

## Augmentasi Data berbasis GAN dan Ekstraksi Fitur EfficientNetB0 dengan XGBoost untuk Meningkatkan Klasifikasi Penyakit Daun Jagung

Ririn Anugerah Ikasatya\*, Cahya Apriliani, Fathir Jannatul Firdaus, Gede Yogi Pratama  
Universitas Bumigora, Mataram, Indonesia

Email Korespondensi: [ririnai2004@gmail.com](mailto:ririnai2004@gmail.com)

Genesis Artikel: Diterima: 20 Maret 2026 Diterbitkan: 31 Maret 2026

**ABSTRACT:** Corn leaf diseases are one of the main factors contributing to decreased corn productivity. Manual identification of leaf diseases remains subjective, time-consuming, and highly dependent on individual experience. This study aims to improve the performance of image-based corn leaf disease classification through the integration of data balancing techniques, deep feature extraction, and machine learning-based classification methods. The dataset consists of four classes with an imbalanced distribution, namely Blight with 802 images, Common Rust with 914 images, Gray Leaf Spot with 401 images, and Healthy with 813 images, where Gray Leaf Spot represents the minority class. Data balancing is performed by generating synthetic images using a convolution-based generative model to increase the number of samples in the minority class. Furthermore, feature extraction is carried out using the EfficientNetB0 architecture, and classification is performed using a gradient boosting-based algorithm. The results show that the proposed approach improves accuracy from 92.49 percent to 93.29 percent and enhances the model's ability to recognize the minority class, as indicated by an increase in recall from 69 percent to 78 percent and an improvement in performance balance from 0.76 to 0.84. These findings indicate that the proposed method is effective in improving classification performance, particularly for the minority class, without reducing performance on majority classes.

**Keyword:** DC-GAN; EfficientNetB0; Image Classification; Corn Leaf Disease; XGBoost.

**ABSTRAK:** Penyakit daun jagung merupakan salah satu faktor utama yang menyebabkan penurunan produktivitas tanaman jagung. Proses identifikasi penyakit yang masih dilakukan secara manual cenderung subjektif, memerlukan waktu yang lama, serta bergantung pada pengalaman individu. Penelitian ini bertujuan untuk meningkatkan kinerja klasifikasi penyakit daun jagung berbasis citra digital melalui integrasi beberapa metode, yaitu penyeimbangan data menggunakan *Deep Convolutional Generative Adversarial Network* (DC-GAN), ekstraksi fitur menggunakan arsitektur EfficientNetB0, serta klasifikasi menggunakan algoritma XGBoost. Dataset yang digunakan terdiri dari empat kelas dengan distribusi tidak seimbang, yaitu *Blight* (802 citra), *Common Rust* (914 citra), *Gray Leaf Spot* (401 citra), dan *Healthy* (813 citra), dengan *Gray Leaf Spot* sebagai kelas minoritas. Penyeimbangan data dilakukan dengan menghasilkan citra sintetis menggunakan DC-GAN untuk meningkatkan jumlah sampel pada kelas minoritas. Hasil penelitian menunjukkan bahwa pendekatan yang diusulkan mampu meningkatkan akurasi dari 92,49% menjadi 93,29%. Selain itu, kemampuan model dalam mengenali kelas minoritas juga meningkat, yang ditunjukkan oleh kenaikan nilai recall dari 69% menjadi 78% serta peningkatan *F1-score* dari 0,76 menjadi 0,84. Implikasi dari penelitian ini adalah bahwa metode yang diusulkan dapat digunakan sebagai solusi dalam deteksi dini penyakit daun jagung secara lebih akurat dan konsisten, sehingga berpotensi mendukung pengambilan keputusan dalam pengelolaan tanaman serta meningkatkan produktivitas pertanian.

**Kata Kunci:** DC-GAN; EfficientNetB0; Klasifikasi Citra; Penyakit Daun Jagung; XGBoost.

Ini adalah artikel akses terbuka dibawah lisensi [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/).



### Cara Sitasi:

Ikasatya, R.A., Apriliani, C., Firdaus, F.J., Pratama, G.Y. (2026). Augmentasi Data berbasis GAN dan Ekstraksi Fitur EfficientNetB0 dengan XGBoost untuk Meningkatkan Klasifikasi Penyakit Daun Jagung. *UPGRADE: Jurnal Pendidikan Teknologi Informasi*, 3(2), 77-88. <https://doi.org/10.30812/upgrade.v3i2.6224>

## PENDAHULUAN

Dalam dua dekade terakhir, produktivitas pertanian di Indonesia semakin mendapat perhatian karena perannya yang strategis sebagai pilar utama ketahanan pangan nasional (Tono et al., 2023). Tanaman jagung merupakan salah satu komoditas unggulan yang dimanfaatkan sebagai bahan pangan, pakan ternak, serta bahan baku industri (Melia et al., 2023). Namun demikian, produktivitas jagung masih sering terganggu oleh serangan penyakit daun seperti *Blight*, *Common Rust*, dan *Gray Leaf Spot* (Yuninsar et al., 2024). Penyakit-penyakit tersebut menyebabkan kerusakan jaringan daun, menurunnya proses fotosintesis, serta berdampak langsung pada penurunan hasil panen (Sarah and Guntaro, 2023). Di tingkat lapangan, proses identifikasi penyakit daun jagung umumnya masih dilakukan secara manual oleh petani sehingga bersifat subjektif, memerlukan waktu, serta sangat bergantung pada pengalaman individu. Kondisi ini menunjukkan perlunya sistem deteksi otomatis berbasis citra digital yang mampu mengidentifikasi penyakit daun jagung secara cepat, akurat, dan efisien sejak tahap awal infeksi (Aryanti et al., 2025).

Perkembangan teknologi kecerdasan buatan, khususnya *machine learning* dan *deep learning*, membuka peluang besar dalam pemanfaatan citra digital untuk mendukung sektor pertanian (Andhika and Bayunanda, 2025). Salah satu metode yang banyak digunakan dalam klasifikasi citra adalah *Convolutional Neural Network* (CNN) karena kemampuannya mengekstraksi fitur visual secara hierarkis dan representatif (Ardianto and Wibisono, 2023; Wirabowo and Susilawati, 2025). Selain itu, pendekatan *transfer learning* memungkinkan pemanfaatan model pra-latih yang telah dilatih menggunakan dataset berskala besar sehingga mampu meningkatkan performa klasifikasi meskipun jumlah data pelatihan terbatas (Purba, 2025). *EfficientNetB0* merupakan salah satu arsitektur CNN yang dirancang untuk mencapai keseimbangan antara akurasi dan efisiensi komputasi, sehingga potensial digunakan sebagai metode ekstraksi fitur pada klasifikasi penyakit daun jagung (Burhanuddin et al., 2025).

Beberapa penelitian sebelumnya telah mengkaji klasifikasi penyakit daun jagung dengan pendekatan yang beragam. Penelitian oleh (Prasetyo and Wulaningrum, 2023) menggunakan metode ekstraksi fitur klasik seperti *Color Moments* dan *Gray Level Co-occurrence Matrix* (GLCM) yang dikombinasikan dengan algoritma *K-Nearest Neighbor*, namun memiliki keterbatasan dalam menghadapi variasi visual citra. Perkembangan selanjutnya menunjukkan bahwa pendekatan berbasis *deep learning* seperti *Convolutional Neural Network* (CNN) mampu memberikan performa yang lebih baik, salah satunya ditunjukkan oleh penelitian yang menggunakan arsitektur ResNet50 dengan akurasi mencapai 98,4% (Putra and Alamsyah, 2022). Selain itu, metode CNN juga terbukti lebih unggul dibandingkan algoritma pembelajaran mesin tradisional seperti *Support Vector Machine* (SVM) dan *Random Forest* dalam menangkap pola kompleks pada citra (Prayesy, 2025).

Meskipun demikian, penggunaan arsitektur *deep learning* berukuran besar seperti VGG16, Xception, dan DenseNet yang banyak digunakan pada penelitian sebelumnya menghasilkan akurasi tinggi tetapi memerlukan sumber daya komputasi yang besar (Aisyah et al., 2025; Ht. Barat et al., 2025; Tirta et al., 2025). Untuk meningkatkan performa, beberapa penelitian mengusulkan pendekatan *ensemble learning* yang mampu meningkatkan akurasi dan stabilitas model hingga 96,02% (Kurniawan et al., 2025), namun pendekatan ini cenderung meningkatkan kompleksitas model. Di sisi lain, permasalahan ketidakseimbangan data menjadi tantangan utama yang menyebabkan model bias terhadap kelas mayoritas. Penelitian terbaru menunjukkan bahwa augmentasi data berbasis *Generative Adversarial Network* (GAN) mampu meningkatkan representasi kelas minoritas serta memperbaiki performa model, khususnya pada nilai recall dan keseimbangan performa (Suhesi et al., 2025).

Selain kompleksitas model, permasalahan ketidakseimbangan jumlah data antar kelas juga menjadi isu penting dalam klasifikasi penyakit daun jagung (Rahmanita et al., 2023). Distribusi data yang tidak merata dapat menyebabkan model cenderung bias terhadap kelas mayoritas dan menurunkan performa pada kelas minoritas (Erlin et al., 2022). Teknik augmentasi konvensional seperti rotasi dan *flipping* memiliki keterbatasan karena hanya menghasilkan variasi dari data yang sama (Husen, 2024). Oleh karena itu, diperlukan pendekatan yang mampu menghasilkan data baru yang lebih representatif untuk meningkatkan keseimbangan dataset. Namun, hingga saat ini penelitian yang secara bersamaan mengatasi permasalahan ketidakseimbangan data, efisiensi komputasi, dan peningkatan performa pada

kelas minoritas dalam satu kerangka pendekatan yang terintegrasi masih terbatas.

Berdasarkan tinjauan tersebut, dapat disimpulkan bahwa sebagian besar penelitian masih berfokus pada peningkatan akurasi atau penggunaan model dengan kompleksitas tinggi, tanpa secara spesifik mengoptimalkan performa pada kelas minoritas sekaligus mempertahankan efisiensi komputasi. Oleh karena itu, penelitian ini mengusulkan integrasi augmentasi data berbasis GAN, ekstraksi fitur menggunakan EfficientNetB0, serta klasifikasi menggunakan algoritma XGBoost untuk meningkatkan kinerja klasifikasi penyakit daun jagung. Kebaruan penelitian ini tidak hanya terletak pada kombinasi metode, tetapi pada pendekatan terintegrasi yang secara khusus dirancang untuk meningkatkan performa pada kelas minoritas melalui augmentasi data sintesis, sekaligus mempertahankan efisiensi komputasi melalui penggunaan arsitektur ringan dan *classifier* berbasis *boosting*. Berbeda dengan penelitian sebelumnya yang umumnya hanya berfokus pada peningkatan akurasi secara keseluruhan, penelitian ini menekankan keseimbangan performa antar kelas sebagai aspek utama dalam evaluasi model. Tujuan penelitian ini adalah mengembangkan model klasifikasi penyakit daun jagung berbasis citra digital yang lebih akurat, seimbang, dan efisien melalui pemanfaatan kombinasi metode tersebut.

## METODE

Penelitian ini bertujuan untuk meningkatkan kinerja klasifikasi penyakit daun jagung melalui integrasi augmentasi data berbasis GAN, ekstraksi fitur menggunakan EfficientNetB0, serta klasifikasi menggunakan algoritma XGBoost. Penelitian ini menggunakan pendekatan eksperimen berbasis *machine learning* yang disusun secara sistematis mulai dari tahap pengumpulan data, pra-pemrosesan citra, augmentasi data, ekstraksi fitur, proses klasifikasi, hingga evaluasi kinerja model.

Dataset yang digunakan dalam penelitian ini berupa citra daun jagung yang diperoleh dari platform Kaggle (*Corn or Maize Leaf Disease Dataset*, 2019). Dataset terdiri dari empat kelas, yaitu *Blight*, *Common Rust*, *Gray Leaf Spot*, dan *Healthy*. Seluruh citra disusun dalam struktur folder berdasarkan kelasnya, kemudian dataset dibagi menjadi data latih, data validasi, dan data uji dengan rasio 70% : 20% : 10% menggunakan library *split-folders* (Annisa et al., 2024). Pada data latih awal sebelum penerapan GAN, jumlah citra pada setiap kelas terdiri dari 802 citra untuk kelas *Blight*, 914 citra untuk kelas *Common Rust*, 401 citra untuk kelas *Gray Leaf Spot*, dan 813 citra untuk kelas *Healthy*. Distribusi tersebut menunjukkan adanya ketidakseimbangan data, terutama pada kelas *Gray Leaf Spot* sebagai kelas minoritas.

Tahap *preprocessing* dilakukan untuk menyiapkan citra agar sesuai dengan kebutuhan model pembelajaran mesin. Setiap citra dibaca menggunakan library OpenCV, kemudian dilakukan konversi ruang warna dari BGR ke RGB. Selanjutnya citra diubah ukurannya menjadi 224×224 piksel agar sesuai dengan ukuran masukan model EfficientNetB0. Nilai piksel citra kemudian dikonversi ke tipe data *float32* dan dinormalisasi menggunakan fungsi *preprocess\_input* dari EfficientNet untuk menyesuaikan distribusi nilai piksel dengan karakteristik model prelatih ImageNet (Hartono et al., 2023). Proses normalisasi dapat dinyatakan pada Persamaan 1.

$$x' = \frac{x}{255} \quad (1)$$

Pada persamaan tersebut,  $x$  merepresentasikan nilai piksel citra sebelum normalisasi, sedangkan  $x'$  merupakan nilai piksel citra setelah normalisasi. Proses ini bertujuan untuk menyamakan skala nilai piksel agar berada pada rentang yang sesuai dengan data latih model prelatih, sehingga dapat mempercepat konvergensi dan meningkatkan stabilitas pembelajaran model. Selain itu, normalisasi juga membantu mengurangi pengaruh variasi pencahayaan antar citra. Seluruh tahapan *preprocessing* diterapkan secara konsisten baik pada skenario tanpa augmentasi maupun dengan augmentasi GAN (Batubara et al., 2020).

Untuk mengatasi ketidakseimbangan data, penelitian ini menerapkan augmentasi menggunakan GAN berbasis arsitektur *Deep Convolutional* GAN. Model ini terdiri dari dua komponen utama, yaitu generator dan *discriminator*. Generator terdiri dari beberapa lapisan *transposed convolution* sebanyak empat layer, dengan fungsi aktivasi ReLU pada *hidden layer* dan *Tanh* pada *layer output* untuk menghasilkan citra sintesis berukuran 64 × 64 piksel dari vektor *noise* berdimensi 100 (You et al., 2022). Sedangkan

*discriminator* terdiri dari empat lapisan convolutional dengan fungsi aktivasi LeakyReLU yang bertugas membedakan citra asli dan sintetis (Ryanto and Sari, 2024). Model dilatih menggunakan *optimizer Adam* dengan *learning rate* 0,0002 dan batch size 32 selama 100 *epoch*. Fungsi *loss* yang digunakan adalah *binary cross entropy*. Proses pelatihan GAN mengikuti fungsi objektif minimaks sebagaimana ditunjukkan pada Persamaan 2.

$$G^{min} D^{max} V(D, G) = E_{x \sim p_{data}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (2)$$

Pada persamaan tersebut,  $x$  merupakan data asli yang berasal dari distribusi data,  $z$  adalah vektor *noise* acak yang digunakan sebagai *input* generator,  $G(z)$  merupakan citra sintetis yang dihasilkan oleh generator,  $D(x)$  menunjukkan probabilitas bahwa data merupakan citra asli, sedangkan  $D(G(z))$  menunjukkan probabilitas bahwa citra sintetis dikenali sebagai data asli oleh *discriminator*. Kualitas citra sintetis dievaluasi secara visual serta dianalisis melalui distribusi data hasil augmentasi terhadap data asli berdasarkan kemampuannya dalam meningkatkan performa klasifikasi pada kelas minoritas.

Sebanyak 200 citra sintetis dihasilkan menggunakan GAN dan ditambahkan ke dalam kelas minoritas *Gray Leaf Spot*, sehingga jumlah data latih meningkat dari 401 menjadi 601 citra. Tahap ekstraksi fitur dilakukan menggunakan EfficientNetB0 sebagai *feature extractor*. Model ini dipilih karena memiliki keseimbangan antara performa dan efisiensi komputasi serta telah dilatih sebelumnya pada dataset ImageNet (Sarasati et al., 2022). Pada penelitian ini EfficientNetB0 digunakan tanpa lapisan klasifikasi (*include\_top = False*) dan ditambahkan lapisan *Global Average Pooling* untuk menghasilkan vektor fitur berdimensi 1280. Proses ekstraksi fitur dapat dinyatakan secara matematis sebagai berikut:

$$f = \text{EfficientNetB0}(x) \quad (3)$$

Pada persamaan tersebut,  $x$  merupakan citra masukan yang telah melalui tahap *preprocessing*, sedangkan  $f$  merupakan vektor fitur hasil ekstraksi yang merepresentasikan karakteristik citra. Tahap klasifikasi dilakukan menggunakan algoritma XGBoost, yaitu metode *ensemble learning* berbasis pohon keputusan. Vektor fitur hasil ekstraksi EfficientNetB0 digunakan sebagai masukan model, sedangkan label kelas dikodekan menggunakan *Label Encoder*. Model XGBoost dilatih menggunakan parameter utama berupa jumlah estimator 300, kedalaman maksimum pohon 6, dan *learning rate* 0,05 dengan fungsi objektif *multi:softprob* untuk klasifikasi multi-kelas Yulianti et al. (2022). Pemilihan parameter tersebut dilakukan berdasarkan hasil eksperimen awal (*trial and error*) untuk memperoleh keseimbangan antara performa model dan kompleksitas komputasi, di mana jumlah estimator yang lebih tinggi meningkatkan kemampuan model dalam mempelajari pola data, sedangkan *learning rate* yang lebih kecil membantu menghindari *overfitting*.

Evaluasi kinerja model dilakukan menggunakan confusion matrix dan *classification report*. *Confusion matrix* digunakan untuk menganalisis distribusi prediksi benar dan salah pada setiap kelas penyakit daun jagung, sedangkan *classification report* digunakan untuk menghitung metrik akurasi, *precision*, *recall*, dan *F1-score*. Nilai akurasi dihitung menggunakan Persamaan 4.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Pada persamaan tersebut, *True Positive* (TP) adalah jumlah data yang diprediksi benar sebagai kelas positif, *True Negative* (TN) adalah jumlah data yang diprediksi benar sebagai kelas negatif, *False Positive* (FP) adalah jumlah data yang salah diprediksi sebagai positif, dan *False Negative* (FN) adalah jumlah data yang salah diprediksi sebagai negatif. *Precision* digunakan untuk mengukur ketepatan prediksi positif, *recall* untuk mengukur kemampuan model dalam mendeteksi kelas positif, dan *F1-score* sebagai rata-rata harmonis antara *precision* dan *recall*. Selain itu, untuk memastikan kestabilan model, dilakukan pengujian menggunakan beberapa variasi *random state* pada proses pembagian data. Evaluasi juga dapat dikembangkan lebih lanjut menggunakan teknik *k-fold cross-validation* untuk memperoleh estimasi performa yang lebih *robust* terhadap variasi data.

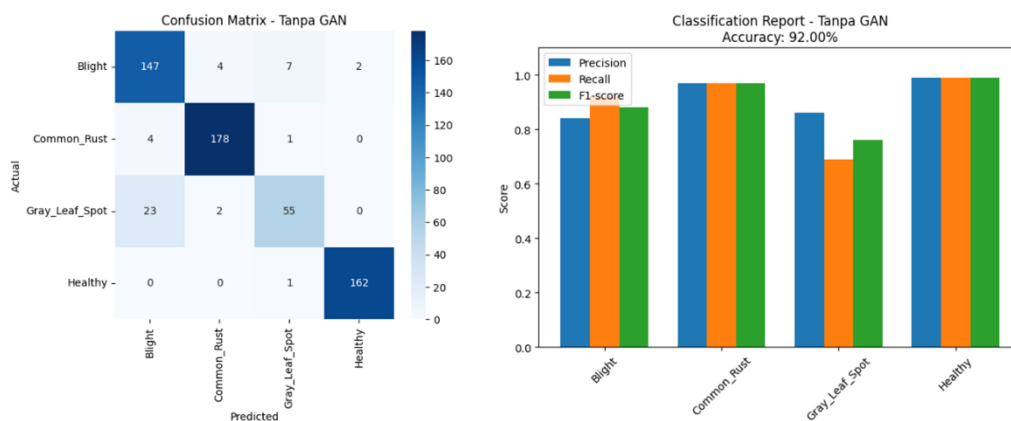
## HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil pengujian model klasifikasi penyakit daun jagung. Selain itu, disajikan pembahasan komprehensif mengenai pengaruh penggunaan GAN dalam meningkatkan performa model. Hasil penelitian ditampilkan dalam bentuk tabel dan visualisasi grafik untuk memudahkan interpretasi, khususnya dalam membandingkan kinerja model dengan dan tanpa penerapan GAN.

### Hasil Klasifikasi Tanpa GAN

Pada skenario pertama, model klasifikasi dibangun tanpa penerapan augmentasi data berbasis GAN. Dataset latih yang digunakan memiliki distribusi data yang tidak seimbang, di mana kelas *Gray Leaf Spot* merupakan kelas minoritas dengan jumlah citra paling sedikit. Seluruh citra diproses melalui tahap ekstraksi fitur menggunakan EfficientNetB0, kemudian hasil ekstraksi fitur tersebut diklasifikasikan menggunakan algoritma XGBoost.

Hasil evaluasi menunjukkan bahwa model klasifikasi tanpa GAN mampu mencapai akurasi sebesar 92,49%. Meskipun demikian, performa model pada kelas *Gray Leaf Spot* masih relatif lebih rendah dibandingkan kelas lainnya. Nilai *recall* untuk kelas *Gray Leaf Spot* hanya mencapai 69%, yang mengindikasikan bahwa sebagian citra pada kelas tersebut masih salah diklasifikasikan ke kelas lain. Gambaran lebih rinci mengenai distribusi hasil prediksi model dapat dilihat pada Gambar 1.



Gambar 1. *Confusion matrix* dan *classification report* model tanpa GAN

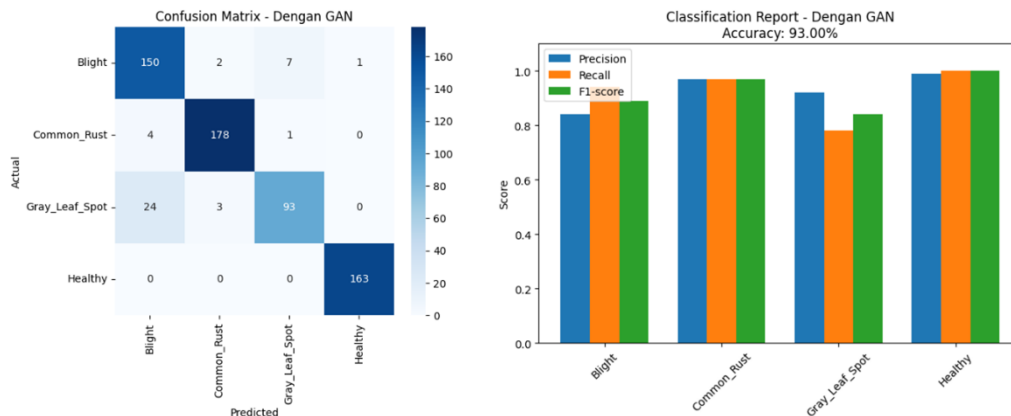
Gambar 1 menampilkan kinerja model klasifikasi tanpa penerapan augmentasi data berbasis GAN yang disajikan melalui *confusion matrix* dan *classification report*. *Confusion matrix* menunjukkan distribusi prediksi model terhadap setiap kelas penyakit daun jagung, di mana kesalahan klasifikasi masih cukup dominan pada kelas *Gray Leaf Spot* sebagai kelas minoritas. Temuan ini diperkuat oleh *classification report* yang memperlihatkan nilai *recall* dan *F1-score* kelas *Gray Leaf Spot* lebih rendah dibandingkan kelas lainnya. Hasil tersebut menunjukkan bahwa ketidakseimbangan data latih berdampak pada keterbatasan model dalam mengenali kelas minoritas secara optimal pada skenario tanpa GAN.

### Hasil Klasifikasi Dengan GAN

Sebelum digunakan dalam proses pelatihan, kualitas citra sintetis hasil GAN divalidasi secara visual dengan membandingkan karakteristik tekstur, pola bercak, dan distribusi warna terhadap citra asli. Hasil pengamatan menunjukkan bahwa citra sintetis memiliki kemiripan visual yang cukup tinggi dengan citra asli, khususnya pada pola khas *Gray Leaf Spot*. Selain itu, distribusi data setelah augmentasi dianalisis untuk memastikan bahwa penambahan data tidak mengubah karakteristik distribusi secara signifikan, melainkan hanya meningkatkan representasi kelas minoritas.

Pada skenario kedua, GAN diterapkan untuk mengatasi ketidakseimbangan data dengan menghasilkan citra sintetis pada kelas *Gray Leaf Spot*. Sebanyak 200 citra sintetis ditambahkan ke dalam data latih, sehingga jumlah citra pada kelas *Gray Leaf Spot* meningkat dari 401 menjadi 601 citra.

Setelah proses augmentasi, fitur citra diekstraksi kembali menggunakan EfficientNetB0 dan selanjutnya diklasifikasikan menggunakan algoritma XGBoost dengan konfigurasi parameter yang sama seperti pada skenario tanpa GAN. Hasil pengujian menunjukkan adanya peningkatan performa model secara keseluruhan dengan nilai akurasi mencapai 93,29%. Peningkatan kinerja yang paling signifikan terjadi pada kelas *Gray Leaf Spot*, di mana nilai *recall* meningkat dari 69% menjadi 78% dan nilai *F1-score* meningkat dari 0,76 menjadi 0,84. Hasil ini menunjukkan bahwa penerapan GAN berhasil meningkatkan kemampuan model dalam mengenali kelas minoritas tanpa menurunkan performa klasifikasi pada kelas mayoritas. Hasil evaluasi secara lebih rinci ditunjukkan pada Gambar 2.



Gambar 2. *Confusion matrix* dan *classification report* model dengan GAN

Gambar 2 menampilkan kinerja model klasifikasi setelah penerapan augmentasi data berbasis GAN yang disajikan dalam bentuk *confusion matrix* dan *classification report*. *Confusion matrix* menunjukkan adanya penurunan kesalahan klasifikasi pada kelas *Gray Leaf Spot*, yang ditandai dengan meningkatnya jumlah prediksi benar dibandingkan dengan skenario tanpa GAN. Hal ini mengindikasikan bahwa citra sintetis hasil GAN mampu meningkatkan sensitivitas model terhadap kelas minoritas.

Peningkatan kinerja tersebut diperkuat oleh *classification report* yang menunjukkan kenaikan nilai *recall* dan *F1-score* pada kelas *Gray Leaf Spot*. Temuan ini menunjukkan bahwa augmentasi data berbasis GAN tidak hanya meningkatkan kemampuan model dalam mendeteksi kelas minoritas, tetapi juga membantu menjaga keseimbangan performa klasifikasi antar kelas tanpa menurunkan kinerja pada kelas lainnya.

### Perbandingan Kinerja Model

Untuk mengetahui pengaruh penerapan augmentasi data berbasis GAN terhadap kinerja klasifikasi penyakit daun jagung, dilakukan perbandingan performa model antara skenario tanpa GAN dan dengan GAN. Perbandingan difokuskan pada metrik evaluasi utama, yaitu akurasi keseluruhan serta nilai *recall* dan *F1-score* pada kelas *Gray Leaf Spot* sebagai kelas minoritas.

### Perbandingan Kinerja Berdasarkan Metrik Evaluasi

Hasil perbandingan performa model pada kedua skenario disajikan dalam bentuk tabel untuk memperlihatkan perbedaan nilai metrik evaluasi secara kuantitatif. Tabel ini menunjukkan bahwa penerapan GAN memberikan peningkatan kinerja model, khususnya pada kemampuan mengenali kelas *Gray Leaf Spot*.

Tabel 1. Perbandingan Performa Model Tanpa GAN dan dengan GAN

Skenario	Akurasi (%)	Recall <i>Gray Leaf Spot</i> (%)	F1-score <i>Gray Leaf Spot</i>
Tanpa GAN	92,49	69	0,76
Dengan GAN	93,29	78	0,84

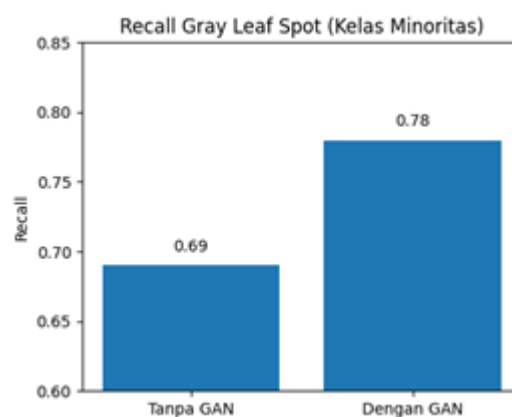
Berdasarkan Tabel 1, terlihat bahwa penerapan GAN meningkatkan akurasi model sebesar 0,8%. Meskipun peningkatan akurasi keseluruhan relatif kecil, dampak yang lebih signifikan terlihat pada kelas *Gray Leaf Spot*. Nilai *recall* meningkat dari 69% menjadi 78%, yang menunjukkan bahwa model dengan GAN lebih mampu mengenali citra penyakit *Gray Leaf Spot* secara benar. Selain itu, nilai *F1-score* juga mengalami peningkatan dari 0,76 menjadi 0,84, yang menandakan keseimbangan yang lebih baik antara *precision* dan *recall* pada kelas minoritas.

Untuk memastikan bahwa peningkatan performa model tidak terjadi secara kebetulan, dilakukan pengujian dengan beberapa variasi *random state* pada proses pembagian data. Hasil pengujian menunjukkan bahwa model dengan penerapan GAN secara konsisten menghasilkan nilai *recall* yang lebih tinggi pada kelas *Gray Leaf Spot* dibandingkan model tanpa GAN. Hal ini mengindikasikan bahwa peningkatan performa yang diperoleh bersifat stabil meskipun belum dilakukan uji statistik formal.

Selain itu, performa pada kelas mayoritas seperti *Healthy* dan *Common Rust* tetap stabil setelah penerapan GAN, yang ditunjukkan oleh nilai *recall* yang tidak mengalami penurunan signifikan pada kedua kelas tersebut. Nilai *recall* pada kelas *Healthy* dan *Common Rust* masing-masing tetap berada pada kisaran di atas 90%, yang menunjukkan bahwa tidak terjadi penurunan performa pada kelas mayoritas. Hal ini menunjukkan bahwa augmentasi data berbasis GAN tidak mengorbankan performa pada kelas mayoritas, melainkan meningkatkan keseimbangan performa antar kelas.

### Visualisasi Perbandingan Performa Model

Untuk memperjelas perbedaan performa antara kedua skenario, dilakukan visualisasi perbandingan dalam bentuk grafik batang. Visualisasi ini digunakan untuk menegaskan pengaruh penerapan augmentasi data berbasis GAN terhadap kinerja model klasifikasi, khususnya pada metrik *recall* kelas *Gray Leaf Spot* sebagai kelas minoritas.



Gambar 3. Grafik perbandingan nilai *recall* pada kelas *Gray Leaf Spot*

Grafik perbandingan *recall* menunjukkan adanya peningkatan yang jelas pada kelas *Gray Leaf Spot* setelah penerapan GAN. Pada skenario tanpa GAN, nilai *recall* masih relatif rendah akibat keterbatasan jumlah data latih. Setelah penambahan citra sintetis hasil GAN, sensitivitas model terhadap kelas *Gray Leaf Spot* meningkat secara signifikan. Hal ini menunjukkan bahwa augmentasi data berbasis GAN mampu mengurangi bias model terhadap kelas mayoritas dan meningkatkan kemampuan deteksi pada kelas minoritas.

### Pembahasan

Hasil penelitian ini menunjukkan adanya kesenjangan (*gap*) pada penelitian sebelumnya yang umumnya hanya berfokus pada penggunaan model CNN untuk meningkatkan akurasi klasifikasi penyakit daun jagung tanpa mempertimbangkan permasalahan ketidakseimbangan data. Sebagian penelitian lebih menitikberatkan pada peningkatan performa model melalui arsitektur *deep learning* yang lebih kompleks,

namun belum secara khusus mengevaluasi kemampuan model dalam mengenali kelas minoritas. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut dengan mengintegrasikan augmentasi data berbasis GAN untuk meningkatkan sensitivitas model terhadap kelas minoritas dalam proses klasifikasi penyakit daun jagung.

Berdasarkan hasil eksperimen yang telah dilakukan, penelitian ini menunjukkan bahwa integrasi GAN, EfficientNetB0, dan XGBoost memberikan dampak positif terhadap kinerja klasifikasi penyakit daun jagung, khususnya pada kelas yang memiliki jumlah data terbatas. Penerapan GAN mampu menghasilkan citra sintetis yang efektif dalam memperkaya variasi data latih, sehingga model menjadi lebih tangguh dalam mengenali pola visual pada kelas minoritas. Temuan ini sejalan dengan referensi (Patmawati et al., 2023) serta (Wildah et al., 2024) yang menyatakan bahwa augmentasi data berbasis GAN dapat meningkatkan kualitas representasi data pada dataset citra yang tidak seimbang.

Peningkatan performa ini terjadi karena citra sintetis yang dihasilkan oleh GAN mampu memperkaya variasi pola visual pada kelas minoritas, sehingga model memiliki representasi fitur yang lebih baik dalam membedakan karakteristik *Gray Leaf Spot* dibandingkan dengan kelas lainnya. Dengan bertambahnya variasi data latih, model tidak lagi terlalu bergantung pada pola terbatas yang sebelumnya menyebabkan kesalahan klasifikasi.

Peningkatan nilai *recall* dan *F1-score* pada kelas *Gray Leaf Spot* menunjukkan bahwa pendekatan yang diusulkan berhasil mengurangi permasalahan bias akibat ketidakseimbangan data. Hasil ini mendukung referensi (Andriani and Iqbal, 2025) yang melaporkan bahwa algoritma XGBoost memiliki keunggulan dalam menangani dataset tidak seimbang serta mampu meningkatkan sensitivitas model terhadap kelas minoritas. Selain itu, penggunaan EfficientNetB0 sebagai metode ekstraksi fitur juga memperkuat referensi (Jiven and Rumini, 2025) yang menunjukkan bahwa arsitektur EfficientNetB0 mampu menghasilkan fitur yang informatif dan efisien ketika dikombinasikan dengan algoritma klasifikasi berbasis *ensemble*.

Selain itu, analisis confusion matrix menunjukkan adanya penurunan jumlah FN pada kelas *Gray Leaf Spot* setelah penerapan GAN. Penurunan FN ini mengindikasikan bahwa model menjadi lebih sensitif dalam mendeteksi penyakit, sehingga mengurangi kesalahan klasifikasi di mana citra penyakit tidak terdeteksi sebagai kelas yang seharusnya. Hal ini menunjukkan bahwa augmentasi berbasis GAN berhasil mengurangi bias model terhadap kelas mayoritas secara lebih terukur.

Jika dibandingkan dengan beberapa penelitian sebelumnya yang menggunakan arsitektur CNN seperti VGG16, Xception, dan DenseNet untuk klasifikasi penyakit tanaman, hasil penelitian ini menunjukkan performa yang kompetitif dengan tingkat akurasi mencapai 93,29%. Meskipun beberapa penelitian melaporkan akurasi yang tinggi, model yang digunakan umumnya memiliki jumlah parameter yang besar sehingga membutuhkan sumber daya komputasi yang lebih tinggi. Pendekatan yang diusulkan dalam penelitian ini menggunakan EfficientNetB0 sebagai *feature extractor* yang lebih ringan serta dikombinasikan dengan algoritma XGBoost sehingga mampu menghasilkan performa klasifikasi yang baik dengan kompleksitas model yang lebih efisien.

Di sisi lain, performa klasifikasi pada kelas mayoritas, seperti *Healthy* dan *Common Rust*, tetap stabil setelah penerapan augmentasi data berbasis GAN. Hal ini menunjukkan bahwa penambahan citra sintetis tidak menyebabkan terjadinya *overfitting* maupun penurunan kinerja model pada kelas mayoritas. Temuan tersebut mengindikasikan bahwa citra sintetis yang dihasilkan oleh GAN mampu menambah keragaman data latih tanpa mengganggu distribusi data asli pada kelas mayoritas.

Kemiripan visual antara citra sintetis dan citra asli juga menunjukkan bahwa GAN mampu menghasilkan data yang realistis dan representatif, sehingga dapat digunakan secara efektif sebagai data tambahan dalam proses pelatihan tanpa menimbulkan distorsi distribusi data.

Namun demikian, hasil penelitian ini menunjukkan kecenderungan yang berbeda dibandingkan dengan beberapa temuan pada referensi (Wildah et al., 2024) yang menyatakan bahwa citra sintetis hasil GAN tidak selalu memberikan peningkatan performa yang signifikan. Perbedaan hasil ini diduga dipengaruhi oleh perbedaan strategi pemanfaatan citra sintetis, di mana pada penelitian ini citra hasil GAN digunakan secara langsung sebagai data tambahan pada proses pelatihan model klasifikasi, sehingga kontribusinya terhadap peningkatan kinerja model menjadi lebih optimal.

Secara keseluruhan, hasil pembahasan ini memperkuat bukti empiris bahwa penggunaan GAN sebagai metode augmentasi data yang dikombinasikan dengan ekstraksi fitur berbasis EfficientNetB0 dan algoritma klasifikasi XGBoost merupakan pendekatan yang efektif dan efisien dalam meningkatkan performa klasifikasi penyakit daun jagung, khususnya pada kelas minoritas, tanpa menurunkan kinerja klasifikasi pada kelas lainnya. Temuan ini menunjukkan bahwa integrasi augmentasi data berbasis GAN dengan ekstraksi fitur EfficientNetB0 dan klasifikasi XGBoost dapat menjadi pendekatan alternatif yang efektif dalam pengembangan sistem deteksi penyakit tanaman berbasis citra digital. Penerapan augmentasi data berbasis GAN mampu meningkatkan nilai *recall* kelas *Gray Leaf Spot* dari 69% menjadi 78% serta meningkatkan akurasi model dari 92,49% menjadi 93,29%.

## KESIMPULAN

Penelitian ini berhasil mengimplementasikan integrasi GAN, EfficientNetB0, dan XGBoost untuk meningkatkan kinerja klasifikasi penyakit daun jagung berbasis citra digital. Penerapan GAN sebagai metode augmentasi data terbukti efektif dalam mengatasi permasalahan ketidakseimbangan data, khususnya pada kelas *Gray Leaf Spot* yang memiliki jumlah data terbatas. Hasil pengujian menunjukkan adanya peningkatan kinerja model, baik dari sisi akurasi maupun kemampuan dalam mengenali kelas minoritas, yang ditunjukkan oleh peningkatan nilai *recall* dan *F1-score* tanpa menurunkan performa pada kelas mayoritas. Kombinasi EfficientNetB0 sebagai *feature extractor* dan XGBoost sebagai algoritma klasifikasi menghasilkan sistem yang efektif dan efisien dalam melakukan klasifikasi pada dataset yang tidak seimbang. Temuan penelitian ini menunjukkan bahwa integrasi augmentasi data berbasis GAN dengan ekstraksi fitur EfficientNetB0 dan klasifikasi XGBoost dapat menjadi pendekatan alternatif yang potensial dalam pengembangan sistem deteksi penyakit tanaman berbasis citra digital. Namun demikian, penelitian ini memiliki beberapa keterbatasan, antara lain penggunaan dataset yang masih terbatas pada data sekunder, belum dilakukannya uji statistik formal untuk mengukur signifikansi peningkatan performa model, serta belum adanya pengujian langsung pada kondisi lapangan. Oleh karena itu, penelitian selanjutnya dapat mengembangkan evaluasi model yang lebih komprehensif serta menguji implementasi sistem pada data nyata untuk meningkatkan validitas dan keandalan model.

## UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Program Studi Ilmu Komputer Universitas Bumigora yang telah memberikan dukungan akademik dalam pelaksanaan penelitian ini. Penulis juga menyampaikan apresiasi kepada dosen pembimbing yaitu Bapak Gede Yogi Pratama, M.Kom yang telah memberikan arahan dan masukan selama proses penelitian serta penulisan artikel ini.

## DEKLARASI

### Taksonomi Peran Kontributor

Penulis 1: Konseptualisasi, Perancangan metodologi, Pengembangan kode program (coding), Analisis formal, Penulisan – Draf Asli, Visualisasi, serta penyusunan keseluruhan bagian artikel. Penulis 2: Pengumpulan dan penelusuran literatur, Penulisan – Review & Editing, serta membantu proses perbaikan dan penyempurnaan penulisan artikel. Penulis 3: Pengumpulan literatur, Peninjauan kode program (coding review), serta kontribusi dalam proses peninjauan dan evaluasi metodologi penelitian. Penulis 4: Koordinasi seluruh penulis, supervisi proses penulisan artikel, serta kontribusi dalam pengembangan dan peninjauan kode program (coding).

### Pernyataan Pendanaan

Penelitian ini tidak menerima pendanaan khusus dari lembaga pendanaan publik, komersial, atau nirlaba manapun.

### Keuntungan Bersaing

Penulis menyatakan bahwa tidak terdapat keuntungan bersaing dalam penelitian ini.

**DAFTAR PUSTAKA**

- Aisyah, A. S., Nadhiroh, A. Y., and Fajri, F. N. (2025). Klasifikasi Penyakit Daun Jagung Menggunakan Convolutional Neural Network Arsitektur Xception Berbasis Web. *AL-ULUM: Multidisciplinary Journal Of Sciences*, 01(03):227–237.
- Andhika, A. R. and Bayunanda, E. (2025). Studi Literature Review Pemanfaatan Machine Learning dalam Multisektor. *Integrative Perspectives of Social and Science Journal (IPSSJ)*, 2(3):6865–6880.
- Andriani, T. and Iqbal, M. (2025). Penggunaan Transfer Learning untuk Peningkatan Akurasi Deteksi Penyakit Tanaman. *Jatilima : Jurnal Multimedia Dan Teknologi Informasi Volume*, 07(02):90–95. <https://doi.org/10.54209/jatilima.v7i02.1448>.
- Annisa, R., Mardiana, Martinus, Putri, R. A. M., Febriyani, C., and Afif, M. (2024). Pengembangan Model Klasifikasi Citra Tanaman Hutan Melicopelatifolia Berbasis CNN dengan Custom-Built Dataset. *JUKI : Jurnal Komputer dan Informatika*, 6:174–181.
- Ardianto, R. and Wibisono, S. K. (2023). Analisis Deep Learning Metode Convolutional Neural Network Dalam Klasifikasi Varietas Gandum. *Jurnal Kolaboratif Sains*, 6(12):2081–2092. <https://doi.org/10.56338/jks.v6i12.4938>.
- Aryanti, A., Apriliani, D., Putri, W. Z., Ramadhani, D., Malinda, T., and Andreansyah, D. (2025). Identifikasi CNN dalam Deteksi Penyakit Daun Jagung Berbasis Pengenalan Gambar. *MIND (Multimedia Artificial Intelligent Networking Database) Journal MIND*, 10(2):195–205. <https://doi.org/10.26760/mindjournal.v10i2.195-205>.
- Batubara, N. P., Widiyanto, D., and Chamidah, N. (2020). Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB Dan Tekstur GLCM Menggunakan Algoritma Naive Bayes. *JURNAL INFORMATIK*, 16(3):156–163. <https://doi.org/10.52958/iftk.v16i3.2196>.
- Burhanuddin, M. I., Syaifullah, A., Jaya, S. P. A., and Somoal, M. G. (2025). Analisis Komparatif Model MobilenetV1 Dan EfficientnetB0 Dalam Klasifikasi Citra Empat Musim Menggunakan Transfer Learning. *Jurnal Teknik Informatika*, 5(2):508–521. <https://doi.org/10.58794/jekin.v5i2.1378>.
- Erlin, Desnelita, Y., Nasution, N., Suryati, L., and Zoromi, F. (2022). Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang Impact of SMOTE on Random Forest Classifier Performance based on Imbalanced Data. *Matrik: Jurnal Managemen, Teknik Informatika, dan Rekayasa Komputer*, 21(3):667–690. <https://doi.org/10.30812/matrik.v21i3.1726>.
- Hartono, A., Fitriyah, H., and Widasari, E. R. (2023). Sistem Deteksi Tomat Matang Hidroponik berdasarkan Warna Hue , Saturation dan Value menggunakan Metode Threshold berbasis Graphical User Interface. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(5):2313–2319. <https://doi.org/10.46764/teknimedia.v5i2.220>.
- Ht. Barat, A. I. R., Astuti, W. S. A., Wanto, A., and Solikhun, S. (2025). Optimisasi VGG16 dengan Transfer Learning dalam Mendeteksi Penyakit Pada Daun Jagung. *Bulletin of Computer Science Research*, 5(5):1049–1058. <https://doi.org/10.47065/bulletincsr.v5i5.631>.
- Husen, D. (2024). Evaluasi Teknik Augmentasi Data Untuk Klasifikasi Tumor Otak Menggunakan CNN Pada Citra MRI. *TEKNIMEDIA*, 5(2):219–227. <https://doi.org/10.46764/teknimedia.v5i2.220>.
- Jiven, F. F. and Rumini (2025). MRI Classification of Brain Tumors Using EfficientNetB0 Feature Extraction and Machine Learning Methods. *Journal of Applied Informatics and Computing (JAIC)*, 9(6):3394–3404. <https://doi.org/10.30871/jaic.v9i6.10363>.

- Kurniawan, I. R., Fakhrizal, I. U., Cahyadewi, F. A., and Saputra, F. D. (2025). Pengaruh Penambahan Arsitektur Model dalam Klasifikasi Citra Bencana Alam Menggunakan Ensemble Learning. *JURNAL FASILKOM*, 15(2):280–289. <https://doi.org/10.37859/jf.v15i2.9103>.
- Melia, F., Aldian, F. M., Pahlevi, M. S. F., Naufal, R., Risquillah, I., and Oktaffiani, S. (2023). Peran Pemerintah dalam Meningkatkan Volume Ekspor Jagung. *Jurnal ekonomina*, 2:270–284. <https://doi.org/10.55681/economina.v2i1.287>.
- Patmawati, Sunyoto, A., and Luthfi, E. T. (2023). Augmentasi Data Menggunakan DCGAN pada Gambar Tanah. *TEKNIMEDIA*, 4(1):45–52. <https://doi.org/10.46764/teknimedia.v4i1.100>.
- Prasetyo, M. B. and Wulaningrum, R. (2023). Deteksi Penyakit Tanaman Jagung Berbasis Pengolahan Citra Digital Menggunakan Metode Color Moments dan GLCM. *INOTEK*, 7:148–156. <https://doi.org/10.29407/inotek.v7i1.3422>.
- Prayesy, P. A. (2025). Studi Perbandingan Metode Support Vector Machine, Random Forest, dan Convolutional Neural Network untuk Klasifikasi Penyakit Kulit. *Jurnal Kecerdasan Buatan dan Teknologi Informasi Vol.*, 4(1):70–76. <https://doi.org/10.69916/jkbti.v4i1.214STUDI>.
- Purba, M. (2025). Klasifikasi Jenis Daun Tanaman Tropis Menggunakan Model ResNet50 Berbasis Transfer Learning dengan Dataset Tropical Plant Leaf. *JCOSIS (Journal Computer Science and Information Syetem)*, (2):51–57. <https://doi.org/10.61567>.
- Putra, I. P. and Alamsyah, D. (2022). Klasifikasi Penyakit Daun Jagung Menggunakan Metode Convolutional Neural Network. *Jurnal Algoritme*, 2(2):102–112. <https://doi.org/10.35957/algoritme.v2i2.2360>.
- Rahmanita, E., Negara, Y. D. P., Kustiyahningsih, Y., Sasmeka, V., and Khotimah, B. K. (2023). Implementasi Metode Naive Bayes dan Information Gain Untuk Klasifikasi Penyakit dan Hama Tanaman Jagung Implementation of Naive Bayes Method and Information Gain for Classification of Diseases and Pests of Corn Plants. *TEKNIKA*, 12(November):198–204. <https://doi.org/10.34148/teknika.v12i3.684>.
- Ryanto, S. S. and Sari, I. P. (2024). Optimasi Rekonstruksi Dan Pengenalan Wajah Berbasis Generative Adversarial Network (GAN) Dan CNN. *JEKIN (Jurnal Teknik Informatika) Generative*, 4(3):769–777. <https://doi.org/10.58794/jekin.v4i3.1081>.
- Sarah, S. and Guntaro (2023). Identifikasi Penyakit Tanaman Jagung Berdasarkan Citra Daun Tinjauan Literatur Sistematis (SLR). *Prosiding-Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMASTER)*, 2(1):278–289.
- Sarasati, F., Nugraha, F. S., and Radiyah, U. (2022). Pemanfaatan Metode Deep Learning untuk Klasifikasi Penyakit pada Tanaman Jagung. *Jurnal Infortech*, 4(2):133–138.
- Suhesi, Kurnia, D. A., Wijaya, Y. A., Anwar, S., and Hayati, U. (2025). Peningkatan Robustness Model Klasifikasi Penyakit Daun Semangka Menggunakan Fine-Tuning Vgg16 dan Augmentasi Data Sintetis Berbasis Generative Adversarial Network. *Journal Of Computer Science and Artificial Intelligence (JCSAI)*, 06(02):1–8. <https://doi.org/10.32485/jcsai.Kata>.
- Tirta, M. S., Kurniawan, R., and Zulus, A. (2025). Model Transfer Learning Dalam Klasifikasi Penyakit Daun Jagung Menggunakan Arsitektur Densenet-201. *Jurnal of Informatics, Electrical and Electronics Engineering*, 4(3):91–99. <https://doi.org/10.47065/jieeee.v4i3.2347>.
- Tono, Ariani, M., and Suryana, A. (2023). Kinerja Ketahanan Pangan Indonesia : Pembelajaran Dari Penilaian dengan Kriteria Global dan Nasional. *Analisis Kebijakan Pertanian*, 21(1):1–20. <https://doi.org/10.21082/akp.v21n1.2023.1-20>.

- Wildah, S. K., Latif, A., and Haryanto, T. (2024). Sintesa Citra Daun Kopi Menggunakan Generative Adversarial Network pada Dataset Penyakit Daun Kopi. *INTI NUSA MANDIRI*, 19(1):23–30. <https://doi.org/10.33480/inti.v19i1.5045>.
- Wirabowo, I. and Susilawati, I. (2025). Implementasi Convolution Neural Network ( CNN ) untuk Deteksi Penyakit pada Daun Jagung Berbasis Citra Digital. *Jurnal Pustaka Data*, 5(1):233–241.
- You, A., Kim, J. K., Ryu, I. H., and Yoo, T. K. (2022). Application of generative adversarial networks ( GAN ) for ophthalmology image domains : a survey. *Eye and Vision*, 9(6):2–19. <https://doi.org/10.1186/s40662-022-00277-3>.
- Yulianti, S. E. H., Soesanto, O., and Sukmawaty, Y. (2022). Penerapan Metode Extreme Gradient Boosting ( XGBOOST ) pada Klasifikasi Nasabah Kartu Kredit. *JOMTA: Journal of Mathematics: Theory and Applications*, 4(1):21–26. <https://doi.org/10.31605/jomta.v4i1.1792>.
- Yuninsar, E., Hasmin, E., and Samsie, I. (2024). Analisis Perbandingan Metode Convolutional Neural Network (CNN) dan Artificial Neural Network (ANN) dalam Klasifikasi Penyakit Daun Jagung. *Jurnal Ilmiah Betrik*, 15(03):321–333. <https://doi.org/10.36050/kx56ex79>.