

Deep Learning-Based Rice Grain Classification with Class Imbalance Handling Using Weighted Sampling and Data Augmentation

Muhammad Aulia Anhar^{1*}, Nova Rijati^{2**}

* Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
111202214848@mhs.dinus.ac.id¹, nova.rijati@dsn.dinus.ac.id²

Article Info

Article history:

Received 2026-05-05

Revised 2026-05-11

Accepted 2026-05-25

Keyword:

Rice Grain Classification,
Class Imbalance,
Deep Learning,
Fine-Grained Classification.

ABSTRACT

Rice grain quality assessment is important for maintaining product consistency and quality standards in the food industry. Manual inspection is still widely used in practice, but it is often subjective and time-consuming especially in fine-grained classification problems where the visual differences between classes are relatively subtle. Another common challenge in rice grain datasets is class imbalance where the number of normal samples is much larger than the number of defective classes. In this paper, we investigate various imbalance handling strategies for deep learning-based rice grain classification using the GrainSet dataset, which contains 30,962 rice grain images divided into eight quality classes with a highly imbalanced class distribution. The proposed approach combines weighted sampling, data augmentation, and hyperparameter optimization to address class imbalance, while Focal Loss and Class-Balanced Loss are used as comparison methods for performance evaluation. Three deep learning architectures, namely MobileNetV3, ResNet50, and ViT-Small, were evaluated using Accuracy, Macro F1-score, and Weighted F1-score metrics. Experimental results show that the proposed approach achieved the most consistent overall performance. Among the tested models, ResNet50 achieved the best result with 98.87% accuracy and a Macro F1-Score 0.9799. The results also show that convolution-based architectures are more stable than transformer-based models for texture-oriented datasets with limited training data. Furthermore, the combination of sampling-based balancing, augmentation and suitable hyperparameter configuration contributed to better recognition performance on minority classes.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Beras merupakan komoditas pangan utama di Indonesia yang memiliki peran strategis dalam menjaga ketahanan pangan nasional. Kualitas butiran beras menjadi faktor penting yang memengaruhi nilai ekonomi, mutu distribusi, serta tingkat penerimaan konsumen. Dalam praktiknya, proses klasifikasi kualitas beras masih banyak dilakukan secara manual melalui pengamatan visual oleh tenaga ahli. Metode ini cenderung subjektif, memerlukan waktu yang lama, serta berpotensi menghasilkan penilaian yang tidak konsisten antar pengamat [2], [3], [5]. Kondisi tersebut menjadi kendala terutama dalam skala industri, sehingga diperlukan sistem klasifikasi otomatis berbasis citra yang

mampu memberikan hasil yang lebih cepat, objektif, dan konsisten.

Perkembangan *computer vision* dan *deep learning* telah mendorong berbagai penelitian pada klasifikasi citra biji-bijian, khususnya klasifikasi kualitas beras berbasis citra dengan perbedaan tekstur yang sangat halus (*fine-grained classification*). Pendekatan berbasis Convolutional Neural Network (CNN) banyak digunakan karena kemampuannya dalam mengekstraksi fitur visual secara otomatis, seperti bentuk, warna, dan tekstur [6], [8], [9]. Selain itu, beberapa penelitian juga telah mengeksplorasi penggunaan machine learning dan *deep learning* untuk klasifikasi kualitas maupun varietas beras [17]–[19]. Hasil penelitian tersebut menunjukkan bahwa arsitektur *deep learning* modern mampu

menghasilkan performa klasifikasi yang tinggi. Namun, pada kasus *fine-grained classification*, model tetap menghadapi tantangan karena perbedaan visual antar kelas cenderung kecil dan sulit dibedakan secara konsisten [7].

Dalam klasifikasi kualitas beras, permasalahan utama tidak hanya terletak pada kompleksitas fitur visual, tetapi juga pada distribusi data yang tidak seimbang. Pada kondisi nyata, jumlah data untuk kelas normal umumnya jauh lebih dominan dibandingkan kelas cacat atau minoritas. Kondisi ini juga terlihat pada dataset GrainSet yang digunakan dalam penelitian ini, yang menyediakan lebih dari 350.000 citra kernel biji-bijian, termasuk sekitar 31.000 citra beras dengan anotasi kualitas yang detail [1]. Ketidakseimbangan distribusi data dapat menyebabkan model cenderung bias terhadap kelas mayoritas sehingga kemampuan model dalam mengenali kelas minoritas menjadi kurang optimal [16], [23]. Pada kasus *fine-grained classification*, kondisi tersebut menjadi lebih kompleks karena perbedaan antar kelas relatif kecil dan sebagian besar karakteristik pembeda hanya terletak pada detail tekstur tertentu.

Penelitian yang memperkenalkan dataset GrainSet menunjukkan bahwa pendekatan berbasis deep learning mampu menghasilkan performa klasifikasi yang tinggi dengan nilai rata-rata F1-score mencapai 94,1% pada klasifikasi kualitas beras [1]. Meskipun demikian, penelitian tersebut masih berfokus pada validasi dataset dan penggunaan model *baseline*, sehingga belum membahas secara mendalam pengaruh strategi penanganan *class imbalance* maupun perbedaan karakteristik arsitektur model terhadap performa klasifikasi, khususnya pada kelas minoritas [16], [23]. Kondisi tersebut menunjukkan bahwa klasifikasi kualitas beras tidak hanya dipengaruhi oleh kemampuan model dalam mengekstraksi fitur visual, tetapi juga oleh kemampuan model dalam menangani distribusi data yang tidak seimbang pada kasus klasifikasi berbasis tekstur.

Di sisi lain, setiap arsitektur *deep learning* memiliki karakteristik yang berbeda dalam menangani klasifikasi citra. Model CNN seperti ResNet dikenal mampu mengekstraksi fitur kompleks melalui jaringan yang dalam dengan memanfaatkan residual learning [12], sedangkan MobileNetV3 dirancang sebagai model ringan yang lebih efisien untuk kebutuhan komputasi terbatas [13]. Sementara itu, Vision Transformer (ViT) menggunakan mekanisme *self-attention* untuk menangkap hubungan global antar bagian citra [14]. Meskipun demikian, performa Vision Transformer umumnya sangat bergantung pada ketersediaan data pelatihan dalam jumlah besar [10], [14]. Perbedaan karakteristik tersebut menunjukkan bahwa pemilihan arsitektur model perlu disesuaikan dengan kondisi dataset, terutama pada kasus data terbatas, berbasis tekstur, dan memiliki distribusi yang tidak seimbang.

Berbagai pendekatan telah digunakan untuk mengatasi permasalahan ketidakseimbangan data, seperti *data augmentation*, *sampling strategy*, dan optimasi *hyperparameter* [9], [16]. Metode seperti *Focal Loss* dan *Class-Balanced Loss* juga telah banyak digunakan untuk

meningkatkan kemampuan model dalam mengenali kelas minoritas [21], [22]. Namun, sebagian besar penelitian sebelumnya masih menerapkan pendekatan tersebut secara terpisah dan belum memberikan analisis yang komprehensif mengenai kontribusi masing-masing metode terhadap peningkatan performa klasifikasi kualitas beras [21]–[23]. Selain itu, kajian yang membandingkan pendekatan berbasis *sampling* dan *loss balancing* pada arsitektur CNN maupun Vision Transformer dalam kondisi *fine-grained classification* masih relatif terbatas.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan klasifikasi kualitas butiran beras berbasis *deep learning* dengan mengombinasikan *weighted sampling*, *data augmentation*, dan optimasi *hyperparameter*. Pendekatan ini dipilih karena ketiga teknik tersebut saling melengkapi dalam mengatasi distribusi data yang tidak seimbang sekaligus meningkatkan stabilitas proses pelatihan model. Selain itu, penelitian ini juga membandingkan pendekatan tersebut dengan metode *Focal Loss* dan *Class-Balanced Loss* menggunakan konfigurasi augmentasi dan hyperparameter yang sama untuk memastikan perbandingan performa dilakukan secara adil. Tiga arsitektur *deep learning*, yaitu MobileNetV3, ResNet50, dan ViT-Small, digunakan untuk merepresentasikan model ringan, model CNN dalam, dan model berbasis transformer.

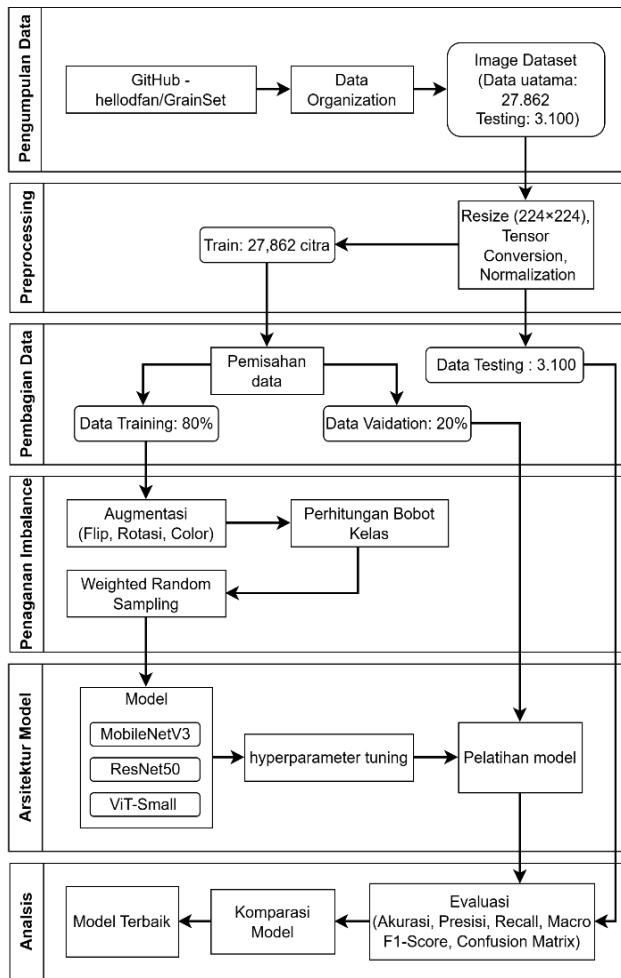
Berbeda dengan penelitian sebelumnya yang umumnya hanya berfokus pada penggunaan satu metode penanganan *class imbalance* [21], [22], penelitian ini mengevaluasi kombinasi *weighted sampling*, *data augmentation*, dan optimasi hyperparameter pada klasifikasi kualitas beras berbasis *deep learning*. Selain itu, penelitian ini juga membandingkan pendekatan usulan dengan *Focal Loss* dan *Class-Balanced Loss* sebagai metode pembanding untuk melihat efektivitas masing-masing metode pada distribusi data yang tidak seimbang. Evaluasi dilakukan tidak hanya berdasarkan nilai akurasi, tetapi juga menggunakan Macro F1-score, confusion matrix, dan Grad-CAM untuk menganalisis kemampuan model dalam mengenali kelas minoritas pada kasus *fine-grained classification* [4], [16].

II. METODE

Penelitian ini mengusulkan pendekatan klasifikasi kualitas butiran beras berbasis *deep learning* dengan mengombinasikan tiga teknik utama, yaitu *data augmentation*, *weighted sampling*, dan optimasi *hyperparameter*. Pendekatan ini dirancang untuk mengatasi permasalahan distribusi data yang tidak seimbang, meningkatkan keragaman fitur pelatihan, serta memperoleh konfigurasi model yang lebih stabil dan optimal pada proses klasifikasi kualitas beras berbasis citra.

Alur penelitian dimulai dari pengolahan dataset dan pembagian data menjadi *training*, *validation*, dan *testing*. Pada data *training* diterapkan *data augmentation* untuk memperkaya variasi data serta *weighted sampling* untuk mengurangi bias terhadap kelas mayoritas. Selanjutnya, model dilatih menggunakan tiga arsitektur berbeda, yaitu

MobileNetV3, ResNet50, dan Vision Transformer (ViT-Small), kemudian dilakukan optimasi *hyperparameter* untuk memperoleh konfigurasi terbaik. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score, serta didukung oleh analisis tambahan untuk melihat kontribusi masing-masing metode. Diagram alur metode penelitian ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Sumber data yang digunakan dalam penelitian ini adalah dataset publik GrainSet, yaitu kumpulan citra biji-bijian teranotasi yang dikembangkan untuk keperluan inspeksi kualitas visual dan dipublikasikan melalui Scientific Data [1]. Dataset ini dipilih karena memiliki anotasi berbasis ahli serta mencakup variasi morfologi biji yang representatif, sehingga relevan untuk digunakan dalam analisis klasifikasi kualitas beras berbasis citra.

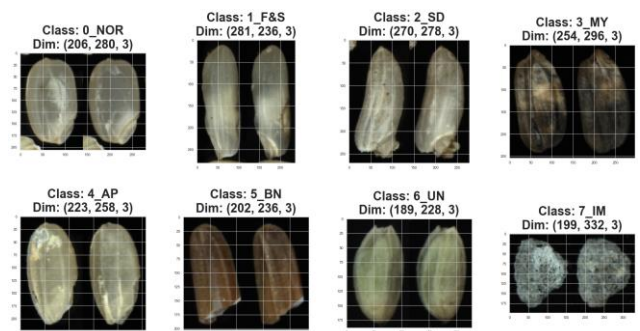
Secara keseluruhan, penelitian ini menggunakan sebanyak 30.962 citra, yang terdiri dari 27.862 citra sebagai data latih (train set) dan 3.100 citra sebagai data uji (test set). Distribusi data latih berdasarkan kelas kualitas beras ditunjukkan pada Tabel I. Dataset terdiri dari delapan kelas dengan distribusi

yang tidak merata, di mana kelas normal (0_NOR) mendominasi sebesar 64,61%, sementara kelas lainnya hanya berada pada kisaran 4–6%. Kondisi ini mencerminkan distribusi data di dunia nyata, namun sekaligus menjadi tantangan dalam proses pelatihan model karena berpotensi menyebabkan bias terhadap kelas mayoritas [16]. Selain pendekatan *weighted sampling*, penelitian ini juga mengevaluasi metode *loss balancing* menggunakan *Focal Loss* dan *Class-Balanced Loss* sebagai pembanding dalam penanganan *class imbalance*.

TABEL I
SAMPEL DATASET

| Kode Kelas | Keterangan Kualitas | Jumlah Citra | Persentase (%) |
|------------|---------------------|--------------|----------------|
| 0_NOR | Normal | 18.000 | 64,61% |
| 1_F&S | Foreign & Seed | 1.336 | 4,80% |
| 2_SD | Slightly Damaged | 1.350 | 4,85% |
| 3_MY | Chalky / Milky | 1.350 | 4,85% |
| 4_AP | Abnormal / Paddy | 1.350 | 4,85% |
| 5_BN | Broken / Neck | 1.341 | 4,81% |
| 6_UN | Unmilled / Husks | 1.335 | 4,79% |
| 7_IM | Immature / Green | 1.800 | 6,44% |
| Total | | 27.862 | 100% |

Ketidakeimbangan distribusi tersebut dapat memengaruhi kemampuan model dalam mengenali kelas minoritas, sehingga performa klasifikasi menjadi tidak merata antar kelas. Rasio antara kelas mayoritas dan kelas minoritas bahkan mencapai lebih dari 13:1, yang menunjukkan tingkat ketidakeimbangan yang cukup tinggi. Kondisi ini menunjukkan bahwa dataset memiliki tingkat *class imbalance* yang cukup tinggi sehingga berpotensi menyebabkan model bias terhadap kelas mayoritas selama proses pelatihan. Oleh karena itu, karakteristik dataset ini menjadi dasar dalam penerapan strategi penanganan *class imbalance* pada tahap pelatihan.



Gambar 2. Sampel Dataset

Berdasarkan Gambar 2, terlihat bahwa perbedaan visual antar kelas relatif kecil, sehingga klasifikasi termasuk dalam kategori *fine-grained classification*. Kondisi ini menuntut model untuk mampu menangkap perbedaan fitur yang sangat halus agar dapat membedakan antar kelas secara akurat.

B. Prapemrosesan Data

Tahap prapemrosesan dilakukan untuk menyiapkan data citra agar memiliki format yang seragam sebelum digunakan dalam proses pelatihan model. Berdasarkan alur penelitian pada Gambar 1, proses ini mencakup perubahan ukuran citra (resizing), konversi ke dalam bentuk tensor, serta normalisasi nilai piksel. Seluruh citra diubah ukurannya menjadi 224×224 piksel untuk menyesuaikan dengan kebutuhan input arsitektur MobileNetV3, ResNet50, dan Vision Transformer (ViT-Small). Selanjutnya, citra dikonversi ke dalam tensor dan dinormalisasi menggunakan nilai mean dan standard deviation dari dataset ImageNet, yaitu $[0,485, 0,456, 0,406]$ dan $[0,229, 0,224, 0,225]$, guna menjaga kestabilan distribusi data selama proses pelatihan.

Proses prapemrosesan diterapkan secara konsisten pada seluruh data, baik data latih maupun data uji. Tahap ini mencakup perubahan ukuran citra, konversi ke dalam bentuk tensor, serta normalisasi tanpa melakukan transformasi tambahan yang dapat mengubah karakteristik asli data. Transformasi augmentasi tidak diterapkan pada tahap prapemrosesan, melainkan secara khusus digunakan pada data *training* dalam tahap penanganan *class imbalance*. Pemisahan ini dilakukan agar proses evaluasi tetap menggunakan data yang merepresentasikan kondisi asli, sehingga hasil pengujian dapat mencerminkan kemampuan generalisasi model secara lebih objektif.

C. Pembagian Data

Pembagian data dilakukan dalam dua tahap. Pertama, dataset GrainSet telah dipisahkan menjadi data latih sebanyak 27.862 citra dan data uji sebanyak 3.100 citra. Kedua, data latih dibagi kembali menjadi data *training* sebesar 80% dan *validation* sebesar 20%.

TABEL II
SAMPEL DATASET

| Jenis Data | Jumlah Citra | Keterangan |
|------------|--------------|--|
| Training | 22.290 | Digunakan untuk pelatihan model |
| Validation | 5.572 | Digunakan untuk pemantauan dan pemilihan model |
| Testing | 3.100 | Digunakan untuk evaluasi akhir |
| Total | 30.962 | |

Pembagian ini bertujuan untuk menjaga agar proses pelatihan dan evaluasi dilakukan secara terpisah. Dengan demikian, performa model yang diperoleh pada data uji dapat lebih mencerminkan kemampuan generalisasi model terhadap data yang belum pernah digunakan selama proses pelatihan.

D. Penanganan Imbalance

Ketidakseimbangan distribusi kelas pada data latih menjadi salah satu permasalahan utama dalam penelitian ini, di mana jumlah citra pada kelas mayoritas jauh lebih besar dibandingkan kelas minoritas. Kondisi ini dapat menyebabkan model cenderung mempelajari pola dari kelas mayoritas secara dominan, sehingga performa klasifikasi pada kelas minoritas menjadi kurang optimal [16]. Berbagai pendekatan telah dikembangkan untuk mengatasi permasalahan ini, seperti teknik resampling berbasis over-

sampling dan under-sampling untuk menyeimbangkan distribusi data [20].

Pendekatan yang digunakan dalam penelitian ini adalah weighted random sampling, di mana bobot setiap kelas ditentukan berdasarkan kebalikan dari jumlah sampel pada kelas tersebut, yang dirumuskan sebagai.

$$w_i = \frac{1}{n_i} \quad (1)$$

di mana w_i merupakan bobot kelas ke- i dan n_i adalah jumlah data pada kelas tersebut. Selanjutnya, setiap sampel x_j dengan label y_j diberikan bobot

$$w(x_j) = \frac{1}{n_{y_j}} \quad (2)$$

Bobot tersebut digunakan untuk menentukan probabilitas pemilihan setiap sampel dalam proses pelatihan, yang dinyatakan sebagai

$$P(x_j) = \frac{w(x_j)}{\sum_{k=1}^N w(x_k)} \quad (3)$$

dengan N adalah jumlah total sampel pada data latih. Melalui mekanisme ini, sampel dari kelas minoritas memiliki peluang lebih besar untuk terpilih, sehingga distribusi data yang diterima model selama pelatihan menjadi lebih seimbang.

Selain metode usulan berbasis weighted sampling, penelitian ini juga mengevaluasi *Focal Loss* sebagai metode pembandingan (*comparison method*) dalam menangani *class imbalance*. Metode ini dirancang untuk mengurangi dominasi sampel yang mudah diklasifikasikan dengan memberikan fokus lebih besar pada sampel yang sulit diprediksi. Dengan demikian, model diharapkan mampu meningkatkan performa pada kelas minoritas dan mengurangi bias terhadap kelas mayoritas [21].

$$FL(p_t) = -(1 - p_t)\gamma \log(p_t) \quad (4)$$

Pada persamaan tersebut, p_t menunjukkan probabilitas prediksi terhadap kelas yang benar, sedangkan γ merupakan focusing parameter yang mengatur tingkat penekanan terhadap sampel yang sulit diklasifikasikan.

Selain itu, *Class-Balanced Loss* juga digunakan sebagai metode pembandingan untuk mengevaluasi efektivitas pendekatan usulan terhadap strategi *loss balancing* lainnya. Metode ini menghitung bobot kelas berdasarkan effective number of samples sehingga kelas dengan jumlah data sedikit memperoleh kontribusi loss yang lebih besar selama proses pelatihan [22].

$$CB(p, y) = \frac{1 - \beta}{1 - \beta^n y} L(p, y) \quad (5)$$

Pada persamaan tersebut, n_y merupakan jumlah sampel pada kelas y , sedangkan β adalah hyperparameter yang digunakan untuk mengontrol pengaruh effective number of samples terhadap pembobotan loss.

Selain itu, digunakan teknik *data augmentation* sebagai pendekatan pendukung untuk meningkatkan variasi data latih. Transformasi yang diterapkan meliputi horizontal flip, vertical flip, rotation, serta penyesuaian tingkat brightness dan contrast. Konfigurasi transformasi augmentasi yang digunakan dalam penelitian ini ditunjukkan pada Tabel II.

TABEL III
KONFIGURASI AUGMENTASI

| Teknik | Parameter |
|-----------------|----------------|
| Horizontal Flip | p=0.5 |
| Vertical Flip | p=0.5 |
| Rotation | $\pm 15^\circ$ |
| Brightness | $\pm 20\%$ |
| Contrast | $\pm 20\%$ |

Kombinasi *weighted sampling* dan *data augmentation* dipilih karena kedua metode tersebut menangani aspek yang berbeda, yaitu distribusi data dan keragaman fitur. Dengan demikian, model diharapkan mampu mempelajari representasi fitur secara lebih seimbang dan meningkatkan kemampuan klasifikasi secara lebih merata pada seluruh kelas, termasuk kelas minoritas.

Seluruh eksperimen pada *weighted sampling*, *Focal Loss*, dan *Class-Balanced Loss* menggunakan konfigurasi augmentasi dan hyperparameter yang sama untuk memastikan proses perbandingan performa dilakukan secara adil.

E. Arsitektur Model

Proses klasifikasi dalam penelitian ini menggunakan tiga arsitektur deep learning, yaitu MobileNetV3, ResNet50, dan Vision Transformer (ViT-Small), untuk membandingkan kemampuan model dalam menangani klasifikasi citra beras dengan karakteristik fine-grained dan distribusi data yang tidak seimbang.

MobileNetV3 digunakan sebagai representasi model ringan yang efisien untuk komputasi terbatas [13]. Sementara itu, ResNet50 dipilih karena kemampuannya dalam mengekstraksi fitur yang lebih kompleks melalui mekanisme residual learning [12]. Berbeda dengan kedua model tersebut, ViT-Small memanfaatkan mekanisme *self-attention* untuk menangkap hubungan global antar bagian citra [14]. ViT-Small dipilih untuk mengevaluasi kemampuan arsitektur transformer pada kasus *fine-grained classification* dengan data terbatas serta membandingkan mekanisme pembelajaran berbasis global attention dengan pendekatan CNN konvensional. Namun, berbeda dengan CNN yang memiliki inductive bias terhadap fitur lokal, model berbasis transformer umumnya membutuhkan jumlah data pelatihan yang lebih besar agar mekanisme *self-attention* dapat bekerja secara optimal [10], [11], [14].

Ketiga model menggunakan bobot awal pretrained dari ImageNet untuk membantu proses *transfer learning*, mempercepat konvergensi pelatihan, serta meningkatkan kemampuan generalisasi model pada dataset dengan jumlah data terbatas [15].

Proses pelatihan dilakukan selama 35 epoch dengan batch size 64. Optimizer yang digunakan adalah AdamW karena mampu memberikan pembaruan bobot yang stabil serta membantu mengurangi overfitting melalui *weight decay*. Fungsi loss yang digunakan adalah CrossEntropyLoss dengan *label smoothing* sebesar 0,1 untuk mengurangi overconfidence model serta membantu proses klasifikasi pada kelas dengan kemiripan visual yang tinggi [16].

Nilai *learning rate* dan *weight decay* disesuaikan untuk setiap model. MobileNetV3 menggunakan *learning rate* 1×10^{-3} karena arsitekturnya relatif ringan. Pada ResNet50, digunakan *learning rate* 5×10^{-4} untuk menjaga stabilitas pelatihan pada jaringan yang lebih dalam. Adapun ViT-Small menggunakan *learning rate* 1×10^{-4} , mengingat model berbasis transformer cenderung lebih sensitif terhadap perubahan parameter [10], [14].

Untuk menjaga stabilitas pelatihan, digunakan *scheduler ReduceLRonPlateau* yang secara adaptif menurunkan *learning rate* ketika nilai *validation loss* tidak mengalami perbaikan.

TABEL IV
KONFIGURASI PELATIHAN MODEL

| Parameter | MobileNetV3 | ResNet50 | ViT-Small |
|---------------|--------------------------------------|--------------------------------------|--------------------------------------|
| Epoch | 35 | 35 | 35 |
| Batch Size | 64 | 64 | 64 |
| Optimizer | AdamW | AdamW | AdamW |
| Learning Rate | 1×10^{-3} | 5×10^{-4} | 1×10^{-4} |
| Weight Decay | 0,01 | 0,02 | 0,05 |
| Loss Function | CrossEntropy + Label Smoothing (0,1) | CrossEntropy + Label Smoothing (0,1) | CrossEntropy + Label Smoothing (0,1) |
| Scheduler | ReduceLRonPlateau | ReduceLRonPlateau | ReduceLRonPlateau |
| Pretrained | ImageNet | ImageNet | ImageNet |

F. Evaluasi Model

Evaluasi performa model dilakukan menggunakan *confusion matrix* untuk menghitung beberapa metrik

klasifikasi, yaitu accuracy, precision, recall, dan F1-score. Penggunaan beberapa metrik ini bertujuan untuk memberikan

gambaran performa model secara lebih menyeluruh, terutama pada kondisi data yang tidak seimbang.

Dalam penelitian ini, fokus utama evaluasi adalah *Macro F1-score*. Metrik ini dipilih karena menghitung rata-rata performa setiap kelas secara setara tanpa mempertimbangkan jumlah *sampel*, sehingga lebih representatif pada dataset yang memiliki distribusi kelas tidak seimbang [4], [16]. Secara matematis, *Macro F1-score* dirumuskan sebagai:

$$\text{Macro F1} = \frac{1}{K} \sum_{i=1}^K F1_i \quad (6)$$

di mana K merupakan jumlah kelas dan $F1_i$ adalah nilai F1-score pada kelas ke- i .

Sebagai pembandingan, digunakan *Weighted F1-score* yang mempertimbangkan proporsi jumlah data pada setiap kelas. Metrik ini memberikan bobot yang lebih besar pada kelas dengan jumlah sampel lebih banyak, sehingga mencerminkan performa model berdasarkan distribusi data aktual. *Weighted F1-score* dirumuskan sebagai:

$$\text{Weighted F1} = \sum_{i=1}^k w_i \cdot F1_i \quad (7)$$

di mana w_i merupakan bobot kelas ke- i yang dihitung berdasarkan proporsi jumlah sampel terhadap total data.

Selain itu, *confusion matrix* digunakan untuk menganalisis distribusi kesalahan prediksi pada setiap kelas. Melalui analisis ini, dapat diketahui pola kesalahan model serta kelas-kelas yang sulit dibedakan.

III. HASIL DAN PEMBAHASAN

A. Hasil Eksperimen

Eksperimen dilakukan untuk mengevaluasi pengaruh strategi penanganan *class imbalance* terhadap performa klasifikasi kualitas beras. Evaluasi dilakukan dengan membandingkan dua skenario utama, yaitu skenario *baseline* tanpa penanganan khusus dan skenario *proposed* yang mengombinasikan *weighted sampling*, *data augmentation*, serta optimasi hyperparameter. Pada tahap awal, model dilatih menggunakan distribusi data asli untuk memperoleh performa dasar sebagai acuan evaluasi.

Hasil evaluasi pada skenario *baseline* ditunjukkan pada Tabel V. Secara umum, seluruh model menghasilkan nilai akurasi di atas 97%, yang menunjukkan bahwa ketiga arsitektur mampu mempelajari pola dasar pada citra beras dengan cukup baik. Namun, nilai Macro F1-score relatif lebih rendah dibandingkan akurasi, yang mengindikasikan bahwa performa model belum merata pada seluruh kelas, terutama pada kelas minoritas dengan jumlah data yang lebih sedikit.

TABEL V
HASIL EVALUASI MODEL PADA SKENARIO BASELINE.

| Model | Accuracy (%) | Macro F1 | Weighted F1 | Time (Min) |
|-------------|--------------|----------|-------------|------------|
| MobileNetV3 | 98.10 | 0.9658 | 0.9813 | 34.82 |
| ResNet50 | 97.42 | 0.9580 | 0.9749 | 78.64 |
| ViT | 98.71 | 0.9765 | 0.9872 | 55.65 |

Pada skenario *baseline*, ViT-Small memperoleh nilai akurasi tertinggi sebesar 98.71% dengan Macro F1-score sebesar 0.9765. Hasil ini menunjukkan bahwa mekanisme *self-attention* pada Vision Transformer mampu menangkap hubungan global antar bagian citra secara efektif [14]. Meskipun demikian, selisih performa antar model relatif kecil, yang menunjukkan bahwa seluruh model masih menghadapi tantangan dalam membedakan kelas dengan karakteristik visual yang sangat mirip pada kasus *fine-grained classification*.

Selanjutnya, hasil evaluasi pada skenario *proposed* ditunjukkan pada Tabel VI. Secara umum, penerapan *weighted sampling*, *data augmentation*, dan optimasi hyperparameter memberikan peningkatan performa pada sebagian besar model, terutama pada metrik Macro F1-score. Peningkatan ini menunjukkan bahwa strategi penanganan *class imbalance* membantu model mempelajari distribusi data secara lebih seimbang, khususnya pada kelas minoritas.

TABEL VI
HASIL EVALUASI MODEL PADA SKENARIO PROPOSED.

| Model | Accuracy (%) | Macro F1 | Weighted F1 | Time (Min) |
|-------------|--------------|----------|-------------|--------------|
| MobileNetV3 | 98.58 | 0.9772 | 0.9860 | 35.80 |
| ResNet50 | 98.87 | 0.9799 | 0.9888 | 60.64 |
| ViT-Small | 97.94 | 0.9683 | 0.9801 | 52.14 |

Pada skenario *proposed*, ResNet50 memperoleh nilai akurasi dan Macro F1-score tertinggi, yaitu masing-masing sebesar 98.87% dan 0.9799. Peningkatan performa ini menunjukkan bahwa arsitektur CNN yang lebih dalam mampu memanfaatkan distribusi data yang lebih seimbang untuk mengekstraksi fitur tekstur secara lebih efektif [14]. MobileNetV3 juga mengalami peningkatan performa dengan nilai Macro F1-score sebesar 0.9772, meskipun kapasitas model yang lebih ringan menyebabkan peningkatan tidak sebesar ResNet50. Sementara itu, ViT-Small memperoleh nilai Macro F1-score sebesar 0.9683, yang menunjukkan bahwa model berbasis transformer masih menghadapi keterbatasan pada dataset dengan jumlah data relatif terbatas dan distribusi yang tidak seimbang [10], [14].

Dari sisi komputasi, MobileNetV3 menunjukkan waktu pelatihan paling rendah, yaitu 35.80 menit, sehingga lebih sesuai untuk implementasi pada perangkat dengan sumber daya terbatas. Sebaliknya, ResNet50 membutuhkan waktu pelatihan lebih lama sebesar 60.64 menit, namun memberikan performa klasifikasi yang paling stabil dan konsisten dengan nilai Macro F1-score tertinggi. ViT-Small memiliki waktu pelatihan sebesar 52.14 menit, tetapi performanya cenderung lebih sensitif terhadap kondisi distribusi data dan konfigurasi pelatihan dibandingkan model berbasis CNN. Hasil ini

menunjukkan adanya trade-off antara efisiensi komputasi dan performa klasifikasi pada masing-masing arsitektur model.

B. Analisis Performa Model

Berdasarkan Tabel IV dan Tabel V, seluruh model menunjukkan akurasi yang tinggi. Namun, nilai *Macro F1-score* memberikan gambaran yang lebih representatif dalam mengevaluasi kemampuan model pada kondisi data tidak seimbang.

Pada skenario *baseline*, ViT-Small memperoleh performa tertinggi, yang menunjukkan bahwa mekanisme *self-attention* mampu menangkap hubungan global antar bagian citra secara efektif [14]. Namun, selisih performa antar model relatif kecil, yang menunjukkan bahwa seluruh arsitektur masih menghadapi kesulitan dalam membedakan kelas dengan karakteristik visual yang sangat mirip pada kasus *fine-grained classification*.

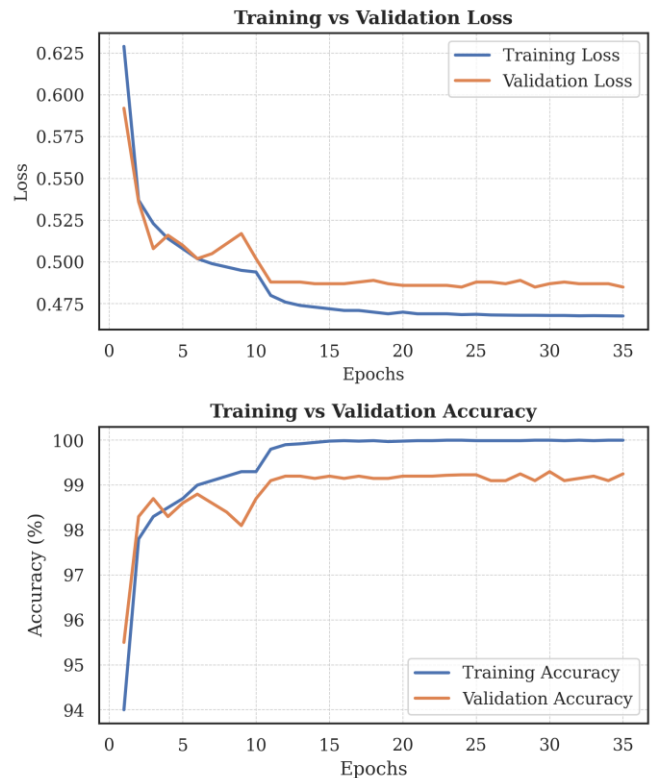
Setelah penerapan metode usulan, peningkatan performa paling signifikan terlihat pada ResNet50 dengan nilai Macro F1-score tertinggi. Hasil ini menunjukkan bahwa arsitektur CNN yang lebih dalam mampu memanfaatkan distribusi data yang lebih seimbang untuk mempelajari fitur tekstur secara lebih efektif pada kasus *fine-grained classification* [12]. Peningkatan performa juga terjadi pada MobileNetV3, meskipun tidak sebesar ResNet50, yang menunjukkan adanya keterbatasan kapasitas representasi fitur pada model dengan arsitektur yang lebih ringan [13].

Berbeda dengan kedua model berbasis CNN, ViT-Small tidak menunjukkan peningkatan performa yang signifikan setelah penerapan metode usulan. Kondisi ini mengindikasikan bahwa model berbasis transformer cenderung lebih sensitif terhadap jumlah data pelatihan dan distribusi kelas yang tidak seimbang [10], [14]. Selain itu, tidak adanya inductive bias terhadap fitur lokal menyebabkan Vision Transformer kurang optimal dalam mempelajari pola tekstur halus pada dataset dengan ukuran terbatas.

Dari sisi komputasi, MobileNetV3 menunjukkan waktu pelatihan paling rendah dibandingkan model lainnya, sehingga lebih sesuai untuk implementasi pada perangkat dengan sumber daya terbatas. Sebaliknya, ResNet50 membutuhkan waktu pelatihan lebih lama, namun memberikan performa klasifikasi yang paling stabil dan konsisten. ViT-Small memiliki waktu pelatihan yang berada di antara kedua model tersebut, tetapi performanya cenderung lebih sensitif terhadap kondisi distribusi data dan konfigurasi pelatihan.

Hasil ini menunjukkan bahwa MobileNetV3 lebih berpotensi digunakan pada perangkat dengan sumber daya terbatas, seperti sistem inspeksi portabel atau *embedded device* di industri pangan. Sementara itu, ResNet50 lebih sesuai digunakan pada sistem dengan kebutuhan akurasi tinggi meskipun memerlukan komputasi yang lebih besar.

ResNet50-Proposed



Gambar 3. Kurva Training dan Validation pada ResNet50-Proposed

Untuk menganalisis stabilitas proses pelatihan dan kemungkinan overfitting, dilakukan pengamatan terhadap kurva *training* dan *validation* pada model ResNet50 dengan konfigurasi *proposed* yang ditunjukkan pada Gambar 3. Hasil pelatihan menunjukkan bahwa nilai *validation* accuracy tetap stabil hingga akhir epoch tanpa mengalami penurunan yang signifikan. Selain itu, *validation* loss juga cenderung stabil tanpa peningkatan yang tajam, sehingga mengindikasikan bahwa model masih mampu melakukan generalisasi dengan baik pada data validasi.

Meskipun terdapat selisih antara *training* dan *validation* performance, kondisi tersebut masih tergolong wajar pada proses *transfer learning* dengan dataset terkontrol tidak menunjukkan indikasi overfitting yang signifikan. Penggunaan *data augmentation* dan *label smoothing* turut membantu meningkatkan kemampuan generalisasi model. Tidak ditemukan indikasi data leakage karena pemisahan data *training*, *validation*, dan *testing* dilakukan sebelum proses augmentasi dan *weighted sampling* diterapkan. Dengan demikian, data pada tahap evaluasi tetap terpisah dari proses pembelajaran model. Sementara itu, evaluasi menggunakan Macro F1-score dan confusion matrix digunakan untuk memastikan performa model tetap seimbang pada seluruh kelas dan tidak bias terhadap kelas mayoritas.

Secara keseluruhan, hasil penelitian menunjukkan bahwa model berbasis CNN, khususnya ResNet50, lebih stabil dan sesuai untuk klasifikasi kualitas beras pada dataset dengan

karakteristik fine-grained dan distribusi data yang tidak seimbang. Temuan ini menunjukkan bahwa kemampuan ekstraksi fitur lokal pada CNN masih lebih efektif dibandingkan pendekatan transformer pada kondisi dataset terbatas dan berbasis tekstur [12], [14].

C. Analisis Metode Penanganan Imbalance

Ketidakseimbangan distribusi kelas ditangani menggunakan *weighted sampling*, *data augmentation*, *Focal Loss*, dan *Class-Balanced Loss*. Pada *weighted sampling*, bobot setiap kelas dihitung berdasarkan kebalikan jumlah sampel sehingga kelas minoritas memiliki probabilitas lebih besar untuk dipilih selama proses pelatihan.

Berdasarkan Tabel VII, kelas minoritas memperoleh bobot lebih besar dibandingkan kelas mayoritas sehingga distribusi pembelajaran menjadi lebih seimbang [16], [20]. Selain itu, *data augmentation* digunakan untuk meningkatkan variasi data latihan dan mengurangi risiko overfitting pada dataset dengan jumlah data terbatas.

TABEL VII
BOBOT KELAS PADA WEIGHTED SAMPLING

| Kelas | Jumlah Data (n_i) | Bobot ($w_i = 1/n_i$) |
|-------|-----------------------|-------------------------|
| 0_NOR | 18000 | 0.0000556 |
| 1_F&S | 1336 | 0.000749 |
| 2_SD | 1350 | 0.000741 |
| 3_MY | 1350 | 0.000741 |
| 4_AP | 1350 | 0.000741 |
| 5_BN | 1341 | 0.000746 |
| 6_UN | 1335 | 0.000749 |
| 7_IM | 1800 | 0.000556 |

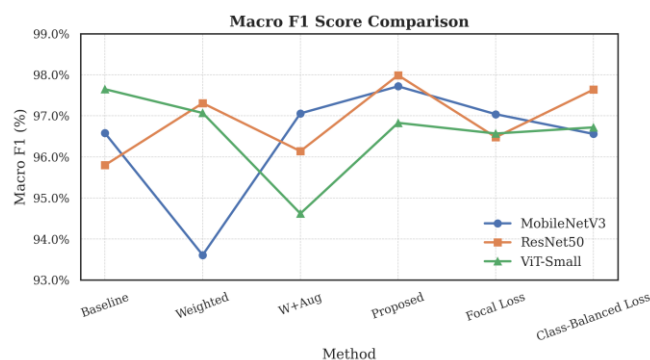
Proposed Method merupakan kombinasi *weighted sampling*, *data augmentation*, dan optimasi hyperparameter. Sementara itu, *Focal Loss-Based* dan *Class-Balanced Loss-Based* menggunakan konfigurasi augmentasi dan hyperparameter yang sama, tetapi dengan pendekatan *loss balancing* yang berbeda. Berdasarkan Tabel VIII, penerapan *weighted sampling* memberikan peningkatan performa pada ResNet50, dengan Macro F1-score naik dari 0.9580 menjadi 0.9731. Namun, penurunan pada MobileNetV3 menunjukkan bahwa model ringan lebih sensitif terhadap perubahan distribusi data. Penambahan *data augmentation* juga belum memberikan hasil yang konsisten, terutama pada ViT-Small yang turun hingga 0.9462, sehingga transformasi augmentasi perlu disesuaikan dengan karakteristik tekstur pada kasus *fine-grained classification*.

TABEL VIII
PERBANDINGAN PERFORMA MODEL PADA SETIAP SKENARIO

| Metode | Model | Accuracy (%) | Macro F1 |
|------------------------|-------------|--------------|----------|
| Baseline | MobileNetV3 | 98.10 | 0.9658 |
| | ResNet50 | 97.42 | 0.9580 |
| | ViT-Small | 98.71 | 0.9765 |
| +Weighted Sampling | MobileNetV3 | 95.42 | 0.9361 |
| | ResNet50 | 98.52 | 0.9731 |
| | ViT-Small | 98.13 | 0.9707 |
| Weighted +Augmentation | MobileNetV3 | 98.39 | 0.9706 |
| | ResNet50 | 97.42 | 0.9614 |
| | ViT-Small | 96.48 | 0.9462 |

| | | | |
|---------------------------|-----------------|--------------|---------------|
| Proposed Method | MobileNetV3 | 98.58 | 0.9772 |
| | ResNet50 | 98.87 | 0.9799 |
| | ViT-Small | 97.94 | 0.9683 |
| Focal Loss-Based | MobileNetV3 | 98.16 | 0.9704 |
| | ResNet50 | 97.90 | 0.9648 |
| | ViT-Small | 97.58 | 0.9657 |
| Class-Balanced Loss-Based | MobileNetV3 | 97.71 | 0.9656 |
| | ResNet50 | 98.55 | 0.9764 |
| | ViT-Small | 97.97 | 0.9672 |

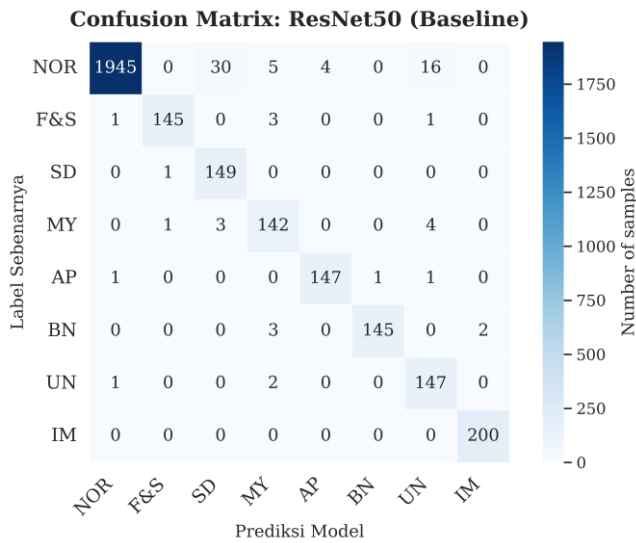
Pada skenario proposed, ResNet50 memperoleh Macro F1-score tertinggi sebesar 0.9799, meningkat 2.19% dibandingkan *baseline*. Hasil ini menunjukkan bahwa kombinasi *weighted sampling*, *data augmentation*, dan optimasi hyperparameter memberikan performa paling stabil dibandingkan *Focal Loss-Based* dan *Class-Balanced Loss-Based*. Meskipun kedua metode *loss balancing* tersebut cukup kompetitif, performanya masih berada di bawah pendekatan *proposed* pada sebagian besar model. Visualisasi pada Gambar 4 juga menunjukkan bahwa ResNet50 memiliki tren performa paling stabil, sedangkan ViT-Small menunjukkan fluktuasi lebih besar akibat sensitivitas model transformer terhadap distribusi data dan sensitivitas terhadap konfigurasi pelatihan.



Gambar 4. Perbandingan Macro F1-score pada Setiap Skenario

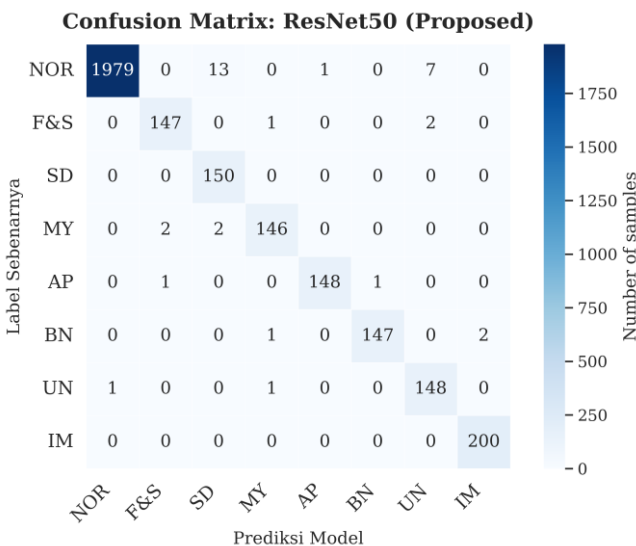
D. Analisis Confusion Matrix

Pada skenario *baseline* Gambar 5, masih ditemukan beberapa kesalahan klasifikasi, terutama pada kelas mayoritas 0_NOR yang masih tertukar dengan kelas 2_SD dan 6_UN. Kondisi ini menunjukkan bahwa sebagian citra beras normal masih memiliki kemiripan visual dengan kelas Slightly Damaged dan Unmilled/Husks, terutama pada karakteristik tekstur dan warna permukaan. Kesalahan klasifikasi juga terlihat pada beberapa kelas minoritas seperti 3_MY dan 5_BN, yang menunjukkan bahwa model belum mampu membedakan fitur antar kelas dengan karakteristik visual serupa secara optimal.



Gambar 5. Confusion Matrix ResNet50 pada Skenario Baseline

Setelah penerapan metode *proposed* Gambar 6, jumlah kesalahan klasifikasi mengalami penurunan dibandingkan *skenario baseline*. Kesalahan prediksi pada kelas 0_NOR terhadap kelas 2_SD dan 6_UN berkurang dibandingkan *skenario baseline*. Selain itu, peningkatan performa juga terlihat pada beberapa kelas minoritas, terutama 3_MY, yang menunjukkan distribusi pembelajaran menjadi lebih seimbang setelah penerapan *weighted sampling*, *data augmentation*, dan optimasi hyperparameter.



Gambar 6. Confusion Matrix ResNet50 pada Skenario Proposed

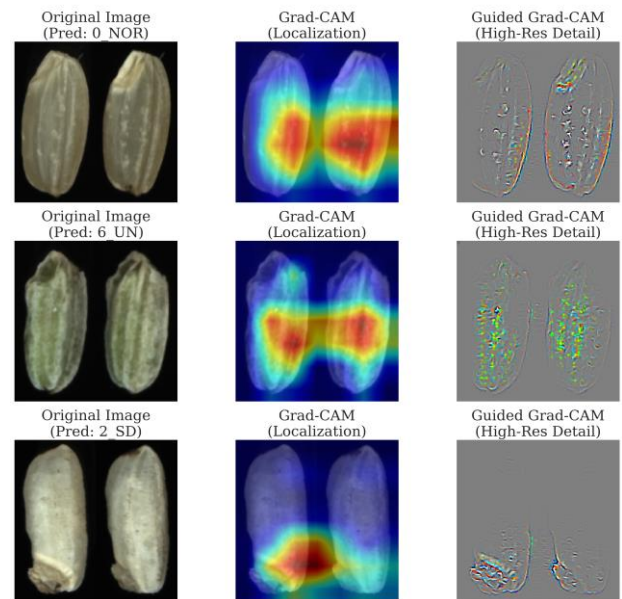
Peningkatan performa pada *skenario proposed* juga terlihat pada nilai precision dan recall beberapa kelas minoritas. Sebagai contoh, kelas 3_MY mengalami peningkatan F1-score dari 0.9311 pada *baseline* menjadi 0.9766 pada *skenario proposed*, sedangkan kelas 6_UN meningkat dari 0.9216 menjadi 0.9642. Hasil ini menunjukkan bahwa pendekatan *proposed* tidak hanya meningkatkan akurasi keseluruhan,

tetapi juga memperbaiki kemampuan model dalam mengenali kelas minoritas secara lebih konsisten.

Distribusi nilai confusion matrix pada *skenario proposed* terlihat lebih terpusat pada diagonal, yang menunjukkan bahwa model semakin akurat dalam melakukan klasifikasi pada setiap kelas.

Meskipun demikian, beberapa kesalahan kecil masih ditemukan pada kelas minoritas seperti 5_BN dan 6_UN, yang menunjukkan bahwa kemiripan karakteristik visual antar kelas masih menjadi tantangan pada kasus *fine-grained classification*.

Analisis Interpretabilitas: Grad-CAM dan Guided Grad-CAM



Gambar 7. Visualisasi Interpretabilitas Model Menggunakan Grad-CAM dan Guided Grad-CAM pada ResNet50-Proposed

Untuk memahami area citra yang menjadi perhatian model selama proses klasifikasi, dilakukan visualisasi menggunakan Grad-CAM seperti ditunjukkan pada Gambar 7. Hasil visualisasi menunjukkan bahwa model cenderung berfokus pada area tekstur permukaan, bagian retakan, dan tepi butiran beras yang memiliki karakteristik visual khas pada masing-masing kelas. Selain itu, model juga mampu mengenali fitur penting pada beberapa kelas minoritas seperti 6_UN. Pada kelas 2_SD, model terlihat memberikan perhatian pada area tekstur dan pola permukaan biji yang berkaitan dengan karakteristik beras *slightly damaged*. Temuan ini menunjukkan bahwa model memanfaatkan fitur tekstur lokal yang relevan pada kasus *fine-grained classification*.

Secara keseluruhan, confusion matrix menunjukkan bahwa kombinasi *weighted sampling*, *data augmentation*, dan optimasi hyperparameter mampu mengurangi kesalahan klasifikasi, terutama pada kelas minoritas. Pendekatan *proposed* juga membantu model mempelajari distribusi data secara lebih seimbang dan meningkatkan kemampuan generalisasi model.

E. Keterbatasan Penelitian

Penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan berasal dari lingkungan terkontrol sehingga variasi pencahayaan, kualitas kamera, dan noise citra pada kondisi nyata belum sepenuhnya terwakili. Selain itu, evaluasi model masih dilakukan pada satu dataset sehingga kemampuan generalisasi terhadap jenis beras lain maupun kondisi citra yang berbeda masih perlu dikaji lebih lanjut.

Model juga belum diuji pada skenario implementasi *real-time* maupun perangkat dengan keterbatasan komputasi secara langsung. Oleh karena itu, penelitian selanjutnya dapat difokuskan pada pengujian menggunakan dataset eksternal, kondisi akuisisi citra yang lebih beragam, serta evaluasi performa model pada lingkungan industri pangan secara nyata.

IV. KESIMPULAN

Penerapan ini menunjukkan bahwa kombinasi *weighted sampling*, *data augmentation*, dan optimasi hyperparameter mampu meningkatkan performa klasifikasi kualitas beras pada kondisi data tidak seimbang. Peningkatan performa paling terlihat pada nilai Macro F1-score, yang menunjukkan bahwa model menjadi lebih seimbang dalam mengenali seluruh kelas, termasuk kelas minoritas.

Berdasarkan hasil eksperimen, ResNet50 pada skenario *proposed* memperoleh performa terbaik dengan nilai Macro F1-score sebesar 0.9799. Hasil ini menunjukkan bahwa arsitektur CNN yang lebih dalam lebih efektif dalam mengekstraksi fitur tekstur pada kasus *fine-grained classification* dibandingkan model berbasis transformer. Sementara itu, ViT-Small menunjukkan performa yang lebih sensitif terhadap distribusi data dan jumlah data pelatihan.

Hasil perbandingan metode juga menunjukkan bahwa pendekatan *proposed* memberikan performa yang lebih stabil dibandingkan *Focal Loss-Based* dan *Class-Balanced Loss-Based* pada sebagian besar skenario pengujian. Selain itu, analisis confusion matrix menunjukkan penurunan kesalahan klasifikasi pada beberapa kelas minoritas setelah penerapan metode *proposed*.

Penelitian selanjutnya dapat difokuskan pada pengujian menggunakan dataset eksternal, kondisi citra yang lebih beragam, serta penerapan teknik augmentasi yang lebih adaptif untuk meningkatkan kemampuan generalisasi model.

DAFTAR PUSTAKA

- [1] L. Fan et al., "An annotated grain kernel image database for visual quality inspection," *Scientific Data*, vol. 10, no. 1, 2023, doi: 10.1038/s41597-023-02660-8.
- [2] S. A.-A. Shuvo et al., "Image-based classification of Bangladeshi rice varieties using deep convolutional neural networks," *Agricultural Products Processing and Storage*, early access, 2024, doi: 10.1007/s44462-026-00061-9.
- [3] W. Agustiono et al., "Rice variety identification based on transfer learning architecture using DENS-INCEP," *IEEE Access*, vol. 13, pp. 78007–78020, 2025.
- [4] F. Farahnakian et al., "A comparative study of state-of-the-art deep learning architectures for rice grain classification," *Journal of Agriculture and Food Research*, vol. 15, 2024.
- [5] R. Setiawan and H. Oumarou, "Classification of rice grain varieties using ensemble learning and image analysis techniques," *Indonesian Journal of Data and Science*, vol. 5, no. 1, pp. 54–63, 2024.
- [6] S. S. Hidayat et al., "Determining the rice seeds quality using convolutional neural network," *International Journal on Informatics Visualization*, vol. 7, no. 2, pp. 527–534, 2023, doi: 10.30630/ijov.7.2.1175.
- [7] L. Fan et al., "GrainSpace: A large-scale dataset for fine-grained and domain-adaptive recognition of cereal grains," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2022.
- [8] P. Xu et al., "Research on maize seed classification and recognition based on machine vision and deep learning," *Agriculture*, vol. 12, no. 2, 2022.
- [9] Y. H. Wang and W. H. Su, "Convolutional neural networks in computer vision for grain crop phenotyping: A review," *Agronomy*, vol. 12, no. 11, 2022.
- [10] Y. Wang et al., "Vision transformers for image classification: A comparative survey," *Technologies*, vol. 13, no. 1, 2025.
- [11] H. Touvron et al., "Training data-efficient image transformers," in *Proc. International Conference on Machine Learning (ICML)*, 2021.
- [12] K. He et al., "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [13] A. Howard et al., "Searching for MobileNetV3," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2019.
- [14] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. International Conference on Learning Representations (ICLR)*, 2021.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [16] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.
- [17] B. Arora et al., "Rice grain classification using image processing and machine learning techniques," in *Proc. International Conference on Inventive Computation Technologies (ICICT)*, 2020.
- [18] D. S. Vishnu and D. L. S. Reddy, "Rice classification using deep neural network," in *Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE)*, Atlantis Press, 2025, doi: 10.2991/978-94-6463-858-5_48.
- [19] L. J. Mawarani et al., "Rice identification using convolutional neural network with YOLOv7 and VGG16," *IPTEK The Journal of Engineering*, vol. 10, no. 3, 2024.
- [20] G. Lemaître, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning," *Journal of Machine Learning Research*, vol. 18, no. 17, pp. 1–5, 2017.
- [21] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2980–2988.
- [22] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 9268–9277.
- [23] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2018.