

Klasifikasi Citra Diabetic Foot Ulcer Menggunakan ResNet50 dengan Evaluasi Confusion Matrix pada Dataset Validasi

Rhiziqo Adjie Syahputra¹, Wahyu Firmantara², dan Sofyano Fadilah Rosyad³

UPN Veteran Jawa Timur¹, UPN Veteran Jawa Timur², UPN Veteran Jawa Timur³

Rhiziqo Adjie Syahputra.21081010332@student.upnjatim.ac.id

ABSTRACT

Diabetic Foot Ulcer (DFU) is a severe complication of diabetes mellitus that requires early detection to prevent extensive tissue damage and possible amputation. Manual identification is time-consuming and requires clinical expertise, while the limited availability of labeled medical images creates challenges in training deep learning models from scratch. This study implements a Transfer Learning approach using the ResNet50 architecture to automatically classify DFU images into Normal and Abnormal categories. The dataset used in this research is sourced from a public Kaggle repository (Laithjj) and consists of 1,734 images divided into training (70%), validation (20%), and testing (10%) sets. Pre-processing includes resizing, normalization, and data augmentation on the training set. The best validation performance achieved an accuracy of 90.48% and a recall of 98% for the Abnormal class. Although the model performs well on test data, further evaluation on real-world images highlights the presence of domain shift. These findings suggest that ResNet50 is effective for DFU feature extraction, yet broader dataset diversity is required to improve clinical generalization.

Kata Kunci

Deep learning;
Klasifikasi citra;
Luka kaki diabetes;
ResNet50;
Transfer learning.

ABSTRAK

Diabetic Foot Ulcer (DFU) merupakan komplikasi serius pada penderita diabetes yang membutuhkan deteksi dini agar tidak berkembang menjadi infeksi berat hingga amputasi. Identifikasi manual memerlukan waktu dan keahlian klinis, sementara jumlah citra medis berlabel yang terbatas menyulitkan pelatihan model dari awal. Penelitian ini menerapkan pendekatan Transfer Learning menggunakan arsitektur ResNet50 untuk klasifikasi citra DFU ke dalam kelas Normal dan Abnormal. Dataset yang digunakan berasal dari repositori publik Kaggle (Laithjj) dengan total 1.734 citra yang dibagi menjadi data latih (70%), validasi (20%), dan uji (10%). Tahapan pra-pemrosesan meliputi normalisasi, pengubahan ukuran citra, serta augmentasi pada data latih. Model memperoleh akurasi validasi terbaik sebesar 90,48% dengan recall 98% pada kelas Abnormal. Meskipun performa pada data uji cukup tinggi, evaluasi pada citra dunia nyata menunjukkan adanya domain shift. Hasil ini mengindikasikan bahwa ResNet50 efektif sebagai ekstraktor fitur DFU, namun masih membutuhkan data yang lebih beragam untuk mencapai generalisasi klinis yang lebih baik.

PENDAHULUAN

Diabetes melitus merupakan penyakit kronis yang dapat menimbulkan komplikasi serius, salah satunya Luka Kaki Diabetes (Diabetic Foot Ulcer/DFU). Kondisi ini muncul akibat neuropati perifer dan gangguan aliran darah, dan bila tidak ditangani dapat berkembang menjadi infeksi berat hingga amputasi. Keterbatasan tenaga medis serta waktu yang diperlukan dalam proses identifikasi manual mendorong perlunya sistem pendukung diagnosis berbasis kecerdasan buatan.

Ketersediaan dataset DFU yang berlabel masih sangat terbatas, sebagaimana diungkapkan oleh Goyal et al. bahwa dataset publik DFU jauh lebih sedikit dibandingkan bidang medis lainnya [1]. Hal ini menghambat pelatihan model CNN dari awal yang membutuhkan jumlah data besar dan bervariasi [2]. Untuk mengatasi keterbatasan tersebut, pendekatan Transfer Learning menjadi solusi efektif karena memungkinkan model memanfaatkan bobot dari pelatihan sebelumnya pada dataset besar seperti ImageNet.

ResNet50 dipilih dalam penelitian ini karena arsitektunya mampu mengatasi masalah vanishing gradient melalui mekanisme residual connection, serta telah terbukti memberikan hasil

yang kompetitif pada klasifikasi luka diabetes [3]. Penelitian lain juga menunjukkan bahwa model CNN memiliki keunggulan signifikan dalam mengekstraksi fitur spasial dibandingkan pendekatan seperti LSTM atau FNN [5], sementara arsitektur yang lebih efisien seperti MobileNetV2 pun menunjukkan kinerja baik pada kasus luka diabetes [4].

Berdasarkan landasan tersebut, penelitian ini bertujuan untuk mengevaluasi performa ResNet50 pada klasifikasi citra DFU serta menganalisis pengaruh domain shift pada proses generalisasi model.

TINJAUAN PUSTAKA

Luka Kaki Diabetes

Luka kaki diabetes (DFU) adalah komplikasi umum yang dialami pasien diabetes akibat neuropati dan iskemia. Deteksi dini sangat krusial untuk mencegah infeksi lanjut yang berujung pada amputasi. Analisis citra medis menjadi salah satu solusi non-invasif untuk membantu diagnosis ini. Namun, variasi visual luka yang tinggi dan kemiripan dengan kondisi kulit normal tertentu menjadi tantangan tersendiri dalam klasifikasi otomatis.

Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan teknik deep learning efektif untuk analisis citra karena dapat mempelajari pola visual secara hierarkis tanpa ekstraksi fitur manual [2]. Beberapa studi menunjukkan CNN lebih unggul dibanding LSTM atau FNN dalam mengenali pola visual pada citra kulit penderita diabetes [5].

Transfer Learning pada Citra Medis

Transfer Learning digunakan untuk mengatasi keterbatasan dataset medis dengan memanfaatkan bobot model besar seperti ImageNet. Penelitian Widyaya dan Budi menekankan pentingnya preprocessing seperti normalisasi dan augmentasi untuk meningkatkan performa model pada domain medis [7].

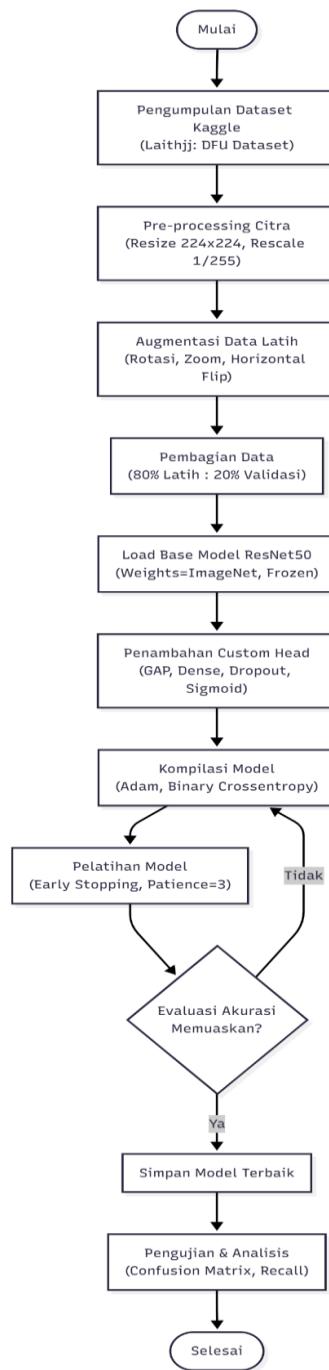
ResNet50 untuk Klasifikasi Luka Diabetes

ResNet50 merupakan arsitektur dengan kedalaman 50 lapisan yang dirancang menggunakan residual blocks untuk menjaga aliran gradien. Tangka et al. menunjukkan bahwa ResNet50 mampu memberikan performa stabil dan kompetitif pada klasifikasi luka diabetes, bahkan dibandingkan model yang lebih modern seperti EfficientNetB1 [3].

METODE

Dataset yang digunakan berasal dari repositori publik Kaggle (Laithjj), yang menyediakan citra kaki dengan dua kategori utama, yaitu Normal dan Abnormal (Ulcer). Total digunakan 1.734 citra, kemudian dibagi menjadi tiga kelompok: 70% untuk pelatihan, 20% untuk validasi, dan 10% untuk pengujian. Struktur dataset mempertahankan pembagian folder asli agar memudahkan proses pemuatannya.

Diagram Alir Penelitian, Tahapan penelitian secara umum ditunjukkan pada Gambar 1, yang memuat alur mulai dari pemuatannya dataset hingga proses evaluasi akhir.



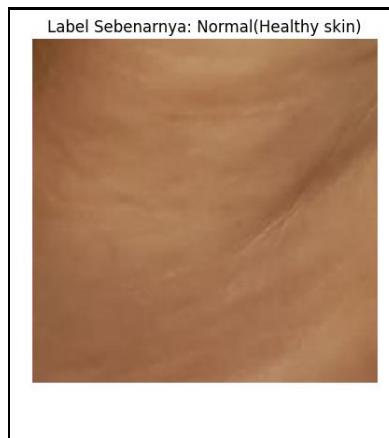
Gambar 1. Diagram alir tahapan klasifikasi luka kaki diabetes menggunakan ResNet50.

Alur penelitian yang ditunjukkan pada Gambar 1 diawali dengan tahap pengumpulan dataset luka kaki diabetes (*Diabetic Foot Ulcer*) dari repositori publik Kaggle (Laithjj). Citra mentah kemudian memasuki tahap *pre-processing* yang meliputi pengubahan ukuran citra (*resize*) menjadi dimensi 224x224 piksel dan normalisasi nilai piksel (*rescaling*) dari rentang [0, 255] menjadi [0, 1] agar sesuai dengan standar input model ResNet50.

Dataset dan Pra-pemrosesan, Dataset diambil dari repositori publik Kaggle (Laithjj), berisi dua kelas citra, yaitu Normal dan Abnormal (Ulcer). Total 1.734 citra digunakan dan dibagi menjadi tiga kelompok: 70% data latih, 20% data validasi, dan 10% data uji. Dataset diambil dari repositori publik Kaggle (Laithjj), berisi dua kelas citra, yaitu Normal dan Abnormal (Ulcer). Total 1.734 citra digunakan dan dibagi menjadi tiga kelompok: 70% data latih, 20% data validasi, dan 10% data uji. Setiap citra diproses agar sesuai dengan kebutuhan model. Tahapannya meliputi:

- Pengubahan ukuran menjadi 224×224 piksel
- Normalisasi intensitas piksel ke rentang 0–1
- Penyesuaian format warna

Augmentasi juga diterapkan pada data latih melalui rotasi, zoom, dan *horizontal flip* untuk meningkatkan variasi citra dan mengurangi overfitting pada dataset medis berukuran kecil. Contoh citra dari kelas Normal dapat dilihat pada Gambar 2.



Gambar 2. Contoh sampel citra dataset pada kelas *Normal (Healthy Skin)*.

Sumber : Dataset publik Kaggle (Laithjj)

Arsitektur Model, Penelitian ini menerapkan Transfer Learning menggunakan arsitektur ResNet50 yang telah dilatih pada dataset ImageNet. Lapisan awal model dibekukan (frozen), kemudian ditambahkan lapisan baru untuk kebutuhan klasifikasi biner:

- GlobalAveragePooling2D
- Dense 256 (ReLU)
- Dropout 0,5
- Dense 1 (Sigmoid)

Aktivasi Sigmoid digunakan pada lapisan output untuk menghasilkan probabilitas kelas biner dan didefinisikan sebagai berikut:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Dimana Z adalah input dari lapisan sebelumnya.

Sedangkan fungsi kerugian yang digunakan adalah Binary Cross-Entropy:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Dimana N adalah jumlah sampel, y_i adalah label sebenarnya, dan \hat{y}_i adalah probabilitas prediksi model.

Skenario Pelatihan, Model dilatih menggunakan optimizer Adam dengan *learning rate* 0,001. Pelatihan direncanakan hingga 15 epoch, namun mekanisme Early Stopping dengan *patience* tiga epoch digunakan untuk menghentikan proses jika akurasi validasi tidak meningkat.

Tabel 1. Konfigurasi Hyperparameter Pelatihan

Parameter	Nilai/Keterangan
Base Model	ResNet50 (ImageNet Weights)
Optimizer	Adam
Learning Rate	0,001
Loss Function	Binary Cross-entropy
Target Epochs	15 (Berhenti otomatis di Epoch 9)
Patience	3 Epochs

Evaluasi Model, Model dievaluasi menggunakan data uji yang tidak terlibat dalam proses pelatihan. Prediksi ditentukan berdasarkan nilai probabilitas dari lapisan output, dengan batas keputusan 0,5. Evaluasi dilakukan menggunakan akurasi, precision, recall, dan F1-score, serta pengamatan terhadap kemungkinan terjadinya domain shift pada citra non-klinis.

HASIL DAN PEMBAHASAN

Pembahasan Data I

Evaluasi awal dilakukan pada 210 citra validasi yang terdiri dari dua kelas: Abnormal(Ulcer) dan Normal(Healthy skin). Hasil pengukuran performa model ditampilkan dalam Tabel 2, yang memuat nilai precision, recall, dan F1-score untuk masing-masing kelas.

Tabel 2. Metrik Evaluasi Model

Kelas	Precision	Recall	F1-Score
Abnormal	0.915	0.961	0.937
Normal	0.961	0.916	0.938

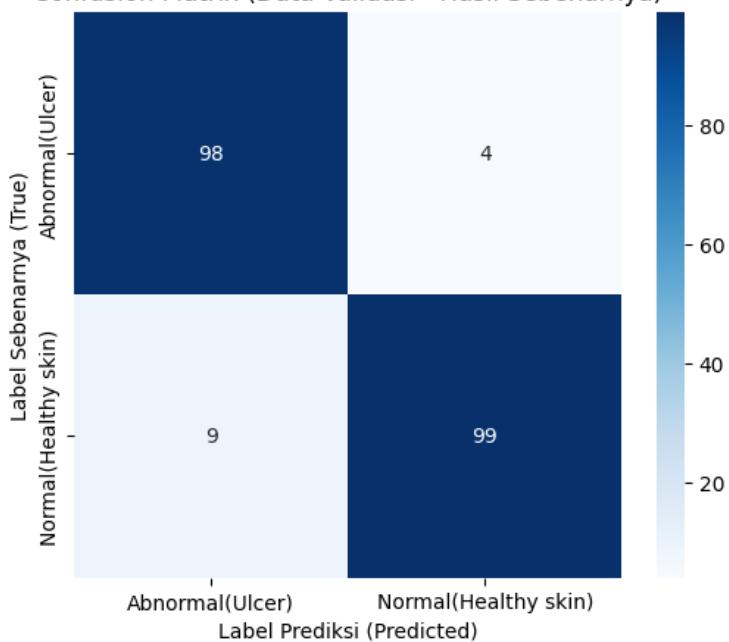
Dari tabel tersebut dapat diamati bahwa kedua kelas memiliki F1-score di atas 0,93. Hal ini menunjukkan bahwa model mencapai keseimbangan yang baik antara ketepatan prediksi dan kemampuan mendeteksi sampel secara benar. Selain itu, nilai precision yang tinggi pada kelas Normal mencerminkan bahwa model jarang memberikan prediksi Normal secara keliru.

Hasil ini mendukung bahwa konfigurasi hyperparameter yang digunakan sebelumnya (Tabel 1) mampu menghasilkan performa stabil pada validasi dan tidak menunjukkan indikasi overfitting yang berlebihan.

Pembahasan Data II

Bagian ini membahas hasil visual dari evaluasi model menggunakan confusion matrix. Visualisasi tersebut memberikan gambaran langsung mengenai distribusi prediksi untuk setiap kelas sehingga memudahkan penulis dalam mengamati pola kesalahan model. Confusion matrix disajikan pada **Gambar 3**.

Confusion Matrix (Data Validasi - Hasil Sebenarnya)



Gambar 3. Confusion matrix hasil klasifikasi pada data validasi.

Dari visualisasi tersebut terlihat bahwa model mampu mengklasifikasikan mayoritas citra Abnormal dengan benar (98 benar dan 4 salah). Sementara itu, pada kelas Normal terdapat 99 prediksi benar dan 9 prediksi salah. Secara keseluruhan, jumlah kesalahan tergolong kecil, sehingga secara praktis model dapat diandalkan pada tahap awal pengenalan luka kulit.

Pengamatan terhadap pola kesalahan menunjukkan bahwa sebagian besar misclassifications terjadi pada citra dengan kondisi pencahayaan yang tidak seragam atau area luka yang tampak samar. Fenomena ini sejalan dengan karakteristik dataset medis yang memiliki keragaman visual cukup tinggi dan dapat mengakibatkan model mengalami kebingungan pada citra yang berada di batas antar kelas.

Hasil ini juga mengindikasikan adanya kemungkinan domain shift, yaitu perbedaan karakteristik antara data pelatihan dan data nyata yang mungkin muncul pada implementasi klinis. Dengan demikian, penggunaan dataset yang lebih variatif akan diperlukan pada tahap pengembangan berikutnya.

KESIMPULAN

Penelitian ini menunjukkan bahwa model ResNet50 berbasis Transfer Learning mampu melakukan klasifikasi citra luka kaki diabetes ke dalam dua kelas (Abnormal dan Normal) dengan performa yang kuat. Evaluasi pada 210 citra validasi menghasilkan akurasi sebesar 94% dengan F1-score sebesar 0.94 pada kedua kelas. Nilai recall pada kelas Abnormal yang mencapai 0.96 menunjukkan bahwa model memiliki sensitivitas tinggi dalam mendeteksi adanya ulkus, sedangkan precision sebesar 0.96 pada kelas Normal menunjukkan kemampuan model dalam memberikan prediksi yang tepat pada citra kulit sehat.

Analisis confusion matrix mengungkapkan bahwa kesalahan prediksi mayoritas terjadi pada citra Normal yang teridentifikasi sebagai Abnormal. Pola ini berkaitan dengan variasi tekstur dan pencahayaan yang menyerupai fitur luka. Walaupun begitu, distribusi kesalahan tetap rendah dan tidak mengganggu performa keseluruhan model.

Hasil penelitian ini mengindikasikan bahwa ResNet50 efektif digunakan sebagai model awal untuk klasifikasi DFU, namun peningkatan keragaman citra masih diperlukan untuk meminimalkan domain shift ketika diterapkan pada citra dunia nyata. Pengembangan lebih lanjut dapat dilakukan dengan menambah variasi data, fine-tuning lebih mendalam, atau memanfaatkan model dengan arsitektur yang lebih efisien.

DAFTAR PUSTAKA

- [1] M. Goyal et al., “DFU2020: A Comprehensive Dataset for Diabetic Foot Ulcer Classification and Localization,” arXiv preprint arXiv:2004.11853, 2020.
- [2] L. Alzubaidi et al., “Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions,” Journal of Big Data, vol. 8, p. 53, 2021.
- [3] G. M. W. Tangka, A. Rantepadang, G. A. Sandag, R. J. Lontaan, and W. G. Mokodaser, “Classification of Diabetic Wounds Using Transfer Learning Model: EfficientNetB1 and ResNet50,” CogITo Smart Journal, vol. 11, no. 1, pp. 1-12, 2025.
- [4] N. Fitriah and S. Sriani, “Classification of Foot Wound Severity in Type 2 Diabetes Mellitus Patients Using MobileNetV2-Based Convolutional Neural Network,” Journal of Applied Intelligent System (JAIC), vol. 9, no. 5, pp. 110-115, 2025.
- [5] S. Arifin, I. Tahyudin, and F. S. Utomo, “Perbandingan Model CNN, LSTM, dan FNN dalam Klasifikasi Kulit Penderita Diabetes,” Jurnal Pendidikan dan Teknologi Indonesia, vol. 5, no. 3, pp. 571-583, 2025.
- [6] A. N. Hidayah, F. Adiputra, and M. Koeshardianto, “Klasifikasi Citra Diabetic Retinopathy Menggunakan CNN dengan Arsitektur VGG-16,” JATI (Jurnal Mahasiswa Teknik Informatika), vol. 9, no. 2, pp. 1308-1315, 2025.
- [7] J. E. Widayaya and S. Budi, “Pengaruh Preprocessing Terhadap Klasifikasi Diabetic Retinopathy dengan Pendekatan Transfer Learning Convolutional Neural Network,” JuTISI (Jurnal Teknik Informatika dan Sistem Informasi), vol. 7, no. 1, pp. 1-10, 2021.