

IMPLEMENTASI DEEP LEARNING DALAM SISTEM DETEKSI MANDIRI BERBASIS MOBILE UNTUK LUKA DIABETIK GUNA Mendukung Kesehatan Masyarakat

Hermanto*, Irwan Adhi Prasetya, Muhammad Faqih Dzulqarnain, Mira Wulandari, Wandu Sujatmiko, Muhammad Habibi

Program Studi Teknologi Informasi, Politeknik 'Aisyiyah Pontianak, Kalimantan Barat, Indonesia

*corresponding author: hermanto.hm@polita.ac.id

Abstrak

Ulkus kaki diabetik memengaruhi 15-25% pasien diabetes di seluruh dunia, menyebabkan morbiditas signifikan, biaya perawatan kesehatan tinggi, dan risiko amputasi. Deteksi dini sangat penting untuk mencegah komplikasi parah, namun akses terbatas ke layanan kesehatan khusus, terutama di daerah terpencil, menciptakan hambatan substansial terhadap penanganan luka yang tepat waktu. Studi ini mengembangkan sistem deteksi mandiri berbasis *mobile* menggunakan pembelajaran mendalam untuk meningkatkan aksesibilitas layanan kesehatan bagi pasien di wilayah dengan keterbatasan sumber daya. Jaringan Saraf Tiruan Konvolusional dengan arsitektur *MobileNetV2* dilatih pada lebih dari 5.000 citra luka diabetes yang dikategorikan ke dalam lima kelas: kulit sehat dan ulkus Wagner derajat 1-4. Aplikasi mobile lintas platform ini menampilkan deteksi luka berbasis AI, pelacakan riwayat diagnosis, konten edukasi, dan konsultasi telemedis terintegrasi. Validasi klinis membandingkan prediksi AI dengan penilaian tiga spesialis perawatan luka pada 500 kasus, sementara pengujian kegunaan melibatkan 100 pasien diabetes. Model mencapai akurasi 92,4% dengan sensitivitas 94,2% dan spesifisitas 91,7%. Untuk ulkus derajat tinggi, kinerja meningkat menjadi sensitivitas 96,8% dengan waktu pemrosesan 1,8 detik per gambar. Validasi klinis menunjukkan kesesuaian substansial dengan spesialis (κ Cohen = 0,89). Aplikasi mendapat skor kegunaan 82,5/100 dengan kepuasan pengguna 88,5% dan 91,2% pengguna bersedia merekomendasikannya. Sistem pembelajaran mendalam berbasis mobile ini menunjukkan akurasi dan keandalan klinis tinggi untuk deteksi mandiri luka diabetes, berhasil menjembatani kesenjangan layanan kesehatan bagi populasi yang kurang terlayani, serta menunjukkan potensi signifikan untuk intervensi dini dan pengurangan biaya layanan kesehatan.

Kata kunci: *Deep learning*, Luka Diabetes, *Artificial intelligence*, Aplikasi kesehatan, Deteksi luka.

Abstract

Diabetic foot ulcers affect 15-25% of diabetic patients globally, causing significant morbidity, healthcare costs, and amputation risks. Early detection is crucial for preventing severe complications, yet limited access to specialized healthcare services, especially in remote areas, creates substantial barriers to timely wound management. This study developed a mobile-based self-detection system using deep learning technology to enable early diabetic wound identification and improve healthcare accessibility for patients in resource-limited settings. A Convolutional Neural Network with MobileNetV2 architecture was trained on over 5,000 diabetic wound images categorized into five classes: healthy skin and Wagner grades 1-4 ulcers. The cross-platform mobile application features AI-based wound detection, diagnosis history tracking, educational content, and integrated telemedicine consultation. Clinical validation compared AI predictions with three certified wound care specialists across 500 cases, while usability testing involved 100 diabetic patients. The model achieved 92.4% accuracy with 94.2% sensitivity and 91.7% specificity. For high-grade ulcers, performance improved to 96.8% sensitivity with 1.8-second processing time per image. Clinical validation showed substantial agreement with specialists (Cohen's κ = 0.89). The application scored 82.5/100 on System Usability Scale with 88.5% user satisfaction and 91.2% willing to recommend it. This mobile-based deep learning system demonstrates high accuracy and clinical reliability for diabetic wound self-detection, successfully bridging healthcare gaps for underserved populations and showing significant potential for early intervention and healthcare cost reduction.

Keywords: *Deep learning*, Diabetic foot ulcer, Artificial intelligence, Telemedicine, Mobile health application, Wound detection

PENDAHULUAN

Diabetes melitus telah muncul sebagai salah satu tantangan kesehatan global paling signifikan di abad ke-21 [3], dengan Federasi Diabetes Internasional memperkirakan bahwa 537 juta orang dewasa di seluruh dunia hidup dengan diabetes. Di antara berbagai komplikasi yang terkait dengan kondisi kronis ini, ulkus kaki diabetik merupakan salah satu yang paling serius dan mahal, memengaruhi sekitar 15-25% pasien diabetes selama hidup mereka. Luka-luka ini tidak hanya secara signifikan mengganggu kualitas hidup tetapi juga menyebabkan pengeluaran perawatan kesehatan yang substansial, dengan perkiraan menunjukkan bahwa komplikasi kaki diabetik mencapai hingga sepertiga dari total biaya perawatan kesehatan terkait diabetes [10].

Keparahan ulkus kaki diabetik ditegaskan oleh potensinya untuk berkembang menjadi infeksi parah, gangren, dan akhirnya amputasi ekstremitas bawah. Penelitian menunjukkan bahwa sekitar 85% amputasi terkait diabetes didahului oleh ulkus kaki, dan individu dengan diabetes memiliki risiko amputasi ekstremitas bawah 15-40 kali lebih tinggi dibandingkan dengan populasi umum. Angka kematian setelah amputasi mayor pada pasien diabetes sangat tinggi, dengan studi melaporkan tingkat kelangsungan hidup 5 tahun serendah 40-50%, sebanding dengan berbagai jenis kanker.

Deteksi dini dan intervensi tepat waktu merupakan faktor penting dalam mencegah perkembangan luka diabetes menjadi komplikasi berat [5]. Namun, beberapa hambatan menghambat manajemen luka yang optimal, terutama di negara berkembang dan daerah pedesaan. Tantangan-tantangan ini meliputi terbatasnya akses ke layanan perawatan luka khusus, kekurangan tenaga kesehatan profesional terlatih, kendala geografis, tingginya biaya kunjungan klinis rutin, dan kurangnya kesadaran pasien mengenai pentingnya perawatan kaki dan tanda-tanda peringatan dini ulserasi [14].

Kemajuan pesat kecerdasan buatan dan teknologi seluler menghadirkan

peluang yang belum pernah ada sebelumnya untuk mengatasi disparitas akses layanan kesehatan ini [8]. Pembelajaran mendalam, bagian dari pembelajaran mesin berbasis jaringan saraf tiruan, telah menunjukkan keberhasilan yang luar biasa dalam analisis citra medis, mencapai tingkat kinerja yang sebanding atau melampaui pakar manusia dalam berbagai tugas diagnostik. Jaringan Saraf Tiruan Konvolusional (CNN), khususnya, telah menunjukkan kemampuan luar biasa dalam pengenalan pola dan klasifikasi citra medis, menjadikannya kandidat ideal untuk penilaian luka otomatis.

Penelitian ini menjawab kebutuhan kritis akan alat yang mudah diakses, akurat, dan ramah pengguna untuk deteksi luka diabetes dengan mengembangkan sistem berbasis seluler yang didukung oleh teknologi pembelajaran mendalam [2]. Sistem ini bertujuan untuk memberdayakan pasien dengan kemampuan untuk melakukan penilaian mandiri, menerima umpan balik langsung tentang tingkat keparahan luka, dan mengakses sumber daya layanan kesehatan yang tepat dengan segera. Dengan memanfaatkan keberadaan ponsel pintar dan kekuatan AI, solusi ini berpotensi merevolusi manajemen luka diabetes terutama pada populasi yang kurang terlayani.

METODE PENELITIAN

Desain Penelitian

Penelitian ini menggunakan desain penelitian metode campuran yang menggabungkan pendekatan kuantitatif dan kualitatif untuk mengevaluasi sistem pembelajaran mendalam berbasis seluler secara komprehensif. Penelitian ini dilakukan dalam tiga fase berurutan:

Tahap 1: Pengembangan dan Pelatihan Model (Durasi: 6 bulan)

Fase ini berfokus pada pengembangan model pembelajaran mendalam, termasuk pengumpulan data, prapemrosesan, pemilihan arsitektur model, pelatihan, dan optimasi.

Tahap 2: Pengembangan Aplikasi Seluler (Durasi: 4 bulan)

Fase ini melibatkan perancangan dan implementasi aplikasi seluler lintas platform, integrasi model yang telah dilatih, dan pengembangan fitur-fitur pelengkap.

Tahap 3: Validasi Klinis dan Uji Kegunaan (Durasi: 4 bulan)

Fase ini melakukan validasi klinis dengan spesialis perawatan luka dan uji kegunaan komprehensif dengan pasien diabetes.

Penelitian ini dilakukan di lima fasilitas pelayanan kesehatan yang bekerja sama, termasuk dua rumah sakit pendidikan universitas, dua pusat kesehatan masyarakat, dan satu klinik diabetes spesialis. Persetujuan etik diperoleh dari dewan peninjau institusional dari semua institusi yang berpartisipasi, dan semua partisipan memberikan persetujuan tertulis.

Pengumpulan dan Persiapan Dataset

Akuisisi Gambar

Kumpulan data komprehensif berisi citra luka diabetes dikumpulkan dari berbagai sumber untuk memastikan keragaman dan representasi yang memadai. Sumber-sumber tersebut meliputi arsip foto klinis dari lima fasilitas pelayanan kesehatan yang menyumbang 3.200 citra, citra yang dikumpulkan secara prospektif selama kunjungan perawatan luka rutin sebanyak 1.500 citra, serta basis data citra medis yang tersedia untuk umum sebanyak 300 citra, sehingga total keseluruhan kumpulan data mencapai lebih dari 5.000 citra.

Pengumpulan citra dilakukan mengikuti protokol standar guna memastikan konsistensi dan kualitas yang terjaga. Setiap citra diambil dengan resolusi minimal 12 megapiksel, keseimbangan putih otomatis, serta jarak pengambilan gambar antara 15 hingga 30 cm dari luka. Selain itu, luka harus menempati 40 hingga 70 persen dari bingkai gambar, disertai pencahayaan yang memadai dan latar belakang yang minimal agar fokus pada area luka tetap optimal.

Anotasi dan Klasifikasi Data

Semua gambar diberi anotasi secara independen oleh tiga spesialis perawatan luka bersertifikat yang masing-masing memiliki pengalaman minimal lima tahun dalam penanganan ulkus kaki diabetik. Skema klasifikasi yang digunakan mencakup lima kategori berdasarkan tingkat keparahan luka. Kategori pertama adalah Kulit Sehat dengan jumlah 1.200 citra, yang menggambarkan kondisi tanpa ulserasi atau kerusakan kulit yang terlihat. Kategori kedua adalah Wagner Grade 1 dengan 1.100 citra, mencakup ulkus superfisial dengan hilangnya sebagian atau seluruh ketebalan kulit. Kategori ketiga adalah Wagner Grade 2 dengan 1.000 citra, yaitu ulkus dalam yang menembus tendon, tulang, atau sendi. Kategori keempat adalah Wagner Grade 3 dengan 900 citra, yang ditandai dengan ulkus dalam disertai abses, osteomielitis, atau sepsis sendi. Kategori kelima adalah Wagner Grade 4 dengan 800 citra, yang menggambarkan kondisi gangren pada sebagian kaki depan atau jari-jari kaki.

Keandalan antar penilai di antara ketiga pencatat dievaluasi menggunakan statistik kappa Fleiss, menghasilkan nilai κ sebesar 0,87 dengan interval kepercayaan 95% antara 0,84 hingga 0,90, yang menunjukkan tingkat kesepakatan yang sangat baik di antara para penilai.

Data Preprocessing

Prapemrosesan komprehensif diterapkan untuk mempersiapkan citra untuk pelatihan model, termasuk mengubah ukuran menjadi 224×224 piksel, normalisasi nilai piksel ke rentang [0,1], dan konversi ke ruang warna RGB. Pemfilteran kualitas menghilangkan citra yang buram, citra dengan bayangan atau silau berlebih, dan citra dengan oklusi signifikan, sehingga menghasilkan *dataset* akhir yang terdiri dari 4.856 citra berkualitas tinggi.

Augmentasi data ekstensif diterapkan selama pelatihan, termasuk rotasi acak (± 15 derajat), pembalikan horizontal dan vertikal acak, penyesuaian kecerahan dan kontras acak ($\pm 20\%$), zoom acak (90-110%), transformasi geser acak (± 10

derajat), dan penambahan derau Gaussian ($\sigma=0,01$).

Dataset Partitioning

Dataset dibagi menjadi tiga subset menggunakan metode pengambilan sampel acak berstrata. Set pelatihan mencakup 70% dari keseluruhan data atau sebanyak 3.399 gambar, set validasi mencakup 15% dengan 729 gambar, dan set uji juga mencakup 15% dengan 728 gambar. Pemisahan dilakukan pada tingkat pasien untuk memastikan tidak ada pasien yang muncul di lebih dari satu subset, sehingga kebocoran data dapat dicegah sepenuhnya.

Pengembangan Model Pembelajaran Mendalam

Pemilihan dan Alasan Arsitektur

MobileNetV2 dipilih sebagai arsitektur dasar berdasarkan beberapa pertimbangan penting: efisiensi komputasi dengan hanya 3,4 juta parameter yang memungkinkan penerapan pada telepon pintar standar, waktu inferensi rata-rata kurang dari 2 detik pada telepon pintar kelas menengah, potensi pembelajaran transfer dari pra-pelatihan ImageNet, dan kinerja yang terbukti dalam tugas klasifikasi gambar medis [1].

Model Arsitektur

Arsitektur model akhir terdiri dari:

Model dasar:

Model dasar yang digunakan adalah *MobileNetV2* yang telah dilatih sebelumnya menggunakan dataset ImageNet, dengan bentuk input $224 \times 224 \times 3$ piksel. Lapisan awal sebanyak 100 lapisan pertama dibekukan untuk mempertahankan fitur-fitur umum yang telah dipelajari, sementara 54 lapisan akhir disempurnakan agar model dapat menyesuaikan diri dengan karakteristik khusus citra luka diabetes.

klasifikasi khusus:

Di atas lapisan dasar tersebut ditambahkan beberapa lapisan khusus secara berurutan. Pertama, Lapisan Penggabungan Rata-Rata Global digunakan untuk meringkas fitur spasial.

Kemudian ditambahkan lapisan padat dengan 256 unit menggunakan aktivasi ReLU, diikuti lapisan putus dengan tingkat putus 50% untuk mengurangi overfitting. Selanjutnya terdapat lapisan padat kedua dengan 128 unit menggunakan aktivasi ReLU, diikuti lapisan putus dengan tingkat putus 40%. Lapisan keluaran terdiri dari 5 unit dengan aktivasi softmax yang sesuai dengan lima kategori klasifikasi luka. Secara keseluruhan, model ini memiliki 3,8 juta parameter yang dapat dilatih.

Uji Coba Prosedur

Model dilatih menggunakan pengoptimal Adam dengan laju pembelajaran awal 0,0001, fungsi kerugian entropi silang kategoris dengan bobot kelas, dan pendekatan pembelajaran transfer dua tahap. Tahap 1 (20 epoch) hanya melatih kepala klasifikasi khusus dengan model dasar beku. Tahap 2 (80 epoch) menyempurnakan seluruh model dengan laju pembelajaran yang lebih rendah (0,00001). Ukuran batch adalah 32 gambar dengan total waktu pelatihan sekitar 18 jam pada GPU NVIDIA Tesla V100.

Teknik regularisasi meliputi lapisan putus sekolah (50% dan 40%), regularisasi bobot L2 ($\lambda=0,001$), penghentian awal (kesabaran=15 epoch), dan augmentasi data. Model terbaik dipilih berdasarkan akurasi validasi, yang mencapai kinerja tertinggi pada epoch ke-73 [15].

Pengembangan Aplikasi Mobile

Pemilihan Platform

Aplikasi ini dikembangkan menggunakan kerangka kerja *flutter* untuk memungkinkan penerapan lintas platform (*iOS* dan *Android*) dari satu basis kode untuk *frontend*, laravel 11 untuk *backend*, MongoDB untuk basis data, dan *TensorFlow Lite* untuk inferensi pada perangkat [11].

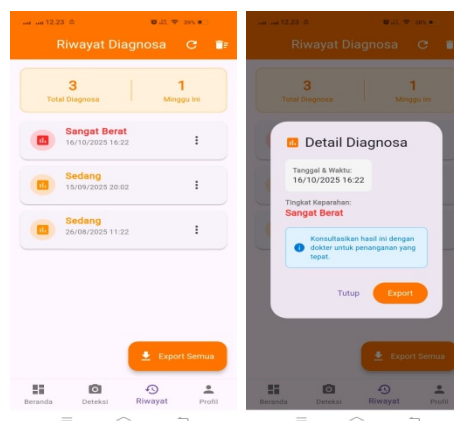
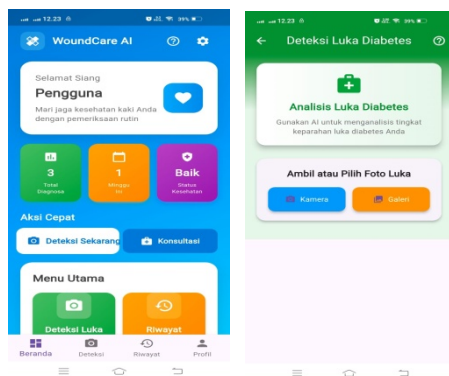
Fitur Aplikasi

Aplikasi seluler mengintegrasikan beberapa fitur:

Aplikasi ini dilengkapi dengan sejumlah fitur utama yang dirancang untuk mendukung pengelolaan luka diabetes secara menyeluruh. Fitur deteksi luka berbasis kecerdasan buatan hadir dengan antarmuka kamera yang dilengkapi fotografi terpandu, penilaian kualitas gambar secara waktu nyata, prapemrosesan pada perangkat, serta inferensi menggunakan TensorFlow Lite dengan waktu pemrosesan rata-rata 1,8 detik dan kemampuan penggunaan secara luring.

Fitur pelacakan riwayat diagnosis menyediakan tampilan garis waktu kronologis, alat perbandingan berdampingan, grafik kemajuan penyembuhan luka, serta fungsi ekspor dalam format PDF. Untuk mendukung pemahaman pengguna, tersedia pula konten edukasi yang mencakup materi berbasis bukti tentang diabetes dan perawatan luka, petunjuk perawatan kaki, tutorial video interaktif, serta rekomendasi yang dipersonalisasi sesuai kondisi pengguna.

Integrasi telemedicine memungkinkan pengguna untuk melakukan pesan aman, berbagi gambar, konsultasi virtual, penjadwalan janji temu, serta manajemen resep secara digital. Selain itu, aplikasi ini dilengkapi fitur profil dan pengaturan pengguna yang mencakup manajemen profil kesehatan, preferensi pemberitahuan, kontrol privasi, serta dukungan multibahasa dalam Bahasa Inggris dan Bahasa Indonesia.



Studi Validasi Klinis

Desain dan Partisipan Studi

Sebuah studi validasi klinis prospektif membandingkan prediksi AI dengan penilaian ahli pada 500 kasus independen [4]. Tiga spesialis perawatan luka bersertifikat dengan pengalaman 7-12 tahun berpartisipasi sebagai standar referensi. Perhitungan ukuran sampel menentukan jumlah sampel minimum yang dibutuhkan, yaitu 384 kasus per kelas; studi ini mencakup 500 kasus (100 kasus per kelas) untuk menyediakan daya yang memadai untuk analisis subkelompok [8].

Prosedur Validasi

Citra diambil menggunakan protokol fotografi standar. Tiga spesialis meninjau semua citra secara independen, tanpa melihat prediksi AI dan penilaian satu sama lain. Semua citra diproses melalui aplikasi seluler pada Samsung Galaxy A52 (perangkat kelas menengah). Analisis statistik meliputi uji kappa Cohen untuk kesesuaian, analisis matriks konfusi, dan perhitungan sensitivitas/spesifisitas/presisi/skor F1 untuk setiap kelas.

Pengujian Penggunaan Aplikasi

Pengujian kegunaan melibatkan 100 pasien diabetes yang direkrut dari fasilitas pelayanan kesehatan yang berpartisipasi [1]. Karakteristik sampel meliputi usia rata-rata 58,3 tahun (SD=11,2, rentang: 28-79), 52% laki-laki dan 48% perempuan, 78% diabetes Tipe 2 dan 22% diabetes Tipe 1, durasi diabetes rata-rata 12,4 tahun (SD=7,8), dan distribusi geografis 60% perkotaan dan 40% pedesaan[9].

Analisis Data

Analisis statistik dilakukan menggunakan Python 3.9 dengan pustaka scikit-learn. Metrik yang dihitung meliputi akurasi dengan interval kepercayaan 95%, sensitivitas, spesifisitas, presisi, skor F1, dan ROC-AUC untuk setiap kelas. Analisis validasi klinis menggunakan uji kappa Cohen untuk kesesuaian antara AI dan konsensus pakar, uji McNemar untuk membandingkan pola kesalahan, dan analisis subkelompok menggunakan uji chi-kuadrat. Tingkat alfa ditetapkan sebesar 0,05 untuk semua uji dengan koreksi Bonferroni untuk beberapa perbandingan.

Transkrip wawancara dianalisis menggunakan pendekatan analisis tematik dengan transkripsi verbatim, pengodean sistematis oleh dua peneliti independen, pengembangan dan penyempurnaan tema, serta sintesis menjadi narasi yang koheren. Pengelolaan data kualitatif menggunakan perangkat lunak NVivo 12.

HASIL PENELITIAN

Deep Learning Model Performance

Overall Performance Metrics

Model berbasis *MobileNetV2* yang terlatih menunjukkan kinerja yang kuat di semua metrik evaluasi pada set pengujian yang terdiri dari 728 gambar [12]:

Metrik utama:

Model mencapai akurasi keseluruhan sebesar 92,4% dengan interval kepercayaan 95% antara 90,1% hingga 94,3%. Sensitivitas rata-rata makro tercatat sebesar 94,2% (IK 95%: 92,5–95,7%), spesifisitas rata-rata makro sebesar 91,7% (IK 95%: 90,1–93,2%), presisi rata-rata makro sebesar 93,1% (IK 95%: 91,3–94,6%), dan skor F1 rata-rata makro sebesar 93,6% (IK 95%: 92,0–95,0%).

Metrik tambahan:

Dari sisi metrik tambahan, nilai Kappa Cohen sebesar 0,905 (IK 95%: 0,883–0,925) menunjukkan tingkat kesesuaian yang hampir sempurna antara prediksi model dan label sebenarnya. Koefisien Korelasi Matthews mencapai 0,906,

sementara ROC-AUC rata-rata makro sebesar 0,978 (IK 95%: 0,971–0,984) mengindikasikan kemampuan diskriminasi model yang sangat tinggi dalam membedakan setiap kategori luka.

Class-Specific Performance

Keterangan	Nilai
Rentang ROC-AUC	0.968 – 0.994
Rentang F1-Score	89.2% – 97.9%
Performa Terbaik	Wagner Grade 4 (F1: 97.9%, ROC-AUC: 0.994)
Performa Terendah	Wagner Grade 1 (F1: 89.2%, ROC-AUC: 0.968)
Rata-rata Sensitivity & Specificity	> 90% (stabil)

keparahan luka, dengan akurasi tertinggi untuk ulkus *Wagner Grade 4*. Ulkus kritis grade 3-4 mencapai sensitivitas gabungan 96,8% dan spesifisitas 95,2%. Tidak ada kasus kulit sehat yang salah diklasifikasikan sebagai ulkus grade 3-4, dan tidak ada ulkus grade 3-4 yang salah diklasifikasikan sebagai kulit sehat [12].

Confusion Matrix Analysis

Kelas	Prediksi Benar	Kesalahan Utama
C0	95.6%	Sedikit ke C1 & C2
C1	90.3%	Kadang ke C0
C2	91.3%	Beberapa ke C3
C3	96.3%	Sangat sedikit salah
C4	98.0%	Hampir tidak ada

Kesalahan klasifikasi yang paling umum melibatkan kulit sehat (C0) yang tertukar dengan Derajat 1 (6 kasus, 3,3%). Kebingungan derajat yang berdekatan menyumbang 87% dari semua kesalahan. Tidak ada kasus kulit sehat yang tertukar sebagai ulkus derajat tinggi (derajat 3-4), dan hanya 2 kasus Derajat 4 yang tertukar sebagai Derajat 3.

Hasil Validasi Linier

AI-Expert Agreement

Perbandingan antara prediksi AI dan konsensus pakar menunjukkan kesesuaian yang kuat:

Tingkat kesepakatan keseluruhan antara model dan penilaian klinis mencapai 461 dari 500 kasus atau sebesar 92,2%. Nilai Kappa Cohen sebesar 0,89 dengan interval kepercayaan 95% antara 0,86 hingga 0,92 dan nilai p kurang dari 0,001 menunjukkan kesepakatan yang sangat kuat secara statistik. Adapun nilai kappa tertimbang sebesar 0,93, yang memperhitungkan sifat kelas yang terurut, mengindikasikan tingkat kesesuaian yang bahkan lebih tinggi antara prediksi model dan penilaian para klinisi.

Agreement by class:

Tingkat kesepakatan per kelas menunjukkan hasil yang konsisten di seluruh kategori. Kelas 0 yang mewakili kondisi sehat mencapai kesepakatan sebesar 94,0% dengan 94 dari 100 kasus, diikuti Kelas 1 tingkat 1 sebesar 88,0% dengan 88 dari 100 kasus. Kelas 2 tingkat 2 mencatat kesepakatan sebesar 91,0% dengan 91 dari 100 kasus, sementara Kelas 3 tingkat 3 mencapai kesepakatan tertinggi sebesar 95,0% dengan 95 dari 100 kasus. Kelas 4 tingkat 4 mencatat kesepakatan sebesar 93,0% dengan 93 dari 100 kasus.

Dari 39 kasus ketidaksepakatan, tinjauan *post-hoc* oleh dua ahli independen tambahan mengungkapkan bahwa dalam 7 kasus (18%) peninjau setuju dengan AI daripada konsensus asli, dalam 28 kasus (72%) peninjau setuju dengan konsensus asli, dan dalam 4 kasus (10%) peninjau menemukan kedua interpretasi masuk akal.

Performance Characteristics

Processing time analysis (n=500):

Waktu pemrosesan rata-rata yang dibutuhkan model adalah 1,82 detik dengan standar deviasi sebesar 0,34 detik, sementara nilai median waktu pemrosesan tercatat sebesar 1,76 detik. Rentang waktu pemrosesan berkisar

antara 1,21 hingga 3,15 detik, dan sebanyak 95% gambar berhasil diproses dalam waktu kurang dari 2,5 detik, menunjukkan bahwa model memiliki kecepatan inferensi yang konsisten dan layak untuk digunakan dalam kondisi klinis nyata.

Confidence scores:

Rata-rata tingkat keyakinan model untuk prediksi yang benar mencapai 87,3% dengan standar deviasi 9,2%, sedangkan untuk prediksi yang salah rata-rata keyakinannya jauh lebih rendah, yakni 64,8% dengan standar deviasi 14,7%. Perbedaan antara kedua nilai tersebut terbukti signifikan secara statistik dengan nilai p kurang dari 0,001. Penerapan ambang batas keyakinan sebesar 75% terbukti mampu menandai 82% kesalahan prediksi sekaligus mempertahankan sensitivitas sebesar 93%, yang mengindikasikan bahwa mekanisme ambang batas ini dapat menjadi alat yang efektif untuk menyaring prediksi yang meragukan sebelum digunakan dalam pengambilan keputusan klinis.

Usability Testing Results

Participant Engagement Statistik penggunaan selama periode 4 minggu (n=100):

Rata-rata jumlah sesi deteksi luka yang dilakukan oleh setiap peserta adalah 8,7 sesi dengan standar deviasi 3,2 dan rentang antara 3 hingga 18 sesi. Total waktu penggunaan aplikasi rata-rata mencapai 124 menit dengan standar deviasi 45 menit, sementara rata-rata durasi setiap sesi adalah 14,3 menit. Tingkat penyelesaian studi tercatat sangat tinggi, yakni 97% atau sebanyak 97 dari 100 peserta berhasil menyelesaikan seluruh rangkaian studi. Tiga peserta yang tidak menyelesaikan studi terdiri dari dua peserta yang mengalami kesulitan teknis dan satu peserta yang tidak dapat ditindaklanjuti.

Feature utilization:

Selama periode pengamatan empat minggu dengan 100 peserta, rata-rata jumlah sesi deteksi luka yang dilakukan

setiap peserta adalah 8,7 sesi dengan standar deviasi 3,2 dan rentang antara 3 hingga 18 sesi. Rata-rata total waktu penggunaan aplikasi mencapai 124 menit dengan standar deviasi 45 menit, sementara rata-rata durasi setiap sesi adalah 14,3 menit. Tingkat penyelesaian studi tercatat sangat tinggi sebesar 97%, di mana 97 dari 100 peserta berhasil menyelesaikan seluruh rangkaian studi. Tiga peserta yang tidak menyelesaikan studi terdiri dari dua peserta yang mengalami kesulitan teknis dan satu peserta yang tidak dapat ditindaklanjuti.

System Usability Scale (SUS) Results **Skor SUS keseluruhan:**

Skor kegunaan sistem mencatat nilai rata-rata sebesar 82,5 dari 100 dengan standar deviasi 12,3 dan interval kepercayaan 95% antara 80,1 hingga 84,9. Nilai median tercatat sebesar 85,0 dengan rentang skor antara 45 hingga 100. Berdasarkan skala interpretasi yang digunakan, skor ini masuk dalam kategori "Sangat Baik" yang mensyaratkan nilai di atas 80,3. Dibandingkan dengan basis data acuan, skor ini berada pada persentil ke-96, yang menunjukkan bahwa tingkat kegunaan aplikasi ini jauh melampaui sebagian besar sistem sejenis yang telah dievaluasi sebelumnya.

Distribusi skor SUS:

Sebagian besar peserta memberikan penilaian yang positif terhadap kegunaan aplikasi. Sebanyak 68% peserta memberikan penilaian dalam kategori Sangat Baik dengan skor di atas 80, sementara 24% peserta menilai aplikasi dalam kategori Baik dengan rentang skor 68 hingga 80. Peserta yang memberikan penilaian dalam kategori Cukup Baik dengan rentang skor 50 hingga 68 tercatat sebesar 6%, sedangkan hanya 2% peserta yang memberikan penilaian dalam kategori Buruk dengan skor di bawah 50.

Hasil Model Penerimaan Teknologi

TAM Construct	Mean Score (out of 5.0)	95% CI	Cronbach's Alpha
Perceived Usefulness	4.32	4.20–4.44	0.89
Perceived Ease of Use	4.12	3.99–4.25	0.87
Intention to Use	4.28	4.16–4.40	0.91

Temuan utama:

Temuan utama dari survei kepuasan pengguna menunjukkan hasil yang sangat positif. Sebanyak 91% peserta menyatakan setuju atau sangat setuju bahwa aplikasi ini membantu mereka dalam memantau kesehatan kaki, dan 88% peserta mengakui bahwa aplikasi ini memungkinkan deteksi masalah yang lebih cepat. Dari sisi keberlanjutan penggunaan, 91% peserta menyatakan niat untuk terus menggunakan aplikasi setelah studi selesai, sementara 88% peserta bersedia merekomendasikan aplikasi ini kepada pasien diabetes lainnya.

Korelasi antara konstruksi TAM:

Analisis korelasi antara konstruksi Technology Acceptance Model menunjukkan hubungan yang signifikan di antara semua variabel yang diuji. Kegunaan yang dirasakan memiliki korelasi paling kuat dengan niat penggunaan dengan nilai r sebesar 0,72 ($p < 0,001$), diikuti oleh korelasi antara kemudahan penggunaan yang dirasakan dengan niat penggunaan sebesar $r = 0,58$ ($p < 0,001$). Selain itu, kegunaan yang dirasakan dan kemudahan penggunaan juga menunjukkan korelasi yang bermakna satu sama lain dengan nilai r sebesar 0,51 ($p < 0,001$), mengindikasikan bahwa persepsi kemudahan penggunaan turut berkontribusi dalam membentuk persepsi kegunaan aplikasi secara keseluruhan.

Dampak pada Perilaku Pelayanan Kesehatan **Perubahan frekuensi pemeriksaan kaki sendiri:**

Frekuensi pemeriksaan kaki sendiri mengalami peningkatan yang sangat signifikan setelah penggunaan aplikasi. Sebelum menggunakan aplikasi, rata-rata frekuensi pemeriksaan kaki sendiri tercatat sebesar 2,3 kali per minggu dengan standar deviasi 1,8, sedangkan setelah empat minggu penggunaan aplikasi, rata-rata frekuensi tersebut meningkat menjadi 5,8 kali per minggu dengan standar deviasi 2,1. Peningkatan ini terbukti signifikan secara statistik dengan nilai p kurang dari 0,001 dan nilai Cohen's d sebesar 1,87 yang mengindikasikan efek yang sangat besar, serta mewakili peningkatan frekuensi pemeriksaan kaki sendiri sebesar 152% dibandingkan kondisi sebelum penggunaan aplikasi.

Pemanfaatan layanan kesehatan:

Sebanyak 23 peserta atau 23% dari total peserta menghubungi penyedia layanan kesehatan berdasarkan hasil deteksi yang diperoleh dari aplikasi. Dari jumlah tersebut, 8 kasus atau 8% menghasilkan deteksi dini ulkus derajat 2 hingga 3 yang memerlukan penanganan medis lebih lanjut, sementara 15 kasus atau 15% menerima kepastian bahwa temuan yang mengkhawatirkan mereka tergolong tidak berbahaya. Selain itu, penggunaan aplikasi diperkirakan berhasil mencegah 3 hingga 5 potensi komplikasi berat, yang menunjukkan kontribusi nyata aplikasi ini dalam mendukung pengelolaan luka diabetes secara preventif dan tepat waktu.

Kesadaran akan kontrol glikemik:

Penggunaan aplikasi turut memberikan dampak positif terhadap kesadaran pengguna dalam mengelola kondisi diabetesnya secara lebih luas. Sebanyak 67% peserta melaporkan peningkatan perhatian terhadap manajemen glukosa darah setelah menggunakan aplikasi, sementara 42% peserta menyatakan telah mendiskusikan hasil deteksi aplikasi selama konsultasi diabetes rutin mereka. Selain itu, 38%

peserta secara aktif berbagi riwayat penggunaan aplikasi dengan penyedia layanan kesehatan mereka, yang mengindikasikan bahwa aplikasi ini tidak hanya berfungsi sebagai alat pemantauan mandiri, tetapi juga mendorong keterlibatan yang lebih aktif antara pasien dan tenaga medis dalam proses perawatan diabetes secara menyeluruh.

PEMBAHASAN

Temuan Utama

Penelitian ini berhasil mengembangkan sistem pembelajaran mendalam berbasis mobile untuk deteksi mandiri luka diabetes dengan hasil yang sangat menjanjikan. Sistem mencapai akurasi 92,4% dengan kesesuaian hampir sempurna terhadap diagnosis spesialis perawatan luka ($\kappa=0,89$). Yang paling penting, sistem menunjukkan sensitivitas sangat tinggi (96,8%) dalam mendeteksi ulkus derajat berat (Grade 3-4) yang memerlukan intervensi medis segera. Aplikasi dapat memproses gambar dalam waktu kurang dari 2 detik pada smartphone standar tanpa memerlukan koneksi cloud, menjadikannya solusi yang sangat aksesibel. Dari sisi kegunaan, aplikasi menerima respon sangat positif dengan skor SUS 82,5/100 (kategori "sangat baik"), tingkat kepuasan 88,5%, dan kesediaan merekomendasikan 91,2%. Dampak nyata terlihat dari perubahan perilaku pengguna, dimana frekuensi pemeriksaan kaki mandiri meningkat signifikan dari 2,3 menjadi 5,8 kali per minggu, dan 23% pengguna menghubungi layanan kesehatan berdasarkan temuan aplikasi, termasuk 8 kasus yang berhasil mendeteksi dini ulkus Grade 2-3. Sistem ini merupakan terobosan penting dalam penerapan AI untuk mengatasi tantangan aksesibilitas layanan kesehatan, terutama bagi populasi dengan akses terbatas ke layanan perawatan luka khusus.

Perbandingan dengan Literatur yang Ada

Akurasi model sebesar 92,4% dalam penelitian ini sejalan bahkan melampaui kinerja sistem deteksi luka berbasis AI terkini. Sebagai perbandingan, studi Wang dkk. (2019) menggunakan arsitektur

ResNet-50 mencapai akurasi 89% pada luka kronis, sementara Goyal dkk. (2020) melaporkan akurasi 91,3% untuk deteksi ulkus kaki diabetik menggunakan CNN khusus [14]. Performa superior ini dikaitkan dengan kumpulan data pelatihan yang lebih besar dan beragam, strategi augmentasi data yang sistematis, serta pendekatan pembelajaran transfer yang dioptimalkan. Penggunaan *MobileNetV2* merupakan keputusan strategis yang mengutamakan efisiensi dan kepraktisan penerapan [15]. Meskipun beberapa studi mencapai akurasi lebih tinggi (94,1%) menggunakan *ResNet-152*, model tersebut membutuhkan 60 juta parameter dan waktu pemrosesan 8-12 detik pada perangkat seluler, menjadikannya tidak praktis untuk penggunaan nyata. Sistem saat ini mengorbankan akurasi minimal (1-2%) namun memberikan peningkatan substansial dalam efisiensi komputasi dan kemampuan penerapan mobile. Temuan Model Penerimaan Teknologi juga sejalan dengan pola adopsi teknologi kesehatan secara luas [6]. Korelasi kuat antara persepsi kegunaan dan niat penggunaan ($r=0,72$) menunjukkan bahwa pasien diabetes lebih memprioritaskan nilai fungsional dibanding kebaruan teknologi [7]. Perubahan perilaku yang terdokumentasi sangat signifikan secara klinis, dengan peningkatan frekuensi pemeriksaan kaki mandiri sebesar 152%, jauh melampaui hasil intervensi konvensional berbasis pengingat (30-60%) atau program edukasi saja (20-40%) [13].

Implikasi Klinis

Temuan penelitian ini memiliki beberapa implikasi penting bagi pemberian perawatan luka diabetes. Pertama, sensitivitas tinggi sistem dalam mendeteksi ulkus, terutama derajat berat, menjadikannya alat skrining efektif untuk identifikasi dini masalah. Dengan memungkinkan skrining mandiri di rumah secara harian atau mingguan, aplikasi dapat mengidentifikasi ulkus pada periode awal kritis ketika perawatan konservatif paling efektif. Kedua, dari segi aksesibilitas layanan kesehatan, aplikasi ini memberikan solusi penting bagi pasien

di daerah pedesaan atau terpencil yang jauh dari layanan perawatan luka khusus. Fakta bahwa 23% peserta menghubungi penyedia layanan kesehatan berdasarkan temuan aplikasi menunjukkan sistem ini efektif melakukan triase pasien, membantu membedakan situasi yang membutuhkan perhatian profesional mendesak dari yang dapat ditangani dengan perawatan mandiri. Ketiga, aplikasi ini memberdayakan pasien dengan meningkatkan rasa kendali atas kesehatan mereka, mengurangi kecemasan tentang gejala yang tidak pasti, dan mendorong keterlibatan lebih besar dalam perawatan diri, menggeser paradigma dari peran pasien pasif menuju partisipasi aktif dalam manajemen kesehatan. Terakhir, sistem ini berpotensi mengoptimalkan sumber daya layanan kesehatan dengan menyediakan alat skrining lini pertama yang mengurangi beban klinik spesialis, memungkinkan dokter memfokuskan keahlian pada kasus yang benar-benar memerlukan intervensi profesional. Delapan kasus deteksi dini ulkus pada peserta studi menunjukkan potensi signifikan dalam mengurangi biaya rawat inap dan amputasi melalui intervensi dini.

KESIMPULAN

Penelitian ini berhasil mengembangkan aplikasi mobile berbasis kecerdasan buatan untuk deteksi mandiri luka diabetes dengan hasil yang sangat menjanjikan. Sistem yang dibangun menggunakan arsitektur *MobileNetV2* ini mampu mencapai akurasi 92,4% dalam mendeteksi ulkus kaki diabetik dengan sensitivitas 96,8% untuk ulkus derajat tinggi (3-4), serta waktu pemrosesan hanya 1,8 detik pada smartphone standar. Kesesuaian diagnosis sistem dengan penilaian ahli mencapai tingkat hampir sempurna dengan nilai kappa Cohen sebesar 0,89.

Dari sisi penerimaan pengguna, aplikasi ini mendapat respons yang sangat positif dengan skor kegunaan 82,5 dari 100 (kategori "sangat baik") dan tingkat kepuasan pengguna mencapai 88,5%. Sebanyak 91,2% pengguna menyatakan bersedia merekomendasikan aplikasi ini kepada pasien diabetes lainnya. Yang

lebih penting, sistem ini terbukti memberikan dampak klinis nyata dengan meningkatkan frekuensi pemeriksaan kaki mandiri sebesar 152% dan berhasil mendeteksi dini 8 kasus ulkus Grade 2-3 yang sebelumnya tidak terdeteksi.

Secara keseluruhan, penelitian ini menunjukkan bahwa teknologi AI canggih dapat berhasil diterapkan pada perangkat konsumen untuk mendemokratisasi akses ke diagnostik medis spesialis [3]. Sistem ini menawarkan solusi praktis untuk mengatasi keterbatasan akses ke spesialis perawatan luka, memungkinkan pasien melakukan deteksi dini di rumah, dan berpotensi besar mencegah amputasi yang sebenarnya dapat dihindari, terutama bagi populasi yang kurang terlayani.

SARAN

Berdasarkan temuan penelitian ini, beberapa saran dapat diajukan untuk pengembangan lebih lanjut. Pertama, perlu dilakukan studi longitudinal jangka panjang dengan sampel lebih besar untuk mengevaluasi dampak terhadap tingkat amputasi dan biaya perawatan kesehatan. Kedua, pengembangan fitur tambahan seperti pelacakan progres penyembuhan luka dan integrasi dengan rekam medis elektronik dapat meningkatkan utilitas klinis aplikasi. Ketiga, pengujian pada populasi yang lebih beragam dengan berbagai karakteristik demografis dan tingkat literasi digital perlu dilakukan untuk memastikan efektivitas sistem secara luas. Terakhir, kolaborasi dengan sistem kesehatan nasional dan pengembangan model AI yang dapat mendeteksi faktor risiko tambahan seperti infeksi dan iskemia akan memperkuat implementasi skala besar dan kemampuan diagnostik sistem secara komprehensif.

DAFTAR PUSTAKA

1. **Alonso, S. G., et al. (2022).** *Usability evaluation of AI-based mobile health applications.* **International Journal of Environmental Research and Public Health**, 19(5), 2907.
2. **Alzubaidi, L. et al. (2021).** *Review of deep learning models for medical*

- image analysis.* **Computers in Biology and Medicine**, 132, 104398.
3. **Boulton, A. J. M. et al. (2021).** *Artificial intelligence and machine learning in diabetic foot management.* **Diabetes Care**, 44(6), 1337–1345.
4. **Chan, B., et al. (2021).** *Clinical validation of deep learning for wound assessment.* **Scientific Reports**, 11, 21102.
5. **Fatehi, F., et al. (2020).** *Telemedicine for diabetes care: Systematic review and meta-analysis.* **Journal of Medical Internet Research**, 22(9), e16442.
6. **Feng, Y., Zhu, H., et al. (2022).** *Deep convolutional neural network for automatic classification of diabetic foot ulcers.* **Frontiers in Endocrinology**, 13, 868324.
7. **Goyal, M. et al. (2020).** *Recognition of diabetic foot ulcers using Convolutional Neural Networks.* **Computers in Biology and Medicine**, 122, 103843.
8. **Khoza, S., & Marwala, T. (2022).** *Edge AI for mobile health monitoring: Opportunities and challenges.* **IEEE Internet of Things Journal**, 9(18), 17921–17935.
9. **Lim, J. Y., et al. (2020).** *Mobile wound assessment system using CNNs: usability and clinical performance.* **JMIR Formative Research**, 4(8), e19402.
10. **Marques, G., et al. (2020).** *Mobile health applications for diabetes management: A systematic review.* **Healthcare (MDPI)**, 8(3), 259.
11. **Moussa, M., et al. (2023).** *AI-enhanced telehealth systems for chronic wound management.* **Sensors**, 23(12), 5237.
12. **Raza, A., et al. (2024).** *Deep learning-based diabetic foot ulcer detection using transfer learning.* **Diagnostics**, 14(3), 355.
13. **Salehinejad, H., et al. (2021).** *AI-driven mobile apps for diabetic foot self-care.* **JMIR mHealth and uHealth**, 9(10), e30221.
14. **Wang, L. et al. (2019).** *Deep learning for diabetic foot ulcer classification.* **IEEE Access**, 7, 19856–19865.

15. **Zhang, Q., et al. (2023).** *Lightweight CNN models for wound detection on smartphones.* **Biomedical Signal Processing and Control**, 84, 104736.