergeseran *horizontal* dan *vertikal* masing-masing 10%, zoom hingga 20%, *flipping horizontal*, serta normalisasi piksel menggunakan *rescale 1/255*. Parameter *fill_mode='nearest'* digunakan untuk mengisi piksel kosong hasil transformasi. Proses augmentasi dilakukan secara *real-time* menggunakan *ImageDataGenerator* sehingga model memperoleh variasi citra yang berbeda pada setiap *epoch* tanpa menambah jumlah file secara fisik. Strategi ini membantu meningkatkan kemampuan generalisasi model serta mengurangi risiko *overfitting*. Pada data validasi dan data uji hanya diterapkan normalisasi tanpa augmentasi tambahan agar proses evaluasi mencerminkan kemampuan model pada data asli, seperti terlihat dalam Gambar 3.



Gambar 3. Hasil Augmentasi Data

1. Eksperimen *Hyperparameter*

Model *VGG16* diuji menggunakan 27 kombinasi *hyperparameter* yang terdiri dari tiga optimizer yaitu Adam, SGD, dan Adamax, tiga nilai *learning rate* yaitu 0.001, 0.01, dan 0.1, serta tiga jumlah *epoch* yaitu 20, 25, dan 50. Seluruh eksperimen menggunakan *batch size* 32 dan ukuran input citra 224×224 piksel yang sesuai dengan arsitektur *VGG16*. Hasil eksperimen menunjukkan bahwa kombinasi *hyperparameter* memiliki pengaruh terhadap performa model. *Learning rate* rendah menghasilkan proses pelatihan yang lebih stabil pada optimizer Adam dan Adamax, sedangkan optimizer SGD menunjukkan performa terbaik pada *learning rate* tinggi dengan jumlah *epoch* yang lebih besar. Kombinasi optimizer SGD dengan *learning rate* 0.1 dan *epoch* 50 menghasilkan performa tertinggi dibandingkan konfigurasi lainnya. Hasil eksperimen diperlihatkan dalam Tabel 2.

Tabel 2. Eksperimen *Hyperparameter*

No	Batch Size	Learning Rate	Epoch	Optimizer
1	32	0.001	20	Adam
2	32	0.001	25	Adam
3	32	0.001	50	Adam
4	32	0.01	20	Adam
5	32	0.01	25	Adam
6	32	0.01	50	Adam
7	32	0.1	20	Adam
8	32	0.1	25	Adam
9	32	0.1	50	Adam
10	32	0.001	20	SGD
11	32	0.001	25	SGD
12	32	0.001	50	SGD
13	32	0.01	20	SGD
14	32	0.01	25	SGD
15	32	0.01	50	SGD
16	32	0.1	20	SGD
17	32	0.1	25	SGD

No	Batch Size	Learning Rate	Epoch	Optimizer
18	32	0.1	50	SGD
19	32	0.001	20	Adamax
20	32	0.001	25	Adamax
21	32	0.001	50	Adamax
22	32	0.01	20	Adamax
23	32	0.01	25	Adamax
24	32	0.01	50	Adamax
25	32	0.1	20	Adamax
26	32	0.1	25	Adamax
27	32	0.1	50	Adamax

2. Hasil Pelatihan Model

Grafik pelatihan model digunakan untuk menganalisis proses pembelajaran selama *training*. Grafik *loss* menunjukkan bahwa *training loss* menurun secara signifikan pada awal epoch kemudian menurun secara bertahap hingga akhir pelatihan. *Validation loss* juga mengalami penurunan yang konsisten tanpa peningkatan yang signifikan, sehingga menunjukkan bahwa model tidak mengalami *overfitting* yang berlebihan, seperti terlihat dalam Gambar 4.



Gambar 4. Grafik Loss

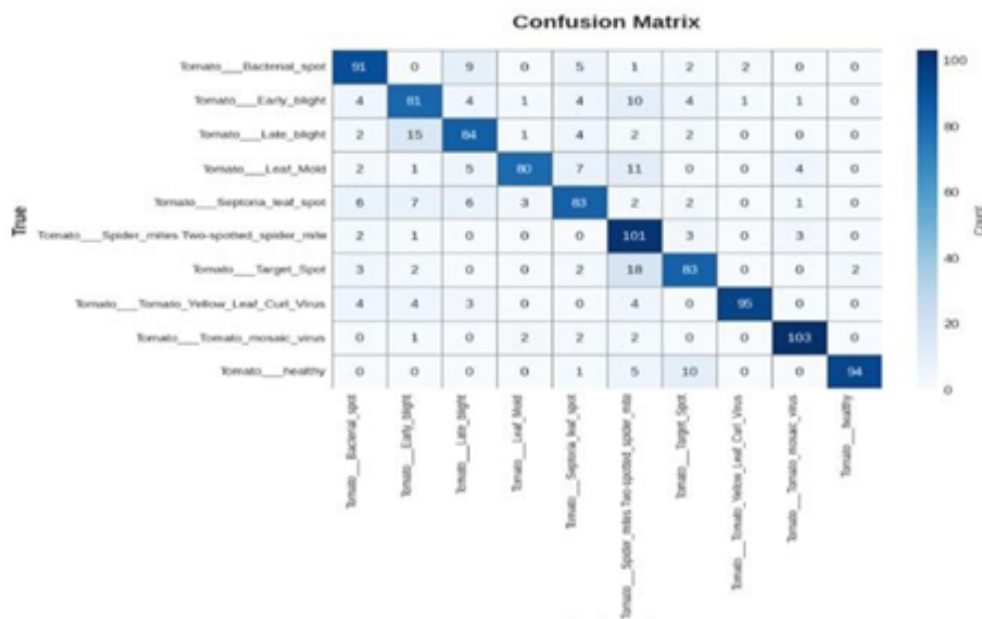
Pada grafik *accuracy* terlihat bahwa *training accuracy* meningkat secara konsisten seiring bertambahnya epoch. *Validation accuracy* meningkat tajam pada epoch awal kemudian stabil pada nilai yang relatif tinggi hingga akhir pelatihan. Kondisi ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data baru, seperti dalam Gambar 5.



Gambar 5. Grafik Accuracy

3. Evaluasi Model

Evaluasi performa model dilakukan menggunakan *confusion matrix* dan *classification report* berdasarkan data uji yang tidak digunakan selama proses pelatihan. Model terbaik diperoleh menggunakan *optimizer* SGD dengan *learning rate* 0.1 dan *epoch* 50 yang menghasilkan akurasi sebesar 81%, *precision* 83%, *recall* 81%, dan *F1-score* 82%. *Confusion matrix* menunjukkan bahwa sebagian besar prediksi berada pada diagonal utama yang menandakan tingkat klasifikasi yang baik pada setiap kelas penyakit daun tomat. Kelas dengan performa tertinggi adalah *Tomato Mosaic Virus* dan *Tomato Yellow Leaf Curl Virus* dengan jumlah prediksi benar yang tinggi. Sementara itu, beberapa kesalahan klasifikasi terjadi pada kelas yang memiliki karakteristik visual serupa seperti *Early Blight* dan *Late Blight*, seperti terlihat dalam Gambar 6.



Gambar 6. Hasil Confusion Matrix

Classification report menunjukkan performa model yang relatif stabil pada seluruh kelas penyakit daun tomat. Nilai *precision* tertinggi diperoleh pada kelas *healthy* dan *Tomato Yellow Leaf Curl Virus*, sedangkan nilai *recall* tertinggi diperoleh pada kelas *Tomato Mosaic Virus*, seperti terlihat dalam Gambar 7.

Classification Report					
No.	Jenis Penyakit	Precision	Recal	F1-Score	Support
1	Bacterial Spot (Bercak Bakteri)	0.80	0.83	0.81	110
2	Early Blight (Pembusukan Dini)	0.72	0.74	0.73	110
3	Late Blight (Busuk Daun)	0.76	0.76	0.76	110
4	Leaf Mold (Bercak Daun)	0.92	0.73	0.81	110
5	Septoria Leaf Spot (bercak daun septoria)	0.77	0.75	0.76	110
6	Target Spot (Bintik bercak)	0.78	0.75	0.77	110
7	Yellow Leaf Curl Virus (Virus keriting kuning daun tomat)	0.97	0.86	0.91	110
8	mosaic virus (Virus mozaik tomat)	0.92	0.94	0.93	110
9	healthy (Tomat sehat)	0.98	0.85	0.91	110
Accuracy Score :					0.81
F1-Score :					0.82
Recall Score :					0.81
Precision Score:					0.83

Gambar 7 Hasil *Classification Report*

Kesalahan klasifikasi yang terjadi sebagian besar dipengaruhi oleh kemiripan pola bercak dan tekstur antar beberapa jenis penyakit daun tomat serta variasi tingkat keparahan gejala pada setiap citra. Secara keseluruhan, model mampu mengklasifikasikan sebagian besar kelas dengan tingkat ketepatan yang baik sehingga dapat digunakan sebagai sistem pendukung identifikasi penyakit tanaman berbasis citra digital.

D. Simpulan

Penelitian ini berhasil mengembangkan model klasifikasi penyakit daun tomat berbasis citra digital menggunakan arsitektur CNN *VGG-16* dengan pendekatan *transfer learning* yang mampu mengklasifikasikan sembilan jenis kondisi daun tomat. Penerapan *transfer learning* terbukti meningkatkan efisiensi pelatihan dan kemampuan generalisasi model. Hasil terbaik diperoleh menggunakan konfigurasi *optimizer* SGD, *learning rate* 0.1, dan *epoch* 50, yang menghasilkan akurasi 81%, *F1-Score* 82%, *recall* 81%, dan *precision* 83%. Pengujian 27 kombinasi *hyperparameter* menunjukkan bahwa pemilihan kombinasi *optimizer* dan *learning rate* yang tepat sangat berpengaruh terhadap keberhasilan model dalam mengidentifikasi berbagai jenis penyakit daun tomat. Untuk penelitian selanjutnya, disarankan meningkatkan jumlah dan keragaman dataset dari berbagai kondisi pencahayaan, sudut pengambilan citra, dan tingkat keparahan gejala. Selain itu, model yang dikembangkan dapat diintegrasikan ke dalam aplikasi berbasis *website* atau *mobile* agar dapat dimanfaatkan langsung oleh petani sebagai sistem pendukung identifikasi penyakit tanaman berbasis pengolahan citra digital.

Daftar Pustaka

- [1] D. Adellia, A. C. Siregar, and S. P. A. Alkadri, "Penerapan Metode Certainty Factor pada Sistem Pakar Diagnosa Hama dan Penyakit pada Tanaman Tomat," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 3, p. 451, 2022, doi: 10.26418/jp.v8i3.56178.
- [2] I. Artikel, "Evaluasi Kualitas Buah, Komponen Hasil dan Hasil Enam Varietas Tomat Calon Tetua Persilangan," vol. 35, no. 3, pp. 413–425, 2024.
- [3] A. Nurdin, D. Satria, Y. Kartika, A. Rezha, and E. Najaf, "Klasifikasi Penyakit Daun Tomat Dengan Metode Convolutional Neural Network Menggunakan Arsitektur Inception-V3," no. 1, pp. 1–6, 2024.

- [4] A. W. Putri, “Implementasi Artificial Neural Network (ANN) Backpropagation Untuk Klasifikasi Jenis Penyakit Pada Daun Tanaman Tomat,” *MATHunesa J. Ilm. Mat.*, vol. 9, no. 2, pp. 344–350, 2021, doi:10.26740/mathunesa.v9n2.p344-350.
- [5] R. A. Kurniawan, A. Sunyoto, and A. Nasiri, “Pengaruh Arsitektur Convolutional Neural Network Untuk Klasifikasi Penyakit Daun Tomat (Effect of Convolutional Neural Network Architecture for Tomato Leaf Disease Classification),” pp. 1–6, 2021, [Online]. Available: <https://www.kaggle.com/datasets/emmarex/plantdiseases>
- [6] M. I. Fathur Rozi, N. O. Adiwijaya, and D. I. Swasono, “Identifikasi Kinerja Arsitektur Transfer Learning Vgg16, Resnet-50, Dan Inception-V3 Dalam Pengklasifikasian Citra Penyakit Daun Tomat,” *J. Ris. Rekayasa Elektro*, vol. 5, no. 2, p. 145, 2023, doi: 10.30595/jrre.v5i2.18050.
- [7] M. A. Mobilenet and D. A. N. Xception, “Analisis perbandingan hasil klasifikasi jenis penyakit tanaman tomat menggunakan arsitektur mobilenet, densenet121, dan xception 1,2,” vol. 3, no. 3, pp. 56–69, 2024.
- [8] P. R. Babu and A. S. Krishna, “Deep Learning- Assisted SVMs for Efficacious Diagnosis of Tomato Leaf Diseases: A Comparative Study of GoogleNet, AlexNet, and ResNet-50,” *Ing. des Syst. d’Information*, vol. 28, no. 3, pp. 639–645, 2023, doi: 10.18280/isi.280312.
- [9] M. I. Wahid, A. Lawi, D. A. Muh, and A. Siddik, “Perbandingan Kinerja Model Ensembled Transfer Learning Pada Klasifikasi Penyakit Daun Tomat,” *Pros. Semin. Nas. Tek. Elektro dan Inform.*, vol. 8, no. 1, pp. 286–291, 2023.
- [10] R. Soekarta, N. Nurdjan, and A. Syah, “Klasifikasi Penyakit Tanaman Tomat Menggunakan Metode Convolutional Neural Network (CNN),” *Insect (Informatics Secur. J. Tek. Inform.*, vol. 8, no. 2, pp. 143–151, 2023, doi: 10.33506/insect.v8i2.2356.
- [11] G. H. A. Panjaitan and F. Simatupang, “Pemodelan Klasifikasi Penyakit Daun Tanaman Tomat dengan Convolutional Neural Network Algorithm,” *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 5, pp. 2667–2675, 2024, doi: 10.30865/klik.v4i5.1646.
- [12] B. N. Surya Adi Laksono, Basuki Rahmat, “Menggunakan Metode Convolutional Neural Network,” *J. Inform. dan Sist. Inf.*, vol. 1, no. 1, pp. 3331–3338, 2020.
- [13] M. A. Salim, “Perbandingan Convolutional Neural Network dan SVM untuk Klasifikasi Penyakit Daun Padi,” *J. Teknol. Dan Sist. Inf.*, vol. 3(2), no. 1, pp. 117–125, 2021.