

IMPLEMENTASI *CLASS WEIGHT*, *DATA AUGMENTATION* DAN *RESNET101* UNTUK MENGATASI *OVERFITTING* DAN *IMBALANCED DATA* PADA KLASIFIKASI *PISTACHIO*

Moch. Raffi Dwi Saktya Rahman ¹⁾, Arie Nugroho ²⁾, Aidina Ristyawan ³⁾

^{1,2,3)} Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri
Email: raffidwi110@gmail.com ¹⁾, arienugroho@unpkediri.ac.id ²⁾, aidinaristi@unpkediri.ac.id ³⁾

Dikirim: 19 Mei 2026; Disetujui: 29 Juni 2026; Dipublikasikan: 31 Juli 2026

ABSTRAK

Klasifikasi kacang *Pistachio* varietas *Kirmizi* dan *Siirt* secara manual rentan terhadap kesalahan akibat kemiripan morfologi antar varietas, sementara penelitian sebelumnya menggunakan *ResNet50* mengalami *overfitting* parah dengan *gap accuracy* 78% dan belum menangani ketidakseimbangan data secara eksplisit. Penelitian ini bertujuan mengimplementasikan *ResNet101* dengan augmentasi data dan *class weighting* untuk meningkatkan akurasi klasifikasi sekaligus mengatasi *overfitting* dan *imbalanced data*. Metode yang digunakan adalah kerangka *CRISP-DM* dengan 2148 citra (rasio 80:10:10) dan empat skenario eksperimen kombinasi data augmentasi dan *class weighting* berbasis *Transfer Learning ResNet101*. Hasil menunjukkan *ResNet101* dengan *class weighting* tanpa augmentasi mencapai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 98,31%, melampaui *Inception-V3* (96%) sebesar 2,31 poin dan *ResNet50* (86%) sebesar 12,31 poin, dengan *class weighting* lebih bermanfaat saat dikombinasikan augmentasi dibanding kondisi data bersih. Augmentasi justru menurunkan akurasi karena latar belakang citra yang *homogen*, menunjukkan efektivitasnya bergantung pada karakteristik dataset dan bukan berlaku universal. Disimpulkan bahwa *ResNet101* dengan *class weighting* tanpa augmentasi merupakan konfigurasi optimal untuk dataset *Pistachio* berkarakteristik *visual homogen*, dan penelitian lanjutan disarankan mengeksplorasi arsitektur *modern* dengan dataset lebih bervariasi.

Kata Kunci : Klasifikasi Pistachio, *Transfer Learning*, *ResNet101*, Augmentasi Data, *Class Weight*.

ABSTRACT

Manual classification of *Kirmizi* and *Siirt* pistachio varieties is prone to errors due to their similar morphology, while previous research using *ResNet50* suffered severe overfitting with a 78% accuracy gap and did not explicitly address class imbalance. This study aims to implement *ResNet101* with data augmentation and class weighting to improve classification accuracy while addressing overfitting and imbalanced data. The *CRISP-DM* framework was adopted using 2,148 images (80:10:10 ratio) and four experimental scenarios combining data augmentation and class weighting based on *ResNet101 Transfer Learning*. Results show that *ResNet101* with class weighting without augmentation achieved accuracy, precision, recall, and F1-score of 98.31%, surpassing *Inception-V3* (96%) by 2.31 points and *ResNet50* (86%) by 12.31 points, with class weighting proving more beneficial when combined with augmentation than on clean homogeneous data. Augmentation reduced accuracy due to the dataset's homogeneous background, indicating its effectiveness depends on dataset characteristics rather than being universally applicable. *ResNet101* with class weighting without augmentation is concluded as the optimal configuration for pistachio datasets with homogeneous visual characteristics, and future research is recommended to explore modern architectures with more varied datasets.

Keywords : *Pistachio Classification*, *Transfer Learning*, *ResNet101*, *Data Augmentation*, *Class Weight*.

1. PENDAHULUAN

Sektor pertanian merupakan aspek penting dalam pembangunan ekonomi suatu negara, di mana inovasi dalam proses pascapanen menjadi semakin dibutuhkan untuk meningkatkan efisiensi dan daya saing global [1]. Kacang Pistachio merupakan komoditas pertanian bernilai ekonomi tinggi dengan Amerika Serikat, Iran, Turki, dan Suriah sebagai produsen utama yang menyumbang hampir 90% produksi dunia [2]. Proses klasifikasi varietas Pistachio pascapanen masih dilakukan secara manual, padahal varietas Kirmizi dan Siirt memiliki kemiripan morfologi yang sangat tinggi sehingga sulit dibedakan secara kasat mata, yang berpotensi menyebabkan kesalahan klasifikasi dan kerugian ekonomi yang signifikan [3].

Perkembangan teknologi *deep learning* telah membuka peluang besar untuk mengotomatisasi proses klasifikasi produk pertanian. *Convolutional Neural Network (CNN)* merupakan salah satu arsitektur *neural network* yang dirancang khusus untuk memproses data citra, dengan kemampuan mengekstraksi fitur *visual* secara hierarkis dan otomatis tanpa memerlukan rekayasa fitur manual [4]. Pendekatan *Transfer Learning* semakin memperkuat kemampuan *CNN* dengan memanfaatkan model yang telah dilatih sebelumnya pada dataset besar seperti *ImageNet*, sehingga proses pelatihan menjadi lebih efisien dan akurasi model meningkat meskipun data yang tersedia terbatas [5].

Salah satu arsitektur *CNN* yang terbukti unggul dalam tugas klasifikasi citra adalah *ResNet101*, yaitu jaringan saraf konvolusional dengan kedalaman 101 lapisan yang mampu menangani masalah *vanishing gradient* pada jaringan yang sangat dalam melalui mekanisme *residual connection*, sehingga mempercepat dan menstabilkan proses pelatihan [6]. Studi komparatif Gocer & Karakas (2020) pada domain klasifikasi lesi kulit menunjukkan bahwa *ResNet101* menghasilkan akurasi tertinggi sebesar 77,72% dibandingkan *VGG16*, *VGG19*, *GoogleNet*, dan *Inception-V3* pada kondisi pengujian yang identik, memberikan landasan empiris pemilihan *ResNet101* pada penelitian ini [7]. Relevansi *ResNet101* juga dikonfirmasi oleh Oktaviana dkk. (2021) yang mengimplementasikan *ResNet101* dengan

metode *transfer learning* untuk klasifikasi penyakit daun padi, di mana model mampu mencapai akurasi validasi 100% bahkan pada dataset berskala kecil, membuktikan kemampuan generalisasi arsitektur ini yang kuat pada domain klasifikasi citra objek natural. Keunggulan tersebut didukung oleh mekanisme *residual connection* yang menjaga stabilitas pelatihan jaringan 101 lapisan sekaligus mencegah *vanishing gradient*, sebagaimana tercermin dari konvergensi kurva pelatihan yang stabil pada eksperimen tersebut [8].

Berbagai penelitian telah dilakukan untuk mengklasifikasikan Pistachio secara otomatis menggunakan *deep learning*. Singh dkk. (2022) menggunakan *AlexNet*, *VGG16*, dan *VGG19* dengan *Transfer Learning* pada 2148 citra, di mana *VGG16* mencapai akurasi tertinggi 98,84% [3]. Lisda dkk. (2023) membandingkan *Inception-V3* dan *ResNet50* pada dataset yang sama, menghasilkan akurasi masing-masing 96% dan 86%, namun *ResNet50* mengalami *overfitting* parah dengan *gap training-validation accuracy* mencapai 78% [1]. Avuçlu (2023) menggunakan *ResNet* dengan *5-fold cross-validation* memperoleh rata-rata akurasi 86,16% [9]. Özkan dkk. (2021) menggunakan metode *KNN* dengan pendekatan *PCA* memperoleh akurasi tertinggi 94% [10].

Selain permasalahan *overfitting*, ketidakseimbangan distribusi data antar kelas juga menjadi tantangan yang sering diabaikan dalam penelitian klasifikasi citra pertanian. Ketidakseimbangan kelas dapat menyebabkan bias model terhadap kelas mayoritas sehingga kemampuan model dalam mengenali kelas minoritas menjadi buruk. Pada model *deep learning* berbasis *gradient descent*, ketidakseimbangan kelas menyebabkan fungsi *loss* didominasi oleh kontribusi kelas mayoritas, sehingga *gradien* pembaruan bobot model lebih condong meminimalkan kesalahan pada kelas mayoritas dan mengabaikan kelas minoritas meskipun secara keseluruhan *accuracy* tetap terlihat tinggi. Salah satu teknik yang efektif untuk mengatasi permasalahan ini adalah *class weighting*, yaitu memberikan bobot penalti yang lebih tinggi pada kelas minoritas selama proses *training* agar model belajar secara lebih adil terhadap seluruh kelas [11]. Efektivitas *class weighting* sebagai teknik penanganan *imbalanced data* telah dibuktikan secara empiris

lintas domain. Mualfah dkk. (2025) menunjukkan bahwa *Random Forest* dengan *class weight* mampu mencapai akurasi 99,98% pada dataset dengan rasio ketidakseimbangan yang sangat ekstrem (120:1), dengan seluruh metrik precision, recall, dan F1-score mendekati nilai sempurna, membuktikan bahwa pemberian bobot pada kelas minoritas selama proses pelatihan mampu mengatasi bias model terhadap kelas mayoritas secara efektif tanpa perlu memanipulasi distribusi data secara fisik [12].

Berdasarkan kajian tersebut, terdapat tiga *research gap* yang perlu diisi: pertama, *ResNet101* belum pernah diuji pada dataset *Pistachio* varietas *Kirmizi* dan *Siirt*; kedua, penelitian sebelumnya belum menerapkan *data augmentation* untuk mengatasi *overfitting* pada *ResNet*; ketiga, ketidakseimbangan data antara kelas *Kirmizi* (1232 sampel) dan *Siirt* (916 sampel) belum ditangani secara *eksplisit* melalui *class weighting*.

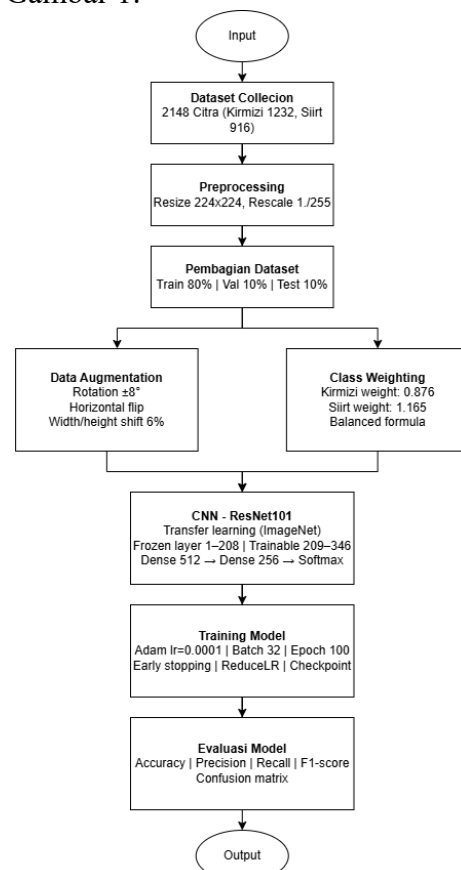
Novelty penelitian ini terletak pada evaluasi sistematis kontribusi *data augmentation* dan *class weighting* secara terpisah maupun kombinasi terhadap *ResNet101* melalui desain empat skenario eksperimen, sehingga dapat diidentifikasi secara empiris teknik mana yang benar-benar berkontribusi terhadap performa model pada dataset dengan karakteristik visual homogen, sebuah analisis yang belum dilakukan pada penelitian klasifikasi *Pistachio* sebelumnya yang umumnya hanya melaporkan hasil dari satu konfigurasi model tunggal.

Dibandingkan dengan tren riset klasifikasi citra terkini yang mulai mengarah pada arsitektur *Vision Transformer* dan *EfficientNet*, penelitian ini memposisikan *ResNet101* sebagai *baseline* arsitektur CNN konvensional yang teruji untuk mengisi gap empiris pada domain klasifikasi *Pistachio*, sekaligus menjadi titik pembandingan bagi penelitian mendatang yang hendak mengevaluasi arsitektur generasi terbaru pada dataset yang sama. Posisi ini relevan karena belum tersedia studi yang secara sistematis mendokumentasikan kontribusi terpisah antara arsitektur, teknik augmentasi, dan penanganan *imbalance* pada klasifikasi varietas *Pistachio*, sehingga hasil penelitian ini dapat menjadi rujukan dasar sebelum eksplorasi arsitektur yang lebih kompleks dilakukan.

Oleh karena itu, penelitian ini bertujuan mengimplementasikan *ResNet101* yang dikombinasikan dengan *data augmentation* dan *class weighting* untuk meningkatkan akurasi klasifikasi sekaligus mengatasi *overfitting* dan *imbalanced data* pada klasifikasi kacang *Pistachio* varietas *Kirmizi* dan *Siirt*.

2. METODE

Penelitian ini mengadaptasi kerangka *CRISP-DM* (*Cross-Industry Standard Process for Data Mining*) untuk konteks deep learning klasifikasi citra *Pistachio*. *CRISP-DM* dipilih karena merupakan metodologi standar industri yang mencakup enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*, sesuai dengan tujuan penelitian untuk menghasilkan model yang siap diimplementasikan dalam sistem produksi [13]. Alur penelitian secara keseluruhan disajikan pada Gambar 1.



Gambar 1. Alur Kerangka Penelitian

Dataset yang digunakan adalah *Pistachio Image Dataset* yang bersumber dari platform Kaggle dengan *lisensi Public Domain* (CC0). Dataset terdiri dari 2148 citra yang terbagi ke dalam dua kelas varietas, yaitu *Kirmizi_Pistachio* sebanyak 1232 citra (57,4%)

dan Siirt_Pistachio sebanyak 916 citra (42,6%), dengan rasio ketidakseimbangan kelas sebesar 1,35:1. Seluruh citra memiliki latar belakang putih atau netral yang konsisten dengan format *JPG/PNG* dalam mode warna *RGB*. Dataset dibagi menggunakan strategi *stratified split* dengan rasio 80:10:10, menghasilkan *data training* sebanyak ± 1718 citra, validasi ± 193 citra, dan *testing* 237 citra.

Preprocessing merupakan tahap penting dalam pengolahan data karena data mentah seringkali belum sempurna dan memerlukan penyesuaian sebelum dapat digunakan secara optimal dalam proses pelatihan model [14]. *Preprocessing* dilakukan dengan *re-size* citra ke ukuran 224×224 piksel sesuai persyaratan input *ResNet101*, dan normalisasi piksel menggunakan *rescaling* $1/255$ untuk mempercepat konvergensi selama *training*. *Data augmentation* diterapkan eksklusif pada *training set* dengan parameter *rotation* $\pm 8^\circ$, *horizontal flip*, dan *width/height shift* 6%.

Pemilihan rentang *rotation* $\pm 8^\circ$ dan *width/height shift* 6% didasarkan pada pertimbangan bahwa morfologi kacang *Pistachio* bersifat *rigid* dan memiliki orientasi alami yang relatif konsisten pada citra hasil pemindaian pascapanen, sehingga rentang transformasi yang terlalu besar berisiko menghasilkan citra yang tidak representatif terhadap kondisi pengambilan gambar yang sesungguhnya. Rentang kecil ini dipilih untuk memperkenalkan variasi posisi yang realistis tanpa mendistorsi fitur *visual* esensial yang membedakan *varietas Kirmizi* dan *Siirt*.

Class weight dihitung menggunakan formula berikut:

$$Class_weight_i = \frac{N_{total}}{K \times N_i}$$

di mana N_{total} adalah total sampel *training*, K adalah jumlah kelas, dan N_i adalah jumlah sampel kelas ke- i . Berdasarkan formula tersebut diperoleh nilai *class weight* untuk kelas *Kirmizi* sebesar 0,876 dan kelas *Siirt* sebesar 1,165.

Penelitian ini merancang empat skenario eksperimen dengan kombinasi teknik yang berbeda untuk menganalisis kontribusi masing-masing teknik secara terpisah maupun kombinasi, sebagaimana disajikan pada Tabel 1.

Tabel 1. Desain Skenario Eksperimen

Skenario	Data Augmentation	Class Weight
1	Ya	Ya
2	Tidak	Ya
3	Ya	Tidak
4	Tidak	Tidak

Model yang digunakan adalah *ResNet101* dengan pendekatan *Transfer Learning* menggunakan bobot *pre-trained ImageNet*. Strategi *fine-tuning* diterapkan dengan membekukan 60% lapisan awal (*layer* 1–208) dan melatih 40% lapisan akhir (*layer* 209–346) untuk menyesuaikan fitur dengan karakteristik citra *Pistachio* [5]. *Classification head* yang ditambahkan terdiri dari *Global Average Pooling 2D*, *Batch Normalization*, *Dense 512* dengan aktivasi *ReLU* dan *L2 regularization* 0,0001, *Dropout* dengan *rate* 0,35, *Batch Normalization*, *Dense 256* dengan aktivasi *ReLU* dan *L2 regularization* 0,0001, *Dropout* dengan *rate* 0,35, dan *Output Dense 2* dengan aktivasi *Softmax*.

Proporsi *fine-tuning* 60:40 dipilih berdasarkan prinsip *transfer learning*, di mana *layer-layer* awal pada jaringan konvolusional bertanggung jawab mengekstraksi fitur generik seperti tepi, tekstur, dan gradien warna yang bersifat *universal* pada berbagai domain citra, sedangkan *layer-layer* akhir mempelajari fitur abstrak yang lebih spesifik terhadap domain data tertentu [5]. Dengan membekukan 60% *layer* awal, fitur generik dari *ImageNet* tetap dipertahankan, sementara 40% *layer* akhir dilatih ulang agar model dapat menyesuaikan representasi fitur abstrak dengan karakteristik *visual* spesifik citra *Pistachio* yang memiliki perbedaan morfologi antarvarietas yang halus. Nilai *dropout rate* 0,35 dan *learning rate* 0,0001 dipilih untuk menjaga keseimbangan antara kapasitas belajar model dan pencegahan *overfitting*, sesuai praktik umum *fine-tuning* pada dataset berskala terbatas.

Training dikonfigurasi menggunakan *optimizer Adam* dengan *learning rate* 0,0001, *categorical crossentropy* sebagai *loss function* yang dikombinasikan dengan *class weight*, *batch size* 32, maksimal 100 *epoch*, dan *random seed*=42 untuk menjamin reproduktibilitas hasil. Tiga *callback* diterapkan selama *training*, yaitu *Early Stopping* dengan *patience*=20 berbasis *val_accuracy* untuk menghentikan *training* jika

tidak ada peningkatan, *ReduceLROnPlateau* dengan *patience*=6 dan *factor*=0,5 untuk mengurangi *learning rate* saat *val_loss* stagnan, serta *ModelCheckpoint* untuk menyimpan bobot model terbaik secara otomatis. Tahap evaluasi akhir berfungsi menggambarkan performa model yang dikembangkan sekaligus memvalidasi tingkat akurasi hasil klasifikasi terhadap data sesungguhnya [15]. Evaluasi dilakukan pada *testing set* sebanyak 237 citra menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix* dengan formula sebagai berikut:

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

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

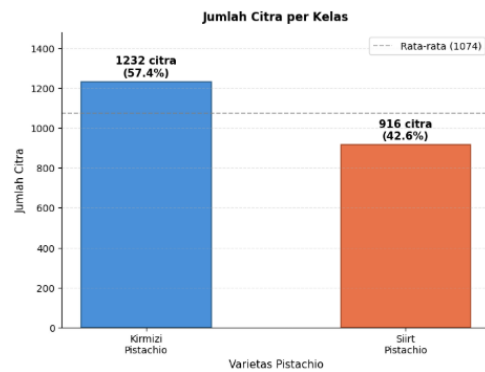
$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian dari empat skenario eksperimen yang dirancang untuk menganalisis kontribusi data augmentation dan class weighting terhadap performa klasifikasi kacang Pistachio varietas Kirmizi dan Siirt menggunakan arsitektur *ResNet101*. Seluruh eksperimen dijalankan menggunakan konfigurasi training yang identik dengan *optimizer Adam*, *learning rate* 0,0001, *batch size* 32, maksimal 100 *epoch*, dan *random seed*=42 untuk menjamin reproduktibilitas hasil. Evaluasi dilakukan pada *testing set* sebanyak 237 citra yang sepenuhnya terpisah dari proses *training* maupun validasi, menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

3.1. Hasil Eksplorasi Data

Dataset Pistachio *Image* terdiri dari 2148 citra dengan distribusi kelas Kirmizi_Pistachio sebanyak 1232 citra (57,4%) dan Siirt_Pistachio sebanyak 916 citra (42,6%), dengan rasio ketidakseimbangan 1,35:1. Distribusi kelas disajikan pada Gambar 2.



Gambar 2. Distribusi Kelas Dataset Pistachio

Hasil eksplorasi visual menunjukkan seluruh citra memiliki latar belakang putih atau netral yang sangat konsisten. Kirmizi memiliki biji berwarna hijau tua dengan bentuk panjang, sedangkan Siirt memiliki biji berwarna kuning kehijauan dengan bentuk oval. Perbedaan morfologi yang halus ini sulit dibedakan secara kasat mata dan menjadi tantangan utama dalam proses klasifikasi [1]

3.2. Ringkasan Hasil Keempat Skenario

Keempat skenario eksperimen dijalankan menggunakan arsitektur dan konfigurasi training yang identik. Ringkasan hasil disajikan pada Tabel 2.

Tabel 2. Ringkasan Hasil Keempat Skenario

S	D A	C W	Epo ch	Test Acc (%)	Val Acc (%)	Gap (%)	Du- rasi (M)
1	Y	Y	100	97,05	98,73	-1,68	92
2	T	Y	41	98,31	98,31	0	23
3	Y	T	85	95,78	97,89	-2,11	56
4	T	T	66	98,31	98,31	0	32

Keterangan:

S = Skenario

DA = *Data Augmentation*

CW = *Class Weight*

Gap = *Test Acc – Val Acc*

Y = Ya

T= Tidak

M = Menit

Skenario 2 dan Skenario 4 menghasilkan *testing accuracy* tertinggi sebesar 98% dengan *gap* 0%, menandakan kemampuan generalisasi yang sangat baik. Skenario 1 dan Skenario 3 yang menggunakan augmentasi menunjukkan *gap negatif* masing-masing -1,68% dan -2,11%, mengindikasikan bahwa model yang terbiasa

dengan data *augmented* mengalami sedikit penurunan performa ketika dihadapkan pada data testing asli tanpa augmentasi.

Perlu dicatat bahwa hasil pada keempat skenario ini diperoleh dari *single experimental run* dengan *random seed* yang tetap (*seed*=42) untuk menjamin reproduisibilitas, namun belum disertai pengulangan eksperimen (*multiple runs*) dengan *seed* berbeda maupun uji signifikansi statistik. Dengan demikian, kestabilan performa antar-*run* belum dapat dipastikan secara statistik, dan perbedaan kecil antara skenario seperti pada Skenario 2 dan Skenario 4 yang menghasilkan akurasi identik (98,31%) perlu diinterpretasikan dengan kehati-hatian sebagai indikasi awal, bukan kesimpulan yang definitif.

3.3. Hasil Per Kelas Setiap Skenario

Detail hasil klasifikasi per kelas disajikan pada Tabel 3.

Tabel 3. Classification Report Keempat Skenario

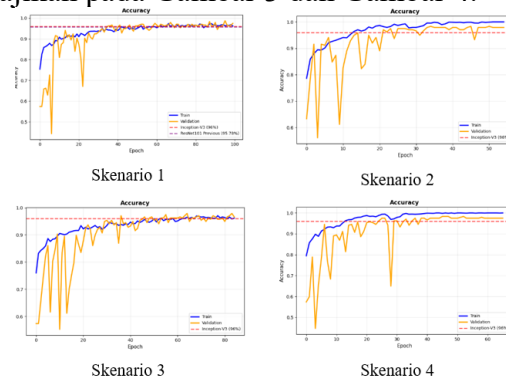
Sken-ario	Kelas	Precis ion (%)	Recall (%)	F1-Score (%)
1	Kirmizi	99,24	95,59	97,38
	Siirt	94,34	99,01	96,62
	Weighte d Avg	97,15	97,05	97,05
2	Kirmizi	98,53	98,53	98,53
	Siirt	98,02	98,02	98,02
	Weighte d Avg	98,31	98,31	98,31
3	Kirmizi	97,01	95,59	96,30
	Siirt	94,17	96,04	95,10
	Weighte d Avg	95,80	95,78	95,79
4	Kirmizi	98,53	98,53	98,53
	Siirt	98,02	98,02	98,02
	Weighte d Avg	98,31	98,31	98,31

Skenario 2 dan Skenario 4 menunjukkan hasil identik dengan *F1-score* Kirmizi 98,53% dan Siirt 98,02%, mengindikasikan *class weighting* tidak memberikan pengaruh signifikan pada dataset bersih dengan ketidakseimbangan ringan. Skenario 3 memperlihatkan selisih *F1-score* terbesar 1,20 poin antara Kirmizi (96,30%) dan Siirt (95,10%), mengkonfirmasi *class weighting*

lebih bermanfaat ketika dikombinasikan dengan augmentasi.

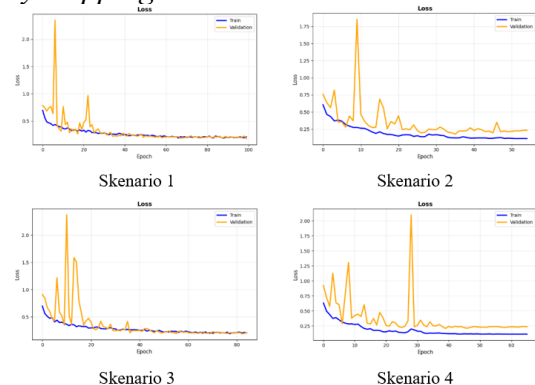
3.4. Analisis Kurva Training

Analisis kurva *training* dilakukan untuk evaluasi perilaku *konvergensi* model pada proses pembelajaran keempat skenario eksperimen. Kurva *training accuracy* dan *training loss* diamati agar mengetahui kestabilan model, kecepatan *konvergensi*, serta ada tidaknya indikasi *overfitting* pada setiap skenario. Hasil *visualisasi* kurva *training accuracy* dan *training loss* keempat skenario disajikan pada Gambar 3 dan Gambar 4.



Gambar 3. Kurva Training Accuracy Keempat Skenario

Skenario 1 dan Skenario 3 yang menggunakan augmentasi menunjukkan kurva *validation accuracy* yang bergerak tidak teratur pada *epoch* awal karena model menerima data yang dimodifikasi secara acak setiap *epoch*. Skenario 1 baru stabil di sekitar *epoch* ke-30 dan berjalan hingga *epoch* ke-100, sedangkan Skenario 3 berhenti di *epoch* ke-85. Sebaliknya, Skenario 2 dan Skenario 4 menunjukkan konvergensi yang jauh lebih cepat dan stabil, dengan Skenario 2 bahkan sudah melampaui baseline *Inception-V3* (96%) pada sekitar *epoch* ke-21 sebelum berhenti melalui mekanisme *early stopping*.

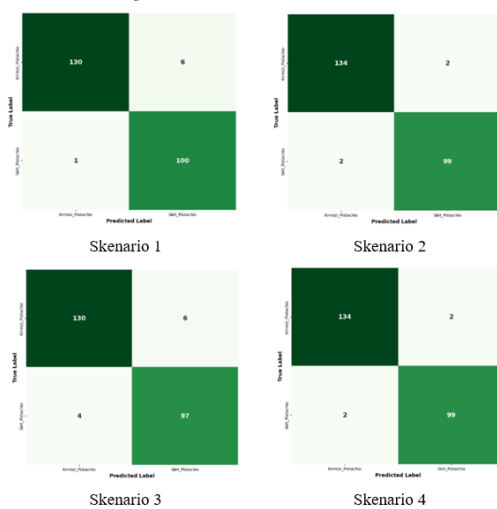


Gambar 4. Kurva Training Loss Keempat Skenario

Pada kurva *loss*, Skenario 1 dan Skenario 3 memperlihatkan lonjakan *validation loss* hingga di atas 1,5 pada *epoch* awal akibat ketidaksesuaian distribusi data *augmented* dengan data validasi asli. Skenario 2 dan Skenario 4 menunjukkan penurunan *loss* yang lebih mulus dengan *training loss* dan *validation loss* yang berdekatan pada akhir *training*, menandakan model berhasil belajar dengan baik tanpa *overfitting*.

3.5. Analisis Confusion Matrix

Analisis *confusion matrix* dilakukan untuk melihat pola kesalahan klasifikasi model pada masing-masing skenario, sebagaimana disajikan pada Gambar 5.



Gambar 5. Confusion Matrix Keempat Skenario

Skenario 2 dan Skenario 4 menunjukkan hasil terbaik dengan hanya 4 citra yang salah diklasifikasikan dari 237 citra *testing*, masing-masing 2 citra Kirmizi salah diprediksi sebagai Siirt dan 2 citra Siirt salah diprediksi sebagai Kirmizi. Skenario 1 memiliki 7 citra yang salah, sedangkan Skenario 3 memiliki kesalahan terbanyak dengan 10 citra yang salah. Pola kesalahan pada Skenario 3 mengkonfirmasi bahwa tanpa *class weighting* model cenderung lebih sering salah pada kelas Siirt yang datanya lebih sedikit [11].

3.6. Pengaruh Data Augmentation

Perbandingan Skenario 1 terhadap Skenario 2 dan Skenario 3 terhadap Skenario 4 menunjukkan data *augmentation* menurunkan *testing accuracy* pada penelitian ini. Skenario 1 menghasilkan 97,05% dibandingkan Skenario 2 sebesar 98,31%, sedangkan Skenario 3 menghasilkan 95,78% dibandingkan Skenario 4

sebesar 98,31%. Penurunan ini disebabkan karakteristik dataset Pistachio yang memiliki latar belakang sangat homogen, sehingga variasi augmentasi seperti rotasi dan pergeseran tidak merepresentasikan kondisi nyata dan justru memperkenalkan noise yang tidak relevan. Hal ini juga tercermin dari durasi *training* yang jauh lebih panjang pada skenario dengan augmentasi yaitu 92 menit untuk Skenario 1 dan 56 menit untuk Skenario 3, dibandingkan 23 menit untuk Skenario 2 dan 32 menit untuk Skenario 4. *ResNet101* dengan bobot *pre-trained ImageNet* yang kaya sudah memiliki kemampuan generalisasi yang cukup kuat sehingga augmentasi tambahan tidak diperlukan pada dataset yang relatif bersih ini [5].

Secara teoritis, fenomena penurunan akurasi akibat augmentasi ini dapat dijelaskan melalui konsep *domain shift*, yaitu kondisi di mana distribusi data hasil augmentasi pada *training set* bergeser dari distribusi data asli pada *testing set*. Pada dataset dengan variabilitas visual alami yang tinggi, *domain shift* ini justru membantu model belajar fitur yang lebih invarian. Namun pada dataset *Pistachio* yang memiliki latar belakang homogen dan orientasi objek yang relatif konsisten, augmentasi rotasi dan pergeseran menciptakan *domain shift* yang tidak merepresentasikan variabilitas nyata, sehingga model yang dioptimalkan pada distribusi *augmented* mengalami sedikit kesulitan generalisasi ketika dievaluasi pada distribusi asli yang lebih sempit dan konsisten.

Temuan ini konsisten dengan perspektif yang diajukan oleh Chazar dkk. (2025), yang menunjukkan bahwa augmentasi terbukti efektif meningkatkan akurasi EfficientNet dari 73,84% menjadi 82,56% pada dataset citra fundus retina yang memiliki variabilitas visual alami tinggi akibat perbedaan pencahayaan dan orientasi pengambilan gambar antar pasien. Keberhasilan augmentasi pada konteks tersebut dikaitkan dengan kemampuannya menghasilkan variasi data yang **merepresentasikan kondisi nyata** domain tersebut. Sebaliknya, pada dataset *Pistachio* yang memiliki latar belakang homogen dan orientasi objek yang sangat konsisten, augmentasi rotasi dan pergeseran yang diterapkan tidak merepresentasikan variabilitas nyata, sehingga justru menciptakan *domain shift* yang menurunkan kemampuan generalisasi model [16]

3.7. Pengaruh *Class Weight*

Penerapan *class weighting* terbukti lebih bermanfaat ketika dikombinasikan dengan augmentasi, terlihat dari pengecilan selisih *F1-score* Skenario 3 sebesar 1,20 poin menjadi 0,76 poin pada Skenario 1, disertai peningkatan *accuracy* dari 95,78% menjadi 97,05%. Namun pada kondisi tanpa augmentasi, Skenario 2 dan Skenario 4 menghasilkan performa identik sebesar 98,31%, mengindikasikan bahwa model *ResNet101* dengan bobot *pre-trained ImageNet* sudah mampu menangani ketidakseimbangan ringan (rasio 1,35:1) secara implisit tanpa pembobotan eksplisit [11].

Keunggulan generalisasi *ResNet101* dibandingkan *ResNet50* pada dataset ini dapat dikaitkan dengan mekanisme residual learning yang lebih dalam. Dengan 101 lapisan, *ResNet101* memiliki kapasitas representasi fitur hierarkis yang lebih kaya melalui lebih banyak *residual block*, sehingga mampu mempelajari fitur morfologi halus yang membedakan *Kirmizi* dan *Siirt* secara lebih efektif. Mekanisme *skip connection* pada setiap *residual block* memungkinkan gradien mengalir langsung ke *layer-layer* awal tanpa mengalami degradasi, sehingga proses optimisasi tetap stabil meskipun jaringan memiliki kedalaman yang signifikan. Kombinasi kedalaman jaringan yang lebih besar dengan bobot *pre-trained ImageNet* yang telah mempelajari representasi visual generik dalam skala besar inilah yang menjadikan *ResNet101* mampu mencapai generalisasi yang lebih baik dibandingkan *ResNet50* pada penelitian Lisda dkk. (2023), yang mengalami *overfitting* parah dengan *gap* mencapai 78%.

3.8. Perbandingan Penelitian Sebelumnya

Tabel 4. Perbandingan dengan Penelitian Sebelumnya

Model	Testing Accuracy	Penelitian
<i>ResNet50</i>	86%	Lisda dkk. (2023)
<i>Inception-V3</i>	96%	Lisda dkk. (2023)
<i>ResNet101</i> Skenario 3	95,78%	Penelitian ini
<i>ResNet101</i> Skenario 1	97,05%	Penelitian ini
<i>ResNet101</i> Skenario 2	98,31%	Penelitian ini

Model	Testing Accuracy	Penelitian
<i>ResNet101</i> Skenario 4	98,31%	Penelitian ini

Tiga dari empat skenario *ResNet101* berhasil melampaui *baseline Inception-V3* (96%) dan *ResNet50* (86%) dari Lisda dkk. (2023). Skenario 2 dan Skenario 4 mencapai akurasi tertinggi 98,31%, melampaui *baseline Inception-V3* sebesar 2,31 poin persentase dan *ResNet50* sebesar 12,31 poin persentase. Selain itu, seluruh skenario berhasil mengatasi *overfitting* parah yang dialami *ResNet50* pada penelitian sebelumnya dengan *gap training-validation accuracy* hingga 78%, di mana Skenario 2 dan Skenario 4 menghasilkan *gap* sebesar 0%. Meskipun demikian, penelitian ini memiliki keterbatasan berupa dataset yang hanya mencakup dua varietas dengan kondisi citra terkontrol dan latar belakang homogen, sehingga performa model pada kondisi pengambilan citra yang lebih bervariasi belum dapat dipastikan, dan *class weighting* terbukti lebih signifikan pada data dengan augmentasi dibanding data yang sudah bersih dan seimbang secara relatif.

4. PENUTUP

4.1. Kesimpulan

Penelitian ini membuktikan bahwa *ResNet101* dengan *class weighting* tanpa *data augmentation* merupakan konfigurasi paling optimal untuk klasifikasi kacang *Pistachio* varietas *Kirmizi* dan *Siirt*, melampaui *baseline Inception-V3* dan *ResNet50* dari penelitian sebelumnya sekaligus mengatasi permasalahan *overfitting* parah yang dialami *ResNet50*. Temuan utama penelitian ini menunjukkan bahwa efektivitas *data augmentation* bersifat kontekstual terhadap karakteristik dataset: pada citra dengan latar belakang homogen dan variabilitas *visual* rendah seperti *Pistachio*, augmentasi rotasi dan pergeseran justru menurunkan akurasi karena menciptakan *domain shift* yang tidak relevan, meskipun tetap efektif menekan *overfitting*. Sebaliknya, *class weighting* terbukti memberikan kontribusi yang lebih signifikan ketika dikombinasikan dengan augmentasi dibandingkan pada kondisi data yang sudah bersih dan relatif seimbang. Implikasi penelitian ini menegaskan pentingnya evaluasi karakteristik dataset sebelum

menerapkan teknik augmentasi secara generik dalam tugas klasifikasi citra pertanian berbasis *deep learning*.

4.2. Saran

Berdasarkan keterbatasan yang ditemukan dalam penelitian ini, penelitian lanjutan disarankan untuk mengeksplorasi arsitektur yang lebih modern seperti *EfficientNet*, *MobileNetV3*, atau *Vision Transformer* dengan kondisi eksperimen yang identik agar perbandingan yang dihasilkan lebih komprehensif. Selain itu, mengingat *data augmentation* terbukti tidak efektif pada dataset dengan latar belakang homogen, penelitian lanjutan dapat mengeksplorasi *synthetic data generation* menggunakan *Generative Adversarial Network (GAN)* sebagai alternatif yang lebih sesuai. Perluasan dataset dengan menambahkan citra yang diambil dalam kondisi pencahayaan berbeda, latar belakang bervariasi, dan berbagai sudut pengambilan gambar juga sangat disarankan untuk meningkatkan kemampuan generalisasi model pada kondisi nyata di lapangan. Selain itu, eksplorasi arsitektur *Vision Transformer (ViT)* dengan kombinasi *class weighting* pada penelitian lanjutan sangat direkomendasikan, mengingat Salsabila dkk. (2024) membuktikan bahwa *ViT* dengan bobot mampu meningkatkan *recall* kelas minoritas secara drastis dari 66,7% menjadi 93,3% pada dataset citra medis yang *imbalanced*, sebuah karakteristik yang juga ditemui pada dataset *Pistachio* dalam penelitian ini [17]. Secara khusus, *EfficientNet-B3* layak diprioritaskan dalam eksplorasi tersebut, mengingat Siadari & Hariyani (2024) membuktikan bahwa *EfficientNet-B3* mengungguli *ResNet-50* sebesar 9,09 poin persentase dalam akurasi klasifikasi citra medis pada kondisi pelatihan yang identik, dengan keunggulan yang dikaitkan pada teknik *compound scaling* yang mengoptimalkan kedalaman, lebar, dan resolusi jaringan secara bersamaan sehingga lebih efisien secara parameter dibanding arsitektur *ResNet* pada dataset berskala terbatas [18].

5. DAFTAR PUSTAKA

- [1] L. Lisda, K. Kusrini, and D. Ariatmanto, "Classification of Pistachio Nut Using Convolutional Neural Network," *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 8, no. 1, pp. 71–77, Jan. 2023, doi: 10.25139/inform.v8i1.5685.
- [2] E. AVUÇLU, "Classification of Pistachio Images Using VGG16 and VGG19 Deep Learning Models," *International Scientific and Vocational Studies Journal*, vol. 7, no. 2, pp. 79–86, Dec. 2023, doi: 10.47897/bilmes.1328313.
- [3] D. Singh *et al.*, "Classification and Analysis of Pistachio Species with Pre-Trained Deep Learning Models," *Electronics (Switzerland)*, vol. 11, no. 7, Apr. 2022, doi: 10.3390/electronics11070981.
- [4] K. Azmi and S. Defit, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," *Jurnal Unitek*, vol. 16, no. 1, p. 2023, 2023, doi: 10.52072/unitek.v16i1.504.
- [5] R. Faturrahman, Y. S. Hariyani, and S. Hadiyoso, "Klasifikasi Jajanan Tradisional Indonesia berbasis Deep Learning dan Metode Transfer Learning," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 11, no. 4, p. 945, Oct. 2023, doi: 10.26760/elkomika.v11i4.945.
- [6] A. Vierisyah and R. Maulana Fajri, "KLASIFIKASI KANKER PARU PARU MENGGUNAKAN CNN DENGAN 5 ARSITEKTUR," *Jini Global : Journal of Intelligent Networks and IoT Global*, 2023, doi: doi.org/10.36982/jinig.v1i2.3643.
- [7] E. Goceri and A. A. Karakas, "COMPARATIVE EVALUATIONS OF CNN BASED NETWORKS FOR SKIN LESION CLASSIFICATION," *International Conferences Computer Graphics, Visualization, Computer Vision and Image Processing 2020*, 2020.
- [8] U. N. Oktaviana, R. Hendrawan, A. D. K. Annas, and G. W. Wicaksono, "Classification of Rice Diseases based on Leaf Image Using Resnet101 Trained Model," *Jurnal RESTI*, vol. 5, no. 6, pp.

- 1216–1222, Dec. 2021, doi: 10.29207/resti.v5i6.3607.
- [9] E. Avuçlu, “Classification Of Pistachio Images With The ResNet Deep Learning Model,” *Selcuk Journal of Agricultural and Food Sciences*, Aug. 2023, doi: 10.15316/sjafs.2023.029.
- [10] I. A. Özkan, M. Köklü, and R. Saraçoğlu, “Classification of pistachio species using improved k-NN classifier,” *Progress in Nutrition*, vol. 23, no. 2, Feb. 2021, doi: 10.23751/pn.v23i2.9686.
- [11] H. Akbar and W. K. Sanjaya, “Kajian Performa Metode Class Weight Random Forest pada Klasifikasi Imbalance Data Kelas Curah Hujan,” *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 3, no. 1, Dec. 2023, doi: 10.20885/snati.v3i1.30.
- [12] D. Mualfah, R. Ardiansyah, and R. Gunawan, “Classification of DDoS attacks using the random forest method and class weight technique on the CICDDoS2019 dataset,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 6, no. 3, pp. 530–535, Dec. 2025, doi: 10.37859/coscitech.v6i3.10731.
- [13] B. Marsheila Islami and A. Nugroho, “Analisis Algoritma KNN dan Penerapan SMOTE dalam Deteksi Dini Kanker Paru-Paru,” *Jurnal Qua Teknika*, vol. 15, no. 2, p. 15, 2025, doi: 10.35457/quateknika.v15i02.4603.
- [14] S. Sucipto, D. Dwi Prasetya, and T. Widiyaningtyas, “Educational Data Mining: Multiple Choice Question Classification in Vocational School,” *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 23, no. 2, pp. 379–388, Mar. 2024, doi: 10.30812/matrik.v23i2.3499.
- [15] M. Shelin Sahira, F. Dessela Putri, A. Ristyawan, and E. Daniati, “Penggunaan Algoritma K-Nearest Neighbor Untuk Klasifikasi Data Diabetes Pada Wanita,” *Prosiding Seminar Nasional Inovasi Teknologi*, vol. 8, pp. 2549–7952, 2024, doi: doi.org/10.29407/inotek.v8i1.4981.
- [16] C. Chazar, M. A. Adli, J. Pardede, and M. Ichwan, “Pendekatan Augmentasi Citra Fundus pada Model EfficientNet untuk Klasifikasi Tingkat Keparahan Retinopati Diabetik dengan Dataset Tidak Seimbang,” *Journal MIND Journal | ISSN*, vol. 10, no. 2, pp. 180–194, 2025, doi: 10.26760/mindjournal.v10i2.180-194.
- [17] J. Khatib Sulaiman, A. Nur Salsabila, M. Liebenlito, D. Urfina Zulkifli, and U. Syarif Hidayatullah Jakarta, “Perbandingan Deteksi Alzheimer: ViT, CNN dan ViT dengan Bobot pada Citra Medis,” *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 1, pp. 2024–1401, 2024.
- [18] T. S. Siadari and Y. S. Hariyani, “Klasifikasi Penyakit Periodontal Pada Citra Gigi Berwarna Menggunakan Arsitektur Berbasis Convolutional Neural Network,” *JETT: Jurnal Elektro dan Telekomunikasi Terapan*, vol. 11, pp. 13–19, 2024, doi: 10.25124/jett.v11i2.8090.