



STRATEGI PENANGANAN KETIDAKSEIMBANGAN DATA CITRA PENYAKIT DAUN KELAPA SAWIT MENGGUNAKAN CLASS WEIGHT DAN GENERATIVE ADVERSARIAL NETWORK

MUHAMMAD FAHREZI MAULANA



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Strategi Penanganan Ketidakseimbangan Data Citra Penyakit Daun Kelapa Sawit Menggunakan *Class weight* dan *Generative Adversarial Network*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dalam penyusunan karya ini, penulis menggunakan bantuan kecerdasan buatan berupa ChatGPT dan Claude sebagai sarana pendukung dalam membantu penyusunan sintaks Python pada proses analisis data dan penyempurnaan naskah. Setelah menggunakan alat/layanan tersebut, penulis meninjau dan menyunting konten sesuai kebutuhan serta bertanggung jawab penuh atas isi karya tugas akhir ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Muhammad Fahrezi Maulana
G1401221103



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

MUHAMMAD FAHREZI MAULANA. Strategi Penanganan Ketidakseimbangan Data Citra Penyakit Daun Kelapa Sawit Menggunakan *Class weight* dan *Generative Adversarial Network*. Dibimbing oleh SACHNAZ DESTA OKTARINA dan PIKA SILVIANTI.

Ketidakseimbangan data masih menjadi tantangan dalam klasifikasi citra karena menyebabkan model cenderung mempelajari pola kelas mayoritas sehingga kemampuan mengenali kelas minoritas menurun. Berbagai strategi penanganan telah dikembangkan, di antaranya pendekatan berbasis algoritma melalui *class weight* dan pendekatan berbasis data melalui pembangkitan citra sintetis menggunakan *Deep Convolutional Generative Adversarial Network* (DCGAN). Penelitian ini bertujuan membandingkan efektivitas kedua strategi tersebut terhadap model *baseline* pada klasifikasi penyakit daun bibit kelapa sawit menggunakan MobileNetV2. Dataset yang digunakan terdiri atas 1.104 citra yang dikelompokkan ke dalam empat kelas, yaitu antraknosa, bercak daun, sehat, dan lainnya. Evaluasi dilakukan menggunakan *Stratified 5-Fold Cross Validation* dengan *macro F1-score* dan *balanced accuracy* sebagai metrik utama. Hasil penelitian menunjukkan bahwa skenario *baseline* menghasilkan rata-rata *macro F1-score* sebesar 89,07% dan *balanced accuracy* sebesar 87,00%, skenario *class weight* sebesar 90,05% dan 91,91%, sedangkan skenario DCGAN sebesar 91,19% dan 89,63%. DCGAN memberikan nilai *macro F1-score* tertinggi, sedangkan *class weight* menghasilkan *balanced accuracy* tertinggi. Namun, berdasarkan uji Kruskal–Wallis, tidak terdapat perbedaan performa yang signifikan antartiga skenario ($p\text{-value} > 0,05$). Dengan demikian, penerapan *class weight* maupun DCGAN memberikan performa yang sebanding dalam menangani ketidakseimbangan data pada dataset yang digunakan.

Kata kunci: *class weight*, DCGAN, ketidakseimbangan data, klasifikasi citra, MobileNetV2.



ABSTRACT

MUHAMMAD FAHREZI MAULANA. Strategies for Handling Imbalanced Oil Palm Leaf Disease Image Data Using Class weight and Generative Adversarial Networks. Supervised by SACHNAZ DESTA OKTARINA and PIKA SILVIANTI.

Data imbalance remains a major challenge in image classification because it causes models to favor learning patterns from majority classes, thereby reducing their ability to recognize minority classes. Various strategies have been developed to address this issue, including algorithm-level approaches through class weighting and data-level approaches through synthetic image generation using Deep Convolutional Generative Adversarial Networks (DCGAN). This study aimed to compare the effectiveness of these strategies with a baseline model for oil palm seedling leaf disease classification using MobileNetV2. The dataset consisted of 1,104 images categorized into four classes: anthracnose, leaf spot, healthy, and others. Model performance was evaluated using Stratified 5-Fold Cross Validation, with macro F1-score and balanced accuracy as the primary evaluation metrics. The results showed that the baseline model achieved an average macro F1-score of 89.07% and a balanced accuracy of 87.00%, while the class weight approach achieved 90.05% and 91.91%, respectively. The DCGAN approach achieved a macro F1-score of 91.19% and a balanced accuracy of 89.63%. DCGAN produced the highest macro F1-score, whereas the class weight approach yielded the highest balanced accuracy. However, the Kruskal–Wallis test indicated no statistically significant difference in performance among the three scenarios (p -value > 0.05). Therefore, both class weighting and DCGAN provided comparable performance in handling data imbalance for the dataset used in this study.

Keywords: class weight, DCGAN, image classification, imbalanced data, MobileNetV2.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

STRATEGI PENANGANAN KETIDAKSEIMBANGAN DATA CITRA PENYAKIT DAUN KELAPA SAWIT MENGGUNAKAN CLASS WEIGHT DAN GENERATIVE ADVERSARIAL NETWORK

MUHAMMAD FAHREZI MAULANA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

IPB University

Tim Penguji pada Ujian Skripsi:

1 Rahma Anisa, S.Stat., M.Si., M.Act.Sc.





Judul Skripsi : Strategi Penanganan Ketidakseimbangan Data Citra Penyakit
Daun Kelapa Sawit Menggunakan *Class weight* dan *Generative
Adversarial Network*

Nama : Muhammad Fahrezi Maulana
NIM : G1401221103

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Sachnaz Desta Oktarina S.Stat., M.Agr., Ph.D.

Pembimbing 2:
Pika Silvianti, S.Si., M.Si.

Diketahui oleh

Ketua Program Studi:
Dr. Bagus Sartono, S.Si., M.Si.
NIP 197804112005011002

Tanggal Ujian:
17 Juli 2026

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2025 sampai bulan Juni 2026 ini ialah Kelapa Sawit, dengan judul “Strategi Penanganan Ketidakseimbangan Data Citra Penyakit Daun Kelapa Sawit Menggunakan *Class weight* dan *Generative Adversarial Network*”.

Terima kasih penulis ucapkan kepada seluruh pihak yang telah memberikan doa, bimbingan, serta dukungan dalam penyusunan karya ilmiah ini, di antaranya:

1. Bapak Rahmad dan Ibu Armayani Hajar Siagian selaku orang tua serta Khairunnisa Maulani dan Muhammad Farhan Maulana selaku saudara kandung yang telah memberikan doa dan dukungan terhadap penulis selama menempuh masa perkuliahan.
2. Ibu Sachnaz Desta Oktarina S.Stat., M.Agr., Ph.D. dan Ibu Pika Silvianti, S.Si., M.Si. yang selalu membimbing penulis dengan memberi saran dan arahan selama penyusunan karya ilmiah ini.
3. Seluruh dosen Program Studi Statistika dan Sains Data SSMI IPB yang telah memberikan ilmu dan pembelajaran yang sangat bermanfaat bagi penulis beserta staf Tata Usaha Program Studi Statistika dan Sains Data SSMI yang turut membantu penulis dalam mempersiapkan administrasi selama perkuliahan hingga penulisan tugas akhir.
4. PT. Perkebunan Nusantara IV yang telah membiayai dan memberikan kesempatan penulis untuk kuliah di IPB University.
5. Mahib, Abyan, Hafidz, Yapi, Ojan, Raihan, Aji, Inteam, dan Keluarga GSB selaku sahabat penulis yang membersamai penulis selama penyusunan karya ilmiah ini.
6. Frisky Muthia Yasifa yang senantiasa memberi semangat dan dukungan emosional terhadap penulis selama penulisan tugas akhir ini.
7. Alm. Indra Mulya Ginting selaku sahabat penulis dari bangku SMA yang memberikan arahan dan semangat untuk berkuliah di Statistika IPB University.
8. Seluruh teman teman Statistika 58, Statistika 59, dan Statistika 60 yang telah menemani penulis selama berkuliah di IPB.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2026

Muhammad Fahrezi Maulana



DAFTAR ISI

DAFTAR TABEL	ix
DAFTAR GAMBAR	ix
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
TINJAUAN PUSTAKA	4
2.1 <i>Convolutional Neural Network (CNN)</i>	4
2.2 <i>Transfer learning</i>	6
2.3 <i>MobileNetV2</i>	7
2.4 <i>Class weight</i>	8
2.5 <i>Generative Adversarial Network (GAN)</i>	9
2.6 <i>Stratified K-fold Cross Validation</i>	11
2.7 Uji Kruskal-Wallis	12
III METODE	13
3.1 Data	13
3.2 Tahapan Analisis	13
IV HASIL DAN PEMBAHASAN	19
4.1 Eksplorasi Data	19
4.2 Praproses Data	20
4.3 <i>Stratified 5-Fold Cross Validation</i>	20
4.4 Pelatihan Model Klasifikasi	21
4.5 Evaluasi Model	29
4.6 Analisis Perbedaan Performa Model	34
V SIMPULAN DAN SARAN	36
5.1 Simpulan	36
5.2 Saran	36
DAFTAR PUSTAKA	37
RIWAYAT HIDUP	40



DAFTAR TABEL

1	Tampilan <i>confusion matrix</i> 4x4	17
2	Distribusi citra setiap <i>fold</i>	21
3	Hasil pelatihan skenario <i>baseline</i>	22
4	Hasil pelatihan skenario <i>class weight</i>	23
5	Jumlah citra hasil pelatihan DCGAN	25
6	Parameter pelatihan DCGAN	25
7	Hasil pelatihan skenario DCGAN	28
8	Hasil evaluasi model skenario <i>baseline</i>	30
9	Hasil evaluasi model skenario <i>class weight</i>	31
10	Hasil evaluasi model skenario DCGAN	32
11	Hasil uji Kruskal-Wallis	34

DAFTAR GAMBAR

1	Arsitektur CNN (Purwono <i>et al.</i> 2022)	4
2	Ilustrasi <i>convolution layer</i> (Purwono <i>et al.</i> 2022)	5
3	Ilustrasi <i>max pooling</i> (Zafar <i>et al.</i> 2024)	6
4	Arsitektur MobileNetV2 (Muhathir <i>et al.</i> 2026)	8
5	Ilustrasi mekanisme <i>stratified k-fold cross validation</i> (Duan 2023)	12
6	Contoh dataset (a) bercak daun (b) sehat (c) kelas lainnya (d) antraknosa	13
7	Tahapan penelitian	14
8	Distribusi masing-masing kelas penyakit	19
9	Hasil praproses (a) data citra aktual (b) data citra hasil praproses	20
10	Distribusi proporsi setiap kelas pada masing-masing fold hasil <i>Stratified 5-Fold Cross Validation</i>	21
11	Kurva <i>validation loss</i> skenario <i>baseline</i>	22
12	Kurva <i>validation loss</i> skenario <i>class weight</i>	24
13	Kurva <i>loss generator</i> dan <i>loss discriminator</i>	26
14	Hasil data sintetis pada <i>fold</i> 1	27
15	Kurva <i>validation loss</i> skenario DCGAN	28
16	Boxplot nilai metrik evaluasi pada setiap skenario	33



@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University