



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 24982-24989

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Komparasi Arsitektur MobileNetV2 dan ResNet50 dalam Klasifikasi Kondisi Permukaan Jalan Raya

Satria Panca Nugroho¹, Dhafi Rathan Rabbani², Fahru Rizal Al Mufariz³, Abdul Ghani Firmansyah⁴ Chaerur Rozikin⁵

Program Studi Informatika, Universitas Singaperbangsa Karawang

2310631170152@student.unsika.ac.id, 2310631170074@student.unsika.ac.id, 2310631170080@student.unsika.ac.id,

2310631170002@student.unsika.ac.id, chaerur.rozikin@staff.unsika.ac.id

Abstrak

Kondisi permukaan jalan yang rusak dapat meningkatkan risiko kecelakaan, menurunkan kenyamanan berkendara, dan memperbesar biaya pemeliharaan kendaraan maupun infrastruktur. Penelitian ini bertujuan membandingkan kinerja arsitektur MobileNetV2 dan ResNet50 dalam mengklasifikasikan kondisi permukaan jalan raya menggunakan pendekatan transfer learning. Dataset yang digunakan terdiri atas 802 citra jalan yang dikelompokkan ke dalam tiga kelas, yaitu Jalan Kategori Baik, Jalan Kurang Baik, dan Jalan Rusak. Data dibagi secara terstratifikasi menjadi 70% data latih, 15% data validasi, dan 15% data uji, kemudian seluruh citra diubah ukurannya menjadi 224×224 piksel dan dinormalisasi sesuai kebutuhan masing-masing arsitektur. Pelatihan dilakukan melalui tahap ekstraksi fitur dan fine-tuning dengan pengoptimal Adam, categorical crossentropy, Early Stopping, dan ReduceLROnPlateau. Evaluasi menggunakan akurasi, presisi, recall, F1-score, confusion matrix, serta pengujian citra jalan di luar dataset. Hasil pengujian menunjukkan ResNet50 memperoleh skor 100% pada data uji, sedangkan MobileNetV2 mencapai akurasi 67,8%, presisi 79,4%, recall 67,8%, dan F1-score 61,9%. Namun, pengujian pada citra baru memperlihatkan bahwa ResNet50 mengalami overfitting dan menghasilkan prediksi yang tidak sesuai, sementara MobileNetV2 memberikan prediksi yang lebih representatif. Temuan ini menunjukkan bahwa arsitektur ringan lebih sesuai untuk dataset terbatas dan berpotensi diterapkan pada sistem pemantauan jalan dengan sumber daya komputasi terbatas. Secara praktis, hasil penelitian dapat menjadi dasar pemilihan model yang mempertimbangkan keseimbangan antara ketepatan klasifikasi, kemampuan generalisasi, dan efisiensi komputasi.

Kata kunci: Klasifikasi Jalan, MobileNetV2, ResNet50, Deep Learning, Convolutional Neural Network

1. Pendahuluan

Infrastruktur jalan merupakan tulang punggung mobilitas masyarakat dan aktivitas ekonomi, terutama di kawasan perkotaan dan industri. Namun kondisi jalan tidak selalu terjaga dengan baik — kerusakan berupa retak, lubang, dan permukaan yang tidak rata kerap terjadi akibat beban lalu lintas yang tinggi, cuaca ekstrem, serta keterbatasan anggaran pemeliharaan. Kerusakan yang tidak segera ditangani dapat berdampak serius, mulai dari peningkatan risiko kecelakaan, kerusakan kendaraan, hingga kerugian ekonomi akibat gangguan distribusi barang.

Selama ini, pemantauan kondisi jalan masih banyak dilakukan secara manual oleh petugas lapangan, yang tentu saja memakan waktu, biaya, dan tenaga yang tidak sedikit. Cakupan inspeksi pun menjadi terbatas, sehingga kerusakan di ruas-ruas tertentu sering kali baru terdeteksi setelah kondisinya sudah cukup parah. Oleh karena itu, dibutuhkan pendekatan yang lebih otomatis dan skalabel untuk memantau serta mengklasifikasikan kondisi permukaan jalan secara efisien.

Perkembangan kecerdasan buatan, khususnya dalam bidang computer vision, membuka peluang besar untuk mengatasi tantangan tersebut. Salah satu pendekatan yang telah banyak digunakan adalah deep learning berbasis Convolutional Neural Network (CNN), yang terbukti mampu mengenali pola visual secara akurat dari data citra. Dua arsitektur CNN yang cukup populer dan sering dipertimbangkan untuk aplikasi serupa adalah MobileNetV2 dan ResNet50. MobileNetV2 dirancang khusus untuk efisiensi komputasi sehingga cocok dijalankan pada perangkat dengan sumber daya terbatas, sedangkan ResNet50 dikenal dengan kedalaman arsitekturnya yang memungkinkan ekstraksi fitur lebih kaya dan akurasi yang lebih tinggi.

Meski keduanya sudah banyak diuji pada berbagai tugas klasifikasi gambar umum, perbandingan langsung antara MobileNetV2 dan ResNet50 secara spesifik pada konteks kondisi permukaan jalan masih terbatas. Padahal, pemilihan arsitektur yang tepat sangat menentukan keberhasilan implementasi sistem pemantauan jalan di dunia nyata. Penelitian ini hadir untuk mengisi celah tersebut dengan melakukan analisis komparatif antara kedua model berdasarkan metrik performa yang relevan, sehingga hasilnya dapat menjadi referensi praktis bagi pengembang sistem maupun pengambil kebijakan infrastruktur.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan eksperimental komputasional yang terdiri dari tahapan pengumpulan data, prapemrosesan (resize dan normalisasi piksel), penyeimbangan distribusi data (stratified split), pelatihan model (transfer learning), dan komparasi arsitektur (MobileNetV2 dan ResNet50).

2.1. Pengumpulan Data (Data Collection)

Data yang digunakan dalam penelitian ini adalah data sekunder berupa "Road Surface Dataset (DatasetJalan)" yang bersumber dari repositori publik Kaggle. Dataset ini memuat 802 citra digital permukaan jalan raya yang dikategorikan ke dalam tiga kelas kondisi, yaitu: Jalan Kategori Baik (251 citra), Jalan Kurang Baik (250 citra), dan Jalan Rusak (301 citra).

2.2. Praproses Data dan Pembagian Dataset (Preprocessing & Data Splitting)

Untuk memastikan distribusi kelas yang seimbang pada setiap subset, pembagian data (data splitting) dilakukan menggunakan metode Stratified Shuffle Split. Dataset dibagi dengan rasio 70% sebagai data latih (560 citra), 15% sebagai data validasi (121 citra), dan 15% sebagai data uji (121 citra). Seluruh citra kemudian diseragamkan ukurannya (resize) menjadi dimensi 224x224 piksel. Tahap normalisasi piksel dilakukan menggunakan fungsi `preprocess_input` bawaan Keras yang spesifik untuk setiap arsitektur; MobileNetV2 melakukan penskalaan piksel ke rentang $[-1, 1]$, sementara ResNet50 mengimplementasikan pertukaran saluran (BGR channel swap) dan pemusatan nilai rata-rata (mean subtraction). Pada rancangan ini, pipeline data digunakan dalam bentuk aslinya tanpa penambahan augmentasi citra.

2.3. Arsitektur Model (Model Architecture)

Penelitian ini menggunakan teknik Transfer Learning dengan membandingkan dua arsitektur base model yang memuat bobot pralatih (pre-trained weights) ImageNet, yakni MobileNetV2 dan ResNet50. Pada bagian atas base model tersebut, ditambahkan lapisan klasifikasi (classification head) buatan yang terdiri dari lapisan Global Average Pooling 2D, Batch Normalization, serta dua lapisan tersembunyi berukuran 256 dan 128 neuron dengan fungsi aktivasi ReLU. Sebagai keluaran akhir, digunakan lapisan Dense dengan 3 neuron menggunakan fungsi aktivasi Softmax. Guna mencegah terjadinya overfitting, diterapkan lapisan Dropout dengan parameter yang disesuaikan terhadap kompleksitas beban jaringan: 0,4 dan 0,3 untuk MobileNetV2, serta 0,6 dan 0,5 untuk ResNet50.

2.4. Skenario Pelatihan (Training Scenario)

Kedua model dikonfigurasi menggunakan fungsi kerugian categorical crossentropy dan pengoptimal Adam dengan learning rate awal sebesar 0,0001. Pelatihan dilakukan dalam dua tahap utama. Tahap pertama, Feature Extraction, dijalankan selama 10 epoch di mana seluruh lapisan pada base model dibekukan (freeze). Tahap kedua, Fine-tuning, dijalankan selama 20 epoch dengan membuka kunci pembelajaran (unfreeze) pada 30 lapisan terakhir untuk MobileNetV2, dan 10 lapisan terakhir untuk ResNet50. Pada tahap ini, learning rate diturunkan menjadi sepersepuluhnya (LR/10). Proses pelatihan juga didukung dengan metode callbacks seperti Early Stopping (patience = 5) dan ReduceLROnPlateau (factor = 0,5, patience = 3).

2.5. Evaluasi Model (Model Evaluation)

Kinerja model dalam melakukan klasifikasi dievaluasi menggunakan Confusion Matrix dan Laporan Klasifikasi (Classification Report) pada subset data uji. Metrik komparasi yang digunakan mencakup Akurasi, Presisi, Recall,

dan F1-Score. Selain itu, dilakukan pula skenario uji coba dunia nyata (real-world testing) menggunakan citra kondisi jalan baru di luar dataset penelitian untuk mengukur tingkat kepercayaan prediksi (confidence score) dan daya generalisasi model di skenario praktis.

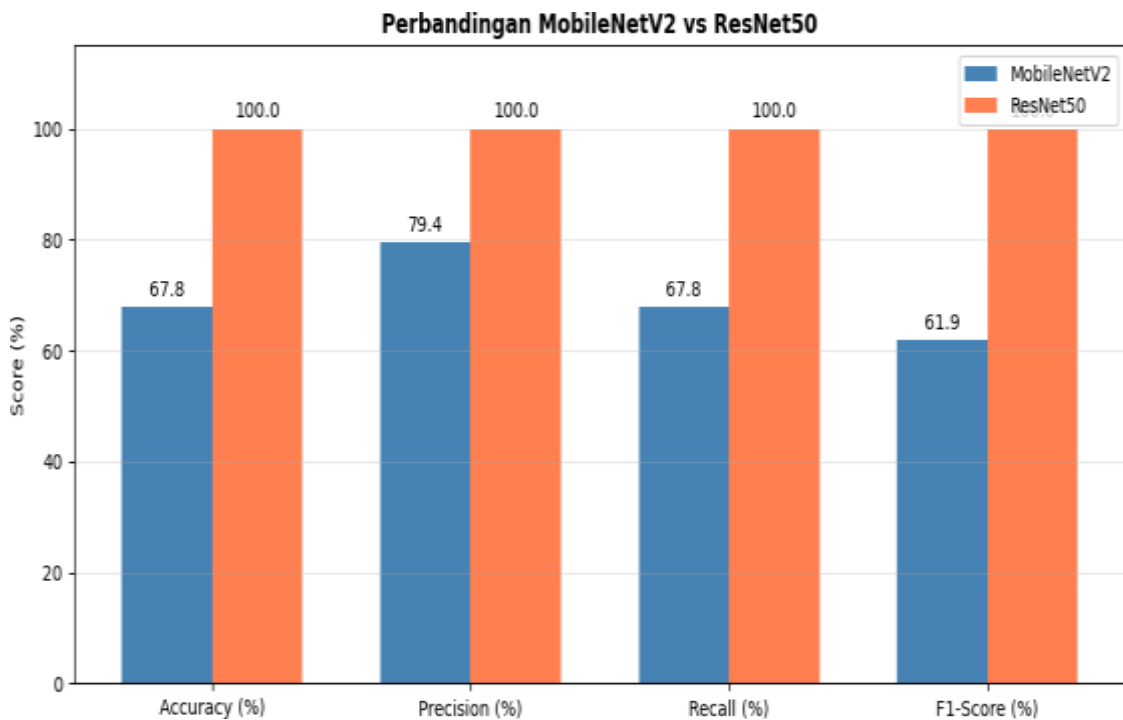
3. Hasil dan Diskusi

Hasil dari penelitian ini memaparkan evaluasi kinerja dua arsitektur Convolutional Neural Network (CNN), yakni MobileNetV2 dan ResNet50, yang telah dilatih menggunakan metode Transfer Learning untuk mengklasifikasikan kondisi permukaan jalan raya ke dalam tiga kategori: Jalan Kategori Baik, Jalan Kurang Baik, dan Jalan Rusak. Proses pengujian dilakukan menggunakan skema stratified split untuk memastikan distribusi kelas yang seimbang pada data latih, data validasi, dan data uji. Evaluasi kinerja dari kedua model diukur berdasarkan metrik Akurasi, Presisi, Recall, dan F1-Score.

Hasil evaluasi metrik klasifikasi dari kedua model pada data uji disajikan pada Tabel 1.

Tabel 1. Perbandingan metrik evaluasi kinerja arsitektur MobileNetV2 dan ResNet50

Model	Accuracy (%)	Precision (%)	Recall (%)
MobileNetV2	67.8	79.4	67.8
ResNet50	100.0	100.0	100.0



Gambar 1. Grafik perbandingan persentase skor metrik MobileNetV2 dan ResNet50

Berdasarkan data yang diperoleh pada Tabel 1 dan Gambar 1, terlihat adanya anomali performa antara kedua arsitektur. Arsitektur ResNet50 secara teoritis merupakan model skala besar dengan kedalaman dan kompleksitas parameter yang sangat tinggi. Namun, hasil evaluasi menunjukkan bahwa model ini mengalami *overfitting* yang ekstrem. Tingginya skor metrik pada ResNet50 mengindikasikan bahwa model tersebut tidak mengekstraksi dan mempelajari fitur visual esensial dari kondisi jalan, melainkan hanya menghafal *noise* serta letak piksel dari citra pada *dataset* latih yang ukurannya relatif terbatas. Ketidakseimbangan antara jumlah parameter model (kurang

lebih 23 juta parameter) dengan kuantitas citra jalan mengakibatkan menurunnya daya generalisasi model terhadap data baru.

Sebaliknya, arsitektur MobileNetV2 yang tergolong dalam skala ringan dengan parameter yang jauh lebih sedikit mampu memberikan daya tahan yang lebih baik terhadap *overfitting*. Arsitektur ini terpaksa melakukan ekstraksi fitur yang bermakna, seperti garis retakan dan tekstur aspal, karena keterbatasan memori untuk menghafal keseluruhan gambar.

Pembahasan dan penafsiran temuan ini semakin diperkuat dengan uji coba model pada lingkungan riil (*real-world testing*) menggunakan citra jalan raya baru yang tidak terdapat di dalam *dataset* awal.



Gambar 2. Hasil pengujian model pada citra jalan raya di luar *dataset* eksperimen

Pada Gambar 2, dapat diobservasi bahwa model ResNet50 gagal mengidentifikasi kerusakan dengan benar (mengklasifikasikan jalan berlubang sebagai Jalan Kategori Baik dengan tingkat kepercayaan 99.2%) akibat overconfidence yang bersumber dari hafalan data latih. Di sisi lain, MobileNetV2 mampu memprediksi kondisi citra dengan jauh lebih representatif sebagai Jalan Kurang Baik.

Temuan utama penelitian ini menunjukkan bahwa skor pengujian yang sangat tinggi tidak selalu identik dengan kemampuan generalisasi yang baik. ResNet50 menghasilkan akurasi, presisi, dan recall sebesar 100% pada subset data uji, tetapi memberikan prediksi yang keliru ketika dihadapkan pada citra jalan di luar dataset. Perbedaan tersebut perlu dibaca sebagai indikasi bahwa model telah menyesuaikan diri terlalu kuat terhadap karakteristik visual data eksperimen. Dalam klasifikasi citra, pola latar, sudut pengambilan gambar, pencahayaan, bayangan, dan komposisi piksel dapat ikut dipelajari sebagai petunjuk kelas apabila variasi data belum memadai. Akibatnya, model mampu memperoleh skor sempurna pada data yang memiliki distribusi serupa dengan data latih, tetapi tidak mempertahankan kinerjanya pada kondisi baru. Kajian klasifikasi dan deteksi kerusakan jalan juga menekankan bahwa ukuran, keragaman, serta representasi kondisi lapangan dalam dataset sangat menentukan kestabilan model [7], [9]. Oleh karena itu, hasil ResNet50 pada penelitian ini lebih tepat dipandang sebagai performa internal yang harus dikonfirmasi melalui pengujian eksternal, bukan sebagai bukti bahwa model telah menyelesaikan permasalahan klasifikasi jalan secara umum.

Kondisi tersebut berkaitan erat dengan perbedaan kapasitas kedua arsitektur. ResNet50 memiliki jaringan yang lebih dalam dan jumlah parameter yang jauh lebih besar dibandingkan MobileNetV2. Kapasitas yang besar memberikan kemampuan untuk membentuk representasi fitur yang kompleks, tetapi sekaligus meningkatkan risiko model mempelajari pola yang terlalu spesifik ketika jumlah sampel terbatas. Dataset penelitian hanya berjumlah 802 citra yang dibagi menjadi data latih, validasi, dan uji. Setelah pembagian, jumlah citra yang benar-benar digunakan untuk memperbarui bobot model menjadi lebih sedikit daripada keseluruhan dataset. Dalam situasi tersebut, lapisan klasifikasi tambahan dan proses fine-tuning dapat membuat model berkapasitas besar cepat mencapai nilai kehilangan yang rendah tanpa menjamin bahwa fitur yang dipelajari bersifat umum. Penelitian komparatif terhadap ResNet dan arsitektur lain menunjukkan bahwa keunggulan model dalam data tertentu sangat bergantung pada ukuran data, strategi pelatihan, dan karakteristik kelas [2], [7]. Dengan demikian, keunggulan teoritis ResNet50 dalam mengekstraksi fitur mendalam tidak otomatis menghasilkan keunggulan praktis ketika rasio antara kompleksitas model dan ketersediaan data tidak seimbang.

MobileNetV2 menunjukkan karakter yang berbeda. Walaupun nilai akurasi 67,8%, presisi 79,4%, recall 67,8%, dan F1-score 61,9% masih perlu ditingkatkan, model ini memberikan respons yang lebih masuk akal pada citra baru. Hal tersebut memperlihatkan bahwa efisiensi arsitektur tidak hanya berhubungan dengan kecepatan inferensi, tetapi juga dapat berfungsi sebagai bentuk pembatas kapasitas model. Depthwise separable convolution yang menjadi karakter utama keluarga MobileNet mengurangi jumlah operasi dan parameter, sehingga model lebih terdorong untuk memanfaatkan pola visual yang benar-benar dominan, seperti retakan, lubang, ketidakteraturan tekstur, serta perubahan warna permukaan. Studi terkini tentang modifikasi MobileNetV2 untuk deteksi lubang jalan juga menunjukkan bahwa arsitektur ringan tetap relevan untuk tugas pemantauan jalan yang membutuhkan efisiensi dan implementasi pada perangkat terbatas [17]. Hasil penelitian ini sejalan dengan arah tersebut karena MobileNetV2 tidak memperoleh skor internal tertinggi, tetapi memperlihatkan perilaku prediksi yang lebih representatif pada pengujian dunia nyata.

Perbedaan antara skor data uji dan hasil pengujian eksternal juga menegaskan pentingnya memeriksa kemungkinan kemiripan antar-citra. Stratified Shuffle Split telah menjaga proporsi kelas pada data latih, validasi, dan uji, tetapi stratifikasi tidak secara otomatis mencegah citra yang berasal dari ruas jalan, video, atau rangkaian pengambilan yang sama tersebar ke beberapa subset. Apabila terdapat citra yang sangat mirip, model dapat mengenali tekstur latar atau pola pencahayaan yang berulang, sehingga skor pengujian menjadi lebih optimistik. Pada penelitian lanjutan, pemisahan data sebaiknya dilakukan berdasarkan kelompok sumber, lokasi, sesi pengambilan, atau urutan video sebelum stratifikasi kelas diterapkan. Pendekatan ini akan menghasilkan evaluasi yang lebih ketat karena citra dari satu kelompok hanya berada pada satu subset. Dataset multinasional RDD2022 dikembangkan untuk memperluas variasi lingkungan, negara, dan tipe kerusakan sehingga pengujian model dapat mencerminkan tantangan generalisasi yang lebih realistis [9]. Prinsip yang sama dapat diterapkan pada dataset penelitian ini dengan menambahkan metadata lokasi dan kondisi pengambilan citra.

Tidak digunakannya augmentasi citra pada rancangan eksperimen merupakan faktor penting lain dalam menafsirkan hasil. Augmentasi dapat memperkenalkan variasi rotasi ringan, pergeseran, perubahan kecerahan, kontras, zoom, atau transformasi perspektif tanpa mengubah label kelas. Pada dataset kecil, variasi tersebut membantu mengurangi ketergantungan model terhadap posisi objek dan karakter visual yang kebetulan muncul pada data latih. Penelitian klasifikasi citra menunjukkan bahwa augmentasi dapat memperkaya representasi data dan menekan overfitting [3], sedangkan penelitian deteksi kerusakan aspal menggunakan model difusi memperlihatkan bahwa data sintesis dapat membantu memperbanyak contoh kerusakan yang jarang muncul [16]. Meskipun demikian, augmentasi untuk citra jalan harus dirancang secara hati-hati. Transformasi vertikal ekstrem, rotasi besar, atau perubahan warna yang berlebihan dapat menghasilkan citra yang tidak realistis. Oleh sebab itu, eksperimen berikutnya dapat membandingkan pelatihan tanpa augmentasi, augmentasi ringan, dan augmentasi adaptif untuk menentukan strategi yang paling efektif bagi MobileNetV2 maupun ResNet50.

Strategi transfer learning yang digunakan sudah tepat untuk kondisi data terbatas karena bobot awal ImageNet menyediakan representasi fitur visual umum. Akan tetapi, besarnya bagian jaringan yang dibuka ketika fine-tuning perlu disesuaikan dengan ukuran dataset dan tingkat kemiripan domain. Penelitian ini membuka 30 lapisan terakhir MobileNetV2 dan 10 lapisan terakhir ResNet50 dengan learning rate yang lebih kecil. Pada ResNet50, meskipun jumlah lapisan yang dibuka lebih sedikit, bagian tersebut tetap dapat membawa banyak parameter yang dapat dilatih. Salah satu alternatif adalah melakukan fine-tuning bertahap, dimulai dari kepala klasifikasi, kemudian membuka blok terakhir satu per satu berdasarkan perkembangan validation loss. Kerangka transfer learning lintas kondisi jalan menunjukkan bahwa adaptasi model terhadap lingkungan baru membutuhkan pengelolaan transfer fitur dan variasi data secara bersama-sama [10]. Penelitian Mask R-CNN pada kerusakan permukaan jalan juga menggunakan bobot pralatih untuk mengatasi keterbatasan jumlah sampel dan menunjukkan pentingnya pemilihan backbone serta learning rate [12].

Penggunaan Early Stopping dan ReduceLROnPlateau membantu mengendalikan pelatihan, tetapi efektivitasnya sangat bergantung pada kualitas data validasi. Apabila data validasi memiliki pola yang sangat mirip dengan data latih, nilai validation loss dapat tetap membaik walaupun kemampuan model pada lingkungan baru tidak meningkat. Untuk mengurangi risiko tersebut, evaluasi dapat menggunakan repeated stratified holdout berbasis kelompok atau cross-validation yang mempertahankan pemisahan lokasi. Setiap fold perlu menyimpan data dari lokasi tertentu sebagai data yang benar-benar tidak terlihat. Nilai rata-rata dan simpangan baku antar-fold kemudian dapat digunakan untuk menilai stabilitas arsitektur. Model yang memperoleh rata-rata sedikit lebih rendah tetapi memiliki variasi kinerja kecil dapat lebih aman untuk implementasi dibandingkan model dengan skor

tertinggi pada satu pembagian data. Pengujian lintas skenario pada penelitian terdahulu menunjukkan bahwa perubahan perangkat, posisi kamera, dan lingkungan dapat menurunkan kinerja model apabila proses transfer tidak dirancang secara khusus [10]. Karena itu, kestabilan antar-kondisi sebaiknya menjadi bagian utama dalam komparasi arsitektur.

Metrik evaluasi dalam penelitian ini juga perlu dibaca secara saling melengkapi. Akurasi menunjukkan proporsi prediksi benar secara keseluruhan, tetapi tidak menjelaskan kelas mana yang paling sering salah. Presisi MobileNetV2 yang lebih tinggi daripada recall mengindikasikan bahwa prediksi positif pada kelas tertentu relatif lebih dapat dipercaya, namun masih terdapat sampel yang gagal dikenali. F1-score sebesar 61,9% pada grafik menguatkan bahwa keseimbangan antara presisi dan recall belum optimal. Confusion matrix per kelas diperlukan untuk melihat apakah kesalahan terutama terjadi antara Jalan Kategori Baik dan Jalan Kurang Baik atau antara Jalan Kurang Baik dan Jalan Rusak. Kedekatan karakter visual antarkelas dapat menyebabkan batas keputusan menjadi kabur, khususnya ketika label didasarkan pada penilaian umum kondisi permukaan. Penelitian klasifikasi kerusakan jalan di Pontianak dan kajian deteksi retakan aspal memperlihatkan bahwa definisi kelas dan pola kerusakan yang terukur berperan besar dalam keberhasilan CNN [4], [6].

Skor kepercayaan yang sangat tinggi pada prediksi ResNet50 yang salah merupakan temuan penting. Nilai confidence 99,2% pada citra jalan berlubang yang diklasifikasikan sebagai Jalan Kategori Baik menunjukkan bahwa keluaran softmax belum terkalibrasi. Softmax dapat menghasilkan probabilitas tinggi meskipun citra berada di luar distribusi data pelatihan. Dalam aplikasi pemantauan jalan, kondisi tersebut berisiko karena sistem dapat menyampaikan keputusan yang salah dengan keyakinan tinggi. Oleh karena itu, selain akurasi, penelitian lanjutan dapat mengukur Expected Calibration Error, Brier Score, atau reliability diagram. Teknik temperature scaling dapat diterapkan setelah pelatihan untuk menyesuaikan probabilitas tanpa mengubah kelas prediksi. Sistem juga dapat menetapkan ambang ketidakpastian; citra dengan probabilitas maksimum rendah atau perbedaan antarprobabilitas kecil diarahkan untuk pemeriksaan manual. Mekanisme tersebut penting agar kecerdasan buatan berfungsi sebagai pendukung inspeksi, bukan menggantikan penilaian teknis secara mutlak.

Pengujian pada citra di luar dataset merupakan kekuatan penelitian karena membuka perbedaan yang tidak terlihat pada evaluasi internal. Walaupun jumlah citra eksternal yang ditampilkan masih terbatas, prosedur ini memperlihatkan bahwa validasi lapangan perlu ditempatkan sebagai bagian formal dari rancangan eksperimen. Pengujian berikutnya dapat mencakup variasi jalan kering dan basah, siang dan sore, permukaan yang tertutup bayangan, sudut kamera berbeda, serta ruas dengan marka dan objek lain. Dataset besar seperti RDD2022 mencakup puluhan ribu citra dari enam negara, sedangkan UAV-PDD2023 menyajikan kerusakan jalan dari perspektif udara dan kondisi yang bervariasi [9], [15]. Kedua contoh tersebut menunjukkan bahwa keragaman sumber data merupakan komponen penting dalam penilaian model. Penelitian ini dapat dikembangkan dengan mempertahankan tiga kelas utama, tetapi menambah citra dari beberapa lokasi dan perangkat agar hasil komparasi MobileNetV2 dan ResNet50 lebih kuat.

Hasil penelitian juga memiliki implikasi terhadap pemilihan bentuk tugas. Klasifikasi tingkat citra memberikan satu label untuk seluruh gambar dan sesuai untuk penyaringan awal kondisi jalan. Namun, satu citra dapat memuat lebih dari satu jenis kerusakan atau hanya menampilkan kerusakan pada area kecil. Dalam kasus tersebut, label global dapat mengabaikan lokasi dan luas kerusakan. Penelitian deteksi multi-kerusakan menunjukkan bahwa model berbasis ResNet dapat digunakan untuk klasifikasi sekaligus pelokalan area bermasalah [7], sedangkan Mask R-CNN memungkinkan deteksi dan segmentasi objek kerusakan secara lebih rinci [12]. Model RoadNet dan DASNet juga dikembangkan untuk mendeteksi serta membedakan tipe kerusakan pada lingkungan jalan [8], [14]. Dengan demikian, hasil MobileNetV2 dalam penelitian ini dapat dijadikan tahap awal sebelum sistem dikembangkan menuju object detection atau semantic segmentation, terutama apabila kebutuhan pengguna tidak hanya menentukan kelas kondisi tetapi juga menemukan posisi kerusakan.

Dari sisi implementasi, MobileNetV2 memiliki peluang lebih besar untuk digunakan pada perangkat bergerak, kamera kendaraan, atau unit komputasi tepi. Model ringan membutuhkan memori dan operasi komputasi lebih rendah, sehingga proses inferensi dapat dilakukan lebih dekat dengan sumber data tanpa selalu mengirimkan seluruh citra ke server. Pendekatan ini mengurangi kebutuhan bandwidth dan memungkinkan pemberian peringatan secara lebih cepat. Penelitian RoadNet menekankan pentingnya model yang efisien untuk mempercepat pelaporan kerusakan jalan [8], sedangkan pengembangan MobileNetV2 termodifikasi menunjukkan bahwa keluarga arsitektur tersebut dapat disesuaikan untuk pengenalan lubang jalan [17]. Walaupun penelitian ini belum

mencatat waktu inferensi, ukuran model, penggunaan memori, dan jumlah operasi, aspek tersebut perlu ditambahkan pada komparasi berikutnya. Pemilihan model untuk sistem nyata seharusnya mempertimbangkan bukan hanya akurasi, tetapi juga latensi, konsumsi energi, ukuran berkas model, dan stabilitas prediksi.

Interpretasi visual dapat membantu memastikan bahwa model benar-benar memperhatikan kerusakan jalan. Metode seperti Grad-CAM dapat menghasilkan peta perhatian yang menunjukkan wilayah citra yang paling berpengaruh terhadap keputusan. Apabila ResNet50 menyoroti pepohonan, tepi jalan, bayangan, atau marka dibandingkan lubang dan retakan, bukti tersebut akan mendukung dugaan bahwa model mempelajari korelasi yang tidak relevan. Sebaliknya, apabila MobileNetV2 memusatkan perhatian pada tekstur aspal dan area kerusakan, hasil tersebut menjelaskan mengapa prediksinya lebih representatif pada citra luar dataset. Penelitian klasifikasi citra perkerasan menggunakan Vision Transformer juga memasukkan interpretasi visual untuk menjelaskan dasar keputusan model [13]. Analisis semacam ini dapat melengkapi confusion matrix dan metrik numerik, karena dua model dengan skor serupa belum tentu menggunakan fitur yang sama. Dalam konteks infrastruktur, kemampuan menjelaskan keputusan penting agar hasil sistem dapat diperiksa oleh pengelola jalan.

Kualitas label juga perlu dipertimbangkan. Kategori Jalan Baik, Jalan Kurang Baik, dan Jalan Rusak bersifat mudah dipahami, tetapi batas antarkategori dapat menjadi subjektif apabila tidak disertai kriteria operasional. Citra dengan satu retakan kecil mungkin dinilai berbeda oleh pelabel yang berbeda, sedangkan jalan dengan tambalan dapat memiliki tekstur kasar tetapi masih layak digunakan. Ketidakpastian label dapat membatasi nilai maksimum yang realistis dan membuat model mempelajari perbedaan yang tidak konsisten. Proses pelabelan berikutnya sebaiknya melibatkan lebih dari satu penilai, menggunakan pedoman visual, serta menghitung kesepakatan antarpemilai. Kelas juga dapat dilengkapi atribut jenis kerusakan, luas kerusakan, atau tingkat keparahan. Penelitian klasifikasi ukuran lubang berdasarkan luas permukaan menunjukkan bahwa informasi tingkat kerusakan dapat memperkaya keluaran model [18]. Dengan label yang lebih terstruktur, perbandingan arsitektur tidak hanya menilai benar atau salah, tetapi juga kesesuaian prediksi dengan kebutuhan pemeliharaan.

Jika ResNet50 tetap digunakan, beberapa tindakan dapat diterapkan untuk mengurangi overfitting. Pertama, dropout pada kepala klasifikasi dapat dikombinasikan dengan weight decay agar bobot tidak tumbuh terlalu besar. Kedua, jumlah neuron pada lapisan tersembunyi dapat dikurangi sehingga kepala klasifikasi lebih sebanding dengan jumlah data. Ketiga, lebih banyak lapisan base model dapat dipertahankan dalam kondisi beku, terutama pada tahap awal, kemudian dibuka secara bertahap setelah model stabil. Keempat, label smoothing dapat mencegah model menjadi terlalu yakin terhadap satu kelas. Kelima, augmentasi dan hard-example mining dapat memperbanyak contoh yang sulit dibedakan. Penelitian deteksi kerusakan jalan berbasis CNN menggunakan fitur grayscale dan HOG menunjukkan bahwa penyederhanaan representasi dapat menekan kebutuhan komputasi [11], sedangkan penelitian deteksi perkerasan lainnya memanfaatkan dropout, normalisasi, dan transfer learning untuk meningkatkan stabilitas [7]. Kombinasi strategi tersebut perlu diuji dengan protokol evaluasi eksternal yang sama agar perbaikannya tidak hanya terlihat pada data internal.

Secara keseluruhan, komparasi ini memperlihatkan bahwa pemilihan arsitektur harus mengikuti konteks data dan tujuan penerapan. ResNet50 berpotensi memberikan ekstraksi fitur yang kuat apabila tersedia dataset besar, beragam, dan protokol regularisasi yang memadai. MobileNetV2 lebih sesuai ketika jumlah data, perangkat keras, dan kebutuhan inferensi cepat menjadi batasan utama. Temuan penelitian tidak berarti bahwa arsitektur dalam selalu lebih buruk, melainkan menunjukkan bahwa kapasitas model harus diseimbangkan dengan jumlah dan variasi data. Penelitian lokal mengenai deteksi kerusakan jalan dengan CNN mendukung penggunaan pengolahan citra otomatis untuk membantu pemantauan infrastruktur [5], sementara penelitian transfer learning terbaru memperlihatkan beragam pilihan backbone untuk kondisi jalan [1], [2]. Berdasarkan hasil data uji dan pengujian dunia nyata, MobileNetV2 menjadi pilihan yang lebih rasional pada konfigurasi penelitian saat ini. Pengembangan selanjutnya perlu memprioritaskan peningkatan kualitas data, pengujian lintas lokasi, kalibrasi probabilitas, serta pengukuran efisiensi komputasi agar kesimpulan arsitektural menjadi lebih kuat dan dapat direplikasi.

4. Kesimpulan

Penelitian ini menyimpulkan bahwa arsitektur berskala ringan seperti MobileNetV2 lebih unggul dan sangat direkomendasikan untuk sistem klasifikasi kondisi jalan dengan dataset terbatas dibandingkan model skala besar seperti ResNet50. ResNet50 terbukti mengalami overfitting ekstrem akibat ketidakseimbangan antara jumlah parameter model (± 23 juta) dengan kuantitas citra, sehingga gagal memprediksi citra jalan raya di dunia nyata.

Sebaliknya, MobileNetV2 lebih efisien, tahan terhadap overfitting, dan akurat di kondisi riil. Disarankan bagi pengembang sistem pemantauan jalan untuk menggunakan arsitektur ringan atau menerapkan teknik augmentasi data secara intensif jika menggunakan model yang kompleks.

Referensi

1. Zaid, M. M. A., Mohammed, A. A., & Sumari, P. (2025). Classification of Road Features Using Convolutional Neural Network (CNN) and Transfer Learning. *International Journal of Computing and Digital Systems*, 17(1), 1-12. <https://doi.org/10.12785/iicds/1571031764>
2. Abdul Ghofur, M., & Ulum, A. B. (2025). Analisis Perbandingan Kinerja Model ResNet50 dan EfficientNet-B0 dalam Klasifikasi Kerusakan Jalan. *Jurnal Media Informatika (JuMin)*, 6(5), 2504-2511. <https://doi.org/10.55338/jumin.v6i5.6843>
3. Fitriani, L., Tresnawati, D., & Sukriyansah, M. B. (2023). Image Classification on Garutan Batik Using Convolutional Neural Network with Data Augmentation. *JUITA: Jurnal Informatika*, 11(1), 107. <https://doi.org/10.30595/juita.v11i1.16166>
4. Mulia, A. W., Ruslianto, I., & Midyanti, D. M. (2023). Klasifikasi Kerusakan Jalan pada Citra Jalan Raya Pontianak dan Sekitarnya dengan Menggunakan Convolutional Neural Network. *Coding: Jurnal Komputer dan Aplikasi*, 11(1), 11-20. <https://doi.org/10.26418/coding.v11i1.57905>
5. Sasmito, B., Setiadi, B. H., & Isnanto, R. (2023). Deteksi Kerusakan Jalan Menggunakan Pengolahan Citra Deep Learning di Kota Semarang. *Teknik*, 44(1), 7-14. <https://doi.org/10.14710/teknik.v44i1.51908>
6. Wibowo, A., & Yulianto, Y. (2023). Deteksi Keretakan Jalan Aspal Menggunakan Metode Convolutional Neural Network. *Jurnal Teknik Sipil Cendekia*, 4(2), 581-593. <https://doi.org/10.51988/jtsc.v4i2.132>
7. Li, D., Duan, Z., Hu, X., Zhang, D., & Zhang, Y. (2023). Automated classification and detection of multiple pavement distress images based on deep learning. *Journal of Traffic and Transportation Engineering (English Edition)*, 10(2), 276-290. <https://doi.org/10.1016/j.jtte.2021.04.008>
8. Alqethami, S., Alghamdi, S., Alsubait, T., & Alhakami, H. (2022). RoadNet: Efficient model to detect and classify road damages. *Applied Sciences*, 12(22), 11529. <https://doi.org/10.3390/app122211529>
9. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2024). RDD2022: A multi-national image dataset for automatic road damage detection. *Geoscience Data Journal*, 11(4), 846-862. <https://doi.org/10.1002/gdj3.260>
10. Li, Y., Che, P., Liu, C., Wu, D., & Du, Y. (2021). Cross-scene pavement distress detection by a novel transfer learning framework. *Computer-Aided Civil and Infrastructure Engineering*, 36(11), 1398-1415. <https://doi.org/10.1111/mice.12674>
11. Chen, G.-H., Ni, J., Chen, Z., Huang, H., Sun, Y.-L., Ip, W.-H., & Yung, K.-L. (2022). Detection of highway pavement damage based on a CNN using grayscale and HOG features. *Sensors*, 22(7), 2455. <https://doi.org/10.3390/s22072455>
12. He, Y., Jin, Z., Zhang, J., Teng, S., Chen, G., Sun, X., & Cui, F. (2022). Pavement surface defect detection using Mask Region-Based Convolutional Neural Networks and transfer learning. *Applied Sciences*, 12(15), 7364. <https://doi.org/10.3390/app12157364>
13. Chen, Y., Gu, X., Liu, Z., & Liang, J. (2022). A fast inference Vision Transformer for automatic pavement image classification and its visual interpretation method. *Remote Sensing*, 14(8), 1877. <https://doi.org/10.3390/rs14081877>
14. Zhao, L., Wu, Y., Luo, X., & Yuan, Y. (2022). Automatic defect detection of pavement diseases. *Remote Sensing*, 14(19), 4836. <https://doi.org/10.3390/rs14194836>
15. Yan, H., & Zhang, J. (2023). UAV-PDD2023: A benchmark dataset for pavement distress detection based on UAV images. *Data in Brief*, 51, 109692. <https://doi.org/10.1016/j.dib.2023.109692>
16. Cano-Ortiz, S., Lloret Iglesias, L., Martinez Ruiz del Arbol, P., & Castro-Fresno, D. (2024). Improving detection of asphalt distresses with deep learning-based diffusion model for intelligent road maintenance. *Developments in the Built Environment*, 17, 100315. <https://doi.org/10.1016/j.dibe.2023.100315>
17. Tanwar, N., & Turukmane, A. V. (2025). Modified MobileNetV2 transfer learning model to detect road potholes. *PeerJ Computer Science*, 11, e2519. <https://doi.org/10.7717/peerj-cs.2519>
18. Ahmad, C. F., Al-Sayegh, A. T., Cheema, A., et al. (2024). Classification of different size of potholes based on surface area using convolutional neural network. *Discover Applied Sciences*, 6, 492. <https://doi.org/10.1007/s42452-024-06207-3>