



Perbandingan Model LightGBM dan CatBoost pada Klasifikasi Penyakit Daun Kacang Tanah Berbasis Ekstraksi Fitur EfficientNet-B2

Safiqur Rohman¹, Eva Yulia Puspaningrum², Ani Dijah Rahajoe³

^{1,2,3}Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia

E-mail: 21081010109@student.upnjatim.ac.id, evapuspaningrum.if@upnjatim.ac.id, anidijah.if@upnjatim.ac.id

Article Info	Abstract
Article History Received: 2025-11-05 Revised: 2025-12-19 Published: 2026-01-16 Keywords: <i>Image Classification;</i> <i>Efficientnet-B2;</i> <i>Lightgbm;</i> <i>Catboost;</i> <i>Peanut Leaf Disease.</i>	Peanut is one of Indonesia's key sources of plant-based protein and oil, yet its productivity continues to decline due to leaf diseases caused by fungal infections and nutrient deficiencies. Rapid image-based disease detection is crucial for preventing yield losses and supporting early intervention. This study compares the performance of two gradient boosting algorithms Light Gradient Boosting Machine and Categorical Boosting in classifying peanut leaf diseases using image features extracted with the EfficientNet-B2 architecture. The dataset consists of 1,876 images across six classes, including healthy leaves and multiple disease types. The preprocessing stage includes image resizing, pixel normalization, and image augmentation to enhance model generalization. Features extracted by EfficientNet-B2 were classified using both algorithms and evaluated through accuracy, precision, recall, and F1-score. Experimental results show that Light Gradient Boosting Machine achieved the highest accuracy of 96.20% with a training time of 58.55 seconds, while Categorical Boosting reached 95.65% accuracy with a training time of 304.80 seconds. These findings highlight that Light Gradient Boosting Machine offers better computational efficiency and slightly superior accuracy. Overall, this hybrid approach demonstrates strong potential for supporting automatic and reliable peanut leaf disease identification, which can aid farmers in early detection and management.

Artikel Info	Abstrak
Sejarah Artikel Diterima: 2025-11-05 Direvisi: 2025-12-19 Dipublikasi: 2026-01-16 Kata kunci: <i>Klasifikasi Citra;</i> <i>Efficientnet-B2;</i> <i>Lightgbm;</i> <i>Catboost;</i> <i>Penyakit Daun Kacang Tanah.</i>	Kacang tanah merupakan salah satu sumber protein dan minyak nabati penting di Indonesia, namun produktivitasnya menurun akibat berbagai penyakit daun yang dipicu oleh infeksi jamur maupun kekurangan unsur hara. Deteksi penyakit berbasis citra secara cepat dan akurat menjadi langkah penting untuk mencegah kerugian hasil panen dan mendukung pengambilan keputusan dini. Penelitian ini membandingkan kinerja dua algoritma gradient boosting Light Gradient Boosting Machine dan Categorical Boosting dalam mengklasifikasikan penyakit daun kacang tanah berdasarkan fitur citra yang diekstraksi menggunakan arsitektur EfficientNet-B2. Dataset terdiri atas 1.876 citra dalam enam kelas, mencakup daun sehat dan beberapa jenis penyakit. Tahapan prapemrosesan meliputi penyesuaian ukuran citra, normalisasi piksel, dan augmentasi citra untuk meningkatkan kemampuan generalisasi model. Fitur yang dihasilkan EfficientNet-B2 diklasifikasikan menggunakan kedua algoritma tersebut dan dievaluasi dengan metrik akurasi, presisi, recall, serta F1-score. Hasil eksperimen menunjukkan bahwa Light Gradient Boosting Machine mencapai akurasi tertinggi sebesar 96,20% dengan waktu pelatihan 58,55 detik, sedangkan Categorical Boosting memperoleh akurasi 95,65% dengan waktu pelatihan 304,80 detik. Temuan ini menegaskan bahwa Light Gradient Boosting Machine lebih efisien secara komputasi sekaligus sedikit lebih unggul dalam akurasi. Secara keseluruhan, pendekatan hibrid ini berpotensi mendukung identifikasi otomatis penyakit daun kacang tanah secara lebih cepat dan andal, sehingga bermanfaat bagi petani dalam pemantauan kesehatan tanaman.

I. PENDAHULUAN

Kacang tanah merupakan tanaman biji-bijian yang banyak dibudidayakan di Indonesia sebagai sumber lemak nabati dan protein. Komoditas ini memiliki nilai ekonomi yang cukup tinggi, di mana pada Juli 2021 harga kacang tanah di Jawa Timur mencapai Rp 26.194/kg, lebih tinggi dibandingkan kedelai (Rp 12.833/kg) dan kacang

hijau (Rp 23.542/kg) (Jawa Timur, 2021). Berdasarkan data Badan Pusat Statistik (2018), produksi kacang tanah menurun dari 638.896 ton pada tahun 2014 menjadi 512.198 ton. Penurunan ini berhubungan dengan serangan patogen, bakteri, dan nematoda pada akar tanaman. Selain itu, kekurangan unsur hara untuk sintesis metabolit sekunder menyebabkan

tanaman menjadi lebih rentan terhadap penyakit (Duaja MD, Aziz SA, 2012).

Penyakit pada tanaman kacang tanah umumnya dapat dikenali melalui gejala yang muncul pada bagian daun, sehingga daun menjadi indikator visual yang paling mudah diamati. Beberapa penyakit daun yang banyak ditemui, seperti karat daun, bercak daun, serta gejala kekurangan nutrisi, dapat menurunkan hasil panen secara signifikan. Penyakit karat daun dan bercak daun bahkan dapat menyebabkan kerugian hingga 50% (Hardaningsih dan Sumartini, 2015). Karena itu, deteksi penyakit secara dini menjadi langkah penting untuk mencegah penurunan produktivitas.

Namun, diagnosis penyakit yang selama ini dilakukan secara manual oleh petani kerap tidak akurat karena kemiripan gejala antar penyakit. Berdasarkan wawancara langsung, banyak petani masih mengalami kesulitan dalam membedakan jenis penyakit daun kacang tanah. Kondisi ini menunjukkan perlunya dukungan teknologi berbasis kecerdasan buatan untuk membantu identifikasi penyakit secara lebih cepat dan tepat.

Pemanfaatan citra digital dan kecerdasan buatan telah menunjukkan potensi besar dalam bidang pertanian, khususnya dalam identifikasi penyakit tanaman. Metode Convolutional Neural Network (CNN) menjadi salah satu pendekatan yang banyak digunakan karena kemampuannya mengekstraksi fitur visual secara otomatis. Salah satunya adalah arsitektur EfficientNet, yang mengoptimalkan kedalaman (depth), lebar (width), dan resolusi (resolution) jaringan secara seimbang. EfficientNet-B2 digunakan dalam penelitian ini untuk mengekstraksi fitur citra daun kacang tanah karena menawarkan keseimbangan yang baik antara akurasi dan efisiensi komputasi (A. Y. Pratama, R. S. Wahyuni, 2022).

Fitur yang diekstraksi tersebut kemudian diklasifikasikan menggunakan dua algoritma gradient boosting, yakni LightGBM dan CatBoost. LightGBM dikenal memiliki keunggulan dalam efisiensi proses pelatihan dan konsumsi memori, serta mampu memberikan akurasi tinggi (T.A.B.W and Others., 2025). Sementara itu, CatBoost memiliki keunggulan dalam menangani fitur kategorikal tanpa encoding yang kompleks serta lebih tahan terhadap overfitting melalui metode ordered boosting (R. Irfannandhy, L. B. Handoko, 2024). Kedua algoritma ini dipilih berdasarkan performa unggul yang telah dibuktikan dalam berbagai studi sebelumnya, terutama untuk kasus klasifikasi yang kompleks.

Penelitian sebelumnya juga menunjukkan relevansi penggunaan EfficientNet dan metode gradient boosting. Misalnya, Anggiratih et al. (2021) menerapkan EfficientNet-B3 dengan pendekatan transfer learning untuk klasifikasi penyakit daun padi, yaitu brown spot dan bacterial leaf, dan memperoleh akurasi 79,53%, lebih tinggi dibandingkan MobileNet-V3 (E. Anggiratih, S. Siswanti, S. K. Octaviani, 2021). Selain itu, Royana et al. (2025) membandingkan LightGBM dan CatBoost pada kasus prediksi customer churn dan menemukan bahwa kedua model memiliki akurasi serupa sebesar 84%, namun CatBoost unggul pada nilai recall sebesar 0,86 dibandingkan LightGBM yang memperoleh 0,84. Keunggulan tersebut berasal dari kemampuannya menangani fitur kategorikal tanpa proses encoding yang rumit dan stabil terhadap overfitting melalui ordered boosting (Akge Ninov Royana, Yisti Vita Via, 2025).

Berdasarkan temuan-temuan tersebut, penelitian ini menggunakan LightGBM dan CatBoost untuk membandingkan performa klasifikasi citra penyakit daun kacang tanah berdasarkan fitur yang diekstraksi menggunakan EfficientNet-B2. Tujuan penelitian ini adalah memperoleh model yang paling akurat dan efisien dalam mendeteksi penyakit daun kacang tanah sehingga dapat mendukung proses identifikasi secara otomatis dan membantu meningkatkan produktivitas pertanian.

II. METODE PENELITIAN

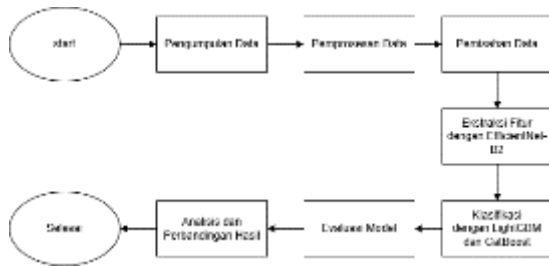
1. Desain Penelitian

Penelitian ini menggunakan metode eksperimen komparatif untuk membandingkan kinerja LightGBM dan CatBoost dalam mengklasifikasikan penyakit daun kacang tanah. Proses penelitian dilakukan melalui beberapa tahapan yang saling berurutan agar hasil dapat diuji dan direplikasi. Tahap pertama adalah pengumpulan data citra, kemudian dilanjutkan dengan preprocessing, yang meliputi penyesuaian ukuran gambar, normalisasi nilai piksel, dan augmentasi untuk menambah variasi data. Setelah itu, dataset dibagi menjadi data latih, validasi, dan uji menggunakan *stratified split* agar proporsi tiap kelas tetap seimbang.

Berikutnya, citra diproses menggunakan EfficientNet-B2 dengan bobot pralatih ImageNet yang digunakan sebagai *feature extractor* tanpa fine-tuning. Vektor fitur hasil ekstraksi kemudian digunakan pada proses pelatihan model LightGBM dan CatBoost. Kedua model dioptimasi dengan Grid Search

menggunakan 3-fold cross-validation, dan parameter terbaik dipilih berdasarkan nilai akurasi.

Tahap akhir adalah evaluasi model menggunakan data uji. Kinerja model diukur dengan akurasi, presisi, recall, dan F1-score yang dihitung dari confusion matrix untuk mengetahui kemampuan kedua algoritma dalam mengidentifikasi penyakit secara akurat.



Gambar 1. Alur Penelitian.

Gambar 1 menunjukkan alur penelitian mulai dari pengumpulan data, preprocessing, pembagian data, ekstraksi fitur menggunakan EfficientNet-B2, proses klasifikasi menggunakan LightGBM dan CatBoost, serta tahap evaluasi model menggunakan metrik akurasi, presisi, recall, dan F1-Score.

2. Dataset

Dataset yang digunakan terdiri dari citra daun kacang tanah yang dikategorikan menjadi enam kelas: bercak daun awal, bercak daun lanjut, daun sehat, karat daun awal, karat daun lanjut, dan kekurangan nutrisi. Seluruh citra diperoleh melalui pengambilan langsung di lahan pertanian menggunakan kamera smartphone dan disimpan dalam format .jpg sebelum diproses lebih lanjut.

3. Preprocessing

Tahapan pertama dalam penelitian ini adalah dengan melakukan preprocessing atau pengolahan data.

a) Ilustrasi citra daun kacang tanah yang terbagi dalam 5 kelas.



Gambar 2. Ilustrasi Citra Daun Kacang Tanah

Pada Gambar 2 diperlihatkan contoh citra daun kacang tanah dari enam kelas penyakit. Ilustrasi tersebut digunakan untuk memberikan gambaran umum mengenai variasi visual dari tiap kelas citra sebelum dilakukan preprocessing.

Tabel 1. Jumlah Dataset

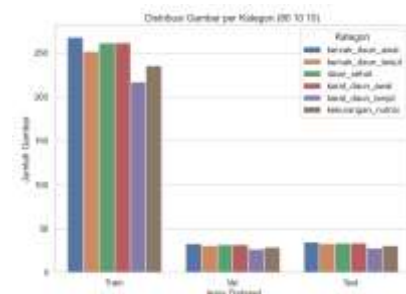
Kelas	Jumlah
Sehat	328
Kekurangan Nutrisi	296
Bercak daun awal	336
Bercak daun lanjut	316
Karat daun awal	328
Karat Daun Lanjut	272
Total	1,876

Tabel 1 menunjukkan distribusi jumlah citra pada setiap kelas penyakit daun kacang tanah, yaitu Sehat, Kekurangan Nutrisi, Bercak Daun Awal, Bercak Daun Lanjut, Karat daun Lanjut, dan Karat Daun Awal. Total keseluruhan data sebanyak 1.876 citra dengan jumlah yang relatif seimbang di tiap kelas, sehingga mendukung proses pelatihan model secara optimal.

Setelah data telah siap, maka selanjutnya dilakuak pengelolaan data dengan beberapa bagian.

b) Pembagian Data

Pembagian data merupakan tahap awal yang penting untuk memastikan model dapat belajar dengan baik. Dataset dibagi menjadi tiga subset menggunakan teknik Stratified Shuffle Split, yaitu 80% untuk data latih, 10% untuk data validasi, dan 10% untuk data uji. Pembagian secara stratified memastikan bahwa proporsi setiap kelas tetap seimbang di seluruh subset, sehingga model dapat belajar secara adil dan tidak bias terhadap kelas tertentu.



Gambar 3. Pembagian Data

Gambar 3 menunjukkan hasil pembagian dataset dengan proporsi 80%

untuk data latih, 10% untuk data validasi, dan 10% untuk data uji. Pembagian ini berfungsi untuk melatih model agar mengenali pola dari data latih, menyesuaikan parameter melalui data validasi, serta menguji kemampuan generalisasi model menggunakan data uji yang belum pernah digunakan sebelumnya

c) Augmentasi

Augmentasi dilakukan untuk memperbanyak variasi data tanpa harus menambah jumlah citra asli. Teknik ini membantu model menjadi lebih robust dan mampu mengenali pola citra dari berbagai sudut, maupun posisi. Pada penelitian ini dilakukan dengan tiga Teknik augmentasi. Augmentasi hanya dilakukan pada data latih dengan Kombinasi augmentasi yaitu, horizontal vlip dengan probabilitas 0.5, vertical vlip dengan probabilitas 0.5, dan rotasi 45° pada gambar secara real-time. Sedangkan untuk subset validasi dan tes tidak dilakukan augmentasi agar data validasi dan tes tetap original.

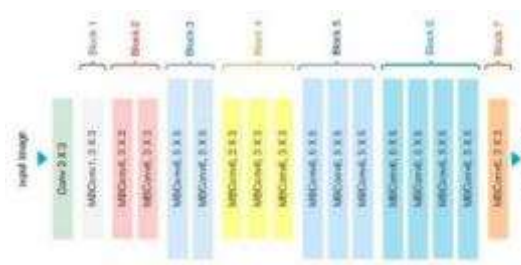
d) Resize dan Normalisasi

Tahap resize dan normalisasi dilakukan untuk menyeragamkan ukuran serta skala nilai piksel pada setiap citra. Langkah ini memastikan data siap diproses oleh model dan meningkatkan efisiensi komputasi selama pelatihan. Citra dilakukan resize ukuran pikselnya menjadi 260x260 sesuai input yang diharapkan oleh arsitektur efficientnet-B2 dan setelah itu dilakukan normalisasi nilai piksel ke rentang [0-1] yang sebelumnya 0-225. Proses normalisasi dilakukan untuk menyesuaikan skala nilai piksel agar jaringan Efficientnet-b2 dapat memproses citra dengan lebih stabil dan efisien.

4. Efficientnet-B2

EfficientNet-B2 merupakan arsitektur Convolutional Neural Network (CNN) yang dirancang dengan pendekatan *compound scaling*, yaitu skema penskalaan yang menyeimbangkan kedalaman, lebar, dan resolusi jaringan secara serentak agar memperoleh kinerja optimal dengan kompleksitas komputasi yang efisien. Dalam penelitian ini, EfficientNet-B2 digunakan sebagai feature extractor dengan memanfaatkan pretrained weights ImageNet, sehingga model tidak dilatih dari awal (random initialization) dan dapat mengekstraksi karakteristik visual daun secara lebih stabil.

Seluruh lapisan convolutional EfficientNet-B2 dibekukan (frozen), sehingga tidak dilakukan fine-tuning selama pelatihan. Pendekatan ini dipilih untuk menjaga konsistensi representasi fitur serta mengurangi risiko overfitting pada dataset berukuran menengah. Setelah melalui proses ekstraksi, keluaran dari lapisan Global Average Pooling berbentuk vektor fitur berdimensi tetap yang kemudian digunakan sebagai input bagi algoritma LightGBM dan CatBoost untuk melakukan klasifikasi penyakit daun kacang tanah.



Gambar 4. Arsitektur Efficientnet-B2

Pada Gambar 4 merupakan struktur arsitektur EfficientNet-B2 yang terdiri dari serangkaian blok konvolusi (MBConv) yang tersusun secara berurutan dari awal hingga tahap akhir feature map. Setiap lapisan blok memiliki fungsi spesifik dalam mengenali dan mengekstraksi ciri-ciri esensial dari gambar daun, meliputi aspek tekstur, warna, serta bentuk dari gejala penyakit (Hafizh Kennandya Maulana, Henni Endah Wahanani2, 2025).

5. Klasifikasi

Setelah proses ekstraksi fitur menggunakan arsitektur EfficientNet-B2 menghasilkan representasi fitur dalam bentuk vektor, tahap berikutnya adalah klasifikasi. Pada tahap ini digunakan dua algoritma *gradient boosting*, yaitu LightGBM dan CatBoost. Kedua algoritma tersebut berfungsi untuk mempelajari pola dari hasil ekstraksi fitur dan mengelompokkannya ke dalam kelas penyakit yang sesuai berdasarkan karakteristik visual daun kacang tanah.

LightGBM merupakan algoritma machine learning berbasis pohon keputusan yang dirancang untuk menangani dataset berukuran besar secara efisien. Dengan memanfaatkan metode gradient boosting serta strategi pertumbuhan pohon secara leaf-wise, algoritma ini mampu memberikan proses pelatihan yang cepat sekaligus hasil prediksi yang akurat. Selain itu, LightGBM

dapat digunakan untuk tugas klasifikasi, regresi, dan evaluasi pentingnya fitur, sehingga sangat sesuai untuk penerapan pada analisis data berskala besar dan kompleks (J. Yan et al, 2021).

Selain itu dilakukan klasifikasi dengan CatBoost. Dimana model ini merupakan algoritma gradient boosting yang berbasis decision tree dan mampu mendukung regresi serta menangani data kategorikal secara otomatis menggunakan pendekatan statistik. Untuk mencegah terjadinya overfitting, algoritma ini melakukan optimasi pada parameter input dan menerapkan teknik permutasi acak serta perhitungan rata-rata label dalam memproses data kategorikal (A. F. Istianto, A. I. Hadiana, 2023).

Sebelum dilakukan proses pelatihan akhir, kedua model melalui tahap optimasi hiperparameter menggunakan metode Grid Search yang dikombinasikan dengan 3-fold cross-validation. Setiap kombinasi parameter diuji pada tiga pembagian data pelatihan yang berbeda untuk memastikan bahwa pemilihan parameter terbaik tidak bergantung pada satu komposisi data tertentu. Pemilihan konfigurasi optimal dilakukan berdasarkan nilai akurasi rata-rata dari seluruh lipatan validasi, sehingga parameter yang dipilih merupakan kombinasi yang memberikan performa paling stabil pada proses klasifikasi.

Tabel 2. Parameter LightGBM

Parameter	Nilai Uji
learning_rate	0.01, 0.1
max_depth	5, 7
n_estimators	50, 100

Tabel 2 merupakan Kombinasi parameter untuk LightGBM, Parameter learning_rate berfungsi untuk mengatur laju pembelajaran model, max_depth berperan dalam mengontrol kedalaman maksimal pohon keputusan, sedangkan n_estimators menentukan jumlah pohon yang akan dibangun. Berbagai kombinasi nilai dari parameter-parameter tersebut diuji untuk mendapatkan performa optimal dari model.

Tabel 3. Parameter CatBoost

Parameter	Nilai Uji
learning_rate	0.01, 0.1
depth	5, 7
iterations	50, 100
subsample	0.8, 1.0

Tabel 3 menunjukkan berbagai kombinasi parameter yang diujicobakan dalam proses optimasi model CatBoost, meliputi learning_rate, depth, iterations, dan subsample, yang masing-masing berpengaruh terhadap kecepatan pembelajaran, kompleksitas struktur, dan tingkat akurasi model. Pengujian ini bertujuan untuk mengidentifikasi konfigurasi parameter optimal yang dapat meningkatkan performa model secara keseluruhan.

6. Evaluasi Model

Evaluasi model merupakan tahap terakhir yang dilakukan untuk mengukur performa dari setiap model klasifikasi. Tahap ini dilakukan terhadap model yang telah dikembangkan untuk memverifikasi bahwa model tersebut telah mencapai tujuan penelitian dan menghasilkan output yang sesuai harapan. Evaluasi dilakukan dengan memanfaatkan confusion matrix untuk menghitung metrik performa, meliputi akurasi, presisi, recall, dan f1-score yang diperoleh dari hasil prediksi klasifikasi menggunakan algoritma gradient boosting. Perhitungan setiap metrik tersebut menggunakan formula sebagai berikut (I. P. Putri, 2021).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1-Score} = \frac{2 \times \text{presisi} \times \text{recall}}{\text{presisi} + \text{recall}} \quad (4)$$

Keterangan:

- a) True Positive (TP): Pelanggan churn yang berhasil memprediksi churn
- b) True Negative (TN): Pelanggan tidak churn yang berhasil memprediksi tidak churn
- c) False Positive (FP): Citra daun yang bukan termasuk kelas penyakit tertentu tetapi diprediksi sebagai penyakit tersebut.
- d) False Negative (FN): Citra daun yang sebenarnya termasuk kelas penyakit tertentu, namun gagal diidentifikasi oleh model.

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

Bagian ini menyajikan hasil klasifikasi penyakit daun kacang tanah menggunakan dua algoritma gradient boosting, yaitu LightGBM dan CatBoost. Kedua model memanfaatkan fitur citra yang diekstraksi menggunakan EfficientNet-B2, kemudian dilakukan pencarian parameter terbaik melalui Grid Search dengan 3-fold cross-

validation. Evaluasi dilakukan berdasarkan akurasi, presisi, recall, F1-score, confusion matrix, serta grafik accuracy dan log-loss untuk menilai stabilitas pembelajaran model.

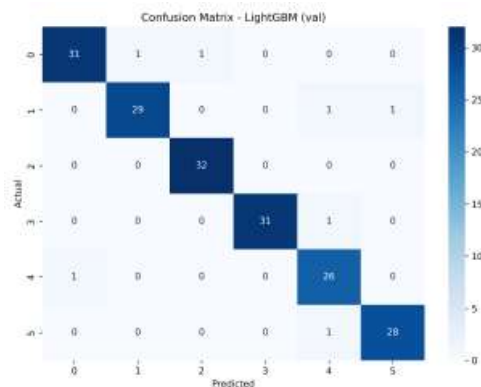
Setelah seluruh kombinasi parameter diuji, diperoleh konfigurasi terbaik untuk masing-masing model sebagaimana ditunjukkan pada Tabel 4 berikut.

Tabel 4. Parameter Terbaik

Metode	Parameter	Nilai
LightGBM	learning_rate	0.1
	max_depth	5
	n_estimators	100
CatBoost	learning_rate	0.1
	depth	7
	iterations	100
	subsample	0.8

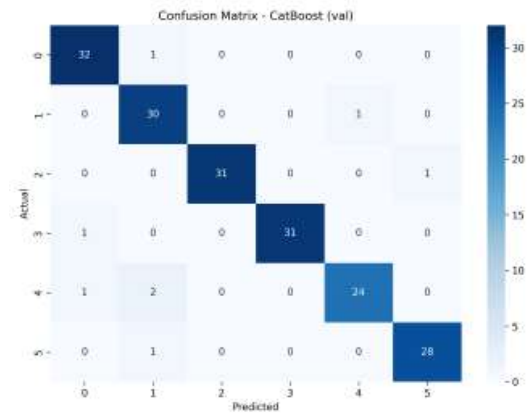
Tabel 4 di atas merupakan Kombinasi parameter yang digunakan dalam pelatihan akhir pada dataset. Model LightGBM dengan parameter tersebut menghasilkan keseimbangan optimal antara kecepatan pelatihan dan akurasi, sementara CatBoost menunjukkan kestabilan hasil yang baik dengan risiko overfitting yang rendah.

Visualisasi hasil klasifikasi ditunjukkan melalui *confusion matrix*.



Gambar 5. Confusion matrix LightGBM

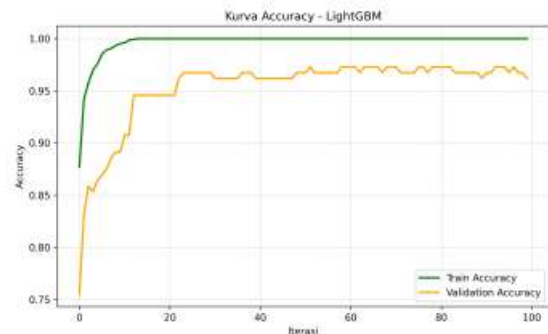
Gambar 5 menunjukkan confusion matrix model LightGBM. LightGBM mampu mengklasifikasikan sebagian besar citra dengan benar pada seluruh kelas. Kesalahan klasifikasi hanya muncul dalam jumlah kecil, terutama pada kelas bercak daun awal dan kekurangan nutrisi, yang memiliki pola visual serupa. Pola diagonal yang dominan menegaskan kemampuan LightGBM dalam mengenali ciri khas tiap penyakit.



Gambar 6. Confusion Matrix CatBoost

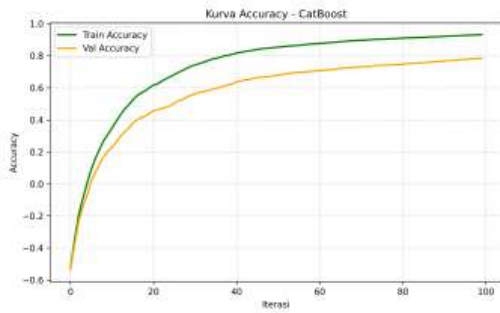
Gambar 6 memperlihatkan hasil confusion matrix model CatBoost yang juga menunjukkan performa tinggi dengan mayoritas prediksi tepat pada diagonal utama. Meski demikian, CatBoost masih mengalami sedikit kesalahan pada beberapa kelas yang memiliki kemiripan pola visual antar daun. Secara keseluruhan, model ini tetap konsisten dan stabil dalam mengklasifikasikan citra pada seluruh kelas penyakit.

Selain itu performa model juga divisualisasikan melalui grafik nilai *accuracy* dan *log-loss* seperti berikut:



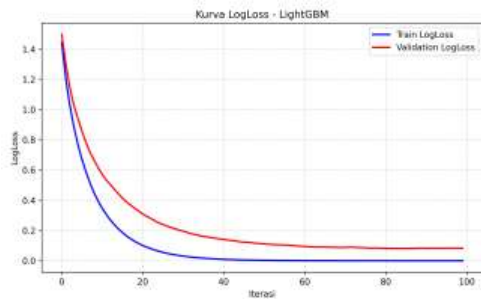
Gambar 7. Accuracy LightGBM

Visualisasi Gambar 7 menunjukkan pola peningkatan akurasi selama proses pelatihan dan validasi model. Kurva akurasi pelatihan meningkat tajam mencapai nilai mendekati 1.0, sementara akurasi validasi stabil di sekitar 0.96 yang menunjukkan model memiliki kemampuan dalam generalisasi yang baik dengan belajar data train tanpa adanya overfitting.



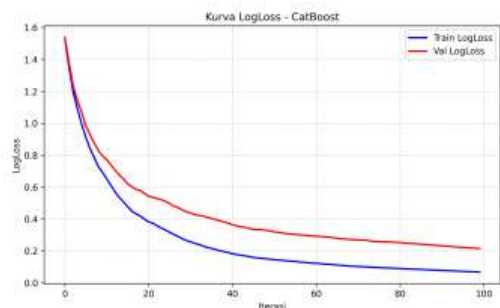
Gambar 8. Accuracy CatBoost

Gambar 8 menampilkan perkembangan akurasi model yang meningkat secara bertahap selama proses iterasi. Tren akurasi train dan validasi meningkat secara konsisten yang menandakan bahwa model mampu mencapai akurasi terbaik seiring meningkatnya iterasi.



Gambar 9. Logloss LightGBM

Gambar 9 di atas merupakan grafik LightGBM dari *logloss* pada data pelatihan dan validasi menurun tajam di iterasi awal kemudian melandai mencapai konvergensi. Penurunan ini mengindikasikan model semakin efektif meminimalkan kesalahan prediksi dengan nilai log loss yang rendah, menunjukkan kestabilan model dalam mempelajari pola data.



Gambar 10. Logloss CatBoost

Visualisasi Gambar 10 menunjukkan penurunan log loss pada data pelatihan dan validasi model secara stabil, mengindikasikan proses pembelajaran yang baik dan

konvergen. Nilai log loss yang menurun secara konsisten menandakan model mampu meminimalkan kesalahan prediksi secara bertahap pada klasifikasi citra daun.

Hasil evaluasi model berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score* dapat dilihat pada Tabel 5 berikut.

Tabel 5. Hasil Evaluasi Model

Model	LightGBM	CatBoost
Akurasi	0.9620	0.9565
Presisi	0.96	0.96
Recall	0.96	0.95
F1-score	0.96	0.96
Waktu pelatihan	58.55 detik	304.80detik

Berdasarkan hasil pada Tabel 5, LightGBM memberikan akurasi tertinggi yaitu 96.20% dengan waktu pelatihan yang jauh lebih cepat, sehingga lebih efisien secara komputasi. CatBoost menghasilkan akurasi 95.65% dengan nilai presisi, recall, dan F1-score yang relatif seimbang, namun membutuhkan waktu pelatihan yang lebih lama. Temuan ini mengindikasikan bahwa LightGBM lebih unggul dalam konteks dataset citra penyakit daun kacang tanah, terutama dari segi efisiensi dan kecepatan pemrosesan.

B. Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi ekstraksi fitur menggunakan EfficientNet-B2 dan algoritma gradient boosting mampu memberikan performa klasifikasi yang tinggi pada identifikasi penyakit daun kacang tanah. Temuan ini sejalan dengan penelitian Anggiratih et al. (2021) yang juga menunjukkan bahwa fitur hasil ekstraksi EfficientNet memiliki kemampuan representasi visual yang kuat pada kasus penyakit tanaman. Selain itu, efektivitas LightGBM dan CatBoost pada penelitian ini konsisten dengan studi Royana et al. (2025) yang melaporkan bahwa kedua algoritma tersebut mampu bekerja optimal pada data dengan karakteristik kompleks melalui pendekatan boosting berbasis pohon.

Secara teknis, Performa LightGBM yang mencapai akurasi tertinggi dipengaruhi oleh distribusi dataset yang seimbang serta kemampuan *leaf-wise tree growth* dalam memanfaatkan fitur terstruktur dari EfficientNet-B2. Hal ini membuat model lebih adaptif dalam membedakan pola tekstur halus antar kelas. Di sisi lain, CatBoost tetap menunjukkan stabilitas tinggi melalui

mekanisme *ordered boosting*, meskipun waktu pelatihannya lebih lama. Kesalahan klasifikasi pada kedua model umumnya muncul pada kelas yang memiliki kemiripan visual, seperti bercak daun awal dan kekurangan nutrisi, yang secara morfologi memang sulit dibedakan.

Perbedaan performa kedua model dipengaruhi oleh karakteristik fitur hasil ekstraksi EfficientNet-B2 dan kompleksitas antar kelas pada dataset. LightGBM dapat memanfaatkan fitur yang kaya secara representatif melalui mekanisme *leaf-wise growth* sehingga menghasilkan akurasi tertinggi. Sebaliknya, CatBoost yang lebih konservatif tetap menunjukkan stabilitas berkat *ordered boosting*, meskipun membutuhkan waktu pelatihan yang lebih panjang. Hal ini menjelaskan mengapa keduanya berkinerja baik namun memiliki keunggulan teknis yang berbeda.

Hasil ini memiliki implikasi praktis bagi pengembangan sistem deteksi penyakit tanaman berbasis citra. Pendekatan hybrid yang memadukan ekstraksi fitur EfficientNet-B2 dengan algoritma boosting terbukti lebih efisien dibandingkan penggunaan CNN end-to-end pada dataset terbatas. Metode ini ringan secara komputasi, mudah diintegrasikan ke aplikasi diagnosis cepat berbasis smartphone, dan memberikan konfirmasi bahwa kombinasi deep feature extraction dan gradient boosting efektif untuk klasifikasi penyakit daun kacang tanah.

IV. SIMPULAN DAN SARAN

A. Simpulan

Penelitian ini bertujuan untuk membandingkan performa LightGBM dan CatBoost dalam klasifikasi penyakit daun kacang tanah berdasarkan fitur yang diekstraksi menggunakan EfficientNet-B2, dan tujuan tersebut telah tercapai. EfficientNet-B2 berhasil menghasilkan representasi fitur yang informatif sehingga kedua model mampu melakukan klasifikasi dengan akurasi tinggi.

Hasil menunjukkan bahwa LightGBM menjadi model terbaik dengan akurasi 96,20% dan waktu pelatihan yang jauh lebih cepat (58,55 detik), sedangkan CatBoost tetap kompetitif dengan akurasi 95,65% dan kestabilan hasil yang baik meskipun memerlukan waktu pelatihan lebih lama (304,80 detik). Temuan ini menegaskan bahwa LightGBM lebih unggul dari sisi efisiensi dan akurasi.

Secara praktis, pendekatan hybrid yang memadukan EfficientNet-B2 dengan algoritma boosting berpotensi diterapkan pada sistem deteksi penyakit tanaman berbasis perangkat mobile atau IoT karena kebutuhan komputasinya yang relatif ringan. Dengan demikian, metode ini dapat mendukung proses identifikasi penyakit secara cepat dan akurat di lingkungan pertanian.

B. Saran

Penelitian ini masih terbatas pada enam kelas penyakit daun kacang tanah, sehingga penelitian selanjutnya disarankan menambah variasi jenis penyakit maupun komoditas tanaman agar model memiliki cakupan generalisasi yang lebih luas. Selain itu, dataset dapat diperbesar agar model mampu mempelajari variasi visual yang lebih kompleks dan mengurangi potensi kesalahan pada kelas yang memiliki kemiripan morfologi.

Dari sisi teknis, pengoptimalan model dapat ditingkatkan dengan menerapkan metode optimasi hiperparameter yang lebih canggih seperti Bayesian Optimization, Random Search, atau Hyperband untuk memperoleh konfigurasi yang lebih optimal daripada Grid Search. Eksplorasi arsitektur ekstraksi fitur lain, seperti EfficientNet varian lain, Vision Transformer (ViT), atau Swin Transformer, juga berpotensi meningkatkan kualitas fitur untuk klasifikasi yang lebih presisi.

Dalam pengembangan aplikasi, integrasi model ke dalam perangkat mobile atau sistem IoT dapat menjadi langkah lanjutan agar deteksi penyakit dapat dilakukan langsung di lapangan secara real-time. Pengembangan antarmuka web atau mobile akan memudahkan petani maupun penyuluh dalam memanfaatkan sistem ini dan mendukung penerapan pertanian presisi.

DAFTAR RUJUKAN

- A. F. Istianto, A. I. Hadiana, and F.R.U. (2023) "Prediksi curah hujan menggunakan metode categorical boosting (Catboost)," JATI (Jurnal Mahasiswa Teknik Informatika), 7(4), pp. 2930–2937.
- A. Y. Pratama, R. S. Wahyuni, dan R.F. (2022) "Klasifikasi Citra Daging Sapi dan Babi Menggunakan EfficientNet," Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), 9, pp. 765–772.

- Akge Ninov Royana, Yisti Vita Via, C.A.P. (2025) "Evaluasi Kinerja LightGBM dan CatBoost untuk Prediksi Churn Berdasarkan Dataset Pelanggan Layanan Streaming Musik," *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(4), pp. 6977-6982.
- Duaja MD, Aziz SA, S.A. (2012) "Pengaruh Pemupukan terhadap Ketahanan Tanaman Kacang Tanah," *Jurnal Penelitian Pertanian Tanaman Pangan*, 31(2), pp. 9-95.
- E. Anggiratih, S. Siswanti, S. K. Octaviani, A.A. (2021) "Klasifikasi Penyakit Tanaman Padi Menggunakan Model Deep Learning EfficientNet-B3 Dengan Transfer Learning," *Jurnal Ilmiah SINUS*, 19(1), p. v. Available at: <https://doi.org/https://doi.org/10.30646/sinus.v19i1.526>.
- Hafizh Kennandya Maulana, Henni Endah Wahanani², M.M.A.H. (2025) "PENERAPAN ARSITEKTUR CNN-EFFICIENTNETB2 DENGAN TRANSFER LEARNING PADA KLASIFIKASI GAMBAR TOKOH WAYANG KULIT," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, 13(1), pp. 429-438. Available at: <https://doi.org/10.23960/jitet.v13i1.5626>.
- Hardaningsih dan Sumartini (2015) "Penyakit-penyakit Penting yang Disebabkan oleh Jamur pada Kacang Tanah dan Cara Pengendaliannya," Monograf Balitkabi Malang. [Preprint].
- I. P. Putri (2021) "Analisis Performa Metode K-Nearest Neighbor (KNN) dan Crossvalidation pada Data Penyakit Cardiovascular," *Indonesian Journal of Data and Science*, 2(1), pp. 21-28.
- J. Yan et al (2021) "LightGBM: accelerated genomically designed crop breeding through ensemble learning," *Genome Biol*, 22, pp. 1-24.
- T.A.B.W and Others. (2025) "Customer Churn Prediction Pada Streaming Musics Platform Menggunakan Ensemble Learning," *eProceedings of Engineering*, 12, p. 1.
- R. Irfannandhy, L. B. Handoko, and N.A. (2024) "Analisis Performa Model Random Forest dan CatBoost dengan Teknik SMOTE dalam Prediksi Risiko Diabetes," *Edumatic: Jurnal Pendidikan Informatika*, 8(2), pp. 714-723.
- Jawa Timur, S.D.P. dan P.P.J. (2021) Laporan Harga Komoditas Pangan di Jawa Timur 2021.
- Anggiratih, E., Siswanti, S. and Octaviani, S.K. (2021) "Klasifikasi Penyakit Daun Padi Menggunakan EfficientNet-B3 dengan Pendekatan Transfer Learning," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 8(3), pp. 475-482.
- Royana, A.N., Via, Y.V. and Others (2025) "Evaluasi Model LightGBM dan CatBoost untuk Prediksi Customer Churn pada Layanan Streaming Musik," *Jurnal Sistem Informasi dan Sains Data*, 4(2), pp. 120-130.