

Klasifikasi buah dan sayuran multi-label menggunakan CNN: mengatasi class imbalance dengan focal loss

Indu Indah Purnomo^{*1}, Gita Ayu Syafarina², Muhammad Hasbi³

Email: ¹indu@uniska-bjm.ac.id, ²gitaayu@uniska-bjm.ac.id, ³m.hasbi@uniska-bjm.ac.id

^{1,2,3}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Islam Muhammad Arsyad AlBanjari

Diterima: 05 April 2020 | Direvisi: 05 Mei 2020 | Disetujui: 27 Mei 2020

©2020 Program Studi Teknik Informatika Fakultas Ilmu Komputer,
Universitas Muhammadiyah Riau, Indonesia

Abstrak

Menginvestigasi efektivitas Focal Loss sebagai solusi untuk masalah ketidakseimbangan kelas (class imbalance) dalam tugas klasifikasi multi-label buah dan sayuran. Dengan menggunakan arsitektur Convolutional Neural Network (CNN) berbasis ResNet50, dua model dilatih dan dievaluasi: satu menggunakan Focal Loss dan satu lagi menggunakan Binary Cross-Entropy (BCE) Loss sebagai baseline. Untuk mengatasi ketersediaan dataset multi-label, sebuah dataset sintesis multi-label dibuat dengan menggabungkan gambar-gambar dari dataset single-label yang sudah ada. Hasil eksperimen menunjukkan bahwa model yang dilatih dengan Focal Loss mencapai akurasi 0.9390 dan F1-score 0.9863, mengungguli model BCE Loss yang hanya mencapai akurasi 0.8850 dan F1-score 0.9718. Analisis komparatif mengindikasikan bahwa Focal Loss, dengan kemampuannya untuk memfokuskan proses pelatihan pada contoh-contoh yang sulit, secara efektif mengatasi class imbalance dan menghasilkan kinerja yang lebih superior. Penelitian ini menyimpulkan bahwa Focal Loss adalah alat yang efektif untuk tugas klasifikasi multi-label dan menyoroti batasan-batasan yang ada, termasuk sifat sintesis dari dataset dan durasi pelatihan yang terbatas, yang menggarisbawahi perlunya penelitian lebih lanjut.

Kata kunci: Klasifikasi Multi-Label, CNN, Focal Loss, ResNet50, Class Imbalance, Klasifikasi Buah dan Sayuran

Multi-label fruit and vegetable classification using CNN: addressing class imbalance with focal loss

Abstract

Investigates the effectiveness of Focal Loss as a solution to the problem of class imbalance in multi-label fruit and vegetable classification tasks. Using a ResNet50-based Convolutional Neural Network (CNN) architecture, two models were trained and evaluated: one using Focal Loss and another using Binary Cross-Entropy (BCE) Loss as a baseline. To address the availability of multi-label datasets, a synthetic multi-label dataset was created by combining images from existing single-label datasets. Experimental results show that the model trained with Focal Loss achieved an accuracy of 0.9390 and an F1-score of 0.9863, outperforming the BCE Loss model which only reached an accuracy of 0.8850 and an F1-score of 0.9718. The comparative analysis indicates that Focal Loss, with its ability to focus the training process on difficult examples, effectively addresses class imbalance and produces superior performance. This study concludes that Focal Loss is an effective tool for multi-label classification tasks and highlights the existing limitations, including the synthetic nature of the dataset and the limited training duration, which underscore the need for further research.

Keywords: Multi-Label Classification, CNN, Focal Loss, ResNet50, Class Imbalance, Fruit and Vegetable Classification

1. PENDAHULUAN

Klasifikasi citra, khususnya untuk objek buah dan sayuran, memiliki relevansi yang tinggi dalam berbagai aplikasi industri, mulai dari sistem pengawasan mutu pertanian, manajemen rantai pasokan, hingga personalisasi pengalaman belanja daring [1]. Secara tradisional, tugas ini sering kali diperlakukan sebagai masalah klasifikasi single-label, di mana setiap citra diasumsikan hanya mengandung satu objek dari satu kelas. Namun, skenario dunia nyata jauh lebih kompleks, sering kali menampilkan citra

yang mengandung banyak objek berbeda. Hal ini memunculkan kebutuhan akan model yang mampu melakukan klasifikasi multi-label, di mana satu citra dapat dikaitkan dengan lebih dari satu kelas secara bersamaan [2], [3].

Dalam tugas klasifikasi multi-label, tantangan signifikan yang sering muncul adalah masalah ketidakseimbangan kelas. Fenomena ini terjadi ketika beberapa kelas memiliki jumlah contoh yang jauh lebih banyak daripada kelas lainnya [4], [5]. Dalam konteks multi-label, masalah ini bisa menjadi lebih rumit karena ketidakseimbangan tidak hanya terjadi antar kelas, tetapi juga dalam konteks example imbalance, di mana sebagian besar contoh negatif (citra yang tidak mengandung label tertentu) sangat mudah untuk diidentifikasi oleh model, sementara contoh positif (citra yang mengandung label) jauh lebih sulit dan beragam. Hal ini dapat menyebabkan model yang dilatih dengan fungsi loss konvensional menjadi bias dan mengabaikan kinerja pada kelas-kelas minoritas [6][7]. Masalah ini diidentifikasi sebagai penyebab utama mengapa detektor objek satu-tahap seringkali memiliki akurasi yang lebih rendah daripada detektor dua-tahap.

Convolutional Neural Network (CNN) adalah kelas jaringan saraf tiruan yang dirancang khusus untuk memproses data piksel, seperti citra digital [8], [9]. Struktur dasar CNN terdiri dari lapisan konvolusi yang mengekstrak fitur spasial, lapisan pooling yang mengurangi dimensi fitur, dan lapisan fully connected yang melakukan klasifikasi berdasarkan fitur yang diekstrak [2]. Salah satu arsitektur CNN yang paling berpengaruh dan sering digunakan adalah ResNet (Residual Network), yang diperkenalkan oleh He et al. pada tahun 2016. ResNet mengatasi masalah vanishing gradient yang menghambat pelatihan jaringan yang sangat dalam melalui pengenalan skip connections atau koneksi residual. Koneksi ini memungkinkan sinyal gradient untuk melewati satu atau lebih lapisan dan langsung mengalir ke lapisan yang lebih dalam, yang pada gilirannya memfasilitasi pelatihan model yang jauh lebih dalam dan akurat. 1 Pemilihan ResNet50 dalam penelitian ini didasarkan pada reputasinya sebagai arsitektur yang efisien dan efektif untuk tugas klasifikasi citra [1].

Fungsi loss adalah komponen fundamental dalam pembelajaran mesin yang mengukur seberapa baik kinerja model [7]. Fungsi ini membandingkan keluaran yang diprediksi dengan label yang sebenarnya dan menghasilkan nilai yang mewakili "kesalahan" model. Tujuan pelatihan adalah untuk meminimalkan nilai loss ini melalui optimisasi. Untuk tugas klasifikasi multi-label, di mana setiap label diperlakukan sebagai tugas klasifikasi biner yang independen, Binary Cross-Entropy (BCE) Loss adalah pilihan standar dan paling umum. BCE Loss mengukur perbedaan antara probabilitas yang diprediksi oleh model (p) dan label yang benar (y) untuk setiap kelas. Fungsi ini memberikan bobot yang sama untuk setiap contoh, terlepas dari apakah contoh tersebut mudah atau sulit untuk diklasifikasikan [10].

Meskipun ResNet50 telah terbukti handal, penerapannya pada dataset multi-label yang dibangun secara sintesis menghadirkan tantangan unik, terutama terkait dominasi fitur latar belakang dibandingkan objek kecil dan ketidakseimbangan frekuensi antar kelas. Oleh karena itu, penelitian ini bertujuan untuk menginvestigasi efektivitas fungsi kerugian Focal Loss dibandingkan dengan Binary Cross-Entropy standar dalam menstabilkan gradien selama pelatihan [10]. Selain itu, penelitian ini mengusulkan metode pembangkitan data sintesis yang efisien untuk mengatasi kelangkaan dataset multi-label di domain pertanian. Klasifikasi multi-label berbeda secara fundamental dari klasifikasi biner atau multi-class tradisional. Dalam konteks citra pertanian, sebuah foto keranjang belanja atau tumpukan hasil panen mungkin berisi apel, wortel, dan bayam secara bersamaan. Secara matematis, jika L adalah himpunan label yang mungkin, tujuan model adalah memprediksi vektor biner $y \in \{0,1\}^{|L|}$ di mana setiap elemen merepresentasikan keberadaan label tertentu [11]. Kompleksitas ruang pencarian ini meningkat secara eksponensial seiring bertambahnya jumlah kelas, membuat model rentan terhadap overfitting jika data latih tidak representative [12], [13]. Masalah ketidakseimbangan kelas (class imbalance) adalah hambatan utama dalam pelatihan Deep Learning. Metode konvensional seperti oversampling atau undersampling seringkali tidak efektif pada kasus multi-label karena mengubah distribusi satu label dapat merusak distribusi label lainnya yang muncul bersamaan (co-occurrence) [5]. Oleh karena itu, pendekatan berbasis algoritma melalui modifikasi fungsi loss menjadi solusi yang lebih elegan [14].

Dalam skenario standar, Binary Cross-Entropy (BCE) digunakan untuk menghitung kerugian rata-rata dari setiap kelas secara independen:

$$CE(p_t) = -\log(p_t)$$

Di mana p_t adalah probabilitas model untuk kelas yang benar. Namun, BCE memiliki kelemahan yaitu memperlakukan semua contoh (baik yang mudah maupun sulit diklasifikasikan) dengan bobot yang sama. Pada dataset yang sangat tidak seimbang, gradien dari ribuan "contoh mudah" (easy negatives) akan menenggelamkan sinyal dari "contoh sulit" (hard positives), menyebabkan model gagal mendeteksi objek minoritas [6].

Formulasi Focal Loss Untuk mengatasi kelemahan BCE, Lin et al. memperkenalkan Focal Loss (FL) yang menambahkan faktor modulasi $(1-p_t)^\gamma$ pada persamaan cross-entropy [10]. Formulasi matematis yang digunakan dalam penelitian ini adalah:

$$FL(p_t) = -\alpha_t(1-p_t)^\gamma \log(p_t)$$

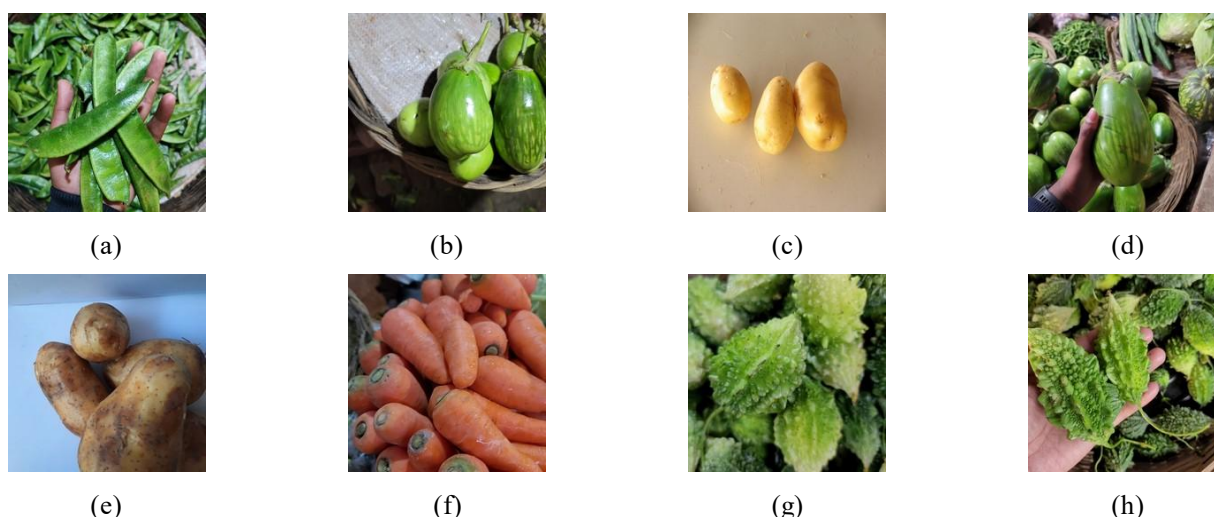
Penelitian ini mengimplementasikan Focal Loss dengan parameter pemokus (focusing parameter) $\gamma=2.0$ dan parameter penyeimbang (balancing parameter) $\alpha=0.25$. Nilai $\gamma>0$ berfungsi mengurangi kontribusi loss dari contoh yang sudah

terklasifikasi dengan baik (di mana $pt \approx 1$), sehingga memaksa model untuk memfokuskan pembaruan bobot pada contoh-contoh yang sulit dan sering salah diprediksi.

2. METODE PENELITIAN

2.1 Persiapan Dataset Multi-Label Sintetis

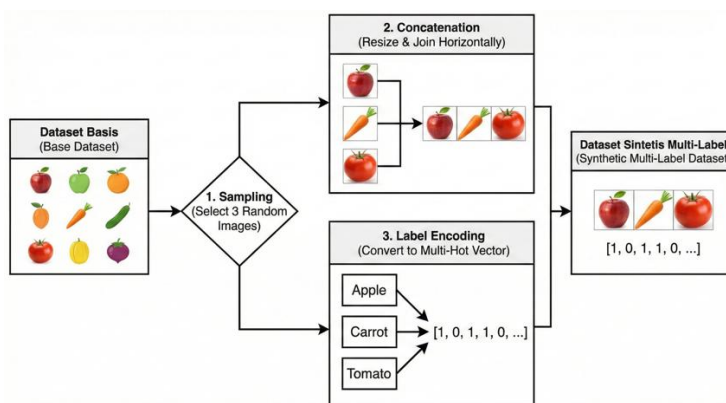
Eksperimen ini menggunakan vegetable-image-dataset sebagai sumber data utama. Dataset ini adalah dataset single-label. Untuk mengubahnya menjadi dataset multi-label, sebuah metodologi inovatif diterapkan. Sebuah kelas dataset kustom (*SyntheticMultiLabelDataset*) dibuat untuk menghasilkan gambar-gambar komposit dengan menggabungkan tiga gambar dasar (single-label) yang dipilih secara acak dari dataset asli ($IMAGES_PER_COMPOSITE = 3$). Proses ini menghasilkan 5000 gambar komposit sintetis untuk pelatihan dan 1000 gambar untuk pengujian. Citra-citra ini digabungkan secara horizontal untuk menciptakan satu citra baru yang mengandung tiga objek berbeda dengan label multi-hot vector [15].



Gambar 3. Sampel representatif dari dataset basis (Vegetable Image Dataset)

Sebelum dilakukan proses sintesis menjadi data multi-label, analisis awal dilakukan terhadap karakteristik visual dataset basis. Seperti yang divisualisasikan pada Gambar 3, dataset ini memiliki kompleksitas fitur yang signifikan yang menjadi tantangan tersendiri bagi model ekstraksi fitur. Pertama, terdapat variasi tekstur permukaan yang ekstrem, mulai dari permukaan halus pada Kentang dan Terong hingga tekstur bergelombang dan kasar pada Pare (Bitter Melon) [1]. Kedua, dataset ini mencakup variasi kondisi pengambilan gambar, di mana beberapa objek difoto secara individual dengan latar belakang bersih, sementara objek lain difoto dalam tumpukan (clutter) atau digenggam tangan manusia.

Keberadaan noise latar belakang dan variasi orientasi objek ini menuntut arsitektur CNN yang robust [16]. Pada penelitian ini, ResNet50 bertugas mengekstrak fitur-fitur invarian dari citra tunggal tersebut sebelum mempelajari korelasi antar-objek pada tahap multi-label [11]. Penggunaan data basis dengan variabilitas tinggi ini menjamin bahwa model sintetis yang dihasilkan tidak hanya menghafal pola posisi, melainkan benar-benar mengenali fitur morfologis dari setiap jenis sayuran. Dikarenakan keterbatasan dataset publik yang menyediakan anotasi multi-label untuk buah dan sayuran, penelitian ini mengembangkan algoritma pembangkit data sintetis menggunakan basis data Vegetable Image Dataset (Single-Label).



Gambar 1 Diagram alir pembuatan dataset sintesis

Gambar 1 mengilustrasikan alur kerja pembuatan dataset sintesis. Proses dimulai dengan dataset sumber single-label (Vegetable Image Dataset). Algoritma memilih tiga citra secara acak ($IMAGES_PER_COMPOSITE = 3$) beserta label aslinya. Citra-citra tersebut kemudian diubah ukurannya dan digabungkan secara horizontal untuk membentuk satu citra komposit baru yang memuat beberapa objek. Secara bersamaan, label-label asli dikonversi menjadi satu vektor target multi-hot di mana nilai '1' menandakan keberadaan kelas objek dalam citra komposit tersebut. Akhirnya, teknik augmentasi data (seperti rotasi dan penyesuaian warna) diterapkan untuk meningkatkan variabilitas data latih sebelum dinormalisasi menjadi tensor [15]. Seperti yang diilustrasikan pada Gambar 1, proses pembuatan dataset melibatkan langkah-langkah berikut:

1. Sampling yaitu Algoritma memilih 3 citra acak ($IMAGES_PER_COMPOSITE = 3$) dari kelas yang berbeda dalam dataset basis.
2. Concatenation yaitu Ketiga citra tersebut diubah ukurannya dan digabungkan secara horizontal untuk membentuk satu citra komposit baru.
3. Label Encoding yaitu Label dari ketiga citra sumber dikonversi menjadi satu vektor multi-hot encoding sepanjang 15 kelas ($num_classes = 15$).

Proses ini menghasilkan total 4.250 citra latih, 750 citra validasi, dan 1.000 citra uji. Pendekatan ini memungkinkan simulasi skenario dunia nyata di mana berbagai objek muncul berdampingan, sekaligus mengatasi masalah ketersediaan data.

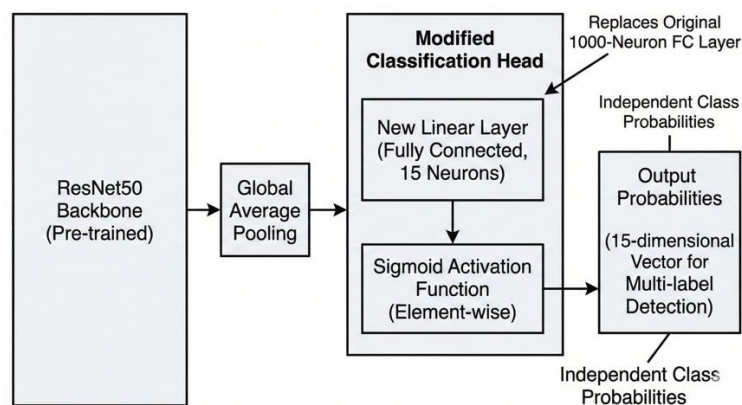
2.2 Pra-pemrosesan Data dan Augmentasi

Sebelum dimasukkan ke dalam jaringan, seluruh citra diubah ukurannya menjadi dimensi standar 224×224 piksel. Untuk meningkatkan kemampuan generalisasi model dan mencegah overfitting, teknik augmentasi data diterapkan secara dinamis pada saat pelatihan (on-the-fly), meliputi:

1. Random Horizontal Flip untuk variasi orientasi.
2. Random Rotation sebesar 15 derajat.
3. Color Jitter (kecerahan, kontras, saturasi) untuk simulasi kondisi pencahayaan yang berbeda [15]

2.3 Arsitektur Model dan Implementasi Loss Function

Arsitektur model didasarkan pada ResNet50 (Residual Network) yang telah dipra-latih pada dataset ImageNet. Pemilihan ResNet50 didasarkan pada arsitektur residual block yang efektif mengatasi masalah vanishing gradient pada jaringan dalam [2], [1]



Gambar 2 Arsitektur model

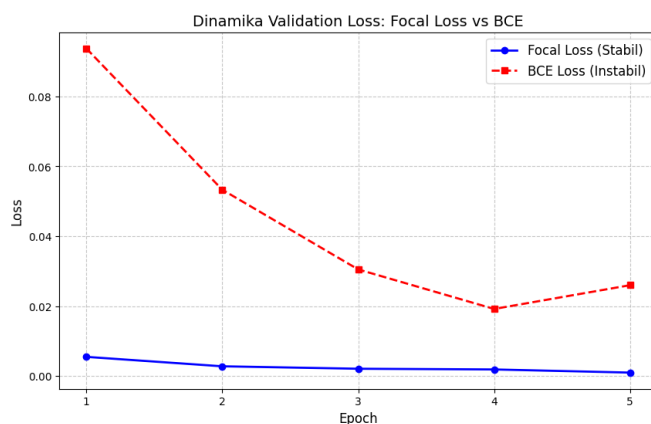
Gambar 2 menunjukkan arsitektur Convolutional Neural Network yang digunakan. Penelitian ini mengadopsi tulang punggung (backbone) ResNet50 yang telah dilatih sebelumnya (pretrained) untuk mengekstrak fitur spasial dari citra komposit berukuran 224×224 piksel. Modifikasi kunci dilakukan pada bagian akhir jaringan ("Head"). Lapisan fully connected (FC) asli yang dirancang untuk 1000 kelas ImageNet diganti dengan lapisan linear baru (nn.Linear). Lapisan baru ini memiliki jumlah neuron keluaran yang setara dengan jumlah kelas unik dalam dataset sintesis ($num_classes = 15$). Berbeda dengan klasifikasi single-label yang menggunakan Softmax, keluaran dari lapisan ini diproses menggunakan fungsi aktivasi Sigmoid secara independen untuk setiap kelas, menghasilkan probabilitas kehadiran masing-masing label di antara 0 dan 1, yang sesuai untuk tugas multi-label [11]. Modifikasi struktural dilakukan pada lapisan terakhir (Fully Connected Layer), di mana lapisan klasifikasi asli (1000 neuron) diganti dengan lapisan linear baru berukuran 15 neuron, sesuai jumlah kelas sayuran dalam dataset. Fungsi aktivasi Sigmoid diterapkan pada setiap neuron output untuk menghasilkan probabilitas independen bagi setiap kelas, memungkinkan deteksi multi-label.

Skenario Pelatihan Eksperimen dilakukan dalam dua skenario terpisah: (1) Pelatihan menggunakan Focal Loss, dan (2) Pelatihan menggunakan Binary Cross-Entropy (BCE) sebagai baseline. Kedua model dilatih menggunakan pengoptimal Adam dengan laju pembelajaran (learning rate) 0.001 dan ukuran batch 32 selama 5 epoch. Lingkungan komputasi menggunakan akselerasi GPU (CUDA) untuk mempercepat proses propagasi balik.

3. HASIL DAN PEMBAHASAN

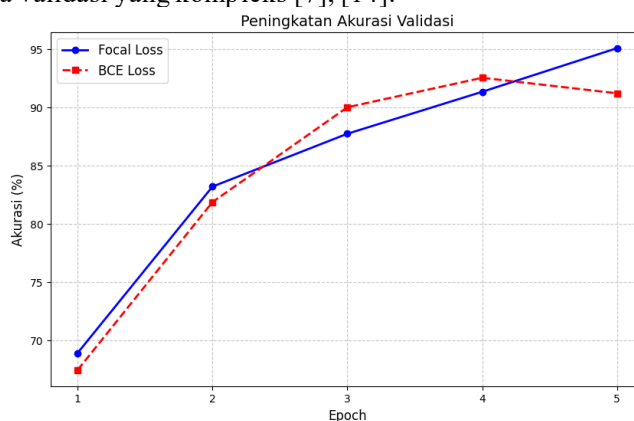
3.1 Dinamika Pelatihan dan Konvergensi

Model Evaluasi proses pelatihan dilakukan dengan memantau metrik loss dan akurasi pada set validasi di setiap epoch. Grafik riwayat pelatihan menunjukkan perbedaan perilaku konvergensi yang signifikan antara kedua fungsi loss.



Gambar 3 Kurva Dinamika Validation Loss

Seperti yang diilustrasikan pada Gambar 3, model yang dilatih dengan Focal Loss (garis biru) menunjukkan penurunan nilai loss yang sangat cepat dan stabil, mendekati nol sejak awal pelatihan. Stabilitas ini bertahan hingga Epoch 5, menandakan bahwa model secara konsisten memperbaiki bobotnya tanpa mengalami fluktuasi besar. Sebaliknya, model dengan BCE Loss (garis merah putus-putus) menunjukkan perilaku yang mengkhawatirkan pada tahap akhir. Meskipun loss menurun hingga Epoch 4 (titik terendah 0.0192), terjadi lonjakan kembali (rebound) yang signifikan pada Epoch 5 menjadi 0.0260. Fenomena "kurva berbentuk V" pada akhir pelatihan BCE ini mengindikasikan awal terjadinya overfitting atau kegagalan model dalam menggeneralisasi fitur pada data validasi yang kompleks [7], [14].



Gambar 4 Kurva Peningkatan Akurasi Validasi

Korelasi antara ketidakstabilan loss dan kinerja prediksi terlihat jelas pada Gambar 4. Model Focal Loss menunjukkan tren peningkatan akurasi yang monoton (terus naik), mencapai puncaknya sebesar 95.07% pada Epoch 5. Hal ini membuktikan bahwa mekanisme pemfokusan (γ) pada Focal Loss efektif menjaga gradien tetap informatif hingga akhir pelatihan. Sebaliknya, kurva akurasi model BCE Loss mengkonfirmasi ketidakstabilan yang terdeteksi pada grafik loss. Setelah mencapai akurasi puncak sementara sebesar 92.53% pada Epoch 4, akurasi justru menurun menjadi 91.20% pada Epoch 5. Penurunan ini menunjukkan bahwa dominasi "contoh mudah" (easy examples) mulai membiarkan keputusan model, sehingga mengurangi kemampuannya dalam membedakan kelas sayuran yang ambigu. Berdasarkan kedua visualisasi ini, Focal Loss terbukti memberikan konvergensi yang lebih robust dan performa akhir yang lebih superior dibandingkan fungsi loss standar.

Berdasarkan data log pelatihan, model yang menggunakan Focal Loss menunjukkan stabilitas yang superior. Pada akhir pelatihan (Epoch 5), Validation Loss Focal Loss mencapai titik terendah yaitu 0.0010 dengan akurasi validasi 95.07%. Kurva loss menurun secara konsisten, menandakan bahwa mekanisme pembobotan pada Focal Loss berhasil menjaga gradien tetap efektif.

Sebaliknya, model BCE Loss menunjukkan gejala instabilitas pada tahap akhir pelatihan. Meskipun mencapai kinerja yang baik pada Epoch 4 (Loss: 0.0192), validation loss justru mengalami peningkatan kembali (rebound) menjadi 0.0260 pada Epoch 5. Fenomena ini mengindikasikan terjadinya overfitting ringan atau kesulitan model BCE dalam menemukan global minima yang stabil ketika dihadapkan pada contoh-contoh yang ambigu di akhir masa pelatihan

4.2 Perbandingan Kinerja pada Data Uji

Pengujian akhir dilakukan pada 1.000 citra uji sintetis yang tidak pernah dilihat model selama pelatihan. Tabel 1 merangkum perbandingan metrik evaluasi antara kedua metode.

Tabel 1. Perbandingan Kinerja Kuantitatif (Test Set)

| Metrik Evaluasi | BCE Loss (Baseline) | Focal Loss (Proposed) | Peningkatan (Improvement) |
|------------------------|----------------------------|------------------------------|----------------------------------|
| Test Loss | 0.0350 | 0.0013 | ↓ 96.2% (Signifikan) |
| Akurasi | 0.8850 (88.50%) | 0.9390 (93.90%) | ↑ 6.10% |
| Precision (Micro) | 0.9795 | 0.9904 | ↑ 1.11% |
| Recall (Micro) | 0.9642 | 0.9823 | ↑ 1.87% |
| F1-Score (Micro) | 0.9718 | 0.9863 | ↑ 1.49% |
| Waktu Konvergensi | Lambat | Cepat & Stabil | - |

Berdasarkan Tabel 1, metode yang diusulkan (Focal Loss) mengungguli metode baseline (BCE Loss) di seluruh metrik pengujian. Peningkatan paling drastis terlihat pada nilai Test Loss, di mana Focal Loss berhasil menekan error hingga 0.0013, jauh lebih rendah dibandingkan BCE (0.0350). Hal ini mengindikasikan bahwa Focal Loss sangat efektif dalam meminimalkan keragu-raguan model pada sampel data yang ambigu. Dari segi Akurasi, terdapat kenaikan margin sebesar 6.10%, yang dalam konteks klasifikasi multi-label dengan 15 kelas adalah peningkatan yang substansial.

Hasil eksperimen menegaskan keunggulan metode yang diusulkan. Model Focal Loss mencapai skor F1 -Micro sebesar 0.9863, mengungguli model BCE yang memperoleh 0.9718. Perbedaan paling mencolok terlihat pada nilai Test Loss, di mana Focal Loss mencatat nilai 0.0013, jauh lebih rendah dibandingkan BCE (0.0350).

Rendahnya nilai loss pada Focal Loss (hampir mendekati nol) membuktikan bahwa model memiliki tingkat keyakinan (confidence) yang sangat tinggi terhadap prediksinya. Hal ini sejalan dengan teori bahwa Focal Loss secara efektif menekan dampak dari easy negatives (background), sehingga model dapat mengalokasikan kapasitas pembelajarannya untuk membedakan fitur-fitur halus antar jenis sayuran dan buah.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan kerangka kerja klasifikasi multi-label untuk buah dan sayuran menggunakan arsitektur ResNet50 yang dimodifikasi. Penggunaan dataset sintetis terbukti menjadi strategi efektif untuk mengatasi keterbatasan data latih di domain ini. Kesimpulan utama dari eksperimen komparatif adalah bahwa Focal Loss secara signifikan lebih unggul dibandingkan Binary Cross-Entropy (BCE) dalam menangani karakteristik data multi-label sintetis. Focal Loss tidak hanya menghasilkan akurasi akhir yang lebih tinggi (93.90% vs 88.50%), tetapi juga menunjukkan stabilitas pelatihan yang lebih baik, mencegah fluktuasi loss yang diamati pada model BCE. Temuan ini merekomendasikan penggunaan Focal Loss sebagai standar loss function untuk pengembangan sistem visi komputer di bidang pertanian cerdas, khususnya pada skenario dengan distribusi kelas yang kompleks.

DAFTAR PUSTAKA

- [1] J.; Zhang *et al.*, “Lightweight and Optimized Multi-Label Fruit Image Classification: A Combined Approach of Knowledge Distillation and Image Enhancement,” *Electronics* 2024, Vol. 13, Page 3267, vol. 13, no. 16, p. 3267, Aug. 2024, doi: 10.3390/ELECTRONICS13163267.
- [2] Y. Wei *et al.*, “HCP: A flexible CNN framework for multi-label image classification,” *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 9, pp. 1901–1907, Sep. 2016, doi: 10.1109/TPAMI.2015.2491929.
- [3] J. Ji, W. Jing, G. Chen, J. Lin, and H. Song, “Multi-Label Remote Sensing Image Classification with Latent Semantic Dependencies,” *Remote Sensing* 2020, Vol. 12, Page 1110, vol. 12, no. 7, p. 1110, Mar. 2020, doi: 10.3390/RS12071110.
- [4] G. Yue, P. Wei, Y. Liu, Y. Luo, J. Du, and T. Wang, “Automated Endoscopic Image Classification via Deep Neural Network With Class Imbalance Loss,” *IEEE Trans Instrum Meas*, vol. 72, 2023, doi: 10.1109/TIM.2023.3264047.
- [5] Y. Liu *et al.*, “Automatic Multi-Label ECG Classification with Category Imbalance and Cost-Sensitive Thresholding” *Biosensors* 2021, Vol. 11, Page 453, vol. 11, no. 11, p. 453, Nov. 2021, doi: 10.3390/BIOS11110453.
- [6] Y. Yao, J. Zhang, P. Zhang, and Y. Sun, “A Dual-branch Learning Model with Gradient-balanced Loss for Long-tailed Multi-label Text Classification,” *ACM Trans Inf Syst*, vol. 42, no. 2, Sep. 2023, doi: 10.1145/3597416.
- [7] J. Du, Y. Zhou, P. Liu, C. M. Vong, and T. Wang, “Parameter-Free Loss for Class-Imbalanced Deep Learning in Image Classification,” *IEEE Trans Neural Netw Learn Syst*, vol. 34, no. 6, pp. 3234–3240, Jun. 2023, doi: 10.1109/TNNLS.2021.3110885.
- [8] A. Pratama, T. Sugiharto, and P. Novantara, “Classification of Avocado Plant Varieties Based on Leaf Shape Using CNN Algorithm,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 6, no. 2, pp. 120–128, Aug. 2025, doi: 10.37859/COSCITECH.V6I2.9474.
- [9] G. Gumelar, T. Sugiharto, and I. Lesmana, “Implentation of CNN for Corn Leaf Disease Identification,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 6, no. 2, pp. 175–180, Aug. 2025, doi: 10.37859/COSCITECH.V6I2.9462.
- [10] Y. Nie, P. Sommella, M. Carratù, M. O’Nils, and J. Lundgren, “A Deep CNN Transformer Hybrid Model for Skin Lesion Classification of Dermoscopic Images Using Focal Loss,” *Diagnostics* 2023, Vol. 13, Page 72, vol. 13, no. 1, p. 72, Dec. 2022, doi: 10.3390/DIAGNOSTICS13010072.
- [11] Z. B. Yu and M. L. Zhang, “Multi-Label Classification with Label-Specific Feature Generation: A Wrapped Approach,” *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 9, pp. 5199–5210, Sep. 2022, doi: 10.1109/TPAMI.2021.3070215.
- [12] C. Liu, J. Wen, Y. Xu, B. Zhang, L. Nie, and M. Zhang, “Reliable Representation Learning for Incomplete Multi-View Missing Multi-Label Classification,” *IEEE Trans Pattern Anal Mach Intell*, vol. 47, no. 6, pp. 4940–4956, 2025, doi: 10.1109/TPAMI.2025.3546356.
- [13] Z. Liu, C. Tang, S. E. Abhadiomhen, X. J. Shen, and Y. Li, “Robust Label and Feature Space Co-Learning for Multi-Label Classification,” *IEEE Trans Knowl Data Eng*, vol. 35, no. 11, pp. 11846–11859, Nov. 2023, doi: 10.1109/TKDE.2022.3232114.
- [14] M. S. Hanif, M. Bilal, A. H. Alsaggaf, and U. M. Al-Saggaf, “Enhancing Multi-Label Chest X-Ray Classification Using an Improved Ranking Loss,” *Bioengineering* 2025, Vol. 12, Page 593, vol. 12, no. 6, p. 593, May 2025, doi: 10.3390/BIOENGINEERING12060593.
- [15] X. Ke, J. Zou, and Y. Niu, “End-to-End Automatic Image Annotation Based on Deep CNN and Multi-Label Data Augmentation,” *IEEE Trans Multimedia*, vol. 21, no. 8, pp. 2093–2106, Aug. 2019, doi: 10.1109/TMM.2019.2895511.
- [16] M. Yanto, A. C. Siregar, and A. Abdullah, “Classification of Cucumber Leaf Diseases Based on Convolutional Neural Network (CNN),” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 6, no. 2, pp. 285–291, Sep. 2025, doi: 10.37859/COSCITECH.V6I2.9982.