

KLASIFIKASI CITRA TANAMAN RIMPANG BERKHASIAT OBAT MENGUNAKAN METODE *HYBRID RESNET-50* DAN *VISION TRANSFORMER (ViT-B/16)*

Alya Anandha^{1*}, Chyquitha Danuputri², Desi Anggreani³

^{1,2,3}Universitas Muhammadiyah Makassar, Indonesia

¹alyaanandha2611@gmail.com

²Chyquithadanuputri@unismuh.ac.id

³desianggreani@unismuh.ac.id

Info Artikel

Pengajuan : 28-08-2026
Peninjauan : 10-09-2026
Revisi : 10-09-2026
Diterima : 18-09-2026
Diterbitkan : 24-09-2026

Kata Kunci:

Rhizome plants, Image
Classification, Hybrid ResNet-50,
Vision Transformer (ViT-B/16),
Computer Vision

Abstrak

Tanaman rimpang berkhasiat obat merupakan bagian penting dari kekayaan hayati Indonesia yang dimanfaatkan sebagai bahan pangan sekaligus obat tradisional. Namun, identifikasi jenis tanaman rimpang masih menjadi tantangan karena banyak kelas memiliki karakteristik visual yang sangat mirip sehingga menimbulkan tingkat kemiripan antar kelas (*inter-class similarity*) yang tinggi. Penelitian ini mengusulkan model klasifikasi citra digital menggunakan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) untuk menggabungkan representasi fitur lokal dan fitur global dalam proses identifikasi tanaman rimpang. Dataset penelitian terdiri atas 5.507 citra digital yang mewakili 10 kelas tanaman rimpang. Tahap praproses meliputi resize citra menjadi 224×224 piksel, normalisasi piksel, augmentasi data, serta pembagian dataset menjadi 70% data training, 15% data validation, dan 15% data testing. Model dilatih menggunakan optimizer AdamW dan fungsi loss CrossEntropyLoss. Hasil pengujian menunjukkan bahwa model Hybrid ResNet-50 dan Vision Transformer memperoleh akurasi sebesar 98,06% pada data testing dengan 809 citra berhasil diklasifikasikan dengan benar dan 16 citra mengalami kesalahan klasifikasi dari total 825 citra uji. Evaluasi dilakukan menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix. Hasil penelitian menunjukkan bahwa kombinasi ResNet-50 dan Vision Transformer mampu meningkatkan kemampuan model dalam membedakan tanaman rimpang yang memiliki kemiripan tekstur, bentuk, dan warna. Model kemudian diimplementasikan pada aplikasi web RIMPANGAI sebagai media identifikasi tanaman rimpang dan penyedia informasi mengenai khasiat serta cara pengolahannya.

How to Cite : Anandha, A., Danuputri, C., & Anggreani, D. (2026). Klasifikasi Citra Tanaman Rimpang Berkhasiat Obat Menggunakan Metode Hybrid Resnet-50 Dan Vision Transformer (ViT-B/16). Journal of Computer Science and Information Technology (JCSIT), 3(4), 500-517.
<https://doi.org/10.70248/jcsit.v3i4.4682>

DOI : <https://doi.org/10.70248/jcsit.v3i4.4682>

This is an open access article under <https://creativecommons.org/licenses/by/4.0/>

Copyright© 2026 by Author. Published by Yayasan Nuraini Ibrahim Mandiri



PENDAHULUAN

Tanaman rimpang merupakan salah satu kelompok tanaman obat yang memiliki peranan penting dalam pengobatan tradisional dan industri pangan di Indonesia. Berbagai jenis rimpang seperti jahe, jahe merah, kunyit, kunyit putih, kencur, lengkuas, lempuyang, temulawak, temu hitam, dan temu kunci mengandung senyawa bioaktif yang dimanfaatkan sebagai bahan obat herbal, minuman kesehatan, serta bumbu masakan. Pemanfaatan tanaman rimpang terus meningkat karena kandungan metabolit sekundernya memiliki aktivitas antioksidan, antiinflamasi, antimikroba, dan

imunomodulator yang bermanfaat bagi kesehatan (Rahayu et al., 2023).

Perkembangan teknologi computer vision dan deep learning memberikan peluang untuk membangun sistem identifikasi tanaman berbasis citra digital secara otomatis. Teknologi ini mampu mengekstraksi karakteristik visual objek secara otomatis sehingga proses klasifikasi menjadi lebih cepat, objektif, dan akurat dibandingkan identifikasi manual. Penerapan pengolahan citra digital pada tanaman obat juga dapat membantu masyarakat memperoleh informasi mengenai jenis tanaman secara lebih mudah dan efisien melalui sistem identifikasi berbasis kecerdasan buatan (Kamilia Zulfiani et al., 2025).

Salah satu arsitektur deep learning yang banyak digunakan dalam klasifikasi citra adalah Residual Network (ResNet-50). ResNet-50 memiliki kemampuan yang baik dalam mengekstraksi fitur lokal seperti tekstur, bentuk, dan pola permukaan objek melalui mekanisme residual connection yang memungkinkan jaringan menjadi lebih dalam tanpa mengalami penurunan performa akibat vanishing gradient (Darmatasia & H, 2025). Di sisi lain, Vision Transformer (ViT-B/16) mampu mempelajari hubungan antar bagian citra secara global menggunakan mekanisme self-attention sehingga menghasilkan representasi fitur yang lebih komprehensif (Dosovitskiy et al., 2021).

Beberapa penelitian sebelumnya menunjukkan bahwa ResNet-50 maupun Vision Transformer mampu menghasilkan performa klasifikasi yang tinggi pada berbagai objek citra. (Kamilia Zulfiani et al., 2025) menerapkan Vision Transformer untuk pengenalan spesies tanaman dan menunjukkan bahwa mekanisme self-attention mampu meningkatkan kemampuan model dalam mengenali karakteristik visual tanaman. Penelitian (Grace & Gunawan, 2025) juga menunjukkan bahwa ResNet-50 dan Vision Transformer memiliki performa yang baik pada klasifikasi citra, namun masing-masing arsitektur memiliki keunggulan yang berbeda dalam mengekstraksi fitur lokal maupun global.

Meskipun demikian, sebagian besar penelitian terdahulu masih menggunakan ResNet-50 atau Vision Transformer sebagai model tunggal sehingga kemampuan model dalam mengatasi kemiripan visual antar kelas belum optimal. ResNet-50 lebih efektif menangkap fitur lokal berupa tekstur dan bentuk objek, sedangkan Vision Transformer lebih unggul dalam memahami hubungan spasial global antar bagian citra. Belum banyak penelitian yang mengintegrasikan kedua arsitektur tersebut untuk klasifikasi tanaman rimpang yang memiliki tingkat inter-class similarity tinggi (Ahmad et al., 2025). Kondisi tersebut menunjukkan adanya research gap yang menjadi dasar pengembangan model hybrid pada penelitian ini.

Berdasarkan research gap tersebut, penelitian ini mengusulkan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) untuk menggabungkan kemampuan ekstraksi fitur lokal dan fitur global dalam satu model klasifikasi citra tanaman rimpang berkhasiat obat. ResNet-50 digunakan sebagai backbone untuk mengekstraksi tekstur, bentuk, dan pola permukaan tanaman rimpang, sedangkan Vision Transformer digunakan untuk mempelajari hubungan spasial global melalui mekanisme Multi-Head Self-Attention (Sugiartawan et al., 2025). Kombinasi kedua arsitektur tersebut diharapkan mampu meningkatkan kemampuan model dalam membedakan sepuluh jenis tanaman rimpang yang memiliki karakteristik visual serupa.

Kontribusi utama penelitian ini adalah mengembangkan model klasifikasi citra tanaman rimpang berkhasiat obat menggunakan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) pada dataset sebanyak 5.507 citra digital yang terdiri atas sepuluh kelas tanaman rimpang. Model dievaluasi menggunakan metrik Accuracy,

Precision, Recall, F1-Score, dan Confusion Matrix untuk memperoleh gambaran performa klasifikasi secara menyeluruh. Selain itu, model diimplementasikan pada aplikasi berbasis web RIMPANGAI sebagai media identifikasi tanaman rimpang dan penyedia informasi mengenai khasiat serta cara pengolahannya.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deep learning untuk membangun sistem identifikasi tanaman tata boga berkhasiat obat berbasis citra digital menggunakan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16). Objek penelitian terdiri atas 10 jenis tanaman rimpang, yaitu jahe, jahe merah, kunyit, kunyit putih, kencur, lengkuas, lempuyang, temulawak, temu hitam, dan temu kunci. Sistem dikembangkan dalam bentuk aplikasi web RIMPANGAI yang mampu mengidentifikasi tanaman rimpang berdasarkan citra digital serta menampilkan informasi mengenai khasiat dan cara pengolahannya.

Dataset Penelitian

Dataset penelitian terdiri atas 5.507 citra digital yang dikelompokkan ke dalam 10 kelas tanaman rimpang, yaitu jahe, jahe merah, kunyit, kunyit putih, kencur, lengkuas, lempuyang, temulawak, temu hitam, dan temu kunci. Seluruh citra dikumpulkan dengan variasi sudut pengambilan gambar, jarak pemotretan, dan kondisi pencahayaan untuk memperoleh karakteristik visual yang beragam pada setiap kelas tanaman rimpang. Dengan karakteristik visual yang beragam, dataset diharapkan mampu meningkatkan kemampuan model dalam mengenali setiap jenis tanaman rimpang secara lebih akurat pada proses pelatihan maupun pengujian.

Tabel 1. Distribusi Dataset Tanaman Rimpang

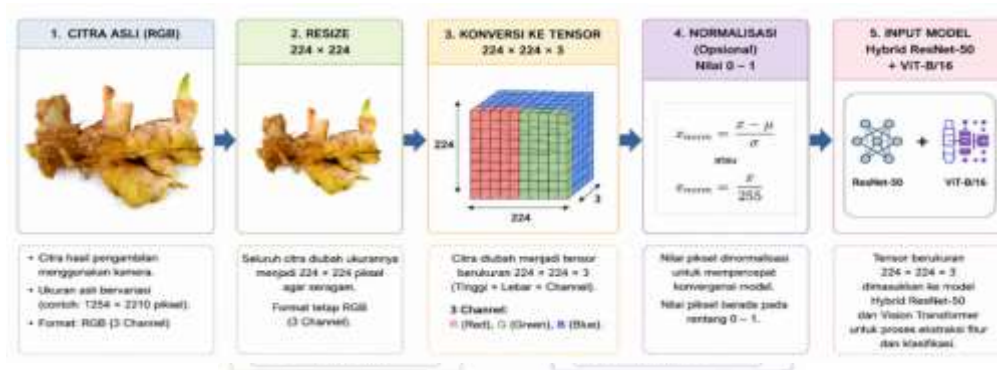
Jenis Tanaman	Jumlah Citra
Jahe	741
Jahe Merah	230
Kencur	741
Kunyit	741
Kunyit Putih	347
Lempuyang	541
Lengkuas	741
Temu Hitam	237
Temu Kunci	541
Temulawak	647

Preprocessing Citra

Sebelum proses pelatihan model, seluruh citra melalui tahapan preprocessing untuk menghasilkan data yang seragam dan meningkatkan kemampuan generalisasi model. Tahapan preprocessing meliputi resize citra, normalisasi piksel, dan augmentasi data. Seluruh citra diubah menjadi ukuran 224×224 piksel agar sesuai dengan ukuran masukan ResNet-50 dan Vision Transformer. Nilai piksel kemudian dinormalisasi sehingga distribusi data menjadi lebih konsisten selama proses pelatihan.

Augmentasi data diterapkan hanya pada data training untuk meningkatkan keragaman citra tanpa mengubah label kelas. Teknik augmentasi yang digunakan meliputi Random Horizontal Flip, Random Rotation, Color Jitter, Random Crop, Gaussian Blur, dan Random Erasing. Kombinasi augmentasi tersebut membantu model mengenali objek pada berbagai variasi orientasi, pencahayaan, dan kondisi visual

sehingga mengurangi risiko overfitting selama proses pelatihan (Kamilia Zulfiani et al., 2025).



Gambar Alur Preprocessing

Pembagian Dataset

Dataset penelitian terdiri atas 5.507 citra digital yang mewakili sepuluh kelas tanaman rimpang berkhasiat obat. Sebelum proses pelatihan model, dataset dibagi menggunakan metode Random Split dengan nilai random seed sebesar 42 untuk memastikan proses pembagian data dapat direproduksi pada setiap percobaan. Pembagian dilakukan dengan rasio 70% sebagai data training, 15% sebagai data validation, dan 15% sebagai data testing.

Metode Random Split dipilih karena pembagian data dilakukan secara acak sehingga setiap citra memiliki peluang yang sama untuk masuk ke dalam data training, validation, maupun testing. Penggunaan random seed bertujuan menjaga konsistensi hasil pembagian dataset pada setiap proses pelatihan model. Dengan pembagian tersebut, data training digunakan untuk melatih model, data validation digunakan untuk memantau performa model selama proses pelatihan, sedangkan data testing digunakan untuk mengevaluasi performa akhir model yang telah dilatih.

Distribusi jumlah citra pada setiap kelas tidak sepenuhnya seimbang, di mana beberapa kelas memiliki jumlah citra lebih banyak dibandingkan kelas lainnya. Untuk mengurangi pengaruh ketidakseimbangan tersebut, penelitian ini menerapkan teknik data augmentation pada data training sehingga variasi citra meningkat tanpa mengubah label kelas. Pendekatan ini membantu model mempelajari karakteristik visual setiap kelas secara lebih beragam dan meningkatkan kemampuan generalisasi model selama proses pelatihan.

Tabel 2. Pembagian Dataset Penelitian

Kelas	Total	Training (70%)	Validation (15%)	Testing (15%)
Jahe	741	519	111	111
Jahe Merah	230	161	34	35
Kencur	741	519	111	111
Kunyit	741	519	111	111
Kunyit Putih	347	243	52	52
Lempuyang	541	379	81	81
Lengkuas	741	519	111	111
Temu Hitam	237	166	35	36
Temu Kunci	541	379	81	81
Temulawak	647	453	97	97

Total	5.507	3.857	825	825
--------------	--------------	--------------	------------	------------

Konfigurasi Pelatihan Model

Model Hybrid ResNet-50 dan Vision Transformer dilatih menggunakan framework PyTorch dengan optimizer AdamW dan fungsi loss CrossEntropyLoss. Optimizer AdamW dipilih karena mampu mengoptimalkan proses fine-tuning model transformer melalui mekanisme weight decay yang lebih stabil dibandingkan Adam konvensional (Sugiartawan et al., 2025).

Learning rate sebesar 0,0001 digunakan sebagai learning rate utama, sedangkan backbone learning rate sebesar 0,00001 diterapkan pada ResNet-50 untuk mempertahankan bobot pretrained selama proses fine-tuning. Batch size sebesar 8 dipilih untuk menyesuaikan kapasitas memori GPU NVIDIA GeForce RTX 3050 yang digunakan selama pelatihan sehingga proses training dapat berjalan secara stabil tanpa mengalami keterbatasan memori.

Jumlah epoch ditetapkan sebanyak 15 berdasarkan hasil pengamatan kurva training accuracy dan validation accuracy. Setelah epoch ke-15, peningkatan akurasi relatif kecil dan kurva validasi telah menunjukkan kondisi konvergen sehingga penambahan epoch diperkirakan tidak memberikan peningkatan performa yang signifikan.

Tabel 3. Hyperparameter

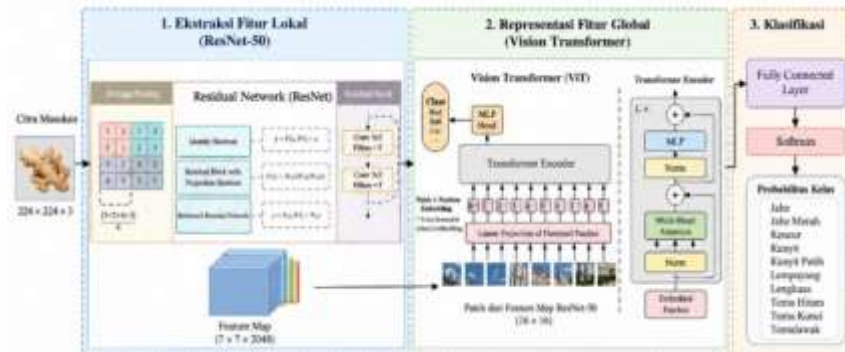
Parameter	Nilai	Justifikasi
Model	Hybrid ResNet-50 + ViT-B/16	Menggabungkan fitur lokal dan global
Image Size	224 × 224	Sesuai input ResNet-50 dan ViT-B/16
Optimizer	AdamW	Stabil untuk fine-tuning Vision Transformer
Learning Rate	0.0001	Menjaga stabilitas proses optimasi
Backbone Learning Rate	0.00001	Fine-tuning bobot pretrained ResNet-50
Batch Size	8	Menyesuaikan kapasitas GPU RTX 3050
Epoch	15	Kurva akurasi telah konvergen pada epoch ke-15
Loss Function	CrossEntropyLoss	Digunakan untuk klasifikasi multikelas

Arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16)

Arsitektur yang digunakan pada penelitian ini merupakan model Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) yang mengintegrasikan kemampuan ekstraksi fitur lokal dan fitur global dalam satu proses klasifikasi citra tanaman rimpang. Pendekatan hybrid dipilih karena karakteristik visual tanaman rimpang memiliki tingkat kemiripan antar kelas (inter-class similarity) yang tinggi, sehingga diperlukan representasi fitur yang lebih komprehensif untuk meningkatkan performa klasifikasi (Sugiartawan et al., 2025).

Pada arsitektur ini, ResNet-50 berfungsi sebagai backbone untuk mengekstraksi fitur lokal berupa tekstur, bentuk, pola permukaan, dan warna objek. Hasil ekstraksi berupa feature map kemudian diproyeksikan menjadi patch embedding yang menjadi masukan bagi Vision Transformer (ViT-B/16). Vision Transformer mempelajari hubungan spasial antar patch menggunakan mekanisme Multi-Head Self-Attention sehingga mampu menghasilkan representasi fitur global yang melengkapi fitur lokal dari ResNet-50. Representasi fitur global selanjutnya diteruskan ke lapisan klasifikasi

(Fully Connected Layer) untuk menghasilkan probabilitas sepuluh kelas tanaman rimpang menggunakan fungsi Softmax (Dosovitskiy et al., 2021).



Gambar Arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16)

Gambar diatas menunjukkan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) yang digunakan pada penelitian ini untuk mengklasifikasikan citra tanaman rimpang berkhasiat obat. Arsitektur terdiri atas tiga tahapan utama, yaitu ekstraksi fitur lokal menggunakan ResNet-50, representasi fitur global menggunakan Vision Transformer, dan proses klasifikasi menggunakan lapisan Fully Connected serta fungsi Softmax.

Pada tahap pertama, citra masukan berukuran $224 \times 224 \times 3$ piksel diproses oleh ResNet-50 untuk menghasilkan feature map berukuran $7 \times 7 \times 2048$. ResNet-50 memanfaatkan residual block dan residual connection untuk mengekstraksi fitur lokal berupa tekstur, bentuk, pola permukaan, dan warna tanaman rimpang secara efektif.

Tahap kedua merupakan proses Vision Transformer (ViT-B/16). Feature map yang dihasilkan ResNet-50 dibagi menjadi beberapa patch berukuran 16×16 piksel, kemudian setiap patch diproyeksikan menjadi patch embedding dan ditambahkan positional encoding. Selanjutnya patch embedding diproses oleh Transformer Encoder yang terdiri atas Multi-Head Self-Attention (MHSA), Layer Normalization, dan Multi-Layer Perceptron (MLP) untuk mempelajari hubungan spasial antar patch sehingga menghasilkan representasi fitur global.

Tahap terakhir adalah klasifikasi. Representasi fitur global dari Vision Transformer diteruskan ke lapisan Fully Connected dan fungsi Softmax untuk menghasilkan probabilitas sepuluh kelas tanaman rimpang, yaitu Jahe, Jahe Merah, Kencur, Kunyit, Kunyit Putih, Lempuyang, Lengkuas, Temu Hitam, Temu Kunci, dan Temulawak. Kombinasi ResNet-50 dan Vision Transformer memungkinkan model memanfaatkan informasi fitur lokal dan global secara bersamaan sehingga meningkatkan kemampuan klasifikasi pada objek yang memiliki kemiripan visual antar kelas (inter-class similarity).

Persamaan Matematika

Persamaan (1) Residual Block ResNet-50

$$y = F(x, W) + x$$

Keterangan:

x = masukan residual block.

$F(x, W)$ = fungsi konvolusi yang mempelajari fitur citra.

W = bobot jaringan.

y = keluaran residual block.

Residual connection memungkinkan informasi dari lapisan sebelumnya diteruskan secara langsung ke lapisan berikutnya sehingga proses pelatihan jaringan yang lebih dalam menjadi lebih stabil dan mengurangi permasalahan vanishing gradient (Darmatasia & H, 2025).

Persamaan (2) Patch Embedding Vision Transformer

Setelah feature map diperoleh dari ResNet-50, citra dibagi menjadi patch berukuran 16×16 piksel dan diproyeksikan menjadi embedding menggunakan persamaan berikut:

$$Z_0 = [x_p^1 E; x_p^2 E; \dots x_p^N E] E_{pos}$$

Keterangan:

x_p = patch citra.

E = matriks proyeksi linear.

E_{pos} = positional encoding.

Z_0 = patch embedding sebagai masukan Vision Transformer.

Positional encoding digunakan agar model tetap mengetahui posisi setiap patch pada citra sehingga informasi spasial tidak hilang selama proses transformasi (Dosovitskiy et al., 2021).

Persamaan (3) Multi-Head Self-Attention

Mekanisme Multi-Head Self-Attention dihitung menggunakan persamaan berikut:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Keterangan:

Q = Query.

K = Key.

V = Value.

d_k = dimensi vektor Key.

Mekanisme self-attention menghitung hubungan antar patch citra sehingga Vision Transformer mampu menangkap informasi global dari seluruh bagian objek tanaman rimpang (Dosovitskiy et al., 2021).

Teknik Evaluasi Model

Evaluasi model dilakukan menggunakan data *testing* untuk mengukur kemampuan model *Hybrid ResNet-50* dan *Vision Transformer (ViT-B/16)* dalam mengklasifikasikan tanaman rimpang. Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap label sebenarnya pada setiap citra. Metrik evaluasi yang digunakan dalam penelitian ini meliputi *Accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

Penggunaan beberapa metrik evaluasi bertujuan untuk memberikan gambaran yang lebih menyeluruh mengenai performa model. Nilai *Accuracy* menunjukkan tingkat ketepatan prediksi model secara keseluruhan, sedangkan *precision*, *recall*, dan *F1-score* digunakan untuk mengevaluasi performa pada setiap kelas tanaman rimpang.

Selain itu, *Confusion Matrix* digunakan untuk mengetahui distribusi prediksi yang benar dan kesalahan klasifikasi pada masing-masing kelas. Hasil evaluasi tersebut menjadi dasar dalam menilai efektivitas model pada proses pengujian. Melalui hasil evaluasi tersebut, dapat diketahui kemampuan model dalam membedakan setiap jenis tanaman rimpang yang memiliki karakteristik visual yang mirip. Analisis ini juga

membantu mengidentifikasi kelas yang masih mengalami kesalahan klasifikasi sehingga performa model dapat dievaluasi secara lebih rinci.

Tabel 4. Metrik Evaluasi Model

Metrik	Fungsi Evaluasi
<i>Accuracy</i>	Mengukur persentase prediksi yang benar terhadap seluruh data pengujian.
<i>Precision</i>	Mengukur ketepatan model dalam memprediksi setiap kelas tanaman rimpang.
<i>Recall</i>	Mengukur kemampuan model mengenali seluruh data yang termasuk dalam suatu kelas.
<i>F1-Score</i>	Mengukur keseimbangan antara nilai precision dan recall.
<i>Confusion Matrix</i>	Menampilkan jumlah prediksi benar dan kesalahan klasifikasi pada setiap kelas tanaman.

Tabel 4, menunjukkan metrik evaluasi yang digunakan untuk mengukur kinerja model *Hybrid ResNet-50* dan *Vision Transformer (ViT-B/16)* dalam mengklasifikasikan tanaman rimpang. *Accuracy* digunakan untuk mengetahui tingkat ketepatan prediksi model secara keseluruhan, sedangkan precision, recall, dan F1-score digunakan untuk mengevaluasi performa model pada setiap kelas tanaman. Selain itu, *Confusion Matrix* digunakan untuk menganalisis distribusi prediksi benar dan kesalahan klasifikasi sehingga dapat diketahui kemampuan model dalam membedakan setiap jenis tanaman rimpang. Evaluasi model dilakukan menggunakan metrik *Accuracy*, precision, recall, F1-score, dan confusion matrix. *Confusion Matrix* digunakan untuk menganalisis jumlah prediksi benar dan kesalahan klasifikasi pada setiap kelas tanaman rimpang.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil Preprocessing Dataset

Tahap preprocessing dilakukan untuk menghasilkan dataset yang siap digunakan pada proses pelatihan model. Seluruh citra diubah menjadi ukuran 224×224 piksel, kemudian dilakukan normalisasi dan augmentasi data agar variasi citra meningkat dan model memiliki kemampuan generalisasi yang lebih baik. Setelah proses preprocessing, dataset terdiri atas 5.507 citra digital yang dikelompokkan ke dalam 10 kelas tanaman rimpang. Dataset kemudian dibagi menjadi data training, validation, dan testing menggunakan metode random split dengan proporsi 70:15:15. Proses preprocessing memastikan seluruh citra memiliki format yang seragam sebelum dimasukkan ke dalam model. Augmentasi data juga membantu meningkatkan keragaman citra sehingga model mampu mengenali objek pada berbagai kondisi pencahayaan.

Tabel 5. Pembagian Dataset Penelitian

Dataset	Jumlah Citra	Presentase
Training	3.857	70%
Validation	825	15%
Testing	825	15%
Total	5.507	100%

Berdasarkan Tabel 5, Data Training digunakan untuk melatih model dalam mengenali karakteristik setiap jenis tanaman rimpang. Data validation digunakan untuk memantau performa model selama proses pelatihan, sedangkan data testing digunakan untuk mengevaluasi kemampuan model pada data yang belum pernah dipelajari sebelumnya.

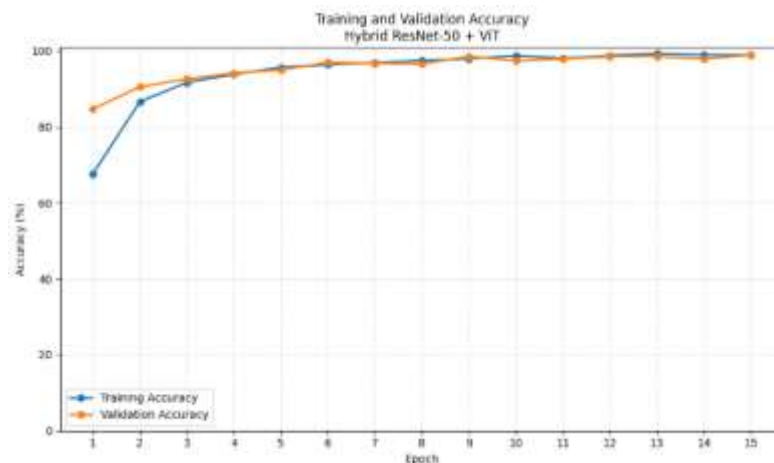
Hasil Pelatihan Model Hybrid ResNet-50 dan Vision Transformer

Proses pelatihan dilakukan menggunakan arsitektur *Hybrid ResNet-50* dan *Vision Transformer (ViT-B/16)* dengan *Optimizer AdamW*, *Batch Size 8*, dan *15 Epoch*. Selama proses pelatihan, model mempelajari karakteristik visual setiap tanaman rimpang berdasarkan tekstur, bentuk, warna, dan pola permukaan objek.

Nilai training *Accuracy* mengalami peningkatan pada setiap *Epoch*, sedangkan training loss menurun secara bertahap. Hal tersebut menunjukkan bahwa model mampu mempelajari pola pada dataset dengan baik dan proses pelatihan berlangsung secara stabil. Peningkatan nilai training *Accuracy* menunjukkan bahwa model semakin mampu mengenali karakteristik setiap kelas tanaman rimpang seiring bertambahnya jumlah *Epoch*.

Nilai validation *Accuracy* yang juga meningkat menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi. Selain itu, penurunan nilai loss mengindikasikan bahwa kesalahan prediksi model semakin berkurang selama proses pelatihan. Hasil ini menunjukkan bahwa konfigurasi pelatihan yang digunakan mampu menghasilkan model dengan performa yang stabil dan siap digunakan pada tahap pengujian menggunakan data testing.

Gambar Grafik Trainin Accuracy dan Validation Accuracy



Gambar menunjukkan perkembangan nilai training accuracy dan validation accuracy selama proses pelatihan sebanyak 15 epoch. Nilai akurasi meningkat secara bertahap hingga mencapai kondisi konvergen pada epoch ke-15, sedangkan nilai validation loss terus menurun tanpa menunjukkan peningkatan yang signifikan pada akhir proses pelatihan. Kondisi tersebut menunjukkan bahwa model mampu mempelajari karakteristik dataset dengan baik dan tidak mengalami gejala overfitting yang berarti selama proses pelatihan.

Sementara itu, validation *Accuracy* juga menunjukkan peningkatan yang konsisten dan berada pada kisaran 84% hingga 98%. Nilai validation *Accuracy* yang hampir sejajar dengan training *Accuracy* menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data validasi dan tidak mengalami overfitting yang signifikan. Hasil ini menunjukkan bahwa kombinasi ResNet-50 dan Vision Transformer mampu menghasilkan performa pelatihan yang stabil pada dataset tanaman rimpang.

Evaluasi Performa Model

Setelah proses pelatihan selesai, model diuji menggunakan 825 citra pada data testing. Pengujian dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan sepuluh jenis tanaman rimpang berdasarkan citra digital. Evaluasi dilakukan menggunakan metrik *Accuracy*, precision, recall, F1-score, dan confusion matrix.

Tabel 6. Hasil Evaluasi Model

Metrik	Nilai
Akurasi	98,06%
Jumlah Data Testing	825
Prediksi Benar	809
Prediksi Salah	16

Hasil pengujian menunjukkan bahwa model Hybrid ResNet-50 dan Vision Transformer berhasil mengklasifikasikan 809 dari 825 citra data testing dengan benar sehingga memperoleh akurasi sebesar 98,06%. Tingkat akurasi tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik pada dataset tanaman rimpang berkhasiat obat. Tingginya akurasi dipengaruhi oleh kombinasi ekstraksi fitur lokal menggunakan ResNet-50 dan representasi fitur global menggunakan Vision Transformer yang mampu mengenali karakteristik visual tanaman rimpang secara lebih komprehensif.

Evaluasi Precision, Recall, dan F1-Score per Kelas

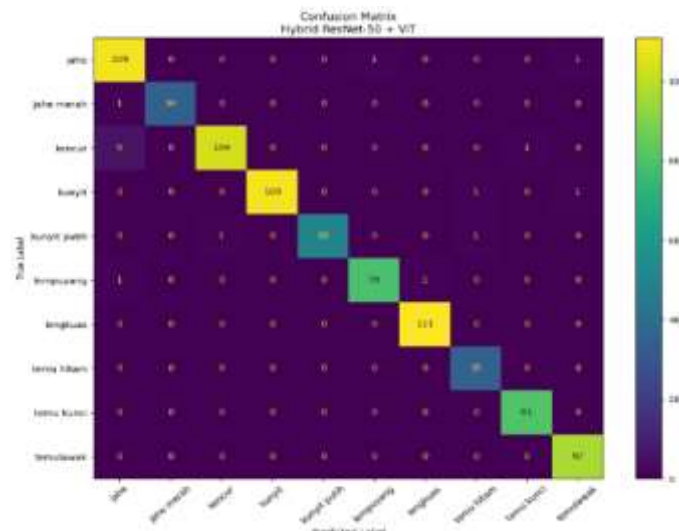
Selain menggunakan metrik akurasi, performa model Hybrid ResNet-50 dan Vision Transformer juga dievaluasi menggunakan metrik Precision, Recall, dan F1-Score untuk setiap kelas tanaman rimpang. Evaluasi ini bertujuan untuk memberikan gambaran kemampuan model dalam mengidentifikasi masing-masing kelas secara lebih rinci, terutama pada kelas yang memiliki jumlah citra lebih sedikit dan karakteristik visual yang mirip.

Tabel 7. Hasil Evaluasi Precision, Recall, dan F1-Score

Kelas Tanaman	Precision	Recall	F1-Score	Support
Jahe	0,9316	0,9820	0,9561	111
Jahe Merah	1,0000	0,9714	0,9855	35
Kencur	0,9905	0,9369	0,9630	111
Kunyit	1,0000	0,9820	0,9909	111
Kunyit Putih	1,0000	0,9615	0,9804	52
Lempuyang	0,9875	0,9753	0,9814	81
Lengkuas	0,9911	1,0000	0,9955	111
Temu Hitam	0,9459	1,0000	0,9722	35
Temu Kunci	0,9878	1,0000	0,9939	81
Temulawak	0,9798	1,0000	0,9898	97
Accuracy	-	-	98,06%	825
Macro Average	0,9814	0,9809	0,9809	825
Weighted Average	0,9812	0,9806	0,9806	825

Tabel 7 menunjukkan bahwa model Hybrid ResNet-50 dan Vision Transformer memperoleh nilai Precision, Recall, dan F1-Score di atas 0,93 pada seluruh kelas tanaman rimpang. Nilai Macro Average sebesar 0,9809 dan Weighted Average sebesar 0,9806 menunjukkan bahwa model memiliki performa yang konsisten pada seluruh kelas, termasuk kelas dengan jumlah citra yang lebih sedikit seperti Temu Hitam.

Analisis Confusion Matrix Dan Misclassification



Gambar Hasil Confusion Matrix

Gambar menunjukkan hasil Confusion Matrix dari model Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) pada data testing sebanyak 825 citra. Berdasarkan hasil pengujian, model berhasil mengklasifikasikan 809 citra dengan benar dan 16 citra mengalami kesalahan klasifikasi sehingga memperoleh akurasi keseluruhan sebesar 98,06%. Dominasi nilai pada diagonal utama confusion matrix menunjukkan bahwa sebagian besar citra berhasil diprediksi sesuai kelas sebenarnya, sedangkan nilai di luar diagonal menunjukkan jumlah citra yang mengalami kesalahan prediksi (misclassification).

Hasil confusion matrix menunjukkan bahwa kesalahan klasifikasi hanya terjadi pada beberapa kelas yang memiliki karakteristik visual yang sangat mirip. Kemiripan bentuk rimpang, tekstur permukaan, warna kulit, dan pola percabangan menyebabkan model menghasilkan probabilitas prediksi yang berdekatan pada beberapa pasangan kelas. Kondisi tersebut menunjukkan bahwa tantangan utama dalam klasifikasi tanaman rimpang adalah inter-class similarity, yaitu kemiripan karakteristik visual antar kelas tanaman rimpang.

Analisis Misclassification Berdasarkan Confusion Matrix

1. Jahe

Kelas Jahe memiliki 111 citra pada data testing dengan 109 citra berhasil diprediksi dengan benar. Dua citra mengalami kesalahan klasifikasi, yaitu satu citra diprediksi sebagai Lempuyang dan satu citra diprediksi sebagai Temulawak. Kesalahan tersebut kemungkinan dipengaruhi oleh kemiripan bentuk memanjang serta warna permukaan rimpang pada kondisi pencahayaan tertentu. Meskipun demikian, nilai Recall kelas Jahe mencapai 98,20% sehingga sebagian besar citra Jahe tetap berhasil dikenali oleh model.

2. Jahe Merah

Kelas Jahe Merah terdiri atas 35 citra data testing dengan 34 citra berhasil diklasifikasikan dengan benar. Satu citra Jahe Merah diprediksi sebagai Jahe. Kesalahan ini menunjukkan bahwa kedua kelas memiliki karakteristik visual yang sangat mirip,

terutama pada tekstur kulit dan bentuk rimpang, sehingga model masih mengalami kesulitan membedakan beberapa sampel yang memiliki warna permukaan hampir sama.

3. Kencur

Kelas Kencur memiliki jumlah kesalahan klasifikasi paling banyak dibandingkan kelas lainnya. Dari 111 citra testing, sebanyak 104 citra berhasil diprediksi dengan benar, sedangkan tujuh citra mengalami kesalahan klasifikasi. Enam citra diprediksi sebagai Jahe dan satu citra diprediksi sebagai Temu Hitam. Hasil ini sejalan dengan nilai Recall kelas Kencur sebesar 93,69%, yang merupakan nilai Recall terendah pada penelitian ini. Kesalahan tersebut dipengaruhi oleh kemiripan tekstur permukaan dan bentuk rimpang pada beberapa citra Kencur dengan Jahe, terutama ketika posisi pengambilan gambar memperlihatkan bagian samping atau ujung rimpang.

4. Kunyit

Kelas Kunyit terdiri atas 111 citra testing dengan 109 citra berhasil diprediksi dengan benar. Dua citra mengalami kesalahan klasifikasi, yaitu satu citra diprediksi sebagai Temu Hitam dan satu citra diprediksi sebagai Temulawak. Kemiripan warna kuning-oranye dan bentuk rimpang menjadi faktor utama terjadinya kesalahan prediksi pada kelas ini. Meskipun demikian, nilai Precision dan F1-Score kelas Kunyit tetap sangat tinggi, yaitu masing-masing sebesar 1,0000 dan 0,9909.

5. Kunyit Putih

Pada kelas Kunyit Putih, sebanyak 50 dari 52 citra berhasil diprediksi dengan benar. Dua citra mengalami kesalahan klasifikasi, masing-masing diprediksi sebagai Kencur dan Temu Hitam. Kesalahan tersebut menunjukkan bahwa warna permukaan yang cenderung pucat dan bentuk rimpang yang relatif kecil masih memiliki kemiripan dengan beberapa kelas lain pada kondisi citra tertentu.

6. Lempuyang

Kelas Lempuyang memiliki 81 citra testing dengan 79 citra berhasil diklasifikasikan dengan benar. Dua citra mengalami kesalahan klasifikasi, yaitu satu citra diprediksi sebagai Jahe dan satu citra diprediksi sebagai Temu Kunci. Kesalahan ini dipengaruhi oleh bentuk rimpang yang sama-sama bercabang dan memiliki ukuran yang relatif kecil sehingga karakteristik visual antar kedua kelas menjadi cukup mirip.

7. Lengkuas

Seluruh 111 citra Lengkuas berhasil diklasifikasikan dengan benar sehingga kelas ini memperoleh nilai Recall sebesar 100%. Hasil tersebut menunjukkan bahwa karakteristik visual Lengkuas memiliki fitur yang cukup berbeda dibandingkan kelas lainnya sehingga mudah dikenali oleh model Hybrid ResNet-50 dan Vision Transformer.

8. Temu Hitam

Kelas Temu Hitam terdiri atas 35 citra testing dan seluruh citra berhasil diprediksi dengan benar. Nilai Recall sebesar 100% menunjukkan bahwa model mampu mengenali seluruh citra Temu Hitam meskipun jumlah dataset pada kelas ini lebih sedikit dibandingkan kelas mayoritas. Hal ini menunjukkan bahwa proses augmentasi data membantu meningkatkan kemampuan generalisasi model pada kelas minoritas.

9. Temu Kunci

Seluruh 81 citra Temu Kunci berhasil diprediksi dengan benar sehingga kelas ini memperoleh nilai Recall sebesar 100% dan F1-Score sebesar 0,9939. Hasil tersebut menunjukkan bahwa model mampu membedakan karakteristik Temu Kunci dengan baik meskipun bentuk rimpangnya relatif kecil dan bercabang.

10. Temulawak

Kelas Temulawak memiliki 97 citra testing dan seluruh citra berhasil diklasifikasikan dengan benar sehingga memperoleh nilai Recall sebesar 100%. Meskipun terdapat beberapa citra dari kelas lain yang diprediksi sebagai Temulawak, model tetap mampu mengenali seluruh citra Temulawak pada data testing.

Tabel 8. Ringkasan Kesalahan Klasifikasi Berdasarkan Confusion Matrix

Kelas Asli	Diprediksi Menjadi	Jumlah Kesalahan
Jahe	Lempuyang	1
Jahe	Temulak	1
Jahe Merah	Jahe	1
Kencur	Jahe	6
Kencur	Temu Hitam	1
Kunyit	Temu Hitam	1
Kunyit	Temulawak	1
Kunyit Putih	Kencur	1
Kunyit Putih	Temu Hitam	1
Lempuyang	Jahe	1
Lempuyang	Temu Kunci	1
Total Kesalahan	11 pasangan kelas	16

Berdasarkan analisis confusion matrix, kesalahan klasifikasi paling banyak terjadi pada kelas Kencur yang diprediksi sebagai Jahe sebanyak enam citra. Hal tersebut menunjukkan bahwa kedua kelas memiliki tingkat kemiripan tekstur dan bentuk yang lebih tinggi dibandingkan pasangan kelas lainnya. Sebaliknya, kelas Lengkuas, Temu Hitam, Temu Kunci, dan Temulawak berhasil dikenali secara sempurna dengan nilai Recall sebesar 100%, sehingga karakteristik visual keempat kelas tersebut lebih mudah dibedakan oleh model.

Secara keseluruhan, hasil confusion matrix menunjukkan bahwa arsitektur Hybrid ResNet-50 dan Vision Transformer mampu mengurangi pengaruh inter-class similarity pada klasifikasi tanaman rimpang. ResNet-50 memberikan representasi fitur lokal berupa tekstur dan bentuk permukaan rimpang, sedangkan Vision Transformer mempelajari hubungan spasial antar bagian citra secara global. Kombinasi kedua representasi fitur tersebut menghasilkan performa klasifikasi yang tinggi dengan tingkat kesalahan hanya 16 citra atau sekitar 1,94% dari seluruh data testing.

Perbandingan Hasil Penelitian dengan Penelitian Terdahulu

Untuk mengetahui posisi performa model yang diusulkan, hasil penelitian ini dibandingkan dengan beberapa penelitian terdahulu yang menggunakan arsitektur Convolutional Neural Network (CNN), ResNet-50, Vision Transformer (ViT), maupun metode transfer learning pada klasifikasi citra tanaman. Perbandingan dilakukan berdasarkan metode yang digunakan, objek penelitian, dan hasil akurasi yang diperoleh. Analisis ini bertujuan untuk menunjukkan kontribusi arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) dalam meningkatkan performa klasifikasi citra tanaman rimpang yang memiliki tingkat kemiripan visual antar kelas (inter-class similarity).

Tabel 9. Perbandingan Hasil Penelitian dengan Penelitian Terdahulu

Penelitian	Metode	Objek Penelitian
Kamilia Zulfiani et al. (2025)	Convolutional Neural Network (CNN)	Membandingkan performa CNN, ResNet-50, dan Vision Transformer pada klasifikasi citra kanker payudara berbasis web.
Grace dan	CNN, ResNet-50, dan	Pengenalan spesies tanaman menggunakan

Penelitian	Metode	Objek Penelitian
Gunawan (2025)	Vision Transformer	arsitektur Vision Transformer dan Top-K Logit Disparity Score.
Ahmad et al. (2025)	Vision Transformer (ViT)	Klasifikasi kualitas daun kelor menggunakan Vision Transformer.
Sugiartawan et al. (2025)	Vision Transformer (ViT)	Identifikasi daun herbal tanaman obat menggunakan metode CNN.
Penelitian Ini	Hybrid ResNet-50 + Vision Transformer (ViT-B/16)	Klasifikasi 10 kelas tanaman rimpang berkehasiat obat menggunakan kombinasi fitur lokal dan global.

Grace & Gunawan, (2025), membandingkan tiga arsitektur deep learning, yaitu CNN, ResNet-50, dan Vision Transformer, pada klasifikasi citra kanker payudara berbasis web. Penelitian tersebut menunjukkan bahwa ResNet-50 memiliki kemampuan yang baik dalam mengekstraksi fitur lokal melalui proses konvolusi, sedangkan Vision Transformer lebih unggul dalam mempelajari hubungan spasial antar bagian citra menggunakan mekanisme self-attention. Hasil penelitian tersebut menunjukkan bahwa kedua arsitektur memiliki karakteristik yang saling melengkapi pada proses klasifikasi citra.

Ahmad et al., (2025), menerapkan Vision Transformer untuk pengenalan spesies tanaman (*Open Plant Species Recognition*). Penelitian tersebut menunjukkan bahwa Vision Transformer mampu mengenali karakteristik tanaman berdasarkan representasi fitur global yang diperoleh dari patch embedding dan mekanisme Multi-Head Self-Attention. Namun, pendekatan tersebut berfokus pada penggunaan Vision Transformer sebagai model tunggal sehingga proses ekstraksi fitur lokal belum dimanfaatkan secara khusus.

Sugiartawan et al., (2025), menggunakan Vision Transformer untuk klasifikasi kualitas daun kelor dan menunjukkan bahwa Vision Transformer efektif dalam membedakan karakteristik visual daun berdasarkan informasi global citra. Penelitian tersebut membuktikan bahwa arsitektur transformer memiliki kemampuan yang baik dalam klasifikasi citra tanaman, terutama ketika objek memiliki variasi tekstur dan bentuk.

Sementara itu, (Kamilia Zulfiani et al., 2025) menggunakan metode Convolutional Neural Network (CNN) untuk identifikasi daun herbal tanaman obat. CNN mampu mengenali pola tekstur, bentuk, dan warna daun secara otomatis melalui proses ekstraksi fitur konvolusi. Akan tetapi, CNN lebih berfokus pada representasi fitur lokal sehingga kemampuan memahami hubungan global antar bagian citra masih terbatas dibandingkan arsitektur transformer.

Berbeda dengan penelitian-penelitian tersebut, penelitian ini mengusulkan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) yang mengintegrasikan kemampuan ekstraksi fitur lokal dan fitur global dalam satu model klasifikasi. ResNet-50 digunakan sebagai backbone untuk mengekstraksi tekstur, bentuk, dan pola permukaan tanaman rimpang, sedangkan Vision Transformer mempelajari hubungan spasial antar patch citra menggunakan mekanisme Multi-Head Self-Attention. Kombinasi kedua representasi fitur tersebut menghasilkan performa klasifikasi yang tinggi dengan akurasi sebesar 98,06%, Precision 0,9812, Recall 0,9806, dan F1-Score 0,9806 pada data testing sebanyak 825 citra.

Hasil tersebut menunjukkan bahwa pendekatan hybrid mampu meningkatkan kemampuan model dalam membedakan tanaman rimpang yang memiliki karakteristik visual sangat mirip, seperti Jahe dengan Jahe Merah maupun Kunyit dengan Temulawak. Dengan demikian, penelitian ini memberikan kontribusi pada pengembangan metode klasifikasi citra tanaman obat melalui integrasi arsitektur ResNet-50 dan Vision

Transformer untuk mengatasi permasalahan inter-class similarity pada klasifikasi tanaman rimpang berkhasiat obat.

Implementasi Model pada Aplikasi Web RIMPANGAI

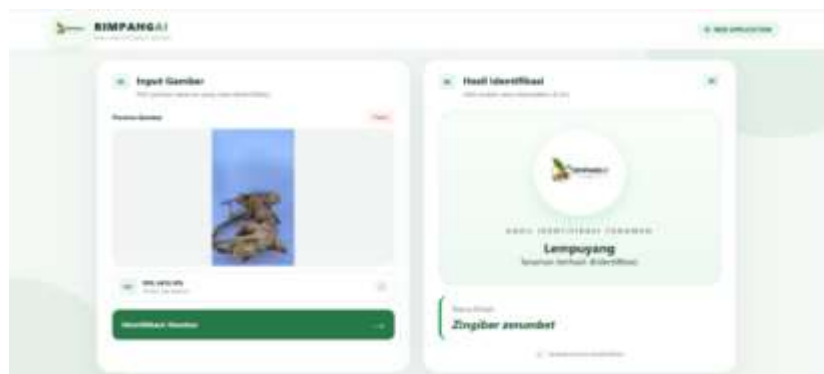
Model Hybrid ResNet-50 dan Vision Transformer (ViT-B/16) yang telah melalui proses pelatihan dan pengujian kemudian diimplementasikan ke dalam aplikasi berbasis web bernama RIMPANGAI. Implementasi ini bertujuan untuk menunjukkan bahwa model klasifikasi yang dihasilkan dapat diterapkan pada sistem identifikasi tanaman rimpang secara praktis. Pengguna dapat mengunggah citra tanaman rimpang melalui antarmuka aplikasi, kemudian sistem melakukan proses inferensi menggunakan model yang telah dilatih untuk menghasilkan prediksi jenis tanaman rimpang beserta tingkat kepercayaan (confidence score).

Selain menampilkan hasil klasifikasi, aplikasi RIMPANGAI juga menyediakan informasi mengenai nama tanaman, khasiat, dan cara pengolahan tanaman rimpang berdasarkan hasil prediksi model. Namun, implementasi aplikasi web pada penelitian ini hanya berfungsi sebagai media penerapan model hasil klasifikasi, sedangkan fokus utama penelitian tetap berada pada evaluasi performa arsitektur Hybrid ResNet-50 dan Vision Transformer dalam mengklasifikasikan citra tanaman rimpang berkhasiat obat.



Gambar Halaman Beranda Aplikasi RIMPANGAI

Gambar menunjukkan halaman beranda aplikasi RIMPANGAI yang menjadi tampilan awal sistem. Halaman ini menyediakan menu utama untuk mengakses fitur identifikasi tanaman rimpang serta informasi singkat mengenai fungsi aplikasi.



Gambar Halaman Identifikasi Tanaman Rimpang

Gambar menunjukkan proses identifikasi tanaman rimpang pada aplikasi RIMPANGAI. Pengguna mengunggah citra tanaman rimpang, kemudian sistem

melakukan preprocessing dan inferensi menggunakan model Hybrid ResNet-50 dan Vision Transformer. Hasil prediksi ditampilkan dalam bentuk nama kelas tanaman beserta nilai confidence sebagai tingkat keyakinan model terhadap hasil klasifikasi.

KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi citra tanaman rimpang berkhasiat obat menggunakan arsitektur Hybrid ResNet-50 dan Vision Transformer (ViT-B/16). Model dikembangkan untuk mengklasifikasikan sepuluh jenis tanaman rimpang, yaitu Jahe, Jahe Merah, Kencur, Kunyit, Kunyit Putih, Lempuyang, Lengkuas, Temu Hitam, Temu Kunci, dan Temulawak menggunakan dataset sebanyak 5.507 citra digital. Berdasarkan hasil pengujian pada 825 citra data testing, model memperoleh akurasi sebesar 98,06%, dengan nilai weighted Precision sebesar 98,12%, Recall sebesar 98,06%, dan F1-Score sebesar 98,06%. Hasil evaluasi tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik pada seluruh kelas tanaman rimpang.

Kontribusi utama penelitian ini terletak pada penerapan arsitektur hybrid yang menggabungkan kemampuan ekstraksi fitur lokal dari ResNet-50 dan representasi fitur global dari Vision Transformer. Kombinasi kedua arsitektur tersebut mampu meningkatkan kemampuan model dalam membedakan tanaman rimpang yang memiliki karakteristik visual serupa (inter-class similarity), sehingga jumlah kesalahan klasifikasi pada data testing hanya sebanyak 16 citra dari total 825 citra pengujian. Hasil ini menunjukkan bahwa integrasi fitur lokal dan global memberikan representasi citra yang lebih komprehensif dibandingkan penggunaan satu arsitektur secara tunggal pada klasifikasi tanaman rimpang.

Penelitian ini juga mengimplementasikan model ke dalam aplikasi berbasis web RIMPANGAI sebagai media identifikasi tanaman rimpang berkhasiat obat. Implementasi tersebut menunjukkan bahwa model dapat diterapkan pada sistem identifikasi berbasis web untuk membantu pengguna mengenali jenis tanaman rimpang secara otomatis berdasarkan citra digital.

Meskipun memperoleh performa yang tinggi, penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan dikumpulkan dengan latar belakang yang relatif seragam dan kondisi pencahayaan yang terkendali sehingga variasi kondisi citra di lingkungan nyata masih terbatas. Selain itu, distribusi jumlah citra pada setiap kelas belum sepenuhnya seimbang, sehingga penelitian selanjutnya dapat menggunakan dataset dengan variasi latar belakang, pencahayaan, dan kondisi pengambilan gambar yang lebih beragam serta mempertimbangkan teknik pembagian data yang mempertahankan proporsi setiap kelas untuk meningkatkan kemampuan generalisasi model pada implementasi di lingkungan nyata.

Pernyataan mengenai Penggunaan Kecerdasan Buatan (AI)

Kecerdasan buatan digunakan dalam penyusunan naskah ini hanya sebagai alat bantu untuk perbaikan tata bahasa, penyuntingan kalimat, dan peningkatan keterbacaan naskah. Seluruh proses penelitian, mulai dari perancangan penelitian, pengumpulan dataset, pengembangan model, analisis hasil, hingga penyusunan isi ilmiah artikel dilakukan dan diverifikasi oleh penulis. Penulis bertanggung jawab sepenuhnya atas keakuratan, keaslian, integritas, dan isi akademik naskah ini.

Jika AI tidak digunakan, Anda dapat menggunakan pernyataan berikut:

Kecerdasan buatan (Artificial Intelligence/AI) digunakan dalam penyusunan naskah ini. Penulis menggunakan perangkat kecerdasan buatan hanya untuk membantu perbaikan

tata bahasa, penyuntingan kalimat, serta meningkatkan kejelasan dan keterbacaan naskah. Seluruh isi, analisis, hasil penelitian, interpretasi, dan kesimpulan merupakan hasil penelitian penulis. Penulis bertanggung jawab penuh atas akurasi, orisinalitas, integritas, dan konten akademis artikel ini.

Konflik kepentingan

Penulis menyatakan bahwa tidak terdapat konflik kepentingan yang berkaitan dengan pelaksanaan penelitian maupun publikasi artikel ini.

Kontribusi penulis

Konseptualisasi dilakukan oleh *Alya Anandha, Chyquitha Danuputri, Dan Desi Anggreani*. Pengumpulan data dilakukan oleh *Alya Anandha*. Validasi data dilakukan oleh *Chyquitha Danuputri dan Desi Anggreani*. Analisis formal dilakukan oleh *Alya Anandha, Chyquitha Danuputri, dan Desi Anggreani*. Penyusunan draf awal naskah dilakukan oleh *Alya Anandha*. Peninjauan dan penyuntingan naskah dilakukan oleh *Chyquitha Danuputri dan Desi Anggreani*. Visualisasi berupa penyusunan tabel, grafik, dan gambar dilakukan oleh *Alya Anandha*. Seluruh penulis telah membaca dan menyetujui versi akhir naskah.

Ucapan terima kasih

“Penulis mengucapkan terima kasih kepada *Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Makassar* atas dukungan, fasilitas, dan kesempatan yang diberikan selama pelaksanaan penelitian ini. Penulis juga menyampaikan terima kasih yang sebesar-besarnya kepada Ibu *Chyquitha Danuputri, S.Kom., M.Kom.* dan Ibu *Ir. Desi Anggreani, S.Kom., M.Kom.* selaku dosen pembimbing yang telah memberikan bimbingan, arahan, masukan, dan motivasi selama proses penelitian hingga penyusunan artikel ini. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah membantu sehingga penelitian ini dapat diselesaikan dengan baik”.

DAFTAR PUSTAKA

- Ahmad, G., Alfarisy, F., Kalinaki, K., Malik, O. A., Putra, R. K., Nuryono, A. A., Alfarisy, G. A. F., Kalinaki, K., Malik, O. A., Putra, R. K., & Nuryono, A. A. (2025). Open Plant Species Recognition Using Vision Transformer Network and Top-K Logit Disparity Score "Open Plant Species Recognition Using Vision Transformer Network and Top-K Logit Disparity Score. *Innovative Informatics and Artificial Intelligence Research (IIAIR)*, 1(1), 12–18.
- Algiffary, A., & Sutabri, T. (2023). Indonesian Journal of Computer Science. *Indonesian Journal of Computer Science*, 12(2), 284–301. <http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135>
- Ario, M., Rahadiano, E., Sejati, P. W., Fauzi, A. I., Atmojo, T., Widayanti, T., & Yudanegara, R. A. (2026). *Klasifikasi Tanaman menggunakan Metode Deep Learning Residual Network (ResNet) Berbasis Data Time Series Penginderaan Jauh di Desa Girimulyo , Lampung Timur*. 15(1), 70–78.
- Bagus Baihaqi, M., Litanianda, Y., & Triyanto, A. (2022). Pada Teachable Untuk Identifikasi Tanaman Aglonema Berbasis Android. *Komputek*, 70–80. <http://studentjournal.umpo.ac.id/index.php/komputek>
- Darmatasia, & H, M. H. (2025). Model Efisien Berbasis Mobile Vision Transformer (Mobilevit) Untuk Klasifikasi Jenis Tanah Dari Citra. *Jurnal INSTEK (Informatika Sains Dan Teknologi)*, 10(2), 518–533. <https://doi.org/10.24252/instek.v10i2.61891>

- Day, N., Mahardika, G., Ramadhana, F., & Eva Yulia Puspaningrum. (2025). Klasifikasi Penyakit Daun Cabai Menggunakan Vision Transformer. *Prosiding Seminar Nasional Informatika Bela Negara*, 5(1), 131–135. <https://doi.org/10.33005/santika.v5i1.751>
- Deep, I., Untuk, L., Citra, K., Menggunakan, U., Convolutional, A., & Network, N. (2021). *Journal of dinda*. 1(2), 52–62.
- DESTIAWAN, S. (2024). *Identifikasi Jenis-Jenis Kayu Menggunakan Convolutional Neural Network Dengan Arsitektur Resnet-50*. 10(1), 1341–1346.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). an Image Is Worth 16X16 Words: Transformers for Image Recognition At Scale. *ICLR 2021 - 9th International Conference on Learning Representations*.
- Fadhil Elrizanda. (2024). *Exploring Residual Networks (ResNet): A Breakthrough in Deep Learning*. 2017, 256. <https://elrizanda.hashnode.dev/exploring-residual-networks-resnet-a-breakthrough-in-deep-learning>
- Fadhila, A., Mabe Parenreng, M., Teknik Elektro, J., & Negeri Ujung Pandang, P. (2023). Pengenalan Tanaman Herbal Daun Merica dan Daun Sirih Menggunakan Metode Convolutional Neural Network (CNN). *Prosiding Seminar Nasional Teknik Elektro Dan Informatika (SNTEI)*, 9(1), 109–113.
- Grace, S. J., & Gunawan, D. (2025). Perbandingan Cnn, Resnet50, Dan Vision Transformer Untuk Klasifikasi Kanker Payudara Berbasis Web. *Rabit: Jurnal Teknologi Dan Sistem Informasi Univrab*, 10(2), 945–956. <https://doi.org/10.36341/rabit.v10i2.6420>
- Hasbi, A., & Tumanggor, B. (2026). *Implementation of the Vision Transformer (ViT) Method in Soil Type Classification*. 5(1), 574–583.
- Imaduddin, Z. (2015). Aplikasi Mobile Untuk Deteksi Dan Klasifikasi Daun Secara Real Time. *Jurnal Teknologi Terpadu*, 1(1). <https://doi.org/10.54914/jtt.v1i1.33>
- Kamilia Zulfiani, H., Sucipto, S., & Abdullah, A. (2025). Identifikasi Daun Herbal Untuk Tanaman Obat Menggunakan Metode Convolutional Neural Network (Cnn). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(6), 10429–10435. <https://doi.org/10.36040/jati.v9i6.16481>
- Pujiati, R., & Rochmawati, N. (2022). Identifikasi Citra Daun Tanaman Herbal Menggunakan Metode Convolutional Neural Network (CNN). *Journal of Informatics and Computer Science (JINACS)*, 3(03), 351–357. <https://doi.org/10.26740/jinacs.v3n03.p351-357>
- Rahayu, M. I., Jaenal, R., & Risyadi, M. H. (2023). Identifikasi Tanaman Obat Herbal Berbasis Citra. *Jurnal Teknologi Informasi Dan Komunikasi*, 12(2), 57–63. <https://doi.org/10.58761/jurtikstmikbandung.v12i2.5763>
- Setiawan, H., Kurniabudi, Sharipuddin, & Gunardi. (2026). Perancangan Aplikasi Computer Vision Berbasis Convolutional Neural Network Untuk Deteksi Kematangan Kelapa Sawit. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 6(1), 1890–1900. <https://doi.org/10.33998/jakakom.2026.6.1.2568>
- Setiyono, B., Arif, M. R., Aini, Q. Q., Soegianto, T. H., Ohanna, J., Gunawan, R. A. F., & Rizkia, A. P. (2023). Identifikasi Tanaman Obat Indonesia Melalui Citra Daun Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 10(2), 385–392. <https://doi.org/10.25126/jtiik.20236809>
- Sugiartawan, P., Ayu, I. D., Murdhani, S., Ayu, P., & Putu, G. (2025). *Classification of Moringa Leaf Quality Using Vision Transformer (ViT)*. 7(June), 128–136.

