

# Model Ensemble Fusion–Stacking untuk Klasifikasi Varietas Salak Berbasis Deep Feature

Bunga Intan<sup>1</sup>, Ahmad Taqwa Martadinata<sup>1</sup>, Abdul Qodir<sup>2,\*</sup>

Fakultas Ilmu Teknik, Program Studi Informatika, Universitas Bina Insan, Lubuklinggau, Indonesia

Email: <sup>1</sup>[bungaintan@univbinainsan.ac.id](mailto:bungaintan@univbinainsan.ac.id), <sup>2,\*</sup>[qodirts2022@gmail.com](mailto:qodirts2022@gmail.com), <sup>3</sup>[taqwa@univbinainsan.ac.id](mailto:taqwa@univbinainsan.ac.id)

Email Penulis Corresponding: [qodirts2022@gmail.com](mailto:qodirts2022@gmail.com)

**Abstrak-** Klasifikasi varietas buah salak secara visual masih menghadapi tantangan akibat kemiripan tekstur, warna, dan karakteristik morfologis antar kelas. Identifikasi manual berpotensi menghasilkan subjektivitas serta inkonsistensi dalam penentuan varietas. Penelitian ini mengusulkan model ensemble berbasis *fusion* dan *stacking* pada ekstraksi fitur *deep learning* untuk meningkatkan akurasi klasifikasi varietas salak. Fitur citra diekstraksi menggunakan dua arsitektur *Convolutional Neural Network* pralatih, yaitu VGG16 dan ResNet50, sebagai *deep feature extractor*. Representasi fitur yang dihasilkan kemudian diklasifikasikan menggunakan algoritma K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM). Probabilitas keluaran kedua model tersebut selanjutnya digabungkan melalui pendekatan *stacking ensemble* dengan Logistic Regression sebagai *meta-learner*. Dataset penelitian terdiri atas 584 citra yang terbagi ke dalam empat varietas salak. Evaluasi dilakukan menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil eksperimen menunjukkan bahwa pendekatan *fusion–stacking* mampu mencapai akurasi sebesar 95%, lebih tinggi dibandingkan model tunggal berbasis CNN dan klasifikasi konvensional. Temuan ini menunjukkan bahwa integrasi *deep feature extraction* dan *ensemble learning* efektif dalam meningkatkan kemampuan diskriminatif model pada klasifikasi citra pertanian.

**Kata Kunci:** Deep Learning; Feature Extraction; Ensemble Learning; Stacking; Klasifikasi Citra; Varietas Salak

**Abstract-** Visual classification of salak (snake fruit) varieties remains challenging due to similarities in texture, color, and morphological characteristics across classes. Manual identification is prone to subjectivity and inconsistency in determining varieties. This study proposes an ensemble model based on fusion and stacking applied to deep learning feature extraction in order to improve the accuracy of salak variety classification. Image features are extracted using two pre-trained Convolutional Neural Network architectures, namely VGG16 and ResNet50, as deep feature extractors. The resulting feature representations are subsequently classified using K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) algorithms. The output probabilities of both classifiers are then combined through a stacking ensemble approach with Logistic Regression as the meta-learner. The dataset consists of 584 images distributed across four salak varieties. Model performance is evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results indicate that the proposed fusion–stacking approach achieves an accuracy of 95%, outperforming single CNN-based models and conventional classification methods. These findings demonstrate that the integration of deep feature extraction and ensemble learning effectively enhances the discriminative capability of the model in agricultural image classification.

**Keywords:** Deep Learning; Feature Extraction; Ensemble Learning; Stacking; Image Classification; Salak Varieties

## 1. PENDAHULUAN

Perkembangan teknologi kecerdasan buatan dalam beberapa tahun terakhir telah mendorong kemajuan signifikan pada bidang *computer vision*, khususnya dalam analisis dan interpretasi citra digital. *Computer vision* bertujuan memungkinkan sistem komputer untuk memahami dan menafsirkan informasi visual secara otomatis sehingga dapat digunakan pada berbagai aplikasi seperti pengenalan objek, klasifikasi citra, serta analisis visual pada sektor industri maupun pertanian [1]; [2]. Selain itu, integrasi teknologi *computer vision* dengan berbagai sistem digital seperti Internet of Things juga semakin berkembang dalam berbagai aplikasi cerdas [3]. Integrasi metode pengolahan citra dengan algoritma pembelajaran mesin telah menciptakan peluang baru dalam pengembangan sistem identifikasi objek berbasis citra yang memiliki tingkat akurasi dan efisiensi yang lebih baik.

Kemajuan tersebut semakin diperkuat dengan hadirnya pendekatan *deep learning*, yang memungkinkan model pembelajaran untuk mengekstraksi fitur dari data secara otomatis tanpa memerlukan proses perancangan fitur secara manual. Arsitektur jaringan saraf dengan banyak lapisan dapat mengekstraksi pola visual secara bertahap dari tingkat sederhana hingga kompleks, sehingga menghasilkan fitur yang lebih informatif dibandingkan teknik tradisional [4]; Dalam bidang *computer vision* modern, arsitektur *Convolutional Neural Network (CNN)* banyak digunakan sebagai metode utama dalam berbagai tugas klasifikasi citra karena kemampuannya dalam menangkap pola dan hubungan spasial pada data gambar [6].

Berbagai studi sebelumnya melaporkan bahwa arsitektur CNN memiliki kemampuan yang tinggi dalam menangani beragam tugas analisis visual, termasuk pengenalan aktivitas manusia, deteksi objek, dan klasifikasi citra pada berbagai domain aplikasi [7]; [8]. Selain itu, perkembangan penelitian terbaru juga menunjukkan bahwa integrasi kecerdasan buatan dengan teknologi *computer vision* semakin luas digunakan dalam berbagai sektor seperti kesehatan, keamanan, hingga sistem pertanian cerdas [9]; [10]. Dalam sektor pertanian modern, *computer vision* berperan penting dalam otomatisasi identifikasi tanaman, pemantauan pertumbuhan tanaman, serta klasifikasi produk pertanian berbasis citra digital [11].

Dalam sektor hortikultura, buah salak merupakan salah satu komoditas yang memiliki nilai ekonomi yang cukup signifikan. Buah ini memiliki berbagai varietas dengan karakteristik visual yang relatif mirip, terutama pada tekstur kulit, pola duri, serta variasi warna permukaan. Kemiripan karakteristik visual tersebut seringkali menyulitkan proses identifikasi varietas secara manual. Berdasarkan kondisi tersebut, diperlukan pengembangan sistem klasifikasi otomatis berbasis *computer vision* yang dapat mengidentifikasi varietas salak secara objektif serta memberikan hasil yang konsisten.

Meskipun model CNN mampu menghasilkan representasi fitur yang kuat, penggunaan satu arsitektur tunggal belum tentu mampu memberikan performa optimal pada dataset dengan tingkat kemiripan visual yang tinggi. Oleh karena itu, beberapa penelitian mulai

mengembangkan pendekatan integrasi beberapa model melalui metode *ensemble learning* untuk meningkatkan kinerja klasifikasi. Ensemble learning merupakan teknik yang menggabungkan beberapa model pembelajaran untuk menghasilkan keputusan yang lebih akurat dibandingkan model tunggal [12].

Salah satu pendekatan ensemble yang banyak digunakan adalah metode *stacking*, di mana prediksi dari beberapa model dasar digunakan sebagai masukan bagi model tingkat lanjut (*meta-learner*) untuk menghasilkan prediksi akhir. Pendekatan ini terbukti mampu meningkatkan performa sistem klasifikasi pada berbagai aplikasi berbasis deep learning [13]; [14]. Selain itu, integrasi fitur dari beberapa arsitektur CNN juga dapat meningkatkan kualitas representasi fitur melalui proses *feature fusion*, sehingga informasi visual yang diperoleh menjadi lebih kaya dan diskriminatif [15]; [16].

Pendekatan serupa juga telah diterapkan pada berbagai penelitian klasifikasi citra berbasis deep feature, seperti pada klasifikasi makanan menggunakan kombinasi ekstraksi fitur mendalam dan algoritma pembelajaran mesin tradisional [17]. Hasil penelitian tersebut menunjukkan bahwa kombinasi antara *deep feature extraction* dan metode ensemble dapat meningkatkan performa klasifikasi secara signifikan dibandingkan pendekatan konvensional.

Meskipun berbagai penelitian telah mengembangkan metode ensemble berbasis deep learning, sebagian besar penelitian masih berfokus pada satu arsitektur ekstraksi fitur atau satu jenis algoritma klasifikasi. Penelitian yang mengintegrasikan beberapa arsitektur CNN melalui proses *feature fusion* dan menggabungkannya dengan pendekatan *stacking ensemble* untuk klasifikasi citra pertanian masih relatif terbatas. Oleh karena itu, diperlukan penelitian yang mengeksplorasi integrasi berbagai metode tersebut untuk meningkatkan performa klasifikasi pada dataset dengan karakteristik visual yang kompleks.

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan model klasifikasi citra berbasis *ensemble fusion-stacking* dengan memanfaatkan dua arsitektur CNN pralatih, yaitu VGG16 dan ResNet50, sebagai *deep feature extractor*. Fitur yang dihasilkan kemudian digabungkan melalui proses *feature fusion* dan diklasifikasikan menggunakan algoritma KNN dan SVM sebagai *base learner*. Selanjutnya, probabilitas prediksi dari kedua model tersebut digunakan sebagai masukan bagi Logistic Regression sebagai *meta-learner* dalam kerangka *stacking ensemble*.

Penelitian ini bertujuan untuk meningkatkan kinerja klasifikasi varietas buah salak dengan memanfaatkan integrasi antara ekstraksi fitur mendalam dan metode *ensemble learning*. Melalui pendekatan tersebut, diharapkan sistem klasifikasi citra yang dikembangkan mampu menghasilkan prediksi yang lebih akurat, stabil, serta mampu beradaptasi terhadap variasi visual pada objek yang memiliki karakteristik yang serupa.

## 2. METODOLOGI PENELITIAN

### 2.1 Desain dan Tahapan Penelitian

Dalam penelitian ini digunakan pendekatan eksperimental untuk mengembangkan model klasifikasi citra berbasis *ensemble learning* melalui integrasi ekstraksi fitur mendalam dan metode *stacking*. Pendekatan ini dirancang untuk meningkatkan kemampuan model dalam membedakan karakteristik visual berbagai varietas salak yang memiliki kemiripan tekstur serta warna permukaan.

Pada penelitian di bidang *computer vision*, proses klasifikasi citra umumnya dilakukan melalui beberapa tahapan utama yang meliputi pengumpulan data, prapemrosesan citra, ekstraksi fitur, pelatihan model klasifikasi, serta evaluasi performa model [1]; [2]. Seiring dengan perkembangan teknologi kecerdasan buatan, pendekatan berbasis *deep learning* menjadi metode yang dominan dalam berbagai aplikasi analisis citra karena kemampuannya dalam mempelajari representasi fitur secara otomatis dari data visual [4]; [8].

Pada penelitian ini, tahapan penelitian meliputi: (1) pengumpulan dan prapemrosesan dataset citra salak, (2) ekstraksi fitur menggunakan arsitektur CNN pralatih, (3) proses klasifikasi menggunakan algoritma pembelajaran mesin, (4) penerapan metode *fusion-stacking ensemble*, serta (5) evaluasi kinerja model klasifikasi.

### 2.2 Dataset Penelitian

Penelitian ini memanfaatkan dataset yang berjumlah 584 citra buah salak yang diklasifikasikan ke dalam empat varietas utama, yaitu *Salacca sumatrana sidempuan*, *Salacca wallichiana*, *Salacca zalacca pondoh*, dan *Salacca affinis*. Distribusi data relatif tidak seimbang antar kelas, sehingga evaluasi performa mempertimbangkan metrik berbasis rata-rata tertimbang (*weighted average*). Setiap citra direpresentasikan dalam format RGB dan diproses pada resolusi seragam untuk menjaga konsistensi dimensi masukan pada arsitektur CNN.



salak affinis



salak sidempuan



salak wallichiana



salak pondoh

Gambar 1. Jenis buah salak

### 2.3 Prapemrosesan Citra

Tahapan prapemrosesan dilakukan untuk meningkatkan kualitas citra dan memastikan kompatibilitas dengan model pralatih. Langkah-langkah yang diterapkan meliputi:

1. *Resizing* citra ke ukuran standar 224×224 piksel.
2. Normalisasi nilai piksel ke rentang [0,1].
3. Augmentasi Meliputi Random rotation, Flip, Transform dan Flip Transform

Normalisasi dilakukan untuk mempercepat konvergensi dan menjaga stabilitas proses ekstraksi fitur [12].

## 2.4 Ekstraksi Fitur Mendalam (*Deep Feature Extraction*)

Penelitian ini menggunakan dua arsitektur CNN pralatih, yaitu VGG16 dan ResNet50, sebagai *feature extractor*. Kedua model telah dilatih pada dataset skala besar ImageNet sehingga mampu menghasilkan representasi fitur visual yang kaya dan hierarkis [8]. Hasil penelitian sebelumnya menunjukkan bahwa kombinasi *deep feature extraction* dengan teknik *ensemble learning* berpotensi meningkatkan akurasi klasifikasi citra dibandingkan model tunggal, karena pendekatan ini memungkinkan pemodelan berbagai karakteristik visual secara lebih komprehensif [18]. Lapisan klasifikasi akhir pada masing-masing model dihapus, dan keluaran dari lapisan *fully connected* terakhir digunakan sebagai vektor fitur. Pendekatan ini memungkinkan pemanfaatan kemampuan representasi CNN tanpa melakukan pelatihan ulang secara penuh (*transfer learning berbasis ekstraksi fitur*).

Jika dinotasikan sebagai:

$F_{VGG}$  = fitur hasil ekstraksi VGG16

$F_{ResNet}$  = fitur hasil ekstraksi ResNet50

Maka proses *feature fusion* dilakukan melalui penggabungan:

$$F_{fusion} = [F_{VGG}, F_{ResNet}]$$

Strategi ini memperkaya informasi diskriminatif karena masing-masing arsitektur memiliki karakteristik representasi yang berbeda [16].

## 2.5 Klasifikasi Dasar (*Base Learners*)

Vektor fitur hasil *fusion* selanjutnya diklasifikasikan menggunakan dua algoritma pembelajaran mesin tradisional:

1. *K-Nearest Neighbor* (KNN)  
KNN termasuk dalam algoritma klasifikasi berbasis jarak yang menentukan kategori suatu data dengan mempertimbangkan kedekatan terhadap beberapa sampel tetangga terdekat dalam ruang fitur. Metode ini banyak digunakan dalam sistem klasifikasi citra karena memiliki proses komputasi yang relatif sederhana serta mampu memodelkan distribusi data secara lokal.
2. *Support Vector Machine* (SVM)  
Support Vector Machine (SVM) adalah metode klasifikasi yang bekerja dengan membangun hiperbidang pemisah optimal sehingga margin antar kelas pada ruang fitur berdimensi tinggi dapat dimaksimalkan. Metode ini sering digunakan pada berbagai aplikasi computer vision karena memiliki kemampuan generalisasi yang baik [12].

Kedua model ini dipilih karena memiliki karakteristik keputusan yang berbeda. SVM unggul dalam pemisahan margin maksimum, sedangkan KNN bersifat berbasis jarak lokal [13].

## 2.6 Stacking Ensemble

Untuk meningkatkan performa klasifikasi, diterapkan metode *stacking ensemble*. Teknik stacking telah banyak digunakan dalam berbagai penelitian karena mampu menggabungkan kekuatan beberapa model pembelajaran sehingga menghasilkan prediksi yang lebih stabil dan akurat [13]. Pada tahap ini, probabilitas keluaran dari KNN dan SVM digunakan sebagai fitur baru yang menjadi masukan bagi model tingkat kedua (*meta-learner*), yaitu Logistic Regression.

Jika:

$P_{KNN}$  = probabilitas prediksi KNN

$P_{SVM}$  = probabilitas prediksi SVM

Maka vektor masukan meta-learner adalah:

$$Z = [P_{KNN}, P_{SVM}]$$

Logistic Regression kemudian mempelajari kombinasi optimal dari kedua probabilitas tersebut untuk menghasilkan prediksi akhir [14]. Pendekatan ini memanfaatkan kekuatan komplementer antar model dasar dan meningkatkan stabilitas prediksi akhir.

## 2.7 Evaluasi Kinerja Model

Evaluasi dilakukan menggunakan metrik: Accuracy, Precision, Recall, F1-Score, Confusion Matrix dan Stratified cross-validation dengan nilai K=10.

Rumus akurasi dinyatakan sebagai:

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

Dengan TP, TN, FP, dan FN masing-masing merepresentasikan *true positive*, *true negative*, *false positive*, dan *false negative*. Penilaian performa model dilakukan menggunakan data uji dari Stratified cross-validation dengan nilai K=10 yang tidak digunakan pada tahap pelatihan guna mengevaluasi kemampuan generalisasi model terhadap data yang belum pernah diproses sebelumnya.

### 3. HASIL DAN PEMBAHASAN

Pada bagian ini disampaikan hasil eksperimen yang diperoleh beserta analisis terhadap performa model klasifikasi yang diusulkan dalam penelitian ini. Evaluasi dilakukan dengan membandingkan beberapa pendekatan klasifikasi yang berbeda, yaitu model berbasis ekstraksi fitur mendalam dari arsitektur CNN pralatih yang dikombinasikan dengan algoritma pembelajaran mesin konvensional serta model ensemble berbasis *fusion-stacking*. Tujuan utama dari eksperimen ini adalah untuk mengevaluasi sejauh mana integrasi beberapa model mampu meningkatkan kemampuan diskriminatif sistem dalam mengklasifikasikan varietas buah salak yang memiliki karakteristik visual yang relatif serupa.

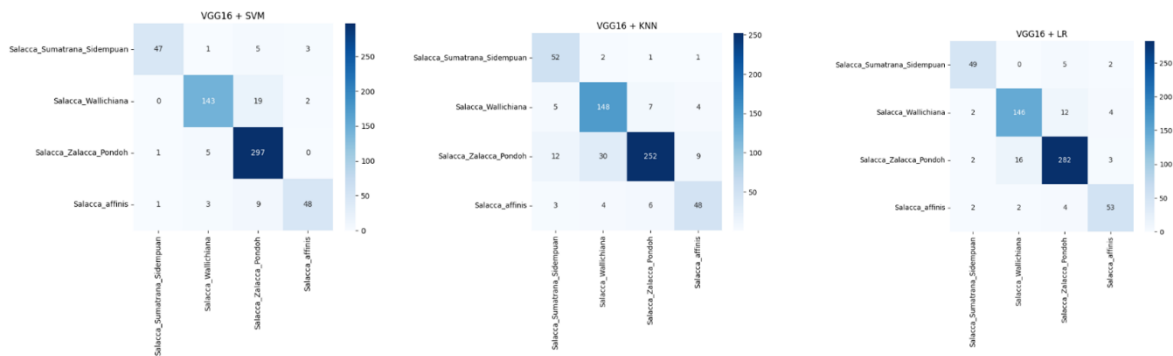
Penelitian ini memanfaatkan dataset yang berjumlah 584 citra buah salak yang diklasifikasikan ke dalam empat varietas, yaitu Salacca sumatrana sidempuan, Salacca wallichiana, Salacca zalacca pondoh, dan Salacca affinis. Seluruh citra terlebih dahulu diproses melalui tahap prapemrosesan sebelum digunakan pada proses ekstraksi fitur dan pelatihan model. Evaluasi kinerja model dilakukan dengan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta analisis *confusion matrix* untuk mengidentifikasi distribusi kesalahan klasifikasi pada masing-masing kelas.

#### 3.1 Hasil Klasifikasi Menggunakan Ekstraksi Fitur VGG16

Eksperimen pertama dilakukan dengan memanfaatkan arsitektur VGG16 sebagai *deep feature extractor*. Arsitektur ini dikenal memiliki struktur jaringan konvolusi yang dalam dan mampu menghasilkan representasi fitur visual yang efektif untuk berbagai tugas klasifikasi citra [8].

Vektor fitur yang diekstraksi menggunakan model VGG16 selanjutnya diproses menggunakan tiga algoritma klasifikasi, yaitu KNN, SVM, dan Logistic Regression. Hasil evaluasi menunjukkan bahwa model VGG16 yang dikombinasikan dengan Logistic Regression memperoleh akurasi sebesar 91%, sedangkan kombinasi dengan SVM mencapai akurasi sebesar 92%. Sementara itu, metode KNN menghasilkan performa yang lebih rendah dengan akurasi sebesar 86%.

Perbedaan kinerja ini menunjukkan bahwa algoritma klasifikasi yang berbeda memiliki kemampuan yang berbeda dalam memproses fitur berdimensi tinggi yang dihasilkan oleh model CNN. SVM dan Logistic Regression cenderung lebih efektif dalam memisahkan kelas pada ruang fitur berdimensi tinggi karena mampu membentuk batas keputusan global yang lebih stabil dibandingkan KNN yang berbasis kedekatan jarak lokal.



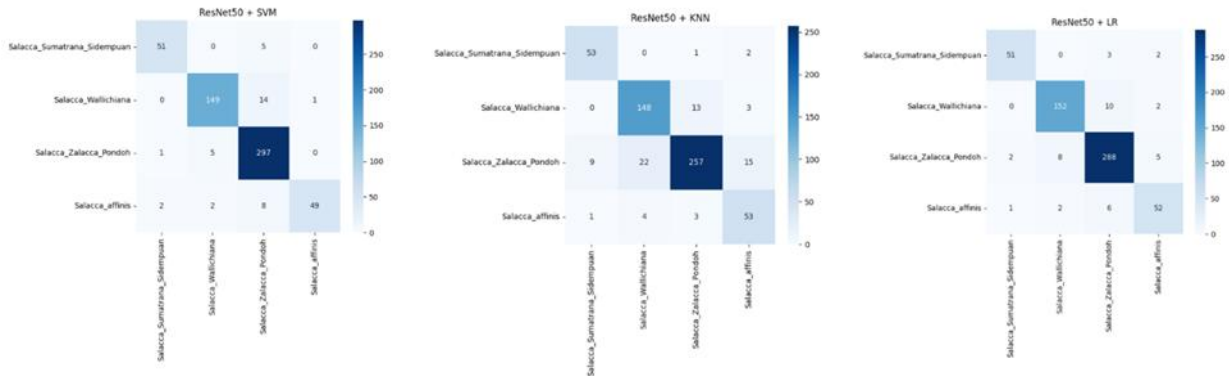
**Gambar 2.** Confusion matrix fitur VGG16

Gambar 2 memperlihatkan *confusion matrix* dari model klasifikasi berbasis VGG16 dengan SVM, KNN dan Logistic Regression. Matriks tersebut menunjukkan bahwa sebagian besar citra berhasil diklasifikasikan dengan benar. Namun demikian, masih terdapat beberapa kesalahan klasifikasi antara kelas *Salacca wallichiana* dan *Salacca zalacca pondoh*. Hal ini menunjukkan bahwa kedua varietas tersebut memiliki karakteristik visual yang relatif mirip sehingga sulit dibedakan oleh model berbasis fitur tunggal.

#### 3.2 Hasil Klasifikasi Menggunakan Ekstraksi Fitur ResNet50

Eksperimen berikutnya dilakukan menggunakan arsitektur ResNet50 sebagai *feature extractor*. ResNet50 merupakan arsitektur CNN yang memperkenalkan konsep *residual learning*, yang memungkinkan jaringan yang lebih dalam untuk dilatih secara lebih stabil dan mengatasi permasalahan degradasi performa pada jaringan yang sangat dalam [9]. Hasil eksperimen menunjukkan bahwa kombinasi ResNet50 dengan Logistic Regression maupun ResNet50 dengan SVM menghasilkan akurasi sebesar 93%, sedangkan kombinasi dengan KNN menghasilkan akurasi sebesar 88%.

Peningkatan performa ini menunjukkan bahwa representasi fitur yang dihasilkan oleh ResNet50 lebih mampu menangkap variasi tekstur dan pola visual pada citra buah salak dibandingkan VGG16. Hal ini menunjukkan bahwa kedalaman jaringan serta mekanisme residual connection berperan penting dalam meningkatkan kualitas representasi fitur pada model CNN.



**Gambar 3.** Confucion matrix fitur ResNet50

Gambar 3 menunjukkan *confusion matrix* dari model ResNet50 dengan SVM, KNN dan Logistic Regression. Berdasarkan matriks tersebut terlihat bahwa sebagian besar citra pada kelas *Salacca zalacca pondoh* berhasil diklasifikasikan dengan sangat baik. Namun demikian, masih terdapat beberapa kesalahan prediksi pada kelas *Salacca affinis*. Hal ini kemungkinan disebabkan oleh jumlah sampel pada kelas tersebut yang relatif lebih sedikit dibandingkan kelas lainnya.

### 3.3 Analisis Perbandingan Model Klasifikasi

Untuk memperoleh gambaran yang lebih komprehensif mengenai kinerja masing-masing model, dilakukan perbandingan performa berdasarkan nilai akurasi yang diperoleh dari setiap kombinasi metode ekstraksi fitur dan algoritma klasifikasi.

**Tabel 1.** Perbandingan Hasil Fitur VGG16, ResNet50 dan Fucion+Stacking

Metode Ekstraksi Fitur	Algoritma Klasifikasi	Accuracy	Precision	Recall	F1-Score
VGG16	KNN	0.86	0.86	0.86	0.86
VGG16	Logistic Regression	0.91	0.91	0.91	0.91
VGG16	SVM	0.92	0.92	0.92	0.92
ResNet50	KNN	0.88	0.88	0.88	0.88
ResNet50	Logistic Regression	0.93	0.93	0.93	0.93
ResNet50	SVM	0.93	0.93	0.93	0.93
Fusion + Stacking	KNN + SVM + Logistic Regression	0.95	0.95	0.95	0.95

### 3.4 Hasil Model Ensemble Fusion–Stacking

Untuk meningkatkan performa klasifikasi, penelitian ini mengusulkan pendekatan *ensemble berbasis fusion dan stacking*. Pada tahap pertama dilakukan proses *feature fusion* dengan menggabungkan fitur yang dihasilkan oleh dua arsitektur CNN, yaitu VGG16 dan ResNet50. Pendekatan ini memungkinkan integrasi berbagai informasi visual sehingga menghasilkan representasi fitur yang lebih kaya [18]. Selanjutnya, vektor fitur hasil *fusion* diproses oleh dua algoritma klasifikasi sebagai *base learner*, yaitu KNN dan SVM. Probabilitas keluaran dari kedua model tersebut kemudian digunakan sebagai masukan bagi Logistic Regression yang berperan sebagai *meta-learner* dalam kerangka *stacking ensemble*. Pendekatan *stacking ensemble* memungkinkan model untuk memanfaatkan kekuatan komplementer dari beberapa algoritma klasifikasi sehingga menghasilkan prediksi yang lebih stabil dan akurat [14]. Hasil eksperimen menunjukkan bahwa model *fusion–stacking* yang diusulkan mampu mencapai akurasi sebesar 95%, yang merupakan performa tertinggi dibandingkan model lainnya.



**Gambar 4.** Confucion matrix fusion-stacking

Gambar 4 memperlihatkan *confusion matrix* dari model ensemble yang diusulkan. Berdasarkan gambar tersebut terlihat bahwa jumlah kesalahan klasifikasi mengalami penurunan dibandingkan model sebelumnya. Sebagian besar citra berhasil diklasifikasikan dengan benar pada setiap kelas varietas salak. Hal ini menunjukkan bahwa pendekatan *fusion-stacking* mampu meningkatkan kemampuan model dalam membedakan karakteristik visual antar varietas.

### 3.5 Pembahasan

Berdasarkan hasil eksperimen yang telah dilakukan, dapat disimpulkan bahwa pendekatan ekstraksi fitur mendalam yang dikombinasikan dengan metode pembelajaran mesin konvensional mampu menghasilkan performa klasifikasi yang cukup baik. Namun demikian, penggunaan satu arsitektur CNN saja belum sepenuhnya mampu menangkap kompleksitas karakteristik visual pada dataset varietas salak.

Integrasi fitur dari dua arsitektur CNN melalui proses *feature fusion* terbukti mampu memperkaya representasi fitur yang digunakan dalam proses klasifikasi. Selain itu, penerapan metode *stacking ensemble* memungkinkan penggabungan kekuatan beberapa model klasifikasi sehingga menghasilkan prediksi yang lebih akurat dan stabil. Temuan penelitian ini sejalan dengan beberapa penelitian sebelumnya yang menunjukkan bahwa integrasi metode *ensemble learning* pada sistem klasifikasi berbasis deep learning mampu meningkatkan performa prediksi dibandingkan model tunggal [13].

Dengan demikian, pendekatan *ensemble fusion-stacking* yang diusulkan dalam penelitian ini terbukti efektif dalam meningkatkan akurasi klasifikasi citra varietas salak dan berpotensi untuk diterapkan pada berbagai sistem klasifikasi citra pertanian lainnya.

## 4. KESIMPULAN

Penelitian ini mengusulkan pendekatan klasifikasi citra berbasis *ensemble fusion-stacking* untuk mengidentifikasi varietas buah salak menggunakan representasi fitur mendalam dari arsitektur CNN pralatih. Proses ekstraksi fitur dilakukan menggunakan dua model *Convolutional Neural Network*, yaitu VGG16 dan ResNet50, yang berfungsi sebagai *deep feature extractor*. Fitur yang dihasilkan dari kedua model tersebut kemudian digabungkan melalui mekanisme *feature fusion* sehingga menghasilkan representasi fitur yang lebih kaya dan informatif. Selanjutnya, proses klasifikasi dilakukan menggunakan algoritma KNN dan SVM sebagai *base learner*, yang kemudian digabungkan melalui pendekatan *stacking ensemble* dengan Logistic Regression sebagai *meta-learner*. Berdasarkan hasil eksperimen yang diperoleh, model yang menggunakan ekstraksi fitur tunggal menunjukkan kinerja yang cukup baik, dengan akurasi tertinggi sebesar 93% pada kombinasi ekstraksi fitur ResNet50 dan algoritma klasifikasi SVM. Namun demikian, penerapan metode *fusion-stacking ensemble* mampu meningkatkan performa sistem klasifikasi secara lebih signifikan. Model ensemble yang diusulkan berhasil mencapai akurasi sebesar 95%, disertai peningkatan nilai *precision*, *recall*, dan *F1-score* dibandingkan model tunggal. Temuan ini menunjukkan bahwa penggabungan beberapa model melalui pendekatan ensemble dapat meningkatkan kemampuan sistem dalam membedakan varietas salak yang memiliki karakteristik visual yang serupa. Hasil penelitian ini menunjukkan bahwa integrasi antara teknik deep feature extraction dan metode ensemble learning dapat menjadi pendekatan yang efektif dalam meningkatkan performa sistem klasifikasi citra, khususnya pada objek yang memiliki kompleksitas visual yang tinggi. Selain itu, pemanfaatan dua arsitektur CNN yang berbeda memungkinkan model memperoleh representasi fitur yang lebih beragam sehingga dapat meningkatkan stabilitas dalam proses prediksi. Untuk penelitian selanjutnya, pendekatan yang diusulkan dapat dikembangkan lebih lanjut dengan mengeksplorasi arsitektur CNN lain yang lebih modern, seperti EfficientNet atau Vision Transformer. Selain itu, penggunaan dataset dengan jumlah sampel yang lebih besar serta penerapan teknik augmentasi citra juga berpotensi meningkatkan kemampuan generalisasi model dalam berbagai kondisi citra yang lebih beragam.

## REFERENCES

- [1] F. Aulia, "Computer Vision dan Pengolahan Citra Digital".
- [2] David Shamoo Excel, "Computer Vision Promising Innovations," *World J. Adv. Res. Rev.*, vol. 23, no. 3, hlm. 610–619, Sep 2024, doi: 10.30574/wjarr.2024.23.3.2725.
- [3] L. Sharma dan M. Carpenter, *Computer Vision and Internet of Things: Technologies and Applications*, 1 ed. Boca Raton: Chapman and Hall/CRC, 2022. doi: 10.1201/9781003244165.
- [4] A. Betti, M. Gori, dan S. Melacci, *Deep Learning to See: Towards New Foundations of Computer Vision*. 2022. doi: 10.1007/978-3-030-90987-1.
- [5] Norbertus Tri Suswanto Saptadi, Hedi Kristiawan, Agung Yuliyanto Nugroho, Nina Rahayu, Suwarniyati, dan Bayu Waseso, "Deep Learning: Teori, Algoritma, dan Aplikasi." Diakses: 3 Maret 2026. [Daring]. Tersedia pada: [https://www.researchgate.net/publication/389428653\\_Deep\\_Learning\\_Teori\\_Algoritma\\_dan\\_Aplikasi?utm\\_source=chatgpt.com](https://www.researchgate.net/publication/389428653_Deep_Learning_Teori_Algoritma_dan_Aplikasi?utm_source=chatgpt.com)
- [6] K. S. Babulal dan A. K. Das, "Deep Learning-Based Object Detection: An Investigation," dalam *Futuristic Trends in Networks and Computing Technologies*, vol. 936, P. K. Singh, S. T. Wierchoń, J. K. Chhabra, dan S. Tanwar, Ed., dalam *Lecture Notes in Electrical Engineering*, vol. 936. , Singapore: Springer Nature Singapore, 2022, hlm. 697–711. doi: 10.1007/978-981-19-5037-7\_50.
- [7] F. X. Gaya-Morey, C. Manresa-Yee, dan J. M. Buades-Rubio, "Deep learning for computer vision based activity recognition and fall detection of the elderly: a systematic review," *Appl Intell*, vol. 54, no. 19, hlm. 8982–9007, Okt 2024, doi: 10.1007/s10489-024-05645-1.
- [8] M. H. Mozaffari, "Deep Learning for Computer Vision Application," *Electronics*, vol. 14, no. 14, hlm. 2874, Jul 2025, doi: 10.3390/electronics14142874.
- [9] A. Wang, H. Wu, dan Y. Iwahori, "Advances in Computer Vision and Deep Learning and Its Applications," *Electronics*, vol. 14, no. 8, hlm. 1551, Apr 2025, doi: 10.3390/electronics14081551.

- [10] B. Liu, L. Yu, C. Che, Q. Lin, H. Hu, dan X. Zhao, "Integration and performance analysis of artificial intelligence and computer vision based on deep learning algorithms," *ACE*, vol. 64, no. 1, hlm. 36–41, Mei 2024, doi: 10.54254/2755-2721/64/20241374.
- [11] Z. Cao, S. Sun, dan X. Bao, "A Review of Computer Vision and Deep Learning Applications in Crop Growth Management," *Applied Sciences*, vol. 15, no. 15, hlm. 8438, Jul 2025, doi: 10.3390/app15158438.
- [12] D. Patil, N. L. Rane, P. Desai, dan J. Rane, "Machine learning and deep learning: Methods, techniques, applications, challenges, and future research opportunities," dalam *Trustworthy Artificial Intelligence in Industry and Society*, Deep Science Publishing, 2024. doi: 10.70593/978-81-981367-4-9\_2.
- [13] G. Naskar, S. Mohiuddin, S. Malakar, E. Cuevas, dan R. Sarkar, "Deepfake detection using deep feature stacking and meta-learning," *Heliyon*, vol. 10, no. 4, hlm. e25933, Feb 2024, doi: 10.1016/j.heliyon.2024.e25933.
- [14] E. Alalwany, B. Alsharif, Y. Alotaibi, A. Alfahaid, I. Mahgoub, dan M. Ilyas, "Stacking Ensemble Deep Learning for Real-Time Intrusion Detection in IoMT Environments," *Sensors*, vol. 25, no. 3, hlm. 624, Jan 2025, doi: 10.3390/s25030624.
- [15] N. Mungoli, "Adaptive Ensemble Learning: Boosting Model Performance through Intelligent Feature Fusion in Deep Neural Networks," 4 April 2023, *arXiv*: arXiv:2304.02653. doi: 10.48550/arXiv.2304.02653.
- [16] B. N. Jyothi dan M. A. Jabbar, "Enhanced Xception Model for Deepfake Detection: Integrating CBAM, Contrastive Learning, and a Stacking Classifier".
- [17] M. Y. Kardawi, F. M. Saragih, L. Rahadiani, dan A. M. Arymurthy, "Indonesian Food Classification Using Deep Feature Extraction and Ensemble Learning for Dietary Assessment," vol. 9, no. 5, 2009.
- [18] K. Pañç dan S. Sekmen, "Multi-CNN Deep Feature Fusion and Stacking Ensemble Classifier for Breast Ultrasound Lesion Classification," *forbes*, Jul 2025, doi: 10.4274/forbes.galenos.2025.02360.
- [19] A. Karim, I. Purnama, and A. Ernawati, "Peningkatan Pengarahan Beam dan Estimasi Sudut Kedatangan Berbasis CNN untuk Sistem Antena MIMO Cerdas," *Explorer (Hayward)*, vol. 6, no. 1, pp. 63–72, Jan. 2026, doi: 10.47065/explorer.v6i1.2592.
- [20] A. Karim, B. Bangun, S. Prayetno, M. Afrendi, and K. Kunci, "Optimasi Prediksi Harga Sawit Menggunakan Teknik Stacking Algoritma Machine Learning dan Deep Learning dengan SMOTE," *Technology and Science (BITS)*, vol. 7, no. 1, 2025, doi: 10.47065/bits.v7i1.7239.