

BAB I PENDAHULUAN

1.1. Latar Belakang

Pengolahan citra digital, khususnya dalam domain *computer vision*, telah menjadi bidang penelitian yang berkembang pesat dan merevolusi berbagai sektor industri. Perkembangan ini didorong oleh kemajuan pesat dalam *deep learning* yang memungkinkan mesin untuk belajar dari data dalam skala besar [1]. Salah satu cabang utamanya adalah pengenalan citra (*image recognition*), sebuah tugas yang bertujuan agar mesin mampu mengidentifikasi dan mengklasifikasikan objek dalam sebuah gambar secara otomatis. Untuk tugas ini, *Convolutional Neural Network* (CNN) telah lama menjadi arsitektur fondasi dalam pengenalan citra karena kemampuannya mengekstraksi fitur visual secara hierarkis, mulai dari tepi dan tekstur pada lapisan awal hingga bagian objek yang lebih kompleks pada lapisan yang lebih dalam [1].

Namun, upaya meningkatkan akurasi dengan memperdalam jaringan (*deeper networks*) justru memunculkan masalah degradasi: penambahan lapisan pada CNN konvensional menyebabkan error pelatihan meningkat. Fenomena ini terbukti bukan disebabkan oleh *overfitting*, melainkan oleh semakin sulitnya jaringan yang sangat dalam untuk dioptimalkan [2]. Inovasi utama ResNet terletak pada mekanisme *shortcut connection* [2], yang memungkinkan *gradient* mengalir secara langsung melewati beberapa lapisan. Hal ini secara efektif mengatasi masalah *vanishing gradient* dan memungkinkan pelatihan jaringan yang jauh lebih dalam dengan tetap menjaga stabilitas dan akurasi.

Keunggulan arsitektur ResNet masih relevan hingga saat ini. Di tengah perkembangan arsitektur baru seperti *Vision Transformers*, ResNet dan variannya tetap banyak digunakan sebagai baseline yang kuat dan efisien dalam berbagai tugas klasifikasi citra [3]. Pendekatan *residual learning* memungkinkan model untuk fokus mempelajari fungsi residual, yaitu selisih antara input dan output lapisan, yang terbukti lebih mudah dioptimalkan daripada mempelajari pemetaan fungsi secara keseluruhan. Hasilnya, ResNet tidak hanya mampu mencapai kedalaman jaringan yang belum pernah ada sebelumnya, tetapi juga menunjukkan kemampuan generalisasi yang lebih baik terhadap data baru.

Di dalam keluarga ResNet sendiri, kedalaman jaringan merupakan pilihan desain yang menghadirkan trade-off mendasar. Varian yang lebih dalam seperti ResNet-50 memiliki kapasitas representasi yang lebih besar dan cenderung lebih akurat, tetapi menuntut biaya komputasi, memori, dan waktu inferensi yang lebih tinggi. Sebaliknya, varian yang lebih dangkal seperti ResNet-34 lebih ringan dan hemat sumber daya. *Trade-off* ini menjadi semakin tidak sederhana ketika model tidak dilatih dari nol, melainkan melalui *transfer learning*, yaitu memanfaatkan bobot pra-latih (misalnya dari ImageNet) lalu menyesuaikannya pada domain sasaran yang datanya terbatas. Dalam kondisi seperti ini, arsitektur yang lebih dalam belum tentu selalu memberikan keuntungan yang sepadan dengan biayanya; model yang lebih dangkal berpotensi menggeneralisasi secara setara dengan biaya jauh lebih rendah. Karena itu, pemilihan kedalaman arsitektur bukan persoalan yang dapat diasumsikan begitu saja, melainkan keputusan praktis yang memerlukan bukti empiris melalui perbandingan yang terkendali.

Persoalannya, perbandingan antararsitektur sering dilakukan secara terlalu sempit, yaitu hanya menilai akurasi global pada data uji yang bersih dan seimbang. Padahal, perbandingan yang benar-benar berguna untuk pengambilan keputusan menuntut ketelitian pada beberapa aspek. Pertama, pada distribusi data nyata yang timpang (*class imbalance*), akurasi global dapat menutupi performa buruk pada kelas minoritas sehingga menimbulkan ilusi akurasi; oleh karena itu diperlukan metrik yang memberi bobot setara antarkelas seperti *macro F1-score* serta analisis *confusion matrix* [7]. Kedua, proses augmentasi dan pembagian data yang tidak dirancang dengan benar dapat menimbulkan data leakage, yang membuat performa tampak tinggi secara artifisial dan tidak mencerminkan kemampuan generalisasi sebenarnya [8]. Ketiga, akurasi semata tidak cukup untuk kebutuhan praktis: arsitektur juga perlu dibandingkan dari sisi efisiensi komputasi, waktu inferensi dan ukuran model karena sistem sasaran kerap berjalan pada perangkat dengan sumber daya terbatas (*edge computing*) seperti ponsel cerdas praktis [9], [10]. Aspek-aspek *robustness*, kalibrasi, dan efisiensi inilah yang justru sering absen dari perbandingan arsitektur, sehingga menjadi celah yang perlu diisi.

Untuk menguji perbandingan tersebut dalam kondisi yang menantang dan realistis, penelitian ini menggunakan studi kasus dataset citra makanan lokal

Indonesia. Kasus ini dipilih secara sengaja karena menekan seluruh aspek di atas sekaligus: ia bersifat *fine-grained* dengan kemiripan antar-kelas yang tinggi misalnya beberapa makanan tradisional yang tampak serupa dan mudah tertukar akibat perbedaan pencahayaan atau sudut pengambilan gambar [4] memiliki distribusi kelas yang timpang secara alami (*long-tail*), serta relevan langsung dengan skenario *deployment* pada perangkat terbatas seperti aplikasi pemantau diet. Selain itu, makanan lokal masih kurang terwakili pada dataset global standar seperti *Food-101*, sehingga menjadi testbed yang otentik dan bukan sekadar adaptasi sederhana dari studi pada dataset umum [5]. Kombinasi ResNet dengan transfer learning sendiri telah terbukti efektif untuk domain semacam ini [6], sehingga cocok dijadikan dasar perbandingan.

Berdasarkan uraian tersebut, membandingkan ResNet-34 dan ResNet-50 secara sistematis dalam kerangka transfer learning mencakup akurasi, performa yang seimbang antar-kelas, ketahanan terhadap gangguan (*robustness*), reliabilitas probabilitas (kalibrasi), serta efisiensi komputasi, dengan tetap menjaga validitas metodologis menjadi penting bukan sekadar untuk mengetahui arsitektur mana yang lebih akurat, melainkan untuk mengidentifikasi arsitektur yang menawarkan keseimbangan operasional terbaik bagi kebutuhan implementasi praktis. Dengan demikian, penelitian ini diposisikan sebagai sebuah studi komparatif, dengan klasifikasi citra makanan lokal Indonesia sebagai studi kasusnya. Berdasarkan latar belakang tersebut, diajukan penelitian dengan judul: "Studi Komparatif Kinerja Arsitektur ResNet-34 dan ResNet-50 Menggunakan Metode Transfer Learning".

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan permasalahan utama dalam penelitian ini sebagai berikut:

1. Bagaimana perbandingan kinerja arsitektur ResNet-34 dan ResNet-50 pada studi kasus citra makanan lokal Indonesia sebelum dan sesudah *finetuning* menggunakan pendekatan *transfer learning*, ditinjau dari akurasi, *macro F1-score*, dan analisis kesalahan per kelas?
2. Bagaimana pengaruh strategi *feature extraction* dan *fine-tuning*, serta peran augmentasi data sebagai mekanisme penyeimbangan kelas, terhadap performa model pada tugas klasifikasi citra makanan lokal Indonesia?

3. Arsitektur dan konfigurasi model manakah yang memberikan *trade-off* terbaik antara performa klasifikasi, stabilitas model, reliabilitas probabilitas, dan efisiensi komputasi, ditinjau dari hasil *clean benchmark*, *robustness*, simulasi *long-tail*, kalibrasi probabilitas, *inference time*, *model size*, serta optimasi deployment?

1.3. Batasan Masalah

Agar penelitian ini lebih terarah dan terfokus, beberapa batasan masalah yang diterapkan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini diuji pada studi kasus dataset citra makanan lokal Indonesia dua dimensi dengan 10 kelas makanan populer Indonesia.
2. Arsitektur yang dibandingkan dibatasi pada ResNet-34 dan ResNet-50 dengan pendekatan *transfer learning* menggunakan bobot pra-latih ImageNet.
3. Penelitian ini tidak melakukan pencarian hyperparameter secara ekstensif. Konfigurasi pelatihan ditetapkan sebagai protokol eksperimen terkontrol agar perbandingan antara ResNet-34 dan ResNet-50 serta antara feature extraction dan fine-tuning dapat dilakukan secara adil dan konsisten.
4. Implementasi penelitian dibatasi pada evaluasi kelayakan deployment model melalui penyimpanan artefak model, pengukuran efisiensi inferensi, serta optimasi model menggunakan *post-training quantization* dan *structured pruning*, tanpa membangun sistem aplikasi produksi penuh.

1.4. Tujuan

Penelitian ini bertujuan untuk membandingkan kinerja arsitektur ResNet-34 dan ResNet-50 menggunakan pendekatan *transfer learning*. Secara spesifik, penelitian ini bertujuan untuk:

1. Melakukan implementasi dan analisis komparatif kinerja arsitektur ResNet-34 dan ResNet-50 pada studi kasus citra makanan lokal menggunakan pendekatan *transfer learning*.
2. Menganalisis pengaruh strategi *feature extraction* dan *fine-tuning*, serta peran augmentasi data sebagai mekanisme penyeimbangan kelas, terhadap performa masing-masing arsitektur.

3. Menganalisis *trade-off* antara performa klasifikasi, stabilitas model, reliabilitas probabilitas, dan efisiensi komputasi melalui evaluasi *clean benchmark*, *robustness*, simulasi *long-tail*, kalibrasi probabilitas, *inference time*, *model size*, serta optimasi deployment menggunakan *post-training quantization* dan *structured pruning*.

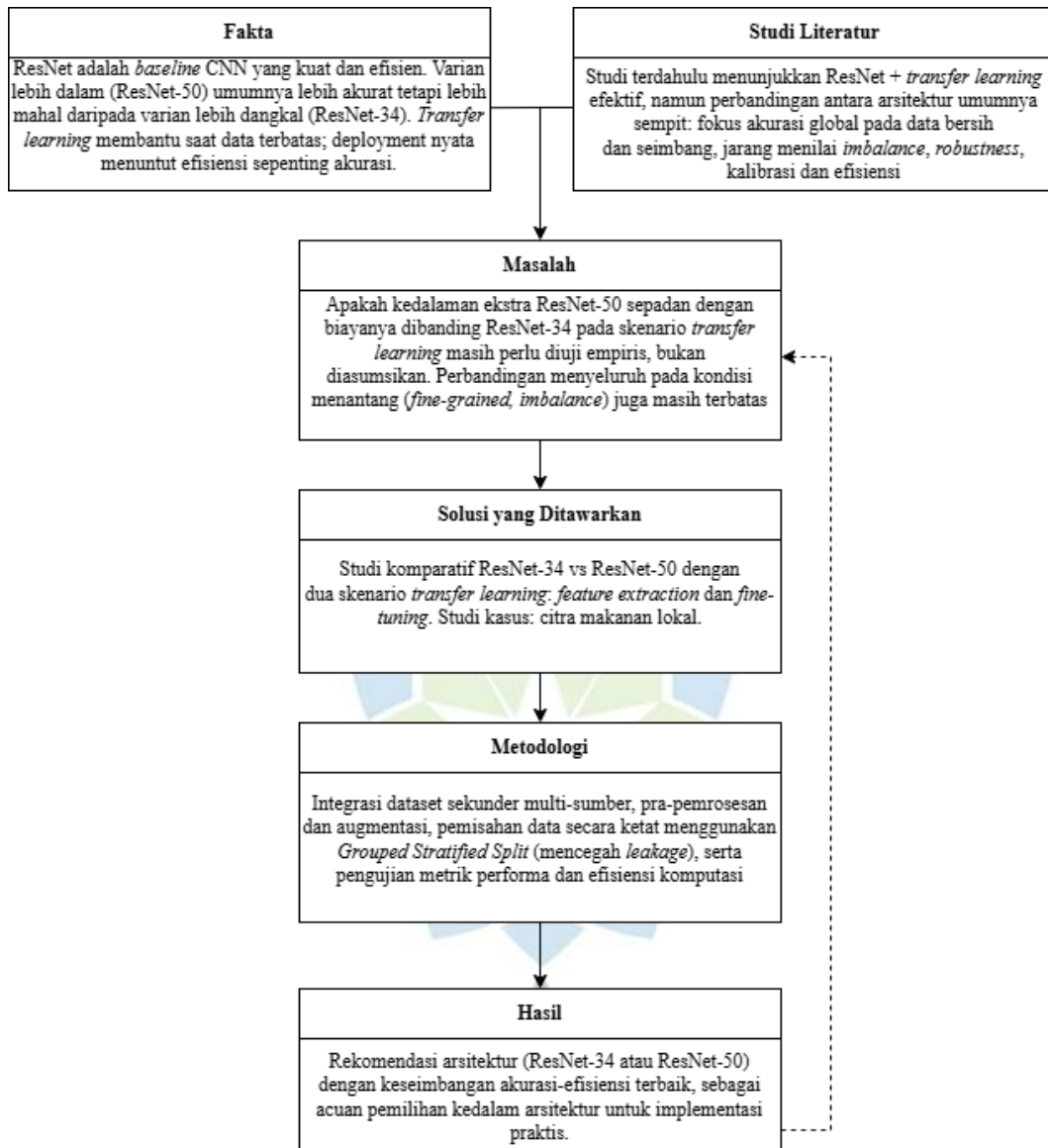
1.5. Manfaat

Penelitian ini diharapkan dapat memberikan beberapa manfaat, antara lain:

1. **Manfaat Teoretis:** memperkaya studi literatur terkait transfer learning dan trade-off kedalaman model (ResNet-34 vs ResNet-50) pada tugas klasifikasi citra fine-grained domain makanan.
2. **Manfaat Praktis:** menghasilkan rekomendasi model yang akurat dan efisien untuk mendukung pengembangan aplikasi.
3. **Manfaat Metodologis:** menyediakan alur eksperimen yang terstruktur (CRISP-DM) beserta praktik pengolahan data (*merging*, *cleaning*, *anti data leakage*) dan evaluasi efisiensi komputasi yang dapat direplikasi.

1.6. Kerangka Berpikir

Kerangka berpikir penelitian ini dirancang untuk memetakan alur logis penelitian secara sistematis, mulai dari identifikasi konteks yang lebih luas hingga luaran spesifik yang diharapkan, sebagaimana divisualisasikan pada Gambar 1.1.



Gambar 1.1 Kerangka Pemikiran

Secara faktual, *Convolutional Neural Network*, khususnya arsitektur ResNet, telah menjadi fondasi yang kuat dan efisien untuk tugas pengenalan citra dan tetap relevan sebagai baseline hingga saat ini. ResNet tersedia dalam beberapa varian kedalaman; varian yang lebih dalam seperti ResNet-50 umumnya memiliki kapasitas representasi dan akurasi yang lebih tinggi, tetapi menuntut biaya komputasi, memori, dan waktu inferensi yang lebih besar dibanding varian yang lebih dangkal seperti ResNet-34. Di sisi lain, *transfer learning* memungkinkan model pra-latih diadaptasi ke domain baru meskipun data berlabelnya terbatas, sementara kebutuhan implementasi nyata kerap menuntut model berjalan pada

perangkat dengan sumber daya terbatas sehingga efisiensi menjadi sama pentingnya dengan akurasi.

Dari fakta tersebut muncul masalah. Pertama, belum jelas apakah tambahan kedalaman ResNet-50 benar-benar sepadan dengan biayanya dibandingkan ResNet-34, terutama pada skenario *transfer learning* dengan data terbatas, sehingga pemilihan arsitektur memerlukan bukti empiris, bukan asumsi. Kedua, perbandingan antararsitektur pada banyak studi masih terlalu sempit karena hanya menilai akurasi global pada data yang bersih dan seimbang, tanpa memperhitungkan ketidakseimbangan kelas (*macro F1-score*), ketahanan terhadap gangguan, kalibrasi probabilitas, dan efisiensi komputasi, serta tanpa menjaga validitas dari risiko *data leakage*. Agar perbandingan diuji dalam kondisi yang benar-benar menantang, diperlukan sebuah studi kasus yang menekan seluruh aspek tersebut; citra makanan lokal Indonesia dipilih karena bersifat *fine-grained*, memiliki distribusi kelas yang timpang, dan realistis untuk skenario *deployment* pada perangkat terbatas.

Untuk menjawab masalah tersebut, penelitian ini menerapkan pendekatan berupa studi komparatif terhadap dua arsitektur residual, yaitu ResNet-34 dan ResNet-50, dengan dua skenario transfer learning: *feature extraction* dan *fine-tuning*. Perbandingan dilakukan pada dataset sekunder multi-sumber yang telah melalui cleaning, label mapping, penyeimbangan kelas melalui augmentasi, serta pembagian data dengan grouped stratified split untuk mencegah *data leakage*. Evaluasi tidak dibatasi pada akurasi, melainkan mencakup *macro F1-score* dan analisis kesalahan per kelas, ketahanan terhadap gangguan visual, simulasi distribusi *long-tail*, kalibrasi probabilitas, waktu inferensi, ukuran model, serta optimasi deployment melalui *post-training quantization* dan *structured pruning*.

Hasil yang diharapkan dari alur tersebut bukan sekadar sebuah model klasifikasi, melainkan rekomendasi arsitektur antara ResNet-34 dan ResNet-50 beserta konfigurasi pelatihannya yang menawarkan keseimbangan operasional terbaik antara akurasi dan efisiensi. Dengan demikian, penelitian ini menghasilkan kesimpulan komparatif yang dapat menjadi acuan dalam memilih kedalaman arsitektur untuk kebutuhan implementasi praktis pada domain serupa.

1.7. Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang proses penulisan yang dilakukan untuk menyelesaikan penelitian ini. Sistematika penulisan penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini terdiri dari latar belakang, perumusan masalah penelitian, batasan masalah penelitian, tujuan penelitian, manfaat penelitian, kerangka pemikiran penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini terdiri dari pembahasan konsep dasar dan teori-teori yang berkaitan dengan topik masalah yang diambil dan hal-hal yang berguna dalam proses analisis permasalahan.

BAB III METODOLOGI PENELITIAN

Bab ini menjabarkan alur metodologi penelitian dari tahap awal hingga akhir menggunakan kerangka CRISP-DM. Setiap tahap dirancang untuk menjawab rumusan masalah pada BAB I dengan berlandaskan teori pada BAB II, dan menjadi acuan pelaksanaan eksperimen yang hasilnya dilaporkan pada BAB IV.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memuat dua hal utama. Pertama, pemaparan hasil penelitian berdasarkan tahapan metodologi pada BAB III. Kedua, pembahasan atas hasil tersebut yang diarahkan untuk menjawab rumusan masalah penelitian secara berurutan.

BAB V PENUTUP

Bab ini berisi kesimpulan yang merangkum jawaban atas rumusan masalah berdasarkan hasil dan pembahasan pada BAB IV, serta saran sebagai bahan pertimbangan untuk pengembangan penelitian selanjutnya.